

# Children spontaneously discover efficient solutions to a difficult sorting task

Received: 2 December 2024

Accepted: 6 August 2025

Published online: 03 October 2025

 Check for updates

Huiwen Alex Yang <sup>1</sup>, Bill D. Thompson <sup>1,2</sup> & Celeste Kidd<sup>1,2</sup>

Efficient algorithms can enhance problem-solving in many cognitive domains but can be difficult to discover and use. For example, classical studies of seriation suggest that children struggle to apply algorithmic strategies in a simple sorting problem. We investigated the spontaneous discovery of algorithmic solutions across development. We gave children a variant of the sorting problem with hidden object ranks: children sort animated bunnies into the right order, from the shortest to the tallest, when the bunnies are standing behind a wall so their heights are not visible. Children performed far above chance on this difficult sorting task, potentially because higher demands in memory and reasoning incentivized strategic behaviours. Children also independently discovered at least two efficient algorithmic solutions to the sorting problem: selection sort and shaker sort. Additionally, our developmental results show that older children were more effective sorters than younger children and used efficient sorting algorithms more frequently. These results indicate that children are far more competent at applying algorithmic solutions to sorting tasks than previous research would suggest, and performance on sorting tasks improves throughout development. Our work demonstrates that children have the ability to spontaneously organize their behaviours and find effective solutions to challenges in the world.

Imagine that you are baking a dozen cookies. One approach is to mix the ingredients and bake each individual cookie 12 times. A more efficient approach is to execute the mixing and baking processes only once for the entire batch. Both methods produce identical results, but the latter represents a more efficient strategy. Strategies are special cases of procedural knowledge that can take the forms of algorithms, which are step-by-step, formulaic techniques that, when followed properly, yield the correct solutions to problems<sup>1,2</sup>. The acquisition and use of cognitive algorithms are important because structured forms of problem-solving can enhance performance at many tasks<sup>3–5</sup>. Some example domains in which people often rely on algorithmically structured knowledge include memory strategies<sup>6–9</sup>, mathematical cognition<sup>10–12</sup> and grammatical rules<sup>10,13</sup>.

The use of strategies to accomplish cognitive tasks is affected by many domain-general cognitive abilities, such as working memory

capacity, inhibition control, cognitive flexibility and processing speed<sup>14–17</sup>. On the one hand, previous research has shown that children start to be strategic even when they are very young<sup>7,18</sup>, and they become increasingly capable of using more efficient strategies as they age<sup>8,19–21</sup>. Research has also shown that children are very competent in choosing adaptively among various algorithms, depending on specific constraints or demands of the tasks<sup>22–24</sup>. On the other hand, young children have been shown to struggle to apply systematic strategies because they have underdeveloped domain-general cognitive abilities<sup>25–28</sup>. In this paper, we provide evidence that shows children's spontaneous discovery of efficient sorting strategies, countering all conventional expectations, according to the current developmental psychology literature. Our work demonstrates that children can spontaneously organize their behaviour in a more systematic manner than previous research suggests, and that they do so to develop efficient solutions to the challenges they encounter.

<sup>1</sup>Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. <sup>2</sup>These authors contributed equally: Bill D. Thompson, Celeste Kidd.

✉ e-mail: [hw.yang@berkeley.edu](mailto:hw.yang@berkeley.edu)

The goal of this paper is to investigate the spontaneous discovery of structured algorithmic strategies by children. In particular, we examined the algorithmic structure in children's behaviours when solving a difficult sorting task. Sorting is considered a fundamental problem in computer science. It is a rich and interesting problem because it requires the use of strategies for executing an appropriate sequence of actions to achieve the correct final ordering of objects<sup>29</sup>. Sorting is the basic building block of many mathematical and programming problems. Its solutions have therefore been studied extensively, and various algorithms with different spatial and temporal efficiency can be used to solve a sorting task<sup>29</sup>. For instance, the well-known bubble sort algorithm sorts effectively by repeatedly iterating through a list, comparing adjacent elements, and swapping them if they are in the wrong order, causing larger elements to 'bubble up' to the top of the list until the entire list is sorted.

Our sorting task tests a skill that Jean Piaget called seriation. The Piagetian version of the task studies the behaviours of children by asking them to arrange a disordered set of sticks of different lengths into the correct order<sup>30,31</sup>. Since then, similar tasks have been used with children to show developmental changes in seriation abilities<sup>32–35</sup>. Children are thought to progress through predictable developmental stages before they can consistently use efficient sorting strategies. Classic work suggests that children under age seven generally struggle with applying structured and efficient strategies: young children tend to rely on unsystematic trial-and-error approaches<sup>36</sup>. Around age seven, they begin to use efficient strategies more consistently and effectively<sup>36,37</sup>. These findings contributed to the widely held assumption that efficient, algorithm-like strategies are largely beyond the grasp of young children. Seriation is an important skill for children to master because it is crucial for the development of early math skills<sup>38</sup>. For instance, it is theorized to be the foundation of the comprehension of relationships between numbers<sup>39</sup> and is predictive of the comprehension of the number line<sup>35</sup>. Learning accurate and efficient algorithms to perform the seriation task is therefore closely linked to the development of more general numerical abilities and mathematics cognition.

Our experiment used a sorting problem with hidden object ranks (Fig. 1). In this case, children were told to rank objects according to their heights, but they could not see the heights of those objects. To make the task engaging and enjoyable, we designed it to be both interactive and fun. Children, ages four to ten, were asked to sort six bunnies from the shortest to the tallest on a touch-screen computer (see Table 1 for the age group breakdown). For each participant and each trial, the initial order of the bunnies was randomly generated (see 'Difficulty' in the Supplementary Information for detailed analyses of the out-of-orderness in the initial arrangements). Children were familiarized with the sorting task by performing three practice trials in which the heights of the bunnies were visible, partially visible and not visible (Fig. 1a). For all six test trials, the heights of the bunnies were not visible (Fig. 1a). In each trial, children performed a series of pairwise comparisons by selecting two bunnies for every comparison until they thought they had achieved the correct final ordering of the bunnies. Only pairs that were out of order would actually switch positions, resulting in a swap (Fig. 1c). If the selected pair was already at the correct relative order, then they would not switch positions, resulting in a non-swap (Fig. 1b). At the end of each trial, children were given feedback regarding their accuracy and efficiency before they moved on to the next trial (see 'Experiment design' and 'Procedure' in the Methods as well as 'Experiment' in the Supplementary Information for details).

We then analysed the algorithmic structure of children's behaviour by identifying key patterns in their behaviour that are foundations of efficient sorting algorithms. Our results show that children achieved high accuracies on these sorting tasks, and they also spontaneously discovered and used at least two efficient sorting algorithms, selection sort and shaker sort (these two sorting algorithms are illustrated in Fig. 2). Our research provides new evidence of children's spontaneous

discovery of strategies. Importantly, our analysis highlights that strategic behaviour can be observed even when children's final answers are incorrect, which suggests that process-based analyses can reveal children's algorithmic thinking that might otherwise go unnoticed.

## Results

### Children are able to sort objects, even when the ordering of the objects is hidden

Our results show that all children (123 out of 123) were able to sort the bunnies in the practice trials in which they were visible (practice trials 1 and 2). This indicates that children as young as four years old have the ability to sort. Our results extend prior findings showing that young children can successfully sort small arrays in which items of the arrays are visible, demonstrating an understanding of relative size when direct comparison is possible<sup>32,40</sup>. In all test trials, we tested children's sorting abilities in a more demanding context, where the items are not visible and their relative ranks must be inferred from the outcomes of pairwise comparisons.

Figure 3a shows the distribution of participants' accuracy, separated by age group. In this figure, percentage accuracy is defined as the proportion of testing trials included in the analysis that resulted in the correct final ordering. We excluded a trial if it contained fewer than five comparisons or more than three standard deviations from the mean number of comparisons (46). As a result, 667 trials (90.38%) were kept for all following analyses (see 'Exclusion criteria' in the Methods for details). A substantial portion of children performed the sorting task correctly even when the heights of the bunnies were not visible: children showed an average 38.68% accuracy in all testing trials, notably higher than chance accuracy (19.93%; see 'Chance accuracy' in the Methods for details). It also shows that older children are more clustered towards higher accuracies. Children's accuracy results are interesting in light of children's reported failures in applying structured and efficient strategies in previous similar and simpler seriation tasks<sup>30,31</sup>. This difference may suggest that the higher memory and reasoning demands of our sorting tasks motivated children to find more efficient algorithmic solutions.

