

# Structure inference in complex environments improves from childhood to adulthood

## Abstract

Early in development, children can infer latent structure in the world from sparse and ambiguous evidence. Through a process known as structure learning, they extract statistical regularities, construct causal models from those regularities, and use those models to arbitrate between exploiting known options and exploring novel alternatives. In turn, each decision and its outcomes refine the model that produced them. Despite the clear reciprocal relationship between structure learning and decision making in the real world, developmental research has largely examined these processes separately. To address this gap, we compared how children, adolescents, and adults behaved in a patch-foraging task designed to reveal how structure learning shapes exploratory decisions in a richly structured, dynamic environment. We found that younger participants left patches sooner than adults, enabling them to explore the environment more broadly within the fixed time window of the study. Computational modeling demonstrated that this difference in exploration arose from differences in participants' causal models of the environments. Younger participants grouped all patches into a single category despite large differences in richness, whereas older participants separated them into distinct categories. Despite differences in representation, participants of all ages used their uncertainty about the environment to guide their decisions. Together, our findings suggest that structure learning undergoes protracted development, but uncertainty-sensitive decision making emerges earlier and can support adaptive behavior even when representations remain imprecise.

**Keywords:** structure learning, development, foraging, exploration

## Research Highlights

- Children and adolescents formed less granular representations of environmental structure than adults.
- Despite these imprecise representations, younger participants showed adult-like sensitivity to uncertainty, planning further ahead when more confident in their internal models of the environment.

## Introduction

Children deftly extract structure from noisy, ambiguous input. They detect statistical regularities, infer associations, and integrate these inferences into causal models (Saffran, Aslin, & Newport, 1996; Fiser & Aslin, 2002; Jung, Walther, & Finn, 2021; Bullock, 1985; Leslie, 1982; Gopnik, Sobel, Schulz, & Glymour, 2001; Lapidow & Walker, 2020). Together, these processes comprise structure learning. Through structure learning, individuals come to understand the world and act effectively within it. The causal models acquired through structure learning guide decisions, and the outcomes of those decisions, in turn, refine those same models. Despite this tight, reciprocal relationship, developmental research has largely examined structure learning and decision making in isolation. Bridging this gap is critical because, with age, young learners increasingly direct their own learning. As they gain autonomy and spend more time away from caregivers, (Borawski, Ievers-Landis, Lovegreen, & Trapl, 2003; Alparone & Pacilli, 2012; Mikkelsen & Christensen, 2009; Steinberg, 2001; Morris, Ratliff, Cosgrove, & Steinberg, 2021) children and adolescents increasingly select for themselves which novel options, situations, and experiences to explore. Understanding developmental change in structure learning and decision making therefore requires studying how these processes interact.

A related question concerns how learners balance pursuing immediate rewards with obtaining information that can improve their causal models. This dilemma, known as the explore-exploit trade-off, captures the tension between exploiting an option already known to be good and exploring an unfamiliar alternative that may prove to be better. Because exploration generates information that can improve future choices, it may be particularly valuable for young learners, who have more time to benefit from what they discover (Gopnik, 2020). Consistent with this, children and adolescents explore more broadly and more stochastically than adults (Liquin & Gopnik, 2022; Blanco & Sloutsky, 2024; Christakou et al., 2013; Jepma, Schaaf, Visser, & Huizenga, 2020; Wan & Sloutsky, 2024; Giron et al., 2023; Dubois et al., 2022). Although stochastic, their exploration is not indiscriminate. By middle childhood, children preferentially sample options that will reduce their uncertainty (de Eccher, Mundry, & Mani, 2024; Blanco & Sloutsky, 2021; Ruggeri, Stanciu, Pelz, Gopnik, & Schulz, 2024; Jepma et al., 2020; Meder, Wu, Schulz, & Ruggeri, 2021), in some cases even more so than adults (Nussenbaum et al., 2023; Tymula et al., 2012; van den Bos & Hertwig, 2017). Strikingly, this uncertainty-directed sampling occurs despite the ongoing development of metacognition and uncertainty monitoring (Ghetti, Lyons, Lazzarin, & Cornoldi, 2008; Weil et al., 2013; Fandakova et al., 2017; Roderer

& Roebers, 2010; Krebs & Roebers, 2010; von der Linden & Roebers, 2006). These developmental differences in exploration shape the information available for learning and could lead children and adolescents to infer different latent structures of the environment than adults. However, very little work has directly examined how age-related differences in exploration relate to differences in structure learning (although see (Liquin & Gopnik, 2022) for a comparison of 4–7-year-olds and adults). This gap is particularly notable given past work in adults demonstrating that uncertainty in structure inference can alter explore/exploit decisions (Harhen & Bornstein, 2023).

One reason these literatures have progressed independently is that most decision-making tasks designed for younger participants place minimal demands on structure learning. For example, in multi-armed bandit tasks, learners repeatedly sample a small, fixed set of options to identify the best one. This design bears little resemblance to real-world exploration. In natural contexts, decision makers face an effectively unbounded set of alternatives, many of which are encountered only once. Patch-foraging tasks capture these features by asking participants to decide whether to continue exploiting a depleting resource or to leave in search of a new patch. Because patches cannot be revisited, effective foraging requires inferring the underlying distribution of patch quality and using those inferences to estimate the richness of novel patches. For these reasons, patch foraging offers a more naturalistic paradigm for studying exploration (Mobbs, Trimmer, Blumstein, & Dayan, 2018; Addicott, Pearson, Sweitzer, Barack, & Platt, 2017; Grima et al., 2025; Lloyd, Viding, McKay, & Furl, 2023). At the same time, most foraging tasks are structured very simply, with patches that are nearly identical in richness. With so little uncertainty, these tasks cannot reveal how learners cope with structural ambiguity or adjust their decisions in response. Consequently, it remains unclear how children, adolescents, and adults explore and infer structure in environments that are more variable and uncertain.

To address these gaps, we had participants aged 8 to 25 completed a patch-foraging task that interleaves structure learning with reward-based decisions in a heterogeneous, dynamic environment (Harhen & Bornstein, 2023). Unbeknownst to participants, the environment contained three patch types—poor, neutral, and rich—each associated with a distinct distribution governing the rate of resource depletion. Participants were instructed to collect as many rewards as possible within a fixed time window. To infer participants’ latent representations of the environment, we fit a Bayesian latent cause model (Harhen & Bornstein, 2023) to their stay/leave decisions. Our model captures individual differences along two key dimensions: (1) the complexity of the inferred environmental structure and (2) how much learners use uncertainty in that structure to adjust their decisions. These dimensions

allow us to ask two developmental questions: how do children and adolescents differ from adults in the structure they infer, and in their use of structural uncertainty to guide decisions? The latent cause model generates distinct predictions for explore-exploit (i.e., stay or leave) decisions across patch types, depending on how a forager represents the environment and responds to structural uncertainty. Foragers who infer a more complex structure should remain longer in rich patches than those relying on a simpler representation. By grouping all patches together, foragers who infer a simpler structure overestimate how quickly rich patches will deplete and thus leave too early. Foragers who adapt their decisions to structural uncertainty should remain longer after entering a patch that differs in richness from the previous one. Because such transitions are rare in our task, they should induce uncertainty. Uncertainty-adaptive foragers respond by discounting future rewards more heavily, resulting in them staying longer in the current patch.

We considered competing hypotheses about how structure learning changes across development. One possibility is that children construct simpler representations of the environment than adults. In passive learning contexts, the ability to detect and explicitly report statistical structure improves into young adulthood (Arciuli & Simpson, 2011; Shufaniya & Arnon, 2018; Raviv & Arnon, 2018). Another possibility is that children infer representations similar to those of adults. Their broader attention, heightened exploration, and less constraining prior knowledge (Lucas, Bridgers, Griffiths, & Gopnik, 2014; Gopnik et al., 2017; Deng & Sloutsky, 2016, 2015; Plebanek & Sloutsky, 2017; Smith & Kemler, 1977) could compensate for still-developing statistical learning abilities. We similarly considered competing hypotheses about uncertainty-guided decision making. On one hand, the protracted development of uncertainty monitoring and metacognition suggests that younger learners may less readily adjust their decisions in response to uncertainty. On the other hand, several studies have shown that even quite young children can successfully use uncertainty to guide their exploration (Nussenbaum et al., 2023; de Eccher et al., 2024; Blanco & Sloutsky, 2021; Ruggeri et al., 2024). Using our task and modeling approach, we adjudicate between these possibilities, clarifying how learners of different ages infer the structure of novel environments and how those inferred structures shape subsequent decisions under uncertainty.

