Cognition 186 (2019) 82–94

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Memory enhancements from active control of learning emerge across development

Azzurra Ruggeri[a,*], Douglas B. Markant[b], Todd M. Gureckis[c], Maria Bretzke[d], Fei Xu[e]

[a] MPRG iSearch, Max Planck Institute for Human Development & School of Education, Technical University Munich, Germany
[b] Department of Psychological Science, University of North Carolina at Charlotte, United States
[c] Department of Psychology, New York University, United States
[d] MPRG iSearch, Max Planck Institute for Human Development & Department of Child and Adolescent Psychiatry, Faculty of Medicine, Technical University Dresden, Germany
[e] Department of Psychology, University of California, Berkeley, United States

ARTICLE INFO

Keywords:
Active learning
Recognition memory
Exploration
Cognitive development

ABSTRACT

This paper investigates whether active control of study leads to enhanced learning in 5- to 11-year-old children. In Experiments 1 and 2, participants played a simple memory game with the instruction to try to remember and later recognize a set of 64 objects. In Experiment 3, the goal was to learn the French names for the same objects. For half of the materials presented, participants could decide the order and pacing of study (Active condition). For the other half, they passively observed the study decisions of a previous participant (Yoked condition). Recognition memory was more accurate for objects studied in the active as compared to the yoked condition. However, the active learning advantage was relatively small among 5-year-olds and increased with age, becoming comparable to adults' by age 8. Our results show that the ability to actively control study develops during early childhood and results in memory benefits that last over a week-long delay. We discuss possible interpretations for the observed developmental change, as well as the implications of these results for educational implementations.

1. Introduction

Educators frequently argue that self-directed, *active learning* situations foster better and deeper learning (Bruner, Jolly, & Sylva, 1976; Kuhn & Kaplan, 2000; Montessori, 1912/1964; Piaget, 1930). Although this idea involves a number of interrelated issues, like motivation or deeper processing, many argue that giving students some degree of independent, volitional control during learning is itself beneficial (Chi, 2009; Gureckis & Markant, 2012; Markant, Ruggeri, Gureckis, & Xu, 2016). An important role for experimental psychology is to assess if and how these bedrock educational principles align with what we know about the basic mechanisms of learning and memory. Indeed, recent studies with adults have shown that minimal forms of volitional control (specifically, allowing learners to select the order and pacing of study) lead to memory improvements compared to situations lacking this control (Harman, Humphrey, & Goodale, 1999; Voss, Galvan, & Gonsalves, 2011; Voss, Gonsalves, Federmeier, Tranel, & Cohen, 2011; Voss, Warren, et al., 2011; Liu, Ward, & Markall, 2007; Markant, DuBrow, Davachi, & Gureckis, 2014; Meijer & Van der Lubbe, 2011; Plancher, Barra, Orriols, & Piolino, 2013).

While interesting, it is unclear how these adult findings might inform educational policies that seek to help developing learners. Self-directed learning requires the coordination of a range of cognitive processes, including decision making, exploration, metacognition, attention, and memory, all of which are subject to critical developmental changes (Kachergis, Rhodes, & Gureckis, 2017). In light of this, one possibility is that only mature learners can effectively leverage self-directed learning. This paper presents a series of experiments designed to trace the emergence and developmental trajectory of self-directed learning as a successful learning modality for children.

1.1. Development of active learning and metamemory

Earlier work suggested that children do not select the most informative evidence to explore until late primary school age (Chen & Klahr, 1999; Kuhn & Brannock, 1977). However, more recent research suggests that children are effective and adaptive active learners from a very early age (Ruggeri, Sim, & Xu, 2017), although the informativeness

* Corresponding author.
E-mail address: ruggeri@mpib-berlin.mpg.de (A. Ruggeri).

https://doi.org/10.1016/j.cognition.2019.01.010
Received 2 June 2018; Received in revised form 8 January 2019; Accepted 14 January 2019
0010-0277/

of their learning strategies undergoes a large developmental change from age 4 to adulthood (see Ruggeri & Feufel, 2015; Ruggeri & Lombrozo, 2015; Ruggeri, Lombrozo, Griffiths, & Xu, 2016). It is not yet fully understood which factors drive developmental changes in active learning effectiveness, how they interact with each other, or how their relative importance changes at different developmental stages. Potential factors include verbal skills, conceptual knowledge, executive functions, formal education, and socioeconomic status (SES). In particular, the effectiveness of active learning strategies is expected to heavily depend on children's metacognitive abilities. Metamemory (the ability to introspect on the accuracy of one's memories) has been shown to impact the implementation of appropriate memory strategies in both adults (Hutchens et al., 2012) and children (Geurten, Catale, & Meulemans, 2015; Grammer, Purtell, Coffman, & Ornstein, 2011). Metamemory monitoring improves considerably over the elementary school years (Roebers, 2017), with older children's confidence judgments showing greater discrimination between accurate and inaccurate memories than younger children's (Fandakova, Shing, & Lindenberger, 2013; Ghetti, Castelli, & Lyons, 2009; Ghetti, Mirandola, Angelini, Cornoldi, & Ciaramelli, 2011). Previous work also suggests that the ability to allocate study time based on the difficulty or familiarity of the materials develops across childhood (Metcalfe, 2002; Metcalfe & Finn, 2013a). For example, early studies found that older (10- and 12-year-olds), but not younger children (6- and 8-year-olds) spent significantly more time learning unrelated, difficult picture pairs (e.g., frog-book) than related, easy pairs (e.g., bat-ball; Dufresne & Kobasigawa, 1988, 1989). Along the same lines, Lockl and Schneider (2003) found that both first and third graders were able to differentiate between easy and difficult picture pairs, but only third graders adjusted their study time accordingly.

These work suggests that the ability to actively organize study behavior strategically and effectively emerges over early childhood, and develops across the lifespan. However, a more relevant question for educators is not whether children are sensitive or aware of their own memory abilities, but whether allowing children to control their own learning process would lead to learning advantages. To this end, merely studying the development of metamemory skills and strategies is not enough. Instead, it is crucial to investigate whether and how active control as a learning modality is beneficial for children as compared to more passive forms of instruction.

1.2. Benefits of active learning on memory

Previous research with adults has investigated the benefits of active learning by using memory tasks. For example, Voss and colleagues (Voss, Galvan, et al., 2011; Voss, Gonsalves, et al., 2011; Voss, Warren, et al., 2011) presented adult participants with a set of objects arranged on a grid, with only one object visible at a time through a moving window, and asked them to memorize as many objects as possible. Participants alternated between active study blocks, in which they controlled the study sequence and timing by deciding how to move the window, and "yoked" study blocks, in which they observed the study sequence that a previous participant had generated in an active study block. By matching the content experienced during study across conditions, yoked designs isolate the effects of active control on learning and memory. These studies have found robust benefits of active control of study on object recognition, meaning that participants were more accurate at recognizing objects that had been actively studied as compared to those studied in the yoked condition. This advantage has been found to persist a week after the initial study session (Voss, Gonsalves, et al., 2011). Similar results have been obtained with a variety of related tasks and materials (Harman et al., 1999; Liu et al., 2007; Markant et al., 2014; Meijer & Van der Lubbe, 2011; Plancher et al., 2013), and with both younger and older adults (Brandstatt & Voss, 2014).

These studies further revealed that the benefits of active study depended on how participants explored the objects. Voss, Galvan, et al. (2011) found that objects studied for longer durations were more likely to be recognized in the active, but not in the yoked, condition (although yoked observers seem to benefit from additional study time when cued to the locations of new stimuli, see Markant et al., 2014). Moreover, revisiting objects within a short period of time also led to better memory performance among younger adults (Brandstatt & Voss, 2014; Voss, Galvan, et al., 2011), but only following active study. Interestingly, the same search pattern was less common among older adults and did not lead to the benefits from active study observed in younger adults (Brandstatt & Voss, 2014). With this type of learning task, it is possible to investigate the link between the search strategies generated during active study (including the sequencing of items and allocation of study time) and the resulting benefits over passive observation of the same information.

A few recent studies have shown that active control can facilitate learning in children as well. For example, Sim, Tanner, Alpert, and Xu (2015) found that 7-year-olds learned categorical rules more effectively when they were free to decide what information to gather, as compared to yoked observations. Active control has also been shown to lead to memory improvements for children in spatial navigation tasks (Feldman & Acredolo, 1979; McComas, Dulberg, & Latter, 1997; Poag, Cohen, & Weatherford, 1983) and when learning novel object-word pairings (Partridge, Mcgovern, Yung, & Kidd, 2015). Although these results suggest that active control would enhance children's episodic memory, it is an open question at what age this benefit would emerge and how it would develop across childhood. Do these advantages reflect developmental progress along the way to adult competence, or a robust and universal advantage of self-directed learning?

The present paper compares the effects of active and yoked study on 5- to 11-year-old children's learning. Experiment 1 replicates the design from previous adult studies (Markant et al., 2014; Voss, Gonsalves, et al., 2011). Experiment 2 and Experiment 3 replicate and extend the results from Experiment 1, exploring the cognitive and metacognitive factors that might impact the benefit of active control of study. In particular, Experiment 2 investigates the effect of differential pre-exposure on children's studying strategies (mimicing the easy-hard manipulations of the metamemory literature). Experiment 3 explores the effect of active study in a paired associate learning task in which children had to learn the French names for the same objects used in Experiments 1 and 2, additionally controlling for children's working memory as a possible moderator for the advantage of active learning.

