Using Scaffolding to Examine The Development of Metacognitive Monitoring and Control

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Allison Paige O'Leary, B.A., M.A.

Graduate Program in Psychology

The Ohio State University

2017

Dissertation Committee:

Vladimir M. Sloutsky, Ph.D., Advisor

Andrew B. Leber, Ph.D.

John E. Opfer, Ph.D.

Abstract

Metacognition describes the ability to represent and access our own cognitive processes. This ability is crucial for understanding and optimizing how we learn and remember, allowing us to avoid strategies that have not worked for us in the past, and to continue using strategies that have. Metacognition has been described as involving two components, monitoring and control, which may have different developmental trajectories. In the current project, we addressed several remaining questions about metacognition and its development. How do monitoring and control develop? How do these components interact? Is metacognitive proficiency malleable in childhood? What kinds of information do children rely on to monitor and control behavior? In 8 experiments, these questions were addressed by using scaffolding such as feedback and strategy instruction to improve metacognitive performance across the lifespan. In Experiment 1, 5-year-olds, 7-year-olds, and adults' metacognitive monitoring and control were tested in a task that required them to initiate these processes spontaneously, demonstrating developmental differences in both monitoring and control. In Experiment 2, 5-year-olds were presented with performance feedback, strategy instruction, or both to assess their effects on the monitoring and control components. Whereas feedback influenced task monitoring, it did not influence metacognitive control. In addition, whereas strategy instruction improved control, it did not influence performance monitoring. These findings were expanded upon in Experiments 3, 4, and 5, wherein 5-

year-olds, 7-year-olds, and adults were provided with no scaffolding, strategy instruction, or feedback, respectively, to assess whether monitoring and control can function independently. Across the age groups, feedback improved performance monitoring, but not metacognitive control. In addition, strategy instruction improved control, but not performance monitoring. These findings suggest a dissociation between the monitoring and control components that persists from early childhood to adulthood. Experiments 6a - 7b addressed whether young children's demonstrated insensitivity to feedback (in terms of metacognitive control) was due to insufficient separation between the task success probabilities used. These findings suggested that young children could rely on feedback to control behavior, but only when reward probabilities were sufficiently separated. On the other hand, young children seemed to rely on differences in effort to monitor the task difficulty. In the prior experiments it was found that young children are able to form and use a strategy to optimize their performance in a task. Experiment 8 assessed whether 5-year-old children are able to transfer a learned strategy to a novel task with different stimulus characteristics. It was found that children who were trained and learned to use a strategy rule transferred this strategy more readily to a novel task than did those who (1) received training but did not learn the strategy or (2) did not receive strategy training. These results are discussed in relation to theories of metacognition development, the role of task success representations in metacognition, and the broader implications of the current findings.

Acknowledgments

First, I want to thank my advisor, Vladimir Sloutsky, for the support and the intellectual and creative freedom to pursue these projects over the last several years. I have learned so much from Vladimir about development but, more importantly, I have learned how to think and evaluate ideas, and how to conduct rigorous scientific research.

I would like to thank Andy Leber for being kind and encouraging, for leading the best seminars, and for teaching me so much about attention.

I would also like to thank John Opfer for never failing to play devil's advocate, which has encouraged the rigor of my science.

I want to thank everyone who has kept the Cognitive Development Lab running smoothly during my time as a grad student. I want to thank Sravani Paladugu, Tracey Miser, Priya Patel, Naomi Polinsky, Alicia Scimeca, and Emily Neer for their help and support with all of the tasks that keep the research flowing. I also want to thank all the undergraduate research assistants who have helped me conduct my research over the years, with a special thanks to Collin Wear, Jasmyn Atalla, Shelby Bitter, and Gloria Ge.

I am especially grateful to my Cog Dev Lab family for sharing in this experience with me. Thanks to Xin Yao for being a source of humor and beautiful photos, to Hyungwook Yim for being tech-savvy and a constant source of wisdom. Thanks to Sophia Deng and Saebyul Lee for being brilliant and strong female role models. Many thanks to Keith Apfelbaum and Sam Rivera for making me laugh so much. Thanks to Tasha Posid for sharing an obsession with Gilmore Girls, to Olivera Savic for keeping our cubicle upbeat, and to Nate Blanco for teaching me to juggle. A special thanks goes to Dan Plebanek for being my partner in cynicism and for hosting trashy TV nights. An extra special thanks goes to Kevin Darby, my lab twin, for being a continual source of friendship, and enchiladas.

I also want to thank Kelly Capatosto, for being my soul twin and best friend and without whom I would have never made it through grad school. Special thanks to my partner Nick Pope, for enduring 2 grueling years of long distance, for loving me and taking care of me in stressful times, and for sharing many adventures with me. Finally, I want to thank my parents for being the most selfless people I know, and for encouraging, supporting, and loving me without condition.

Vita

2007 ....East Burke High School, Connelly Springs, NC

2011 ....B.A. University of North Carolina at Greensboro, Greensboro, NC

2011-2012 ....Graduate Fellow, The Ohio State University

2014 ....M.A. Psychology, The Ohio State University, Columbus, OH

2012-2016 ....Graduate Teaching Associate, Department of Psychology, The Ohio State University

2016-Present ....Graduate Research Associate, Department of Psychology, The Ohio State University

Publications

O'Leary, A. P. & Sloutsky, V. M. (2016). Carving metacognition at its joints: Protracted development of component processes. *Child Development*.

O'Leary, A. P., & Sloutsky, V. M. (2013) Auditory overshadowing in preschoolers: A Preference for the input, the system, or both? *In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth. (Eds.), Proceedings of the XXXV Annual Conference of the Cognitive Science Society* 1091-1096

Fields of Study

Major Field: Psychology

Specialization: Developmental Psychology

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

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*).

The goal of this dissertation research is to better understand the development of metacognitive monitoring and control, the kinds of information people rely on to monitor and control behavior, and how these two metacognitive components interact across the lifespan. To address these questions, I describe four lines of evidence: (1) the malleability of metacognitive monitoring and control in young children; (2) the independence of metacognitive monitoring and control; (3) effects of feedback and task experience in metacognition; and (4) transfer of metacognitive strategy use in young children. Following this, I present four studies (with a total of eight experiments) addressing each of these issues in turn. Finally, I conclude by summarizing the key findings of the reported studies and discussing their implications for the study of the development of monitoring and control and their interactions.

1.1 Malleability of Metacognitive Monitoring and Control in Young Children

There are a number of important findings pertaining to the development of metacognition and its components. These findings, however, have presented 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). The following sections review some of these findings, and suggest possible explanations for the discrepancies between them.

1.1.1 The Developmental Trajectory of Metacognitive Monitoring and Control: Early Versus Late Onset

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 (Roebers, 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 approach 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 even 3-year-old children may exhibit evidence of metacognitive control, 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.

1.1.2 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, who are often aware of their mistakes 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 possibility that metacognition is malleable and that scaffolding will improve young children's performance is tested in Chapter 3.

1.2 Independence of Metacognitive Monitoring and Control

Another topic of debate in the study of metacognition has been the nature of the interaction between metacognitive monitoring and control. One set of theories assumes a unidirectional, feed-forward relation between monitoring and control, such that proficient monitoring is a prerequisite for proficient control (e.g., Nelson & Narens, 1990; Son, 2004). Under this interpretation, one needs to detect (i.e., *monitor*) that one set is more difficult before they can formulate and execute the strategy to study the difficult set longer (i.e., *control* behavior). This possibility is supported by the developmental work discussed above which suggests that young children could identify the more difficult of two sets before they could control their behavior by prioritizing that set for study, with the former occurring around 6 years of age and the latter occurring around 8 years of age (Dufresne & Kobasigawa, 1989). This approach suggests that (a) metacognitive monitoring is necessary for metacognitive control and, therefore, (b) improvements in monitoring should precede improvements in control.

Another set of theories proposed an opposite unidirectional effect, such that proficient control underlies and leads to proficient monitoring (Koriat & Ackerman, 2010; Koriat, Ackerman, Adiv, Lockl, & Schneider, 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.

It is also possible that monitoring and control can function independently. 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. If this is the case, neither component is necessary or sufficient for the other component. There is some recent evidence pointing to possible independence of monitoring and control in 5-year-old children, in that improvements in one component did not always correspond to improvements in the other (O'Leary & Sloutsky, 2016). The next section focuses on the evidence for each of these possible interaction patterns.

1.2.1 *Evidence for Interactions Between Monitoring and Control*

Monitoring and control are frequently discussed together, but how do they actually interact? Some evidence supports a feed-forward interaction, in that proficiency in metacognitive control relies on proficiency in metacognitive monitoring. This relation has been previously referred to as the MC (Monitoring → Control) model (Koriat, Ma'ayan, & Nussinson, 2006; Koriat, Ackerman, Adiv, Lockl, & Schneider, 2014), and the same designation will be used throughout. Support for this model includes the finding that judgments of learning made by adults during an initial study phase predicted which items were later selected for re-study (i.e., items rated as more poorly learned were re-studied longer; Kornell & Metcalfe, 2006). Koriat et al. (2014) found that even older children displayed this correlation when 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). This model is also consistent with the discrepancy-reduction model of metacognition (Dunlosky & Hertzog, 1998), which suggests that, in a learning context, metacognitive control is based on the monitoring of encoding strength, and adjusted until an encoding threshold is reached.

Evidence also exists for an opposite feed-forward model in which metacognitive control guides metacognitive monitoring (i.e., the CM, or Control → Monitoring, model). This model is supported by studies showing that participants' judgments of learning 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 remember), 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 that this relation is weaker (Hoffmann-Biencourt, et al., 2010) or non-existent (Koriat, Ackerman, Lockl, & Schneider, 2009) in younger children.

Another possibility is that monitoring and control can operate relatively independently (this possibility is heretofore described as the *independence model*). Previous work has shown that, in young children, metacognition is malleable, and improvements in monitoring can occur without corresponding improvements in control (O'Leary & Sloutsky, 2016). For example, feedback improved 5-year-olds' monitoring of an easier task (relative to when no feedback was provided), but did not improve their control performance (i.e., their ability to actually *select* the easier task to improve performance). This suggests that improvements in monitoring need not correspond to improvements in control, at least for young children. In addition, providing 5-year-olds with a strategy increased their selection of an easier game (i.e., improved control performance), but did not improve their ability to monitor their accuracy in the task. This provides evidence that improvements in control can also transpire without corresponding improvements in monitoring. Taken together, these findings suggest that monitoring and control can operate independently in early childhood.

1.2.2 Using Scaffolding to Directly Investigate Component Independence

Both the MC and CM models predict that improvements in the first component (e.g., monitoring under the MC account) should lead to improvements in the second (e.g., control under the MC account). As described above, one way to directly investigate this possibility is to scaffold each component, and observe corresponding improvements in the other. For example, if a manipulation (e.g., feedback) improves monitoring, but not control, this would provide evidence against the CM model which states that control improvements should precede and give rise to monitoring improvements. In addition, if another manipulation (e.g., providing a strategy) improves metacognitive control, but not monitoring, this would provide evidence against the MC model which states that monitoring improvements should precede control improvements. This scaffolding method has been used to investigate component independence in 5-year-old children (O'Leary & Sloutsky, 2016), providing initial support for the independence model. These initial findings suggested that a systematic developmental investigation is needed to reach more definitive conclusions. Perhaps monitoring and control operate independently in young childhood, but become more coupled with experience and development. Below, we describe the evidence supporting the use of feedback and strategy instruction scaffolding to directly investigate component independence.

Feedback. Intuitively, giving a participant feedback about their performance should help them appraise how well they are performing the task. Indeed, receiving feedback after each response in a numerical discrimination task helped children identify which of two sets was easier to discriminate (O'Leary & Sloutsky, 2016). However, receiving performance feedback alone did not encourage young children to select the easier of two tasks to obtain a higher reward. Thus, performance feedback may improve children's metacognitive monitoring without improving their control. This pattern suggests that monitoring is indeed malleable, and that the two components may be sensitive to different types of information, suggesting independence.

Strategy instruction. Whereas feedback draws attention to one's performance, strategy instruction provides participants with an approach for completing a task with ease. This is thought to eliminate the need for the participant to formulate a strategy – meaning instead that the participant needs to only execute the given strategy. For example, in a task in which high performance was incentivized, participants were instructed to select the easier of two task options, without telling them which of the two tasks was easier (O'Leary & Sloutsky, 2016). Providing this strategy increased the proportion of easy task choices that 5-year-olds made, thus suggesting that the control component is also malleable in early childhood. In Chapter 4, a scaffolding approach is taken to directly investigate the independence of metacognitive monitoring and control across the lifespan.

1.3 Effects of Feedback and Task Experience in Metacognition

As described above, the interactions between monitoring and control remain a topic of debate, with some claiming the two components can operate independently (e.g., O'Leary & Sloutsky, 2016; O'Leary & Sloutsky, submitted), in that some interventions improve one component without affecting the other. For example, in a task where participants could choose an easy or a difficult game, performance feedback improved children's task and performance monitoring, whereas their metacognitive control (e.g., selecting a more beneficial strategy) was unaffected. In other words, providing performance feedback helped children 1) more accurately estimate their performance, and 2) identify an easier game, but did not lead them to adjust their strategy to select an easier game.

At first glance, these findings may also suggest that children's metacognitive control, unlike their metacognitive monitoring, is not sensitive to differences in explicit feedback. In other words, feedback information may affect the two components differently (e.g., see Miller & Geraci, 2011). However, there are a number of potential reasons feedback did not improve metacognitive control in young children, which do not reflect total insensitivity to feedback. Instead, due to young children's imperfect ability to accumulate feedback probabilities over time, there may be uncertainties surrounding their representations of success in each task, which could prevent them from successfully monitoring and controlling behavior.

1.3.1 Reducing Uncertainty in Representations of Task Success

For feedback to influence metacognitive control, individuals need to 1) remember and accumulate the probabilities of success in the two games, 2) sample from those representations of task success formed in memory, and 3) make a decision on the basis of that sampling. A lack of effect of feedback may indicate a breakdown of the process at any of these three steps. For example, young children may struggle to accumulate probabilities of success among different types of trials with variable probabilities of success, which take place over an extended amount of time. This may lead to representations of performance that have a high level of uncertainty. As such, it may be difficult for young participants to detect that performance differs in the two games because the sampled probabilities of success are 1) not sufficiently distant, or 2) not sufficiently precise.

To illustrate, consider that young children's accuracy in previous experiments has been approximately 60% in an easier game, and 90% in a more difficult game (O'Leary & Sloutsky, 2016). This creates a situation where the accuracy difference between the two games is only 30%. If the total number of trials is 30 and children select the easy and difficult game equally often (which they do under typical circumstances), then the average number of successes on the easy game is about 13.5 and on the difficult game is about 9. Therefore, one possibility is that that the ratio of successes in the two conditions (i.e., 1.5) is not large enough for young children to detect the difference. According to this possibility, the *proportion* of separation is not sufficient to drive young children to select the easier game, and the reward probabilities may need greater proportional separation to affect young children's strategic decision making. One way to test this possibility is to increase the distance between the success probabilities in the two games – ideally, to 0% vs. 100% – and evaluate the effects on children's selection of the easier game. We evaluate this possibility in Experiments 6a and 6b.

It is also possible that young children's performance estimations for the two games lacked sufficient *precision*. Because it may be difficult for young children to aggregate and keep track of their successes in two different games across many trials, their representations of success in the two games may be noisy. If this is the case, the problem is that the absolute *amount* of separation between the two games (i.e., 13.5 – 9 = 4.5 trials) may not be sufficient. One way of testing this possibility is to retain the proportion of successes in the two games, but increase the experience with two games, thus increasing the amount of separation, while retaining the ratio. For example, simply doubling the number of trials (assuming the same probability of success as before) will result in 27 successful trials in the easy game and 18 successful trials in the difficult game, with the difference between the two equaling 9. Therefore, increasing participants' sample of easy and difficult trials may reduce the uncertainty in their estimates of success in each game, resulting in more robust selection of the easy game. We evaluate this possibility in Experiments 7a and 7b.

It should be noted that the *proportion* and *amount* of separation hypotheses are not mutually exclusive. If young children show facilitation across the experiments, this would only indicate that their ability to benefit from performance feedback is dependent on both the ratio and the amount of distance between their performance representations.

1.3.2 Different Sources of Information About Task Success

Both of the hypotheses described above make assumptions about success in the base-level task, with the underlying assumption that the goal is to optimize task success. However, there are at least two ways in which participants can optimize task success. First, participants may aim to *maximize performance*. In other words, the primary goal may be to maximize the instances of positive feedback (corresponding to points/stickers) acquired in the game. If this were the case, children should select the game that leads to the most positive feedback, an external signal of performance, regardless of the ease of the two tasks. Prior work suggests that adult's allocation of study is influenced by the reward structure of the task, in that they choose to select more highly rewarded items (i.e., those worth more points on an upcoming test) regardless of item difficulty (Ariel, Dunlosky, & Bailey, 2009). Sixth-graders also demonstrate a preference for studying high reward items (Li, Ji, Li, Li, Zhang, & Li, 2016) and 9-year-olds prioritize study of difficult items when accuracy is emphasized. The role of reward structure on the task selections of young children (e.g., 5-year-olds), however, remains uninvestigated.

