

- 1.给本文拟定一个abstract（30分）；
  - 2.讨论本研究的贡献、意义和不足之处（30分）；
  - 3.讨论未来可能的研究方向（40分）。
- （第一题用英文作答，后两题中英文作答皆可）

**A**cross animal species, an organism's survival depends on its ability to adapt to the conditions of its environment (1). In humans, experiences in childhood and adolescence provoke strategies of decision-making that are adaptive in those environmental conditions and that can persist into adulthood (2). Early experiences of adversity have been associated with significant negative psychosocial outcomes (3). However, these outcomes may result from strategies of decision-making, which individuals adopt to cope with their early caregiving environments (4). In the present study, we examine how exposure to early adversity affects exploration and sensitivity to reward feedback while foraging.

Adverse childhood experiences (ACEs) are events that are extreme stressors experienced by an individual during development (ages 0 to 18 y old) (3). ACEs can be categorized into three broad groups: threatening events, which are directly experienced by the individual (e.g., physical abuse), neglect (e.g., emotional neglect), or household adversity, which refers to circumstances in the individuals' environment that can cause high levels of stress (e.g., parental divorce) (5). Notably, higher rates of ACEs are associated with poorer health and social outcomes in adulthood, such as substance misuse and antisocial behavior (6, 7).

A theoretical account known as life history theory proposes that adverse rearing conditions direct individuals to adopt strategies that maximize short-term benefits (8). This behavior may manifest through reduced exploration and greater delay discounting, both of which indicate a preference for rewards that are immediately available, compared to greater but more-delayed rewards. Behaviorally, greater delay discounting is negatively correlated with exploration, suggesting that these preferences both reflect cognitive processes related to the individual's temporal horizon (9). Consistent with life history theory, adolescent rats exposed to early stress explore their environment less compared

to controls (10). Moreover, in human adolescents, adversity related to childhood poverty is associated with a preference for immediate rather than delayed rewards in delay discounting paradigms (11). These findings are notable as decision-making that prioritizes short-term rewards can lead to poorer socioeconomic outcomes (12) and has also been linked to problematic health behaviors (e.g., substance misuse) (13). While individuals might adapt to their surroundings in their formative years by adopting this decision-making strategy, its continued use can lead to poorer psychosocial outcomes later in life.

The adoption of a decision-making strategy that is focused on exploiting immediate rewards from the environment may be beneficial in resource-scarce conditions, which are characteristic of the rearing context of ACE-exposed individuals (14). This benefit has been demonstrated in adolescents with experience of early life stress (specifically events leading to institutionalization) using a variant of the Balloon Analog Risk Task (BART) (15). In a study by Humphreys and colleagues (4), adolescents who had experienced institutionalization explored less when the environment favored less exploration, meaning they collected more rewards in this condition compared to adolescents with more-stable upbringings (4). Furthermore, in this same study, adolescents with experience of institutionalization explored less when the optimal decision-making strategy was to explore more, meaning that they collected fewer rewards in this condition compared to adolescents from stable backgrounds. Yet, a potential limitation associated with the variant of the BART used by Humphreys and colleagues (4) is that choosing to explore can lead the participant to lose any unbanked points they have accumulated if they pump beyond the balloon's limit. This may be problematic as early adversity is associated with a heightened sensitivity toward negative feedback (16), and

therefore, previous research demonstrating reduced exploration following early stress may be confounded by heightened loss aversion in this population (17). Moreover, while the effects of early adversity on exploration have been investigated in adolescents, it is unknown whether ACEs in humans lead to reduced levels of exploration in adulthood.

This previous literature poses new hypotheses that we tested here. First, we examined whether forms of early stress other than institutionalization led to a reduction in exploration in adulthood. Humphreys and colleagues' (4) focus on institutionalization parallels the rodent literature, which has manipulated early stress through mother–infant separation. Here, we adopted a wider definition of early adversity that encompasses experiences that are more common in the general human population (18, 19). Identifying how ACEs affect exploration can contribute toward understanding how more prevalent forms of early stress affect decision-making across the lifespan. In addition, we examined whether adults who had experienced early adversity also adopted exploration strategies that are optimal in environments that favor exploitation.

A canonical paradigm for studying how organisms explore their environment is patch-foraging. When foraging, an agent decides whether to remain with a known patch to exploit rewards from it or to explore a novel patch that has a fresh distribution of rewards (20, 21). Whether the organism exploits a current patch or explores a novel one should depend on the richness of the environment, which refers to the average number of rewards accumulated while foraging (22). This is formalized in a computational account of foraging known as Marginal Value Theorem (MVT). The theorem proves that to maximize reward intake, the forager should opt to explore when the rewards expected from exploiting the present patch fall below the average reward rate for the environment (20). Previous research has found that human adults adjust their foraging strategies according to qualities of the environment so as to maximize rewards (23). Nevertheless, adults explore less than an optimal foraging strategy would dictate (24). Here, we examine whether this tendency to under-explore is particularly pronounced in ACE-exposed individuals.

If early adversity indeed reduces exploration later in life, such an adaptive decision-making strategy might arise from the computational mechanisms that ACE-exposed individuals use to learn from reward feedback. Specifically, the rate at which individuals learn associations between stimuli and outcomes may provide insight into why ACE-exposed individuals prefer to exploit immediate rewards rather than potentially delaying reward by exploring. In adverse household conditions, there is inconsistency in caregiver behavior and the individual learns that positive and negative behaviors occur at random (25). Indeed, a recent paper has characterized early adversity as a violation of environmental predictability, which has profound consequences for sensitive periods of development (26). On this account, it is argued that experiences that the child should expect, such as parental care, are either unreliable or atypical in adverse households. For example, adverse households can involve frequent and unpredictable threats to survival, such as instances of physical assault (3). It has been argued that such conditions of adversity can lead to schemas of unpredictability, in which the world is perceived as unstable (27). Consistent with this view, individuals exposed to childhood adversity develop neurobiological and behavioral adaptations to navigate changeable environments, such as rapidly shifting attention (28). Together, this evidence suggests that individuals who have been exposed to ACEs perceive the environment as relatively volatile, which can impact neurobiological and behavioral outcomes (3).

