DOI: 10.1111/cdev.13929

CHILD DEVELOPMENT

EMPIRICAL ARTICLE

"Plan chunking" expands 3-year-olds' ability to complete multiple-step plans

Tashauna L. Blankenship[1] | **Melissa M. Kibbe**[2]

[1]Department of Psychology, University of Massachusetts, Boston, Massachusetts, USA

[2]Department of Psychological & Brain Sciences, Boston University, Boston, Massachusetts, USA

Correspondence
Tashauna L. Blankenship, Department of Psychology, University of Massachusetts Boston, Boston, MA 02125, USA.
Email: tashauna.blankenship@umb.edu

Funding information
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: F32HD094554

Abstract
The ability to use knowledge to guide the completion of goals is a critical cognitive skill, but 3-year-olds struggle to complete goals that require multiple steps. This study asked whether 3-year-olds could benefit from "plan chunking" to complete multistep goals. Thirty-two U.S. children (range = 35.75–46.59 months; 18 girls; 9 white, 3 mixed race, 20 unknown; tested between July 2020 and April 2021) were asked to complete "treasure maps," retrieving four colored map pieces by pressing specific buttons on a "rainbow box." Children completed more of the four-step sequence correctly when the steps were presented in a way that encouraged chunking the steps into pairs. These findings suggest a potential mechanism supporting memory-guided planning abilities in early childhood.

Executing a plan is a dynamic process that requires coordination between accessing the relevant knowledge or memories required for the plan and the executive control systems that allow one to apply that knowledge to carry out the relevant steps in the required order. Take the example of making a peanut butter and jelly sandwich. To accomplish this goal, one must first retrieve the relevant knowledge (e.g., the ingredients in the sandwich, where the ingredients are located in the kitchen), and then plan the steps needed to use that knowledge (e.g., take out the ingredients, place a slice of bread on a plate, and put peanut butter on the bread). Critically, the steps of the plan must be accomplished *in a specific order* for the end goal to be reached; spreading peanut butter on a plate and then putting the bread on top does not result in a sandwich. If the plan is executed correctly, using the relevant knowledge and applying that knowledge in the correct order, the goal will be achieved, and one can eat one's peanut butter and jelly sandwich. This is an example of memory-guided planning (Blankenship & Kibbe, 2019), which involves coordination between both memory retrieval (ingredients needed, order in which the ingredients need to be assembled) and planning (how to execute the steps to put the sandwich together).

While the cognitive systems required for memory-guided planning emerge early in life, including explicit memory (Bauer et al., 1999; Hayne & Herbert, 2004; Richmond & Nelson, 2009; Rovee-Collier et al., 1980), episodic memory (Drummey & Newcombe, 2002; Hayne & Imuta, 2011; King & Markant, 2022; Scarf et al., 2013), and executive control (Espy, 1997; Rajan et al., 2014; Rajan & Bell, 2015; Willoughby et al., 2012), young children often struggle to carry out plans that require drawing on relevant memories to complete multiple steps (Atance & Jackson, 2009; Prabhakar & Ghetti, 2020; Prabhakar & Hudson, 2014; Suddendorf et al., 2011). Blankenship and Kibbe (2019) found that the primary limiting factor on young children's memory-guided planning abilities was the ability to execute a multiple-step plan in the correct sequence. In their task, 3-year-olds were asked to retrieve beads from drawers in a box in a fixed order by interacting with unique action objects on the box. To succeed, children had to recall which relevant actions to take to open the drawers (memory retrieval) and then produce the actions in the correct order (planning). They found that while 3-year-olds could retrieve the relevant memories required for the different steps of the plan, they were only able to execute the actions *in the correct order* for plans requiring up to two steps (see also Prabhakar and Hudson (2014) who observed a similar two-step limit in a semantic memory-guided planning task in

3-year-olds). By age 4, children were able to effectively carry out plans requiring four steps.

While previous work has hypothesized that the shift between 3 and 4 years is driven primarily by the development of future-oriented thinking and planning abilities (Atance & O'Neill, 2005; Gott & Lah, 2014), Blankenship and Kibbe (2019) speculated that 3-year-old children's difficulty carrying out multiple-step memory-guided plans could be due to limitations on working memory. Children may have difficulty keeping track of the relevant steps in the correct order, and/or keeping track of which steps have already been completed, because their working memory is extremely limited (to about two items reliably stored in working memory at once; Kibbe & Applin, 2022; Simmering, 2012). If true, this would suggest that 3-year-old children may have the competence for coordinating memory and planning to complete complex goals with multiple steps, but that this competence may be masked by limitations on working memory, limitations which subsequently ease with development (to about a more adult-like ~4 items by age 5; Cheng & Kibbe, 2022; Simmering, 2012).

Young children benefit when to-be-remembered information is grouped or organized (Cowan et al., 2010; Kibbe & Feigenson, 2014; Miller, 1956; Solopchuk et al., 2016). For example, when four objects are shown grouped into pairs based on spatial proximity or shared perceptual features, infants, children, and adults remember more objects in working memory compared to when these cues are absent (Feigenson & Halberda, 2004; Rosenberg & Feigenson, 2013; Jiménez et al., 2011; Kibbe & Feigenson, 2016). This process, known as "chunking," can effectively help infants and children expand their otherwise extremely limited working memory capacities.

Furthermore, 3-year-olds benefit from experience that allows them to group actions into subgoals. In an imitation task, Loucks et al. (2017) showed 3-year-olds a series of novel actions that were either grouped into separate subgoals or interleaved between different subgoals. Children were then asked to imitate the actions. Critically, prior to test, one group of children was exposed to the steps of one of the subgoals while another group of children was exposed to unrelated steps. They found children who were previously exposed to the goal structure of one of the subgoals performed better in the interleaved condition than those who were not, suggesting that 3-year-olds were able to reorganize the steps of the higher level goal into subgoals when learning to imitate the actions of an adult.

