nature human behaviour

Article

https://doi.org/10.1038/s41562-024-01992-8

Exposure to detectable inaccuracies makes children more diligent fact-checkers of novel claims

Received: 25 October 2023

Accepted: 25 August 2024

Published online: 10 October 2024

Check for updates

Evan Orticio[1], Martin Meyer[1,2] & Celeste Kidd[1]

How do children decide when to believe a claim? Here we show that children fact-check claims more and are better able to catch misinformation when they have been exposed to detectable inaccuracies. In two experiments ($N = 122$), 4–7-year-old children exposed to falsity (as opposed to all true information) sampled more evidence before verifying a test claim in a novel domain. Children's evidentiary standards were graded: fact-checking increased with higher proportions of false statements heard during exposure. A simulation suggests that children's behaviour is adaptive, because increased fact-checking in more dubious environments supports the discovery of potential misinformation. Importantly, children were least diligent at fact-checking a new claim when all prior information was true, suggesting that sanitizing children's informational environments may inadvertently dampen their natural scepticism. Instead, these findings support the counterintuitive possibility that exposing children to some nonsense may scaffold vigilance towards more subtle misinformation in the future.

Children have unprecedented access to information on their phones and computers. This fact represents a very recent shift—one of both promise and problem. The internet leaves users exposed to unprecedented amounts of misinformation. Exposure to misinformation can lead to the long-term adoption of false beliefs in both adults[1] and children[2]. This is true even when the learner is aware of this bias[3]. Misinformation exposure is also expected to increase with the widespread adoption of generative artificial intelligence models such as ChatGPT and Bard. When these models produce fabricated information in their outputs, they can transmit them to users[4]. Children are probably most vulnerable because they have less world knowledge[5] and are biased to trust information[6], particularly under conditions of uncertainty[7]. Indeed, even when preschoolers directly observe data that conflict with testimony, they rarely seek additional data and struggle to disregard the misleading information[8]. Despite these unique vulnerabilities, the overwhelming majority of work on misinformation centres on adults[9].

What we know of children's media habits suggests that they are immersed. A third of American children are on at least one social media platform by 9 years of age[10]. A majority of American teens get their news from social media or YouTube[11]. Also, children who have used ChatGPT for schoolwork report using it in place of traditional search[12]. Thus, children's media diets are rife with dubious sources. How do we best prepare children to navigate this complex informational sea?

The preeminent solution has been to shield children from misinformation via sanitized platforms. YouTube Kids, for example, curates a small selection of child-focused content through a combination of automated filters and human review[13]. This solution is limited by its reactive nature. As an example, YouTube Kids received widespread criticism when a Guardian article reported on a multitude of videos featuring themes that were not appropriate for children (for example, violent and sexual situations) that were inaccurately labelled as child-friendly by the platform's filters, probably because they contained characters from children's movies and shows[14,15]. Efforts to sanitize content for children are resource intensive and subjective (who decides what is age-appropriate?). Moreover, automated curation approaches are

[1]Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. [2]Department of Psychology, Yale University, New Haven, CT, USA.
e-mail: eorticio@gmail.com

easily gamed, and human curation approaches cannot scale as rapidly as new content is produced[16].

Another proposed strategy for safeguarding people from misinformation comes from inoculation theory (for reviews, see refs. 17–19). Inspired by an analogy to biomedical inoculation, the theory postulates that preemptively exposing learners to a weakened form of a misleading argument can confer immunity to its persuasiveness later on. This process involves 'prebunking' the argument by refuting false information in advance, and/or deconstructing misleading argumentation techniques more broadly. Researchers claim to have successfully inoculated adults against misinformation spanning many topics[20,21], including climate change[22], vaccination[23] and extremist ideology[24]. However, inoculation interventions are fragile, ephemeral and difficult to scale[25]. Inoculation interventions are short-acting, and ineffective after 48 h without direct and immediate reinforcement[26]. These interventions have also been criticized for fatal methodological weaknesses in the assessment of their efficacy[27–29]. For example, a recent analysis found no evidence that inoculation improves discrimination, but rather that it induces a potentially counterproductive, negative response bias[30].

Further, misinformation inoculation techniques are rarely tested in young children—and there is reason to expect they may not achieve even the modest, ephemeral effects seen in adults because of differences in children's decision-making and metacognition. Children have less developed metacognitive skills than adults, a tendency towards overconfidence[31–33] and less executive function[34]. Thus, it may be more difficult to find interventions that effectively lower overconfidence and slow decision-making about factual accuracy for children.

Despite this, research suggests that children can accurately assess the epistemic quality of human sources (see ref. 35 for a review, and ref. 36 for a recent meta-analysis). When given the choice to learn from two informants about an unfamiliar topic, 4-year-olds choose the one who provided more reliable information, for example, by accurately labelling familiar words[37] or providing the full extent of relevant data about a toy[38]. Similarly, children of this age will endorse a statement from a knowledgeable informant over a conflicting statement from one who is less so[39,40]. These assessments of informants are nuanced. Children as old as 4 endorse statements from informants in accordance with the proportion of inaccurate statements they make (for example, 80% versus 20%), demonstrating that judgements of trust extend beyond a binary of trustworthy or untrustworthy[41]. Furthermore, children change their endorsement of an informant's statement if later empirical checks show it to be inaccurate[42].

Children may leverage this capacity to differentiate between high- and low-quality sources in digital media environments. However, two issues still limit our understanding of how children might fare in these more complex informational ecosystems. For one, the sources on the internet are seldom transparent and often cannot be traced back to a particular agent. Rather, the statements children encounter derive from something more nebulous: webpages, posts or a series of such entities supplied by an algorithm. It is unclear how children will perform in such cases when informant-specific cues are unavailable. Second, measurements of selective social learning have traditionally relied on a forced choice between rejecting or endorsing statements from a given source. In reality, trust is more nuanced and has downstream consequences on what information people choose to seek out or not. Given uncertainty regarding a novel statement, children may opt to fact-check it. Fine-grained evaluation of children's investigatory response would provide a more informative picture of their scepticism. Moreover, investigation rather than immediate rejection is a solid strategy for the effective use of a source that provides partial misinformation.

