CHILD DEVELOPMENT

Child Development, xxxx 2021, Volume 00, Number 0, Pages 1–18

Developmental Differences in Subjective Recollection and Its Role in Decision Making

Diana Selmeczy, Alireza Kazemi, and Simona Ghetti
University of California, Davis

We examined how subjective assessments of recollection guide decision making. Subjective recollection was dissociated from accuracy during a forced-choice recognition task. Distracters were either similar to targets (match condition) or to other studied, but untested items (nonmatch condition). We assessed 223 participants (112 males) across three experiments (137 White, 37 Asian-American, 7 African-American, 4 American-Indian, 32 mixed race, 6 undisclosed). In Experiment 1, 6- to 10-year-olds and adults (N = 119) were less accurate (d = 0.70), but more likely to claim subjective recollection and make memory selections in anticipation of a reward in the nonmatch condition (ds = 0.64–0.70). This pattern was eliminated in 6- to 7-year-olds when we limited the number of selections (Experiment 2, N = 52), but was replicated when we required the selections to be counted (Experiment 3, N = 52), underscoring the effects of decision complexity on children's self-reflections.

Imagine a teacher asking her students "What is a group of owls called?" after giving a lesson on animals. A student may feel that she knows the answer, but where does this subjective feeling come from? It might be that she recalls something specific about her teacher's previous lesson, such as the teacher telling the class that the answer was a term related to government. Or the student finds owls interesting and thinks she remembers everything that was said about them. In both cases, her subjective experience of remembering motivated her decision to raise her hand and give the right answer (i.e., parliament). In this scenario, the student assessed the quality of her retrieval, a process known as metamemory monitoring, and used the result of this assessment to decide to volunteer an answer, a process known as metamemory control (Nelson & Narens, 1990). These metamemory processes have important consequences throughout development, ranging from supporting children's learning behaviors in the classroom (Bruin & Gog, 2012) to their reliability as eyewitnesses (Roebers & Schneider, 2005). However, relatively little is known about developmental differences in the factors that influence metamemory experiences and how these guide decision making. We are particularly interested in metamemory monitoring of the experience of recollection (i.e., subjective recollection) and its consequences. Recollection refers to the process by which we remember specific details about an event; this process develops substantially in childhood (Ghetti & Angelini, 2008). Recollection is considered to provide a particularly compelling basis for decision making (Hembacher & Ghetti, 2013; Kelley & Sahakyan, 2003), because it is often highly diagnostic of successful memory retrieval (Gardiner, Ramponi, & Richardson-Klavehn, 2002). Thus, we sought to further characterize the development of subjective recollection and its links to decision making. Overall, this research provides new insight into the conditions under which children can or cannot rely on subjective evaluations to make memory-based decisions, joining a growing literature examining the development of decision-making processes.

Support for this research was provided by the National Institute of Child Health and Human Development Grant F32HD084313 to Diana Selmeczy and a Scholar Award from the James S. McDonnell Foundation to Simona Ghetti. We thank Emily Hembacher for her contribution to data collection.

Correspondence concerning this article should be addressed to Diana Selmeczy, Center for Mind and Brain, University of California, Davis, Davis, CA 95616. Electronic mail may be sent to diana.selmeczy@uccs.edu.

Development of Metamemory Monitoring and Control

There is a wealth of evidence demonstrating developmental improvements in both metamemory

 0009-3920/2021/xxxx-xxxx
DOI: 10.1111/cdev.13611

assessments of ongoing memory processes (i.e., monitoring) and their influence on decision making (i.e., control). Metamemory monitoring is evident by around preschool age (Balcomb & Gerken, 2008; Hembacher & Ghetti, 2014), but this ability continues to improve throughout childhood and adolescence (Fandakova et al., 2017). Additionally, theories of metamemory posit that individuals rely on multiple cues that contribute to metamemory monitoring (e.g., ease of processing (Koriat & Ma'ayan, 2005), accessibility (Dunlosky & Nelson, 1992)) independent of objective memory strength (Koriat, 1997), and research with children demonstrates developmental improvements in children's use of these types of cues in metamemory judgments (Koriat & Ackerman, 2010; Koriat, Ackerman, Lockl, & Schneider, 2009). When it comes specifically to subjective recollection, by 6 years of age children retrieve specific details about their past experiences more accurately when they report subjective recollection compared to when they claim to experience familiarity (Ghetti, Mirandola, Angelini, Cornoldi, & Ciaramelli, 2011). This capacity is accompanied by an explicit understanding that the key difference between the experience of recollection and a general sense of familiarity has to do with whether one can retrieve clear and specific details about past events. Despite these remarkable skills, developmental differences are found not only in the calibration between metamemory assessments and objective memory accuracy but also in the type of details children experience recollecting. For example, metamemory assessments of recollection in younger children are more likely to rely on perceptual features than semantic features, whereas the opposite is true in adults (Ghetti et al., 2011; Hembacher & Ghetti, 2013). Overall, research suggests developmental differences in the types of contextual details that children experience as remembering (or consider valuable cues to remembering), but does not elucidate whether children can monitor differences in the retrieval process (e.g., recollection vs. familiarity) in addition to the content of the information retrieved (e.g., precise vs. general details). The present research will probe this question further by requiring children and adults to report experiences of recollection and familiarity under conditions that either do or do not encourage retrieval of precise details.

Turning to the development of metamemory control, current evidence has pointed to it lagging behind that of metamemory monitoring (Destan, Hembacher, Ghetti, & Roebers, 2014), suggesting that children may experience difficulties translating the outcome of monitoring processes into effective decisions and actions. For example, in a study in which 6- to 10-year-olds appropriately claimed subjective recollection when they could accurately report contextual details, only 9- to 10-year-olds were more likely to select which items they wanted volunteered toward a future reward based on this subjective experience (Hembacher & Ghetti, 2013; see also Koriat, Goldsmith, Schneider, & Nakash-Dura, 2001; Roebers, Schmid, & Roderer, 2009). Additionally, Metcalfe and Finn (2013) demonstrated that although third graders were as able as fifth graders to accurately monitor which items they would remember on an upcoming memory test, only fifth graders used this metamemory monitoring assessment to guide effective study strategies. Intriguingly, however, under some circumstances, even preschoolers appear to use monitoring outputs to regulate their memory decisions. For example, preschoolers successfully sorted memory responses on the basis of confidence when the memory task involved a simple forced-choice response between two options and they could sort each of their responses in a reward or discard box (Hembacher & Ghetti, 2014), as opposed to select a subset of responses to be submitted to a reward (Hembacher & Ghetti, 2013). Thus, one possible factor contributing to the protracted developmental trajectory of metamemory control may be related to the complexity of decision demands.

The Role of Decision Complexity in Metamemory

Metacognition likely plays an important role in children's and adults' responses to complex decision demands. For example, adults with higher metacognitive awareness appropriately choose tasks with lower compared to higher decision demands in order to optimize performance (Desender, Calderon, Opstal, & den Bussche, 2017). Moreover, metacognition may determine when and how efficiently children engage in demanding proactive strategies (Chevalier & Blaye, 2016; Chevalier, Martis, Curran, & Munakata, 2015), highlighting the link between developmental improvements in metacognition and executive function (Roebers, 2017).

There is initial evidence that these links extend to memory performance. Decision demands may constrain whether a relation between metamemory monitoring and control is observed, such that increasing demands may interfere with children's ability to use the output of metacognitive monitoring processes to engage in metacognitive control over memory decisions. For example, 6-year-olds experience accuracy gains when incentivized to focus on individual aspects of memory decisions (e.g., receive a token for each correct answer; Roebers & Schneider, 2005), but they struggle to achieve similar gains when the reward structure is less apparent at the individual item level (e.g., cross out potentially inaccurate responses after taking a test; Roebers et al., 2009). Although this suggests that increasing the complexity of the decision-making process may impact the extent to which children engage in metamemory control, we also consider the possibility that increased decision demands may only alter the extent to which metamemory monitoring processes produce reliable information in the first place. Indeed, the ability to rely on experience-based cues during metamemory monitoring judgments (e.g., retrieval time, Koriat & Ackerman, 2010) improves during childhood, and added decision demands may disrupt younger children's sensitivity to these types of cues.

Previous research has not examined whether and how decision demands impact the emergence or developmental differences of subjective recollection. The present research manipulates decision complexity to investigate the extent to which children's metamemory monitoring and control are affected. In the current task, decision complexity was increased by requiring participants to select a limited number of responses to be submitted for a reward (as opposed to letting participants select as many responses as they pleased with no limitations). Critically, the interpretability of any finding linking metamemory monitoring and decision-making rests on the ability of these assessments to capture unique information that is not already carried by other indicators of performance. However, subjective recollection and objective memory accuracy are typically correlated, making it difficult to establish a causal role of subjective experiences for decision making. This fact makes it imperative to experimentally dissociate subjective and objective memory measures in order to examine the unique contribution of metamemory.

