Journal of Experimental Child Psychology 237 (2024) 105765

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Children's use of reasoning by exclusion to infer objects' identities in working memory

Chen Cheng [a,*], Melissa M. Kibbe [b]

[a] *Division of Social Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*
[b] *Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215, USA*

ARTICLE INFO

Article history:
Received 2 December 2022
Revised 13 July 2023
Accepted 9 August 2023

Keywords:
Working memory
Reasoning by exclusion
Logical inference
Object representations
Representational uncertainty

ABSTRACT

Reasoning by exclusion allows us to form more complete representations of our environments, "filling in" inaccessible information by ruling out known alternatives. In two experiments (Experiment 1: *N* = 34 4- to 6-year-olds; Experiment 2: *N* = 85 4- to 8-year-olds), we examined children's ability to use reasoning by exclusion to infer the identity of an unknown object and investigated the role of working memory in this ability. Children were asked to encode a set of objects that were then hidden, and after a brief retention interval children were asked to select the identity of the object hidden in one of the locations from two alternatives. On some trials, all the images were visible during encoding, so selecting the correct identity when probed required successful working memory storage and retrieval. On other trials, all but one of the images was visible during encoding, so selecting the correct identity when probed also required maintaining a representation of an *unknown* object in working memory and then using reasoning by exclusion to fill in the missing information retroactively to complete that representation by ruling out known alternatives. To investigate the working memory cost of exclusive reasoning, we manipulated the working memory demands of the task. Our results suggest that children can use reasoning by exclusion to retroactively assign an identity to an incomplete object representation at least by 4 years of age but that this ability incurs some cognitive cost, which eases

* Corresponding author.
E-mail address: chencheng@ust.hk (C. Cheng).

https://doi.org/10.1016/j.jecp.2023.105765

with development. These results provide new insights into children's representational capacities and on the foundational building blocks of fully developed exclusive reasoning.

Introduction

We live in an ever-changing world in which the amount of access we have to information varies considerably across space and time. To achieve cognitive goals in the face of incomplete information, we can use logical reasoning to make inferences about what we do not know based on what we already know. Reasoning by exclusion is the process of inferring unobserved information by eliminating known alternatives (Premack, 1995). For example, imagine you are at a buffet dinner where there are two covered dishes, both unlabeled. You lift the lid of one of the dishes and see chicken. You are then told that there is one chicken option and one vegetable option. Without lifting the lid of the other dish, you immediately make the inference that the other dish must contain the vegetable option. You were able to make this inference using knowledge of the situation at hand to exclude possibilities for the contents of the unknown dish, effectively reasoning about its likely identity. Reasoning by exclusion therefore is a powerful means of resolving representational uncertainty without needing to exert excess physical effort (such as walking over to the other dish and lifting the lid) in order to do so.

Recent work has shown that the foundations of our exclusive reasoning abilities about objects and their spatial locations are laid early in development (e.g., Aust, Range, Steurer, & Huber, 2008; Call, 2004; Feiman, Mody, & Carey, 2022; Ferrigno, Huang, & Cantlon, 2021; Hill, Collier-Baker, & Suddendorf, 2012; Leahy & Carey, 2020; O'Hara, Schwing, Federspiel, Gajdon, & Huber, 2016; Völter, Sentís, & Call, 2016). For example, around 3 years of age, children who know a reward is hidden in one of two locations, and who observe that one of the locations is empty, will search for the reward in the other location (Mody & Carey, 2016; see also Gautam et al., 2021; Grigoroglou et al., 2019; Hill et al., 2012), suggesting that they are using the observed information (the emptiness of the revealed location) to reason about the possible contents of the unobserved location. Younger children also can make inferences about the contents of a location when given verbal input about another location (e.g. "It's not in this bucket"; Austin et al., 2014; Feiman et al., 2017; Grigoroglou et al., 2019) and can infer the identity of an ambiguous object when given information rules out alternative identities (Cesana-Arlotti et al., 2018).

Successful reasoning by exclusion requires children to rely on working memory. Consider the child who is tasked with finding a toy in one of two locations (e.g., Mody & Carey, 2016). The child sees two possible locations, which are then occluded, and then sees that a toy is placed somewhere behind the occluder. The child is now faced with the maximum amount of uncertainty they will encounter in this particular task: Although the two *possible* locations of the toy are known, the exact location of the toy is unknown. The child is then given information that reduces that uncertainty: One of the locations is revealed to be empty. In this scenario, the child needs to represent the two locations, and the object (the toy) unbound (or ambiguously bound) to a particular location, and store those representations in working memory. After getting new information about the contents of the scene (the toy is not in Location A), the child can then *update* their stored representations, binding the representation of the toy to its only possible location in space (Location B). Thus, successful exclusive reasoning requires that children have sufficient working memory resources to track known information and to update uncertain or unknown representations.

In children, working memory is severely capacity limited and undergoes substantial developmental increases in capacity between infancy and late childhood (Cheng & Kibbe, 2022; Cowan, 2001, 2016; Cowan, Saults, & Clark, 2015; Kibbe, 2015; Leslie, Xu, Tremoulet, & Scholl, 1998; Pailian, Libertus, Feigenson, & Halberda, 2016; Simmering, 2012). Despite the fact that reasoning by exclusion relies on working memory, little is known about how working memory and reasoning by exclusion abilities may interact across development. What is the cognitive cost of reasoning by exclusion, and how might working memory capacity constrain reasoning by exclusion abilities in young children?

On the one hand, the ability to reason by exclusion about uncertain/unknown object locations or identities may impose greater demands on working memory than simply storing representations of a known array of items. This is because reasoning by exclusion tasks often require children to store uncertain or unknown representations in working memory and to then *update* those representations once they receive the relevant disambiguating information. Previous work has shown that when children need to update the contents of working memory (i.e., to reflect real-world changes to object locations), they make more errors and can store fewer items in memory than when they are only asked to store information in working memory without updating (Cheng & Kibbe, 2022; Pailian, Carey, Halberda, & Pepperberg, 2020). The reasoning by exclusion process itself also may incur some cognitive cost above and beyond the costs of updating in working memory because children may need to expend cognitive effort to make inferences about unknown information from known alternatives. Previous work with adults has shown that reasoning tasks (e.g., syllogistic reasoning) draw substantially on working memory resources (Gilhooly et al., 1993), and children with higher working memory capacity do better on analogical reasoning tasks (Simms et al., 2018; see also Richland et al., 2006). Inferring unknown object properties from known alternatives therefore may be more demanding and more error prone than storing known information in working memory, and as working memory load increases children's reasoning by exclusion abilities may be more limited. However, as children develop greater working memory capacity, their reasoning by exclusion abilities also should increase.

On the other hand, previous work with adults has shown that reasoning by exclusion may be deployed automatically when working memory is taxed, reducing the overall burden on working memory (see, e.g., Emrich et al., 2017; van den Berg & Ma, 2018). For example, as working memory load increases, an efficient strategy is maintaining a subset of to-be-remembered items rather than the entire set and then using the remembered subset to infer the identities of the remaining object (s). Like other working memory strategies that occur automatically when working memory is taxed (e.g., chunking, recoding; Chase & Simon, 1973; Miller, 1956; Thalmann et al., 2019), reasoning by exclusion may be fundamental to working memory. Given that young children have been shown to use a variety of efficient working memory processes (including chunking, recoding, and metacognitive awareness; see, e.g., Applin & Kibbe, 2021; Kibbe & Feigenson, 2014), reasoning by exclusion as an efficient working memory process could be operational even in young children. Under this possibility, reasoning by exclusion might not impose significant additional demands on working memory, and reasoning by exclusion should not be negatively affected by increasing working memory load (or indeed, reasoning by exclusion may even become more reliable as working memory is taxed), and we would be unlikely to see improvements in reasoning by exclusion abilities as working memory capacity increases with development.

The current experiments

The goal of the current experiments was to systematically examine the cognitive cost of reasoning by exclusion by investigating the role of working memory in reasoning by exclusion abilities across development. In two experiments, we used a modified version of an identity–location binding working memory task (Cheng & Kibbe, 2022) in which sets of images of objects were occluded and children were asked to remember the locations of specific objects. In our modified task, children completed trials in which one of the object's identities was *unknown* and children needed to infer its identity using reasoning by exclusion to eliminate known alternatives. We manipulated working memory load in both experiments and measured the impact of these manipulations on children's reasoning by exclusion abilities.