2. Experiment 1

2.1. Participants

Participants in Experiment 1 were 51 5- to 8-year-old children (24 female, $M_{age} = 82.59$ months; $SD = 13.72$ months; range: 60 to 105 months) tested in the laboratory and recruited from Berkeley, California and surrounding communities. Nine participants (18%) did not return for the retest session, but the data from their first test session were included in the analyses. The mean interval between first and second sessions was 8.3days ($SD = 2.63$ days; range: 5 to 15 days). Sample size was determined based on previous active learning studies with adults (see Markant et al., 2014) and children (Metcalfe & Finn, 2013a; Partridge et al., 2015; Ruggeri & Lombrozo, 2015; Ruggeri et al., 2016). Before the children participated in the study, written informed consent was obtained from participants' parents and the local ethical review board at the University of California, Berkeley approved the study protocol (#2010-03-1013). Children received a small present to thank them for their participation.

2.2. Materials

The stimuli set consisted of 200 line drawings of the most frequent objects mentioned by 2- to 5-year-old children in their everyday conversations with adults, as recorded by the CHILDES corpus (Child

Language Data Exchange System; MacWhinney and Snow, 1985). We further ensured that children from all countries (Germany, Italy, and the U.S.) recognized and were familiar with the stimuli by asking them, at the end of the recognition task, if there were any objects among those presented they did not recognize and know the name of. Participants across all countries successfully recognized all the stimuli used. To minimize perceptual differences, all objects were re-drawn using the same stroke, color palette and coloring style. Eight of the 200 drawings were used as training stimuli for the familiarization trials and 192 drawings were used as stimuli for the first and second experimental sessions (see Fig. S1 of the Supplementary Online Materials, SOM). The experimental materials were presented on a 9.7” iPad touchscreen using custom software.

2.3. Design and procedure

Children were tested individually using a simple memory game. In the game, children were presented with a grid of objects and asked to remember as many of the presented objects as possible. The design and procedure were modeled after Experiment 2 in Markant et al. (2014), with modifications aimed at making it more suitable for children (including fewer study rounds, fewer objects to be studied and shorter study duration per round).

2.3.1. Familiarization

Participants were first presented with two familiarization trials aimed at introducing the goal of the game, the study procedures, and making children comfortable using the touchscreen. During each familiarization trial, children were presented with four objects arranged in a 2 × 2 grid. The objects were shown on the screen for two seconds before disappearing under occluders (same as for the main experimental session, see Fig. 1, top). Participants were instructed that the goal of the game was to remember all the objects presented on the screen. The first familiarization trial introduced the study procedure of the active blocks. Participants were told that in some rounds they could decide which occluder button to touch in order to view the object hidden beneath. After a touch, a red frame appeared for 500 ms, followed by the removal of the occluder that revealed the hidden object. Children were instructed that, before studying another object, they had to touch the object currently displayed once more to make it disappear behind the occluder. The experimenter modeled the touching actions while explaining the procedure. Children then had the opportunity to practice the active study procedure. If necessary, the experimenter provided feedback and repeated the instructions (e.g., ”remember, you have to touch the object again to make it disappear, before you can study another object”). Once children were familiar with the active study procedure, they moved on to the second familiarization trial, which introduced children to the study procedure of the yoked blocks. They were told that in other rounds the game would decide what objects they would see and for how long. Children were then presented with a randomly generated study sequence. As in the active blocks, a red frame preceded each object for 500 ms so that children had time to allocate their attention to the new study location before the object appeared. To keep engagement and attention level comparable to the active blocks, during yoked blocks children were asked to touch the objects as soon as they appeared, although this touch had no effect on the display. There were no time constraints for the familiarization trials.

2.3.2. Study phase

The main experimental session consisted of two active and two yoked study blocks (four blocks total), presented in alternating order (i.e., active, yoked, active, yoked). The active block was always presented first, so that children’s initial active study pattern would not be influenced by the study pattern observed in the yoked blocks. Each study block presented children with 16 objects arranged in a 4 × 4 grid. All 16 objects were visible on the screen for 2 s at the beginning of each

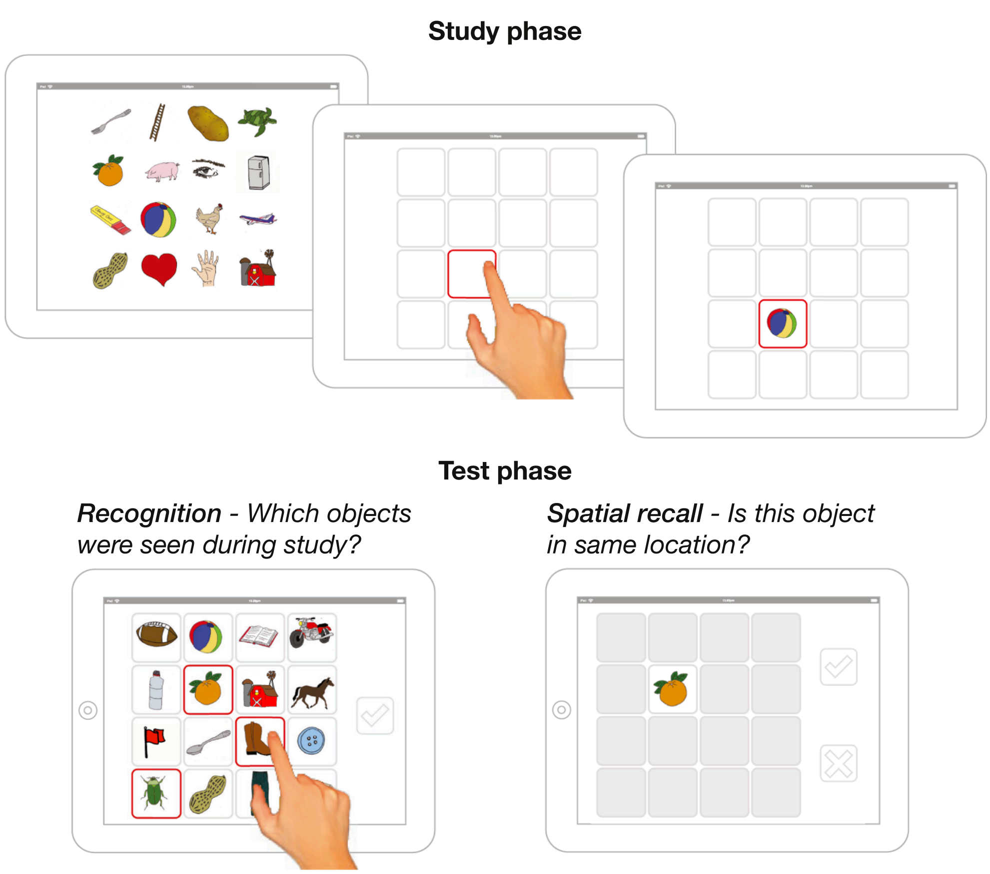

Fig. 1. Procedure used in Experiments 1 and 2. **Top:** Each study round began with all objects displayed for two seconds. After the objects disappeared, participants either selected a location to study (Active condition), causing a red frame to appear, followed by the object, or touched the location where the object appeared (Yoked condition), preceded by a red frame. **Bottom left:** During each test block, participants selected the objects that they recognized from the study phase. **Bottom right:** Spatial recognition test (Experiment 2 only). For objects that were recognized, participants judged whether they were presented on the test grid in the same location as where they had appeared during study.

Table 1
Estimated effects from logistic regression model of recognition responses to studied items.

	Experiment 1	Experiment 2
(Intercept)	1.12 (0.12)***	1.24 (0.15)***
Age	0.06 (0.10)	0.69 (0.13)***
Condition [yoked]	−0.28 (0.08)***	−0.68 (0.09)***
Test [retest]	−0.48 (0.09)***	−0.66 (0.10)***
Study duration	0.45 (0.08)***	0.72 (0.09)***
No. visits	0.29 (0.08)***	−0.05 (0.08)
FA rate	0.16 (0.05)***	0.33 (0.07)***
Half [second]	−0.03 (0.06)	0.20 (0.07)**
Age X Condition [yoked]	−0.22 (0.06)***	−0.28 (0.07)***
Age X Test [retest]	0.27 (0.06)***	0.16 (0.07)*
Condition [yoked] X Test[retest]	−0.01 (0.12)	0.04 (0.13)
Condition [yoked] X Study duration	−0.13 (0.10)	−0.30 (0.11)**
Condition [yoked] X No. visits	−0.02 (0.10)	0.40 (0.10)***
Marginal R^2	0.12	0.22
Conditional R^2	0.27	0.40

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

study block, before disappearing under occluders (see Fig. 1, top). Across the four blocks, children were asked to memorize 64 objects (see Materials section and Fig. S1 of the SOM). In the active blocks, children had 90 s to select and study the objects in order to memorize them. In the yoked blocks, children were presented with the 90-s study sequence (i.e., same objects and pacing) of one of the previous participants' active learning blocks. In between blocks, there was a 20-s break in which children were briefly reminded of the study procedure for the next block.

2.3.3. Test phase

The study phase was immediately followed by a test phase consisting of 8 blocks. In each test block, 16 objects were again presented in a 4 × 4 grid (see see Fig. 1, bottom left). Across the 8 test blocks, 64 of the objects had appeared during the study phase (*old* objects) and 64 were objects that were not presented during study (*new* objects). The number of old objects from active and yoked blocks randomly varied across test blocks (active: $M = 4.23$, $SD = 2.16$; yoked: $M = 4.3$, $SD = 2.25$). All objects were displayed in random locations on the grid.

For each block, children were asked to indicate the objects they had studied earlier by touching them on the screen. Selected objects were framed in red to help participants keep track of the objects selected as recognized. Children could deselect any of the previously selected objects by touching them again on the screen and making the red frame disappear. After selecting all the objects they recognized from the study phase, children were prompted to touch a button to proceed to the next test block. Children were not given any feedback about their performance during or after the test phase.

About one week later, children returned to the lab for a second session, in which they were asked to complete 8 new test blocks. The 64 objects studied in the first session were randomly mixed with 64 new objects (i.e., objects that were not used during the first experimental session, neither as study nor as test objects). The testing procedure was identical to the test phase from the first experimental session.