### Older children are more accurate sorters

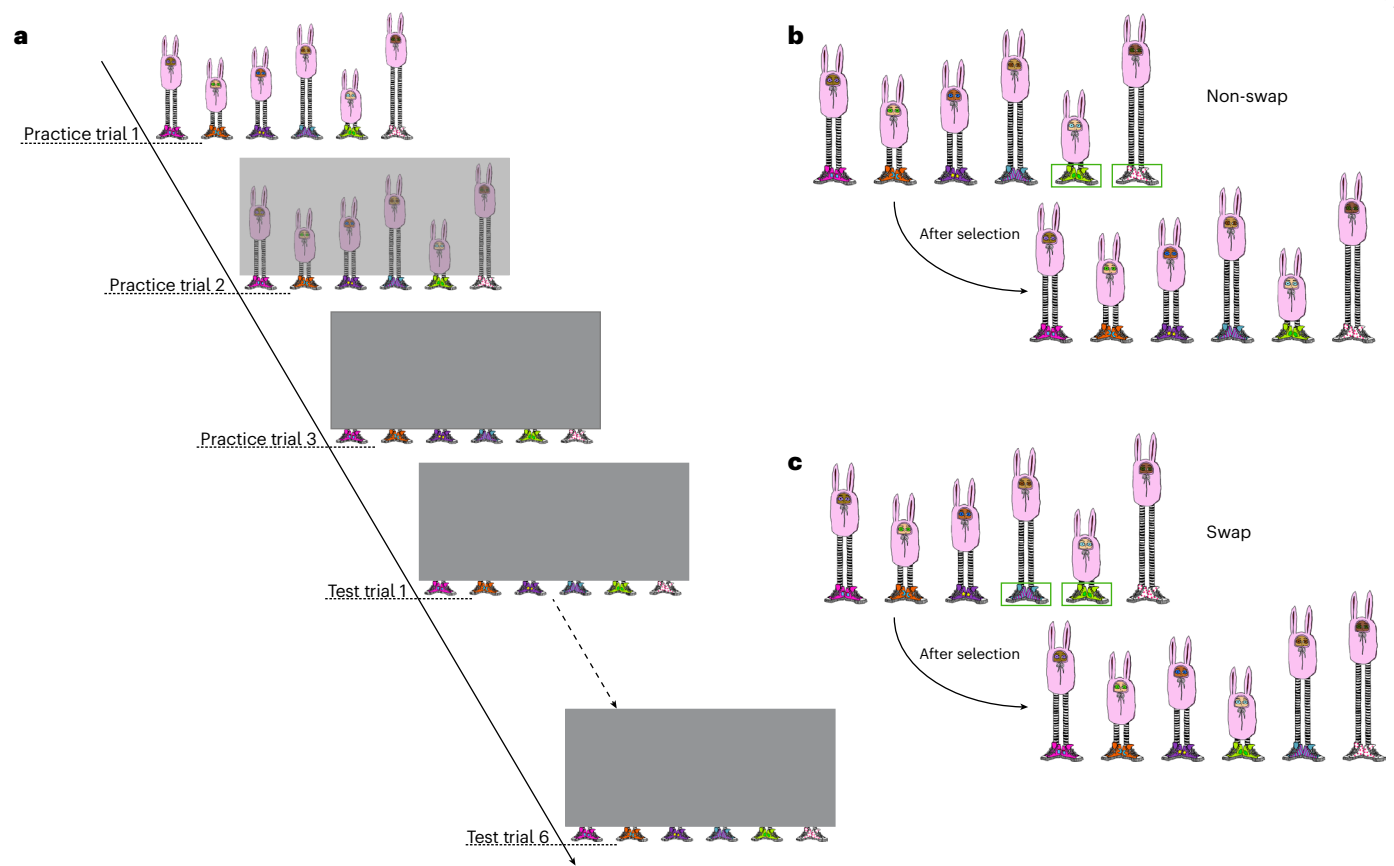
Older children performed better on our sorting task: their accuracy across six test trials is higher than that of younger children. Figure 3b shows children's percentage accuracy as a function of age. The percentage of trials that older children performed correctly is significantly greater than that of younger children ( $\beta = 0.35$ ;  $P < 0.001$ ;  $z = 4.61$ ; 95% confidence interval (CI), (0.20, 0.50)).

Figure 3c shows test trial accuracies averaged across participants and separated by age groups. It shows that participants' average sorting accuracies do not fluctuate throughout the experiment. Younger children's accuracies appear to decrease with practice, potentially due to a decrease in attention and interest over time. Older children's accuracies appear to increase with practice, suggesting that their performance on later trials may benefit from their experiences with previous trials in the experiment. The positive practice effect in older children is shown by the positive interaction between age and trial number ( $\beta = 0.19$ ;  $P = 0.023$ ;  $z = 2.27$ ; 95% CI, (0.03, 0.36); see 'Accuracy' in the Supplementary Information for more information).

It is also worth noting that older children were not using more comparisons than younger children on correctly performed test trials ( $\beta = 0.39$ ;  $P = 0.285$ ;  $t_{256} = 1.07$ ; 95% CI, (−0.32, 1.09); Bayes factor ( $BF_{10}$ ), 0.24). One possible explanation is that older children's sorting behaviours might be more efficient than those of younger children, and this allows them to have a higher accuracy while not using more comparisons.

### Older children's sorting behaviour shifts from spatial proximity to temporal proximity

To facilitate a greater understanding of the developmental results, we analysed the behavioural structure of participants' responses.



**Fig. 1 | Study design and procedure. a**, Illustration of the study procedure, including practice trials and test trials. During each trial, the participants sorted six bunnies in the correct order. In practice trial 1, the heights were fully visible;

in practice trial 2, the heights were partially visible; and in practice trial 3 and the test trials, the heights were not visible. **b**, An example of a comparison that resulted in a non-swap. **c**, An example of a comparison that resulted in a swap.

**Table 1 | Age groups and the number of participants in each age group**

| Age group | Number of participants |
|-----------|------------------------|
| 4         | 16                     |
| 5         | 18                     |
| 6         | 21                     |
| 7         | 22                     |
| 8         | 21                     |
| 9         | 25                     |

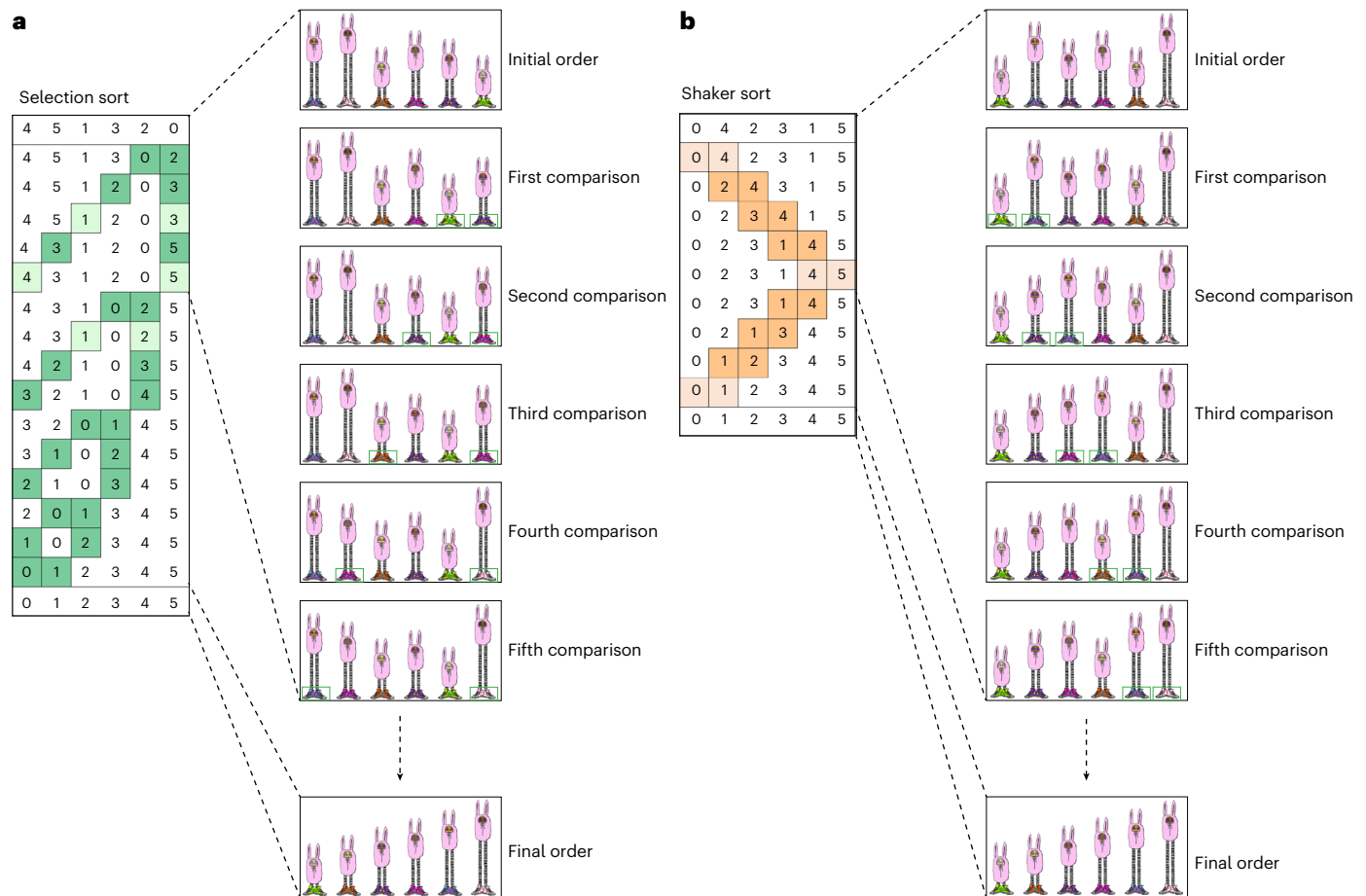
We looked at two types of comparisons participants can perform: adjacent comparisons and consecutive comparisons. An adjacent comparison is a selection in which the two bunnies chosen occupy neighbouring positions. A consecutive comparison is a selection in which one of the positions matches a position chosen in the previous comparison. Note that adjacent and consecutive comparisons are not mutually exclusive; a single comparison may qualify as both, either or neither. Both types of comparisons are essential building blocks of many identifiable sorting algorithms. For instance, if a participant is using shaker sort (details below), then they will frequently use adjacent comparisons, and if a participant is using selection sort, they will frequently use consecutive comparisons and adjacent comparisons.

