



TITLE:

Contribution of executive functions to learning sequential actions in young children

AUTHOR(S):

Yanaoka, Kaichi; Saito, Satoru

CITATION:

Yanaoka, Kaichi ...[et al]. Contribution of executive functions to learning sequential actions in young children. *Child Development* 2021, 92(4): e581-e598

ISSUE DATE:

2021-07

URL:

<http://hdl.handle.net/2433/265323>

RIGHT:

This is the peer reviewed version of the following article: [Kaichi Yanaoka, Satoru Saito (2021). Contribution of Executive Functions to Learning Sequential Actions in Young Children. *Child Development*, 92(4), e581-e598.], which has been published in final form at <https://doi.org/10.1111/cdev.13489>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and webs ...

Childe Development

Contribution of executive functions to learning sequential actions in young children

Kaichi Yanaoka^{1,2} and Satoru Saito³

- 1) The University of Tokyo, Japan
- 2) Japan Society for the Promotion of Science
- 3) Kyoto University, Japan

Address correspondence:

Kaichi Yanaoka

Graduate School of Education, The University of University,

Hongo 7-3-1, Bunkyo-ku, Kyoto, Japan

E-mail. makifactor@gmail.com

Yanaoka, K., & Saito, S. (2021). Contribution of executive functions to learning sequential actions in young children. *Child Development*, 92(4), e581–e598.
<https://doi.org/10.1111/cdev.13489>

Abstract

This study examined whether executive functions impact how flexibly children represent task context in performing repeated sequential actions. Children in Experiments 1 ($N = 52$; 3–6 years) and 2 ($N = 50$, 4–6 years) performed sequential actions repeatedly; one group received reminders. Experiment 1 indicated that reminders promote flexible changes of contextual representations. Experiment 2 observed such effects in younger children and showed that executive functions were associated with the flexible representation of task context. Reminders did not perfectly compensate for the role of executive functions but wiped out individual differences of executive functions that contribute to children's acquisition of routines. Therefore, setting goals before context-dependent actions is necessary, but not sufficient, to modulate contextual representations in routines.

Keywords: routine, executive functions, preschoolers

Contribution of Executive Functions to Learning Sequential Actions in Young Children

In their everyday lives, even young children learn to perform sequential actions toward achieving a goal. For example, when young children prepare for school in the morning, they must keep many actions in mind, such as changing their clothes, eating breakfast, brushing their teeth, and packing a school bag. One experience may not be sufficient for them to master sequential actions; however, as they repeatedly perform those actions, they gradually learn them as routines, which involve relatively automatic processes toward goals. This has been evidenced by developmental research demonstrating that young children can imitate novel arbitrary sequential actions, report what happens in familiar events, and accurately perform sequential actions without models through repetition (e.g., taking a bath, making a sandwich; Bauer & Hertsgaard, 1993; Fivush, 1984; Freier, Cooper, & Mareschal, 2015; Loucks, Mutschler, & Meltzoff, 2017; Loucks & Price, 2019; Yanaoka & Saito, 2017). Although previous studies showed such competence in young children, little attention has been paid to how children learn routines and what cognitive abilities support this process.

Among the promising approaches to this issue is theorization with a computational model, which explains how adults learn routines and focuses on the supportive roles of executive control systems (Cooper, Ruh, & Mareschal, 2014). In early childhood, children experience pronounced developmental improvements in executive functions, which are regulatory processes involved in the conscious control of goal-directed behaviors (e.g., Diamond, 2013; Garon, Bryson, & Smith, 2008; Miyake & Friedman, 2012; Zelazo et al., 2013). Therefore, although it can be assumed that the development of executive functions would help preschoolers learn routines, developmental studies on learning routines and on executive functions have been conducted separately. Thus, in the present study, by extending the recent computational model

for adults (Cooper et al., 2014), we aimed to investigate the contributions of executive functions to the acquisition of routines among young children.

Executive Functions and Learning Routines

To consider the interaction between learning routines and executive functions, we reviewed Norman and Shallice's (1986) dual systems theory. In this theory, conceived contention scheduling and the supervisory attentional system are assumed to account for goal-directed behaviors. In routine situations, contention scheduling is held to select one activated action schema automatically in response to either environmental cues or task-specific ordering constraints. Contrastingly, the operation of contention scheduling is not enough to control actions in novel or less familiar situations, and the supervisory attentional system needs to modulate the function of contention scheduling via top-down control.

Based on the dual systems theory, Cooper et al. (2014) developed the goal circuit model. This model updated a simple recurrent network (Botvinick & Plaut, 2004), which maps perceptual input units to action output units via a set of internal hidden units, by adding units of the supervisory attentional system and hierarchical goal representations to the network. In the present study, following Botvinick and Plaut (2006), we defined a goal as an explicit representation of a desired outcome that is matched against action effects as a part of a process of mean-ends analysis. The inclusion of goal units allows for what we consider to be a more psychologically plausible approach to learning. When learning novel sequential actions, the model takes a representation of the current goal as input and aims to generate an action and a representation of the subsequent goal at each hierarchical level (e.g., one node represents the goal of making coffee, while another represents the sub-goal of adding sugar). Goal representations are possibly modulated by the supervisory attentional system and fed back into

the model in the next processing step. Thus, the model is trained to encode sequential actions to achieve goals. In line with this, previous research demonstrated that children develop the ability to control sequential actions through maintaining hierarchical goal representations from early childhood to middle childhood (Freier, Cooper, & Mareschal, 2017; Perone, Anderson, & Youatt, 2020; Yanaoka & Saito, 2017).

With repetition of a recurrent process, an implicit representation of task context is gradually accumulated in the model's internal hidden units. This implicit contextual representation might comprise information regarding what has been seen and done previously and which goals have been active. Contextual representations might allow the model to select the next appropriate action, resulting in influences from the supervisory attentional system presumably being reduced and routines being learned. Consistent with this model, Ruh, Cooper, and Mareschal (2010) demonstrated the role of repetition in learning routines. Adult participants were asked to select actions to make beverages in a computerized task and were found to select context-dependent actions more slowly when concurrently performing a secondary, attentionally demanding task. Importantly, interference from the secondary task was reduced with increasing practice, suggesting that the repetition of sequential actions gradually decreases demands on the supervisory attentional system.

Therefore, according to Cooper et al. (2014), while learning sequential actions is goal-based and requires top-down control at the beginning, the need for top-down control gradually reduces as sequential actions are repeatedly performed and contextual representations take over the control of actions.

Flexible Modulation of Contextual Representations

Although Cooper et al. (2014) assumed that implicit contextual representations underlie routines, it was still not fully understood how contextual information was represented over the course of routines. Botvinick and Plaut (2004), who provided the basis for Cooper et al.'s study (2014), directly examined this issue. They argued that flexibly representing contextual information depending on its necessity in selecting actions leads to the acquisition of routines. Botvinick and Plaut (2002) provided an example of routines for making coffee or tea. Although routines for making both coffee and tea are composed of common subtasks—such as boiling hot water, preparing a cup, and pouring hot water—some unique actions rely on specific contextual information (i.e., coffee context or tea context), such as adding cream to coffee but not tea. To select the appropriate context-dependent action, it is necessary to access the contextual information about which goals have been active (e.g., making coffee or tea). Thus, contextual information is learned to be clearly represented immediately before choosing whether it is necessary to add cream. In contrast, contextual information is learned to be vaguely represented during common subtasks, because it is not necessary when performing them. Thus, contextual representations flexibly change over the course of routines, based on whether the information is relevant to a particular action. According to Botvinick and Plaut (2002, 2004), this flexible modulation of contextual representations can account for many phenomena related to routines, such as error-free performance, action slips, and generalization.

Botvinick and Bylsma (2005) conducted the first experimental study to test how contextual representations could flexibly change over the course of repeated sequential actions by requiring participants to repeatedly make a cup of coffee. They implemented a short interrupting task, then examined the frequency of action slips that occurred after participants returned to making coffee. Making a cup of coffee involved putting coffee grounds in a filter,

adding sugar, and adding milk. These processes relied on contextual information, such as whether the sugar had already been added, in which case there would be no further need to add sugar. If the routine had been acquired, contextual information on adding sugar would be clearly represented immediately before choosing whether to add sugar; however, while adding milk, the contextual information would be represented vaguely because it was not relevant to action selection. Thus, it can be anticipated that the vaguer the distinction between two representations becomes (i.e., context for adding or not adding sugar), the easier it becomes to forget whether a task should be performed, which is more easily affected by interrupting tasks. In the experiment, brief interruptions were inserted during the subtasks of adding milk or sugar. The interrupting task within the subtask was either in the middle (e.g., while pouring ingredients into the cup) or the end (e.g., when stirring). The results demonstrated that inserting an interrupting task in the middle of the subtask generated more errors in selecting a context-dependent action than at the end. This finding clearly indicated that contextual representations become attenuated in the middle of a subtask sequence, but recover at the end, suggesting that contextual information is flexibly represented depending on its necessity in selecting actions.