2.4. Results

The database containing the raw data for all the three Experiments included in this manuscript is archived in a public repository, linked in the Supplementary Online Materials (Database: Ruggeri, Markant, Gureckis, Bretzke, & Xu, 2019).

2.4.1. Recognition of studied objects

Mixed effects logistic regression was used to model recognition responses ("old" versus "new") for objects presented during the study phase. All mixed effects models were fit using the *lme4* library in *R* (Bates, Mächler, Bolker, & Walker, 2014). A separate analysis of false alarms, that is, the number of objects marked as recognized at test that had not been presented during the study blocks, is presented in the Section S1 of the SOM. This analysis showed that the rate of false alarms was relatively low ($M = .09$, $SD = .09$) and did not differ as a function of condition or age.

The model included fixed effects for age (continuous), study condition (active vs. yoked), half (first two active/yoked study blocks vs. second two active/yoked study blocks), and testing session (test vs. retest) as well as pairwise interactions between age, condition, and testing session. The number of false alarms was included as a fixed effect to control for participants' overall tendency to respond "old" in each testing session. Random effects terms were included for participants and items (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). Two additional predictors were included to assess item-level effects of study behavior: cumulative study duration (i.e., the total amount of time a certain object was studied) and number of visits (i.e., the number of times the object was visited). Finally, we included terms for the interaction between each study behavior measure and study condition, to assess whether the effects of increased study duration or number of visits differed under active and yoked study blocks. Study duration, number of visits, and false alarm rates were square-root transformed to correct for positive skew. Non-categorical predictors were scaled and centered prior to model fitting.

The parameters of the fitted model are shown in Table 1. Contrast-based hypothesis tests were conducted using the *multcomp R* library (Hothorn, Bretz, & Westfall, 2008). Effect sizes and 95% confidence intervals for significant effects are reported in terms of relative odds ratio (*OR*) which indicate the multiplicative change in the odds of recognizing a studied object that is associated with a unit change in a given predictor. A 95% confidence interval that excludes $OR = 1$ indicates a significant effect. Additional model comparisons were conducted to assess whether the distance between the location in which the objects were presented on the study grid and their location on the test grid impacted recognition accuracy, but including distance as a predictor did not improve the fit of the model ($\chi^2(1) = .11$, $p = .74$).

Recognition accuracy was higher for objects studied in the active condition in both test (Active: $M = .70$, $SD = .16$; Yoked: $M = .67$, $SD = .16$; $OR = 1.32$ [1.12, 1.56]) and retest (Active: $M = .61$, $SD = .17$; Yoked: $M = .57$, $SD = .17$; $OR = 1.34$ [1.12, 1.59]). Recognition accuracy declined significantly from test to retest in both the active ($OR = .62$ [.52, .74]) and yoked condition ($OR = .61$ [.52, .73]).

There was no overall effect of age on recognition ($OR = 1.06$ [.87, 1.30]), but there were significant age × condition ($OR = .80$ [.71, .91]) and age × testing session ($OR = 1.31$ [1.16, 1.49]) interactions: The positive effect of active study was larger for older children, and the decrease in performance in the retest was smaller for older children (see Fig. 2A). Fig. 2B shows the within-subjects difference in recognition for active study and yoked study as a function of age. Although absolute levels of performance in both conditions changed from test to retest, the enhancement from active study was relatively consistent across the two sessions. Table 2 shows descriptive statistics for accuracy based on a median split on age.

The fitted model was used to contrast the predicted effects of study condition at ages 6 and 8 (approximately equal to the 25% and 75% quartiles of the age distribution, respectively). At age 6, there was no difference in recognition for objects studied in active versus yoked conditions, for neither the test ($OR = 1.12$ [.93, 1.35]) nor the retest

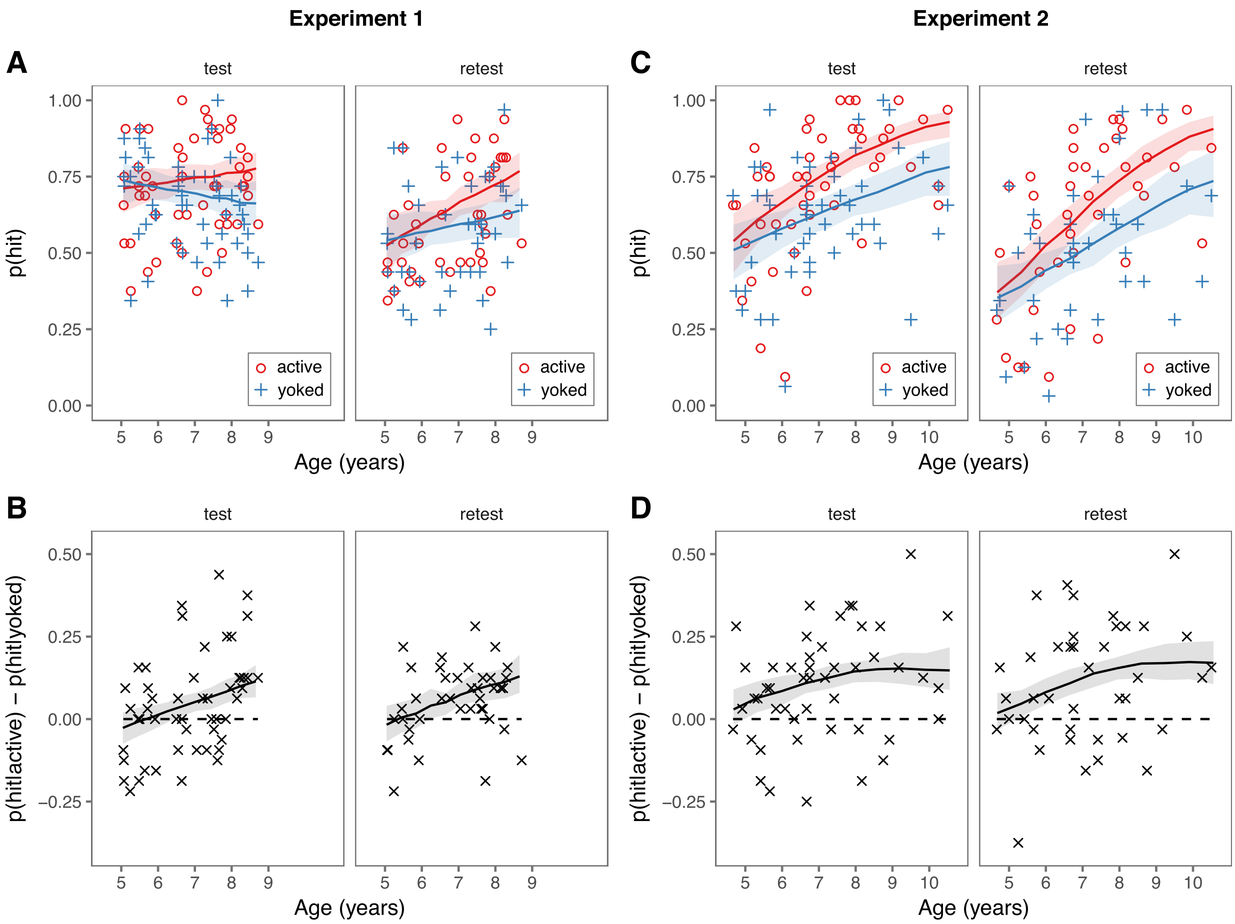

Fig. 2. Recognition accuracy for studied objects in Experiments 1 and 2. Individual points represent participants, solid line is fitted regression line, and shaded regions indicate 95% confidence intervals calculated through parametric bootstrapping. **A, C:** Recognition of objects studied in active (red) and yoked (blue) conditions as a function of age. **B, D:** Within-subjects difference between active and yoked recognition as a function of age. Dashed lines indicate no difference between study conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sessions (*OR* = 1.13 [.93, 1.37]). At age 8, actively studied objects were more likely to be recognized than yoked objects in both the test (*OR* = 1.66 [1.35, 2.05]) and retest sessions (*OR* = 1.68 [1.35, 2.10]). This 2-years age difference was associated with a significant increase in the magnitude of the advantage from active study (*OR* = 1.49 [1.19, 1.85]).

Study behavior at the object level had a strong impact on recognition performance in both study conditions. Objects were more likely to be recognized when studied for a longer duration in both the active (*OR* = 1.56 [1.34, 1.82]) and yoked (*OR* = 1.38 [1.21, 1.57]) conditions. In addition, objects were more likely to be recognized when visited more frequently for both the active (*OR* = 1.34 [1.15, 1.57]) and the yoked (*OR* = 1.31 [1.14, 1.51]) conditions. Thus, additional study effort, either measured as number of visits or time spent studying an object, led to improved recognition memory regardless of condition.

2.4.2. Study behavior in the active study blocks

Participants studied an average of 30.7 (*SD* = 2.72) of the 32 objects presented (96%) during the active study blocks. In addition to the object-level study behavior presented above, we examined how participants explored the grid during active blocks and whether their search

Table 2
Accuracy (Mean and SD) by study condition and test session (based on median split on age).

		Active	Yoked	Active − Yoked
Experiment 1				
Younger (age < 7 years)	Test	0.69 (0.16)	0.69 (0.14)	−0.004 (0.15)
	Retest	0.54 (0.15)	0.52 (0.18)	0.02 (0.12)
Older (age ⩾ 7 years)	Test	0.73 (0.16)	0.64 (0.16)	0.09 (0.15)
	Retest	0.68 (0.17)	0.61 (0.17)	0.07 (0.10)
Experiment 2				
Younger (age < 6.75 years)	Test	0.59 (0.21)	0.56 (0.22)	0.03 (0.14)
	Retest	0.44 (0.23)	0.37 (0.20)	0.07 (0.19)
Older (age ⩾ 6.75 years)	Test	0.83 (0.13)	0.67 (0.17)	0.16 (0.16)
	Retest	0.76 (0.18)	0.62 (0.21)	0.13 (0.18)

Table 3
Analysis of study behavior during active blocks.

