Received: 14 May 2019 | Revised: 5 March 2020 | Accepted: 11 June 2020

DOI: 10.1111/desc.13017

PAPER

Developmental Science WILEY

The value of proactive goal setting and choice in 3- to 7-year-olds' use of working memory gating strategies in a naturalistic task

Livia Freier[1] | **Pankaj Gupta**[1] | **David Badre**[1,2] | **Dima Amso**[3]

[1]Department of Cognitive, Linguistic, and Psychological Science, Brown University, Providence, RI, USA

[2]Brown Institute for Brain Sciences, Brown University, Providence, RI, USA

[3]Psychology Department, Columbia University, New York, NY, USA

Correspondence
Dima Amso, Department of Cognitive, Linguistic, and Psychological Science, Brown University, Providence, RI, USA.
Email: dima_amso@brown.edu

Funding information
National Institutes of Health, Grant/Award Number: R01 MH099078; ONR MURI, Grant/Award Number: N00014-16-1-2832

Abstract

Rule-guided behavior depends on the ability to strategically update and act on content held in working memory. Proactive and reactive control strategies were contrasted across two experiments using an adapted input/output gating paradigm (Neuron, 81, 2014 and 930). Behavioral accuracies of 3-, 5-, and 7-year-olds were higher when a contextual cue appeared at the beginning of the task (input gating) rather than at the end (output gating). This finding supports prior work in older children, suggesting that children are better when input gating but rely on the more effortful output gating strategy for goal-oriented action selection (Cognition, 155, 2016 and 8). A manipulation was added to investigate whether children's use of working memory strategies becomes more flexible when task goals are specified internally rather than externally provided by the experimenter. A shift toward more proactive control was observed when children chose the task goal among two alternatives. Scan path analyses of saccadic eye movement indicated that giving children agency and choice over the task goal resulted in less use of a reactive strategy than when the goal was determined by the experimenter.

KEYWORDS

choice, cognitive control, cognitive development, working memory

1 | INTRODUCTION

1.1 | Overview

Children's everyday behavior is highly contextual. For instance, the choice of what items to wear may vary considerably depending on the weather, the time of the day, and the day of the week. Working memory (WM) resources are limited in capacity. Thus, children must use some strategy to filter choice options in a manner consistent with current goals (e.g., Awh, Vogel, & Oh, 2006; Braver & Cohen, 2000; Chatham, Frank, & Badre, 2014; Fallon, Zokaei, & Husain, 2016; Frank, Loughry, & O'Reilly, 2001; Gruber, Dayan, Gutkin, & Solla, 2006; Sörqvist, Stenfelt, & Rönnberg, 2012).

One strategy that can be used to achieve a distant goal is to update into WM only items relevant to the goal, eliminating irrelevant items as distraction. This is called selective *input gating* (e.g., O'Reilly & Frank, 2006). A second strategy is to accumulate all information and to select, from the now cumulative contents of WM, only items relevant to one's goal. This is called *output gating* (e.g., Chatham et al., 2014). Output gating as a control strategy places more demands on WM memory (more items actively maintained at once) and attentional selection (more distractors from which to choose the target goal response). Here, we examine the development of both WM updating strategies in 3- to 7-year-old children, and whether one strategy is more commonly used to guide goal-oriented action.

Few studies have examined how the temporal dynamics of goal setting impact children's performance on cognitive control tasks. Munakata, Snyder, and Chatham (2012) argued that there is a developmental shift from reactive to proactive cognitive control in childhood. Even when children can use a proactive preparatory strategy, they tend to be reactive. For example, Chatham, Frank, and Munakata (2009) found that 3.5- but not 8-year-old children often use reactive approaches even when a more efficient proactive strategy is possible. They used an AX-CPT task, where children were asked to make a target response to the letter X only if it is preceded by the letter A, allowing proactive control, or control preparation, when any letter A is presented. Responses to nontarget stimuli, including "AY", "BX", and "BY" sequences (30% of trials) revealed that 8-year-olds, but not 3.5-year-olds, made fewer "BX" than "AX" errors because they were able to prepare for the "X" when presented with "A" cues. The results of this study suggest that older children use a more proactive strategy than 3.5-year-olds when they have learned that a context "A" strongly predicts a target response "X".

While proactive and reactive control tasks focus on the temporal dynamics of control (Chevalier, Chatham, & Munakata, 2014; Doebel, Barker, Chevalier, Michaelson, Fisher, & Munakata, 2017), they do not explicitly manipulate which input and output gating strategy is used. The aim of the current study was to investigate the development of WM gating strategies using a task that manipulates the demands for input and output gaiting. In order to measure gating strategies, we measured both behavioral accuracy, and in a subset of children, used scan path analyses of eye tracking data while children were selecting the target to make inferences about which input or output gating strategies (or a mixture) contributed to performance. Based on reviewed work, we predicted that even though input gating will be more efficient than output gating, children will rely on inefficient output gating strategies. We also probed conditions in which younger children might be biased to use an input gating strategy. We predicted that if children were given the opportunity to *choose* a context from two alternatives, input gating performance would improve as a result of strengthened task representations in WM, thus preventing the use of an inefficient output gating strategy. In the following sections we review the literature on WM gating and its measurement and on the value of choice in decision making.

1.1.1 | WM input and output gating

Input and output gating strategy use has been studied using rule-guided behavior tasks in adults (e.g., Badre & Frank, 2012; Badre, Kayser, & D'Esposito, 2010; Bhandari & Badre, 2018; Chatham et al., 2014; Frank & Badre, 2012) and children aged 7–17 years (e.g., Unger, Ackerman, Chatham, Amso, & Badre, 2016). In a version of the task adapted in this work, participants are taught a single rule but are given an ordering of contexts and items consistent with either an input or output gating strategy. For example, the rule might state that a number stimulus, or context, cues whether a letter, a symbol, or both will be the target for response. Next, a number, letter, or symbol is presented sequentially, and the items are maintained in WM until participants are prompted for a response. At the prompt, participants press a button to identity the correct target based on the context. By varying when the context is presented, before or after the items, this paradigm allows manipulation of the demands for input and output gating WM strategies separately. On context First trials (CF), the context is presented before the possible target options, allowing for the updating of only the relevant lower level item to memory, either the letter, the symbol, or both, while ignoring the irrelevant one (input gating). In contrast, on context Last trials (CL), the lower level letters and symbols are presented first, and the context that determines which target to select is presented last. This CL manipulation requires participants to hold all lower level items in WM until context is presented. Then, participants select the relevant target from within WM (output gating).

Research highlights

- Children use more proactive control when choosing a task goal among two alternatives.
- Scan path analyses of saccadic eye movement suggest that goal setting increases the efficiency of working memory gating strategies.

When the context is presented last (CL), only an output gating strategy is possible. When the context is presented first (CF), children could use the input gating approach, could rely on output gating, or some mixture of the two strategies on a trial-by-trial basis. Unger et al. (2016) found that input gating performance was more efficient than output gating earlier in development. Output gating, measured on CL trials, increased individual response times relative to when the context was presented first, suggesting that this process was less efficient. Using an analysis of reaction time (RT) distributions for CF and CL trials, Unger et al. (2016) also found that children relied more on output gating, even if input gating was more efficient. Specifically, RT distributions revealed that 7-year-olds use a mixture of input and output gating strategies on CF trials. Thus, children are better at input gating but rely on the more effortful output gating strategy for goal-oriented action selection.

1.1.2 | Value of choice on WM strategies

While output gating may provide a default strategy for children to tackle many everyday tasks, a shift to a more flexible use of strategies may be achieved through interventions targeting a greater reliance on input gating. There may be many possible ways to shift strategies to be more proactive. One important consideration for the study of input and output strategy use in the service of efficient cognitive control in 3- to 7-year-old children relates to the sense of agency of choice in goal setting. Mechanistically, choice has been shown to impact adult decision making (Pedersen, Frank, & Biele, 2017). Using a reinforcement learning framework, Collins and Frank (2014) demonstrated the critical role of dopaminergic pathways for both learning and choice. To our knowledge, it has not been specifically investigated how the act of choosing a relevant context affects which WM strategies children adopt.

