Memory enhancements from active control of learning emerge across development

Azzurra Ruggeri

MPRG iSearch, Max Planck Institute for Human Development & School of Education, Technical
University Munich

Douglas B. Markant

Department of Psychological Science, University of North Carolina at Charlotte

Todd M. Gureckis

Department of Psychology, New York University

Fei Xu

Department of Psychology, University of California, Berkeley

Abstract

This paper investigates whether active control of study leads to enhanced learning in 5- to 10-year-olds. Children played a simple memory game. In Experiments 1 and 2, the goal was to remember as many as possible from 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 suggest that the ability to effectively control study emerges and develops during early childhood and leads to memory benefits over a week delay.

Keywords: active learning, recognition memory, exploration, cognitive development

Memory enhancements from active control of learning emerge across development

Introduction

Educators frequently argue that self-directed, *active learning* situations foster deeper learning (Bruner, Jolly, & Sylva, 1976; Kuhn & Kaplan, 2000; Montessori, 1912/1964; Piaget, 1930). Although this idea often draws from a constellation of interrelated issues, like motivation or deeper processing, many argue that giving students some degree of independent 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, a number of recent experiments 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 study situations lacking this control (Harman, Humphrey, & Goodale, 1999; Voss, Warren, et al., 2011; Voss, Galvan, & Gonsalves, 2011; Voss, Gonsalves, Federmeier, Tranel, & Cohen, 2011; Liu, Ward, & Markall, 2007; Meijer & Van der Lubbe, 2011; Markant, DuBrow, Davachi, & Gureckis, 2014; Plancher, Barra, Orriols, & Piolino, 2013).

While provocative, it is unclear how these adult findings might inform educational policies which 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—most of which are subject to critical developmental changes (Kachergis, Rhodes, & Gureckis, 2017). In this sense, one possibility is that learners can more effectively leverage self-directed learning as they mature. This paper presents a series of laboratory experiments designed to trace the emergence and developmental trajectory of self-directed learning as a successful learning modality for children.

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, Warren, et al., 2011; Voss, Galvan, & Gonsalves, 2011; Voss, Gonsalves, 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” 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 study on learning and memory. These studies have found robust benefits of active study on object recognition, meaning that participants were more accurate at recognizing objects that had been actively studied from novel objects 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).

These studies further revealed that the benefits of active study depended on *how* participants studied the objects. Objects studied for longer durations (Voss, Galvan, & Gonsalves, 2011) and revisited within a short period of time (Voss, Gonsalves, et al., 2011) were more likely to be recognized, but only in the active learning condition. Similar results have been obtained with a variety of related tasks and materials (Harman et al., 1999; Liu et al., 2007; Meijer & Van der Lubbe, 2011; Markant et al., 2016; Plancher et al., 2013) and with both younger and older adults (Brandstatt & Voss, 2014).

A few recent studies have shown that active control can facilitate learning in children as well. For example, Sim, Tanner, Alpert, and Xu (2015) provided evidence that 7-year-olds learn categorical rules more effectively when they are free to decide what information to gather, as compared to yoked observation. 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 in learning novel object-word pairings (Partridge, McGovern, Yung, & Kidd, 2015). Do these advantages reflect developmental progress along the way to adult competence, or a robust and universal advantage of self-directed learning?

The benefit of active control is expected to heavily depend on children’s metacognitive abilities. Metamemory monitoring, that is, the ability to introspect on the accuracy of one’s memories, improves considerably over the elementary school years, 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 (Metcalf, 2002; Metcalf & Finn, 2013). For instance, although 6-graders demonstrate sensitivity to the strength of their own memories, they are

inefficient in controlling how long to study particular objects in order to achieve the best level of recall (Metcalf, 2002). 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.

The present paper compares the effects of active and yoked study on 5- to 10-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 replicates and extends the results from Experiment 1 to further explore the effects of differential pre-exposure on children's studying strategies. Experiment 3 explores the effects of active study in a paired associate learning task in which children have to learn the French names for the same objects used in Experiments 1 and 2.