Executive Functions and Flexible Modulation of Contextual Representations in Preschoolers

According to Cooper et al.'s (2014) model, executive functions promote how efficiently we learn routines, that is, how efficiently contextual information is flexibly represented during the performance of repeated sequential actions. To date, few studies have directly examined the relation between executive functions and flexible modulation of contextual representations. In one such pioneering study, Yanaoka and Saito (2019) created a routine acquisition task for young children, in which they were asked to repeatedly make toast, following one recipe for a

mouse and another for a cat. Thus, there were two task contexts for attaining task goals, that is, making toast for a mouse and a cat. Importantly, as in the example from Botvinick and Plaut's study (2002), some common subtasks (e.g., spreading butter on toast) were performed in both contexts, whereas one unique subtask (e.g., putting cheese on toast) was specific to only one context. The point where transition occurs between common subtasks and a context-dependent subtask is called a *branch point*. One might expect that the representations of task contexts for a mouse and a cat would become degraded in the middle of the common subtasks; however, contextual representations were assumed to recover and be more salient at the end of the common subtasks (i.e., just before the context-dependent subtask). Following Botvinick and Bylsma's study (2005), interrupting tasks were inserted either after completion of the task immediately preceding the branch point (end interruption) or during the task immediately preceding the branch point (middle interruption). The results indicated that young children with high executive functions were more likely to be more affected by the middle interruption than the end interruption, whereas young children with low executive function tended to be vulnerable to both interruptions. This finding suggested that the representational flexibility of task context underlying children's performance of repeated sequential actions is associated with executive function development.

Yet, how executive functions contribute to the mechanisms by which young children learn routines remained unclear. Consistent with Cooper et al.'s (2014) model, one can predict that executive functions, which are strongly associated with maintenance of goal representations, promote learning routines. Therefore, to examine theoretically driven hypotheses, the present study aimed to address whether and how executive functions directly influence efficient flexible modulation of contextual representations in young children.

The Present Study

According to Cooper et al. (2014), learning sequential actions is initially directed toward achieving main goals and sub-goals; thus, goal representations need to be maintained over the course of sequential actions. Particularly at a branch point, having access to task goals is essential in selecting context-dependent actions at the beginning of learning sequential actions. Then, repeating experiences that task goals could be remembered just before a branch point and a context-dependent action is appropriately selected would lead to efficient accumulation of representations of task context, resulting in the acquisition of routines. The maintenance of task goal representations has been shown to be one common construct among executive functions (Miyake & Friedman, 2012) and a strong driver of executive function development (Munakata, Snyder, & Chatham, 2012). Thus, we expected that executive functions would play an essential role in learning routines through having access to goal representations at a branch point.

In the present study, to manipulate cognitive demands on executive functions, especially goal maintenance, we verbally and visually reminded children of a task goal before a branch point. It has been acknowledged that reminders are effective for activating goal representations in young children. For example, Barker and Munakata (2015) demonstrated that prompting the activation of goal representations through reminders played a causal role in improving inhibitory control. Kliegel and Jäger (2007) also found that visual reminders based on the content of a future intention (e.g., an apple to remind children to look for an apple cue) were effective for helping even 3-year-olds remember to carry out intended actions. Thus, reminders are expected to compensate for activating goal representations at a branch point in children with low executive functions and improve their ability to learn routines. As we were primarily interested in the acquisition of routines, we inserted reminders into only the practice phase, in which children

were first asked to make toast for either a mouse or a cat six times without interruptions, during which they repeatedly experienced sequential actions for making toast. It was expected that children would thus acquire the routine at an early stage of the task. Yanaoka and Saito (2019), who conducted the toast-making task across two days, found no main effects of the day on how children flexibly represented task context. Thus, if executive functions promote the flexible modulation of contextual representations, they might play a critical role in the early stages of a task. Therefore, we expected that reminders would be effective, even if they were inserted into the first six practice trials.

Experiment 1

Using the toast-making task (Yanaoka & Saito, 2019), Experiment 1 aimed to examine the direct influence of executive functions on representational flexibility of contextual information in young children, through reminder manipulation in the practice phase (with reminders vs. without reminders; the former called the “reminder group” and the latter the “no reminder group”). We had two hypotheses. First, young children would be affected by interruptions in the middle and at the end of a subtask in the no reminder group (i.e., no effect of interruption timing), whereas children in the reminder group would produce more branch point errors as a result of an interruption in the middle of a subtask rather than at the end (Hypothesis 1). Second, executive functions would modulate the effects of reminders on how task context was flexibly represented; thus, reminders would be effective only for children with low executive functions (Hypothesis 2). Yanaoka and Saito (2019) reported that the representational flexibility of task context changes at around age 4 to 5; thus, we collected data from 3-year-olds to find clear effects of reminders. Additionally, we included 4- and 5-year-olds to examine the interaction between executive functions and flexible modulations of contextual representations.

Methods

Participants. Generalized linear mixed models (GLMMs) were used for our analyses; thus, we carried out Monte Carlo simulations using the `powerSim` function from the `simr` package in R (Green & Macleod, 2016) and determined the minimum sample size for sufficient statistical power. The simulation revealed that a sample of $N = 48$ participants yielded a power of $\beta = 0.95$ (95% CI = 0.93, 0.96) to detect medium fixed-effects of size 0.5 for the interaction between the reminder factor, the timing factor, and advanced Dimensional Change Card Sort (DCCS) task performance, which was our main interest, at an α level of .05. A total of 52 children (28 boys and 24 girls) attending preschool in Japan participated in Experiment 1; however, one five-year-old ($n = 1$), one four-year-old ($n = 1$), and two three-year-olds ($n = 2$) were excluded because they could not understand instructions and complete all tasks. Thus, 48 children were included in the final analyses and divided into either the reminder group ($M = 55.5$ months, $SD = 9.45$ months, age range: 40–72 months) or the no reminder group ($M = 56.2$ months, $SD = 9.76$ months, age range: 40–74 months). None of the participants had any history of neurological disorders or neurodevelopmental delays, and all were native Japanese speakers. Their socioeconomic background was predominantly middle class. Informed consent was obtained from the parents of all children prior to their participation. This and subsequent experiments were approved by the appropriate institutional ethics committee.

Procedure. All participants performed the toast-making task and the following four tasks, which are often used for assessing executive functions: the red-blue task (Simpson & Riggs, 2005, revised), standard DCCS task (Zelazo, 2006), advanced DCCS task (Chevalier & Blaye, 2009), and visual counting span test (Case, Kurland, & Goldberg, 1982). These tasks were conducted across two days. The toast-making task was conducted on both days. The two types of

DCCS tasks were conducted on the same day, and the other two tasks were conducted on another day. The order of executive function tasks was counterbalanced across participants between the two days. To make the children feel relaxed, testing was conducted in a quiet room at the kindergarten on both days. The experimenter spent several days establishing rapport with the children prior to the experiment. On both days, testing required approximately thirty minutes.

Measures

Toast-Making Task

Materials. We used the same materials used in Experiment 2 of Yanaoka and Saito's (2019) study. To represent ingredients necessary for making toast, we prepared toy bread, cream-colored clay, yellow felt, and configured paper clay, which acted as substitutes for toast, butter, cheese, and fried eggs, respectively. Each ingredient was placed in a white plastic box. The boxes were arranged in a specific sequence—from left to right, toast, butter, cheese, and fried eggs—and a toaster oven made of cardboard was adjacent to the box of fried eggs. There were also paper plates where the ingredients were placed. A laptop computer (Microsoft Surface Pro 4) was set up on an adjacent desk. In the trials with interruptions, each child would hear a telephone ringing through the computer in the middle of making toast. After the child touched an image of a telephone on the computer screen to stop the sound, three digits between 1 and 9 were auditorily presented, at a rate of one per second; no digit could appear twice in the same trial. The child was then asked to recall the digits.

Procedure. The procedure for the toast-making task closely followed that described by Yanaoka and Saito (2019). A child was asked to play the role of baker and make toast for either a “mouse” or a “cat,” using the provided materials. The experimenter demonstrated how to make toast for either a mouse or a cat using different recipes with verbal instructions. For a mouse, the

experimenter picked a plate, placed two pieces of toy bread on the plate, spread yellow clay on the toast, put yellow felt on the clay, put configured paper clay on the toast, placed the toast in the oven and closed the door, counted for five seconds until it warmed up, then set the completed toast on an adjacent desk. The process of making toast for a cat was similar to that for a mouse, with the exception that the toast was not placed in the oven. When demonstrating how to make toast, the experimenter told the child that the reason for the difference in the recipes was that a mouse liked hot toast and a cat could not eat hot toast. The order of the demonstration of the recipe for a mouse and a cat was counterbalanced. After the demonstrations, the experimenter asked the child to verbally describe how to make toast for a mouse or a cat, and all the children were able to answer the question.

First, children engaged in six practice trials (three mouse and three cat trials). The mouse and cat trials were performed in a quasi-random sequence in which one type of recipe did not appear three times consecutively (e.g., cat, cat, mouse, cat, mouse, mouse). At the beginning of each trial, each child was presented with the picture of a mouse or a cat and verbally encouraged to make toast for that animal. To prevent the child from checking a task goal again while making toast, the pictures were hidden. After finishing the practice trials, children practiced how to deal with the interruptions with the experimenter for a short time. Specifically, children were told that as soon as they heard the phone ring during the preparation of toast, it was necessary to stop what they were doing and then answer the call. Touching the image of a telephone appeared on the laptop computer, and they were asked to recall three lists of the three digits each, in the correct order. Immediately after the interruption, they were instructed to resume making toast. The experimenter verbally explained the task procedure with interruptions, and then the children experienced such trials twice.

