UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Think outside the box: Making up causal hypotheses from unreliable evidence

Permalink
https://escholarship.org/uc/item/7mw6988v

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors
Dolek, Simal
Radovanovic, Mia
Gonzales, Robie
et al.

Publication Date
2025

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Think outside the box: Making up causal hypotheses from unreliable evidence

Simal Dolek (Simal.Dolek@dal.ca)
Psychology and Neuroscience
Dalhousie University

Mia Radovanovic
(m.radovanovic@mail.utoronto.ca)
Department of Psychology
University of Toronto

Robie Gonzales
(robie.gonzales@dal.ca)
Faculty of Computer Science
Dalhousie University

Jessica Sommerville
(jessica.sommerville@mail.utoronto.ca)
Department of Psychology
University of Toronto

Marta Kryven (marta.kryven@dal.ca)
Faculty of Computer Science
Department of Psychology and Neuroscience
Dalhousie University

Abstract

Human adults and children think of the natural world as orchestrated by rules, yet many of them are neither equally rigid nor clearly evident. Some are beset by exceptions, and others are not intuitive. What cognitive mechanisms underpin human learning of how the world works? We propose a computational model for formulating and testing hypotheses in naturalistic contexts, that combines Bayesian inference under uncertainty over self-generated and social evidence with causal hypotheses based on formal rules and perceptual heuristics. We validate our model experimentally, showing that it explains how 7- to 10-year-olds' solve a physical puzzle, including the distribution and the types of evidence children sampled. The proposed model outperforms both a purely rule-based Bayesian hypothesis search and a resource-rational random sampling approach. Our results suggest that children implement an internal mechanism for generating and testing a limited number of hypotheses, including formal programmatic rules and heuristics generated from salient problem features to seek more evidence when formal rule generation fails.

Keywords: exploration; information-seeking; hypotheses testing; inference; developmental psychology; computational model; Bayesian model

Introduction

In 'Alice's Adventures in Wonderland', little Alice finds a mysterious bottle labeled *'Drink me'*, and decides to taste it, conjecturing that as long as the bottle is not marked 'poison', it is safe to consume. This reasoning is as funny as it is a perfectly reasonable inference from observation. Like most children, Alice engages with her limited prior knowledge of the world to formulate a causal hypothesis (Gopnik & Schulz, 2004), which in-turn guides her exploration and learning. But where do such hypotheses come from? And how can children use them toward the eventual epistemological goal of learning complex models of the natural world?

Children demonstrate nuanced exploratory strategies and learn through exploration (Bonawitz et al., 2014; Cook et al., 2011; Denison et al., 2013) while simultaneously attending to social cues from expert adults (Legare & Nielsen, 2015). However, even adults may struggle to override their intuitions about salient perceptual features to discover counterintuitive causes (Shtulman & Valcarcel, 2012). For instance, we tend to hold strong (incorrect) intuitions about buoyancy: the density of objects determines whether they will sink or float, but weight is a highly salient feature which hinders the judgments of both children and adults (Brault Foisy et al., 2021).

Figure 1: Experimental procedure used to inform our model. (a) Lineup of boxes and keys during the experiment. (b) The 13 key-tag designs, along with examples of hypothetical rules they may be associated with.

In these cases, being able to learn from examples of incorrect hypotheses failing to produce desired effects is particularly important. Studies with simple learning tasks suggest that children update their beliefs and generate new hypotheses when their own ideas (Bonawitz et al., 2012) or teaching fail (Radovanovic et al., 2024). However, high confidence in the incorrect hypothesis may lead children to attribute failure to personal error rather than to the hypothesis itself. For example, even infants attribute errors to themselves when they believe that a solution works (e.g., it has been successful for others), leading them to seek help rather than explore alternative solutions (Gweon & Schulz, 2011; Lucca et al., 2020). Thus, children's beliefs about the reliability of self-generated evidence could create critical challenges for learning.

Another challenge for learning is generating alternative hypotheses to test. As incorrect intuitions may prevent learners from seeking informative evidence (Markant & Gureckis, 2014; Markant et al., 2016), children may perseverate in par-

tially correct hypotheses to produce interesting effects when new hypothesis generation fails (Meng et al., 2018; Sobel et al., 2022). To facilitate inferences about children's exploration, previous work often primes probabilities of candidate hypotheses through its experimental design (Gopnik & Sobel, 2000). However, the cognitive mechanisms which children use to generate hypotheses and seek evidence in naturalistic contexts containing complex causes remain less explored.

One strategy for guiding exploration away from incorrect solutions, common in machine learning (Kirkpatrick et al., 1983) and developmental models (Giron et al., 2023; Gopnik, 2009), is to inject randomness. This strategy, often referred to as hot temperature exploration, assumes that learners produce a number of truly 'random' actions as they explore. However, psychological and neuroscientific evidence suggests that exploration is influenced by various cognitive biases rather than by true randomness (Kidd et al., 2012; Nickerson, 2002). If children's exploration is not truly random, what drives their apparent stochasticity?

Another approach common in machine learning adopts optimistic heuristics in the face of uncertainty (e.g. (Sutton, 2018; Tang et al., 2024)). By operating on the assumption that latent causes are discoverable, such heuristics can guide a learner to start collecting evidence. They can empower a learner to make progress toward the goal by generating evidence, even if the initial belief is overly optimistic and later requires updating. Indeed, even when adults cannot efficiently formulate a good hypothesis of their own, they can still uncover latent causes from passive evidence, attesting to the potential usefulness of such heuristics (Markant & Gureckis, 2014).

Inspired by extensive, multi-domain evidence for optimism in childhood (Leonard & Sommerville, 2025), we adopt this second perspective, and propose that children's causal hypotheses integrate formal rules with optimistic heuristics. These heuristics may be generated from salient perceptual features (i.e., analogies) by a process similar to how adults use analogy to predict unfamiliar phenomena (e.g. predicting the duration of pharaohs' reign based on modern monarchs; Griffiths and Tenenbaum, 2006). These heuristic hypotheses ensure that children can at least start collecting evidence to progress towards generating formal rules.

Here we propose a Bayesian computational model of a cognitive mechanism by which children formulate and test causal hypotheses in real-world tasks. The model learns the reliability of social and self-generated evidence while integrating optimistic information-seeking heuristics and formal rules. We validate our hypothesis on data from a problem-solving task involving both physical exploration and social information, showing that our model explains children's information-seeking behavior, including (1) an initial phase of testing rules acquired through social learning, (2) persistently testing identical evidence to estimate its reliability, and (3) heuristic evidence search that informs rule generation.

Table 1: Specification of boxes.

Order	Color	Shape	Number of shapes
1	red	moon	1
2	pink	cloud	2
3	cream	diamond	4
4	purple	heart	3
5	teal	triangle	5

Methods

Behavioral data: The Box Task

We base our model on data from a recent study with $N = 100$ 7- to 10-year-old children, who solved a physical puzzle that required opening 5 boxes with 13 keys in limited time (see Radovanovic et al., 2024 for full details). Children were shown a lineup of colored boxes marked with shapes and numbers as well as a set of keys in no specific order whose key-tags were marked with colors and symbols, as shown in Fig.1a. The full set of key-tags is shown in Fig.1b., and box specifications are given in Table 1. Children then viewed a video, in which a teacher demonstrated that the red key opens the red box, and instructed children that all boxes were opened by keys which matched their color. As we sought to understand how children navigate uncertainty in both self-generated and social information, this instruction was designed to be misleading. In reality, the red key was the only colour-matched combination that worked, while **true rule** to open the boxes required matching the number on the key-tag to the number of shapes on the box. Children were given 5 minutes to open all boxes, and were not given performance-based incentives. In the end of the experiment, children completed four **generalization trials** in which they guessed which of four keys opens an unseen box. Three of the suggested keys matched only irrelevant perceptual features of the box, and the fourth key matched the true rule.