Unlike most everyday behaviors in which goals are voluntarily chosen by the acting individual, prior work investigated cognitive control using a range of paradigms in which the goal is always announced by the experimenter (e.g., Cepeda, Kramer, & Gonzalez de Sather, 2001; Chevalier & Blaye, 2009; Chevalier, Huber, Wiebe, & Espy, 2013; Zelazo, Frye, & Rapus, 1996). Translating externally specified goals into a corresponding representation of the task in WM may be particularly difficult for young children because this relies, at least to some degree, on verbal rehearsal (e.g., Miyake, Emerson, Padilla, & Ahn, 2004), a process that becomes efficient only in later childhood (e.g., Pressley & Hilden, 2006). Consistent with this notion is the observation that preschoolers perform better on card-sorting tasks when prompted to verbalize relevant stimulus features (e.g., Kirkham, Cruess, & diamond, A., 2003) and on proactive control engagement when linguistic labels are applied to target objects (Doebel, Dickerson, Hoover, & Munakata, 2018). In both cases, children exert some agency over task-relevant information, possibly strengthening the representation of the relevant dimension in WM at the outset and biasing children to use an input gating strategy.

1.1.3 | Current study

In the current study we used a version of the input/output gating paradigm (Chatham et al., 2014; Unger et al., 2016) to further investigate the development of WM gating strategies in 3- to 7-year-old children. Our specific task design was driven by considerations specific to this age group. First, prior works suggest that the abstractness of items in a task affects the precision with which items are recalled by adults (e.g., Ricker & Cowan, 2010) and children (e.g., Boucher et al., 2016). Similarly, preschoolers' ability to activate task goals might reflect difficulties in translating arbitrary cue-task associations (e.g., Chevalier & Blaye, 2009). We therefore adapted the original task to provide children with a more naturalistic wooden block game. Rather than a number or symbol, the context was represented by a house that had to be matched with a corresponding door block based on a two-dimensional fit (color and shape).

Houses were either presented before (CF) or after (CL) the door options. As noted, previous studies applied a mixture model of RT distributions to relate gating strategies to proactive and reactive modes of control, respectively (Unger et al., 2016). The naturalistic nature of the present task complicates the use of RT distributions. Determining definitively whether performance in the CF condition reflects input gating, output gating, or some mixture of strategies is difficult as children simply make one response per trial and it is either correct or incorrect. Recent work linked participants' eye movement sequences during the decision interval to performance on a Ravens Matrices test (Hayes, Petrov, & Sederberg, 2011). Inspired by this work, and in order to gain insight into gating strategies on CF trials, we examined similarity of scan paths during the decision interval on CF trials to scan paths used on output gating CL trials. As noted, only output gating is possible on CL trials. We reasoned that an examination of the similarity of eye movement sequences during the decision interval on CF relative to CL trials may reveal the extent to which children used a mixture of input and output gating on CF trials.

Finally, in the current study we included a choice condition that allowed children to select the contextual cue on each trial. We predicted that this manipulation would facilitate goal setting, and thus the efficiency with which gating strategies are implemented by young children. In summary, we adapted an input/output gating task for use in very young children. Based on previous work, we predicted that children may be biased to use an inefficient output gating strategy, even as input gating may be more efficient. We also probed the hypothesis that choice might strengthen WM representations, and thus bias children to use an input gating strategy on CF trials.

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants

Participants ($N = 60$) were recruited from the Developmental Cognitive Neuroscience Lab database. An additional 10 children were tested but excluded due to poor quality eye tracking calibration. The final sample included 20 children in each age group: 3-year-olds ($M = 3.61$, $SD = 0.34$, 10 females), 5-year-olds ($M = 5.29$, $SD = 0.37$, 10 females), and 7-year-olds ($M = 7.34$, $SD = 0.46$, 13 females). 93% of the children in this sample were White Caucasian, 3% Asian American, and 3% African American. Families were compensated $15 for taking part in the study and children received toy stickers after they completed the game. All children were prescreened prior to scheduling for known birth complications and preterm birth, known neurodevelopmental disorder, and colorblindness. Parents who reported any of these were invited to have their child participate in a different study. All families were consented in accord with the standards set by the Brown University IRB. General intelligence scores were obtained from 58 children during a second session using the Woodcock Johnson test of cognitive abilities (WJ-III-COG). No children were excluded because of an IQ more than 2 *SD*s above or below the group means (3-year-olds $M = 105.8$, $SD = 13.5$, 5-year-olds $M = 100.4$, $SD = 11.5$, and 7-year-olds $M = 115.1$, $SD = 13.3$).

2.1.2 | Eye-tracking data acquisition

Children were fitted with a portable eye-tracking headgear (Positive Science, LLC). The infrared eye camera was adjusted above each participant's right eye. The scene camera was attached to the flexible head cap and positioned central above the eyes. Eye-tracking data were captured at a sampling rate of 30 Hz. The resolution for eye and scene recordings was set to 320 × 240 and 640 × 480, respectively.

First, calibration data were recorded in order to map the position of the eye to the reference scene. Children were instructed to look and point to the location of a toy that would appear on a board 77 inch in front of them (26° visual angle). The toy popped up at five fixed locations. These locations served offline calibration of both eye and scene cameras. Additional calibration was included in between trials as needed. Audio was captured using a microphone recording at 44,100 Hz. Recordings were analyzed offline using Yarbus eye-tracking software (version 2.3.0). Eye recordings were synchronized and scaled relative to scene recordings. Calibration was always performed using a minimum of five calibration points and corneal reflection location was estimated. In addition to pupil threshold, a feature detection algorithm was selected for pupil tracking.

2.1.3 | Task design and materials

The logic of the task was adapted from a WM task previously used in both adults and children (Chatham & Badre, 2015; Chatham et al., 2014; Unger et al., 2016). The higher level context was represented by a wooden house with a hole in place of the door (3.5 inches wide by 5 inches high by 3 inches deep). Each house was uniquely characterized by its color (blue/red/green) and the shape of its door (circle/rectangle/triangle). A matching wooden door block (2.5 inches wide by 2.5 inches high by 3 inches deep) had to be selected for each house (target block) while ignoring all other door blocks (distractors). Based on color and shape, each house was matched with a corresponding door block (Figure 1). The higher level context (house) specified which lower order door block, that is, a color and shape matching door block, would be the correct response on each trial. Thus, there was only one correct answer on each trial; the block matching both dimensions. This differentiates the block task from previous paradigms. Critically, on half the trials the context was presented at the beginning of the trial (CF condition), allowing participants to input gate the location of the matching block while ignoring all irrelevant blocks (Figure 2). All CF sequences included the following: (a) presentation of the house; (b) removal of the house; (c) presentation of a door block, (d) which was then covered and moved into a serial lineup that remained in the child's sight on the table; (e) presentation and lineup of the next door block; and (f) response. On the remainder of the trials all blocks had to be maintained in WM until the context was presented at the end of the trial (CL condition). CL sequences were composed of (a) presentation of a door block, (b) that was covered and moved into the serial lineup; (c) presentation and lineup of the next door block; (d) presentation of the house; (e) removal of the house; and (f) response. Each house context was used in three low and three high WM load trials, corresponding to a target plus one or two distractor door blocks, respectively.

In a second choice condition, we repeated the same basic procedure but rather than have the experimenter choose the higher level context, children could choose which of two houses served as the context and selected the house of their choice to play the game with. Blocks of choice trials and standard trials were presented in randomized order. Children played the game while being seated at a child-sized table (25.5 inches by 25.5 inches). Test trials were preceded by a familiarization phase and eight practice trials (four per condition). Children were fitted with a Positive Science portable eye tracker, calibrated using a five-point array, and the task began.

2.1.4 | Familiarization procedure

Participants explored all stimuli during this initial familiarization phase. If the child did not spontaneously match the house and door blocks based on shape and color, the experimenter prompted the child to do so. Children were asked to name all colors and shapes. The experimenter retrieved a house and two small door blocks, one matching the house on both dimensions (shape/color) and the other matching either in terms of shape or color. The experimenter asked the child to pick the door block that belonged to the presented house. This was repeated with a second house. All children correctly selected door blocks based on a two-dimensional fit, and were further able to name all colors and shapes.