In this study, we asked whether 3-year-old children could use "plan chunking" to help them complete multiple-step memory-guided plans. We created a novel online task in which we asked children to help a monkey character complete "treasure maps" by retrieving red and blue treasure map pieces from a cartoon "rainbow treasure box." The treasure box had two perceptually distinct buttons on its top that, when "pressed," caused the box to produce different color map pieces (red or blue). During test trials, children were shown incomplete maps and were asked to retrieve four map pieces in a fixed order to complete the maps. Crucially, the incomplete map remained visible during each test trial, so that children did not have to hold the required sequence in working memory, but children were not able to see which map pieces they had already retrieved. This meant that children had to maintain their place in the planning sequence by keeping track of which actions they had already taken and which actions they still needed to take. Thus, the task required memory retrieval (which button should be pressed to retrieve a particular color), planning (when should each button be pressed to achieve the goal), and plan maintenance (keeping track of one's place in the sequence).

We manipulated whether the four steps were presented in *color-matched pairs* (e.g., retrieve red, red, blue, blue), promoting chunking, or in *color alternating pairs* (e.g., retrieve red, blue, red, blue), which does not typically promote chunking in young children (e.g., Feigenson & Halberda, 2008; Kibbe & Feigenson, 2016). We measured children's ability to retrieve the correct colors within each pair in the correct order, and the kinds of errors children made when they did not do so. We reasoned that, if 3-year-olds are given cues that would allow them to meaningfully group the steps of a multiple-step plan, they may be able to carry out more complex, multiple-step plans to completion than they otherwise could in the absence of such grouping cues.

METHOD

Participants

Thirty-two 3-year-old children (M=42.12 months, range=35.75–46.59 months; 18 girls) participated online via Zoom. Sample size was determined prior to data collection based on similar studies with this age range (e.g., Blankenship & Kibbe, 2019; see also Loucks et al., 2017). Children were recruited from the greater Boston area through recruitment events, online ads, and word of mouth. Of the 32 children tested, caregivers of 12 children responded to our demographic questionnaire, with nine identified as White (of whom one was identified as Hispanic or Latinx) and the remaining three identified as mixed race by their caregivers. All caregivers who responded to the questionnaire self-reported that they had completed either a high school (n=1) or college degree (n=11). An additional 14 children were tested but excluded from analyses due to technical errors (n=8), unclear responses (n=1), or refusal to complete the task (n=5). All study procedures were approved by the Boston University Charles River Campus Institutional Review Board (protocol #3594E). Data were collected between

July 2020 and April 2021. All participants received a $10 Amazon gift card.

Stimuli

Due to the COVID-19 pandemic, all participants completed the study via teleconference using the Zoom platform. Participants were seated approximately 25 inches from the screen to allow the experimenter to observe their pointing behaviors, and all completed the study using a laptop or desktop computer. Stimuli consisted of cartoon images of a rainbow-colored rectangular box, treasure maps, a treasure chest, and a monkey with a pirate hat (Figure 1). All stimuli were presented in Microsoft PowerPoint. The experimenter could interact with elements within the PowerPoint slides using animations and slide anchoring. Full stimuli and counterbalancing details are available at https://osf.io/dexat/?view_only=e300a8beb2654e69af22172fde519143

Procedure

Online experiment setup and calibration

The experimenter greeted the child and caregiver in the Zoom session. The experimenter then shared their screen and instructed parents on how to arrange the video panel so that the session was open in a full screen, the caregiver's camera view was hidden, and the experimenter's video was visible in the top center of the screen. All Zoom sessions were recorded for subsequent coding.

Children then completed a pointing calibration. Children were shown two stars, one red and one blue, which appeared in positions that would be relevant in the experiment, and were asked to point to each one (i.e., "Can you point to the red star?"). When children pointed to a star, the experimenter triggered an animation of the star moving side to side. This calibration allowed us to get a sense of individual children's pointing behaviors in relation to their own screens, which we used to aid in later coding of where children pointed to the screen during the experiment.

The experiment then consisted of three phases: an Encoding phase, a Task Introduction phase, and a Test phase (Figure 1).

Encoding phase

In the Encoding phase, children formed the memories they would later need to use during the Application and Test phases. A schematic of the Encoding phase is shown in Figure 1, left panel. First, the experimenter told children that the goal of the game was to help Monkey finish her treasure maps and find her treasure. Children were then shown a treasure map that Monkey needed to complete. The first map had a single red or blue (order counterbalanced) outline of a square. Children were then told that Monkey needed that color map piece to finish the map and find her treasure. The experimenter