Dissociating Memory From Metamemory

Subjective memory assessments and memory accuracy are typically correlated. It may therefore be difficult to rule out the possibility that subjective assessments are merely epiphenomenal (Dijksterhuis & Aarts, 2010) and establish, instead, that they contribute to decision making on their own. Indeed, there is evidence that certain types of decisions occur without explicit intentions or awareness (Gaal, de Lange, & Cohen, 2012) and that neural measures can indicate decisions well before conscious awareness of intentions (Soon, Brass, Heinze, & Haynes, 2008). To isolate metamemory ability from memory ability, researchers have developed metrics that account for the effects of the latter (Fleming & Lau, 2014). A complementary approach is to experimentally dissociate metamemory assessments from objective memory performance (Dobbins, Kroll, & Liu, 1998; Metcalfe & Finn, 2008) to determine which aspect contributes to decision making. Recent work in adults demonstrated that decision making followed subjective assessments above and beyond memory accuracy (Desender, Boldt, & Yeung, 2018; Fandakova, Johnson, & Ghetti, 2021; Hembacher & Ghetti, 2016), but whether this is also true for children is currently unknown. Critically, since developmental improvements in metamemory co-occur with improvements in objective memory performance (Ghetti & Fandakova, 2020), it is particularly important to dissociate these processes in order to better assess the unique contribution of metamemory processes.

The Present Research

This research was designed to investigate developmental differences in subjective recollection (i.e., our indicator of metacognitive monitoring) between 6- to 7-year-olds, 9- to 10-year-olds, and adults, and its role in guiding decisions to sort memory responses into those to be selected toward a future reward and those to be discarded (i.e., our indicator of metacognitive control). Previous research examining developmental differences in subjective recollection has not examined how encouraging retrieval of diagnostic features versus global information affects subjective recollection. Furthermore, previous studies employed paradigms in which subjective assessments and memory performance were correlated (Destan et al., 2014; Hembacher & Ghetti, 2013). Finally, the few studies conducted so far varied in decision demands across studies (Hembacher & Ghetti, 2013, 2014; Metcalfe & Finn, 2013), which may in part account for discrepancies in patterns of results. The current experiments addressed these issues by examining developmental differences in (a) subjective recollection as a function of whether or not the task encouraged retrieval of diagnostic features; (b) the role of subjective recollection in guiding sorting decisions above and beyond memory accuracy, and (c) role of decision demands in altering subjective recollection and sorting decisions.

To address these questions, we adopted a paradigm previously used in adults in which subjective recollection and accuracy were dissociated (Hembacher & Ghetti, 2016). The task involved encoding a series of images of concrete objects. Memory was tested with a forced-choice recognition test, in which the perceptual similarity of distractors was manipulated to create a match condition (i.e., the distractor was a different exemplar of the target, resulting in a high similarity between the test probes) and nonmatch condition (i.e., the distractor was a different exemplar of another studied, but untested item, resulting in low similarity between the test probes; Figure 1). Compared to the nonmatch condition, the match condition encourages participants to engage in a more detail-oriented retrieval process in order to identify the most diagnostic features to differentiate studied items from distracters. Therefore, we expected objective memory performance to be higher for match compared to nonmatch trials (Hembacher & Ghetti, 2016), and we anticipated a similar size effect in all age groups since young children's and adults' memory accuracy has been shown to respond similarly to manipulations of perceptual similarity (Ngo, Newcombe, & Olson, 2017). Additionally, we predicted normative developmental improvements in overall objective memory accuracy (Ghetti & Lee, 2010). In contrast, we expected subjective recollection to be lower for the match compared to nonmatch condition. This is because participants may consider that they did not identify the most diagnostic feature. This consideration may discourage claims of subjective recollection which are typically based on more global retrieval experiences (Dobbins et al., 1998; Hembacher & Ghetti, 2016). The high level of performance in the match condition suggests that participants recall the key diagnostic details and thus it is unlikely that lower subjective recollection rates simply reflect a general reliance on familiarity in this condition. Conversely, the nonmatch condition encourages a more global assessment of the identity of the test probes, resulting in more errors, but also a stronger sense of subjective recollection if, for example, the memory for an apple is stronger than that of a butterfly (regardless of the specific exemplar shown in the test). In this condition, greater familiarity for one of the members of the pair (e.g., the apple generates more familiarity than the butterfly) could lead to the selection of that item; yet subjective recollection is the more frequent experience (see also Hembacher & Ghetti, 2016).

We examined subjective recollection using the remember-familiar paradigm, which is thought to capture distinctions in experiences of recollection (i.e., retrieval of contextual information) versus familiarity (i.e., general feeling of oldness in the absence of contextual details; Gardiner et al., 2002). We chose to focus on subjective recollection because the current task required children to recall specific diagnostic features in order to identify a target, and this process requires recollection (Gardiner et al., 2002). Confidence ratings are another measure of metamemory monitoring typically used in children (Destan et al., 2014; Hembacher & Ghetti, 2013). Although high confidence responses are correlated with the retrieval of specific details (Selmeczy & Dobbins, 2014), high confidence judgments may also reflect strong familiarity (Ingram, Mickes, & Wixted, 2012), making them less appropriate for capturing subjective experiences associated with specific details. Children as young as 6-year-old can reliably report subjective experiences of recollection (Ghetti et al., 2011), and therefore we predicted that all age groups would exhibit lower remember reports for match versus nonmatch conditions. However, given children's tendency to overclaim recollection (Ghetti et al., 2011; Rollins & Riggins, 2017), we expected children to report recollection more frequently than adults. Children's tendency to overclaim recollection compared to adults may be particularly strong for match trials because children may attend less to the retrieval of precise details and there is more room to demonstrate increases when compared with adults. Overall, despite a predicted attenuation in younger children, we anticipated that all age groups would show a dissociation between subjective recollection and objective memory performance, allowing us to examine the role of subjective recollection on decision making.

We also probed decision making, or metacognitive control (Nelson & Narens, 1990), by having participants sort which answers they wanted counted toward their final score, similar to previous developmental work (Destan et al., 2014; Hembacher & Ghetti, 2013). After each trial, participants were asked to sort their answer into a treasure box —if they thought it ought to be counted toward a reward, and into a trash can—if they thought it ought to be discarded. We considered two alternative developmental hypotheses about the role of subjective recollection on decision making. If metacognitive monitoring lags behind control processes (Destan et al., 2014; Schneider & Lockl, 2008), especially when the metamnemonic assessment is more sophisticated as is the case for subjective recollection (Hembacher & Ghetti, 2013), we would predict that subjective recollection guides sorting decisions in older children and adults, but not in younger children. This would be demonstrated by higher treasure rates for nonmatch compared to match conditions in older children and adults, mirroring the pattern observed for remember rates, but no such distinction in younger children. Alternatively, if making a decision on every trial (instead of selecting subsets of trials) facilitates children's translating the results of their assessments of subjective recollection into decisions to retain or discard responses (Hembacher & Ghetti, 2014), then even younger children may select answers to be rewarded on the basis of their subjective recollection. These predictions were examined in Experiment 1.

Experiment 1

The goal of Experiment 1 was to examine developmental differences in the dissociation between subjective recollection and memory accuracy, and the extent to which sorting decisions followed the pattern of subjective recollection or memory accuracy. We included children ages 6- to 7-year-old and 9- to 10-year-old since subjective recollection is reliably observed as early as 6 years of age but continues to improve throughout middle childhood (Ghetti et al., 2011). Young adults were also included as a comparison group and to replicate previous findings using the current paradigm (Hembacher & Ghetti, 2016).

Method

Participants

A total of 119 participants across three age groups were examined including forty-one 6- and 7-year-olds ($N = 20$ males, age: $M = 7.07$, $SD = 0.61$), thirty-eight 9- and 10-year-olds ($N = 21$ males, age: $M = 9.98$, $SD = 0.64$), and 40 young adults ($N = 19$ males, age: $M = 20.79$, $SD = 1.18$). Our sample size is sufficient to detect a 3 (age group: 6- to 7-year-olds, vs. 9- to 10-year-olds vs. adults) × 2 (trial type: match vs. nonmatch) interaction of medium to small effect size ($f = .14$; $\eta_p^2 = .02$) with 80% power (Faul, Erdfelder, Lang, & Buchner, 2007).

Participants' race was distributed as White ($N = 73$), Asian ($N = 17$), African American ($N = 2$), American Indian ($N = 3$), other ($N = 3$), mixed race ($N = 18$), and not reported ($N = 3$). Twenty-six participants reported being Hispanic or Latino. Participant's family income was distributed as income < \$15,000 ($N = 3$), between \$15,000 and \$25,000 ($N = 2$), between \$25,000 and \$40,000 ($N = 7$), between \$40,000 and \$60,000 ($N = 12$), between \$60,000 and \$90,000 ($N = 31$), more than \$90,000 ($N = 59$), and not reported ($N = 5$). Participant demographics conformed with those of the local community. Data from an additional three participants (adults: $N = 1$; 6-year-olds: $N = 2$) were collected and removed due to below chance memory accuracy (<50% percent correct). Young adults included students recruited through the University of California, Davis experiment pool and were compensated with course credit for their participation. Families were recruited through flyers in the community and given monetary compensation for their participation. All participants provided informed consent in accordance with the University of California, Davis institutional review board.

