



Patch Foraging in Adolescence and Adulthood: Examining the Role of  
Developmental, Cognitive and Social Factors in Foraging Behaviour

Alexander James Lloyd

Thesis submitted in fulfilment of the degree of Doctor of Philosophy

December 2021

Department of Psychology

Royal Holloway, University of London

Egham, Surrey, TW20 0EX

## **Abstract**

Adolescence, a period of significant physical, social, and neurobiological change, coincides with a rise in novelty-seeking behaviours. Recent theories propose that seeking novel experiences supports adolescents to explore their surroundings and gain the experiential knowledge necessary for adulthood. In Chapters 3, 4 and 5 I address the first aim of this thesis, which was to identify the computational mechanisms that contribute to the rise of novelty-seeking in adolescence. I utilised a patch foraging paradigm which measures the individual's preference for exploiting rewards that are immediately available or exploring novel patches in their environment to search for a potentially larger reward. In Chapters 3 and 4, I demonstrate that adolescents aged 16-17 explored more and integrated reward feedback faster than adults aged 21+. In Chapter 5, I demonstrate that stochasticity is a critical mechanism that facilitates novelty seeking in volatile environments for adolescents aged 16-17 and adults aged 24+. The second aim of this thesis was to test whether social influence differentially affects novelty seeking in adolescence compared to adulthood. However, in Chapter 6 I found that both adolescents' and adults' exploration behaviours are susceptible to social influence. The final aim of this thesis was to test the predictions of a theory that explains how experiences of adversity during childhood impact decision-making in adulthood. In Chapter 7, I demonstrate that adults who have been exposed to adversity explore their surroundings less and underweight reward feedback compared to individuals without these experiences, supporting the predictions of a theory explaining how adversity impacts decision-making. The findings of this thesis demonstrate that heightened exploration in adolescence is associated with positive outcomes, though these behaviours are susceptible to social influence and adverse experiences, highlighting potential avenues for intervention to protect adolescents from harmful outcomes and promote positive outcomes for this age group.

## Acknowledgements

First and foremost, I am deeply grateful to my supervisors, Nick, Ryan, Josh, and Cat. I would like to thank Nick, who I doubt expected to be taking on a new PhD student at the beginning of a global pandemic! You have gone above and beyond as a supervisor, and I can't thank you enough for all the time you have spent supporting me to produce this work. To Ryan, thank you for your impeccable attention to detail, good humour, and continuous encouragement throughout my PhD. To Josh, thank you for your patience in working through many code errors with me early on in my studies. Finally, to Cat, thank you for taking the chance on me when I emailed you enquiring about PhDs back in 2017 and for your continued support through my studies. I feel truly fortunate to have worked with such inspirational academics and supportive mentors over the last three years.

I would also like to thank the schools, colleges, and charities who I have worked with. The studies in this thesis were only made possible through the support of collaborators who gave their time freely and generously to help me complete this research. Thank you to the staff at Farnborough College; Holmesdale School; Glyn School; I'm A Scientist, Get me Out of Here; National Association for People Abused in Childhood; Survivors South West Yorkshire; The Survivor's Trust; and one anonymous support group. I would also like to thank each of the participants who gave their time to support this research.

I would like to thank my fellow PhD students, who I was fortunate to share this unique experience with. The PhD community was apparent from very early on and has been invaluable in providing the peer support that helped during coding issues, recruitment challenges and desk rejections, often ending in much needed trips to the pub. I look forward to continuing the friendships made through the PhD community as everyone takes their next steps.

I am deeply grateful to my mum and sister, whose unwavering love and support through my studies has meant everything to me. You have been my foundation throughout my education and always there to celebrate the successes and weather the challenges, without which I wouldn't be who I am today. Finally, to my partner Beatrice, who has been with me from the very first day of this PhD until the very last. Thank you for continually inspiring me, challenging me, and never failing to find the humour in any situation. Your encouragement over the past three years has kept me going and I couldn't imagine sharing this experience with anyone else.

## Table of Contents

### Contents

Abstract.....	2
Acknowledgements.....	3
Table of Contents.....	5
Bibliography of Work Completed During PhD.....	8
List of Figures.....	9
List of Tables.....	18
Glossary and Abbreviations.....	21
Chapter 1: Introduction and Literature Review.....	22
1.1. General Introduction.....	22
1.2. Adolescence: A Sensitive Period of Development.....	24
1.3. Adolescence and Risk-Taking Behaviour: A Public Health Concern.....	28
1.4. Why Do Adolescents Take Risks? Models of Adolescent Decision-Making.....	29
1.5. Adolescent Decision-Making in Conditions of Ambiguity.....	41
1.6. Exploration Across the Lifespan.....	45
1.7. Reinforcement Learning for Optimal Exploration.....	49
1.8. Human and Animal Models of Patch Foraging.....	56
1.9. Peer Influence in Adolescence.....	62
1.10. Early Life Stress and Decision-Making in Adulthood.....	68
1.11. The Current Thesis.....	71
Chapter 2: Methodology.....	75
2.1. Computational Modelling of Cognition.....	75
2.2. Modelling Assumptions and Variable Choices.....	89
2.3. Interim Conclusion.....	93
Chapter 3: Are Adolescents More Optimal Decision-Makers in Novel Environments? Examining the Benefits of Heightened Exploration in a Patch Foraging Paradigm'.....	94
3.1. Abstract.....	94
3.2. Introduction.....	95
3.3. Methods.....	100
3.4. Results.....	109
3.5. Discussion.....	113
3.6. Supplementary Materials.....	120

Chapter 4: Adolescents Utilise a Faster Learning Rate Whilst Foraging Compared to Adults	124
4.1. Abstract .....	124
4.2. Introduction.....	125
4.3. Methods.....	130
4.4. Results.....	135
4.5. Discussion .....	141
4.6. Supplementary Materials .....	147
Chapter 5: Optimal Exploration in Stable and Volatile Foraging Environments Depends on Response Stochasticity and Anxiety .....	151
5.1. Abstract .....	151
5.2. Introduction.....	152
5.3. Method .....	155
5.4. Results.....	163
5.5. Discussion .....	173
5.6. Supplementary Material.....	180
Chapter 6: Adolescents' and Adults' Exploration Behaviour is Susceptible to Social Influence While Foraging .....	193
6.1. Abstract .....	193
6.2. Introduction.....	194
6.3. Methods.....	198
6.4. Results.....	210
6.5. Discussion .....	218
6.6. Supplementary Material.....	226
Chapter 7: Individuals with Adverse Childhood Experiences Explore Less and Underweight Reward Feedback: .....	250
7.1. Abstract .....	250
7.2. Introduction.....	252
7.3. Results.....	258
7.4. Discussion .....	263
7.5. Methods.....	268
7.6. Supplementary Materials .....	277
Chapter 8: General Discussion.....	292
8.1. Summary of Findings.....	292
8.2. Adolescents Explore Their Environment More Than Adults .....	294
8.3. Social Influence in Adolescence .....	304

8.4. Adversity, Exploration and Reward Processing .....	309
8.5. Limitations, Implications and Future Directions .....	313
8.6. Conclusion .....	316
References.....	318

## Bibliography of Work Completed During PhD

- Lloyd, A., McKay, R., & Furl, N. (2022). Individuals with Adverse Childhood Experiences Explore Less and Underweight Reward Feedback. *Proceedings of the National Academy of Sciences* 119 (4), e2109373119. doi: 10.1073/pnas.2109373119
- Lloyd, A., McKay, R., Hartman, T., Vincent, B., Murphy, J., Gibson-Miller, J., ... Mason, L. (2021). Delay discounting and under-valuing of recent information predict lower adherence to social distancing measures during the COVID-19 pandemic. *Scientific Reports*, 11. doi: 10.1038/s41598-021-98772-5 [Not in thesis]
- Bentall, R., Lloyd, A., Bennett, K., McKay, R., Mason, L., Murphy, J., McBride, O., ... & Shevlin, M. (2021). Pandemic buying: Testing a psychological model of over-purchasing and panic buying using data from the United Kingdom and the Republic of Ireland during the early phase of the COVID-19 pandemic. *PLOS ONE*, 16(1), e0246339. doi: 10.31234/osf.io/u7vqp [Not in thesis]
- Lloyd, A., McKay, R., Sebastian, C.L., & Balsters, J.H. (2021). Are adolescents more optimal decision-makers in novel environments? Examining the benefits of heightened exploration in a patch foraging paradigm. *Developmental Science*, 24, e13075. doi: 10.31234/osf.io/ypknq
- Lloyd, A., & Borrill, J. (2020). Examining the effectiveness of Restorative Justice in reducing victims' post-traumatic stress. *Psychological Injury and Law*, 13, 77–89. doi: 10.1007/s12207-019-09363-9 [Not in thesis]
- Lloyd, A., & Döring, A. (2019). When do peers influence adolescent males' risk taking? Examining decision-making under conditions of risk and ambiguity. *Journal of Behavioral Decision Making*, 32(5), 613-626 doi: 10.1002/bdm.2141J [Not in thesis]

## List of Figures

Chapter	Figure Description	Page
<i>Chapter 1</i>		
	Figure 1.1: Simulation of a Rescorla-Wagner RL agent estimating the reward schedules of two probabilistic bandits. Bandit 1 (left) has a 75% chance of yielding a reward and bandit 2 (right) has a 25% chance of yielding reward, which are indicated by the horizontal dashed lines. Blue lines indicate a decision-making policy that only uses the most recent occurrences (i.e., a high learning rate), whereas red lines indicate a decision-making policy that integrates information over a longer series of trials (i.e., a low learning rate).	53
<i>Chapter 2</i>		
	Figure 2.1: A curve fitting demonstrating models that have been underfitted (left) and overfitted (right) to the dataset. Each point on the plot indicates a single datapoint and lines indicate how well the model fits to the data. The middle panel is an example of a model that includes an appropriate number of parameters to explain the dataset. Note, this figure does not reflect models used in the current thesis but was simulated for illustrative purposes.	79
	Figure 2.2: Plots demonstrating the simulated agent's estimate of the environmental reward rate (top two panels) and how this affects the agent's foraging behaviour (bottom two panels). The panels on the left demonstrate a forager who utilises a high learning rate and therefore updates their estimate of the average reward rate of the environment quickly in response to recent trials. In contrast, the panels on the right demonstrate a forager who updates their estimate of the average reward rate too slowly to adjust their behaviour to changes in the environment.	83
	Figure 2.3: Simulated plot of the minimisation algorithm implemented by the optim function. The x-axis presents the	87

algorithm's estimate of the learning rate parameter. The grey vertical lines a1-3 indicate different starting values provided to the algorithm. Point B is the global minima and the best estimate of the learning rate parameter in this example. Points A and C are local minima that the algorithm has identified.

### Chapter 3

Figure 3.1: Schematic outline of the patch foraging paradigm. 102

Figure 3.2: Plot of the mean leaving threshold for adolescents and adults for Environment 1 (fast depletion and short travel), Environment 2 (slow depletion and short travel), Environment 3 (fast depletion and long travel) and Environment 4 (slow depletion and long travel). Green bars indicate the optimal leaving value for that environment. 110

Supplementary Figure 3.1: Plots demonstrating the leaving threshold that led the simulated agent to accumulate the greatest number of rewards in each environment. Plot A corresponds to Environment 1 which had a fast depletion rate and short travel time. The optimal leaving threshold in this environment was 6.46. Plot B corresponds to Environment 2 which had a slow depletion rate and short travel time. The optimal leaving threshold for this environment was 7.04. Plot C corresponds to Environment 3 which had a fast depletion rate and long travel time. The optimal leaving threshold for this environment was 5.07. Plot D corresponds to Environment 3 which had a slow depletion rate and long travel time. The optimal leaving threshold for this environment was 5.90. 120

Supplementary Figure 3.2: Mean leaving threshold across the four environments for adult and adolescent participants. 121

### Chapter 4

Figure 4.1: Figure demonstrating the difference between absolute values of the AIC and BIC for the dual learning rate 136

model and the single learning rate model. The left plot demonstrates the difference between values of the AIC and the right plot demonstrates the difference between values of the BIC. Lower values on the y-axis indicate the single learning rate model is a better fit to the data.

Figure 4.2: Simulation results demonstrating foraging behaviour for simulated adolescents and simulated adults. The y-axis denotes the average number of rewards experienced on that trial and the x-axis is the trial number, which corresponds to each explore/exploit decision. The red dashed line is the leaving threshold adolescents utilised in each environment whereas the dashed turquoise line is the leaving threshold adults used in each environment. 139-140

## Chapter 5

Figure 5.1: Violin plot demonstrating participants' learning rate in each environment. The middle bar of each boxplot indicates the mean for each age group and the upper and lower bars indicate the upper and lower quartiles, respectively. Green bars indicate the optimal learning rate. 164

Figure 5.2: Violin plot demonstrating that adolescents had a significantly higher leaving threshold compared to adults. Boxplots indicate the mean leaving threshold along with upper and lower quartiles. Horizontal green bar indicates the optimal leaving threshold in the stable environment. 166

Figure 5.3: Raincloud plot of estimates for the  $\beta$  parameter, which measures stochasticity in participants' decision-making, with lower values on the y-axis denoting more stochastic choices and higher values denote more deterministic choices. Horizontal green lines indicate the optimal value of beta for each environment. 168

Figure 5.4: Plot demonstrating the association between anxiety (x-axis) and the degree to which participants adjusted their decision stochasticity (y-axis) between the two environments (calculated through subtracting stochasticity in the stable environment from stochasticity in the volatile environment). Values above 0 on the y-axis indicate participants had lower values on  $\beta$  in the volatile environment relative to the stable environment whereas values below 0 indicate participants who had higher values on  $\beta$  in the stable environment. The horizontal green line indicates the optimal amount of adjustment in stochasticity between stable and volatile environments, determined by our simulations.

Figure 5.5: Plot demonstrating the association between anxiety and summed rewards collected in both environments.

Figure 5.6: Violin plot demonstrating the total number of rewards accrued in each environment. The boxplot demonstrates the upper and lower quartile.

Supplementary Figure 5.1. Plots demonstrating the total number of points available to collect in the stable environment (left) and volatile environment (right). Each environment affords the participant the opportunity to collect a maximum of approximately 800 rewards (i.e., apples).

Supplementary Figure 5.2: Results of model comparison for each participant. Values below zero indicate the model with a single learning rate was a better fit to the data than the model with separate learning rates for positive and negative feedback. The y-axis denotes the number of participants who are represented by each value on the x-axis.

Supplementary Figure 5.3: Plots of the optimal values on alpha (top) and beta (bottom) parameters in the stable environment (red lines) and volatile environment (blue lines). Vertical lines indicate the value of the parameter that yielded the greatest

number of points (in their respective colours). Gray shaded areas indicate the confidence interval around the number of rewards collected for each parameter value.

Supplementary Figure 5.4: Participants' experience of rewards 187-188  
in each environment, averaged separately for each age-group. The y-axis is the number of rewards (apples) experienced on each trial. Local peaks indicate when participants travelled to a new patch and subsequently experienced a large number of rewards. The top plot is the average reward experienced in the stable environment, the middle plot is the average reward experienced in the volatile environment where the rich portion was presented first and the bottom plot is the average reward experienced in the volatile environment where the poor portion was presented first. In the volatile environments, dark grey blocks indicate the rich portions of the condition and light grey blocks indicate the poor portions of the environment.

Supplementary Figure 5.5: Simulation of participants' behaviour 189  
using mean parameter estimates for adolescents (red) and adults (blue). The y-axis indicates the average number of rewards collected on each trial (x-axis).

Supplementary Figure 5.6: Plot demonstrating the fixed 190  
parameter values used to simulate the synthetic data presented in Supplementary Figure 4 (circular points). Vertical lines indicate  $\pm 1$  SD from the participants' mean parameter estimate. Diamond points indicate the mean parameters estimated by the model when run over these simulated data.

## Chapter 6

Figure 6.1: Schematic of the Social Patch Foraging task. When 201  
participants choose to stay (top panel), they are presented with the number of apples they accrued along with their total score. When participants choose to leave (bottom panel), they are presented with a travel screen which included an image of the

peer and text indicating whether they recommended staying for more, less, or the same amount of time on the next tree (patch).

Figure 6.2: Violin plots demonstrating participants' leaving thresholds in the asocial condition of the task. Middle bars indicate the mean leaving threshold for that age group with upper and lower bars reflecting the upper and lower quartiles, respectively. Horizontal green bars indicate the optimal leaving threshold for that environment. 211

Figure 6.3: Plot demonstrating participants' leaving thresholds in each of the social influence conditions for both foraging environments. Error bars indicate standard deviations for each condition. The green dashed line indicates the optimal leaving threshold averaged across both foraging environments. \*  $p < .050$ , \*\*  $p < .010$ , \*\*\* $p < .001$  214

Figure 6.4: Estimates of the social influence parameter ( $\psi$ ) for each age group. Error bars indicate standard deviations in each age group. Larger values indicate that participants are more susceptible to social influence. 217

Supplementary Figure 6.1: Bar plot demonstrating the leaving threshold (y-axis) for each of the social influence conditions (x-axis). Higher values denote more exploration. Red bars are the rich quality environment and turquoise bars indicate the poor-quality environment. 232

Supplementary Figure 6.2: Plot demonstrating the mean and standard deviation of participants' leaving thresholds (y-axis) for each age group across the social influence conditions (x-axis), with individual datapoints in gray. The left plot is for the rich environment and the right plot is for the poor environment. The dashed green line indicates the optimal leaving threshold for that environment. 233

Supplementary Figure 6.3: Bar plot demonstrating the mean and standard error of parameter estimates for  $\psi$  (y-axis) across the three social influence conditions (x-axis). Error bars indicate standard deviations in each condition. Higher values on the y-axis indicate the participant was more susceptible to social influence. The asocial condition did not have a parameter estimate for susceptibility to social feedback and was omitted from this analysis.

234

Supplementary Figure 6.4: Plot demonstrating the interaction between social influence condition and environment in the younger adolescent group.

239

Supplementary Figure 6.5: Plots of the social model (left) and asocial model (right) using means of participants' parameter estimates for each social influence condition. Horizontal dashed lines indicate the average leaving threshold observed in our behavioural data for each condition.

248

## Chapter 7

Figure 7.1: 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 quartile. Horizontal green bars indicate the optimal leaving thresholds for each environment. Higher leaving threshold indicates more exploration and less exploitation.

259

Figure 7.2: 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.

260

Figure 7.3: 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 quartile, respectively.	263
Figure 7.4: 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.	271
Supplementary Figure 7.1: Order with which participants completed the foraging task. The screens detail (in order): the initial instructions, the practice environment, final instructions, first environment (note, the order with which the colours were presented was counterbalanced), interim screen signalling the change in environment, the second environment and finally a break screen notifying participants about the upcoming ACE questionnaire.	279
Supplementary Figure 7.2: Heatmap demonstrating the correlations between simulated (i.e., variables with the 'hardcode' prefix) and recovered parameters. Figures in each grid square indicate the Pearson's $r$ correlation coefficient.	281
Supplementary Figure 7.3: Plots demonstrating the relative fit of the single learning rate model compared to the dual learning rate model. Values below zero indicate that the single learning rate model was a better fit to participants' data whereas values above zero indicate the dual learning rate model was a better fit to the data.	283

Supplementary Figure 7.4: Plots demonstrating the average reward experienced on each trial for the simulated foraging agents in the rich environment (left) and poor environment (right). Local troughs indicate the simulated agents' leaving thresholds. 284

Supplementary Figure 7.5: Plot demonstrating the average patch residency time (y-axis) in the two foraging environments (x-axis). These results are comparable to our analyses of participants' leaving thresholds detailed in Figure 1 in the main manuscript. 287

## List of Tables

<b>Chapter</b>	<b>Table Description</b>	<b>Page</b>
<i>Chapter 3</i>		
	Table 1.1: Parameters and optimal leaving thresholds of the four environments in the patch foraging task.	105
	Table 2: Mean (SD) of participants' deviation from the optimum and number of apples accumulated in each environment.	112
<i>Chapter 4</i>		
	Table 4.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents' highest level of education.	130-131
	Table 4.2: Features of the four foraging environments from Lloyd et al. (2021b).	132
	Table 4.3: Mean (SD) estimates for the free parameters in the single learning rate model.	137
	Table 4.4: Average of the model fit indices used to assess which of the models provided a better explanation of participants' foraging behaviour.	138
<i>Chapter 5</i>		
	Table 5.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents' highest level of education.	156-157
	Supplementary Table 5.1: Mean (SD) of participants' leaving thresholds in the rich portion of the volatile condition.	192
<i>Chapter 6</i>		

Table 6.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents' highest level of education.	198-199
Supplementary Table 6.1: Mean (SD) of participants' average patch residency time in seconds.	228
Supplementary Table 6.2: Mean (SD) of the degree to which participants deviated from their peer's advice in each environment. Negative values indicate the participant explored more than the peer advised whereas positive values indicate the participant exploited patches for longer than the peer advised.	230
Supplementary Table 6.3: Mean (SD) estimates of the $\psi$ parameter for younger adolescents, older adolescents and adults in each of the social influence conditions (excluding the asocial condition).	235-236
Supplementary Table 6.4: Descriptive statistics demonstrating the degree to which younger adolescents deviated from the optimal leaving threshold in each foraging environment.	238
Supplementary Table 6.5: Descriptive statistics demonstrating the degree to which older adolescents deviated from the optimal leaving threshold in each foraging environment.	239-240
Supplementary Table 6.6: Descriptive statistics demonstrating the degree to which adults deviated from the optimal leaving threshold in each foraging environment.	241
Supplementary Table 6.7: Mean (SD) parameter estimates for the free parameters $\alpha$ , $\beta$ and $c$ in the winning social learning model.	246

Supplementary Table 6.8: Correlations between free parameter estimates for the social learning model. 247

Supplementary Table 6.9: Model fit indices for the model recovery analyses. 249

## *Chapter 7*

Table 7.1: Prevalence of each Adverse Childhood Experience in the current sample. 269-270

Supplementary Table 7.1: Average of the indices of model fits from participants' data for the single and dual learning rate models, in each environment. 282-283

Supplementary Table 7.2: Table demonstrating regression model statistics for the subscales of the DOSPERT. Bold cells indicate results that were statistically significant at the .05 level. 285

## Glossary and Abbreviations

ACC	Anterior cingulate cortex
ACE	Adverse childhood experience
ADHD	Attention deficit hyperactivity disorder
AIC	Akaike information criterion
BART	Balloon analogue risk task
BAS	Behavioural approach system
BIC	Bayesian information criterion
BIS	Behavioural inhibition system
dIPFC	Dorsolateral prefrontal cortex
DOSPRT	Domain-specific risk-taking
fMRI	Functional magnetic resonance imaging
GAD	Generalised anxiety disorder
MVT	Marginal value theorem
NAPAC	National Association for People Abused in Childhood
OLS	Ordinary least squares
OSF	Open Science Framework
PFC	Prefrontal cortex
RL	Reinforcement learning
RT	Reaction time
vIPFC	Ventrolateral prefrontal cortex
VS	Ventral striatum

# Chapter 1: Introduction and Literature Review

## 1.1. General Introduction

Adolescence is a developmental period during which the individual experiences significant physical, psychological, and neurobiological changes (Spear, 2000). Coinciding with these changes, adolescents exhibit increased novelty-seeking behaviour relative to adults and children (Shulman et al., 2016). Several theories have been developed to understand these behavioural differences, the situational factors that can influence adolescents' novelty-seeking, and how experiences during adolescence can influence decision-making in adulthood. The present thesis aims to test predictions of theories of development to contribute to our understanding of the role of adolescence within the lifespan. To achieve this, the current thesis is organised around three key objectives:

The first objective of this thesis *is to test the predictions of theories that account for the rise of novelty-seeking at the onset of adolescence*, which I examine in Chapters 3, 4 and 5. These theories propose that adolescents' goal-directed behaviour is different from adults' due to the development of neural regions associated with processing reward. However, these theories diverge in their accounts of whether novelty-seeking is primarily expressed as risk taking or exploration. Theories that focus on the link between novelty-seeking and risk-taking highlight the potential for adolescents' behaviour to lead to negative outcomes, with little utility for future decision-making. In contrast, theories that link novelty-seeking to exploration highlight the role for these behaviours to contribute to adolescents' experiential knowledge, which can be used to guide future decision-making and may therefore have utility within the lifespan. As such, I utilise behavioural paradigms to examine differences between adolescents'

and adults' reward decision-making, which contributes to understanding whether novelty-seeking during adolescence serves a functional role within the lifespan.