In Experiment 1, 4- to 6-year-olds viewed sets of virtual "cards" depicting images of different animals, which were then occluded. We then probed a location and asked children which image was hidden behind that occluder. Children responded by selecting one of the two images: either the correct target image or an incorrect distractor image of another animal that was hidden on that trial. In one block of trials, children performed a straightforward working memory task in which all the cards in the set were presented "face up" before becoming occluded, such that the images on their faces were visible for children to encode. To succeed on these trials, children needed to encode the location of each image, hold the location–identity bound representations in working memory during the maintenance period, and then recall the information when prompted for a response. In the other block of trials, all but one of the card faces were visible during the encoding period and the remaining card was presented "face down," such that the image on its face was *unknown* during encoding (e.g., children observed a card with a bear on it and a face-down card). After the cards were occluded, children were probed on either a face-up card or the face-down card. Success on trials in which children were probed on a face-down card required them to use reasoning by exclusion; children needed to remember the location and the identity of the card(s) that was (were) face up during encoding and then use that information to rule out known alternatives and infer the identity of the probed card (e.g., "It's not the bear, so it has to be the cat!"). We manipulated working memory load by varying the number of cards in the set.

In Experiment 2, we expanded our age range to include 4- to 8-year-old children and modified the method used in Experiment 1 to increase working memory load. Children again completed two blocks of trials (a straightforward working memory task block and a reasoning by exclusion block), but the locations of the objects swapped places during the maintenance period, requiring children to *actively update* the contents of working memory by updating their representations of the locations of the objects. On trials with a card facing down during encoding, this meant tracking and updating the location of the card with an unknown identity.

Together, the two experiments were designed to shed light on the cognitive cost of reasoning by exclusion in working memory across development. Specifically, we predicted different performance depending on the two possibilities we presented above for how reasoning by exclusion and working memory may interact. If reasoning by exclusion is an automatic strategy that children can deploy to make working memory more efficient, then once children show the ability to reason by exclusion, we would expect that (1) children should perform *better* on trials that require reasoning by exclusion compared with same-set-size trials that only tap working memory storage or updating and (2) children should perform similarly on trials with a face-down object regardless of whether targets were presented face-up or face-down during encoding. Alternatively, if reasoning by exclusion imposes cognitive costs, then we would predict that (1) children should perform worse on trials that require reasoning by exclusion compared with same-set-size trials that only tap working memory storage or updating and (2) children should perform worse on trials with a face-down object when probed on the face-down target compared with when probed on a target that was available during encoding. We also predicted that if reasoning by exclusion and working memory share a common pool of cognitive resources, we should observe poorer performance on trials that require reasoning by exclusion as working memory load increases.

Finally, although previous work has studied the emergence of children's ability to reason by exclusion about object locations (e.g., Mody & Carey, 2016; see also Gautam et al., 2021; Grigoroglou et al., 2019; Hill et al., 2012), less is known about children's ability to reason by exclusion about unknown objects' *identities* and whether children's reasoning by exclusion abilities may change with development. An additional goal of the current experiments was to examine whether children can indeed reason by exclusion about unknown objects' identities and whether the ability to use reasoning by exclusion to infer unknown identities may emerge across our age range.

Experiment 1

Method

Participants

Participants were 34 4- to 6-year-old children (mean age = 5 years 6 months, range = 4 years 1 month to 6 years 9 months; 17 girls). The size of the sample was determined prior to data collection based on a power analysis using G*Power 3.1 and was sufficient to yield 80% power to detect a medium-sized effect (d = .50) of a comparison between the unknown-object and face-up blocks on children's performance using a paired-samples t test (alpha = .05, two-tailed). Due to COVID-19 pandemic restrictions, participants were tested individually online via Zoom videoconferencing software (see Cheng & Kibbe, 2022, for a similar procedure). Children were recruited via recruitment events in the greater Boston area and participated in the current experiment after completing a separate unrelated study. All participating families received a $10 Amazon gift card for their participation. The study was approved by the Boston University Charles River Campus Institutional Review Board.

Apparatus and stimuli

Children participated remotely from a quiet room in their own home using their own device with a screen at least 10 inches (32 children used a laptop or desktop computer, 1 used an iPad, and 1 used a Chromebook). Parents were asked before the start of the experiment to ensure that siblings were not present in the room during testing. Stimuli were designed and presented in Keynote presentation software running on the experimenter's computer and were displayed to children using Zoom's screen sharing feature. Before the study began, parents were asked to hide the self-view video window and to position the experimenter's video window at the top center of the screen.

On each trial, children were asked to remember the animal characters depicted on two or three virtual "cards," which were selected from a total of 12 possible unique cards. The animal characters were images from the World of Eric Carle Mini Memory Match Game (Mudpuppy Toys) (see Fig. 1; similar stimuli were used by Cheng & Kibbe, 2022). Sessions were recorded and saved to a secure campus server for later coding. The stimuli, data, and R analysis code for Experiments 1 and 2 can be found at the Open Science Framework (https://osf.io/m6qsb/).

Design

Test trials were divided into two blocks: a *face-up* block and an *unknown-object* block. On each test trial, children viewed sets of either two cards (Set Size 2 trials) or three cards (Set Size 3 trials) presented in a horizontal row. We chose these set sizes because previous work showed that children performed worse on Set Size 3 trials compared with Set Size 2 trials where there was a working memory updating component to the task (Cheng & Kibbe, 2022). We reasoned that if reasoning by exclusion is affected by working memory load, these set sizes would be sufficient to reveal those effects without inducing floor-level performance.

In the face-up block, all the cards were presented with their images visible (see Fig. 1A). In the unknown-object block, all but one of the cards in the set had visible images and the remaining card was presented "face down," such that the image on the card was not visible (see Fig. 1B and C). The face-up block was a straightforward measure of working memory ability; children needed to encode which object was hidden in which location and to maintain that information in working memory over a brief retention interval. The unknown-object block, by contrast, included trials that also required reasoning by exclusion; the to-be-remembered set included an unknown image at encoding, and children were probed either on a previously observed image in the array or on the unknown image, the identity of which could be inferred by ruling out the known alternatives.

Children were allowed to view the array for a total duration of 1 s per face-up card, after which the cards were hidden by occluders that descended from the top of the screen. The cards remained occluded for 1 s. Children were then probed to report the identity of one of the cards; an animated hand pointed to one of the locations and two images appeared above the probed location: the image that was hidden in that location (target) and an image that was hidden elsewhere in the array (distractor). The card on the left was labeled with a 1 and the card on the right was labeled with a 2. Children were asked to select which of the two cards was hidden in that location by verbally responding "1" or "2". Whether the target appeared on the left or the right of the distractor was counterbalanced across trials.

Children completed one practice trial at the beginning of the experiment and two additional practice trials before the unknown-object block. Block order was counterbalanced across children; half the children completed the face-up block first and the other half completed the unknown-object block first. The entire task took about 10 to 15 min to complete.

Procedure

Initial practice trial. The experimenter told children, "We are going to play a hide-and-seek game." She showed children the entire set of 12 possible images and told children, "Each time, a few of my friends will appear, then they will hide behind blocks. Your job is to help me figure out who is hiding where." Children then saw two cards appear on the screen for 2 s. The experimenter then said, "Now they are