Materials

We used 160 colorful images of common items selected from Yassa et al., 2011. The set contained 80 pairs of unique items that were semantically and perceptually similar (i.e., distinct versions of the same object). We created four blocks of 20 pairs for counter balancing purposes. Blocks were counterbalanced across status (target vs. distracter) and similarity (similar vs. dissimilar) to create four testing orders. The same materials were used across all experiments.

Procedure

Encoding task. Participants completed an incidental encoding task during which they made indoor–outdoor judgments. The task began with a short practice phase followed by 80 study items presented sequentially for 2,000 ms each (see Figure 1A).

Retrieval task. The key experimental, within subject manipulation was delivered during retrieval. The retrieval task consisted of 40 two-alternative forced choice trials, including 20 match and 20 nonmatch trials. Match trials consisted of a previously studied target (A) and a different exemplar of the same object depicted in the target picture (A′). Nonmatch trials consisted of a previously studied target (A) and a different exemplar of another picture presented during encoding (B′). The target picture that corresponded with nonmatch distractors (B) was never tested. Participants

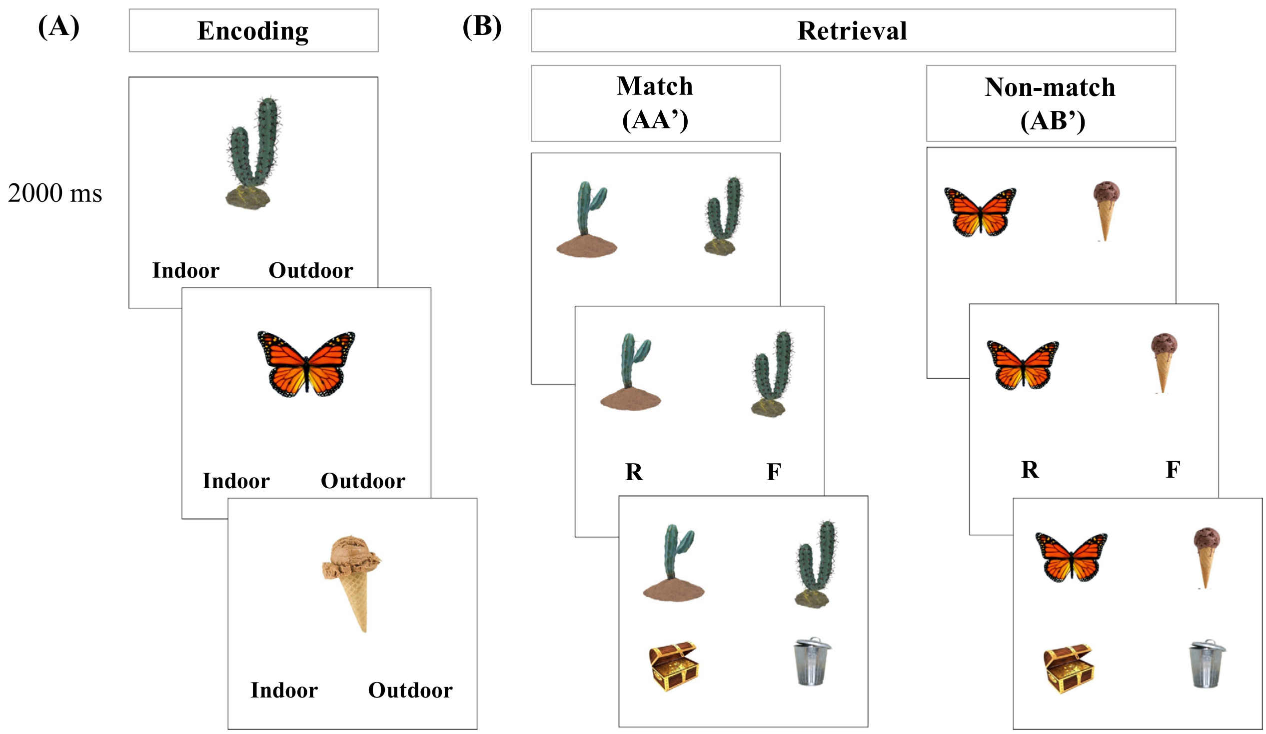

Figure 1. Experimental design. During encoding, participants completed indoor–outdoor judgments on pictures presented for 2,000 ms. During retrieval, participants completed a two-alternative forced choice recognition test consisting of match trials (target (A) and distractor version of the same target picture (A′)) and nonmatch trials (target (A) and distractor version of another picture presented during encoding (B′)). After their memory judgment, participants indicated whether the picture was experienced as remembered or familiar. Finally, participants submitted their sorting decisions by indicating whether to count their answer toward their final score by choosing a treasure box for volunteered responses and a trash can for withheld responses. The location of the treasure box and trash can was counterbalanced across participants.

were instructed to identify the target out of the two images presented. Then, participants completed a metamemory assessment, namely they provided a Remember versus Familiar judgment. Specifically, participants were instructed to select Remember if they could think back to when they first saw the picture and could retrieve certain specific aspects about the picture or what the picture made them think of. Participants were told to select a Familiar response if they could tell that they had seen the picture before but could not think back to when they first saw it or could not come up with any specific details about it (see Supporting Information for full instructions). After the Remember-Familiar judgment, participants made a sorting decision. Participants were instructed to place an answer in the treasure box option if they wanted that answer to be counted toward a reward or to place the answer in the trash can if they did not want that answer to be counted (see Figure 1B). To help account for the possibility of participants automatically mapping a Remember response to a treasure selection, we counterbalanced the location of these options on the screen between participants.

To motivate participants, adults were told that their score was going to be compared with other participants and their ranking would be displayed on the screen at the end of the task whereas children were told that they will receive a prize based on their final score. Participants completed a short series of practice trials followed by the testing phase. All responses were self-paced and answers were submitted via touch screen monitor by choosing corresponding graphical icons. All participants received the same compensation at the end of the task and final scores were never shown or compared to others at the end of the study.

Results and Discussion

Confirmatory analyses included examining accuracy, remember rates, and treasure rates using a 3 (age group: 6- to 7-year-olds, vs. 9- to 10-year-olds vs. adults) × 2 (trial type: match vs. nonmatch) mixed analysis of variance (ANOVA) with trial type varied within participants. Additional analyses should be considered as exploratory or post hoc.

Accuracy

Accuracy was calculated as the percentage of correctly answered responses. Results revealed a main effect of age group, $F(2, 116) = 4.50$, $p = .01$, $\eta_p^2 = .07$. Overall accuracy was lower for 6- to 7-year-olds ($M = 0.67$, $SD = 0.08$) compared to both 9- to 10-year-olds ($M = 0.72$, $SD = 0.08$, $p = .014$, Cohen's $d = 0.57$, $p_{bonf} = .041$) and adults ($M = 0.72$, $SD = 0.09$, $p = .009$, Cohen's $d = 0.60$, $p_{bonf} = .026$), who did not differ from each other ($p = .816$), revealing typical developmental improvements in overall memory, despite well above chance performance even in the youngest group ($t(40) = 13.42$, $p < .001$). The main effect of trial type was significant $F(1, 116) = 58.33$, $p < .001$, $\eta_p^2 = .34$, such that accuracy was higher for match ($M = 0.76$, $SD = 0.11$) versus nonmatch trials ($M = 0.65$, $SD = 0.13$, $p < .001$, Cohen's $d = 0.70$). The age group by trial type interaction was not significant, $F(2, 116) = 0.64$, $p = .53$, $\eta_p^2 = .01$ (see Figure 2A). Thus, all age groups showed the expected higher performance for match compared to nonmatch trials and this effect did not vary across age groups. Even children as young as 6 benefitted from being confronted with a more constrained retrieval process that encouraged identifying the most diagnostic features of previously viewed items.

Remember Rates

Remember rates were calculated as proportion of trials for which participants selected the Remember option. Results revealed a main effect of age group, $F(2, 116) = 5.58$, $p = .005$, $\eta_p^2 = .09$, such that, as expected, overall Remember rates were higher for 6- to 7-year-olds ($M = 0.77$, $SD = 0.20$, $p = .021$, Cohen's $d = 0.52$, $p_{bonf} = .062$) and 9- to 10-year-olds ($M = 0.80$, $SD = 0.15$, $p < .001$, Cohen's $d =$ 0.81, $p_{bonf} = .002$) compared to adults ($M = 0.68$, $SD = 0.15$). The main effect of trial type was significant $F(1, 116) = 67.56$, $p < .001$, $\eta_p^2 = .37$, such that Remember rates were higher for nonmatch ($M = 0.82$, $SD = 0.17$) compared to match trials ($M = 0.68$, $SD = 0.22$, $p < .001$, Cohen's $d = 0.70$, $p_{bonf} < .001$). Finally, we also found a significant age group by trial type interaction, $F(2, 116) = 5.84$, $p = .004$, $\eta_p^2 = .09$ (see Figure 2B). The interaction resulted from higher Remember rates for match trials in 6- to 7-year-olds ($M = 0.71$, $SD = 0.25$, $p = .008$, Cohen's $d = 0.60$, $p_{bonf} = .024$) and 9- to 10-year-olds ($M = 0.77$, $SD = 0.19$, $p < .001$, Cohen's $d = 1.04$, $p_{bonf} < .001$) compared to adults ($M = 0.58$, $SD = 0.19$). No age differences were observed in Remember rates for nonmatch trials, $F(2, 116) = 1.16$, $p = .32$, $\eta_p^2 = .02$. Thus, all age groups showed higher Remember rates for nonmatch versus match trials (ps < .01, $ps_{bonf} < .03$), but children were more likely to report remembering for match trials than adults.