Here, we propose an approach for motivating adaptive scepticism towards digital misinformation that builds upon children's known capacities to use the statistical properties of information in their environment—capacities for which we have strong evidence even in infancy[43–45]. The idea leverages the fact that children attend to statistical properties of their environments in order to form expectations that then modulate their learning and behaviour[43,46]. Existing empirical work shows that children wait longer in a delay-of-gratification task when given evidence that waiting will pay off[47,48]. We hypothesize that, in a similar fashion, children will use the prior reliability of information in a given context to adjust their a priori scepticism towards new claims. In two experiments and a simulation, we test whether controlled but imperfect informational environments may serve as useful scaffolding for children's abilities to detect misinformation. Exposure to detectable inaccuracies may provide critical opportunities for children to express their scepticism and practice key critical thinking skills.

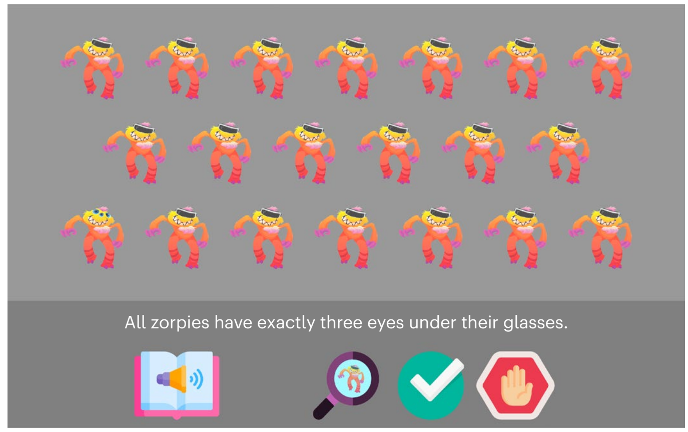

Fig. 1 | Test phase, identical in studies 1 and 2. After checking a zorpie (for example, bottom row, first from left), children could choose to accept the statement (green button), reject the statement (red button) or check another zorpie before deciding (magnifying glass). Credits: alien icon, adapted from image by upklyak on Freepik; stop icon, Freepik; inquiry icon, adapted from image created by Freepik - Flaticon; Foursquare check-in icons, Pixel Buddha - Flaticon; listen icon, Freepik.

Study 1

Study 1 asks whether children use the prior reliability of information in their environment to shape their standards of evidence for a novel claim. Do children increase their evidentiary standards for a claim selectively after exposure to misinformation? To test this, children were randomly assigned to judge the veracity of a set of animal facts that were either all true (reliable condition) or partially false (unreliable condition). Following this, children judged a novel claim about aliens, and were given the opportunity to freely sample evidence about the claim before making their final decision (Fig. 1). We hypothesized that children would sample more evidence before trusting the claim in the unreliable condition.

Results and discussion

Manipulation check. Children reliably discerned between true and false statements in the exposure phase of the experiment. Children's accuracy in evaluating statements as true or false was above chance in both the reliable (mean (M) = 9.40 of 10 correct, two-sided $t(29) = 21.88$, $P < 0.001$, Cohen's $d = 3.99$, 95% confidence interval (CI) 8.99 to 9.81) and unreliable ($M = 8.43$ of 10 correct, two-sided $t(29) = 8.64$, $P < 0.001$, $d = 1.58$, 95% CI 7.62 to 9.25) conditions, indicating that we successfully manipulated the perceived reliability of information in the exposure phase. Nine of the 60 participants failed to achieve 80% accuracy, but their exclusion does not affect any results, so we retain their data for all future analyses. In addition, all but three children (95%) correctly judged the test claim to be true, suggesting that children were generally tracking the evidence appropriately. Of the three participants who rejected the test claim, two were in the unreliable condition.

https://doi.org/10.1038/s41562-024-01992-8

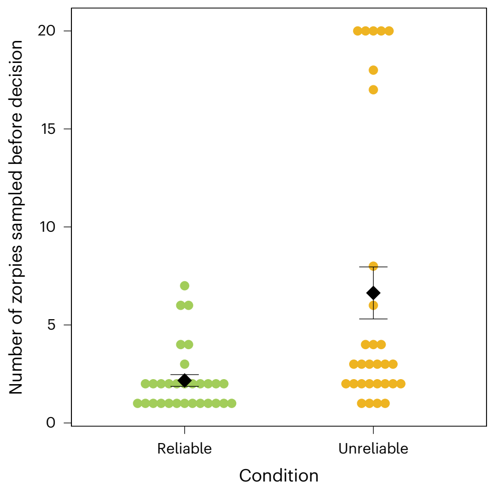

Fig. 2 | Children sampled more evidence in the unreliable condition (*n* = 30 per condition). The dots are individual data points, the diamonds are condition means and the error bars represent one s.e.m. The effect of condition remains robust after winsorization, ensuring that the highest sample values do not drive the effect.

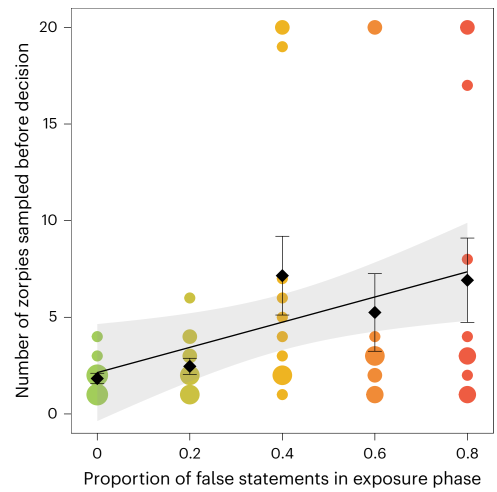

Fig. 3 | Children (*N* = 62) sampled more evidence as the reliability of their environments decreased. The amount of evidence sampled (out of a possible 20 zorpies) versus the proportion of false statements in the exposure phase in experiment 2. The size of the dot represents the number of data points. The diamonds are conditional means, and the error bars represent one s.e.m. The line is the linear regression fit with a 95% confidence interval.

Children seek more evidence in unreliable environments. Children increase their standards of evidence for new claims in an environment containing some misinformation. Figure 2 shows the number of zorpies children sampled before deciding to accept or reject the test claim by condition. We used the rstatix package in R to run a non-parametric Wilcoxon paired signed-rank test assessing the effect of condition on the amount of evidence sampled. On average, children in the unreliable condition sampled more evidence than those in the reliable condition ($M = 6.63$ versus 2.17 zorpies, location parameter −1, two-sided Wilcoxon $W = 233$, $P < 0.001$, Wilcoxon effect size (r) = 0.43, 95% CI −2 to −1). When exposed to some misinformation, children sought out more evidence before deciding whether to accept the test claim. In the unreliable condition, a number of children even opted for an exhaustive or near-exhaustive sampling strategy, checking up to 20 zorpies in a row even though all the prior evidence was identical. Children were thus able to leverage the prior quality of information in a known domain (animal facts) in order to adapt their scepticism and subsequent information search about a novel claim about which they had no prior knowledge.