These two types of comparisons, however, differ in their demands on participants' cognitive abilities. To perform adjacent comparisons, participants do not need to remember which positions they selected in the previous comparison. However, to perform consecutive

comparisons, participants have to remember at least one selected position from the previous comparison to select the same position again. Consecutive comparisons therefore have a higher demand for participants' abilities to memorize and track what they have selected or to use a strategy that encodes this regularity. Consecutive comparisons are also more demanding because they require participants to click on two bunnies that are not necessarily close to each other. Furthermore, consecutive comparisons demand a higher level of conceptual understanding of object relations in the sense that for participants to select two bunnies that are not adjacent, they would have to understand that two objects that are not placed immediately next to each other may also be related and have properties that can be compared.

We performed a generalized linear regression analysis using a binomial link function with random participant effects to assess the effect of age on the types of comparisons participants used. The results show that while age does not have a significant effect on the percentage of adjacent comparisons participants used ( $\beta = -0.04$ ;  $P = 0.457$ ;  $z = -0.74$ ; 95% CI,  $(-0.14, 0.06)$ ; Fig. 4a), older children were more likely to perform consecutive comparisons ( $\beta = 0.22$ ;  $P < 0.001$ ;  $z = 5.15$ ; 95% CI,  $(0.14, 0.30)$ ; Fig. 4b; see 'Comparison analysis' in the Supplementary Information for more information). Our results provide supporting evidence for the theory that consecutive comparisons require increased cognitive resources.

We also performed an additional set of generalized linear regression analyses using a binomial link function with random participant effects to assess whether being further into the experiment or a trial influenced the types of comparisons participants used. The results show that children were less likely to perform adjacent comparisons when they were further into a trial ( $\beta = -0.12$ ;  $P < 0.001$ ;  $z = -5.194$ ;



**Fig. 2 | Illustrations of two different sorting algorithms: selection sort and shaker sort. a**, An example of using the selection sort algorithm to solve the sorting task. **b**, An example of using the shaker sort algorithm to solve the sorting task. The first row indicates the initial order of the bunnies, and the last row indicates the order of the bunnies when the trial is finished. Every row in between

indicates a comparison, with two shaded positions being the two bunnies selected by the participant in the comparison. Squares that are darkly shaded are comparisons that resulted in a swap. Squares that are lightly shaded are comparisons that resulted in a non-swap.

95% CI,  $(-0.16, -0.074)$ ; Fig. 4c). In contrast, children were more likely to perform consecutive comparisons when they were further into a trial ( $\beta = 0.31$ ;  $P < 0.001$ ;  $z = 12.348$ ; 95% CI,  $(0.26, 0.35)$ ; Fig. 4d). Our results show that the more experience and practice children had, the more likely they were to realize the flexibility and utility of consecutive comparisons and start incorporating them into their sorting strategies. As a result, children who heavily relied on adjacent comparisons gradually transitioned to increase their use of consecutive comparisons, which may contribute to the increase in efficiency that we see in Fig. 3b.

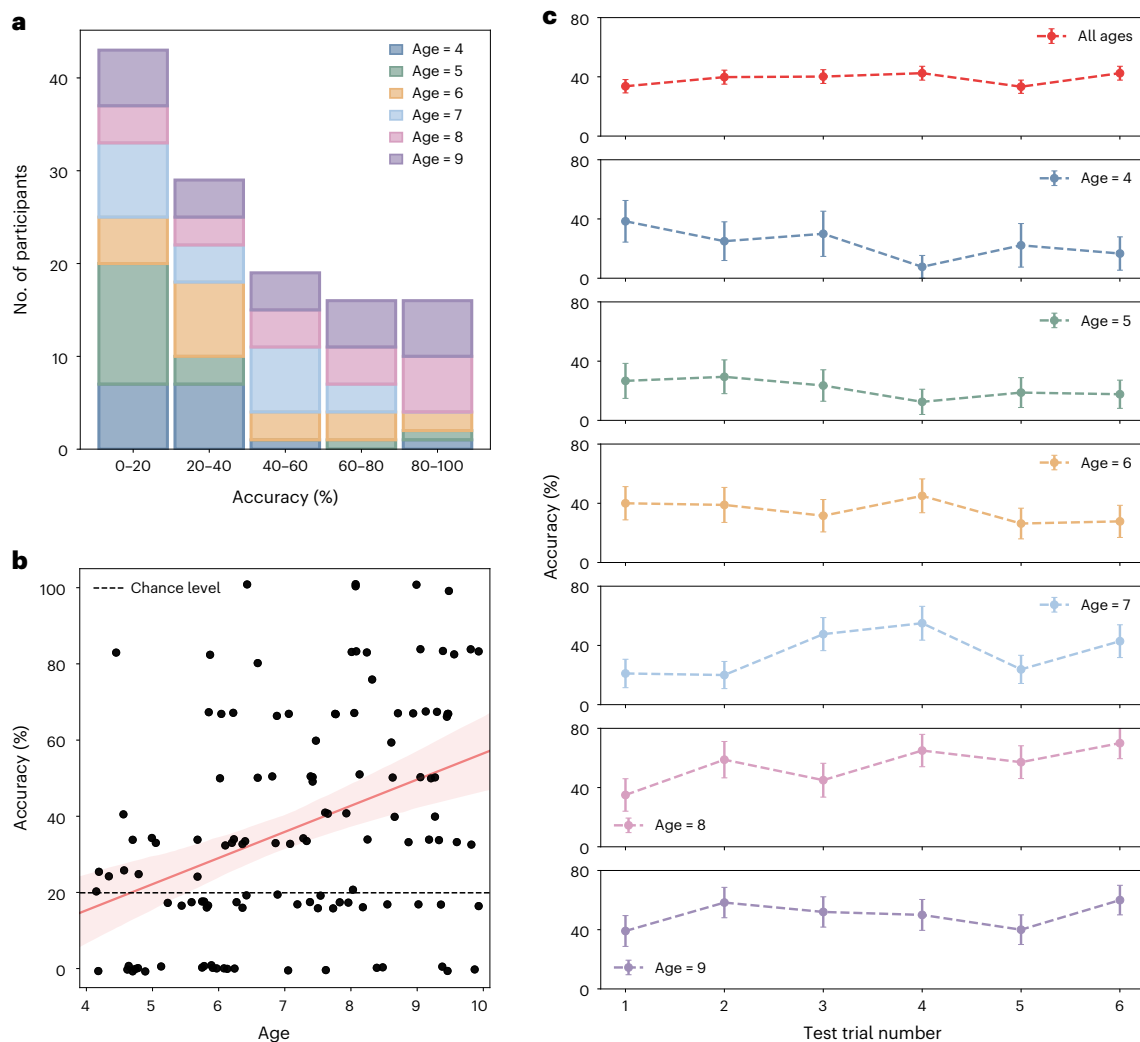
Together, these results point to behavioural differences between more efficient sorters and less efficient sorters. They also offer a possible explanation as to why older children were more efficient, and why children's efficiency increased during the experiment.

### Children discover and use various sorting algorithms

We analysed the algorithmic structure of the strategies children used to understand these behavioural differences in more detail. We performed a pattern-matching analysis on the sequences of comparisons that they performed in each trial to identify specific algorithms. As Thompson et al.<sup>41</sup> demonstrated, participants who use different algorithms exhibit different identifiable signature behavioural patterns. Children appeared to engage more frequently with exploratory behaviours, patterns of comparisons that do not align with any identifiable sorting algorithm but still reflect engagement with the task, rather

than precisely following a specific algorithm, in contrast to adults (see 'Exploratory behavior' in the Supplementary Information for some examples). It is therefore challenging to align their full sequences of comparisons with a specific algorithm without substantial errors in classification. Even when children are using algorithms in problem-solving, they also frequently deviate from the algorithms that they are using, producing incomplete implementations of strategies. We identified key signature patterns for the two sorting algorithms that children discovered to overcome this difficulty in data analysis. We classified a participant as using a particular algorithm if the participant produced a sequence of comparisons that contains the signature pattern anywhere throughout the trial.