	First block	Second block	Block		Age		Block X Age	
	M (*SD*)	*M* (*SD*)	*F*	*p*	*F*	*p*	*F*	*p*
Visitation rate								
Exp 1	1.90 (0.52)	2.32 (0.63)	49.4	< 0.001*	4.90	0.03*	0.14	0.71
Exp 2	2.14 (0.56)	2.42 (0.51)	17.2	< 0.001*	2.28	0.14	0.07	0.79
Exp 3	1.31 (0.47)	1.44 (0.47)	8.65	< 0.01*	20.	< 0.001*	1.32	0.25
Movement distance								
Exp 1	1.43 (0.23)	1.38 (0.24)	1.56	0.22	3.06	0.09	0.38	0.54
Exp 2	1.34 (0.20)	1.38 (0.21)	2.98	0.09	8.04	0.01*	0.87	0.34
Exp 3	1.47 (0.18)	1.49 (0.15)	0.22	0.64	0.37	0.54	0.	0.98
Sequence entropy								
Exp 1	1.54 (0.44)	1.61 (0.49)	0.92	0.34	3.30	0.08	2.77	0.10
Exp 2	1.38 (0.50)	1.35 (0.46)	0.11	0.73	3.33	0.07	2.44	0.12
Exp 3	1.15 (0.55)	1.29 (0.61)	3.19	0.08	5.99	0.02*	0.22	0.64

Note: The denominator degrees of freedom for the *F* tests were 49 (Experiment 1), 47 (Experiment 2), and 67 (Experiment 3). * $p < 0.05$, ** $p < 0.01$.

strategies were related to overall performance in the recognition test. We used the following measures to assess exploration behavior: (1) the visitation rate, i.e. the number of visits per object,[1] averaged across all objects within a block; (2) mean movement distance between consecutive objects (i.e., how far the participants moved to select the next object, measured in euclidean distance); and (3) sequence entropy, a measure of how systematically participants explored the grid. Sequence entropy was calculated as follows. For each block, a 4 × 4 transition matrix was constructed to represent all possible transitions from the object currently studied (positioned at the top left corner of the matrix) to different locations in the grid, such that all transitions were in the positive x and y directions. We then measured the proportion of transitions that occurred in each observed study sequence and calculated the Shannon entropy (Shannon, 1948) of this distribution. For example, if a participant repeatedly selected the same object to study for an entire block, sequence entropy would be zero; if she followed a consistent search pattern (e.g., always moving left-to-right within each row), sequence entropy would be relatively low; if new locations were chosen at random, sequence entropy would be high (with a maximum value of 2.77 nats if all transitions were made with equal frequency).

First, we ran two-way ANOVAs to assess the impact of task experience (i.e., first or second active learning block) and age on each search measure (Table 3). The analyses revealed an increase in average number of visits per object from the first to the second active block, and a positive main effect of age, such that older children had a higher visitation rate. There were no other differences in movement distance or sequence entropy.

Second, we modeled the effects of each search measure on recognition performance (proportion of objects that were correctly recognized) using mixed effects logistic regression. In addition to predictors for condition, age, and test session, the model included predictors for three measures of aggregate study behavior: visitation rate, movement distance, and sequence entropy. Additional predictors were included to test for interactions between each measure and study condition. Random intercept terms were included for participants. The parameters of the fitted models are shown in Table 4. There were no effects of visitation rate (Active: (*OR* = .98 [.88, 1.09]; Yoked: *OR* = 1.03 [.93, 1.15]) or sequence entropy (Active: (*OR* = .1.03 [.94, 1.13]; Yoked: *OR* = .96 [.87, 1.06]). Average movement distance had no effect in the active condition (*OR* = .95 [.85, 1.06]), but had a significant negative effect in the yoked condition (*OR* = .87 [.78, .96]): yoked performance decreased when observing study sequences that involved many transitions between distant locations in the grid.

Finally, we examined whether the within-subjects difference between active and yoked study was attributable to differences in study behavior between participants and the partners to whom they were yoked. For instance, a child who implements a very systematic search strategy in her own active blocks might have difficulty when following more random search sequences generated by her partner. This model included as predictors the between-subjects differences in the average number of visits, movement distance, and sequence entropy. The analysis showed no relationships between these predictors and the observed differences between active and yoked conditions (all *ps* > .05).

3. Experiment 2

In Experiment 2 we replicate and extend Experiment 1 in two significant ways. First, we introduce a pre-exposure manipulation to investigate whether the advantage of active control for memory depends on the efficacy of children's metamemory. At the beginning of each block, some objects were displayed on the screen for a longer time as compared to others, before disappearing under the occluders. If children recognize that they had less time to study the short-exposure objects, they may strategically devote more of their study effort to these objects.

Second, we introduce an explicit spatial memory test to assess if active control also improves spatial memory. Voss, Warren, et al. (2011), Voss, Galvan, et al. (2011), Voss, Gonsalves, et al. (2011) found that, following active study, adult participants were better able to recall the locations on the grid where objects had been presented during study (but see Markant et al., 2014; Brandstatt & Voss, 2014, for variable evidence on this effect).

3.1. Participants

Participants in Experiment 2 were 48 4- to 10-year-old children (30 female, M_{age} = 84.53 months; *SD* = 18.87 months; range: 56 to 126 months) recruited from a day care center in Livorno, Italy. Eight participants (16%) did not return for the retest, but the data from the first test session were included in the analyses. The mean interval between first and second sessions was 5.9 days (*SD* = 2.6 days; range: 3 to 9 days). As in Experiment 1, written informed consent was obtained from participants' parents and the ethical review board at the Max Planck Institute for Human Development in Berlin approved the study protocol for the project "Active Learning and Memory." Children

[1] We also examined the total number of study episodes and the median duration of study episodes (averaged across objects). However, because these measures were strongly correlated with the average number of visits, we do not report them here.

Table 4
Estimated effects of search behavior on test performance.

	Experiment 1	Experiment 2	Experiment 3
Intercept	0.99 (0.10)***	1.13 (0.13)***	−0.29 (0.05)***
Age	0.14 (0.12)	0.70 (0.13)***	0.18 (0.06)**
Condition [yoked]	−0.28 (0.08)***	−0.56 (0.08)***	−0.18 (0.06)**
Test [retest]	−0.36 (0.09)***	−0.56 (0.09)***	
False alarms	0.14 (0.04)**	0.31 (0.06)***	
Working memory			0.15 (0.05)***
Visitation rate	−0.02 (0.05)	−0.16 (0.06)**	0.07 (0.06)
Sequence entropy	0.03 (0.05)	−0.16 (0.06)*	0.08 (0.05)
Movement distance	−0.06 (0.06)	−0.03 (0.06)	−0.05 (0.05)
Age X Condition [yoked]	−0.29 (0.08)***	−0.30 (0.07)***	0.07 (0.07)
Age X Test [retest]	0.30 (0.07)***	0.14 (0.07)*	
Condition [yoked] X Test [retest]	−0.03 (0.11)	0.00 (0.12)	
Condition [yoked] X Visitation rate	0.05 (0.07)	0.23 (0.08)**	−0.06 (0.08)
Condition [yoked] X Sequence entropy	−0.07 (0.07)	0.16 (0.09)	−0.04 (0.07)
Condition [yoked] X Movement distance	−0.09 (0.08)	0.12 (0.09)	−0.04 (0.07)
Marginal R^2	0.03	0.14	0.03
Conditional R^2	0.12	0.27	0.05

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

received a small present as compensation for their participation.

3.2. Materials, design and procedure

The materials used in Experiment 2 were identical to those used in Experiment 1. The design and procedures of Experiment 2 were the same as in Experiment 1, with two crucial differences: First, during the study phase we manipulated children's pre-exposure to the 16 objects presented on the grid. The study grid was subdivided into three regions: a corner quadrant comprised of 4 objects; an inner border comprised of 5 objects immediately adjacent to the corner region; and an outer border including the remaining 7 objects. At the beginning of each study block, all objects were displayed, then the corner quadrant was occluded after 2s, the inner border was occluded after 4s, and the outer border region was occluded after 6s. Once all objects were occluded, the study phase continued following the same procedure as in Experiment 1.

Second, we added a test of spatial memory for those objects that participants recognized from the study phase. After participants selected the objects they recognized, they were asked whether each object was currently in the same position on the grid as where it had appeared during the study phase (see Fig. 1, bottom right). As in Experiment 1, after about one week participants were tested in a second session, in which they were asked to complete 8 new test blocks.

3.3. Results

3.3.1. Recognition of studied objects

Recognition responses to studied objects were modeled using the same method as in Experiment 1 (see Table 1 for the parameters of the fitted model). False alarm rates were again relatively low ($M = .08$, $SD = .15$). False alarms increased from the test to the retest, but there were no effects of age or condition (see Section S1 of the SOM).

Recognition accuracy was higher for objects studied in active blocks in both the test (Active: $M = .71$, $SD = .21$; Yoked: $M = .61$, $SD = .21$; $OR = 1.98$ [1.66, 2.37]) and the retest (Active: $M = .61$, $SD = .26$; Yoked: $M = .50$, $SD = .25$; $OR = 1.91$ [1.58, 2.31]) sessions. Recognition accuracy declined from the test to the retest for both the active ($OR = .52$ [.43, .63]) and the yoked conditions ($OR = .54$ [.45, .64]).