Finally, the experimenter verbally instructed the process of making toast for a mouse and a cat again, and then each child performed 12 test trials per day across two days. After every sixth trial, the children were given a short break, during which they received stickers as a reward. Two-thirds of the test trials were interrupted using a similar procedure to that used by Botvinick and Bylsma (2005). Each of the interrupted trials was interrupted either at the middle or the end of putting fried eggs on toast, which was followed by the branch point of placing toast in the oven for a mouse, but not for a cat. The middle interruption occurred immediately after beginning to put fried eggs on toast, whereas the end interruption occurred as the child finished putting fried eggs on all four pieces of toast. Test trials were composed of four blocks with one mouse trial and one cat trial in three conditions: no-interruption control, middle interruption, and end interruption. In line with the practice trials, the test trials were also performed in a quasi-random sequence in which trials with one type of recipe did not appear three times in a row.

We conducted the experimental manipulation during the procedure of the practice trials. In both the reminder and no reminder groups, feedback was provided after each practice trial. If the child succeeded in making toast in accordance with task context for attaining a goal (i.e., making toast for a mouse or a cat), the experimenter said, “very good work. You made toast for a mouse (cat).” However, if the child failed, the experimenter said, “OK. You needed (not) to put the toast in the oven when you took an order from a mouse (cat).” In the middle of the practice trial, the experimenter did not intervene when the child made a mistake. It was a unique feature of the reminder group that the child was prompted to remember a task goal during the practice trials. Specifically, the experimenter showed a picture of a mouse or a cat to the child and said “you are now making toast for a mouse (cat)” just before and after the fried-eggs step. We decided to provide the reminder at this step because putting fried eggs on the plate was

performed directly before the branch point at which task goals should be accessed. The experimenter did not remind the children in the no reminder group whom they were making toast for. The reminder was not included in the test trials for either group, but only in the practice trials for the reminder condition.

Scoring. Our index was the rate of action selection errors that occurred at the transition from putting fried eggs on toast to placing it in the oven and closing the door across test trials. We called these errors “branch point errors,” and they corresponded to the intrusion from the other context to the currently relevant context, resulting in context-inappropriate action. Branch point errors were divided into two specific patterns. One was that the child placed toast in the oven and closed the door, or almost placed it in the oven, when making toast for a cat. The other one was that the child took toast to the adjacent desk without placing it in the oven and closing the door, or almost took it to the adjacent desk but stopped, when making toast for a mouse. Following the scoring system used in previous studies (Botvinick & Bylsma, 2005; Yanaoka & Saito, 2019), to maximize error detection sensitivity, not only intrusion errors but also microslips (Schwartz, Reed, Montgomery, Palmer, & Mayer 1991; Giovannetti, Schwartz, & Buxbaum, 2007) were included as branch point errors. In the current study, intrusion errors were defined as failures to select context-appropriate actions without stopping them, and microslips were defined as the initiation and termination of an incorrect action before the error was completed. For example, in the context of making toast for a cat, if a child finished placing toast in the oven and closing the door, the action would be an intrusion error, whereas if a child stopped putting toast in the oven and brought it to the adjacent desk, the action would be a microslip.

All trials were coded by the experimenter and a research assistant—an undergraduate psychology major—based on a video recording. The two coders identified 12 trials with

imprecise timing for interruptions, which were excluded from later analysis. The 12 trials were broken down into four trials for 3-year-olds, two trials for 4-year-olds, and six trials for 5-year-olds. Each child experienced only one dropped trial, except for one 5-year-old who experienced the trial twice. Furthermore, 12 trials were divided into 7 middle interruption trials and 5 end interruption trials. The remaining trials were evaluated for branch point errors (i.e., no error, intrusion error, and microslip). Interrater agreement for the coding was good (Cohen's kappa = 0.80), and coding disagreements were resolved through discussion.

Red-blue task. This task was conducted to measure children's inhibition abilities, and the procedure was essentially the same as that of the black-white task used by Simpson and Riggs (2005). During the warm-up phase, each child was presented with a red card and a blue card and asked to name the card's color. All the participants were able to answer correctly. Then, each child was instructed to point to the red card when the experimenter said "blue" and the blue card when he said "red." In the practice phase, two trials were given to each child, and if the child responded correctly to two trials in succession, the experimenter praised the child and the test phase began. If the child answered incorrectly in the practice phase, the task instruction was repeated. In the test phase, the children engaged in 16 trials, during which the experimenter said "red" and "blue" eight times each, in a pseudorandom sequence. No feedback was provided during the test phase. The dependent measure was the number of correct trials during a session (0–16).

Standard DCCS. This task was conducted to measure switching abilities. The procedure closely followed the protocol described by Zelazo (2006). The experimenter showed each child two trays and two target cards and named the target cards using two dimensions (i.e., color: yellow and green; shape: cup and ship). This task had three phases: demonstration, pre-switch,

and post-switch. In the demonstration phase, either the shape game or color game was introduced, and the experimenter presented the child with two practice trials. For example, in the color (or shape) game, each child was instructed to place all of the yellow cards (or all of the cup cards) in the box with the yellow cup, and to place all of the green cards (or all of the ship cards) in the box with the green ship. If the child sorted two consecutive trials correctly, the demonstration phase ended, and the pre-switch phase began immediately; otherwise, the task instruction and practice trials were repeated. In the pre-switch phase, each child was reminded of the pre-switch rule again and asked to sort six test cards according to the pre-switch rule. The test cards were presented in a pseudorandom sequence, and no feedback was provided during the pre-switch phase. Once the pre-switch phase was completed, the experimenter announced that the rule had changed, and each child was instructed to sort the test cards according to the new sorting rule. In the post-switch phase, each child was presented with six more test cards and asked to sort them. The rule was counterbalanced across participants. The total score represented the proportion of correct trials only during the post-switch phase.

Advanced DCCS. This task was also conducted to measure shifting abilities, using a more advanced version of the Standard DCCS task. After finishing all the phases of the Standard DCCS task, the experimenter introduced one more new game to each child, in which the child was asked to sort cards according to one of two dimensions (color and shape) at random. Previous studies (Chevalier & Blaye, 2009; Zelazo et al., 2013) indicated that processing arbitrary visual cues regarding when to switch rules was too demanding for 3- to 4-year-olds, and explicit verbal cues could help them realize when to switch. Thus, for each trial, the experimenter explicitly told children which dimension they were required to sort. Each

participant completed eight trials in a pseudorandom sequence, including four color trials and four shape trials. Their scores represented the number of accurate trials.

Visual counting span task. This task was conducted to measure working memory and was originally designed by Case, Kurland, and Goldberg (1982). In this task, several sheets of paper, on which two types of objects—frogs and cats, stars and rainbows, or suns and hearts—were depicted, were placed face down on a desk. The experiment turned over one sheet at a time, and each child was asked to count one type of object (e.g., frog) before turning the sheet face down again. The maximum number of the target objects was five and the minimum was one. After all the sheets were counted, the child was asked to recall the counts for each set in the correct order. First, children were given a practice trial, in which they had to count and recall two sheets. In the test phase, there were three trials at each length sequence (2, 3, 4, and 5 span). In the test phase, if the child perfectly recalled more than one trial in each span, he/she proceeded to the next sequence length; otherwise, the testing ended. Following Conway, Kane, Bunting, Hambrick, Wilhelm, and Engle (2005), we calculated the sum of the proportion of elements recalled in the correct serial position within each trial, across all trials.

Results and Discussion

Data analysis. In the toast-making task, a trial was scored as incorrect if the child made a branch point error. Children were more likely to make intrusion errors than microslips, and the number of microslips was small. The patterns of microslips in each condition were not entirely different from those of intrusion errors in each condition, thus we reported the results from the total score (intrusion error + microslip). We analyzed binary data with mixed-effects logistic regressions using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in the R project for statistical computing (R Core Team, 2013). With the rate of branch point errors as a

dependent variable, we were interested in two comparisons. First, we compared the error rate in the control condition with the average error rate in the two interruption conditions. This comparison was identified as an *interruption factor* and indicated whether the child was affected by the interruptions that occurred while making toast. Another comparison was made between the error rates in the middle and end interruption conditions. This comparison was identified as a *timing factor* and indicated that the degree to which the interruptions induced branch point errors depended on the timepoint (middle or end). Thus, we conducted planned comparisons and adapted Helmert coding using an interruption timing factor (control, middle, end). Contrast coding provides a series of precise comparisons among the group means. Helmert coding is one kind of contrast coding and compares each level of a categorical variable to the mean of the subsequent levels. To identify the interruption factor, we coded the control condition as 0.67, the middle interruption condition as -0.33, and the end interruption condition as -0.33. To identify the timing factor, we coded the control condition as 0, the middle interruption condition as -0.5, and the end interruption condition as 0.5. The following variables were coded to be centered: reminder factor (reminder group = 1, no reminder group = -1), animal (mouse = 1, cat = -1), day (day 1 = -1, day 2 = 1), and gender (boy = 1, girl = -1).