Experiment 1

Participants

Participants in Experiment 1 were 52 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. Ten participants (19%) did not show up for the retest session, but the data from their first test session were included in the analyses. Sample size was determined based on previous active learning studies with adults (see Markant et al., 2014) and children (Metcalf & Finn, 2013; Partridge et al., 2015; Ruggeri & Lombrozo, 2015; Ruggeri, Lombrozo, Griffiths, & Xu, 2016).

Materials

To ensure young children were familiar with the 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, as recorded by the CHILDES corpus (Child Language Data Exchange System; MacWhinney and Snow, 1985). 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 used as stimuli during the first and second experimental sessions (see Figure S1 of the Supplementary Online Materials, SOM). The experimental materials were presented on a 9.7" iPad touchscreen using custom software.

Figure 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 appeared during study.

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).

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 2x2 grid. The objects were shown on the screen for two seconds before disappearing under occluders (same as for the main experimental session, see Figure 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 would reveal 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.

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 4x4 grid. All 16 objects were visible on the screen for 2 seconds at the beginning of each study block, before disappearing under occluders (see Figure 1, top). Across the four blocks, children were asked to memorize 64 objects (see Materials section and Figure S1 of the SOM). In the active blocks, children had 90 seconds to select and study the objects in order to memorize them. In the yoked blocks, children were presented with the 90-second study sequence (i.e., same objects and pacing) of one of the previous

participant's active learning blocks. In between blocks, there was a 20-second break in which children were briefly reminded of the study procedure for the next block.

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 4x4 grid (see Figure 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 in each block was randomly varied between 1 and 15. 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 arranged 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 ($M = 8.26$ days; $SD = 2.63$ days), 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.

Results

Recognition of studied objects. Mixed effects logistic regression was used to model recognition responses ("old" versus "new") for objects presented during the study phase. Mixed effects models were fit using random intercepts for participants and objects (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). A full description of the model can be found in the Section S1 of the SOM (see also Table S1). Effect sizes are reported as relative odds ratios (*OR*) that indicate the multiplicative change in the odds of recognizing a studied item that is associated with a unit change in a given predictor. A 95% confidence interval that excludes $OR = 1$ indicates a significant effect. An analysis of false alarms, that is, the number of objects

Figure 2. Recognition accuracy for the 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.

marked as recognized at test that had not been presented during the study blocks, is presented in the Section S2 of the SOM. This analysis shows that the rate of false alarms was relatively low ($M = .09$, $SD = .09$) and did not differ as a function of condition or age.

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 ($OR = .62$ [$.55, .70$]) for both conditions.

There was no overall effect of age on recognition ($OR = 1.06$ [$.87, 1.30$]), but there were significant age $\times$ condition ($OR = .80$ [$.71, .91$]) and age $\times$ testing session ($OR = 1.31$ [$1.16, 1.49$]) interactions: The

negative effect of yoked study was larger for older children, and the decrease in performance in the retest was smaller for older children (see Figure 2A). Figure 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 1 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 is 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 sessions ($OR = 1.13 [.93, 1.37]$). At age 8, actively studied objects are more likely to be recognized than yoked objects in both the test ($OR = 1.66 [1.35, 2.05]$) and retest ($OR = 1.68 [1.35, 2.10]$). This 2-years age difference is 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.

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

Table 1

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	.69 (.16)	.69 (.14)	-.004 (.15)
	Retest	.54 (.15)	.52 (.18)	.02 (.12)
Older (age ≥ 7 years)	Test	.73 (.16)	.64 (.16)	.09 (.15)
	Retest	.68 (.17)	.61 (.17)	.07 (.10)
Experiment 2				
Younger (age < 6.75 years)	Test	.59 (.21)	.56 (.22)	.03 (.14)
	Retest	.44 (.23)	.37 (.20)	.07 (.19)
Older (age ≥ 6.75 years)	Test	.83 (.13)	.67 (.17)	.16 (.16)
	Retest	.76 (.18)	.62 (.21)	.13 (.18)

behavior presented above, we examined how participants explored the grid during active blocks and whether their search 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 object¹, averaged across all objects within a block; 2) mean movement distance between successive 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 x 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).

¹We 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 2
Analysis of study behavior during active blocks.