## Methods

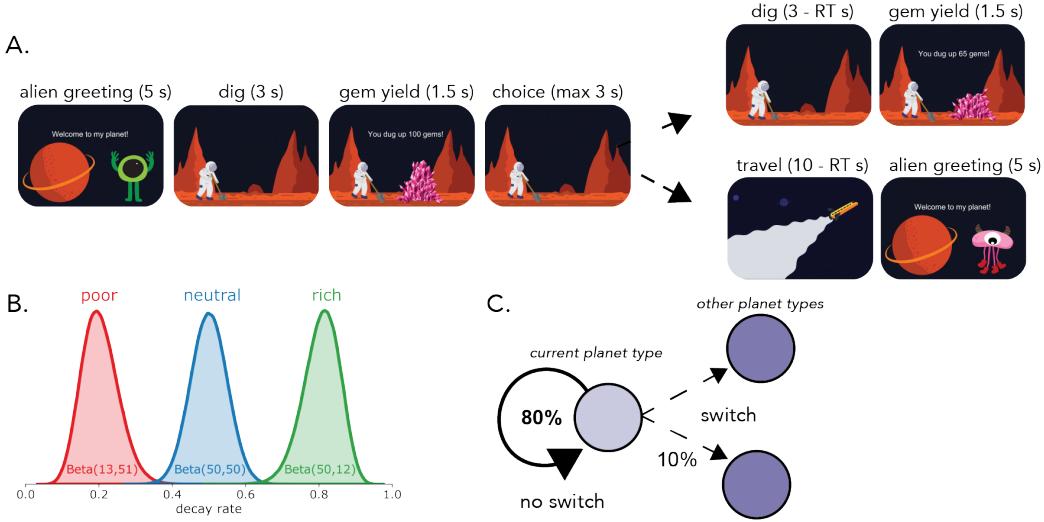
### Participants

Our final sample consisted of 252 participants, aged 8 to 25 ( $M = 17.11$  years,  $SD = 5.29$ , 128 females, 124 males). This sample included 70 children (8.08 - 12.94 years;  $M = 10.49$ , 36 females), 68 adolescents (13.07 - 17.94 years;  $M = 15.47$ , 35 females), and 114 adults (18 - 25.83 years;  $M = 22.14$ , 57 females). This sample size exceeds those of many prior developmental studies investigating value-guided learning and decision making (Cohen, Nussenbaum, Dorfman, Gershman, & Hartley, 2020; Nussenbaum, Scheuplein, Phaneuf, Evans, & Hartley, 2020; Nussenbaum et al., 2023). We selected this sample size to achieve at least 80% power (with  $\alpha = 0.05$ ) to detect a small effect size. The effect size estimate ( $\rho = 0.16$ ) was taken from the original adult study (Harhen & Bornstein, 2023). Participants were recruited through the lab's database, for which we solicited sign-ups via Facebook and Instagram ads, local science fairs and events, and fliers on New York University's campus. Lab researchers verified each participant's age and identity prior to their participation in the online study.

All participants reported normal or corrected-to-normal vision and no history of psychiatric or learning disorders. Based on self- or parent-report, 42% of participants were White, 31% were Asian, 13.1% were mixed race, 12.7% were Black, less than 1% were Pacific Islander or Native Hawaiian, and less than 1% were Native American. Additionally, 11.8% of the sample identified as Hispanic. Participants' annual household incomes ranged from less than \$20,000 to over \$500,000. The distribution was as follows: 9.3% earned less than \$20,000; 7.7% earned \$20,000–39,999; 15.9% earned \$40,000–59,999; 11.4% earned \$60,000–79,999; 11.8% earned \$80,000–99,999; 18.3% earned \$100,000–199,999; 4.9% earned \$200,000–299,999; and fewer than 1% each reported incomes of \$300,000–399,999, \$400,000–499,999, or above \$500,000.

An additional 45 participants completed the study but were excluded based on the following criteria: difficulty understanding task instructions (failing the instruction comprehension check more than twice,  $n=4$ ), unusually quick responses (mean reaction time < 200 ms,  $n=12$ ), and extreme strategies ( $n = 14$  for mean planet residence time  $\pm 2 SD$  from group mean;  $n=4$  for fully depleting gem mines on more than 75% of visited planets;  $n=11$  for leaving more than 75% of visited planets immediately after the initial dig). Participants received a \$10 Amazon gift card for completing the study and had the opportunity to earn an additional performance-based bonus of up to \$2.

## Task



**A. Task design.** Participants traveled to various planets to dig for space gems. On each trial, they decided between continuing to dig on the current planet or traveling to a new one. Both options had their costs: digging depleted the mine, progressively reducing its gem yield, while traveling took a substantial amount of time. **B. Environment structure.** Planets belonged to one of three types—poor, neutral, or rich—differing in how quickly they depleted as mined. Each type was characterized by a distinct distribution over decay rates. **C. Environment dynamics.** Planet richness was correlated in time. A new planet had an 80% probability of being the same type as the previous planet (“no switch”) and a 20% probability of transitioning (“switch”) to a different type.

Participants completed a child-friendly variant of a patch foraging task previously used to examine structure learning in adults (Harhen & Bornstein, 2023). Modifications were made to accommodate younger participants. These modifications included shortening the length of the task (from five blocks to four), enhancing the clarity of the instructions, augmenting the instruction comprehension quiz, and increasing the maximum decision time (from two seconds to three). Participants were instructed to complete the task in a quiet, comfortable space of their choosing and without assistance from another person. For child and adolescent participants, parents were explicitly instructed not to assist their child during the task.

In the task, participants acted as miners collecting “space gems” across various planets (Fig 1). Bonus payments were tied directly to the total amount of gems collected. Upon landing on a planet, participants dug once and received a gem yield sampled from a Normal distribution ( $\mu = 100$ ,  $\sigma = 5$ ). On subsequent trials, participants decided whether to stay and continue digging on the current planet or to leave and travel to a new planet. Each

choice was associated with a cost. The gem yield diminished with each dig at a rate that depended on the planet type (described below), and traveling to a new planet took substantially more time than digging on the current planet. Participants had three seconds to decide between staying and leaving. If they chose to stay, a short animation of their avatar digging was displayed for the remainder of the three-second decision window (3 seconds minus response time), followed by a gem yield (1.5 sec). If they chose to leave, a longer animation of a flying rocket ship was displayed for the remainder of the ten-second travel window (10 seconds minus response time), followed by an alien welcoming them to the new planet (5.5 seconds). Animation durations were linked to the participant's response time on the previous trial to ensure that their decision speed did not influence the average reward rate of the overall environment. Because the average reward rate is a central decision variable in patch-foraging tasks, this design feature ensures that individual differences in stay/leave decisions cannot be attributed to differences in response time. If participants did not respond within three seconds, a red "X" and a prompt encouraging faster responses appeared, after which they could re-attempt the decision. In total, the task consisted of four blocks, each lasting six minutes.

Before beginning the main task, participants received instructions that were simultaneously presented in written form on the screen and read aloud via an audio track. Participants were told that they would explore a series of planets to search for "space gems," and that their goal was to collect as many gems as possible to maximize their bonus payment. They were instructed that digging on a planet yielded diminishing returns and that once a planet began to deplete, they could choose to leave for a new planet. Participants were told that travel between planets incurred a delay and that they could never return to a planet after leaving it. Upon arriving at a new planet, participants would encounter a novel alien whose appearance provided no information about the amount of gems available. Participants were also informed that the task would last 20 minutes regardless of their choices and that they would periodically return to a "home base" for breaks (for complete instructions, see S1.1). To reinforce the instructions, participants completed a brief practice game that allowed them to experience digging and traveling to new planets. To prevent early learning about the distribution of gems, the number of gems obtained during the practice game was not revealed. Instead, participants only saw a barrel that was described as containing gems. Participants then completed an instruction comprehension quiz (for quiz questions, see S1.1). Those who answered all questions correctly on the first attempt progressed to the main task. Those who missed any questions were given a second attempt with corrective feedback. Participants who failed the quiz more than once were excluded from all analyses.

The task environment featured three planet types. Rich planets depleted the most slowly (gem yield decreasing at a rate sampled from a Beta distribution with parameters  $\alpha=50$ ,  $\beta=12$ ,  $M=0.8$ ,  $SD=0.05$ ), poor the most rapidly ( $\alpha=13$ ,  $\beta=51$ ,  $M=0.2$ ,  $SD=0.05$ ), and neutral planets at an intermediate rate ( $\alpha=50$ ,  $\beta=50$ ,  $M=0.5$ ,  $SD=0.05$ ). On each trial, the depletion rate was newly sampled from the planet's respective distribution. For example, if a participant dug four times on a rich planet, they might observe a reward sequence such as 100, 83, 65, and 53 gems, but on a poor planet the sequence they might instead observe 100, 18, 4, and 1 gems. To mimic the structure of natural environments, which often exhibit high autocorrelation in resources (Boyce, 2006; Legendre & Fortin, 1989; Legendre, 1993), we designed the task environment so that planet richness was correlated across time, with switches occurring infrequently enough to support learning. With these considerations in mind, we set the transition probabilities such that the next planet was the same type as the previous one with 80% probability and a different type with 20%. Importantly, these differences between planets were not communicated to participants, requiring them to infer this information from the observed sequence of rewards alone. In addition, we designed the task to measure participants' representational biases independent of task demands. To achieve this, we structured the sequence of planets such that using a more simple representation, one that grouped all planets together, achieved comparable overall rewards to using a more complex representation, one distinguishing between different planet types. Simulations confirmed that the two strategies yielded similar outcomes on average (see S1.4).

## Analysis approach

**Marginal Value Theorem** To assess the extent to which individuals over- or under-harvested, we compared their planet (patch) residence times to the predictions of the Marginal Value Theorem (Charnov, 1976), which prescribes optimal behavior in patch foraging tasks. Here, we define patch residence time as the number digs a forager performs within a patch. We focus on residence time rather than rewards obtained on a single planet because residence time is sensitive to both the within-planet reward rate and the global reward rate of the task environment. According to MVT, an optimal agent decides whether to stay or leave by comparing the immediate expected returns from staying ( $V_{stay}$ ), to the opportunity cost of staying on the current planet ( $V_{leave}$ ).

$V_{stay}$  is the reward expected from the next dig, the previous reward multiplied by the predicted depletion rate. An MVT-optimal forager accurately identifies the planet's type and uses the true mean of its depletion distribution for their prediction.

$$V_{stay} = r_t * \hat{d} \quad (1)$$

$$\hat{d} = \begin{cases} 0.2 & \text{if planet is poor} \\ 0.5 & \text{if planet is neutral} \\ 0.8 & \text{if planet is rich} \end{cases}$$

Where  $r_t$  is the reward received on the last dig, and  $\hat{d}$  is the predicted depletion.