The second objective of this thesis is *to test the predictions of theories that explain how social influence affects adolescents' decision-making*, which I examine in Chapter 6. Specifically, I examine theories that explain how social influence affects adolescents' novelty-seeking behaviour. Adolescents' susceptibility to social influence has typically been associated with negative outcomes, such as reckless behaviour (Smith et al, 2014), yet recent research has also highlighted how social influence can promote adolescents' prosocial behaviours (Foulkes et al., 2018). Through testing theories of social influence in adolescence, these findings will contribute to our understanding of how adolescents' behaviour can be affected by situational factors and when such factors can increase the risk for negative outcomes at this point in the lifespan.

The third and final objective of this thesis is *to test the predictions of a theory that explains how experiences of adversity during childhood and adolescence can impact decision-making in adulthood*, which I examined in Chapter 7. Experiences of adversity during childhood and adolescence are associated with negative psychosocial outcomes in adulthood (Danese & McEwen, 2012). This thesis will contribute to understanding the cognitive mechanisms that may link early adversity to negative outcomes in adulthood and in doing so will test the predictions of a theory that purports to explain this link. Through understanding the features of decision-making that are affected by early adversity, future research can seek to develop interventions to support individuals with adverse childhood experiences.

In addressing these three objectives, the current thesis aims to contribute to our understanding of adolescence as a key period of development within the lifespan. These findings will highlight the conditions under which adolescents' decision-making differs from

that of adults, which can contribute to our understanding of the role of novelty-seeking behaviours in the lifespan. Moreover, these findings can provide insight into the conditions under which adolescents are at risk of engaging in harmful or hazardous behaviours.

This literature review will provide an overview of existing research related to these three objectives. First, an overview of adolescence as a developmental period will be presented (Section 1.2.-1.3.), followed by an evaluation of key theories of adolescent decision-making (Section 1.4.). Next, a review of the experimental literature that has examined the conditions under which adolescents exhibit heightened novelty-seeking behaviour relative to other age groups will be presented (Section 1.5.-1.8.). This content will form the background to the first objective of this thesis. Following this, I will evaluate theories of why adolescents exhibit heightened susceptibility to social influence (Section 1.9.), providing background to the second objective of the current thesis. I will then present a review of evidence regarding how experiences of adversity in childhood and adolescence affect decision-making in adulthood (Section 1.10.), which will form the background for the third objective of this thesis. Based on the literature reviewed, key research questions will be identified that will be addressed in this thesis (Section 1.11.).

## **1.2. Adolescence: A Sensitive Period of Development**

Adolescence is a period of significant development within the lifespan, during which individuals experience physical, social, and cognitive changes (Ellis et al., 2012; Spear, 2000). These changes are complemented by extensive neurobiological development, which involves substantial reorganisation of regions implicated in social and cognitive function (Somerville et al., 2019). During this period, adolescents exhibit marked changes in their behaviour relative to adults and children (Steinberg, 2008), which have been associated with heightened rates of exploration and risk-taking (Romer et al., 2017). The combination of these social, cognitive,

and neurobiological changes supports the individual in developing the skills necessary for adulthood (Ellis et al., 2012; Spear, 2000).

The psychological changes that occur during adolescence are associated with hormonal changes that mark the onset of puberty (Hall, 1904). During puberty, gonadal hormones are released, triggering a series of physiological changes. In male adolescents, this is driven by the hormone testosterone, which drives muscle growth, a deepening of the voice and the development of body hair (Mehta et al., 2015). In female adolescents, oestrogen is released, causing the onset of menarche, physical growth, and the development of body hair (Alonso & Rosenfield, 2002). These physiological changes can be affected by one's childhood environment, and the onset of puberty is now earlier than it was 150 years ago, which is considered to reflect improved child rearing conditions (Gamble, 2017). The onset of puberty also varies within and between countries, with puberty beginning earlier in industrialised countries (Sawyer et al., 2018) and in more urban areas within nations (Said-Mohamed et al., 2018). In recognition of the advanced onset of puberty relative to previous generations, most researchers consider adolescence to include individuals aged ten and above (Sawyer et al., 2018).

While the onset of adolescence is marked by physical and hormonal changes, the end of adolescence is less rigidly defined. Early definitions of adolescence identified the end of this period as the age of 18, when the individual was expected to have reached autonomy in society, having secured financial independence and separation from caregivers (Arnett & Taber, 1994; Hall 1904). However, recent evidence has indicated that the neurobiological developments that begin in the second decade of life do not end until the early to mid-20s (Larsen & Luna, 2018). Culturally, milestones that were considered to mark the beginning of adulthood, such as marriage and parenthood, have become less common in the second decade of life and are now more prevalent in the third and fourth decades of life (Patton et al., 2018). In recognition of the

timeline with which these cognitive and cultural milestones occur, recent definitions of adolescence have expanded the upper bound of this stage of the lifespan to the age of 24 (Sawyer et al., 2018). A robust body of interdisciplinary evidence supports the existence of a unique developmental period in the second decade of life which is marked by the interplay of neural, cognitive, behavioural, and social changes (Aoki et al., 2017; Steinberg, 2005).

Neuroimaging research has provided insight into the functional and structural changes that occur to the brain during adolescence. From childhood through to adulthood, the volume of white matter increases linearly (Bethlehem et al., 2021; Coupe et al., 2017; Giedd et al., 1999; Tamnes & Mills, 2020). Structurally, there are increases in fractional anisotropy within white matter (i.e., increasingly unequal diffusion of water molecules; Barnea-Goraly et al., 2005), suggesting increased organisation and specialisation which facilitates more efficient communication between neural regions (Blakemore et al., 2010; Ladouceur et al., 2012). In contrast, changes to gray matter are region-specific and are typically non-linear (Mills et al., 2016; Tamnes et al 2010; Shaw et al., 2008). Many regions demonstrate an inverted-U pattern of development, increasing in gray matter volume in childhood and then reducing in volume from the beginning of adolescence through to adulthood (Bethlehem et al., 2021; Lenroot & Giedd, 2006; Mills et al., 2016; Tamnes & Mills, 2020). For example, gray matter volume in the frontal lobe is highest at ages 11-12 and declines thereafter (Giedd et al., 1999). This pattern might reflect a process of dendritic outgrowth followed by synaptic pruning, which aids in the specialisation of these regions (Blakemore et al., 2010).

Due to the large-scale reorganisation of the brain at this point in the lifespan, adolescence has been identified as a sensitive period of development (Fuhrmann et al., 2015). Sensitive periods are windows during development when the individual requires input from relevant stimuli to facilitate the acquisition of abilities (Knudsen, 2004). For example, childhood contains sensitive periods for language acquisition, and deprivation of exposure to auditory

speech can lead to long-term difficulties in language production (Neville et al., 1992). In adolescence, there is rapid reorganisation of regions responsible for cognitive abilities (e.g., cognitive control; Casey et al., 2008) and the processing of social and emotional information (Blakemore & Mills, 2014; Dumontheil, 2016). It has therefore been suggested that a lack of exposure to scenarios that aid the development of social, emotional, and cognitive skills during adolescence (e.g., social isolation) can cause detrimental outcomes, which persist into adulthood (Orben et al., 2020; Tomova et al., 2021). For example, a longitudinal study found that self-reported social isolation in adolescence predicted mental health problems associated with emotion processing, such as depression, in adulthood (Goosby et al., 2013). Though, it should be noted that similar associations between self-reported isolation and depression have also found in adult populations (e.g., Robb et al., 2020), which may refute the suggestion that adolescence is a uniquely sensitive period for social development. Nevertheless, there is evidence that chronic levels of isolation from peers during adolescence is associated with structural changes to regions of the brain responsible for social cognition, such as the dorsomedial prefrontal cortex (Martin et al., 2017; Nakagawa et al., 2015). This evidence suggests that deprivation of social stimuli during sensitive periods in adolescence can increase the short and long-term risk for negative outcomes, such as mental health problems.

Converging psychological, neuroscientific, and biological evidence has identified adolescence as a distinct period of development within the lifespan (Jaworska & MacQueen 2015; Sawyer et al., 2018). Adolescence contains sensitive periods of development, with neural regions related to cognitive abilities (e.g., the dorsal medial prefrontal cortex and temporoparietal junction) and social processing (e.g., the anterior temporal cortex) demonstrating structural and functional reorganisation that reflect the maturation of the brain (Blakemore & Mills, 2014; Casey et al., 2008; Giedd et al., 1999; Mills et al., 2016). Therefore,

it is important to ensure adolescents are exposed to experiences that encourage social and cognitive development to facilitate the transition from childhood to adulthood.

### **1.3. Adolescence and Risk-Taking Behaviour: A Public Health Concern**

Along with being a period of physical and neurobiological development, adolescence is also identified as a period of heightened risk-taking behaviour (Steinberg, 2008). In early psychological studies involving adolescents, it was noted that these individuals were more likely to be involved in crime and other activities outside of adult social norms (Hall, 1904). Indeed, the conception of adolescence as a period of ‘storm and stress’ has continued in the psychological sciences into the 21<sup>st</sup> Century (Casey et al., 2010) and is reflected in adults’ attitudes towards this age group (Protzko & Schooler, 2019).

These attitudes towards adolescents may not be unwarranted; national and international reports indicate that adolescents are overrepresented in statistics related to hazardous behaviours. For example, adolescents are more likely than adults and children to be involved in substance misuse (Willoughby et al., 2014), dangerous driving (Swedler et al., 2012) and risky sexual practices (Potard et al., 2008). In addition, rates of delinquency and antisocial behaviours increase dramatically in the adolescent years, but steadily decline as the individual reaches adulthood (Moffitt, 1993; 2018). These behaviours put the individual at risk of negative outcomes and in the most extreme circumstances can lead to serious injury or mortality (Eaton et al., 2012).

Epidemiological statistics are important as they indicate that risky behaviours are often heightened in adolescence but do not persist into adulthood (Moffitt, 1993; 2018). It has been suggested that the psychological and biological developments that occur at this point in the lifespan predispose adolescents to engage in behaviours that heighten the risk for harmful

outcomes (Steinberg, 2008). Given the potential for severely negative outcomes to the individual and society, adolescents' risk-taking behaviour has been identified as a public health concern (World Health Organisation, 2018), which has motivated research into the aetiology of these behaviours.

#### **1.4. Why Do Adolescents Take Risks? Models of Adolescent Decision-Making**

To understand why adolescents engage in higher levels of risk taking behaviour relative to other age groups, a number of verbal theories have been proposed that describe the processes underlying adolescents' decision-making. Verbal theories provide descriptive or explanatory frameworks for complex evidence and can be used to generate testable predictions (Smaldino, 2020; van Rooij & Baggio, 2020). The first objective of this thesis is to test the predictions of these theories of adolescent decision-making, which offer explanations about why this age group exhibits heightened novelty-seeking and risk-taking behaviours compared to adults. Novelty-seeking describes the individual's propensity to seek new experiences or sensations that contain some degree of uncertainty, and the pursuit of novelty can lead the individual to uncover information about the surroundings (Wahlstrom et al., 2010). Risk-taking can be considered a subset of novelty-seeking behaviours, wherein the novel experience has an opportunity for reward but also contains the potential for a negative outcome (Steinberg, 2007). Indeed, empirical studies have demonstrated that risk taking and novelty-seeking are distinct phenomena (Mallet & Vignoli, 2007). The following section will review and critically evaluate these models drawing on evidence from psychology and neuroscience, which will inform my hypotheses about how decision-making changes across the lifespan.

#### *1.4.1 Imbalance Models*

The earliest neurobiological models of adolescent decision-making are the Dual Systems Model (Steinberg, 2005) and Maturational Imbalance Model (Casey et al., 2008). These models draw on evidence that there is asymmetric development between areas of the brain responsible for reward processing and areas responsible for cognitive control during adolescence (Steinberg, 2008; 2010). Termed ‘imbalance models’ (Casey et al., 2011), proponents of these models argue that regions of the brain responsible for reward processing develop dramatically at the onset of puberty (Galvan et al., 2006). By contrast, the development of areas responsible for cognitive control<sup>1</sup> are more protracted, and these regions do not reach full maturity until approximately the age of 24 (Ordaz et al., 2013; Somerville et al., 2010). Consequently, in scenarios with the potential for both reward and harm, adolescents are predisposed to seek rewards whilst lacking the self-regulation to avoid negative outcomes.

Rewards play a significant role in guiding adolescents’ decision-making about whether to engage with risky scenarios (Van Leijenhorst et al., 2009). In the context of risk taking, reward sensitivity refers to the degree to which the individual is motivated to pursue reward-laden stimuli, which can lead them to select the risky option and ignore the potential for negative outcomes that can result from selecting a risky option (Kim et al., 2015). According to imbalance models, adolescents’ reward processing is hyper-active compared to adults’, such that neural responses in adolescents’ ventral striatum (VS) are more sensitive to rewards relative to losses that could be incurred following a decision (Cohen et al., 2010; Fareri et al., 2008). Due to their heightened reward sensitivity, rewards present a greater incentive for adolescents compared to adults, increasing risk-seeking behaviour (Van Leijenhorst et al., 2009). Consistent with this account, several studies found that the VS exhibits greater

---

<sup>1</sup> Note, cognitive control is also referred to as ‘executive function’ or ‘inhibitory control’ by some authors (see McKenna et al., 2017). For ease of comparison across different models, cognitive control is used to describe these capabilities throughout this thesis.

activation during adolescence in the anticipation and receipt of reward compared to adulthood, with VS activity peaking at age 15 (Casey et al., 2008; Del Giacco et al., 2021; Galvan et al., 2006; Hoogendam et al., 2013; Forbes et al., 2010; Schreuders et al., 2018; Qu et al., 2015). Further, a meta-analysis of 26 studies that measured reward-based decision-making found that adolescents demonstrate heightened activity of limbic, frontolimbic and striatal regions in the anticipation and receipt of rewards relative to adults (Silverman et al., 2015). As such, it is widely recognised that adolescents exhibit hyperactive responses to reward, consistent with the predictions of imbalance models (Blakemore & Robbins, 2012; Somerville et al., 2010; Spear, 2013; Telzer, 2016).

Studies examining the development of the dopaminergic reward system, which contains the VS, have found that the hyperactivity of this system during adolescence predicts increased risk taking at this point in the lifespan (van Duijvenvoorde et al., 2016; Del Giacco et al., 2021). For example, heightened neural activation in response to reward is positively associated with the propensity to select risky options in an adolescent sample (van Duijvenvoorde et al., 2014). In the same study, the authors found that self-reported reward sensitivity was related to VS activation over a 2-year period. Further lab-based research has demonstrated that risk taking behaviour exhibits a quadratic pattern of development, peaking at approximately age 15 and declining into adulthood (Braams et al., 2015; Qu et al., 2015), corroborating trends observed in epidemiological statistics (e.g., Moffitt, 2018). Behavioural measures of risk taking were associated with activity in the nucleus accumbens (Braams et al., 2015), a region that forms part of the dopaminergic reward system (Delago, 2007). Together, these findings are consistent with the predictions of imbalance models that adolescents' heightened risk-taking behaviour, relative to other age groups, is associated with the rapid development of the dopaminergic reward system at the onset of puberty.

According to imbalance models, the influence of the hyperactive reward system is exacerbated by the relatively slow maturation of regions responsible for inhibiting or regulating behaviour (Casey et al., 2008). These ‘cognitive control’ capacities are associated with the ventrolateral prefrontal cortex (vlPFC) and dorsolateral prefrontal cortex (dlPFC), which are responsible for complex cognition (Hooper et al., 2004). The prefrontal cortex (PFC) undergoes a protracted structural development, with gray matter volume across the PFC peaking at age 11 for females and 12 for males, and declining thereafter (Blakemore & Choudhury, 2006; Blakemore & Robbins, 2012; Shaw et al., 2008). The relative immaturity of the dlPFC in adolescence has been associated with diminished working memory relative to adulthood, which reduces adolescents’ ability to keep a specific objective in mind during goal-directed decision-making (Luciana et al., 2005). Moreover, structural immaturities in adolescents’ vlPFC, relative to adults’, have been associated with diminished response inhibition during the go-no-go (Somerville et al., 2011; Tamm et al., 2002) and Stroop tasks (Marsh et al., 2006). It has been suggested that adolescents’ poorer performance on these tasks, relative to adults, is due to their diminished ability to suppress irrelevant information to make task-appropriate responses (Luna et al., 2010). Therefore, proponents of imbalance models argue that adolescents’ reduced working memory and response inhibition, combined with heightened reward sensitivity, means that adolescents are more likely to pursue rewards that can be gained through risk taking rather than employing cognitive control to inhibit behaviour and avoid negative outcomes that are associated with the risky option (Albert & Steinberg, 2011). As such, compared to adults, adolescents are more likely to engage in risk-seeking behaviour (Luna et al., 2010).

Support for imbalance models has also come from evidence that adolescents exhibit a preference for immediate rather than delayed rewards. This preference has been demonstrated in delay discounting paradigms, which ask the participants to select between an immediate,

smaller reward, or a delayed, but larger reward (e.g., £5 now or £20 in 1-week; Odum, 2011). The preference for immediate rewards (i.e., steeper delay discounting), has been associated with lower levels of cognitive control (Shamosh et al., 2008). Adolescents exhibit steeper delay discounting relative to adults, indicating this age group has immature cognitive control abilities in the presence of reward (de Water et al., 2014; Olson et al., 2007; Steinberg et al., 2009; van den Bos et al., 2015).

Strong empirical evidence that adolescent risk-seeking behaviour is driven by the imbalance between neural regions responsible for processing reward and those responsible for cognitive control has been found in animal (Marshall et al., 2020) and human studies (Crone, 2009; Duell et al., 2016; Duell & Steinberg, 2020; Ellingson et al., 2019; Shulman et al., 2015; Shulman et al., 2016; Steinberg et al., 2008; van Duijvenvoorde et al., 2016; van Leijenhorst et al., 2010a). One notable study using a multinational sample of 5,404 adolescents found distinct developmental trajectories for reward and cognitive control systems using a range of behavioural and self-report measures (Steinberg et al., 2018). This study demonstrated that sensitivity to reward during decision-making increased rapidly at the onset of puberty, whereas the development of cognitive control was more protracted, with the disparity between these systems being greatest at ages 15-16. Together, these studies provide robust evidence for the predictions of imbalance models including a quadratic pattern of risk-taking behaviour that peaks at ages 15-16 and declines into adulthood.

However, imbalance models have been criticised as a simplistic account of adolescent decision-making. For these models to have explanatory validity at the neural level, one might expect that developmental trajectories observed at the behavioural level (e.g., Steinberg et al., 2018) should be reflected in the trajectories of the developing brain. However, not all studies have found that the disparity between developmental trajectories of gray matter in brain regions responsible for reward processing and cognitive control is a predictor of self-reported risk

taking (e.g., Mills et al., 2014). Additionally, even at a behavioural level, Pfeifer and Allen (2012) argue that reducing the broad range of influences on adolescents' behaviour to the disparity between reward sensitivity and cognitive control fails to capture the complex factors that can predispose adolescents to engage in risk-taking. For example, research has highlighted the contributions of genetics (Anokhin et al., 2009), situational factors such as the home environment (Holmes et al., 2019) and psychopathology (Fairchild et al., 2009) in adolescents' risk taking behaviour, suggesting there are a broader range of factors and individual differences that contribute to the extent to which risk taking increases during adolescence (Foulkes & Blakemore, 2018).

Further, a number of statistics about real-world risk-taking behaviour have indicated that these behaviours peak towards late adolescence or early adulthood (i.e., age 18-19; Steinberg, 2013; Willoughby et al., 2021). Such findings are contrary to predictions of the imbalance models, which suggest risk taking should peak in middle-adolescence (i.e., ages 15-16) when the disparity between regions responsible for processing reward and cognitive control are at their greatest (Li, 2017). These findings have been attributed to the increased independence at these ages, which afford access to opportunities to engage in risk-taking behaviour, such as having access to a car (Willoughby et al., 2014). Consistent with Pfeifer and Allen (2012), this evidence would suggest that situational influences play a greater role in risk-taking behaviour than neurobiological developments. As such, while imbalance models are an important development in the understanding adolescent decision-making and are supported to some extent by lab-based studies, there are limitations to the predictions they make about adolescents' risk taking in real world contexts.

#### *1.4.2. The Lifespan Wisdom Model*

Through focussing on the rise in risk taking behaviour in adolescence, imbalance models only account for harmful behaviours exhibited by this age group, neglecting the fact

that most individuals progress through this developmental period without experiencing the negative outcomes highlighted by these models (Romer et al., 2017). In doing so, it has been suggested that imbalance models inadvertently pathologise adolescence (Ellis et al., 2012). Rather, the 'Lifespan Wisdom Model' argues that risk taking is one way in which novelty-seeking manifests in adolescence and that such behaviours allow this age group to explore their surroundings and gain experiential knowledge (Romer et al., 2017).

Like imbalance models, the Lifespan Wisdom model can explain the corpus of data that finds the VS is hyperresponsive to reward in adolescence relative to adulthood (e.g., van Leijenhorst et al., 2010b), as it suggests this sensitivity to rewards promotes novelty-seeking and exploration. Indeed, novelty can also serve as a cue for potential reward, which would be consistent with the view that adolescents' heightened novelty-seeking relative to adults is related to the hyperactivity of the dopaminergic reward system at this point in the lifespan. Exploration and novelty-seeking behaviours are expected to peak at ages 16-18, which is around the time that the sensitivity of the VS to reward is greatest (Romer et al., 2017; Qu et al., 2015) but also coincides with the period that adolescents begin to experience increased independence from caregivers (Willoughby et al., 2014). However, the Lifespan Wisdom Model deviates from imbalance models in suggesting that through trialling novel and potentially risky scenarios, the individual can learn the likelihood that their decisions will lead to positive or negative outcomes. These experiences allow adolescents to develop 'wisdom', or knowledge of their surroundings that can support them to make decisions based on prior experience. Novelty-seeking is then directed towards scenarios that have utility for learning about one's surroundings to facilitate the development of wisdom. As the individual matures and develops their experience, novelty-seeking decreases and decisions are increasingly made from knowledge gained from previous encounters (Romer et al., 2017).

However, it is clear from statistics related to substance abuse and delinquency that risk-taking behaviour can have negative, as well as positive consequences for adolescents or other age groups (Swedler et al., 2012; Willoughby et al., 2014). To reconcile this observation with the apparent functional role of novelty-seeking in adolescence, the Lifespan Wisdom Model differentiates between two trajectories of decision-making in adolescence. The first, beneficial style of decision-making, is one that is driven by novelty-seeking and promotes exploration of the environment. In contrast, there are a subset of adolescents who engage in maladaptive risk taking, which is driven by impulsivity (Romer et al., 2017). Impulsivity describes the individual's propensity to act without planning or consideration of potential consequences (Steinberg et al., 2008) and shows a distinct developmental trajectory to novelty-seeking (Harden & Trucker-Drob, 2011). Individuals who have higher levels of impulsivity do not exhibit the same learning from experience that characterises typical adolescent decision-making. Consistent with this view, within adolescent populations, higher levels of impulsivity predict poorer performance on decision-making tasks that rely on strategic learning, such as the Iowa Gambling Task (Olson et al., 2007). Adolescents who have higher levels of impulsivity are prone to engage in persistent risk taking, which may indicate an impoverished ability to learn from their experiences to avoid negative outcomes, such as persistent substance misuse (Khurana et al., 2018). Thus, the Lifespan Wisdom model can account for differences in decision-making and behaviour that exist within adolescent populations. The model suggests that exploration can be beneficial to development yet recognises the potential that persistent risk-taking behaviours can lead to harm.