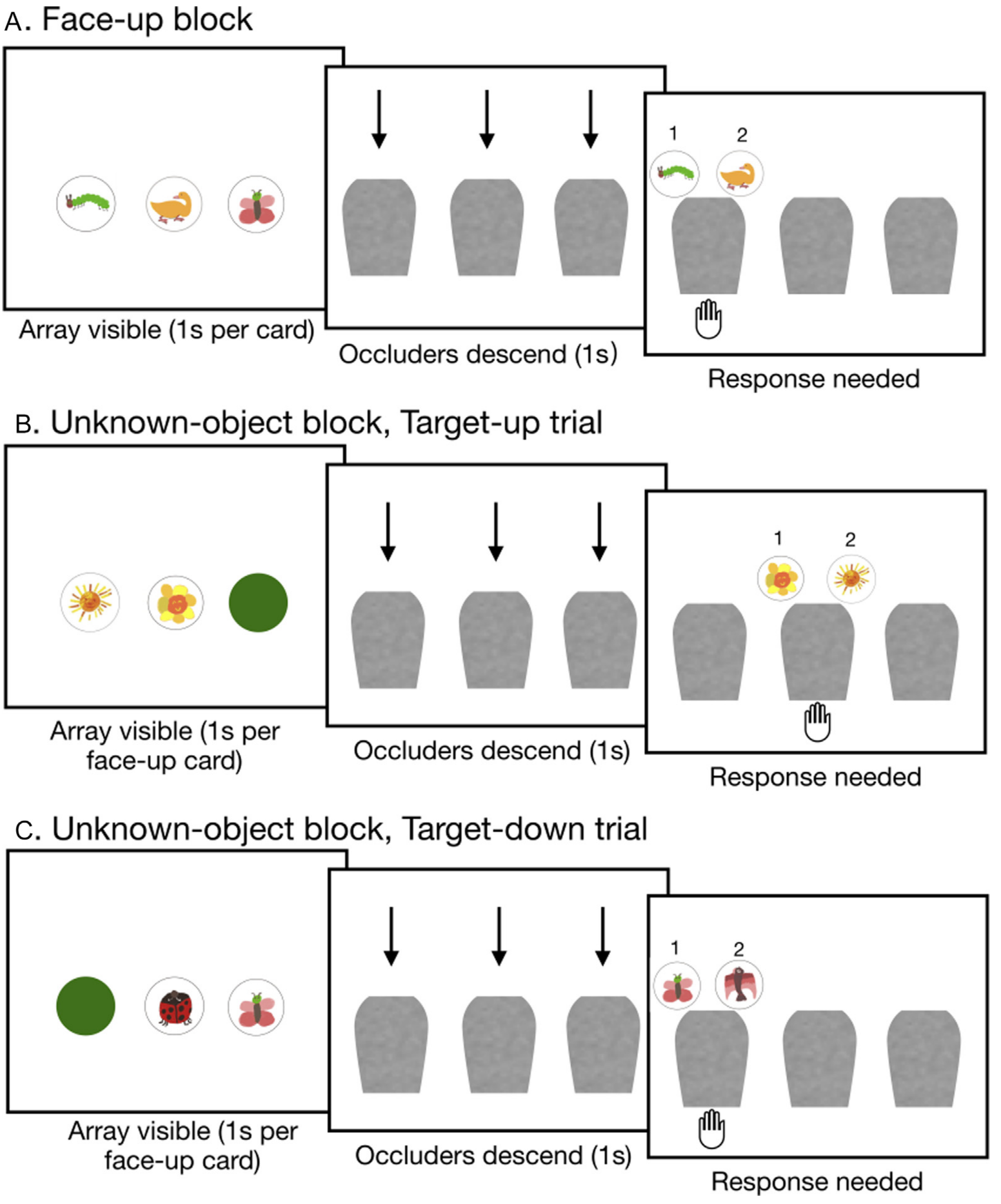

Fig. 1. Example of Set Size 3 trials in the face-up block (A) and the unknown-object block (B,C) of Experiment 1. In the unknown-object block, the target was presented face up during encoding on half the trials (target-up trial) (B) and face down during encoding on the other half (target-down trial) (C), requiring reasoning by exclusion to respond correctly. (In this example, if children remembered the known location of the butterfly, they could use that knowledge to rule out the butterfly and select the bird.) Children also completed Set Size 2 trials (not pictured). Note that the actual images used in the experiments were not granted copyrights for publication; therefore, the images displayed in Figs. 1 and 3 were recreated by the first author to be similar to the actual images.

going to hide!", after which children saw the occluders descend from the top of the screen. After 1 s, one of the locations was probed and the experimenter asked, "Which one hides here?" Children selected their response from two images: the target image that had been hidden in that location or the image that had been hidden in the other location. The experimenter provided feedback on children's responses by revealing the hidden target. If children answered correctly, the experimenter proceeded to the test trials; otherwise, the experimenter repeated the practice trial to ensure that children understood the task. Of the 34 children, 30 responded correctly on the first try. The remaining 4 children succeeded after repeating the trial once.

Face-up block. The test trials in the face-up block were similar to the initial practice trial except that after the experimenter revealed the face of the card (giving children feedback on whether they responded correctly), children were not given the opportunity to repeat trials if they responded incorrectly. Children completed eight test trials: four Set Size 2 trials, followed by four Set Size 3 trials (see Fig. 1 for an example trial). The location of the probed card was counterbalanced across trials. When probed, children were always given a choice between two objects that had been observed on that trial.

Unknown-object block. The unknown-object block began with two additional practice trials to familiarize children with the fact that the virtual cards could be presented face up or face down. First, children saw an array of two cards, both facing down. The experimenter said, "Here are my two friends. But this time only one of them will turn around," and children saw one of the cards flip over to reveal an animal image. The experimenter continued, "The other friend is going to be sneaky and he won't turn around. He will only turn around once he hides behind the block. Now they are going to hide like this!" The cards were then occluded, and children were probed on the occluded face-up card, choosing between the target and a previously unseen image (32 of the 34 children succeeded the first time; for the remaining 2 children, the experimenter repeated the trial an additional two times before the children responded correctly). This "flipping over" animation only occurred in the practice trial.

Next, the experimenter told children, "Sometimes we are going to look for the one that we did not see before." Children then saw two different cards, one facing up and one facing down, which were then occluded. Children were first probed on the occluded face-up card, choosing between a target and a previously unseen image (32 of the 34 children answered correctly, and the remaining 2 children succeeded after one repetition). Children were then probed on the occluded face-down card with the same two image choices (all the children responded correctly the first time). The experimenter then explained, "Because we already knew that the [face-up card] is right over here [the experimenter circled the location of the face-up card with her mouse cursor], so it cannot be here under this block [she circled the location of the face-down card], so it had to be the other one, and that's the [face-down card]!"

Children then completed eight unknown-object test trials: four Set Size 2 trials then four Set Size 3 trials. Each trial proceeded similarly to the face-up trials except that one of the cards was presented face down and the rest were presented face up. Within each set size, children were probed on the face-up card on half the trials (target-up trials) and were probed on the face-down card on the other half (target-down trials). On Set Size 2 target-up trials, children always chose between the target image and the unobserved image. On Set Size 3 target-up trials, children chose between the target image and the unobserved image in one trial and between the target image and the other visible image on the other trial.[1] After children responded, the experimenter revealed the card face, effectively giving children feedback on whether they responded correctly. Children completed one of two trial orders: Set Size 2 target-up, target-down, target-down, target-up followed by Set Size 3 target-down, target-up, target-up, target-down or Set Size 2 target-down, target-up, target-up, target-down followed by Set Size 3 target-up, target-down, target-down, target-up.

[1] There was no difference in children's performance on Set Size 3 target-up trials as a function of which image was the distractor, the unobserved image (27 of the 34 children correctly chose the target), or the other visible image (29 of the 34 children correctly chose the target) (Fisher's exact test, p = .75).

Results

Analyses were conducted on children's responses on each trial (correct responses were coded as 1 and incorrect responses were coded as 0).

We first confirmed that children in our task could use reasoning by exclusion to infer the identity of the unknown object in the unknown-object block. Children's mean proportion correct across trials was significantly above chance (.50) at both set sizes on target-up trials (in which they saw the image on the target card at encoding; mean proportion correct: M_{ss2} = .97, M_{ss3} = .82) and on target-down trials (in which the target image was hidden during encoding and therefore needed to be inferred; M_{ss2} = .85, M_{ss3} = .85) (all p < .001, BF_{10} > 1000), suggesting that children were successfully using reasoning by exclusion to infer the identity of the unknown object. Further analyses confirmed that even the youngest children in our sample (4-year-olds) successfully inferred the identity of the unknown object in the target-down trials (p < .001, BF_{10} = 231.54). All children also were significantly above chance in the face-up block (M_{ss2} = .80, M_{ss3} = .79, both p < .001, all BF_{10} > 1000). See Table S1 in the online supplementary material for full results.

We next asked whether children's performance varied as a function of working memory load (Set Size 2 or 3), block (face-up or unknown-object), or age (in years, continuous) using a generalized linear mixed-effects model (GLMM; R "lme4" package; Bates et al., 2015) in which we entered block, block order (face-up block first or unknown-object block first), age, and working memory load as fixed factors and participant and trial number as random factors. The best fit model included the interaction between block and block order (see Tables S2 and S3 full GLMM results in supplementary material). We observed a main effect of age, $\chi^2(1)$ = 17.12, p < .001, β = .66, SE = .16, but no main effect of working memory load, $\chi^2(1)$ = 1.82, p = .178, β = −0.33, SE = 0.25; children's overall performance in the task increased across our age range (see Fig. 2, top panel). We also observed a main effect of block, $\chi^2(1)$ = 13.71, p < .001, β = −1.41, SE = 0.38), a main effect of block order, $\chi^2(1)$ = 7.55, p = .006, β = −1.21, SE = 0.44), and an interaction between block and block order, $\chi^2(1)$ = 8.82, p = .003, β = 1.53, SE = 0.52. Children performed significantly better in the unknown-object block when they completed this block *after* the face-up block compared with children who completed the unknown-object block first, $t(32)$ = −2.201, p = .035, d = 0.78, whereas children performed similarly on the face-up block regardless of trial order, $t(32)$ = 1.08, p = .288, d = 0.38 (see Fig. 2, bottom left panel).

Finally, we examined the extent to which reasoning by exclusion affected performance in the unknown-object block. The design of the unknown-object block allowed us to directly measure the impact of reasoning by exclusion on children's ability to report the identity of the probed target. We compared children's performance on trials in which the target object had been observed before occlusion (target-up trials) with their performance on trials in which the identity of the target object needed to be inferred (target-down trials). A central aspect of the design is that, in both trial types, children did not know ahead of time which object would be probed during the response period. If reasoning by exclusion does not impose additional demands on working memory, we predicted that children should perform similarly on target-down and target-up trials (because both trial types require encoding one unknown object and either one (Set Size 2) or two (Set Size 3) known objects). If children show poorer performance on the target-down trials, it would suggest that reasoning by exclusion may be more challenging than simply recalling the presented information.