Computational Modeling

We now turn to considering how the Box Task may be solved, and introduce definitions central to our modeling framework.

Random sampling with replacement (RSWR) The simplest solution to the Box Task is to randomly sample with replacement attempts from $13 \times 5 = 65$ key-box combinations until all boxes are opened. This solution minimizes cognitive effort, and may be resource-rational if the number of attempts it requires to open all boxes with a high confidence is within the five minutes provided for the experiment.

We can estimate the number of required attempts, assuming perfect box-opening skill, using a Binomial distribution: Let random variable X, the number of successes in n trials. $X \sim \text{Binomial}(n, p)$, where $p = \frac{1}{13}$ is the probability of success in a single trial. We compute the probability of observing at least 5 successes in n trials, $P(X \geq 5) \geq 0.90$. Since $P(X \geq 5) = 1 - P(X \leq 5)$, we will solve an equivalent equa-

tion for n: $P(X < 5) \leq 0.10$, where $P(X < 5)$ is the cumulative probability of having 0, 1, 2, 3, or 4 successes in n trials. So, $P(X < 5) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$, where each element is computed using the binomial probability formula:
$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = p^k (1-p)^{n-k} \frac{n!}{k!(n-k)!}$. Solving this equation for n gives 102 attempts.

Given the 5-minute long experiment this solution is resource-rational assuming someone can spend less than 3 seconds per attempt, and has a **motor error rate** of zero, defined as the probability that the correct key-box combination is unsuccessful (e.g. due to someone's box opening skill). A motor error rate of zero is likely within the ability of 50 of 100 children, who had zero error rates (see Fig.2f. RSWR makes two predictions. First, it predicts that children should exhibit no preferences during generalization, such as preferences to generalize toward the true rule, or toward matching perceptual features. Second, the rates of sampling different box-and key combinations should not be statistically different.

Bayesian Experiment Design (BED) Next, we describe a BED framework that specifies the optimal experimental design (Lindley, 1956; Rainforth et al., 2024) for how a Bayesian agent should solve the Box Task.
Let $\mathcal{H} = \{H_1, H_2, \ldots, H_n\}$, the set of all possible hypotheses with prior probabilities $\{P(H_1), P(H_2), \ldots, P(H_n)\}$. A BED agent iteratively updates the probabilities of the hypotheses based on evidence, where the posterior probability of H_i given evidence $\mathbf{e} = (e_1, e_2, \ldots, e_n)$, is given by the prior probability of $P(H_i)$, and the likelihood of observing each piece of evidence sampled under H_i: $P(H_i|\mathbf{e}) \propto P(H_i)P(\mathbf{e}|H_i)$.

In the context of the Box Task, a hypothesis H_i is given by a list of tuples, $(\text{key}_k, \text{box}_b, p)$, where key_k is a key, box_b is a box, and p is the probability that the key opens the box. Each piece of evidence $e_j = (\text{key}_k, \text{box}_b, \text{outcome})$ is a tuple of key, box, and outcome (TRUE or FALSE). Without the loss of generality, we assume $\forall b \in [1..5]\ \exists! k \in [1..13] :\ key_k \rightarrow box_b$ and $p = 1$. That is, for each box there exists one and only one deterministic key that opens it. Instead of separately modeling the reliability of individual keys, we model inference about the reliability of the participant's *self-generated evidence, or own skill* parameterized by θ. We model the belief that a correct key successfully opens the corresponding box in a single attempt: $P(H_i|\mathbf{e}) \propto P(H_i)P(\mathbf{e}|H_i)P(\theta)$. The skill parameter $P(\theta)$ is learned jointly with learning posteriors of the hypotheses. The likelihood of observing e_j under H_i depends on whether H_i suggests key k should open box b:

$$P(e_j \mid H_i, \theta) = \begin{cases} \theta & \text{if success} \wedge \text{ by } H_i\ k_{e_j} \rightarrow b_{e_j}, \\ 1-\theta & \text{if fail} \wedge \text{ by } H_i\ k_{e_j} \rightarrow b_{e_j}, \\ 0 & \text{if success} \wedge \text{ by } H_i\ k_{e_j} \not\rightarrow b_{e_j}, \\ 1 & \text{if fail} \wedge \text{ by } H_i\ k_{e_j} \not\rightarrow b_{e_j}. \end{cases}$$

Then, the posterior for θ and $P(H_i)\ \forall\ H_i \in \mathcal{H}$, is computed as follows.
First, $P(\theta_i \mid \mathbf{e}, H_i) \propto \text{Beta}(\alpha + \text{successes}_{H_i}, \beta + \text{failures}_{H_i})$, computes what θ would be if each hypothesis were true.
Next, the posterior probability of each hypothesis is given by:
$P(H_i, \theta \mid \mathbf{e}) \propto P(\mathbf{e} \mid H_i, \theta)P(H_i)P(\theta)$, where
$P(\mathbf{e} \mid H_i, \theta) = \prod_{j=1}^{n} P(e_j \mid H_i, \theta)$.
Finally, the combined posterior for θ:
$P(\theta \mid \mathbf{e}) = \sum_i P(\theta_i \mid \mathbf{e}, H_i)P(H_i \mid \mathbf{e})$

Modeling Reliability. This defines a Bayesian learning problem, where the learner formulates hypotheses about possible box-key pairings, and conducts experiments to identify the correct one. The choice of priors reflects the salience of each hypotheses, and affects learning efficiency. We model the **reliability of social evidence** by a free parameter ω to control the relative salience of the taught rule.

Further, we model θ as a random variable with a prior distribution $P(\theta) \sim Beta(\alpha, \beta)$. For example, initializing $\alpha = 19, \beta = 1$ reflects a prior belief that the correct key-box combination is successful 0.95 of the time. While θ controls the agent's persistence at testing a failing key-box combination, it is not equivalent to the agent's actual motor skill. A misalignment between the subjective skill and the actual skill can potentially impair learning by causing excessive testing of failing hypotheses or insufficient persistence when correct key-box combinations fail.

Sampling Choices. The BED agent samples evidence by maximizing information gain of a sample as the Kullback-Leibler divergence between the prior and posterior distributions over hypotheses (the expected reduction in entropy that would result from each piece of evidence):
$IG(e_j) = H(\text{Prior}) - \mathbf{E}_{P(e_j)}[H(\text{Posterior} \mid e_j)]$, where
$H(\text{Prior}) = -\sum_{i=1}^{n} P(H_i \mid \theta) \log P(H_i \mid \theta)$,
$H(\text{Posterior} \mid e_j) = -\sum_{i=1}^{n} P(H_i \mid e_j, \theta) \log P(H_i \mid e_j, \theta)$,
where $\mathbf{E}_{P(e_j)}$ is the expectation of the posterior entropy across the possible outcomes of the e_j (success, fail).