Alternatively, participants may aim to *minimize effort*, or expend the least amount of effort possible to complete the task. If this were the case, participants should select the easier game regardless of the feedback they receive, requiring participants to monitor a kind of internal feedback regarding the amount of effort expended in each task. Previous work has shown that adults tend to avoid cognitive demand, and prefer tasks that require less cognitive effort (Kool, McGuire, Rosen, & Botvinick, 2010; O'Leary & Sloutsky, 2016; O'Leary & Sloutsky, submitted). In addition, adults often select items for re-study based on the difficulty of learning those items (Metcalfe, 2002; Metcalfe & Kornell, 2005), indicating that they spontaneously take internal signals into account, and make strategic decisions on the basis of this information. Young children, however, have shown little to no propensity to minimize effort, in that they choose indiscriminately when selecting between tasks of different difficulty levels (O'Leary & Sloutsky, 2016) and allocate study time equally across easy and difficult items (Dufresne & Kobasigawa, 1989). One goal of the study reported in Chapter 5 was to directly test what kinds of information (i.e., performance feedback or effort) young children rely on to monitor and control behavior.

1.4 Transfer of Metacognitive Strategy Use in Young Children

As described above, scaffolding can help young children develop and apply a strategy rule to improve their performance in a task (O'Leary & Sloutsky, 2016; O'Leary & Sloutsky, submitted). However, the nature of the strategy that young children form has yet to be investigated. It is possible that young children form a strategy that is specific to the current task at hand (e.g., to select the *blue* game). Alternatively, they may form a strategy that is easily transferred to a novel task (e.g., to select the *easy* game).

Previous work suggests that people should be more likely to generalize strategies employed by means of a metacognitive mechanism than those prompted by use of an associative mechanism. For example, older children with greater metacognitive awareness of a strategy were more likely to employ that strategy in novel situations

(Siegler & Jenkins, 1989). However, the majority of these studies have focused on older children or adults. Very little is known about metacognitive strategy transfer in young children.

Investigating strategy transfer can aid in understanding whether young children's strategies are sufficiently abstract to generalize across tasks. Importantly, if children are able to transfer a strategy, this will suggest that the strategy they form is not tied to the specific stimuli in the task in which the strategy was learned. In particular, if children can be trained to identify and select a task set of a certain difficulty level (regardless of the superficial task features), this holds implications for training young children's strategy use in a number of domains. For example, this kind of training could transfer to benefit young children's ability to identify a more difficult set for study, to recognize that a task is too difficult and that they should seek help, or to optimize timed test performance by identifying and completing easier items first. The goal of the study reported in Chapter 6 is to take a first step to address these issues, by assessing whether young children transfer a strategy across similar tasks.

Chapter 2: Overview of the Experiments

The goal of this dissertation is to examine the development of metacognitive monitoring and control by focusing on four issues: (1) the malleability of metacognition in young children, (2) independence of metacognitive monitoring and control, (3) the role of task success representations in metacognition, and (4) transfer of metacognitive strategy use in young children. To address these issues, eight experiments were conducted.

Experiment 1 in Chapter 3 examined the developmental trajectory of metacognitive monitoring and control using a numerical discrimination task as the "base-level" task (i.e., the task that gives rise to meta-level representations). Five-year-olds, 7-year-olds, and adults were presented with a numerical discrimination task in two levels of difficulty (i.e., easy and difficult; Halberda & Feigenson, 2008), which mapped onto a particular color (e.g., easy discriminations may be presented in blue and difficult discriminations may be presented in red). Importantly, in Experiment 1 participants were not told which game was easier or that the two games differed in difficulty at all. Because participants were incentivized to perform as accurately as possible, the proportion of trials on which participants selected the easier game served as the measure of metacognitive control. Participants' performance and task monitoring were also assessed at the end of the experiment. Participants were asked to estimate the proportion of their correct responses in the task (performance monitoring), which were then compared to their actual performance to assess the accuracy of their estimations. Furthermore, participants were asked to indicate (a) whether they noticed the tasks' differential difficulty and (b) which task was easier (task monitoring). Late onset theories of metacognition would predict that even 5-year-olds, and potentially 7-year-olds would perform poorly, especially in terms of metacognitive control. In contrast, early onset theories would predict that even 5-year-olds would be proficient in terms of both metacognitive monitoring and control.

Provided that the developmental trajectory observed in Experiment 1 was consistent with theories describing a late onset, Experiment 2 was conducted to test whether metacognition is malleable in young children. In particular, participants were provided with performance feedback (to improve monitoring), strategy instruction (to improve control), or both. Facilitative effects of feedback and/or strategy instruction would provide a conclusive explanation for why some studies of metacognition, which have provided these kinds of scaffolding, have demonstrated an earlier onset of metacognitive proficiency.

In Chapter 4, we replicated the findings from Chapter 3 (Experiment 3) and extended the findings to include 7-year-olds and adults. In particular, the issue of whether metacognitive monitoring and control operate independently was directly addressed. One set of theories assumes a unidirectional, feed-forward relation between monitoring and control, such that proficient monitoring is a prerequisite for proficient control (MC account; e.g., Nelson & Narens, 1990; Son, 2004). Another set of theories proposed an opposite unidirectional effect, such that proficient control underlies and leads to proficient monitoring (CM account; Koriat & Ackerman, 2010; Koriat, Ackerman, Adiv, Lockl, & Schneider, 2014). It is also possible that monitoring and control can function independently. In Experiments 4 and 5, these competing accounts were tested by providing 5-year-olds, 7-year-olds, and adults with strategy instruction and performance feedback, respectively, to assess their effects on metacognitive monitoring and control.

Experiments 6a-7b in Chapter 5 aimed to address the finding that feedback had influenced young children's monitoring, but not control, performance (e.g., in Experiments 2 and 5). In particular, it was hypothesized that uncertainty in young children's representations of task success prevented them from using this information to adjust their behavior. In addition, young children may rely on different kinds of information about task success (e.g., performance feedback vs. internal signals of effort) to monitor and control their behavior. In Experiment 6a, young children received feedback tied to their task selection, rather than their actual performance, to assess whether their performance improved relative to when veridical performance feedback was given. In Experiment 6b, differences in difficulty were eliminated, so children could only detect the differences between the two tasks by relying on the differential feedback.

In Experiment 7a, participants received additional exposure to the base-level task (with feedback) to evaluate whether additional task experience improved metacognition by making representations of task success more precise. In Experiment 7b, performance feedback was removed, meaning that participants could only rely on the difference in required effort to detect and/or act upon any differences in the two tasks. Overall, Chapter 5 addresses whether the separation in task success representations influences metacognitive monitoring and control, and whether young children rely on different types of information to monitor and control behavior.

Finally, Experiment 8 in Chapter 4 was conducted to test the extent to which strategy learning in the aforementioned tasks can transfer to novel situations. In particular, a pre-/post-test design was used to assess whether successful training to use a strategy in one base-level task (i.e., numerical discrimination) transfers to a similar task (i.e., line length discrimination). One possibility is that the strategies formed and employed by participants in the previous studies were based on task-specific information (i.e., to select the blue game), rather than an abstract strategy rule that can be applied across tasks (i.e., to select the easy game). Evidence for transfer would suggest that even young children are able to form and use an abstract strategy across differing task contexts, which holds implications for training young children's metacognition.

Chapter 3: Malleability of Metacognitive Monitoring and Control In Young Children

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 an 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. This age range was chosen 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). An adult sample was included to identify components of metacognition that change between childhood and adulthood.

To address the second aim, Experiment 2 was conducted 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, of interest was *how* these effects transpire. For example, if feedback improves children's metacognitive monitoring, but not their control, this would provide some 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. These predictions are further discussed in the introduction of Experiment 2.

It was predicted that performance feedback would improve children's metacognitive *monitoring* by providing an external cue to their performance. It was 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, it was predicted that improvements in both monitoring and control would transpire when children are provided with both feedback *and* strategy instruction.

EXPERIMENT 1

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. Five-year-olds were recruited through local daycares and preschools in Columbus, Ohio. Seven-year-olds were recruited through local elementary schools. 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, participants' performance and task monitoring were assessed. To evaluate *performance monitoring*, participants were asked 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 made it possible 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, participants were asked 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 gender on any of our measures (all ps > .08). There was no effect of gender 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 the questions of interest, so the data were collapsed across gender 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 *t*s > 8.8, *p*s < .001, *d*s > 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 *t*s > 3.25, *p*s < .005, *d*s > 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. Metacognitive control was assessed by examining 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 (*t*s < 1, *p*s > .6, *d*s < .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).

The proportion of *individuals* who systematically chose the easy task was also assessed. 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. Participants were asked 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, a composite score was calculated 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 the task difficulty (see Figure 4). 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, $p = .42$.

The proportion of individuals who answered all 3 questions correctly (i.e., those who showed highly proficient monitoring) were identified. 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, X^2 (1, N = 90) = 22.81, $p < .001$ (see Figure 4). 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 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. Participants were asked 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 their performance was higher 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 (see Figure 5). Figure 6 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 6). 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, X^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 the 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, Experiment 1 demonstrated 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. This difference may have transpired, at least in part, due to the differences in the tasks described in the introduction. Experiment 2 investigates 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. Performance feedback may also provide external cues about one's performance and may have a similar effect on children's monitoring (Butler & Winne, 1995).

On the basis of these considerations, it was 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. It was 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.

Experiment 2 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. Only 5-year-olds were included in this experiment to be able to make more direct comparisons to studies reporting early metacognitive proficiency.

Method

Participants

Ninety 5-year-olds participated in this experiment: 30 in the Feedback Only condition (12 girls, *M* = 5.37 years, *SD* = 0.23 years), 30 in the Instruction Only condition (11 girls, *M* = 5.23 years, *SD* = 0.17 years), and 30 in the Feedback + Instruction condition (13 girls, *M* = 5.42 years, *SD* = 0.25 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, data from the 5-year-olds in both Experiments 1 and 2 were used. This provided 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 made it possible 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 (see Figure 7). 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, $F(1, 116) = 0.62$, $p = .43$, $\eta^2 = .01$. 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%, $t(29) = 2.87$, $p < .01$, $d = 1.07$).

The proportion of optimizers (i.e., individual children who chose the easier task on at least 20 trials) in each condition was also assessed (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 $(1, N = 120) = 0.96$, $p = .327$. 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, a composite task monitoring score was computed. 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 (see Figure 8). 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 *p*s > .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. Relative performance monitoring -- the proportion of children who rated their performance as higher on easy relative to difficult trials – was also examined. The proportion of correct responders did not differ as a function of feedback or additional instruction, both *p*s > .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.

Chapter 4: Independence of Metacognitive Monitoring and Control

Both the MC (Monitoring → Control) and CM (Control → Monitoring) models of metacognition discussed in the introduction predict that improvements in the first component (e.g., monitoring under the MC account) should lead to improvements in the second (e.g., control under the MC account). One way to directly investigate this possibility is to scaffold each component, and observe corresponding improvements in the other. For example, if a manipulation (e.g., feedback) improves monitoring, but not control, this would provide evidence against the CM model which states that control improvements should precede and give rise to monitoring improvements. In addition, if another manipulation (e.g., providing a strategy) improves metacognitive control, but not monitoring, this would provide evidence against the MC model which states that monitoring improvements should precede control improvements. This scaffolding method has been used to investigate component independence in 5-year-old children (O'Leary & Sloutsky, 2016), providing initial support for the independence model. These initial findings suggested that a systematic developmental investigation was needed to reach more definitive conclusions. Perhaps monitoring and control operate independently in young childhood, but become more coupled with experience and development.

Experiment 3 examined the development of metacognitive monitoring and control (with the goal of replicating prior findings) by measuring these abilities in 5-year-olds, 7-

year-olds, and adults, in the absence of any scaffolding. In Experiments 4 and 5, a scaffolding approach was taken to investigate component independence, by providing participants with strategy instruction and performance feedback, respectively, and comparing performance to that in Experiment 3. By observing the effects of scaffolding, one can directly assess whether the components are interdependent or independent and whether this independence/interdependence changes with development.

EXPERIMENT 3

Method

Participants

A sample of 5-year-olds (*N* = 30, 18 females, *M* = 5.35 years, *SD* = 0.25 years), 7-year-olds (*N* = 30, 12 females, *M* = 7.41 years, *SD* = 0.28 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 in this chapter, 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, Design, and Procedure

The stimuli, design, and procedure were identical to that of Experiment 1 in Chapter 3. As in Experiment 1, we measured discrimination accuracy, metacognitive control (i.e., the proportion of easy task choices made by each participant), "absolute" performance monitoring (i.e., estimates of overall performance compared to actual performance), and "relative" performance monitoring (i.e., estimates of performance in the easy game relative to that in the difficult game).

Results and Discussion

Discrimination Accuracy

To verify that the two games were differentially difficult, participants' performance in the easy and difficult games were compared before proceeding with the main analyses. Indeed, participants of all age groups were more accurate in the easy game than the difficult game, all ts > 4.7, ps, < .001, ds > 1.1 (see Table 3).

Metacognitive Control

The proportion of participants' easy task choices was used as a measure of metacognitive control. As predicted, adults chose the easy game more often than would be expected by chance (M = 76%), $t(29) = 6.41$, $p < .001$, $d = 2.38$. 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%), $p = .48$ (see Figure 11). 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, $F(2, 87) = 14.4$, $p < .001$, $\eta^2 = .25$. Five- and 7-year-olds did not significantly differ, $p = .14$ (see Table 3).

To assess individual differences in the task, participants were classified as "optimizers" if they systematically chose the easy task. To determine this, a moving window of 12 trials (across the 30 trials in the task) was used 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). Twenty adults (66% of the sample) consistently chose the easy task, as did 8 7-year-olds (27% of the sample) and 2 5-year-olds (7% of the sample; see Table 1). These proportions differed across the age groups, X^2 (2, N = 90) = 25.20, $p < .001$. Post-hoc comparisons revealed that more adults optimized than 7-year-olds, X^2 (1, N = 60) = 9.64, $p < .005$, and 5-year-olds, X^2 (1, N = 60) = 23.25, $p < .001$. In addition, more 7-year-olds optimized than 5-year-olds, X^2 (1, N = 60) = 4.32, $p < .05$. Although 5- and 7-year-olds did not differ in their overall proportion of easy task choices, there were a greater proportion of 7-year-olds than 5-year-olds who strategically chose the easy game. Taken together, these findings suggest some development of metacognitive control between 5- and 7-years-of-age, and clear evidence of development between 7-years-of-age and adulthood.

Since both adults and 7-year-olds chose the easy task more than would be expected by chance, we calculated backward learning curves to identify whether participants learned and adjusted their strategy gradually or abruptly (Hayes, 1953). To construct the backward learning curves, the first trial at which each participant began to systematically choose the easier game was identified. To do this, a moving window of 12 trials was used again to identify the earliest window at which each participant chose the easier game on at least 11 of those trials (p = .052, according to binomial probability). The first trial of this window was designated as Trial 0 (T_0). Identifying this trial made it possible to assess the rate of optimization by aligning participants along the trial at which they began showing systematic metacognitive control. Participants' performance was then analyzed in blocks preceding and subsequent to T_0 (with 5 trials per block; see Figure 2).

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 participants discovered and applied a strategy rule in an all-or-nothing fashion (see Rehder & Hoffman, 2005, for related arguments). In other words, this pattern would suggest that once participants discovered which game would lead to optimal performance, they abruptly adjusted their strategy to choose that game consistently. This would also result in a shallow slope following optimization.

In contrast, comparable slopes before, at, *and* after $Block_0$ would indicate that control took place via gradual associative learning (which is perhaps more "implicit" than rule discovery) rather than all-or-nothing rule or strategy discovery. This pattern of data would suggest that, over time, participants learned the associations between the color of the game and its corresponding outcome. As they learned that one game was more likely to result in correct responses (or in greater confidence), they began to choose that option more often. As such, they would choose that option increasingly often, but would achieve this via a gradual increase in easy task choices, rather than an all-or-nothing switch in strategy.

For each participant, three slopes were calculated to indicate 1) learning before optimality was achieved (the slope between B_{-2} and B_{-1}); 2) learning at T_0 (the slope between B_{-1} and B_0); and 3) 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 participants did not have trials in some blocks (primarily blocks B_{-2} and B_{-1}). Because these blocks represented the very beginning of the task for these individuals (i.e., there were no blocks before they showed optimal behavior), we assumed 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 comparisons across age groups and conditions.

Adults who optimized (*N* = 20) in Experiment 3 demonstrated no evidence of learning, 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, $t(13) = .90$, $p = .385$. At the same time, there was a steep slope at T_0 (B_{-1} to B_0; slope = .39, $t(19) = 10.35$, $p < .001$; see Figure 9), suggesting that adults' learning was rather abrupt (and thus more consistent with rule or strategy discovery than with more gradual associative learning). Overall, this pattern of data suggests that adults did not learn to control behavior through slowly acquired associations, but instead through applying a strategy rule (i.e., selecting the easier game because it results in greater accuracy).

The 7-year-olds who optimized (*N* = 8) displayed a very similar pattern of results. Seven-year-olds only showed substantial improvement at T_0 (B_{-1} to B_0; slope = .45, $t(7) = 6.18$, $p < .001$), indicating that they, too, optimized behavior by applying an adaptive strategy. Slopes before, and following, T_0 did not differ from zero (both *p*s > .18). In sum, adult and 7-year-old optimizers showed a near identical pattern of findings, in that they showed no evidence of learning prior to discovering the strategy, followed by a steep slope when the strategy was discovered and initially applied. In addition, there was no evidence of additional learning following the initial strategy application, suggesting that once the strategy was discovered, participants in both age groups applied it consistently through the remainder of the task.

