AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0001562

Creativity and Flexibility in Young Children's Use of External Cognitive Strategies

Kristy L. Armitage[1], Thomas Suddendorf[1], Adam Bulley[2, 3, 4], Amalia P. M. Bastos[5, 6], Alex H. Taylor[7, 8], and Jonathan Redshaw[1]

[1] School of Psychology, The University of Queensland
[2] School of Psychology, The University of Sydney
[3] Brain and Mind Centre, The University of Sydney
[4] Department of Psychology, Harvard University
[5] School of Psychology, The University of Auckland
[6] Department of Cognitive Science, University of California, San Diego
[7] Institut de Neurociències, Universitat Autònoma de Barcelona
[8] ICREA, Barcelona, Spain

A cardinal feature of adult cognition is the awareness of our own cognitive struggles and the capacity to draw upon this awareness to offload internal demand into the environment. In this preregistered study conducted in Australia, we investigated whether 3–8-year-olds ($N = 72$, 36 male, 36 female, mostly White) could self-initiate such an external metacognitive strategy and transfer it across contexts. Children watched as an experimenter demonstrated how to mark the location of a hidden prize, thus helping them successfully retrieve that prize in the future. Children were then given the opportunity to spontaneously adopt an external marking strategy across six test trials. Children who did so at least once were then introduced to a conceptually similar but structurally distinct transfer task. Although most 3-year-olds deployed the demonstrated strategy in the initial test phase, none of them modified that strategy to solve the transfer task. By contrast, many children aged 4 years and older spontaneously devised more than one previously unseen reminder-setting strategy across the six transfer trials, with this tendency increasing with age. From age 6, children deployed effective external strategies on most trials, with the number, combination, and order of unique strategies used varying widely both within and across the older age groups. These results demonstrate young children's remarkable flexibility in the transferral of external strategies across contexts and point to pronounced individual differences in the strategies children devise.

Public Significance Statement

We investigated whether 3–8-year-old children could adopt a reminder-setting strategy to help them find and retrieve a hidden prize and then, without any prompting, transfer the logic behind the strategy to solve a related problem in a novel context. From age 4, children were using previously unseen reminder-setting strategies in the transfer task that followed the same logic as the initial task, with this tendency increasing markedly with age. The number, combination, and order of strategies used varied widely both within and across age groups, indicating remarkable creativity and flexibility in children's external strategy use.

Kristy L. Armitage https://orcid.org/0000-0002-1898-8622
Thomas Suddendorf https://orcid.org/0000-0003-3328-7442
Adam Bulley https://orcid.org/0000-0001-8014-0078
Amalia P. M. Bastos https://orcid.org/0000-0003-3037-687X
Alex H. Taylor https://orcid.org/0000-0003-3492-7667
Jonathan Redshaw https://orcid.org/0000-0002-7729-1577

This work was supported by an ARC Discovery Early Career Researcher Award awarded to Jonathan Redshaw (DE210100005), an ARC Discovery Project to Thomas Suddendorf, Jonathan Redshaw, and Adam Bulley (DP210101572), and a Royal Society Te Apārangi Marsden Fund to Alex H. Taylor and Thomas Suddendorf. Adam Bulley is supported by an Australian National Health and Medical Research Council CJ Martin Biomedical Fellowship APP1162811 (Grant GNT1162811).

The authors hereby declare no conflicts of interest. The data that support the findings of this study are openly available on the Open Science Framework at https://osf.io/4hctd/. Ethical approval was obtained through the University of Queensland's Faculty of Health and Behavioural Science's Ethics Committee (Clearance ID: 2019000267).

Kristy L. Armitage assisted with study design, collected and analyzed data, wrote the first draft, and edited the manuscript. Thomas Suddendorf and Adam Bulley assisted with study design and edited the manuscript. Alex H. Taylor and Amalia P. M. Bastos designed the study and edited the manuscript. Jonathan Redshaw assisted with study design, supervised the project, and edited the manuscript.

Correspondence concerning this article should be addressed to Kristy L. Armitage, School of Psychology, The University of Queensland, Sir Fred Schonell Drive, St Lucia, QLD, 4072, Brisbane, Australia. Email: kristy.armitage@uq.net.au

Keywords: cognitive offloading, reminder setting, strategy selection, metacognition, problem solving

Supplemental materials: https://doi.org/10.1037/dev0001562.supp

Humans have a long history of storing to-be-remembered information in their environment. Inscribed tokens used for documenting trade in 8000 B.C. in Uruk, Mesopotamia, are a frequently cited example of this phenomenon (Chatterjee, 2017; Schmandt-Besserat, 1978), but there were many other ancient mnemonic devices used for keeping records. For example, Incas (Ascher & Ascher, 1981) and Polynesians (E. Best, 1921) used quipus, or knotted cords, to store and convey numerical information, and the Eastern Woodlands tribes of Native Americans used beads fashioned from the shells of clams to record important treaties and historical events (Haas, 2008). With the progressive emergence of technologies like writing, printing, audio recording, computers, and smartphones, external memory strategies have become ever more central to managing our lives.

Consider, for example, the need to remember to attend an important meeting, turn off the stove after cooking, take medication, or study for an upcoming exam. Adults (Smith et al., 2000) and children (Kliegel & Jäger, 2007; Mahy et al., 2018; Mazachowsky et al., 2021) often fail to carry out the right action at the right time. To prevent such prospective memory failures, adults frequently set themselves reminders (Henry et al., 2012; Kim & Mayhorn, 2008), and are more likely to do so in response to (a) increases in the amount of to-be-remembered information (Gilbert, 2015a), (b) poorer objective unaided memory abilities (Gilbert, 2015b), and (c) lower task-specific metacognitive confidence in unaided memory abilities (independent of objective unaided ability; Gilbert, 2015b). After watching a demonstration, children can likewise set reminders to compensate for internal cognitive limits from as early as 4 or 5 years of age (Bulley et al., 2020; Heisel & Ritter, 1981).

Although some evidence indicates that *spontaneous* uses of external cognitive strategies are rare until at least 6 or 7 years (Armitage & Redshaw, 2022; Bulley et al., 2020), one recent study revealed that, in narrow circumstances, even 4- and 5-year-old children could spontaneously devise an effective reminder-setting strategy (Armitage et al., 2022). Children watched an experimenter place a target into one of three identical opaque containers and then shuffle the containers out of view, after which children were asked to guess where the target was hidden. In the test phase, children were asked to mark one of the three containers by placing an object into a transparent jar attached to the containers prior to shuffling, meaning they were able to perform perfectly if they simply marked the rewarded container and then selected that marked container after shuffling. Children aged 3, 4, and 5 years used this external strategy above chance levels if they had seen it demonstrated to them, but only 4- and 5-year-olds were able to spontaneously devise the strategy to improve their future recall. A caveat of this design, however, is that children were required to mark either the correct container or an incorrect container on every trial and could not refrain from marking (such that chance performance was 33.3%). Therefore, children's reminder-setting behaviors in this study were not entirely self-initiated.

Despite recent studies documenting the basic emergence of external cognitive strategies, it is currently unknown whether young children can transfer such strategies across contexts. It would be inefficient and cognitively taxing for an individual to devise new strategies in response to each new task—consider, for example, using a calendar only to track upcoming doctors' appointments and trying to devise entirely new strategies for reminding oneself of other upcoming meetings or events. Instead, people benefit from recognizing the general relevance of a strategy across contexts and implementing it accordingly (Clerc et al., 2014; Day & Goldstone, 2012). Proficiency with strategy transfer may be especially beneficial for children, who are frequently exposed to new academic and everyday challenges. Children that can recognize the generalizability of an external cognitive strategy to different but relevant contexts and flexibly adapt the strategy in accordance with new task parameters will therefore be better equipped to deal with these challenges, which may lead to improved learning and problem-solving performances. Understanding how cognitive offloading develops, and particularly the degree to which children can generalize external cognitive strategies, is therefore of high interest from both pure and applied research perspectives.