We submitted block order, age (in years, continuous), working memory load (Set Size 2 or 3), and the target's availability during encoding (target-up or target-down) as fixed factors, and participant and trial number as random factors, to a generalized linear mixed-effects model. The best fit model included interactions between target availability and working memory load (full model results are shown in Tables S4 and S5 in the supplementary material). Again, we observed a main effect of age, $\chi^2(1)$ = 10.64, p < .001, β = 0.74, SE = 0.23, with children's performance increasing with age, a main effect of block order, $\chi^2(1)$ = 5.16, p = .023, β = −0.99, SE = 0.44), with children who completed the face-up block first performing better overall, and no main effect of working memory load, $\chi^2(1)$ < .001, p > .99, β < |0.001|, SE = .51, with children performing similarly well in both the Set Size 2 and Set Size 3 blocks. Crucially, we observed a main effect of target availability, $\chi^2(1)$ = 5.08, p = .024, β = 1.83, SE = 0.81, and an interaction effect between target availability and working memory load, $\chi^2(1)$ = 4.75, p = .029, β = −2.07, SE = .95: at Set Size 2, children performed better on target-up trials

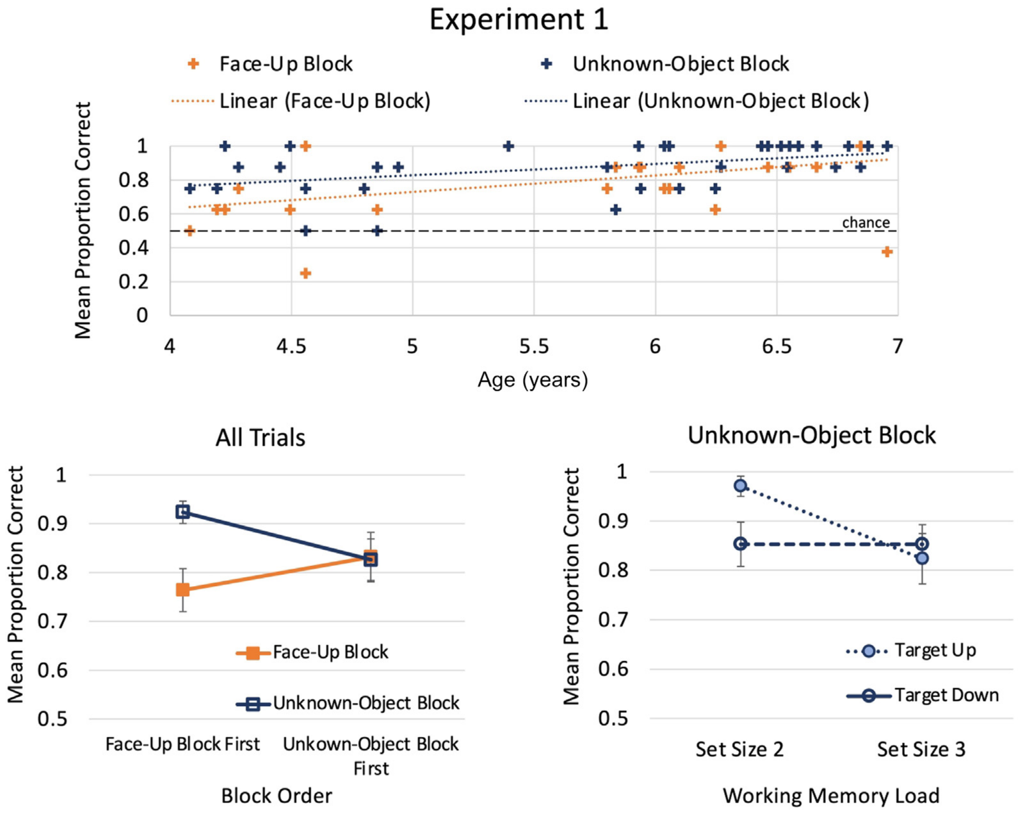

Fig. 2. Summary of the results from Experiment 1. The top panel shows individual children's mean proportion correct in the face-up and unknown-object blocks (averaged across working memory load) as a function of age. The bottom left panel shows children's mean proportion correct on the face-up and unknown-object blocks as a function of block order. The bottom right panel shows children's mean proportion correct on target-up and target-down trials in the unknown-object block as a function of working memory load. Error bars represent ±1 standard error of the mean.

(M = .97) compared with target-down trials (M = .85), paired-samples $t(33) = 2.48$, $p = .019$, $d = 0.86$; at Set Size 3, children performed similarly on target-up trials (M = .82) and target-down trials (M = .85), paired-samples $t(33) = 0.57$, $p = .57$, $d = 0.20$. On target-up trials, children performed worse at Set Size 3 compared with Set Size 2, paired-samples $t(33) = 2.96$, $p = .006$, $d = 1.03$, whereas children's performance on target-down trials was similar between Set Size 2 and Set Size 3, paired-samples $t(33) = 0$, $p = 1$, $d = 0$. These results are depicted in Fig. 2 (bottom right panel).

Discussion

Experiment 1 required 4- to 6-year-old children to encode and maintain a set of images in working memory for a brief retention interval and then to report on the identity of one of the objects by choosing from two alternative choices. Children completed two blocks of trials: a face-up block in which all the objects were visible during encoding (a straightforward working memory task) and an unknown-object block in which one of the objects was not visible during encoding, requiring children either to recall a previously observed object (face-up trials) or to infer an unknown object's identity using reasoning by exclusion (unknown-object trials).

We found that children were able to select the correct identity of the probed object at rates significantly above chance both when they were probed on an object that was available during encoding and when the identity of the object needed to be inferred, suggesting that they were able to use reasoning by exclusion to infer unknown object identities. We also found that children's overall performance was not affected by increasing memory load in either block. Because children's accuracy was high at both set sizes regardless of whether children were completing the straightforward working memory task or whether they needed to use reasoning by exclusion, we speculate that the working memory loads we chose might not have been sufficient to elicit differences in performance.

Importantly, we found that when the task required reasoning by exclusion, children performed differently than when they were simply asked to store information in working memory, in a few ways. First, we found that children who completed the face-up block *before* completing the unknown-object block performed better in the unknown-object block compared with the face-up block, whereas children who completed the unknown-object block first showed similar performance across both blocks. We speculate that children who completed the face-up block first may have benefitted from getting practice with the straightforward working memory task before needing to deploy reasoning strategies in the unknown-object block. For example, children may have benefitted from having experience with the effort involved in encoding and maintaining sets of objects (see Applin & Kibbe, 2021) and may have been more apt to make efficient use of reasoning by exclusion as a strategy following that experience.

Second, children's performance in the unknown-object block yielded important insights into how reasoning by exclusion may interact with working memory in children. Children's ability to select the correct target on target-up trials (no reasoning by exclusion required to succeed) and target-down trials (reasoning by exclusion required to succeed) was dependent on working memory load. When memory load was lower (Set Size 2 trials), children performed near perfectly on trials in which the target was presented face up at encoding but took a hit to performance when the target was presented face down. However, when memory load was higher (Set Size 3 trials), children's performance on target-up trials was similar to that on target-down trials. Furthermore, children's performance on target-up trials was affected by set size, consistent with previous work that examined working memory in children this age (e.g., Cheng & Kibbe, 2022), whereas children's performance on target-down trials was not affected by the increase in working memory load.

There may be several possibilities that could account for these results. One possibility is that these results may be a product of the fixed order in which the set sizes were presented within each block (Set Size 2 followed by Set Size 3). By completing Set Size 2 first, children had the opportunity to "practice" solving the target-down trials at a relatively lower cognitive load, and this practice allowed them to carry this effective strategy over to Set Size 3 trials with less cost.

A second possibility is that these results may suggest a strategy for how children are solving the task. Specifically, in the context of our task, the process of ruling out known alternatives to identify the unknown target may be dependent on two factors: (a) keeping track of the *location* of the unknown object and (b) keeping track of the *identities* of the known objects (but not necessarily their specific locations). At Set Size 2, where there was only one possible location for the known identity, children were nearly perfect at selecting the target identity when probed on the object that was face up at encoding but took a hit to performance when probed on the object that was face down at encoding, suggesting that the reasoning by exclusion process incurred some cognitive cost. At Set Size 3, where there were two possible locations for the known objects, children performed worse than at Set Size 2 when the objects were face up at encoding but performed similarly to Set Size 2 trials when probed on the object that was face down at encoding.

We speculate that at Set Size 3 children could have used a strategy of *quarantining* the unknown location from the known locations, and when probed on that location they used what they remembered about the known identities (unbound to spatial location) to reason by exclusion about the identity of the unknown object, incurring a similar cognitive cost as in Set Size 2 trials. However, when children were probed to recall the *specific bindings* between identities and locations of the face-up objects, they may have produced binding errors that resulted in slightly lower performance—essentially a set size effect, but only for the objects that were *visible* at encoding.