Performance Monitoring

To measure *absolute performance monitoring*, we compared each participant's performance estimates to their actual performance. To do this, each participant's actual performance was subtracted from their estimated performance, after which the absolute value was taken and adjusted to account for our use of a discrete scale (as in Experiment 1). Here, scores further away from 0 indicated greater misestimation. Even adults significantly misestimated their performance (M = -.14, above 0), $t(29) = 3.38$, $p < .005$, $d = 1.26$, demonstrating a strong tendency to underestimate (as demonstrated by their unadjusted estimated-actual performance scores; see Figure 3). Five-year-olds showed the opposite pattern, in that they significantly overestimated their performance (M = .20, above 0; see Figure 3), $t(29) = 7.50$, $p < .001$, $d = 2.79$. Seven-year-olds also misestimated performance (M = .05, above 0), $t(29) = 2.35$, $p < .05$, $d = .87$, 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 10).

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) was also assessed. Four 5-year-olds (13% of the sample), 9 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, X^2

(2, $N = 90) = 29.83$, $p < .001$. To assess the source(s) of the effect, post-hoc comparisons were performed among the age groups. These analyses revealed that more adults monitored relative performance than either 5-year-olds, X^2 (1, $N = 60) = 26.79$, $p < .001$, or 7-year-olds, X^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. Finally, the current data suggest that absolute and relative performance monitoring may undergo asynchronous development, in that adults showed more accurate *relative* performance monitoring than 7-year-olds, but 7-year-olds demonstrated less biased *absolute* performance monitoring than adults.

Summary of Findings

Experiment 3 replicated and extended the findings from Experiment 1. 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. When it came to absolute performance monitoring, however, 7-year-olds were more accurate than both 5-year-olds and adults. Importantly, this difference reflected a transition between performance overestimation in early childhood, to performance underestimation in adulthood. Finally, relative performance monitoring showed some improvement between 5- and 7-years of age, as well as substantial improvement between childhood and adulthood.

The process through which participants learned to engage in metacognitive control, and systematically select an easier task, was also measured. Using backward learning curves, it was revealed that both 7-year-olds and adults adjusted behavior by discovering and applying a rule or a strategy, rather than gradually learning associations between the color of the game and their performance outcome. Specifically, there was no evidence of learning prior to the initial discovery and application of the strategy. In addition, once the strategy was discovered it was applied abruptly (in that there was a steep spike in easy task choices) and consistently (this rate remained stable for the remainder of the task). These findings highlight the process of optimization when engaged spontaneously, in the absence of any cue or instruction to perform optimally.

In Experiment 4, a scaffolding approach was taken to more directly investigate component independence, and participants were supplied with a strategy that would optimize their performance (i.e., to choose the easier game). Previous work has shown that providing such a strategy can improve the metacognitive control of 5-year-olds, but the effect of this instruction on the metacognitive performance of 7-year-olds and adults remain unknown. If improvements in control transpire in the absence of improvements in monitoring, this would provide evidence against the MC (Monitoring → Control) model of metacognition, which suggests that improvements in monitoring should underlie improvements in control.

EXPERIMENT 4

Method

Participants

Thirty 5-year-olds (N = 30, 9 females, M = 5.25 years, SD = 0.22 years), 30 7-year-olds (N = 30, 21 females, M = 7.52 years, SD = 0.29 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 experiments.

Materials, Design, and Procedure

The stimuli and procedure were similar to that of Experiment 3, with one exception. In this experiment, participants were told at the beginning of the task that the two games differed in difficulty, and were instructed to select the easier game on each trial (prior to the choice opportunity). Crucially, participants were not told which of the two games was easier.

Results and Discussion

Discrimination Accuracy

As in Experiment 3, participants of all ages were more accurate in the easy game than the difficult game, all ts > 3.23, ps, < .01, ds > 1.23 (see Table 4).

Metacognitive Control

Similar to Experiment 3, 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 significantly outperformed 7-year-olds ($p < .001$), who significantly outperformed 5-year-olds ($p < .005$), according to post-hoc LSD comparisons (see Figure 11).

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, X^2 (2, N = 90) = 33.69, $p < .001$. Post-hoc comparisons showed that there were more adult optimizers (M = 93%) than 7-year-old optimizers (M = 63%), X^2 (1, N = 60) = 7.95, $p < .01$, and more 7-year-olds who optimized than 5-year-olds (M = 20%), X^2 (1, N = 60) = 11.59, $p < .005$ (see Figure 12).

Backward learning curves were again used to assess the rate at which 7-year-olds and adults controlled behavior in Experiment 4 (see Figure 13). Similar to Experiment 3, adults 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, $p < .001$; See Figure 13), with no other slopes reaching non-zero significance. Therefore, similar to Experiment 3, both adults and 7-year-olds exhibited abrupt rather than gradual change, which is more consistent with strategy change or rule discovery than with associative learning. Taken together, these findings suggest that the process of metacognitive control occurred similarly when participants were provided with a strategy (in Experiment 4) and when they had to formulate one spontaneously (in Experiment 3).

Performance Monitoring

Participants in all age groups misestimated performance (with adjusted performance monitoring scores greater than 0; all *p*s $< .001$). There were no age differences in participants' absolute performance monitoring across the three age groups in Experiment 4 ($p = .61$). At the same time, relative performance monitoring did differ by age, X^2 (2, $N = 90$) = 9.00, $p < .05$. Post-hoc comparisons revealed that more adults (M = 67%) successfully monitored their relative performance than did 7-year-olds (M = 33%) and 5-year-olds (M = 33%; both *p*s $< .05$). There was no difference between the proportions of 7-year-olds and 5-year-olds that successfully monitored their relative performance ($p = 1.00$).

Cross-Experiment Comparisons

Metacognitive Control. To directly assess effects of instruction on performance relative to Experiment 3, we conducted a 2 (Instruction: No Instruction vs. Instruction) x 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA using data from both Experiment 3 (where no instruction was provided) and Experiment 4 (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 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 ($p < .005$). There was also a main effect of instruction, $F(1, 174) = 17.64$, $p < .001$, $\eta^2 = .09$, in that participants in Experiment 4 ($M = .74$) outperformed those in Experiment 3 ($M = .62$). The interaction was not significant ($p = .25$).

In line with this finding, assessing the proportion of optimizers in Experiments 3 and 4 revealed an effect of age, X^2 (2, N = 180) = 53.70, $p < .001$, in that more adults (93% of the sample) optimized than 7-year-olds (63% of the sample, $p < .001$) and 5-year-olds (20% of the sample, $p < .005$), and more 7-year-olds optimized than 5-year-olds ($p < .001$). There was also an effect of instruction, X^2 (2, N = 180) = 11.83, $p < .005$, suggesting that more participants optimized when provided with an adaptive strategy (see Figure 5).

Performance Monitoring. A 2 (Instruction: No Instruction vs. Instruction) x 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA was conducted on participants' absolute performance monitoring scores in Experiments 3 and 4. 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 (see Figure 14).

The proportion of participants who exhibited accurate relative performance monitoring (i.e., correctly detected that they had been more accurate in the easy game) was also assessed. Relative performance monitoring differed significantly across the age groups in Experiments 3 and 4, X^2 (2, N = 180) = 35.18, $p < .001$, in that more adults 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 that monitored relative performance did not differ ($p = .31$). There was no overall effect of instruction on participants' relative performance monitoring, $p = .65$ (see Figure 16).

Summary of Findings

In Experiment 4, it was found that providing a strategy facilitated metacognitive control across the age groups, resulting in more systematic selection of the easier game. Similar to Experiment 3, 7-year-olds' and adults' optimization was more consistent with abrupt strategy change than with associative learning (as evidenced by the profile of the backward learning curves).

Although providing a strategy clearly facilitated control across the age groups, there were no improvements in absolute or relative performance monitoring. In other words, increases in performance estimation accuracy could not have given rise to the increases in easy task choices. This finding provides evidence against the MC model of metacognition.

To address the possibility of the CM model, participants were provided with performance feedback in Experiment 5. Previous work has shown that 5-year-olds' monitoring benefited from feedback (O'Leary & Sloutsky, 2016), likely by providing an external signal of performance in the task. As suggested by the findings of Experiments 1 and 3, 5-year-olds have difficulty estimating performance on the basis of self-generated error signals, which may lead them to rely on 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 will further support the hypothesis that the two processes can operate independently.

In addition to providing an external signal of performance, feedback makes it perhaps more likely that participants will rely on associative learning to learn and 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 3 and 4.

EXPERIMENT 5

Method

Participants

Thirty 5-year-olds (11 females, *M* = 5.41 years, *SD* = .28 years), 30 7-year-olds (14 females, *M* = 7.50 years, *SD* = .32 years), and 30 undergraduate students (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 that of Experiment 3, 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 500ms. If they answered incorrectly, they saw a sad face and heard a low tone for 500ms.

Results and Discussion

Discrimination Accuracy

As in Experiment 3, participants of all ages were more accurate in the easy game than in the difficult game, all *t*s > 5.57, *p*s, < .001, *d*s > 1.51 (see Table 5).

Metacognitive Control

In Experiment 5, only adults chose the easy task more than would be expected by chance, 85%, $t(59) = 12.43$, $p < .001$, $d = 3.24$, whereas 7-year-olds ($M = 53\%$) and 5-year-olds ($M = 54\%$) were not different from chance (both *p*s > .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 proportion of participants who optimized by systematically selecting the easier game. A chi-square analysis revealed an effect of age, X^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, backward learning curves were calculated for the adults who systematically chose the easier game (see Figure 17). As in previous experiments, the slope at T_0 (B_{-1} to B_0) was steep (slope = .39, $p > .001$), suggesting abrupt strategy discovery. However, in contrast to Experiments 3-4, there was a small, yet non-zero slope after T_0 (B_0 to B_4; slope = .07, above 0, $p < .05$). Though this slope was non-zero, 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 3-4.

Performance Monitoring. As in previous experiments, participants of all age groups misestimated performance (with performance monitoring scores above 0; all *p*s < .01). 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 this was due to the fact that both adults ($M = .06$) and 7-year-olds ($M = .04$) more accurately estimated their performance than 5-year-olds ($M = .11$; both *p*s < .05). A chi-square analysis also revealed an effect of age on the proportion of participants who correctly monitored their performance in the easy game relative to the difficult game, X^2 $(2, N = 90) = 9.63$, $p < .01$. Post-hoc comparisons revealed that adults outperformed 7-year-olds ($p < .05$), and 5-year-olds ($p < .005$). However, 5- and 7-year-olds' performance did not differ ($p = .61$).

Cross-Experiment Comparisons

Metacognitive Control. As in Experiment 4, we conducted a 2 (Feedback: No Feedback vs. Feedback) by 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA on the proportion of easy task choices selected in Experiments 3 (where no feedback was provided) and 5 (where feedback was provided). 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 ($p < .001$). 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 ($p = .09$, $\eta^2 = .03$; see Figure 11).

The pattern described above also held for the proportion of optimizers, in that there was no effect of feedback (i.e., comparing Experiments 3 and 5), $p = .88$. The proportion of optimizers (see Figure 5) did differ by age (collapsed across Experiments 3 and 5), X^2 (2, $N = 180$) = 58.67, $p < .001$, in that more adults optimized than 5-year-olds ($p < .001$) or 7-year-olds ($p < .001$), between which there was no difference ($p = .13$; see Figure 12).

Performance Monitoring. We conducted a 2 (Feedback: No Feedback vs. Feedback) x 3 (Age: 5-year-olds vs. 7-year-olds vs. Adults) ANOVA on participants' absolute performance monitoring scores in Experiments 3 and 5. 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 5 ($M = .07$) than in Experiment 3 ($M = .13$). 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 (see Figure 14).

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, X^2 (1, $N = 180$) = 7.20, $p < .01$. There was also a significant overall effect of age, X^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 ($p = .18$; see Figure 16).

Summary of Findings

In Experiment 4, 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, the opposite tendency was observed in Experiment 5. In this experiment, participants' performance monitoring improved with performance feedback, whereas control performance was unaffected. This finding provides some evidence against the CM model, in that increased precision of performance monitoring did not rely on changes in control performance. This held true for both absolute and relative performance monitoring, suggesting that increases in control performance were not necessary for either of the monitoring components to improve.

How did adults learn to control their behavior in Experiment 5? 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). Backward learning curves indicated, however, that feedback did not encourage associative learning as a means to achieve metacognitive control. Instead, as in the previous experiments, adults continued to control behavior by means of abrupt strategy rule application.

Chapter 5: Effects of Feedback and Task Experience on Metacognition

The current study had two primary aims. The first aim was to identify whether young children's metacognitive monitoring and control rely on the separation of their representations of success in the two tasks. The second aim was to assess what kinds of information young children use to monitor and control behavior: those based on external signals of performance (i.e., feedback) or internal signals of effort. Achieving this aim will reveal whether young children's monitoring and control can be improved by increasing the separation between representations of performance and/or effort.

In Experiment 6a, children were provided with feedback having a large success to failure ratio of separation between the two games. Instead of receiving feedback reflecting their actual performance (i.e., receiving positive feedback approximately 60% of the time in the difficult game and 90% of the time in the easy game), participants received positive feedback 0% of the time in the difficult game (regardless of their actual performance) and 100% of the time in the easy game (also regardless of their actual performance). If young children benefit from the larger separation ratio, their metacognitive performance should improve relative to that with veridical (i.e., 60/90 percent) performance feedback (O'Leary & Sloutsky, 2016). In Experiment 6b, it was assessed whether increases in performance are due to the feedback per se, or due to the feedback potentially providing a *cue* about the amount of effort required in the two games. To test this possibility, differences in objective difficulty between the two tasks were eliminated, meaning that children could only rely on the feedback to differentiate between the two games. Thus, any differences in task selection or monitoring between Experiments 6a and 6b could be attributed to task effort.

In Experiment 7a, participants received more experience in the two tasks, to increase the amount of separation (and, potentially, the precision of their representations of success across the tasks). To do this, veridical (i.e., 60/90 percent, on average) feedback was used, but children were provided with more trials of the easy and difficult games. Improvements in performance relative to the standard veridical feedback condition (with only 30 trials; i.e., the Feedback only condition from Experiment 2) would suggest that task experience reduced the noise in participants' task success representations. In Experiment 7b, no performance feedback was provided, to test whether this effect was due to more precise representations of internal or external signals of performance. Differences in performance relative to Experiment 7a would provide an estimate of children's reliance on external signals of performance.

EXPERIMENT 6a

Method

Participants

Thirty 5-year-olds (10 females; $M = 5.25$ years, $SD = 0.15$ years) participated in this experiment. Participants were recruited through local daycares and preschools in Columbus, Ohio.

Stimuli and Procedure

The stimuli and procedure were similar to those used in the "Feedback only" condition of Experiment 2. At the experiment's onset, children were incentivized to complete the task as accurately as possible. Each trial consisted of a choice opportunity, fixation, test stimulus, and response screen. During the choice opportunity, participants were allowed to choose that trial's difficulty by selecting between the two corresponding dot colors – our measure of metacognitive control. Participants were presented with two sets of dots for 500ms and were asked to identify which set had been larger. After each discrimination response, participants were provided with feedback. Unlike the "Feedback only" condition of Experiment 2, participants always received positive feedback (i.e., a smiley face and a high tone presented for 500ms) after each easy task response, and always received negative feedback (i.e., a sad face and a low tone presented for 500ms) after each difficult response.

Following the test trials, three questions were used to assess participants' *task monitoring*. First, participants were asked whether they thought the red game and the blue game were the same or different. Second, they were asked whether they thought one game was easier than the other. If they answered 'no' to this question, the experiment ended. If they answered 'yes,' they were asked a third question -- *which* game they thought was easier.

Results and Discussion

Discrimination Accuracy

As expected, participants were more accurate in the easy game (M = 82%) than the difficult game (M = 58%; $t(29) = 4.61$, $p < .001$, $d = .84$).

Metacognitive Control

Participants chose the easier game more than would be expected by chance (M = 67%, $t(29) = 5.12$, $p < .001$, $d = 1.90$). As in Experiment 1, individual children were categorized as "optimizers" if they selected the easier game on at least 20 of 30 trials ($p < .05$, according to binomial probability). Doing so revealed that 47% of participants chose the easier game systematically.

Task Monitoring

To evaluate performance monitoring, a composite score of children's responses to the following questions was calculated: 1) "was one game easier than the other?" and 2) "which game was easier?" Participants earned a point for each of the questions answered correctly, with a maximum of two points. However, if participants answered "no" to the first question, they were not offered the second question and received a score of 0. In Experiment 6a, participants' mean task monitoring score was 1.57.

Comparisons to Previous Data

To assess effects of the separated feedback, we compared performance in this experiment to that of participants in the "Feedback only" condition of Experiment 2 (heretofore referred to as the "feedback baseline;" O'Leary & Sloutsky, 2016), which has since been replicated (in Experiment 5; O'Leary & Sloutsky, submitted). There, participants received veridical feedback regarding their performance in the two games. The only difference between those experiments and the current experiment was that feedback was separated (i.e., 0% positive feedback in the difficult game an 100% positive feedback in the easy game) in the current experiment. Any improvements in performance here, relative to that in the feedback baseline, would indicate a facilitative effect of greater reward probability separation.