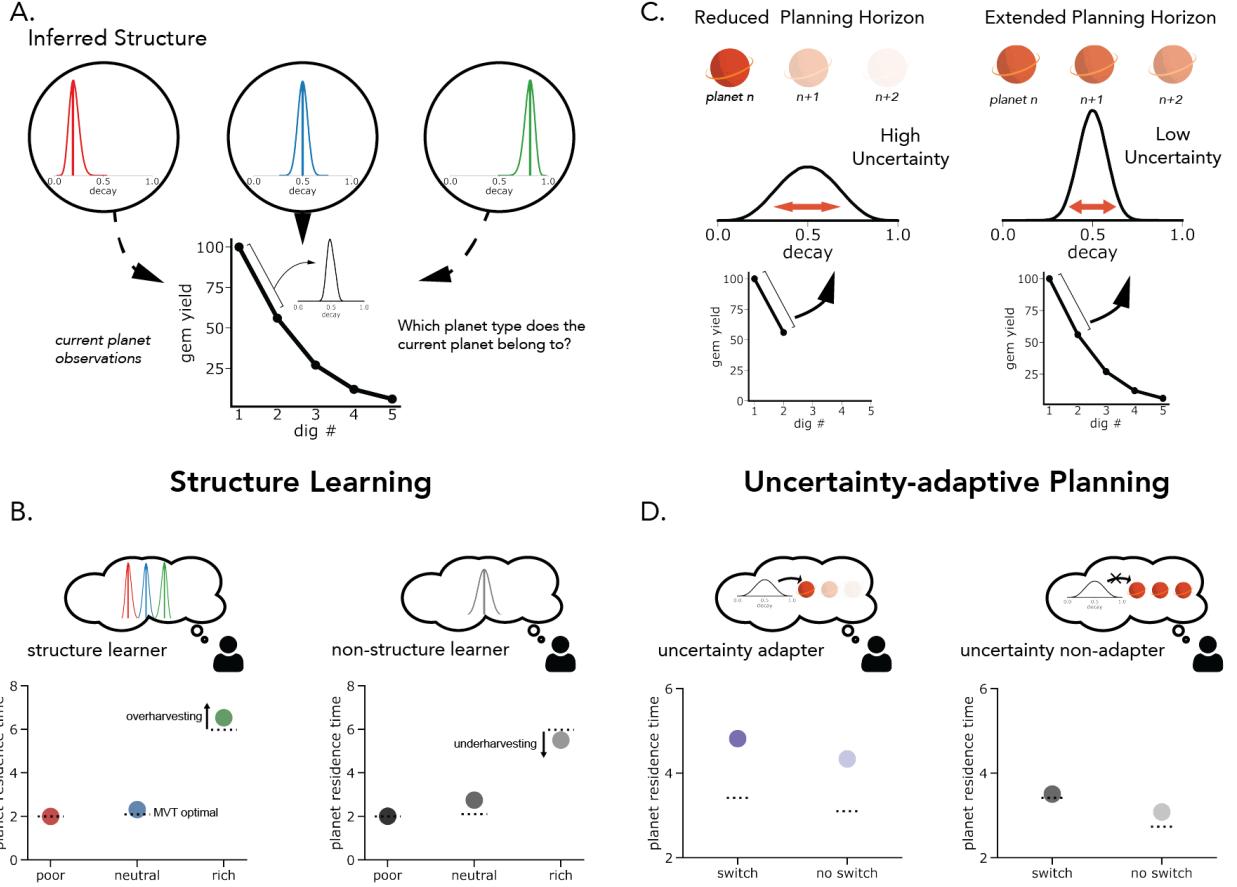
They estimate  $V_{leave}$ , the expected reward from digging on an alternative planet using the global reward rate, the total rewards received ( $r_{total}$ ) divided by the total time spent foraging ( $t_{total}$ ). Note that under this definition  $V_{leave}$  will change with each decision to harvest or leave a planet. By multiplying the global reward rate by the time required to dig ( $t_{dig}$ ),  $V_{leave}$  reflects the opportunity cost of digging on the current planet over an unknown alternative planet.

$$V_{leave} = \frac{r_{total}}{t_{total}} * t_{dig} \quad (2)$$

The forager compares these values and chooses greedily, always selecting the higher-valued action.

**Structure learning and uncertainty-adaptive planning model** Our model relaxes MVT’s assumption of perfect knowledge and instead assumes that foragers rationally learn the structure of their environment and adapt their planning in response to uncertainty about that structure (Fig. 2). In this way, our modeling approach is akin to that of bounded rationality (Simon, 1990), explicitly accounting for the limited information available to decision makers.

Foragers must navigate the environment without knowing the true number of planet types, the classification of individual planets, or the decay rate distribution associated with each type (Fig 2A). To model how foragers infer this latent structure, we use the Chinese Restaurant Process (Aldous, 1985). Although developed in statistics, the Chinese Restaurant Process (CRP) has been used to model numerous psychological processes including category



**Figure 2: A. Structure learning computation.** The forager makes two simultaneous inferences: (1) identifying the current planet’s type based on observed rewards and (2) determining the total number of planet types in the environment. We model these two inferences as a Chinese Restaurant Process. **B. Structure learning predictions.** The forager’s inference of planet types is governed by the parameter,  $\alpha$ . The model predicts distinct patterns of over- and under-harvesting depending on the forager’s representation of the environment and the number of planet types they consider. The non-structure learner’s behavior is simulated with  $\alpha = 0$  and the structure learner’s behavior is simulated with  $\alpha = 0.2$ , the value that produced the most veridical representation in simulation. Throughout, we refer to a model with  $\alpha = 0.2$  as  $\alpha^*$ . The markers show model-simulated planet residence times (PRTs), while dotted lines indicate the Marginal Value Theorem (MVT)-optimal PRTs for reference. **C. Uncertainty adaptive planning computation.** Foragers adjust their planning horizon based on their uncertainty about the current planet’s type. Less uncertainty encourages planning further into the future, while more uncertainty discourages planning. We modulate the planning horizon through adjusting the extent future rewards are discounted. **D. Uncertainty adaptive planning predictions.** Foragers whose planning incorporates their uncertainty should overharvest more when the planet type switches, because these switches are rare and uncued and so should, on balance, increase local uncertainty. In contrast, foragers who do not incorporate uncertainty into their planning (or who do not detect the change) should overharvest to a similar extent regardless of whether the planet type has changed. The uncertainty adapter’s behavior is simulated with  $\gamma_{coef} = -0.3$  and the non-uncertainty adapter’s behavior is simulated with  $\gamma_{coef} = 0$ .

learning (Griffiths, Navarro, & Sanborn, 2006; Perfors & Tenenbaum, 2009; Kemp, Perfors, & Tenenbaum, 2007), state space inference (Gershman, Blei, & Niv, 2010), and the organization of episodes in memory (Shin & DuBrow, 2021). The CRP prior is governed by two principles: (1) the probability of assigning a planet to an existing type increases with the number of planets already assigned to that type (2) there remains some probability, proportional to the concentration parameter,  $\alpha$ , of assigning a new planet to a new type. This model allows the forager's representational complexity to grow as they accumulate experience

Critically, our model predicts different stay-leave behavior for different values of  $\alpha$  (Fig 2B). These differences are most pronounced on rich planets. On rich planets, individuals with  $\alpha > 0$  overharvest, while those with  $\alpha = 0$  underharvest. These deviations from MVT-optimality arise from uncertainty about the current planet's type and misestimation of its decay rate. Unpacking this further, an agent with  $\alpha = 0$  lumps all previously observed depletions into a single type, which inflates their estimate of how quickly rich planets deplete, leading them to leave too early.

$$P(k) = \begin{cases} \frac{n_k}{N+\alpha} & \text{if } k \text{ is old} \\ \frac{\alpha}{N+\alpha} & \text{if } k \text{ is new} \end{cases}$$

Where  $n_k$  is the number of planets assigned to type  $k$ ,  $\alpha$  is the concentration parameter, and  $N$  is the total number of planets encountered.

After observing a depletion on a planet, the forager computes the posterior probability of the planet being a type as:

$$P(k|D) = \frac{P(D|k)P(k)}{\sum_{j=1}^J P(D|j)P(j)} \quad (3)$$

Where  $J$  is the number of clusters created up until the current planet,  $D$  is a vector of all the depletions observed on the current planet, and all probabilities are conditioned on prior type assignments of planets,  $p_{1:N}$ .

Because computing the exact posterior probability of planet type assignments is computationally intractable, we approximate it using a particle filter. Each particle maintains a hypothetical set of assignments and is weighted

based on how well it explains the observed data. When a forager leaves a planet, the particles are resampled. A new pool of particles is generated by sampling with replacement from the previous pool, with the likelihood of a particle being selected proportional to its weight. This process favors particles that better explain the data, increasing their probability of persisting in subsequent iterations. We use a moderate number of particles, 200, as this amount allows for psychological plausibility while still approximating the posterior well (Griffiths et al., 2006).

To predict the next dig's yield, we use a three-step Monte Carlo sampling procedure. First, a particle is selected based on its weight. Second, a planet type is sampled from the selected particle's posterior. Third, a decay rate is drawn from the selected planet type's decay rate distribution. This process is repeated 1,000 times, and the decay rates are averaged over to produce the final prediction.

Each decay rate distribution is initialized to a Gaussian with  $\mu=0.5$  and  $\sigma=0.5$ . While the true decay rates follow a Beta distribution, the model assumes normally distributed observations to allow for analytic updates using a Normal-Gamma prior.