2.1.5 | Training and test trial procedure

Choice and standard trials were identical except that, instead of the experimenter presenting a house, children were allowed to select from one of two houses to serve as the higher level context. Each CF training trial began with the experimenter presenting a house (or a choice of 2 houses; higher level context) in the center of the

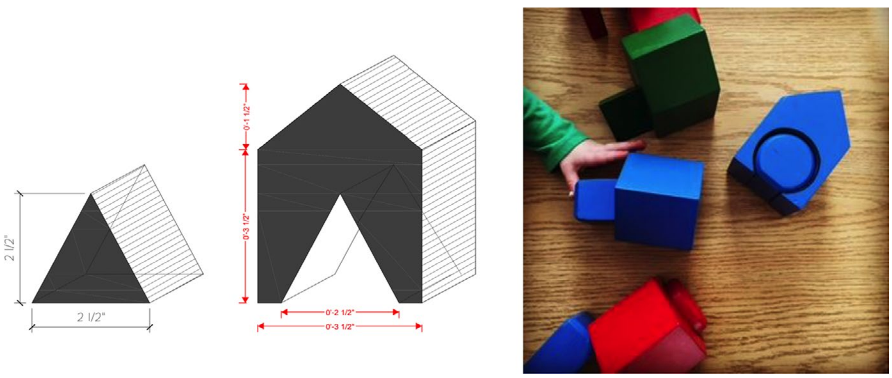

FIGURE 1 Example of stimuli (five houses total)

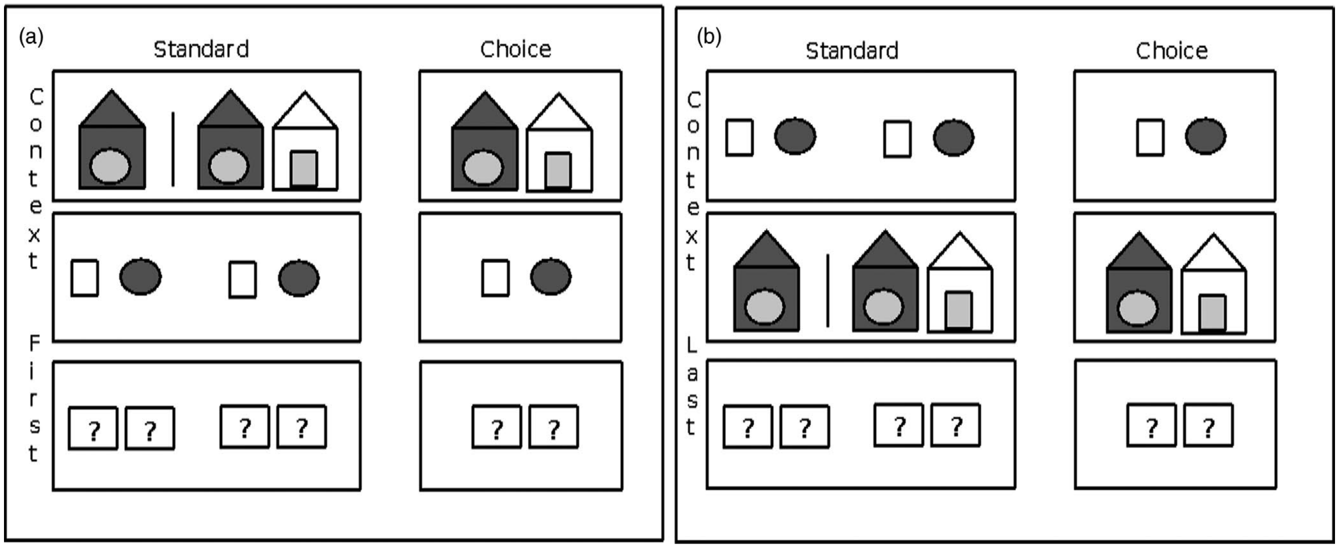

FIGURE 2 Schematic of the "block game" (low WM trial). A house serves as the higher level context to match a corresponding door block (the selected house and target block are illustrated in dark grey). The context is presented (a) at the beginning of the trial (CF) or (b) at the end (CL). Door blocks, (a) center panel and (b) top panel, are presented in random order one at a time. Each door block is covered by a box and placed in a serial lineup on the table surface. Participants point to the box that holds the target block. Both CF and CL are composed of two standard no choice conditions and one choice condition. In the choice task participants select a context from two alternatives. On standard trials, one (experiment 1) or two houses (experiment 2) are presented but the context is always predetermined

table at a 24° visual angle. On the standard condition, the child was told that this would be the house they would be playing with on this trial of the game. On the choice condition, the child was asked to choose which house they wanted to play with for that trial. The house was then removed from the child's view. The experimenter placed either the target door block (the one that matched in color and shape) or one of the distractor blocks in the center of the table. This door block was immediately covered with a box and moved into a line-up where it remained until the end of the trial. This covering process forces children to maintain the locations of the items in the line-up in WM. Upon presentation of all door block options (two in the low WM load trials and three in the high WM load trials), the experimenter asked the child "*Can you point to the box with the block that belongs to the house?*" The selected door block was then revealed and all items were cleared from the table. The combination of colors and shapes serving as the second-order context was counterbalanced across participants. CL training trials were identical except the context house(s) was presented after the presentation of the door blocks. All children passed a minimum of four training trials (eight total). At the end of the training phase, the eye tracker was recalibrated and the test phase began without further delay.

Test trials were identical to training trials with two exceptions. Two blocks of choice trials and two blocks of standard trials were presented in randomized order. Each context first and context last condition was presented six consecutive times for the choice and six consecutive times for the standard blocks, corresponding to three low and three high WM load trials.

TABLE 1 Performance accuracy by age group across experiment 1 and 2

	M (SD)							
	CF standard low WM	CF choice low WM	CF standard high WM	CF choice high WM	CL standard low WM	CL choice low WM	CL standard high WM	CL choice high WM
Experiment 1								
3-year-olds	0.60 (0.32)	0.73 (0.28)*	0.48 (0.28)*	0.55 (0.35)*	0.63 (0.24)*	0.57 (0.39)	0.30 (0.26)	0.38 (0.27)
5-year-olds	0.85 (0.20)*	0.88 (0.20)*	0.73 (0.26)*	0.75 (0.24)*	0.87 (0.17)*	0.62(0.35)	0.53 0.31)*	0.53 (0.29)*
7-year-olds	0.93 (0.14)*	0.95 (0.12)*	0.88 (0.22)*	0.82 (0.23)*	0.88 (0.20)*	0.88 (0.20)*	0.73 (0.23)*	0.88 (0.20)*
Experiment 2								
3-year-olds	0.63 (0.32)	0.52 (0.33)	0.35 (0.35)	0.37 (0.26)	0.47 (0.33)	0.62 (0.22)*	0.30 (0.32)	0.38 (0.25)
5-year-olds	0.8 (0.23)*	0.9 (0.16)*	0.7 (0.36)*	0.78 (0.29)*	0.68 (0.25)*	0.8 (20)*	0.47 (0.20)*	0.60 (0.34)*

Note: $N = 20$ per age group per experiment (total $N = 100$).

*Percent correct performance significantly above chance.