Metacognitive Control. Using an independent samples t-test, the proportion of easy task choices made in Experiment 6a were compared to that of the feedback baseline condition. Participants chose the easy game significantly more when provided with highly separated feedback (M = 66%) than in the feedback baseline (M = 52%; not different from chance, p = .42), $t(50.55) = -3.80$, d = 1.05 (see Figure 19A). This pattern was also reflected when considering the proportion of optimizers in the two data sets. A chi-square analysis revealed that significantly more children optimized by systematically choosing the easier task when provided with separated feedback (M = 46%) than in the feedback baseline condition (M = 7%), X^2 (1, N = 60) = 12.27, $p < .001$ (see Figure 19B).

Task Monitoring. Next, the task monitoring scores of participants in Experiment 6a were compared with those of participants in the feedback baseline condition. There was no significant difference in the task monitoring performance of these two groups (p = .11, d = .47), indicating that greater separation of success probability did not affect young children's ability to identify an easier task (see Figure 20).

In sum, separating the difference in reward probabilities from around 30% (in the feedback baseline) to 100% (thus, also increasing the ratio of separation) substantially improved 5-year-olds' metacognitive control, but not their metacognitive monitoring. This provides some evidence that young children can optimize their behavior on the basis of explicit performance feedback, given that the reward probabilities are sufficiently separated. In addition, children may rely more heavily on these external cues about performance to make task selection decisions, than they do to monitor the task difficulty. It is possible, however, that providing separated feedback helped children identify which of the two tasks had been more *effortful*. In other words, the external feedback may have provided a cue about internal processes, like the amount of response conflict present in the easy and difficult games.

To test this possibility, we eliminated any difference in the amount of effort required by the two games in Experiment 6b. To this end, children were provided only with discriminations in the more difficult ratios, regardless of which game they had chosen (red or blue). Thus, in Experiment 6b, the separated feedback provided the only cue to differentiate between the perceived successes in the two games. If children rely solely on feedback, an external signal of difficulty, performance should be identical in Experiments 6a and 6b. However, if children rely on more internal signals of effort (these were eliminated in Experiment 6b), proportion of easy task choices should decrease compared to Experiment 6a.

EXPERIMENT 6B

Method

Participants

Thirty 5-year-olds (13 females; M = 5.23 years, SD = .19 years) participated in this experiment, none of which participated in any of the previous experiments. Participants were recruited through local daycares and preschools in Columbus, Ohio.

Materials, Design, and Procedure

The stimuli and procedure were similar to that of Experiment 6a, with one exception. In this experiment, all discrimination trials included dots displayed in the "difficult" ratios (9 vs. 10, 10 vs. 11, 11 vs. 12, 12 vs. 13, and 13 vs. 14). This made it impossible for participants to rely on differences in expended effort to engage in metacognitive monitoring and control. Thus, all they could rely on to optimize performance and appraise task difficulty was the "separated" feedback given after each discrimination response.

Results and Discussion

Discrimination Accuracy

As expected, participants' performance in the "easy" (M = 55%) and "difficult" (M = 57%) games did not differ, because all discriminations were difficult (p = .70). The feedback provided was the only difference between the two games in Experiment 6b. As such, the game in which performance was reinforced 100% of the time will heretofore be referred to as the easy game, and the game in which performance was reinforced 0% of the time will be referred to as the difficult game.

Metacognitive Control

As in Experiment 6a, the proportion of easy task choices were compared to chance performance (50%). Children in Experiment 6b chose the "easier" (i.e., more highly rewarded) game significantly more than would be expected by chance, M = 63%, $t(29) = 3.57$, $p < .005$, $d = 1.33$. In addition, 37% percent of the sample was categorized as optimizers (i.e., individuals who chose the easier game on at least 20 of 30 trials).

Task Monitoring

A composite score was calculated to represent participants' success at monitoring task difficulty. In the absence of differential effort in the two games, participants' composite task monitoring score was only .68 out of a possible 2 points.

Cross-Experiment Comparisons

Metacognitive Control. To assess the effect of internal signals of effort on participants' metacognitive control, performance in Experiment 6b was compared to that in Experiment 6a (see Figure 19A). Similarities in performance would indicate that children's control behavior was unaffected by eliminating the differences in effort between the two games. As such, this would suggest that young children rely on external signals of performance to optimize their behavior. Indeed, this is what was found, in that there was no significant difference between the proportions of easy task choices in the two experiments ($p = .44$).

This finding was further corroborated when we compared the proportions of optimizers in the two experiments. There was no significant difference between the proportions of optimizers ($p = .43$), indicating that young children relied on external, rather than internal, signals of performance to control behavior (see Figure 19B).

Task Monitoring. We also compared children's task monitoring scores in Experiments 6a ($M = 1.57$) and 6b ($M = .68$). Though there were no differences in control performance, there were significant differences in task monitoring ($t(49) = 4.73$, $p < .001$, $d = 1.32$), in that children's task monitoring was significantly impaired by the lack of differential effort in the two games (see Figure 20). In other words, children were much less likely to say that the two tasks differed in difficulty, or to identify the more highly reinforced game as the easier game, when the rate of feedback was the only thing that differed between the two games.

In sum, the present findings suggest that young children rely on external signals of performance, like explicit feedback, to optimize behavior by choosing an easier game. They even did so when there were no differences in the effort required to complete the two games, and the rate of reward was the only thing that differed between them. As such, young children may not be sensitive to differences in internal error signal when it comes to making behavioral adjustments to optimize performance. On the other hand, when children were asked to explicitly appraise the difficulty of the two tasks, they did so on the basis of the required effort, rather than performance feedback. As such, children were much less likely to say that the two games differed in difficulty, or that the more highly reinforced game was easier, in Experiment 6b. This suggests that children are sensitive to difference in internal effort, and use this when making explicit appraisals of task difficulty. However, children either cannot or do not use this information to control behavior, as evidenced by lack of a facilitative effect in Experiment 6a relative to Experiment 6b.

Contrary to previous work with smaller differences in the probability of success across the two tasks (in Experiments 2 and 5), both Experiments 6a and 6b indicated that young children could control behavior on the basis of external signals of performance. However, it remains unknown whether children are simply *unable* to optimize on the basis of less separated reward contingencies, or whether children can take advantage of them with more experience. Consider that, to track these reward probabilities, children must accumulate the positive and negative feedback given, in the red and the blue games, across many trials. With smaller separation in the reward probabilities, it may be difficult for children to recognize that one leads to higher overall reward than the other (i.e., there may be a kind of sampling error leading to perceived overlap of success in the two games). However, with more trials, this sampling error can be reduced, leading to more accurate representations of success in the two games. In Experiment 7a, children received 60 trials of task exposure with veridical performance feedback before they were allowed to select between the red and blue games. This allowed participants to observe the reward contingencies of the two games over many trials. If performance in Experiment 7a improves relative to that observed with fewer trials of veridical feedback in previous work (the "feedback only" condition of Experiment 2), this would indicate that children can capitalize on smaller differences in reward probabilities to engage in metacognition, albeit with a large enough "sample size."

EXPERIMENT 7A

Method

Participants

Thirty 5-year-olds (18 females; *M* = 5.38, *SD* = .26) participated in this experiment, none of whom participated in either of the previous experiments. Children were tested in local daycares and preschools in the Columbus, Ohio area.

Materials, Design, and Procedure

Stimuli were identical to those in the previous experiments. However, there were several crucial differences in the design and procedure of Experiment 7a. In this experiment, similar to previously reported studies and in contrast to Experiments 6a and 6b, participants received veridical feedback that reflected their accuracy in the task. In addition, participants received 60 “exposure” trials, in which they completed discrimination trials in the two levels of difficulty. These exposure trials included 30 trials of the easy (e.g., blue) game and 30 trials of the difficult (e.g., red) game, and the order of these trials was randomized. During the exposure phase, participants were not given a choice and completed 30 easy and 30 difficult trials. Following each discrimination decision, participants received feedback about the accuracy of each discrimination response. Following the exposure trials, participants completed 30 test trials in which they were required to choose between the blue and red games at the onset of each trial. Participants also received performance feedback on each trial during the test phase.

Results and Discussion

Discrimination Accuracy

Participants’ performance was higher in the easy (M = 90%) game than the difficult game (M = 52%; $t(27) = 9.68$, $p < .001$, $d = 1.83$).

Metacognitive Control

The proportion of easy task choices was compared to chance performance (50%) finding that children in Experiment 7a chose the easy task more than would be expected by chance, 64%, $t(29) = 3.90$, $p < .005$, $d = 1.45$. Individual participants were classified as optimizers if they chose the easy game on at least 20 out of 30 trials, identifying 30% of participants as optimizers.

Task Monitoring

Similar to the previous experiments reported here, we computed a task monitoring score to assess how well children noticed that the two tasks differed in difficulty. Children in Experiment 7a had an average task monitoring score of 1.6.

Comparisons to Previous Data

To assess the effects of exposure on children's performance, we compared the performance of children in Experiment 7a to those collected in the "feedback baseline" condition of Experiment 2. The only difference in the procedure given to these two samples was that children in Experiment 7a received 60 trials of exposure to the task, with the test phase of Experiment 7a being identical to the task presented to participants in the feedback baseline.

Metacognitive Control. Relative to children in the feedback baseline condition ($M = 52\%$), children in Experiment 7a chose the easier game significantly more often, $M = 64\%$, $t(58) = -2.88$, $p < .01$, $d = .76$ (see Figure 21A). In addition, a higher proportion of children optimized performance by selecting an easier game in Experiment 7a, $M = 30\%$, $X^2 (1, N = 60) = 5.46$, $p < .05$, than in the same task, without pre-exposure, $M = 7\%$ (see Figure 21B). Both of these findings indicate that children more readily controlled their behavior when they had more experience receiving feedback in the two tasks.

Task Monitoring. Children in Experiment 7a also marginally outperformed children in the feedback baseline condition in terms of task monitoring, $t(50) = -1.87$, $p = .07$, $d = .53$. Children were more likely to notice the difference in difficulty and identify which of the two games was easier when they had more experience with the two games (Figure 22).

In Experiment 7a, participants received feedback after every trial, in both the exposure and test phases. This provided more experience with the reward contingencies in the two games, and may have made it easier for children to notice and act on the differences in the two games on the basis of that feedback. It is also possible, however, that the exposure trials helped children notice the difference in *effort* expended in the two games, or that feedback merely served to help participants pick up on the difference in effort. In Experiment 7b, this possibility was directly assessed by providing participants with increased experience with the task, in the absence of performance feedback. If performance is lower than that of Experiment 7a, this would indicate that children rely primarily on external signals of performance (i.e., feedback) and that they have difficulty acting on internal signals of effort. If performance is similar to that in Experiment 7a, that would indicate that children can rely on the internal signals of effort to monitor and/or make decisions about the two games.

EXPERIMENT 7B

Method

Participants

Thirty 5-year-olds (15 females; M = 5.28, SD = .22) participated in this experiment. None of the participants participated in any of the previous experiments. Children were recruited from local preschools and daycares in the Columbus, Ohio area.

Materials, Design, and Procedure

The stimuli, design, and procedure were similar to that of Experiment 7a, with one exception. No feedback was provided in either the exposure phase or the test phase.

Results and Discussion

Discrimination Accuracy

As expected, participants accuracy was higher in the easy game (M = 79%) than the difficult game (M = 58%; $t(27) = 4.95$, $p < .001$, $d = .94$).

Metacognitive Control

Unlike Experiment 7a, participants in Experiment 7b did not select the easier game more than would be expected by chance, M = 53%, $p = .45$. Only 20% of the sample was classified as optimizers (i.e., selected the easier game on at least 20 of 30 trials).

Task Monitoring

Composite task monitoring scores were calculated for participants in Experiment 7b. The average task monitoring score was 1.43 for this sample of 5-year-olds.

Cross-Experiment Comparisons

Control and monitoring indicators in Experiment 7b were compared to those in Experiment 7a, to assess the effect of feedback on children's metacognition. Poorer optimization and monitoring relative to Experiment 7a suggest a heavy reliance on performance feedback. Similar performance across the experiments would indicate that children can rely on internal signals of effort.

Metacognitive Control. Children made substantially fewer easy task choices in Experiment 7b, M = 53%, than Experiment 7a, M = 64%, $t(58) = 1.96$, $p = .05$, $d = .53$, indicating that they rely on feedback to optimize behavior, and return to chance performance in the absence of feedback (Figure 21A).

Task Monitoring. In contrast to measures of control, the task monitoring accuracy across Experiments 7a and 7b revealed no difference, $p = .43$ (see Figure 22). In other words, children successfully monitored the task difficulty whether or not they received feedback about their performance. In terms of task monitoring, it seems that providing additional experience with the task improved performance by reducing uncertainty about the amounts of *effort* required in the two games. Feedback did not facilitate performance over and above the effect of experience alone.

As in Experiments 6a and 6b, the results from Experiments 7a and 7b suggest that young children rely on external signals of performance (i.e., feedback) to successfully control behavior by selecting an easier task. However, they seem to rely on internal signals of effort to make appraisals about the difficulty of a task. This suggests that children rely on altogether different types of information to monitor and control behavior, supporting the idea that the two components of metacognition are dissociable.

Chapter 6: Transfer of Strategy Learning

Experiment 8 was designed to assess whether young children are able to learn and transfer a strategy to a novel task. A pre-/post-test design was used, between which some participants were trained to use a strategy (i.e., to select the easier task) through feedback and instruction, in a numerical discrimination task. Participants in the control condition completed a numerical discrimination task with no scaffolding in the training phase. Of interest was whether participants who were successfully trained to use the strategy would transfer this strategy to a line-length discrimination task with new stimuli (tested in the pre- and post-test phases). Successful transfer would provide evidence that young children can form a strategy that that is not dependent on the specific stimuli used during learning (i.e., to select the *blue* game), and which is sufficiently abstract to apply to a novel task (e.g., to select the *easier* game).

EXPERIMENT 8

Method

Participants

Thirty 5-year-olds (M = 5.24 years, SD = .24) participated in the control condition, and 51 children (M = 5.32 years, SD = .26) participated in the strategy training condition. We tested children in the strategy training condition until 30 children had learned from the strategy training, in that they selected the easier game on at least 24 of

32 trials in the training phase ($p < .05$, according to binomial probability). Non-learners were not dropped from the sample, to illustrate the effect of learning on strategy transfer.

Stimuli

Pre-/Post-Test Phase. Children were tested on the transfer task in the pre- and post-test phases of the experiment. In the transfer task, children could complete line length discriminations at two levels of difficulty: easy (lines presented in a 1:2 length ratio), or difficult (lines presented in a 9:10 length ratio, on average). These two levels of difficulty mapped onto a color (yellow or green) in which the lines were presented. This mapping was randomly assigned at the beginning of the task, and held in both the pre- and post-test phases.

Training Phase. The stimuli in the training phase were similar to those in the pre-/post-test phases. However, instead of line discriminations, children discriminated quantities of dots (Halberda & Feigenson, 2008; O'Leary & Sloutsky, 2016). Stimuli in the training condition were identical to those used in the Feedback + Instruction condition of Experiment 2. Stimuli in the control condition were identical to those used in Experiment 1. Stimuli in the retention phase were identical to those used in Experiment 1.

Procedure

Overall, the procedure was similar to that used in Experiment 1. All participants completed 4 phases of the task: pre-test, training, retention, and post-test. During the pre- and post-test phases, participants completed the line discrimination (i.e., *transfer*) task, whereas during the training and retention phase they completed the dot discrimination (i.e., *training*) task. During all of the phases, a trial proceeded similarly. First, children were presented with a choice opportunity, in which they were allowed to choose between two "games" (i.e., yellow or green in the transfer task and red or blue in the training task). Next, children completed a discrimination trial in which they had to identify the longer line or the larger dot set. Below, we describe the differences between the phases and task conditions.

In the *pre-test phase*, participants across the conditions completed 2 practice trials and 16 pre-test trials of the transfer (i.e., line discrimination) task. Importantly, participants did not receive feedback or instruction as to how to perform optimally. In the *training* phase, participants were randomly assigned to either a control or strategy instruction condition. All participants in this phase completed 2 practice trials and 32 test trials of the training (e.g., dot discrimination) task. Participants in the strategy training condition were given feedback on every trial of the training phase and were reminded to select the easier task at the onset of each trial. Participants in the control condition were given no feedback or strategy instruction. Thus, they only gained exposure to the training task, but received no scaffolding to encourage their metacognitive performance. In the *retention* phase, all participants regardless of condition completed 16 trials of the training task in the absence of any feedback or strategy instruction, to evaluate whether they continued to use a strategy in the same task once scaffolding was removed. The post-test phase was identical to the pre-test phase, except that all participants completed 32 test trials. Again, no feedback or strategy was provided in this phase. We expected that participants who learned to use a strategy in the training phase would transfer this strategy to a novel task (i.e., in the post-test phase) more so than those who did not learn to use the strategy (i.e., non-learners in the strategy training condition), or those who received no training (i.e., in the control condition).

Results and Discussion

Pre-Test Phase

Discrimination Accuracy. Participants completed the line discrimination task in the pre-test phase. A paired-samples t-test was conducted to ensure that performance in fact differed between the easy and difficult trial types. Participants' accuracy was higher in the easy game relative to the difficult game, in the pre-test phase of both the control, $t(27) = 6.76$, $p < .001$, $d = 1.24$, and strategy instruction, $t(48) = 7.05$, $p < .001$, $d = 1.28$, conditions.