Unlike MVT-optimal foragers, who have perfect knowledge, foragers acting according to our model make decisions under epistemic uncertainty. They can never be certain that they have an accurate model of the environment. Theoretical work in reinforcement learning suggests that under uncertainty, reducing how far out into the future the individual plans, the planning horizon, can improve performance (Jiang, Kulesza, Singh, & Lewis, 2015). We implement this concept in our structure learning model with an adaptive factor for discounting future rewards,  $\gamma_{\text{effective}}$ . Note that while the planning horizon technically defines how far into the future the consequences of a decision are evaluated, discounting future consequences produces the same behavior in our task. We define  $\gamma_{\text{effective}}$  to be a function of an individual's baseline discounting rate ( $\gamma_{\text{base}}$ ), their uncertainty over the current planet's type ( $U$ ), and the degree to which uncertainty influences their planning ( $\gamma_{\text{coef}}$ , Fig 2C). We compute uncertainty ( $U$ ) as the entropy of the Multinomial distribution over planet types, using the same three-step Monte Carlo sampling procedure described above.

$$\gamma_{\text{effective}} = \frac{1}{1 + e^{-(\gamma_{\text{base}} + \gamma_{\text{coef}} * U)}} \quad (4)$$

Thus, lower values of  $\gamma_{\text{coef}}$  correspond to heavier discounting of future rewards in response to uncertainty.  $\gamma_{\text{coef}}$  provides an additional measure of the extent to which a forager uses their structural knowledge—or lack thereof—to

adjust their decision making. We include this parameter because the learning and use of structural knowledge has been shown to dissociate developmentally (Nussenbaum et al., 2020; Potter, Bryce, & Hartley, 2017; Decker, Otto, Daw, & Hartley, 2016; Cohen et al., 2020; Schlichting, Guarino, Schapiro, Turk-Browne, & Preston, 2017). The model predicts that individuals with lower  $\gamma_{coef}$  should not only increase their overharvesting overall but also increase it particularly following a rare switch in planet type, as these are points of heightened uncertainty, (Fig 2D).

To model action selection, we use a softmax function, incorporating a lapse rate ( $\epsilon$ ) to account for occasional inattention:

$$p_{stay} = (1 - \epsilon) \frac{1}{1 + e^{-\beta(v_{stay} - v_{leave})}} + \frac{\epsilon}{2} \quad (5)$$

With  $\beta$  being the inverse temperature and  $\epsilon$  being the lapse rate. Including a lapse rate helps to separate systematic behavior from noise, thus reducing the misestimation of our main parameters of interest.

**Computational model-agnostic analyses** Before fitting computational models to participants' choices, we first used linear mixed-effects regression models to analyze participants' behavior. We compared their behavior to the qualitative patterns predicted by variants of our structure learning and uncertainty-adaptive planning model (Fig 2). Our primary dependent variable was the difference between a participant's actual patch residence time and the MVT-prescribed residence time (i.e., over- or underharvesting). We first regressed the patch residence time relative to MVT on age to test for overall age-related differences. We then fit a model including planet type (poor, neutral, rich), planet number (a proxy for experience), age, and their interactions. This analysis allowed us to evaluate two key predictions of the structure-learning models: (1) if participants acquire the environment's structure over the course of the task, their behavior should become more MVT-aligned, reflected in an effect of planet number and (2) the  $\alpha = 0$  and  $\alpha > 0$  models diverge on rich planets, with  $\alpha > 0$  predicting underharvesting and  $\alpha > 0$  predicting overharvesting (Fig 2B). Thus, the sign of the planet-type effect on overharvesting indicates which model better captures participants' qualitative behavior, and interactions with age test for developmental differences in representations of the environment (i.e., more complex with many inferred planet types or simpler with a single planet type).

To test the model's uncertainty-adaptive planning predictions, we fit a separate regression examining how planet type switches, planet number, and age interacted to predict deviations from MVT. Learners who rationally adapt

their decisions to uncertainty should increase their overharvesting following rare switches, yielding a positive effect of switch type (Fig 2D), with age interactions capturing developmental variation in this adjustment. Finally, we regressed response times on switches in planet type, planet number, and age as an additional index of structure sensitivity. Slower responses following switches would indicate that participants detected changes in planet richness. An age interaction would suggest developmental differences in structure sensitivity.

In all our model-agnostic analyses, we used the “lme4” package for R (Bates, Kliegl, Vasishth, & Baayen, 2018) to fit the mixed-effects models to choice and reaction time data. Except where noted, models included participant-level random intercepts and random slopes across within-participant fixed effects. To minimize Type I error, we initially specified the maximal model (Barr, Levy, Scheepers, & Tily, 2013). If the model failed to converge, we iteratively simplified the model by first removing interactions between random slopes, followed by random slopes themselves, until convergence was achieved. We used the ‘bobyqa’ optimizer and set the number of model iterations to 10,000. Continuous variables—age, planet number, and reaction time—were z-scored prior to their inclusion. Age was z-scored across participants while planet number and reaction times were z-scored within. Reaction times were log-transformed before z-scoring.

**Computational model-fitting** We compared two versions of the model. In one version,  $\alpha$  was set to 0. When  $\alpha = 0$ , the forager assumes that all planets belong to a single type. In the other version,  $\alpha$  was set to 0.2, the value that, in simulation, produced the most veridical representation across a range of the model’s other parameters (see S1.4). Throughout, we refer to this latter model as  $\alpha^*$ .

We fit  $\alpha$  to behavior using discrete models rather than as a free parameter because parameter recovery was moderately poor when  $\alpha$  we allowed  $\alpha$  to be a free continuous parameter (S1.3). However, model recovery was excellent for the discrete models (S1.2). The difficulty in recovering continuous values of  $\alpha$  likely stems from both the short task duration (which was shortened by a block from the version run in young adults, Harhen & Bornstein, 2023) and from the fact that smooth changes in  $\alpha$  do not produce smooth changes in the inferred representation. Because representations consist of discrete clusters, different values of  $\alpha$  can be aliased when they generate identical latent structures. The remaining parameters— $\gamma_{base}$ ,  $\gamma_{coef}$ ,  $\beta$ , and  $\epsilon$ —were fit as free parameters because their recovery ranged from moderately good to good (S1.2).

We fit participants’ data on a choice-by-choice basis. Free parameters and their bounds are detailed in the

supplementary materials (Table S1). To identify the parameter values that minimized the negative log likelihood of participants' choices, we used Bayesian Adaptive Direct Search (BADS, Acerbi & Ji, 2017), an optimization algorithm suited for stochastic and computationally-expensive functions. To increase the likelihood of finding the global minimum, we initialized the optimization with different starting points generated from a Sobol sequence. Sobol sequences are quasi-random and have been shown to be more effective than grid or random search (Bergstra & Bengio, 2012), while offering greater computational efficiency than Latin hypercube sampling (Renardy, Joslyn, Millar, & Kirschner, 2021). Starting points were generated until the convergence criteria were met, defined as five consecutive iterations without improvement to the overall minimum. The final parameter values were those that yielded the lowest negative log likelihood across all starting points. Parameter recoverability analyses for both models are included in the supplementary materials (S1.2).

The models' likelihoods are stochastic due to the approximation of the posterior distribution over planet type assignments. To address this noise, we repeated the cluster assignment process 1,000 times. We computed the log likelihood of participants' choices for each of these repetitions and marginalized over them. We then negated this value to obtain the input to the optimization algorithm.

**Model comparison** We assessed model fit using Akaike Information Criteria (AIC), which penalizes for model complexity. For age-group-level comparisons, we used protected exceedance probabilities (PXP). PXPs estimate the likelihood that a given model is the most frequent best-fitting model within a group while accounting for chance differences in model frequencies. Model recoverability analyses are included in the supplementary materials (S1.2).

**Computational model validation** Finally, we sought to determine whether our models ( $\alpha = 0$  or  $\alpha^*$ ) captured key qualitative features of behavior and provided a close quantitative fit to the data. To do so, we simulated data using the best-fitting model and parameters for each participant and compared this simulated behavior to participants' actual behavior.

**Data and code availability statement** All data and code are available at <https://github.com/noraharhen/Harhen-Budiono-Hartley-Bornstein-2025-Foraging>.

# Results

## Computational model-agnostic

**Overharvesting increases with age** We first asked whether overharvesting varied with age. We quantified the extent to which participants' planet residence times deviated from those prescribed by the Marginal Value Theorem (MVT) and found that participants, on average, overharvested, staying longer than the MVT-optimal residence time ( $b_0 = 0.81$ ,  $SE = 0.11$ ,  $t(248.49) = 7.47$ ,  $p < .001$ ). This tendency diminished with age. Younger participants left planets earlier, thus aligning them more closely with MVT ( $b_{age} = 0.22$ ,  $SE = 0.11$ ,  $t(248.33) = 2.01$ ,  $p = .045$ ).

**Use of structure knowledge strengthens with age** We next tested two key predictions of our structure learning model (Fig. 2). The first concerns how overharvesting should vary across planet types. The model predicts that all foragers will overharvest on neutral planets and, to a lesser extent, on poor planets. However, behavior on rich planets should depend on how the forager represents the environment (Fig. 2B). Foragers who distinguish between planet types should overharvest the least on poor planets, the most on neutral planets, and to a moderate extent on rich planets. Those who do *not* distinguish between types should instead *underharvest* on rich planets, leaving too early because they overestimate how quickly these planets deplete. The second prediction concerns how overharvesting should change over time. With increasing task experience, foragers should form more accurate representations of the environment and consequently behave more closely in line with MVT. To test these two predictions, we examined how overharvesting varied with planet type, task experience (indexed by planet number), age, and their interactions.