The distribution of sampling behaviour in the unreliable condition was bimodal, so we also winsorized the data such that the maximum value of zorpies sampled was 8 (the maximum of the other mode). The fact that an exhaustive sampling strategy leads children to sample exactly 20 zorpies is the result of a design choice, so replacing extreme values with the maximum value of the other mode provides a more stringent and design-neutral test of our hypothesis. The effect remained robust after winsorization (location parameter −1, two-sided Wilcoxon $W = 233$, $P < 0.001$, $r = 0.43$, 95% CI −2 to −1), suggesting that it was not driven by the subset of exhaustive samplers in the unreliable condition.

Finally, we tested whether children's sensitivity to the reliability of their informational environments changes with age. We used the betabin function from the aod package in R to run a beta-binomial regression using condition and standardized age to predict the proportion of zorpies sampled. This analysis replicated the main effect of the unreliable condition (regression coefficient $\beta = 0.84$, Wald test statistic $z(55) = 2.78$, $P = 0.005$, 95% CI 0.25 to 1.43) but revealed no main effect of age ($\beta = 0.07$, $z(55) = 0.30$, $P = 0.764$, 95% CI −0.37 to 0.50) and no interaction ($\beta = -0.02$, $z(55) = -0.07$, $P = 0.942$, 95% CI −0.61 to 0.56). There was no evidence of changes in how children responded to the reliability of information from ages 4 through 6 in our sample.

Study 2

The reliability of a body of information is not all-or-nothing. Do children appreciate nuances in the reliability of their broader informational environments and adapt their level of scepticism accordingly? To address this question, study 2 introduced five between-subjects conditions of varying reliability, ranging from 0% to 80% false statements in the exposure phase. In addition, children probably assumed that the information in study 1 came from a single, cohesive source. The task was framed as an ebook, and children heard all statements in audio recordings using the same voice. Can children still make smart inferences about claims derived from a more complex environment composed of many distinct sources? In study 2, we presented statements as individual search results, read in distinct voices, to test whether children make more abstract generalizations about their informational environment to adapt their information seeking.

Results and discussion

Manipulation check. Children reliably discerned between true and false statements in the exposure phase of the experiment in study 2. Accuracy in the exposure phase was above chance ($M = 9.24$ of 10 correct, $t(61) = 33.3$, two-sided $P < 0.001$, $d = 4.23$, 95% CI 8.99 to 9.50). Four of the 62 participants failed to achieve 80% accuracy, but their exclusion does not affect any results, so we retain their data for all future analyses. In addition, all but four children (93.5%) correctly judged the test claim to be true. The children who rejected the test claim were in the two most unreliable conditions (three in the 80% false condition, one in the 60% false condition).

Graded sensitivity to reliability. Figure 3 shows the number of zorpies children sampled before deciding to accept or reject the test claim as a function of the proportion of false statements encountered in the exposure phase. A beta-binomial regression revealed that the proportion of false statements from the exposure phase positively predicted the proportion of the 20 zorpies that participants sampled in the test phase ($\beta = 1.10$, $z(59) = 2.04$, $P = 0.041$, 95% CI 0.04 to 2.16). Scepticism increased with increases in the number of false statements in the exposure phase, manifesting in more extensive information search in the test phase. A linear model finds the same effect and supports the same conclusions (Supplement A). Children are thus able to make sophisticated, graded judgments about the reliability of their current informational environments, and use that to guide future learning. Note that this sensitivity was observed in a simulated search engine context composed of distinct sources—each statement was heard from a different voice. This suggests that children went beyond speaker-based heuristics and tracked the cumulative quality of information throughout the exposure phase.

While we observed that increases in the proportion of misinformation led to increases in the sampling of evidence, it is unclear whether this would generalize across all tasks. In our task, the information to be gained by additional sampling was maximally transparent. The outcome was binary and directly related to the claim in question (the next zorpie is three-eyed or not), and the full space of available evidence was clearly delineated. Other studies with low-risk exploration have found linear associations between low certainty and information seeking in children[49] and adults[50].

However, some evidence suggests that environments characterized by variable expected information gain induce a U-shaped relationship between curiosity and information seeking[51,52]. This pattern of results is consistent with a dual-process account of metacognition, in which information seeking is guided not only by certainty but also by an appraisal of the potential information gain afforded by the environment[53,54]. It is speculated that the bimodal distribution of sampling strategies even in the most unreliable conditions of study 2 may represent two distinct interpretations of the environment. Some of the children who checked only a few zorpies in highly unreliable environments may have been highly sceptical, but doubted that the available evidence would provide accurate information in the first place. The effect we observe may therefore be the combination of two patterns of responses: a monotonic, positive relationship dampened by a subgroup exhibiting a U-shaped pattern. Alternatively, the pattern may be influenced by variations in children's metacognition or executive function.

Sensitivity may increase with age. Does children's scepticism become more finely attuned to the reliability of their informational environment as they age? We ran a beta-binomial regression using standardized age and the standardized proportion of false statements in the exposure phase to predict the proportion of zorpies sampled. First, we replicated the main effect of environmental reliability (that is, proportion of false statements in the exposure phase ($\beta = 0.37$, $z(57) = 2.48$, $P = 0.013$, 95% CI 0.08 to 0.67). In addition, this analysis revealed a significant main effect of age ($\beta = 0.43$, $z(57) = 2.81$, $P = 0.005$, 95% CI 0.13 to 0.73) and a significant interaction ($\beta = 0.35$, $z(57) = 2.14$, $P = 0.032$, 95% CI 0.03 to 0.67). The main effect of age suggests that older children sought out more evidence than younger children overall. The reliability by age interaction suggests older children were more sensitive than younger children to variation in environmental reliability. Older children in our sample, and particularly the 7-year-olds, were more likely than younger children to sample a high number of zorpies when they had encountered a high proportion of false information in the past.

Study 3 (simulation)

In studies 1 and 2, the test claims were true. Yet, the selective scepticism that children exhibit in these studies is theoretically adaptive because increased information sampling facilitates the discovery of counterevidence. If the test claim was actually false, what kind of environment would best prepare children to discover that? We ran a simulation to determine whether experience learning in an unreliable environment enables children to identify misinformation more easily.