Using Akaike's Information Criterion (AIC), we sequentially deleted fixed terms from a full model (main effects and first- and second-order interaction effects) to find the most parsimonious model that fit the data. To establish the significance of the effect of predictors, we ran likelihood ratio tests comparing the full model with the respective reduced model (without the predictor to be tested; Barr, Levy, Scheepers, & Tily, 2013).

Executive functions and the toast-making task practice trials. We examined whether executive function development was matched between the reminder and no reminder groups;

thus, Welch's t-tests using group (between: reminder or no reminder) were conducted. As shown in Table 1, we did not identify significant differences between the two groups in any executive function tasks. Next, to find the effect of reminders on performance in practice trials during the toast-making task, we also conducted a Welch's t-test with performance in practice trials. It revealed a significant effect of reminders, indicating that reminders effectively decreased branch point errors during the practice trials. Indeed, children in the reminder group exhibited nearly perfect performance in the toast-making task (see Table 1). In summary, we could match the baselines of the two conditions and confirm the reminders functioned, as expected.

[Table 1 goes here]

Reminders and the acquisition of routines. Preliminary analyses indicated no main effect of day ($b = 0.08, z = 1.03, \chi^2 = 1.03, df = 1, p = .309$) or gender ($b = -0.05, z = -0.45, \chi^2 = 0.43, df = 1, p = .445$), with no differences in performance between Day 1 and Day 2 or between boys and girls; thus, day and gender factors were excluded from subsequent analyses. Our first analysis explored whether the reminder factor could moderate the extent to which both interruption and timing factors affected branch point error rates. For fixed effects, the model included a timing factor, an interruption factor, a reminder factor, and two-way interactions (Reminder factor x Timing factor, Reminder factor x Interruption factor). As random factors, random intercepts for participant and animal were included.

The best model (AIC = 1034.3) fitted significantly better than the null model (AIC = 1053.9, $p < .001$). The main effect of the interruption factor was shown to be significant ($b = 0.80, z = 4.20, \chi^2 = 18.87, df = 1, p < .001$), which indicated that both types of interruptions led to more branch point errors, compared to the control condition. The main effect of the timing factor was shown not to be significant ($b = 0.25, z = 1.34, \chi^2 = 1.76, df = 1, p = .185$); however, it was

more important that we found a significant interaction between the reminder and timing factors ($b = 0.44, z = 2.38, \chi^2 = 5.61, df = 1, p = .018$). As shown in Figure 1, the error rate in the middle interruption condition was significantly higher than in the end interruption condition for the reminder group ($b = 0.69, z = 2.63, p = .009$), whereas the difference of interruption timing was not significant in the no reminder group ($b = -0.06, z = -0.24, p = .809$). This finding supported Hypothesis 1 and suggested that reminding children of task goals before the branch point during the practice phase could promote their acquisition of a routine, as evidenced by the flexible modulation of the contextual representations for the toast-making task.

[Figure 1 goes here]

Reminders, acquisition of routines, and executive functions. Our second analysis aimed to examine whether executive functions moderated the interaction between reminders and the differential effects of the two types of interruptions on branch point errors. Thus, we added the factor of age in months, which we treated as a continuous variable, and related factors for executive function tasks to the above model that investigated reminder effects. To examine the relation between the four executive function tasks and the toast-making task, correlation analyses for each group, controlling for age, were conducted (see Table 2). For the no reminder group, advanced DCCS task performance was significantly correlated with toast-making task performance only for the end interruption condition. This pattern, indicating the close relation between advanced DCCS task performance and overcoming only the end task interruption, was consistent with Yanaoka and Saito's finding (2019); thus, we reported a model including the interactions related to the advanced DCCS task. Our analysis model included the main effects of age in months, four executive function tasks, the interruption factor, the timing factor, two-way interactions (Advanced DCCS x Timing factor, Advanced DCCS x Interruption factor), and

three-way interactions (Reminder factor x Advanced DCCS x Timing factor, Reminder factor x Advanced DCCS x Interruption factor) as fixed-effect variables. The random effect variables consisted of a random intercept for participant and animal.

[Table 2 goes here]

To examine the specific roles of executive functions, we reported the parsimonious model including the factor of age in months as a control variable (see Table 3). The model (AIC = 1018.1) fitted significantly better than the null model (AIC = 1053.9, $p < .001$). Red-blue task performance significantly affected the rate of branch point error. This is consistent with the results of Experiment 1 of Yanaoka and Saito's study (2019). The red-blue task was also shown to be related to the execution of familiar sequential actions based on hierarchical goal representations (Yanaoka & Saito, 2017). The red-blue task potentially reflects children's abilities to maintain goal representations during the performance of the toast making task and their abilities to inhibit the goal that is not relevant in the current trial. Furthermore, both a main effect of the interruption factor and a significant interaction between the reminder factor and the timing factor were confirmed again, even when controlling for red-blue task performance. Unlike the results of Yanaoka and Saito (2019), the two-way interaction between the advanced DCCS performance and the timing factor was not significant and was therefore excluded from the best-fitting model (more discussions in Supporting Information S1). Importantly, contrary to Hypothesis 2, the three-way interaction between the reminder factor, advanced DCCS task performance, and the timing factor was excluded from the best-fitting model. In terms of this predictor, a likelihood ratio test revealed that the three-way interaction was marginally significant ($\chi^2 = 3.13$, $df = 1$, $p = .077$); thus, this evidence was not strong enough to improve the model fitting.

However, executive functions might be related to the acquisition of routines. Based on correlation analyses (see Table 2), we identified different patterns between the reminder group and the no reminder group. For the no reminder group, red-blue task performance was related to the correct rates of the toast-making task for all conditions, and advanced DCCS task performance was also related to toast-making task performance for the end interruption condition. In contrast, for the reminder group, there was only one significant correlation between the red-blue task and the toast-making task for the end interruption condition. These results suggested that reminders could decrease heavy demands on executive functions needed in the toast-making task.

Taken together, to the best of our knowledge, Experiment 1 provided the first evidence of the effects of reminders on learning routines—that is, the flexible modulation of contextual representations. However, contrary to our prediction, the reminder effect did not interact with advanced DCCS task performance. Thus, we could not specify how reminders promoted the acquisition of routines and draw a strong conclusion that executive functions promote learning routines.

[Table 3 goes here]

Experiment 2

One possible reason for the results of Experiment 1 that did not support our predictions was that the sample in Experiment 1 included a high proportion of 3-year-olds who exhibited floor effects on advanced DCCS performance; thus, the variance in advanced DCCS performance became too small (see Table S2 in Supporting Information S2). Therefore, Experiment 1 could not replicate even the two-way interaction between advanced DCCS performance and the timing factor observed by Yanaoka and Saito (2019) (but see Supporting

Information S1). Thus, as in Yanaoka and Saito's study (2019), Experiment 2 targeted children over the age of four, with the expectation that advanced DCCS performance in older age groups would have more variation. Furthermore, it was expected that advanced DCCS performance would moderate the effects of reminders on susceptibility to two kinds of interruptions. Overall, Experiment 2 aimed to replicate the effects of reminders on learning routines and further demonstrate the interaction between the reminder effects and executive functions.

Methods

Participants. Following the estimated sample size of Experiment 1, we included a total of 50 four- to six-year-olds in Experiment 2; however, two children were excluded because they could not finish all trials in the toast-making task. Thus, the final analysis included 24 children in the reminder group ($M = 62.9$ months, $SD = 6.84$ months, age range: 52–72 months) and 24 children in the no reminder group ($M = 61.5$ months, $SD = 6.94$ months, age range: 51–72 months), none of whom took part in Experiment 1. Their predominant socioeconomic background was also middle class.

Procedure. In Experiment 2, participants were given the same tasks as in Experiment 1. The tasks were conducted in the same quiet room at the kindergarten in 30-minute sessions across two days.

Measures.

Executive function tasks. We conducted the same four tasks to measure executive functions as in Experiment 1. In terms of the advanced DCCS task, we increased the number of trials, because the change could make children's performance evenly distributed enough to accurately measure developmental differences in their switching abilities. Thus, in the advanced DCCS task, each participant underwent 16 trials, including eight color and eight shape trials.

Toast-making task. We used the same procedures, materials, and scoring system as in Experiment 1. We identified 14 trials as having imprecise interruption times, and these trials were excluded. The 14 trials were broken down into five trials for 4-year-olds, seven trials for 5-year-olds, and two trials for 6-year-olds. Each child experienced only one dropped trial, except for two 5-year-olds who experienced the trial twice. Furthermore, 14 trials were divided into 7 middle interruption trials and 7 end interruption trials. The remaining trials were evaluated for branch point errors. We scored branch point error rates, including microslips, across 24 trials. We randomly selected 30 children among all the participants before we started coding (15 children from the reminder group, and 15 children from the no reminder group). All the trials completed by them were coded by the experimenter and a different coder from Experiment 1, who also majored in psychology. Interrater agreement for the coding was excellent (Cohen's $\kappa = 0.93$), and coding disagreements were resolved through discussion. The trials completed by the other 18 children were coded by the experimenter alone.