Results from the linear regression model supported both predictions. Participants overharvested the least on poor planets, the most on neutral planets, and to a moderate extent on rich planets (Fig. 3;  $b_0 = 1.30$ ,  $SE = 0.081$ ,  $t(245.63) = 15.96$ ,  $p < .001$ ;  $b_{poor} = -0.63$ ,  $SE = 0.052$ ,  $t(307.32) = -12.05$ ,  $p < .001$ ;  $b_{rich} = -0.42$ ,  $SE = 0.13$ ,  $t(245.17) = -3.15$ ,  $p = .0018$ ). This follows the pattern predicted by our structure learning model that distinguishes between planet types (Fig. 2B;  $\alpha > 0$ ). Participants' overharvesting also decreased with experience, particularly on rich planets ( $b_{planet\ number} = -0.24$ ,  $SE = 0.05$ ,  $t(238.88) = -4.85$ ,  $p < .001$ ;  $b_{poorXplanet\ number} = 0.067$ ,  $SE = 0.047$ ,  $t(346.64) = 1.43$ ,  $p = .15$ ;  $b_{richXplanet\ number} = -0.26$ ,  $SE = 0.060$ ,  $t(229.28) = -4.29$ ,  $p < .001$ ). This is consistent with participants using increasingly accurate representations of the environment as they accumulate

task experience.

These effects varied with age. Older participants overharvested more than younger participants, specifically on rich planets ( $b_{ageXrich} = 0.36$ ,  $SE = 0.13$ ,  $t(245.01) = 2.68$ ,  $p = .0078$ ), with no significant age-related effects on poor or neutral planets ( $b_{age} = 0.059$ ,  $SE = 0.082$ ,  $t(245.48) = 0.72$ ,  $p = .47$ ;  $b_{ageXpoor} = -0.045$ ,  $SE = 0.052$ ,  $t = (303.047) = -0.86$ ,  $p = .39$ ). Under the structure learning model, behavior diverges most prominently between “structure learners” ( $\alpha > 0$ ) and “non-structure learners” ( $\alpha = 0$ ) on rich planets. Older participants’ behavior more closely matched the behavior of “structure learners” than did younger participants, consistent with age-related improvements in the use of complex structural knowledge.

**Implicit structure knowledge is present across all ages** To measure structure learning more implicitly, we analyzed reaction times following switches in planet type. We focused on participants’ second decisions on each planet, as the first depletion provides the initial cue to a planet’s type. If participants inferred that planets of the same type cluster together in time, a switch should surprise them and slow their responding. To test this prediction, we used a mixed-effects model to examine how response times varied with planet type switches, planet number, age, and their interactions. The model included participant-level random intercepts and slopes for the planet switch regressor. Consistent with our prediction, participants responded more slowly following a switch in planet type ( $b_{switch\ point} = 0.049$ ,  $SE = 0.023$ ,  $t(255.5) = 2.09$ ,  $p = .037$ ), providing additional evidence that they were sensitive to the environment’s structure and dynamics. Response times decreased across the task overall ( $b_{planet\ number} = -0.049$ ,  $SE = 0.012$ ,  $t(8551) = -4.072$ ,  $p < .001$ ), but switch-related slowing did not diminish with experience ( $b_{switch\ pointXplanet\ number} = 0.014$ ,  $SE = 0.024$ ,  $t(8711) = 0.60$ ,  $p = .55$ ). Importantly, age did not moderate the effect of switches on response times ( $b_{ageXswitch\ point} = 0.0088$ ,  $SE = 0.023$ ,  $t(256.3) = 0.38$ ,  $p = .71$ ). Younger participants slowed after switches just as adults did, suggesting an implicit awareness of planet differences across our full age range. These results point to a dissociation between detecting structure and using that structure to guide decisions. While participants of all ages were sensitive to the environment’s dynamics, only older participants leveraged this knowledge to inform their choices.

**Uncertainty adaptive planning emerges early in development** We next tested whether participants adjusted their planning horizons in response to uncertainty about a planet’s type, the key feature of our structure learning model’s uncertainty-adaptive planning computation (Fig. 2C). Foragers using this strategy should overharvest

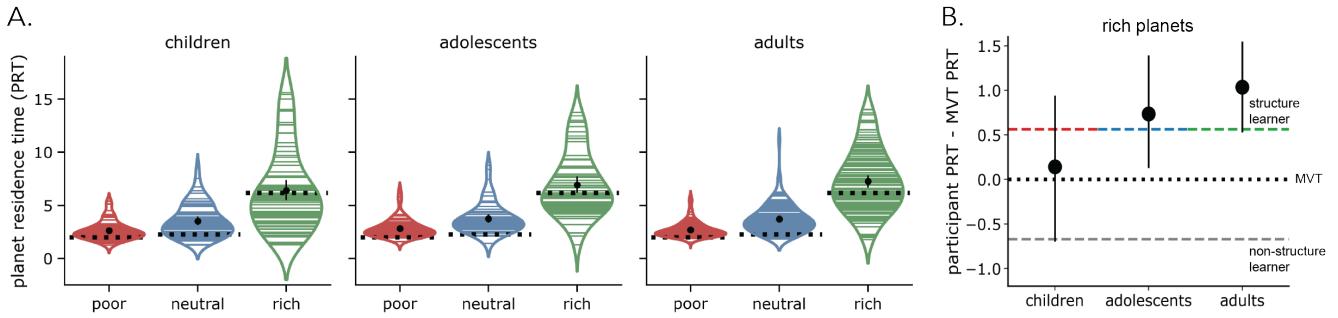
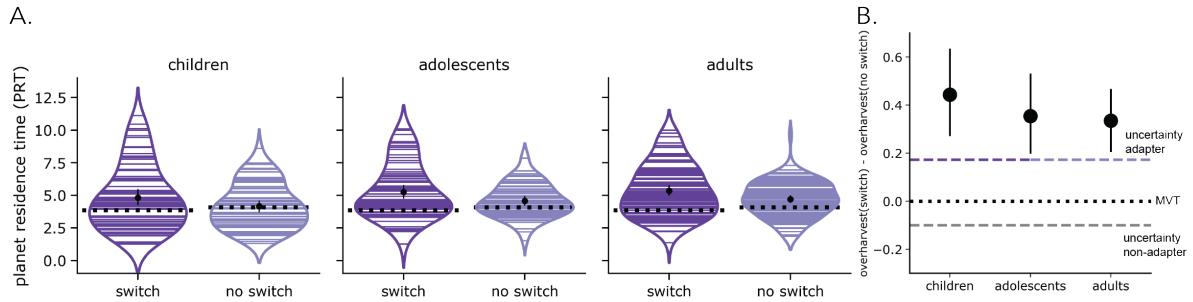


Figure 3: **Model-agnostic signature of structure learning.** **A.** Planet residence times (PRTs) relative to Marginal Value Theorem (MVT)-optimal PRTs across planet types and age groups. Violin plots show the distribution of PRTs for each age group. Thin lines show individuals' mean PRT for the planet type, markers indicate group means, and error bars denote 95% confidence intervals. Dotted lines indicate the MVT-optimal PRTs for each planet type. Across all planet types, participants overharvest relative to MVT. **B.** Age-related differences were the most pronounced on rich planets. Importantly, our two models of interest predict that structure learners ( $\alpha^*$ ) and non-structure learners ( $\alpha = 0$ ) should diverge most strongly in their behavior on these planets. For this reason, we examined differences between age groups and models more closely by plotting the deviation from MVT-optimal PRT, whereas children's PRTs were close to the MVT-optimal value. Children's behavior aligned more closely with the predictions of the non-structure learner ( $\alpha = 0$ ; gray dotted line), whereas adolescents' and adults' behavior more closely matched the predictions of the structure learner model ( $\alpha^*$ ; colored dotted line). This pattern of results suggests age-related differences in the inferred structure of the environment.

more following a switch in planet type, when uncertainty is highest (Fig. 2D). To test this, we examined how overharvesting varied with switches, task experience, age, and their interactions. The mixed-effects model included participant-level random intercepts and slopes for planet number. As predicted, participants overharvested more following switches in planet type (Fig 4;  $b_{switch\ point} = 0.31$ ,  $SE = 0.039$ ,  $t(8410) = 7.844$ ,  $p < .001$ ). This effect marginally decreased with experience ( $b_{switch\ point \times planet\ number} = -0.078$ ,  $SE = 0.042$ ,  $t(8414) = -1.85$ ,  $p = .065$ ), suggesting that switches became less uncertainty-inducing as participants gained familiarity with the environment. Early in the task, switch-related overharvesting did not significantly vary with age ( $b_{age \times switch\ point} = -0.0094$ ,  $SE = 0.039$ ,  $t(8410) = -0.24$ ,  $p = .81$ ). However, older participants showed a greater reduction in switch-related overharvesting with experience ( $b_{age \times switch\ point \times planet\ number} = -0.11$ ,  $SE = 0.042$ ,  $t(8412) = -2.65$ ,  $p = .0082$ ). These findings indicate that uncertainty-adaptive planning emerges early in development, but that older participants more effectively incorporate their experience of the environment into their decisions over time.



**Figure 4: Model-agnostic signature of uncertainty-adaptive planning.** **A.** Planet residence times (PRT) relative to Marginal Value Theorem (MVT)-optimal PRTs across planet types and age groups. Violin plots show the distribution of PRTs within each age group. Thin lines indicate individuals' mean PRT for the planet type, markers indicate group means, and error bars denote 95% confidence intervals. Dotted lines show the MVT-optimal PRTs for reference. **B.** To assess how participants adjusted their behavior when uncertainty increased, we computed the difference in overharvesting (participant PRT minus MVT-optimal PRT) for trials in which the planet "switched" type relative to the previous planet versus trials with no switch. Because these switches are relatively rare, they are moments of heightened uncertainty. An individual who adapts to uncertainty (uncertainty adapter model;  $\gamma_{coef} < 0$ ) should shorten their planning horizon in these cases and consequently are predicted to increase the extent of their overharvesting. In contrast, an individual who does not adapt is predicted to show less adjustment in their overharvesting in response to these switches (uncertainty non-adapter;  $\gamma_{coef} = 0$ ).