There was a positive overall effect of age on recognition ($OR = 2.00$ [1.54, 2.60]) and a positive age × testing session interaction ($OR = 1.17$ [1.02, 1.34]), such that the decline in performance in the retest was smaller for older children (Fig. 2C). In addition, there was a negative age × condition interaction ($OR = 0.76$ [0.66, 0.87]), such that the positive effect of active study was larger for older children (Fig. 2D). Table 2 reports descriptive statistics for recognition accuracy based on a median split on age. We again used the fitted model to contrast the predicted effects of study condition at ages 6 and 8. Recognition accuracy was higher for actively studied objects at both age 6 (test: $OR = 1.63$ [1.34, 1.97]; retest: $OR = 1.57$ [1.28, 1.93]) and age 8 (test: $OR = 2.33$ [1.90, 2.87]; retest: $OR = 2.25$ [1.83, 2.78]). This 2-year age difference was associated with a significant increase in the advantage from active study ($OR = 1.43$ [1.20, 1.71]).

Objects studied for longer duration were more likely to be recognized in both the active ($OR = 2.06$ [1.75, 2.44]) and yoked ($OR = 1.53$ [1.34, 1.76]) conditions. There was no effect of the number of visits in the active condition ($OR = 0.95$ [.82, 1.10]), but there was a significant interaction between condition and number of visits, such that increased visits to an object had a positive effect on recognition for objects studied in the yoked ($OR = 1.41$ [1.23, 1.62]) condition. Including the distance between the study location and location on the test grid did not improve the fit of the model ($\chi^2(1) < .01$, $p = .98$).

3.3.2. Spatial recall of correctly recognized objects

Mixed effects logistic regression was used to analyze the accuracy of participants' spatial judgments regarding the locations of correctly recognized objects. In the spatial recall test, participants judged whether an object appeared in the same or different location as in the study phase. Random effects terms were included for participants and items. The same set of predictors used to model recognition responses was included here. In addition, the model included a predictor for the object's true location relative to its location in the study phase (same vs. different), and the interactions of the location with age, condition, and test session. The coefficients of the fitted model are shown in Table 5.

Accuracy for objects appearing in the same location as during study was high ($M = .90$, $SD = .17$). However, spatial accuracy was significantly lower for objects that were in different locations at test as compared to the study phase ($M = .23$, $SD = .29$; $OR = .03$ [.02, .04]). There was no effect of age on accuracy for objects that appeared in the same location ($OR = .93$ [.73, 1.18]), but spatial recognition accuracy for objects that appeared in different locations increased with age ($OR = 1.77$ [1.42, 2.18]).

Table 5
Estimated effects from logistic regression model of spatial recall test for correctly recognized items.

	Experiment 2
(Intercept)	2.20 (0.15)***
Age	−0.08 (0.12)
Location [diff]	−3.53 (0.16)***
Condition [yoked]	0.07 (0.18)
Test [retest]	−0.20 (0.18)
No. visits	0.24 (0.10)*
Study duration	−0.10 (0.07)
Age X Condition	−0.07 (0.10)
Age X Test [retest]	−0.19 (0.10)
Age X Location [diff]	0.64 (0.10)***
Location [diff] X Condition [yoked]	−0.04 (0.20)
Location [diff] X Test [retest]	−0.14 (0.20)
Condition [yoked] X Test [retest]	−0.10 (0.19)
Condition [yoked] X No. visits	−0.23 (0.14)
Condition [yoked] X Study duration	0.23 (0.11)*
Marginal R^2	0.48
Conditional R^2	0.51

** $p < 0.01$.
* $p < 0.05$.
*** $p < 0.001$.

There were no effects of testing session (OR = .81 [.57, 1.15]), study condition (OR = 1.06 [.74, 1.51]) or condition × age interaction (OR = .94 [.77, 1.13]) on spatial accuracy. Finally, there was a positive effect of the number of visits on spatial accuracy in the active condition (OR = 1.28 [1.06, 1.55]), but not in the yoked condition (OR = 1.01 [.83, 1.23]). Total study duration did not affect accuracy in either the active (OR = .90 [.77, 1.05]) or yoked conditions (OR = 1.13 [.96, 1.34]).

3.3.3. Study behavior in active study blocks

Participants studied an average of 31 (SD = 2.2) of the 32 objects (97%) during the active study blocks. We examined study behavior across the two active blocks using the same measures and analyses described in Experiment 1 (see Table 3). The visitation rate increased from the first block to the second block. In addition, movement distance increased with age such that older children tended to transition to objects that were more distant in the grid.

We modeled the effects of each search measure on recognition performance (proportion of studied objects that were correctly recognized) using the same mixed effects logistic regression described in Experiment 1 (see Table 4). Increased visitation rate led to lower overall accuracy in the active condition (OR = .85 [.76, .95]), but had no effect in the yoked condition (OR = 1.07 [.97, 1.19]). Increased sequence entropy also led to lower overall accuracy in the active condition (OR = .85 [.75, .97]) but had no effect in the yoked condition (OR = 1.00 [.89, 1.12]). Thus, active participants who followed more systematic search patterns and who had a lower overall rate of visits tended to have higher recognition performance. However, these effects were isolated to the active condition and had no corresponding impact on yoked performance. There were no effects of movement distance in the active (OR = .97 [.86, 1.10]) or yoked (OR = 1.10 [.98, 1.23]) conditions. We examined whether the within-subjects difference between active and yoked conditions was attributable to differences in search behavior between participants and the partners they were yoked to. This model included as predictors for the between-subjects differences the average number of visits, movement distance, and sequence entropy. The analysis showed again no relationships between these predictors and the observed advantage for active study (all $ps > .05$).

Finally, we investigated whether participants allocated study effort (i.e., more frequent visits) to objects depending on how long they were pre-exposed at the beginning of the study blocks, and whether this pre-exposure had an impact on recognition accuracy. We first used ordinal logistic regression to model the relative proportion of visits to T_{pre} = 2, 4, and 6 s objects. Then, to assess any effects of pre-exposure on recognition, we refitted the model described in Experiment 1 with additional predictors for pre-exposure time. We found no notable effects of pre-exposure time, age or condition. Details of the model and results can be found in Section 2 of the SOM.

4. Experiment 3

In Experiment 3, we examine the benefits of active control in a task similar to that used in Experiments 1 and 2, but modified to be more similar to the learning situations children encounter in school. Instead of being tasked with simply remembering a set of objects, children had to learn the French names of the same objects given in Experiments 1 and 2. In addition, in Experiment 3, we explore how the benefit of active learning is linked to the development of other cognitive resources such as executive function and working memory (Roebers, 2017). It is well documented that between five and seven years of age children undergo a dramatic improvement on tasks that measure the ability to control attention and other aspects of behavior. To explore the degree to which individual differences in executive function mediate the advantage of self-directed learning, children were administered a working memory task, the animal sorting task, that has been shown to correlate with reasoning skills, perceptual abilities and primary memory (Kray & Lindenberger, 2000). Our basic hypothesis was that individual differences in working memory may mediate the magnitude of the active learning advantage. Along this line, a recent study suggested that the relationship between children's developmental stage and their internal memory strategy knowledge may be partially mediated by working memory skills (Geurten, Catale, & Meulemans, 2016).

4.1. Participants

Because Experiment 3 required children to be able to read the presented French names, participants were slightly older than those recruited in the previous experiments. Participants were 72 6- to 11-year-old children (39 female, M_{age} = 105.35 months; SD = 15.61 months; range: 78 to 135 months) tested at the Museum of Natural History and at the FEZ Alice museum for children, in Berlin, Germany. Three additional participants were excluded due to incomplete test sessions, ten because they were studying French at school or already knew some French words, one due to an ASD diagnosis that was not disclosed before the experiment started. As in the other experiments, written informed consent was obtained from participants' parents and the local ethical review board at the Max Planck Institute for Human Development in Berlin approved the study protocol for the project "Active Learning and Memory." Children received a small present as compensation for their participation.

4.2. Materials and procedure

The materials used in Experiment 3 were identical to those used in Experiment 1 and 2. However, instead of being tasked with remembering as many of the presented objects as possible, children had to learn their French names. In both the familiarization and the study phase, the objects were arranged in a 2 × 2 (familiarization) or in a 4 × 4 (study) matrix, as in the previous experiments, but they did not disappear under occluders (see Fig. 3, top left). When children touched each object's button, the object's French label appeared, preceded by the appearance of a blue frame for 500 ms (see Fig. 3, top right). Children were instructed that, before seeing the label of another object, they had to touch the label currently displayed a second time so it would be covered by the corresponding object. There was no initial pre-exposure of the object labels.

