Received: 14 June 2019 | Revised: 27 July 2020 | Accepted: 28 July 2020
DOI: 10.1111/desc.13026

SHORT REPORT

Developmental Science WILEY

Systematic exploration and uncertainty dominate young children's choices

Nathaniel J. Blanco | Vladimir M. Sloutsky

Department of Psychology, The Ohio State University, Columbus, OH, USA

Correspondence
Vladimir M. Sloutsky, Department of Psychology, 1835 Neil Ave., The Ohio State University, Columbus, OH 43210, USA.
Email: Sloutsky.1@osu.edu

Funding information
National Institutes of Health, Grant/Award Number: R01HD078545 and P01HD080679

Abstract

Organisms need to constantly balance the competing demands of gathering information and using previously acquired information to obtain rewarding outcomes (i.e., the "exploration–exploitation" dilemma). Exploration is critical to obtain information to discover how the world works, which should be particularly important for young children. While studies have shown that young children explore in response to surprising events, little is known about how they balance exploration and exploitation across multiple decisions or about how this process changes with development. In this study, we compare decision-making patterns of children and adults and evaluate the relative influences of reward seeking, random exploration, and systematic switching (which approximates uncertainty-directed exploration). In a second experiment, we directly test the effect of uncertainty on children's choices. Influential models of decision-making generally describe systematic exploration as a computationally refined capacity that relies on top-down cognitive control. We demonstrate that (a) systematic patterns dominate young children's behavior (facilitating exploration), despite protracted development of cognitive control; and (b) that uncertainty plays a major, but complicated, role in determining children's choices. We conclude that while young children's immature top-down control should hinder adult-like systematic exploration, other mechanisms may pick up the slack, facilitating broad information gathering in a systematic fashion to build a foundation of knowledge for use later in life.

KEYWORDS
attention, cognitive development, decision-making, exploration, information seeking, reward

1 | INTRODUCTION

Exploration is a critical activity in which organisms seek information that they can use to make effective and rewarding decisions. Although exploration is often advantageous, it is not cost free. The "exploration–exploitation" dilemma (see Hills, Todd, Lazer, Redish, & Couzin, 2015; Mehlhorn et al., 2015 for reviews) is a trade-off: gather information and forgo rewards or forgo acquiring information and use existing knowledge to get rewards. Effective decision-making requires appropriately balancing exploration and exploitation, with the ideal balance depending on the environment and the knowledge state of the individual. Higher levels of exploration are optimal when there is greater uncertainty (i.e., when one knows little about the environment, or when the environment is variable or dynamic), and therefore, exploration should be especially valuable for young children who often have little knowledge of how the world works.

Indeed, a number of studies support the idea that young children engage in exploration in response to surprising or ambiguous events (Bonawitz, van Schijndel, Friel, & Schulz, 2012; Cook, Goodman, & Schulz, 2011; Schulz & Bonawitz, 2007). Even infants show evidence of exploration, approaching or interacting longer with objects that violated expectations (Sim & Xu, 2017; Stahl & Feigenson, 2015).

Their importance notwithstanding, most previous studies examine exploration as preferences for things that are less known or surprising, while leaving some other important issues unexplored. In particular, little is known about how children mediate the competing demands of exploration and exploitation when faced with a series of decisions. Single choices are typically exploratory or exploitative, but the relative balance achieved between these two types of choices, and the mechanisms supporting that balance, can only be observed across many decisions. By examining children's patterns of choices across multiple decisions, and comparing them to adults, we can begin to examine questions such as the relative balance of exploration and exploitation across development, the types of exploration strategies available to young children, and the ways exploration–exploitation strategies change with development.

Theories suggest that children should explore more than adults, or at least more broadly (Gopnik et al., 2017), but this prediction has only very recently received empirical support (Schulz, Wu, Ruggeri, & Meder, 2019; Sumner, Steyvers, & Sarnecka, 2019). Furthermore, this expectation is not uncontroversial: Influential models of decision-making describe exploration as a computationally refined capacity, requiring top-down regulation (Badre, Doll, Long, & Frank, 2012; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Frank, Doll, & Oas-Terpstra, 2009). In addition, mounting evidence points to an important role of executive control process mediated by prefrontal cortex (PFC) in *systematic* exploration (Badre et al., 2012; Blanco et al., 2015; Frank et al., 2009; Otto, Knox, Markman, & Love, 2014). Ideally, exploration should be systematically directed toward parts of the environment that have greater uncertainty. As things change over time, uncertainty increases in areas that have less recently been checked. This type of directed exploration leads to sequential structure in choice patterns (Blanco et al., 2015; Knox, Otto, Stone, & Love, 2012). Given the protracted development of top-down cognitive control and PFC (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; Case y, Giedd, & Thomas, 2000; Sowell, Thompson, Holmes, Jernigan, & Toga, 1999), current theories predict that young children's exploration should be largely unsystematic, and there is evidence that directed exploration may not emerge until adolescence (Somerville et al., 2017). Similarly, development is largely conceptualized as a decrease in the stochasticity (or noisiness) of behavior over time. For example, in Gopnik et al. (2017), development is likened to decreasing "temperature" (moving from broader to narrower sampling) in a random sampling process.

In what follows, we present two experiments in which children and adults make a series of choices, and we evaluate the strategies driving their sequences of actions. Our results indicate that early in development exploration is predominant and largely systematic, and that uncertainty plays an important, but complicated, role. We then discuss how such systematic exploration could be implemented in the absence of mature top-down control.

Highlights

- Little is known about how children balance exploration and exploitation across multiple decisions, or how this process changes over development.
- Exploration should be especially useful for young children, who need information to understand how the world works.
- In two reported experiments, adults and 4-year-olds perform a decision-making task in which different options are worth different amounts of reward.
- While adults maximized reward, children exhibited heightened levels of exploration and were characterized by sequential patterns in their choices that approximate uncertainty-based exploration.

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants

Experiment 1 included thirty-two 4-year-olds (mean age = 54.8 months, 15 girls) and 34 adults (19 women). Four-year-old children were selected because at his age, children have known immaturities in selective attention and cognitive control (Bunge et al., 2002; Casey et al., 2000), while being able to understand the task and make independent decisions. Sample size was chosen to be consistent with previous studies comparing children of this age to adults in cognitive tasks (e.g., Deng & Sloutsky, 2015, 2016). Children were recruited from preschools and childcare centers on the basis of returned parental permission forms. Adults were undergraduate students participating for course credit.