The development of external strategy transfer might be expected to follow a similar trajectory to the (more well-documented) development of internal strategy transfer. The evidence for this development, however, is mixed. In one study, 4–5-year-olds were presented with a memory task, and the children either spontaneously devised or were prompted to adopt the strategy of only studying to-be-recalled items (and avoid studying items that were not to be recalled). Children were then able to transfer this strategy into a new memorization context, albeit with decreased recall performance (Clerc & Miller, 2013). Other evidence suggests that children become more likely to generalize the use of an internal strategy to new contexts with increasing age (Schwenck et al., 2009), with proficient transfer only demonstrated around 7 or 8 years and often facilitated by explicit instructions or prompting (Chen & Klahr, 1999; Klahr & Nigam, 2004; O'Sullivan & Pressley, 1984; Ringel & Springer, 1980).

Here, we examined whether young children could engage in fully self-initiated reminder setting and also explored whether they could transfer that reminder-setting strategy to a novel context. We adapted Armitage et al.'s (2022) paradigm but removed the requirement for children to mark one of three containers on each trial, such that we measured children's spontaneous reminder-setting behaviors rather than comparing their performance against an a priori chance level. In this new design, 3–8-year-old children were given the opportunity to evaluate their unaided competence, decide that a reminder was necessary, and self-initiate a reminder-setting action in order to ensure they would retrieve the reward, akin to real-world compensatory behaviors. Children who successfully self-initiated an external strategy were then introduced to the transfer task, which was structurally distinct from the original task but with the same underlying conceptual demands. We reasoned that, if children truly understood the logic of the task, they should have been able to transfer the general principle of the reminder-setting behavior across contexts and solve the second task. Our preregistered hypotheses outlined our general expectation that children would set more self-initiated reminders and would be more likely to transfer this behavior across contexts with increasing age.

Method

Participants

This study was preregistered on Open Science Framework before commencing data collection (https://osf.io/4hctd/). In our preregistration, we specified that we would test 5–10-year-olds, but as even 5-year-olds performed relatively proficiently during pilot testing, we instead tested 3–8-year-olds. Our anticipated sample size was 72 children, as it is comparable to the sample size used in another study of children's use of external reminders (Redshaw et al., 2018) and allowed for two children in each age group to complete each counterbalancing condition (see S1 in the online supplemental materials for complete counterbalancing information). A post hoc power analysis revealed that this sample size provided a 99.8% chance of detecting hypothesized large age effects (equivalent to $r = .50$). We also specified that children would be recruited and tested at a public museum but given changes to the test protocol resulting from COVID-19, all participants were instead recruited through an existing university database. Recruitment ceased when the intended sample size for each age was reached.

The final sample of 72 children (36 males, 36 females) aged between 3.03 and 8.94 years ($M = 6.01$, $SD = 1.71$) was made up of 12 children (six males, six females) from each of the following ages: 3-year-olds ($M = 3.64$, $SD = 0.25$), 4-year-olds ($M = 4.46$, $SD = 0.30$), 5-year-olds ($M = 5.51$, $SD = 0.27$), 6-year-olds ($M = 6.49$, $SD = 0.18$), 7-year-olds ($M = 7.50$, $SD = 0.23$), and 8-year-olds ($M = 8.48$, $SD = 0.31$). Twelve additional children were excluded due to either experimenter error ($n = 5$), having a clinical diagnosis ($n = 2$), being of an age for which data collection had already been completed ($n = 2$), being instructed to set reminders by an observer ($n = 1$), and for passing fewer than two of the three training trials ($n = 1$). All caregivers provided informed consent before testing commenced. Ethical approval was obtained prior to beginning recruitment.

Participating families resided in the Brisbane region and surrounding areas in southeast Queensland, Australia. The Australian Bureau of Statistics Socio-Economic Indices for Areas (SEIFA) was used to estimate socioeconomic status of each family according to their household postcode. SEIFA provides ranks for geographical areas from *low* (1) to *high* (10) based on characteristics such as employment and public resources (Australian Bureau of Statistics, 2018). SEIFA ranks for our sample ranged from 3 to 10 but were skewed toward socioeconomic advantage ($M = 8.69$, $SD = 1.59$). This skew was observed in 3-year-olds ($M = 8.89$, $SD = 1.05$), 4-year-olds ($M = 7.2$, $SD = 2.95$), 5-year-olds ($M = 9.00$, $SD = 0.00$), 6-year-olds ($M = 8.80$, $SD = 1.30$), 7-year-olds ($M = 9.25$, $SD = 0.50$), and 8-year-olds ($M = 9.50$, $SD = 0.71$). We also collected data on the highest parental education level in each family. This ranged from a high school certificate or equivalent (1.69%), a trade or other apprenticeship (1.69%), a posthigh school certificate or diploma (20.34%), an undergraduate university degree (27.12%), and a postgraduate university degree (47.46%), which indicated a disproportionate number of families with at least one highly educated parent.

Design and Procedure

Experience Task

As in Armitage et al. (2022), children were presented with three identical opaque containers, each attached to a smaller transparent jar, placed on the ground between themselves and the experimenter. On each of six trials, the experimenter would remove the container lids, place a sticker into one container, and return the lids, such that the sticker was no longer visible (see Figure 1 Phase 1). The experimenter then concealed the containers behind a screen and told the child that they would "shuffle the containers all around, and then you have to guess where the sticker has gone." During shuffling, the experimenter would rearrange the containers into one of six possible orders (where one matched the starting position of the containers, such that the sticker was found in the same location as it was hidden) before removing the screen and asking children to guess which container held the sticker. Across the six trials, all six possible orders for the final position of containers were used, with the stickers being hidden twice in each container and found twice in each final position (see S1 in the online supplemental materials for complete counterbalancing information). After shuffling, children had no indication of where the sticker was hidden and had to simply guess between the three options.