Formal rules and optimistic heuristics **Generating Hypotheses.** In theory BED is computationally intractable, as the set of all possible key-box hypotheses is infinite. Our simplifying assumption of one and only one key per box still leaves us with $\binom{13}{5} = \frac{13!}{5!(13-5)!} = \frac{13!}{5!\cdot 8!} = 154440$ hypotheses. Previous work proposes a Sampling Hypothesis (Denison et al., 2013), which suggests that children sequentially sample a small number of new hypotheses from an internal proposal distribution with a probability propositional to their posterior. Below, we consider applications of the Sampling Hypothesis to our task.

One way to approach hypothesis generation is to sample hypotheses as **formal rules** – that is, program-like expressions applied to all boxes – `box.color = key.color; box.shape = key.shape; box.order = key.number; count(box.shape) = key.number`[1]. Another approach is to sample hypotheses as sequences of key-box mappings generated based on perceptual similarities(Griffiths & Tenenbaum, 2006). To generate a heuristic we sample a key for each unopened box with a probability pro-

[1]The last rule in this list is the true rule. The first rule is the color rule taught to the participants.

portional to the pair's similarity – for example – `red box` → `red`; `pink box` → `grey2`; `cream box` → `orangediamond`; `purple box` → `greenheart`; `teal box` → `yellowtriangle`.

Sampling Choices. We implement our model as approximate inference by a variant of Adaptive Sequential Markov Chain Monte Carlo Sampling (Del Moral et al., 2006), using BED with a small working set of hypotheses, while modeling reliability of self-generated and social evidence as described above. To do this, we initialize the working set with just two hypotheses: a hypothesis sampled from the proposal distribution, and one *other* hypothesis. The *other* hypothesis is defined in contrast to the hypotheses currently in the working set, implying that none of the working hypotheses are true, and controlling the probability that a new rule will be generated. The probability of the *other* hypothesis increases with negative evidence against the working hypotheses, leading the model to periodically samples new hypotheses from the proposal distribution.

The proposal distribution can generate an infinite number of hypotheses, including formal rules and feature-matching heuristics. To model social priors, our implementation manually defines formal rules along with their prior probabilities. The proposal distribution returns a formal rule with a probability $1-r$, and with probability r returns a heuristic generated by matching perceptual features. The likelihood that a heuristic includes a given key-box combination is proportional to the perceptual similarity, defined as a salience-weighted linear combination of features. For example, the combination ($key = red_1$, $box = red_1$) is more likely than ($key = red_1$, $box = teal$). The generated proposals are *accepted* if they agree with existing evidence, and *rejected* otherwise. In practice, this leads the model to initially exploring highly likely and salient rules, and increasingly switch to heuristics until either (1) all boxes are opened by a heuristic, or (2) the correct rule is discovered, leading the model to quickly open the remaining boxes.

Combining social knowledge, skill, and heuristics. Different settings of subjective skill θ, social prior ω, and heuristic generator r produce different types of behavior, as shown in Fig.3. Setting $r = 1$ reduces the model to a biased random sampler that reasons about own skill, but can not reason about formal, generalizable rules. This model would require more trials to unlock the boxes than RSWR, given that most feature matching combinations are incorrect. Setting $r = 0$ reduces the model to reasoning about formal rules only. The model configuration $r = 0, \mathbb{E}[\theta] \approx 1, \omega = 1$ will be analogous to the Sampling Hypotheses (Denison et al., 2013). Different settings of ω control the effect of teaching, that is, the model's tendency to initially test the taught rule. Given incorrect instructions, model configurations with a lower ω will need fewer attempts to open all the boxes than configurations with a higher ω. Setting $0 < \theta < 1$ will cause the model to persistently test failing key-box combinations, and as $\mathbb{E}[\theta] \to 1$ the model will increasingly trust self-generated evidence.

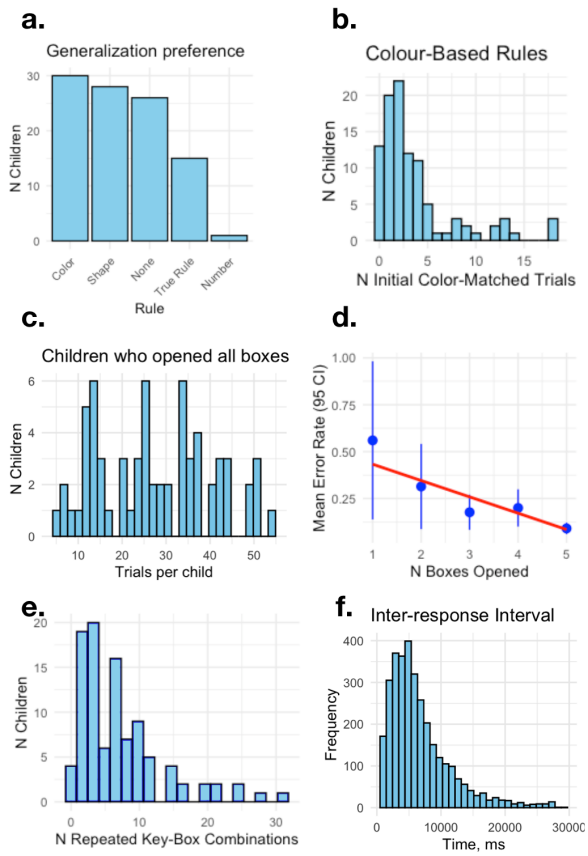

Figure 2: Descriptive statistics over children's' performance in the Box task. (a) Preferences during generalization. (b) The histogram of the number of color-matched opening attempts at the beginning of the experiment. (c) The histogram of the number of trials in the experiment, for children who opened all boxes (N=66). (d) Individual motor error rates plotted against the number of boxes opened. (e) (f) Distribution of inter-response intervals between opening attempts, according to manually coded videos timestamps.

Behavioral metrics

We validate the models' ability to predict behavior using three metrics. First, we consider children's responses to generalization trials. We consider children to express a significant **generalization preference**, if they selected the same type of response (out of 4 possible response types) on 3 or more generalization trials, as the probability of this repeated choice is significantly smaller than chance ($0.25^3 = 0.015 < 0.05$). Figure 2a shows the distribution of generalization preferences across children, according to which 15 of the 100 children generalize to the true rule, and another 60 generalize to one of the perceptual features. Second, it needs to predict **the number of initial color-matched attempts** at the beginning of the experiment (Fig.2b.), which are associated with seeking evidence toward the taught rule. Lastly, the model needs to predict **the number of trials needed to open all boxes** (Fig.2c.), which reflects children's overall performance in the

Box Task.

Results

We find that children's behavior cannot be explained by RSRW, since the majority of children express a significant preference during generalization either toward perceptual features or toward the true rule. Further, 87 of 100 of children begin the experiment by a color-marched key-box combination, which is significantly greater ($\chi^2(1) = 98.09, p < .0001$) than the 16 of 100 children predicted by RSRW.

Next, we note that when restricted to only formal rules ($r = 0$) and assuming perfect subjective skill ($\mathbb{E}[\theta] \approx 1$) our proposed model produces behavior equivalent to that of the Sampling Hypothesis (Denison et al., 2013), while also introducing a free parameter ω that controls how social priors influence subjective probabilities of the hypotheses. That is, our model configured with ($r = 0, \mathbb{E}[\theta] \approx 1, \omega > 0$) will test a small number of formal rules, where each rule is sampled with a probability proportional to its subjective and social salience. When given a high social prior, this model configuration explains children's tendency to begin the experiment with color-matched key-box combinations. That is, its predictions align with the observation that 87 of 100 children begin the experiment by attempting a color-matched combination, and persist in testing color-matched combinations during the initial trials (see panels Fig.3a. and Fig.3b.). We find that parametrized with $\omega = 10$, the model's distribution of initial color-matching attempts aligns with children (Welch Two Sample t-test $t(63.27) = 0.84, p = .4$), however, the model still needs fewer trials in total compared to the $N = 40$ children who opened all boxes with zero error rate (Welch Two Sample t-test $t(40.47) = 8.1, p < .0001$). Therefore, sampling formal rules with a probability biased by a social prior can explain children's initial evidence seeking, but does not explain what they do after the taught rule fails.