To give an example of this strategy, take the case of a child who is completing the trial depicted in Fig. 1C. On this trial, the child viewed a face-down card (on the left) and two face-up cards: a ladybug and a butterfly. Faced with this information, the child could maintain in working memory the *location* of the face-down card ("the leftmost spot") and the identities of the face-up cards ("the ladybug and the butterfly") but might not reliably encode the locations of the face-up cards (e.g., the child does not encode that the ladybug is in the middle and the butterfly is on the right). When probed on the face-down card, children can successfully and easily rule out the ladybug or the butterfly, performing similarly to Set Size 2 target-down trials. However, if probed on one of the face-up cards, children may be more likely to answer incorrectly because they might not have encoded the precise locations of the cards in working memory. Therefore, this strategy would yield a "set size effect" between Set Size 2 and Set Size 3 trials, but only the cards that were face-up at encoding.

In Experiment 2, we examined the impact of working memory load on children's ability to use reasoning by exclusion to infer unknown identities using a task that made such a quarantining strategy much more difficult to deploy. The task was similar to that in Experiment 1 except we introduced dynamic change in the objects' locations (Cheng et al., 2019; Cheng & Kibbe, 2022; Feigenson & Yamaguchi, 2009; Kibbe & Leslie, 2013; Pailian et al., 2020): After the images were occluded, the occluders swapped locations one or two times, requiring children not only to store information about which image was hidden in which location and which location contained an unknown identity but also to update that information as the locations shifted. Because previous work showed that children's working memory updating undergoes protracted development from 6 to 8 years of age (Cheng & Kibbe, 2022; Pailian et al., 2020), to capture a more complete developmental picture we expanded our age range to include 7- and 8-year-olds in addition to 4-, 5-, and 6-year-olds, and increased our sample size within each age group. Our goal for Experiment 2 therefore was to extend the results of Experiment 1, modifying the task to make greater demands on working memory and to make relying on alternative strategies more difficult.

Experiment 2

Method

Participants

A total of 85 4- to 8-year-old children (mean age = 6 years 3 months, range = 4 years 0 months to 8 years 11 months; 43 girls) participated in the experiment. We first recruited a sample of 4- to 6-year-olds with a sample size determined to yield 80% power to detect small effects (f = .20) on children's performance using linear multiple regression with block (unknown-object or face-up), working memory load (Set Size 2 two swaps, Set Size 3 one swap, or Set Size 3 two swaps), and age (in years, continuous) as factors and participant as a random factor (suggested sample size N = 59). After data collection began, we decided to expand the age range to include 7- and 8-year-olds in order to better understand any developmental trends that we might observe in the data, aiming to have roughly equal numbers of children in each age year. The final sample included 19 4-year-olds, 18 5-year-olds, 16 6-year-olds, 16 7-year-olds, and 16 8-year-olds. We excluded an additional 2 children from analysis because they did not complete the study. Due to COVID-19 pandemic restrictions, children were tested online using Zoom videoconferencing software. Children completed this study after completing a separate unrelated study and received a $10 Amazon gift card for their participation. The study was approved by the Boston University Charles River Campus Institutional Review Board.

Apparatus and stimuli

The stimuli were similar to those used in Experiment 1. Children participated in the study from their homes. In total, 82 children used a laptop or desktop computer and 3 children used a tablet.

Design

Children completed a face-up block and an unknown-object block, with block order counterbalanced across participants. The task was similar to that in Experiment 1 except that during the maintenance period the occluders swapped locations by physically moving across the screen (Fig. 3). In Set Size 2 trials, the two occluders swapped places with each other. In Set Size 3 trials, for each swap a subset of occluders was chosen pseudorandomly to swap.

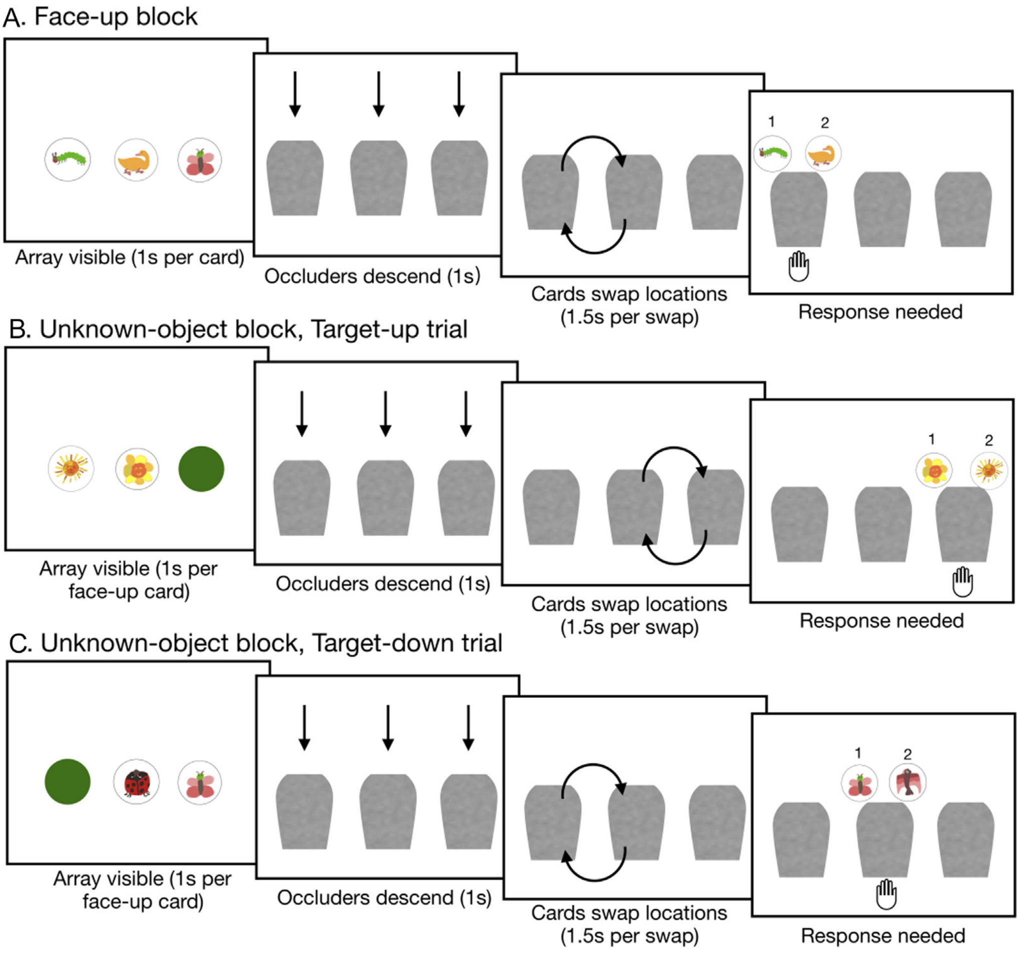

Fig. 3. Examples of Set Size 3 one swap trials in the face-up block (A) and the unknown-object blocks (B, C) of Experiment 2. In the unknown-object block, children completed trials in which the probed card was visible during encoding (target-up trial) (B) and trials in which the probed card was face down during encoding (target-down trial) (C).

We varied the working memory load within each block by varying both the set size and the number of times the cards swapped places. Working memory load was chosen based on preliminary testing of 36 4- to 6-year-olds in the laboratory (prior to the onset of COVID-19 restrictions) on a face-up version of the task using a range of updating loads. We used these preliminary data to determine three updating loads that were not too easy nor too difficult for children in this age range but that also appeared to yield some variability in performance. These working memory loads were Set Size 2 two swaps, Set Size 3 one swap, and Set Size 3 two swaps. Critically, because the objects swapped places during the maintenance period, children in Experiment 2 would have more difficulty in using a "quarantining" strategy in which they kept their representation of the unknown object separate from their representations of the known objects.

Procedure

Initial practice trials. The first practice trial was the same as in Experiment 1 (81 of the 85 children chose correctly the first time, and the remaining 4 children succeeded after repeating the trial once). In the second practice trial, children were introduced to the swap movement. The experimenter explained, "Let's try this one. Here are my two friends; they are going to hide again, but this time after they hide they are going to move, and we have to keep track of where they are hiding. Here we go!" Children saw two cards appear on the screen, which were then occluded. The occluders then swapped locations once (taking the cards with them). Children were then probed to select the identity of the card hidden behind one of the occluders from two alternatives and were given feedback on their responses. Of the 85 children, 74 succeeded the first time, and the remaining 11 children succeeded after one repetition of the trial.

Face-up block. Test trials in the face-up block were similar to the second practice trial except that after the experimenter revealed the face of the probed card, children did not have the opportunity to repeat trials if they had answered incorrectly. There were 12 face-up trials in total. Children first completed four Set Size 2 two swaps trials, followed by four Set Size 3 one swap trials and finally four Set Size 3 two swaps trials. The location of the probed image and whether the target image appeared on the left or the right of the distractor image were counterbalanced across trials. Fig. 3A shows an example Set Size 2 two swaps face-up trial.