The two most common algorithms were selection sort (also the most common algorithm used by adults) and shaker sort (less common among adults, but attested). Figure 5 shows the signature sequences of comparisons generated by using these two algorithms. Selection sort generates a fixed sequence of 15 comparisons that are guaranteed to establish the correct order. It can also be separated into forward selection sort and backward selection sort, depending on whether the participant starts from the left or the right. The left panel of Fig. 5a shows an example of the sequence of comparisons that would be generated if a participant were using the selection sort algorithm. In this figure, the top row represents the initial ordering of the bunnies, such that 0 represents the shortest bunny and 5 represents the tallest bunny. The bottom row represents the final ordering. Every row represents



**Fig. 3 | Results. a**, Distribution of participants' accuracy by age group. The stacked bar chart shows the number of participants across different accuracy ranges (0–100%) for each age group, with younger participants generally having lower accuracies and older participants having higher accuracies. **b**, Percentage accuracy as a function of age (red solid line), with each dot representing an individual participant's performance. The line is the linear regression fit with a 95% CI. A significant portion of the red line is above the chance level

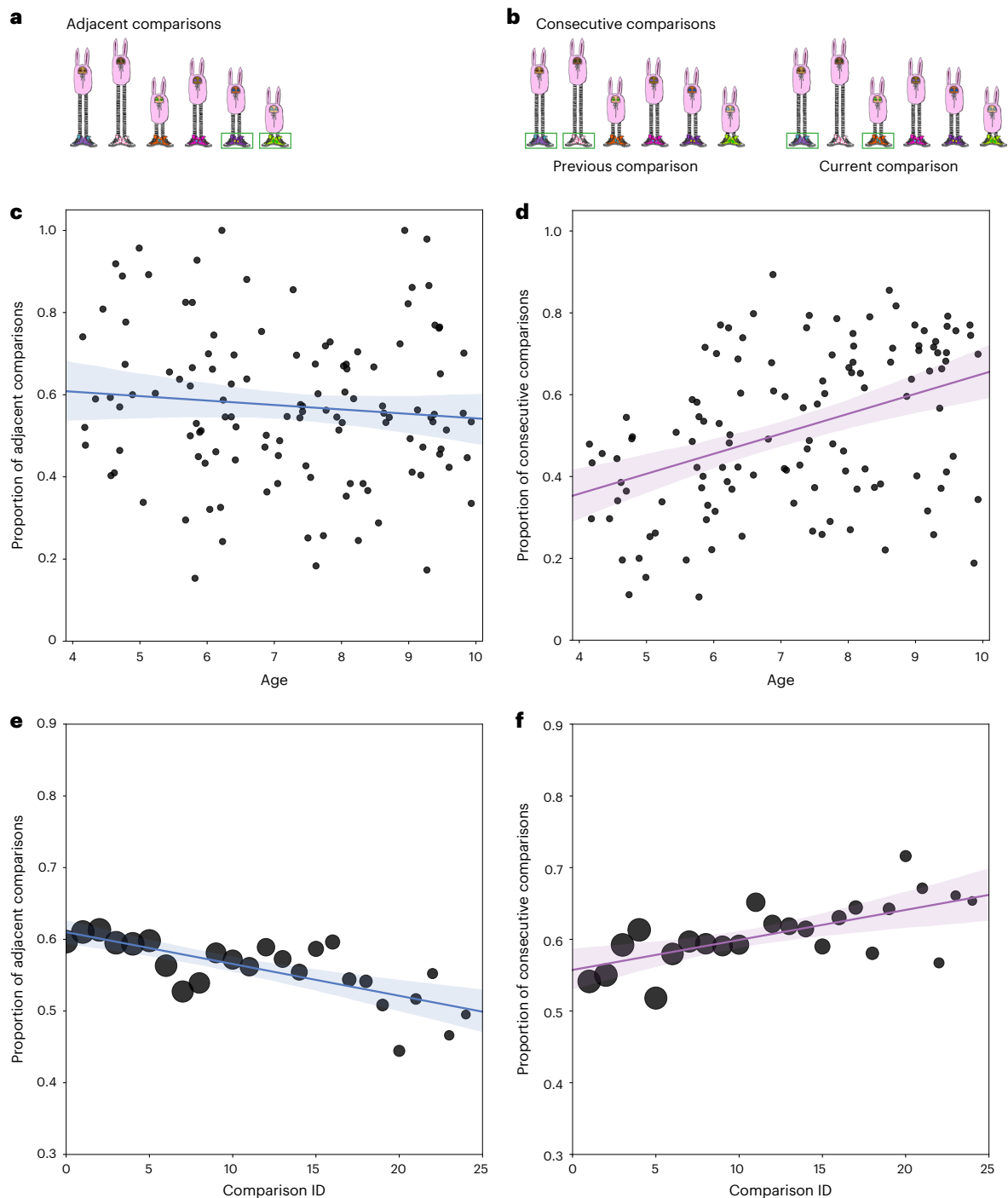
(black dashed line), indicating that children's performances are significantly better than chance. The positive slope of our fit line also shows that older children have higher accuracy. **c**, Accuracy across test trials for all age groups. The top panel shows the average accuracy trend for all participants, while the lower panels break it down by individual age groups. Each plot shows fluctuations in accuracy over six test trials, with only older children showing practice effects. The error bars represent one s.e.m.

a comparison, with two selected positions shaded. Darker positions indicate a swap, and lightly shaded positions indicate a non-swap. The numbers in every row represent the ordering of the bunnies after each comparison. The middle panel of Fig. 5a shows behavioural patterns generated by a participant when they faced a trial with the same initial orderings as shown in the left panel. Consistent with our expectation, children's actual behaviour shows redundant comparison attempts and deviations from the most efficient implementation of the algorithm, but nonetheless clear signs of algorithmic structure. The pattern we used to identify the use of selection sort is shown in the right panel of Fig. 5a.

Shaker sort is achieved by selecting two adjacent bunnies at a time and moving the selection from left to right, and then from right to left, until the participant arrives at the correct ordering (imagine shaking a horizontal bottle, side to side). It can also be separated into forward shaker sort and backward shaker sort, depending on which side the participant started at. Shaker sort also guarantees accuracy if applied for enough iterations (shakes). However, the number of iterations required does depend on the specific initial ordering of the bunnies.

Figure 5b shows an example of using the shaker sort algorithm with maximum efficiency, as well as behavioural results from a participant using this algorithm. The pattern we used to identify the use of shaker sort is shown in the right panel of Fig. 5b.

Our pattern-matching analysis shows that for a notable percentage of trials (34.48%, 230 out of 667 trials; Fig. 6), participants performed a sequence of comparisons that were consistent with them using at least one identifiable sorting algorithm. There are 110 trials in which participants' behaviour aligns with the pattern of using selection sort (chance level, 0.003%; examples shown in Fig. 6; see Supplementary Information), 141 trials in which participants' behaviour aligns with the pattern of using shaker sort (chance level, 0.13%; examples shown in Fig. 6; see Supplementary Information) and 21 trials in which participants' behaviour aligns with the pattern of using both algorithms (Table 2 shows the number of trials and participants in each algorithm category). On the participant level, our pattern-matching analysis shows that 67 participants used at least one identifiable algorithm (54.47%). Figure 6 also indicates that older children used these two identified algorithms more frequently (selection sort:  $\beta = 0.45$ ;  $P < 0.001$ ;