We ran four simulations of 100,000 hypothetical experiments each, in which a randomly sampled proportion of the zorpies in the sample space served as hypothetical counterevidence to the test claim (that is, zorpies without three eyes). Then, using participants' real sampling behaviour from study 1, we computed the proportion of times each individual would have successfully found one or more of the

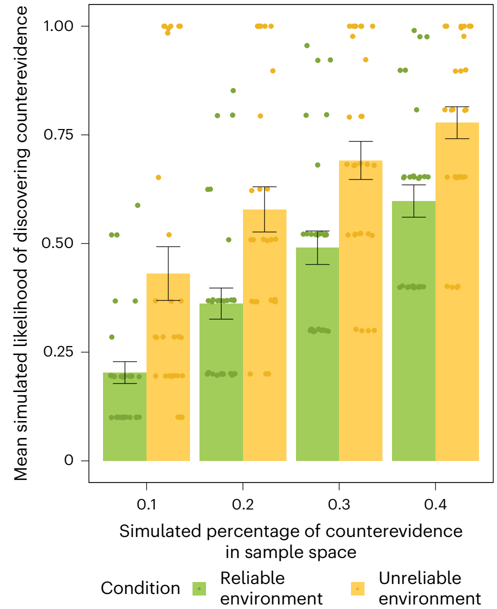

Fig. 4 | Misinformation is more likely to be detected in unreliable environments. Simulation results reveal that children in the unreliable condition of study 1 ($n = 30$ per condition for each simulation) would have been more likely to discover counterevidence than those in the reliable condition. This pattern holds when counterevidence is both rare and relatively common. The dots are likelihoods for each participant in study 1. The error bars represent one s.e.m.

counterevidence zorpies during the task across the 100,000 simulated experiments. This proportion represents the simulated likelihood of a given participant to discover counterevidence for the claim and, in doing so, debunk misinformation in the test phase. We repeated this simulation procedure for four proportions of counterevidence in the sample space such that 10%, 20%, 30% and 40% of the available zorpies represented evidence against the test claim. Figure 4 visualizes the mean likelihood of discovering counterevidence across the 100,000 simulations, according to the participant's experimental condition from study 1 and the simulated proportion of counterevidence within the sample space.

To test whether children in the unreliable condition were more likely to discover counterevidence during sampling, we ran a beta-binomial regression using experimental condition (reliable versus unreliable) to predict the proportion of simulation runs in which the participants discovered one or more pieces of counterevidence during sampling. Since we used participants' actual sampling behaviour from study 1, there is dependence within each participant's outcomes at each of the four simulated proportions of counterevidence. Therefore, we analysed the results of the four simulations (with 10%, 20%, 30% and 40% counterevidence in the sample space) separately with identical beta-binomial models. The results indicated a significant effect of condition in the 10% counterevidence condition ($\beta = 1.40$, $z = 4.05$, $P < 0.001$, dispersion parameter $\phi = 0.47$, 95% CI 0.72 to 2.08), the 20% counterevidence condition ($\beta = 1.73$, $z = 5.06$, $P < 0.001$, $\phi = 0.41$, 95% CI 1.06 to 2.40), the 30% counterevidence condition ($\beta = 1.72$, $z = 5.31$, $P < 0.001$, $\phi = 0.36$, 95% CI 1.08 to 2.35) and the 40% counterevidence condition ($\beta = 1.63$, $z = 5.24$, $P < 0.001$, $\phi = 0.32$, 95% CI 1.02 to 2.23), such that participants in the unreliable condition discovered counterevidence in a higher proportion of simulations compared with those in the reliable condition. Thus, across varying proportions of counterevidence in the sample space, children who were exposed to more misinformation in the exposure phase were more likely to discover counterevidence about a novel, verifiable claim in the subsequent test phase (if such counterevidence existed). The simulation results replicate the expected probabilities of each participant sampling one or more pieces of counterevidence without replacement based on the hypergeometric distribution (Supplement B). These convergent findings support the commonsense conclusion that unreliable informational environments elicit increased scepticism and fact-checking behaviour, which enables children to debunk misinformation more readily.

General discussion

To learn accurately and efficiently, children must have an adaptive policy for deciding which claims to trust on the spot, and which to seek more evidence for. In two experiments, we investigated whether children use the reliability of their informational environment to make rational inferences about whether a new claim warrants fact-checking. Study 1 demonstrated that children seek more evidence for a novel claim about aliens that arises in a context containing some misinformation about animals. Study 2 showed that this evidence-seeking behaviour increased in proportion to the number of false claims children were previously exposed to. Moreover, this effect held when information was presented in a search engine context, in which each claim derived from a distinct source. Children thus adjusted their level of scepticism not in accordance with speaker-specific cues but with the reliability of a broad informational environment. They made fine-grained assessments of the reliability of incoming information in a known domain, inferred that this reliability would generalize to another domain and chose a graded evidentiary standard corresponding to that reliability. Finally, we showed with a simulation (study 3) that this behaviour is adaptive: learners have the greatest opportunity to discover counterevidence and debunk misinformation in the most unreliable environments, where misinformation is most likely to be present.

Children's ability to calibrate their evidentiary standards to the reliability of their environments helps them confront the challenge of balancing speed and accuracy during learning. Children wasted little time verifying a claim within an environment with established reliability in a known domain. Instead, they reserved more extensive information seeking for more questionable informational contexts, modulating their evidentiary standards according to nuanced changes in reliability. While this strategy is certainly not infallible, it gives children a sensible policy for information seeking in line with resource-rational decision-making[55]. Even when they lack domain-relevant knowledge to judge a claim's content, children leverage sophisticated attributes of their context to guide their scepticism and exploration selectively. These findings dovetail with recent work showing that adults' beliefs about naturalistic news headlines adapt to the prior base rate of true headlines in their news feeds[56], which emphasizes the relevance of this environmental adaptation for understanding real-world beliefs.

Our experiments used an open information sampling task to capture a graded sense of children's level of scepticism or evidentiary standards. This continuous measure allowed us to capture the quantity of evidence children searched for, which corresponded to degrees of belief in a given claim. This approach confers several advantages over the forced-choice paradigm that is common in the selective trust literature. The continuous measure indexes the strength or uncertainty of children's individual beliefs, rather than demonstrating only a relative belief in one informant's testimony over another's. It also provides insight into children's strategies for translating their level of trust into rational information-seeking behaviour, which is crucial to arrive at true beliefs in complex environments. Further, the measure we employed is implicit, which makes it more suitable for use in younger children than explicit reports[57]. This work thus builds upon literature that demonstrates that infants' and children's information-seeking behaviour is sensitive to uncertainty[58–60]. We show that information seeking is sensitive to environmental certainty, as well as content-specific certainty.