2.1.2 | Stimuli and procedure

The task was a simplified *n*-armed bandit task framed as a game in which participants collected virtual candy from alien creatures (Figure 1). Four choice options were each marked by a different creature. Each creature gave a fixed number of candies that remained the same on every trial: 1, 2, 3, and 10 candies, respectively. These values were chosen so that one option was clearly superior to evaluate how often participants exploit it compared to exploring other options. The locations of the creatures and reward values were fixed across the experiment, but were randomly determined for each participant. Note that while reward values were stable, participants were not informed of this fact, and so they should still have subjective uncertainty about the outcomes of their choices and should need to explore to some extent. But, because adults learn and exploit rewarding options even in situations with noisy or changing

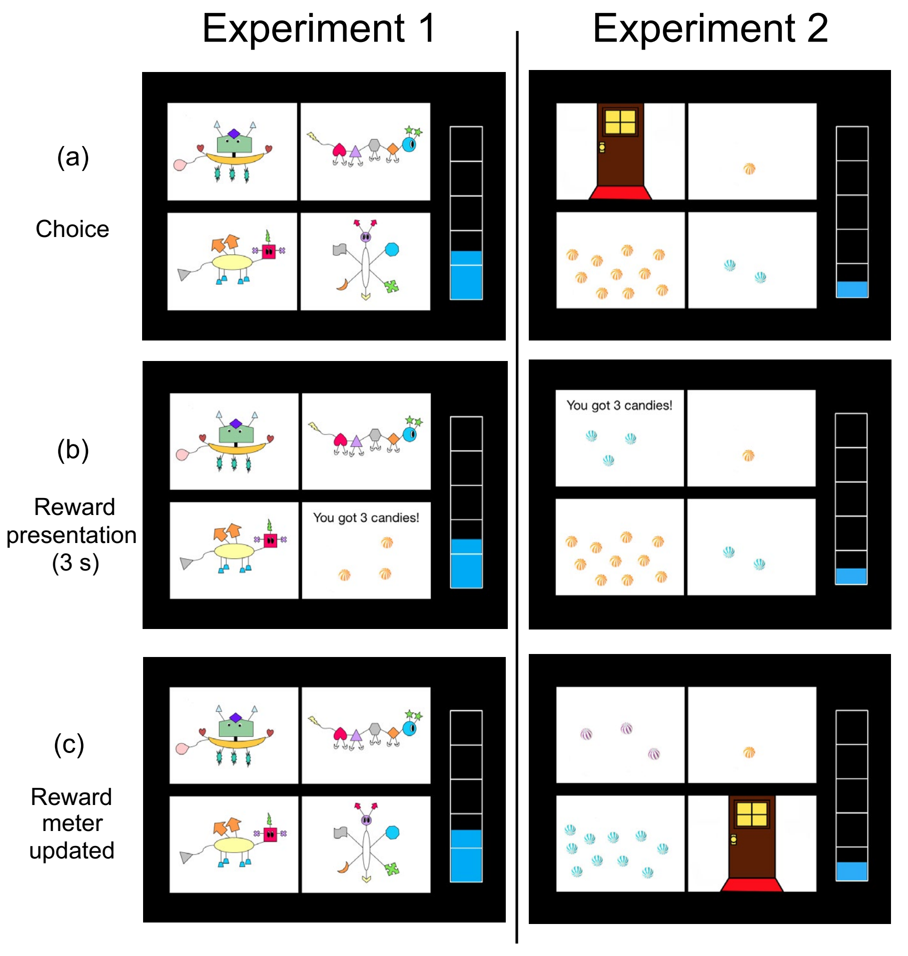

FIGURE 1 Trial dynamics. (a) Participants chose one of four creatures. (b) They received virtual candy for their choice. (c) That candy was added to the meter tracking total accumulated candy, and the next trial began. Benchmarks on the meter (every 180 candies) indicated the number of stickers earned. In Experiment 2, the reward for one option was hidden behind a door while the other three were visible prior to the choice

rewards (e.g., Blanco et al., 2015), they should primarily exploit the best option, with little exploration.

The goal of the game was to earn as much candy as possible. On each of 100 trials, children chose one of the four creatures using a touch screen (adults used a computer mouse). The reward amount was then displayed (for 3 s), and a meter tracking total accumulated reward was updated. Children earned stickers for each 180 candies, with benchmarks on the meter indicating these goals. Adults did not earn stickers.

Following the main experiment, a subset ($N = 22$) of child participants completed a follow-up memory test.[1] The experimenter pointed to each of the four locations in turn and asked the child how many candies they would get if they chose that option.

The primary planned behavioral analyses were to examine (a) the proportion of the highest value option choices over the course of the experiment as an indicator of exploitation; and (b) the proportion of trials that participants switched between different options as an indicator of exploratory tendencies. Computational modeling analyses were planned to evaluate the strategies that drive children's and adults' choices.

2.1.3 | Computational modeling

To investigate participants' exploration strategies, we developed a Reinforcement Learning model (Sutton & Barto, 1998) that included both systematic and random exploration. Critically, the choice probabilities were a combination of the expected values of the options and their choice lags. Choice lag was simply the number of trials since an option was last chosen. Choice lag serves as a proxy for uncertainty, because uncertainty increases as a function of time since an option was last checked. Although reward values are stable, because participants are unaware of this fact, their subjective uncertainty should still follow this pattern.

The model learns reward values by updating expected values for each option based on prediction error using the following equation:

$$V_{i,t+1} = V_{i,t} + \alpha(R_{i,t} - V_{i,t}), \tag{1}$$

where $V_{i,t}$ is the expected value of option i on trial t, $R_{i,t}$ is the reward on trial t earned for choosing option i, and α is the learning rate (a free parameter). Only the chosen option is updated on each trial. The model is initialized with all V_i set to 0. Choice probabilities were determined by the following function:

$$P(a_{i,t}) = \frac{e^{\beta*[V_{i,t}*(1-\phi)+L_{i,t}*\phi]}}{\sum_{j=1}^{n} e^{\beta*[V_{j,t}*(1-\phi)+L_{j,t}*\phi]}}, \tag{2}$$

where $P(a_{i,t})$ is the probability of choosing option i on trial t. $L_{i,t}$ is the lag term encoding the number of trials since option i was last chosen.