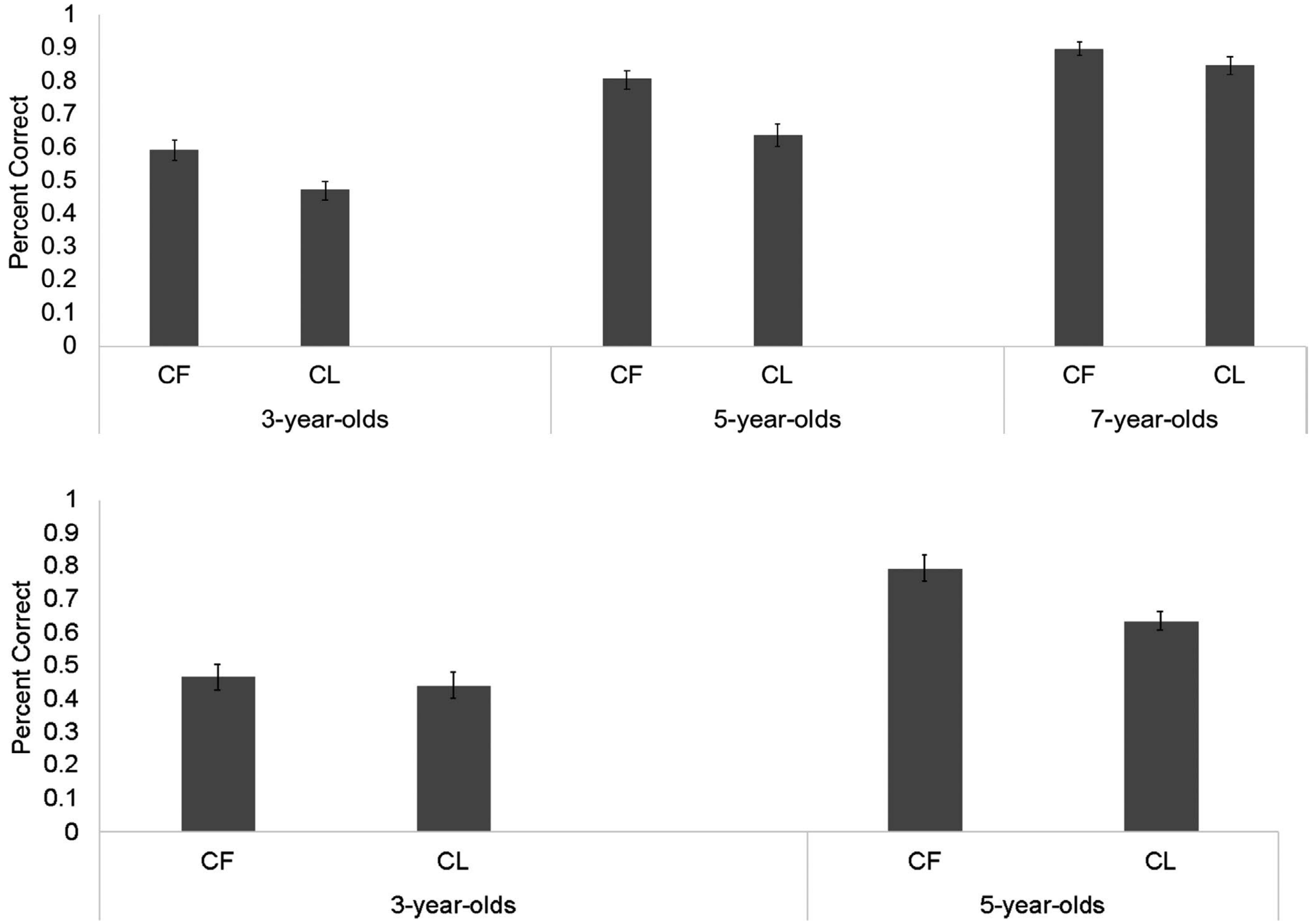

FIGURE 3 Performance accuracy (percent correct trials) for context first (CF) and context last (CL) trial in experiment 1 (top panel) and experiment 2 (bottom panel)

2.2 | Results

Table 1 shows descriptive statistics by age group (means and *SD*s). We conducted an omnibus analysis, including context (CF/CL), WM load (low/high), and target choice (standard/choice) as within-subject factors, and age (3 years/5 years/7 years) as a between-subjects factor. The analysis resulted in a main effect of WM load, $F(1,57) = 32.16$, $p < .001$, $\eta_p^2 = 0.36$, where performance was better on low WM load ($M = 78.33\%$, $SD = 18.8\%$) relative to high WM load trials ($M = 63.2\%$, $SD = 22\%$). There was also a main effect of age, $F(2,57) = 49.30$, $p < .001$, $\eta_p^2 = 0.63$. Contrasts showed that 3-year-olds ($M = 53.13\%$, $SD = 10.28\%$) performed more poorly than 5- ($M = 72.08$, $SD = 12.24\%$) and 7- ($M = 87.08\%$, $SD = 9.83\%$) year-old children (all *p*s $< .001$). Moreover, 5-year-olds performed more poorly than 7-year-olds ($p < .001$). Most relevant to our questions, the analysis resulted in a main effect of context, $F(1,57) = 42.77$, $p < .001$, $\eta_p^2 = 0.43$. Accuracy was higher on CF ($M = 76.39\%$, $SD = 17.5\%$) relative to CL trials ($M = 65.14\%$, $SD = 20.27\%$).

There were additionally three significant interactions relevant to the context condition. There was a significant context by age interaction, $F(2,57) = 3.89$, $p = .03$, $\eta_p^2 = 0.12$, a significant target choice by context by age interaction, $F(2,57) = 4.05$, $p = .02$, $\eta_p^2 = 0.12$, and a significant target choice by context by WM load interaction, $F(1,57) = 8.13$, $p = .006$; $\eta_p^2 = 0.13$. We follow-up on each of these using within and across specific age group analyses of context, choice, and WM load.

Relevant to the context by age interaction, we found that while 3-year-old, $F(1,19) = 9.87$, $p = .005$, $\eta_p^2 = 0.34$, 5-year-old, $F(1,19) = 36.19$, $p < .001$, $\eta_p^2 = 0.66$, and 7-year-old, $F(1,19) = 6$, $p = .02$, $\eta_p^2 = 0.24$, children performed better on CF relative to CL trials, Figure 3 indicates that the interaction of context and age reflects the context manipulation having the relatively smallest impact on 7-year-olds. However, the circumstances under which these results obtained were clarified in the omnibus with an additional target choice by context by age interaction, $F(2,57) = 4.05$, $p = .02$, $\eta_p^2 = 0.12$. A follow-up analysis of data from the standard blocks resulted in a main effect of context, $F(1,57) = 12.68$, $p = .001$, $\eta_p^2 = 0.18$, but no context by age interaction, $F(2,57) = 0.09$, $p = .92$, $\eta_p^2 = 0.003$. Figure 4 shows that CF performance was better than CL performance on standard blocks for all three groups. This result is interpreted to mean that at baseline, there is not an improvement in the relative contributions of proactive to reactive control strategies from 3 to 7 years, but rather consistently better CF relative to CL performance.

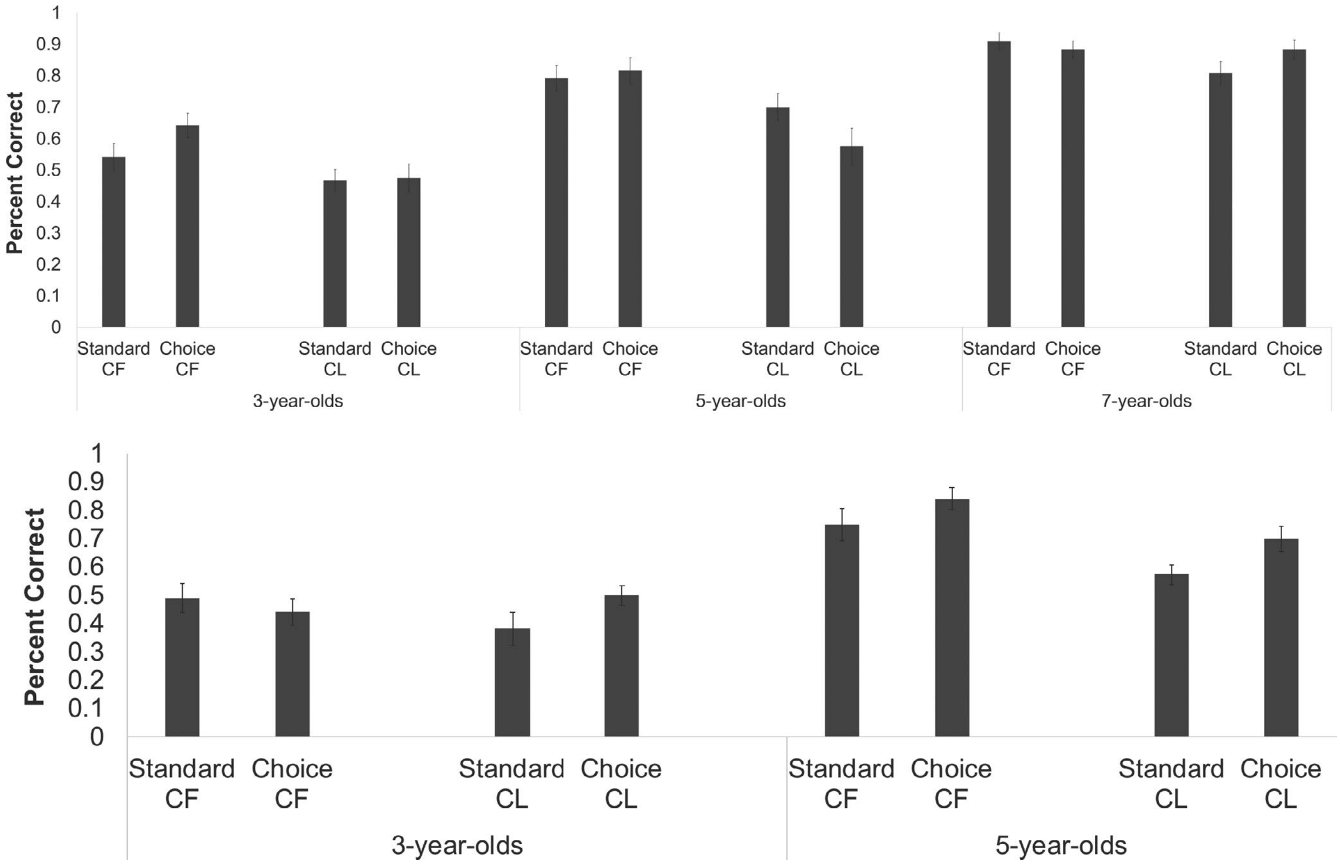

FIGURE 4 Performance accuracy (percent correct trials) for context first (CF) and context last (CL) conditions for each standard and choice block in experiment 1 (top panel) and experiment 2 (bottom panel)