In rapidly changeable or volatile environments, knowledge of the more distant history of reward feedback has less utility in

predicting future outcomes than recent feedback (29). As such, decisions in volatile environments should be based on more recent feedback about whether actions will be rewarded (30). As a real-world example, during the COVID-19 pandemic, information regarding health behaviors rapidly changed in response to the spread of the virus, and as such, adaptive decision-making involved utilizing the most-recent information to make health-related choices (31). The utility of relying on recent feedback in volatile environments has been demonstrated in instrumental learning paradigms, in which participants learn the likelihood of receiving rewards from sampling stimuli with probabilistic reward schedules. In environments where the probabilities that relate stimuli, actions, and reward to each other are volatile, the decision-maker should update knowledge quickly in response to recent feedback (29). This ability—the “learning rate”—is quantified in most traditional, formal reinforcement-learning models by a parameter known as alpha (in the specific implementation we use here, the learning rate equals  $1 - \alpha$ ). The learning rate measures how highly individuals weight recent feedback relative to more historic feedback, with higher rates indicating greater emphasis on recent events (32). By contrast, in environments where the probability of receiving rewards from each action remains stable, the optimal learning strategy is to utilize a wider range of historic experience during decision-making to avoid overweighting rare events (i.e., use a lower learning rate) (29).

To our knowledge, no research to date has examined whether early life stress is associated with differences in how highly individuals weight recent feedback (i.e., how high their learning rate is). Certain disorders can lead individuals to overestimate the volatility of stable environments, which can lead to heightened emphasis placed on recent events (e.g., Autism Spectrum Disorder) (33). Recently, it has been suggested that adverse experiences might lead to atypical learning strategies, which could explain why early adversity is linked to the onset of emotional disorders such as anxiety and depression (34). Neuroimaging research has demonstrated that ACEs lead to reduced volume of the anterior cingulate cortex (ACC) (35), a region that has been implicated in tracking environmental volatility (29) and the value of exploring new patches while foraging (36, 37) (however, refer to ref. 38). As such, we hypothesized that ACEs would impact ACC-mediated learning mechanisms capable of adjusting learning rates to the (in)stability of the environment. Specifically, as exposure to ACEs disposes the individual to perceive the environment as unstable (3), we predicted that in adulthood, these experiences will be associated with overestimating the volatility of stable environments, reflected by a higher learning rate.

The current study investigated how early experiences of adversity impact decision-making. We measured exploration behavior (leaving thresholds) on a patch-foraging task in individuals with more or fewer ACEs and fit a reinforcement-learning model (23) to their behavior to estimate their rate of learning from reward feedback. We preregistered three hypotheses. Our first hypothesis had two parts: Hypothesis 1a was that participants with a high number of ACEs would explore less (i.e., exploit patches for longer) compared to those in the low-ACE group. Hypothesis 1b was that participants with high rates of ACEs would weight recent evidence higher (as represented by their learning rates) than participants with fewer ACEs. Hypothesis 1b also predicted that higher learning rates would be associated with lower leaving thresholds in patches (i.e., less exploration). Our second hypothesis (Hypothesis 2) was that ACE-related decision strategies would lead to real-world problematic outcomes in the form of a positive relationship between ACEs and self-reported risk-taking.

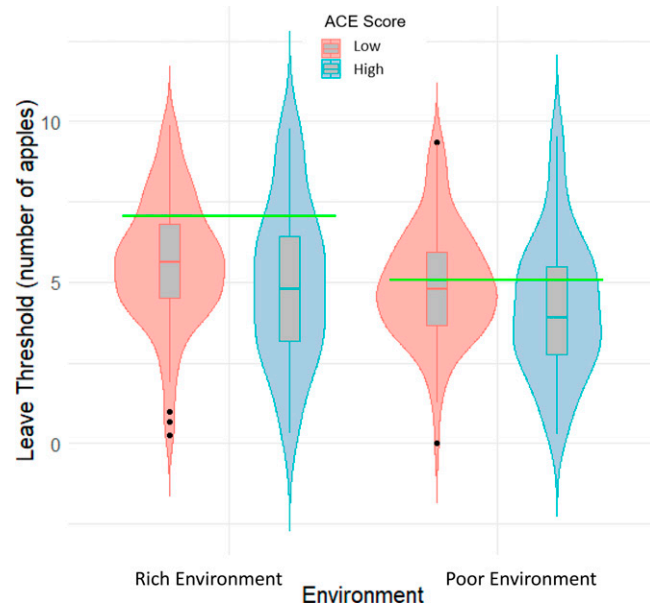
As adults exposed to ACEs are expected to explore less (Hypothesis 1a), they should be closer to optimal in conditions

where exploitation garners greater rewards (4). As such, the first part of our third hypothesis (3a) was that participants who reported higher levels of ACEs would demonstrate more optimal exploration in the “poorer” task environment, where the better strategy is to explore less, compared to participants with lower levels of ACEs. Complementing this, the second part of the third hypothesis (3b) was that participants with lower levels of ACEs would demonstrate more optimal exploration in the “richer” task environment, where the better strategy was to explore more, compared to participants with higher levels of ACEs. Addressing these questions can inform our understanding of the computational mechanisms underlying different decision-making strategies associated with early adversity and their relationship with risk-taking behaviors.