Two free parameters, ϕ and β, determine the levels of exploitation, systematic exploration, and random exploration. Specifically, ϕ ($0 \le \phi \le 1$) mediated the relative weights of expected value (i.e., exploitation) and lag (i.e., systematic exploration) on choices. Higher values of ϕ indicate greater influence of the lag, and hence more systematic exploration. When ϕ is 0, the model chooses based only on expected value; when ϕ is 1, it chooses only based on the lag.

β is the inverse temperature parameter that controls the extent that choices are deterministic or stochastic (Sutton & Barto, 1998). Lower values of β capture more "random" choices (i.e., random exploration) and greater values of β capture greater consistency of choices (regardless of whether they are driven by value or by lag). This model is similar to "exploration bonus" models (Daw et al., 2006; Kakade & Dayan, 2002), but with lag serving as a proxy for uncertainty. The model was fit to each participant by finding the set of parameters that maximized the likelihood of the participant's data given the model. Note that this model is used primarily as an analysis tool to disentangle factors contributing to participants' decision-making strategies and should not be taken as a direct model of the cognitive processes involved.

2.2 | Results

Participants' choices across the experiment are presented in Figure 2a. Adults quickly learned the best option and exploited it heavily (86.3% of trials), with only occasional exploration of other options. Children exploited the best option much less than adults (42.9% of trials), $t(64) = 10.05$, $p < .001$, 95% CI: $0.348 \le (\mu_1 - \mu_2) \le 0.520$, $d = 2.48$. Even within the first 20 trials (Figure 2b), when adults' exploration should be highest, they exploited the best option more than children (74.3% vs. 30.6%), $t(64) = 13.11$, $p < .001$, 95% CI: $0.370 \le (\mu_1 - \mu_2) \le 0.501$, $d = 3.23$.

Interestingly, there were striking individual differences between children (Figure 3). While a minority of children were adult-like (i.e., greatly preferring the highest value option), most chose all options approximately equally often. Using a lenient criterion of exploiting the best option on >50% of trials, only 11 of 32 of children made "adult-like" choices.

Memory test analyses confirmed that most children learned the reward values, despite not exploiting the most valuable option. Of the 22 children who completed the memory test, only two misidentified the best option. Eighteen correctly recalled the best option

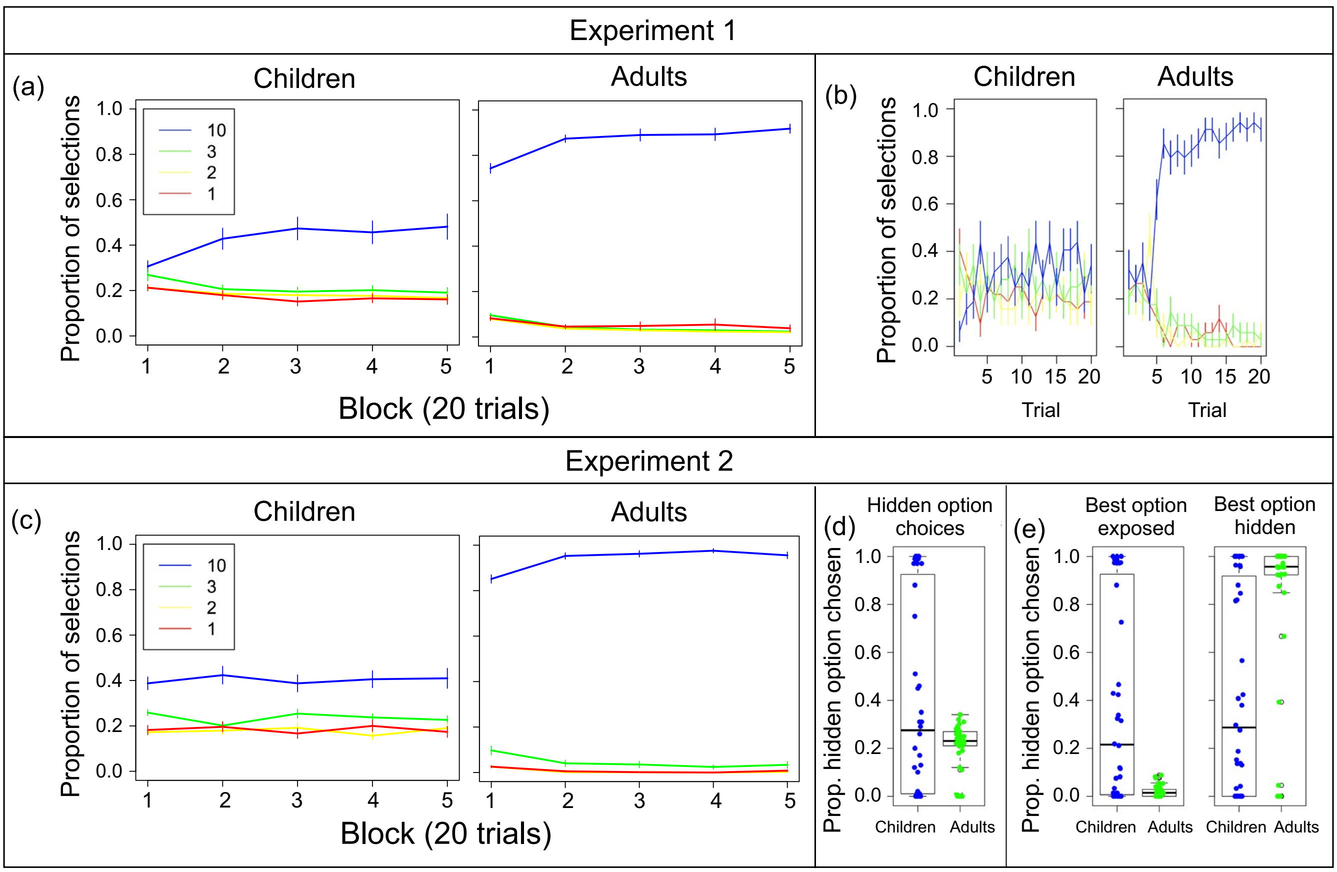

FIGURE 2 Choice proportions over time. (a, c) Adults exploited the most valuable option on a majority of trials, while children chose it substantially less often in both Experiments 1 and 2. (b) Analysis of the first 20 trials in Experiment 1 indicates that adults shifted to high levels of exploitation after a very short period of exploring all options. (d) Adults chose the hidden option in Experiment 2 at approximately chance levels, while children tended to either highly prefer or to avoid the hidden option. (e) Adults tended to only choose the hidden option when it was the highest value option, and largely avoided it otherwise, while most children's choices were unaffected by whether the hidden option was the highest value or not. Error bars reflect standard errors of the mean