One factor that could contribute to children needing more trials compared to the models is that they may attribute failure to their own lack of skill, regardless of their actual motor error rate. Indeed, all but 4 children repeatedly attempt identical key-box combinations on consecutive trials (M = 7.12, SD = 6.27, see Fig.2e.). To capture this behavior, we parametrize our model with a subjective skill parameter θ and set $\omega = 7$ (fitted to the children's rate of testing a color-matched combination on the first trial). For demonstration purposes, let $\theta = Beta(5,1)$, giving the subjective probability of success $\mathbb{E}[\theta] = 0.8$. We find that our model configured with ($\omega = 7, \theta = Beta(5,1), r = 0$) needs more attempts than children to discard incorrect color-matched rules (Welch Two Sample t-test $t(186.77) = 5.96, p < .0001$), but still needs fewer attempts in total to complete the task (Welch Two Sample t-test $t(85.79) = 4.92, p < .0001$) (see Fig.3c). Further, any model configuration limited to formal rules only ($r = 0$) predicts generalization to the true rule in every instance where it opens all boxes. In contrast, only 13 of the 66 children who opened all of the boxes generalized to the true rule, suggesting that to reproduce behavior we need $r > 0$.

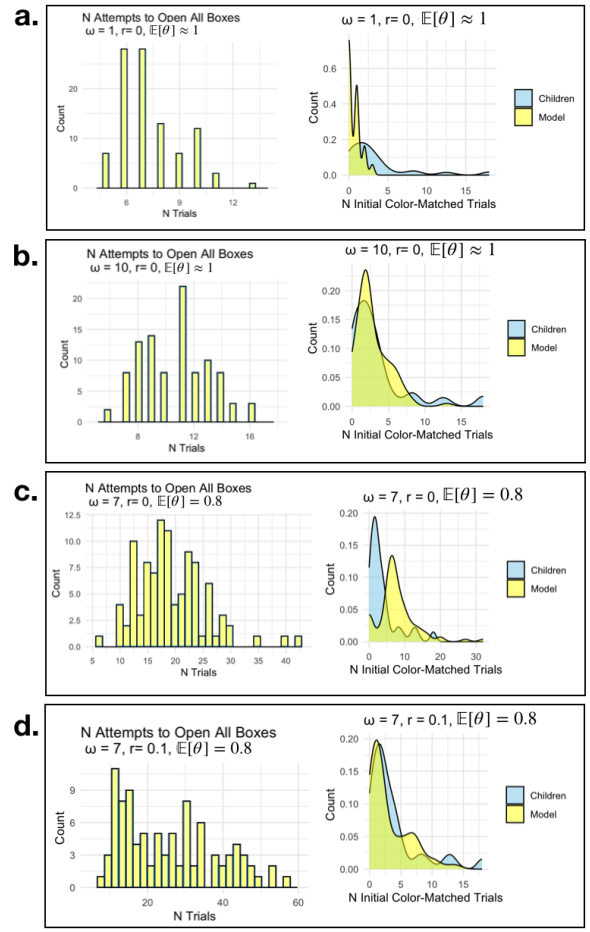

Figure 3: Simulated distributions of number of attempts needed to open all boxes, and the distributions of initial color-matched attempts across model configurations. (a) $\omega = 10$, $\mathbb{E}[\theta] \approx 1, r = 0$, (b) $\omega = 1$, $\mathbb{E}[\theta] \approx 1, r = 0$. (c) $\omega = 7$, $\theta = Beta(5,1), r = 0$. (d) $\omega = 7$, $\theta = Beta(5,1), r = 0.1$. Plots (a,b) compare models to children with zero error rates who opened all boxes (N=40), and (c,d) compare to all (N=66) children who opened all boxes.

Finally, we observe that while children attempt all keys on all boxes, including combinations inconsistent with any formal rule (e.g. *whiteseven key, pink box*), children's information seeking and generalization are consistent with feature-matching heuristics rather than hot temperature exploration. First, children continue to test combinations that match a given feature (for example, shape) after finding positive evidence matching *a different feature* (e.g., box order or color). This behavior is inconsistent with searching over formal rules, since positive evidence for a given rule would lead a Bayesian agent to discard inconsistent alternatives. Second, during generalization the majority of children express a preference for feature-matched combinations, including children who opened all boxes. We find that our model parametrized

with $\omega = 7, r = 0.1, \theta = Beta(5,1)$ (see Fig.3d) is consistent with both the number of trials needed to open all boxes (Welch Two Sample t-test t(148.9) = 0.76, p = 0.5), and with the number of initial color-matched attempts (Welch Two Sample t-test t(184.43) = 1.2, p = 0.2), suggesting that all three parameters are necessary to accurately model children's behavior in the Box Task. More information can be found in Supplemental Materials: https://osf.io/qw2e7/

Discussion

We describe a computational model for forming hypotheses and searching for information in a problem-solving task mirroring counterintuitive, real-world problems. We validate our model on experimental data from 7-to-10 year old children, showing that its predictions closely align with children's behavior in terms of types of evidence sampled, and the amount of evidence needed to complete the task. Our results reveal that children use sophisticated exploratory strategies to learn latent causal hypotheses, by combining social knowledge, optimistic information-seeking heuristics, formal rules that summarize partial information, and inductive inference under uncertainty over the social and self-generated evidence.

Our results suggest that children are able to make progress toward learning complex latent rules by testing heuristic hypotheses based of perceptual features, in contrast to the popular view that children tend to engage in high levels of random exploration (Giron et al., 2023; Gopnik, 2020). This behavior resembles optimistic learning objectives, widely applied in machine learning, where they are shown to theoretically and empirically promote goal-directed exploration (e.g. see Tang et al., 2024 for a recent example). Here we argue that children's tendency to come up with causal hypotheses based on perceptual analogies may arise from a similar optimistic heuristic objective.

Our results suggests that in addition to representing the learned posterior by sampling from it (Sanborn & Chater, 2016), children may be employing a metacognitive generative approach to come up with formal rules by summarizing partial evidence (e.g. see recent work alluding to such a process in human adults Ellis, 2023). As a proof of concept of this view, our modeling approach builds on the computational principles of the mind as a Bayesian sampler (Denison et al., 2013; Sanborn & Chater, 2016) with the hypothesis generation part in our implementation currently limited to heuristics. In an ongoing work we also examine ways by which both formal rules and heuristic may be generated online during inference, using a program synthesis approach inspired by (Piriyakulkij et al., 2024).

Our modeling approach explains children's persistence – that is, repeated attempts to test a failing hypothesis – by a desire to estimate reliability of evidence. Informally, our model captures the intuition that if a specific key-box combination fails, participants might attribute the outcome not only to the tested hypothesis, but also to unreliable evidence. Our result is consistent with a previously observed preference to stick with current hypotheses (Bonawitz et al., 2014), and is similar to the widely reported phenomenon of confirmation bias – or repeated demonstration of one's current hypothesis. Importantly, we find that participants repeat positive tests only rarely, while exhibiting increasing persistence while repeatedly testing negative combinations (see SI). While this behavior may seem opposed to confirmation bias, the differences in engagement with positive and negative evidence in our task could be explained by time constraints. Consistent with our interpretation, confirmation bias has been previously explained by search for invariance, where positive tests generate information about the degree to which an observed relationship holds across time and contexts (Lapidow & Walker, 2020).