Recent longitudinal studies have demonstrated that there are heterogeneous trajectories of risk-taking behaviour in adolescence, providing empirical support for the Lifespan Wisdom Model. Using latent growth curve modelling, which characterises subgroups within a population, Peeters and colleagues (2019) found that only a small proportion of participants

exhibited increased problematic substance use across adolescence. The majority of participants belonged to a group that had low baseline substance misuse and exhibited little or no increase in this behaviour. The authors suggest that while experimenting with a new drug can be characterised as an exploratory behaviour, persistent substance misuse reflects a preference for immediate rewards without consideration of the longer-term outcomes (i.e., is an impulsive behaviour; Peeters et al., 2019).

Examining the cognitive predictors of risk-taking trajectories, Khurana et al. (2018) found that a group characterised as exhibiting ‘weak cognitive control’ demonstrated high-stable levels of delay discounting (i.e., showed greater devaluing of larger yet delayed rewards relative to smaller, immediate rewards) and increasing levels of acting-without-thinking, which predicted the onset of substance misuse disorder at age 18. In contrast, heightened novelty-seeking was not associated with negative risk-taking outcomes and characterised a second, larger subgroup of adolescents (Khurana et al., 2018). Furthermore, similar heterogeneous patterns of development have been found for psychological regulation, which measures the inverse of impulsivity (Weller et al., 2021). Together, these findings provide empirical support that trajectories of risk-taking behaviour are heterogeneous across adolescents. Whilst a subgroup of individuals do engage in excessive risk taking, the majority of adolescents do not exhibit these problematic behaviours. Rather, the majority of adolescents exhibit typical levels of novelty-seeking for this age group that promotes exploration of the individual’s surroundings.

The Lifespan Wisdom Model also suggests that cognitive control actively promotes exploratory behaviour in adolescence, rather than adolescents’ behaviour arising from the absence of cognitive control as argued by imbalance models (Romer et al., 2017). Risk taking is often an effortful process compared to the relatively safe option, and subsequently requires cognitive control to enact (Sharp et al., 2021; Shenhav et al., 2013). For example, trialling a

drug for the first time may require cognitive control to override one's habitual response to avoid risks in the pursuit of a valued outcome (e.g., the approval of peers; Do et al., 2020). Further, to learn from information gained from exploration requires cognitive control (Otto et al., 2015) and there is evidence that the ability to use cognitive control for this purpose increases from childhood to adolescence (Dubois et al., 2020). As such, heightened cognitive control in late adolescence (Ferguson et al., 2021; Steinberg et al., 2018) may support this age group to learn from effortful, exploratory behaviours, which can be utilised in future decision-making.

Theoretical accounts such as the Lifespan Wisdom Model are complemented by evidence that strategic, planned risk taking is dissociable from impulsive, reactive risk taking (Gibbons et al., 2009). It has been proposed that risk taking can be employed strategically to explore the environment and facilitate the development of 'wisdom' or experiential knowledge, otherwise known as 'reasoned risk taking' (Maslowsky et al., 2019). The pursuit of these reasoned risks are driven by heightened motivational drives (i.e., increased novelty seeking; Romer et al., 2011). In contrast, reactive risk taking is impulsive and driven by a diminished ability to inhibit responses (i.e., poorer cognitive control; Luna, 2010). Indeed, higher levels of reasoned risk taking relative to reactive risk taking have been associated with better working memory (Romer et al., 2011) and the ability to set future plans and goals, known as future orientation. These findings indicate that relative to reactive risk taking, reasoned risk taking is associated with better cognitive control abilities and that these abilities can be deployed by adolescents to explore their environment (Maslowsky et al., 2019).

The Lifespan Wisdom Model presents a novel account for the increase in exploratory and risk-taking behaviours during middle-adolescence (i.e., ages 16-17). The model can explain evidence that adolescents' dopaminergic neural responses are hyperactive to the receipt of reward, relative to adults, and offers an advance on imbalance models of adolescent decision-making by recognising that there are different trajectories of risk taking during this

period. The model also acknowledges that adolescents' novelty seeking behaviours can have a functional purpose in the lifespan, presenting adolescents with the opportunity to develop knowledge of their surroundings through experience and supporting them in the transition from childhood to adulthood.

#### *1.4.3. Limitations of Models of Adolescent Decision-Making*

Imbalance models and the Lifespan Wisdom Model present extensive accounts of adolescent development that can explain much of the psychological and neuroscientific evidence in this area. However, there are some data these models cannot explain, and critics have argued that a limitation of both imbalance models and the Lifespan Wisdom Model is that they fail to acknowledge the extent to which individual differences influence the degree to which adolescents engage in risk taking (Crone et al., 2016; Foulkes & Blakemore, 2018). For example, there is a large degree of individual variation in the structural development of the adolescent brain (e.g., Mills et al., 2014; 2021; Wierenga et al., 2014) and it has been suggested that there are similar individual differences in risk-taking behaviour during adolescence, beyond those described by the Lifespan Wisdom Model (e.g., Bjork & Pardini, 2015). This has been highlighted in work that has demonstrated lower socioeconomic status and lower levels of closeness with parents are associated with increased risk taking in adolescence (Brieant et al., 2021; Potter & Font, 2019). These differences are important to acknowledge, as the majority of adolescents only engage in a limited amount of risk taking (albeit more than children or adults; Peeters et al., 2019; Romer et al., 2017). Adolescents who engage in persistent risk taking are also more likely to exhibit symptoms consistent with Conduct Disorder or Oppositional Defiant Disorder (Bjork & Pardini, 2015; Hopfer et al., 2013). Therefore, to broadly characterise adolescence as a period of heightened risk taking does not adequately recognise individual differences in development, nor the contribution of additional confounding variables, such as psychopathology.

An additional criticism of imbalance models and the Lifespan Wisdom Model are that they overemphasise the role of neurodevelopment in adolescents' risk taking. Rather, adolescents' behaviour is influenced by a range of factors, including the situational factors that are present during decision-making (Lloyd & Doering, 2019). For example, whether a decision is made in the presence of affective stimuli, such as similar-aged peers (Smith et al., 2014), can also influence adolescents' decision-making. In scenarios where there is an absence of affective stimuli (i.e., affectively 'cold' situations), adolescents exhibit decision-making that is similar to adults as they utilise on rational, deliberative processes (Figner & Weber, 2011; Icenogle & Cauffman, 2021). However, when an emotive stimulus is introduced (i.e., affectively 'hot' situations), adolescents exhibit heightened risk-taking behaviour relative to other age groups (Albert et al., 2013; Figner & Weber, 2011).

Finally, not all research has demonstrated that adolescents engage in more risk taking compared to adults or children, which poses a challenge to imbalance models and the Lifespan Wisdom Model, which both predict a rise in risk taking beginning at ages 12-13, which coincides with early adolescence (e.g., Casey et al., 2008; Romer et al., 2017; Steinberg, 2008). A meta-analysis of lab-based studies by Defoe and colleagues (2015) found that children take similar levels of risk to adolescents, which were both higher than adults' rates of risk taking. While they did find that adolescents engaged in more risk taking than adults in affectively 'hot' scenarios (Figner & Weber, 2011), adolescents took fewer risks than children if they could select a safe option (Defoe et al., 2015). As such, developmental trends in risk taking behaviour may be driven by paradigms that utilise affectively hot scenarios to measure decision-making.

The evidence that adolescents are averse to risk taking when they can select a safe option (Defoe et al., 2015) may be due to this age group overestimating the likelihood of the negative outcome occurring. Millstein and Halpern-Felsher (2002) found that when adolescents aged 10-16 were asked to evaluate the dangers associated with hazardous activities,

such as the likelihood of becoming addicted to nicotine after smoking a cigarette, they overestimated the risk involved in these actions relative to adults (though adolescents typically exhibit good probabilistic reasoning; Donati et al., 2015). These findings present a more nuanced picture of adolescent decision-making, as lab-based studies do not reproduce the curvilinear pattern of risk-taking observed in epidemiological statistics, which demonstrate reckless behaviours peak in adolescence around ages 16-17 (e.g., Moffitt, 2018; Swedler et al., 2012). This inconsistency may be attributed to the paradigms used to study adolescent decision-making, as experimental conditions of ‘risk’ may not be valid measures of the behaviour that is colloquially referred to as ‘risk-taking’ (Blankenstein et al., 2021; see 1.5. Adolescent Decision-Making in Conditions of Ambiguity). As such, it is important to consider the experimental conditions under which adolescents exhibit heightened novelty-seeking compared to adults, which I review in the next section.

### **1.5. Adolescent Decision-Making in Conditions of Ambiguity**

The research reviewed above suggests that adolescents are not indiscriminately more risk-seeking than other age groups and that situational factors play a significant role in adolescents’ decision-making. In the following section, I evaluate the experimental evidence that has examined the conditions under which adolescents’ exhibit greater novelty-seeking relative to adults and children. Based on this review, I suggest that compared to adults, adolescents are more likely to gamble in conditions of ambiguity and less likely to gamble in conditions of risk. This evidence is considered in context of the theories of adolescent decision-making reviewed in the previous section.

A problem with experimental paradigms that measure risk-taking behaviour is that these conditions rarely reflect real-world scenarios, as there are few decisions where the probabilities

associated with a choice are fully available to the decision-maker. For example, choosing to drive above the speed limit is a reckless behaviour where the probabilities associated with negative outcomes (e.g., a crash) are not fully available to the decision-maker. While the individual might have knowledge of the layout of the road, they do not know the probability that a car will be driving in the other direction, which influences the likelihood of a negative outcome. These scenarios may be more appropriately described as being conditions of ambiguity (Kahneman & Tversky, 2000; Romer, 2010). Under conditions of ambiguity, the probabilities involved in a decision are partially or completely unknown to the decision-maker (Tymula et al., 2012). In contrast, risk refers to conditions in which the probabilities involved in a choice are made explicit to the decision-maker to evaluate (Tversky & Kahnmann, 1980). For example, in choosing to gamble on a roulette wheel, the individual can clearly see the probabilities associated with their decision. As such, the epidemiological statistics that demonstrate risky behaviour exhibits an inverted U pattern, peaking in adolescence (e.g., Moffitt et al., 2018; Willoughby et al., 2014), may be measuring behaviours that are more adequately measured by conditions of ambiguity rather than risk.

Adults are typically averse to ambiguity, such that they prefer options with known probability distributions in favour of those with unknown probability distributions. Notably, these preferences extend to scenarios when selecting the ambiguous option is economically rational, as demonstrated by the 'Ellsberg paradox' (Ellsberg, 1961; Kahneman & Tversky, 2000). It has been argued that conditions of ambiguity are more relevant to adolescents' decision-making compared to risk, as their inexperience means that they have less prior experience to draw on in unfamiliar scenarios relative to adults (Crone & Dahl, 2012). Hence, a greater number of scenarios are processed as ambiguous at this point in the lifespan (Romer, 2010).

The importance of distinguishing between risk and ambiguity has been highlighted by evidence that these conditions influence decision-making in dissociable ways and are represented by distinct neural regions. Conditions of risk activate the dorsomedial prefrontal cortex and VS, whereas conditions of ambiguity recruit regions including the dlPFC and inferior parietal lobe, which has been found in both adult (Huettel et al., 2006; Wu et al., 2021) and adolescent samples (Blankenstein et al., 2018). Furthermore, risk and ambiguity attitudes have been found to be uncorrelated behaviourally (Huettel et al., 2006; van den Bos & Hertwig, 2017).

Recent empirical evidence has suggested that adolescents are more likely than adults to gamble when probability distributions are obscured and are therefore ambiguous (Blankenstein et al., 2016; Lloyd & Doering, 2019; Osmont et al., 2017; van den Bos & Hertwig, 2017; van Duijvenvoorde et al., 2012; however, see Blankenstein et al., 2018). Yet, when probability distributions are made explicit, adolescents are more averse to risk taking than adults (Reiter et al., 2019; Tymula et al., 2012). Though, it is important to recognise that not all studies have found that adolescents are more risk averse compared to adults when probabilities are explicit, as one study found that risk taking peaked at 14.38 years and declined into adulthood (Burnett et al., 2010), consistent with predictions of imbalance models. However, this study only recruited male adolescents, who exhibit higher rates of risk taking relative to female adolescents and could account for elevated rates of risk taking observed in this sample (van Leijenhorst et al., 2008). In studies recruiting male and female adolescents, there is evidence that adolescents exhibit similar or greater levels of risk aversion compared to adults (Millstein & Halpern-Fischer, 2002; Reiter et al., 2019; Tymula et al., 2012; van Leijenhorst et al., 2008). Indeed, Reiter and colleagues (2019) found that male adolescents were more risk seeking than female adolescents, which could suggest the findings of Burnett et al. (2010) were partially due to the absence of females in their sample.

Findings that adolescents gamble more than adults in conditions of ambiguity relative to risk are consistent with predictions made by the Lifespan Wisdom Model, but not those of imbalance models. Imbalance models predict motivation for reward should peak around ages 15-16 and therefore adolescents should gamble in conditions of both risk and ambiguity more frequently than adults. In contrast, the Lifespan Wisdom Model proposes that adolescents will only gamble with a variable outcome if it has some utility for learning about their surroundings. Through trialling ambiguous stimuli, individuals are able to gain experience with likelihoods that their actions will lead to positive or negative outcomes (Moreno & Rosokha, 2016), thereby learning the structure of their environment (Do, et al., 2020). These behaviours facilitate the development of ‘wisdom’, or experience-based reasoning that is later utilised for decision-making in adulthood, and thus serves a functional purpose within the lifespan (Khurana, et al., 2018).

Yet, to date, there is little empirical evidence that demonstrates developmental differences in the degree to which feedback from sampling ambiguous scenarios is utilised in future decision-making. The majority of studies comparing differences between conditions of risk and ambiguity have utilised economic forced choice tasks. In these paradigms, trial-by-trial choices are independent of one another and therefore feedback does not have utility for future decisions (Garrett & Daw, 2020), which limits the ability of these tasks to measure exploration. While there is evidence that gambling in conditions of ambiguity activates neural regions associated with encoding stimulus-outcome contingencies (Blankenstein et al., 2019), there is no direct evidence that receiving feedback from ambiguous stimuli informs adolescents’ future decision making, which is a key feature of exploration.

## **1.6. Exploration Across the Lifespan**

I have reviewed evidence that adolescents are more likely than adults to gamble in conditions of ambiguity compared to risk and have suggested this evidence is consistent with predictions of the Lifespan Wisdom Model. However, it is important to consider paradigms where feedback can be used to guide future decision-making to obtain valid measures of exploration behaviour (Garrett & Daw, 2020). Measuring exploration, where feedback informs future decisions, will allow me to explicitly test the predictions of the Lifespan Wisdom Model and compare this to the predictions of imbalance models, thereby addressing the first objective of this thesis. In the following section, I review existing literature about whether adolescents explore more than adults, and whether this evidence is consistent with the predictions of imbalance models and the Lifespan Wisdom Model.

It is common for humans to encounter scenarios that require them to search across different options with the aim of selecting the one that will yield the best outcome. For example, when deciding which restaurant to go to for dinner, you may choose between your favourite restaurant, which you are guaranteed to like, or exploring a new option that you have not tried before. If you choose to explore somewhere new, there is ambiguity that comes with this decision; you may discover a menu that is more enjoyable than your usual option or be disappointed that the food was of poorer quality compared to the option you were familiar with. The dinner date example is a classic illustration of the explore/exploit trade-off (Song et al., 2019). In this context of decision-making, the individual must trade-off between exploring a novel, unknown option or exploiting an option that is familiar to the individual. While adolescence is characterised as a period of heightened novelty-seeking (Crone & Dahl, 2012; Romer et al., 2017), which is typically associated with greater levels of exploration (Costa et al., 2014), there is mixed evidence regarding whether this age group consistently explore more than adults or children in lab-based tasks.

Explore/exploit dilemmas are a particularly complex condition of decision-making, as they incorporate several cognitive processes that are absent from other decision-making tasks (Addicott et al., 2017). In deciding to explore an unfamiliar option, the decision-maker is forfeiting the highest value option they currently know about (Addicott et al., 2021). This sacrifice can, however, lead to discovering an option with an even greater potential for reward in the long-term. Therefore, it has been suggested that the explore/exploit dilemma incorporates intertemporal decision-making (i.e., deciding between options with either an immediate or delayed reward; Shenhav et al., 2013). Indeed, it has been demonstrated in an adult sample that exploration is negatively correlated with delay discounting, suggesting that individuals who show greater discounting of future rewards also discount the longer-term rewards that can be gained from exploring unknown options (Sadeghiyeh et al., 2020). Moreover, exploration is a cognitively effortful action, as it requires the individual to deviate from the more habitual response of exploiting the known option (Otto et al., 2015). As such, motivation is required to enact exploratory behaviours (Addicott et al., 2017). Consistent with this account, disorders linked with diminished motivation to seek rewards are associated with fewer exploratory behaviours relative to healthy controls (such as alcohol use disorder; Edwards & Koob, 2010; Morris et al., 2016).

To navigate these complex scenarios, human beings have developed similarly complex strategies to maximise reward in the explore/exploit trade-off. One such strategy, known as directed exploration, emphasises how information gained from sampling unknown options acts as an intrinsic reward (Gershman, 2018). This ‘information bonus’ can motivate individuals to explore unknown options and subsequently utilise this information to guide future decisions. Indeed, there is evidence that exploration of novel options recruits the reward network (specifically the VS; Wittmann et al., 2008), which is consistent with the view that the information bonus associated with exploration is processed as a reward.

A second strategy, known as random exploration, operates through introducing randomness - otherwise known as 'stochasticity' - into the individual's decision-making. Stochasticity leads the individual to occasionally select an option that, to the decision-maker's knowledge, does not have the highest expected value of reward. Random exploration can be an adaptive strategy when the associations between stimuli, outcomes and rewards are changing, as this strategy can lead to faster discovery of such changes to the environment (Wilson et al., 2021; Gershman, 2018). Though the use of random exploration is less computationally costly, it is also less precise than directed exploration and therefore can introduce noise into decision-making (Schwartenbeck et al., 2019). The utility of directed and random exploration to resolve the explore/exploit dilemma has been demonstrated in lab-based tasks. For example, Wilson and colleagues (2014) demonstrated that adults used a combination of directed and random exploration strategies to maximise their reward intake in the explore/exploit trade-off.

Many behaviours observed during adolescence have been attributed to the tendency for this age group to explore more than older and younger individuals (Romer et al., 2017). However, research examining the development of exploration has found that strategies for exploring are differentially employed across the lifespan. Using the Horizon task, Somerville and colleagues (2017) found that random exploration strategies were employed across childhood, adolescence and adulthood. In addition, the use of directed exploration increased with age and adolescents aged 16-17 demonstrated similar use of this strategy to adults when uncovering information about their environment. The increased use of directed exploration in older adolescents could reflect the development of 'wisdom' proposed by the Lifespan Wisdom Model (Romer et al., 2017), as adolescents increasingly use prior experience to identify options that have utility for future decision-making. However, Somerville and colleagues (2017) found that when there was a conflict between an exploratory option with a high information value

and an option with an immediate reward, adolescents were more likely to select the option with an immediate reward whereas adults selected the option with a higher information value. This finding is consistent with adolescents' preference for immediate rewards in delay discounting paradigms (e.g., van den Bos et al., 2015) and suggests that while adolescents do utilise strategic exploration, this can be tempered by their preference for immediate rewards.

The refinement of exploratory tendencies from childhood to adulthood is supported by studies that examine exploratory play in childhood. Using a game that required children to learn the food preferences of an avatar, Pelz and Kidd (2020) found that exploratory play increased across childhood and became increasingly complex, allowing older children to navigate the task more successfully than younger children. This increase was predicted by the development of cognitive abilities, which allowed older participants to engage in more complex reasoning and explore novel food combinations (Pelz & Kidd, 2020). The importance of cognitive abilities for effective exploration may also explain why adolescents use less complex exploration strategies relative to adults (i.e., random exploration strategies), as their cognitive capacities are still maturing (Steinberg et al., 2018). Consistent with this view, recent research has found that children and younger adolescents rely on less computationally costly exploration strategies relative to older adolescents (Dubois et al., 2020). Together, this evidence suggests that relative to adults, adolescents may use less complex exploration strategies that can be suboptimal when navigating their environment.

However, adolescents' use of random exploration may be a beneficial strategy within the lifespan. The use of random exploration strategies can lead the individual to trial a broader range of actions, whereas directed exploration is suited to testing a narrower range of predictions (Gopnick, 2020). While random exploration can be time and energy costly when the individual has good predictions about their environment (Schwartenbeck et al., 2019), adolescents are relatively naïve about their surroundings compared to adults (Romer, 2010).

The increased reliance on random exploration in adolescence, relative to adulthood, is consistent with the Lifespan Wisdom Model, as it suggests that younger individuals utilise strategies that direct them towards trialling new actions, which can help them to learn about their surroundings. Through receiving feedback from these actions, adolescents are able to develop experiential knowledge that can be utilised later in life (Gopnick et al., 2015; Romer et al., 2017).

Overall, there is mixed evidence regarding whether adolescents are more likely to opt for immediate rewards by exploiting known resources or exploring their environment for longer-term gains. Research investigating the development of exploration strategies across the lifespan has found that adolescents exhibit different exploration strategies compared to adults (Somerville et al., 2017). However, the strategies utilised by adolescents may be suited to gaining diverse experience of their surroundings, supporting predictions of the Lifespan Wisdom Model (Gopnick, 2020).

### **1.7. Reinforcement Learning for Optimal Exploration**

A key claim of the Lifespan Wisdom Model is that feedback gained through exploration is utilised to learn about the statistics of the adolescent's surroundings. Specifically, adolescents should be motivated to trial new experiences and through these actions learn which scenarios to avoid and which to approach. The psychological process of learning from rewarding and aversive stimuli can be explained by reinforcement learning models. In this section, I review the existing literature that has examined how adolescents learn from reward feedback and how developmental differences in this behaviour may relate to adolescents' ability to learn about the statistics of their surroundings.

To maximise reward during the explore/exploit trade-off, the decision-maker must use feedback gained from previous experience to select the option that will yield the highest reward. How the decision-maker learns the statistics of their environment is explained by reinforcement learning (RL) frameworks. These frameworks specify the interdependence of stimulus, action, and outcome, which are used to guide decision-making (Sutton & Barto, 2018). RL has been an important framework for understanding decision-making across a range of domains, including social decision-making (Lockwood et al, 2020), online decision-making (Lindström et al., 2021) and human intelligence (Silver et al., 2021).

Basic RL models propose that upon encountering a novel environment, the naïve learner selects from one of the available stimuli, which produces either a positive or negative outcome. The feedback from this action is then used to update the learner's expectation about the association between that stimulus and the recently experienced outcome. After sampling these stimuli over a period of time, the learner should use their experience to identify which option(s) to approach and which option(s) to avoid (Dayan & Daw, 2008).

Building on this basic framework, more complex models have aimed to understand the processes that influence how the learner integrates feedback from sampled stimuli. One such model is the Rescorla-Wagner model (Rescorla & Wagner, 1972), which has provided a robust explanation of how humans learn stimulus-outcome contingencies (Dayan & Daw, 2008; Miller et al., 1995; Siegel & Allan, 1996). In this model, the reward expected on the next trial is explained by the agent's experience of reward on the previous trial and how much this deviates from their expectation (known as 'prediction error'; Niv & Schoenbaum, 2008). The degree to which information on this trial is integrated into the learner's expectation that the stimulus will yield a reward is modulated by how much emphasis the learner places on recent information, which is known as their 'learning rate' (Rescorla & Wagner, 1972). Higher learning rates indicate the learner rapidly integrates recent reward feedback into their valuation

of a stimuli and lower values indicating reward feedback is integrated over a longer period (Sutton & Barto, 2018). As the Rescorla-Wagner model relies only on direct associations between stimuli, actions and outcomes, this class of RL has been described as ‘model-free’ learning (Doll et al., 2012).