The test phase consisted of a cued recall test for all 64 objects from

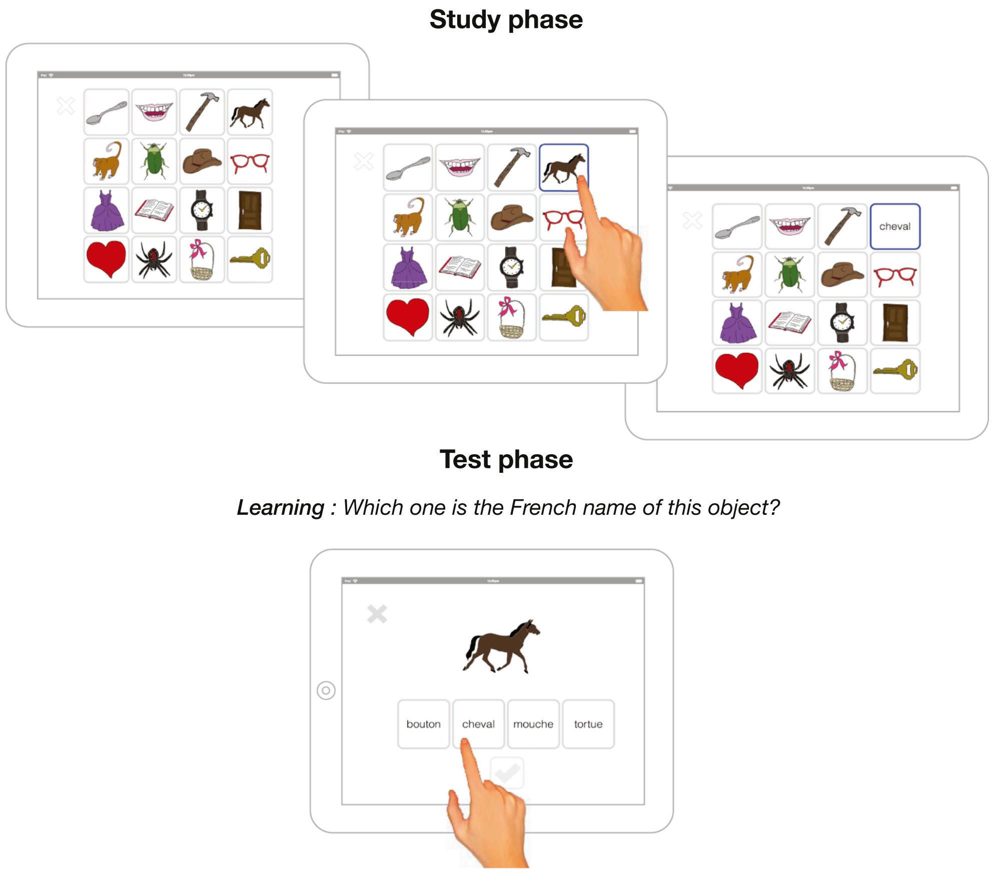

Fig. 3. Top: The objects were displayed on the screen, arranged on a 4 × 4 grid. The participant either selected a location to study (Active condition), causing a blue frame to appear, followed by the object's French label, or touched the location where the object label appeared (Yoked condition), anticipated by a blue frame. **Bottom:** During each recognition test block, participants selected among four possible alternatives the label matching the object presented.

the study phase, presented in random order. On each test trial, an object was presented together with four French labels, approximately matched for word length (see Fig. 3, bottom), among which children had to select the correct one. All the presented words (both targets and distractors) were part of the learning set.

Because children were initially tested in museums, we attempted at retesting children online: A week after the museum testing session, their parents were sent a personalized link via email and children were asked to complete the retest phase by selecting, among four possible candidates, the correct french label for the 64 objects studied in the first testing session. Instructions were provided in the text of the email to remind parents and children of the procedure. Despite the several reminders sent to parents, less than 50% of the participants completed the online retest session. Of those who did, the retest was completed between 7 and 18 days after the first testing session (M = 13.8 days; SD = 7.2 days), an extremely wide range. We therefore excluded the retest data from the analyses.

At the end of the first test session, children were administered the Animal Sorting task (Kray & Lindenberger, 2000) as a measure of working memory capacity. Participants were first asked to sort nine cards, each representing a different animal (e.g., flea, ant, snail, mouse, hamster, cat, horse, elephant, whale), according to their realistic body size (smallest to biggest). The experimenter then read aloud to the child a list of animals (increasing across trials from three to eight animals; e.g., "mouse, horse, flea"), and the child was asked to repeat the animals listed by sorting them in order of their body size (e.g., "flea, mouse, horse"). The experimenter interrupted the task after two consecutive unsuccessful trials.

4.3. Results

4.3.1. Working memory

The Animal Sorting task was scored by counting the number of correct trials before the task was interrupted (i.e., before two consecutive wrong trials). The mean WM score was 6.61 (SD = 2.1). WM was positively correlated with age, n = 69, r = .31, p = .008.

4.3.2. Cued recall

Fig. 4A shows cued recall accuracy in the active and yoked conditions as a function of age. Fig. 4B shows the within-subjects difference between active and yoked performance by participant. Mixed effects logistic regression was used to model responses ("correct" versus "incorrect") for the object labels selected during test blocks. The accuracy of cued recall responses (correct vs. incorrect) for items that appeared during study blocks were analyzed using mixed effects logistic regression. Random effects terms were included for participants and items. The model included fixed effects for age (continuous), study condition (active vs. yoked) and testing session (test vs. retest), as well as the interactions between age, condition, and testing session. The model included the number of total visits and cumulative study duration (both square root transformed) as measures of search effort per item, as well as the interaction of each measure with condition. Lastly, performance on the working memory test was included as a fixed effect. Non-categorical predictors were scaled and centered prior to model fitting. The parameters of the fitted model are shown in Table 6.

Recall accuracy was higher for objects studied in the active condition than the yoked condition in the first test session (Active: M = .43, SD = .13; Yoked: M = .39, SD = .13;OR = 1.21 [1.06, 1.37]). There was an overall positive effect of age on performance (OR = 1.21 [1.07, 1.37]), but no significant interaction between age and condition (OR = 1.05 [.92, 1.20]). Additionally, we found that performance in the working memory task was positively related to recall accuracy (OR = 1.18 [1.07, 1.30]). Additional model comparisons showed that the fit of the model did not improve when including a predictor for the interaction between working memory and condition ($\chi^2(1)$ = 2.71, p = .10) or the interaction between working memory and

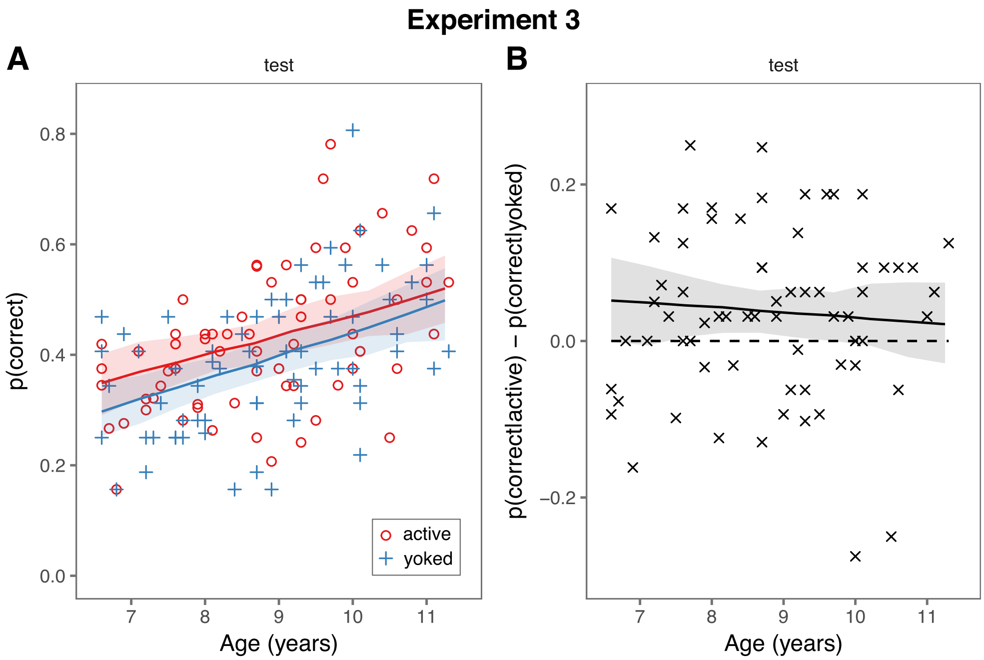

Fig. 4. Cued recall test for the objects studied in Experiment 3. Individual points represent participants, solid line is fitted regression line, and shaded regions indicate 95% confidence intervals calculated through parametric bootstrapping. **A:** Recall of labels for objects studied in active (red) and yoked (blue) conditions as a function of age. **B:** Within-subjects difference between active and yoked recall as a function of age. Dashed line indicates no difference between conditions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 6
Estimated effects from logistic regression model of cued recall in Experiment 3.

	Experiment 3
(Intercept)	−0.35 (0.09)***
Condition [yoked]	−0.19 (0.07)**
Age	0.19 (0.06)**
Working memory	0.16 (0.05)**
Half [second]	0.10 (0.07)
Study duration	0.00 (0.05)
No. visits	0.21 (0.06)***
Condition [yoked] X Age	0.05 (0.07)
Condition [yoked] X Study duration	0.07 (0.08)
Condition [yoked] X No. visits	−0.04 (0.08)
Marginal R^2	0.04
Conditional R^2	0.13

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

age ($\chi^2(1) = .29$, $p = .59$).

Object-level study behavior impacted recall performance in a similar manner for both study conditions. Object labels were more likely to be correctly identified when objects were visited more frequently in both the active ($OR = 1.22$ [1.10, 1.38]) and yoked ($OR = 1.19$ [1.07, 1.32]) conditions. There was no effect of total study duration on recall performance in either condition (Active: $OR = 1.00$ [.90, 1.11], Yoked: $OR = 1.07$ [.96, 1.20]).

4.3.3. *Study behavior in active study blocks*

During active blocks, participants studied an average of 30.7 ($SD = 2.9$) of the 32 labels presented (96%). The visitation rate increased from the first to the second active block (Table 3). There were positive main effects of age such that older children visited objects more frequently and had less systematic search patterns (i.e., with higher sequence entropy). There were no differences in movement distance between blocks or as a function of age.

Finally, we modeled the effects of average number of visits, movement distance and sequence entropy on the accuracy of cued recall of studied objects using the same mixed effects linear regression described in Experiment 1 (see Table 4), but the analysis revealed no significant effects.

5. General discussion

The present experiments investigated whether active control over study leads to advantages in memory encoding and learning across childhood. We found that episodic memory (Experiments 1 and 2) and word learning (Experiment 3) is more accurate for objects/labels studied in an active as compared to a yoked condition, where participants merely observed the active study pattern of a previous participant. This comparison carefully controls for content and timing of study materials, isolating the effects of active control on learning. Notably, the advantage from active study was not transitory: Despite the expected performance decline between test and retest, we found that the benefit of active control persisted in the delayed retest, several days after the initial study session. This lends crucial evidence that active control improves long-term retention, a principal educational aim of self-directed learning.

For older children the magnitude of the active study advantage for recognition accuracy (Exp. 1: 7.5%; Exp 2: 15%) and word learning (Exp. 3: 4%) was comparable to effects seen in adults (6% to 10%; see Markant et al., 2014). While these effects may seem modest, a 5–10% improvement in retention could be practically important when played out across an entire curriculum.