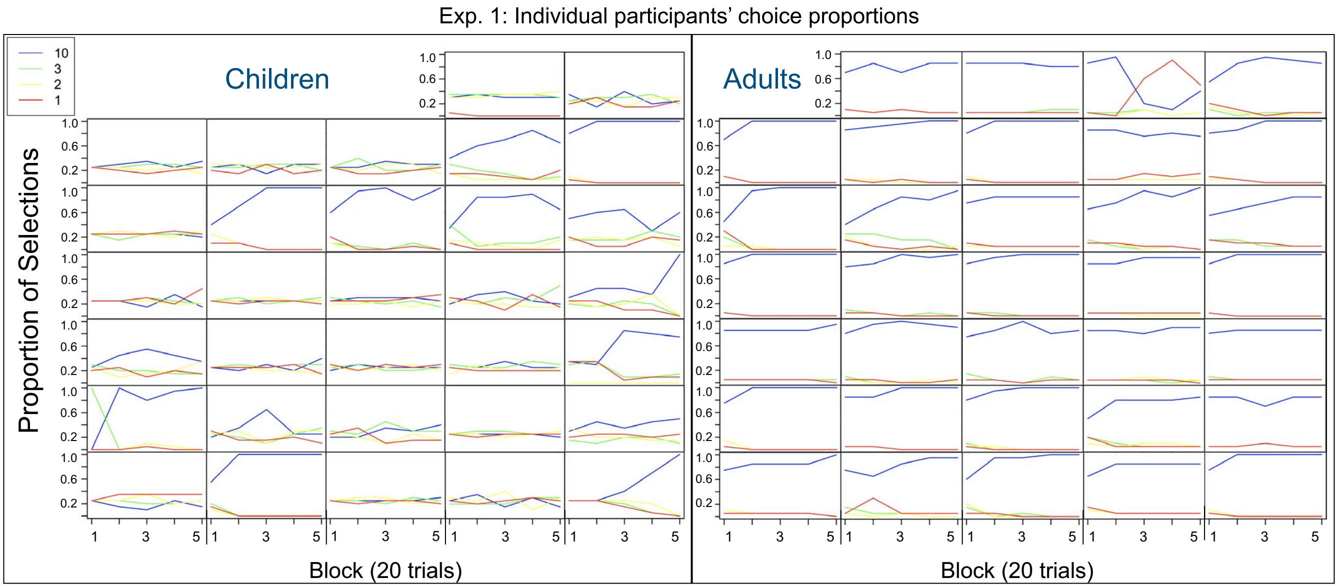

FIGURE 3 Individual participants' choices in Experiment 1. Children exhibited striking individual differences in their choices. While some children chose similar to adults, the majority of children chose all options with approximately equal proportions. Few children were between these two extremes. Almost all adults chose the highest value option on a large majority of trials

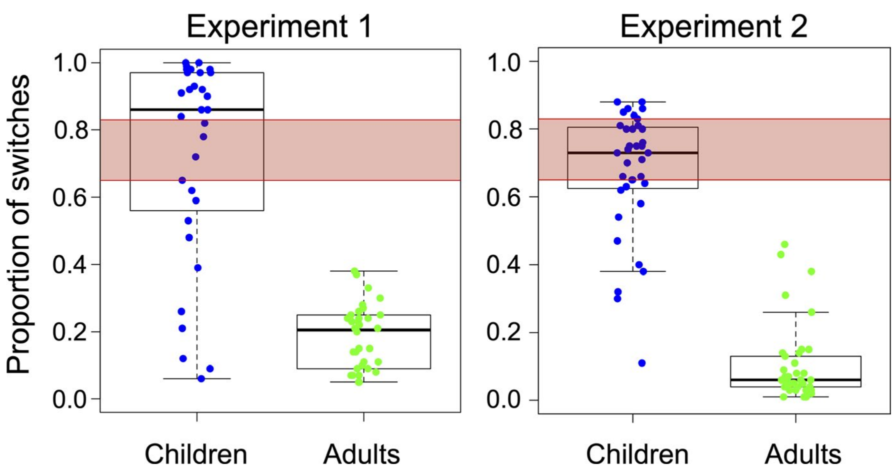

FIGURE 4 Response switching. The probability of response switching was much higher for children than adults. Dots represent individual participants, and the shaded region indicates the 95% probability density of switch proportions given random responding. The majority of children (17/32) were well above this region, indicating non-random responding and elevated exploration. In Experiment 2, children's switching was consistent with chance since their choices were primarily determined by which option was hidden on each trial. See also Figure 6 for a breakdown of children's switching behavior by best-fitting model for Experiment 1

and its value, and two recalled the best option but misremembered its exact value. Mean accuracy across all memory items was 77.2%. This finding also parallels a recent study in which, despite children accurately indicating the best option following a probabilistic learning task, they were less likely than adults to maximize their choices toward that option (Plate, Fulvio, Shutts, Green, & Pollak, 2018). Together, these findings suggest that children's choices are not motivated by achieving maximum reward to the extent that adults are.

Most importantly, children's choices were far from random. In contrast to adults (who were well below chance), many children switched their responses more frequently than would be expected by chance (Figure 4). These patterns of choices indicate heightened levels of exploration in children, seemingly even at the expense of forgoing exploiting a valuable option. Children's tendency to explore *systematically* is supported by modeling results presented in the next section.

2.2.1 | Modeling analyses

In order to validate inclusion of systematic exploration in the model, we compared it to a reduced model without systematic exploration (i.e., with ϕ fixed at 0) using the Akaike information criterion (AIC; Akaike, 1974). The full model better fit 23 of 32 children (total AIC_{full} = 6,391 vs. total $AIC_{\phi=0}$ = 7,466) and 21 of 34 adults (total AIC_{full} = 3,058 vs. total $AIC_{\phi=0}$ = 3,316), suggesting most participants engaged in systematic exploration to some extent.