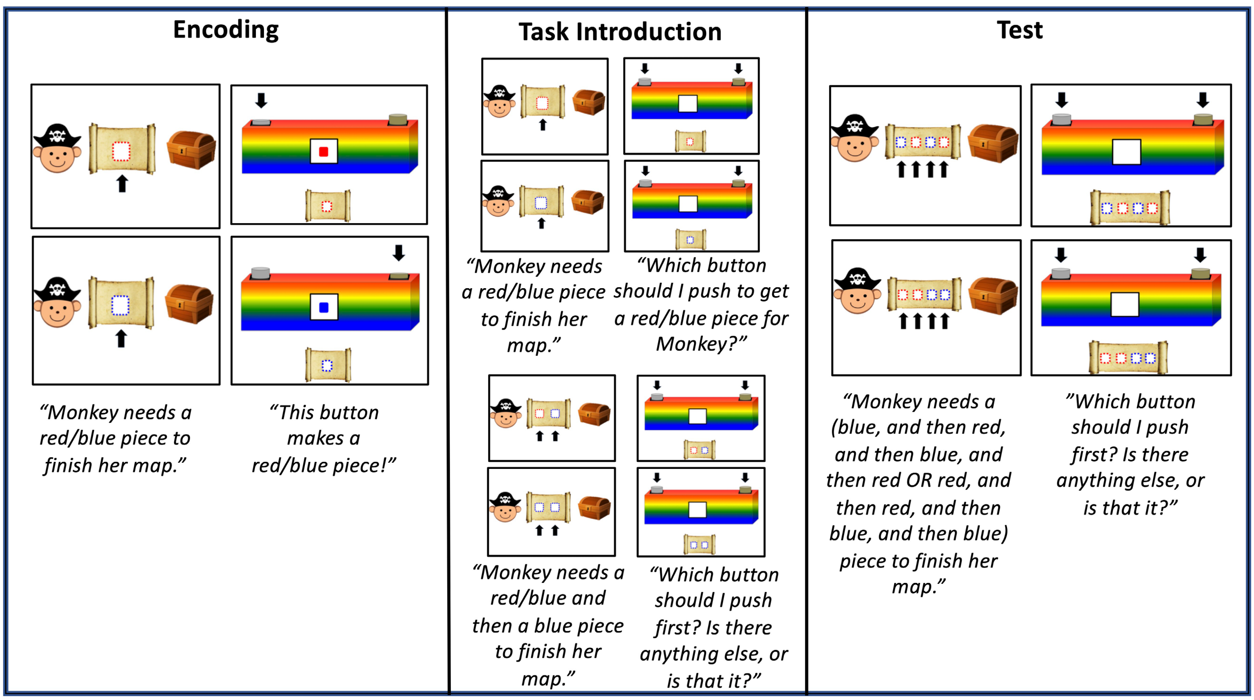

FIGURE 1 Example trials from Encoding (left panel), Task Introduction (middle panel), and Test (right panel) phases. The Test phase included both Alternating (top) and Chunking (bottom) trials.

then showed children a rainbow box with two different colored (tan and silver) buttons on its top and a square black "opening" at its bottom. Children were told that the box made red and blue treasure map pieces. The incomplete map was then displayed underneath the box and the experimenter told children, "Remember, Monkey needs a red piece to finish her treasure map."

The experimenter proceeded to demonstrate how the box produced a red map piece (Video S1). She told children "I'm going to show you which of these buttons to push to make a red treasure map piece to finish Monkey's Treasure Map! Watch this!" The experimenter then clicked the button (whether the tan or silver button produced the red piece was counterbalanced across children). The slide was animated so the button appeared to move down when "pressed" by the mouse cursor, and the red treasure map piece appeared in the "opening" of the box. The experimenter then said, "Wow! This button (hovering over button pushed) made a red piece come out! Let's finish Monkey's treasure map." Next, children saw the images of Monkey, the incomplete map, and the treasure chest, and saw an animation of the red piece move over the red outlined square on the map, after which the treasure chest opened.

This procedure was then repeated for the other button/color map piece association ("Monkey needs a blue piece to finish her treasure map and find her treasure"). At the end of the Encoding phase, children had observed which button caused the rainbow box to produce which piece, and how these pieces could be used to complete "treasure maps."

Task Introduction phase

In the Task Introduction phase, children got practice applying what they learned in the Encoding phase to complete simple plans (see schematic, Figure 1).

First, children got experience selecting buttons and completing one-step treasure maps. The experimenter first displayed another treasure map with one red or blue (counterbalanced) outlined square (see Video S2 for an example of a one-step red piece trial). Children were told, "Monkey needs a red piece to finish her treasure map and find her treasure." Children then saw the same rainbow box from the Encoding phase, with the incomplete treasure map underneath, and were asked to point to the button that made the map piece needed to complete the map: "See Monkey's treasure map. Which button should I push to make a red piece for Monkey? Can you point and show me?" The incomplete map was always visible while children made their response, just as in later Test trials (see below). After children gave their response, the experimenter clicked on the chosen button and the associated map piece appeared in the opening. If the child chose the correct button, the experimenter said "Great! Pushing this button (hovering cursor over button) made a red piece come out (hovering over map piece). Let's finish Monkey's treasure map!" The experimenter then advanced to an animated slide where the red piece moved onto the map and the treasure chest opened. If the child chose incorrectly, the experimenter said "No, this button (hovering cursor over button) made a blue piece come out (hovering over map piece), but we need a red piece to finish the map see (hovering cursor over map at bottom). Let's try again!" Children were then prompted to try again ("See Monkey's treasure map. Monkey needs a red piece to finish her treasure map and find her treasure."). Once children successfully selected the correct button, the experimenter repeated the procedure for the other button/color map piece association ("Monkey needs a blue piece to finish her treasure map and find her treasure."). Children were given a maximum of four repeat attempts per button/color map piece association, but no children exceeded that criterion. Twenty-five children responded correctly on both trials on the first try, six repeated one trial, and one repeated both trials.

Next, children were introduced to how the Test trials would proceed, but with simpler multiple step plans and corrective feedback throughout (see Video S3 for an example trial). The experimenter showed children a map with two colored outlined squares "Monkey's next treasure map needs more than one piece, see how there is a red square and then a blue square. That means that Monkey needs a red piece first [black arrow appeared under red piece] and then a blue piece [black arrow appeared under blue piece] to finish her treasure map and find her treasure." Children then saw a new slide with the rainbow box and the incomplete treasure map underneath, again visible throughout. The experimenter said, "See Monkey's Treasure Map (hovering cursor around map)? Which button should I push *first*?" Once the child responded, the experimenter clicked on the button they chose and the associated color map piece appeared in the box's opening. If children chose correctly, the experimenter said, "Great! Pushing this button (hovering cursor over button) makes a red piece come out (hovering over map piece)." The piece then vanished, and the experimenter proceeded to prompt children to retrieve the next color piece by saying, "Is there anything else, or is that it?" If the child chose incorrectly for either piece in the sequence, the color piece they selected appeared and the experimenter said, "No, this button (hovering cursor over button) made a blue/red piece (hovering cursor over map piece), but we need a red/blue piece to finish the map see (hovering cursor around map)? Let's try again!" Children were then prompted to select again. The prompt depended on where the child was in the sequence when the mistake was made. If they incorrectly guessed the first step, the experimenter said "Which button should I push first?" and if they incorrectly guessed the second step, the experimenter said "Is there anything else, or is that it?" Children also were corrected if they incorrectly said the map was completed with a single piece ("No, remember you got the red/blue piece but we still need the blue/red piece). Which button should I push? Can you show me?" or if they attempted to make additional responses beyond two pieces ("No, this button made a red/blue piece but we already have all the pieces. We finished the map!"). Once children selected the correct second action, and indicated that the map was complete, the experimenter advanced to the next slide and showed children the correct pieces completing the treasure map. Thus, children did not see the completed treasure map until they had selected both pieces.