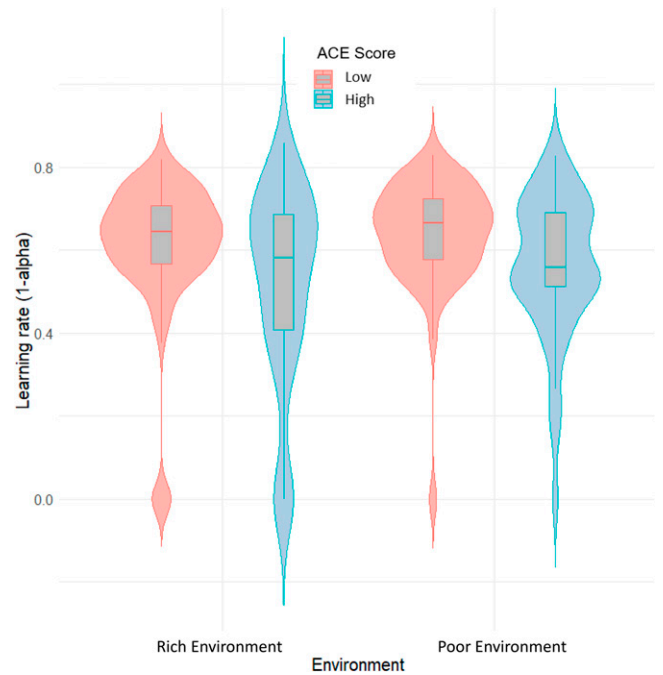
## Results

To test Hypothesis 1a that individuals with a high number of ACEs would explore less (i.e., would have a lower leaving threshold) compared to individuals with a low number of ACEs, we ran a mixed ANOVA with the foraging environment (rich or poor) as the within-subject factor and ACE exposure (high or low) as the between-subject factor. We replicated findings (23, 24) that in the rich environment, participants had a higher leaving threshold than in the poor-quality environment [ $F_{(1,137)} = 28.26, P < 0.001$ , and  $\eta^2 = 0.03$ ]. Furthermore, participants in the high-ACE group remained in patches significantly longer (i.e., explored less) than participants with less exposure to ACEs [ $F_{(1,137)} = 4.46, P = 0.037$ , and  $\eta^2 = 0.03$ ] (Fig. 1). There was no interaction between environment type and ACE exposure [ $F_{(1,137)} = 0.63, P = 0.429$ , and  $\eta^2 < 0.001$ ]. These effects were robust to the addition of gender as a covariate, an analysis designed to account for the overrepresentation of women in the high-ACE group.

Examining Hypothesis 1b, a mixed ANOVA demonstrated that individuals in the high-ACE group had a lower mean learning rate across the two environments compared to the low-ACE



**Fig. 1.** Violin and boxplots demonstrating the leaving thresholds for each environment. The central line in each boxplot refers to the mean. *Upper* and *Lower* lines reflect the upper and lower quartiles, respectively. Black dots identify outliers that are greater than 1.5 times the interquartile range from the upper and lower quartiles. Horizontal green bars indicate the optimal leaving thresholds for each environment. Higher leaving threshold indicates more exploration and less exploitation.



**Fig. 2.** Violin plot demonstrating the difference in mean learning rate (y-axis) between the two environments (x-axis) for both the high-ACE and low-ACE groups. The central line in each boxplot refers to the mean. *Upper* and *Lower* lines reflect the upper and lower quartiles, respectively.

group [ $F_{(1,137)} = 8.92, P = 0.003$ , and  $\eta^2 = 0.05$ ]. This finding was in the opposite direction to our hypothesis and suggested that those in the high-ACE group weighted recent feedback lower than those in the low-ACE group (Fig. 2). This analysis also revealed that participants adjusted their learning rate between the two environments [ $F_{(1,137)} = 9.63, P = 0.002$ , and  $\eta^2 = 0.01$ ], as participants utilized a lower learning rate in the rich environment compared to the poor environment. However, there was no interaction between environment and ACE score with respect to participants' learning rate [ $F_{(1,137)} = 0.40, P = 0.527$ , and  $\eta^2 < 0.001$ ]. This finding was also robust to the addition of gender as a covariate. We found no significant differences between the ACE groups on the other two free parameters, associated with the reinforcement-learning model, including the beta parameter [ $F_{(1,137)} = 0.25, P = 0.621$ , and  $\eta^2 < 0.001$ ] and the intercept parameter  $c$  [ $F_{(1,137)} = 0.76, P = 0.377$ , and  $\eta^2 = 0.003$ ].

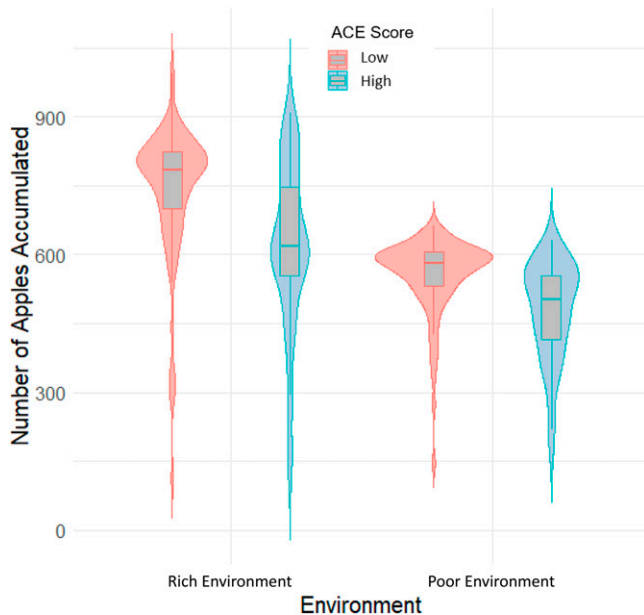
Inconsistent with the second part of Hypothesis 1b, there was a significant negative correlation between the alpha parameter and leaving thresholds in both the rich-quality environment [ $r(149) = -0.91, P < 0.001$ ] and the poor-quality environment [ $r(149) = -0.71, P < 0.001$ ]. As the learning rate was equal to  $1 - \alpha$ , this suggested that in both environments, weighting recent feedback higher was associated with higher rates of exploration. We also simulated data using the mean parameter estimate from the high and low-ACE groups to examine whether we could recapitulate the trends observed in participants' data. These simulations demonstrated that the free parameters estimated from participants' behavior were able to reproduce the differences in leaving thresholds that we observed between the high- and low-ACE groups as well as between environments.