In contrast, ‘model-based’ RL strategies are used for more complex behaviours which require the learner to generate an internal map of events (Gläscher et al., 2010). These internal models of the world can be useful in scenarios where complex associations exist between stimuli and outcomes (Dayan & Berridge, 2014). A common task used to distinguish between model free versus model-based strategies is a two-stage paradigm (Daw et al., 2005). In this task, the learner chooses between two stimuli which probabilistically transition to one of two final stages. During the final stage, the learner selects from a further two stimuli, which yield a probabilistic reward. The successful model-based learner should identify which of the final-stage stimuli yields the highest reward, whereas a model-free learner would opt for the first stage stimulus that produced a reward of any magnitude (da Silva & Hare, 2020). Adults typically utilise model-based inference, rather than model free strategies to effectively navigate this task (da Silva & Hare, 2020).

In both model free and model based RL, learning rates should be calibrated to the structure of the environment to navigate one’s surroundings. In typical day-to-day experiences, it is reasonable to expect a certain degree of stability, though there may be small variations in these experiences (i.e., expected uncertainty; Soltani & Izquierdo, 2019). For example, when an individual purchases a tin of soup from the supermarket, they could reasonably expect to find the tin filled with the soup they had chosen. If this individual opened a tin and found it to be empty, this would violate their expectation about the contents of the tin (i.e., be a negative prediction error). However, as they know the likelihood of tins being filled with food to be a stable property about the world, they should not place too much emphasis on this unlikely event

and forego purchasing any tins of soup in the future. Placing less emphasis on recent events means the individual is using a lower learning rate and in the ‘tin’ example, is a rational decision-making policy.

The utility of this policy has been demonstrated empirically using tasks such as the multi armed bandit task (Schulz et al., 2020). In this paradigm, the participant must sample two slot machines or ‘bandits’ that have probabilistic reward schedules, with the aim of accruing as many points as possible within a limited number of trials (Katehakis & Veinott, 1987). If the probabilities associated with these bandits remain stable for the duration of the task (for example, one has a probability of 0.75 whereas the other has a probability of 0.25 of yielding a reward), then integrating information over a longer period can be a valuable strategy for learning the true reward schedules associated with each bandit (Nussenbaum & Hartley, 2019). Under these conditions, integrating information too quickly can lead to the learner overweighting rare events (e.g., if the agent loses on the bandit with a higher probability of yielding a reward, this learner might falsely believe the bandit has a poor likelihood of yielding a reward). This rapid integration of information can prevent the learner from identifying the true probabilities associated with a bandit and therefore lead to suboptimal decision-making (see Figure 1.1). As such, it is appropriate for the decision-maker to utilise a low learning rate.

1)  
75% chance  
of winning



2)  
25% chance  
of winning

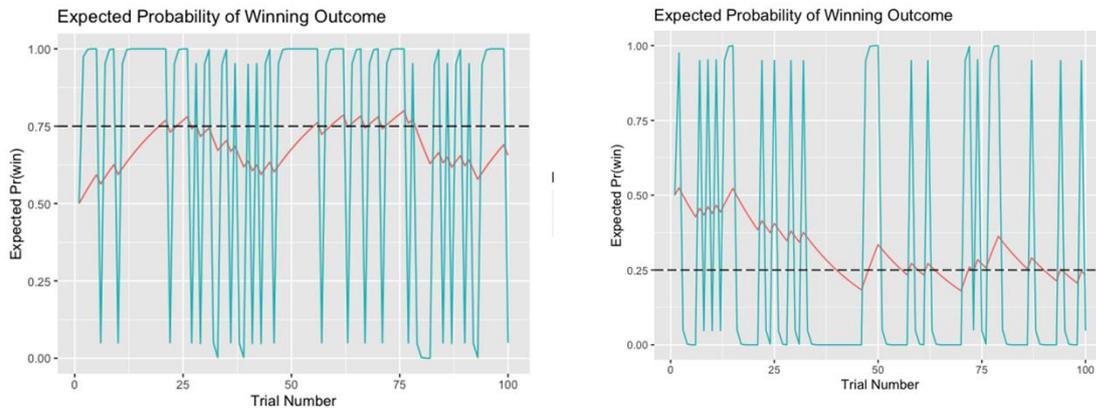


Figure 1.1: Simulation of a Rescorla-Wagner RL agent estimating the reward schedules of two probabilistic bandits. Bandit 1 (left) has a 75% chance of yielding a reward and bandit 2 (right) has a 25% chance of yielding reward, which are indicated by the horizontal dashed lines. Blue lines indicate a decision-making policy that only uses the most recent occurrences (i.e., a high learning rate), whereas red lines indicate a decision-making policy that integrates information over a longer series of trials (i.e., a low learning rate).

However, using a lower learning rate can also be suboptimal in environments that are rapidly changing, such as the COVID-19 pandemic (Lloyd et al., 2021a). During the pandemic, health policies were rapidly changing to mitigate the spread of coronavirus and as such, it was beneficial to utilise the latest information released by the Government to update health behaviour. Consistent with this, one study found that individuals who had a higher learning rate on an experimental task were better able to adopt novel health behaviours during the pandemic (Lloyd et al., 2021a). Experimentally, the ability to adapt to changeable environments has been demonstrated in a manipulation of the multi armed bandit task. Behrens and colleagues (2007) introduced a ‘volatile’ block, where the probabilities associated with each bandit would shift during the task. In this condition, a policy that relies on more recent information is advantageous (Browning et al., 2015). Adult participants were able to adapt their

learning rate between conditions of stability and volatility when accruing rewards (Behrens et al., 2007), suggesting they can effectively adjust RL strategies to different environments.

Examining how the use of RL strategies develop across the lifespan, several studies have identified a shift from simple model free strategies in childhood to increasingly complex model-based approaches in adolescence and adulthood (Decker et al., 2016; Dubois et al., 2020). Cohen and colleagues (2020) developed a RL task where participants had to learn to select one of two probabilistic stimuli to acquire a reward. They also included a 'hidden agent', who intervened to generate positive, negative, or random outcomes. Participants would be more accurate if they correctly estimated the influence of the hidden agent on the likelihood of receiving a reward. Both adolescents' (aged 13-17) and adults' (aged 18-25) behaviour was best fit by a Bayesian RL model that included beliefs about the hidden agent's influence on the outcomes, indicating a model-based strategy. However, children's (aged 7-12) behaviour was best explained by a model which only incorporated information from trial feedback. This model assumed no inference about the hidden agent and indicates that children used a simpler, model free learning to guide their decision-making. Older participants' more matured cognitive control may account for their greater use of the model-based strategy (Otto et al., 2015). For adolescents aged 13 and above, the increased use of model-based strategies may be supported by the strengthening of functional connectivity between neural regions responsible for cognitive control and learning (DePasque & Galvan, 2017; McCormick et al., 2021; van Duijvenvoorde et al., 2016).

To analyse RL strategies, model parameters can be computed through a search process known as maximum likelihood estimation, which maximises the likelihood of the data by testing different values of parameters (Farrell & Lewandowsky, 2010). These quantities have psychological interpretations, such as the rate at which individuals integrate recent information, represented by the learning rate parameter (Sutton & Barto, 2018). Examining developmental

differences in RL parameters, there is mixed evidence regarding changes in the learning rate from childhood to adulthood. Separate studies have found learning rate to increase (Davidow et al., 2016; Master et al., 2020; Xia et al., 2020), decrease (Decker et al., 2015) or not change with age (Javadi et al., 2014). However, critiques have noted that these inconsistent results may be attributed to the different tasks used in these studies, as the optimal learning rate is governed by the statistics of the task (Eckstein et al., 2021; Nussenbaum & Hartley, 2019). These critiques are notable as it could suggest that developmental differences (or lack thereof) in learning rate are governed by the paradigm used in these studies rather than changes to cognition across the lifespan.

To further understand how learning from feedback changes across development, it is also important to distinguish between better-than-expected and poorer-than-expected outcomes (i.e., positive and negative prediction errors; Niv & Schoenbaum, 2008). It has been demonstrated that humans track positive and negative prediction errors separately during RL (Lefebvre et al., 2017), and these types of feedback are represented by different neural regions (Spoomaker et al., 2011). Compared to adults, adolescents overweighted negative prediction errors, exhibiting greater neural activation in the anterior insula, which supported this age group to flexibly learn about changes to their environment (Hauser et al., 2015). However, the authors noted that these differences in responses to negative prediction errors did not translate to behavioural differences on the task.

Further research has examined how adolescents and adults learn from punishment, in addition to reward. Palminteri and colleagues (2016) demonstrated that adults learned from both reward and punishment, which allowed them to learn which stimuli to select to accrue points or avoid losses. In contrast, adolescents learned from reward, but did not demonstrate learning in the punishment condition. As such, adolescents exhibited an impoverished ability to select appropriate stimuli to avoid losses (Palminteri et al., 2016). These findings are

consistent with imbalance models that posit adolescents are motivated by reward without the ability to restrict behaviour to avoid negative outcomes (Shulman et al., 2016).

Another parameter that contributes to RL measures how stochastic participants' decisions are (i.e., how often they deviate from the option that is expected to yield the highest reward; Gershman, 2018; Wilson et al., 2021). This tendency is measured by the SoftMax temperature, with higher values on this parameter being associated with the learner more reliably selecting the option with the higher expected value (Nussenbaum & Hartley, 2019). Consistent with developmental research on the explore/exploit trade-off, several studies have demonstrated that adolescents have lower values on the SoftMax temperate and subsequently exhibit a greater degree of stochasticity during RL compared to adults (Decker et al., 2015; Javadi et al., 2014; Palminteri et al., 2016). As higher levels of stochasticity are associated with testing diverse predictions about one's environment (Gopnik, 2020), these findings are consistent with the Lifespan Wisdom Model that adolescents maximise novelty-seeking behaviours (Romer et al., 2017).

### **1.8. Human and Animal Models of Patch Foraging**

Foraging for resources is a near-ubiquitous behaviour across the animal kingdom (Mobbs et al., 2018). In traditional foraging scenarios, the organism searches for resources that are directly linked to adaptive pressures, such as items of food or mates (Prokopy & Roitberg, 1984; Timberlake et al., 1987). Human beings have spent the bulk of their evolutionary history in foraging economies and the success of this species has been dependent on effectively navigating foraging environments (Johnson & Earle, 2000). While foraging in hunter-gatherer societies was directly related to needs necessary for survival, this context of decision-making has also been utilised to understand a wider range of behaviours, including panic buying

(Bentall et al., 2021), cannabis consumption (Phillips et al., 2020), and gathering information online (Pirolli & Card, 1999).

Patch foraging is a particular type of foraging where the organism searches for resources across discrete patches (Charnov, 1976). These patches are separated by a distance that incurs a time cost, requiring the organism to travel before they can begin collecting the resource. Once at a patch, the animal expends time and energy accumulating rewards from this source. However, the longer that the organism stays within a patch, the fewer rewards are available for it to collect. Therefore, the longer the forager remains with a single patch, a greater expenditure of time and energy is required for a relatively smaller return of rewards (Stephen & Krebs, 1986). Alternatively, the forager can choose to explore their environment to find a new patch, which has a fresh distribution of rewards. Yet, choosing to explore incurs a time cost without any immediate return of rewards (McNamara & Houston, 1985). Therefore, effective foraging requires the organism to optimise the explore/exploit trade-off.

As effective foraging improves an organism's likelihood of survival, several species developed adaptations to increase foraging success. For example, bumblebees (*Bombus terrestris*) have developed a heightened sensitivity towards colours that indicate the quantity of pollen a flower produces, and bee colonies with a stronger sensitivity towards these colours outperform colonies without this sensitivity (Raine & Chittka, 2007). In primates (*Macaca mulatta*), neurons in the dorsal anterior cingulate cortex encode a reward value that signals the value of exploring a novel patch and this is sensitive to changes in the foraging environment, indicating a neural basis for foraging decisions in primates (Hayden et al., 2011). As human beings have spent the majority of their evolutionary history in conditions where foraging was necessary for the organism's survival (Smith et al., 1983), it is argued that this species' cognitive abilities have evolved around the adaptive pressures to optimise reward intake while foraging (Kolling & Akam, 2017). Specifically, it has been suggested that the distributed

circuits across the prefrontal and anterior cingulate cortex evolved to reduce errors in foraging choices in primates (Hunt, 2021; Passingham & Wise, 2012)

Ecological work into patch foraging has led to the development of a formal model to explain optimal foraging behaviour (Charnov, 1976). According to the 'Marginal Value Theorem' (MVT), the optimal forager should leave the current patch when the rewards expected from exploiting the current patch fall below the average reward rate for the environment. Through use of this policy, animals foraging for food can maximise intake whilst minimising energy expenditure, allowing them to maintain homeostasis (Korn & Bach, 2015). According to MVT, the forager must track two separate values when deciding whether to exploit or explore. The first is the reward expected from exploiting the current patch, which is known as the foreground reward rate. The second is the average reward rate for the environment, known as the background reward rate (Gabay & Apps, 2021). Classical MVT expects the forager has a complete knowledge of the background rate of rewards, which is a rational assumption for animals who have prior knowledge of the environment (Charnov, 1976). Empirically, it has been demonstrated that animals conform to MVT while foraging across a range of domains (e.g., Bettinger & Grote, 2016; Pyke, 1978; Wajnberg et al., 2000) and MVT has been found to outperform other models when explaining animal foraging (e.g., Bayesian models; Marshall et al., 2013). Anthropological studies have found that human hunter gatherer groups broadly conform to MVT when deciding when to move campsites (Venkataraman et al., 2017) and while foraging for food (Hawkes et al., 1982), demonstrating the importance of the MVT framework for understanding foraging behaviour across species.

Several lab-based studies have examined how humans approach foraging scenarios. In a seminal study examining the neural mechanisms of foraging, Kolling and colleagues (2012) measured neural activation while participants made decisions between exploiting one of two known stimuli or exploring a new set of stimuli. The latter of these options incurred either a

low, medium, or high search cost. Consistent with animal studies (e.g., Hayden et al., 2011), functional magnetic resonance imaging (fMRI) analysis suggested that higher activation of the anterior cingulate cortex (ACC) was positively correlated with the value of searching the environment. The authors computed the value of exploring as the average value of the other available options, which was known to participants (Kolling et al., 2012). Moreover, activity in the ACC was negatively correlated with the value of exploiting the current options. Notably, the ACC is also involved in motivating exploration in the explore/exploit trade-off (Blanchard & Gershman, 2018). As such, these findings may indicate that the ACC is associated with exploration through its role in comparing the values of exploiting the current resource against the value of the options that can be explored (however, see Shenhav et al., 2014).

Further research has expanded on these findings by considering how humans are able to adapt their foraging behaviour according to changes in the quantity of rewards available in an environment. The availability of rewards in a foraging environment affects the background reward rate, which according to MVT, should dictate when an organism leaves a patch (Stephen & Krebs, 1986). In a novel apple foraging task, Constantino and Daw (2015) manipulated the time incurred when travelling to new patches along with the rate at which rewards depleted from patches being exploited. The authors found adults broadly conformed to the predictions of MVT, staying for longer in poorer environments and exploring more frequently in richer environments.

To explain how participants learned the reward statistics of each environment, Constantino & Daw (2015) developed a model based on RL, which suggests the forager uses previous reward history to estimate the value of the background reward rate. Like traditional RL models, this includes the free parameter measuring learning rate, which governs how highly the participant weighs recent information during their decision-making, and a free parameter measuring stochasticity in participants' decision-making (Sutton & Barto, 2018). Using a

Bayesian model comparison technique, Constantino and Daw (2015) demonstrated that this simpler model provided a better explanation of the data compared to several more complex models. However, further work has suggested a more nuanced account where the forager has separate learning rates for positive (i.e., better than expected) and negative (i.e., worse than expected) outcomes while foraging (Garrett & Daw, 2020). Compared to a model with a single learning rate, the model that separated learning rates for positive and negative prediction errors was found to provide a better explanation of participants' behaviour. Together, these findings highlight that human adults estimate both the foreground (i.e., within patch) reward rate and background (i.e., average environmental) reward rate and use this to guide their foraging behaviour.

Although adults are able to adjust their foraging behaviour to environments that are richer or poorer in quality, these participants exhibit a consistent bias to exploit patches for longer than an optimal foraging agent (Constantino & Daw, 2015). Computationally, this was captured using a parameter that measured participants' risk sensitivity, suggesting participants perceived exploration as a riskier decision compared to exploiting the known option. Indeed, adults' overexploitation bias has been replicated in several other foraging studies (e.g., Lenow et al., 2017; Le Heron et al., 2020), which have also implicated the mesolimbic dopaminergic system in tracking the background rate of rewards (Le Heron et al., 2020). It has been suggested that older organisms explore less due to a combination of greater experience, which allows organisms to identify resource rich patches more efficiently, and physical decline, which necessitates more efficient foraging (Gurven et al., 2006; Walker et al., 2002). Corroborating this account, one study found older adults explored less than younger adults across two lab-based foraging tasks (Mata et al., 2013).

To my knowledge, only one study to date has examined human foraging in a developmental sample. Bach and colleagues (2020) recruited 781 adolescents aged between

14-24 who completed a risky foraging task, where they searched for rewarding tokens. In the foraging phase the participant explored grid spaces to acquire reward tokens. Then, in a second chase phase, the participant's avatar was pursued by a 'predator', which would remove all their collected tokens if it landed on the same space as the participant. Their results demonstrated that there were significant sex differences, such that male participants accumulated significantly more rewards and were willing to approach closer to the predator relative to female participants. However, they observed no age differences on task performance as measured by the number of tokens collected. This finding was further corroborated by an absence of maturation effects in an accelerated longitudinal design, where multiple single cohorts, each starting at a different age, are tested at several timepoints (Bach et al., 2020). While these findings appear inconsistent with work demonstrating that risk-taking behaviour peaks at ages 15-16 (e.g., Steinberg, 2008) and exploration behaviours peak at ages 16-17 (Xia et al., 2020), with both declining into adulthood, the paradigm used in this task did not operationalise foraging in a traditional sense, as participants could not exploit patches for additional rewards once they had located individual tokens.

The first objective of the current thesis is to investigate how adolescents approach patch foraging tasks, which I use to test the predictions of imbalance models and the Lifespan Wisdom Model. Specifically, I will examine whether adolescents exhibit different foraging strategies compared to adults (Chapters 3 and 4), which I initially frame as a two-tailed hypothesis based on contrasting research that demonstrated adolescents explore more often than adults (Xia et al., 2020) and less often than adults (Somerville et al., 2017). As part of this objective, I examine how adolescents and adults adapt their learning strategies between stable and volatile foraging environments (Chapter 5) to understand how adolescents utilise exploration to learn about their surroundings. Through examining the behavioural and computational differences between adolescents' and adults' foraging strategies, I aim to

provide evidence that can inform our understanding of the role of exploration and learning at this point in the lifespan. These findings have implications for theories of adolescent development (e.g., Casey, 2008; Romer et al., 2017; Steinberg, 2005) and can aid in understanding phenomena that emerge during this period, such as novelty-seeking and risk taking.

### **1.9. Peer Influence in Adolescence**

The second objective of the current thesis is to examine how social variables, specifically the presence and views of peers, influence adolescents' decision-making. While a body of research has substantiated effects of peer presence on adolescent risk taking in both lab-based and real-world settings (e.g., Chein et al., 2011; Giletta et al., 2021; Simons-Morton et al., 2005), there are debates regarding the mechanisms through which peers influence adolescents' behaviour (Andrews et al., 2021; Ciranka & van den Bos, 2019). Moreover, there is a paucity of research on how peers influence adolescents' decisions to explore new environments when seeking rewards. In Chapter 6, I use a foraging task to test theories of how social influence affects adolescents' behaviour and whether this age group is more susceptible to the views of their peers compared to adults. Understanding the role of social influence and how this interacts with exploration and learning during adolescence is important, as peers can be both a protective (Foulkes et al., 2018) and risk factor (Knoll et al., 2015) for hazardous behaviours at this point in the lifespan. In the following section, I will critically review models of how adolescents' decision-making is affected by their peers and how this can be expected to influence their behaviour while foraging.

'Distraction' theories claim that the presence of peers disrupts adolescents' ability to engage in goal-directed behaviour (Dumontheil, 2016). On this account, the emotional salience

of peer presence reduces adolescents' self-control and cognitive skills (e.g., Ciranka & van den Bos, 2019; Rosenbaum et al., 2017). In reward-based decision-making, lower levels of cognitive control are associated with diverting from the option with the highest expected value, (i.e., exhibiting more stochastic behaviour; Dubois et al., 2021), presenting one mechanism through which peers might distract adolescents' decision-making. In support of this account, Tymula (2019) found that, when adolescents make decisions under conditions of risk and ambiguity that had the same expected value, adolescents under observation by a similar-aged peer were less consistent in their decision-making compared to adolescents who completed the task alone. As such, the influence of peers on adolescents' behaviour may be through increasing stochasticity during decision-making compared to other age groups, possibly mediated via distraction (Ciranka & van den Bos, 2019).

In contrast, 'reward sensitivity' theories propose that similarly aged peers sensitise adolescents' motivational system to the potential rewards that can be obtained through risky behaviour. This stimulus further exacerbates adolescents' sensitivity to rewards, which is already heightened due to the rapid development of the dopaminergic reward system at the onset of puberty (Spear, 2011). Evidence for this theory has been derived from fMRI studies that observe heightened activation in the ventral striatum and orbitofrontal cortex when adolescents aged 14-18 receive rewards under the observation of a peer during a driving simulation task, relative to conditions where adolescents complete the task alone (Chein et al., 2011). Due to their heightened sensitivity to rewards adolescents are more likely to opt for the choice with the higher reward compared to the less rewarding, but safer option, when under the observation of their peers (Chien et al., 2013; O'Brien et al., 2011). These 'reward sensitivity' theories of peer influence draw on evidence that adolescence is a period of imbalance between networks responsible for processing reward and those responsible for cognitive control (Steinberg, 2008).

Behavioural evidence for reward sensitivity theories have been derived from risky decision-making paradigms where the presence or absence of a same-aged peer is manipulated. For example, Gardner and Steinberg (2005) asked adolescents aged 13-16 and adults aged 18-22 to complete a driving simulation, manipulating whether they were observed by a same-aged peer or completed the task alone. Consistent with statistics related to driving accidents, adolescents were more likely to take risky decisions in the simulation compared to adults, but only when observed by a peer. These findings support reward sensitivity theories as reward seeking behaviour was heightened in the presence of peers relative to when they completed the task alone. Corroborating this account, Smith and colleagues (2014) found that adolescents exhibited greater risk taking during a wheel of fortune task where the probabilities associated with each outcome were made explicit. Notably, observation by a peer increased risk taking when the probability of losing points outweighed the probability of winning points, suggesting that adolescents' behaviour was driven by their sensitivity to the rewards that could be obtained through sampling the risky option when in the presence of peers. These findings highlight how peer presence increases the salience of rewards under conditions of risk. Moreover, these effects are observed in adolescents, but not adults (Chein et al., 2011) or children (Large et al., 2019), which is consistent with the predictions of reward sensitivity models (Ciranka & van den Bos, 2019).

However, in real world contexts peer interactions are not static, with peers passively observing adolescents during risk taking. Rather, these scenarios are dynamic, with peers expressing attitudes that may either endorse or discourage risky behaviour (Lloyd & Doering, 2019). The attitudes that peers express can have important implications for adolescents' behaviour. For example, associating with deviant peers predicts adolescents' consumption of alcohol, tobacco, and cannabis (Greenwood et al., 2021). 'Social motivation' theories propose that adolescents become increasingly oriented towards peers as they progress through

adolescence and therefore the views and attitudes expressed by these peers have greater importance for adolescents relative to children and adults (Braams et al., 2019). These views subsequently influence adolescents' own attitudes and behaviours, as the social rewards gained through conforming to the views of a peer outweigh the potential risks to one's health that can result from risky behaviour (Blakemore & Mills, 2014).