Children completed a total of two two-step Task Introduction trials. In one of the trials, children always completed a map that needed two pieces of the same color (e.g., a red piece and another red piece). In the other trial, children always completed a map that needed two pieces of different colors (e.g., a red piece and a blue piece). Which colors appeared on the trials (red/red or blue/blue, red/blue or blue/red) and the order in which the trials were completed were counterbalanced across children. Children were given a maximum of four attempts per step to succeed, but none of the children exceeded four attempts. Across the two practice trials, 14 children needed one additional attempt, 10 children needed two additional attempts, and one child needed four additional attempts across the two actions. There was no difference in the total number of attempts needed for alternating (19 total additional attempts) or chunking (19 total additional attempts) two-step Task Introduction trials.

Test phase

In the critical Test trials, children were asked to complete four-step maps. A schematic of the Test trials is shown in Figure 1. Children were again told they would be helping Monkey finish her treasure maps, but that this time the maps would need more pieces (see Video S4 for a demo of this introduction and a four-step alternating Test trial). The experimenter then showed children a map with four outlined squares (two red and two blue). The experimenter then named each color square in the order that they needed to be retrieved (e.g., "Monkey needs a blue piece then a red piece then a blue piece and then a red piece to finish her treasure map and find her treasure"; black arrows appeared under each square). Children were then shown a slide with the rainbow box in the middle and the incomplete map at the bottom, and were asked to select a button ("Which button should I push first?"). Once the child selected a button, the experimenter clicked on the button and the associated color map piece appeared in the opening where it was visible for approximately 2 s, and then was removed by the experimenter. The experimenter then said, "Is there anything else, or is that it?" Children were not given any feedback regarding the accuracy of their responses: the experimenter continued to prompt children until children said they were done retrieving pieces. Once children indicated they were done selecting pieces, the experimenter said, "Great, I am going to hide Monkey's treasure map for now, but I will let you know if she found her treasure when we are done. Let's try another one!" and then proceeded to the next trial. Thus, children did not get any feedback about whether they had selected the correct sequence on any of the Test trials. At the end of the experiment, children were shown a slide depicting Monkey surrounded by four open treasure chests, and were given positive feedback ("Thank you for helping Monkey find her treasure! You did a great job!").

Crucially, on each trial, children did not have to hold the map sequence itself in working memory, since they could view the entire incomplete map at the bottom of the screen. We chose to make the map visible throughout in order to disentangle the working memory required to maintain one's position in one's action plan (our primary interest) from working memory required to store the map sequence. Thus, our task design meant that children did not have to expend working memory to remember the full sequence. Children also could not view the pieces they had already selected, since each piece vanished after it was retrieved from the box. Instead, children had to use working memory to keep track of where they were in the sequence as they retrieved each piece, and execute the correct series of actions accordingly.

Children completed a total of four Test trials. Half of the trials were Chunking trials, in which the colors were presented in *pairs of the same color* (Red, Red, Blue, Blue; or Blue, Blue, Red Red). The other half of trials were Alternating trials, in which the colors were presented in *pairs of alternating colors* (Red, Blue, Red, Blue; Blue, Red, Blue Red). The order of presentation of Chunking and Alternating trials was counterbalanced across participants (either Chunking, Alternating, Alternating, Chunking or Alternating, Chunking, Chunking, Alternating).

We coded which button children pointed to when prompted to choose an action. The directions of children's points were first coded by the experimenter, and were then coded by a second independent observer. Inter-rater reliability was high (Kappa = .92).

RESULTS

On each trial, children were asked to complete four-step plans consisting of pairs of either chunked (e.g., red, red, blue, blue) or alternating (e.g., red, blue, red, blue) steps. We were therefore interested in children's success at executing the steps within each pair, and how their performance differs when the steps were presented in a way that encourages chunking compared to an alternating pattern. We defined "successful" execution of each pair of steps as completing *both steps in the pair in the correct* *order*. For example, if a trial consisted of the steps *red, red, blue, blue*, and a child responded with *red, red, red, blue*, they would be coded as succeeding on the first pair and failing on the second pair. This allowed us to examine how the plan pattern (chunked versus alternating) impacted 3-year-olds' ability to execute multiple-step plans, with a particular interest in children's performance on the second pair, since previous work showed that 3-year-olds can complete two-step plans successfully (Blankenship & Kibbe, 2019; Prabhakar & Hudson, 2014). A figure of children's performance on each individual action within each pair can be found in the Supplement (Figure S1).

Three trials were not included in analyses: one child completed only the first two trials (one Alternating and one Chunking trial) and declined to respond on the remaining two trials, and one child completed the first three trials (two Alternating and one Chunking trial) but was not able to complete the final trial due to a technical issue. The analyses were conducted on 125 remaining trials.