To examine whether ACEs are associated with heightened risk-taking (Hypothesis 2), we conducted several regression analyses. Each subscale of the Domain-Specific Risk-Taking scale (DOSPERT) was entered as an outcome variable, and ACE score, gender, and age were entered as predictors. We did not find support for the hypothesis that ACEs were associated with

more risk-taking. However, across all risk domains (with the exception of the social domain), being male was significantly positively associated with risk-taking.

Our third set of hypotheses predicted that because of their unstable backgrounds, individuals with high levels of ACEs would adopt closer to optimal leaving thresholds in the poorer environment, where the optimal strategy was to explore less than participants with low levels of ACEs (Hypothesis 3a). In contrast, we predicted that the high-ACE group would be less optimal than the low-ACE group in the rich environment (Hypothesis 3b). Inconsistent with our predictions, individuals with high ACE scores were further from the optimal leaving threshold in both the rich-quality ( $M = 2.04$ ,  $SD = 2.46$ ) and poor-quality environments ( $M = 0.79$ ,  $SD = 2.16$ ) compared to individuals with low ACE scores [ $M_{rich} = 1.43$ ,  $SD_{rich} = 1.94$ ,  $M_{poor} = 0.03$ ,  $SD_{poor} = 1.68$ ,  $F_{(1,137)} = 4.46$ ,  $P = 0.037$ , and  $\eta^2 = 0.03$ ]. Participants across the board exhibited more optimal behavior in the poor-quality environment than in the rich-quality one [ $F_{(1,137)} = 107.31$ ,  $P < 0.001$ , and  $\eta^2 = 0.09$ ]. There was no interaction between ACE score and environment [ $F_{(1,137)} = 0.63$ ,  $P = 0.429$ , and  $\eta^2 < 0.001$ ].

We conducted exploratory analyses to examine whether ACEs affected the number of rewards accumulated during the task (i.e., the number of apples they harvested in each environment). Results of a mixed ANOVA demonstrated that the high-ACE group collected fewer apples and, hence, were less optimal foragers than the low-ACE group [ $F_{(1,137)} = 24.39$ ,  $P < 0.001$ , and  $\eta^2 = 0.09$ ; Fig. 3]. Further, participants accumulated more points in the rich environment compared to the poor environment, as demonstrated by a main effect of environment type [ $F_{(1,137)} = 187.14$ ,  $P < 0.001$ , and  $\eta^2 = 0.24$ ]. There was not a significant interaction between environment type and ACE group [ $F_{(1,137)} = 3.79$ ,  $P = 0.053$ , and  $\eta^2 = 0.01$ ].



**Fig. 3.** Apples accumulated in each environment by ACE group. The central line in each boxplot refers to the mean. Lines above and below the mean reflect the upper and lower quartiles, respectively.

## Materials and Methods

**Participants.** To selectively recruit participants who had been exposed to ACEs, we advertised among four international charities and support groups for adult survivors of childhood trauma. These were the following: Survivors South West Yorkshire, the National Association for People Abused in Childhood (NAPAC), The Survivor's Trust, and one anonymous support group. Control participants were recruited from a recruitment platform (Sona Systems; <https://www.sona-systems.com/>) hosted by a United Kingdom-based university and through Prolific (<https://www.prolific.co/>). The Prolific sample was recruited from the same regions that the charities were based in the United Kingdom and Europe. We recruited a total of 151 participants ( $M_{age} = 38.91$ ,  $SD_{age} = 11.09$ ), with a mean of 2.66 ACEs ( $SD = 2.76$ ). For group-level analyses, we categorized participants as having experienced a high number of ACEs if they reported  $\geq 4$  ACEs. This threshold was determined based on previous research (e.g., refs. 5 and 19). Six participants did not provide answers for the ACE questionnaire and were excluded from the analyses. Of the final sample, 47 participants met the threshold for the high-ACE group, and 98 were included in the low-ACE group. Age did not significantly differ between the two groups,  $t(143) = 1.28$ ,  $P = 0.202$ , nor did level of education,  $t(141) = 0.01$ ,  $P = 0.991$ . Because there were more females in the high-ACE group compared to the control group  $\chi^2(1, n = 143) = 13.10$ ,  $P < 0.001$ , we controlled for gender in the analyses. Ethical approval for this study was received from Royal Holloway, University of London's ethical review board (reference: Full-Review-2128-2020-04-07-13-13-PFJT001).

### Materials.

**ACEs.** The Adverse Childhood Events Scale (51) is a self-report measure of the number of ACEs that an individual has experienced. The scale details 10 items referring to different categories of ACEs, such as physical abuse, neglect, and parental imprisonment. For each category of ACE, the participant reported whether they experienced this during childhood (between the ages of 0 and 18), which is coded as a binary option (yes/no). Total scores on the measure range from 0 to 10, with higher values denoting the individual has been exposed to more ACEs (Table 1).