Unknown-object block. Children first completed two unknown-object practice trials, which proceeded similarly to the unknown-object practice trials in Experiment 1 except that the objects swapped locations once following occlusion. Feedback was provided and the trial was repeated if children's responses were incorrect. In the first unknown-object practice trial, children were probed to select the identity of the face-up card (72 of the 85 children succeeded the first time, and the remaining 13 children succeeded after one repetition). In the second practice trial, the experimenter first probed the location where the face-up card was hidden (78 of the 85 children succeeded the first time, and the remaining 7 children responded correctly after the second presentation) and then the location of the face-down card with the same two alternatives (all children correctly chose the previously unseen image the first time).

The test trials proceeded similarly to the practice trials except that after children were given feedback they were not allowed to repeat a trial if they responded incorrectly. As in the face-up block, the location of the probed card and whether the target image appeared on the right or the left of the distractor image were counterbalanced across trials. Children completed 12 trials—four trials each of Set Size 2 two swaps, Set Size 3 one swap, and Set Size 3 two swaps—presented in blocks, with block order fixed across participants (children always completed Set Size 2 two swaps first, followed by Set Size 3 one swap and then Set Size 3 two swaps). Within each block, two trials were target-up trials (see Fig. 3B for an example) and two were target-down trials (see Fig. 3C for an example). In Set Size 3 target-up trials, children chose between the target image and the unobserved image on half the trials and between the target image and the other visible image on the other half.[2] The probed card was always involved in at least one swap movement, but in Set Size 3 two-swap trials, the probed card sometimes swapped during the first swap movement and sometimes swapped during the second swap movement. The order of the trials within each block was counterbalanced following the ABBA or BAAB pattern.

Results

We first confirmed that children could use reasoning by exclusion to infer the identity of the unknown object in a task that made more demands on working memory. We found that, overall, children performed significantly above chance in the unknown-object block on both target-up and target-down trials, suggesting successful reasoning by exclusion (all $p < .001$, $BF_{10} > 1000$; similar results were observed for the face-up block). However, when we binned children into groups based on their age in years and compared each group's performance on target-down trials with chance, we found

[2] There was no difference in children's performance on Set Size 3 target-up trials as a function of which image was the distractor. In one swap trials, 65 of the 85 children correctly chose the target when paired with the unobserved image and 55 of the 85 children correctly chose the target when paired with the visible image (Fisher's exact test, $p = .13$); in two swaps trials, 57 of the 85 children correctly chose the target when paired with the unobserved image and 66 of the 85 children correctly chose the target when paired with the visible image (Fisher's exact test, $p = .17$).

that only children aged 6 years and older showed above-chance performance (all $p < .001$, $BF_{10} > 69$), whereas 4- and 5-year-olds' performance on target-down trials was not different from chance (see Table S6 in supplementary material).

Next, we ran a GLMM with block (face-up or unknown-object), block order (face-up block first or unknown-object block first), age (in years, continuous), and working memory load (Set Size 2 two swaps, Set Size 3 one swap, or Set Size 3 two swaps) entered as fixed factors and participant and trial number entered as random factors. The best fit model included an interaction between age and block (see Tables S7 and S8 in supplementary material for full GLMM results). We observed a main effect of age, $\chi^2(1) = 70.45$, $p < .001$, $\beta = 0.60$, SE = 0.07, a main effect of block, $\chi^2(1) = 6.14$, $p = .013$, $\beta = 1.19$, $SE = 0.48$, and an interaction between block and age, $\chi^2(1) = 5.23$, $p = .022$, $\beta = -0.18$, $SE = 0.078$), and no other main effects or interactions were significant; whereas children performed better overall in the face-up block compared with the unknown-object block (regardless of working memory load), performance in both blocks converged with age (see Fig. 4, top panel).

Finally, we analyzed children's performance in the crucial unknown-object block. We entered working memory load (Set Size 2 two swaps, Set Size 3 one swap, or Set Size 3 two swaps), target availability at encoding (target-up trial or target-down trial), age (in years, continuous), and order (face-up block first or unknown-object block first) as fixed factors and participant and trial number as random factors into a GLMM. The best fit model did not include any interactions between variables (see Tables S9 and S10 in supplementary material for full GLMM results). We again observed a main effect of age, $\chi^2(1) = 69.24$, $p < .001$, $\beta = 0.60$, $SE = 0.072$, but also a main effect of working memory load, $\chi^2(2) = 8.16$, $p = .017$, with children performing better overall on Set Size 2 two swaps trials compared with both Set Size 3 trials. Crucially, we observed a main effect of target availability, $\chi^2(1) = 4.10$, $p = .043$, $\beta = 0.31$, $SE = 0.15$; children's performance was higher on trials where the probed card was visible during encoding compared with trials where the probed card was not visible during the encoding period (Fig. 4, bottom panel), and this pattern was consistent across our age range. Indeed, inspection of Fig. 4 shows that children's performance on target-up trials was consistently higher than that on target-down trials across working memory loads.

Discussion

In Experiment 2, we increased working memory load by introducing dynamic change in the locations of the objects during the retention interval. In addition to making the working memory task more demanding, the dynamic motion of the objects during the maintenance period made it more difficult for children to use a strategy of quarantining the unknown object from the known objects; children needed to bind both the *known* and *unknown* identities of objects to specific locations and then track and update those locations to successfully identify the target either by retrieving it from memory (face-up block; unknown-object block target-up trials) or by inferring its identity by ruling out known alternatives (target-down trials). We also expanded our age range to include children aged 4 to 8 years.

We found that, overall, children performed better in the face-up block than in the unknown-object block, suggesting that maintaining an array of known identities in working memory is less error prone than maintaining an array with an unknown identity. It is important to note that children needed to remember *more* identities at the outset in the face-up block compared with the unknown-object block. Children's poorer performance in the unknown-object block suggests that the cost of maintaining/inferring an unknown identity in working memory is higher than the cost of maintaining an additional *known* identity in working memory. We also failed to replicate the block order effect from Experiment 1; children in Experiment 2 no longer benefitted from completing the face-up block first, perhaps due to the higher working memory loads in Experiment 2.

The expansion of our age range in Experiment 2 revealed a developmental trend: Performance in the face-up and unknown-object blocks began to converge around 6 and 7 years of age, with older children achieving similar performance across the two blocks. This result suggests that the costs associated with maintaining/inferring an unknown object's identity ease with development, perhaps as children develop further executive control abilities (see General Discussion for further discussion).

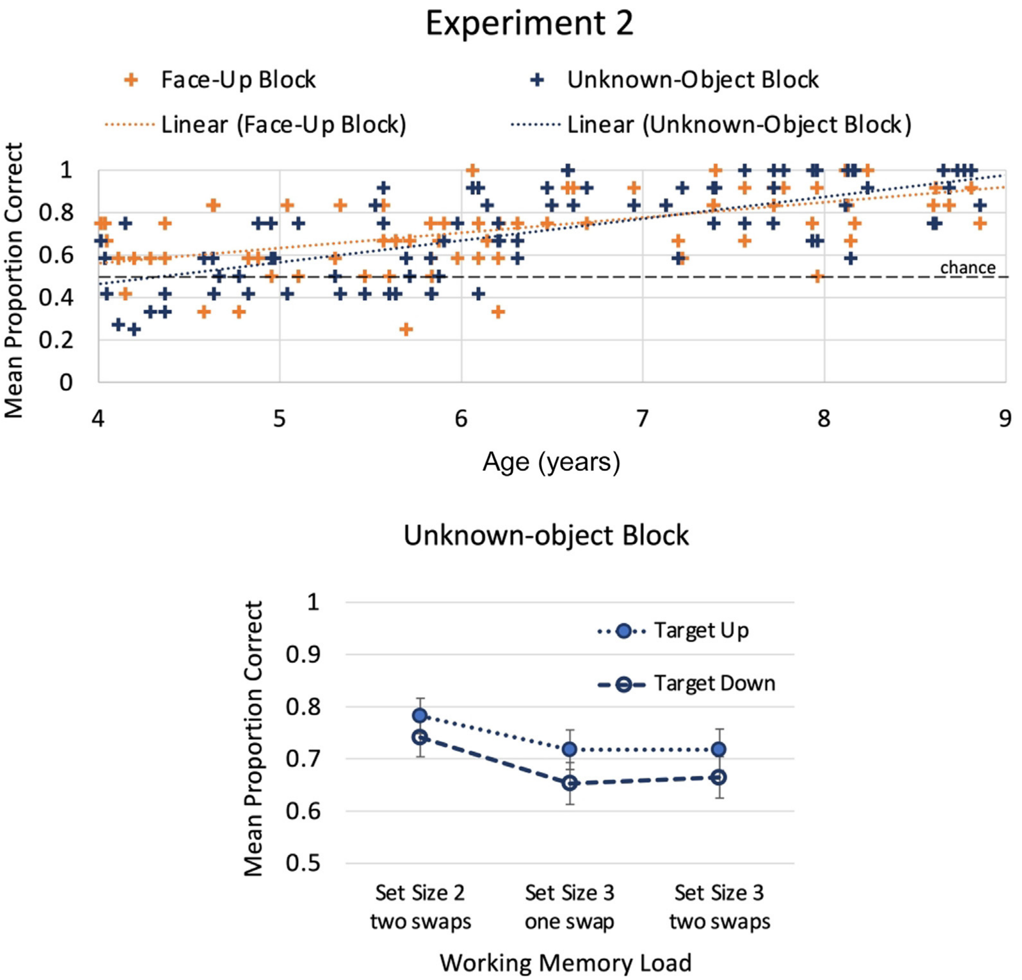

Fig. 4. Summary of the results from Experiment 2. The top panel shows individual children's mean proportion correct in the face-up and unknown-object blocks (averaged across working memory load) as a function of age. The bottom panel shows children's mean proportion correct in the target-up and target-down trials in the unknown-object block as a function of working memory load. Error bars represent ±1 standard error of the mean.