Prior to data collection, we predicted that children in the chunking condition would be able to complete more steps in the correct order within each pair than children in the alternating condition. Therefore, we conducted confirmatory analyses on children's performance within each pair on Chunking and Alternating trials. We did not have a priori hypotheses related to the types of errors children would make when they responded incorrectly, nor whether there would be differences in error types across the two conditions, so our analyses on children's errors were exploratory. Data are available at https://osf.io/dexat/?view_only=e300a8beb2654e69af22172fde519143

We ran a generalized linear mixed model on children's mean proportion correct with Trial Type (Chunking or Alternating) and Trial Pair (1st or 2nd) as fixed effects and Participant as a random effect. This revealed no main effect of Trial Type (X^2 = .85, p = .356), a significant main effect of Trial Pair (X^2 = 9.43, p = .002) and a significant Trial Type X Trial Pair interaction (X^2 = 5.60, p = .018). These results are illustrated in Figure 2, left panel. Children performed similarly in the first trial pair in both Chunking and Alternating trials (paired samples t(31) = −.19, p = .851, 95% CI [−.38 .31]), but performed significantly better in the second trial pair when the sequence was chunked compared with the alternating sequence (paired samples t(31) = 2.06, p = .048, 95% CI [.004, .72]), suggesting that chunking helped children sustain their performance beyond two steps.

We also conducted a series of planned one-sample t-tests to compare children's mean proportion correct across the first and second pairs in each sequence in Chunking and Alternating trials to children's expected performance if they were choosing the buttons at random (chance = .5 × .5 = .25). When the steps in the sequence were chunked, children were significantly above chance for both the first (M = .44, t(31) = 2.55, p = .016) and second (M = .41, t(31) = 2.27, p = .031) pairs in the sequence. When the steps were alternating, children were significantly above chance for the first pair (M = .45, t(31) = 2.81, p = .008), but were not different from chance

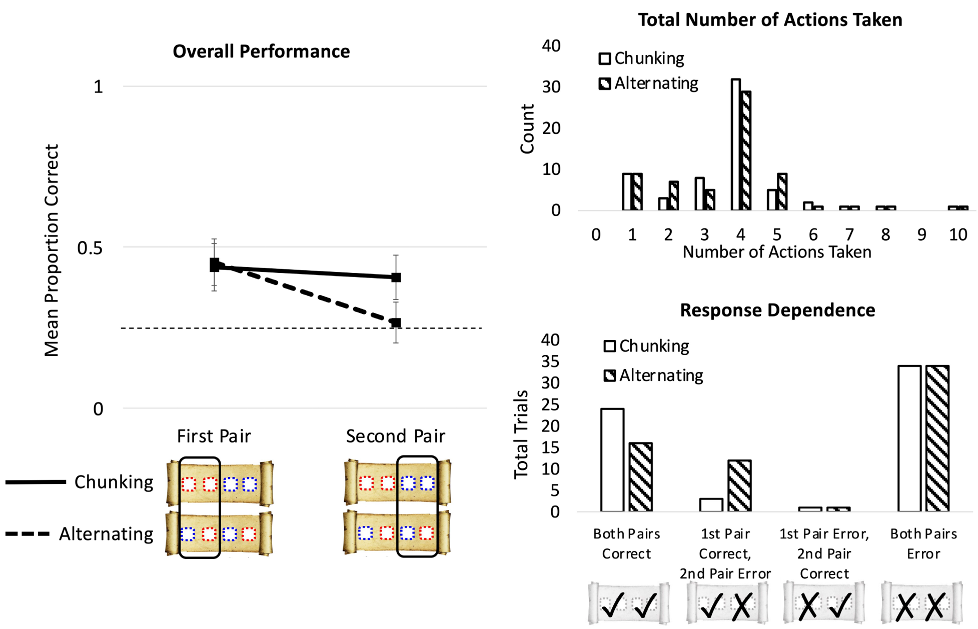

FIGURE 2 Left panel: Children's mean proportion correct for the first and second pairs in the sequence in both Chunking and Alternating trials. The legend shows two example treasure maps of the four possible maps that children were asked to complete. The horizontal dashed line represents chance-level performance. Right top panel: The total number of actions children took on each trial (regardless of whether the actions were correct or incorrect). Right bottom panel: Response dependence between the first and second pairs in the sequence in both Chunking and Alternating trials.

for the second pair in the sequence ($M = .27$, $t(31) = .25$, $p = .807$). See Figure 2, left panel.

We next took a finer-grained look at children's patterns of responses in the task, with particular interest in the kinds of errors children made. We first looked at the total number of actions children took on each trial before indicating they were done retrieving map pieces (regardless of whether those actions were correct or incorrect). Children stopped after four actions on 49% of the trials, with the remaining 51% of trials distributed between 1 and 10 actions, and there was no significant difference between the number of actions taken on Chunking and Alternating trials (paired samples $t(61) = .363$, $p = .718$), Figure 3, top right panel.

Next, we investigated dependence between children's responses in the first and second pairs in the sequence. Children's accuracy at completing the sequence in the first trial pair was significantly correlated with their accuracy at completing the sequence in the second trial pair in both the Chunking ($r = .87$, $p < .001$) and Alternating conditions ($r = .62$, $p < .001$), controlling for trial. As Figure 2, bottom right panel, illustrates, children most frequently responded correctly in both pairs or incorrectly in both pairs. To investigate differences in response dependence between Chunking and Alternating conditions, we compared the frequency with which children responded correctly on both pairs, incorrectly on both pairs, or correctly in the first pair and incorrectly in the second pair (we did not include responding incorrectly in the first pair and correctly in the second pair, as this pattern occurred too infrequently, see Figure 2, bottom right panel). We found that the frequency of these responses was different between Chunking and Alternating conditions ($X^2 = 6.99$, $p = .030$); children more frequently responded correctly in both pairs in the Chunking condition, and children more frequently responded correctly in the first pair and incorrectly in the second pair in the Alternating condition.