**Risk-taking.** The DOSPERT (52) is a 30-item self-report measure assessing an individual's risk-taking propensity. The scale measures five domains of risk-taking the following: financial, health/safety, recreational, ethical, and social risks. Each domain is measured by six items ranging from 1 to 7, in which 1 denotes that the individual would be highly unlikely to engage in that behavior and 7 denotes that the individual is highly likely to engage in that behavior. In the current sample, mean scores for each domain of risk-taking ranged from 1 to 6. The reliability of the scales computed from the current sample were as follows: social  $\alpha = 0.56$ , recreational  $\alpha = 0.85$ , financial  $\alpha = 0.80$ , health/safety  $\alpha = 0.66$ , and ethical  $\alpha = 0.67$ .

**Patch-foraging.** We used a patch-foraging task (Fig. 4) in which participants harvested apples (rewards) from trees (patches) (23). In this task, the decision-maker must decide between exploiting a known patch that gradually yields fewer rewards or exploring a novel patch with a fresh distribution of rewards. We designed the paradigm such that the time it took to exploit apples (i.e., the harvest time) was always 3 s, regardless of participants' reaction times. The 3 s included participants' reaction time and the presentation of rewards on that trial. This design feature was implemented to ensure that faster



**Table 1. Prevalence of each ACE in the current sample**

ACE	Percentage of total sample
Threatening events	
Emotional abuse	44.08
Physical abuse	30.26
Sexual abuse	36.25
Neglect	
Emotional neglect	44.08
Physical neglect	14.47
Family adversity	
Divorce	35.52
Witnessing domestic abuse	15.13
Substance abuse within the household	20.39
Mental illness within the household	35.52
Incarcerated relative	7.24

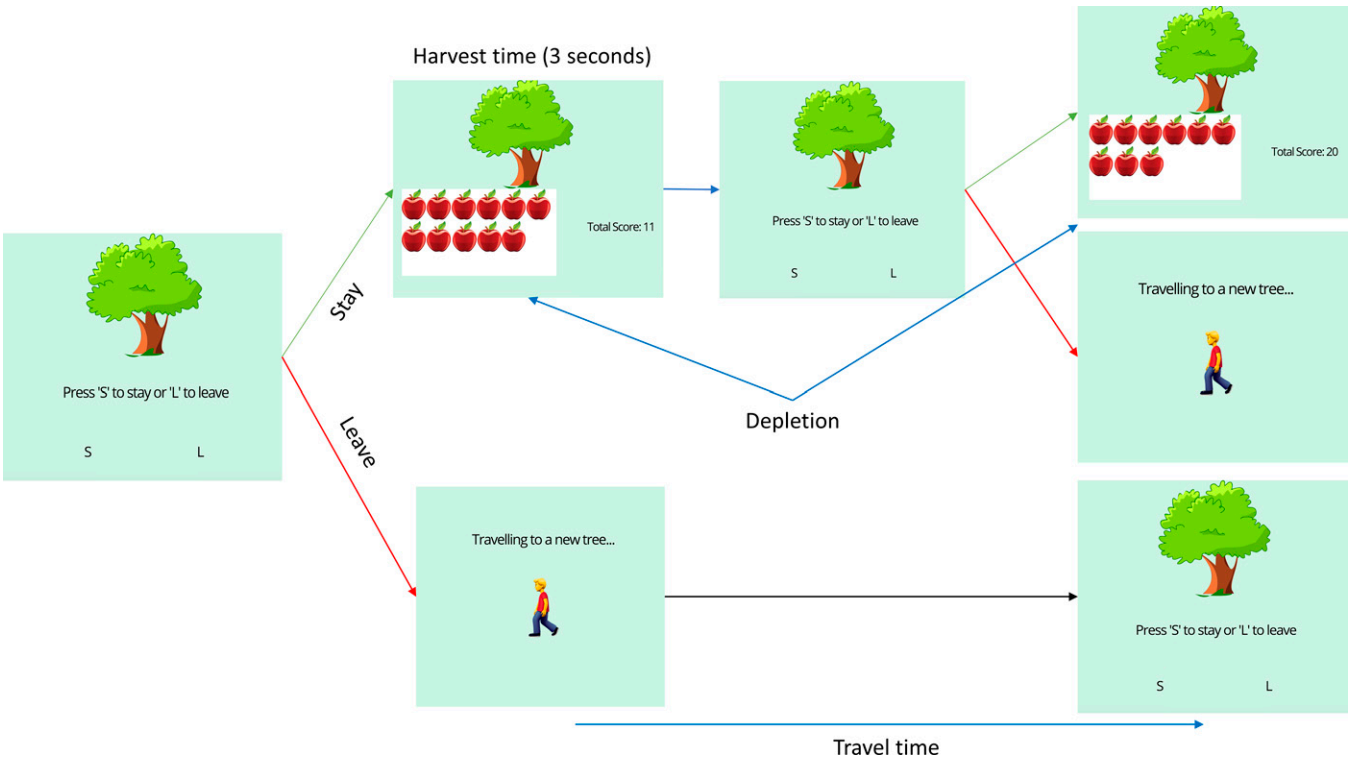
reaction times did not impact how quickly participants could accumulate rewards. For example, a participant who responded in 0.6 s would be presented with their score for 2.4 s. Participants had up to 2 s to make a decision before they were presented with a timeout screen. Timeout trials were subsequently excluded from further analyses.