The views of same-aged peers may have greater weight for adolescents compared to other age groups as adolescents are acutely sensitive to social rejection. In a Cyberball task, adolescents who were excluded by two fictional peers reported higher levels of anxiety relative to adults (Sebastian et al., 2010). In a separate study adults exhibited greater activation of the ventrolateral prefrontal cortex during exclusion compared with inclusion, relative to adolescents (Sebastian et al., 2011). Activation in this region had previously been associated with reduced distress during social exclusion in adults (Eisenberger et al., 2003), potentially indicating self-regulation. Reduced ventrolateral PFC response may therefore indicate reduced self-regulation during social rejection in adolescence, although there was no direct inverse association between response in this region and self-reported distress.

Perhaps to avoid social exclusion, adolescents conform to the views expressed by same-aged peers. For example, younger adolescents (aged 12-14) adjust their attitudes towards risky scenarios to reflect those of same aged peers but show less adjustment to the views of adults (Knoll et al., 2015; 2017). Notably, this effect is bidirectional, such that the attitude of peers could lead to both increased and decreased tolerance to risk. However, older adolescents (aged 16-18) did not adjust their views to reflect those of same-aged peers, suggesting susceptibility to social influence is heightened in early adolescence relative to later adolescence (Knoll et al., 2015). Further, an additional study has demonstrated that relative to the views of adults, the views of same-aged peers can influence adolescents' attitudes towards prosocial behaviours, such as raising money for charity (Foulkes et al., 2018). Together, these findings suggest that

peers who discourage risk taking and endorse prosocial behaviours may be an effective means of guiding adolescents away from harmful behaviours. Indeed, utilising peers to reduce risk-taking behaviour has been recommended to combat adolescent risk taking in novel contexts, such as the COVID-19 pandemic (Andrews et al., 2020).

Several studies have also demonstrated that the views expressed by a peer can lead to either increases or decreases in adolescents' risk-taking behaviour (de Boer & Harakeh, 2017; Harakeh & de Boer, 2019; Pei et al., 2020; Reynolds et al., 2014). Braams and colleagues (2019) demonstrated that late adolescents (aged 15-17) were more likely to follow the advice of a peer when this endorsed the safe, rather than risky option on the Wheel of Fortune task (however, see Haddad et al., 2014). In a computational reanalysis of Braams et al. (2019), Ciranka and van den Bos (2019) demonstrated that a computational parameter capturing the 'social bonus' of following a peer's advice was higher when the peer endorsed the safe option. In contrast to Knoll and colleagues (2015), evidence from Braams et al. (2019) suggests that late adolescents are sensitive to social influence that endorses risk aversion, whereas early adolescents may be more susceptible to social influence that endorses both risk seeking and risk averse behaviour.

There have been mixed findings with regards to whether the level of ambiguity influences adherence to peer feedback in adolescence. Braams and colleagues (2019) found no difference between conditions of risk and ambiguity on whether participants adhered to the peer's advice, which is consistent with the findings of several other studies (Braams et al., 2020; Blankenstein et al., 2016). However, a contrasting body of evidence has demonstrated that peer influence is stronger in conditions of ambiguity compared to risk (Lloyd & Doering, 2019; Osmont et al., 2021). Recently, it has been demonstrated that adolescents utilise social information when they are uncertain about their choice preferences (Reiter et al., 2021), which is heightened in conditions of ambiguity relative to risk (Braams et al., 2019). For example,

van Hoorn and colleagues (2017) demonstrated that adolescents conformed more to the views of peers as the uncertainty of a gamble increased. In novel foraging contexts, the decision maker is initially naïve about the rewards statistics of the environment, meaning there is uncertainty in the values of exploiting and exploring (Sadeghiyeh et al., 2019). While information about the environment can be learned from trialling different actions, social information may also provide guidance on the appropriate levels of exploration and exploitation – an effect observed in other foraging species (e.g., Baude et al., 2008).

There has been limited research examining how peers influence exploration across the lifespan. Using the Iowa Gambling Task, it has been demonstrated that adolescents aged 13-15 initially follow advice from a same-aged peer about which options to exploit, but this age group exhibits heightened exploration in later trials, allowing adolescents to discover other advantageous options that were available to select during the task (Buritica et al., 2019). In contrast, adults aged 18-22 consistently followed the advice of a peer, which led them to forfeit discovering other advantageous decks throughout the task. Adolescents' reliance on their independent experience may be supported by increases in metacognition – the ability to evaluate quality of one's own knowledge – from childhood to adolescence (Moses-Payne et al., 2021). Yet, whether adolescents show a preference for following safe rather than risky peer feedback whilst exploring, as they do in conditions of risk (Braams et al., 2019; 2021) and ambiguity (van Hoorn et al., 2017), has not been examined to date.

Distraction, reward sensitivity and social motivation theories propose three mechanisms through which peers influence adolescents' behaviour. While these models are not mutually exclusive, they motivate different hypotheses about how adolescents' foraging behaviour will be affected by their peers and possible age-related differences. As such, the second objective of this thesis is to consider how social influence affects adolescents' and adults' foraging behaviour, which is examined empirically in Chapter 6.

### **1.10. Early Life Stress and Decision-Making in Adulthood**

The third objective of the current thesis is to test a theory of development that explains how different strategies of decision-making are adopted based on the individual's early rearing environment. Life History Theory proposes two trajectories of development that are either suited to stable rearing conditions or adverse (i.e., unstable, volatile) rearing conditions (Ellis et al., 2011). I aim to examine whether an individual's early caregiving environment is associated with changes to their foraging strategies, along with the computational mechanisms that are linked to these decision-making styles.

According to Life History Theory, in typical households that are relatively stable, adolescents adopt a decision-making style that prioritises long-term rewards (i.e., those that can even span over several decades; Brumbach et al., 2009), which persist into adulthood. Adolescents who have been raised in stable households should use this setting as a base to explore their surroundings and develop their knowledge, which can be used in adulthood later in life to gain rewards (Figueredo et al., 2006). For example, having a stable home environment may allow adolescents to explore novel activities or educational pursuits that can yield positive outcomes for the individual in adulthood.

In contrast, individuals exposed to unstable or volatile rearing environments do not experience the same caregiving consistency (Nelson III & Gabard-Durnam, 2020) and such conditions act as a cue that resources are inadequate or scarce (Belsky et al., 1991). According to Life History Theory, individuals exposed to these circumstances adopt a 'fast' decision-making style, which prioritises the maximisation of short-term rewards. This style of decision-making is aimed at ensuring earlier reproductive success when the chances for parenthood later in life are uncertain (Brumbach et al., 2009). The aim of ensuring short-term reproductive

success may be reflected on a biological and behavioural level, as childhood adversity is associated with an earlier onset of menarche and an earlier sexual debut in adolescence (Alley & Diamond, 2021; Jorm, et al., 2003). Due to the robust association between experiences of instability in early development and the maximisation of short-term rewards (Lee et al., 2018), it has been proposed that individuals who utilise this strategy engage in behaviours that appear problematic, such as poorer dietary choices (Kaplan et al., 2000). However, understood within the framework of Life History Theory, these behaviours are aimed at prioritising short-term outcomes, such as high calorie intake, over delayed outcomes, such as good long-term physical health (Gunstad et al., 2006). Examining how early adversity affects decision-making strategies can provide insight into the link between early life stress and negative outcomes in adulthood.

Early life stress has a profound impact on the developing brain and related networks, which may contribute to behavioural patterns that develop following childhood trauma (Danese & McEwen, 2012). It has been consistently evidenced that individuals with experience of early life stress exhibit greater activation of the amygdala in negative emotional contexts relative to controls (Cohen et al., 2013; Hanson et al., 2015; Tottenham et al., 2011; Van Tieghem & Tottenham, 2017). The amygdala has been implicated in learning about the relative safety or danger of a scenario (Davis & Whalen, 2001) and greater sensitivity of this region may be adaptive in conditions of adversity to identify threatening situations (Tottenham & Sheridan 2010). In contrast, regions responsible for processing reward exhibit less activation in individuals with a history of early life stress relative to those without these experiences (Boecker et al., 2011; Dillon et al., 2009; Kasperek et al., 2020; Vidal-Ribas et al., 2019). For example, Hanson and colleagues (2015) found that greater levels of early life stress were associated with lower levels of VS activity when participants received positive reward feedback, though the authors found no effects of early life stress on VS activation in response

to negative reward feedback. These findings indicate that early life stress can alter neural responses to reward feedback during decision-making.

Importantly, early life stress may also impact the connectivity between the amygdala and VS, which has implications for reward learning (Fareri & Tottenham, 2016). Along with processing threat, the amygdala is responsible for coding the value of a given stimulus-outcome pairing (Hart et al., 2014). This value is subsequently communicated to the VS, which is responsible for coordinating motor responses when a valued reward is at stake (i.e., whether to pursue or avoid the stimulus; Bissonette et al., 2013; van der Meer & Redish, 2011). As such, amygdala-striatal interactions may be crucial for updating stimulus values and value-based learning (Costa et al., 2016; Fareri & Tottenham, 2016; Niv & Schoenbaum, 2008). Indeed, several studies in typical adults have demonstrated the importance of amygdala-striatal interactions for updating stimulus values during reward learning (Camara et al., 2008; Cohen et al., 2008). Therefore, early life stress may affect how the individual learns from reward feedback during the explore/exploit trade-off.

Supporting this account, early life stress is associated with a poorer ability to learn associations between stimuli and outcomes. Hanson et al. (2017) tested adolescents aged between 12-17 with and without exposure to adversity (specifically physical abuse) on an associative learning task. In this paradigm, participants were required to learn the probabilities associated with two stimuli to accrue rewards. The authors found that adolescents with exposure to early stress made noisier decisions, captured by a parameter measuring stochasticity in decision-making (see 1.6. Exploration Across the Lifespan; Hanson et al., 2017), which is associated with the use of random exploration strategies (Gershman, 2018). Further, another study demonstrated that women who had experienced childhood sexual abuse demonstrated poorer reinforcement learning. These differences were driven by a diminished ability to utilise positive reward feedback, but not punishment, to guide future behaviour

(Pechtel & Pizzagalli, 2013). Together, these differences in learning stimulus-outcome contingencies between individuals with and without exposure to early stress indicate that reward-learning may be one computational mechanism impacted by these experiences.

Utilising reward feedback is an important ability for navigating the explore/exploit trade-off, yet limited research has examined how early life stress impacts exploratory behaviour in adulthood. In animal studies, exposing juvenile rats to stress causes less exploration once they reach adolescence (Spivey et al., 2008). In a foraging paradigm, rats exposed to stress in adolescence exhibited higher rates of vigilance behaviours and lower rates of foraging behaviours compared to rats without exposure to stress (Chaby et al., 2015; 2016). Notably, the authors found that rats exposed to stress were more efficient at foraging under high threat conditions, though this did not translate to differences in rates of exploration or exploitation.

However, to my knowledge, no research has examined how early life stress affects humans' ability to navigate the explore/exploit trade-off, which I examine empirically in Chapter 7. These findings have theoretical implications for understanding how the preference for immediate rewards following early life stress (i.e., steeper delay discounting; Lee et al., 2018) translates to other behavioural domains, such as the explore/exploit trade-off. This can inform our understanding of the link between early life stress and behaviours that emerge following these experiences (e.g., risk taking; Belsky et al., 1991)

### **1.11. The Current Thesis**

The current thesis will use a series of foraging paradigms to test theories of decision-making across the lifespan, using a combination of behavioural measures, questionnaires, and computational modelling. The first objective is to test predictions from imbalance models and the Lifespan Wisdom Model about developmental differences in goal-directed behaviour. To

achieve this, I examine whether adolescents aged 16-17 exhibit differences in exploration and learning compared to adults aged 21 and above (Chapters 3, 4 and 5). In Chapter 3, I examine whether adolescents explore more frequently in a foraging task compared to adults. I predict that adolescents will demonstrate differences in foraging behaviour relative to adults, though I do not predict a direction for this difference as there is conflicting research regarding whether adolescents explore more or less than adults (reviewed in 1.5. Exploration Across the Lifespan). These findings will provide evidence that arbitrates between the competing predictions of imbalance models and the Lifespan Wisdom Model. Imbalance models would predict that, relative to adults, adolescents' heightened sensitivity to reward would lead them to pursue immediate rewards at the expense of larger, yet delayed rewards (Casey, 2008; Steinberg, 2008). In contrast, the Lifespan Wisdom Model would predict that adolescents' heightened novelty-seeking would motivate them to explore more than adults, to support adolescents to learn about their surroundings (Romer et al., 2017). As such, these findings can provide evidence about the functional role exploration and novelty-seeking in adolescent development.

To address the first objective of this thesis, I also compare adolescents' and adults' learning strategies on a patch foraging task in Chapter 4. As the Lifespan Wisdom Model predicts that adolescents engage in heightened exploration relative to adults as a means of learning about their surroundings, I examine whether there are developmental differences in how these age groups learn from reward feedback. This finding will contribute to debates about how model parameters that measure reward-learning develop across the lifespan. For example, whether individuals' learning rate increases or decreases from adolescence to adulthood (Nussenbaum & Hartley, 2019). In doing so, these findings will provide a mechanistic explanation of how adolescents and adults utilise information to guide future decision-making.

The first objective of this thesis is also addressed in Chapter 5, where I examine how adolescents aged 16-17 and adults aged 24 and above adjust their learning strategies according to changes in their environment. Specifically, I predict that adolescents will show better adaption of their learning strategies between conditions of stability and volatility compared to adults (Crone & Dahl, 2012). The findings of this study will contribute to our understanding of how adolescents adjust their behaviour according to changes in their environment relative to adults.

The second objective of the current thesis is to examine theories of how social influence affects adolescents' behaviour. I use a foraging task which includes feedback from a same-aged peer that either promotes exploration, exploitation, or the correct foraging behaviour as defined by MVT. This will allow us to test hypotheses about the mechanisms through which social influence affects adolescents' reward-seeking behaviours. Specifically, if peers make adolescents hypersensitive to reward (Chein et al., 2011), I expect adolescents' behaviour to be affected uniformly across the peer influence conditions. However, if peers influence adolescents' behaviour through endorsing certain behaviours (Ciranka & van den Bos, 2019; Knoll et al., 2015), I expect adolescents to adjust their behaviour to conform to the views of their peer. Finally, in leveraging computational models of foraging behaviour, I can test whether peers increase the stochasticity in adolescents' decision-making, which would provide support for distraction theories (Dumontheil, 2016). I further compare the behaviour of adolescents aged 13-17 and adults aged 21 and above to examine whether sensitivity towards peers is heightened in adolescence relative to adulthood.

The third objective of the current thesis is to test the predictions of Life History Theory that experiences of adversity in childhood and adolescence will direct individuals to adopt strategies of decision-making that either prioritise short- or long-term rewards (Lee et al., 2018). I predict that adults who have experienced early stress will demonstrate a preference for

short term rewards and subsequently exploit patches for longer than individuals without experience of early life stress. Further, I examine how participants' learning rate is associated with their exposure to early life stress and whether this influences their foraging behaviour. These findings will contribute to the literature that has examined deficits in reward processing following early stress (e.g., Hanson et al., 2015) through highlighting how a novel computational mechanism, learning rate, is affected by these experiences.

## Chapter 2: Methodology

### 2.1. Computational Modelling of Cognition

Verbal theories have been useful in guiding our understanding of decision-making in adolescence and adulthood. However, critics have highlighted the limitations of verbal theories to explain the full complexity of cognitive development at this point in the lifespan (e.g., Pfeifer & Allen, 2012). Verbal theories are a heuristic for describing patterns in complex datasets, but do not capture exact processes that can be described using formal models (Smaldino, 2020). By nature, verbal theories rely on terms that can have ambiguous or multiple meanings (e.g., ‘cognitive control’), which in the most extreme cases can render a theory unfalsifiable (Popper, 1963; Gigerenzer, 1998).

Verbal theories can also influence researchers’ interpretations of findings, such that results are interpreted within a specific model and ignore other possible explanations (Casey, 2015). This poses a problem for scientific advancement as findings are interpreted within existing frameworks, rather than refining or re-evaluating them (Smaldino, 2020). For example, some verbal theories that fail to be substantiated by attempted replications have been defended by referring to ‘hidden moderators’ that were not measured during the replication study, rather than questioning the theory’s validity (Fried, 2020; Gershman, 2019). To counter these issues, it has been suggested that formal (i.e., mathematical, or computational) models should be developed to make the assumptions of verbal theories explicit. Formal models can be used to explain cognitive development across the lifespan (for one example, see Kievit, 2020). Throughout this thesis, I adopt a computational modelling approach to understand developmental differences in patch foraging behaviour. In the following section, I present the theoretical and methodological background to computational modelling how it is employed in

the current thesis, and key design choices made about the computational model utilised throughout the thesis.

### *2.1.1. Computational Models: An Overview*

Computational models provide a formal framework with which to understand human cognition (Lewandowsky & Farrell, 2010). The aim of computational models within psychology and neuroscience is to convey the processes through which cognition operates in precise, quantifiable terms, which are typically expressed in mathematical language (Guest & Martin, 2021). These formal models can explicate processes that are often assumed or implied in verbal theories, allowing for these mechanisms to be inspected and compared against competing explanations of these cognitive processes (van den Bos & Eppinger, 2016). The use of computational modelling has become increasingly prevalent in developmental science and has provided important insight into cognitive development across the lifespan (van den Bos et al., 2018).

Computational modelling approaches conceptualise cognition as an information processing system that is designed around solving specific problems (Kraft & Griffiths, 2018). Marr (1977) proposed that any information processing system can be understood at three levels. The first is the computational problem the system is trying to solve, or the goal it is trying to achieve (Marr & Poggio, 1976). In human cognition, one such goal might be to learn the types of bushes that contain the highest density of edible fruits while foraging (Passingham & Wise, 2012). The second level is the algorithm the system uses to solve such a problem. If one's aim is to learn which bushes are the densest, using reward feedback to update one's expectation of the number of fruits on a particular species of bush is one such example of an algorithm a system might use (Sutton & Barto, 2018). The third level is the implementation, which is how the system enacts the algorithm. In human cognition, this may be the physical or neural adaptations to that goal, such as a specialised region of the brain that encodes prediction error if

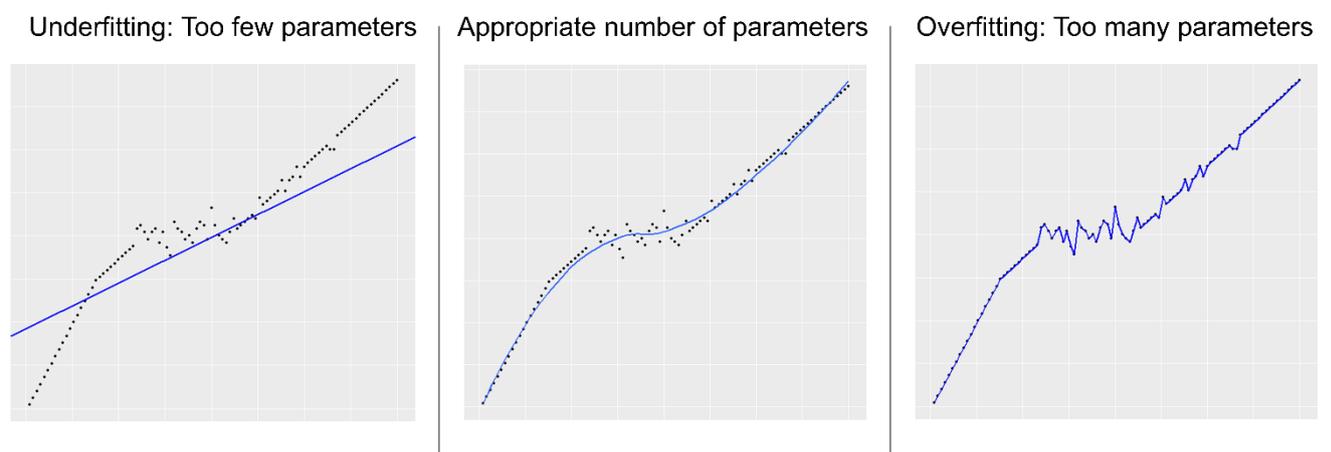
the density of fruits on a bush is better or worse than expected (Lockwood et al., 2020). In this framework, computational models form the second level of the information processing system, specifying the algorithm through which a cognitive task (in this example, learning which bushes to forage) is solved.

The structure of formal computational models is, in part, defined by the structure of the tasks that are used to measure cognition. Fixed parameters reflect aspects of the task that are not influenced by participants' behaviour. For example, in my foraging tasks, one fixed parameter is the time cost associated with travelling to a new patch, which I manipulated in several of the foraging paradigms. However, as the goal of computational modelling is to characterise cognitive processes, these formal models also have parameters that vary between individuals and between task conditions (Wilson & Collins, 2019). These 'free parameters' are latent variables that have psychological interpretations. A type of model called Reinforcement Learning (RL; see 1.7. Reinforcement Learning for Optimal Exploration) models decision-making using the learning rate, which measures how much weight the individual places on recent reward feedback versus more historic reward feedback, as a free parameter. The values of free parameters are computed through fitting computational models to behavioural data and estimating the parameters that best explain participants' behaviour. This parameter estimation is achieved through trialling different values of each parameter until the values that maximise the likelihood of the data are identified (Lewandowsky & Farrell, 2010). The best estimates of the free parameters can be subsequently utilised as variables in statistical analyses and compared across populations or task conditions to provide a mechanistic insight into decision-making (e.g., Dubois et al., 2020; Eckstein et al., 2020; van den Bos & Hertwig, 2017). In the current thesis, I compare free parameters estimated from adolescents' and adults' patch foraging behaviour.

Computational modelling can also be used to compare different algorithmic explanations of human cognition to determine which provides the best explanation of participants' behaviour. This process is known as 'model comparison' and involves the specification of several computational models, each with a different mathematical formulation. Comparing these models to determine which provides a more parsimonious explanation of behaviour can yield insight into the cognitive mechanisms that are involved in decision-making. I use this model comparison technique in Chapters 4, 5 and 7 to test a proposal from recent work that foraging behaviour is best explained by a model that separates learning rates for positive and negative prediction errors compared to a model which includes a single learning rate for all prediction errors (Garret & Daw, 2020). Moreover, comparing model fits across populations can provide insight into the strategies different age groups utilise to approach cognitive tasks (Wilson & Collins, 2019). For example, Cohen and colleagues (2020) found that children used a strategy that was best explained by a simple RL model that only involved learning stimulus-action contingencies, whereas adolescents and adults used a strategy that was best fit by a more complex RL model, which involved developing an internal model of the task structure (i.e., a model-based strategy; Daw et al., 2005). In the current thesis, I will use model comparison to arbitrate between how well different models explain participants foraging data. Moreover, I will use this method to assess whether there are systematic differences between adolescents' and adults' cognitive strategies when navigating foraging environments.

While computational models can provide mechanistic insight into human cognition, these formalisations should be recognised to have some important limitations. The inclusion of additional free parameters can lead models to reach an unrealistic level of specificity that allows these models to explain a single dataset perfectly, which is problematic because these highly specific models are not generalisable to other individuals or datasets. Models that are

‘overfitted’ to the data have limited utility, as they do not inform our understanding of features of cognition that can be generalised to other individuals but instead explain specific patterns in a single dataset (Daw, 2011). To counter the issue of overfitting, computational modelling relies on the principle of Occam’s Razor, which states that models of cognition should be as simple as possible, but no simpler (Myung & Pitt, 1997). In addition, some indices used to assess how well these models fit participants’ data include penalisations for each additional free parameter that is included in the model (Lockwood & Klein-Flügge, 2021), providing a safeguard against models with unrealistic levels of specificity. I utilise such measures of model fit in the current thesis, which I describe in greater detail below. In contrast, models with too few parameters can fail to adequately explain participants’ behaviour, known as ‘underfitting’. For a visual example of under and overfitting, see Figure 2.1.