Finally, we examined the *kinds* of errors children made when they failed to execute the steps in each pair in the sequence in the correct order. Within each pair, we found that the errors children made could be uniquely categorized as one of three possible types of errors: children either failed to take the correct number of actions within each pair (i.e., taking 1 or 0 actions instead of 2), or children took two actions but with the wrong "pattern" (e.g., on an Alternating trial, selecting [red, red] instead of [red, blue]; on a Chunking trial, selecting [red, blue] instead of [red, red]), or children used the correct "pattern" but committed "swap errors" within the pattern (e.g., on an Alternating trial, selecting [blue, red] instead of [red, blue]; on a Chunking trial, selecting [blue, blue] instead of [red, red]). On a small subset of trials (23 of 125 trials (18%), see Figure 3), children continued to retrieve color pieces beyond four. For the vast majority of these trials (seven Chunking trials, 11 Alternating trials), children had already made one or more errors in the sequence; there were only five trials (three Chunking, two Alternating) in which children continued taking actions after they had produced the correct four-step sequence.

The distribution of these errors is shown in Figure 3. While children made slightly more "swap errors" in Alternating compared to Chunking trials, there were no significant differences between the distributions of errors across the different trial types in either the first trial pair ($X^2 = 2.92$, $p = .232$) or the second trial pair ($X^2 = .58$, $p = .746$). Combined with the above results, these results suggest that, while children made more errors in Alternating trials compared with Chunking trials in the second trial pair (see Figure 2), the types of errors children made were not significantly impacted by the presentation of the sequence. The implications of these results are discussed below.

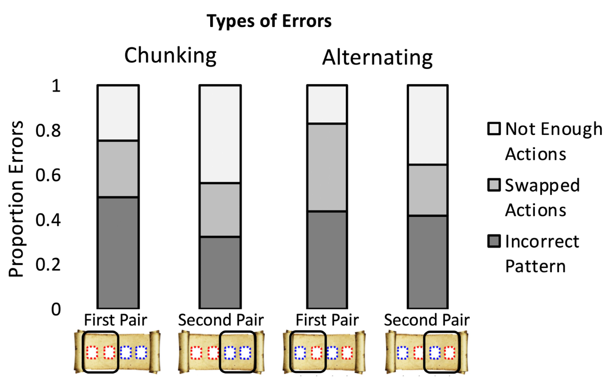

FIGURE 3 Proportion of different types of errors children made in the in the first and second pairs in the sequence in both Chunking and Alternating trials.

DISCUSSION

Previous work suggested that children younger than 4 years struggle with multi-step memory-guided planning. Although they can retrieve the relevant knowledge they need to complete a sequence of steps from memory, they have difficulty planning and executing those steps (see, e.g., Blankenship & Kibbe, 2019, 2022; Prabhakar & Hudson, 2014), which has led to the suggestion that memory-guided planning competence develops between the ages of 3 and 4 (see, e.g., Atance & O'Neill, 2005; Blankenship & Kibbe, 2019; Prabhakar & Ghetti, 2020; Russell et al., 2010). We found that, when the necessary steps of a multistep plan were presented in a way that encouraged chunking, 3-year-olds were able to complete four-step plans at above-chance levels. When chunking cues were absent, 3-year-olds were only able to reliably complete two steps of the four-step sequence, similar to previous work (Blankenship & Kibbe, 2019; Prabhakar & Hudson, 2014).

The current results suggest one potential explanation for why memory-guided planning is more difficult for younger children. Planning and executing an ordered sequence requires children to maintain multiple subgoals, retrieve the relevant long-term memories, and then engage in the appropriate actions in the correct order. This places a great deal of demand on a still-developing working memory system. Importantly, in our task, children did not have to maintain the entire sequence in working memory, since it was always visible on the bottom of the screen. The chunking cues were therefore not needed to aid in storing the sequence in working memory, but instead in keeping track of their position in the action sequence (what they have done, what they still need to do) *as they took actions to retrieve specific pieces*—at each step, children had to recall which color was needed, which button produced that color, and take the correct action to select the button. This meant that we could examine the impact of chunking cues on children's ability to execute a sequence of planned actions without placing additional burden on children's working memory by also asking them to store the sequence. Our results suggest that 3-year-olds may struggle with planning because they have a difficult time maintaining the order of subgoals within the sequence, or keeping track of where they currently are in the sequence as they complete each subgoal. Plan chunking may therefore ease working memory load by allowing children to group together subgoals, thus improving children's ability to execute multiple-step plans.

Children's pattern of errors also hints at how plan chunking may be supporting children's memory-guided planning abilities. In both Chunking and Alternating conditions, children most frequently retrieved four total pieces, suggesting that in both conditions children were able to keep track of the *how many* actions they had taken and *how many* actions they should take. However, in the Chunking condition compared with the Alternating condition, children were more likely to take all four actions in the correct order, and were less likely to make an error on the second pair of actions in the sequence after successfully completing the first pair of actions in the sequence, compared to the Alternating condition. These results suggest that children were able to use the chunking cues to help them keep track of *which* actions they had taken and *which* actions they should take. When children did lose track of which actions they should take, they made similar types of errors across both conditions, suggesting the possibility that once children lost track of which actions they took, they no longer benefitted from the chunking cues.

Our results suggest that younger children may be more competent planners than previously thought, but that their competence may be masked by working memory limitations that make it difficult for them to maintain their place in a sequence of actions. We suggest that working memory development may be a key factor (but not likely the only factor; see e.g., Gott & Lah, 2014; Prabhakar & Ghetti, 2020; Schacter et al., 2017) driving the development of memory-guided planning abilities and potentially supporting other types of sequence-based actions (including imitation; Loucks et al., 2017). This hypothesis leads to several predictions that should be investigated in future research. For example, if plan chunking supports children's working memory for their position in a planned action sequence, then reducing working memory demands further (by, e.g., showing children which actions they have already taken) should decrease the benefit of plan chunking. We could also expect to see a reduction in the benefits of plan chunking (at least for plans like those tested here) as children age and their working memory capacity increases. Furthermore, the extent to which children benefit from plan chunking may be limited by the extent to which chunking benefits children's static working memory storage for sets of items. Here, we showed children two sets of two steps presented in symmetrical chunks, but young children are less likely to benefit from chunking cues when chunks include more individuals than children's working memory capacity (e.g., three individuals; Rosenberg & Feigenson, 2013). Yet, some previous work suggests that 3-year-olds may be able to group three actions together when imitating a goal that requires all three actions (Loucks et al., 2017). Future work would investigate how plan chunking may be distinct from working memory for static arrays.