	First block <i>M(SD)</i>	Second block <i>M(SD)</i>	Block <i>F</i>	<i>p</i>	Age <i>F</i>	<i>p</i>	Block X Age <i>F</i>	<i>p</i>
Visitation rate								
Exp 1	1.90 (.52)	2.32 (.63)	49.4	<.001*	4.90	.03*	.14	.71
Exp 2	2.14 (.56)	2.42 (.51)	17.2	<.001*	2.28	.14	.07	.79
Exp 3	1.31 (.47)	1.44 (.47)	8.65	<.01*	20.	<.001*	1.32	.25
Movement distance								
Exp 1	1.43 (.23)	1.38 (.24)	1.56	.22	3.06	.09	.38	.54
Exp 2	1.34 (.20)	1.38 (.21)	2.98	.09	8.04	.01*	.87	.34
Exp 3	1.47 (.18)	1.49 (.15)	.22	.64	.37	.54	0.	.98
Sequence entropy								
Exp 1	1.54 (.44)	1.61 (.49)	.92	.34	3.30	.08	2.77	.10
Exp 2	1.38 (.50)	1.35 (.46)	.11	.73	3.33	.07	2.44	.12
Exp 3	1.15 (.55)	1.29 (.61)	3.19	.08	5.99	.02*	.22	.64

Note: The denominator degrees of freedom for the F tests were 49 (Experiment 1), 47 (Experiment 2), and 67 (Experiment 3). * $p < .05$, ** $p < .01$.

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 2). 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 (see Section S3 of SOM). 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 $ps > .05$).

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 have 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, et al. (Voss, Warren, et al., 2011; Voss, Galvan, & Gonsalves, 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).

Participants

Participants in Experiment 2 were 49 4- to 10-year-old children (25 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.

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 comprised of 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 Figure 1, bottom right).

As in Experiment 1, after about one week ($M = 5.93$ days; $SD = 2.62$ days), participants were tested in a second session, in which they were asked to complete 8 new test blocks.

Results

Recognition of studied objects. Recognition responses to studied objects were modeled using the same method as in Experiment 1 (see Section S1 of the SOM). False alarm rates were again relatively low

($M = .08$, $SD = .15$; see Section S2 of the SOM for the full analysis).

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: $M = .50$, $SD = .25$; $OR = 1.91$ [$1.58, 2.31$]). Recognition accuracy declined from the test to the retest ($OR = .53$ [$.46, .61$]), for both the active and the yoked conditions.

There was a positive overall effect of age on recognition ($OR = 2.00$ [$1.54, 2.60$]) and a positive age $\times$ testing session interaction ($OR = 1.17$ [$1.02, 1.34$]), such that the decline in performance in the retest was smaller for older children (Figure 2C). In addition, there was a negative age $\times$ condition interaction ($OR = 0.76$ [$0.66, 0.87$]), such that the negative effect of yoked study was larger for older children (Figure 2D). Table 1 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.

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 (see Section S4 of the SOM for full details of the model). 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 from 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 ($OR = 1.77$ [$1.42, 2.18$]).

There were no effects of testing session ($OR = .81$ [$.57, 1.15$]), study condition

($OR = 1.06 [.74, 1.51]$) or condition $\times$ 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 ($OR = 1.13 [.96, 1.34]$).

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 2). 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; see Section S3 of SOM). 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 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 refit the model described in SOM Section S1 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 5 of the SOM.

Experiment 3

The memory task used of Experiment 1 and 2, while well controlled, is somewhat artificial and divorced from the learning situations children encounter in school. In Experiment 3 we examine the benefits of active control in a task where children are tasked to learn the French labels of given objects. Children were additionally given a working memory task, to explore whether individual differences in working memory mediate the benefit of active control on learning.

Participants

Because Experiment 3 required children to be able to read the presented French labels, 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 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 knew some French words, one because affected by Asperger's syndrome.

Materials and procedure

The materials and design 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 labels.

In both the familiarization and the study phase the objects were arranged in a 2x2 (familiarization) or in a 4x4 (study) matrix, as in the previous experiments, but they did not disappear under occluders (see Figure 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 Figure 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 the study phase, presented in random order. On each test trial, an object was presented together with four French labels, approximately matched on word length (see Figure 3, bottom), among which children had to select the correct one.

Since children were initially tested in museums, we attempted retesting children online: A week following the museum testing session, their parents were sent a personalized link via email and 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 decided to exclude 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").