*Figure 2.1: A curve fitting demonstrating models that have been underfitted (left) and overfitted (right) to the dataset. Each point on the plot indicates a single datapoint and lines indicate how well the model fits to the data. The middle panel is an example of a model that includes an appropriate number of parameters to explain the dataset. Note, this figure does not reflect models used in the current thesis but was simulated for illustrative purposes.*

In the current thesis, I use computational modelling to gain a mechanistic insight into adolescents’ and adults’ patch foraging behaviour. Through comparing free parameters

between these age groups, I examine how cognition develops from adolescence to adulthood. In the following section, I describe the steps undertaken to generate and test the computational models used throughout the thesis.

### *2.1.2. Implementation of Computational Modelling*

To contextualise the implementation of the computational modelling methodology used in this thesis, I first describe the basic structure of the patch foraging paradigms used, and the core features of the computational models that I employ. Broadly stated, the patch foraging paradigms I utilise presents the decision-maker with a series of trials where they are required to choose between exploiting a patch to harvest immediate rewards and exploring a novel patch which has a fresh distribution of rewards. The longer participants remain with their current patch, the fewer rewards are available to collect, as these rewards deplete exponentially. However, exploration decisions incur a time cost without any immediate return of rewards. When participants arrive at a new patch, the initial value of rewards that are harvested from the patch is drawn from a probabilistic distribution and is known as the ‘initial richness’ of the patch (Constantino & Daw, 2015). These features of the task should be utilised by the forager to evaluate the relative value of explore and exploit decisions.

The approach I take to modelling participants’ behaviour is drawn from previous studies that have adopted a RL framework to explain patch foraging (Constantino & Daw, 2015; Garrett & Daw, 2020; Kolling & Akam, 2017; Niv et al., 2002). In this framework, reward feedback from explore and exploit decision are integrated into the forager’s estimate of the average reward rate of the environment. An account that explains optimal foraging behaviour (Marginal Value Theorem (MVT); Charnov, 1976) prescribes that the decision-maker should leave their current patch when the rewards expected from exploiting the current patch fall below the average reward rate of the environment. Accordingly, in my computational

models, the expected value of the next exploit decision and the participant's estimate of the average reward rate are entered into an equation that calculates the probability that participants will explore or exploit on any given trial (Constantino & Daw, 2015). These models are described in greater detail below (see 2.2.1. Modelling Outline and Assumptions) and in the corresponding empirical chapters of this thesis (Chapters 4, 5, 6 and 7).

The first stage to implementing computational models is to develop a paradigm that is amenable to this methodology. In the current thesis, I adapted a paradigm that has previously been analysed using computational modelling, originally developed by Constantino & Daw (2015). Importantly, this is a sequential decision-making task which allows us to model learning, as previous trials inform the participant's expectation about the rewards on the following trials. One notable design feature is that decisions are presented in discrete trials, where participants decide to stay or leave patches. In contrast, some research has used continuous trials, where the behaviour of interest is participants' patch residency time (e.g., Le Heron et al., 2020). However, RL approaches typically compute the probability that a participant will select a particular action based on their experience of the reward history associated with each available option (in the case of foraging, whether to explore or exploit). As such, utilising discrete trials allowed us to model choice behaviour according to this framework, rather than continuous trials which model the time spent within patches (Botvinick et al., 2009).

A further design consideration with the foraging task was to ensure that the paradigm included a sufficient number of trials where participants decided whether to explore or exploit to reliably estimate the free parameters (Daw, 2011). An endemic issue with foraging paradigms is that there is variability in the number of trials each participant experiences, which results from participants having autonomy over the length of time they spend in each patch (Gabay & Apps, 2021). Foraging environments are set at a fixed length of time to measure how

successful participants are at collecting rewards within a limited period. However, as foraging explore/exploit decisions have different time costs associated with each action, the participants' preference for these actions can influence the number of trials experienced within the fixed block. For example, a participant who opts for a highly exploratory strategy will experience fewer trials than a participant who opts to exploit patches until they are fully depleted, as exploratory decisions incur a greater time-cost than exploitation decisions. To anticipate this issue, I employed simulations of the computational models to ensure that each environment included a sufficient number of trials to model both exploit and explore decisions.

Simulation is also important as it allows for verification that differences in model parameters translate to behavioural differences on the task, which can inform hypotheses about group-level differences (e.g., differences between adolescents and adults) on free parameters that vary according to the fitted data in participants from different groups. For example, in Chapter 5, I simulated an RL agent's behaviour on a novel foraging task where the statistics of the environment alternated between being rich in rewards and being poor in rewards (i.e., were volatile; Behrens et al., 2007). I ran two simulations using this RL model, one with a high learning rate where feedback is rapidly integrated into the estimate of the background rate of rewards and one with a low learning rate where feedback is slowly integrated into the estimate of the background rate of rewards. Plotting these estimates alongside one another demonstrated that utilising a higher learning rate provided a more accurate reflection of changes to the quality of the volatile foraging environment, which translated to more optimal behaviour (see Figure 2.2). These simulations informed my predictions about how participants would adapt their learning rate between the stable and volatile foraging environments. This example is one instance of how simulation was used to inform the design of my studies, such as the number of trials I included in each environment, and my predictions about the behavioural effects I expected to observe. Throughout the thesis, I simulate computational models to evaluate

whether they can capture differences in foraging behaviour prior to fitting these models to participants' data.

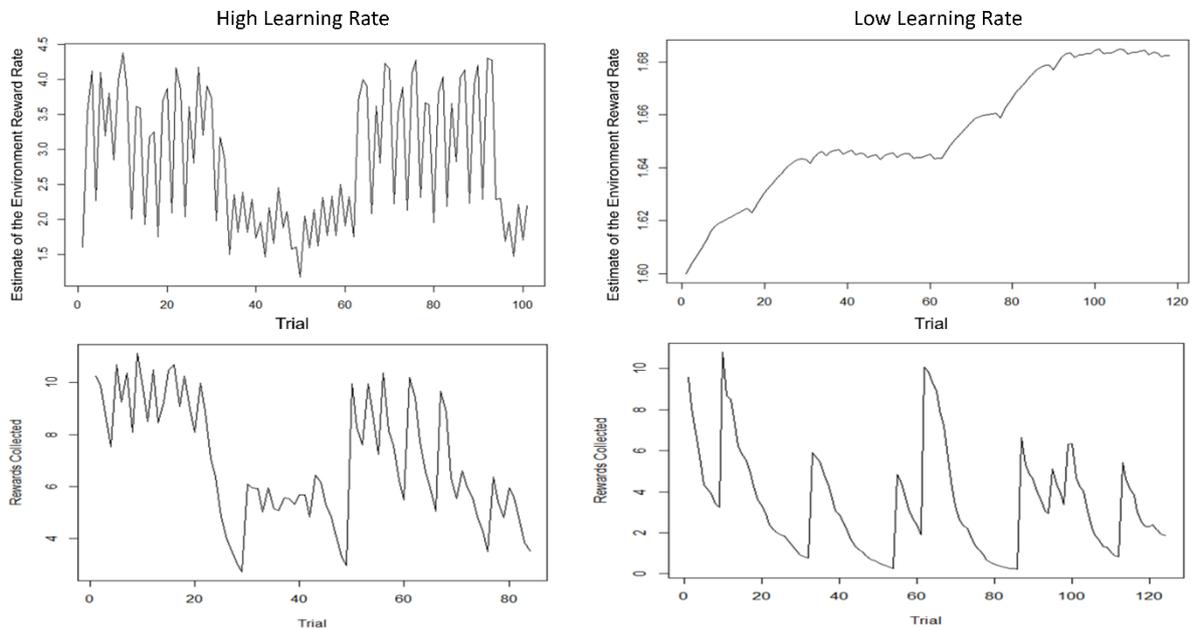


Figure 2.2: Plots demonstrating the simulated agent's estimate of the environmental reward rate (top two panels) and how this affects the agent's foraging behaviour (bottom two panels). The panels on the left demonstrate a forager who utilises a high learning rate and therefore updates their estimate of the average reward rate of the environment quickly in response to recent trials. In contrast, the panels on the right demonstrate a forager who updates their estimate of the average reward rate too slowly to adjust their behaviour to changes in the environment.

Once I was confident that my model simulations were adequately able to capture differences in foraging behaviour, I then conducted parameter recovery on these models. The process of parameter recovery involves simulating data where all parameters in the model are fixed and therefore the true values are known. Once the data have been simulated, the key parameters of psychological interest (e.g., the learning rate) are freed and the model is fit to the data to estimate the value of these parameters (Wilson & Collins, 2019). Models with good parameter identifiability should return parameter estimates close to the hardcoded values entered into the simulation (Lockwood & Klein-Flügge, 2021). I ran parameter recovery for a

wide range of values of each free parameter to ensure that identifiability is not limited to a narrow subset of parameter values. For example, recovery for the learning rate parameter should be run on values ranging from 0-1 in increments of 0.01, which is the typical range for this free parameter in human RL task data (Sutton & Barto, 2018). For the  $\beta$  parameter, parameter recovery was conducted on values ranging from 0-5 in increments of 0.5 and for the overexploitation parameter, parameter recovery was conducted on values ranging from -1-3 in increments of 0.5. Positive correlations between hardcoded and fitted parameters indicate successful parameter recovery (Wilson & Collins, 2019). While there is no formal threshold for correlations to be deemed indicative of successful parameter recovery, stronger positive correlations indicate the parameter can be estimated more robustly (Lockwood & Klein-Flügge, 2021). Throughout the thesis, I use these positive correlations as evidence that my parameters were recoverable, though I find variation in the strength of these associations. As I have different task statistics in each of the empirical chapters, I conducted parameter recovery for each chapter separately to ensure the model could reliably estimate free parameters in each study. In chapters where I compare several models (e.g., Chapters 4, 5, 6 and 7), I conduct parameter recovery on each of the models tested in those studies.

After recovering my parameters, I was able to estimate the parameters for each participant. In the current thesis, I run parameter estimation over the sum of the negative log likelihood that participants will decide to explore or exploit, which is an approach adopted in previous RL studies (e.g., Constantino & Daw, 2015; Eckstein et al., 2020; Davidow et al., 2016; Wilson & Collins, 2019). This value is supplied to the optim function (Nash, 1990), which is included in R's stats package (specifically versions 3.5.1 and 3.6.2; R Core Team, 2018; 2020). The optim function uses an optimisation process based on the Nelder-Mead algorithm to find the values that minimise the sum of the negative log likelihood (Nelder & Mead, 1965). The algorithm trials different parameter values to identify those that make the

observed data most likely given the assumptions of a computational or statistical model (Myung, 2003). Bounds are set on the parameter values the algorithm can trial to ensure that the function does not use nonsensical or invalid values. In the model I used in this thesis, minimum and maximum bounds were set at 0 and 1 for the learning rate parameter, which is the typical range for this parameter (Sutton & Barto, 2018). The bounds for the  $\beta$  parameter were set at 0 and 5, which were selected based on previous research (e.g., Lockwood & Klein-Flügge, 2021). The bounds for the overexploitation parameter were set at -1 and 3, which were determined by simulation, as values less than -1 led the agent to continually explore, which is a nonsensical foraging strategy (Constantino & Daw, 2015). Starting from free parameter values that are provided to the algorithm, the function trials values of parameters that incrementally increase or decrease from this starting value, computing whether changing the value of these parameters improves or worsens the fit of the model to the data (Steenbergen, 2006). The iterative process of trialling different parameter values continues until adjusting the values of the free parameters no longer leads to an improvement of the model's fit to the data, as indicated by lower values of the summed negative log likelihood. The free parameter values that minimise the sum of the negative log likelihood are those that best explain the behavioural data.

However, one constraint of the Nelder-Mead algorithm is that the starting values utilised to begin the optimisation process are supplied to the optim function, rather than starting values being trialled across the entire parameter space (Nash & Varadhan, 2011). As such, there is the potential that the algorithm will only trial a narrow range of parameter values and identify the best set of parameter values within this range, failing to test the full range of viable values (Speekenbrink & Konstantinidis, 2015). These local troughs in the parameter estimation process are known as 'local minima' and divert the algorithm from finding the estimates of the free parameters that provide the best fit to the data across the entire parameter space, otherwise

known as the ‘global minimum’ (Wilson & Collins, 2019). One method of avoiding local minima is to supply `optim` with a range of starting values (i.e., conduct a grid search across starting values) and identify the best fitting parameter estimates across the entire parameter space (e.g., Burrítica et al., 2019; Speekenbrink & Konstantinidis, 2015; see Figure 2.3). I utilise this approach throughout this thesis to estimate the parameter values that best fit participants’ behavioural data. Specifically, I created a matrix containing different combinations of starting values for each of the parameters. Each combination of starting values is then trialled by the algorithm to identify the parameter values that best fit the data. The starting values used to create the matrix were generated using the `seq()` function for R, which generates a sequence of numbers with a specified length that would determine the intervals between each value in the sequence. For example, in a sequence between 0.2 and 0.9 with a length of 5, the function would return the values 0.200, 0.375, 0.550, 0.725 and 0.900. This function was utilised to generate a 5\*5\*5 matrix of starting values with the ranges of 0.2 and 0.9 for the learning rate parameter, 0.5 and 3.5 for the  $\beta$  parameter, and 0.5 and 1.5 for the overexploitation parameter. In utilising this method, I ensured that a wide range of starting values, and different combination of these starting values were trialled to ensure I identified the global minima during parameter estimation. In addition to using this method to estimate parameters from my behavioural data, I apply this same method of conducting a grid search across multiple starting values during parameter recovery, as I also use the `optim` function during this process.

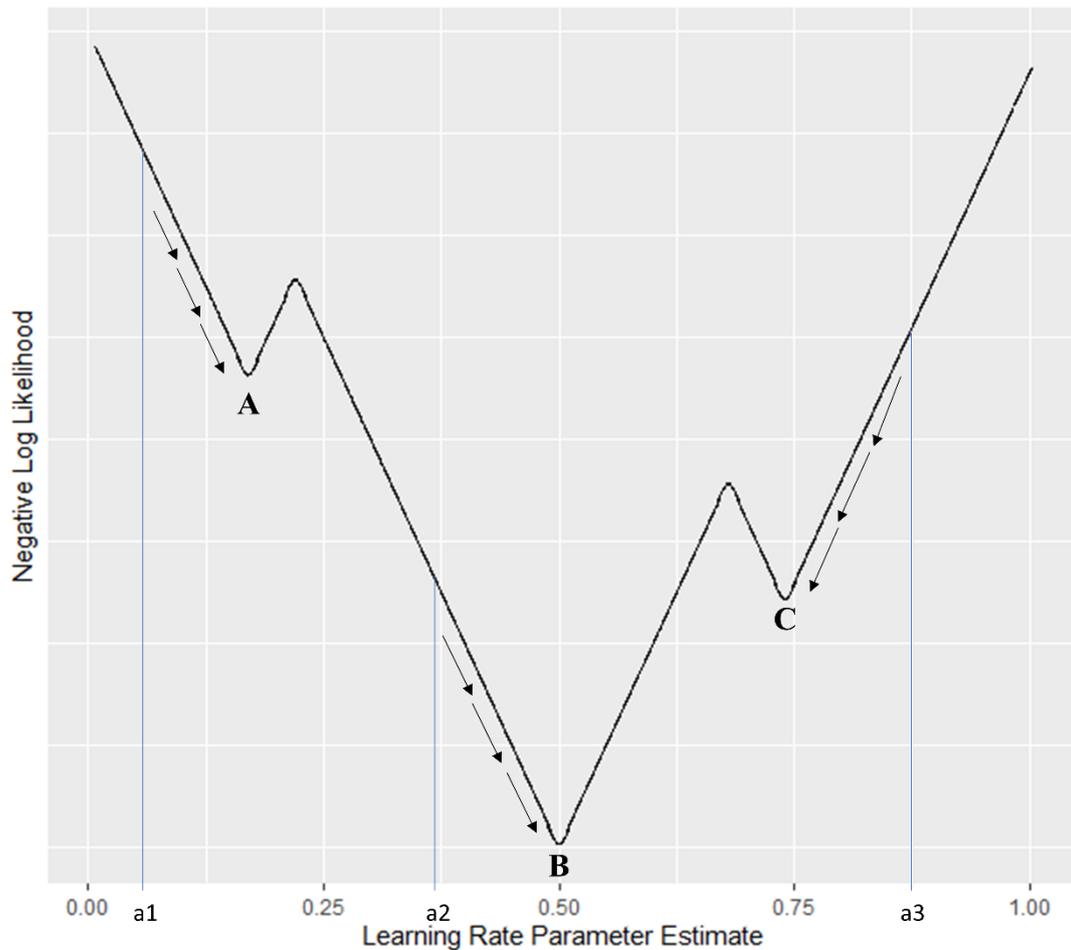


Figure 2.3: Simulated plot of the minimisation algorithm implemented by the optim function. The x-axis presents the algorithm's estimate of the learning rate parameter. The blue vertical lines a1-3 indicate different starting values provided to the algorithm. Point B is the global minima and the best estimate of the learning rate parameter in this example. Points A and C are local minima that the algorithm has identified.

Once the algorithm has estimated the best fitting values of the free parameters, one can calculate indices of model fit, which provide a quantitative measure of how well the model explains the data. Two commonly used indices of model fit are the Bayesian Information Criterion (BIC; Schwarz, 1978) and Akaike Information Criterion (AIC, Akaike, 1998). While these metrics share some commonalities, a notable difference between the BIC and AIC is the extent to which they penalise or favour more complex models. The BIC is computed as  $-2 \cdot \log(\text{likelihood}) + n\text{Params} \cdot \log(n\text{Trials})$ , where  $n\text{Params}$  is the number of free parameters in the model and  $n\text{Trials}$  is the number of trials in the dataset. As greater values of  $n\text{Params}$  increases

the final value of the BIC, this metric penalises more complex models with a higher number of free parameters. In contrast, the AIC, calculated as:  $-2 \cdot \log\text{-likelihood} + 2 \cdot n\text{Params}$  and favours models that include a greater number of free parameters (Lockwood & Klein-Flügge, 2021). Lower values on both criteria indicate that the model is a better fit to the data (van den Bos et al., 2018). In the current thesis, I use both metrics to assess how the computational models fit participants' behavioural data.

Once the AIC and BIC have been calculated for a set of candidate models, one can then compare the fits of these models to identify which provides the most parsimonious explanation of the data. A simple way to identify the winning model is to compare which model has lower values on the AIC and BIC, which is the method of model comparison I employ throughout this thesis (however, there are also more complex methods such as k-fold cross-validation or leave-one-out procedures; Farrell & Lewandowsky, 2011; Friedman et al., 2001; van den Bos et al., 2018). When the same model has lower values on both criteria, this is an indication of the winning model. However, due to the formulaic differences between these criteria, they do not always identify the same model as the best fit to the data (as I find in Chapter 6). In this case, it has been suggested that the winning model can be determined by one's *a priori* hypotheses, which is the approach I adopt in Chapter 6 (Lockwood & Klein-Flügge, 2021). Specifically, if one has hypotheses about which parameters are expected to differ between populations, and statistical models evidence that such differences exist, these results can be considered adequate reason to favour a model using only one criterion of model fit (Lockwood & Klein-Flügge, 2021). In the current thesis, I adopt the AIC and BIC as indices of model fit and use these measures to identify models that best explain participants' patch foraging behaviour.

Finally, it is important to conduct posterior predictive checks on the winning model to ensure this can recapitulate participants' behaviour and any group-level differences observed

in the data. Simulation can again be utilised to replicate trends found in behavioural data through taking the central tendency of free parameter estimates for each group and simulating data according to these parameter values (e.g., Burrítica et al., 2019). These simulations should reflect patterns in the data and a failure to find similarities between behavioural and simulated data can be indicative of a poorly fitted model (van den Bos et al., 2018). Posterior predictive checks can be complemented by model recovery analyses, where each of the candidate models are fitted to the simulated data. The winning model that was used to generate the data should provide a better explanation of these data, indicated by lower values on the AIC and BIC (Wilson & Collins, 2019).

I have described the steps taken to implement and validate the computational models utilised throughout this thesis. Through completing these steps, I can be confident that the computational models utilised in this thesis are adequately able to capture participants' behaviour. Analyses of candidate models and free parameters within these models can arbitrate between competing explanations of how participants approach foraging problems. In the current thesis, I use this methodology to test hypotheses about adolescent development and how specific features of decision-making change across the lifespan.

## **2.2. Modelling Assumptions and Variable Choices**

### *2.2.1. Modelling Outline and Assumptions*

In this thesis, I make several choices in the formulation of my models that affect how participants' behaviour is explained. To contextualise these decisions, I will provide a basic outline of the computational model I use throughout the thesis before discussing specific features of the core model. Design choices made with regards to specific models can be found in their respective chapters (e.g., the social learning model described in Chapter 6).

The model I use throughout this thesis is derived from Marginal Value Theorem (MVT), which prescribes that the optimal forager should leave their current patch when the rewards expected from the next exploit decision fall below the average reward rate for the environment. As the forager does not know the reward statistics of the environment *a priori*, they must learn the average reward rate. I utilise an equation developed by Constantino and Daw (2015) to model how the forager learns the average reward rate:

$$p_i = (1 - \alpha)^{T_i} \frac{s_i}{\tau_i} + (1 - (1 - \alpha)^{T_i}) p_{i-1} \quad (1)$$

In Equation 1,  $p_i$  is the running estimate of the average reward rate, which is subject to the reward ( $s$ ) received on any  $i$ th trial, along with the time associated with a stay or leave decision ( $\tau_i$ ). The forager's estimate of the average reward rate is also influenced by the free parameter  $\alpha$ , which measures the learning rate (see 1.7. Reinforcement Learning for Optimal Exploration for a description of the learning rate). The forager's estimate of the average reward rate is then entered into a SoftMax equation, which calculates the probability the participant will stay on each trial:

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

In Equation 2,  $c$  measures participants' overexploitation bias and  $\beta$  measures stochasticity in participants' decision-making. The value  $k$  is the participants' estimate of the depletion rate within their current environment and  $h$  is the time associated with exploiting a patch, which was fixed at three seconds in all versions of the foraging task used in this thesis.

Within this core model are several assumptions that should be justified. In the current formulation of the model, participants' estimate of the depletion rate ( $k$ ) is calculated as a running average of the true depletion experienced on each trial (i.e.,  $s_i / s_{i-1}$ ). This approach assumes that participants utilise a simple running estimate of the depletion rate rather than a more complex approach, such as one that includes a free parameter measuring the degree to

which participants utilise recent trials to update their estimate of the depletion rate (i.e., a learning rate for the depletion rate). This modelling choice was made as previous studies have found that more complex RL models provide a poorer explanation of foraging behaviour relative to the simpler model currently used in this thesis (Constantino & Daw, 2015). Further, adding a free parameter to measure the rate at which participants updated their estimate of the depletion rate could risk making the models too complex and unable to generalise to other datasets or individuals (i.e., be overfitted to the data; Wilson & Collins, 2019). Moreover, whenever possible, we adhered to the formulation of the model that has been well-tested in previous research (e.g., Constantino & Daw, 2015) which facilitated comparison between our findings to these prior studies.

I also needed to choose initial values (before participants experienced the first trial in each environment) of the average reward rate and depletion rate variables. As in previous research (Constantino & Daw, 2015), we used simulation to determine the maximum average reward rate that could be acquired using the optimal foraging strategy and utilised this value as the initial value for the average reward rate. The initial value for the depletion rate was set as a null variable and was populated throughout the duration of the environment.