5.1. *Factors driving the emergence of the active learning advantage*

Critically, our results show that the advantage from active control emerges around age six and continues to improve into late childhood. Identifying those factors mediating the benefits of active control (Markant et al., 2016) and understanding the mechanisms underlying the development of the active advantage for episodic memory is an important goal for future work.

The later emergence of the benefit for active learning may be linked to the development of cognitive resources, including processes related to executive functioning and self-regulation (Roebers, 2017). Remarkable improvements can be observed in children's cognitive abilities during the so-called "five-to-seven year shift", particularly in their ability to exert control over their attention and behavior (i.e., executive functioning). Indeed, the results of Experiment 3 showed that working memory increased with age and was associated with improved learning in the word learning task. However, individual differences in WM could not account for the gap between active and yoked performance.

In addition to general improvements in executive function, it is crucial to investigate further how the advantage from active learning depends on children's developing metacognitive abilities. Indeed, our results are consistent with other findings documenting a developmental shift at age 6 years in the ability to control study. For example, Destan, Hembacher, Ghetti, and Roebers (2014) tasked 5- to 7-year-olds to learn the meanings of 16 Japanese characters. Children first learned the meaning of each character during a fixed-length encoding phase and were asked to provide a judgment of learning for each character, that is, to answer how sure they were that they will remember the meaning of that particular character later on, on a scale from 0 (*very unsure*) to 4 (*very sure*). Children were then given the opportunity to study the meaning of the characters again for as long as they liked and had to complete a recognition test, where they had to pick the correct meaning for each of the studied characters out of three options. Children were finally asked to provide a confidence judgment for each response on a four-point scale. Although all children gave significantly higher confidence judgments for correct answers, suggesting that 5-year-olds have already developed robust metacognitive monitoring skills, only 6- and 7-year-olds differentially allocated more study time on more uncertain items (as assessed by their own judgments of learning ratings). Metcalfe and Finn (2013b) reported a similar distinction between monitoring and control, such that both third- and fifth-graders could accurately monitor their uncertainty but only fifth-graders would use their judgments of learning to guide study.

The emergence of an advantage of active control at age six may also reflect the transition to formal schooling. In a recent longitudinal study, Brod, Bunge, and Shing (2017) followed 5-year-olds born close to the official school-entry cutoff date who did or did not enter school that year. Children who entered school showed bigger improvements in accuracy on an EF test as compared to those who stayed in kindergarten. This behavioral improvement correlated with an increase in activation of right posterior parietal cortex, involved in sustained attention, suggesting that formal schooling may be shaping brain mechanisms underlying cognitive improvements. The entrance age for school in the US, Italy and Germany is approximately 6 years. To disentangle the extent to which the benefit of active learning is driven by brain and cognitive maturation versus exposure to formal schooling, future research might replicate our design in a Western country where children enter school at age five (e.g., UK, Ireland or Australia) or seven (e.g., Bulgaria, Croatia or Finland), or implement a design similar to Brod et al.'s (2017).

The effect of active learning was robust across different types of tasks and even populations of learners of different nationalities. It would be interesting to extend our design to other active learning tasks (e.g., question-asking or spatial exploration tasks) to further probe the robustness of the active learning advantage, and to investigate whether such benefits have different onsets across tasks depending on their developmental requirements. For example, the benefits of active learning in spatial exploration may already be evident in younger children, whereas the advantage of active learning in a 20-questions game might depend on vocabulary or cognitive skills acquired only later in development (Ruggeri & Lombrozo, 2015; Ruggeri et al., 2016). In addition, although the cross-cultural robustness observed in our experiments is surprising considering that the American, Italian and German school systems differ substantially in terms of curricula, classroom policies and size, time spent at school, grading and testing methods, it would be important to replicate our design in a non-western country, to further explore the extent to which active learning and its advantage over more passive forms of instruction might be culturally mediated.

5.2. The value of active control: Adaptive search vs. agency

Our experimental approach allows a more fine-grained analysis of behavior that connects the way children study with their later memory accuracy. In our experiments, some aspects of search behavior had consistent relationships to memory performance. In particular, allocating greater study effort, either in terms of number of visits to an object or the total duration spent studying an object, led to more accurate recognition (Experiments 1 and 2), spatial recall (Experiment 2), and cued recall (Experiment 3).

We also found evidence that children change their active study strategies across blocks, favoring a faster pace and more revisits per object, and that older children tended to visit objects more frequently than younger children. This is inverse to the trend observed by Brandstatt and Voss (2014) with older adults revisiting objects less often. This suggests that general age-related memory differences might result partly from different strategies for controlling the study experience. We did not find evidence that children allocated study time depending on how long they were pre-exposed to the objects to be studied. This lack of strategy change may be due to the fact that we gave them a relatively long time for studying the objects (i.e., 90 s), such that there was enough time to visit and revisit all the objects in the array. It is possible that if children had less study time available (e.g., 30 s), they would devote more study time on the short pre-exposure objects.

The results from Experiment 2 offer some clues about the relationship between exploration in the active condition and recognition performance. Performance was higher among participants with lower visitation rates and lower sequence entropy, consistent with more regular search patterns. However, these results should be interpreted with caution, given that the same pattern was not observed in Experiment 1. In addition, older children in Experiment 3 had higher performance despite a tendency toward higher visitation rates and sequence entropy. Further work is necessary to establish how differences in search behavior relate to performance in these tasks. Importantly, however, we found that differences in search behavior within active–yoked pairs of participants did not predict the advantage of active control of study. This result speaks to one of the crucial questions left open from previous studies: The benefit of active control of study is not explained by a mismatch of expected versus actual study patterns in the yoked condition. That is, participants did not perform worse in the yoked condition just because they were tuned to the same study pattern they had implemented in the active condition.

Thus, while suggestive, differences in search behavior alone cannot explain the differences between active and yoked study across all three studies, including its developmental emergence. Indeed, the results largely support the conclusion that additional study benefited both conditions, indicating that yoked observers still had better memory for items that were revisited often or studied longer, in contrast to the results of Voss, Galvan, et al. (2011) and Voss, Warren, et al. (2011)and aligned with the results of (Markant et al., 2014).

These findings suggest that the ability to adaptively control study is subject to developmental changes. Although we did not separately measure metacognitive monitoring and control of study, this is consistent with the work described above, showing that effective control emerges later in development than accurate monitoring. However, there is another possibility unaccounted for in the present study: that the mere opportunity to exert control improves performance, independently of any changes in metacognitive processing or search behavior. Feelings of agency in other contexts has been shown to enhance motivation (Bandura, 1993; Cordova & Lepper, 1996; Leotti & Delgado, 2011; Perlmuter & Monty, 1977), episodic memory (Murty, DuBrow, & Davachi, 2015) and learning (Cordova & Lepper, 1996). Agency beliefs are also changing during this period of development and begin to predict cognitive performance between 2nd and 4th grade (Chapman, Skinner, & Baltes, 1990). Further work is needed to disentangle how perceived agency and its consequences underlie the advantage from active control of study. In this sense, implementing more fine-grained methodologies to investigate children's search strategies, such as eye-tracking or neuroimaging techniques, could offer further insights. Moreover, future work should examine more thoroughly the role of motivational aspects on the emergence of the active learning benefit for memory encoding, and on active learning effectiveness more generally. Along these lines, recent research suggests that introducing difficulties (e.g., self-testing) into the learning process can affect learning and memory in *desirable* ways (Schmidt and Bjork, 1992; Bjork, 1994). Active control might be one of those desirable difficulties, leading to superior long-term retention compared to other, supposedly less effortful or cognitively stimulating learning activities. At the same time, the desirability of such difficulty might be mediated by age and individual differences in general motivation to learn (e.g., curiosity), cognitive abilities and ability to cope with more stressful situations.

5.3. Conclusions

In this paper we demonstrated that active control of study leads to robust and long-lasting advantages in memory encoding, and that these advantages translate into enhanced learning of educationally relevant materials. One major lesson from these results is that self-directed learning may benefit older more than younger children, although it is still an open question the extent to which different factors impact the later emergence of this advantage (e.g., the necessary cognitive machinery for coordinating active information acquisition and learning; transition to formal schooling; motivation). By examining the basic learning and memory consequences of volitional control across childhood, these findings have general implications for informing educational practice and pave the way for developing more individualized school interventions.

6. Author contributions

All authors developed the study concept and contributed to the study design. D. Markant developed the software used for testing. Testing and data collection were performed by A. Ruggeri and M. Bretzke. D. Markant performed the data analysis. All authors interpreted the results. A. Ruggeri drafted the manuscript, and the other authors provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Celina Vicuna, Sana Alimohamed, Lesley Blair Winchester, Mandana Mostofi, Sarah De La Vega, Minh-Thy Nguyen, Gregor Caregnato, and Chiara Cunzolo for assistance in data collection and coding, as well as Susana Herrera and Kritika Shrestha for drawing and coloring the stimuli. This research was supported by Grant No. 1640816 from the National Science Foundation to FX, by Grant No. BCS- 1255538 from the National Science Foundation, the John Templeton Foundation "Varieties of Understanding" project, and a James S. McDonnell Foundation Scholar Award to TMG. We thank the Naturkunde Museum, Labyrinth Museum and Alice FEZ in Berlin, Germany, as well as the Koala Ludo center in Livorno, Italy, for offering testing space.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.cognition.2019.01.010.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148. https://doi.org/10.1207/s15326985ep2802_3.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting linear mixed-effects models using lme4*. arXiv preprint arXiv:1406.5823.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. Shimamura (Eds.). *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Brandstatt, K. L., & Voss, J. L. (2014). Age-related impairments in active learning and strategic visual exploration. *Frontiers in Aging Neuroscience, 6*(FEB), 1–11. https://doi.org/10.3389/fnagi.2014.00019.