While our current results compare stochastic aspects of the model's behavior with children as a population, more insights can be gained by fitting this model to individuals to analyze the underlying computational phenotypes(Schurr et al., 2024). For instance, in future work we intend to examine the relationship between children's learning effectiveness, and the extent of alignment between their subjective skill and their actual motor error rate. On the one hand, optimistically overestimating one's own abilities may help children by maintaining motivation and creating opportunities for learning as children attempt difficult tasks (e.g., Chang et al., 2009). Indeed, children are more likely to persist on difficult tasks when their effort is expected to lead to success (e.g., Lucca et al., 2020; Solby et al., 2021). On the other hand, children who are overly optimistic about their performance may learn less from negative outcomes (Habicht et al., 2022) which may ultimately hinder them in identifying hidden rules.

Further individual analysis could consider whether children's ability to synthesize rules from evidence and make accurate predictions about their performance (e.g., Leonard and Sommerville, 2025) improves with age, and examine factors that lead children to place a higher priors on the teacher's instructions, as suggested in the original publication of the Box TaskRadovanovic et al., 2024. The validity of our modeling approach may be further tested by adapting it to other tasks, with and without time constraints, and by comparing children with human adults.

In summary, our computational model provides a unified framework for understanding how children form causal hypotheses in complex problem-solving tasks with unreliable evidence. While we do not argue that this model tells us all there is to learning world-models, it takes a novel step toward understanding how children skillfully balance social priors, optimistic heuristics, and inductive reasoning over latent causes to navigate uncertainty. Our approach highlights the sophistication of children's hypothesis forming strategies, which may resemble optimistic objectives in machine learning adapted to developmental constraints.

Acknowledgments

We are grateful to Top Piriyakulkij (Cornell University) for insightful discussions.

CRediT Author Contributions

M.K: Conceptualization, Methodology, Formal Analysis, Visualization, Investigation, Writing: Original Draft & Editing, Project Administration, Supervision
M.R Writing Review & Editing
J.S.: Conceptualization, Writing Review
S.D.: Visualization (Supplementary Figure S5)
R.G: Poster presentation

References

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive psychology*, *74*, 35–65.

Bonawitz, E., van Schijndel, T. J., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*(4), 215–234.

Brault Foisy, L.-M., Ahr, E., Blanchette Sarrasin, J., Potvin, P., Houdé, O., Masson, S., & Borst, G. (2021). Inhibitory control and the understanding of buoyancy from childhood to adulthood. *Journal of Experimental Child Psychology*, *208*, 105155.

Chang, E. C., Chang, R., & Sanna, L. J. (2009). Optimism, pessimism, and motivation: Relations to adjustment. *Social and Personality Psychology Compass*, *3*(4), 494–506.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*(3), 341–349.

Del Moral, P., Doucet, A., & Jasra, A. (2006). Sequential monte carlo samplers. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *68*(3), 411–436.

Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children's causal inferences: The sampling hypothesis. *Cognition*, *126*(2), 285–300.

Ellis, K. (2023). Human-like few-shot learning via bayesian reasoning over natural language. *Advances in Neural Information Processing Systems*, *36*, 13149–13178.

Giron, A. P., Ciranka, S., Schulz, E., van den Bos, W., Ruggeri, A., Meder, B., & Wu, C. M. (2023). Developmental changes in exploration resemble stochastic optimization. *Nature Human Behaviour*, *7*(11), 1955–1967.

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. Random House.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190502.

Gopnik, A., & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in cognitive sciences*, *8*(8), 371–377.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child development*, *71*(5), 1205–1222.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological science*, *17*(9), 767–773.

Gweon, H., & Schulz, L. (2011). 16-month-olds rationally infer causes of failed actions. *Science*, *332*, 1524.

Habicht, J., Bowler, A., Moses-Payne, M. E., & Hauser, T. U. (2022). Children are full of optimism, but those rose-tinted glasses are fading—reduced learning from negative outcomes drives hyperoptimism in children. *Journal of Experimental Psychology: General*, *151*(8), 1843–1853.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.

Kirkpatrick, S., Gelatt Jr, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, *220*(4598), 671–680.

Lapidow, E., & Walker, C. M. (2020). *The search for invariance: Repeated positive testing serves the goals of causal learning*. Springer.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Science*, *19*(11), 688–699.

Leonard, J. A., & Sommerville, J. A. (2025). A unified account of why optimism declines in childhood. *Nature Reviews Psychology*, *4*(1), 35–48.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, *27*(4), 986–1005.

Lucca, K., Horton, R., & Sommerville, J. A. (2020). Infants rationally decide when and how to deploy effort. *Nature Human Behaviour*, *4*, 372–379.

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94.

Markant, D. B., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive science*, *40*(1), 100–120.

Meng, Y., Bramley, N. R., & Xu, F. (2018). Children's causal interventions combine discrimination and confirmation. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, 762–767.

Nickerson, R. S. (2002). The production and perception of randomness. *Psychological review*, *109*(2), 330.

Piriyakulkij, W. T., Langenfeld, C., Le, T. A., & Ellis, K. (2024). Doing experiments and revising rules with natural language and probabilistic reasoning. *arXiv preprint arXiv:2402.06025*.

Radovanovic, M., Yucer, E., & Sommerville, J. A. (2024). Girls persist more but divest less from ineffective teaching than boys. *Journal of Experimental Psychology: General*.

Rainforth, T., Foster, A., Ivanova, D. R., & Bickford Smith, F. (2024). Modern bayesian experimental design. *Statistical Science*, *39*(1), 100–114.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in cognitive sciences*, *20*(12), 883–893.

Schurr, R., Reznik, D., Hillman, H., Bhui, R., & Gershman, S. J. (2024). Dynamic computational phenotyping of human cognition. *Nature Human Behaviour*, 1–15.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, *124*, 209–215.

Sobel, D. M., Benton, D., Finiasz, Z., Taylor, Y., & Weisberg, D. S. (2022). The influence of children's first action when learning causal structure from exploratory play. *Cognitive Development*, *63*, 101194.

Solby, H., Radovanovic, M., & Sommerville, J. A. (2021). A new look at infant problem-solving: Using deeplabcut to investigate exploratory problem-solving approaches. *Frontiers in Psychology*, *12*.

Sutton, R. S. (2018). Reinforcement learning: An introduction. *A Bradford Book*.

Tang, H., Key, D., & Ellis, K. (2024). Worldcoder, a model-based llm agent: Building world models by writing code and interacting with the environment. *arXiv preprint arXiv:2402.12275*.

Supplementary Information

Table S1: Specifications of all formal rules considered in our model. The keys are listed for boxes in order they appear in the line up: red, pink, cream, purple, teal. We consider multiple rules based on color similarity, given that the colors of boxes and keys did not match exactly. The teal box could be seen as matching either the blue or a green key tag, and the white box as matching either the white or a yellow key.