Metacognitive Control. The proportion of participants' easy task choices served as the measure of metacognitive control. Three subjects chose the difficult game more than would be expected by chance (on at least 13 of 16 trials; $p < .05$) during the pre-test phase. This indicated that they were systematically choosing the more challenging game, or that they showed a color preference. This makes their training and post-test data difficult to interpret, so these 3 subjects were excluded for these and all following analyses.

As expected, neither the control group nor the strategy instruction group chose the easier game more than would be expected by chance (50%, both *p*s > .2) during the pre-test phase, nor did the two groups differ from one another ($p = .96$; see Table 6).

Individuals who chose the easy game more than would be expected by chance (on at least 13 of 16 trials, p < .05) were classified as optimizers. Only 4% ($N = 1$) of the sample in the control condition and 2% (*N* = 1) of the sample in the strategy instruction condition were categorized as optimizers. According to a chi-square analysis, these proportions did not significantly differ ($p = .69$).

Training Phase

Discrimination Accuracy. Participants completed the dots discrimination task in the training phase. Participants in the strategy instruction condition received performance feedback and were reminded to select the easier game on each trial. Participants in the control condition received no feedback or strategy instruction. As in the pre-test, participants in both the control and strategy instruction conditions were more accurate in the easy game than the difficult game (both *p*s < .001).

Metacognitive Control. As expected, participants in the control condition did not choose the easier game more than would be expected by chance, $p = .36$. In the strategy instruction condition, on the other hand, participants chose the easier game more than would be expected by chance, $t(48) = 7.91$, $p < .001$, $d = 1.13$.

In the training phase, only 4% (*N* = 1) of participants in the control condition were categorized as optimizers. We tested participants in the strategy instruction condition until 30 participants learned to choose the easier game (i.e., more than chance). This resulted in a sample in which 61% (*N* = 30) of participants chose the easy game more than would be expected by chance. Those individuals who optimized in the training phase of the strategy instruction condition will be heretofore referred to as *learners* (i.e., those individuals who were effectively trained). Those in the strategy instruction condition who did not optimize will be referred to as *non-learners* (*N* = 19).

Retention Phase

Discrimination Accuracy. Participants completed the dot discrimination task in the retention phase, in the absence of any feedback or strategy instruction. As in previous phases, participants were more accurate in the easy game than the difficult game in both the control and strategy instruction conditions (both ps < .001).

Metacognitive Control. In the control condition, participants chose the easier game on 54% of trials, which was not significantly different from chance (p = .24). In the strategy instruction condition, participants chose the easier game on 79% of trials, which differed significantly from chance, even in the retention phase where no feedback or strategy instruction were provided. We then assessed the proportion of easy task choices for learners and non-learners. Learners chose the easy task more than would be expected by chance in the retention phase (M = 98%, $t(29)$ = 57.26, p = .96, d = 10.45), whereas non-learners did not (M = 50%, p = .96; see Table 6).

Ten percent of participants in the control condition (N = 3) were categorized as optimizers in the retention phase, as were 67% (N = 33) of individuals in the strategy instruction condition. When participants in the strategy instruction condition were broken down by learner status, 100% (N = 30) of learners optimized by continuing to systematically select the easy game in the retention phase. However, only 15% (N = 3) of non-learners optimized. This suggests that the individuals who learned to select the easy game in the training phase continued to do so even when the feedback and repeated strategy instruction were removed.

Post-Test Phase

Discrimination Accuracy. Participants completed the line discrimination task in the post-test phase. Accuracy was higher in the easy game than the difficult game in both the control condition and the strategy instruction condition (both *p*s $< .001$).

Metacognitive Control. As in all the prior phases, participants in the control condition did not select the easy game more than would be expected by chance in the post-test phase, $M = 55\%$, $p = .20$. In the strategy instruction condition, however, participants selected the easy game more than would be expected by chance, $M = 62\%$, $t(48) = 3.63$, $p < .005$, $d = .52$. Separating participants based on learner status revealed that those who learned in the training phase chose the easy game 67% of the time (above 50%; $t(29) = 3.68$, $p < .005$, $d = .67$), indicating that they successfully transferred a strategy to the post-test phase (see Table 6). Non-learners, however, did not choose the easy game more than would be expected by chance ($M = 53\%$; $p = .32$).

Eighteen percent ($N = 5$) of participants in the control condition, and 30% ($N = 15$) of participants in the strategy instruction condition were categorized as optimizers in the post-test phase. Separating participants from the strategy instruction condition into learners and non-learners revealed that 43% ($N = 13$) of learners systematically chose the easier game, whereas only 11% ($N = 2$) of non-learners did so.

Effects of Strategy Training

Of interest was whether participants who learned to employ a strategy in the training phase readily transferred that strategy to a novel task in the post-test phase. First, to assess the effects of successful training on strategy transfer, we entered the proportion of easy task choices in the training condition into a 2 (Test phase: Pre vs. post) x 3

(Learner status: Non-learners vs. learners) repeated measures ANOVA. There was a main effect of test phase, $F(1, 56) = 9.23$, $p < .005$, $\eta^2 = .14$, and a significant test phase by learner status interaction, $F(1, 56) = 4.46$, $p < .05$, $\eta^2 = .07$, suggesting that the effect of training on transfer was larger in learners than non-learners (see Figure 23).

Similarly, a 2 (Test phase: Pre vs. post) x 2 (Learner status: Control vs. learners) repeated-measures ANOVA revealed a main effect of test phase, $F(1, 47) = 5.37$, $p < .05$, $\eta^2 = .10$, and a significant test phase by learner status interaction, $F(1, 47) = 4.34$, $p < .05$, $\eta^2 = .09$. The significant interaction indicates that the training effect was larger for learners than in individuals in the control group.

Finally, a 2 (Test phase: Pre vs. post) x 2 (Learner status: Control vs. non-learners) repeated-measures ANOVA indicated no main effect of test phase ($p = .52$) or learner status ($p = .92$), and no significant interaction ($p = .73$), suggesting that there was no difference between the training effect for non-learners and individuals in the control condition.

Overall, these findings suggest that young children who were able to learn and apply a strategy in the training phase also applied the strategy (1) when the scaffolding was removed and, more importantly, (2) when presented with a novel task.

Chapter 7: General Discussion

The studies reported in Chapters 3-6 introduced several novel findings pertaining to the development of metacognitive monitoring and control, as well as how these components interact throughout development, and the kinds of information young children use to monitor and control behavior.

In Chapter 3, it was demonstrated that, in contrast to adults, young children are not cognitive misers and do not minimize effort under typical circumstances. Perhaps more importantly, findings suggested that metacognitive monitoring and control could be distinct processes, with each undergoing protracted development. At the same time, the 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.

In Chapter 4, prior findings were extended, further suggesting that metacognition undergoes protracted development and its components exhibit asynchronous and independent developmental trajectories. Most importantly, there were dissociations between monitoring and control across development, thus providing evidence against both the CM (Control → Monitoring) and MC (Monitoring → Control) models of metacognition. These findings suggest that monitoring and control can function independently from early childhood, and that the two components do not seem to become more coupled with experience and development.

In Chapter 5, it was demonstrated that young children's metacognitive control could benefit from performance feedback, provided that they are able to form representations of performance that are sufficiently separated and precise. In addition, whereas children relied on external cues of performance (i.e., feedback) to select the more beneficial of two tasks, they relied on internal cues of effort to make appraisals of the task difficulty. These findings suggest that the monitoring and control processes rely on different sources of information in early childhood, and that young children either cannot or do not use monitored differences in effort to control behavior.

Finally, in Chapter 6, young children who learned to apply a strategy rule transferred that rule to a novel task. This suggests that the strategy formed was not dependent on the feedback and frequent strategy reminders used in training, as children continued to apply the strategy even when this scaffolding was removed. In addition, the findings suggested that the strategy formed was not tied to the specific stimuli in the task in which they learned. On the contrary, young children formed a strategy that was sufficiently abstract to identify and select the easier game in a task with unique stimuli.

These findings have provided insight into the development of monitoring, control, and their interactions. In addition, it has elucidated the role of children's representations of task success in their metacognitive performance. Finally, the studies presented in this dissertation have implications for the training of metacognition. In what follows, I discuss each of these topics.

7.1 The Development of Monitoring, Control, and Their Interactions

In Experiments 1, 3, 4, and 5, we demonstrated differences in metacognitive performance between 5-year-olds, 7-year-olds, and adults. For example, in Experiment 3, 5-year-olds did not systematically select an easier task, whereas 7-year-olds and adults did so spontaneously. This highlights the development of strategy formation and execution between 5 and 7 years of age. Seven-year-olds and adults were able to both formulate and execute a strategy for optimal performance, whereas 5-year-olds were not. One possibility is that 5-year-olds were not able to *formulate* a strategy (i.e., to select the easier game for greater reward). This could be due to (at least) two reasons: 1) they did not recognize their performance could be improved (e.g., due to overconfidence), or 2) they lacked experience using that particular strategy, or recalling how it could be applied to the task.

Alternatively, 5-year-olds may have successfully formulated the strategy but failed to *execute*, or actually carry out, the strategy. This would reflect a difficulty in switching strategies, or inhibiting a prepotent strategy (selecting randomly) in favor of a more optimal one (selecting the easier game more). This kind of attentional switching indeed develops rapidly during the preschool period (Diamond, 2002), and likely reflects maturation of the prefrontal cortex (Morton & Munakata, 2002). In contrast to younger children, older children and adults may also have greater working memory capacity allowing them to maintain a strategy across trials (Eenshuistra, Weidema, & Hommel, 2004; Kane & Engle, 2003). In addition, they likely have more proficient cognitive flexibility, allowing them to more easily adjust their strategy in the moment (Chevalier & Blaye, 2009).

The backward learning curves calculated in Experiments 3, 4, and 5 can shed some light on how metacognitive control was carried out. For those individuals who did optimize by selecting an easier task, the rate at which they discovered and applied the strategy rule was analyzed. In 7-year-olds and adults, there was little evidence of gradual learning, suggesting instead sudden transitions from one strategy to another. This pattern may represent a kind of insight learning as the strategy rule was discovered and/or applied (Siegler & Araya, 2005). This is in contrast to other types of learning, in which the probability of success is slowly accumulated across trials (Ahissar & Hochstein, 1993). These differences in learning are also likely to transpire in differences in transfer: gradual learning may result in more narrow transfer than rule or strategy discovery (Siegler & Jenkins, 1989).

In terms of performance monitoring, a different developmental trajectory was observed. In Experiments 1 and 3, it was demonstrated that 5-year-olds tended to overestimate their performance, adults tended to underestimate, and 7-year-olds showed more accurate estimations than either of those groups. This may reflect a general bias in how these age groups appraise performance. As discussed in the introduction, there is evidence that young children overestimate their performance, even in the face of counterevidence (e.g., negative feedback). As such, 7-year-olds may be free from such a bias, having outgrown an overestimation bias, but not yet developed an underestimation bias. Future work should investigate the sources of such biases and their change over time.

Although 7-year-olds more precisely estimated their overall performance, adults outperformed them when it came to estimating relative performance (i.e., estimating higher accuracy in the easy game than the difficult game). As such, precision in one's overall estimations does not necessarily translate to precision in *relative* estimations, which may place greater demands on selective attention (i.e., keeping track of successes and failures in both the red and blue games separately). Finally, although overall performance differed across the age groups, scaffolding affected the age groups similarly. In other words, the facilitative effects of feedback and strategy instruction remained stable throughout development.

Taken together, these findings suggest a unique conception of the interaction between monitoring and control. Under under some circumstances, monitoring and control operate independently, and improvements in one component can transpire without improvements in the other. Importantly, our findings do not dispute the possibility that monitoring and control can be coupled, as was demonstrated in previous studies. Our findings do demonstrate that the two components are not *necessarily* coupled, and that they can operate independently in some contexts. As such, any interactions between monitoring and control likely cannot be explained by a simple feed-forward model, and may be moderated by other factors (e.g., conscious awareness).

7.2 Effects of Feedback and Task Experience in Metacognition

7.2.1 Reducing Uncertainty in Task Success Promotes Metacognitive Control

Previous work has shown that performance feedback improved young children's monitoring, but not their control performance. In Chapter 5, two possibilities were tested to explain why children's control performance was unaffected by feedback. The first possibility was that children's representations of task success were not sufficiently *separated* in the previous experiments, leading to overlap in their representations of the two tasks. In Experiments 6a and 6b, metacognitive performance did improve when the reward probabilities were sufficiently separated (i.e., using 0% vs. 100% rather than 60% vs. 90%). Supplying enough separation between the reward probabilities reduced the overlap between representations of performance in the two tasks, likely making it easier for young children to recognize and select the game that led to higher task success. In an Iowa Gambling Task suited for young children, even 4-year-olds showed a preference for a more highly rewarded option (Kerr & Zelazo, 2004), and this preference was stronger when the gain/loss schedules of the two options were more distinct (Gao, Wei, Bai, Lin, & Li, 2009). Until now, however, this preference has not been demonstrated for reward given on the basis of actual performance in a task.

The second (not mutually exclusive) possibility was that the feedback received by children in earlier studies did not allow for *precise* representations of task success. In other words, there may have been uncertainty surrounding the representations of success in the two games. In Experiments 7a and 7b, that uncertainty was reduced by providing children with more experience with the two games. Additional experience receiving feedback in the two games led young children to systematically select the easier game, but only if they received feedback. Similarly, previous findings have shown that adults were more likely to use an adaptive memory strategy after pre-learning the word pairs (Hines, Hertzog, & Touron, 2013), indicating that the extent of learning may be a factor contributing to adaptive strategy use. In the current study, however, young children's control only benefited if they received feedback during pre-exposure. Below, the effects of external and internal cues on young children's monitoring and control are discussed.

7.2.2 Internal and External Signals of Task Success

What kinds of information do young children rely on to monitor and control behavior? It was hypothesized that the experimental manipulations used in Chapter 5 would influence behavior by reducing uncertainty in representations of 1) task performance (i.e., from feedback) or 2) amount of effort expended in the tasks. Across the reported experiments, there was dissociation between the types of information used for monitoring and control. To control behavior, young children seemed to rely on external signals of task success (i.e., feedback). However, they appeared to use internal signals of effort to make task difficulty judgments. Below, we discuss each of these findings in turn.

The fact that young children relied on feedback to make task selections is somewhat consistent with previous work conducted with adults. In particular, adults have demonstrated a tendency to restudy more highly rewarded items (i.e., those worth more points on an upcoming test), even if these items are easier to learn (or potentially already learned; Ariel, Dunlosky, & Bailey, 2009). However, though adults can rely on reward if it is provided, they can rely on item difficulty in the absence of reward information. Young children in Experiment 7b, in contrast, did not make task selections on the basis of difficulty, even with additional experience with the task. Thus, unlike adults, young children do not spontaneously rely on internal signals of effort to control behavior. Instead, they only make strategy adjustments after receiving an external cue, like feedback, about their task performance. Future work should assess whether some types of feedback (i.e., positive or negative) are more effective at driving young children's behavior than others. For example, they may be more motivated to avoid negative feedback, than to gain positive feedback.

That children make task selections on the basis of feedback is not to say that young children are wholly insensitive to information about effort. On the contrary, young children did base their appraisals of task difficulty on the amount of effort expended in the tasks, rather than their actual performance. This was evidenced by the fact that task monitoring accuracy suffered in Experiment 6b relative to 6a, when differences in effort were eliminated and children could only rely on the feedback information. In addition, young children's task monitoring was unaffected when feedback was removed in Experiment 7b (relative to Experiment 7a), indicating that their ability to identify an easier task was not dependent on receiving differential feedback. This suggests that, in some ways, young children are sensitive to differences in effort. Importantly, however, they do not or cannot use this information to drive their task selections.

If young children *cannot* use internal signals of effort to control behavior, this would mean that they do not have access to these representations, or that the representations are too noisy to drive behavioral adjustments. On the other hand, it is possible children simply do not *spontaneously* use the information, and that they can access it under some circumstances (e.g., if explicitly accessed by task monitoring questions, or by promoting the selection of an easier task). Future work should assess whether young children are *capable* of using internal signals of effort to control behavior, in the absence of external signals of performance. This work indicates that they do not do so spontaneously.

7.3 Implications for the Training of Metacognition

The present work has implications for intervening on metacognition both in and outside the lab. These data indicated that young children have a difficult time prioritizing a task based on the task difficulty alone, and that they rely on external cues of performance to adjust to a more adaptive strategy. This suggests that receiving explicit feedback is crucial for young children's strategy use, and that negative feedback may be as necessary as positive feedback to this end. Other work has established that providing children with unconditional positive praise may backfire in other domains, especially if it refers to ability rather than effort (Mueller & Dweck, 1998), encourages social comparison (Corpus, Ogle, & Love-Geiger, 2006), or if children have low self-esteem (Brummelman, Thomaes, de Castro, & Bushman, 2014). This work, which shows that children need greater separation between reward probabilities to prompt the engagement of metacognitive control, also highlights the importance of giving children negative feedback in addition to positive feedback to encourage adaptive strategy use.

In addition, the results from the training study in Experiment 8 highlight that metacognition can be trained in young children in a way that transfers to a novel task. Though only near transfer was tested here, this creates an important first step toward evaluating whether learning from a simple training will transfer to a more dissimilar task, or tasks with greater real-world significance. Young children were trained to identify and select an easier task. As such, similar training may help children identify items that are more difficult to learn, and to select them for further study. Metacognitive training could also improve children's ability to recognize a difficult task, to undercut overconfidence and promote help-seeking behavior. In addition, it could help children optimize their time in a testing context, selecting and completing easier items before difficult items to save time. The current evidence of strategy transfer is preliminary, however, and these specific possibilities should be directly investigated in future research.