Let us consider some limitations. While these data demonstrate childrens' sensitivity to their informational environments, it is unclear precisely how they conceptualize the information source in the first place. Children may have directed their scepticism towards the content of the particular ebook (study 1) or search engine (study 2) they were exposed to, towards any information presented on the touchscreen computer or towards any information presented by the experimenter, among other possibilities. The present experiments cannot distinguish between these possibilities, but evidence suggests that children of this age can limit the scope of their inferences at least to a particular technological device. Preschoolers and kindergarteners express selective trust for computers[40], social robots[61] and internet sources[62] on the basis of their past accuracy. Still, future work should clarify the limits of children's scepticism in this paradigm by manipulating whether the test phase occurs in the same context as the exposure phase.

The nature of the misinformation that children encounter probably impacts the scepticism they express. In our studies, misinformation was maximally detectable: errors in the animal facts were blatant and contradicted the accompanying pictures, which were always consistent with reality. It was important to establish that pictures in the environment were trustworthy because pictures served as the evidence to be sampled in the test phase. However, these elements of our task are unrepresentative of typical misinformation that children might encounter online. Misinformation is often more subtle, and the ground truth is rarely directly accessible. If misinformation is undetectable, it probably will not impact children's subsequent scepticism and fact-checking behaviour. Moreover, the errors in our task were related to key properties of the animals (for example, 'Zebras have red and green stripes'). Children are more vigilant against speakers who make semantic errors like these compared with speakers who make episodic errors, potentially because the former 'resist folk-psychological explanation'[63]. Although it is unclear whether children make the same folk-psychological inferences about these broader, depersonalized informational environments, there is evidence that preschoolers use the accuracy of text-based sources to make inferences about their authors[64]. At the same time, children tend to attribute a computer's errors to problems with the device itself rather than with its human user[40]. Thus, it remains unclear whether children's evidentiary standards would scale up in the same way in response to other forms of misinformation in digital contexts. Still, the fact that children's scepticism shows graded adaptation to the rate of misinformation in an environment composed of many distinct voices suggests that these inferences need not be tethered to a representation of a single agent. Future work should assess how qualitatively different forms of misinformation correspond to children's environment-level scepticism.

A central insight of this work is that children's approach towards novel information is shaped by expectations that are formed through experience with their informational environment. This suggests that efforts to expose children only to curated informational environments may be misguided. Early experiences with overly sanitized environments may lead children to develop overly trusting priors and rob them of opportunities to develop critical thinking skills. By the same token, early exposure to more heterogeneous informational environments may allow children to 'flex their scepticism muscles' and build upon their existing capacities for adaptive information seeking. This idea is consistent with evidence that exposing adults to blatant misinformation makes them less susceptible to more subtle misinformation compared with a control condition[65], although recent direct and conceptual replications failed to find this effect[66].

Crucially, the form of scepticism engendered by our paradigm was not overhasty rejection of all novel information, but critical assessment of that information. This stands in contrast to work that suggests that adaptation to a largely false news feed causes adults to disproportionately misidentify true headlines as false[56,67]. Real-world circumstances often justify this kind of adaptation. For example, a long history of medical discrimination and abuses such as the Tuskegee Syphilis Study are justified reasons for adaptive scepticism that explain why

Black Americans are more likely than other racial groups to endorse health-specific coronavirus disease 2019 conspiracy beliefs[68] and to refuse vaccination[69]. However, given that most media consumed by the public is true[70], an overgeneralized resistance towards reliable information has potential to harm people's discrete beliefs and to erode trust in public institutions more broadly. In our study, when presented with misinformation, children did not show any indication that they would immediately discard true claims from a novel domain. Instead, when uncertain of a claim's truth, children opted to perform a sometimes tedious empirical check of its veracity. In preparing children for an informational environment where misinformation is present but not omnipresent, we contend that the critical engagement with novel claims fostered by our paradigm is preferable to mere acceptance or rejection. Operating in environments where detectable misinformation is present but checkable seems to be a promising means of scaffolding this analytic disposition in children.

Overall, these findings demonstrate that the knowledge that novel claims require evidence is early emerging and context sensitive. Even so, scepticism is a useful signal only if children know how to act upon it in rational and informative ways. While our task provides the full space of evidence to verify the test claim (in the form of the 20 zorpies), fact-checking in the real world is substantially more complex. This suggests that the most fruitful avenue of intervention may not be on scepticism itself, but on children's more specific capacities to know where to look for relevant evidence in a given domain and to evaluate how different kinds of evidence bear on complex claims. Indeed, research suggests that children are not sensitive to the relative strengths of explanations until early school age[71]. In sum, intervention efforts should focus on helping children develop a broad skill set for evaluating information, rather than attempting to control their information diets.

Methods

Experiments were approved by the institutional review board at the University of California, Berkeley (protocol number 2018-12-11653). Informed consent was obtained by a legal guardian of all participants before participation. All children provided verbal assent, and 7-year-old children additionally signed an assent form. None of the studies was pre-registered. Data collection and analysis were not performed blind to the conditions of the experiments. Participants were pseudo-randomly assigned to experimental conditions.

Study 1

Participants. Sixty 4- to 6-year-old children (M_{age} = 5.51, standard deviation 0.89, 47% white, 58% female, 40% male, 2% gender not declared) were recruited from parks in the San Francisco Bay Area. Three additional children were excluded from analysis because they had watched another child participate or were too distracted to complete the study. Children received a small toy valued at $1–2 as compensation for their participation. No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those reported in previous publications[38]. The sample size was sufficiently large that parametric tests should be robust to violations of the normality assumption[72].

Procedure. Children used a touchscreen computer to play a game created in PsychoPy. In an exposure phase, the experimenter asked children to determine whether a set of statements about animals in an e-book were right or wrong. On each of 12 exposure trials, the tablet displayed a statement (for example, 'Zebras have black and white stripes') and an accompanying picture. Children first tapped a button to hear an audio recording of the statement and then indicated whether the fact was right (by tapping a green button) or wrong (by tapping a red button). The facts varied by condition (between-subjects, n = 30 per condition). In the reliable condition, all 12 animal statements were true. In the unreliable condition, 4 of the 12 animal statements were clearly false (for example, 'Zebras have red and green stripes'). Pictures were identical across conditions, so children could judge the statements using real-world knowledge or the pictures alone. The first two trials were considered practice trials, and children were given feedback if they were wrong. No feedback was provided on the remaining ten trials.