	Rule	List of Keys
1	colour1	red_1, pink_6, white_7, purple_arrow, blue_star
2	colour2	red_1, pink_6, white_7, purple_arrow, green_heart
3	colour3	red_1, pink_6, white_7, purple_arrow, green_3
4	colour4	red_1, pink6, yellow_triangle, purple_arrow, blue_star
5	colour5	red_1, pink_6, yellow_triangle, purple_arrow, green_heart
6	colour6	red_1, pink_6, yellow_triangle, purple_arrow, green_3
7	colour7	red_1, pink_6, yellow_5, purple_arrow, blue_star
8	colour8	red_1, pink_6, yellow_5, purple_arrow, green_heart
9	colour9	red_1, pink_6, yellow_5, purple_arrow, green_3
10	colour10	red_1, pink_6, grey_2, purple_arrow, blue_star
11	colour11	red_1, pink_6, grey_2, purple_arrow, green_heart
12	colour12	red_1, pink_6, grey_2, purple_arrow, green_3
13	colour13	red_1, pink_6, grey_cloud, purple_arrow, blue_star
14	colour14	red_1, pink_6, grey_cloud, purple_arrow, green_heart
15	colour15	red_1, pink_6, grey_cloud, purple_arrow, green_3
16	order	red_1, grey_2, green_3, orange_4, yellow_5
17	shape	– , grey_cloud, orange_diamond, green_heart, yellow_triangle
18	ground truth	red_1, grey_2, orange_4, green_3, yellow_5

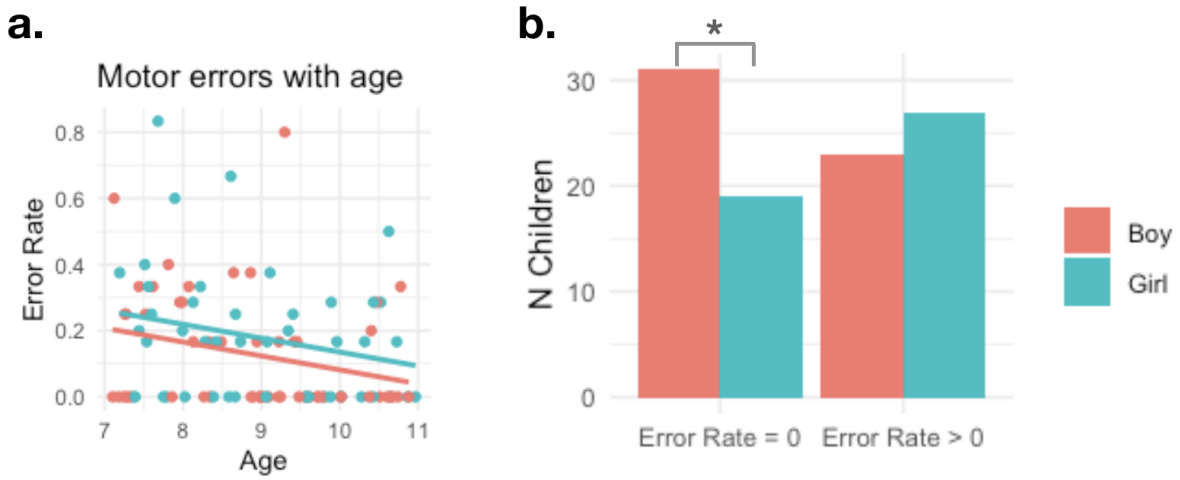

Figure S1: Error rates of children in the experiment when opening a box with its correct key with age and gender. (a.) Individual error rates decrease with age (Pearson $r = -.25, p = .1$). Half of the children (N=50) have a zero error rate, and the remaining children have error rates between 0.2 to 0.8. (b.) More boys than girls had a zero error rate (proportions: 0.62 and 0.38, $\chi^2(1,50) = 4.84$, $p = 0.028$.).

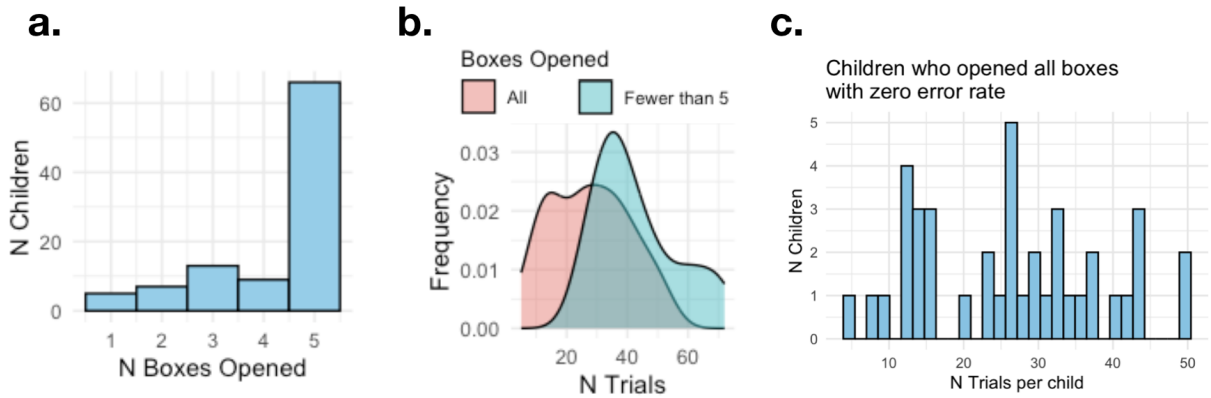

Figure S2: Children's performance statistics. (a) The distribution of children's performance across the experiment measured as the total number of boxes opened within 5 minutes. (b) The distribution of the total number of opening attempts made by children during the entire experiment. Children who opened all boxes (solved the puzzle) and those who did not are shown separately. Children who did not open all boxes made on average more attempts. (c) The number of attempts by children who opened all boxes, and hade a zero error rate (N=40).