Relatedly, future work should also examine the role of attentional control on memory-guided planning across development. Attentional control has been associated with individual differences in both working (Unsworth et al., 2021) and episodic memory (Blankenship et al., 2015), and therefore is likely critical for successful memory-guided planning, which relies on both types of memory. It is possible that the ability to take advantage of plan chunking cues may be moderated by attentional control, such that children with lower attentional control could benefit more from plan chunking. Future work will aim to better understand the role of individual differences in executive functions and working memory during memory-guided planning.

The current study has some limitations. Due to the ongoing COVID-19 pandemic, the study was conducted online via Zoom. While previous work has shown that in-person and online studies produce similar results in both infants and young children (Cheng & Kibbe, 2022; Scott et al., 2017; but see also Lapidow et al., 2021, for further considerations), and while our results dovetail with similar in-person studies (Blankenship & Kibbe, 2019; Prabhakar & Hudson, 2014), our task was novel, and further work is needed to examine the extent to which plan chunking impacts memory-guided planning abilities in different contexts. We also chose to focus on 3-year-olds, since previous work has shown that this is a developmentally critical period for memory-guided planning abilities. It is currently unknown when in development children may begin to benefit from plan chunking, or the extent to which plan chunking continues to support memory-guided planning beyond age 3.

Future work would investigate these questions. Finally, it is unclear to what extent having an end-goal of "completing a treasure map" impacted children's ability to take advantage of chunking cues to complete the sequence. It is possible that framing the task as having a larger goal prompted children to attend to and track the subgoals (which made chunking easier). However, it is also possible that children may take advantage of structure when taking multiple sequential actions any time that structure is present, even if the action steps are not embedded in a larger goal structure. Further work is needed to examine these possibilities.

ACKNOWLEDGMENTS

We are grateful to the families for their participation in our research and to our research team for their assistance with participant recruitment, data collection, and coding. We would like to give a special thanks to Gracie Blades and Sara Jade Wainwright for their help with data collection.

FUNDING INFORMATION

This research was supported by grant F32HD094554 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) awarded to Tashauna L. Blankenship. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or the National Institutes of Health.

DATA AVAILABILITY STATEMENT

Data were analyzed using the built-in tools in SPSS and JASP. The analytic code necessary to reproduce the analyses presented in this paper is not publicly accessible. The materials necessary to replicate the study, and the data necessary to reproduce the analyses presented here, are publicly accessible. The analyses were not preregistered. Data and materials are available at https://osf.io/dexat/?view_only=e300a8beb2654e69af22172fde519143.

ORCID

Tashauna L. Blankenship https://orcid.org/0000-0002-6649-8732
Melissa M. Kibbe https://orcid.org/0000-0002-9088-2523

REFERENCES

Atance, C. M., & O'Neill, D. K. (2005). The emergence of episodic future thinking in humans. *Learning and Motivation*, *36*(2), 126–144. https://doi.org/10.1016/j.lmot.2005.02.003

Atance, C. M., & Jackson, L. K. (2009). The development and coherence of future-oriented behaviors during the preschool years. *Journal of Experimental Child Psychology*, *102*(4), 379–391. https://doi.org/10.1016/j.jecp.2009.01.001

Bauer, P. J. (1996). What do infants recall of their lives? Memory for specific events by one-to two-year-olds. *American Psychologist*, *51*(1), 29. https://doi.org/10.1037/0003-066X.51.1.29

Blankenship, T. L., & Kibbe, M. M. (2019). Examining the limits of memory-guided planning in 3 and 4- year olds. *Cognitive Development*, *52*, 100820. https://doi.org/10.1016/j.cogdev.2019.100820

Blankenship, T. L., & Kibbe, M. M. (2022). Two-year-olds apply episodic memories to accomplish novel goals. *Journal of Experimental Child Psychology*, *214*, 105286. https://doi.org/10.1016/j.jecp.2021.105286

Blankenship, T. L., & Bell, M. A. (2015). Frontotemporal coherence and executive functions contribute to episodic memory during middle childhood. *Developmental Neuropsychology*, *40*(7-8), 430–444. https://doi.org/10.1080/87565641.2016.1153099

Cheng, C., & Kibbe, M. M. (2022). Development of updating in working memory in 4–7-year-old children. *Developmental Psychology*, *58*, 902–912. https://doi.org/10.1037/dev0001337

Cowan, N., Hismjatullina, A., AuBuchon, A. M., Saults, J. S., Horton, N., Leadbitter, K., & Towse, J. (2010). With development, list recall includes more chunks, not just larger ones. *Developmental Psychology*, *46*(5), 1119–1131. https://doi.org/10.1037/a0020618

Drummey, A. B., & Newcombe, N. S. (2002). Developmental changes in source memory. *Developmental Science*, *5*(4), 502–513. https://doi.org/10.1111/1467-7687.00243

Espy, K. A. (1997). The shape school: Assessing executive function in preschool children. *Developmental Neuropsychology*, *13*(4), 495–499. https://doi.org/10.1080/87565649709540690

Feigenson, L., & Halberda, J. (2004). Infants chunk object arrays into sets of individuals. *Cognition*, *91*(2), 173–190.