### *2.2.2. Outline of Key Variables*

Throughout the empirical chapters in this thesis, I consider both participants' leaving threshold, which is the number of apples participants select as a threshold to determine when to leave their current patch, and how much this leaving threshold deviates from an optimal forager based on simulations derived from the MVT leaving rule (Charnov, 1976). These variables are employed differently within each empirical chapter, as they provide information that can address particular research questions. Specifically, the leaving threshold is a measure of participants' exploration behaviour, as this variable measures the number of apples that the

next harvest must fall below before participants are willing to leave and explore a novel patch. As such, the leaving threshold can be utilised to address developmental questions about whether adolescents exhibit heightened exploration relative to adults, which can inform theories that posit adolescence is a period of novelty-seeking relative to adulthood (such as in Chapter 3). In addition, through providing information about developmental differences in exploitation behaviour (i.e., a preference for an immediate, known reward rather than a delayed yet unknown reward), the leaving threshold can inform predictions about participants' preference for immediate versus delayed rewards (as utilised in Chapter 7).

However, the leaving threshold is limited insofar as this variable cannot be used to draw inferences about the *appropriateness* of participants' exploration behaviour in their foraging environment. This distinction is important, as previous research has found that adolescents' heightened exploration can be suboptimal relative to adults on tasks such as the Iowa Gambling Task (e.g., Almy et al., 2020). To address this limitation of the leaving threshold measure, I also utilise a measure of how participants' leaving threshold compares to a behavioural optimum, defined computationally through simulation. The 'deviation from the optimal leaving threshold' variable involves calculating the difference between participants' leaving threshold and the optimal leaving threshold for the foraging environment. This variable is employed throughout the thesis to address research questions about whether adolescents' heightened exploration is associated with positive outcomes (Chapter 3), as opposed to suboptimal outcomes as found in previous research (Almy et al., 2020). Therefore, identifying whether adolescents' exploration can lead to optimal outcomes for this age group would be a novel finding in this literature and one that could not be detected without defining the optimal leaving thresholds in each foraging environment.

Further, in Chapter 6, I use the deviation from the optimal leaving threshold variable to address a novel research question about how peers can affect the optimality of adolescents' exploration behaviour, which builds on previous research that has examined how peers affect risk-taking and exploration behaviour (e.g., Braams et al., 2019; Burritica et al., 2019). Moreover, using the deviation from the optimal leaving threshold variable allows me to infer whether adults' bias to explore less than adolescents (Almy et al., 2020) might be beneficial in poor quality environments relative to rich quality environments (e.g., in Chapter 7), which could not be concluded using the leaving threshold alone. Nevertheless, the deviation from the optimal leaving threshold is limited insofar as it does not provide information about differences in exploration, as this variable is standardised to the parameters of the foraging environment. Therefore, the use of both the leaving threshold variable and the deviation from the optimal leaving threshold variable are complementary, as they each provide information about rates of exploration and the *appropriateness* of such exploration behaviours, which can be used to address different research questions. I use these variables throughout the thesis to address predictions about adolescent development, age-related trends in social susceptibility and how early adversity affects adult decision-making.

### **2.3. Interim Conclusion**

I have described the methodological approach adopted in the current thesis. Specifically, I have described the theoretical and practical background to the computational methodology employed through the thesis. In addition, I have outlined key choices made in the modelling of participants' foraging behaviour, and the justification for distinguishing between participants' leaving thresholds and the degree to which their leaving thresholds deviate from a behavioural optimum. Using these methods, I will address the three key objectives of this thesis that test theories of decision-making across the lifespan and how this feature of cognition develops from adolescence to adulthood.

# **Chapter 3: Are Adolescents More Optimal Decision-Makers in Novel Environments? Examining the Benefits of Heightened Exploration in a Patch Foraging Paradigm<sup>2,3</sup>**

## **3.1. Abstract**

Adolescence is a period of heightened exploration relative to adulthood and childhood. This predisposition has been linked with negative behaviours related to risk-taking, including dangerous driving, substance misuse and risky sexual practices. However, recent models have argued that adolescents' heightened exploration serves a functional purpose within the lifespan, allowing adolescents to develop experiential knowledge of their surroundings. Yet, there is limited evidence that heightened exploration in adolescence is associated with positive outcomes. To address this, the present pre-registered study utilised a foraging paradigm with a sample of adolescents aged 16-17 ( $N=68$ ) and of adults aged 21 and above ( $N=69$ ). Participants completed a patch foraging task, which required them to choose between exploiting a known resource which gradually yields fewer rewards, and exploring a novel, unknown resource with a fresh distribution of rewards. Findings demonstrated that adolescents explored more than adults, which – in the context of the current task – represented more optimal patch foraging behaviour. These findings indicate that adolescents' heightened exploration can be beneficial, as they were able to effectively navigate unknown environments and accrue rewards more successfully than adults. This provides evidence that heightened exploration in adolescence, relative to adulthood, can lead to positive outcomes and contributes to our understanding of the role increased novelty-seeking plays at this point in the lifespan.

---

<sup>2</sup> This chapter is now published in *Developmental Science* as: Lloyd, A., McKay, R., Sebastian, C.L., & Balsters, J.H. (2020). Are adolescents more optimal decision-makers in novel environments? Examining the benefits of heightened exploration in a patch foraging paradigm. *Developmental Science*, e13075. doi: 10.31234/osf.io/ypknq

<sup>3</sup> OSF project page: <https://osf.io/nu627/>

### 3.2. Introduction

Adolescence is a period of heightened exploration of novel scenarios compared to childhood and adulthood (Spear, 2000; Steinberg, 2008). Adolescents' predisposition for novelty-seeking has typically been associated with negative outcomes in this population, such as increased rates of substance misuse, delinquency and injury (e.g. Arnett, 2002; Eaton et al., 2012; Moffitt, 2018). However, exploring novel opportunities is not exclusively associated with negative outcomes. For example, by trying out a new restaurant one can discover a new favourite meal, but this also involves exploring a possibility with unknown outcomes. Recent models have proposed that heightened exploration in adolescence, relative to adulthood and childhood, plays a functional role in development and is beneficial at this point in the lifespan (Romer, Reyna & Satterthwaite, 2017). The present study will examine whether adolescents' novelty-seeking can in fact produce more optimal navigation of novel environments.

While much of the literature has focussed on negative outcomes resulting from adolescents' decision-making, more recent research has considered positive outcomes that arise from novelty-seeking. Indeed, a proposed 'imbalance' between sensation-seeking and inhibitory control that occurs during adolescence (Casey, Getz & Galvan, 2008; Shulman et al., 2016; Steinberg, 2008) has been linked to 'positive' forms of risk taking, or forms that promote socially valued goals (Duell & Steinberg, 2019). For example, defending a peer who is being victimized can put an adolescent at risk of negative outcomes, but is also a positive, prosocial behaviour. Such acts of 'prosocial risk taking' recruit similar reward networks to those implicated in non-social forms of risk taking and demonstrate how neurobiological developments during this period can lead to positive behavioural outcomes (Do et al., 2017). Consistent with this, a recent study found that higher levels of sensation-seeking in adolescence predicted both positive and negative forms of risk taking equally (Patterson et al., 2019), suggesting that features of adolescent decision-making are not exclusively unfavourable.

The Lifespan Wisdom Model (Romer et al., 2017) builds on this by suggesting that adolescents' heightened novelty-seeking serves a functional role in development. As adolescence is a period marked by achieving autonomy from caregivers (Ellis et al., 2012), the Lifespan Wisdom Model proposes that adolescents' heightened predisposition for novelty-seeking, relative to other age groups, provides them with opportunities to gain experiential knowledge of their surroundings (Murty, Calabro & Luna, 2016). However, it is argued that a subset of adolescents do not adequately integrate their experience to inform future decisions and – as a result – persistently engage in risky behaviour. These adolescents are characterised by high levels of impulsivity, which drives them to seek immediate reward despite experience with negative outcomes associated with an action (e.g. persistent substance misuse; Khurana et al., 2015). Risk taking driven by impulsivity therefore does not support adolescents to gain experiential knowledge of their surroundings but is driven by a desire for short term rewards. For the majority of adolescents, however, exploration and novelty-seeking arguably support the transition from childhood dependence to adult independence.

Exploration is an effortful process requiring cognitive control (Daw, O'Doherty, Dayan, Seymour & Dolan, 2006; Otto, Skatova, Madlon-Kay & Daw, 2014). Recently, it has been suggested that adolescents are more motivated than other age groups to employ cognitive control to explore novel scenarios in order to acquire information about their surroundings (Do, Sharp & Telzer, 2019). Consistent with this, compared to adults, adolescents explore uncertain outcomes in conditions of ambiguity where probabilistic information is unknown, but not in conditions of risk where this information is explicit (Tymula et al., 2012; van den Bos & Hertwig, 2017). According to the Lifespan Wisdom Model, adolescents' ambiguity tolerance is due to the information that can be learned through sampling the unknown probabilities in these options (Romer et al., 2017). However, previous studies that have investigated ambiguity tolerance in adolescence have predominantly employed economic choice tasks (van den Bos

& Hertwig, 2017), which are restricted in their ability to measure exploration, as feedback does not have utility for later trials (Garrett & Daw, 2020). Furthermore, parameters measuring ambiguity tolerance are currently unable to distinguish whether this tolerance is driven by optimism that the lottery will yield a reward or by a motivation to learn the probability structure underlying the ambiguous option (Tymula et al., 2012). Therefore, there is limited evidence that adolescents' heightened ambiguity tolerance reflects an information gathering exercise.

The emphasis on exploration in the Lifespan Wisdom Model is noteworthy as exploration is utilised across species to maximise adaptive outcomes, such as those that support the health, development and safety of an organism (Ellis et al., 2012). For example, a number of animal species, including humans, utilise exploration to forage for resources such as food and mates (Cohen & Todd, 2018; Marques, Li, Schaak, Robson & Li, 2019). Indeed, human beings have spent the bulk of their evolutionary history in patch foraging economies (Smith, 1983; Venkataraman, Kraft, Dominy & Endicott, 2017). Patch foraging therefore exemplifies the significance of exploration in human decision-making. Deciding when to engage in exploration requires evaluating the trade-off between exploiting a known patch of land with gradually diminishing returns, and exploring a novel patch with a fresh distribution of supplies (i.e. the explore/exploit trade off; Stephen & Krebs, 1986). In order to maximise rewards and minimise energy costs, the decision-maker must learn when the opportunity cost of exploiting rewards from the current resource exceeds the opportunity cost of exploring a novel resource – at which point exploration is rational (Charnov, 1976).

Yet, recent lab-based studies have found that human adults are not optimal patch foragers (Constantino & Daw, 2015; Lenow et al., 2017; Le Heron et al., 2020). Constantino and Daw (2015) developed an apple picking patch foraging paradigm, requiring adults to maximise the number of apples collected from trees. Their paradigm had four 'environments' containing patches of different qualities. The authors predicted that participants would adjust

their foraging behaviour according to the richness of the respective environments to maximise rewards. Consistent with this, Constantino and Daw (2015) found that participants altered their leaving threshold according to the quality of the environments. However, throughout their task Constantino and Daw (2015) observed that participants consistently overexploited, staying with individual patches for longer than the optimal foraging strategy would dictate. Using a computational modelling approach, the authors demonstrated that overexploiting was explained by a parameter capturing participants' risk sensitivity, resulting in too much value being placed on the immediate rewards that could be exploited. Thus, despite the adaptive pressures to maximise reward while foraging, human adults do not explore optimally in patch foraging paradigms.

Two prominent theories have been developed to explain human exploration strategies. The first of these suggests that exploration is a stochastic process, driven by introducing randomness into decision-making that leads the agent to divert from the most rewarding option known to them (Daw et al., 2006). In an environment with changing reward contingencies, this strategy can accelerate the rate at which the decision-maker learns of changes in the structure of the environment (Gershman, 2018). The second, directed exploration, proposes that the decision-maker selects options to gain information about the underlying reward distribution of the environment, which increases the subjective value of the uncertain option in an 'information bonus' (Gershman, 2018; Wilson et al., 2014). The use of directed exploration emerges in adolescence, but is only utilised when exploration is consistent with the more rewarding option (Somerville et al., 2017). In cases where there is conflict between an option that has a higher reward value and another that has a higher information value, adolescents preferred the high reward option, whereas adults preferred the high information option (Somerville et al., 2017). This finding is consistent with evidence that, compared to other age groups, adolescents prefer immediate rewards in favour of delayed ones (van den Bos et al.,

2015). However, Somerville and colleagues (2017) also found that the use of random exploration did not change with development and overall, there were no developmental differences in task performance as measured by the number of points earned in the task. These findings are consistent with an emerging body of research has demonstrated that increasingly complex model-based exploration strategies develop in late adolescence (Dubois et al., 2020; Jepma et al., 2020; Xia et al., 2020), though these may be tempered by adolescents' preference for immediate reward relative to adults (Somerville et al., 2017).

To date, research has demonstrated that adolescents explore options with unknown information more than adults and children do. However, there is limited evidence that heightened exploration of novel scenarios in adolescence is associated with more optimal outcomes compared to other age groups. Here, we investigate exploration/exploitation trade-offs in adolescents and adults in an ecologically valid paradigm, patch foraging. Unlike other paradigms, patch foraging performance has been computationally defined across humans and other animal species such that performance can be compared to an objective behavioural optimum. While research has demonstrated that adolescents prefer to exploit the option that yields an immediate reward in an explore/exploit paradigm (Somerville et al., 2017), an alternative body of work has suggested that adolescents are motivated to explore novel scenarios more than adults (Do et al., 2019; Romer et al., 2017; Steinberg, 2008). As such, we predicted that 1) there will be a difference in foraging behaviour between adolescents and adults, measured by the point at which they leave patches to explore new ones.

To examine individual differences that contribute to foraging strategies, we also measured impulsivity and risk sensitivity. Impulsive traits are associated with a preference for immediate rather than delayed rewards (Khurana et al., 2015) and the prevalence of these traits varies across adult and adolescent populations (Chamorro et al., 2012; Romer, 2010). In a patch foraging context, we predict that heightened impulsivity will be associated with a greater

tendency to select the option that yields an immediate reward, which is to exploit the current patch. As such, we predicted that 2) higher levels of impulsivity will predict greater exploitation of patches. Finally, as overexploiting has been explained by a parameter capturing risk sensitivity, we predicted that 3) participants with higher levels of risk sensitivity will exploit individual patches to a greater degree.

### 3.3. Methods

#### 3.3.1. Design

The present study utilised a 2×2×2 mixed factorial design: Age (between-subject: adolescents, adults) x Depletion rate (within-subject: fast vs. slow patch depletion rate) x Travel time (within-subject: short vs. long travel time). Depletion rate and travel time variables are explained in greater detail in the Patch Foraging section below.

#### 3.3.2. Participants

*Pilot:* Seven participants were recruited for an initial pilot study: three adolescents aged between 16-18 ( $M = 17$ ,  $SD = 1$ ) and four adults aged between 24-26 ( $M = 24.66$ ,  $SD = 1.15$ ). Participants were recruited from a convenience sample who had previously expressed interest in research participation.

*Main study:* A total of 138 participants were recruited in the main study based on an a priori power calculation ( $\alpha = 0.05$ , power = 0.8, Cohen's  $f$  effect size = 0.25) using G\*Power 3.1 (Faul, Erdfelder, Lang & Buchner, 2007). Adolescent participants were aged 16-17, to capture the period in which novelty-seeking peaks (Romer, 2010). Sixty-eight adolescents were recruited in total, with a mean age of 16.57 ( $SD = 0.53$ , 77% female). The adult sample consisted of individuals aged 21 and above, as this is the period by which cognitive control is thought to have plateaued (Andrews-Hanna et al., 2011). During analysis,

it emerged that one participant was not in this age bracket, so this person was subsequently excluded. As such, the final adult sample was comprised of 69 participants aged between 21-50 with a mean age of 30.77 ( $SD = 7.96$ , 46% female). The study was approved by the Ethics Committee at the authors' institution.

### **3.3.3. Task**

*Patch Foraging:* The patch foraging task was adapted from a similar paradigm used in previous lab-based studies (Constantino & Daw, 2015; Lenow et al., 2017). The goal for participants was to collect as many apples as possible over the duration of the study. Patches that were available to forage were represented as single trees; on each trial participants could choose to stay and exploit a given patch (tree), or could leave to explore a new patch with a fresh distribution of apples. Exploiting a single patch would yield gradually diminishing rewards over time, such that the longer participants stayed with each patch, the fewer rewards they received from each exploit decision. When participants chose to exploit, they were presented with the number of apples they had collected on that harvest, alongside their cumulative score. Exploit decisions were scripted such that these would take three seconds including the decision phase and reward phase (i.e. when participants received the apples from harvesting). To ensure that faster reaction times did not permit participants to collect more apples within the time limit, the presentation of the reward phase was determined based on how quickly participants responded. For example, a participant who responded in 1.3 seconds would be presented with their score for 1.7 seconds. Participants were given up to two seconds to decide, after which they received a timeout message. Timeout trials were subsequently excluded from further analysis. If participants chose to explore, they would wait for a block-specific interval to elapse before arriving at a new patch, which had a fresh distribution of rewards available (see Figure 3.1).

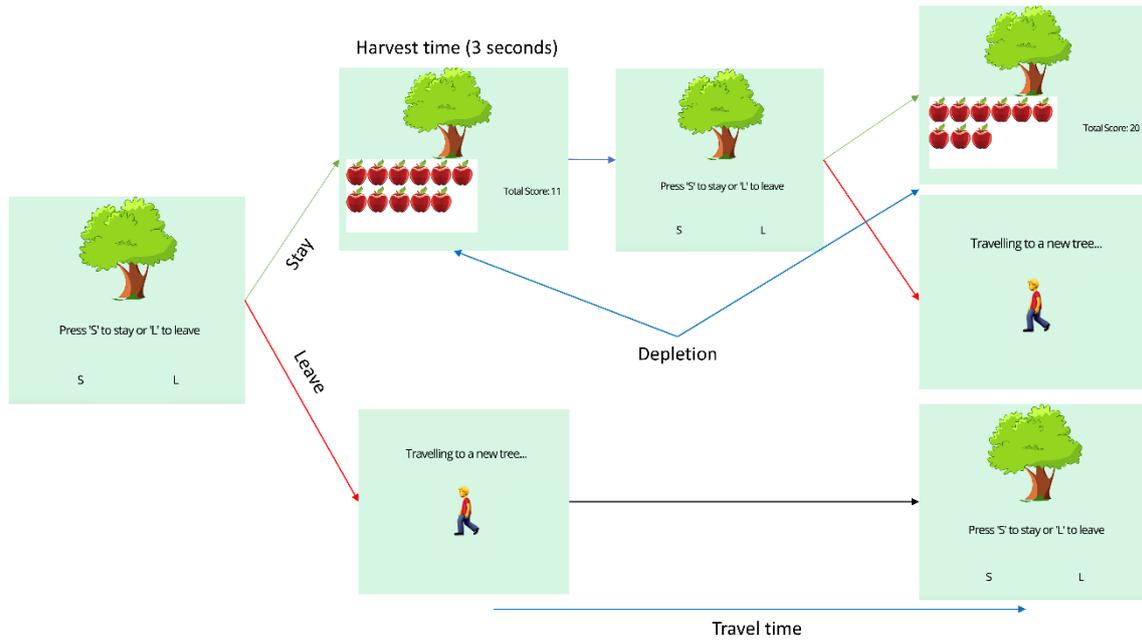


Figure 3.1: Schematic outline of the patch foraging paradigm

There were four unique environments in the present study, corresponding to four blocks. Participants completed the environments in counterbalanced order, to control for order effects. Parameters of the environments were manipulated to create environments that were either richer, yielding a higher average number of rewards over time, or poorer, yielding a lower average of rewards. Notably, participants did not know the features of each environment prior to encountering that block, requiring them to learn the state of each environment to maximise their average reward intake.

The two parameters that were manipulated in the present study were the rate at which rewards depleted in a single patch and the length of time it took to travel to a new patch. Depletion rate was manipulated such that the rate at which the apples diminished on each successive harvest was either fast or slow. As in Constantino and Daw (2015), the depletion rate applied after each exploit decision was randomly drawn from a Gaussian distribution. In environments with a fast depletion rate, the rewards from the previous harvest were multiplied by a figure drawn from a distribution with a mean of 0.88 ( $SD = 0.07$ ). In environments with a

slow depletion rate, the mean of the distribution was 0.94 ( $SD = 0.07$ ). To ensure that applying the depletion rate in this way did not lead to an increase in apples, the depletion rate was set to a maximum of 1, meaning the distribution of the depletion rate was negatively skewed. In environments where the depletion rate was slow, participants experienced a greater number of harvests with rich yields. As such, participants could accumulate a greater average number of rewards per unit of time compared to environments with a fast depletion rate.

The second parameter that was manipulated was the time it took to travel to a new patch if participants chose to explore. In environments with shorter travel times between patches, there is less cost in time (and thus resources) for the forager, meaning they can afford to explore more often to maximise average reward intake. However, in patches with longer travel times, there is a greater expenditure of time and resources. As such, given long travel times the forager should spend longer with each individual patch before they explore a new one.

Two dependent variables were calculated from the patch foraging task. The first was leaving threshold, i.e. the smallest reward (number of apples) that signalled to the participant the current patch had less value than exploring a novel patch (thus a high leaving threshold meant the participant left the patch earlier). The leaving threshold was operationalised as the average of the previous two rewards participants had observed in that patch prior to leaving, to provide an upper and lower bound estimate of participants' leaving threshold. We also included a second dependent variable, which was deviation from the optimal leaving policy. This was calculated through subtracting participants' actual leaving thresholds from the previously determined optimal leaving threshold in that environment (see Marginal Value Theorem below). As such, positive figures indicate participants overexploited the patch, whereas negative figures indicate participants underexploited that patch. This provided a measure of the degree to which participants deviated from the optimal foraging strategy.

*Marginal Value Theorem:* Optimal patch foraging behaviour is explained by an equation called Marginal Value Theorem (MVT), developed by Charnov (1976). This rule states that the forager should leave their current patch when the reward expected from the next exploit decision falls below the average reward experienced in that environment. To predict the value of exploiting the current patch, the forager should multiply the last known state of the tree ( $s_i$ ) by the block-specific depletion rate ( $\kappa$ ). This value is compared against the average reward rate per second ( $p$ ), which is multiplied by the time taken to harvest the patch in seconds ( $h$ ). If the expected value of exploiting the current patch ( $\kappa * s_i$ ) falls below the value of  $ph$ , the forager should explore the next patch. As our paradigm utilised discrete, rather than continuous trials, we utilised the MVT equation developed by Constantino and Daw (2015) to model foraging behaviour. Simplified, the equation states that the decision-maker should leave their current patch when the following inequality is satisfied:

$$\kappa s_i < ph \quad (1)$$

The MVT equation demonstrates that the optimal foraging strategy follows a leaving threshold rule (see Constantino & Daw, 2015, for a full proof of the MVT equation as it relates to this task). In order to estimate the optimum threshold for the parameters in each environment used in the current study, a simulation of the task was built in R (R Core Team, 2018). For each environment, the simulation was set to run with a leaving threshold from 1-10 in increments of 0.001. This produced 9001 simulations per environment and a total of 36,004 simulations for all four environments. The total score accrued using each of these leaving thresholds was then calculated, which was subsequently divided by the units of time to empirically demonstrate that the optimum leaving threshold for each environment followed the predictions of MVT (see Supplementary Material). For each environment, we ran each simulation ten times and averaged the optimum leaving threshold across these simulations to account for stochasticity

in the rewards accumulated due to the probabilistic nature of the depletion rate and initial richness of patches. For a list of features of each environment, see Table 1.