In summary, children tended to be more liberal in reporting Remember responses than adults for match trials, which is consistent with previous work (Ghetti et al., 2011). This result suggests that children may not as readily consider the identification of specific diagnostic features that distinguish the target from the distractor in their subjective recollection assessment. Critically, all age groups demonstrated a dissociation between subjective recollection and objective performance.

Treasure Rates

Treasure rates were calculated as proportion of trials selected for a reward. There was no main effect of age group, $F(2, 116) = 1.28$, $p = .28$, $\eta_p^2 = .02$, but there was a significant main effect of trial type $F(1, 116) = 49.99$, $p < .001$, $\eta_p^2 = .30$, such that treasure rates were higher for nonmatch ($M = 0.85$, $SD = 0.16$) compared to match trials ($M = 0.73$, $SD = 0.23$, $p < .001$, Cohen's $d = 0.64$). This effect was qualified by a significant age group by trial type interaction, $F(2, 116) = 3.28$, $p = .04$, $\eta_p^2 = .05$ (see Figure 2C). This interaction was due to significantly higher treasure rates for match trials in 6- to 7-year-olds ($M = 0.78$, $SD = 0.24$, $p = .031$, Cohen's $d = 0.49$, $p_{bonf} = .092$) and numerically higher in 9- to 10-year-olds ($M = 0.75$, $SD = 0.24$, $p = .104$, Cohen's $d = 0.37$, $p_{bonf} = .312$) compared to adults ($M = 0.67$, $SD = 0.20$). For nonmatch trials, treasure rates did not differ between age groups, $F(2, 116) = 0.26$, $p = .77$, $\eta_p^2 = .00$. Importantly, each age group demonstrated significantly higher treasure rates for nonmatch versus match trials, (ps < .003, $ps_{bonf} < .010$), and this was also true in the youngest age group, $p = .001$, $p_{bonf} = .004$.

One could argue that sorting decisions showed the same pattern as subjective recollection because participants may have treated these two responses as the same judgment. Although the instructions did not draw any explicit connections between the two judgments, participants may have adopted the strategy to always choose treasure after a Remember response and trash after a Familiar response. If this were the case, we would expect all Remember responses and no Familiar responses to be treasured. Figures 2B and 2C show that Remember and treasure rates were not identical, making it unlikely that participants adopted this strategy. Additionally, we directly assessed the correspondence between subjective experience and treasuring rates separately for Remember and Familiar responses. Treasure rates were analyzed using a 3 (age group: 6- to 7-year-olds, vs. 9- to 10-year-olds vs. adults) × 2 (response type: remember vs. familiar) mixed ANOVA with response type varied within participants. We collapsed across trial type in order to increase the number of trials and avoid case-wise deletions for participants who did not report all combination of responses. For the current analysis, 11 participants (6- to 7-year-olds: $N = 4$; 9- to 10-year-olds: $N = 6$; adults: $N = 1$) were removed due to never treasuring a Familiar response. Results revealed a main effect of response type $F(2, 105) = 263.24$, $p < .001$, $\eta_p^2 = .72$, such that participants were more likely to treasure a Remember response ($M = 0.93$, $SD = 0.11$) compared to a Familiar response ($M = 0.36$, $SD = 0.36$, Cohen's $d = 1.56$). No other effects were significant, *p*s > .153. Critically, although treasure rates for Remember responses were high in all groups, (6- to 7-year-olds $M = 0.93$, $SD = 0.12$; 9- to 10-year-olds $M = 0.92$, $SD = 0.12$; adults $M = 0.95$, $SD = 0.07$), they were significantly lower than 100% (*p*s < .001). Furthermore, the treasure rates for Familiar responses were significantly above 0% in all age groups (6- to 7-year-olds $M = 0.43$, $SD = 0.43$; 9- to 10-year-olds $M = 0.35$, $SD = 0.35$; adults $M = 0.29$, $SD = 0.30$, *p*s < .001). Thus, participants did not merely adopt a matching strategy such that they treasured Remember responses and trashed Familiar responses.

Finally, we conducted an additional follow-up analysis in order to examine the simultaneous contribution of accuracy and subjective recollection to participants' treasure selection decision at the individual item-level using a multi-level logistic model. This analysis allowed us to further investigate the role of subjective recollection above and beyond memory accuracy in participants' individual treasure decisions. Treasure decisions (Treasure = 1; Trash = 0) were predicted using fixed effects of accuracy (Correct = 1, Incorrect = 0), Remember response (Remember = 1, Familiar = 0), and trial type (1—match, 0—nonmatch), and subject was included as a random intercept (see Table S1). Results revealed that accuracy significantly predicted treasure decisions in 6- and 7-year-olds and 9-and 10-year-olds (6- and 7-year-olds: $B = 0.95$, $SE = .20$, OR = 2.59, $p < .001$; 9- and 10-year-olds: $B = 0.84$, $SE = .20$, OR = 2.31, $p < .001$) and marginally predicted treasure decisions in adults ($B = 0.35$, $SE = .20$, OR = 1.42, $p = .09$). Critically, remember responses predicted treasure decisions above and beyond accuracy and trial type across all age groups (6- and 7-year-olds: $B = 3.52$, $SE = .23$, OR = 33.72, $p < .001$; 9- and 10-year-olds: $B = 3.65$, $SE = .24$, OR = 38.50, $p < .001$; Adults: $B = 4.82$, $SE = .28$, OR = 123.81, $p < .001$), and the Remember response regression weights were much larger than those of accuracy. Overall, these analyses are consistent with the previously reported treasure rate results in suggesting that subjective recollection is a significant contributor to decisions to select or discard test probes, and furthermore, we demonstrate that children and adults may rely more heavily on subjective experiences compared to memory accuracy.

Overall, we observed that treasure decisions followed the same general pattern as subjective experiences for all age groups. This result is consistent with previous work demonstrating that young children's decision making is guided by subjective experiences measured via confidence ratings (Hembacher & Ghetti, 2014), and here we demonstrate that this is also present when using subjective recollection ratings. However, previous work by Hembacher and Ghetti (2013) failed to find this relation in 6- to 7-year-olds; in that study, however, children were instructed to select the treasure response option on a subset of the trials (i.e., 30 times out of 96 trials). In other words, they did not only have to assess which items were best remembered in absolute terms, but also attempt a relative comparison across trials to determine the worthiest memories or, at a minimum, children ought to withhold selections in order to have enough treasure responses available across the entire task in order to maximize performance. If the request to select a limited number of responses results in more complex decision demands, this could interfere with children's ability to use their subjective experiences effectively during their decision making (Chevalier et al., 2015). Thus, the introduction of the same request in this paradigm may reproduce the pattern of results observed in Hembacher and Ghetti (2013). In Experiment 2, we tested this hypothesis by limiting the number of times participants could select the treasure response option.