We also found that in the unknown-object block, when the dynamic change in the objects' locations made using a quarantining strategy difficult, children performed better when they were probed on a known identity target than when they were probed on a target whose identity they needed to infer, and this difference in performance was consistent across working memory loads and age. We discuss the broader implications of these results in the General Discussion.

General discussion

Reasoning by exclusion often necessitates reliance on limited working memory; to successfully rule out known alternatives, children must hold those alternatives in mind along with a representation of the unknown object or location. In two experiments, we examined the cognitive cost of reasoning by exclusion in working memory by investigating the impact of working memory load on reasoning by exclusion abilities during a period of childhood in which working memory is undergoing considerable development. Our results suggest three main takeaways about the cognitive cost of reasoning by exclusion in working memory across development.

First, we found that reasoning by exclusion to infer unknown identities by ruling out stored alternatives in working memory is not cost free for children. Whereas previous work suggested that adults may use reasoning by exclusion to make working memory *more* efficient and *less* effortful (e.g., Emrich et al., 2017; van den Berg & Ma, 2018), children appeared to take a ''hit'' to performance when they needed to reason by exclusion compared with when they were asked to simply report the identity of a stored item in working memory. And children's overall worse performance on unknown-object trials in which the target object's identity needed to be inferred compared with when they were probed on a known object suggested that, even if children remember the identity (or identities) of the known alternative(s), they do not necessarily use those representations reliably to make the correct inference. Our results suggest that the reasoning by exclusion may incur a one-time cognitive cost; working memory load appeared to affect performance on both target-up and target-down trials similarly (particularly evident in Experiment 2 where a strategy of quarantining the unknown object from the known objects was significantly more difficult).

Second, we also found that children's performance on trials that required reasoning by exclusion to infer an unknown identity appeared to improve with development. In Experiment 2, where we extended our age range to include children aged 4 to 8 years, we found that around 6 and 7 years children's performance on the straightforward working memory task (the face-up block) and their performance on the reasoning by exclusion task (the unknown-object block) began to converge. Indeed, whereas the development of working memory capacity itself may be one limiting factor in reasoning by exclusion across development, these results suggest that the sources of children's limitations on reasoning by exclusion may be more nuanced.

One possible source of the cognitive cost of reasoning by exclusion may come from the way children distribute limited working memory resources across the items in the array. Whereas younger children may have distributed their working memory resources evenly across all the items in the display, older children may be able to use a strategy of tracking only the objects for which they observed identities, which would require children to proactively plan which item(s) they should attend to (Chevalier et al., 2014) while inhibiting and actively not tracking the unknown object (Zelazo et al., 2003) and then using reasoning by exclusion only when probed on the unknown (and untracked) object. Previous work suggests that children shift from more reactive to more proactive working memory recall strategies (Chevalier et al., 2014) and show improved metacognitive awareness of working memory limitations at 5 to 7 years of age (Applin & Kibbe, 2021), consistent with a possible strategy shift around 6 and 7 years of age.

In fact, children's pattern of performance in Experiment 1 shows some hints that such strategies may be deployed even by younger children under the right circumstances. In Experiment 1 (where children did not need to dynamically update their representations in working memory), children performed *better* in the unknown-object block when they completed this block after completing the more straightforward working memory task block (the face-up block). Given that both the face-up and unknown-object blocks required children to maintain sets of static objects, children may more easily have been able to recognize the similarity between the tasks; thus, children may have been able to use the static working memory task to get their feet under them, allowing them to subsequently realize that they could deploy a more efficient strategy for reasoning by exclusion in the unknown-object block. In addition, the interaction between working memory load and target availability in Experiment 1 suggested that children may indeed have used a strategy of inhibiting or quarantining the unknown object when the demands of the task made it relatively easy to do so (as when children only needed to maintain static arrays). Further work is needed to understand the strategies that children use when reasoning by exclusion under varying task demands, and how those strategies may shift with the development of working memory or cognitive control.

Another possible source of cognitive effort in our task is related to how children represent the face-down card: What do children represent about an object that has an unknown identity? In our study, younger children's ability to track and update working memory for known object identities may be less robust under conditions with unknown identities (see Kovács et al., 2021, for similar evidence in an uncertain belief updating task), particularly in Experiment 2 where the location of the unknown object needs to be dynamically tracked along with other known objects. On unknown-object trials, children may have deployed object file representations for each object (known and unknown) in order to track all these objects as they moved through space. Younger children's ignorance of the identity of one of the object file representations may have interfered with maintaining and tracking identity-bound object representations of the other object(s) as they moved (see Ma & Flombaum, 2013, for evidence of an unknown number of targets affecting multiple objects tracking in adults). Children's difficulty in dynamically tracking the locations of object files with unknown identities in working memory may be resolved by the maturation of attentional or representational resources with age, as evidenced by the age-related improvement in children's performance in the unknown-object block of Experiment 2. More work is needed to understand how representing object files with unknown identities affects object tracking, storage, and updating in working memory, and how this ability might change with development.

To summarize, we observed a developmental increase in children's ability to infer an unknown identity using reasoning by exclusion. Yet, there are many potential sources of the observed developmental increase, including but not limited to children's ability to remember, track, and update known and unknown object identities, improved flexibility in planning, and development in the ability to reason by exclusion. Future work is needed to identify the role of a range of cognitive factors in the developmental improvement we observed.

Third, our study provides an additional data point on early reasoning by exclusion abilities in children. Previous work that investigated the emergence of reasoning by exclusion abilities in development suggested that 3-year-old children can use reasoning by exclusion to infer the location of a hidden object (e.g., Grigoroglou et al., 2019; Mody & Carey, 2016), and infants can resolve ambiguity in the identity of a partially visible object prospectively (Cesana-Arlotti et al., 2018, 2020) when all possible alternatives are already known at the outset. Our results extend this previous work by showing that, by at least 4 years (the youngest age tested here), children can exclude known alternatives to infer the identity of an unknown object (albeit at a cost) and that these abilities interact with working memory load and undergo developmental change across early to middle childhood. Future work should examine the emergence of this reasoning by exclusion about unknown identities in younger children. Future work also should examine the development of reasoning by exclusion to make inferences about locations or other aspects of objects. Given that previous work has shown that working memory for object locations may be less effortful and may follow a different developmental trajectory than working memory for identity–location bound objects (Kibbe, 2015; Kibbe & Applin, 2022), reasoning by exclusion over these different types of information in working memory also may undergo different developmental trajectories.

Could children have been using strategies other than reasoning by exclusion to do the task? Whenever children were probed an unknown object, they were presented with two alternative choices: an image that had appeared on that trial and an image that had not appeared on that trial. Children could potentially have succeeded at selecting the correct image using a strategy of choosing novelty (or avoiding familiarity) rather than reasoning by exclusion. However, although possible, there are reasons to think that children were not doing this. First, if children were using such a strategy, their performance on target-up trials and target-down trials should be similar given that the ability to choose novelty (or avoid familiarity) would be entirely dependent on the ability to remember the other targets that appeared on that trial, regardless of working memory load, and should be relatively easy to do if one can remember the other objects. However, we observed poorer performance when children were probed on the unknown object compared with the known object, inconsistent with such a strategy. Furthermore, if children were choosing novelty or avoiding familiarity, the ability to do so should become more difficult across trials; because objects were drawn from a finite set, the likelihood of proactive interference from previous trials increased across trials (Hamilton et al., 2022) and objects became less novel overall. Thus, if children were using a "choose novelty" strategy, we might expect children to perform worse when the unknown-object block was presented second. However, we found that in Experiment 1 children performed *better* in the unknown-object block after completing the face-up block first, and we did not observe any reliable order effects in Experiment 2. Although it is still possible that children may have used a mix of strategies on our task, we think it is unlikely that they were using purely non-exclusion-based strategies. Future work would investigate the kinds of strategies children may use when approaching a cognitively effortful task involving reasoning or inference.