Results and Discussion

Executive functions and the toast-making task practice trials. Using Welch's t-tests, we confirmed that executive function task performance was matched between the reminder and no reminder groups (see Table 1). Regarding the effects of reminders on practice trials for the toast-making task, we also conducted a Welch's t-test, which showed that performance in the reminder group was significantly higher than in the no reminder group. Thus, 4- to 6-year-olds were shown to benefit from reminders in performing sequential actions, in accordance with task context for making toast for a mouse or a cat.

Reminders and acquisition of routines. Preliminary analyses indicated no main effect of day ($b = 0.01, z = 0.12, \chi^2 = 0.01, df = 1, p = .903$) or gender ($b = -0.03, z = -0.32, \chi^2 = 0.02, df = 1, p = .771$); thus, day and gender factors were excluded from subsequent analyses.

As in Experiment 1, we conducted analyses using mixed-effects logistic regression with branch point error rates in the toast-making task as a dependent variable. To examine the effects of reminders on branch points error rates in both the middle and end interruption conditions, we tested the same model as in Experiment 1, which included the reminder factor, the timing factor, the interruption factor, and their two-way interactions (Reminder factor \times Timing factor, Reminder factor \times Interruption factor). As random effect variables, random intercepts for participant and animal were included. Additionally, the best-fitting model was determined based on AIC, and a likelihood ratio test was used for significance testing, as in Experiment 1.

The best model (AIC = 1134.1) fit significantly better than the null model (AIC = 1153.5, $p < .001$). It was revealed that the average rates of branch point errors in the middle and end interruption conditions were higher than those in the control condition (Interruption factor: $b = 0.72, z = 4.17, \chi^2 = 18.48, df = 1, p < .001$). Unlike in Experiment 1, the timing factor was also significant ($b = 0.52, z = 2.93, \chi^2 = 8.66, df = 1, p = .003$), which indicated that the middle interruption condition led to more branch point errors than the end interruption condition. Additionally, the two-way interaction between the reminder and timing factors was not significant; thus, the flexible modulation of contextual representations did not interact with the presence of reminders in 4- to 6-year-olds (see Figure 2).

[Figure 2 goes here]

Reminders, acquisition of routines, and executive functions. To investigate the possibility that the two-way interaction between the reminder and timing factors interacted with

the development of executive functions, we added the factor of age in months and factors related to executive function tasks to the model demonstrating reminder effects. Particularly, in line with Experiment 1, we reported the results of the model, adding the main effects of age in months and the four executive function tasks, two-way interactions (Advanced DCCS x Timing factor, Advanced DCCS x Interruption factor), and three-way interactions (Reminder factor x Advanced DCCS x Timing factor, Reminder factor x Advanced DCCS x Interruption factor). Random intercepts for participant and animal were included for random effect variables.

As in Experiment 1, we reported the parsimonious model including the factor of age in months as a control variable (see Table 3). The model (AIC = 1122.6) fitted better than the null model (AIC = 1053.9, $p < .001$). It was shown that red-blue task performance and visual counting span task performance were positively associated with toast-making task performance. It is possible that the relations reflect the processes of keeping goal relevant information in working memory and of inhibiting currently irrelevant goal (Marcovitch et al., 2010; Yanaoka, Moriguchi, & Saito, 2020). The relation of the red-blue task to the toast making task was consistent with Experiment 1, but that of the visual counting span task was first observed in Experiment 2 (more discussions in Supporting Information S3). Furthermore, both the timing and interruption factors remained significant. It is noteworthy that the two-way interaction between advanced DCCS performance and the timing factor was significant. As shown in Figure 3, simple effects revealed that children with high advanced DCCS performance showed fewer branch point errors in the end interruption condition than the middle interruption condition ($b = 0.87, z = 3.25, p = .001$), whereas children with low advanced DCCS performance were not sensitive to the timing of the interruptions ($b = 0.13, z = 0.51, p = .610$). It is important to note that this finding replicates Yanaoka and Saito's finding (2019); however, our main prediction

was not strongly supported. There was a trend of a significant three-way interaction between the reminder factor, advanced DCCS task performance, and the timing factor. Although we acknowledged that the effect was insufficient, we conducted pairwise tests to see the pattern of results (see Figure 4). Consistent with our prediction, the two-way interaction between advanced DCCS performance and the timing factor was significant in the no reminder group ($b = 0.27, z = 2.80, p = .005$), though it was not significant in the reminder group ($b = 0.01, z = -0.10, p = .917$). Specifically, in the no reminder group, children with high advanced DCCS performance were more vulnerable to the middle interruptions than the end interruptions ($b = 1.25, z = 3.38, p < .001$), but children with low advanced DCCS performance did not show such a tendency ($b = -0.25, z = -0.67, p = .505$). However, contrary to our expectations, the differences of interruption timing were not significant in both children with high and low advanced DCCS in the reminder group (High advanced DCCS: $b = 0.52, z = 1.33, p = .183$, Low advanced DCCS: $b = 0.50, z = 1.52, p = .129$). These findings suggest that the reminder factor tended to modulate the relation between the timing factor and advanced DCCS performance, but the flexible modulation of contextual representations was not fully promoted by reminders for children with low advanced DCCS and rather was diminished by reminders for children with high advanced DCCS.

Regarding correlations between the executive function tasks and the toast-making task (see Table 2), there were few differences between the reminder group and the no reminder group, which is consistent with the absence of the interaction between reminders and the timing factor. Taken the results of Experiment 1 together, it can be expected that the reminder effect would remove heavy demands on executive functions needed to perform the toast-making task.

[Figure 3 goes here]

[Figure 4 goes here]

General Discussion

The aim of the present study was to examine whether executive functions interacted with how contextual representations flexibly change though repeatedly performing sequential actions. To attain this goal, Experiments 1 and 2 manipulated the presence of reminders, which were assumed to support children's abilities to maintain goal representations. As expected, Experiment 1 found that reminders can help 3- to 6-year-olds flexibly represent task context. Experiment 2, which targeted 4- to 6-year-olds, did not yield such reminder effects; instead, executive functions were associated with how task context was flexibly represented. Additionally, the effects of reminders on acquiring routines did not interact with executive functions in Experiment 1 but tended to interact with them in Experiment 2.

Effects of Reminders on Learning Routines

Previous studies have shown that the external environment affects how accurately adults and children perform sequential actions. For example, offering spatial cues (i.e., arranging relevant items in the appropriate order) has been shown to decrease action selection errors (Freier et al., 2015; Hudson & Fivush, 1991), while presenting distractions disrupts the selection of context-dependent actions (Freier et al., 2015; Giovanetti, Libon, Buxbaum, & Schwartz, 2002). Consistent with this, our findings indicated that external support from reminders encouraged children to decrease branch point errors in performing sequential actions during practice trials. In addition, to the best of our knowledge, we provided the first evidence that external support enables young children to represent task context flexibly. Thus, reminding a child of a task goal before a branch point only in practice trials can promote learning routines and make their internal representations change flexibly. This finding is in line with that of Cooper et al. (2014), which indicated that performance of sequential actions transitions from relying on explicit goal

representations to relying on contextual representations through repetition. Additionally, Yanaoka and Saito (2019) reported that the flexible modulation of contextual representations was not found in 4-year-olds, much less 3-year-olds; however, the present study demonstrated that even 3-year-olds could flexibly represent task context when reminded of a task goal.

The effects of reminders on the flexible modulation of contextual representations were observed in Experiment 1, but not Experiment 2. As a critical difference between the two experiments, Experiment 1 included 3- to 6-year-olds, while Experiment 2 did not include 3-year-olds. This difference likely influenced the interaction between reminders and the timing factor. That is, one can expect that the effectiveness of reminders would be seen in relatively young children between 3 and 6 years of age. To directly examine this possibility, for Experiments 1 and 2, we conducted exploratory analyses in which age-related interactions with the timing factor were included (see detail information about the analyses in Supporting Information S4). In Experiment 1, the two-way interaction between reminders and the timing factor remained significant and age in months did not moderate the interaction, whereas a significant three-way interaction among age in months, the reminder factor, and the timing factor was found in Experiment 2. In terms of the three-way interactions, it was revealed that younger children in the reminder group showed sensitivity to the timing of interruptions, whereas those in the no reminder group did not show such sensitivity. In contrast, older children in the no reminder group were significantly influenced by the timing of interruptions, but those in the reminder group were not. As mentioned previously, consistent with these results, Yanaoka and Saito (2019) revealed that before the age of four, children have difficulty with spontaneous flexible modulation of contextual representations, whereas 5-year-olds can represent task context flexibly. More importantly, reminders were shown to be effective for younger children who have