Feigenson, L., & Halberda, J. (2008). Conceptual knowledge increases infants' memory capacity. *Proceedings of the National Academy of Sciences*, *105*(29), 9926–9930.

Gott, C., & Lah, S. (2014). Episodic future thinking in children compared to adolescents. *Child Neuropsychology*, *20*(5), 625–640. https://doi.org/10.1080/09297049.2013.840362

Gott, C., & Lah, S. (2014). Episodic future thinking in children compared to adolescents. *Child Neuropsychology*, *20*(5), 625–640. https://doi.org/10.1080/09297049.2013.840362

Hayne, H., & Herbert, J. (2004). Verbal cues facilitate memory retrieval during infancy. *Journal of Experimental Child Psychology*, *89*(2), 127–139. https://doi.org/10.1016/j.jecp.2004.06.002

Hayne, H., & Imuta, K. (2011). Episodic memory in 3-and 4-year-old children. *Developmental Psychobiology*, *53*(3), 317–322. https://doi.org/10.1002/dev.20527

Jiménez, L., Méndez, A., Pasquali, A., Abrahamse, E., & Verwey, W. (2011). Chunking by colors: Assessing discrete learning in a continuous serial reaction-time task. *Acta Psychologica*, *137*(3), 318–329. https://doi.org/10.1016/j.actpsy.2011.03.013

Kibbe, M. M., & Feigenson, L. (2014). Developmental origins of recoding and decoding in memory. *Cognitive Psychology*, *75*, 55–79. https://doi.org/10.1016/j.cogpsych.2014.08.001

Kibbe, M. M., & Applin, J. B. (2022). Tracking what went where across toddlerhood: Feature-location bound object representations in 2-to 3-year-olds' working memory. *Child Development*, *93*(6). 1713–1726 https://doi.org/10.1111/cdev.13813

Kibbe, M. M., & Feigenson, L. (2016). Infants use temporal regularities to chunk objects in memory. *Cognition*, *146*, 251–263. https://doi.org/10.1016/j.cognition.2015.09.022

King, J., & Markant, J. (2022). Selective attention to lesson-relevant contextual information promotes 3-to 5-year-old children's learning. *Developmental Science*, *25*(4), e13237. https://doi.org/10.1111/desc.13237

Lapidow, E., Tandon, T., Goddu, M., & Walker, C. M. (2021). A tale of three platforms: Investigating preschoolers' second-order inferences using in-person, Zoom, and Lookit methodologies. *Frontiers in Psychology*, *12*, 731404. https://doi.org/10.3389/fpsyg.2021.731404

Loucks, J., Mutschler, C., & 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*(7), 1904–1933. https://doi.org/10.1111/cogs.12446

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81. https://doi.org/10.1037/h0043158

Prabhakar, J., & Ghetti, S. (2020). Connecting the dots between past and future: Constraints in episodic future thinking in early childhood. *Child Development*, *91*(2), e315–e330. https://doi.org/10.1111/cdev.13212

Prabhakar, J., & Hudson, J. A. (2014). The development of future thinking: Young children's ability to construct event sequences to achieve future goals. *Journal of Experimental Child Psychology*, *127*, 95–109. https://doi.org/10.1016/j.jecp.2014.02.004

Rajan, V., & Bell, M. A. (2015). Developmental changes in fact and source recall: Contributions from executive function and brain electrical activity. *Developmental Cognitive Neuroscience*, *12*, 1–11. https://doi.org/10.1016/j.dcn.2014.10.001

Rajan, V., Cuevas, K., & Bell, M. A. (2014). The contribution of executive function to source memory development in early childhood. *Journal of Cognition and Development*, *15*(2), 304–324. https://doi.org/10.1080/15248372.2013.763809

Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science*, *12*(4), 549–556. https://doi.org/10.1111/j.1467-7687.2009.00795.x

Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science*, *16*(4), 610–621. https://doi.org/10.1111/desc.12055

Rovee-Collier, C. K., Sullivan, M. W., Enright, M., Lucas, D., & Fagen, J. W. (1980). Reactivation of infant memory. *Science*, *208*(4448), 1159–1161. https://doi.org/10.1126/science.7375924

Russell, J., Alexis, D., & Clayton, N. (2010). Episodic future thinking in 3-to 5-year-old children: The ability to think of what will be needed from a different point of view. *Cognition*, *114*(1), 56–71. https://doi.org/10.1016/j.cognition.2009.08.013

Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory in 3-and 4-year-old children. *Developmental Psychobiology*, *55*(2), 125–132. https://doi.org/10.1002/dev.21004

Schacter, D. L., Benoit, R. G., & Szpunar, K. K. (2017). Episodic future thinking: Mechanisms and functions. *Current Opinion in Behavioral Sciences*, *17*, 41–50. https://doi.org/10.1016/j.cobeha.2017.06.002

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29. https://doi.org/10.1162/OPMI_a_00001

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

Solopchuk, O., Alamia, A., Olivier, E., & Zénon, A. (2016). Chunking improves symbolic sequence processing and relies on working memory gating mechanisms. *Learning & Memory*, *23*(3), 108–112. 10.1101%2Flm.041277.115

Suddendorf, T., Nielsen, M., & Von Gehlen, R. (2011). Children's capacity to remember a novel problem and to secure its future solution. *Developmental Science*, *14*(1), 26–33. https://doi.org/10.1111/j.1467-7687.2010.00950.x

Unsworth, N., & Miller, A. L. (2021). Individual differences in the intensity and consistency of attention. *Current Directions in Psychological Science*, *30*(5), 391–400. https://doi.org/10.1177/09637214211030266

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*(2), 418–431. https://doi.org/10.1037/a0025779

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Blankenship, T. L., & Kibbe, M. M. (2023). "Plan chunking" expands 3-year-olds' ability to complete multiple-step plans. *Child Development*, *00*, 1–10. https://doi.org/10.1111/cdev.13929