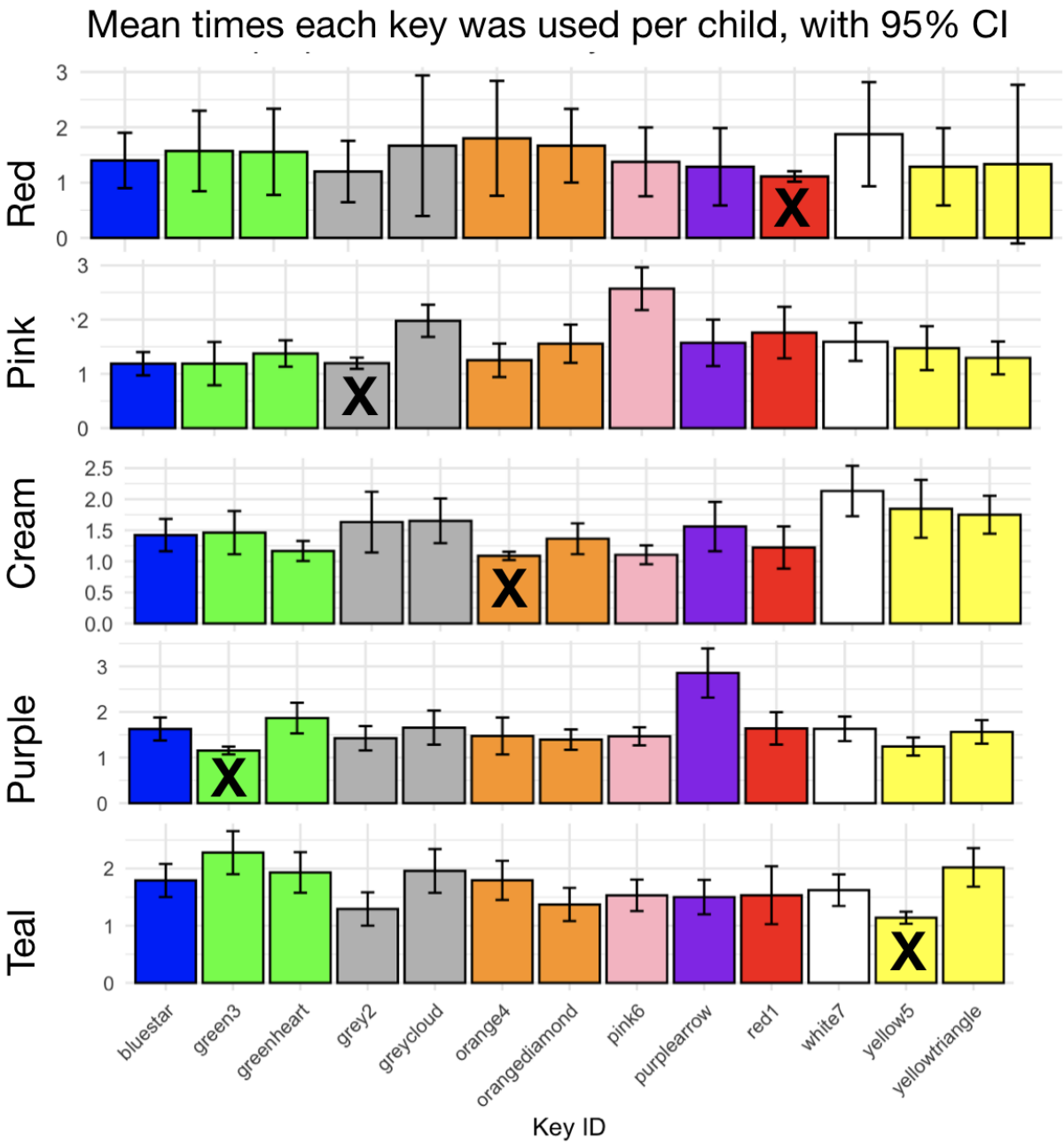

Figure S3: The distribution of keys tried by children on each of the five boxes. The correct key for each box is labeled with an ”X”.

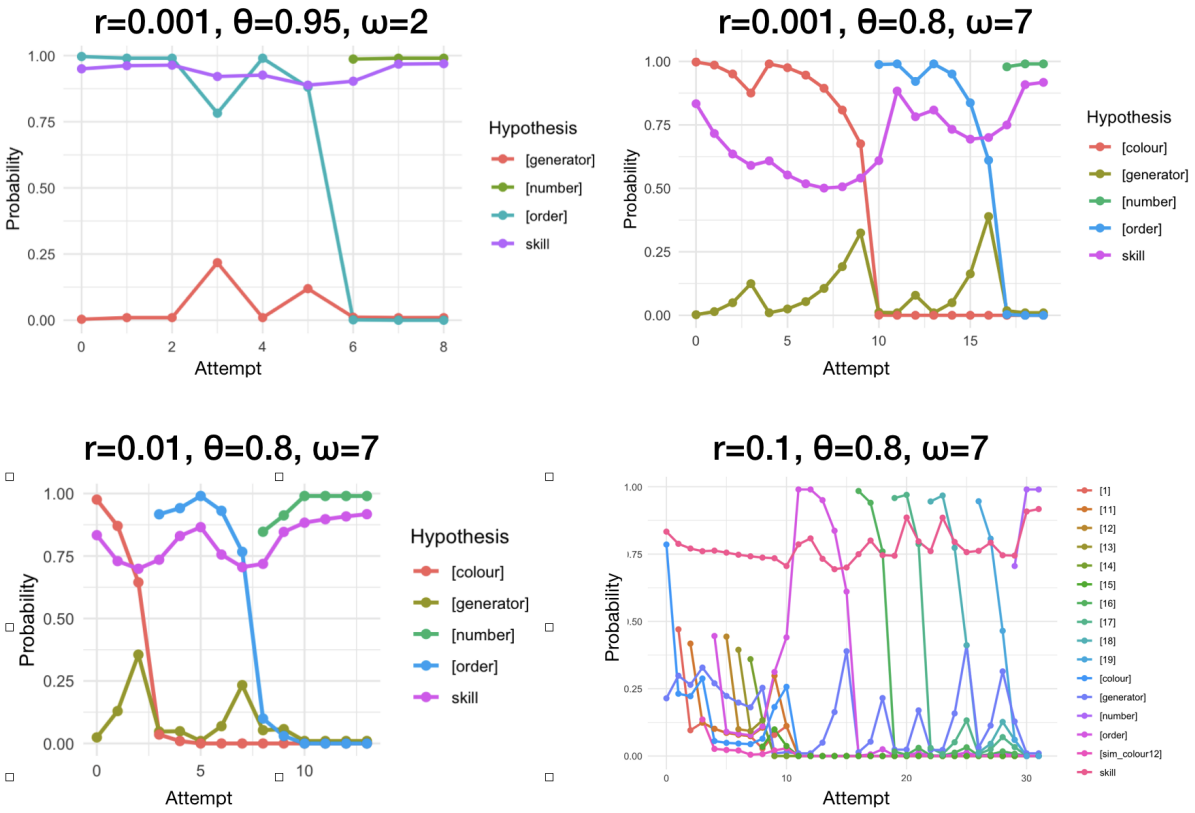

Figure S4: The probabilities of various hypotheses, subjective skill, and the probability of generating a new rule over time, produced by the model across four simulations with different parameters. Heuristic rules are labeled with numbers. Formal rules are labeled with names that describe their function. The legend colors change between plots, as each simulation is stochastic, with a different and unique set of sampled rules.

Reasoning in language

Here we consider an alternative to approximate Bayesian inference: an intuitive process of reasoning in natural language, that may not involve explicitly articulating formal rules. To model such a reasoning, we prompted GPT-4 with the description of all keys, boxes, and evidence, asking it to respond with new (key, box) samples. We then simulate an opening attempt using the ground truth rule, and include the history of all opening attempts, along with their results in subsequent prompts. This process continues until all 5 boxes are opened, resulting in $M = 13.5, SD = 3.2$ opening attempts (see Fig.S7).

Importantly, the prompt gives LLM access to full descriptions of boxes and keys, unlike children, who may not notice all numbers or shapes on the boxes. While in theory LLM could attempt any key-box combination, in practice it produces behavior consistent with our model limited to formal rules, a high social prior, and a high skill ($\theta = 0.95, r = 0, \omega = 10$).

We compared the number of trials used by GPT-4 to children who opened all boxes and had a zero error rate (N = 40) to control for motor error rates possibly affecting the number of trials that children need. We found that children needed significantly more attempts compared to this model (Welch Two Sample t-test $t(46.16) = -6.44, p < .0001$, means 26.5 and 13.5 respectively).

Table S2: A sequence of LLM responses listing (key, box) attempts and their outcomes in one simulation.