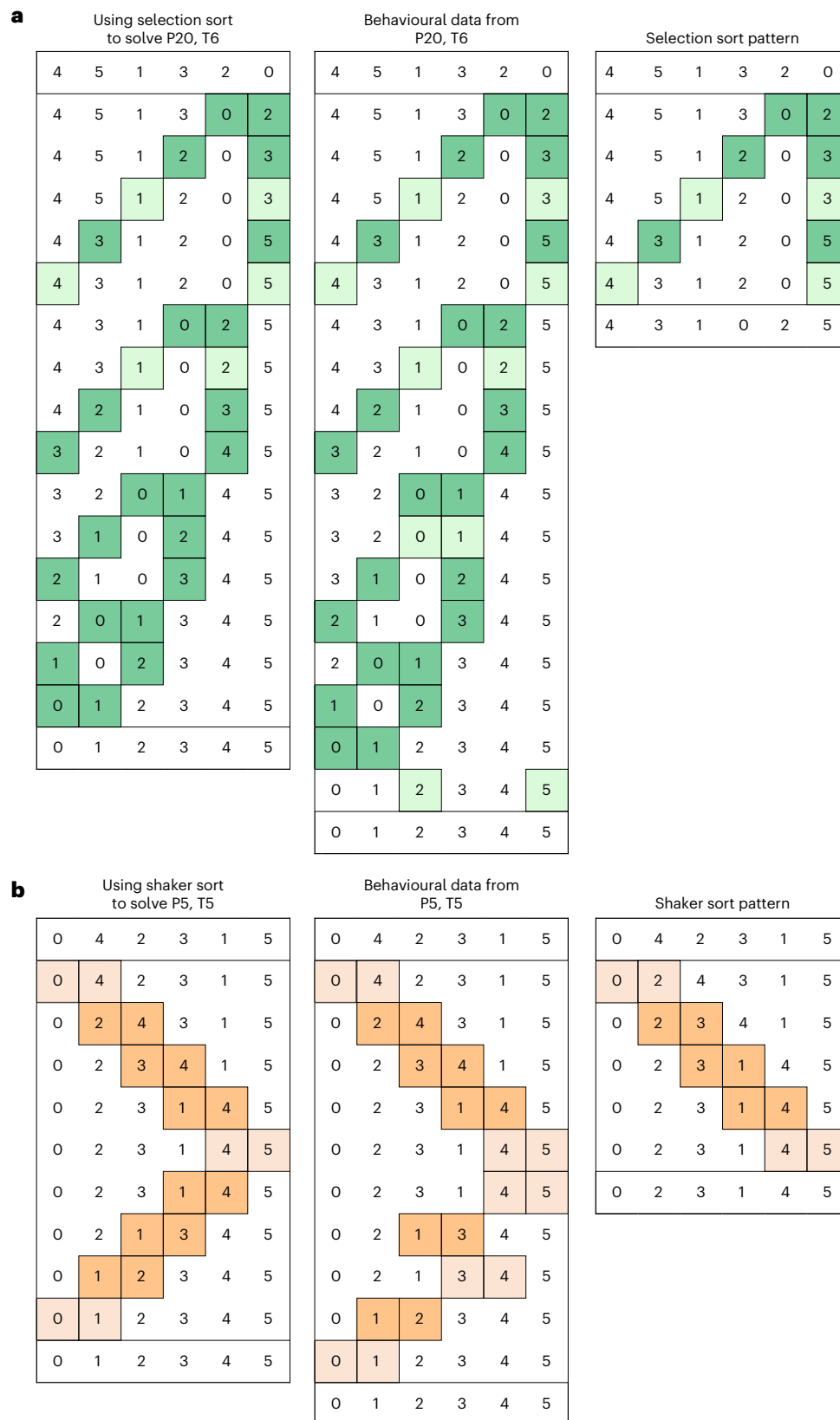


**Fig. 4 | Children's use of adjacent and consecutive comparisons during sorting.** **a**, An example of an adjacent comparison. An adjacent comparison is a selection in which the two bunnies chosen occupy neighbouring positions. **b**, An example of a consecutive comparison. A consecutive comparison is a selection in which one of the positions matches a position chosen in the previous comparison. **c**, Proportion of comparisons that are adjacent comparisons as a function of age. The line is the linear regression fit with a 95% CI. The flat trend indicates that older and younger children used similar proportions of adjacent comparisons. **d**, Proportion of comparisons that are consecutive comparisons as a function of age. The line is the linear regression fit with a 95% CI. The positive

trend demonstrates that older children used a greater proportion of consecutive comparisons. **e**, Proportion of comparisons that are adjacent comparisons as a function of comparison number. The line is the linear regression fit with a 95% CI. The size of each dot represents the number of data points. The negative trend demonstrates that the deeper participants get into a trial, the less likely they are to use adjacent comparisons. **f**, Proportion of comparisons that are consecutive comparisons as a function of comparison number. The line is the linear regression fit with a 95% CI. The size of each dot represents the number of data points. The positive trend demonstrates that the deeper participants get into a trial, the more likely they are to use consecutive comparisons.

$z = 5.92$ ; 95% CI, (0.30, 0.60); shaker sort:  $\beta = 0.30$ ;  $P < 0.001$ ;  $z = 4.74$ ; 95% CI, (0.18, 0.43)). Younger children nonetheless showed the ability to use these algorithmic solutions: even the youngest children in our participant pool (four-year-olds and five-year-olds) showed some

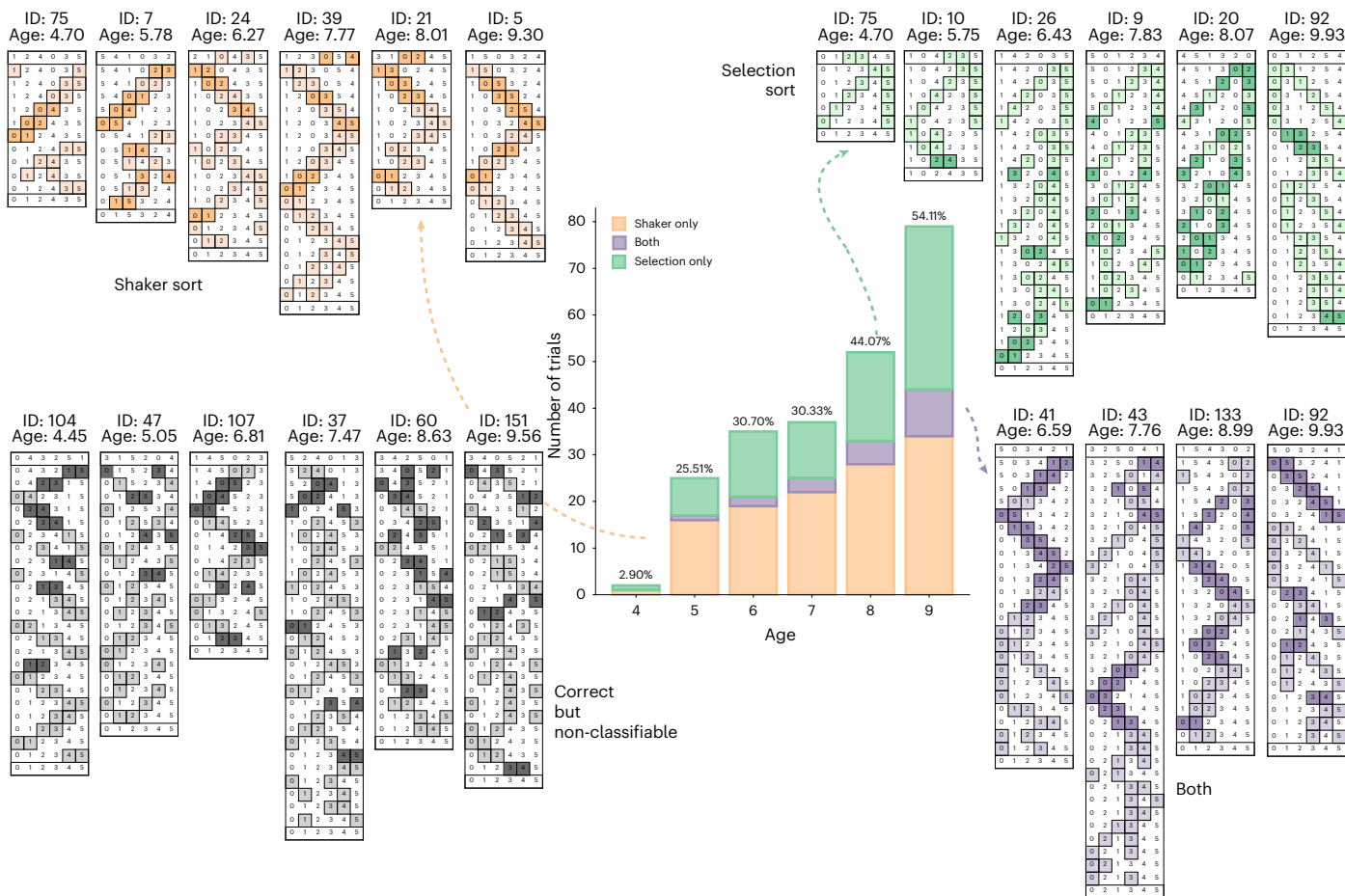
capacity to use the identified algorithms. To illustrate the range of successful but unclassified strategies, we also include a panel showing examples of correctly completed trials that were not classified by our algorithm-based analysis (Fig. 6).



**Fig. 5 | Sorting algorithms and the patterns we used to identify them.**

**a**, Selection sort and the pattern we used to identify it. Left, sequences of comparisons generated by using selection sort. Middle, behavioural pattern generated by participant 20 on test trial 6 (P20, T6) using selection sort. Right, the pattern we used to identify the use of selection sort. **b**, Shaker sort and the

pattern we used to identify it. Left, sequences of comparisons generated by using shaker sort. Middle, behavioural pattern generated by participant 5 on test trial 5 (P5, T5) using shaker sort. Right, the pattern we used to identify the use of shaker sort. Squares that are darkly shaded are comparisons that resulted in a swap. Squares that are lightly shaded are comparisons that resulted in a non-swap.



**Fig. 6 | Children's use of sorting algorithms.** Examples of children's raw behaviours as they employed different sorting algorithms. The stacked bar plot in the middle shows the number of trials identified using at least one algorithm, separated by age group and algorithm used. The percentage of trials identified using at least one algorithm for each age group is annotated at the top of each bar. Trials are categorized into three types: shaker only (orange), selection only

(green) and both (purple). The data show that older children increasingly use both strategies. The surrounding panels show illustrative examples of children employing shaker sort (orange), selection sort (green) and both (purple), alongside correctly sorted trials where no identifiable algorithm was detected (grey). The examples are arranged by age, with data from younger children on the left within each algorithm category.