Results

Cued recall. Mixed effects logistic regression was used to model responses ("correct" versus "incorrect") for the object labels selected during test blocks. Figure 5A shows cued recall accuracy in the active and yoked conditions as a function of age. Figure 5B shows the within-subjects difference between active and yoked performance by participant. A full description of the model can be found in Section S6 of the SOM (see also Table S6). 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 ($OR = 1.18$ [$1.07, 1.30$]).

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

Figure 3. Top: The objects were displayed on the screen, arranged on a 4x4 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.

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: $OR = 1.07 [.96, 1.20]$).

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 active block to the second active block (Table 2). 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 mixed effects linear regression (see Section

Figure 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.

S3), but the analysis revealed no significant effects.

General discussion

The present paper investigated whether active control over the study experience 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 observe the active study pattern of a previous participant. This comparison carefully controls for content and timing of study materials, isolating the effects of active learning.

We found that the advantage from active control emerges only around age six, and continues to improve into late childhood. Identifying the exact mechanisms underlying this developmental change requires further study (e.g., it might reflect formal schooling or developmental differences in cognitive resources). However, our results do suggest that self-directed study is not a universally advantageous learning strategy. 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 the effect found with 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.

The advantage for active learning was not short-lived and transitory. Despite the expected performance decline between test and retest, we found that the benefit of active control persisted a week after the initial study session. This is crucial because improving long-term retention is one of the educational aims of self-directed learning. The effect is also extremely robust across different types of tasks and even populations of learners, given our international samples.

The experimental approach we adopted allows a more fine-grained analysis of behavior that helps us connect the *way* children study and their later memory accuracy. Overall, revisiting objects reliably led to more accurate recognition. In contrast to the results of Voss et al. (2011a; 2011c) and aligned with the results of (Markant et al., 2014), we found that these effects were mostly consistent across study conditions. We additionally found evidence that children change their active study strategies across blocks, favoring a faster pace and more revisits per item. We also found 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. It is possible that if children have had less study time available (e.g., 30 s), they might have devoted more study time on the short pre-exposure objects.

The results from Experiment 2 also suggest a general relationship between search systematicity in the active study condition, as measured by sequence entropy, and recognition performance. Most interestingly, we found that sequence entropy does not mediate the active learning advantage. 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 do not perform worse in the yoked condition just because they were *tuned* to the same study pattern they had implemented in the active condition.

In conclusion, in this paper we demonstrated that active control of study leads to robust and long-lasting advantages in memory encoding starting from age six, and that these advantages translate into enhanced learning of educationally relevant materials. These findings contribute to the discussion on best educational practices by examining the basic learning and memory consequences of volitional control from a developmental perspective. One major lesson from these results is that self-directed learning may benefit older children more than younger ones, as the necessary cognitive machinery for coordinating active information acquisition and learning becomes available.

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.
- 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. doi: 10.3389/fnagi.2014.00019
- Bruner, J., Jolly, A., & Sylva, K. (1976). *Play - Its role in development and evolution*. New York: Basic Books.
- Chi, M. T. H. (2009, jan). Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science*, 1(1), 73–105. Retrieved from <http://doi.wiley.com/10.1111/j.1756-8765.2008.01005.x> doi: 10.1111/j.1756-8765.2008.01005.x
- 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. Retrieved from <http://doi.apa.org/getdoi.cfm?doi=10.1037/a0031361> doi: 10.1037/a0031361
- Feldman, A., & Acredolo, L. (1979, sep). The Effect of Active versus Passive Exploration on Memory for Spatial Location in Children. *Child Development*, 50(3), 698. doi: 10.2307/1128935
- Ghetti, S., Castelli, P., & Lyons, K. E. (2009, aug). Knowing about not remembering: developmental dissociations in lack-of-memory monitoring. *Developmental Science*, 13(4), 611–621. Retrieved from <http://doi.wiley.com/10.1111/j.1467-7687.2009.00908.x> doi: 10.1111/j.1467-7687.2009.00908.x
- Ghetti, S., Mirandola, C., Angelini, L., Cornoldi, C., & Ciaramelli, E. (2011, nov). Development of Subjective Recollection: Understanding of and Introspection on Memory States. *Child Development*, 82(6), 1954–1969. Retrieved from <http://doi.wiley.com/10.1111/j.1467-8624.2011.01645.x> doi: 10.1111/j.1467-8624.2011.01645.x
- Gureckis, T. M., & Markant, D. (2012). Self-Directed Learning: A Cognitive and Computational Perspective. *Perspectives on Psychological Science*, 7(5), 464–481.
- Harman, K., Humphrey, k., & Goodale, M. (1999). Active manual control of object views facilitates visual