Preliminary Training

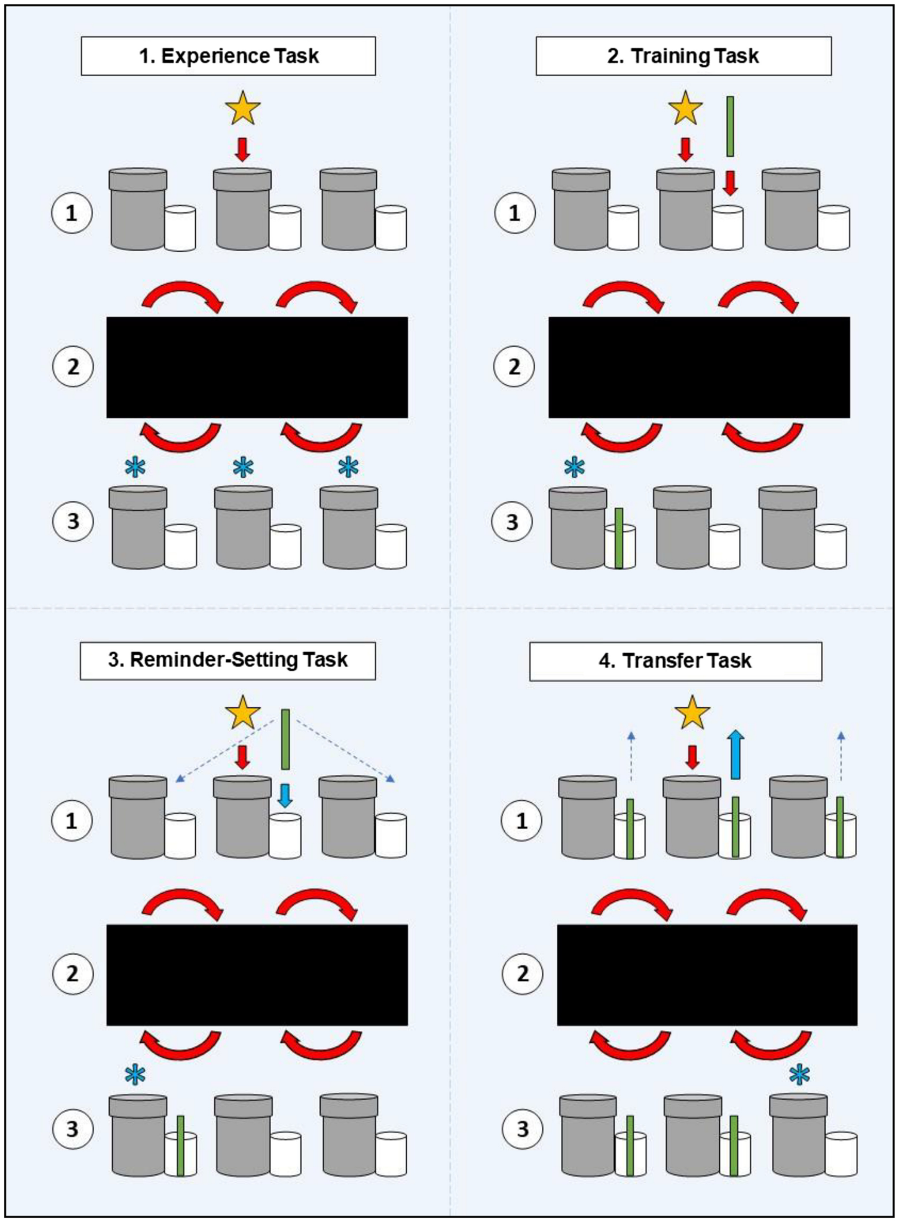
The experimenter replaced the apparatus with three differently shaped transparent jars. They then held three identical sticks in front of the child and explained, "We're now going to use these sticks, and we're going to call them 'markers'. In this game, you can always do whatever you want with your markers." The experimenter then placed one marker into each of the transparent jars, while saying "You can put your markers into containers, just like this." They then removed the markers from the jars, passed them to the child, and asked, "Can you try putting a marker into each of the containers?" Once the child had done this, the experimenter removed the markers from the jars, while saying "You can also take your markers out of containers, just like this." They returned the markers to the jars and asked, "Can you try taking the markers out of the containers?" This process allowed children to gain permission to add and remove markers throughout the task, and to practice these behaviors while avoiding the basic conditioning effects of performing this training using the test apparatus. Once the child had completed this stage, the experimenter replaced the three jars with the original apparatus used in the experience task.

Training Task

The experimenter then told the child that they would "start by using one marker." The child completed three trials of the training task, which followed the same procedure as the experience task, except that prior to shuffling, the experimenter placed one marker into the transparent jar attached to the rewarded container (see Figure 1 Phase 2). The marker therefore indicated the presence of the sticker, so that after the containers were shuffled, children simply had to locate the marker and search in the attached container (as for half of the children in Armitage et al., 2022). Across the three trials, three possible orders for the

Figure 1

A Representation of All Four Phases of the Task, With Each Phase Comprising Three Steps



Note. All children completed the first three phases of the task, but only those that used an external strategy in the third phase also completed the fourth phase. The experimenter's actions are represented in red, and the child's actions in blue. In each phase, the experimenter always hid a sticker (denoted by the yellow stars) and shuffled the containers behind a screen (denoted by the black bars), and in Phase 2, the experimenter also placed the marker (denoted by green bars). In Phase 3, the marker was given to the child, and the bolded blue arrow indicates where they should have placed the marker in order to indicate the location of the reward. In Phase 4, all three containers were marked, and the bolded blue arrow indicates which marker the child should have removed in order to indicate the location of the reward. In each phase, the blue asterisks above the containers in step 3 represent where the child should search for the sticker if they understand the task (and have marked correctly). Note that the actual target hiding locations and shuffled container locations were counterbalanced across children and trials. See the online article for the color version of this figure.

final position of containers were chosen, with the stickers being hidden once in each container and found once in each final position. The data for each child were only included in analyses if they passed at least two of the three training trials, as evidenced by correctly selecting the marked container after shuffling.

Reminder-Setting Task

Children then completed six trials of the reminder-setting task, which also followed the same procedure as the experience task, except that prior to shuffling, the experimenter placed one marker on the ground in front of the child while saying "Remember, you can do whatever you want with this one." The experimenter covered their eyes and asked the child to let them know when they were ready for the containers to be shuffled (see Figure 1 Phase 3). Once the child told the experimenter that they were ready, the shuffling procedure commenced. The reminder-setting task followed the same counterbalancing procedures as outlined in the experience task. All variables were coded using video recordings of testing sessions.

We were interested in whether children would spontaneously place the available marker into the jar attached to the rewarded container, such that the marker differentiated the rewarded container from the unrewarded containers, as observed during training. After the shuffling procedure, they would then simply need to locate the marker and search in the attached container. By contrast, if children did not use the available marker, or placed it into a jar attached to an unmarked container, such that the marker did not differentiate the rewarded container from the unrewarded containers, they would be left guessing the answer. In other words, in the reminder-setting task children could not use the mere *absence* of a marker as a cue to find the target. If children did not use a marker at all in this task, they were thanked for their time, compensated with a small prize, and the caregiver was provided with a written and verbal debrief.

Transfer Task

Children who used at least one marker in the previous task (either correctly or incorrectly) were administered six trials of the transfer task. This task followed the same procedure as the reminder-setting task, except that the experimenter placed one marker into each of the three transparent jars, while saying "This time, we're going to use all three markers." The experimenter then pointed to the markers and said, "Remember that you can do whatever you want with these." Although this task could be solved using the same logic as the reminder-setting task, by differentiating the rewarded container from the unrewarded container, children had to spontaneously devise a novel means of achieving this. For example, rather than *adding* a marker to an empty transparent jar, as in the reminder-setting task, children could instead choose to *remove* the marker from the rewarded container, and leave the unrewarded containers marked (see Figure 1 Phase 4). Thus, the absence of a marker would indicate the presence of the sticker, so that after the containers were shuffled, children would simply have to locate the empty transparent jar and search in the attached container. The transfer task followed the same counterbalancing procedures as the experience task. Once children had completed their six trials, they were thanked and compensated, and their caregiver was debriefed.

Results

Children's responses across all six trials in the relevant phases were analyzed using generalized estimating equations (GEEs) that accounted for covariance between each individual's responses. Our preregistration specified an analysis plan for children's marking behaviors, but we also report additional analyses to provide a fuller picture of children's performance in each task. Any analyses that were not included in our preregistration have been explicitly labeled as exploratory. All analyses were performed using the statistical programs R (R Core Team, 2019) and SAS 9.4 (SAS Institute Inc., 2013). No significant sex differences were detected across any dependent variables in the data set, and so we do not report effects of sex in the main text. All model details (including modeled interactions and sex effects) for both preregistered and exploratory analyses can be found in S2–S5 in the online supplemental materials. All data and syntax have been made publicly available at the Open Science Framework and can be accessed at https://osf.io/4hctd/.