**Table 2 | Pattern-matching results**

| Sorting algorithm   | Number of trials | Number of participants |
|---------------------|------------------|------------------------|
| Selection sort only | 89 (13.34%)      | 16 (13.01%)            |
| Shaker sort only    | 120 (18.00%)     | 21 (17.07%)            |
| Both                | 21 (3.15%)       | 30 (24.40%)            |
| Neither             | 437 (65.51%)     | 56 (45.53%)            |
| All                 | 667              | 123                    |

The number of trials and the number of participants that used selection sort, shaker sort, both sorting algorithms and neither.

We also found that the use of identifiable sorting algorithms positively influences children's sorting performance. Trials in which at least one of the two identifiable algorithms was used were significantly more likely to be completed correctly ( $\beta = 1.65$ ;  $P < 0.001$ ;  $z = 8.01$ ; 95% CI, (1.25, 2.05); see 'Algorithm use and performance' in the Supplementary Information for more information).

Our results indicate that not only do children perform well at this more challenging version of a sorting task, but they also independently discover and apply at least two efficient sorting algorithms. These algorithms are efficient because they are easy to implement and are guaranteed to give the correct final result,

and they both reflect behaviour that is structured, systematic and goal-oriented.

## Discussion

We show that children start to use efficient sorting algorithms earlier than previous research showed. By presenting children with a modified sorting task in which object ranks were hidden, this study demonstrates that even young children can successfully complete the task, relying on inferred rather than visible ordering. We show that older children's sorting behaviour shifted from spatial proximity to temporal proximity. We also show that accuracy improved with age, and our analysis suggests that this improvement is partly driven by an increased use of identifiable and efficient sorting algorithms. Finally, we found that a significant proportion of children employed at least one such algorithm during the task. We show that many children flexibly used at least one of the two sorting strategies across trials; notably, 30 children (24.4%) used both strategies at different points in the experiment. This indicates that algorithmic behaviour is not confined to a few particularly skilled individuals but instead reflects a broader developmental trend in spontaneous strategic behaviour.

Together, these findings contribute to our understanding of cognitive development and strategy acquisition. Beyond demonstrating children's early competence in strategy use, our findings offer new insights into the development of logical reasoning and strategy

acquisition. By identifying children's spontaneous use of structured sorting algorithms without prior instruction, our results contribute to broader theories of how goal-directed behaviour and abstract strategy use emerge in early cognitive development. Methodologically, our pattern-matching approach demonstrates how fine-grained behavioural signatures, rather than only final accuracy, can reveal underlying cognitive strategies that would otherwise go undetected. This offers a promising framework for studying other domains of early problem-solving. Practically, our results underscore the potential of introducing foundational algorithmic thinking in early childhood, even before formal instruction, which has important implications for designing educational interventions that support mathematical and computational reasoning skills from a young age.

Our experiments show that children demonstrated surprisingly high competence in these sorting tasks. In a study with adults done by Thompson et al.<sup>41</sup>, participants were asked to sort six images with hidden ranks. Participants in the asocial condition, in which they received no guiding information on efficient sorting algorithms, yielded a 64.5% average accuracy on these sorting tasks. Children in our experiment produced an average accuracy of 38.68%. These results support the theory that children start to discover and use efficient algorithms from a very young age. While Piaget<sup>30</sup> and Piaget and Inhelder<sup>31</sup> pointed out that children develop the ability to spontaneously seriate around the age of seven, our results show that spontaneous systematic sorting behaviours can be observed in children as young as four years old. Our pattern-matching analysis also provides a possible explanation.

We found that children can spontaneously discover and use systematic algorithms far younger than Piaget and subsequent psychologists thought. These instances of algorithm use potentially went undetected because they do not always yield correct performance. For example, some children showed a tendency to use selection sort but deviated from the algorithm during implementation, resulting in an error in the final ordering. Another example is children using the shaker sort algorithm but making only one pass, resulting in the final ordering sometimes being correct and sometimes incorrect, depending on the initial ordering of the bunnies. Multiple domain-general cognitive abilities, such as memory and motor capacities, might influence whether children are able to implement sorting algorithms and achieve the correct final ordering. Older children may be better able to handle the cognitive demands of more efficient strategies, such as selection and shaker sort, which involve consecutive comparisons. This trade-off between efficiency and cognitive cost may explain why we did not observe younger children using more complex algorithms, such as gnome sort, which require a higher cognitive load due to their reliance on constant backtracking, working memory demands and the need to monitor multiple conditional steps in sequence. Importantly, the presence of structured, process-driven behaviour even among younger participants highlights that foundational strategic abilities are already emerging at this stage. By shifting analytic focus from outcomes to the structure of behaviour, our approach highlights how younger children's cognition may have been underestimated when evaluated solely through accuracy. This perspective provides a richer view of early algorithmic thinking and opens new directions for understanding how cognitive strategies emerge and stabilize across development.

Our classification method was designed to detect meaningful structure while allowing for partial algorithm use, which is especially important when behaviour is noisy. As a result, our pattern-matching analysis allowed us to identify participants who used identifiable algorithms, even when they did not produce the correct final ordering from a complete implementation of the algorithm. For example, trials were classified as selection sort with just five consistent comparisons, rather than requiring full execution. Our motivation to apply a flexible classifier was that for a participant to produce the signature pattern to the degree of accuracy required by our analysis, the participant must have a working understanding of the structure of the algorithm and the

sequences of procedures it implies. While we did not test how relaxing criteria might affect classification rates, especially in younger children, this is a promising direction for future work. It could help determine whether age differences reflect execution noise or access to different strategies, as proposed by previous research<sup>21</sup>.

One limitation of our current approach is that we focused on two specific, easily identifiable sorting algorithms. Other plausible strategies, such as gnome sort (which has been observed in adult behaviour), may also be present in children's behaviour but are more difficult to detect reliably using our current pattern-matching method, especially given the short trial lengths and variable implementation typical of younger participants. As a result, non-classifiable trials, such as those presented in Fig. 6 may reflect either unrecognized but structured behaviour or genuinely inconsistent strategies. Future work could expand the analytic framework to include a broader set of algorithmic patterns or use unsupervised techniques to better characterize the full range of children's sorting behaviours.

Another limitation of the current study is that we cannot directly test whether increased task demands promote more strategic behaviour. While our findings are consistent with this idea, future work should systematically vary task demands and apply similar analytic methods to determine how information constraints shape children's spontaneous strategy use.

Previous research has shown that infants can seriate only under particular circumstances and only after they have observed a demonstration<sup>42–44</sup>. Studies have also shown that young children's seriation performances can be improved through training<sup>45–47</sup>. However, these training methods all involve letting children observe the proper procedure of putting a set of objects in order, and they work best when children can already seriate, to some degree, before the training.

Our experiment makes a unique contribution in showing that children can spontaneously discover and use efficient sorting algorithms, without receiving guidance or observing demonstrations of these sorting algorithms. Here we use the term 'discovery' to refer to children's spontaneous, self-guided generation of structured strategies in the absence of explicit instruction. However, children's learning of algorithms needs to be incentivized by particular circumstances. For instance, our sorting task adds cognitive demands that require children to have a conceptual understanding of the task by removing visual access to the ranks of objects. This characteristic might motivate children to use more structured sorting behaviours that reduce cognitive demands and thus promote the discovery of algorithmic sorting solutions. Since seriation is a fundamental building block for many early math abilities and seriation performance is predictive of the development of future math skills, our work has important implications for educational interventions that aim to promote young children's math abilities.

## Methods

### Participants

All experiments were approved by the institutional review board of the University of California at Berkeley under IRB protocol 2018-12-11653: Learning, Attention, and Decision-Making Throughout Development. All participants gave informed consent. A convenience sample of 154 children, ages four to nine, participated in this study in a museum, a lab or a classroom setting. We removed 31 participants from the analysis due to technical errors or failure to complete the experiment. The analysis was done on the remaining 123 participants (mean age, 7.21; s.d., 1.64). Table 1 shows the age group breakdown.

### Experiment design

We designed a sorting game that children can play on a touch-screen computer. The game contains three practice trials and six test trials. Figure 1 shows our experimental design. In every trial, children see six animated bunnies. The bunnies have different heights, and the

objective is to sort them into the correct order, with the shortest bunny on the left and the tallest bunny on the right. In addition to having different heights, the bunnies all wear different coloured shoes. To prevent order effects, both the initial arrangement and the shoe pattern of the bunnies were randomly generated for each participant and each trial.

During each trial, children performed a sequence of comparisons until they were satisfied with the ordering. In each selection, they chose two bunnies, and the two selected bunnies performed a pairwise comparison. If the bunnies were out of relative rank order (that is, the leftmost selected bunny was taller than the rightmost selected bunny), they changed places, resulting in a swap, as shown in Fig. 1c. If the selected bunnies were already in the right relative order, as shown in Fig. 1b, with the shorter one on the left and the taller one on the right, then their positions did not change, resulting in a non-swap. The number of comparisons used on a given trial is the sum of swaps and non-swaps.

We used six-item arrays to strike a balance between complexity and feasibility. This format introduces enough of a challenge to engage children's reasoning, while remaining manageable in our sequential comparison paradigm where item heights are not visible and must be inferred through interaction.