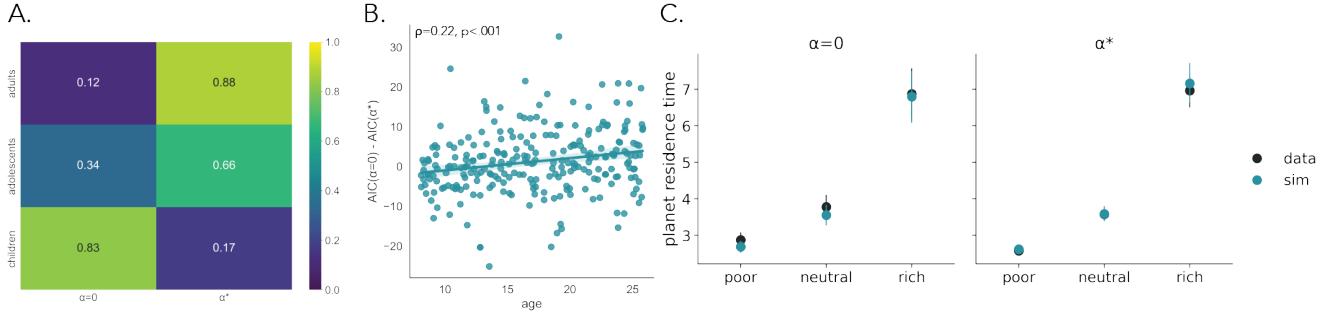
## Computational model-based results

Our central question was whether structure learning differs with age. To address this, we fit two models to participants' choices. In one model,  $\alpha$  was fixed at 0, corresponding to an undifferentiated representation of the environment. In the other,  $\alpha$  was fixed at 0.2—the value that most consistently produced a veridical representation in simulations (S1.4), which we refer to as the  $\alpha^*$  model. Within each age group, individuals varied in which model best captured their choices (Fig. S5). However, our primary aim was to determine whether the model that most frequently best fit participants' choices differed across age groups. To determine this, we calculated the protected exceedance probability (PXP) for each model within each age group (Fig. 5A). A PXP quantifies the probability that a given model is more frequent in a group than the alternatives, while accounting for the possibility that observed differences in model evidence arise by chance. We found that adults' behavior was best captured by the  $\alpha^*$  model (PXP = 0.88), suggesting that they inferred and used the environment's latent structure to guide their decisions. In contrast, children's behavior was more often best fit by the  $\alpha = 0$  model (PXP = 0.83), indicating that they did not differentiate between planet types. Adolescents fell between these groups, with no model clearly dominating ( $\alpha^*$  PXP = 0.66,  $\alpha=0$  PXP = 0.34). There was considerable variability within each age group, in part because two of the groups included participants spanning years of rapid developmental change. We therefore examined developmental differences in model fit continuously. To do so, we computed the difference in AIC scores between the two models for each participant. Positive values indicated being better fit by the  $\alpha^*$  model. Although substantial variability remained across participants of similar ages, on average, the AIC difference increased with age (Fig. 5B; Spearman's  $\rho = .22, p < .001$ ), recapitulating the pattern observed in our model-agnostic analyses. Together, these findings indicate that the ability to infer and use structural knowledge strengthens across development.

Under the  $\alpha^*$  model, no free parameters significantly varied with age ( $\gamma_{base}: \rho = -0.049, p = .44$ ;  $\gamma_{coef}: \rho = -0.068, p = .28$ ;  $\beta: \rho = 0.056, p = .37$ ;  $\epsilon: \rho = -0.092, p = .15$ ). Parameters from the  $\alpha=0$  model yielded the same results ( $\gamma_{base}: \rho = -0.041, p = .52$ ;  $\beta: \rho = 0.0095, p = .88$ ;  $\epsilon: \rho = -0.087, p = .17$ ).

## Discussion

Here, we investigated how the ability to infer and exploit latent structure changes across development. In a heterogeneous, dynamic patch-foraging task, younger participants left patches earlier than adults, particularly on



**Figure 5: Model-based results.** **A.** Protected exceedance probabilities (PXP) reveal differences in the most frequent best-fitting model across age groups. Adults' choices were better captured by the  $\alpha^*$  model, while children's choices were better captured by the  $\alpha=0$  model. For adolescents, however, neither model completely dominated, though the  $\alpha^*$  had a somewhat higher PXP than the  $\alpha=0$  model. **B.** To examine continuous age-related differences in the model fit, we computed the Akaike Information Criterion (AIC) difference between the two models for each participant. Positive values indicate a better fit with the  $\alpha^*$  model. The AIC difference increased with age (Spearman's  $\rho = 0.22, p < .001$ ), indicating that the  $\alpha^*$  model provided a progressively better account of behavior with increasing age. **C.** Comparison of participants' data to model-generated data. Each subpanel includes only the participants best fit by that model ( $\alpha = 0$  or  $\alpha^*$ ). For each subpanel, we compare the participants' mean planet residence times across the three planet types to the model residence times generated using their best-fitting parameters. Error bars denote 95% confidence intervals.

rich planets. Computational modeling indicates that these behavioral differences stem from age-related differences in how participants represented the environment. Younger participants' behavior more closely aligned with a model assuming all patches belonged to a single type, whereas adults' choices were better captured by a model inferring multiple types, supporting fine-grained distinctions between planets. This finding extends prior work showing that, in passive learning contexts, children's and adolescents' structure learning abilities lag behind adults' (Arciuli & Simpson, 2011; Shufaniya & Arnon, 2018; Raviv & Arnon, 2018). Despite these differences in structure learning, however, participants across the age range adapted their planning to uncertainty, remaining in patches longer following rare, unexpected switches in patch type. This result is consistent with past studies demonstrating that even young children use uncertainty to guide their exploration (Nussenbaum et al., 2023; de Eccher et al., 2024; Blanco & Sloutsky, 2021; Ruggeri et al., 2024). Together, our findings suggest that while the ability to infer latent structure continues to develop into adulthood, learners even in middle childhood can flexibly adjust their decision making in response to the environment's dynamics.

Children's choices suggest that, as a group, they adopted a simpler representation of the environment, treating

all patches as belonging to a single type. However, children's response times add nuance to this finding. Children responded more slowly following a rare switch in patch type, indicating surprise at the change. Such an effect implies sensitivity to differences between patch types at some level. This dissociation between response times and choices suggests that children acquired structural knowledge that they could not yet use to guide their decisions. One possibility is that children's knowledge remained only implicit, and thus inaccessible for strategic control and decision-making. Similar dissociations in implicit and explicit knowledge have been observed in younger children in numerous domains, including numerical cognition and language (Karmiloff-Smith, 1995).

Another, not mutually exclusive, possibility is that learners maintain multiple representations that vary in granularity and flexibly draw on them depending on the decision at hand (Koechlin, Ody, & Kouneiher, 2003; Simon, 1975; Stachenfeld, Botvinick, & Gershman, 2017; Ho et al., 2022). From this perspective, participants of all ages may have acquired both simpler and more refined representations of the environment but differed in which representation they used to guide their choices. Although our task was not explicitly designed to test for this dissociation, prior studies have observed strikingly similar results, finding that children and adolescents can acquire structural knowledge but fail to apply it in their planning and reasoning (Nussenbaum et al., 2020; Potter et al., 2017; Decker et al., 2016; Cohen et al., 2020; Schlichting et al., 2017). Resource rationality offers a useful lens for understanding this dissociation (Persaud, Bass, Colantonio, Macias, & Bonawitz, 2020). Under this framework, cognition is shaped by the pressure to make efficient use of limited computational resources (Lieder & Griffiths, 2019). Because drawing on more complex representations places greater demands on working memory, future, simulation, and episodic memory—abilities that mature into young adulthood (Hartley, Nussenbaum, & Cohen, 2021; Coughlin, Robins, & Ghetti, 2019; Keresztes, Ngo, Lindenberger, Werkle-Bergner, & Newcombe, 2018; Bunge & Wright, 2007)—children may favor simpler representations that are better suited to their developing cognitive abilities. As these capacities improve with age, learners may increasingly rely on more detailed representations to guide their decisions.

Yet another possibility is that children rely on simpler representations not because of cognitive constraints, but because doing so is a rational choice in the context of our task. We designed the task such that both simple and complex representations yield equivalent rewards on average. Under these conditions, a simpler representation is more efficient, achieving the same outcomes at a lower computational cost. A more puzzling question, then, is why adults adopt the costlier representation. This result stands in contrast to a large and influential literature

demonstrating that adults often rely on heuristics, strategies that ignore information to make decisions more quickly or frugally, across a wide range of decision contexts (Tversky & Kahneman, 1974; Gigerenzer & Gaissmaier, 2011; Shah & Oppenheimer, 2008). When such heuristics are well matched to the environment's structure, decision makers will ignore irrelevant features and only attend to the information necessary to make accurate and efficient choices. With that said, a small number of studies do observe a similar pattern to ours'. They find that adults prefer to plan over an internal model of the environment even when a simpler, model-free strategy would yield comparable payoffs, whereas children adopt a simpler strategy (Nussenbaum et al., 2020; Potter et al., 2017; Decker et al., 2016). One explanation is that adults generalize the decision strategies they use in uncertain real-world situations to the novel task environment. In these cases, maintaining more granular representations and being sensitive to uncertainty may be beneficial. As a result, adults may rely on more complex representations not because the task explicitly incentivizes them to do so, but because past experience has taught them that it is an advantageous strategy.