difficulty modulating flexible contextual representations, suggesting that reminders could encourage younger children to start flexibly representing task context. This interpretation can also account for the results observed in Experiment 1. Participants in Experiment 1 were partly composed of 3-year-olds, and they were younger in total than those in Experiment 2; thus, reminders easily benefit them, leading to no significant three-way interactions in Experiment 1. These findings are also consistent with those of previous studies showing that reminders are highly effective, especially for younger children (Barker & Munakata, 2015; Kliegel & Jäger, 2007; Mahy, Mazachowsky, & Pagobo, 2018). However, unexpectedly, reminders may have a disruptive effect on older children and weaken how efficiently they modulate flexible contextual representations. These results might be attributed to the use of a verbal strategy. When inserting reminders, the experimenter presented a task goal verbally, and the children were possibly encouraged to access it verbally as well. In Experiment 2, 15 out of 24 children in the reminder group (four 4-year-olds, seven 5-year-olds, and four 6-year-olds) verbalized a task goal more than once as a monologue during test trials, whereas 6 out of 24 children (two 4-year-olds, three 5-year-olds, and one 6-year-old) in the no reminder group did. In particular, 5- to 6-year-olds develop the ability to maintain goal representations proactively (e.g., Munakata et al., 2012); thus, it is possible that older children in the reminder group were more likely to continue maintaining a task goal and having access to it verbally during test trials than those in the no reminder group. As a result, their performance of sequential actions might rely on explicit goal representations for a longer period and shift slowly to relying on contextual representations; thereby, it is logical that their branch point errors were not greatly influenced by the timing of interruptions. In contrast, older children in the no reminder group did not use nor keep using the verbal strategy during test trials, resulting in that their performance of sequential actions would

change from relying on goal representations to relying on contextual representations efficiently. To validate this possibility, future studies need to demonstrate whether encouraging adults and/or children to verbalize a task goal during test trials would weaken how efficiently they modulate flexible contextual representations. Despite this reservation, these findings from exploratory analyses clarify the effectiveness and limitations of reminders on learning routines in young children.

Contribution of Executive Functions to Acquisition of Routines

To specify the role of reminders in how task context is flexibly represented in young children, we need to consider its interaction with executive function development. If reminders play a substitute role for executive functions, they would logically modulate the interaction between executive functions and flexible modulation of contextual representations. First, Experiment 1 revealed that executive functions did not interact with other factors, and Experiment 2 revealed that advanced DCCS performance interacted with how task context is flexibly represented. Although these findings seemed to be inconsistent, when highlighting the performance of 4- to 6-year-olds in Experiment 1, we could confirm consistent patterns among the findings of Yanaoka and Saito (2019), Experiment 1, and Experiment 2 (see detailed information about the analyses in Supporting Information S1 and S5). In relation to our main predictions, Experiment 1 showed no interactions between the reminder effects and executive functions, and Experiment 2 showed that reminders tended to modulate the two-way interaction between the advanced DCCS and the timing factor. Furthermore, according to the results of correlation analyses (see Table 2), Experiment 1 confirmed that reminders can weaken the correlation between executive functions and overall performance in the toast-making task, whereas Experiment 2 confirmed that the correlation pattern had few differences between the

reminder and no reminder groups. These findings suggest that the support from reminders is likely to be involved in the roles of executive functions and play a role in wiping out individual differences in executive functions. However, reminders did not perfectly compensate for the role of executive functions in flexibly representing task context during repeated sequential actions. Indeed, in Experiment 2, reminders did not fully encourage children with low advanced DCCS to exhibit flexible modulation of contextual representations. However, it was confirmed that reminders helped children learn routines across the two experiments. Therefore, we propose that reminders can reduce the demands of executive functions on performing sequential actions, and such a reduction of demands might partly contribute to young children's abilities to learn routines. The involvement of one subprocess of executive functions, i.e., having access to a task goal, at a branch point is necessary, but not sufficient, to flexibly modulate contextual representations in performing repeated sequential actions. These findings are consistent with the goal circuit model (Cooper et al., 2014), in which goal representations are activated under the supervision of the executive control system to learn routines. Additionally, beyond Yanaoka and Saito (2019) revealing the association between executive functions and the acquisition of routines, our results constitute convincing evidence clarifying the contribution of executive functions to routine acquisition by the application of the goal circuit model to young children. Indeed, developmental evidence on executive functions and how young children learn routines has been accumulated independently. Thus, our findings provide the opportunity to integrate these two research areas.

Limitations and Future Directions

Despite the study's novelty, three issues are left as limitations. We will list those below with possible future directions of the current study.

First, more investigations into replicating the consistent pattern with Experiments 1 and 2 should be addressed. Although we highlighted 4- to 6-year-olds in Experiment 1 and could confirm that the data from Experiment 1 are consistent with the results of Experiment 2 (see Supplemental Information S5), the results of the analyses could not reach to be significant, possibly due to a small sample size and a biased distribution of the age groups. We recognize that a factor of age group is one of integral factors affecting the flexible modulation of contextual representations. Therefore, future studies with an appropriate sample size need to match the number of participants in each age group and replicate the finding from Experiments 1 and 2.

Second, the processes of learning routines in young children remained unclear. The current study examined whether children learned routines throughout test trials, after implementing 6 practice trials. Thus, in the case where children could acquire routines evidenced by the flexible modulation of contextual representations, we could not capture when they started to form routines. It might have been the case that they quickly acquired routines during practice trials or gradually learned them during the test trials. Furthermore, if the children did not show the flexible modulation of contextual representations, the number of practice trials might be insufficient or their inability to maintain task goals might prevent them from learning routines even if they experience the performance of sequential actions many times. Given that previous studies for adults have tackled these issues regarding learning processes (Giovannetti et al., 2007; Ruh et al., 2010), future studies with children also should address them from a developmental perspective.

Third, the precise role of executive functions in learning routines remained unclear. It was mentioned earlier that executive functions are composed of correlated and dissociable factors such as updating, inhibition, and shifting, and active maintenance of goal representations

has been shown to be the common constructs among these factors (Miyake & Friedman, 2012; Munakata et al., 2012). This is true of the executive function tasks used in this study; it has already been established that goal maintenance is associated with the red-blue task (Towse, Lewis, & Knowles, 2007), the advanced DCCS task (Chevalier & Blaye, 2009), and the visual counting span task (Marcovitch, Boseovski, Knapp, & Kane, 2010). Therefore, it is likely that goal maintenance commonly underlies executive functions captured by our measures. Moving onto our findings, we indicated that overall performance of the toast making task was related to the red-blue task (Experiments 1 and 2), the visual counting span task (Experiment 2), and the advanced DCCS task (Experiment 2). Importantly, contributions of the red-blue task to the toast making task were reduced by reminders in Experiment 1, and reminders tended to moderate the interaction between the advanced DCCS and the timing of interruptions in Experiment 2. These findings suggest that goal maintenance was associated with performing the toast making task and reminders wiped out individual differences in goal maintenance through activating a task goal before a branch point. According to the goal circuit model (Cooper et al., 2014), maintaining goal representations is required to learn routines. The finding that reminders promoted learning routines supports this view. However, the results of the present study also indicated that advanced DCCS performance has specific relations to the timing of interruptions but not to red-blue task performance and visual counting span performance. These findings suggest that goal maintenance measured by our executive function tasks is necessary, but not sufficient, to the acquisition of routines evidenced by the representational flexibility of contextual information. Rather, the ability to switch task rules, measured by the advanced DCCS task, might be another critical factor that enables children represent task context flexibly. Specifically, switching abilities could play an essential role in switching from performing common subtasks where

contextual information is not necessary to performing unique subtasks at a branch point where contextual information is necessary. The repetition of switching from common to unique subtasks would lead to the flexible modulation of contextual representations. However, a number of developmental studies using confirmatory factor analysis provided evidence regarding the structure of executive functions in young children, revealing that a factor of switching abilities cannot be separated from other factors (i.e., one-factor model: Wiebe, Espy, & Charak, 2008, or two-factor model: Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012). Thus, it is difficult to conclude that switching ability plays a unique role in flexibly representing task context. Yet, it is known that switching ability can be separated from the factors of inhibition and working memory in later childhood (e.g., Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Rose, Feldman, & Jankowski, 2011). Future studies that target older children or adults should examine how shifting ability plays a role in learning routines. As the goal circuit model regarded the executive control system as a unitary one and does not assume exact roles such as inhibition, switching, and updating, this next step should also serve to advance this pioneering, promising model.

Conclusion

The current findings provide the first evidence for the roles of executive functions in the acquisition of routines in young children. We replicated the findings of Yanaoka and Saito (2019) and further demonstrated that reminders removed demands from executive functions and encouraged 3- to 6-year-olds to flexibly represent task context according to task goals. These findings offer important insights into the relation between executive functions and how young children learn routines.