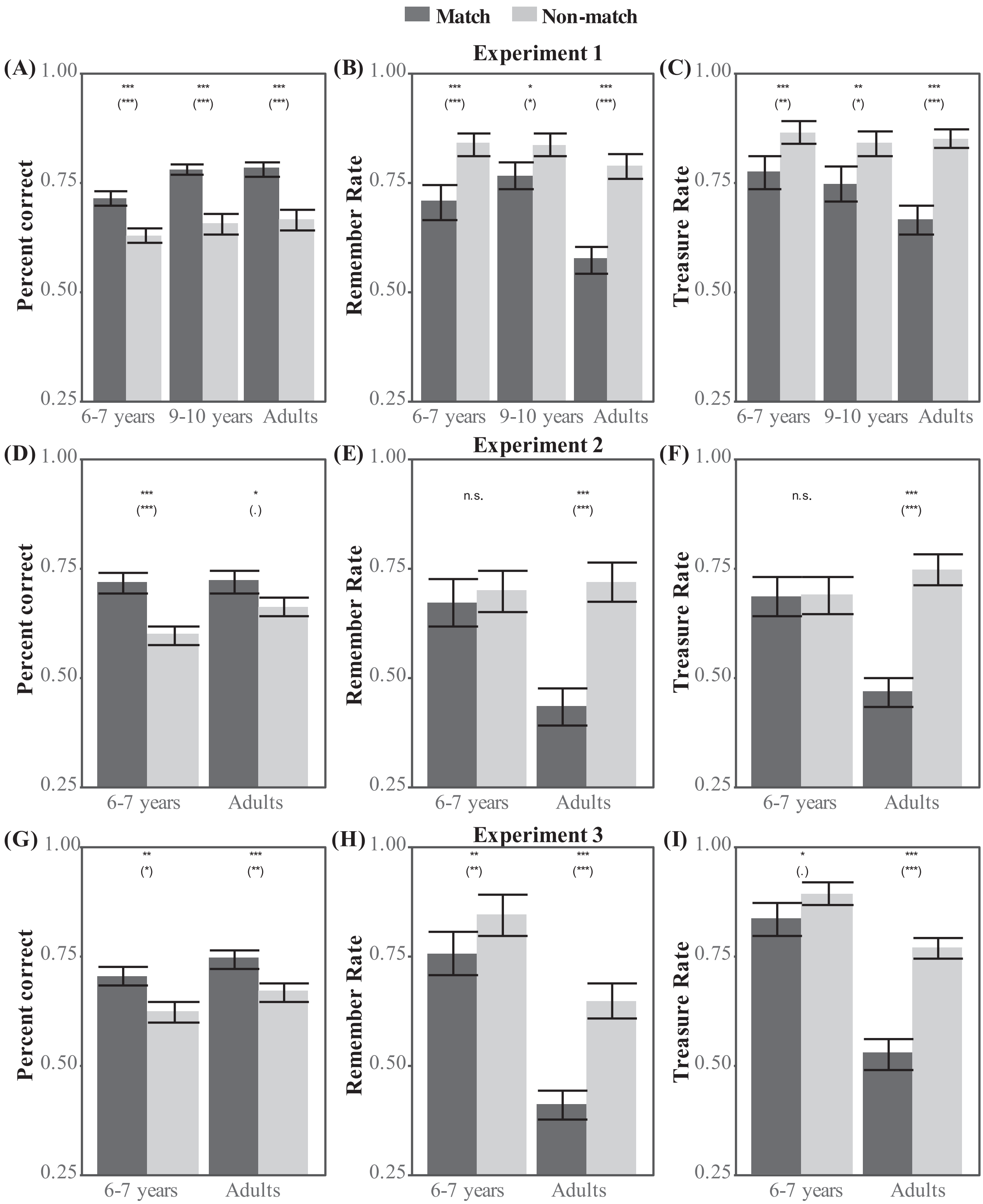

Figure 2. Accuracy, remember rates, and treasure rates as a function or age and retrieval condition. Results are reported for Experiment 1 (A–C), Experiment 2 (D–F), and Experiment 3 (G–I). Error bars represent ± 1 *SE* around the mean. Pairwise and Bonferroni corrected comparisons (in parenthesis) are reported as (.)$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. *ns*: not significant.

Experiment 2

The goal of Experiment 2 was to examine the role of increased decision demands on metamemory monitoring and control by limiting the number of answers participants could treasure and select toward maximizing their performance. By limiting the number of items that could be selected toward a future reward, we expected 6- to 7-year-olds to have increased difficulty with their sorting decision. This would be demonstrated by young children exhibiting similar treasure rates for nonmatch versus match trials, despite their higher remember rates for nonmatch trials. Alternatively, it is possible that the added decision demands may impact both subjective recollection and treasure decisions. This would occur if added decision complexity taxes children's ability to engage in metamemory processes in general, including both monitoring and control. This prediction would be consistent with research suggesting that engaging in metacognition is resource demanding for older adults (Stine-Morrow, Shake, Miles, & Noh, 2006) and added cognitive demands decreases metacognitive performance in young adults (Schwartz, 2008). Additionally, executive functioning has been correlated with both metacognitive monitoring (Chevalier & Blaye, 2016) and control (Roebers, Cimeli, Röthlisberger, & Neuenschwander, 2012) in children, making it possible that both would be influenced by increased decision demands.

Finally, to verify that participants' conceptual understanding of subjective recollection was not responsible for the developmental differences observed in Experiment 1, we included a short questionnaire asking participants to classify descriptions of Remember versus Familiar reports. Since we were particularly interested in younger children's performance in Experiment 1 and previous research (Hembacher & Ghetti, 2013), Experiment 2 included children ages 6- to 7-year-olds and young adults as a comparison group.

Method

Participants

Experiment 2 included a total of 52 participants across two age groups including twenty-six 6- and 7-year-olds ($N = 13$ males, age: $M = 7.08$, $SD = 0.49$) and 26 young adults ($N = 13$ males, age: $M = 20.46$, $SD = 1.66$). This sample size is sufficient to detect a 2 (age group: 6- to 7-year-olds vs. adults) × 2 (trial type: match vs. nonmatch) interaction of medium effect size ($f = .20$; $\eta^2_p = .04$) with 80% power (Faul et al., 2007).

Participants' race was distributed as White ($N = 34$), Asian ($N = 10$), African American ($N = 3$), other ($N = 2$), mixed race ($N = 2$), and not reported ($N = 1$). Nine participants reported being Hispanic or Latino. Family reported income was distributed as <\$15,000 ($N = 1$), between \$15,000 and \$25,000 ($N = 6$), between \$25,000 and \$40,000 ($N = 1$), between \$40,000 and \$60,000 ($N = 6$), between \$60,000 and \$90,000 ($N = 9$), more than \$90,000 ($N = 27$), and not reported ($N = 2$). Participant demographics conformed with those of the local community. Data from an additional two participants (6- and 7-year-olds: $N = 2$) were collected and removed due to below chance accuracy (<50% percent correct based on all trials). Young adults included students recruited through the University of California, Davis experiment pool and were compensated with course credit for their participation. Families were recruited through flyers in the community area and given monetary compensation for their participation. All participants provided informed consent in accordance with the University of California, Davis institutional review board.

Procedure

The procedure was identical to Experiment 1, with the exception that participants were only allowed to select the treasure response option 20 times out of 40 total trials. In Experiment 1, the treasure response option was selected approximately 33 times for 6- and 7-year-olds and 30 times for adults. Thus, restricting the treasure response option to 20 trials was expected to increase the decision demands for both age groups. In order to track the number of times the treasure response was selected, a clear bowl containing 40 gold coins, corresponding to each memory trial, was placed between a small treasure box and trash container. After each trial, the experimenter put a gold coin in the treasure box or trash can corresponding to the participant's sorting decision. After every 10 trials, the experiment would count the number of coins in the treasure box out loud and inform participants of how many items they selected and how many items they have left to select as a treasure response. Once 20 treasure trials occurred, the experimenter would move the treasure box and tell participants they could no longer pick the treasure response option.

After the memory task, participants completed a Remember-Familiar questionnaire consisting of eight statements (four remember, four familiar responses; see Supporting Information). Participants were read the statements by the experimenter and asked to classify each statement as either a Remember or Familiar response.

Results and Discussion

Confirmatory analyses of accuracy, recollection rates, and treasure rates were examined using a 2 (age group: 6- to 7-year-olds, vs. adults) × 2 (trial type: match vs. nonmatch) mixed ANOVA with trial type varied within participants. Additional analyses should be considered as exploratory or post hoc.

Preliminary Analyses

Trial selection. Since participants could only select 20 trials toward a treasure response, treasure responses after the 20^{th} treasure trial selection were forced to be trashed. Therefore, analyses were conducted using only trials that occurred before the 20th treasure response. The average number of trials completed before the 20th treasure trial selection was $M = 33.65$ ($SD = 6.34$) trials for adults and $M = 30.27$ ($SD = 7.30$) trials for children. Nevertheless, we verified that all of our results held when we analyzed the full set of test trials.

Results

Accuracy

Results replicated Experiment 1 and revealed a main effect of trial type, $F(1, 50) = 21.58$, $p < .001$, $\eta_p^2 = .30$, such that accuracy was higher for match ($M = 0.72$, $SD = 0.13$) versus nonmatch trials ($M = 0.63$, $SD = 0.12$, Cohen's $d = 0.64$). The main effect of age group and the age group by trial type interaction was not significant, ps > .13 (see Figure 2D). Thus, all age groups showed the expected pattern with higher performance for match compared to mon-match trials and this effect did not vary across age groups.

Remember Rates

As in Experiment 1, we found a significant age group by trial type interaction, $F(1, 50) = 16.39$, $p < .001$, $\eta_p^2 = .25$ (see Figure 2E). However, in contrast to Experiment 1, for 6- and 7-year-olds Remember rates did not differ between conditions (nonmatch: $M = 0.70$, $SD = 0.25$, match: $M = 0.67$, $SD = 0.28$, $p = .62$). For adults, Remember rates were significantly higher for nonmatch ($M = 0.72$, $SD = 0.24$) compared to match trials ($M = 0.43$, $SD = 0.22$, $p < .001$, Cohen's $d = 1.56$, $p_{bonf} < .001$), replicating Experiment 1 findings. Specifically, this interaction was driven by higher Remember rates in 6- to 7-year-olds compared to adults for match trials $t(50) = 3.38$, $p = .001$, Cohen's $d = 0.94$, $p_{bonf} = .003$, but not nonmatch trials, $t(50) = 0.30$, $p = .766$. These results suggest that added decision demands altered children's ability to experience or report subjective recollection such that their Remember rates did not differentiate between trial types. To examine whether the introduction of decision demands eliminated children's ability to engage effectively in any form of metamemory process, we examined whether children could monitor the accuracy of their memories regardless of trial type. Children's Remember rates were significantly higher for correct ($M = 0.72$, $SD = 0.22$) relative to incorrect responses ($M = 0.61$, $SD = 0.28$, $p = .001$, Cohen's $d = 0.71$; see Table S4 for additional comparisons). Thus, young children's ability to monitor differences between specific aspects of their retrieval process (i.e., making a fine-grained comparison or not) was eliminated with added demands; however, their ability to monitor general accuracy of their memories persisted.