While we observed age-related differences in structure inference, we found no differences in uncertainty-adaptive planning. Even the youngest participants remained in patches longer following rare switches in patch type, indicating sensitivity to uncertainty according to our model. Although children generally plan less deeply than adults (Nussenbaum et al., 2020; Potter et al., 2017; Decker et al., 2016; Albert & Steinberg, 2011; Luciana, Collins, Olson, & Schissel, 2009), they may engage in uncertainty-adaptive planning because it is a resource-efficient strategy. Planning far into the future is only advantageous when the planner can accurately predict action outcomes. Reserving planning for states of relative certainty allows child and adolescent decision makers to make more efficient use of their limited cognitive resources. Importantly, this form of planning primarily relies on reactive rather than proactive control, meaning that control is recruited in response to unexpected changes rather than being continuously deployed in anticipation. Indeed, young children are known to favor reactive over proactive control (Chevalier, Dauvier, & Blaye, 2018; Chevalier, Meaney, Traut, & Munakata, 2020; Niebaum, Chevalier, Guild, & Munakata, 2021). Moreover, uncertainty-adaptive planning depends on the ability to monitor internal uncertainty, an ability that has been observed in children far younger than those in our sample (Baer & Kidd, 2022; Lapidow, Killeen, & Walker, 2022; Schulz, Wu, Ruggeri, & Meder, 2019; de Eccher et al., 2024). Collectively, these findings suggest that even young learners possess the metacognitive abilities needed to flexibly adjust their planning in response to uncertainty.

While all of our key findings are supported by both computational model-agnostic and model-based analyses, our model-based approach is limited by the moderate recoverability of some model parameters, most notably the

parameter controlling uncertainty-adaptive planning,  $\gamma_{coef}$ . Limited parameter recoverability may have reduced our ability to detect age-related differences in  $\gamma_{coef}$ , as well as in other parameters such as the softmax temperature and lapse rate. In addition, we did not fit the structure inference parameter,  $\alpha$ , as a continuous free parameter. Instead, we compared models with low versus high values of  $\alpha$ , as continuous estimation yielded poor parameter recovery. This discrete approach yielded good model recovery, but at the cost of limiting our ability to detect more gradual age-related changes in structure inference. Future work could address these limitations by increasing the number of choices participants made. Indeed, we find that adding an additional block improves parameter recovery (S1.2). In developmental samples, however, collecting more choices is challenging due to children's limited attentional stamina. One promising solution could be to administer the task across multiple shorter sessions. In addition to improving recovery, this approach would allow for direct assessment of parameter reliability across time, providing evidence that the estimated parameters capture stable, trait-like differences between individuals.

Although we observed an overall age-related difference in structure inference, substantial variability remained among participants of similar ages. Among children and adolescents, such heterogeneity may reflect differences in developmental trajectories. However, we also observed considerable inter-individual variability in adults, suggesting that development alone does not fully account for differences in structural learning. We did not directly assess cognitive capacities such as working memory or cognitive control, but variation in these abilities may help explain the individual differences we observed. Prior work indicates that these abilities support the acquisition and use of internal models of the environment (Otto, Raio, Chiang, Phelps, & Daw, 2013; Otto, Skatova, Madlon-Kay, & Daw, 2015).

Prior work has shown that children can discover latent structure that adults cannot, often in cases where the structure is unusual (Gopnik et al., 2017; Lucas et al., 2014). This advantage has been attributed to children's broader exploration. Although children in our task did explore more broadly, leaving patches sooner, this did not provide them with more veridical representations of the environment. One possibility is that broad exploration is not well-suited to structure discovery in every environment. Under standard definitions, exploratory actions are those that sacrifice immediate reward to gain information (Schulz & Gershman, 2019). In patch foraging tasks, remaining longer under uncertainty fits this definition. Doing so allows learners to gather additional observations that refine estimates of local depletion rates, even as it entails deviating from the reward-maximizing behavior prescribed by the Marginal Value Theorem. Importantly, different forms of exploration may support learning about

different aspects of the environment. Leaving patches sooner provides the opportunity to explore a greater number of patches, improving the forager's estimate of the global reward rate. Staying longer, meanwhile, refines knowledge of local patch dynamics. Both are forms of exploration but ones operating at different scales. Leaving sooner is a form of broad exploration while leaving later, though often interpreted as over-exploitation, may in fact be a form of deep exploration. Developmental differences in the balance between these exploratory strategies could provide one potential explanation for why foragers of different ages may acquire different forms of structural knowledge.

Together, our findings provide new insight into how structure learning and its use in complex environments change across development. Through computational modeling, we inferred how participants represented the environment. Notably, children's broader exploration did not confer a structure-learning advantage in this task. Instead, adults' tendency toward deep exploration supported structure inference. These results point toward a more nuanced view of developmental differences in not only structure learning but also exploration. The present work motivates future studies examining how age-related differences in information-seeking interact with the true structure of the environment to produce developmental differences in mental models of the environment.

## References

Acerbi, L., & Ji, W. (2017). Practical bayesian optimization for model fitting with bayesian adaptive direct search. *Adv. Neural Inf. Process. Syst.*, 1836–1846.

Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L., & Platt, M. L. (2017). A primer on foraging and the Explore/Exploit Trade-Off for psychiatry research. *Neuropsychopharmacology*, 42(10), 1931–1939.

Albert, D., & Steinberg, L. (2011, September). Age differences in strategic planning as indexed by the tower of london: Strategic planning. *Child Dev.*, 82(5), 1501–1517.

Aldous, D. J. (1985). Exchangeability and related topics. In D. J. Aldous, I. A. Ibragimov, & J. Jacod (Eds.), *École d'été de probabilités de Saint-Flour XIII — 1983* (pp. 1–198). Springer Berlin Heidelberg.

Alparone, F. R., & Pacilli, M. G. (2012, February). On children's independent mobility: the interplay of demographic, environmental, and psychosocial factors. *Child. Geogr.*, 10(1), 109–122.

Arciuli, J., & Simpson, I. C. (2011, May). Statistical learning in typically developing children: the role of age and speed of stimulus presentation: Statistical learning. *Dev. Sci.*, 14(3), 464–473.

Baer, C., & Kidd, C. (2022). Learning with certainty in childhood. *Trends Cogn. Sci.*, 26(10), 887–896.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.*, 68(3).

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models.

Bergstra, J., & Bengio, Y. (2012). *Random search for hyper-parameter optimization* (Vol. 13).

Blanco, N. J., & Sloutsky, V. M. (2021, March). Systematic exploration and uncertainty dominate young children's choices. *Dev. Sci.*, 24(2), e13026.

Blanco, N. J., & Sloutsky, V. M. (2024). Exploration, exploitation, and development: Developmental shifts in decision-making. *Child Dev.*.

Borawski, E. A., Ievers-Landis, C. E., Lovegreen, L. D., & Trapl, E. S. (2003, August). Parental monitoring, negotiated unsupervised time, and parental trust: the role of perceived parenting practices in adolescent health risk behaviors. *J. Adolesc. Health*, 33(2), 60–70.

Boyce, M. S. (2006, May). Scale for resource selection functions. *Divers. Distrib.*, 12(3), 269–276.

Bullock, M. (1985). Causal reasoning and developmental change over the preschool years. *Hum. Dev.*, 28(4), 169–191.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Curr. Opin. Neurobiol.*, 17(2), 243–250.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.*, 9(2), 129–136.

Chevalier, N., Dauvier, B., & Blaye, A. (2018). From prioritizing objects to prioritizing cues: a developmental shift for cognitive control. *Dev. Sci.*, 21(2).

Chevalier, N., Meaney, J. A., Traut, H. J., & Munakata, Y. (2020). Adaptiveness in proactive control engagement in children and adults. *Dev. Cogn. Neurosci.*, 46(100870), 100870.

Christakou, A., Gershman, S. J., Niv, Y., Simmons, A., Brammer, M., & Rubia, K. (2013). Neural and psychological maturation of decision-making in adolescence and young adulthood. *J. Cogn. Neurosci.*, 25(11), 1807–1823.

Cohen, A. O., Nussenbaum, K., Dorfman, H. M., Gershman, S. J., & Hartley, C. A. (2020). The rational use of causal inference to guide reinforcement learning strengthens with age. *NPJ Sci Learn*, 5, 16.

Coughlin, C., Robins, R. W., & Ghetti, S. (2019). Development of episodic prospection: Factors underlying improvements in middle and late childhood. *Child Dev.*, 90(4), 1109–1122.

Decker, J. H., Otto, A. R., Daw, N. D., & Hartley, C. A. (2016). From creatures of habit to Goal-Directed learners: Tracking the developmental emergence of Model-Based reinforcement learning. *Psychol. Sci.*, 27(6), 848–858.

de Eccher, M., Mundry, R., & Mani, N. (2024). Children's subjective uncertainty-driven sampling behaviour. *R. Soc. Open Sci.*, 11(4), 231283.

Deng, W. S., & Sloutsky, V. M. (2015, March). The development of categorization: effects of classification and inference training on category representation. *Dev. Psychol.*, 51(3), 392–405.

Deng, W. S., & Sloutsky, V. M. (2016, December). Selective attention, diffused attention, and the development of categorization. *Cogn. Psychol.*, 91, 24–62.

Dubois, M., Bowler, A., Moses-Payne, M. E., Habicht, J., Moran, R., Steinbeis, N., & Hauser, T. U. (2022). Exploration heuristics decrease during youth. *Cogn. Affect. Behav. Neurosci.*, 22(5), 969–983.

Fandakova, Y., Selmeczy, D., Leckey, S., Grimm, K. J., Wendelken, C., Bunge, S. A., & Ghetti, S. (2017, July). Changes in ventromedial prefrontal and insular cortex support the development of metamemory from childhood into adolescence. *Proc. Natl. Acad. Sci. U. S. A.*, 114(29), 7582–7587.