References

- Barker, J. E., & Munakata, Y. (2015). Time isn't of the essence: Activating goals rather than imposing delays improves inhibitory control in children. *Psychological Science*, *26*, 1898-1908. doi: [10.1177/0956797615604625](https://doi.org/10.1177/0956797615604625)
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278. doi: [10.1016/j.jml.2012.11.001](https://doi.org/10.1016/j.jml.2012.11.001)
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-9. <<https://CRAN.R-project.org/package=lme4>>.
- Bauer, P. J., & Hertzgaard, L. A. (1993). Increasing Steps in Recall of Events: Factors Facilitating Immediate and Long-Term Memory in 13.5- and 16.5-Month-Old Children. *Child Development*, *64*, 1204-1223. doi: [10.1111/j.1467-8624.1993.tb04196.x](https://doi.org/10.1111/j.1467-8624.1993.tb04196.x)
- Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: Evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review*, *12*, 1011-1017. doi: [10.3758/BF03206436](https://doi.org/10.3758/BF03206436)
- Botvinick, M., & Plaut, D. C. (2002). Representing task context: proposals based on a connectionist model of action. *Psychological Research*, *66*, 298-311. doi: [10.1007/s00426-002-0103-8](https://doi.org/10.1007/s00426-002-0103-8)
- Botvinick, M. M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, *111*, 395-429. doi: [10.1037/0033-295X.111.2.395](https://doi.org/10.1037/0033-295X.111.2.395)
- Botvinick M, M., & Plaut, D. C. (2006). Such stuff as habits are made on: a reply to Cooper and Shallice (2006). *Psychological Review*. *113*, 917-928. doi: [10.1037/0033-295X.113.4.917](https://doi.org/10.1037/0033-295X.113.4.917)

- Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386-404.
doi: 10.1016/0022-0965(82)90054-6
- 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.
doi: 10.1037/a0015409
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786. doi: 10.3758/BF03196772
- Cooper, R. P., Ruh, N., & Mareschal, D. (2014). The goal circuit model: A hierarchical multi route model of the acquisition and control of routine sequential action in humans. *Cognitive Science*, *38*, 244-274. doi: [10.1111/cogs.12067](https://doi.org/10.1111/cogs.12067)
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168.
doi: 10.1146/annurev-psych-113011-143750
- Fivush, R. (1984). Learning about school: the development of kindergartners' school scripts. *Child Development*, *55*, 1697-1709. doi: [10.2307/1129917](https://doi.org/10.2307/1129917)
- Freier, L., Cooper, R. P., & Mareschal, D. (2015). The planning and execution of natural sequence actions in preschool years. *Cognition*, *144*, 58-66.
doi: 10.1016/j.cognition.2015.07.005
- Freier, L., Cooper, R. P., & Mareschal, D. (2017). Preschool children's control of action outcomes. *Developmental science*, *20*, e12354. doi: 10.1111/desc.12354
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review

- using an integrative framework. *Psychological Bulletin*, *134*, 31-60. doi: 10.1037/0033-2909.134.1.31
- Giovannetti, T., Libon, D. J., Buxbaum, L. J., & Schwartz, M. F. (2002). Naturalistic action impairments in dementia. *Neuropsychologia*, *40*, 1220-1232. doi: 10.1016/S0028-3932(01)00229-9
- Giovannetti, T., Schwartz, M. F., & Buxbaum, L. J. (2007). The coffee challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology*, *29*, 690-705. doi: 10.1080/13803390600932286
- Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*, 493-498. doi: 10.1111/2041-210X.12504
- Hudson, J. A., & Fivush, R. (1991). Planning in the preschool years: The emergence of plans from general event knowledge. *Cognitive Development*, *6*, 393-415. doi: 10.1016/0885-2014(91)90046-G
- Kliegel, M., & Jäger, T. (2007). The effects of age and cue-action reminders on event-based prospective memory performance in preschoolers. *Cognitive Development*, *22*, 33-46. doi: 10.1016/j.cogdev.2006.08.003
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59-80. doi: 10.1348/026151003321164627
- Loucks, J., & Meltzoff, A. N. (2013). Goals influence memory and imitation for dynamic human action in 36-month-old children. *Scandinavian Journal of Psychology*, *54*, 41-50. doi: [10.1111/sjop.12004](https://doi.org/10.1111/sjop.12004)

- Loucks, J., Mutschler, J., & Meltzoff, A. N. (2017). Children's representation and imitation of events: How goal organization influences 3-year-old children's memory for action sequences. *Cognitive Science, 41*, 1904-1933. doi: 10.1111/cogs.12446
- Loucks, J., & Price, H. L. (2019). Memory for temporal order in action is slow developing, sensitive to deviant input, and supported by foundational cognitive processes. *Developmental Psychology, 55*, 263-273. doi: 10.1037/dev0000637
- Mahy, C. E., Mazachowsky, T. R., & Pagobo, J. R. (2018). Do verbal reminders improve preschoolers' prospective memory performance? it depends on age and individual differences. *Cognitive Development, 47*, 158-167. doi: 10.1016/j.cogdev.2018.06.004
- Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4-to 6-Year-Old Children. *Child Development, 81*, 1687-1695. doi: 10.1111/j.1467-8624.2010.01503.x
- 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. doi: 10.1177/0963721411429458
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. Davidson, G. Schwarz & D. Shapiro (Eds.), *Consciousness and Self Regulation* (Vol. 4, pp. 1-18). New York, NY: Plenum.
- Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development, 13*, 395-423. doi: 10.1080/15248372.2011.585478

- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: three key transitions. *Current Directions in Psychological Science*, *21*, 71-77.
doi: 10.1177/0963721412436807
- Perone, S., Anderson, A. J., & Youatt, E. A. (2020). Don't forget your lunch: Age and individual differences in how children perform everyday tasks. *Cognitive Development*, *54*, 100879.
doi: 10.1016/j.cogdev.2020.100879
- R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science*, *14*, 1161-1175. doi: 10.1111/j.1467-7687.2011.01068.x
- Ruh, N., Cooper, R. P., & Mareschal, D. (2010). Action selection in complex routinized behaviors. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 955-975. doi: 10.1037/a0017608
- Schwartz, M. F., Reed, E. S., Montgomery, M., Palmer, C., & Mayer, N. H. (1991). The quantitative description of action disorganization after brain damage: A case study. *Cognitive Neuropsychology*, *8*, 381-414. doi: 10.1080/02643299108253379
- Simpson, A., & Riggs, K. J. (2005). Inhibitory and working memory demands of the day-night task in children. *British Journal of Developmental Psychology*, *10*, 1-17.
doi: 10.1348/026151005X28712

- Towse, J. N., Lewis, C., & Knowles, M. (2007). When knowledge is not enough: the phenomenon of goal neglect in preschool children. *Journal of Experimental Child Psychology, 96*, 320-332. doi: 10.1016/j.jecp.2006.12.007
- Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44*, 575-587. doi: 10.1037/0012-1649.44.2.575
- Yanaoka, K., & Saito, S. (2017). Developing control over the execution of scripts: The role of maintained hierarchical goal representations. *Journal of Experimental Child Psychology, 163*, 87-106. doi: 10.1016/j.jecp.2017.06.008
- Yanaoka, K., & Saito, S. (2019). Repeated sequential action by young children: Developmental changes in representational flexibility of task context. *Developmental Psychology, 55*, 780-792. doi: 10.1037/dev0000678
- Yanaoka, K., Moriguchi, Y., & Saito, S. (2020). Cognitive and neural underpinnings of goal maintenance in young children. *Cognition, 203*, 104378.
doi: 10.1016/j.cognition.2020.104378
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297-301. doi: 10.1038/nprot.2006.46
- Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development, 78*, 16-33.
doi: 10.1111/mono.12032

Table 1. Executive Functions and Practice Trials Performance in the Toast-Making Task in the Reminder and No Reminder Groups

	Reminder	No reminder	Difference between groups
Experiment 1			
Red-blue task	12.13	13.08	$t = -1.13, df = 45.99, p = .263$
Standard DCCS	0.75	0.71	$t = 0.30, df = 45.89, p = .752$
Advanced DCCS	4.88	5.17	$t = -0.64, df = 45.61, p = .528$
Visual counting span	5.25	5.92	$t = -0.84, df = 45.95, p = .406$
Toast-making task (Practice)	5.88	5.21	$t = 3.16, df = 33.12, p = .003$
Experiment 2			
Red-blue task	13.79	13.70	$t = 0.11, df = 45.57, p = .907$
Standard DCCS	0.88	0.88	$t = 0, df = 46, p = 1$
Advanced DCCS	11.38	11.83	$t = -0.58, df = 46.00, p = .563$
Visual counting span	6.17	6.00	$t = 0.25, df = 45.09, p = .806$
Toast-making task (Practice)	5.83	5.42	$t = 2.13, df = 31.31, p = .041$

Table 2. Correlation Coefficients Between Four Executive Functions Tasks and Correct Rates in the Toast-Making Task in Each Condition

	Reminder group (Correct rate)			No reminder group (Correct rate)		
	Control	Middle	End	Control	Middle	End
Experiment 1						
Red-blue task	.35	.40*	.21	.54**	.58**	.63**
Standard DCCS	.11	-.19	-.01	.38	.16	.08
Advanced DCCS	.19	.31	.27	.27	.15	.57**
Counting span	-.01	.14	.19	.07	.06	.14
Experiment 2						
Red-blue task	.22	.17	.08	.35	-.20	.29
Standard DCCS	-.07	.36	.10	.07	-.35	.09
Advanced DCCS	.25	.01	.22	.06	-.55**	.13
Counting span	.28	.18	.46*	.40	-.05	.35

 * $p < .05$, ** $p < .01$

Table 3. Mixed-Effects Logistic Regression Results in the Best-fitting Model for Executive Functions Effects