Treasure Rates

We found a significant age group by trial type interaction, $F(1, 50) = 21.17$, $p < .001$, $\eta_p^2 = .30$, such that treasure rates were significantly higher for nonmatch ($M = 0.75$, $SD = 0.19$) compared to match trials ($M = 0.47$, $SD = 0.16$, $p < .001$, Cohen's $d = 1.40$, $p_{bonf} < .001$) in adults, but not 6- and 7-year-olds (nonmatch: $M = 0.69$, $SD = 0.22$, match: $M = 0.69$, $SD = 0.23$, $p = .95$; see Figure 2F). Thus, increased decision demands influenced children's treasure reports, such that they were no longer sensitive to differences in trial type. Specifically, this interaction was driven by higher treasure rates in 6- to 7-year-olds compared to adults for match trials $t(50) = 4.02$, $p < .001$, Cohen's $d = 1.12$, $p_{bonf} < .001$, but not nonmatch trials, $t(50) = 1.04$, $p = .301$. However, we confirmed that young children's treasure rates were still sensitive to memory accuracy such that treasure rates were significantly higher for correct ($M = 0.74$, $SD = 0.18$) relative to incorrect responses ($M = 0.59$, $SD = 0.24$, $p < .001$, Cohen's $d = 0.87$; see Table S5 for additional comparisons). Added decision demands also did not alter children's ability to select memories for reward when they reported recollection (see Supporting Information Results) and Remember rates predicted trial-level treasure decisions above and beyond memory accuracy (see Table S2). Taken together, these findings suggest that added decision demands did not eliminate children's metamemory monitoring and control over accuracy, but it did alter their ability to effectively monitor and control more fine-grained aspects of the retrieval process.

Remember-Familiar Comprehension

We compared the overall accuracy of categorizing Remember and Familiar descriptions to verify participant's understanding of these experiences. Results revealed that 6- to 7-year-olds did not exhibit significantly lower comprehension ($M = 0.68$, $SD = 0.19$) compared to adults ($M = 0.73$, $SD = 0.26$, $p = .50$). Additionally, comprehension accuracy did not correlate with overall performance or Remember rates for either children (ps > .25) or adults (ps > .35). Therefore, these results suggest that developmental and individual differences in comprehension accuracy are unlikely to account for the results presented earlier.

Overall, Experiment 2 demonstrated that increasing decision demands influenced children's metamemory processes evaluating retrieval of diagnostic details. However, in Experiment 2 participants were told how many trials they had selected toward a reward, which did not occur in Experiment 1. This procedure was introduced so that participants could know how many trials they could still select without keeping count themselves. However, this procedure may have had the unintended consequence of adding a cognitive burden involved in processing the count on every 10th trial. In order to rule out this potential confound, in Experiment 3 participants could once again treasure an *unlimited* number of times, as in Experiment 1, but they were informed of how many trials they had treasured throughout the task, as in Experiment 2.

Experiment 3

The goal of Experiment 3 was to rule out the possibility that that act of counting treasure responses in itself was responsible for our findings in Experiment 2. In Experiment 3, participants were told the number of trials they had treasured throughout the task as in Experiment 2, but they were *not limited* in the number of trials they could treasure and therefore decision complexity should be simpler and comparable to Experiment 1. We predicted that without the increased demand of selectivity, we would observe the same pattern of results as in Experiment 1. Thus, we anticipated that trial type would influence both subjective recollection and treasure decisions in all age groups. We included children ages 6- to 7-year-olds and young adults as a comparison group.

Method

Participants

Experiment 3 included a total of 52 participants across two age groups including twenty-six 6- and 7-year-olds ($N = 13$ males, age: $M = 7.05$, $SD = 0.50$) and 26 young adults ($N = 13$ males, age: $M = 20.59$, $SD = 1.97$). Participants' race was distributed as White ($N = 30$), Asian ($N = 10$), African American ($N = 2$), American Indian ($N = 1$), other ($N = 4$), mixed race ($N = 3$), and not reported ($N = 2$). Eighteen participants reported being Hispanic or Latino. Family income was distributed as less than between \$15,000 and \$25,000 ($N = 2$), between \$25,000 and \$40,000 ($N = 6$), between \$40,000 and \$60,000 ($N = 6$), between \$60,000 and \$90,000 ($N = 12$), more than \$90,000 ($N = 25$), and not reported ($N = 1$). Participant demographics conformed with those of the local community. Data from an additional two participants (adults: $N = 1$; 6- to 7-year-olds: $N = 1$) were collected and removed due to below chance accuracy (< 50% percent correct). Young adults included students recruited through the University of California, Davis experiment pool and were compensated with course credit for their participation. Families were recruited through flyers in the Davis, CA area and given monetary compensation for their participation. All participants provided informed consent in accordance with the University of California, Davis institutional review board.

Procedure

The procedure was identical to Experiment 1, in that participants were not restricted in the number of times they could select the treasure response. Similar to Experiment 2, however, participants received feedback after every 10 trials regarding the number of answers they chose as a treasure response and were told that the number of responses would be counted.

Results and Discussion

Confirmatory analyses of accuracy, recollection rates, and treasure rates were examined using a 2 (age group: 6- to 7-year-olds, vs. adults) × 2 (trial type: match vs. nonmatch) mixed ANOVA with trial type varied within participants. Additional reported analyses should be considered as exploratory or post hoc analyses.

Accuracy

Replicating Experiments 1 and 2, results revealed a main effect of trial type, $F(1, 50) = 20.70$, $p < .001$, $\eta^2_p = .29$ such that accuracy was higher for match ($M = 0.72$, $SD = 0.12$) compared to nonmatch trials ($M = 0.65$, $SD = 0.11$, Cohen's $d = 0.64$). The main effect of age group and the age group by trial type interaction was not significant, $ps > .12$ (see Figure 2G). Thus, all age groups showed the expected pattern with higher performance for match compared to nonmatch trials and this effect did not vary across age groups.

Remember Rates

The pattern of results was similar to those observed in Experiment 1. The age group by trial type interaction was significant, $F(1, 50) = 9.07$, $p = .004$, $\eta^2_p = .15$, such that difference in Remember rates between nonmatch compared to match trials was significantly smaller in 6- to 7-year-olds ($M_{diff} = 0.09$, $SD_{diff} = 0.14$, Cohen's $d = 0.61$) compared to adults ($M_{diff} = 0.24$, $SD_{diff} = 0.21$, Cohen's $d = 1.11$, $p = .004$); however, the difference was significantly different from zero in both age groups ($ps < .005$; see Figure 2H). Thus, Remember rates distinguished between trial type for both adults and children when eliminating additional decision demands, despite the added ongoing count of the frequency of treasuring. This suggests that the added decision demands, and not the addition of counting, were responsible for Experiment 2 findings.

Treasure Rates

Similar to remember rates, treasure rates also showed a significant age group by trial type interaction, $F(1, 50) = 17.56$, $p < .001$, $\eta^2_p = .26$ (see Figure 2I). The difference in treasure rates between nonmatch and match trials was significantly smaller in 6- and 7-year-olds ($M_{diff} = 0.06$, $SD_{diff} = 0.13$, Cohen's $d = 0.46$) compared to adults ($M_{diff} = 0.24$, $SD_{diff} = 0.18$, Cohen's $d = 1.33$, $p < .001$); however, the difference was significantly different from zero in both age groups ($ps < .03$). Thus, similar to Experiment 1 and in contrast to Experiment 2, children's treasure rates distinguished between trial type when selection demands were not increased, even when counting the frequency of treasuring. Children's ability to select memories for reward when they reported recollection was similar to previous Experiments and Remember rates predicted trial-level treasure decisions above and beyond memory accuracy (see Table S3).

Remember-Familiar Comprehension

We compared the overall accuracy of categorizing Remember and Familiar descriptions to verify participant's understanding of these experiences. No significant differences were observed for comprehension accuracy between 6- to 7-year-olds ($M = 0.70$, $SD = 0.22$) and adults ($M = 0.67$, $SD = 0.22$, $p = .63$), suggesting similar interpretations of Remember-Familiar experiences. Additionally, comprehension accuracy did not correlate with overall performance or Remember rates for either children ($ps > .66$) or adults ($ps > .43$). Thus, individual differences in comprehension accuracy are unlikely to explain our results.

Overall, Experiment 3 suggests that the increased selectivity demands in Experiment 2, and not the addition of counting treasure trials, resulted in changes in children' subjective experience and decision making such that trial type no longer influenced remember and treasure rates.