Experience Task

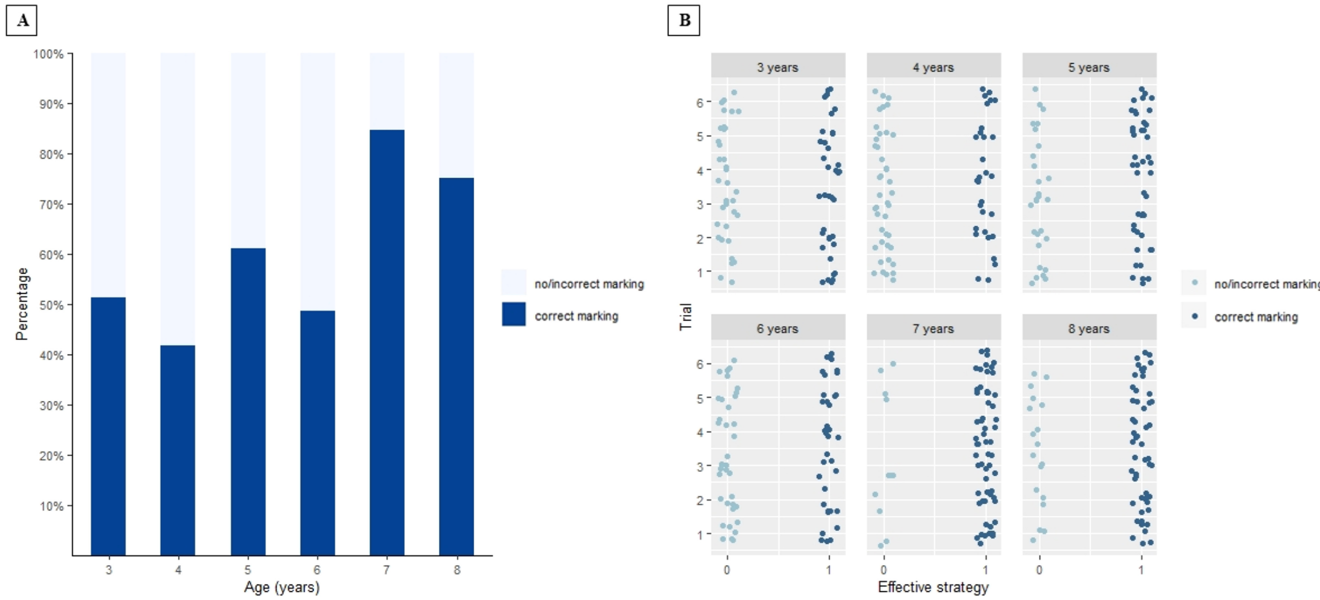
A one-sample t-test indicated that children's average search accuracy ($M = 2.19$) did not significantly differ from chance (2/6), $t(71) = 1.18$, $p = .243$, substantiating the assumption that they were simply guessing the target locations.

Reminder-Setting Task

We preregistered our intent to measure and analyze *any* use of marking, which includes instances where children marked an incorrect container, and the *correct* use of marking, which only includes instances in which the marker was correctly placed into the jar attached to the rewarded container (results for each dependent variable can be found in S3 in the online supplemental materials). Any marking and correct marking in the reminder-setting task were both modeled as a function of age (linear) and trial (linear). For any marking, there was a significant main effect of age, $\chi^2(1, N = 72) = 8.48$, $p = .004$, $w = 0.34$ (Cohen, 1992), with 58.33% of 3- and 4-year-olds, 83.33% of 5-year-olds, 66.67% of 6-year-olds, and 100% of 7- and 8-year-olds spontaneously marking at least once across their six trials. Overall, 3-year-old children spontaneously marked on 58.33% of trials, 4-year-olds on 54.17% of trials, 5-year-olds on 73.61% of trials, 6-year-olds on 63.89% of trials, 7-year-olds on 98.61% of trials, and 8-year-olds on 90.28% of trials. A post hoc test indicated that the apparent drop in performance between 5- and 6-year-olds was not statistically significant, $z = 1.26$, $p = .104$. There was no significant main effect of trial, $\chi^2(1, N = 72) = 2.77$, $p = .096$, $w = 0.20$, such that children were not more likely to spontaneously mark on later trials. For correct marking, there was also a significant main effect of age, $\chi^2(1, N = 72) = 5.76$, $p = .016$, $w = 0.28$, as seen in Figure 2, and no significant main effect of trial, $\chi^2(1, N = 72) = 2.58$, $p = .108$, $w = 0.19$. Again, a post hoc analysis revealed no significant difference between the marking behaviors of 5- and 6-year-old children, $z = 1.54$, $p = .062$.

An exploratory analysis assessed the effect of marking correctly on children's search accuracy. Accuracy was modeled as a function of age (linear), trial (linear), and correct marking (binary). When children marked correctly, their search accuracy was significantly increased, $\chi^2(1, N = 72) = 46.88$, $p < .001$, $w = 0.81$, suggesting that children did indeed recognize the utility of the reminders in guiding their search behavior. While descriptive data suggest that the average number of trials answered correctly increased roughly with age across 3-year-olds ($M = 3.50$, $SD = 2.24$), 4-year-olds ($M = 3.42$, $SD = 2.23$), 5-year-olds ($M = 3.91$, $SD = 2.07$), 6-year-olds ($M = 4.25$, $SD = 1.29$), 7-year-olds ($M = 5.33$, $SD = 1.50$), and 8-year-olds ($M = 5.00$, $SD = 1.35$), there was no significant unique effect of

Figure 2
Children's Marking Behavior in the Reminder-Setting Task



Note. (A) Percentage of trials (*y*-axis) on which children showed no/incorrect marking or correct marking, split according to age group (*x*-axis). (B) Each participant's use of no/incorrect marking or correct marking (*x*-axis) on each trial (*y*-axis), split according to age group. See the online article for the color version of this figure.

age in the model, $\chi^2(1, N = 72) = 2.35, p = .125, w = 0.18$. This finding indicates that, when accounting for reminder use, neither older nor younger children were more likely to find another systematic means to pass the task. There was also no significant main effect of trial, $\chi^2(1, N = 72) = 2.07$, $p = .150$, $w = 0.17$, and no age by correct marking interaction, $\chi^2(1, N = 72) = 0.64$, $p = .423$, $w = 0.09$, indicating that reminders were similarly useful to children across ages.