### Analysis

For all regression analyses, we report effect sizes appropriate to the model type. Specifically, for logistic regression models, we report log odds ratios as the effect sizes, reflecting the change in the log odds of the outcome per unit increase in the predictor. For linear regression models, we report unstandardized  $\beta$  coefficients to facilitate intuitive interpretation of the magnitude and direction of effects.

### Procedure

In the first practice trial, children could see the bunnies, as shown in the first panel in Fig. 1a. They advanced to the next practice trial only after they successfully put the bunnies in the right order. However, children received no instructions on which bunnies they should select to achieve the goal. In the second practice trial, depicted in the second panel in Fig. 1a, children could see the bunnies standing behind a grey glass window. They could see the different coloured shoes fully, but they could not see the heights of the bunnies clearly. The purpose of these two practice trials was to familiarize children with the demand and the goal of a sorting task. The second practice trial also prepared them for the following trials, in which they could not see the heights of the bunnies at all. In the final practice trial and all the test trials, as shown in the remaining panels in Fig. 1a, children were not able to see the bunnies' different heights. They were told that these bunnies were standing behind a grey curtain and that they could only see the bunnies' shoes. At the end of each trial, children verbally signalled to the experimenter when they thought the bunnies were in the right order, and they received a summary of whether the bunnies were in the correct final order and the number of comparisons they had performed in the trial.

There were two main reasons for allowing children to verbally signal the experimenter to end a trial, rather than using an automated system that would terminate the trial once the bunnies were correctly sorted. First, we wanted children to feel free to end a trial whenever they felt it was appropriate—even if the bunnies were not yet in the correct order—and to avoid pressuring them to continue making comparisons simply because the program had not yet signalled completion. Second, we aimed to encourage children to reflect on their own sorting behaviour, including their decisions about when to stop. This approach allowed us to observe whether children would continue applying a correct sorting algorithm—even after the bunnies appeared to be sorted—demonstrating an understanding that the algorithm ensures the correct final order regardless of the initial configuration.

### Exclusion criteria

An exclusion criterion was applied to remove trials in which children either appeared to exert insufficient effort to solve the task or used an unusually high number of comparisons relative to their peers. The principle underlying this exclusion was established prior to data collection, while the specific numerical threshold was determined on the basis of the observed data.

We excluded a trial if it contained fewer than five comparisons or more than three standard deviations from the mean number of comparisons (46). This exclusion criterion was chosen because the minimum reliable solution to a pairwise comparison sorting problem of six items requires ten comparisons, and we consider a trial invalid when the child participant did not use at least half of the necessary minimum number of comparisons. The mean length across all participants across all trials without exclusion is 15.82. The standard deviation of the length across all participants across all trials without exclusion is 10.29. The cut-off is 46 comparisons. After we applied this exclusion criterion, 71 trials were excluded, and 667 trials (90.38%) were kept for all the following analyses.

### Chance accuracy

We conducted an exploratory analysis to assess the reliability of our results on performance accuracy and to ensure that the observed accuracies were not simply due to children randomly selecting pairs of bunnies for each comparison. We calculated a baseline chance accuracy for all participants as the probability that a participant performs correctly on a trial by randomly selecting two bunnies in every comparison. Chance accuracy is calculated by using a permutation test: for every valid trial, given the initial positions of the bunnies and the number of comparisons performed by the participant during this trial, we calculated the chance of a participant achieving a correct final ordering when they perform every comparison by randomly selecting two bunnies. The chance accuracy of 1,000 permutations per trial is 19.93%. This means that, on average, if participants were selecting pairs of bunnies randomly, their accuracy should be 19.93%. Given that the average percentage accuracy for all participants is 38.68%, higher than the chance accuracy, we show that participants were not randomly selecting pairs of bunnies.

### Efficiency

We denote efficiency as the number of comparisons (swap and non-swap) that a participant performed in a trial.

### Algorithm identification

We selected two identifiable and efficient sorting algorithms: selection sort and shaker sort. We used a pattern-matching analysis method to identify whether a trial used one, both or neither of these two algorithms.

The selection sort algorithm sorts an array by repeatedly selecting the smallest (or largest) element from the unsorted portion. The pattern we used to identify the use of selection sort is a five-comparison pattern involving the participant establishing either the smallest element or the largest element of the list by repeatedly comparing the leftmost or the rightmost bunny with all other bunnies. A trial was categorized as using selection sort if the participant's comparison sequence for a trial contained this pattern. To allow for partial or imperfect implementations, especially among younger children, we chose a five-comparison threshold for detecting each algorithm. This threshold captures structured patterns that are unlikely to arise by chance, without requiring a full execution of the algorithm. While stricter or more lenient criteria could affect classification rates, especially developmentally, we opted for a balanced approach to enable meaningful yet flexible detection of strategy use.

The shaker sort algorithm sorts an array by selecting two consecutive elements at a time and moving the selection from left to right and then from right to left until the participant arrives at the correct

ordering. The pattern we used to identify the use of shaker sort is a five-comparison pattern involving the participant selecting two adjacent bunnies either from left to right or from right to left. A trial was categorized as using selection sort if the participant's comparison sequence for a trial contained this pattern.

Note that a trial can be categorized as using both of these two sorting algorithms. Our analysis focused on two algorithms with clear and compact behavioural signatures that are detectable in short trials. Other algorithms, such as gnome sort, may be more difficult to distinguish reliably in child behaviour due to shorter sequences and more variable implementation, and were therefore not included in this initial classification scheme.

Our pattern-matching analysis results are shown in Table 2. They show that 34.48% of trials and 54.47% of participants used at least one of these sorting algorithms. Our results also show that, in general, participants used shaker sort more frequently than selection sort. We also show that a significant percentage of participants can use both sorting algorithms, providing supportive evidence for the variety of strategies that children could employ to solve the sorting task.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

The data from all experiments are available at <https://osf.io/sepzw/>.

### Code availability

The code for the analyses is available at <https://osf.io/sepzw/>.

### References

- Alexander, P. A., Graham, S. & Harris, K. R. A perspective on strategy research: progress and prospects. *Educ. Psychol. Rev.* **10**, 129–154 (1998).
- Chi, M. T. H. in *Thinking and Learning Skills* Vol. 2 (eds Chipman, S. F., Segal, J. W. & Glaser, R.) 457–483 (Routledge, 2013).
- Pressley, M. et al. A primer of research on cognitive strategy instruction: the important issues and how to address them. *Educ. Psychol. Rev.* **2**, 1–58 (1990).
- Bjorklund, D. F. (ed.) *Children's Strategies: Contemporary Views of Cognitive Development* (Taylor and Francis, 1990).
- Harris, K. R. & Graham, S. *Making the Writing Process Work: Strategies for Composition and Self-Regulation* 2nd edn, Cognitive Strategy Training Series (Brookline Books, 1996).
- Ornstein, P. A. & Naus, M. J. Effects of the knowledge base on children's memory strategies. *Adv. Child Dev. Behav.* **19**, 113–148 (1985).
- Wellman, H. M. in *Memory Development: Universal Changes and Individual Differences* (eds Weinert, F. E. & Perlmutter, M.) 3–29 (Lawrence Erlbaum Associates, 1988).
- Baker-Ward, L., Ornstein, P. A. & Holden, D. J. The expression of memorization in early childhood. *J. Exp. Child Psychol.* **37**, 555–575 (1984).
- Ornstein, P. A., Haden, C. A. & San Souci, P. in *Learning and Memory: A Comprehensive Reference* Vol. 3 (ed. Byrne, J. H.) 715–744 (Academic Press, 2008).
- Anderson, J. R. *The Architecture of Cognition* (Psychology Press, 1996).
- Braithwaite, D. W., Pyke, A. A. & Siegler, R. S. A computational model of fraction arithmetic. *Psychol. Rev.* **124**, 603–625 (2017).
- Fuson, K. C. et al. Children's conceptual structures for multidigit numbers and methods of multidigit addition and subtraction. *J. Res. Math. Educ.* **28**, 130–162 (1997).
- Byrnes, J. P. The conceptual basis of procedural learning. *Cogn. Dev.* **7**, 235–257 (1992).
- Imbo, I. & Vandierendonck, A. The development of strategy use in elementary school children: working memory and individual differences. *J. Exp. Child Psychol.* **96**, 284–309 (2007).
- Hodzik, S. & Lemaire, P. Inhibition and shifting capacities mediate adults' age-related differences in strategy selection and repertoire. *Acta Psychol.* **137**, 335–344 (2011).
- Clerc, J., Leclercq, M., Paik, J. & Miller, P. H. Cognitive flexibility and strategy training allow young children to overcome transfer-utilization deficiencies. *Cogn. Dev.* **57**, 100997 (2021).
- Hedge, C., Powell, G., Bompas, A. & Sumner, P. Strategy and processing speed eclipse individual differences in control ability in conflict tasks. *J. Exp. Psychol. Learn. Mem. Cogn.* **48**, 1448–1469 (2022).
- Wellman, H. M., Ritter, K. & Flavell, J. H. Deliberate memory behavior in the delayed reactions of very young children. *Dev. Psychol.* **11**, 780–787 (1975).
- Gholson, B. *The Cognitive-Developmental Basis of Human Learning: Studies in Hypothesis Testing* Developmental Psychology Series (Academic Press, 1980).
- Ornstein, P. A., Baker-Ward, L. & Naus, M. J. in *Memory Development: Universal Changes and Individual Differences* (eds Weinert, F. E. & Perlmutter, M.) 31–50 (Lawrence Erlbaum Associates, 1988).
- Siegler, R. S. *Emerging Minds: The Process of Change in Children's Thinking* (Oxford Univ. Press, 1998).
- Shrager, J. & Siegler, R. S. SCADS: a model of children's strategy choices and strategy discoveries. *Psychol. Sci.* **9**, 405–410 (1998).
- Siegler, R. S. The perils of averaging data over strategies: an example from children's addition. *J. Exp. Psychol. Gen.* **116**, 250–264 (1987).
- Crowley, K., Shrager, J. & Siegler, R. S. Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Dev. Rev.* **17**, 462–489 (1997).
- Reynolds, M. R., Niileksela, C. R., Gignac, G. E. & Sevilano, C. N. Working memory capacity development through childhood: a longitudinal analysis. *Dev. Psychol.* **58**, 1254–1263 (2022).
- Bell, J. A. & Livesey, P. J. Cue significance and response regulation in 3 to 6 year old children's learning of multiple choice discrimination tasks. *Dev. Psychobiol.* **18**, 229–245 (1985).
- Huizinga, M. & Van Der Molen, M. W. Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Dev. Neuropsychol.* **31**, 193–215 (2007).
- Kail, R. in *Advances in Child Development and Behavior* Vol. 23 (ed. Reese, H. W.) 151–185 (Academic Press, 1991).
- Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. (eds) *Introduction to Algorithms* 3rd edn (MIT Press, 2009).
- Piaget, J. *The Child's Conception of Number* (W. W. Norton, 1965).
- Piaget, J. & Inhelder, B. *The Psychology of the Child* (Basic Books, 1969).
- McGonigle Chalmers, M. & Kusel, I. The development of size sequencing skills: an empirical and computational analysis. *Monogr. Soc. Res. Child Dev.* **84**, 7–202 (2019).
- Chapman, M. *Constructive Evolution: Origins and Development of Piaget's Thought* (Cambridge Univ. Press, 1988).
- Kingma, J. The range of seriation training effects in young kindergarten children. *Contemp. Educ. Psychol.* **11**, 276–289 (1986).
- Kingma, J. & Reuvelkamp, J. The construction of a developmental scale for seriation. *Educ. Psychol. Meas.* **44**, 1–23 (1984).
- Inhelder, B. & Piaget, J. *The Growth of Logical Thinking: From Childhood to Adolescence* (Basic Books, 1958).
- Bjorklund, D. F. & Harnishfeger, K. K. Developmental differences in the mental effort requirements for the use of an organizational strategy in free recall. *J. Exp. Child Psychol.* **44**, 109–125 (1987).