We then analyzed the best-fitting parameter values to examine the relative influences of reward value, random exploration, and systematic exploration on participants' choices. Best-fitting values of all three free parameters were compared using Wilcoxon rank-sum tests because the parameters were not normally distributed (Figure 5a). Children and adults did not differ in the β parameter that controls the level of random exploration, $W = 480$, $p = .418$, or in the learning rate α, $W = 498$, $p = .562$.

Children did, however, have a higher value than adults for ϕ, $W = 751$, $p = .007$, indicating a greater influence of systematic exploration on children's choices compared to adults. Importantly, in the reduced model without systematic exploration, best-fitting β was lower for children than adults, $W = 264$, $p < .001$—highlighting the importance of accounting for systematic exploration. Without it we may have concluded that children's choices were simply noisier than adults. Furthermore, the lag-based choices in children are unlikely to be explained by a mere adherence to a spatial pattern (e.g., 1→2→3→4; see supplemental analyses at osf.io/6thfn/ for details.)

2.2.2 | Single-process models

Because large individual differences between children suggest the presence of distinct strategies, we fit simplified versions of the model that utilized only value ($\phi = 0$) or only choice lag ($\phi = 1$) to determine the primary factor determining each participant's choices.

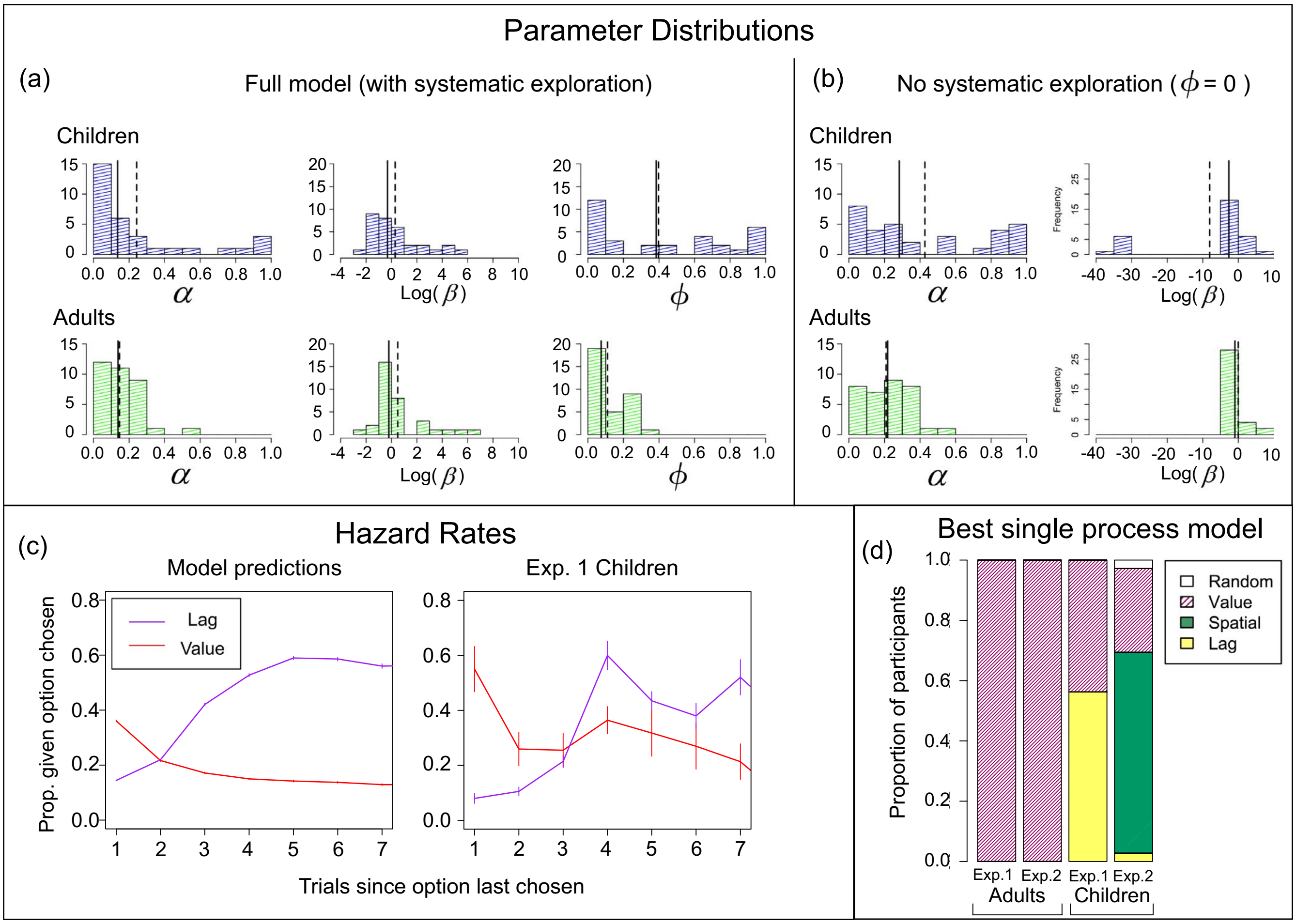

FIGURE 5 Modeling predictions and results. (a) The best-fitting parameter values of the Full model show that adults and children do not differ in the learning rate (α) or the level of random exploration (β), but that children show a greater relative influence of systematic exploration (ϕ) than adults. (b) Parameter estimates from the reduced model (with no systematic exploration) suggest that if systematic exploration is not accounted for children appear to be more random (indicated by lower β), highlighting the importance of including systematic exploration in understanding children's choices. Vertical bars represent the medians (solid) and means (dashed). (c) The hazard rates of choices, which are the probability of choosing an option given the amount of time since it was last chosen, demonstrate a critical difference in the two single-process models' predictions. The Lag model predicts an increasing probability of picking an option the longer it has been since it was last selected. Predictions were produced from simulations using best-fitting parameters from children best-fit by each model. (d) The proportions of participants best-fit by the Value, Lag, and Spatial uncertainty (Exp. 2) models are plotted. All adults were value-based. About half of children relied on choice lag when all options were uncertain (Exp. 1) and most relied directly on uncertainty when it was concentrated on a single option (Exp. 2), but note that this includes both uncertainty seekers and uncertainty avoiders