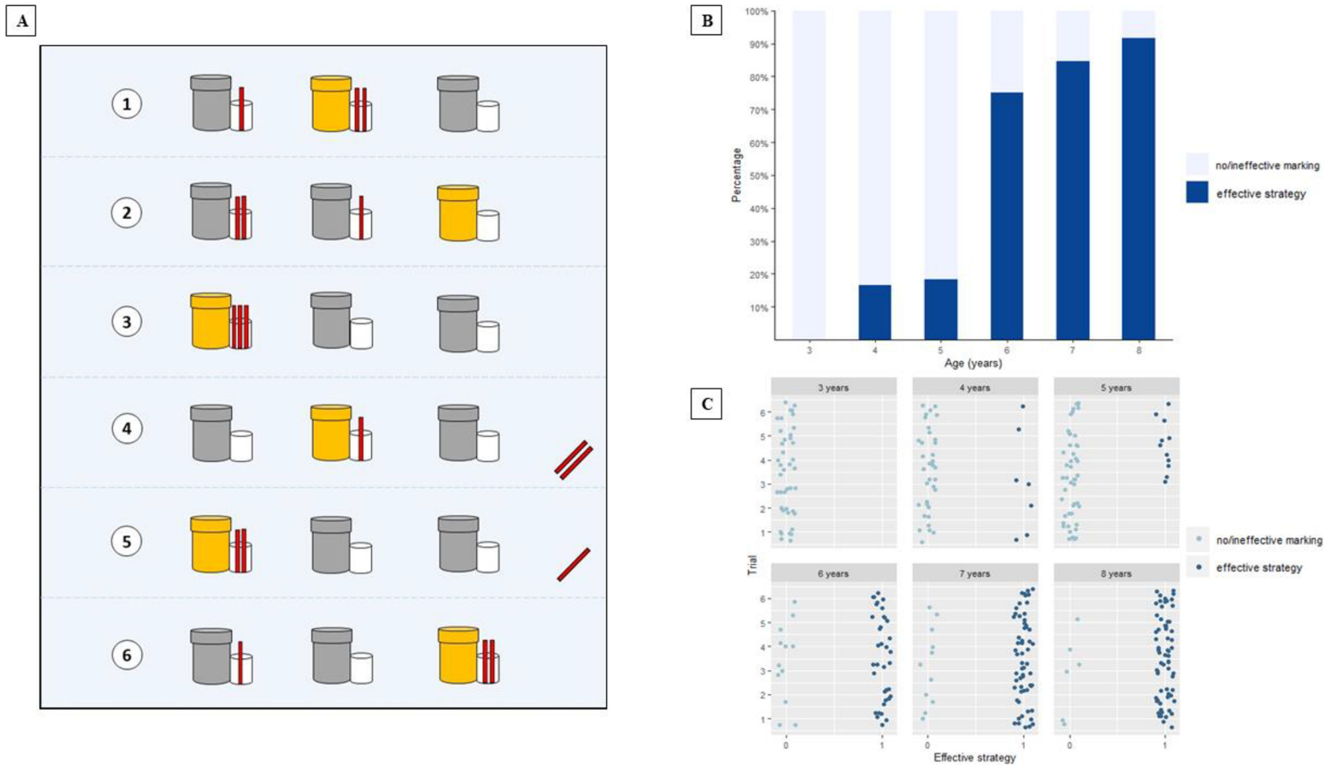
Transfer Task

Of the 72 participants in our original sample, 56 used a marker (either correctly or incorrectly) at least once in the reminder-setting task and thus met the criteria for inclusion in the transfer task. The transfer sample included seven 3-year-olds (58.3%), seven 4-year-olds (58.3%), ten 5-year-olds (83.3%), eight 6-year-olds (66.7%), twelve 7-year-olds (100.0%), and twelve 8-year-olds (100.0%) from the original sample. The preregistered transfer task analyses were run first with only this limited sample, and then again with a full sample where scores of 0 were imputed for children who did not complete the transfer task (to simulate a random sample from the population based on the assumption that children who did not mark in the reminder-setting task would also fail to mark in the transfer task). As the pattern of significant findings was consistent across both samples, we only report the results from the limited sample of children who qualified for and completed the transfer task in text (results from both samples can be found in S4 and S5 in the online supplemental materials).

As for the reminder-setting task, we also measured and analyzed *any* use of marking and the *correct* use of marking in the transfer task (results for each dependent variable can be found in S4 in the online supplemental materials). Correct marking in the transfer task was defined as any marking that differentiated the target container from both of the other containers (see Figure 3 for example). Any marking and correct marking were both modeled as a function of age (linear) and trial (linear). For any marking, there was a significant main effect of age, $\chi^2(1, N = 56) = 32.94$, $p < .001$, $w = 0.77$, with 0.00% of 3-year-olds, 57.14% of 4-year-olds, 30.00% of 5-year-olds, 87.50% of 6-year-olds, 91.67% of 7-year-olds, and 100% of 8-year-olds spontaneously marking at least once across their six trials. While 3-year-olds never marked spontaneously, 4-year-olds marked on 19.05% of trials, 5-year-olds on 21.67% of trials, 6-year-olds on 81.25% of trials, 7-year-olds on 88.89% of trials, and 8-year-olds on 97.22% of trials. A post hoc test indicated that the apparent increase in performance between 5- and 6-year-olds was statistically significant, $z = 6.15$, $p < .001$. There was no significant main effect of trial, $\chi^2(1, N = 56) = 0.55$, $p = .458$, $w = 0.10$. For correct marking, there was also a significant main effect of age, $\chi^2(1, N = 56) = 32.18$, $p < .001$, $w = 0.76$, as seen in Figure 3, and no significant main effect of trial, $\chi^2(1, N = 56) = 1.88$, $p = .170$, $w = 0.18$. Again, a post hoc analysis revealed a significant difference between the marking behaviors of 5- and 6-year-old children, $z = 5.73$, $p < .001$.

An exploratory analysis was conducted to assess the effect of marking correctly on children's search accuracy. Children's search accuracy was modeled as a function of age (linear), trial (linear), and correct marking (binary). A significant main effect of correct marking indicated that this behavior did indeed improve children's performance on the transfer task, $\chi^2(1, N = 56) = 19.43$, $p < .001$, $w = 0.59$ (see Figure 3 Panels B and C for a visual representation of this age effect). As in the reminder-setting task, descriptive data suggested an age-related increase in the average number of trials

Figure 3
Children's Marking Behavior in the Transfer Task



Note. (A) Representation of one participant's use of five different effective strategies across the six transfer trials (note that the same strategy was used on the first and sixth trials). (B) The percentage of trials (*y*-axis) on which children set effective or no/ineffective reminders, split according to age group (*x*-axis). (C) Each participant's use of effective and no/ineffective reminders (*x*-axis) on each trial (*y*-axis), split according to age group. See the online article for the color version of this figure.

answered correctly across 3-year-olds ($M = 2.14$, $SD = 1.07$), 4-year-olds ($M = 2.00$, $SD = 1.41$), 5-year-olds ($M = 2.60$, $SD = 1.35$), 6-year-olds ($M = 5.13$, $SD = 1.46$), 7-year-olds ($M = 4.83$, $SD = 1.64$), and 8-year-olds ($M = 5.25$, $SD = 1.06$). Again, however, there was no significant unique effect of age in the model, $\chi^2(1, N = 56) = 0.14$, $p = .707$, $w = 0.05$. There was also no significant main effect of trial, $\chi^2(1, N = 56) = 0.71$, $p = .399$, $w = 0.11$, and no age by correct marking interaction, $\chi^2(1, N = 56) = 0.11$, $p = .743$, $w = 0.04$, indicating that children across ages similarly benefited from marking correctly.