General Discussion

The current set of experiments examined developmental differences in subjective recollection and its role in decision making. From the early days of theorizing about episodic memory (Tulving, 1985), it has been suggested that the experience of recollection is particularly compelling and thus may motivate decision making. Despite its importance, little is known about the development of subjective recollection and its role in decision making. Moreover, the contribution of metamemory monitoring to metamemory control can be difficult to determine from previous literature, as objective performance and subjective experiences are correlated in the majority of paradigms used in children (Hembacher & Ghetti, 2013; Roebers et al., 2009).

The Development of Subjective Recollection

We successfully induced a dissociation between subjective recollection and objective performance across age groups. Participants reported higher Remember rates for nonmatch compared to match trials, despite higher objective memory accuracy for match trials. While this was true for both adults and children ages 6- to 10-year-olds, children reported higher rates of remembering particularly for match trials. These age differences in subjective recollection were not driven by differences in conceptual understanding of subjective experiences, as children and adults performed similarly when classifying Remember-Familiar responses, and classification accuracy was not associated with overall memory performance or rates of subjective recollection. Furthermore, these findings were replicated across two experiments, when complex decision demands were not added (Experiment 1 and Experiment 3), with similar effect sizes.

Our findings are consistent with previous research showing that younger children tend to overclaim states of recollection (Ghetti et al., 2011; Rollins & Riggins, 2017). There are several potential factors that may underlie this developmental difference. Metacognitive research shows developmental improvements in the cues or heuristics that children rely on for their subjective evaluations (Koriat & Ackerman, 2010), such as effort investment as a cue for how well material has been learned (Koriat et al., 2009) or distinctiveness as a cue for rejecting false events (Geurten, Meulemans, & Willems, 2018). Previous research examining subjective recollection demonstrates that young children's Remember reports are more diagnostic of accuracy when retrieving color as opposed to semantic details, despite better or similar objective memory for semantic details (Ghetti et al., 2011; Hembacher & Ghetti, 2013). This research suggests that developmental differences exist in the diagnostic cues that children attend to when reporting recollective experiences. Interestingly, the difference in accuracy between match and nonmatch trials was robust and similar across age groups. This result is consistent with previous reports of age-invariance in 6-year-old children's and adults' ability to distinguish between similar targets and lures involving concrete objects drawn from the same sample as our own stimuli (Ngo et al., 2017). However, research has also shown developmental improvements beyond early childhood (Keresztes et al., 2017; Rollins & Cloude, 2018). Moreover, these studies do not measure metamemory, making it difficult to draw firm conclusions about relations between accuracy and metamemory in these types of recognition tasks.

In the current task, we manipulated whether or not the test trial promoted the search for a diagnostic feature. The stronger developmental difference in the match condition suggests that children may prioritize global retrieval success (e.g., I remember seeing an apple) over the retrieval of criterial features (e.g., I remember this apple with a longer stem) even when the task clearly demands it resulting in higher accuracy. Moreover, in the current task match and nonmatch trials were tested using a randomized sequence, which requires participants to switch criterion for subjective recollection on every trial. Studies in adults show that changes in decision criteria on a trial-wise basis may pose challenges because of added noise or variability in the decision process (Benjamin, Tullis, & Lee, 2013). Children suffer switching costs more than adults do (Cepeda, Kramer, & Gonzalez de Sather, 2001) and this may be particularly true in the context of a general tendency to overclaim recollection (Ghetti et al., 2011; Hembacher & Ghetti, 2013). Future research may reduce the demand for switching by grouping match and nonmatch trials during test to examine whether children's attending to specific diagnostics features might be stronger if the need for a certain retrieval mode is induced and then sustained across multiple trials.

Future research would also benefit from exploring the types of diagnostic features that may promote children's subjective awareness of their retrieval processes. For example, younger children's subjective recollection may be more similar to that of adults if specific perceptual features that are particularly salient to children, such as color, were diagnostic of memory or if encoding tasks promoted attention toward diagnostic perceptual features. Additionally, eye-movement patterns have been shown to reveal important aspects of memory retrieval processes (Pathman & Ghetti, 2014). Future examination of eye-movements in the current task can provide new insight on the specific features that children attend to during their memory and subjective judgments. For example, eye-movement transitions between the target and distractor pictures during perceptually similar trials can reveal whether there are developmental differences in how children compare and assess items, and how this in turn may lead to differences in subjective experiences. Eye-movements may also help reveal when participants identify the most diagnostic feature in the target before endorsing it and whether they additionally identify features in the distracters that help reject it using a recall-to-reject strategy in their memory decision (Lampinen, Odegard, & Neuschatz, 2004).

Decision Making and the Role of Selection Demands

The dissociation between subjective recollection and memory accuracy allowed us to examine whether subjective experiences contributed to decision making beyond memory performance. We found that both children and adults' sorting decisions followed their subjective experiences such that they volunteered nonmatch trials more frequently than match trials. Our item level-analysis confirmed this trend. When we simultaneously included subjective recollection and memory accuracy to predict sorting decisions in our multilevel model, we found strong evidence that subjective recollection contributed to decision making above and beyond accuracy. Critically, when we increased decision selectivity by limiting the number of responses that could be volunteered, we observed that in children, but not adults, subjective recollection and sorting decisions no longer distinguished between match and nonmatch trials. We confirmed that these findings were not driven by counting of the number of trials volunteered throughout the task, but instead resulted from the added decision selectivity. This result suggests that adding decision complexity alters children's ability to attend to specific aspects of their retrieval experience.

Although developmental research on the connection between metacognition and executive function remains limited, longitudinal research demonstrates a link between these constructs in children (Roebers et al., 2012). Furthermore, cognitive demands have been shown to decrease metacognitive performance in adults (Schwartz, 2008), although impairments can be reduced with additional training (Coutinho et al., 2015) suggesting flexibility in this process. In this study, we found that added decision demands reduced sensitivity to differences in match compared to nonmatch trials, which manifested in children increasing their Remember rates during match trials. Although even younger children typically recognized that match trials required retrieval of fine-grained details, with added decision demands their ability to attend to this aspect of the retrieval process suffered. We also confirmed that children maintained the ability to discriminate between accurate and inaccurate memories in their sorting decisions, again underscoring the specificity of these findings. Taken together, these results suggest that added decision demands impact children's sensitivity to contexts that do or do not promote retrieval of specific details, but the ability to distinguish internal memory strength remained intact. These results are consistent with previous research suggesting that children's metamemory experiences can be based on different types of diagnostic cues (Geurten et al., 2018; Koriat et al., 2009), and these cues may be differentially manipulated by different decision demands. Furthermore, our findings are consistent with recent research demonstrating that increased decision demands in the form of limiting response selection can alter younger, but not older, children's ability to seek out helpful information during learning (Mills, Sands, Rowles, & Campbell, 2019), suggesting decision demands may also play an important role in learning as well as retrieval.

Future research should provide a deeper understanding of the conditions under which children's metamemory decisions may be improved. Limited research on the effects of metacognitive training in adults suggests that feedback about effective metacognitive strategies improves help seeking behaviors (Roll, Aleven, McLaren, & Koedinger, 2011), but may also result in inflated confidence (Huff & Nietfeld, 2009). In children, performance feedback and metacognitive feedback (i.e., feedback about the accuracy of monitoring) have been shown to reduce overconfidence resulting in more accurate decision making (van Loon & Roebers, 2017). Additionally, in adults the provision of metamemory judgments (compared to providing a response without judgments) improves memory retrieval under certain conditions (Myers, Rhodes, & Hausman, 2020). However, the direct impact of encouraging metamemory judgments on memory retrieval and decision making in children has been largely unexplored. Our study demonstrates that simplifying decision demands may scaffold children's metacognitive abilities and is an important factor to consider when determining how to benefit children's metamemory decisions. Furthermore, although we did our best to recruit a diverse sample of participants, it was primarily representative of a mid to upper class suburb community. Thus, we can draw limited conclusions on how changes in decision demands affect children from wider socioeconomic strata.

Children face many situations in which they ought to evaluate their memory evidence and make decisions on the basis of this evidence. This work provides new evidence that both children and adults use subjective experiences to guide their decision making above and beyond objective memory performance, underscoring the importance of characterizing the processes influencing these experiences. The present findings paint a complex picture in which children's ability to assess their states are affected by decision demands, highlighting that our understanding of children's memory decisions requires an account of both children's regulation abilities and contextual factors.