- recognition. *Current Biology*, 9(22), 1315–1318.
- 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.
- Kachergis, G., Rhodes, M., & Gureckis, T. (2017). Desirable difficulties in the development of active inquiry skills. *Cognition*, 166, 407–417.
- Kray, J., & Lindenberger, U. (2000, mar). Adult age differences in task switching. *Psychology and aging*, 15(1), 126–47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10755295> doi: 10.1037//OS82-797-U5.1.126
- Kuhn, B. J. K. A., D., & Kaplan, D. (2000). The Development of Cognitive Skills To Support Inquiry Learning. *Cognition*, 18(4), 495–523. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/S1532690XCI1804_4#.VeBsh0YprQQ doi: 10.1207/S1532690XCI1804
- 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.
- 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.
- Markant, D., Ruggeri, A., Gureckis, T., & Xu, F. (2016, sep). Enhanced Memory as a Common Effect of Active Learning. *Mind, Brain, and Education*, 10(3), 142–152. Retrieved from <http://doi.wiley.com/10.1111/mbe.12117> doi: 10.1111/mbe.12117
- McComas, J., Dulberg, C., & Latter, J. (1997, sep). Children's Memory for Locations Visited: Importance of Movement and Choice. *Journal of Motor Behavior*, 29(3), 223–229. doi: 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.
- Metcalfe, J. (2002, sep). Is study time allocated selectively to a region of proximal learning? *Journal of experimental psychology. General*, 131(3), 349–63. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12214751> doi: 10.1037/0096-3445.131.3.349
- Metcalfe, J., & Finn, B. (2013, apr). Metacognition and control of study choice in children. *Metacognition*

- and Learning*, 8(1), 19–46. Retrieved from
<http://link.springer.com/10.1007/s11409-013-9094-7> doi: 10.1007/s11409-013-9094-7
- Montessori, M. (1912/1964). *The Montessori Method*. New York: Schocken.
- 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 (Ed.), *Proceedings of the 37th annual conference of the cognitive science society*. Austin, TX. Cognitive Science Society.
- 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.
- Poag, C. K., Cohen, R., & Weatherford, D. L. (1983, feb). Spatial representations of young children: the role of self- versus adult-directed movement and viewing. *Journal of experimental child psychology*, 35(1), 172–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6827216>
- Ruggeri, A., & Lombrozo, T. (2015, oct). Children adapt their questions to achieve efficient search. *Cognition*, 143, 203–216. Retrieved from
<http://linkinghub.elsevier.com/retrieve/pii/S0010027715300317> doi:
 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. doi:
 10.1037/dev0000240
- Shannon, C. E. (1948). A mathematical theory of communication, part i, part ii. *Bell Syst. Tech. J.*, 27, 623-656. Retrieved from <http://ci.nii.ac.jp/naid/10014857095/en/>
- 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 (Ed.), . 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.
- 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.

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.

Supplementary Online Materials

Figure S1. Study/Test objects used in Experiments 1, 2 and 3

S1. Recognition

Recognition responses (“old” versus “new”) for items that appeared during study blocks were analyzed using mixed effects logistic regression. 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, that is, the number of objects recognized in the test blocks that had not been presented in the study blocks, 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.

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 models are shown in Table S1. The classification rate of the best-fit model was .73 for Experiment 1 and .76 for Experiment 2. Contrast-based hypothesis tests were conducted using the *multcomp* library in *R*. Effect sizes and 95% confidence intervals for significant effects are reported in terms of relative odds ratio (*OR*) associated with each comparison. 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 (Exp. 1: $\chi^2(1) = .11, p = .74$; Exp. 2: $\chi^2(1) = .02, p = .89$).