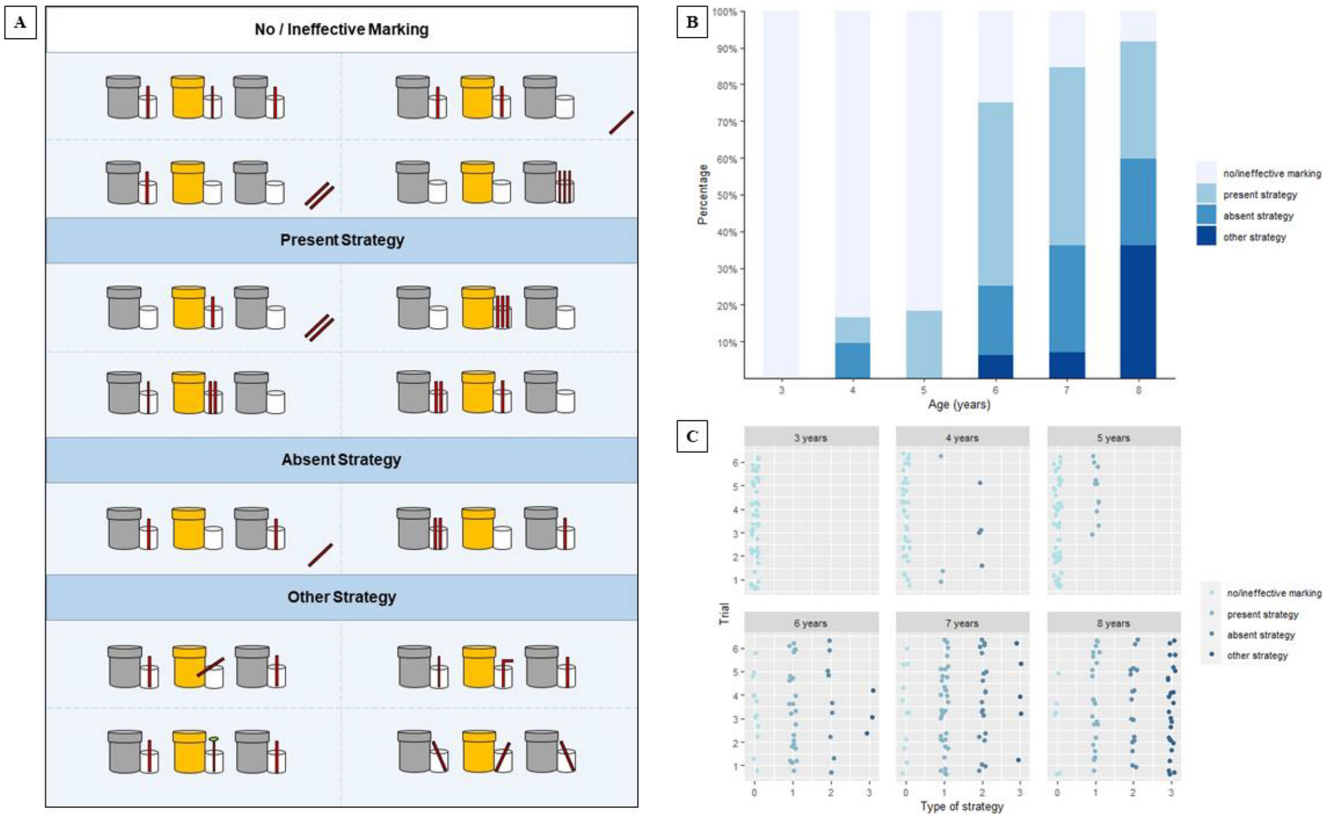
Flexibility of Marking Behaviors

Unlike in the reminder-setting task, there were multiple ways for children to mark correctly in the transfer task (see Figure 3 for examples), and many children showed unexpected flexibility in finding solutions. We therefore explored the number of different effective marking strategies devised by children in this final phase. A child would score a 1 if they only devised one effective strategy (regardless of whether they used it once or 6 times), and a 6 if they devised six different effective strategies (i.e., a different strategy on each trial). The number of effective strategies devised by the children in our sample ranged from 0 to 5, and increased with age, $r(54) = .602$, $p < .001$, with an average of 0.00 strategies for 3-year-olds, 1.14 strategies for 4-year-olds, 1.40 strategies for 5-year-olds, 1.88 strategies for 6-year-olds, 2.75 strategies for 7-year-olds, and 2.92 strategies for 8-year-olds (see Figure 3 Panel A for an example of the marking behavior of an 8-year-old child who devised five different effective strategies and successfully located all six stickers across the six trials).

We then organized children's marking behaviors into four categories. *No/ineffective marking* included trials on which children either refrained from marking or marked in a way that did not differentiate the rewarded container from the unrewarded containers. A *present* strategy included trials on which the presence of one, two or three marker/s differentiated the rewarded container from the unrewarded containers. An *absent* strategy included trials on which the absence of a marker differentiated the rewarded container from the unrewarded containers. Some marking strategies did not meet the criteria of present or absent strategies because they did not involve the actions of adding or removing markers. These were termed *other* strategies and may be considered more innovative as they do not resemble behaviors demonstrated throughout the experiment (see Figure 4 Panel A for examples of each type of strategy).

To analyze these data, we created a separate variable for each category, where a score of 1 indicated that the marking behavior of interest was used, and a score of 0 indicated that another marking behavior was used. We then modeled each category as a function

Figure 4
Children's Strategy Use in the Transfer Task



Note. (A) Examples of marking behaviors that were categorized as either *no/ineffective* marking, or as an effective strategy for differentiating the rewarded container from the unrewarded container, classified as either a *present*, *absent,* or *other* strategy. Note that in the *other* category, the top-left diagram shows a marker being placed across rather than inside the glass jar, the top-right diagram shows a marker being bent out of shape, the bottom-left diagram shows fluff from a shirt or carpet being placed on top of a marker, and the bottom-right diagram shows a marker being pointed in the opposite direction to the other markers. (B) The percentage of trials (y-axis) on which no/ineffective marking, and present, absent, and other strategies were used, split according to age group (x-axis). (C) The percentage of participants' use of no/ineffective marking, and present, absent, and other strategies (x-axis) on each trial (y-axis), split according to age group. See the online article for the color version of this figure.

of age (linear) and trial (linear). Significant main effects of age were detected in all models. Children became *less* likely to refrain from marking or use their markers ineffectively with increasing age, $\chi^2(1, N = 56) = 32.58$, $p < .001$, $w = -0.76$, with 3-year-olds meeting these criteria on 100% of trials, 4-year-olds on 83.33% of trials, 5-year-olds on 81.67% of trials, 6-year-olds on 25.00% of trials, 7-year-olds on 15.28% of trials, and 8-year-olds on 8.33% of trials. By contrast, for the present strategy use, $\chi^2(1, N = 56) = 9.96$, $p = .002$, $w = 0.42$, absent strategy use, $\chi^2(1, N = 56) = 8.53$, $p = .004$, $w = 0.39$, and other strategy use, $\chi^2(1, N = 56) = 6.43$, $p = .011$, $w = 0.34$, children became *more* likely to use these strategies with increasing age (see Figure 4 Panels B and C for a visual depiction of these effects). Three-year-olds did not use present strategies on any trials, but 4-year-olds did on 7.14% of trials, 5-year-olds on 18.30% of trials, 6-year-olds on 50.00% of trials, 7-year-olds on 48.61% of trials, and 8-year-olds on 31.94% of trials. Three- and 5-year-olds also did not use absent strategies on any trials, but 4-year-olds did on 12.50% of trials, 6-year-olds on 18.75% of trials, 7-year-olds on 29.17% of trials, and 8-year-olds on 23.61% of trials. Finally, 3-, 4-, and 5-year-olds did not use other strategies on any trials, but 6-year-olds did on 6.25% of trials, 7-year-olds on 6.94% of trials, and 8-year-olds on 36.11% of trials. Across the four models, there were no main effects of trial (all χ^2s < 2.54, all ps $> .111$, all ws < 0.21), indicating that each category of marking behaviors was used relatively consistently across earlier and later trials.