7.4 Conclusion

This dissertation research presents novel evidence regarding the development of metacognitive monitoring and control, and the interactions between these components. First, metacognition is malleable in early childhood, as well as adulthood, suggesting that the kind of information people receive (e.g., performance feedback and/or strategy instruction) can influence their metacognitive proficiency. Second, metacognitive monitoring and control are able to operate independently across the lifespan. This project has provided evidence against a simple feed-forward model of metacognition, indicating that monitoring can improve without improvements in control, and vice versa. Third, young children's ability to monitor and control behavior is sensitive to how they represent the probability of their success in a task. In particular, more highly separated task representations can lead to improved monitoring and control. More specifically, young children's monitoring relies on differences in levels of perceived effort, whereas their control relies on differences in external signals of performance, like explicit performance feedback. Finally, young children are able to learn and transfer a strategy rule across tasks with different stimuli characteristics. This suggests that even young children are able to form abstract strategies that can be applied across multiple task contexts. These results have important implications for theories of metacognition, as well as for developing training regimens for metacognition that can be applied in various domains (e.g., study time allocation and help-seeking).

References

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences*, *90*, 5718-5722. http://dx.doi.org/10.1073/pnas.90.12.5718

Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: when agendas override item-based monitoring. *Journal of Experimental Psychology: General*, *138*, 432-447. http://dx.doi.org/10.1037/a0015928

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444. http://dx.doi.org/ 10.1146/annurev-psych-113011-143823

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.

Borkowski, J. G., Waters, H. S., & Schneider, S. (2009). *Metacognition, strategy use, and instruction*. New York, NY: Guilford Press.

Brummelman, E., Thomaes, S., de Castro, B. O., Overbeek, G., & Bushman, B. J. (2014). "That's not just beautiful—that's incredibly beautiful!" The adverse impact of inflated praise on children with low self-esteem. *Psychological Science*, 25, 728-735. https://doi.org/10.1177/0956797613514251

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.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, *45*, 782. http://dx.doi.org/10.1037/a0015409

Corpus, J. H., Ogle, C. M., & Love-Geiger, K. E. (2006). The effects of social-comparison versus mastery praise on children's intrinsic motivation. *Motivation and Emotion*, *30*, 333-343. http://dx.doi.org/10.1007/s11031-006-9039-4

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.

Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review*, *17*, 462-489. http://dx.doi.org/10.1006/drev.1997.0442

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.

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

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*(2), 274-296.

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

Eenshuistra, R. M., Weidema, M. A., & Hommel, B. (2004). Development of the acquisition and control of action–effect associations. *Acta Psychologica*, *115*, 185-209. http://dx.doi.org/10.1016/j.actpsy.2003.12.006

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906-911.

Gao, S., Wei, Y., Bai, J., Lin C., & Lo, H. (2009). Young children's affective decision-making in a gambling task: Does difficulty in learning the gain/loss schedule matter? *Cognitive Development*, *24*, 183-191. http://dx.doi.org/10.1016/j.cogdev.2008.07.006

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. 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. 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*(9), 1768-1776.

Hertzog, C., Hines, J. C., & Touron, D. R. (2013). Judgments of learning are influenced by multiple cues in addition to memory for past test accuracy. *Archives of Scientific Psychology*, *1*, 23-32. http://dx.doi.org/10.1037/arc0000003

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*(4), 679-709.

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://doi.org/ 10.1348/026151009X479042

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.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*, 47. http://dx.doi.org/10.1037/0096-3445.132.1.47

Kendler, T. S. (1972). An ontogeny of meditational deficiency. *Child Development, 43*(1), 1-17.

Kerr, A., & Zelazo, P. D. (2004). Development of "hot" executive function: The children's gambling task. *Brain and Cognition*, *55*, 148-157. http://dx.doi.org/10.1016/S0278-2626(03)00275-6

Kool, W., & Botvinick, M. (2014). A labor/leisure tradeoff in cognitive control. *Journal of Experimental Psychology: General*, *143*(1), 131.

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*(4), 665-682.

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. 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. 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. 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. 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

Li, W., Ji, H., Li, F., Li, P., Zhang, Y., & Li, X. (2016). Metacognitive control: Shifting from habitual to agenda processes on item selection during study in elementary school children. *The British Psychological Society, 86,* 657-670.

http://dx.doi.org/10.1111/bjep.12130

Liberty, C., & Ornstein, P. A. (1973). Age differences in organization and recall: The effects of training in categorization. *Journal of Experimental Child Psychology*, *15*(1), 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*(2), 153-169.

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child Development*, *82*(6), 1778-1787.

Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! Introspection on uncertainty supports early strategic behavior. *Child Development*, *84*(2), 726-736.

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. (2002). Is study time allocated selectively to a region of proximal learning?. *Journal of Experimental Psychology: General*, *131*, 349-363. http://dx.doi.org/10.1037//0096-3445.131.3.349

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*, 463-477. http://dx.doi.org/10.1016/j.jml.2004.12.001

Metcalfe, J. E., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing*. The MIT Press.

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: the influence of incentives and feedback on exam predictions. *Metacognition and Learning*, *6*, 303-314. http://dx.doi.org/10.1007/s11409-011-9083-7

Moely, B. E., Olson, F. A., Halwes, T. G., & Flavell, J. H. (1969). Production deficiency in young children's clustered recall. *Developmental Psychology*, *1*(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. http://dx.doi.org/10.1016/j.neuroimage.2009.01.037

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://dx.doi.org/10.1002/dev.10033

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*, 33-52. http://dx.doi.org/10.1037/0022-3514.75.1.33

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.

O’Leary, A. P., & Sloutsky, V. M. (2016). Carving metacognition at its joints: Protracted development of component processes. *Child Development*. http://dx.doi.org/10.1111/cdev.12644

O'Leary, A. P., & Sloutsky, V. M. (submitted). Metacognitive monitoring and control can function independently: Evidence from 5-year-olds, 7-year-olds, and adults.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, *19*(1), 90-138.

Reder, L. M. (1996). *Implicit memory and metacognition*. Psychology Press.

Rehder, B., & Hoffman, A. B. (2005). Eye-tracking 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 adult's event recall and suggestibility. *Developmental Psychology*, *38*(6), 1052.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*(2), 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. http://dx.doi.org/10.1111/1467-9280.00076

Siegler, R., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior*, *33*, 1-42. http://dx.doi.org/10.1016/S0065-2407(05)80003-5

Siegler, R. S., & Jenkins, E. A. (1989). How children discover new strategies. Hillsdale, NJ: Erlbaum.

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

Son, L. K., & Kornell, N. (2008). Research on the allocation of study time: Key studies from 1890 to the present (and beyond). In *A Handbook of Memory and Metamemory* (pp. 333-351). New York, NY: Psychology Press.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 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*(9), 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*(8), 1649-

1657.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, *8*, 327-353. http://dx.doi.org/10.1016/S1041-6080(96)90022-9

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. (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*(1), 37-63.

Appendix: Tables and Figures

The tables and figures for Experiments 1-8 are provided on the following pages.

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)

Table 3. Summary of findings in Experiment 3.

	5-Year-Olds	7-Year-Olds	Adults
Discrimination accuracy			
Overall	0.66	0.86	0.93
Easy trials	0.75	0.99	1.00
Difficult trials	0.57	0.68	0.69
Discrimination RT (ms)			
Overall	900	806	726
Easy trials	805	625	698
Difficult trials	1204	937	973
Control			
Easy task choices	0.51	0.58	0.75
Optimizers (out of 30)	7% (*N* = 2)	27% (*N* = 8)	66% (*N* =20)
Performance monitoring			
Absolute (error)	0.20	0.05	0.14
Relative (out of 30)	13% (*N* = 4)	30% (*N* = 9)	80% (*N* = 24)

Table 4. Summary of findings in Experiment 4.

	5-Year-Olds	7-Year-Olds	Adults
Discrimination accuracy			
Overall	0.71	0.89	0.94
Easy trials	0.80	0.98	0.97
Difficult trials	0.55	0.60	0.63
Discrimination RT (ms)			
Overall	1378	830	698
Easy trials	1293	683	677
Difficult trials	1516	1125	1160
Control			
Easy task choices	0.57	0.73	0.92
Optimizers (out of 30)	20% (*N* = 6)	63% (*N* = 19)	93% (*N* =28)
Performance monitoring			
Absolute (error)	0.14	0.10	0.12
Relative (out of 30)	33% (*N* = 10)	33% (*N* = 10)	67% (*N* =20)

Table 5. Summary of findings from Experiment 5.

	5-Year-Olds	7-Year-Olds	Adults
Discrimination accuracy			
Overall	0.76	0.83	0.96
Easy trials	0.91	0.95	1.00
Difficult trials	0.58	0.70	0.74
Discrimination RT (ms)			
Overall	1230	1092	653
Easy trials	1050	936	612
Difficult trials	1322	1166	932
Control			
Easy task choices	0.55	0.53	0.85
Optimizers (out of 30)	13% ($N = 4$)	13% ($N = 4$)	77% ($N =23$)
Performance monitoring			
Absolute (error)	0.11	0.04	0.06
Relative (out of 30)	47% ($N = 14$)	53% ($N = 16$)	83% ($N = 25$)

Table 6. Proportions of easy task choices and optimizers in each phase of Experiment 8.

	Control ($N = 28$)	Training (Non-Learners; $N = 19$)	Training (Learners; $N = 30$)
Pre-Test Phase			
Easy task choices	0.52	0.52	0.52
Optimizers	0.04	0.00	0.03
Training Phase			
Easy task choices	0.52	0.50	0.95
Optimizers	0.04	0.00	1.00
Retention Phase			
Easy task choices	0.54	0.50	0.98
Optimizers	0.11	0.09	1.00
Post-Test Phase			
Easy task choices	0.55	0.53	0.67
Optimizers	0.18	0.07	0.43

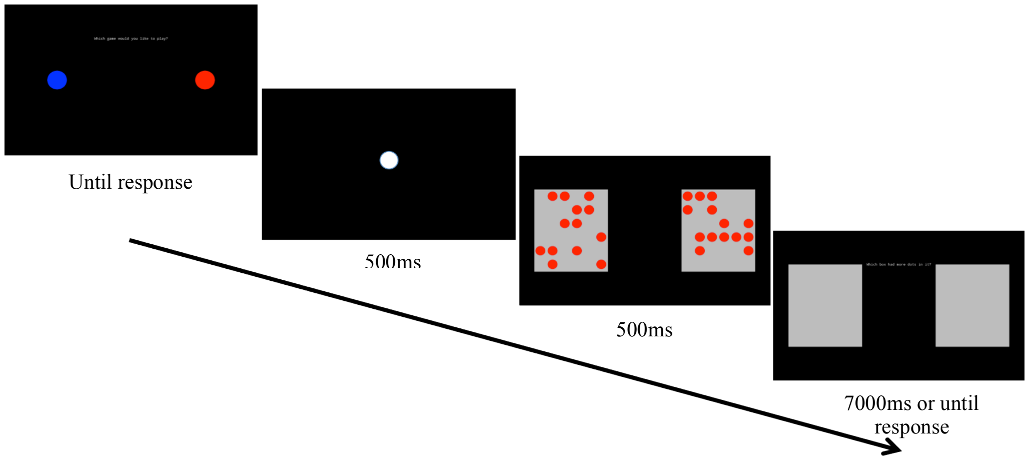


Figure 1. The task sequence including choice opportunity, fixation, test stimulus, and response screen.

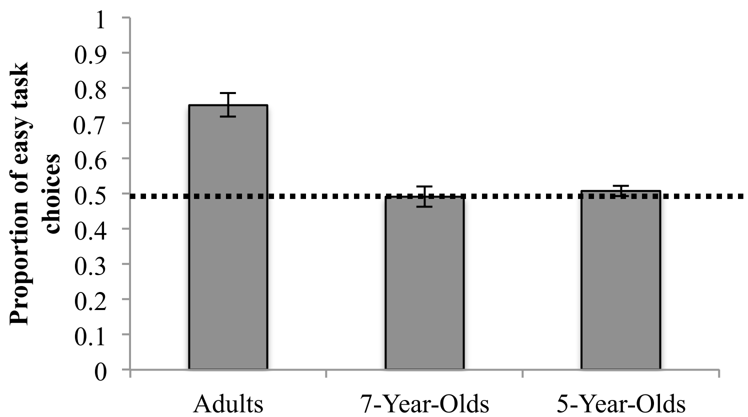


Figure 2. Proportion of easy task choices in Experiment 1 by age group.

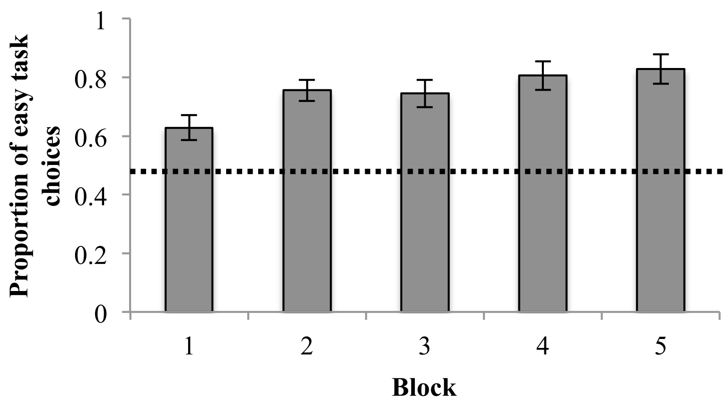


Figure 3. Proportion of adults' easy task choices in Experiment 1 by block.

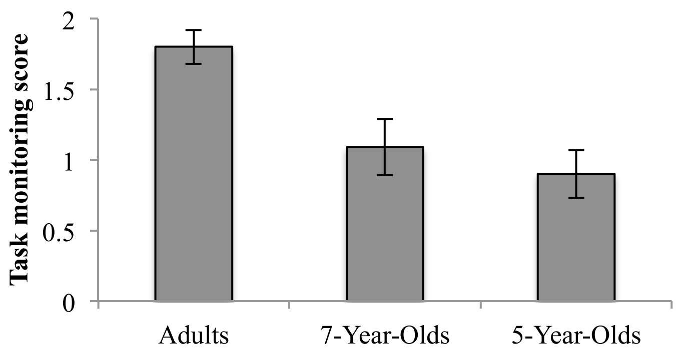


Figure 4. Composite task monitoring scores in Experiment 1 by age group.

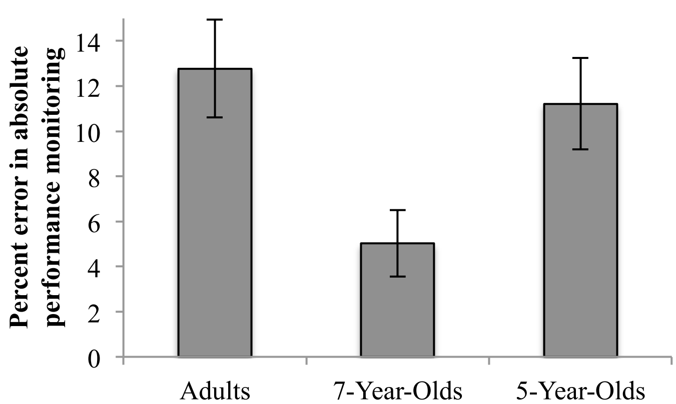


Figure 5. Absolute performance monitoring scores in Experiment 1 by age group (lower scores indicate more accurate performance estimation).

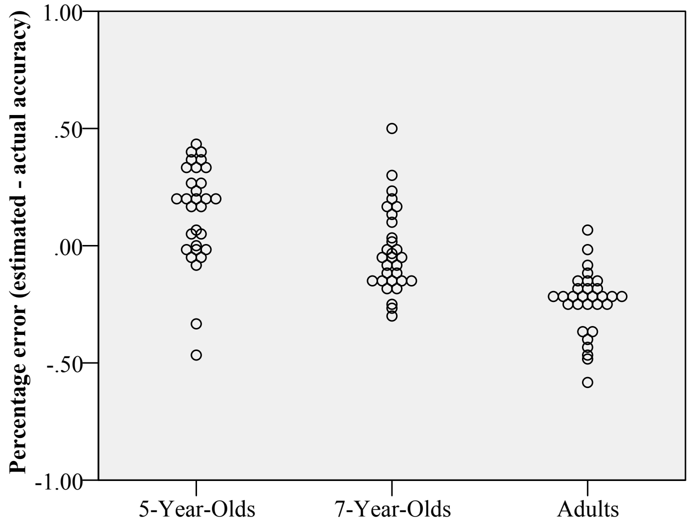


Figure 6. The direction of performance estimations for each age group in Experiment 1. Scores above zero indicate overestimation.