However, developmental differences emerged when children were allowed to choose the target house. A follow-up analysis of data on choice blocks resulted in a context by age interaction, $F(2,57) = 7.57$, $p = .001$, $\eta_p^2 = 0.21$. There was no difference for CF/CL performance by age for 3-year-olds versus 5-year-olds, $F(1,38) = 1.12$, $p = .30$, $\eta_p^2 = 0.03$. Figure 4 shows both groups performed better on CF relative to CL trials. However, 7-year-olds were significantly different from 3- to 5-year-olds, $F(1,38) = 6.33$, $p = .02$, $\eta_p^2 = 0.14$. Figure 4 indicates that the target choice manipulation had a smaller impact on the difference between CF and CL trials in 7- relative to 3- to 5-year-olds. Indeed, between 5 and 7 years, there was improvement on CL, $t(38) = 4.69$, $p < .001$ (Table 1) but not CF trials in the choice blocks, $t(38) = 1.38$, $p = .18$. Moreover, 7-year-olds showed a trend toward better performance on CL choice than CL standard trials, $t(19) = 1.92$, $p = .07$.

Relatedly, there was a significant target choice by context by WM load interaction, $F(1,57) = 8.13$, $p = .006$; $\eta_p^2 = 0.13$. Follow-up analyses probed the impact of target choice and WM load on CF and CL trials separately. There is not an impact of target choice by WM load on CF trials, $F(1,57) = 0.82$, $p = .37$, $\eta_p^2 = 0.01$. There was an interaction of target choice and WM load on CL trials, $F(1,57) = 7.86$, $p = .007$, $\eta_p^2 = 0.12$. Worse performance on choice relative to standard CL trials was specific to low WM (standard $M = 79\%$, $SD = 23\%$; choice $M = 69\%$, $SD = 35\%$), $t(59) = 2.07$, $p = .04$, but not statistically for high (standard $M = 52\%$, $SD = 32\%$; choice $M = 60\%$, $SD = 33\%$) WM load trials, $t(59) = 1.70$, $p = .10$.

2.2.1 | Experiment 1 conclusions

Taken together, these data indicate that CF performance is better than CL performance at 3, 5, and 7 years (Figure 3). The target choice manipulation had a positive impact on CF trials in 3- to 5-year-old children and a positive impact on CL performance at 7 years. Specifically on choice blocks, by the age of 7, the magnitude of the difference in accuracy on CF and CL trials decreases, relative to the difference score earlier in development. This is due to a sharp improvement in CL performance from 5 to 7 years and a smaller but significant increase in CF performance from 5 to 7 years.

Together, these data point to two conclusions. First, there are differences in efficiency for input and output gating information into WM in 3- to 7-year-olds. CF trials were performed with higher accuracy than CL trials for all three age groups. Even 3-year-olds performed well on CF trials, but were at chance on CL trials (Table 1). CL ostensibly requires output gating, whereas CF reflects a possible mixture of input and output gating. As such, we can infer that the involvement of some level of input gating on CF trials makes it such that performance is *more* efficient than output gating in this 3- to 7-year-old sample.

Second, the target choice manipulation seemed to have an impact on the strength of WM representations. However, this impact differed by age. Choice improved accuracy over standard for CF trials in 3- to 5-year-olds and CL trials in 7-year-olds. The magnitude of the difference in accuracy on CF and CL trials was similar for 3- and 5-year-olds but decreased by 7 years only for the choice blocks. We are cautious, however, to conclude that the target choice manipulation did not impact CF performance in 7-year-olds. It may be that 7-year-olds were at ceiling for CF standard performance (Table 1; Figure 4).

Using eye movement scan path analyses during the decision interval (Petrov et al., 2011), we will later examine the prediction that agency on the choice blocks, relative to standard blocks where the experimenter chose the house, may have strengthened WM representations when the house context was presented, thereby biasing younger children to use a more efficient input gating strategy on CF trials.

However, we must first consider one procedural confound that may have also driven better CF relative to CL performance on choice relative to standard blocks in younger children. The standard and choice trials in this task differed with respect to the number of house stimuli presented at the beginning of each trial (1 on standard and 2 on choice). This procedural difference may have resulted in children differentiating house targets on the choice blocks in a manner that strengthened WM representations for the relevant feature dimensions, theoretically allowing for better input gating on CF choice than standard trials, but not solely because of the agency involved. Rather, it could be that seeing two houses allowed children an opportunity to discern the target features (e.g., blue rectangle door shape) but also the nontarget features (e.g., red circle door shape). Additional clues are offered by the interaction of context, target choice, and WM. This interaction revealed overall *worse* performance on choice CL than standard trials. On CL choice trials, children must remember all of the dimensions of the target blocks and then are shown two houses, possibly increasing demand on WM maintenance as children must search for the chosen target in the contents of WM. In this way, a two-house presentation may have made the manageable CL low WM trials no longer manageable by increasing general demand on WM. By the age of 7 years, children were on average showing higher accuracy, albeit only a statistical trend, on choice CL than CF trials, perhaps because they were able to manage the WM load demand on choice CL trials.

We therefore ran a second experiment, designed to (a) replicate the effect of context on WM performance, and (b) examine the possibility that the target choice manipulation may have had an impact not only because of the agency component but also because of a difference in the number of houses presented on choice relative to standard blocks. We focused the second experiment on 3- to 5-year-olds, largely because of the near ceiling effects of CF in 7-year-old children. Experiment 2 differs from experiment 1 only in terms of the number of houses that were presented in the standard condition. It is important to note that a single house served as contextual cue and was chosen not by the child but by the experimenter (Figure 2). All remaining procedures and analyses were identical to experiment 1.

The predictions are as follows. First, if the results of experiment 1 were entirely driven by agency, we would expect a direct replication of experiment 1 results. However, if only the two-house aspect of the choice blocks was contributing to better CF than CL performance in younger children, we would expect equivalent performance for choice and standard in a follow-up experiment where two houses were presented for both standard and choice blocks. Finally, if *both* agency of the child and the two-house presentation played an additive role in strengthening WM representations, we would still expect better performance on CF trials for the choice relative to the standard blocks, albeit smaller than in experiment 1. Alternatively, if the two-house presentation masked an effect of choice on CL performance by overloading WM, we would expect to see an emerging effect of target choice even on CL trials in the younger children in experiment 2.

We note here chance performance on all CL conditions in 3-year-old children (Table 1). While relative CF/CL and choice/standard performance did not differ in 3- and 5-year-old children, it is clear from these data that 3-year-olds were indeed having trouble with the task and especially CL trials. Experiment 2 will offer us a second opportunity to examine the conditions under which 3-year-olds can succeed.

3 | EXPERIMENT 2

3.1 | Participants

A total of 44 preschoolers were tested. Four children were excluded from further analyses due to poor quality of eye-tracking data. Twenty 3-year-olds (M = 3.32, SD = 0.15) and 20 5-year-olds (M = 5.38, SD = 0.30) made up the final sample (83% White Caucasian, 10% Asian American, 3% African American, 2% Hispanic, and 2% African American). Participants were prescreened, caregivers consented in accord with Brown University IRB, and were compensated for taking part in this study.