Finally, we explored whether the use of a strategy that facilitated correct retrieval on one trial increased the likelihood of using the same type of strategy on the following trial. The type of strategy used by children on trials 2–6 of the transfer task was coded as either the *same* (1) or *different* (0) from the type of strategy used on the preceding trial (note that trial 1 was therefore excluded from this analysis). This dependent variable was then modeled as a function of age (linear), trial (linear), and accuracy on the preceding trial (binary). A significant main effect of age indicated an age-related decrease in the use of the same type of strategy on consecutive trials, $\chi^2(1, N = 56) = 5.41$, $p = .020$, $w = -0.31$, across 3-year-olds ($M = 1.00$, $SD = 0.00$), 4-year-olds ($M = 0.77$, $SD = 0.43$), 5-year-olds ($M = 0.94$, $SD = 0.24$), 6-year-olds ($M = 0.80$, $SD = 0.41$), 7-year-olds ($M = 0.70$, $SD = 0.46$), and 8-year-olds ($M = 0.67$,

$SD = 0.48$), substantiating the claim that older children showed greater flexibility in strategy use across trials. However, there was no significant main effect of accuracy on preceding trials, $\chi^2(1, N = 56) = 0.05$, $p = .819$, $w = -0.03$, indicating that children were not more likely to continue using strategies that proved to be effective, and this effect did not vary significantly with age, $\chi^2(1, N = 56) = 3.68$, $p = .055$, $w = 0.25$. There was also no significant main effect of trial, $\chi^2(1, N = 56) = 2.16$, $p = .142$, $w = 0.19$.

Discussion

We presented 3–8-year-old children with a paradigm that required self-initiated reminder-setting to successfully locate a reward on every trial. While all 7- and 8-year-old children demonstrated a proclivity to engage in self-initiated reminder setting, in line with the consistent finding that children of this age can readily recognize and compensate for their cognitive deficits (Armitage et al., 2020; Armitage & Redshaw, 2022; Bulley et al., 2020), we also found evidence of this behavior in the youngest children. Most 3-year-old children set at least one correct reminder in the initial reminder-setting task, suggesting that, even when the requirement to mark on every trial is removed (cf. Armitage et al., 2022), young children are still able to self-initiate their own external reminder-setting strategies.

It is likely, however, that the success of 3-year-old children in the reminder-setting task was due to being proficient imitators (Nielsen & Blank, 2011), such that the brief demonstration of the target behavior during the training phase was sufficient to guide subsequent reminder setting. Indeed, none of the 3-year-olds used any external strategy (effective or ineffective) in the transfer task, where doing so required a different set of actions guided by the same logic as the initial task. By contrast, some children aged 4 years and older were able to spontaneously transfer the logic of differentiating the marked container from unmarked containers to devise novel solutions to the transfer task, and by age 6 children were spontaneously using effective external strategies on the majority of transfer trials. The competence demonstrated by children in the transfer task extends beyond what can be attributed to imitation, providing compelling evidence that they understood the reasoning behind their marking behaviors in the reminder-setting task.

The results from the transfer task provide further evidence that external cognitive strategies become far more frequent from around 6 years (Armitage et al., 2020; Armitage & Redshaw, 2022; Bulley et al., 2020), potentially due to transitions in executive functions (J. R. Best & Miller, 2010) or metacognitive control (Schneider et al., 2022). Executive functions refer to the skills necessary for purposeful goal-oriented action (Anderson, 1998), such as problem solving, reasoning, and planning, and include the abilities to think creatively and flexibly (Diamond & Lee, 2011; Mahy et al., 2014). Many of these skills are directly relevant to the demands of the transfer task, and accordingly, the notable increases in external cognitive strategy use around 6 years correspond to marked developmental improvements in executive functions after age 5 (J. R. Best et al., 2009; J. R. Best & Miller, 2010). A critical avenue for future research is therefore to empirically test this hypothesis, to ascertain whether any variance in children's external cognitive strategy use can be explained by developmental improvements in executive functions.

The finding that at least some 4- and 5-year-old children transferred an external cognitive strategy aligns with research suggesting that children of this age can similarly transfer internal cognitive strategies across contexts (Clerc & Miller, 2013). Other studies, however, have not found a transfer of internal strategies until later in development (Chen & Klahr, 1999; Klahr & Nigam, 2004; O'Sullivan & Pressley, 1984; Ringel & Springer, 1980). Our finding of an early emergence may be due to a close "transfer distance" (Klahr & Chen, 2011) between the reminder-setting and transfer tasks. Indeed, while younger children are able to transfer internal cognitive strategies across contextually similar environments with minimal time delays, as in the current study, more remote transfer may only become possible later in childhood (Klahr & Chen, 2011). Future research should therefore investigate far transfer of external reminder-setting strategies, perhaps by creating a more visually and procedurally distinct transfer task and introducing a time delay between the learning and transfer contexts.

A particularly unexpected finding was the number of unique effective strategies devised by children in the transfer task. Children of all ages (aside from 3-year-old children who did not set any reminders in this task) showed some degree of flexibility in the application of the previously acquired strategy to a new context. Although there is no ostensible functional benefit of devising more than one effective reminder-setting strategy in this context, 4-year-old children were, on average, already devising more than one effective strategy across the six trials, possibly in an attempt to make the task more enjoyable or show off to the experimenter. This tendency increased with age, with many of the older children demonstrating remarkable flexibility in their strategy use and greater reluctance to deploy the same strategy on subsequent trials regardless of whether it was previously effective. Furthermore, many strategies devised by older children involved behaviors other than simply adding or removing markers, showing creativity beyond what they had observed in training (e.g., differentiating the rewarded container from the unrewarded containers by bending the relevant marker out of shape or pointing it toward a different direction; see Figure 4 for an illustration of these behaviors). This developmental pattern of increased use of multiple external reminder-setting strategies during early and middle childhood mirrors the developmental pattern seen for internal memory strategies (Schwenck et al., 2009). The pattern also aligns with more general age-related improvements in innovation, as measured by children's ability to create a novel tool to retrieve a reward (Beck et al., 2011), and divergent thinking, or the ability to generate multiple relevant and original alternative solutions to a single problem (Guilford, 1975; Said-Metwaly et al., 2021).

There were also pronounced individual differences in the types of strategies devised by children. While young children often refrained from marking or used their markers ineffectively, many older children deployed a range of effective strategy types classified as present, absent, or other. Within each age group (excluding the 3-year-olds), there was variability not only in the number of strategy types used, but also the combination and order of strategies across the six trials, with no single strategy predominating on earlier or later trials. While similar variability in strategy use has been found in same-age children after prompting them to use a specific internal memory strategy (Schwenck et al., 2009), our results demonstrate that such variability can even emerge spontaneously, while holding training procedures and instructions constant across all participants.

We have already highlighted several limitations of this study, including being unable to confirm the hypothesized relationship between executive functions and external cognitive strategy use, and only exploring near transfer rather than far transfer. Another notable limitation, however, is the generalizability of our findings. Our sample was drawn from a White, Educated, Industrialized, Rich, and Democratic (WEIRD; Heinrich et al., 2010) population and was skewed toward socioeconomic advantage. We also recruited from an opt-in university database, resulting in a sample of families able and willing to volunteer their time to tertiary research. In line with this sampling strategy, which was necessary due to COVID-related restrictions on testing in public locations (see Participants section), the majority of participating families reported being highly educated, with almost half having at least one parent with a higher tertiary qualification. Further research might therefore explore whether these findings generalize to families with more representative socioeconomic status and education levels, as well as across cultures to non-WEIRD populations.

Here, we have presented evidence suggesting that while 3-year-old children can copy and benefit from reminder-setting strategies, they may not yet possess the understanding required to devise their own. By contrast, from age 4, some children were able to engage in logical self-initiated reminder setting, as demonstrated by their ability to spontaneously devise previously unseen reminder-setting strategies in the transfer task that followed the same logic as the initial task, with this tendency increasing markedly with age. Unexpectedly, children aged 4 and older typically devised more than one effective external reminder-setting strategy in the transfer task, with the number, combination, and order of strategies used varying widely both within and across age groups. Our results therefore capture young children's remarkable flexibility in the transferal of external strategies across contexts and highlight the emergence of individual differences in spontaneous external strategy deployment.