38. Schminke, C. W., Maertens, N. & Arnold, W. *Teaching the Child Mathematics* (Dryden, 1973).
39. Ginsburg, H. *Children's Arithmetic: The Learning Process* (D. Van Nostrand, 1977).
40. Koslowski, B. Quantitative and qualitative changes in the development of seriation. *Merrill Palmer Q. Behav. Dev.* **26**, 391–405 (1980).
41. Thompson, B., Van Opheusden, B., Sumers, T. & Griffiths, T. L. Complex cognitive algorithms preserved by selective social learning in experimental populations. *Science* **376**, 95–98 (2022).
42. Calvert, S. L., Richards, M. N. & Kent, C. C. Personalized interactive characters for toddlers' learning of seriation from a video presentation. *J. Appl. Dev. Psychol.* **35**, 148–155 (2014).
43. Frigaszy, D. M., Galloway, A. T., Johnson, P. J. & Brakke, K. The sources of skill in seriating cups in children, monkeys and apes. *Dev. Sci.* **5**, 118–131 (2002).
44. Greenfield, P. M., Nelson, K. & Saltzman, E. The development of rulebound strategies for manipulating seriated cups: a parallel between action and grammar. *Cogn. Psychol.* **3**, 291–310 (1972).
45. Coxford, A. F. The effects of instruction on the stage placement of children in Piaget's seriation experiments. *Arithmetic Teach.* **11**, 4–9 (1964).
46. Kidd, J. K., Pasnak, R., Gadzichowski, M., Ferral-Like, M. & Gallington, D. Enhancing early numeracy by promoting the abstract thought involved in the oddity principle, seriation, and conservation. *J. Adv. Acad.* **19**, 164–200 (2008).
47. Kingma, J. Training of seriation in young kindergarteners. *J. Genet. Psychol.* **148**, 167–181 (1987).

## Acknowledgements

We thank C. Cooke, K. Murugesu, S. Benedict, I. Ha, K. Avalos-De La Cruz, A. Dhakal and H. Palmeri for their assistance with data collection; members of the Kidd Lab and the Computational Cognitive Science Lab for helpful discussions; and the children and families who made this research possible. This work was supported by the Walton Family Foundation (award no. 00106041, C.K.). The funders had no role in

study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

All authors jointly designed the study. H.A.Y. implemented the computerized experiment, collected the data and analysed the data. H.A.Y. wrote the initial draft, and all authors edited subsequent drafts. All authors approved the final paper for publication.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41562-025-02302-6>.

**Correspondence and requests for materials** should be addressed to Huiwen Alex Yang.

**Peer review information** *Nature Human Behaviour* thanks Shari Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

|                                     |                                                                                                                                                                                                                                                                                                |
|-------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| n/a                                 | Confirmed                                                                                                                                                                                                                                                                                      |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( <i>n</i> ) for each experimental group/condition, given as a discrete number and unit of measurement                                                                                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly                                                                                                                                    |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>                                                               |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of all covariates tested                                                                                                                                                                                                                     |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons                                                                                                                                        |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i> ) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted<br><i>Give P values as exact values whenever suitable.</i>                     |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings                                                                                                                                                                      |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes                                                                                                                                                |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i> ), indicating how they were calculated                                                                                                                                               |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

|                 |                                                                                                                                                                                                                 |
|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data collection | Data was collected on a touchscreen computer using custom PsychoPy code. The open-source code base for our data collection procedures is available at <a href="https://osf.io/sepzw/">https://osf.io/sepzw/</a> |
| Data analysis   | The open-source code base for our data analysis procedures is available at <a href="https://osf.io/sepzw/">https://osf.io/sepzw/</a>                                                                            |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

- All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
  - A description of any restrictions on data availability
  - For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data we collected and analyzed is available at <https://osf.io/sepzw/>

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

### Reporting on sex and gender

Sex and gender were not considered in the study design, and were not entered in any analyses because the authors cannot identify any a priori reasons why gender would affect the cognitive processes of interest. Participants optionally provided gender data. Participants were 51% female, 45% male, 2% other and 2% gender not declared.

### Reporting on race, ethnicity, or other socially relevant groupings

Participants' guardians optionally reported their child's race and ethnicity according to the groupings used in the US census. This data was not analyzed. Sample was 52% White, 20% Asian, 2% Black or African American, 19% Multiracial, and 8% Other or unidentified. 17% identified as Hispanic, of any race.

### Population characteristics

Participants were aged 4-9 years.

### Recruitment

Samples were collected from families in the San Francisco Bay Area. Participants were recruited through one of three methods: at a local children's museum, in a research laboratory, or within a classroom setting.

### Ethics oversight

The study was approved by the Institutional Review Board of University of California, Berkeley.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences

☒ Behavioural & social sciences

☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

### Study description

The study is quantitative and experimental.

### Research sample

Our research samples were children aged 4-9 from the San Francisco Bay Area. The gender and racial demographics of our samples are broadly representative of the local population (see above).

### Sampling strategy

We obtained a convenience sample of 154 children aged 4 to 9 years from families in the San Francisco Bay Area. Participants were recruited through one of three methods: at a local children's museum, in a research laboratory, or within a classroom setting. To ensure the reliability of age-related comparisons, we included at least 20 valid samples for each age group of interest. This criterion helps mitigate biases and improves the reliability of any observed age-related trends or patterns.

### Data collection

Data was collected on a touchscreen computer. The participants touched the screen to provide their responses. If parents were present, they were seated out of view of the screen or instructed not to interfere with the study.

### Timing

Data were collected over multiple recruitment sessions between 04/23/2023 and 07/27/2024.

### Data exclusions

31 participants were removed from the analysis due to technical errors or failure to complete the experiment.

### Non-participation

7 participants did not complete the experiment.

### Randomization

Participants were not allocated into different experimental groups.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials &amp; experimental systems

|                                     |                                                        |
|-------------------------------------|--------------------------------------------------------|
| n/a                                 | Involved in the study                                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                    |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                 |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                        |

## Methods

|                                     |                                                 |
|-------------------------------------|-------------------------------------------------|
| n/a                                 | Involved in the study                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Plants

## Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

## Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

## Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.