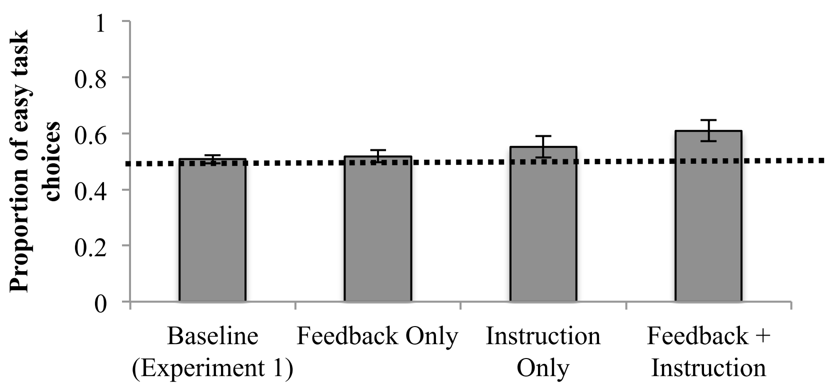


Figure 7. Proportions of 5-year-olds' easy task choices in Experiments 1 and 2 by condition.

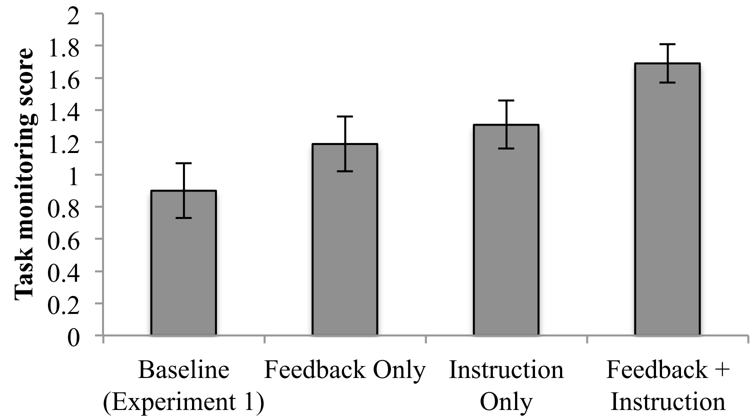


Figure 8. Participants' task monitoring scores in Experiments 1 and 2 by condition.

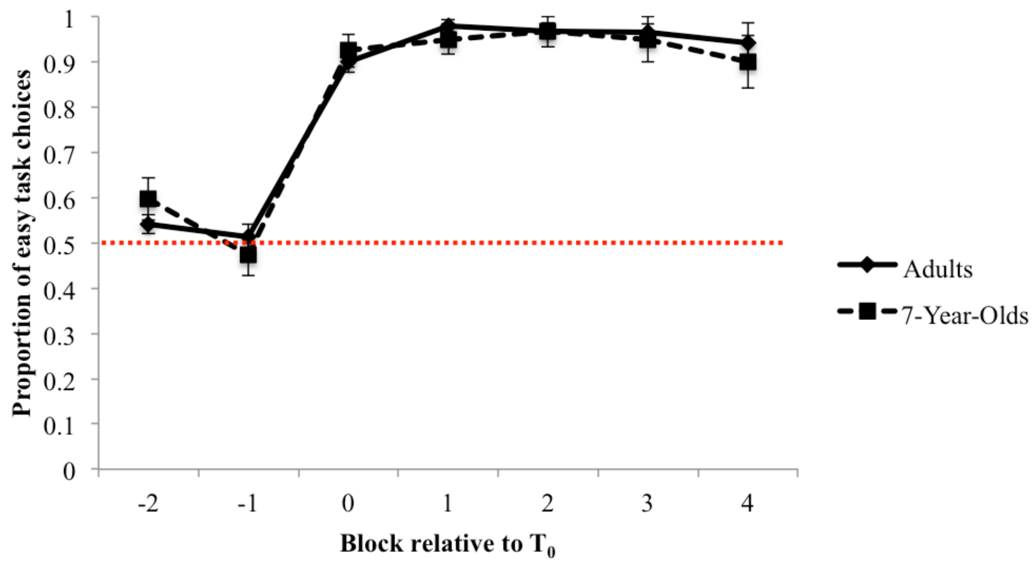


Figure 9. Backward learning curves for adults and 7-year-olds in Experiment 3.

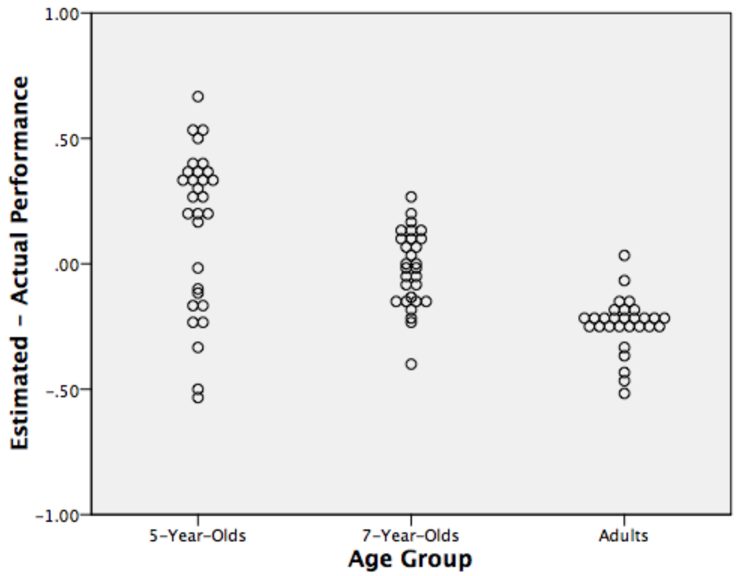


Figure 10. Estimated - actual performance for 5-year-olds, 7-year-olds, and adults in Experiment 3.

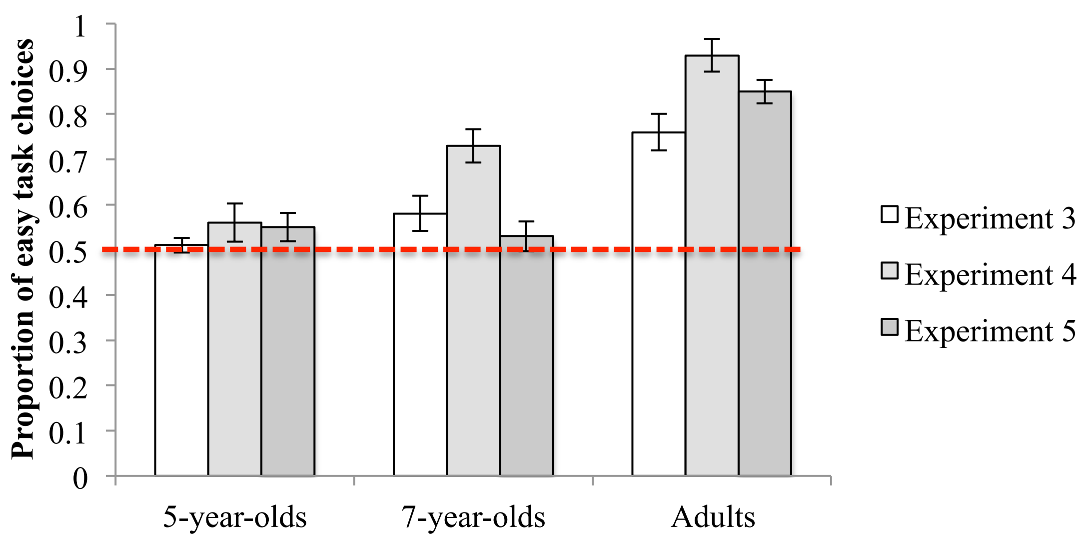


Figure 11. Proportion of easy task choices across Experiments 3, 4, and 5.

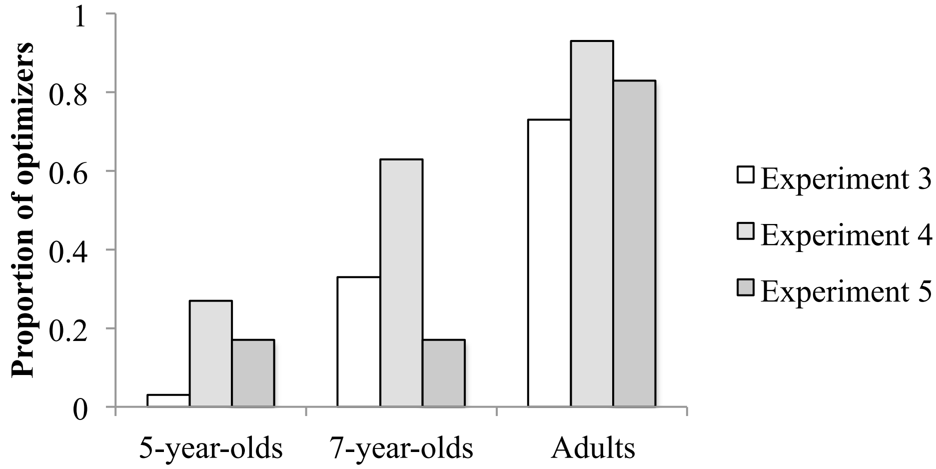


Figure 12. Proportion of optimizers across Experiments 3, 4, and 5.

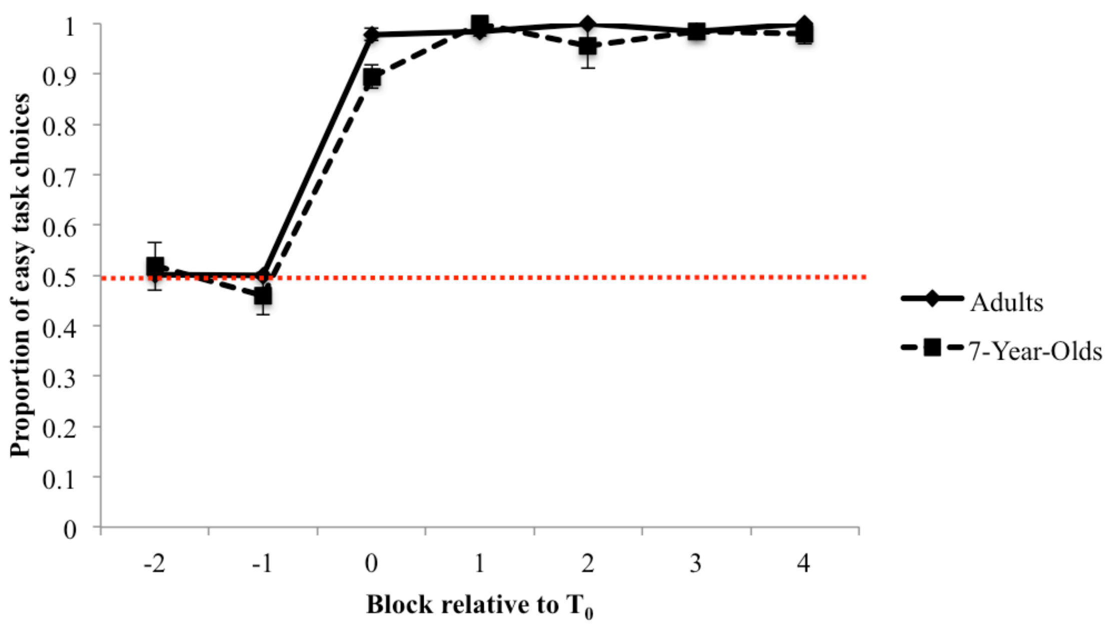


Figure 13. Backward learning curves for 7-year-olds and adults in Experiment 4.

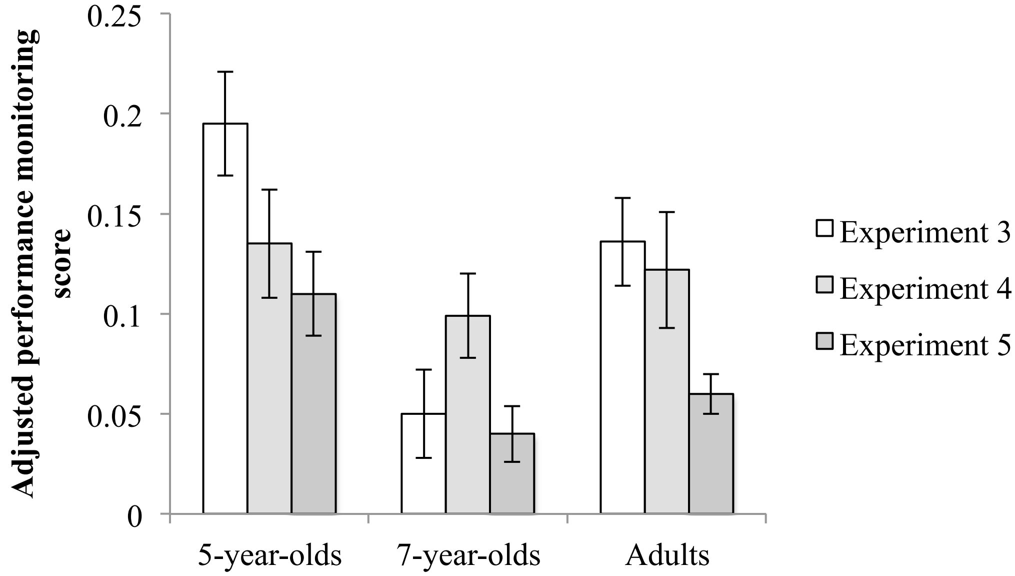


Figure 14. Adjusted performance monitoring scores across Experiments 3, 4, and 5 (lower scores mean more precise estimations).

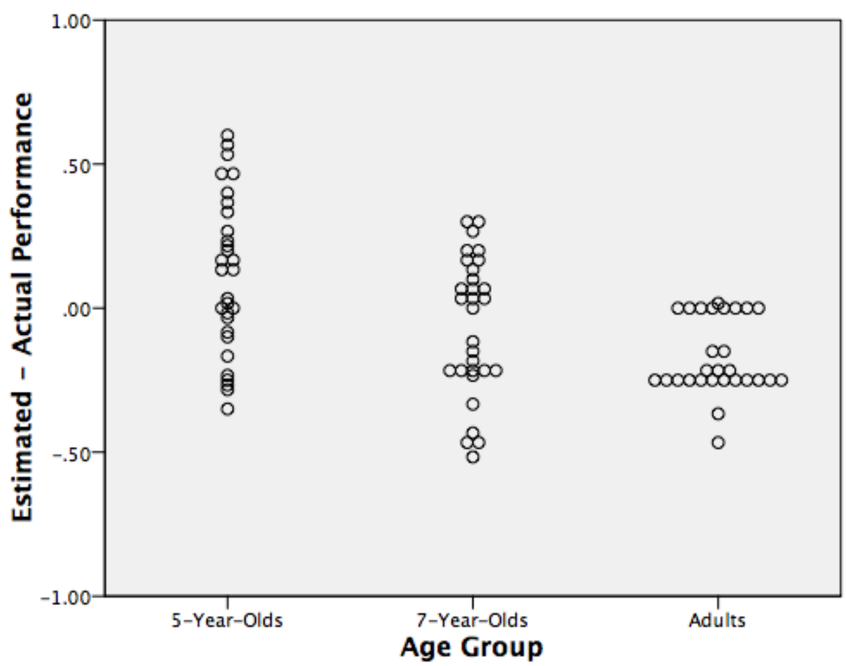


Figure 15. Estimated - actual performance for 5-year-olds, 7-year-olds, and adults in Experiment 4.

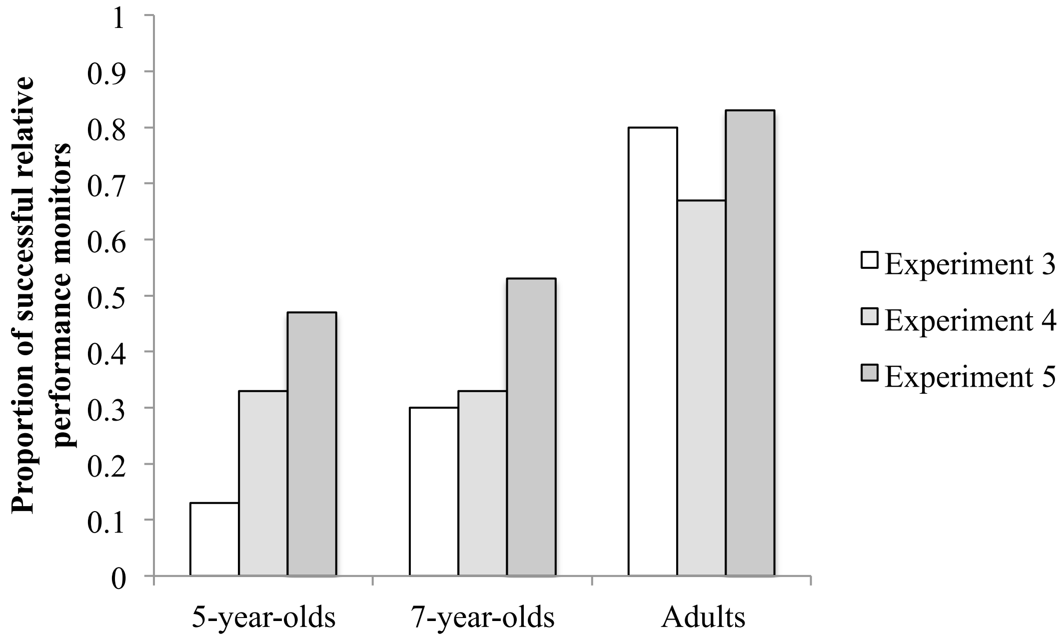


Figure 16. The proportion of participants who correctly monitored their relative performance across Experiments 3, 4, and 5.