References

Anderson, V. (1998). Assessing executive functions in children: Biological, psychological, and developmental considerations. *Neuropsychological Rehabilitation*, *8*(3), 319–349. https://doi.org/10.1080/713755568

Armitage, K. L., Bulley, A., & Redshaw, J. (2020). Developmental origins of cognitive offloading. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1928), Article 20192927. https://doi.org/10.1098/rspb.2019.2927

Armitage, K. L., & Redshaw, J. (2022). Children boost their cognitive performance with a novel offloading technique. *Child Development*, *93*(1), 25–38. https://doi.org/10.1111/cdev.13664

Armitage, K. L., Taylor, A. H., Suddendorf, T., & Redshaw, J. (2022). Young children spontaneously devise an optimal external solution to a cognitive problem. *Developmental Science*, *25*(3), Article e13204. https://doi.org/10.1111/desc.13204

Ascher, M., & Ascher, R. (1981). *Code of the Quipu: A study in media, mathematics, and culture*. University of Michigan Press.

Australian Bureau of Statistics. (2018). *2033.0.55.001—Census of population and housing: Socio-economic indexes for areas (SEIFA), Australia, 2016*. https://www.abs.gov.au/ausstats/abs@.nsf/mf/2033.0.55.001

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition*, *119*(2), 301–306. https://doi.org/10.1016/j.cognition.2011.01.003

Best, E. (1921). Polynesian mnemonics: Notes on the use of quipus in Polynesia in former times; also some account of the introduction to the art of writing. *New Zealand Journal of Science and Technology*, *4*(2), 67–74.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180–200. https://doi.org/10.1016/j.dr.2009.05.002

Bulley, A., McCarthy, T., Gilbert, S. J., Suddendorf, T., & Redshaw, J. (2020). Children devise and selectively use tools to offload cognition. *Current Biology*, *30*(17), 3457–3464. https://doi.org/10.1016/j.cub.2020.06.035

Chatterjee, S. (2017). Record keeping through the ages. *IQ: The RIM Quarterly*, *33*(1), 40–45.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098–1120. https://doi.org/10.1111/1467-8624.00081

Clerc, J., & Miller, P. H. (2013). Utilization deficiencies and transfer of strategies in preschoolers. *Cognitive Development*, *28*(1), 76–93. https://doi.org/10.1016/j.cogdev.2012.09.003

Clerc, J., Miller, P. H., & Cosnefroy, L. (2014). Young children's transfer of strategies: Utilization deficiencies, executive function, and metacognition. *Developmental Review*, *34*(4), 378–393. https://doi.org/10.1016/j.dr.2014.10.002

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, *47*(3), 153–176. https://doi.org/10.1080/00461520.2012.696438

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959–964. https://doi.org/10.1126/science.1204529

Gilbert, S. J. (2015a). Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology*, *68*(5), 971–992. https://doi.org/10.1080/17470218.2014.972963

Gilbert, S. J. (2015b). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, *33*, 245–260. https://doi.org/10.1016/j.concog.2015.01.006

Guilford, J. P. (1975). Varieties of creative giftedness, their measurement and development. *Gifted Child Quarterly*, *19*(2), 107–121. https://doi.org/10.1177/001698627501900216

Haas, A. M. (2008). Wampum as hypertext: An American Indian intellectual tradition of multimedia theory and practice. *Studies in American Indian Literatures*, *19*(4), 77–100. https://doi.org/10.1353/ail.2008.0005

Heinrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Heisel, B. E., & Ritter, K. (1981). Young children's storage behaviour in a memory-for-location task. *Journal of Experimental Child Psychology*, *31*(2), 350–364. https://doi.org/10.1016/0022-0965(81)90023-0

Henry, J. D., Rendell, P. G., Phillips, L. H., Dunlop, L., & Kliegel, M. (2012). Prospective memory reminders: A laboratory investigation of initiation source and age effects. *The Quarterly Journal of Experimental Psychology*, *65*(7), 1274–1287. https://doi.org/10.1080/17470218.2011.651091

Kim, P. Y., & Mayhorn, C. B. (2008). Exploring students' prospective memory inside and outside the lab. *American Journal of Psychology*, *121*(2), 241–254. https://doi.org/10.2307/20445459

Klahr, D., & Chen, Z. (2011). Finding one's place in transfer space. *Child Development Perspectives*, *5*(3), 196–204. https://doi.org/10.1111/j.1750-8606.2011.00171.x

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction. *Psychological Science*, *15*(10), 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Kliegel, M., & Jäger, T. (2007). The effects of age and cue-action reminders on event-based prospective memory performance in preschoolers. *Cognitive Development*, *22*(1), 33–46. https://doi.org/10.1016/j.cogdev.2006.08.003

Mahy, C. E. V., Mazachowsky, T. R., & Pagobo, J. R. (2018). Do verbal reminders improve preschoolers' prospective memory performance? It depends on age and individual differences. *Cognitive Development*, *47*, 158–167. https://doi.org/10.1016/j.cogdev.2018.06.004

Mahy, C. E. V., Moses, L. J., & Kliegel, M. (2014). The development of prospective memory in children: An executive framework. *Developmental Review*, *34*(4), 305–326. https://doi.org/10.1016/j.dr.2014.08.001

Mazachowsky, T. R., Hamilton, C., & Mahy, C. E. V. (2021). What supports the development of children's prospective memory? Examining the relation between children's prospective memory, memory strategy use, and parent scaffolding. *Journal of Cognition and Development*, *22*(5), 721–743. https://doi.org/10.1080/15248372.2021.1939352

Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important that what gets copied. *Developmental Psychology*, *47*(4), 1050–1053. https://doi.org/10.1037/a0023866

O'Sullivan, J. T., & Pressley, M. (1984). Completeness of instruction and strategy transfer. *Journal of Experimental Child Psychology*, *38*(2), 275–288. https://doi.org/10.1016/0022-0965(84)90126-7

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Redshaw, J., Vandersee, J., Bulley, A., & Gilbert, S. J. (2018). Development of children's use of external reminders for hard-to-remember intentions. *Child Development*, *89*(6), 2099–2108. https://doi.org/10.1111/cdev.13040

Ringel, B. A., & Springer, C. J. (1980). On knowing how well one is remembering: The persistence of strategy use during transfer. *Journal of Experimental Child Psychology*, *29*(2), 322–333. https://doi.org/10.1016/0022-0965(80)90023-5

Said-Metwaly, S., Fernandez-Castilla, B., Kyndt, E., Van den Noortgate, W., & Barbot, B. (2021). Does the fourth-grade slump in creativity actually exist? A meta-analysis of the development of divergent thinking in school-age children and adolescence. *Educational Psychology Review*, *33*(1), 275–298. https://doi.org/10.1007/s10648-020-09547-9

SAS Institute Inc. (2013). *SAS 9.4 statements: Reference*.

Schmandt-Besserat, D. (1978). The earliest precursor of writing. *Scientific American*, *238*(6), 50–59. https://doi.org/10.1038/scientificamerican0678-50

Schneider, W., Tibken, C., & Richter, T. (2022). The development of metacognitive knowledge from childhood to young adulthood: Major trends and educational implications. *Advances in Child Development and Behavior*, *63*, 273–307. https://doi.org/10.1016/bs.acdb.2022.04.006

Schwenck, C., Bjorklund, D. F., & Schneider, W. (2009). Developmental and individual differences in young children's use and maintenance of a selective memory strategy. *Developmental Psychology*, *45*(4), 1034–1050. https://doi.org/10.1037/a0015597

Smith, G., Della Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory*, *8*(5), 311–321. https://doi.org/10.1080/09658210050117735

Received October 28, 2022
Revision received March 2, 2023
Accepted March 17, 2023 ■