Brod, G., Bunge, S. A., & Shing, Y. L. (2017). Does one year of schooling improve children's cognitive control and alter associated brain activation? *Psychological Science, 28*(7), 967–978. https://doi.org/10.1177/0956797617699838.

Bruner, J., Jolly, A., & Sylva, K. (1976). *Play - Its role in development and evolution.* New York: Basic Books.

Chapman, M., Skinner, E. A., & Baltes, P. B. (1990). Interpreting correlations between children's perceived control and cognitive performance: Control, agency, or means-ends beliefs? *Developmental Psychology, 26*(2), 246. https://doi.org/10.1037/0012-1649.26.2.246.

Chen, Z., & Klahr, D. (1999). All other things being equal: acquisition and transfer of the control of variables strategy. *Child Development, 70*(5), 1098–1120. https://doi.org/10.1111/1467-8624.00081.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science, 1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology, 88*(4), 715. https://doi.org/10.1037/0022-0663.88.4.715.

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.

Dufresne, A., & Kobasigawa, A. (1988). Developmental differences in children's spontaneous allocation of study time. *The Journal of Genetic Psychology, 149*(1), 87–92. https://doi.org/10.1080/00221325.1988.10532142.

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. https://doi.org/10.1016/0022-0965(89)90033-7.

Fandakova, Y., Shing, Y. L., & Lindenberger, U. (2013). Differences in binding and monitoring mechanisms contribute to lifespan age differences in false memory. *Developmental Psychology, 49*(10), 1822–1832. https://doi.org/10.1037/a0031361.

Feldman, A., & Acredolo, L. (1979). The effect of active versus passive exploration on memory for spatial location in children. *Child Development, 50*(3), 698. https://doi.org/10.2307/1128935.

Geurten, M., Catale, C., & Meulemans, T. (2015). When children's knowledge of memory improves children's performance in memory. *Applied Cognitive Psychology, 29*(2), 244–252. https://doi.org/10.1002/acp.3102.

Geurten, M., Catale, C., & Meulemans, T. (2016). Involvement of executive functions in children's metamemory. *Applied Cognitive Psychology, 30*(1), 70–80. https://doi.org/10.1002/acp.3168.

Ghetti, S., Castelli, P., & Lyons, K. E. (2009). Knowing about not remembering: Developmental dissociations in lack-of-memory monitoring. *Developmental Science, 13*(4), 611–621. https://doi.org/10.1111/j.1467-7687.2009.00908.x.

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*(6), 1954–1969. https://doi.org/10.1111/j.1467-8624.2011.01645.x.

Grammer, J. K., Purtell, K. M., Coffman, J. L., & Ornstein, P. A. (2011). Relations between children's metamemory and strategic performance: Time-varying covariates in early elementary school. *Journal of Experimental Child Psychology, 108*(1), 139–155. https://doi.org/10.1016/j.jecp.2010.08.001.

Gureckis, T. M., & Markant, D. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science, 7*(5), 464–481. https://doi.org/10.1177/1745691612454304.

Harman, K., Humphrey, k., & Goodale, M. (1999). Active manual control of object views facilitates visual recognition. *Current Biology, 9*(22), 1315–1318. https://doi.org/10.1016/S0960-9822(00)80053-6.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346–363. https://doi.org/10.1002/bimj.200810425.

Hutchens, R. L., Kinsella, G. J., Ong, B., Pike, K. E., Parsons, S., Storey, E., ... Clare, L. (2012). Knowledge and use of memory strategies in amnestic mild cognitive impairment. *Psychology and Aging, 27*(3), 768–777. https://doi.org/10.1037/a0026256.

Jaeger, T. F. (2008). Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007.

Kachergis, G., Rhodes, M., & Gureckis, T. (2017). Desirable difficulties in the development of active inquiry skills. *Cognition, 166*, 407–417. https://doi.org/10.1016/j.cognition.2017.05.021.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*(1), 126–147. https://doi.org/10.1037//OS82-797-U5.1.126.

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and natural experiment contexts. *Developmental Psychology, 13*(1), 9–14. https://doi.org/10.1037/0012-1649.13.1.9.

Kuhn, B. J. K. A. D., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition, 18*(4), 495–523. https://doi.org/10.1207/S1532690XCI1804.
Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological Science, 22*(10), 1310–1318. https://doi.org/10.1177/0956797611417005.
Liu, C. H., Ward, J., & Markall, H. (2007). The role of active exploration of 3d face stimuli on recognition memory of facial information. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 895. https://doi.org/10.1037/0096-1523.33.4.895.
Lockl, K., & Schneider, W. (2003). Metakognitive Überwachungs- und Selbstkontrollprozesse bei der Lernzeiteinteilung von Kindern. *Zeitschrift für Pädagogische Psychologie, 17*(3/4), 173–183. https://doi.org/10.1024//1010-0652.17.3.173.
Markant, D., DuBrow, S., Davachi, L., & Gureckis, T. M. (2014). Deconstructing the effect of self-directed study on episodic memory. *Memory & Cognition, 42*(8), 1211–1224. https://doi.org/10.3758/s13421-014-0435-9.
Markant, D., Ruggeri, A., Gureckis, T., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education, 10*(3), 142–152. https://doi.org/10.1111/mbe.12117.
MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of Child Language, 12*(2), 271–295. https://doi.org/10.1017/S0305000900006449.
McComas, J., Dulberg, C., & Latter, J. (1997). Children's memory for locations visited: importance of movement and choice. *Journal of Motor Behavior, 29*(3), 223–229. https://doi.org/10.1080/00222899709600837.
Meijer, F., & Van der Lubbe, R. H. (2011). Active exploration improves perceptual sensitivity for virtual 3d objects in visual memory. *Vision Research, 51*, 2431–2439. https://doi.org/10.1016/j.visres.2011.09.013.
Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning? Journal of experimental psychology. *General, 131*(3), 349–363. https://doi.org/10.1037/0096-3445.131.3.349.
Metcalfe, J., & Finn, B. (2013a). Metacognition and control of study choice in children. *Metacognition and Learning, 8*(1), 19–46. https://doi.org/10.1007/s11409-013-9094-7.
Metcalfe, J., & Finn, B. (2013b). Metacognition and control of study choice in children. *Metacognition and Learning, 8*(1), 19–46. https://doi.org/10.1007/s11409-013-9094-7.
Montessori, M. (1912/1964). *The Montessori Method.* New York: Schocken.
Murty, V. P., DuBrow, S., & Davachi, L. (2015). The simple act of choosing influences declarative memory. *The Journal of Neuroscience, 35*(16), 6255–6264. https://doi.org/10.1523/JNEUROSCI.4181-14.2015.
Partridge, E., Mcgovern, M. G., Yung, A., & Kidd, C. (2015). Young children ' s self-directed information gathering on touchscreens. In J. Y. R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, & A. Warlaumont (Eds.). *Proceedings of the 37th annual conference of the cognitive science society*. Austin, TX: Cognitive Science Society.
Perlmuter, L. C., & Monty, R. A. (1977). The importance of perceived control: Fact or fantasy? experiments with both humans and animals indicate that the mere illusion of control significantly improves performance in a variety of situations. *American Scientist, 65*(6), 759–765 http://www.jstor.org/stable/27848177.
Piaget, J. (1930). *The child's conception of physical causality.* New York: Harcourt, Brace.
Plancher, G., Barra, J., Orriols, E., & Piolino, P. (2013). The influence of action on episodic memory: A virtual reality study. *The Quarterly Journal of Experimental Psychology, 66*(5), 895–909. https://doi.org/10.1080/17470218.2012.722657.
Poag, C. K., Cohen, R., & Weatherford, D. L. (1983). Spatial representations of young children: The role of self- versus adult-directed movement and viewing. *Journal of Experimental Child Psychology, 35*(1), 172–179. https://doi.org/10.1016/0022-0965(83)90077-2.
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.
Ruggeri, A., Markant, D. B., Gureckis, T. M., Bretzke, M., & Xu, F. (2019). *Memory enhancements from active control of learning emerge across development.* doi:https://doi.org/10.17605/OSF.IO/87M9E.
Ruggeri, A., & Feufel, M. A. (2015). How basic-level objects facilitate question-asking in a categorization task. *Frontiers in Psychology, 6*(July), 1–13. https://doi.org/10.3389/fpsyg.2015.00918.
Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition, 143*, 203–216. https://doi.org/10.1016/j.cognition.2015.07.004.
Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology, 52*(12), 2159–2173. https://doi.org/10.1037/dev0000240.
Ruggeri, A., Sim, Z. L., & Xu, F. (2017). Why is Toma late to school again? Preschoolers identify the most informative questions. *Developmental Psychology, 53*(9), 1620–1632. https://doi.org/10.1037/dev0000340.
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–218. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x.
Shannon, C. E. (1948). A mathematical theory of communication, part i, part ii. *Bell System Technical Journal, 27*, 623–656.
Sim, Z. L., Tanner, M., Alpert, N., & Xu, F. (2015). Children learn better when they select their own data. In J. Y. R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, & A. Warlaumont (Eds.). Austin, TX: Cognitive Science Society.
Voss, J., Galvan, A., & Gonsalves, B. (2011). Cortical regions recruited for complex active-learning strategies and action planning exhibit rapid reactivation during memory retrieval. *Neuropsychologia, 49*, 3956–3966. https://doi.org/10.1016/j.neuropsychologia.2011.10.012.
Voss, J., Gonsalves, B., Federmeier, K., Tranel, D., & Cohen, N. (2011). Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. *Nature Neuroscience, 14*(1), 115–120. https://doi.org/10.1038/nn.2693.
Voss, J., Warren, D., Gonsalves, B., Federmeier, K., Tranel, D., & Cohen, N. (2011). Spontaneous revisitation during visual exploration as a link among strategic behavior, learning, and the hippocampus. *Proceedings of the National Academy of Sciences, 108*(31), E402–E409.