We presented participants with two foraging environments, which differed in the number of rewards obtainable. In the “poor-quality” environment, the optimal forager should explore less and exploit each patch more (i.e., use a lower leaving threshold) relative to the “rich-quality” environment, where the optimal forager should explore more and exploit each patch less (i.e., use a higher leaving threshold). To maximize the difference between optimal leaving thresholds in the rich and poor foraging environments, we manipulated environment quality both by varying the rate at which rewards exponentially depleted from individual patches and by varying the travel time between patches. Both manipulations were based on previous research (23, 24). The depletion rate was applied using the following formula:  $s_{j+1} = s_j * N_i$ , where  $s$  refers to the reward experienced on each trial ( $i$ ), and  $N_i$  refers to a value drawn on each trial from a Gaussian distribution. In the rich-quality

environment, depletion rates were drawn from a Gaussian with a mean of 0.94 ( $SD = 0.07$ ) and a travel time of 6 s between patches. In the poor-quality environment, depletion rates were drawn from a Gaussian with a mean of 0.88 ( $SD = 0.07$ ) and a 12-s travel time between patches. In both environments, the initial distribution of rewards on each patch ( $s_0$ ) was drawn from a Gaussian distribution with a mean of 10 ( $SD = 1$ ). Participants completed each patch-foraging environment for 7 min in a counterbalanced order.

We utilized two behavioral variables from this task in addition to the learning-rate parameter derived from our computational model (see *Computational Modeling* below). The first was the participants’ leaving thresholds for each environment. As in previous research (23), we assume that participants select an expected value of apples as their leaving threshold and only leave the patch when they expect fewer apples than this value in future harvests. As such, participants who choose to explore earlier will leave patches when there is a relatively high number of apples still expected from the next exploit decision, whereas participants who choose to exploit more will persist in the same patch until the expectation of apples is relatively low. Higher values of leaving threshold denote greater exploration, and lower values denote greater exploitation. We took the average number of apples from the last two harvests when calculating this variable, as in previous research (23). The second dependent variable measured how well participants performed: the difference between participants’ leaving threshold and the optimal leaving threshold. Negative values of this variable suggest that the participant remained in patches for less time than was optimal, whereas positive values suggest that participants remained for longer than was optimal. Following previous research (24, 50), we determined the optimal leave threshold by running a grid search across leaving thresholds between 1 and 10 in increments of 0.001 and summing the total number of apples accrued. The simulation was run for both the rich and poor environments, allowing us to identify the leaving threshold that yielded the highest number of apples and was therefore optimal in that environment. The optimal leaving threshold in the rich-quality environment was 7.04 apples, and in the poor-quality environment, the optimal leaving threshold was 5.07 apples (refer to the green horizontal lines in Fig. 1).

**Procedure.** Participants who chose to take part were provided with a link to the study (or were transferred automatically by Prolific), which was hosted on Gorilla.sc, an online behavioral study platform (53). Participants completed a



**Fig. 4.** Outline of the patch-foraging paradigm. The leftmost screen is presented to participants when they first enter the environment. They decide between staying or leaving the current patch. If participants choose to stay (Top screen), they are presented with the number of apples they have harvested and their cumulative score. They are then returned to the patch to decide whether to stay or leave. If participants choose to leave (Bottom screen), they wait for an environment-specific interval before reaching the new patch.

consent form, which informed them that they would be asked to complete a computerized task. They were also informed that they would be asked questions about their childhood, which might be stressful. Participants were recommended not to take part if answering such questions might cause them a high level of distress. Contact details for the researcher were provided to ensure that participants had the opportunity to ask questions about the study. Participants were paid a base rate of £3 (or equivalent) but were also informed that they could earn a performance-based bonus (up to an additional £2 or the equivalent amount in the participant's currency). Participants were then directed to a page in which they filled out demographic information. Following this, participants were provided with the task instructions and completed a 2-min practice of the patch-foraging task. The depletion rate of the practice environment was drawn from a Gaussian distribution with a mean of 0.90 ( $SD = 0.07$ ), and the travel time was set at 9 s. As such, the parameters of the practice were different to those used in the main task.

After completing the practice, participants completed both the rich- and poor-quality foraging environments. Once both environments were completed, participants were taken to a break screen, informing them they were about to be asked questions that were sensitive in nature, and were reminded they were free to omit any questions they did not wish to answer. Participants then completed the ACE questionnaire and DOSPERT. Finally, participants were provided with a debrief, which included information about support pages for survivors of childhood trauma.

**Computational Modeling.** According to a prominent theory of foraging behavior (MVT) (20), the rational agent aiming to maximize reward intake should leave the current patch when the reward expected from staying within that patch falls below the average reward rate for that environment. However, as such an agent does not know the state of the environment a priori, this must be learned.

To model participants' learning rate, we used an equation developed by Constantino and Daw (23). The model explains how participants estimate the average reward rate ( $p_i$ ) in each environment. Crucially, this estimate depends on a free parameter ( $\alpha$ ), which varies across individuals (32). This parameter captures the degree to which participants weight recent feedback to guide their decision-making:

$$p_i = (1 - \alpha)^T \frac{S_i}{T_i} + \left(1 - (1 - \alpha)^T\right) p_{i-1}. \quad [1]$$

Although traditional reinforcement-learning models directly equate the alpha parameter with the learning rate, it can be seen from Eq. 1 that the model developed by Constantino and Daw (23) parameterizes alpha as the complement of the learning rate. Thus, lower values of alpha mean higher values of

$(1 - \alpha)$  and, hence, a higher learning rate. Participants' estimate of the average reward rate is also subject to the time cost ( $T_i$ ) associated with the participant's explore or exploit decision and the reward experienced on each trial ( $s_i$ ). Participants' estimate of the average reward rate is then entered into a Softmax function, which produces the probability that the participant will stay on each trial:

$$P(a_i = \text{stay}) = 1 / (1 + \exp(-(c + \beta[\kappa_{i-1} s_{i-1} - p_{i-1} h])). \quad [2]$$