Conclusions

The ability to make inferences about unknown or uncertain information by ruling out known alternatives is a powerful means of resolving representational uncertainty in our environments, and this ability makes working memory operate more efficiently during adulthood. Previous work suggested that the ability to use reasoning by exclusion to infer object locations emerges early in development. Our results suggest that the ability to reason by exclusion about object identities may operate at a cost early in development. Although even young children can retroactively assign an identity to an unknown object by ruling out known identities stored in working memory, this ability was more error-prone than recalling information stored in working memory. With age, children's ability to reason by exclusion in working memory converged with their working memory recall abilities.

Data availability

I have shared the link to my data in OSF.

Acknowledgments

This research was supported by a National Science Foundation grant (BCS 1844155) awarded to M. M.K. We thank the caregivers and their children who participated in our study and thank Shiba Esfand and Jiaqi Zhao for their assistance in data collection. A portion of the data from Experiment 2 was presented at the annual meeting of the Cognitive Sciences Society in July 2021.

Data availability

Stimuli and data for Experiments 1 and 2 can be found at the Open Science Framework (https://osf.io/m6qsb/?view_only=45d3934fbff2443da25e1c4f3fb28c8f).

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jecp.2023.105765.

References

Applin, J. B., & Kibbe, M. M. (2021). Young children monitor the fidelity of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(5), 808–819. https://doi.org/10.1037/xlm0000971.

Aust, U., Range, F., Steurer, M., & Huber, L. (2008). Inferential reasoning by exclusion in pigeons, dogs, and humans. *Animal Cognition, 11*(4), 587–597. https://doi.org/10.1007/s10071-008-0149-0.

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology, 50*(8), 2061–2070. https://doi.org/10.1037/a0037179.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Call, J. (2004). Inferences about the location of food in the great apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, and Pongo pygmaeus). *Journal of Comparative Psychology, 118*(2), 232–241. https://doi.org/10.1037/0735-7036.118.2.232.

Cesana-Arlotti, N., Kovács, Á. M., & Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications, 11*(1). https://doi.org/10.1038/s41467-020-19734-5 5999.

Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science, 359*(6381), 1263–1266. https://doi.org/10.1126/science.aao3539.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2.

Cheng, C., Káldy, Z., & Blaser, E. (2019). Two-year-olds succeed at MIT: Multiple identity tracking in 20- and 25-month-old infants. *Journal of Experimental Child Psychology, 187*. https://doi.org/10.1016/j.jecp.2019.06.002 104649.

Cheng, C., & Kibbe, M. (2022). Development of updating in working memory in 4-7-year-old children. *Developmental Psychology, 58*(5), 902–912. https://doi.org/10.1037/dev0001337.

Chevalier, N., James, T. D., Wiebe, S. A., Nelson, J. M., & Espy, K. A. (2014). Contribution of reactive and proactive control to children's working memory performance: Insight from item recall durations in response sequence planning. *Developmental Psychology, 50*(7), 1999–2008. https://doi.org/10.1037/a0036644.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87–114. https://doi.org/10.1017/s0140525x01003922.

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science, 11*(2), 239–264. https://doi.org/10.1177/1745691615621279.

Cowan, N., Saults, J. S., & Clark, K. M. (2015). Exploring age differences in visual working memory capacity: Is there a contribution of memory for configuration? *Journal of Experimental Child Psychology, 135*, 72–85. https://doi.org/10.1016/j.jecp.2015.03.002.

Emrich, S. M., Lockhart, H. A., & Al-Aidroos, N. (2017). Attention mediates the flexible allocation of visual working memory resources. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1454–1465. https://doi.org/10.1037/xhp0000398.

Feigenson, L., & Yamaguchi, M. (2009). Limits on infants' ability to dynamically update object representations. *Infancy, 14*(2), 244–262. https://doi.org/10.1080/15250000802707096.

Feiman, R., Mody, S., & Carey, S. (2022). The development of reasoning by exclusion in infancy. *Cognitive Psychology, 135*. https://doi.org/10.1016/j.cogpsych.2022.101473 101473.

Feiman, R., Mody, S., Sanborn, S., & Carey, S. (2017). What do you mean, no? Toddlers' comprehension of logical "no" and "not". *Language Learning and Development, 13*(4), 430–450. https://doi.org/10.1080/15475441.2017.1317253.

Ferrigno, S., Huang, Y., & Cantlon, J. F. (2021). Reasoning through the disjunctive syllogism in monkeys. *Psychological Science, 32*(2), 292–300. https://doi.org/10.1177/0956797620971653.

Gautam, S., Suddendorf, T., & Redshaw, J. (2021). When can young children reason about an exclusive disjunction? A follow-up to Mody and Carey (2016). *Cognition, 207*. https://doi.org/10.1016/j.cognition.2020.104507 104507.

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition, 21*, 115–124. https://doi.org/10.3758/BF03211170.

Grigoroglou, M., Chan, S., & Ganea, P. A. (2019). Toddlers' understanding and use of verbal negation in inferential reasoning search tasks. *Journal of Experimental Child Psychology, 183*, 222–241. https://doi.org/10.1016/j.jecp.2019.02.004.

Hamilton, M., Ross, A., Blaser, E., & Kaldy, Z. (2022). Proactive interference and the development of working memory. *Wiley Interdisciplinary Reviews Cognitive Science, 13*(3). https://doi.org/10.1002/wcs.1593 e1593.

Hill, A., Collier-Baker, E., & Suddendorf, T. (2012). Inferential reasoning by exclusion in children (Homo sapiens). *Journal of Comparative Psychology, 126*(3), 243–254. https://doi.org/10.1037/a0024449.

Kibbe, M. M. (2015). Varieties of visual working memory representation in infancy and beyond. *Current Directions in Psychological Science, 24*(6), 433–439. https://doi.org/10.1177/0963721415605831.

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. (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., & Leslie, A. M. (2013). What's the object of object working memory in infancy? Unravelling "what" and "how many". *Cognitive Psychology, 66*(4), 380–404. https://doi.org/10.1016/j.cogpsych.2013.05.001.

Kovács, Á. M., Téglás, E., & Csibra, G. (2021). Can infants adopt underspecified contents into attributed beliefs? Representational prerequisites of theory of mind. *Cognition, 213*. https://doi.org/10.1016/j.cognition.2021.104640 104640.

Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences, 24*(1), 65–78. https://doi.org/10.1016/j.tics.2019.11.004.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences, 2*(1), 10–18. https://doi.org/10.1016/s1364-6613(97)01113-3.

Ma, Z., & Flombaum, J. I. (2013). Off to a bad start: Uncertainty about the number of targets at the onset of multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance, 39*(5), 1421–1432. https://doi.org/10.1037/a0031353.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158.

Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition, 154*, 40–48. https://doi.org/10.1016/j.cognition.2016.05.012.

O'Hara, M., Schwing, R., Federspiel, I., Gajdon, G. K., & Huber, L. (2016). Reasoning by exclusion in the kea (Nestor notabilis). *Animal Cognition, 19*(5), 965–975. https://doi.org/10.1007/s10071-016-0998-x.

Pailian, H., Carey, S. E., Halberda, J., & Pepperberg, I. M. (2020). Age and species comparisons of visual mental manipulation ability as evidence for its development and evolution. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-64666-1 7689.

Pailian, H., Libertus, M. E., Feigenson, L., & Halberda, J. (2016). Visual working memory capacity increases between ages 3 and 8 years, controlling for gains in attention, perception, and executive control. *Attention, Perception, & Psychophysics, 78*(6), 1556–1573. https://doi.org/10.3758/s13414-016-1140-5.

Premack, D. (1995). Cause/induced motion: Intention/spontaneous motion. In J. P. Changeux & J. Chavaillon (Eds.), *Origins of the human brain* (pp. 286–308). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198523901.003.0018.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002.

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.

Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology, 166*, 160–177. https://doi.org/10.1016/j.jecp.2017.08.005.

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(1), 37–55. https://doi.org/10.1037/xlm0000578.

van den Berg, R., & Ma, W. J. (2018). A resource-rational theory of set size effects in human visual working memory. *eLife, 7*. https://doi.org/10.7554/eLife.34963 e34963.

Völter, C. J., Sentís, I., & Call, J. (2016). Great apes and children infer causal relations from patterns of variation and covariation. *Cognition, 155*, 30–43. https://doi.org/10.1016/j.cognition.2016.06.009.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., Chiang, J. K., Hongwanishkul, D., Schuster, B. V., Sutherland, A., ... Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), vii-137. https://doi.org/10.1111/j.0037-976x.2003.00260.x.