In the subsequent test phase (Fig. 1), children moved on to a second chapter of the e-book, which was about a novel alien species called 'zorpies'. They were asked to evaluate a new statement about zorpies: 'All zorpies have exactly three eyes under their glasses'. The screen displayed the fact alongside 20 zorpies wearing opaque sunglasses. After tapping a button to hear the fact, children were told that they could tap any zorpie to remove its glasses and reveal its eyes. All zorpies were identical and had three eyes, so any evidence the child sampled supported the test statement. Once the child tapped a zorpie, they had to decide to tap a button to accept the statement as true, reject the statement as false or check another zorpie first. This procedure repeated such that children could check as many zorpies as they wished (from 1 to all 20) before indicating whether the fact was right or wrong and completing the study. The task was designed to produce different information-seeking behaviours depending on one's level of scepticism towards the claim. A fully trusting learner might see that all the zorpies are identical and be satisfied after checking only one, while a highly sceptical learner might feel the need to check all 20 zorpies because the statement refers to 'all zorpies'.

Study 2

Participants. Sixty-two 4- to 7-year-old children (M_{age} = 5.88, standard deviation 1.06, 52% white, 40% female, 50% male, 10% other gender or undeclared) were recruited from parks in the San Francisco Bay Area. Participants were pseudo-randomly assigned to five between-subjects conditions with sample sizes of n = 12, 13, 13, 12 and 12, respectively. Four additional participants were excluded from analysis because they had watched another child participate. None of the study 2 participants had completed study 1 previously. Children received a small toy valued at $1–2 as compensation for their participation. Data collection stopped at the end of a testing day when the sample size had reached or surpassed 60 children, which was the sample size of study 1.

Procedure. The procedure was identical to study 1 aside from two main changes. First, we created five between-subjects conditions such that 0%, 20%, 40%, 60% or 80% of the ten exposure trials were false statements. Exposure trials were presented in a randomized order. Second, the activity was reframed so that the statements appeared to originate from distinct search engine results. The experimenter typed 'Animal facts' into a search bar to generate a simulated results page in the exposure phase. The experimenter tapped on a search result to begin a trial. On each trial, the style of the picture and the voice of the audio recording were different. The test phase followed. In the test phase, the experimenter input 'Alien facts' into the search bar and tapped a result to display the page of zorpies. The audio recording on the zorpie test trial featured another distinct voice.

Reporting summary

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

Data availability

Data for all experiments are available at https://osf.io/7hxkt/.

Code availability

Code for the simulation and analyses is available at https://osf.io/7hxkt/.