Fiser, J., & Aslin, R. N. (2002, November). Statistical learning of new visual feature combinations by infants. *Proc. Natl. Acad. Sci. U. S. A.*, 99(24), 15822–15826.

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychol. Rev.*, 117(1), 197–209.

Ghetti, S., Lyons, K. E., Lazzarin, F., & Cornoldi, C. (2008, March). The development of metamemory monitoring during retrieval: the case of memory strength and memory absence. *J. Exp. Child Psychol.*, 99(3), 157–181.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annu. Rev. Psychol.*, 62(1), 451–482.

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. *Nat Hum Behav.*

Gopnik, A. (2020). Childhood as a solution to explore-exploit tensions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 375(1803), 20190502.

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. *Proc. Natl. Acad. Sci. U. S. A.*, 114(30), 7892–7899.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Dev. Psychol.*,

37(5), 620–629.

Griffiths, T. L., Navarro, D. J., & Sanborn, A. N. (2006). A more rational model of categorization. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 28(28).

Grima, L. L., Haberkern, H., Mohanta, R., Morimoto, M. M., Rajagopalan, A. E., & Scholey, E. V. (2025, October). Foraging as an ethological framework for neuroscience. *Trends Neurosci.*, 48(11), 877–890.

Harhen, N. C., & Bornstein, A. M. (2023). Overharvesting in human patch foraging reflects rational structure learning and adaptive planning. *Proc. Natl. Acad. Sci. U. S. A.*, 120(13), e2216524120.

Hartley, C. A., Nussenbaum, K., & Cohen, A. O. (2021). Interactive development of adaptive learning and memory. *Annual Review of Developmental Psychology*, 3, 59–85.

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022, June). People construct simplified mental representations to plan. *Nature*, 606(7912), 129–136.

Jepma, M., Schaaf, J. V., Visser, I., & Huizenga, H. M. (2020). Uncertainty-driven regulation of learning and exploration in adolescents: A computational account. *PLoS Comput. Biol.*, 16(9), e1008276.

Jiang, N., Kulesza, A., Singh, S., & Lewis, R. L. (2015). The dependence of effective planning horizon on model accuracy. *Adapt Agent Multi-agent Syst*, 1181–1189.

Jung, Y., Walther, D. B., & Finn, A. S. (2021, September). Children automatically abstract categorical regularities during statistical learning. *Dev. Sci.*, 24(5), e13072.

Karmiloff-Smith, A. (1995). *Beyond modularity*. The MIT Press, Massachusetts Institute of Technology.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007, May). Learning overhypotheses with hierarchical bayesian models. *Dev. Sci.*, 10(3), 307–321.

Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M., & Newcombe, N. S. (2018, August). Hippocampal maturation drives memory from generalization to specificity. *Trends Cogn. Sci.*, 22(8), 676–686.

Koechlin, E., Ody, C., & Kouneiher, F. (2003, November). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648), 1181–1185.

Krebs, S. S., & Roebers, C. M. (2010, September). Children's strategic regulation, metacognitive monitoring, and control processes during test taking. *Br. J. Educ. Psychol.*, 80(Pt 3), 325–340.

Lapidow, E., Killeen, I., & Walker, C. M. (2022). Learning to recognize uncertainty vs. recognizing uncertainty to learn: Confidence judgments and exploration decisions in preschoolers. *Dev. Sci.*, 25(2), e13178.

Lapidow, E., & Walker, C. M. (2020). The search for invariance: Repeated positive testing serves the goals of causal learning. In *Language and concept acquisition from infancy through childhood* (pp. 197–219). Cham: Springer International Publishing.

Legendre, P. (1993, September). Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74(6), 1659–1673.

Legendre, P., & Fortin, M. J. (1989, June). Spatial pattern and ecological analysis. *Vegetatio*, 80(2), 107–138.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, 11(2), 173–186.

Lieder, F., & Griffiths, T. L. (2019, February). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behav. Brain Sci.*, 43, e1.

Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, 218, 104940.

Lloyd, A., Viding, E., McKay, R., & Furl, N. (2023, July). Understanding patch foraging strategies across development. *Trends Cogn. Sci.*.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: developmental differences in learning the forms of causal relationships. *Cognition*, 131(2), 284–299.

Luciana, M., Collins, P. F., Olson, E. A., & Schissel, A. M. (2009). Tower of london performance in healthy adolescents: the development of planning skills and associations with self-reported inattention and impulsivity. *Dev. Neuropsychol.*, 34(4), 461–475.

Meder, B., Wu, C. M., Schulz, E., & Ruggeri, A. (2021). Development of directed and random exploration in children. *Dev. Sci.*, 24(4), e13095.

Mikkelsen, M. R., & Christensen, P. (2009, March). Is children's independent mobility really independent? a study of children's mobility combining ethnography and GPS/mobile phone Technologies1. *Mobilities*, 4(1), 37–58.

Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: insights from ethology. *Nat. Rev. Neurosci.*, 19(7), 419–427.

Morris, A. S., Ratliff, E. L., Cosgrove, K. T., & Steinberg, L. (2021, December). We know even more things: A decade review of parenting research. *J. Res. Adolesc.*, 31(4), 870–888.

Niebaum, J. C., Chevalier, N., Guild, R. M., & Munakata, Y. (2021). Developing adaptive control: Age-related differences in task choices and awareness of proactive and reactive control demands. *Cogn. Affect. Behav.*

*Neurosci.*, 21(3), 561–572.

Nussenbaum, K., Martin, R. E., Maulhardt, S., Yang, Y. J., Bizzell-Hatcher, G., Bhatt, N. S., ... Hartley, C. A. (2023). Novelty and uncertainty differentially drive exploration across development. *Elife*, 12.

Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra Psychol.*, 6(1).

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013, December). Working-memory capacity protects model-based learning from stress. *Proc. Natl. Acad. Sci. U. S. A.*, 110(52), 20941–20946.

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015, February). Cognitive control predicts use of model-based reinforcement learning. *J. Cogn. Neurosci.*, 27(2), 319–333.

Perfors, A., & Tenenbaum, J. (2009). Learning to learn categories. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 31(31).

Persaud, K., Bass, I., Colantonio, J., Macias, C., & Bonawitz, E. (2020, March). Opportunities and challenges integrating resource-rational analysis with developmental perspectives. *Behav. Brain Sci.*, 43(e18), e18.

Plebanek, D. J., & Sloutsky, V. M. (2017, June). Costs of selective attention: When children notice what adults miss. *Psychol. Sci.*, 28(6), 723–732.

Potter, T. C. S., Bryce, N. V., & Hartley, C. A. (2017). Cognitive components underpinning the development of model-based learning. *Dev. Cogn. Neurosci.*, 25, 272–280.

Raviv, L., & Arnon, I. (2018, July). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Dev. Sci.*, 21(4), e12593.

Renardy, M., Joslyn, L. R., Millar, J. A., & Kirschner, D. E. (2021). To sobol or not to sobol? the effects of sampling schemes in systems biology applications. *Math. Biosci.*, 337, 108593.

Roderer, T., & Roebers, C. M. (2010, December). Explicit and implicit confidence judgments and developmental differences in metamemory: an eye-tracking approach. *Metacogn. Learn.*, 5(3), 229–250.

Ruggeri, A., Stanciu, O., Pelz, M., Gopnik, A., & Schulz, E. (2024, January). Preschoolers search longer when there is more information to be gained. *Dev. Sci.*, 27(1), e13411.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *J. Cogn. Neurosci.*, *29*(1), 37–51.

Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Curr. Opin. Neurobiol.*, *55*, 7–14.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychol. Sci.*, *30*(11), 1561–1572.

Shah, A. K., & Oppenheimer, D. M. (2008, March). Heuristics made easy: an effort-reduction framework. *Psychol. Bull.*, *134*(2), 207–222.

Shin, Y. S., & DuBrow, S. (2021, January). Structuring memory through inference-based event segmentation. *Top. Cogn. Sci.*, *13*(1), 106–127.

Shufaniya, A., & Arnon, I. (2018). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cogn. Sci.*, *42*(8), 3100–3115.

Simon, H. A. (1975, April). The functional equivalence of problem solving skills. *Cogn. Psychol.*, *7*(2), 268–288.

Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15–18). London: Palgrave Macmillan UK.

Smith, L. B., & Kemler, D. G. (1977, October). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *J. Exp. Child Psychol.*, *24*(2), 279–298.

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017, November). The hippocampus as a predictive map. *Nat. Neurosci.*, *20*(11), 1643–1653.

Steinberg, L. (2001, March). We know some things: Parent–adolescent relationships in retrospect and prospect. *J. Res. Adolesc.*, *11*(1), 1–19.

Tversky, A., & Kahneman, D. (1974, September). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Tymula, A., Rosenberg Belmaker, L. A., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012, October). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proc. Natl. Acad. Sci. U. S. A.*, *109*(42), 17135–17140.

van den Bos, W., & Hertwig, R. (2017, January). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Sci. Rep.*, *7*(1), 40962.

von der Linden, N., & Roebers, C. M. (2006, December). Developmental changes in uncertainty monitoring during an event recall task. *Metacogn. Learn.*, 1(3), 213–228.

Wan, Q., & Sloutsky, V. M. (2024). Exploration, distributed attention, and development of category learning. *Psychol. Sci.*, 9567976241258146.

Weil, L. G., Fleming, S. M., Dumontheil, I., Kilford, E. J., Weil, R. S., Rees, G., ... Blakemore, S.-J. (2013, March). The development of metacognitive ability in adolescence. *Conscious. Cogn.*, 22(1), 264–271.