References

Balcomb, F., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science*, *11*, 760. https://doi.org/10.1111/j.1467-7687.2008.00725.x

Benjamin, A. S., Tullis, J. G., & Lee, J. H. (2013). Criterion noise in ratings-based recognition: Evidence from the effects of response scale length on recognition accuracy. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1601–1608. https://doi.org/10.1037/a0031849

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, *37*, 715–729. https://doi.org/10.1037/0012-1649.37.5.715

Chevalier, N., & Blaye, A. (2016). Metacognitive monitoring of executive control engagement during childhood. *Child Development*, *87*, 1264–1276. https://doi.org/10.1111/cdev.12537

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *47*, 1–12. https://doi.org/10.1162/jocn_a_00782

Coutinho, M. V. C., Redford, J. S., Church, B. A., Zakrzewski, A. C., Couchman, J. J., & Smith, J. D. (2015). The interplay between uncertainty monitoring and working memory: Can metacognition become automatic? *Memory & Cognition*, *43*, 990–1006. https://doi.org/10.3758/s13421-015-0527-1

de Bruin, A. B. H., & van Gog, T. (2012). Improving self-monitoring and self-regulation: From cognitive psychology to the classroom. *Learning and Instruction*, *22*, 245–252. https://doi.org/10.1016/j.learninstruc.2012.01.003

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological Science*, *29*, 761–778. https://doi.org/10.1177/0956797617744771

Desender, K., Calderon, C. B., Opstal, F. V., & den Bussche, E. V. (2017). Avoiding the conflict: Metacognitive awareness drives the selection of low-demand contexts. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1397. https://doi.org/10.1037/xhp0000391

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. https://doi.org/10.1016/j.jecp.2014.04.001

Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology*, *61*, 467–490. https://doi.org/10.1146/annurev.psych.093008.100445

Dobbins, I. G., Kroll, N., & Liu, Q. (1998). Confidence-accuracy inversions in scene recognition: A remember-know analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1306–1315. https://doi.org/10.1037/0278-7393.24.5.1306

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition*, *20*, 374–380. https://doi.org/10.3758/bf03210921

Fandakova, Y., Johnson, E. G., & Ghetti, S. (2021). Distinct neural mechanisms underlie subjective and objective recollection and guide memory-based decision making. *eLife*, *10*. https://doi.org/10.7554/eLife.62520

Fandakova, Y., Selmeczy, D., Leckey, S., Grimm, K. J., Wendelken, C., Bunge, S. A., & Ghetti, S. (2017). Changes in ventromedial prefrontal and insular cortex support the development of metamemory from childhood into adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 7582–7587. https://doi.org/10.1073/pnas.1703079114

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/bf03193146

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. https://doi.org/10.3389/fnhum.2014.00443

Gardiner, J. M., Ramponi, C., & Richardson-Klavehn, A. (2002). Recognition memory and decision processes: A meta-analysis of remember, know, and guess responses. *Memory*, *10*, 83–98. https://doi.org/10.1080/09658210143000281

Geurten, M., Meulemans, T., & Willems, S. (2018). A closer look at children's metacognitive skills: The case of the distinctiveness heuristic. *Journal of Experimental Child Psychology*, *172*, 130–148. https://doi.org/10.1016/j.jecp.2018.03.007

Ghetti, S., & Angelini, L. (2008). The development of recollection and familiarity in childhood and adolescence: Evidence from the dual-process signal detection model. *Child Development*, *79*, 339–358. https://doi.org/10.1111/j.1467-8624.2007.01129.x

Ghetti, S., & Fandakova, Y. (2020). Neural development of memory and metamemory in childhood and adolescence: Toward an integrative model of the development of episodic recollection. *Annual Review of Developmental Psychology*, *2*, 1–24. https://doi.org/10.1146/annurev-devpsych-060320-085634

Ghetti, S., & Lee, J. (2010). Children's episodic memory. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 365–373. https://doi.org/10.1002/wcs.114

Ghetti, S., Mirandola, C., Angelini, L., Cornoldi, C., & Ciaramelli, E. (2011). Development of subjective recollection: understanding of and introspection on memory States. *Child Development*, *82*, 1954–1969. https://doi.org/10.1111/j.1467-8624.2011.01645.x

Hembacher, E., & Ghetti, S. (2013). How to bet on a memory: Developmental linkages between subjective recollection and decision making. *Journal of Experimental Child Psychology*, *115*, 436–452. https://doi.org/10.1016/j.jecp.2013.03.010

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science*, *25*. https://doi.org/10.1177/0956797614542273

Hembacher, E., & Ghetti, S. (2016). Subjective experience guides betting decisions beyond accuracy: Evidence from a metamemory illusion. *Memory*, *25*, 1–11. https://doi.org/10.1080/09658211.2016.1197946

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning*, *4*, 161–176. https://doi.org/10.1007/s11409-009-9042-8

Ingram, K. M., Mickes, L., & Wixted, J. T. (2012). Recollection can be weak and familiarity can be strong. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 325–339. https://doi.org/10.1037/a0025483

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language*, *48*, 704–721. https://doi.org/10.1016/s0749-596x(02)00504-1

Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 9212–9217. https://doi.org/10.1073/pnas.1710654114

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*, 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science*, *13*, 441–453. https://doi.org/10.1111/j.1467-7687.2009.00907.x

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. https://doi.org/10.1016/j.jecp.2008.10.005

Koriat, A., Goldsmith, M., Schneider, W., & Nakash-Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, *79*, 405–437. https://doi.org/10.1006/jecp.2000.2612

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*, 478–492. https://doi.org/10.1016/j.jml.2005.01.001

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 332–342. https://doi.org/10.1037/0278-7393.30.2.332

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*, 174–179. https://doi.org/10.3758/pbr.15.1.174

Metcalfe, J., & Finn, B. (2013). Metacognition and control of study choice in children. *Metacognition and Learning*, *8*, 19–46. https://doi.org/10.1007/s11409-013-9094-7

Mills, C. M., Sands, K. R., Rowles, S. P., & Campbell, I. L. (2019). "I want to know more!": Children are sensitive to explanation quality when exploring new information. *Cognitive Science*, *43*, e12706. https://doi.org/10.1111/cogs.12706

Myers, S. J., Rhodes, M. G., & Hausman, H. E. (2020). Judgments of learning (JOLs) selectively improve memory depending on the type of test. *Memory & Cognition*, *48*, 745–758. https://doi.org/10.3758/s13421-020-01025-5

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation* (Vol. *26*, pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2017). The ontogeny of relational memory and pattern separation. *Developmental Science*, *21*, e12556. https://doi.org/10.1111/desc.12556

Pathman, T., & Ghetti, S. (2014). The eyes know time: A novel paradigm to reveal the development of temporal memory. *Child Development*, *85*, 792–807. https://doi.org/10.1111/cdev.12152

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review*, *45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: an explorative study on their interrelations and their role for school achievement. *Metacognition and Learning*, *7*, 151–173. https://doi.org/10.1007/s11409-012-9089-9

Roebers, C. M., Schmid, C., & Roderer, T. (2009). Metacognitive monitoring and control processes involved in primary school children's test performance. *British Journal of Educational Psychology*, *79*, 749–767. https://doi.org/10.1348/978185409x429842

Roebers, C. M., & Schneider, W. (2005). The strategic regulation of children's memory performance and suggestibility. *Journal of Experimental Child Psychology*, *91*, 24–44. https://doi.org/10.1016/j.jecp.2005.01.001

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, *21*, 267–280. https://doi.org/10.1016/j.learninstruc.2010.07.004

Rollins, L., & Cloude, E. B. (2018). Development of mnemonic discrimination during childhood. *Learning & Memory*, *25*, 294–297. https://doi.org/10.1101/lm.047142.117

Rollins, L., & Riggins, T. (2017). Age-related differences in subjective recollection: ERP studies of encoding and retrieval. *Developmental Science*, *82*, e12583. https://doi.org/10.1111/desc.12583

Schneider, W., & Lockl, K. (2008). *Procedural metacognition in children* (J. Dunlosky & R. A. Bjork, Eds.). Psychology Press. https://doi.org/10.4324/9780203805503.ch20

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. *Memory & Cognition*, *36*, 9–19. https://doi.org/10.3758/mc.36.1.9

Selmeczy, D., & Dobbins, I. G. (2014). Relating the content and confidence of recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 66–85. https://doi.org/10.1037/a0034059

Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543–545. https://doi.org/10.1038/nn.2112

Stine-Morrow, E. A. L., Shake, M. C., Miles, J. R., & Noh, S. R. (2006). Adult age differences in the effects of goals on self-regulated sentence processing. *Psychology and Aging*, *21*, 790–803. https://doi.org/10.1037/0882-7974.21.4.790

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*, 1–12. https://doi.org/10.1037/h0080017

van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00121

van Loon, M. H., & Roebers, C. M. (2017). Effects of feedback on self-evaluations and self-regulation in elementary school. *Applied Cognitive Psychology*, *31*, 508–519. https://doi.org/10.1002/acp.3347

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, *21*, 968–979. https://doi.org/10.1002/hipo.20808

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Table S1. Experiment 1 Mixed Logistic Regression Model Predicting Trial-Level Treasure Decision

Table S2. Experiment 2 Mixed Logistic Regression Model Predicting Trial-Level Treasure Decision

Table S3. Experiment 3 Mixed Logistic Regression Model Predicting Trial-Level Treasure Decision

Table S4. Pairwise Comparisons for Correct Versus Incorrect Remember Rates for All Age Groups

Table S5. Pairwise Comparisons for Correct Versus Incorrect Treasure Rates for All Age Groups

Appendix S1. Supplementary Methods and Results