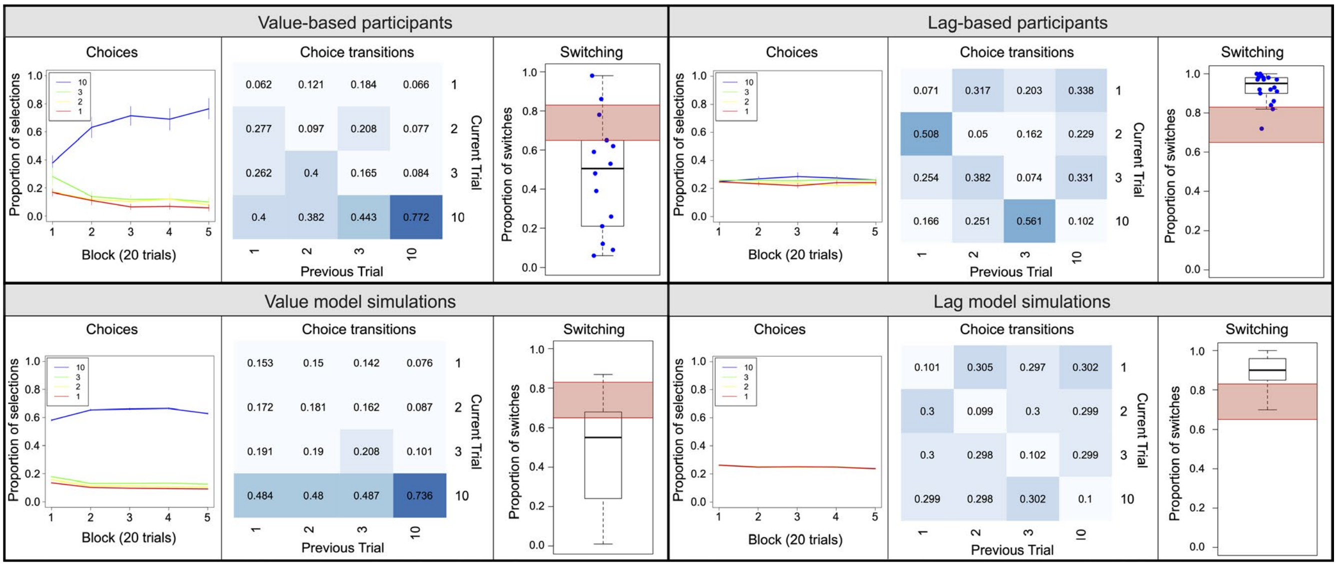

FIGURE 6 Children's strategies. Children in Experiment 1 seemed to exhibit two distinct strategies, well captured by the Value and Lag models, respectively. These strategies differed in a few key ways. Value-based participants ($N = 15$) exploited the best option often, explored little, and otherwise showed no particular pattern in their sequences of choices. Lag-based participants ($N = 17$) chose all options equally often, and switched responses at extremely high rates, almost never picking the same option twice in a row. Darker colors represent higher transition probabilities. The predictions for each model are based on simulations using the best-fitting parameters for children that were best-fit by that model

α (i.e., learning rate) was set to 1 for this analysis to better equate the two single-process models. A random model was also considered that chose all options with equal probability regardless of rewards or lags, but no participants were best fit by the random model in Experiment 1. As shown in Figure 5c, the key difference between the two models is that the Lag model predicts that options become more likely to be chosen the longer it has been since they were last selected, whereas the Value model does not.

As shown in Figure 5d, all adults were best fit by the Value model, whereas about half of children (17/32) were better fit by the model capturing systematic exploration—the Lag model. These proportions were significantly different X^2 (1; N = 66) = 26.30, $p < .001$, Cramer's $V = 0.631$. To examine whether these models accurately characterize children's behavior, we compared children best fit by each model to simulations from the models. The results presented in Figure 6 indicate that these single-process models capture children's choices quite well, further suggesting that children's responses are not random but instead follow one of two distinct systematic strategies.

Together these results indicate that, whereas adults' choices were driven by value, many children used a strategy that approximates uncertainty-based exploration. To better understand what role uncertainty might play in children's choice behavior, we conducted Experiment 2.

3 | EXPERIMENT 2

In Experiment 1, children exhibited systematic patterns of choices that suggest they may be engaging in a form of uncertainty-based exploration, with very little exploitation. However, it remained unclear whether their choices were actually driven by uncertainty, and if so, to what extent. In Experiment 2, we manipulated uncertainty directly by hiding only one option's reward value while exposing the other options. Since the hidden option's value is unknown, there is more uncertainty surrounding its outcome than the options with visible rewards. This effectively eliminates temporal (i.e., lag-based) uncertainty, while increasing differences in uncertainty between options. If children's choices are unaffected by uncertainty, the pattern observed in Experiment 1 (i.e., lag-based switching) should persist. In contrast, if their choices are driven by uncertainty, the pattern observed in Experiment 1 should disappear since there is no lag-based uncertainty, and choices should instead revolve around the hidden option.

3.1 | Method

Experiment 2 included 36 4- and 5-year-old children (mean age = 58.9 months, 22 girls) and 37 adults (20 women). Procedures were identical to Experiment 1, except that the rewards available for choosing three of the options were visible prior to the choice (Figure 1). The reward for the remaining option was hidden by a picture of a door and was revealed only if that option was chosen. The option that was hidden was randomly determined on each trial.

3.2 | Results

Like Experiment 1, adults chose the best option on almost every trial (93.9% of trials), significantly more often than children (40.3%

of trials), $t(71) = 15.37$, $p < .001$, 95% CI: $0.467 \le (\mu_1 - \mu_2) \le 0.606$, $d = 3.60$ (Figure 2c). On average, children chose the hidden option on 39.5% of trials, significantly above chance levels, $t(35) = 2.15$, $p = .039$, 95% CI: $0.258 \le \mu \le 0.532$, $d = 0.35$ (Figure 2d). Adults chose it significantly less often (21.5% of trials) than children, $t(71) = 2.63$, $p = .010$, 95% CI: $0.044 \le (\mu_1 - \mu_2) \le 0.315$, $d = 0.62$. When the hidden option was the highest value option, adults chose it on the majority of trials (84.3% of trials), but otherwise they almost never selected it (2.2% of trials; Figure 2e). Children chose the hidden option equally often whether it was best or not (42.8% vs. 37.9%, respectively)—less than adults when it was best, $t(71) = 4.81$, $p < .001$, 95% CI: $0.243 \le (\mu_1 - \mu_2) \le 0.587$, $d = 1.13$, but more than adults otherwise, $t(71) = 5.29$, $p < .001$, 95% CI: $0.222 \le (\mu_1 - \mu_2) \le 0.492$, $d = 1.24$.