This Softmax function [2] contains a further two free parameters. The first,  $\beta$ , captures stochasticity in decision-making. Higher values on this parameter denote that the participant acts more deterministically according to the MVT leaving rule, whereas lower values indicate that the participant is more likely to divert from this decision rule. The second free parameter,  $c$ , is an intercept that estimates participants' exploitation bias. In this equation,  $\kappa_i$  refers to the rate at which apples deplete from patches. As in previous research (23), we assumed that participants use a running estimate of the depletion rate, which we calculated through averaging the depletion rate experienced on previous trials in that environment. On each trial, the true depletion rate was calculated as  $d_i = s_i / s_{i-1}$ . We then calculated participants' running estimate of the average reward rate ( $k_i$ ) through averaging across all values of the depletion rate ( $d_i$ ) experienced in the environment. Participants estimate the reward expected on the next trial by multiplying the last-known reward value ( $s_i$ ) by their estimate of the depletion rate ( $\kappa_i$ ). This value is compared against the estimate of the average reward rate ( $p_i$ ) multiplied by the opportunity cost associated with exploit decisions ( $h$ ), which was fixed at 3 s in both environments. As such, the term  $\kappa_i s_i - p_i h$  captures the difference between the reward that participants expect from the next exploit decision and their current estimate of the average reward rate (23). Parameter recovery indicated that all parameters could be identified uniquely without parameter correlations, though we found that the parameter  $c$  was less well recovered compared to the other parameters in the model. We compared this model, which uses only a single learning rate for all outcomes, to a model in which the learning rate was split for better-than-expected and poorer-than-expected outcomes (54).

**Data Availability.** Anonymized CSV data and analysis code have been deposited in Open Science Framework (<https://osf.io/8znyx/>) (55).

**ACKNOWLEDGMENTS.** We would like to thank the NAPAC, Survivors South West Yorkshire, The Survivor's Trust, and one anonymous support group for their assistance with recruitment for this study.

1. J. M. McNamara, A. I. Houston, Optimal foraging and learning. *J. Theor. Biol.* **117**, 231–249 (1985).
2. W. T. Boyce, B. J. Ellis, Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Dev. Psychopathol.* **17**, 271–301 (2005).
3. A. Danese, B. S. McEwen, Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiol. Behav.* **106**, 29–39 (2012).
4. K. L. Humphreys et al., Exploration-exploitation strategy is dependent on early experience. *Dev. Psychobiol.* **57**, 313–321 (2015).
5. K. Hughes et al., The effect of multiple adverse childhood experiences on health: A systematic review and meta-analysis. *Lancet Public Health* **2**, e356–e366 (2017).
6. M. T. Baglivio, K. T. Wolff, A. R. Piquero, N. Epps, The relationship between adverse childhood experiences (ACE) and juvenile offending trajectories in a juvenile offenders sample. *J. Crim. Justice* **43**, 229–241 (2015).
7. J. P. Mersky, J. Topitzes, A. J. Reynolds, Impacts of adverse childhood experiences on health, mental health, and substance use in early adulthood: A cohort study of an urban, minority sample in the U.S. *Child Abuse Negl.* **37**, 917–925 (2013).
8. B. J. Ellis, M. Del Giudice, Developmental adaptation to stress: An evolutionary perspective. *Annu. Rev. Psychol.* **70**, 111–139 (2019).
9. H. Sadeghiyeh et al., Temporal discounting correlates with directed exploration but not with random exploration. *Sci. Rep.* **10**, 4020 (2020).
10. J. Spivey, D. Barrett, E. Padilla, F. Gonzalez-Lima, Mother-infant separation leads to hypoactive behavior in adolescent Holtzman rats. *Behav. Processes* **79**, 59–65 (2008).
11. A. Oshri et al., Socioeconomic hardship and delayed reward discounting: Associations with working memory and emotional reactivity. *Dev. Cogn. Neurosci.* **37**, 100642 (2019).
12. J. Haushofer, E. Fehr, On the psychology of poverty. *Science* **344**, 862–867 (2014).
13. S. H. Kollins, Delay discounting is associated with substance use in college students. *Addict. Behav.* **28**, 1167–1173 (2003).
14. J. Belsky, L. Steinberg, P. Draper, Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Dev.* **62**, 647–670 (1991).
15. C. W. Lejuez et al., Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *J. Exp. Psychol. Appl.* **8**, 75–84 (2002).
16. A. C. Miu, M. I. Bîlc, I. Bunea, A. Szentágotai-Tátrai, Childhood trauma and sensitivity to reward and punishment: Implications for depressive and anxiety symptoms. *Pers. Individ. Dif.* **119**, 134–140 (2017).
17. H. J. Huh, K. Baek, J.-H. Kwon, J. Jeong, J.-H. Chae, Impact of childhood trauma and cognitive emotion regulation strategies on risk-averse and loss-averse patterns of decision-making in patients with depression. *Cogn. Neuropsychiatry* **21**, 447–461 (2016).
18. Office of National Statistics. *Children Looked after in England (Including Adoptions and Care Leavers): 2019–2020* (Department of Education, 2020).
19. M. A. Bellis et al., *Adverse Childhood Experiences and Their Impact on Health-harming Behaviours in the Welsh Adult Population: Alcohol Use, Drug Use, Violence, Sexual Behaviour, Incarceration, Smoking and Poor Diet*. (Public Health Wales, 2015).
20. E. L. Charnov, Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9**, 129–136 (1976).
21. D. Stephens, J. Krebs, *Foraging Theory* (Princeton University Press, 1986).
22. R. L. Bettinger, M. N. Grote, Marginal value theorem, patch choice, and human foraging response in varying environments. *J. Anthropol. Archaeol.* **42**, 79–87 (2016).
23. S. M. Constantino, N. D. Daw, Learning the opportunity cost of time in a patch-foraging task. *Cogn. Affect. Behav. Neurosci.* **15**, 837–853 (2015).
24. A. Lloyd, R. McKay, C. L. Sebastian, J. H. Balsters, Are adolescents more optimal decision-makers in novel environments? Examining the benefits of heightened exploration in a patch foraging paradigm. *Dev. Sci.* **24**, e13075 (2020).
25. A. M. Novick et al., The effects of early life stress on reward processing. *J. Psychiatr. Res.* **101**, 80–103 (2018).
26. C. A. Nelson III, L. J. Gabard-Durnam, Early adversity and critical periods: Neurodevelopmental consequences of violating the expectable environment. *Trends Neurosci.* **43**, 133–143 (2020).
27. L. Ross, E. Hill, Childhood unpredictability, schemas for unpredictability, and risk taking. *Soc. Behav. Personal.* **30**, 453–473 (2002).
28. P. S. Jensen et al., Evolution and revolution in child psychiatry: ADHD as a disorder of adaptation. *J. Am. Acad. Child. Adolesc. Psychiatry* **12**, 1672–1681 (1997).
29. T. E. J. Behrens, M. W. Woolrich, M. E. Walton, M. F. S. Rushworth, Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).