	Estimate	SE	Z	χ^2	Pr ($>\chi^2$)
Experiment 1					
Intercept	-0.65	0.51	-1.28		
Age in months	-0.02	0.01	-1.39	1.87	.171
Red-blue task	0.19	0.04	4.57	17.55	<.001
Reminder	0.05	0.10	0.49	0.24	.625
Timing	0.26	0.19	1.42	2.02	.156
Interruption	0.79	0.19	4.12	18.52	<.001
Reminder \times Timing	0.44	0.19	2.35	5.58	.018
Experiment 2					
Intercept	0.22	0.67	0.34		
Age in months	-0.01	0.01	-0.37	0.14	.710
Red-blue task	0.09	0.04	2.06	4.18	.041
Advanced DCCS	-0.08	0.04	-1.88	3.48	.062
Visual counting span	0.14	0.04	3.36	10.26	.001
Interruption	0.71	0.18	4.05	17.53	<.001
Timing	0.50	0.18	2.75	7.66	.006
Reminder	-0.02	0.08	-0.25	0.06	.803
Timing \times Reminder	0.02	0.18	0.10	0.02	.918
Reminder \times Advanced DCCS	0.03	0.03	0.93	0.87	.351
Timing \times Advanced DCCS	0.14	0.07	2.08	4.35	.037
Reminder \times Timing \times Advanced DCCS	-0.13	0.07	-1.92	3.71	.054

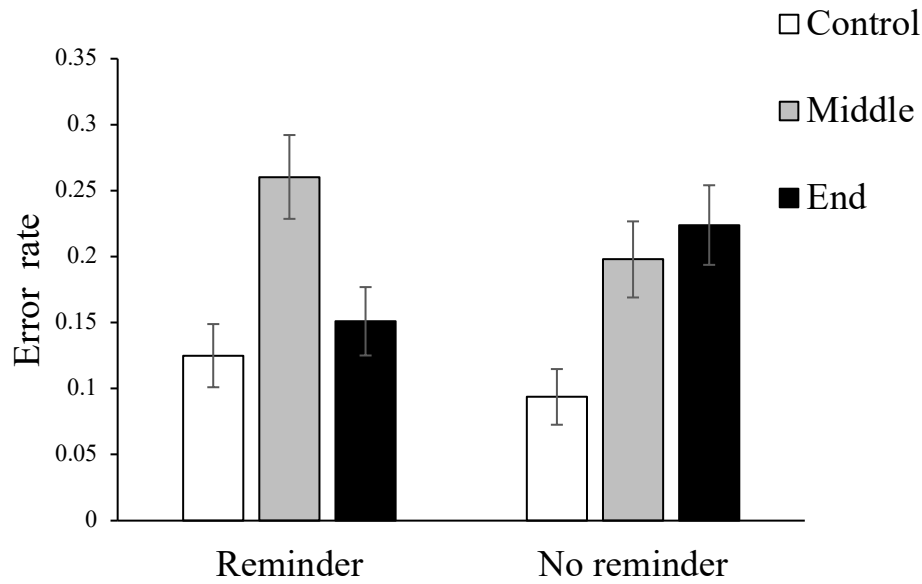


Figure 1. Mean proportions of branch point errors for the control, middle interruption, and end interruption conditions in the reminder and no reminder groups in Experiment 1.

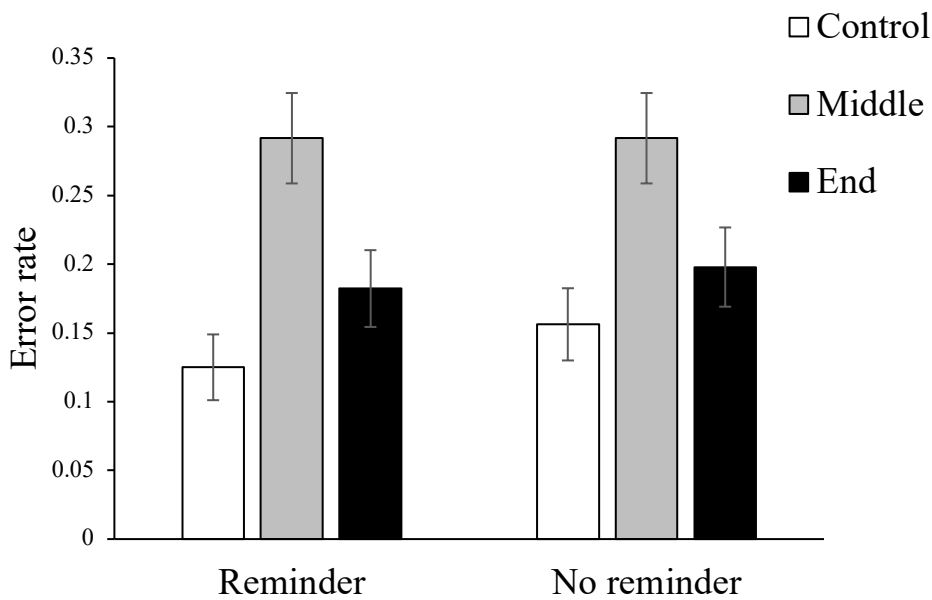


Figure 2. Mean proportions of branch point errors for the control, middle interruption, and end interruption conditions in the reminder and no reminder groups in Experiment 2.

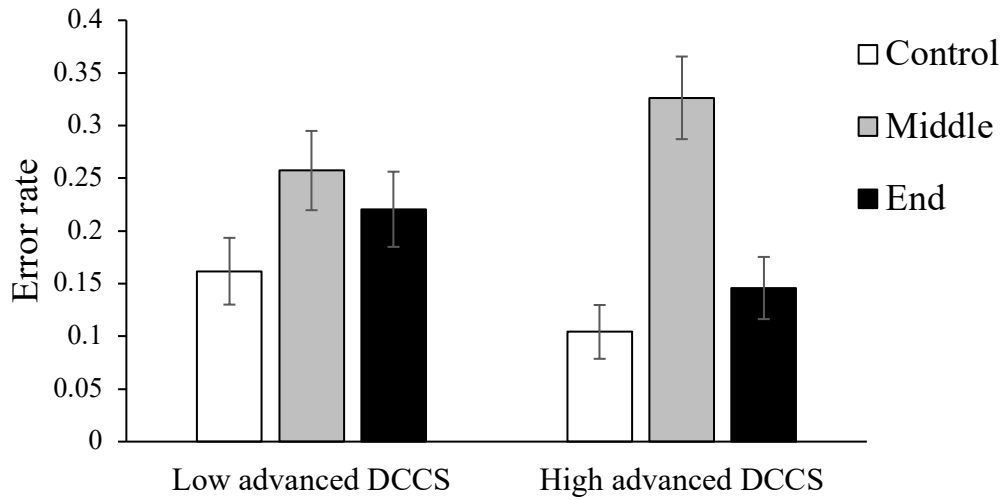


Figure 3. Mean proportions of branch point errors for the control, middle interruption, and end interruption conditions in children with high and low performance on the advanced DCCS task in Experiment 2.

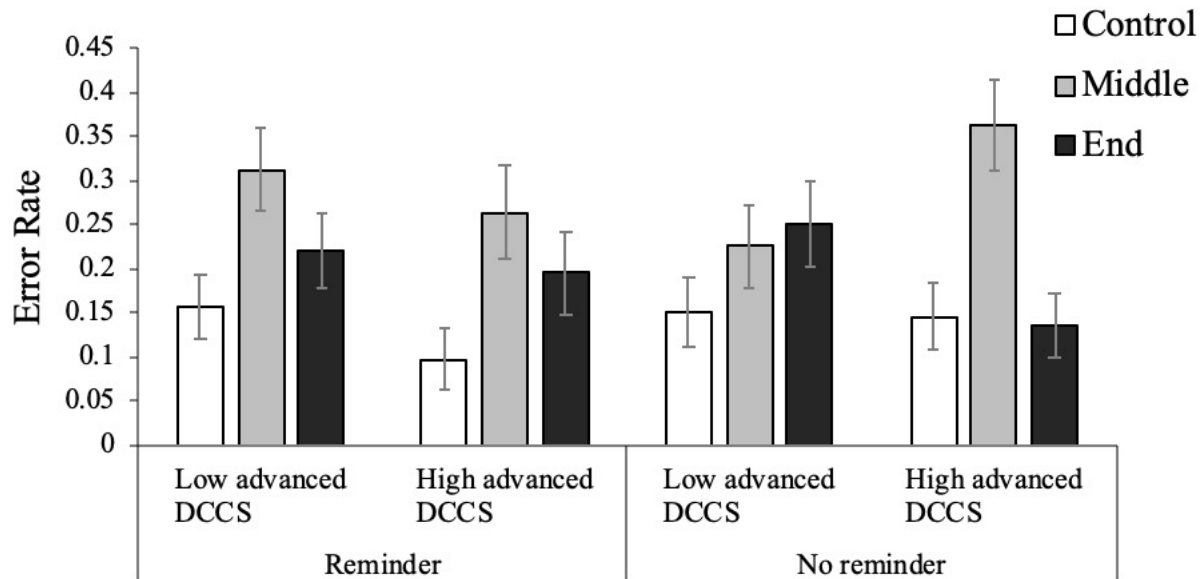


Figure 4. Mean proportions of branch point errors for the control, middle interruption, and end interruption conditions in children with high and low performance on the advanced DCCS task in the reminder and no reminder groups in Experiment 2.