Children exhibited large individual differences: Some overwhelmingly selected the hidden option, while others avoided it. Binomial tests on each child's number of hidden option selections showed that 15 of 36 children were significantly above chance levels, with nine children choosing the hidden option on almost every trial (≥97%). Another 15 children were below chance levels, with 12 choosing the hidden option on less than 3% of trials.

3.2.1 | Modeling analyses

The single-process models were fit to participants' data to determine the primary factor driving each participant's choices. In addition to the Lag, Value, and random models, a Spatial uncertainty model was fit. The latter model considers only whether each option's reward was exposed or hidden. Its choices were determined using the following equation:

$$P(a_{i,t}) = \frac{e^{\beta_u * U_{i,t}}}{\sum_{j=1}^{n} e^{\beta_u * U_{j,t}}}, \tag{3}$$

where $U_{i,t}$ was simply coded as 1 if the reward for option i was hidden and 0 if it was visible. Importantly, β_u could be positive or negative, and so it could capture both uncertainty seekers and avoiders. The learning rate parameter α was set to 1 for the Value model.

Similar to Experiment 1, all adults were best fit by the Value model (Figure 5d). Unlike in Experiment 1, the large majority of children were best fit by the Spatial uncertainty model ($N = 24$), while only one was best-fit by the Lag model. Ten children were best fit by the Value model, and one by the random model. This is in sharp contrast to Experiment 1, where more than half of children were best characterized by the Lag model.

These results suggest that when uncertainty was concentrated on a single option, it was the main factor influencing choices for the majority of children, replacing the systematic exploration pattern (i.e., lag-based responding) seen in Experiment 1. That the systematic pattern disappeared when there was no lag-based uncertainty suggests that children's choice patterns in Experiment 1 were driven at least partially by uncertainty. At the same time, many children were affected by uncertainty in an unexpected way—they avoided the hidden option altogether—something that may require further research.

4 | GENERAL DISCUSSION

Experiment 1 presents novel evidence of heightened exploration early in development, suggesting that children tend to explore more than adults even in situations when they may benefit from exploitation, and that their choices follow systematic patterns which approximate uncertainty-directed exploration. Even in a highly simplified task, where everything was stable with no variability in choice outcomes (which normally results in very little exploration), children explored at extreme levels. These surprising results demonstrate how much there is still to learn about young children's decision process.

Importantly, when uncertainty was concentrated spatially in Experiment 2 (with no temporal uncertainty), the sequential patterns that dominated children's strategies in Experiment 1 disappeared. This further suggests that children's choice patterns in Experiment 1 were motivated by uncertainty-based exploration. Interestingly, though, when uncertainty was concentrated on one option, some children were drawn strongly toward it, whereas others avoided it. This suggests a complex role of uncertainty in children's choices that warrants further investigation. One possibility is that preference for uncertainty follows an inverted U-shape favoring some intermediate level of uncertainty (and avoidance of higher uncertainty), but that the ideal level varies between children. Another possibility is that being visible increases the appeal of the rewards just enough to outweigh the lure of the uncertain option for some children. An additional (but not mutually exclusive) possibility is that children are not directly motivated to reduce uncertainty, but that various simpler processes (e.g., novelty preference) are at work, together producing broad, but systematic, sampling of the environment. Consistent with this idea, recent work shows that children continue searching after it is unnecessary for their current goals and additional information will not help complete their task (Ruggeri, Lombrozo, Griffiths, & Xu, 2016), suggesting that their exploration/search tendencies are driven by factors beyond task performance or reward seeking. Future work is needed to further clarify the roles and interactions of uncertainty and other mechanisms in motivating children's choices.

The reported results encourage a reassessment of the role of top-down control in systematic exploration. Some type of systematic exploration clearly dominates choices early in development when control processes are immature. While young children's immature top-down control should hinder adult-like systematic exploration, other mechanisms seem to pick up some of the slack, facilitating broad information gathering in a systematic fashion. How is this systematic exploration achieved?

We suggest that one particularly promising candidate may be children's immature attention allocation. Growing evidence suggests that (unless a highly salient stimulus captures young children's attention, e.g., Robinson & Sloutsky, 2010; Sloutsky & Napolitano, 2003), young children tend to distribute attention broadly, even when focused attention is warranted (Coch, Sanders, & Neville, 2005; Deng & Sloutsky, 2016; Plebanek & Sloutsky, 2017, 2018). Consistent with research suggesting a tight link between attention and decision-making (Gottlieb, 2012; Konovalov & Krajbich, 2016; Smith & Krajbich, 2018), and more specifically between attending selectively to the reward predictor and exploitation (Leong, Radulescu, Daniel, DeWoskin, & Niv, 2017), we posit that young children's tendency to distribute attention broadly is an important mechanism supporting exploration. For example, it may enable systematic exploration through graded novelty preference (the longer an option was not selected, the more it attracts attention).

Exploration seems to be a major driving force during early childhood—even outweighing the importance of immediate rewards. Starting with broad exploration and progressing to narrower search and increasing exploitation is an effective strategy across multiple domains (e.g., machine learning algorithms or resource foraging), and human development may be no exception (Gopnik et al., 2017). Put simply, elevated exploratory behavior is advantageous for learning (but not necessarily for performance) as it affords a broad sampling of the environment, which is useful when one knows little. This developmental progression has previously been conceptualized as analogous to simulated annealing (Gopnik et al., 2017), decreasing temperature over time in a random sampling algorithm. While this idea is generally consistent with our findings, the specific analogy may be somewhat misleading since children's choices do not conform to random sampling. Furthermore, the temperature parameter in our model did not reliably differ between adults and children.

A salient aspect of the reported results is that there were large individual differences between children. An important avenue of future research will be to examine why some children adopt "adult-like" reward-based strategies and others do not. Some possible factors worth investigating include early academic instruction and differences in maturation of the brain areas involved in decision-making.

One challenge in this research is achieving equal motivation between different age groups. That the prospect of winning stickers may make children more motivated by rewards in the task than adults is a potential limitation of the current studies. It is important to note, however, that this difference alone cannot account for the reported results. In fact, greater motivation should result in greater focus on rewards and higher levels of exploitation, which is the opposite of what we find.