References

1. Brown, A. S. & Nix, L. A. Turning lies into truths: referential validation of falsehoods. *J. Exp. Psychol. Learn., Mem. Cogn.* **22**, 1088 (1996).
2. Fazio, L. K. & Sherry, C. L. The effect of repetition on truth judgments across development. *Psychol. Sci.* **31**, 1150–1160 (2020).
3. Fazio, L. K., Brashier, N. M., Payne, B. K. & Marsh, E. J. Knowledge does not protect against illusory truth. *J. Exp. Psychol. Gen.* **144**, 993 (2015).
4. Kidd, C. & Birhane, A. How AI can distort human beliefs. *Science* **380**, 1222–1223 (2023).
5. Xu, S., Shtulman, A. & Young, A. G. Can children detect fake news? In *Proc. Annual Meeting of the Cognitive Science Society* (eds Culbertson, J. et al.) 2988–2993 (Cognitive Science Society, 2022).
6. Jaswal, V. K., Croft, A. C., Setia, A. R. & Cole, C. A. Young children have a specific, highly robust bias to trust testimony. *Psychol. Sci.* **21**, 1541–1547 (2010).
7. Plate, R. C., Shutts, K., Cochrane, A., Green, C. S. & Pollak, S. D. Testimony bias lingers across development under uncertainty. *Dev. Psychol.* **57**, 2150 (2021).
8. Hermansen, T. K., Ronfard, S., Harris, P. L. & Zambrana, I. M. Preschool children rarely seek empirical data that could help them complete a task when observation and testimony conflict. *Child Dev.* **92**, 2546–2562 (2021).
9. Ecker, U. K. et al. The psychological drivers of misinformation belief and its resistance to correction. *Nat. Rev. Psychol.* **1**, 13–29 (2022).
10. Poll, M. Sharing too soon? Children and social media apps. *C.S. Mott Children's Hospital* https://mottpoll.org/reports/sharing-too-soon-children-and-social-media-apps/ (2021).
11. *New Survey Reveals Teens Get Their News from Social Media and YouTube* (Common Sense Media, 2019).
12. *New Poll Finds Parents Lag Behind Kids on AI and Want Rules and Reliable Information to Help Them* (Common Sense Media, 2023).
13. Rodriguez, A. YouTube Kids is giving parents more control over what their kids watch. *Quartz* https://qz.com/1262977/youtube-kids-is-launching-a-mode-curated-by-humans-not-just-algorithms/ (2018).
14. The disturbing YouTube videos that are tricking children. *BBC* https://www.bbc.com/news/blogs-trending-39381889 (2017).
15. Maheshwari, S. On YouTube Kids, startling videos slip past filters. *The New York Times* https://www.nytimes.com/2017/11/04/business/media/youtube-kids-paw-patrol.html/ (2017).
16. Kallioniemi, P. The role of human curation at the age of algorithms. *J. Digit. Media Interact.* **4**, 7–20 (2021).
17. Lewandowsky, S. & Van Der Linden, S. Countering misinformation and fake news through inoculation and prebunking. *Eur. Rev. Soc. Psychol.* **32**, 348–384 (2021).
18. Compton, J., van der Linden, S., Cook, J. & Basol, M. Inoculation theory in the post-truth era: extant findings and new frontiers for contested science, misinformation, and conspiracy theories. *Soc. Pers. Psychol. Compass* **15**, e12602 (2021).
19. van der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nat. Med.* **28**, 460–467 (2022).
20. Basol, M., Roozenbeek, J. & Van der Linden, S. Good news about bad news: gamified inoculation boosts confidence and cognitive immunity against fake news. *J. Cogn.* **3**, 2 (2020).
21. Roozenbeek, J., van der Linden, S. & Nygren, T. Prebunking interventions based on the psychological theory of 'inoculation' can reduce susceptibility to misinformation across cultures. *Harvard Kennedy School Misinformation Review* https://doi.org/10.37016//mr-2020-008 (2020).
22. van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the public against misinformation about climate change. *Glob. Chall.* **1**, 1600008 (2017).
23. Wong, N. C. H. 'Vaccinations are safe and effective': inoculating positive HPV vaccine attitudes against antivaccination attack messages. *Commun. Rep.* **29**, 127–138 (2016).
24. Braddock, K. Vaccinating against hate: using attitudinal inoculation to confer resistance to persuasion by extremist propaganda. *Terror. Polit. Violence* **34**, 240–262 (2022).
25. Maertens, R., Roozenbeek, J., Basol, M. & van der Linden, S. Long-term effectiveness of inoculation against misinformation: three longitudinal experiments. *J. Exp. Psychol. Appl.* **27**, 1–16 (2021).
26. Capewell, G. et al. Misinformation interventions decay rapidly without an immediate posttest. *J. Appl. Social Psychol.* **54**, 441–454 (2023).
27. Guay, B., Berinsky, A. J., Pennycook, G. & Rand, D. How to think about whether misinformation interventions work. *Nat. Hum. Behav.* **7**, 1231–1233 (2023).
28. Williams, D. The fake news about fake news. *Boston Review* https://www.bostonreview.net/articles/the-fake-news-about-fake-news/ (2023).
29. Chan, M. P. S. & Albarracín, D. A meta-analysis of correction effects in science-relevant misinformation. *Nat. Hum. Behav.* **7**, 1514–1525 (2023).
30. Modirrousta-Galian, A. & Higham, P. A. Gamified inoculation interventions do not improve discrimination between true and fake news: reanalyzing existing research with receiver operating characteristic analysis. *J. Exp. Psychol. Gen.* **152**, 2411–2437 (2023).
31. Finn, B. & Metcalfe, J. Overconfidence in children's multi-trial judgments of learning. *Learn. Instr.* **32**, 1–9 (2014).
32. Lipko, A. R. et al. Using standards to improve middle school students' accuracy at evaluating the quality of their recall. *J. Exp. Psychol. Appl.* **15**, 307–318 (2009).
33. Salles, A., Ais, J., Semelman, M., Sigman, M. & Calero, C. I. The metacognitive abilities of children and adults. *Cogn. Dev.* **40**, 101–110 (2016).
34. Best, J. R. & Miller, P. H. A developmental perspective on executive function. *Child Dev.* **81**, 1641–1660 (2010).
35. Harris, P. L., Koenig, M. A., Corriveau, K. H. & Jaswal, V. K. Cognitive foundations of learning from testimony. *Annu. Rev. Psychol.* **69**, 251–273 (2018).
36. Tong, Y., Wang, F. & Danovitch, J. The role of epistemic and social characteristics in children's selective trust: three meta-analyses. *Dev. Sci.* **23**, e12895 (2020).
37. Koenig, M. A., Clément, F. & Harris, P. L. Trust in testimony: children's use of true and false statements. *Psychol. Sci.* **15**, 694–698 (2004).
38. Gweon, H., Pelton, H., Konopka, J. A. & Schulz, L. E. Sins of omission: children selectively explore when teachers are under-informative. *Cognition* **132**, 335–341 (2014).
39. Corriveau, K. H. & Kurkul, K. E. 'Why does rain fall?': children prefer to learn from an informant who uses noncircular explanations. *Child Dev.* **85**, 1827–1835 (2014).
40. Danovitch, J. H. & Alzahabi, R. Children show selective trust in technological informants. *J. Cogn. Dev.* **14**, 499–513 (2013).
41. Pasquini, E. S., Corriveau, K. H., Koenig, M. & Harris, P. L. Preschoolers monitor the relative accuracy of informants. *Dev. Psychol.* **43**, 1216–1226 (2007).
42. Hermansen, T. K., Ronfard, S., Harris, P. L., Pons, F. & Zambrana, I. M. Young children update their trust in an informant's claim when experience tells them otherwise. *J. Exp. Child Psychol.* **205**, 105063 (2021).
43. Saffran, J. R., Aslin, R. N. & Newport, E. L. Statistical learning by 8-month-old infants. *Science* **274**, 1926–1928 (1996).