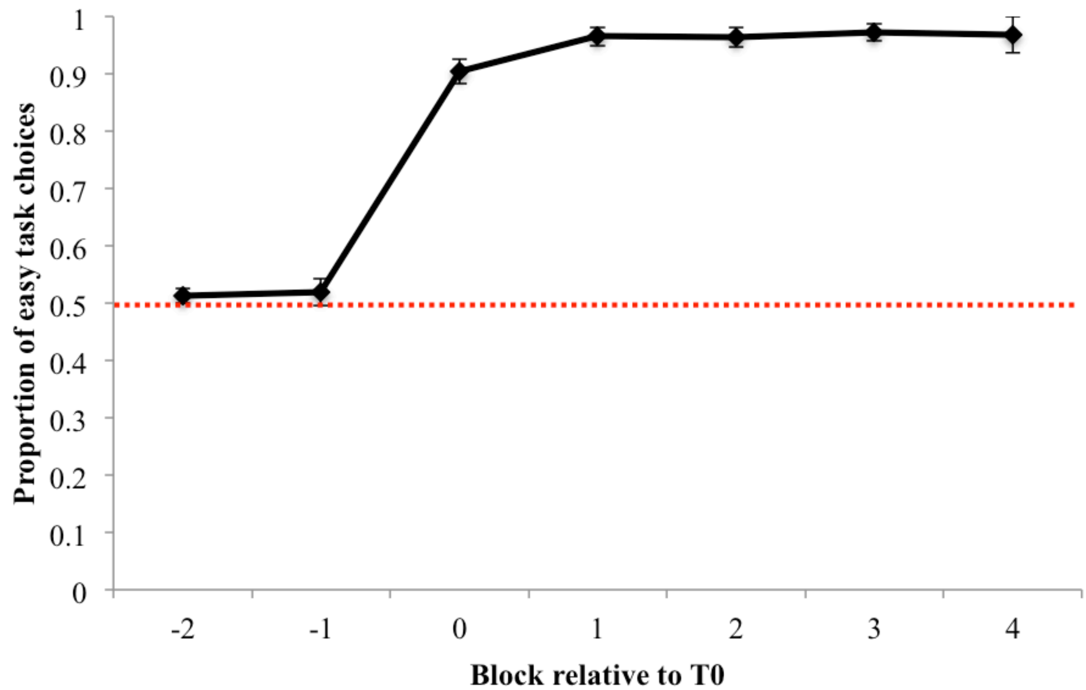


Figure 17. Backward learning curve for adults in Experiment 5.

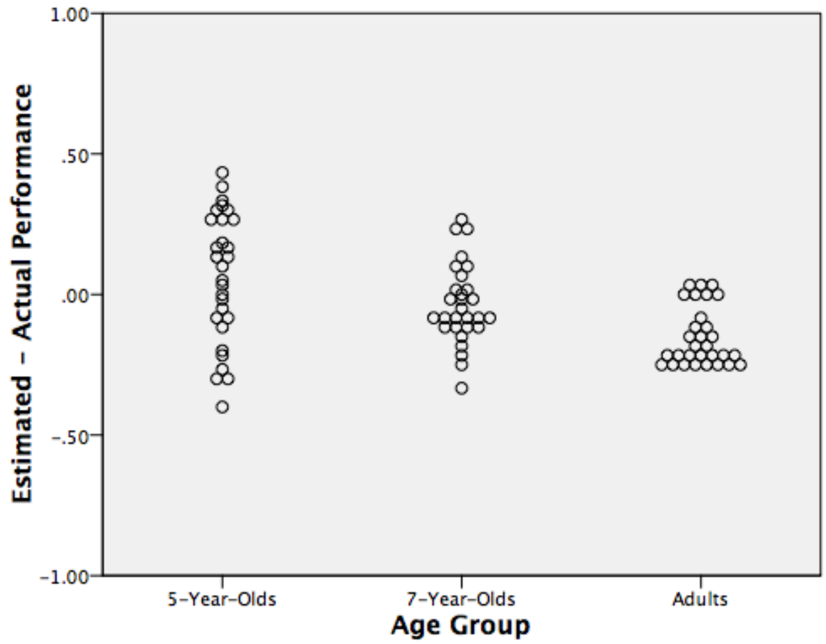
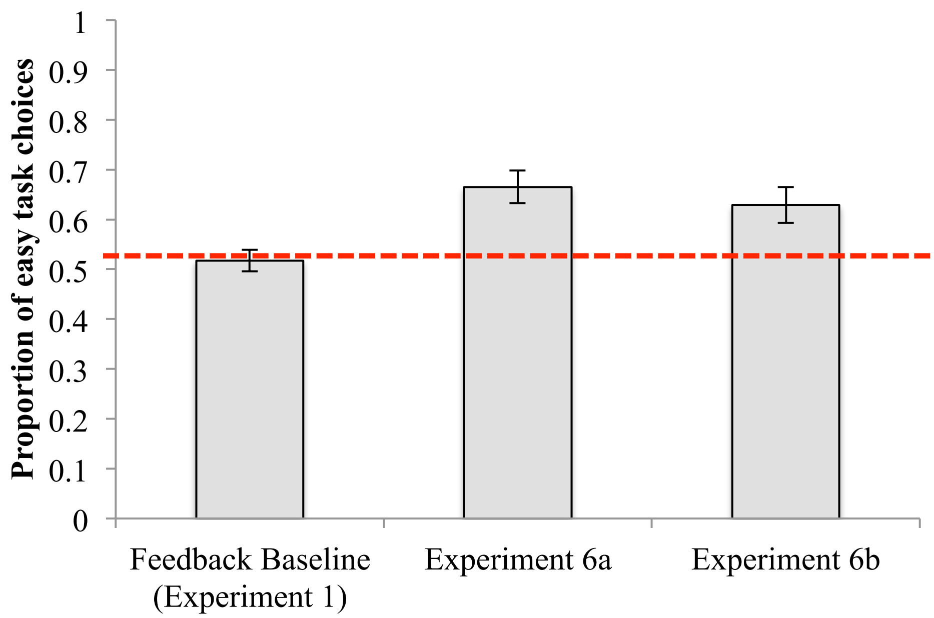


Figure 18. Estimated - actual performance for 5-year-olds, 7-year-olds, and adults in Experiment 5.

A. Proportion of easy task choices



B. Proportion of optimizers

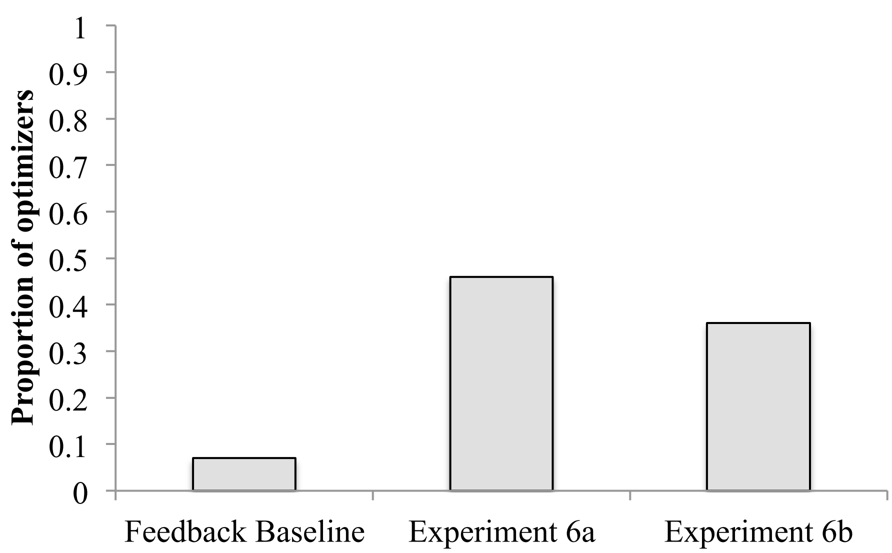


Figure 19. Proportion of easy task choices (A) and optimizers (B) in Experiments 6a and 6b.

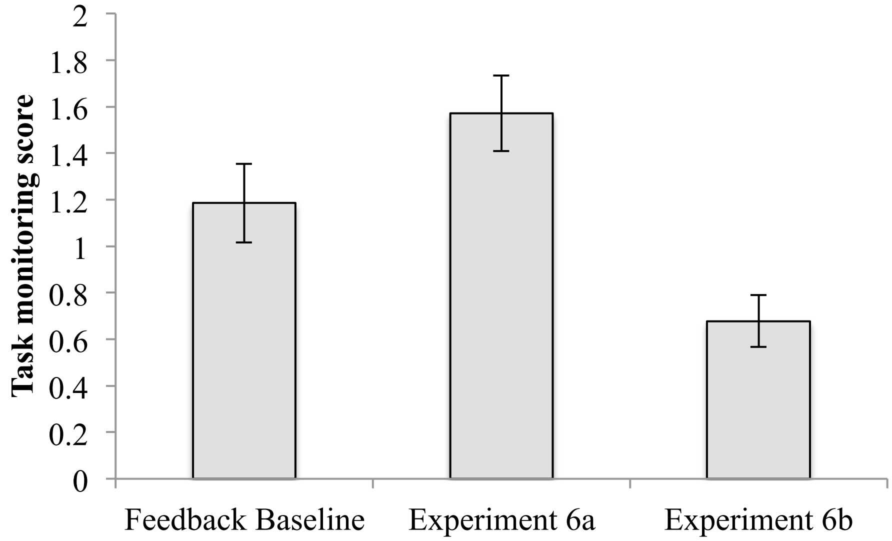
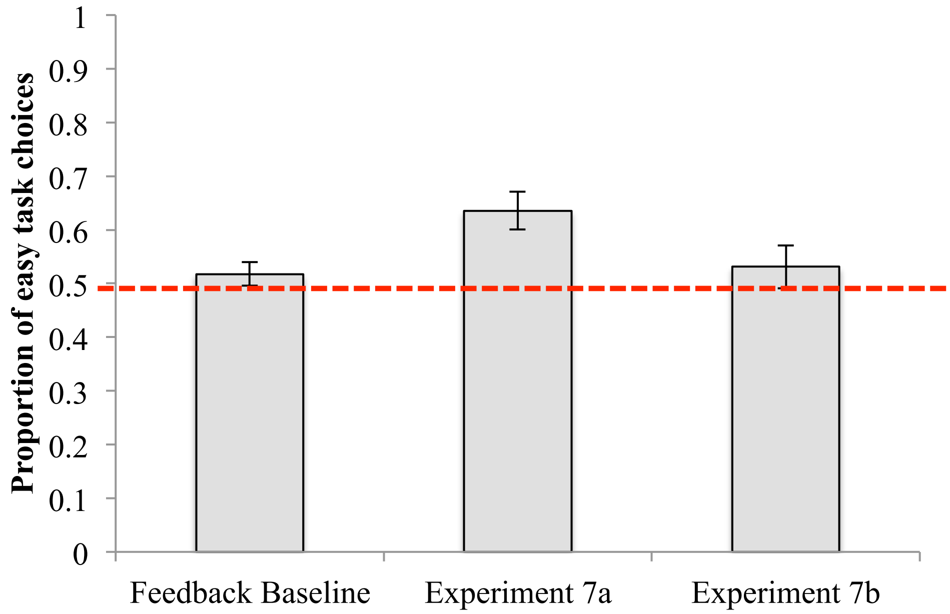


Figure 20. Task monitoring scores in Experiments 6a and 6b.

A. Proportion of easy task choices



B. Proportion of optimizers

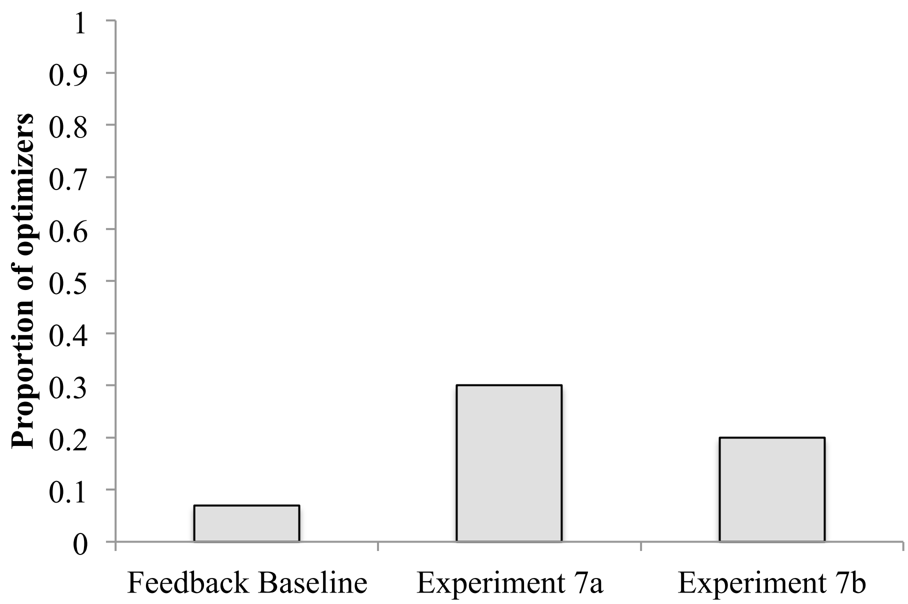


Figure 21. Proportion of easy task choices (A) and optimizers (B) in Experiments 7a and 7b.

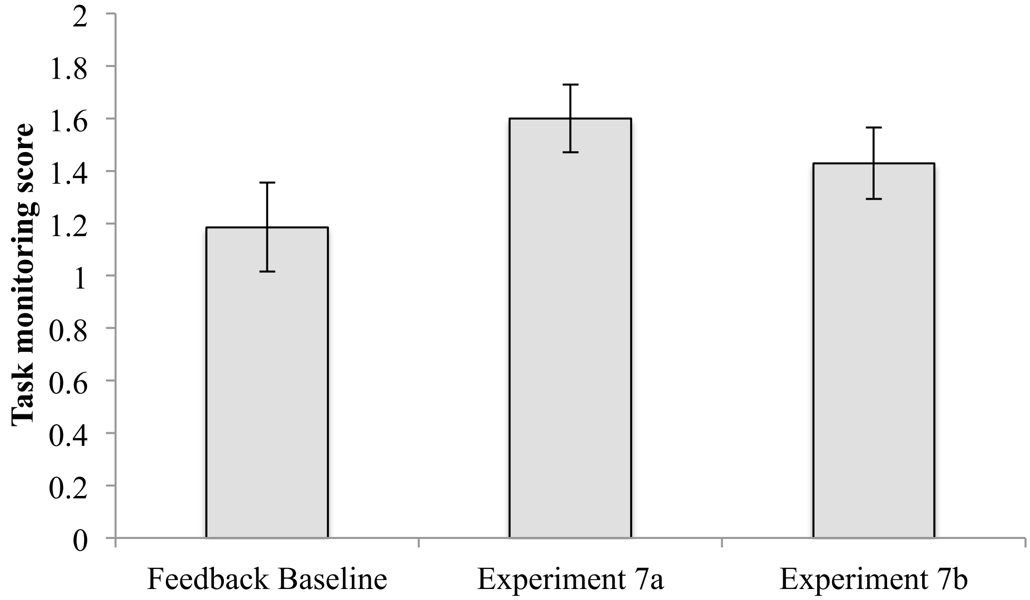


Figure 22. Task monitoring scores in Experiments 7a and 7b.

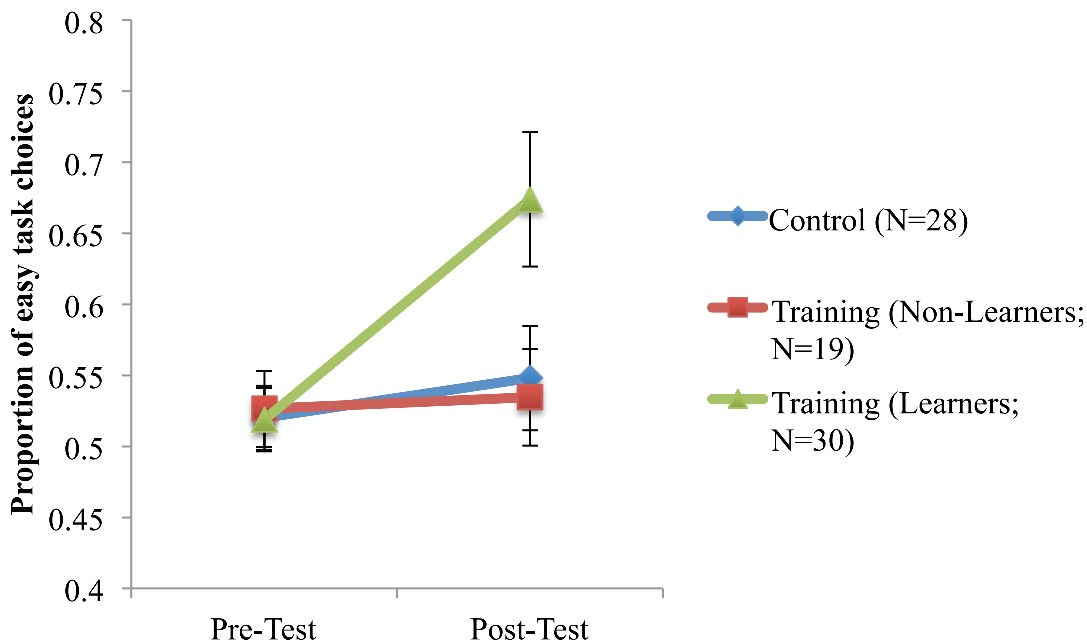


Figure 23. Proportion of easy task choices in the pre- and post-test phases of Experiment 8.