In sum, this research demonstrates that systematic exploratory behavior can stem from other mechanisms than top-down cognitive control, and that these mechanisms may be specifically tuned to achieve broad information gathering in young children. If this is the case, some of the early immaturities of top-down control (e.g., distributed attention) may be advantageous for children: these immaturities may optimize learning (rather than performance) by subserving broad information gathering.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health grants R01HD078545 and P01HD080679 to V.M.S. Some of the data presented in this paper were previously presented in a poster at the Cognitive Development Society Biennial Conference, 2017. We thank Ian Krajbich, Jess Cooper, and Brad Love for helpful feedback.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data reported in this paper are archived and available on Open Science Framework at osf.io/tcjb8/.

ORCID

Nathaniel J. Blanco https://orcid.org/0000-0001-9899-8048
Vladimir M. Sloutsky https://orcid.org/0000-0003-4054-3360

ENDNOTE

[1] The memory test was added to the study after data collection had started, so the first 10 child participants did not receive the memory test.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. https://doi.org/10.1109/TAC.1974.1100705

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron*, *73*, 595–607. https://doi.org/10.1016/j.neuron.2011.12.025

Blanco, N. J., Love, B. C., Cooper, J. A., McGeary, J. E., Knopik, V. S., & Maddox, W. T. (2015). A frontal dopamine system for reflective exploratory behavior. *Neurobiology of Learning and Memory*, *123*, 84–91. https://doi.org/10.1016/j.nlm.2015.05.004

Bonawitz, E. B., van Schijndel, T. J., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, *64*, 215–234. https://doi.org/10.1016/j.cogpsych.2011.12.002

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301–311. https://doi.org/10.1016/S0896-6273(01)00583-9

Casey, B. J., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology*, *54*, 241–257. https://doi.org/10.1016/S0301-0511(00)00058-2

Coch, D., Sanders, L. D., & Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *Journal of Cognitive Neuroscience*, *17*, 605–622. https://doi.org/10.1162/0898929053467631

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition*, *120*, 341–349. https://doi.org/10.1016/j.cognition.2011.03.003

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876–879. https://doi.org/10.1038/nature04766

Deng, W. S., & Sloutsky, V. M. (2015). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology*, *51*, 392–405. https://doi.org/10.1037/a0038749

Deng, W., & Sloutsky, V. M. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62. https://doi.org/10.1016/j.cogpsych.2016.09.002

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience*, *12*, 1062–1068. https://doi.org/10.1038/nn.2342

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., … Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 7892–7899. https://doi.org/10.1073/pnas.1700811114

Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*, *76*, 281–295. https://doi.org/10.1016/j.neuron.2012.09.034

Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., & Cognitive Search Research Group. (2015). Exploration versus exploitation in space, mind, and society. *Trends in Cognitive Sciences*, *19*, 46–54. https://doi.org/10.1016/j.tics.2014.10.004

Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, *15*, 549–559. https://doi.org/10.1016/S0893-6080(02)00048-5

Knox, W. B., Otto, A. R., Stone, P., & Love, B. (2012). The nature of belief-directed exploratory choice in human decision-making. *Frontiers in Psychology*, *2*, 398. https://doi.org/10.3389/fpsyg.2011.00398

Konovalov, A., & Krajbich, I. (2016). Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. *Nature Communications*, *7*, 12438. https://doi.org/10.1038/ncomms12438

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron*, *93*, 451–463. https://doi.org/10.1016/j.neuron.2016.12.040

Mehlhorn, K., Newell, B. R., Todd, P. M., Lee, M. D., Morgan, K., Braithwaite, V. A., … Gonzalez, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, *2*, 191–215. https://doi.org/10.1037/dec0000033

Otto, A. R., Knox, W. B., Markman, A. B., & Love, B. C. (2014). Physiological and behavioral signatures of reflective exploratory choice. *Cognitive, Affective, & Behavioral Neuroscience*, *14*, 1167–1183. https://doi.org/10.3758/s13415-014-0260-4

Plate, R. C., Fulvio, J. M., Shutts, K., Green, C. S., & Pollak, S. D. (2018). Probability learning: Changes in behavior across time and development. *Child Development*, *89*, 205–218. https://doi.org/10.1111/cdev.12718

Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, *28*, 723–732. https://doi.org/10.1177/0956797617693005

Plebanek, D. J., & Sloutsky, V. M. (2018). Selective attention, filtering, and the development of working memory. *Developmental Science*, *164*, e12727–e12812. https://doi.org/10.1111/desc.12727

Robinson, C. W., & Sloutsky, V. M. (2010). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology*, *107*(3), 351–358. https://doi.org/10.1016/j.jecp.2010.04.006

Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology*, *52*(12), 2159–2173. https://doi.org/10.1037/dev0000240

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychological Science*, *30*(11), 1561–1572. https://doi.org/10.1177/0956797619863663

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, *43*, 1045–1050. https://doi.org/10.1037/0012-1649.43.4.1045

Sim, Z. L., & Xu, F. (2017). Infants preferentially approach and explore the unexpected. *British Journal of Developmental Psychology*, *35*, 596–608. https://doi.org/10.1111/bjdp.1219

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, *74*, 822–833. https://doi.org/10.1111/1467-8624.00570

Smith, S. M., & Krajbich, I. (2018). Attention and choice across domains. *Journal of Experimental Psychology: General*, *147*, 1810–1826. https://doi.org/10.1037/xge0000482

Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Abi Akar, N., Insel, C., & Wilson, R. C. (2017). Charting the expansion of strategic exploratory behavior during adolescence. *Journal of Experimental Psychology: General*, *146*, 155–164. https://doi.org/10.1037/xge0000250

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). *In vivo* evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, *2*, 859–861. https://doi.org/10.1038/13154

Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. *Science*, *348*, 91–94. https://doi.org/10.1126/science.aaa3799

Sumner, E., Steyvers, M., & Sarnecka, B. W. (2019). It's not the treasure, it's the hunt: Children are more explorative on an explore/exploit task than adults. In A.K. Goel C.M. Seifert & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society*. Montreal, QB: Cognitive Science Society.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT press.

How to cite this article: Blanco NJ, Sloutsky VM. Systematic exploration and uncertainty dominate young children's choices. *Dev Sci*. 2020;00:e13026. https://doi.org/10.1111/desc.13026