44. Fiser, J. & Aslin, R. N. Statistical learning of new visual feature combinations by infants. *Proc. Natl Acad. Sci. USA* **99**, 15822–15826 (2002).
45. Xu, F. & Garcia, V. Intuitive statistics by 8-month-old infants. *Proc. Natl Acad. Sci. USA* **105**, 5012–5015 (2008).
46. Kidd, C., Piantadosi, S. T. & Aslin, R. N. The Goldilocks effect: human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS ONE* **7**, e36399 (2012).
47. Watts, T. W., Duncan, G. J. & Quan, H. Revisiting the marshmallow test: a conceptual replication investigating links between early delay of gratification and later outcomes. *Psychol. Sci.* **29**, 1159–1177 (2018).
48. Kidd, C., Palmeri, H. & Aslin, R. N. Rational snacking: young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition* **126**, 109–114 (2013).
49. Coughlin, C., Hembacher, E., Lyons, K. E. & Ghetti, S. Introspection on uncertainty and judicious help-seeking during the preschool years. *Dev. Sci.* **18**, 957–971 (2015).
50. Desender, K., Boldt, A. & Yeung, N. Subjective confidence predicts information seeking in decision making. *Psychol. Sci.* **29**, 761–778 (2018).
51. Baranes, A. F., Oudeyer, P. Y. & Gottlieb, J. The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Front. Neurosci.* **8**, 317 (2014).
52. Wang, J., Yang, Y., Macias, C. & Bonawitz, E. Children with more uncertainty in their intuitive theories seek domain-relevant information. *Psychol. Sci.* **32**, 1147–1156 (2021).
53. Goupil, L. & Proust, J. Curiosity as a metacognitive feeling. *Cognition* **231**, 105325 (2023).
54. Baer, C. & Kidd, C. Learning with certainty in childhood. *Trends Cogn. Sci.* **26**, 887–896 (2022).
55. Bhui, R., Lai, L. & Gershman, S. J. Resource-rational decision making. *Curr. Opin. Behav. Sci.* **41**, 15–21 (2021).
56. Orchinik, R., Martel, C., Rand, D. G. & Bhui, R. Uncommon errors: adaptive intuitions in high-quality media environments increase susceptibility to misinformation. *OSF* https://doi.org/10.31234/osf.io/q7r58 (2023).
57. Goupil, L. & Kouider, S. Developing a reflective mind: from core metacognition to explicit self-reflection. *Curr. Dir. Psychol. Sci.* **28**, 403–408 (2019).
58. Langenhoff, A. F., Engelmann, J. M. & Srinivasan, M. Children's developing ability to adjust their beliefs reasonably in light of disagreement. *Child Dev.* **94**, 44–59 (2023).
59. Lapidow, E., Killeen, I. & Walker, C. M. Learning to recognize uncertainty vs. recognizing uncertainty to learn: confidence judgments and exploration decisions in preschoolers. *Dev. Sci.* **25**, e13178 (2022).
60. Goupil, L., Romand-Monnier, M. & Kouider, S. Infants ask for help when they know they don't know. *Proc. Natl Acad. Sci. USA* **113**, 3492–3496 (2016).
61. Brink, K. A. & Wellman, H. M. Robot teachers for children? Young children trust robots depending on their perceived accuracy and agency. *Dev. Psychol.* **56**, 1268–1277 (2020).
62. Tong, Y., Wang, F., Danovitch, J. & Wang, W. When the internet is wrong: children's trust in an inaccurate internet or human source. *Br. J. Dev. Psychol.* **40**, 320–333 (2022).
63. Stephens, E. C. & Koenig, M. A. Varieties of testimony: children's selective learning in semantic versus episodic domains. *Cognition* **137**, 182–188 (2015).
64. Vanderbilt, K. E., Ochoa, K. D. & Heilbrun, J. Consider the source: children link the accuracy of text-based sources to the accuracy of the author. *Br. J. Dev. Psychol.* **36**, 634–651 (2018).
65. Loftus, E. F. Reactions to blatantly contradictory information. *Mem. Cogn.* **7**, 368–374 (1979).
66. O'Donnell, R. & Chan, J. C. Does blatantly contradictory information reduce the misinformation effect? A registered report replication of Loftus (1979). *Legal Criminol. Psychol.* https://doi.org/10.1111/lcrp.12242 (2023).
67. Altay, S., Lyons, B. A. & Modirrousta-Galian, A. Exposure to higher rates of false news erodes media trust and fuels overconfidence. *Mass Commun. Soc.* https://doi.org/10.1080/15205436.2024.2382776 (2024).
68. Smith, A. C., Woerner, J., Perera, R., Haeny, A. M. & Cox, J. M. An investigation of associations between race, ethnicity, and past experiences of discrimination with medical mistrust and COVID-19 protective strategies. *J. Racial Ethn. Health Disparities* **9**, 1430–1442 (2022).
69. Majee, W., Anakwe, A., Onyeaka, K. & Harvey, I. S. The past is so present: understanding COVID-19 vaccine hesitancy among African American adults using qualitative data. *J. Racial Ethn. Health Disparities* **10**, 462–474 (2023).
70. Acerbi, A., Altay, S. & Mercier, H. Research note: fighting misinformation or fighting for information? *Harvard Kennedy School Misinformation Review* https://doi.org/10.37016/mr-2020-87 (2022).
71. Danovitch, J. H., Mills, C. M., Sands, K. R. & Williams, A. J. Mind the gap: how incomplete explanations influence children's interest and learning behaviors. *Cogn. Psychol.* **130**, 101421 (2021).
72. Elliott, A. C. *Statistical Analysis Quick Reference Guidebook: With SPSS Examples* (Sage Publications, 2007).

Acknowledgements

We thank C. Jiang and J. Witkowski for their assistance with data collection, S. Piantadosi and members of the Kidd 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.), Jacobs Foundation (award no. 048412, C.K.), John Templeton Foundation (award no. 61475, C.K.) and the Berkeley Center for New Media (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. E.O. and M.M. collected the data. E.O. analysed the data. E.O. 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/s41562-024-01992-8.

Correspondence and requests for materials should be addressed to Evan Orticio.

Peer review information *Nature Human Behaviour* thanks Paul Harris, Melissa Koenig and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Evan Orticio

Last updated by author(s): Jul 31, 2024

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	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	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)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), 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 custom code and materials are available at https://osf.io/7hxkt/.
Data analysis	Data analysis was conducted in R version 4.3.1. All code is available at https://osf.io/7hxkt/.

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

Data for all experiments are available at https://osf.io/7hxkt/.

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 identity any a priori reasons why gender would affect the cognitive processes of interest. Participants optionally provided gender data. In Study 1, participants were 58% female, 40% male, and 2% gender not declared. In Study 2, participants were 40% female, 50% male, and 10% other gender or undeclared.
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. Study 1 sample was 47% White, 15% Asian, 7% Black or African American, 13% Multiracial, and 18% Other or unidentified. 20% identified as Hispanic, of any race. Study 2 sample was 52% White, 11% Asian, 11% Black or African American, 11% Multiracial, 2% American Indian or Alaska Native, and 13% Other or unidentified. 16% identified as Hispanic, of any race.
Population characteristics	Participants were aged 4-7 years.
Recruitment	Convenience samples were collected at local parks in the San Francisco Bay Area. Recruitment took place in free public parks with the aim of recruiting a socioeconomically diverse sample, relative to other common recruitment strategies (e.g. museums, in-lab testing) which skew toward more economically advantaged populations and populations who are more familiar with research.
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

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. Participants were assigned to between-subjects conditions.
Research sample	Our research samples were children aged 4-7 from the San Francisco Bay Area. The gender and racial demographics of our samples are broadly representative of the local population (see above). Recruitment took place in free public parks with the aim of recruiting a socioeconomically diverse sample, relative to other recruitment strategies (e.g. museums, in-lab testing) which skew toward more economically advantaged populations.
Sampling strategy	For Study 1, we obtained a convenience sample of 60 children recruited from local parks. We chose 60 participants (30 per cell) as a rule of thumb for moderate effect sizes (Cohen, 1988) and due to the difficulty of recruiting children. For Study 2, our stopping rule was to halt data collection at the end of a recruitment day if the total sample equaled or surpassed 60 children.
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. Researchers were not blind to the experimental condition or hypotheses.
Timing	Data collection for Study 1 occurred from 30 June 2022 to 7 August 2022. Data collection for Study 2 occurred from 8 October 2022 to 23 October 2022.
Data exclusions	In Study 1, two children were excluded because they had watched another child participate and one was excluded for being too distracted to complete the study. In Study 2, four children were excluded because they had watched another child participate.
Non-participation	One child in Study 1 failed to complete the study due to distraction. They were recruited in a public park and decided to continue playing on the playground instead.
Randomization	Participants were pseudo-randomly assigned to each experimental condition. Experimenters alternated experimental conditions with each new participant that was approached.

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 & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ 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, mosiacism, off-target gene editing) were examined.