3.2 | Results

Table 1 shows that performance for 3-year-olds was at chance on all CF trials and above chance on only one CL condition (CL standard low WM). Three-year-old children's data are interpreted accordingly and with caution. We conducted an omnibus analysis including context (CF/CL), WM load (low/high), and target choice (standard/choice) as within-subject factors. As in experiment 1, the analyses yielded significant effects for WM load $F(1,38) = 39.15$, $p < .001$, $\eta_p^2 = 0.51$. Also, as in experiment 1, we found a significant main effect of context, $F(1,38) = 14.07$, $p = .001$, $\eta_p^2 = 0.27$. Children performed significantly better on CF than on CL trials (Figure 3). There was additionally a context by age interaction, $F(1,38) = 8.40$, $p = .006$, $\eta_p^2 = 0.18$. Figure 4 shows that only 5-year-olds showed the expected effect of context, $t(19) = 4.50$, $p < .001$. In contrast 3-year-olds performed at chance on the majority of both CF and CL trials, $t(19) = 0.70$, $p = .49$ (Table 1). There was also a significant main effect of target choice, $F(1,38) = 4.97$, $p = .03$, $\eta_p^2 = 0.12$. The main effect of target choice indicates that relative to standard *both* CF and CL performance was better on the choice trials manipulation (Figure 4). All other interactions were not significant (all $ps > .08$). The emergence of better performance for choice on CL data is consistent with the interpretation that the two-house presentation may have exacted a WM burden on CL blocks in experiment 1, thereby masking an effect of choice.

3.2.1 | Cross-task comparisons in 5-year-olds

To further subject this interpretation of experiment 2 data to analytic scrutiny, we ran an analysis of context (CF/CL), WM load (low/high), and target choice (standard/choice) as within-subject factors and experiment as a between-subjects factor (experiment 1, experiment 2) on 5-year-olds from both experiments. Five-year-olds were chosen because 3-year-old performance in experiment 2 was noisy. This analysis is designed to address two critical questions regarding the specific effect of choice on WM. First, does equating the two-house presentation on standard and choice CF trials impact the size of the difference in accuracy for CF choice relative to standard in experiment 2 relative to experiment 1? If there is a reduction in the size of the effect of choice on CF trials in experiment 2, then this would indicate that the size of the difference in CF trial performance by the target choice manipulation was indeed at least partially driven by the number of houses presented on standard (one house) relative to choice (two houses) CF trials. Second, does equating the two-house presentation on standard and choice trials statistically increase the effect of choice on CL performance across experiments.

The cross-experiment analysis resulted in an expected context by target choice by experiment interaction, $F(1,38) = 4.28$, $p = .05$, $\eta_p^2 = 0.10$. A follow-up analysis indicated no difference in the effect of performance on choice relative to standard CF trials across Experiments, $F(1,38) = 0.68$, $p = .42$, $\eta_p^2 = 0.02$. We can infer that the effect of choice on CF trials is consistent across experiments and was not due to the difference in the number of houses presented on CF standard trials.

Rather the interaction reflected a difference in the effect of the target choice manipulation on CL trial performance across experiments, $F(1,38) = 6.77$, $p = .01$, $\eta_p^2 = 0.15$. There was no difference in CL choice performance across experiments, $t(38) = 1.68$, $p = .10$. However, there was a difference for CL *standard* performance $t(38) = 2.41$, $p = .02$, with better performance for experiment 1 relative to experiment 2. This is likely because of the addition of the two-houses on experiment 2. Moreover, we observed better performance on CL choice relative to standard trials in experiment 2, $t(20) = 2.18$, $p = .04$, that was not evident in experiment 1 in this age group but was evident in 7-year-olds. These data indicate that the effect of the choice manipulation on CL trials may have been masked by the addition of two houses only on standard trials in experiment 1.

3.2.2 | Experiment 2 conclusions

In conclusion, the second experiment replicated the finding that CF performance is better than CL performance, and confirmed the unique value of choice in improving WM performance. A difference between the two experiments was performance by the 3-year-olds. Three-year-olds in experiment 2 performed generally very poorly. It is not clear whether this is a cohort effect or whether the addition of two houses on all trials on the standard condition changed their general understanding of the task.

These results indicate that the target choice manipulation does indeed support WM performance on CF trials through agency. However, the addition of the two houses had masked its effect on CL trials on the young children, most likely because of the additional load demands the two-house option placed on WM. However, the mechanisms by which the choice manipulation impacted WM performance remain unclear.

One hypothesis is that choice strengthens WM representations for the relevant feature dimensions. A possible test of this hypothesis, that we raised in the introduction, is specific to strategies engaged on CF trials. The act of choice, or agency, may have biased young children to adopt an efficient input gating strategy on CF trials, the only manipulation where they could use either input or output gating. As such, it is possible that agency plays a role in biasing strategy to be proactive, and thus engaging relatively more efficient input gating on CF trials, especially on the choice blocks. In the following section, we use an analysis of eye movement scan paths to offer evidence for this prediction.

4 | SCAN PATH DATA COMBINED ACROSS EXPERIMENTS 1 AND 2

Comparing scan path sequences across the standard and choice blocks can offer insight into the WM strategies used in each task, and whether children approached the CF task less reactively when allowed to choose a context from two alternatives. Specifically, we aim to test whether participants were *more likely* to use a scan path on CF standard trials that was more similar to the output gating scan path they used on CL trials, and conversely whether they were less likely to use an output gating scan path on CF choice trials. As a manipulation check, we also generated data where classification was done using CF trials for training and CL trials for test.

Full details on the scan path fixation extraction procedure, the Collins (2002) structured perceptron for sequence classification algorithm used can be found in Supplementary Materials. Across the two experiments 82 participants (31 three-year-olds, 33 five-year-olds, and 18 seven-year-olds) provided eye-tracking data for all eight conditions. We coded the sequence of eye movements during the decision interval, after the house(s) were removed and before the choice was made, allowing for multiple fixations to the presented block options. The coded sequence maintained the order and location of fixations. Four classifiers were trained corresponding to each CL condition (standard high WM, standard low WM, choice high WM, and choice low WM). Correct trials for each of the four analogous CF conditions were tested separately (e.g., train on correct and incorrect CL low WM and test on correct CF low WM).

For classification of a sequence produced by a subject, we used the remaining subjects in each age group for training the system and the current subject for testing. This method follows leave-one-out paradigm of classification problems. Classification was done using discriminative training methods for Hidden Markov Models (HMM). This is a binary classifier, per trial, with the values averaged across the three trials per condition to generate a 0%–100% value per participant per condition. A 0 indicates that the scan path does not classify as the analogous CL condition scan path and a 100% indicates that it strongly classifies as the analogous CL condition scan path.

5 | RESULTS

We are using scan path data to test the prediction that children are more likely to use a mixture of output and input gating strategies on the standard blocks, but are biased toward efficient input gating on the choice blocks. The logic of this analysis assumes that there is *relatively* more opportunity for input gating on CF trials and relatively more output gating on CL trials. The idea is that this would be expressed in how fixations are sequenced during the decision interval, when all blocks are presented for selection. Note, there is a second feature that could drive scan path similarity. Specifically, WM load reflects the number of physical wooden blocks available for fixating and is shared among the training CL trials and test CF trials. This means that our classifier may always perform above chance. Thus, we also included the reversal, trained on CF and tested on CL, scan path sequences during the decision interval to analytically check for the influence of WM load. We conducted an omnibus analysis examining classifier strength as a function of training context (CF/CL) target choice condition (standard/choice), WM load (low/high), and age (3 years/5 years/7 years). We also included experiment (1/2) as a between-subjects factor to examine any differences in classification strength as a function of the number of houses presented in the standard blocks across experiments. Any effects of target choice and context (trained on CL and tested on CF or vice versa) should address the prediction of interest, that is, that relative to standard, the choice manipulation forces a more proactive input gating strategy and less output gating on CF trials. Thus, we report results relevant to the interaction of choice and context.

The analysis resulted in a target choice by context interaction, $F(1,73) = 25.57$, $p < .001$, $\eta_p^2 = 0.26$, and a target choice by context by WM interaction, $F(1,73) = 7.69$, $p = .007$, $\eta_p^2 = 0.09$. We note here that experiment did not further interact with either the effect of context by target choice, $F(1,73) = 1.44$, $p = .24$, or the effect of context by target choice by WM, $F(1,73) = 3.66$, $p = .06$, indicating that the following results are not impacted by the one-house (experiment 1) and two-house (experiment 2) presentation that only affected standard trials.