Attempt	Key	Box	Opened
1	red_1	red	1
2	pink_6	pink	0
3	grey_cloud	pink	0
4	grey_2	pink	1
5	orange_4	cream	1
6	purplearrow	purple	0
7	green_3	purple	1
8	triangle_yellow	teal	0
9	diamond_orange	cream	0
10	yellow_5	teal	1

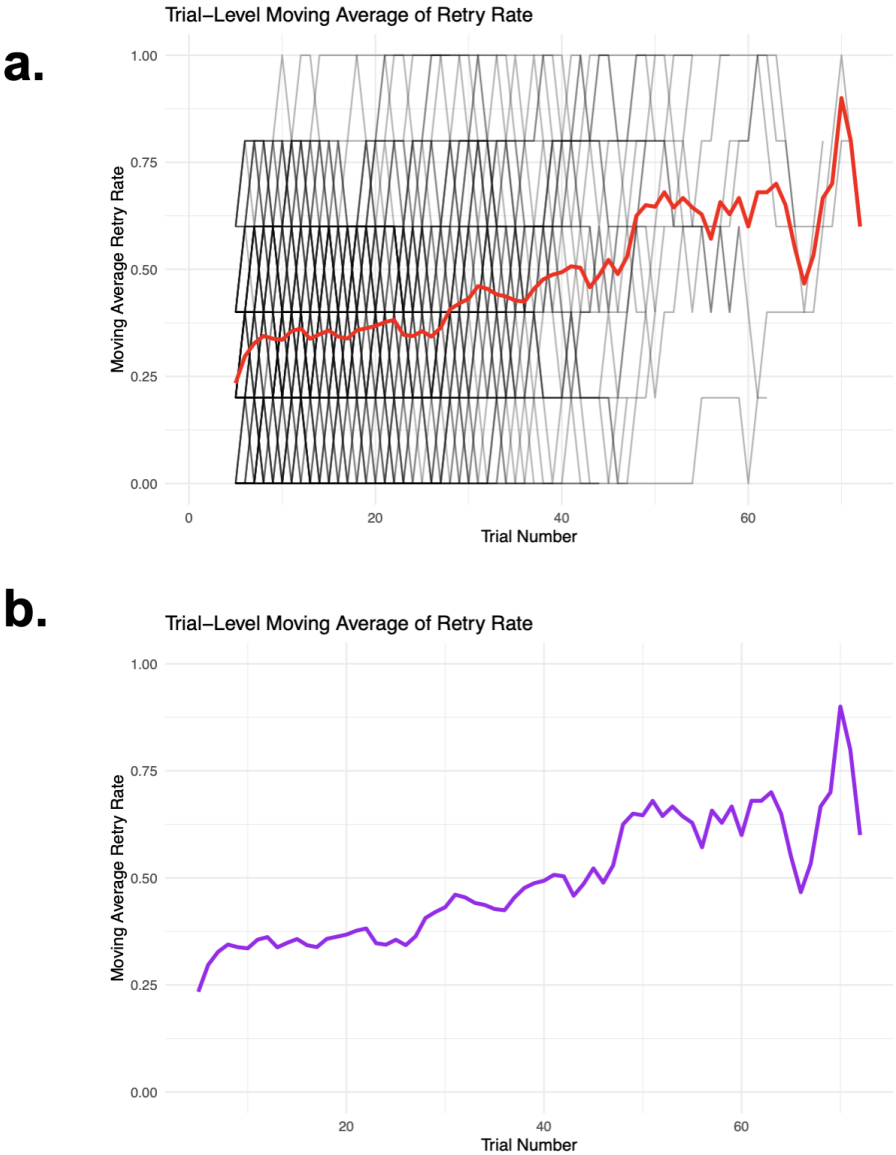

Figure S5: Trial-level moving average of key-box retry rates using a 5-trial window. **(a)** Each thin black line represents a single participant's retry rate, smoothed across trials using a 5-trial moving average. The red line indicates the group-level average across participants at each trial index. (**b**) The same group-level moving average is shown without individual trajectories.

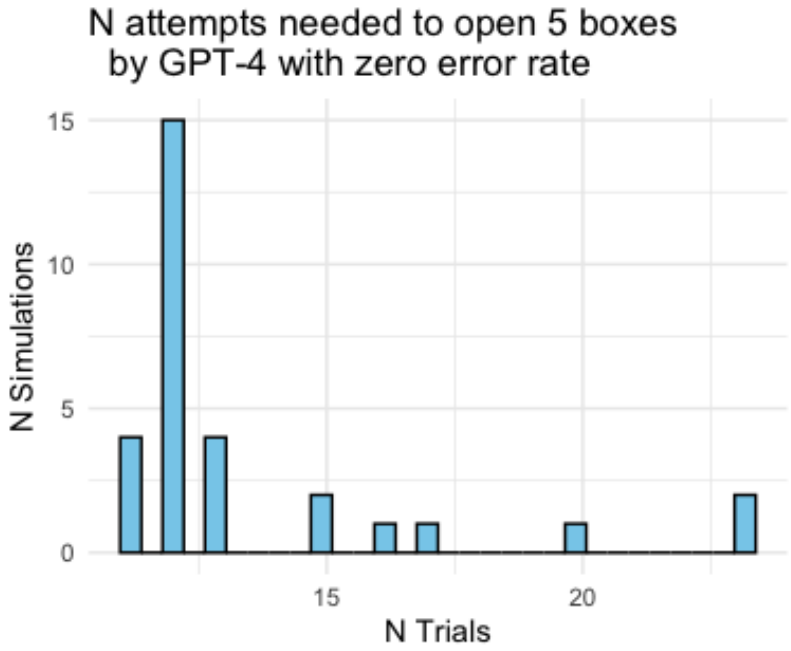

Figure S6: Number of opening attempts needed by GPT-4 to open 5 boxes with a zero error rate.

System Prompt:

You are an intelligent agent playing a game.
Your task is to open 5 boxes using 13 keys in fewest attempts.
You do not need special skills to play this game. This game can be played by an 8-12 year old child.

User Prompt:

For each box there is a key that opens it, so the goal of the game is to find the right key for each box.

You have a demonstration video from a teacher telling you how to open all boxes. In the video, the teacher says: ”I’m going to show you the right way to unlock the doors. To open the doors, you have to use a key that matches the color of the box. So, to open this red box, I’m going to use this red key. Great, now you can open all the doors!”

Here are the boxes:

The red box has 1 moon shape.
The pink box has 2 cloud shapes. Each cloud is numbered from 1 to 2.
The cream (a color between yellow and white) box has 4 diamond shapes. Each diamond is numbered from 1 to 4.
The purple box has 3 heart shapes. Each heart is numbered from 1 to 3.
The teal (a color between green and blue) box has 5 triangle shapes. Each triangle is numbered from 1 to 5.

Here are the 13 keys (in no specific order):

The red1 key has a red key-tag, with number 1.
The pink6 key has a pink key-tag, with number 6.
The grey2 key has a gray key-tag, with number 2.
The greycloud key has a gray key-tag, with a cloud shape.
The orange4 key has an orange key-tag, with number 4.
The green3 key has a green key-tag, with number 3.
The bluestar key has a blue key-tag, with a star shape.
The yellow5 key has a yellow key-tag, with number 5.
The greenheart key has a green key-tag, with heart shape.
The white7 key has a white key-tag, with number 7.
The triangleyellow key has a yellow key-tag, with a triangle shape.
The diamondorange key has an orange key-tag, with a diamond shape.
The purplearrow key has a purple key-tag, with an arrow shape.
Here are the key box combinations you already tried and their outcomes:

[Evidence appended here]

Now is your turn to open the boxes. Which key will you try next?
Please respond in the format ”key, box” (e.g. ”red1, red”) and do not include any other text in the response.

Figure S7: Prompt given to GPT-4 to elicit (key, box) responses.