30. M. Browning, T. E. Behrens, G. Jocham, J. X. O'Reilly, S. J. Bishop, Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).
31. A. Lloyd *et al.*, Delay discounting and under-valuing of recent information predict poorer adherence to social distancing measures during the COVID-19 pandemic. *Sci. Rep.* **11**, 19237 (2021).
32. R. Sutton, A. Barto, *Reinforcement Learning: An Introduction* (MIT Press, 2018).
33. R. P. Lawson, C. Mathys, G. Rees, Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* **20**, 1293–1299 (2017).
34. E. Pulcu, M. Browning, The misestimation of uncertainty in affective disorders. *Trends Cogn. Sci.* **23**, 865–875 (2019).
35. R. A. Cohen *et al.*, Early life stress and morphometry of the adult anterior cingulate cortex and caudate nuclei. *Biol. Psychiatry* **59**, 975–982 (2006).
36. N. Koling, T. E. J. Behrens, R. B. Mars, M. F. S. Rushworth, Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
37. T. C. Blanchard, B. Y. Hayden, Neurons in dorsal anterior cingulate cortex signal post-decisional variables in a foraging task. *J. Neurosci.* **34**, 646–655 (2014).
38. A. Shenhav, M. A. Straccia, J. D. Cohen, M. M. Botvinick, Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat. Neurosci.* **17**, 1249–1254 (2014).
39. M. Metzler, M. T. Merrick, J. Kleven, K. A. Ports, D. C. Ford, Adverse childhood experiences and life opportunities: Shifting the narrative. *Child. Youth Serv. Rev.* **72**, 141–149 (2017).
40. C. Le Heron *et al.*, Dopamine modulates dynamic decision-making during foraging. *J. Neurosci.* **40**, 5273–5282 (2020).
41. D. Romer, V. F. Reyna, T. D. Satterthwaite, Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Dev. Cogn. Neurosci.* **27**, 19–34 (2017).
42. B. Larsen, B. Luna, Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neurosci. Biobehav. Rev.* **94**, 179–195 (2018).
43. B. J. Ellis, W. T. Boyce, J. Belsky, M. J. Bakermans-Kranenburg, M. H. van Ijzendoorn, Differential susceptibility to the environment: An evolutionary-neurodevelopmental theory. *Dev. Psychopathol.* **23**, 7–28 (2011).
44. J. L. Hanson, A. R. Hariri, D. E. Williamson, Blunted ventral striatum development in adolescence reflects emotional neglect and predicts depressive symptoms. *Biol. Psychiatry* **78**, 598–605 (2015).
45. P. Pechtel, D. A. Pizzagalli, Disrupted reinforcement learning and maladaptive behavior in women with a history of childhood sexual abuse: A high-density event-related potential study. *JAMA Psychiatry* **70**, 499–507 (2013).
46. R. Boecker *et al.*, Impact of early life adversity on reward processing in young adults: EEG-fMRI results from a prospective study over 25 years. *PLoS One* **9**, e104185 (2014).
47. C. A. Taswell, V. D. Costa, E. A. Murray, B. B. Averbeck, Ventral striatum's role in learning from gains and losses. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E12398–E12406 (2018).
48. J. L. Hanson *et al.*, Early adversity and learning: Implications for typical and atypical behavioral development. *J. Child Psychol. Psychiatry* **58**, 770–778 (2017).
49. K. A. McLaughlin *et al.*, Childhood social environment, emotional reactivity to stress, and mood and anxiety disorders across the life course. *Depress. Anxiety* **27**, 1087–1094 (2010).
50. J. K. Lenow, S. M. Constantino, N. D. Daw, E. A. Phelps, Chronic and acute stress promote overexploitation in serial decision making. *J. Neurosci.* **37**, 5681–5689 (2017).
51. V. J. Felitti *et al.*, Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. *Am. J. Prev. Med.* **14**, 245–258 (1998).
52. A.-R. Blais, E. U. Weber, A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgm. Decis. Mak.* **1**, 15 (2006).
53. A. L. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, J. K. Evershed, Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Methods.* **52**, 388–407 (2020).
54. N. Garrett, N. D. Daw, Biased belief updating and suboptimal choice in foraging decisions. *Nat. Commun.* **11**, 3417 (2020).
55. A. Lloyd, R. T. McKay, N. Furl, Adverse Childhood Experiences and patch foraging. Open Science Framework. <https://osf.io/8znyx/>. Deposited 23 December 2021.