Follow-up analyses by context show that when the classifier was trained on scan path sequences on CL trials, and tested on CF trials, we found that children's scan path strategy approximated CL scan paths more on the standard than on the choice blocks, $F(1,76) = 35.27$, $p < .001$, $\eta_p^2 = 0.32$. Again, this effect was consistent

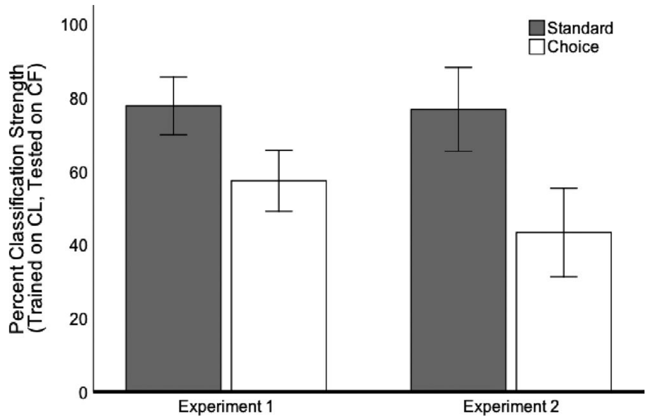

FIGURE 5 Scan path classifications of eye movements during the decision intervals of the standard and choice blocks for experiment 1 and experiment 2

across experiments, $F(1,76) = 1.21$, $p = .28$, $\eta_p^2 = 0.02$, as shown in Figure 5. This finding supports our initial prediction. These results reflect the finding that when the classifier was trained on scan path sequences on CL trials, and tested on CF trials, children's scan path strategy similarly approximated more CL on standard relative to choice blocks for low, $t(80) = 4.34$, $p < .001$, and high WM load, $t(80) = 4.52$, $p < .001$.

In contrast, there was not a significant target choice effect for the control analysis, when the classifier was trained on CF but tested on CL, $F(1,79) = 3.04$, $p = .09$, $\eta_p^2 = 0.04$. However, the context by target choice by WM interaction in the omnibus did stem uniquely from this control analysis. There was a significant interaction of context by target choice by WM, $F(1,79) = 11.77$, $p = .001$, $\eta_p^2 = 0.13$, when the classifier was trained on CF and tested on CL trials.

When the classifier was trained on CF and tested on CL trials, there was an effect of target choice on high WM load, $t(83) = 3.52$, $p = .001$, but not on low WM trials, $t(84) = 0.25$, $p = .80$. Note this is the reverse of the direction of results when the classifier was trained on CL and tested on CF. When WM load was high, CL scan paths approximated CF scan paths on choice ($M = 46.41$, $SD = 45.41$) more than on standard trials ($M = 70.98$, $SD = 42.67$). This is ostensibly because standard CF trials involved a higher mixture of input/output gating than did the choice CF trials. As such, when the training was done on the CF trials, scan paths on standard blocks were more similar to the CL (output gating) than scan paths from the choice blocks.

Regardless of training and test set, scan paths across CL and CF trials approximated each other, and critically indicated less output gating, on choice blocks. Taken together, these data indicate that the success of the choice manipulation in bootstrapping CF performance may be in that it prevents children from choosing to employ an effortful output gating strategy.

6 | DISCUSSION

Across experiments 1 and 2 children of all age groups demonstrated higher behavioral accuracy on CF relative to CL trials. This observation is in good agreement with Unger et al. (2016) and suggests that the ability to update task-relevant information into WM through a selective input gate is stronger earlier in development than output gating. We employed scan path analyses of saccadic eye movement to tease apart these cognitive strategies and to relate them to performance on CF trials. Our results support the idea that children use a mixture of reactive and proactive control when provided with standard CF trials and less so when provided with the choice of which context to use on each trial (Figure 5). These results did not interact with age. As such, the choice manipulation may be supporting less of a reactive/proactive mixture and more proactive control—which the behavioral data in this and other studies (Munakata et al., 2012; Unger et al., 2016) suggest is more efficient.

Previous behavioral studies indicate that a shift in the mode of control unfolds around 5 years of age (e.g., Lucenet & Blaye, 2014). Scan path analyses in this study indicated that children as young as 3 years of age can engage control proactively when choosing a goal at the outset of each trial (Table 1, experiment 1 data). However, 3-year-olds WM performance was variable across the experiments. While performance in both experiments 1 and 2 was poor for CL trials, performance in experiment 1 indicated that they were able to succeed at CF trials. However, accuracy data from experiment 2 was variable, largely at chance, but still reflecting the same relative patterns as 5-year-olds. The striking difference is actually on CF trials (Figure 4). It is not clear whether the addition of two houses on all trials changed the nature of the task for this young group. For example, in experiment 1, it may have been that they were able to bootstrap their understanding of the rules of the game by the multiple cues offered on choice and standard trials. On standard trials, the experimenter simply presented a house (experimenter's turn, the child does nothing) and on choice trials, it was the child's turn to choose among two alternatives. It is possible that on experiment 2, the addition of two houses even when the experimenter chose the target house on standard trials may have confused children's understanding of when they were supposed to choose versus when the experimenter was supposed to choose the house to play with, thereby distracting them in a way that detracted from WM maintenance of relevant target features. Future work should consider this possibility.

Introducing a choice manipulation to the original paradigm revealed a shift to more proactive control when children selected the contextual cue themselves. Although no previous studies have related WM gating to goal setting, our results suggest that goal setting may facilitate the use of an input gating strategy when available. Actively choosing a goal has been argued to strengthen task-goal representations, a critical base for executive control functions (e.g., Baddeley, Chincotta, & Adlam, 2001; Chevalier et al., 2012; Friedman et al., 2008; Miyake & Friedman, 2012; Rubinstein, Meyer, & Evans, 2001).

Output gating allows us to select WM representations reactively and in a goal-directed manner. However, output gating appears to be a costly process even for adults as it places demands on memory and selective attention (Badre & Frank, 2012; Bhandari & Badre, 2018; Chatham et al., 2014). In addition to capacity demands posed by CL trials in this study, stored WM representations may have become subject to considerable interference, thus compromising target selection at the stage of retrieval. Fallon, Zokaei, Norbury, Manohar, and Husain (2017) investigated the involvement of dopaminergic mechanisms in the storage and retrieval of visual information and observed modulatory effects on the fidelity with which stimuli features were recalled by adults. Comparing memory capacity across the lifespan children have been found to be more susceptible to confusion between items than adults (e.g., Rodríguez-Villagra, Göthe, Oberauer, & Kliegl, 2013). Preschoolers, in particular, are prone to attend irrelevant features of task stimuli (e.g., Chevalier, Blaye, Dufau, & Lucenet, 2010). As such, shape and color features may have become confused or swapped on CL trials.

The notion that output gating may become an effective WM strategy only in later childhood is consistent with developmental evidence using a range of semantic updating tasks. These tasks typically require children to make some form of comparison among stored WM representations (e.g., recalling the smallest item or the largest number). While the demand to output gate is not specifically manipulated, these tasks nonetheless call for serial updating of information into WM followed by the selective retrieval of a single target item. Taken together these studies indicate a linear increase in performance from early childhood to late adolescence (e.g., Balacchi, Carretti, & Cornoldi, 2010; Carriedo, Corral, Montoro, Herrero, & Rucían, 2016; Lendínez, Pelegrina, & Lechuga, 2015).

In agreement with the wider developmental research on rule use in young children (e.g., Blackwell & Munakata, 2014; Unger et al., 2016; Zelazo, 2004), we also observed a decrease in behavioral accuracy in the standard and choice task with increasing WM load. However, there were not meaningful interactions among WM strategy use in the service of rule-guided behavior and WM load. In line with this finding, recent developmental studies demonstrate that WM capacity alone does not explain developmental change in rule-guided behavior (Amso, Haas, McShane, & Badre, 2014; Unger et al., 2016). Similarly, WM load does not explain why choice bootstraps the use of an input gating strategy in 5- and 7-year-olds, as well as output gating in 7-year-olds. Age-related effects in terms of CF choice scan paths suggest that the behavioral age differences in this study reflect a development in the use of input gating between the ages of 3 and 7 years.

A promising avenue for future research lies in further understanding the developmental trajectories of WM gating with the aim to generate adaptive interventions. Empirical evidence shows that WM functioning is critically linked to the development of other complex cognitive abilities such as learning novel concepts (e.g., Pickering, 2006), literacy (e.g., St Clair-Thompson & Gathercole, 2006), and math skills (e.g., DeMarie & López, 2013). Yet WM interventions typically aim to improve capacity rather than the ability to implement WM strategies in respect to specific task demands. Taking on a broader perspective, interventions may benefit from teaching children about the effectiveness of different cognitive strategies, thus providing transferable skills to tackle a wide range of tasks independently. In addition, allowing children to actively choose among task options may increase attentiveness to task requirements, therefore facilitating the selection of goal-directed cognitive strategies.

ACKNOWLEDGMENTS

We are grateful to the families who took part in this study. We also thank Kelley Gunther and Diego Placido from the DCNlab for their support with testing in this study. This work was funded by the National Institutes of Health (grant R01 MH099078 to DA and DB). DB was supported by an ONR MURI (N00014-16-1-2832).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author.