	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) ^{***}

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S1

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

S2. False alarms

A mixed Logit model was fit to false alarms, which included fixed effects for age (continuous), testing session (test vs. retest), and the interaction between age and testing session. A random intercept term was included for participants. The resulting coefficients are shown in Table S2. In Experiment 1, false alarm rates did not differ between the test ($M = .08$, $SD = .12$) and the retest ($M = .09$, $SD = .11$). There were no significant effects of age, test, or age $\times$ test interaction. In Experiment 2, false alarms increased from test to retest (test: $M = .07$, $SD = .17$; retest: $M = .09$, $SD = .17$; $OR = 1.71$ [$1.31, 2.22$]). There were no effects of age or the age $\times$ test interaction.

	Experiment 1	Experiment 2
(Intercept)	-2.96 (0.19)***	-3.99 (0.31)***
Age	-0.06 (0.18)	-0.49 (0.30)
Test[retest]	0.19 (0.10)	0.53 (0.13)***
Age X Test	0.09 (0.09)	0.16 (0.13)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S2

Estimated effects from logistic regression model of false alarms.

S3. Aggregate study behavior

A mixed effects logit model was used to evaluate the affects of aggregate study behavior on overall recognition performance (the proportion of correctly recognized items from the study phase). In addition to predictors for condition, age, test session, and working memory score (Experiment 3 only), the model included predictors for three measures of aggregate study behavior: visitation rate, movement distance, and sequence entropy (see main text for description). 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 S3.

	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)***	−0.59 (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)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S3

Estimated effects of search behavior on recognition.

S4. Spatial memory of recognized objects (Experiment 2)

The accuracy of spatial recall judgments (correct vs. incorrect) for recognized items was modeled using mixed effects logistic regression. 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 fixed effects used to model recognition responses (Section S1) were included here. In addition, the model included a fixed effect of 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 S4.

	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) [*]

^{***} $p < 0.001$, ^{**} $p < 0.01$, ^{*} $p < 0.05$

Table S4

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

S5. Pre-exposure manipulation (Experiment 2)

In Experiment 2, items were pre-exposed for $T_{pre} = 2$ s, 4 s, or 6 s. We examined whether the proportion of visits to items within each level of pre-exposure differed as a function of age or study block ². 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 (Figure S2). The model included fixed effects for age (continuous) and study block (first vs. second). A random intercept term was included for participants. The resulting coefficients are shown in Table S5. This analysis revealed a significant change in the proportion of visits from the first active block to the second active block ($OR = .93$ [.87, .99]). Compared to the first block ($T_{pre} = 2$: $M = .27$, $SD = .05$, $T_{pre} = 4$: $M = .31$, $SD = .06$, $T_{pre} = 6$: $M = .42$, $SD = .07$), in the second block a smaller proportion of visits were made to objects exposed for a short duration while more visits were made to those exposed for the longer duration ($T_{pre} = 2$: $M = .23$, $SD = .05$, $T_{pre} = 4$: $M = .31$, $SD = .04$, $T_{pre} = 6$: $M = .46$, $SD = .07$). There were no additional effects of age or its interaction with study block on relative visitation rate.

To assess any effects of pre-exposure on recognition, we additionally refit the model described in SOM Section S1 with additional predictors for pre-exposure time, the interaction between pre-exposure time and age, and the interaction between pre-exposure time and condition. There were no effects of pre-exposure time ($OR = 1.02$ [.92, 1.12]), pre-exposure time $\times$ age ($OR = 1.02$ [.95, 1.09]), or pre-exposure time $\times$ condition ($OR = 1.02$ [.89, 1.15]) on recognition performance.

	Experiment 2
Age	−0.02 (0.03)
Study block[second]	−0.08 (0.03)*
Age*Study block[second]	−0.04 (0.03)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S5

Estimated effects from ordered logit regression model of visitation rate to items of varying pre-exposure.

²An analysis of median study duration produced a similar pattern of results so we only report the results for proportion of visits here.

Figure S2. Proportion of visits allocated to items as a function of pre-exposure duration. Dashed line indicates the proportion expected from random search. Due to the different numbers of items in each pre-exposure category, random search would be associated with an increase in visits with increasing pre-exposure time. Solid black lines indicate the average proportion allocated by participants.

S6. Cued recall (Experiment 3)

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 S6.

	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)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table S6

Estimated effects from logistic regression model of cued recall in Study 3.