REFERENCES

Amso, D., Haas, S., McShane, L., & Badre, D. (2014). Working memory updating and the development of rule-guided behavior. *Cognition*, *133*(1), 201–210. https://doi.org/10.1016/j.cognition.2014.06.012

Awh, E., Vogel, E. K., & Oh, S.-H. (2006). Interactions between attention and working memory. *Neuroscience*, *139*, 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*, 641–657. https://doi.org/10.1037/0096-3445.130.4.641

Badre, D., & Frank, M. J. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 2: Evidence from fMRI. *Cerebral Cortex*, *22*, 527–536. https://doi.org/10.1093/cercor/bhr117

Badre, D., Kayser, A., & D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. *Neuron*, *66*, 315–326. https://doi.org/10.1016/j.neuron.2010.03.025

Balacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in Coloured Raven Matrices performance in typically developing children. *European Journal of Cognitive Psychology*, *22*, 1010–1020. https://doi.org/10.1080/09541440903184617

Bhandari, A., & Badre, D. (2018). Learning and transfer of working memory gating policies. *Cognition*, *172*, 89–100. https://doi.org/10.1016/j.cognition.2017.12.001

Blackwell, K. A., & Munakata, Y. (2014). Costs and benefits linked to developments in cognitive control. *Developmental Science*, *17*, 203–211. https://doi.org/10.1111/desc.12113

Boucher, O., Chouinard-Leclaire, C., Muckle, G., Westerlund, A., Burden, M. J., Jacobson, S. W., & Jacobson, J. L. (2016). An ERP study of recognition memory for concrete and abstract pictures in school-aged children. *International Journal of Psychophysiology*, *106*, 106–114. https://doi.org/10.1016/j.ijpsycho.2016.06.009

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 713–737). Cambridge: MIT Press.

Carriedo, N., Corral, A., Montoro, P. R., Herrero, L., & Rucían, M. (2016). Development of the updating executive function: From 7-year-olds to young adults. *Developmental Psychology*, *52*(4), 666–678. https://doi.org/10.1037/dev0000091

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. C. M. (2001). Changes in executive control across the life span: Examination of task-switching performance. *Developmental Psychology*, *37*(5), 715–729. https://doi.org/10.1037/0012-1649.37.5.715

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Current Opinion in Behavioral Sciences*, *1*, 23–31. https://doi.org/10.1016/j.cobeha.2014.08.001

Chatham, C. H., Frank, M. J., & Badre, D. (2014). Corticostriatal output gating during selection from working memory. *Neuron*, *81*, 930–942. https://doi.org/10.1016/j.neuron.2014.01.002

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 5529–5533. https://doi.org/10.1073/pnas.0810002106

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, *45*, 782–797. https://doi.org/10.1037/a0015409

Chevalier, N., Blaye, A., Dufau, S., & Lucenet, J. (2010). What visual information do preschoolers and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Developmental Psychology*, *46*, 955–972.

Chevalier, N., Chatham, C., & Munakata, Y. (2014). The practice of going helps children to stop: the importance of context monitoring in inhibitory control. *Journal of Experimental Psychology: General*, *143*(3), 959–965.

Chevalier, N., Huber, K. L., Wiebe, S. A., & Espy, K. A. (2013). Qualitative change in executive control during childhood and adulthood. *Cognition*, *128*, 1–12. https://doi.org/10.1016/j.cognition.2013.02.012

Chevalier, N., Sheffield, T. D., Nelson, J. M., Clark, C. A. C., Wiebe, S. A., & Espy, K. A. (2012). Underpinnings of the cost of flexibility in preschool children: The roles of inhibition and working memory. *Developmental Neuropsychology*, *37*(2), 99–118.

Collins, A. G. E., & Frank, M. J. (2014). Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice. *Psychological Review*, *121*, 337–366.

Collins, M. (2002). Discriminative training methods for hidden Markov models: theory and experiments with perceptron algorithms. In Association for Computational Linguistics (Eds.), *Proceedings of the ACL-02 conference on Empirical methods in natural language processing* (Vol. 2, pp. 1–8). Stroudsburg, PA.

De Marie, D., & López, L. M. (2013). Memory in schools. In P. J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (vol. II, pp. 836–864). Chichester: John Wiley & Sons Ltd.

Doebel, S., Barker, J., Chevalier, N., Michaelson, L., Fisher, A., & Munakata, Y. (2017). Getting ready to use control: Advances in the measurement of young children's use of proactive control. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0175072.

Doebel, S., Dickerson, J. P., Hoover, J. D., & Munakata, Y. (2018). Using language to get ready: Familiar labels help children engage proactive control. *Journal of experimental child psychology*, *166*, 147–159.

Fallon, S. J., Zokaei, N., & Husain, M. (2016). Causes and consequences of limitations in visual working memory. *Annals of the New York Academy of Sciences*, *1369*, 40–54.

Fallon, S. J., Zokaei, N., Norbury, A., Manohar, S. G., & Husain, M. (2017). Dopamine alters the fidelity of working memory representations according to attentional demands. *Journal of Cognitive Neuroscience*, *29*(4), 728–738. https://doi.org/10.1162/jocn_a_01073

Frank, M. J., & Badre, D. (2012). Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: Computational analysis. *Cerebral Cortex*, *22*, 509–526. https://doi.org/10.1093/cercor/bhr114

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, & Behavioral Neuroscience*, *1*, 137–160. https://doi.org/10.3758/CABN.1.2.137

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Gruber, A. J., Dayan, P., Gutkin, B. S., & Solla, S. A. (2006). Dopamine modulation in the basal ganglia locks the gate to working memory. *Journal of Computational Neuroscience*, *20*, 153–166. https://doi.org/10.1007/s10827-005-5705-x

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2011). A novel method for analyzing sequential eye movements reveals strategic influence on Raven's Advanced Progressive Matrices. *Journal of Vision*, *11*, 1–11.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimensional card-switching task. *Developmental Science*, *6*, 449–467.

Lendínez, C., Pelegrina, S., & Lechuga, M. T. (2015). Age differences in working memory updating: The role of interference, focus, switching, and substituting information. *Acta Psychologica*, *157*, 106–113. https://doi.org/10.1016/j.actpsy.2015.02.015

Lucenet, J., & Blaye, A. (2014). age-related changes in the temporal dynamics of executive control: A study in 5- and 6-year-old children. *Frontiers in Psychology*, *5*, 1–11. https://doi.org/10.3389/fpsyg.2014.00831

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: The effect of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica*, *115*, 123–142.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14. https://doi.org/10.1177/0963721411429458

Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, *21*(2), 71–77. https://doi.org/10.1177/0963721412436807

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328. https://doi.org/10.1162/089976606775093909

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*, 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

Pickering, S. J. (2006). Working memory in dyslexia. In T. P. Alloway & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorder* (pp. 7–40). New York, NY: Psychology Press.

Pressley, M., & Hilden, K. (2006). Cognitive strategies: Production deficiencies and successful strategy instruction everywhere. In W. Damon, R. Lerner (Series Eds.), D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology: Vol. 2 Cognition, perception, and language* (6th ed., pp. 511–556). Hoboken, NJ: Wiley.

Ricker, T. J., & Cowan, N. (2010). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(6), 1355–1368.

Rodríguez-Villagra, O. A., Göthe, K., Oberauer, K., & Kliegl, R. (2013). Working memory capacity in a go/no-go task: Age differences in interference, processing speed, and attentional control. *Developmental Psychology*, *49*, 1683–1696. https://doi.org/10.1037/a0030883

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 763–797.

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: Toward a unified view of attention. *Journal of Cognitive Neuroscience*, *24*(11), 2147–2154. https://doi.org/10.1162/jocn_a_00275

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology*, *59*, 745–759. https://doi.org/10.1080/17470210500162854

Unger, K., Ackerman, L., Chatham, C. H., Amso, D., & Badre, D. (2016). Working memory gating mechanisms explain developmental change in rule-guided behaviour. *Cognition*, *155*, 8–22.

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences*, *8*, 12–17. https://doi.org/10.1016/j.tics.2003.11.001

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

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Freier L, Gupta P, Badre D, Amso D. The value of proactive goal setting and choice in 3- to 7-year-olds' use of working memory gating strategies in a naturalistic task. *Dev Sci*. 2020;00:e13017. https://doi.org/10.1111/desc.13017