*Table 3.1: Parameters and optimal leaving thresholds of the four environments in the patch foraging task. The optimal leaving threshold is that which, if utilised by the participant, would lead them to collect the greatest number of apples in that environment.*

	Environment 1: fast depletion, short travel	Environment 2: slow depletion, short travel	Environment 3: fast depletion, long travel	Environment 4: slow depletion, long travel
Depletion rate ( <i>M, SD</i> )	0.88, 0.07	0.94, 0.07	0.88, 0.07	0.94, 0.07
Travel time (s)	6	6	12	12
Optimal leaving threshold – i.e. leave when the expected number of apples is lower than this	6.46	7.04	5.07	5.90

MVT also provides an explanation for changes to the optimum leaving threshold between environments that are richer or poorer quality. Environments with greater richness raise the value that must be expected from exploiting the current patch to make it more time-valuable than exploring a new resource-rich patch. For example, environments that have slower depleting patches will take longer to reach the environmental average compared to fast

depleting patches (Gabay & Apps, 2021). Consequently, in patches with a slower depletion rate, participants receive a higher yield of rewards for the time they invest in individual patches. This raises the average reward rate ( $p$ ), which, when entered into the MVT equation (1), signals a high leaving threshold. In fast depleting patches, participants receive a lower yield of rewards for the same time investment in patches, decreasing the average reward rate and signalling a lower leaving threshold. Following this rule, participants should remain for longer in environments with a slow depletion rate and should also have a higher leaving threshold compared to environments with a fast depletion rate (Gabay & Apps, 2021).

Similarly, in environments with shorter travel times between patches, the average rewards accumulated are higher, as less time is required to reach novel patches with a fresh distribution of rewards. This subsequently raises the average reward rate ( $p$ ) in environments with short travel times compared to long between-patch travel times. Therefore, the leaving threshold signalling when the forager should explore is higher when travel time between patches is shorter (Marshall, Carter, Ashford Rowcliffe & Cowlshaw, 2013). The predictions of MVT across different quality environments have been mathematically and empirically evidenced across a range of human and animal studies, providing a robust account of optimal foraging behaviour (Bettinger & Grote, 2016; Charnov, 1976; Stephen & Krebs, 1986)

#### **3.3.4. Questionnaire Measures**

*Demographic Measures:* Participants were asked to report the following demographic variables: age, gender (male or female), ethnicity and socioeconomic status. Ethnicity options were based on recommendations from the UK Office for National Statistics (Office for National Statistics, 2009). Participants that did not fall into one of these categories were asked to report their ethnicity in a text-entry box. Adolescents' socioeconomic status was approximated with a measure of parental education, with options ranging from no formal

qualifications to university degree (Steinberg et al., 2018). Adults' socioeconomic status was measured by their own highest level of education, comprised of the same options.

*Pubertal Development:* The Pubertal Status scale (Carskadon & Acebo, 1993) is a measure of adolescents' pubertal development. As development is subject to individual differences, this measure provides information about an individual's development that is distinct from chronological age. The scale is comprised of five questions which ask whether particular developmental milestones, such as their physical growth, are underway or complete. The individuals' pubertal development is calculated as a mean of their scores on the scale. The mean scores range from 1-4 with higher values denoting more highly developed individuals.

*Behavioural Inhibition and Behavioural Activation:* The BIS/BAS (Carver & White, 1994) is a combination of 20 items capturing behavioural inhibition scores (BIS) and behavioural approach scores (BAS). Higher average scores on the BIS subscale (ranging from 1-4) denote higher sensitivity towards non-rewarding or unpleasant stimuli, for example: "I worry about making mistakes". The BAS is divided into three subscales measuring drive, fun seeking and reward responsiveness. Higher average scores on the BAS subscales (ranging from 1-4 per subscale) indicate participants' responsiveness to rewarding stimuli, for example: "When I get something I want, I feel excited and energized". Together, these items measure participants' responsiveness to rewarding or aversive stimuli and predict risk sensitivity across development (Urošević, Collins, Muetzel, Lim, & Luciana 2012). The reliability of the subscales ranged from acceptable to good ( $\alpha = .66-.81$ ).

*Impulsivity:* The Barratt-11 (Patton, Stanford & Barratt, 1995) measures the degree to which the individual acts impulsively. The scale is comprised of 30 items, each presented on a Likert scale ranging from 1-4. The items capture the degree to which individuals act without thinking, or do not engage in self-regulation. The individual's impulsivity is calculated as their

mean score on the items in the scale. Mean scores range from 1-4, with higher values indicating greater impulsivity. The scale demonstrated good reliability ( $\alpha = .82$ ).

### **3.3.5. Procedure**

Participants were invited in groups to an IT suite where they were provided with either a laptop or desktop computer. Participants were briefed as to the aims of the study and provided with the opportunity to ask questions. Following this, participants were asked to complete a consent form, demographic information, and were provided with task instructions that specified participants should only focus on their own screen. Participants were informed that they would be awarded a gift card for taking part that would be loaded with a base reward of £4. They were also informed that a bonus of up to £3 could be earned, which was determined by the number of points they accumulated during the task. As such, they should aim to collect as many apples as possible throughout the task. All incentives were transparent, and no deception was used. The average amount awarded to participants was £5.62. The task was run on Gorilla.sc, an online platform for running behavioural experiments (Anwyl-Irvine, Massonnie, Flitton, Kirkham & Evershed, 2019).

Participants completed a practice task, which used different parameters to the main task. After the practice, participants completed each of the four unique patch foraging environments. Each environment was scripted to last for 7 minutes, which was determined based on previous research (Lenow et al., 2017). Once participants completed each environment, they were directed to a screen explaining that they were now in a new environment, which might have different properties to the one they had just completed. After the task ended, participants completed the BIS/BAS, the Barratt-11 and, for adolescent participants only, the Pubertal Status Scale. Participants were then provided with their final score and debriefed as to the aims of the study.

## 3.4. Results

### 3.4.1. Confirmatory Analyses

Analyses were conducted in R v3.5.1 and JASP v0.10.2 (JASP Team, 2019). Results of the pilot study indicated that participants' behaviour was qualitatively consistent with the predictions of MVT. Descriptive statistics indicated that participants adjusted their leaving threshold according to the different qualities of the environment, which was considered adequate support to run the main study.

Results of the main analyses demonstrated that across the different quality environments, adolescents consistently explored more compared to adults. A mixed 2 (age group: adolescents vs adults)  $\times$  2 (patch depletion rate: fast vs slow)  $\times$  2 (travel time: short vs long) ANOVA demonstrated a main effect of age on the leaving threshold  $F_{(1, 122)} = 13.56, p < .001, \eta^2 = .10, BF_{01} = 2.117e-7$ . Compared to adults, adolescents had a higher average leaving threshold, meaning the number of apples required to stay at a patch was higher and thus they explored more often. Across the four environments, adolescents' mean leaving threshold was 5.82 ( $SD = 1.79$ ) apples, whereas adults' mean leaving threshold was 4.78 ( $SD = 1.66$ ) apples (see Figure 3.2). Furthermore, results of this analysis revealed a main effect of travel time  $F_{(1, 122)} = 15.57, p < .001, \eta^2 = .01, BF_{01} = 6.985e-7$ . This suggests that participants explored more in environments where travel times between patches was short compared to environments when travelling between patches incurred a greater time cost, consistent with a rational foraging strategy. Further, there was a main effect of changes to the depletion rate between environments  $F_{(1, 122)} = 32.28, p < .001, \eta^2 = .01, BF_{01} = 2.936e-5$ . In environments where the patch depletion rate was fast participants had a lower leaving threshold, whereas in environments where the patch depletion was slow, participants had a higher leaving threshold, which is consistent with optimal foraging behaviour. However, there was an interaction between age group and patch depletion rate  $F_{(1, 122)} = 6.13, p = .015, \eta^2 = .002, BF_{01} = 0.002$ . Post hoc tests using the

Bonferroni correction demonstrated that there was not a significant difference between adolescents' leaving thresholds in the fast depletion conditions compared to the slow depletion conditions ( $p_{bonf} = .172$ ). However, there was a significant difference between adults' leaving thresholds in the slow depletion conditions compared to the fast depletion conditions ( $p_{bonf} < .001$ ). These findings indicate that adults were more responsive to changes in the depletion rate compared to adolescents and adjusted their leaving thresholds accordingly. As this was an ordinal interaction, it was still valid to interpret the main effects of this model. Results of the ANOVA revealed no interaction effect between travel time and depletion rate  $F_{(1,122)} = 0.91$ ,  $p = .343$ , age group and travel time  $F_{(1,122)} = 1.149e-9$ ,  $p = .991$ , nor between age group, depletion rate and travel time  $F_{(1,122)} = 1.26$ ,  $p = .264$ .

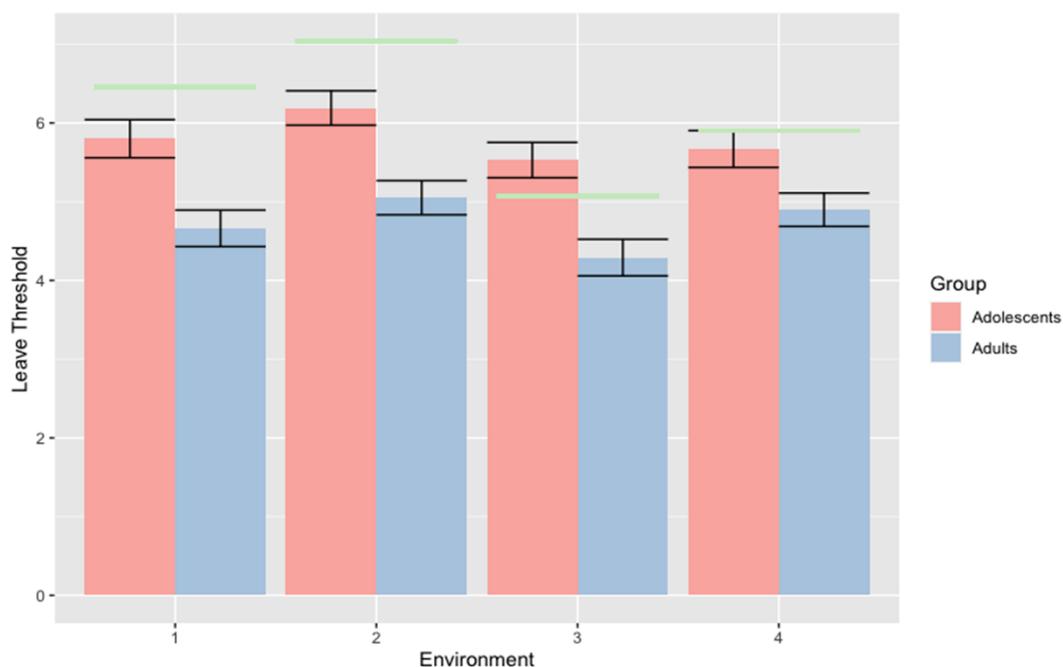


Figure 3.2: Plot of the mean leaving threshold for adolescents and adults for Environment 1 (fast depletion and short travel), Environment 2 (slow depletion and short travel), Environment 3 (fast depletion and long travel) and Environment 4 (slow depletion and long travel). Green bars indicate the optimal leaving value for that environment

Although adolescents' leaving threshold was higher than adults', this does not confirm that their behaviour was more optimal. To address this, we conducted a mixed 2x2x2 ANOVA

using average deviation from the optimum leaving threshold as the dependent variable. Results demonstrated a main effect of age on deviation from the optimum  $F_{(1, 122)} = 13.56, p < .001, \eta^2 = .08, BF_{01} = 1.324e-27$ , with adolescents' behaviour closer to the optimum leaving threshold compared to adults'. There was a main effect of travel time on deviation from the optimum  $F_{(1,122)} = 135.43, p < .001, \eta^2 = .06, BF_{01} = 2.472e-4$ , with participants being closer to the optimum leaving threshold in the long travel time environments. In addition, there was a main effect of depletion rate on deviation from the optimum  $F_{(1,122)} = 23.76, p < .001, \eta^2 = .01, BF_{01} = 6.426e-28$ . Participants were closer to the optimum leaving threshold in environments with a fast depletion rate compared to environments with a slow depletion rate. There was also a significant interaction effect between depletion and age group  $F_{(1,122)} = 6.13, p = .015, \eta^2 = .002, BF_{01} = 4.393e-26$ . Post hoc tests using the Bonferroni correction demonstrated that adults' deviation from the optimum did not change significantly between fast and slow depleting environments ( $p_{bonf} = .508$ ), whereas adolescents were significantly less optimal in slow depleting environments compared to fast depleting environments ( $p_{bonf} < .001$ ). As adolescents did not adjust their leaving threshold to the same extent as adults in response to changes to the depletion rate, they exhibited greater suboptimality in the slow depletion environments compared to the fast depletion environments. However, there were no interactions between travel time and depletion rate  $F_{(1,122)} = 0.36, p = .552$ , travel time and age group  $F_{(1,122)} = 1.149e-4, p = .991$ , nor travel time, depletion rate and age group  $F_{(1,122)} = 1.26, p = .264$  (see Table 3.2).

Consistent with these findings, adolescents accumulated a higher average number of apples in each environment ( $M = 650.82, SD = 99.50$ ) compared to adults ( $M = 626.71, SD = 95.05$ ). However, this difference was not statistically significant as demonstrated by a mixed ANOVA  $F_{(1, 130)} = 1.04, p = .310$ , though the Bayes factor demonstrated extreme evidence for the alternative hypothesis (i.e. that adolescents accumulated a higher average number of

rewards compared to adults)  $BF_{01} = 4.342e-86$ . This inconsistency could be attributed to the probabilistic nature of the reward schedule. Consistent with MVT, there was a main effect of depletion rate  $F_{(1,130)} = 216.27, p < .001, BF_{01} = 2.280e-67$  and travel time  $F_{(1,130)} = 443.62, p < .001, BF_{01} = 9.617e-39$ , with participants accumulating a higher number of rewards in richer environments compared to poorer environments (see Table 3.2). In addition, there was an interaction effect between travel time and depletion rate  $F_{(1,130)} = 13.72, p < .001, BF_{01} = 0.027$ , but no interaction between depletion rate and age group  $F_{(1,130)} = 3.26, p = .073$ , travel time and age group  $F_{(1,130)} = 0.02, p = .890$ , nor depletion rate, travel time and age group  $F_{(1,130)} = 0.75, p = .389$ .

Table 3.2: Mean (SD) of participants' deviation from the optimum and number of apples accumulated in each environment

	Adolescent deviation from the optimum	Adult deviation from the optimum	Adolescent number of apples accumulated	Adult number of apples accumulated
Environment 1	0.66 (1.98)	1.80 (1.89)	675.01 (105.35)	652.36 (110.64)
Environment 2	0.85 (1.72)	1.99 (1.76)	724.60 (135.12)	711.77 (108.73)
Environment 3	-0.46 (1.85)	0.78 (1.95)	588.77 (91.59)	529.94 (76.45)
Environment 4	0.23 (1.82)	1.00 (1.74)	633.13 (113.16)	629.25 (105.12)
Total	0.30 (1.84)	1.37 (1.84)	650.82 (99.50)	626.71 (95.05)

To examine whether patch foraging behaviour was predicted by any of the disposition measures, we conducted a linear regression. When the BAS subscales, the BIS and the Barratt-11 were all entered as predictors of participants' deviation from the optimum threshold, the model was not significant  $F_{(5,132)} = 1.05, p = .393, R^2 = 0.04$ , nor were any individual predictors significant ( $ps > .05$ ). Following this, a second regression model was conducted with the addition of interaction terms to account for the interaction between age group and the disposition measures. Consistent with our previous analysis, none of the predictors were significant and we did not find evidence of an interaction effect between participants' age group and any of the disposition measures ( $ps > .05$ ).

### **3.4.2. Exploratory Analyses**

As the gender split between the adult and adolescent sample was imbalanced, we first ran our model including gender as a covariate in a mixed 2x2x2 ANCOVA. Following this, we entered gender as an additional fixed factor into a 2x2x2x2 ANOVA to examine whether foraging performance differed as a main effect of gender. These analyses suggested gender did not appreciably affect the results, as foraging behaviour differed between age groups but not genders (see Supplementary Material).

## **3.5. Discussion**

### **3.5.1. Adolescents Exhibit Heightened Exploration During Patch Foraging**

Although risk and novelty-seeking during adolescence have often been associated with negative consequences, recent theories have proposed that these tendencies can in fact be beneficial. The present study utilised a realistic decision-making problem that is almost universal across animal species: patch foraging. Adolescents had higher leaving thresholds across all four patch foraging environments, which meant this group explored more than adults.

This behaviour was also closer to the optimal foraging behaviour as defined computationally using MVT. In line with previous research (Constantino & Daw, 2015; Le Heron et al., 2020; Lenow et al., 2017), our results demonstrated that adults remained with individual patches for longer than the optimal strategy would dictate. These findings indicate that heightened exploration during adolescence can be beneficial when exploration is associated with rewarding outcomes, which can support this age group to gain experiential knowledge of their surroundings.

The present findings extend upon the current patch foraging literature, which has previously demonstrated that human adults are suboptimal patch foragers (e.g. Le Heron et al., 2020). Whilst multiple studies have shown that human adults overexploit patches, our findings indicate that adolescents did not demonstrate this pattern of overexploiting to the same degree as adults, meaning their behaviour was closer to the optimal patch foraging strategy.

Adolescents' predisposition for exploration has been attributed to the hyperactivity of the dopaminergic reward system, motivating this age group to seek novel sensations and experiences (van Leijenhorst et al., 2010b). In the current patch foraging paradigm, new patches represented novel opportunities, as each tree was only visited once. Adolescents' behaviour was therefore consistent with a pattern of heightened novelty-seeking relative to adult participants. Importantly, however, adolescents' predisposition for exploration was not indiscriminate, as they adjusted their leaving thresholds according to changes in the inter-patch travel time between the different environments. While the difference was not statistically significant, we also found that adolescents adjusted their leaving thresholds according to changes in the depletion rate. Adolescents' suboptimality between conditions with fast and slow depletion rates was primarily driven by overexploiting in the richest quality environment (environment 2), which had the highest optimal leaving threshold. However, in environment 4, which also had a slow depletion rate, adolescents were close to the optimal leaving threshold.

As such, the finding that adolescents did not adjust their behaviour to changes in the depletion rate could be due to the high leaving threshold required in the richest environment, rather than an insensitivity towards this change in the environment.

Heightened novelty-seeking in the second decade of life has also been associated with the increased rates of risk behaviour observed in adolescence relative to other age groups (Steinberg, 2008). In the context of patch foraging, previous research has demonstrated that adults' tendency to overexploit is explained by a parameter capturing risk sensitivity (Constantino & Daw, 2015). That adolescents explored more than adults could indicate that heightened levels of novelty-seeking were associated with diminished risk sensitivity in this age group, but this tolerance for risk led to more optimal outcomes compared to risk-averse adults. In contrast to recent research (Bach, Moutoussis, Bowler, Neuroscience in Psychiatry Network consortium & Dolan, 2020), we did not find gender differences in foraging behaviour in a developmental sample. The current study suggests that when exploration is associated with opportunity, adolescents' propensity for heightened exploration can be beneficial. In scenarios where exploration can lead to negative outcomes (as in some behavioural tasks e.g. the Balloon Analogue Risk Task; Lejuez et al., 2002), the same predisposition for exploration can lead to suboptimal outcomes (Peper, Braams, Blankenstein, Bos & Crone, 2018). Taken together, these findings demonstrate how adolescents' novelty-seeking can lead to both positive and negative outcomes for this age group (Patterson et al., 2019).

These findings build on evidence of developmental differences in ambiguity tolerance. Compared to other age groups, adolescents are more tolerant to variable outcomes and thus riskier choices where the probabilities involved in the decisions are unknown (van den Bos & Hertwig, 2017). In the present patch foraging paradigm, participants were not aware of the distribution of rewards in novel patches and therefore exploration required tolerance of an ambiguous outcome. Adolescents invested less time in the known option, instead choosing to

explore the unknown, ambiguous opportunity. Through using a serial decision-making task, our findings contribute evidence that adolescents are more motivated than adults to explore ambiguous options and that this can lead to more optimal outcomes.

The use of complex exploration strategies develops in later adolescence and remains into adulthood, which can support individuals to effectively navigate their environment (Jepma et al., 2020; Somerville et al., 2017). In contrast to our findings, previous work has demonstrated that adolescents prefer to exploit a known option for an immediate reward rather than explore a novel one to gain information about their surroundings (Somerville et al., 2017). However, in this study, adolescents only preferred the high reward option when this was incongruent with the high information option. Further work examining the development of explore/exploit strategies has found that adolescents explore more than is optimal, leading them to accrue fewer rewards than adults (Xia et al., 2020). In the present patch foraging paradigm, relative to adults' performance, heightened levels of exploration were congruent with the high reward option, as this raised the average reward rate according to MVT.

When examining the development of strategic exploration, previous research has demonstrated that directed exploration emerges in late adolescence and stabilises in adulthood, whereas random exploration is age-invariant (Somerville et al., 2017). Random exploration involves introducing stochasticity into one's decisions, which can lead the agent to select the option perceived to have the lower reward (Gershman, 2018). While this exploration strategy can be beneficial for uncovering information about the environment (Wilson et al., 2014), it is less precisely tuned to the structure of that environment (Somerville et al., 2017). The use of random exploration strategies in adolescence may have contributed to the finding that this age group adjusted their leaving threshold to a lesser extent in response to changes in the depletion rate, as their behaviour was influenced by a stochastic process that was less finely calibrated to the structure of the environments.

Taken in context of previous studies, our findings highlight that adolescents' reward seeking behaviour and heightened propensity for exploration can complement one another when exploration is associated with reward. This may afford adolescents opportunities that are forfeited by other age groups that are averse to uncertain outcomes (for example, trialling out for a new sports team or going to a summer camp). As such, our findings contribute to understanding when adolescents' heightened exploration can lead to positive outcomes for this age group.

### **3.5.2. The Benefits of Exploration in Adolescence**

Heightened exploration during adolescence supports this age group to gain experience of novel scenarios, which can be utilised to guide future decisions (Romer et al., 2017). At the onset of puberty, there is increased activation of the dopaminergic reward network (van Leijenhorst et al., 2010b), which promotes novelty-seeking behaviour (Costa, Tran, Turchi & Averbeck, 2014). Further, adolescents are motivated to employ cognitive control resources to explore unknown scenarios that can yield information about the probability structure of their environment (Do et al., 2019). Consistent with this account, adolescents in the present study were motivated to seek out novel patches, which required the deployment of cognitive resources to engage in effortful exploration rather than choosing the comparatively less taxing option of exploiting the known option (Daw et al., 2006; Do et al., 2019; Otto et al., 2011). According to the Lifespan Wisdom Model, this decision-making strategy is employed across adolescence to aid adolescents to learn about their surrounding environment (Romer et al., 2017). The information gained through foraging is then subsequently exploited later in life (Mata, Wilke & Czienskowski, 2013). Findings from the present study support this account, as when adolescents were faced with an unknown environment, these participants were motivated to explore in order to effectively navigate this novel scenario compared to adults. Heightened exploration in adolescence may therefore be considered adaptive within the lifespan, as it

supports the development of the individual's experiential knowledge in unknown scenarios (Ellis et al., 2012).

There are limitations to the present study that are important to consider. This study only measured exploration in relation to positive reward feedback. Previous research has demonstrated that adolescents respond differently to positive and negative feedback (McCormick & Telzer, 2017). Future studies should consider whether adolescents are able to utilise optimal decision-making strategies when exploring novel scenarios to avoid negative outcomes, as well as acquire rewards. Further, the environmental parameters utilised in the present study were more suited to strategies that required higher levels of exploration rather than exploitation. As such, adolescents' more optimal performance could be due to the structure of the environments rather than their ability to effectively navigate patches. However, we note that adolescents were more optimal in environments with fast depletion rates, where the optimal strategy is to explore less, suggesting this age group were better able to estimate the optimal leaving threshold when less exploration was required. Nevertheless, future research could consider designing environments with a wider range of environmental parameters to test whether adolescents retain their optimality in foraging scenarios that favour less exploration.

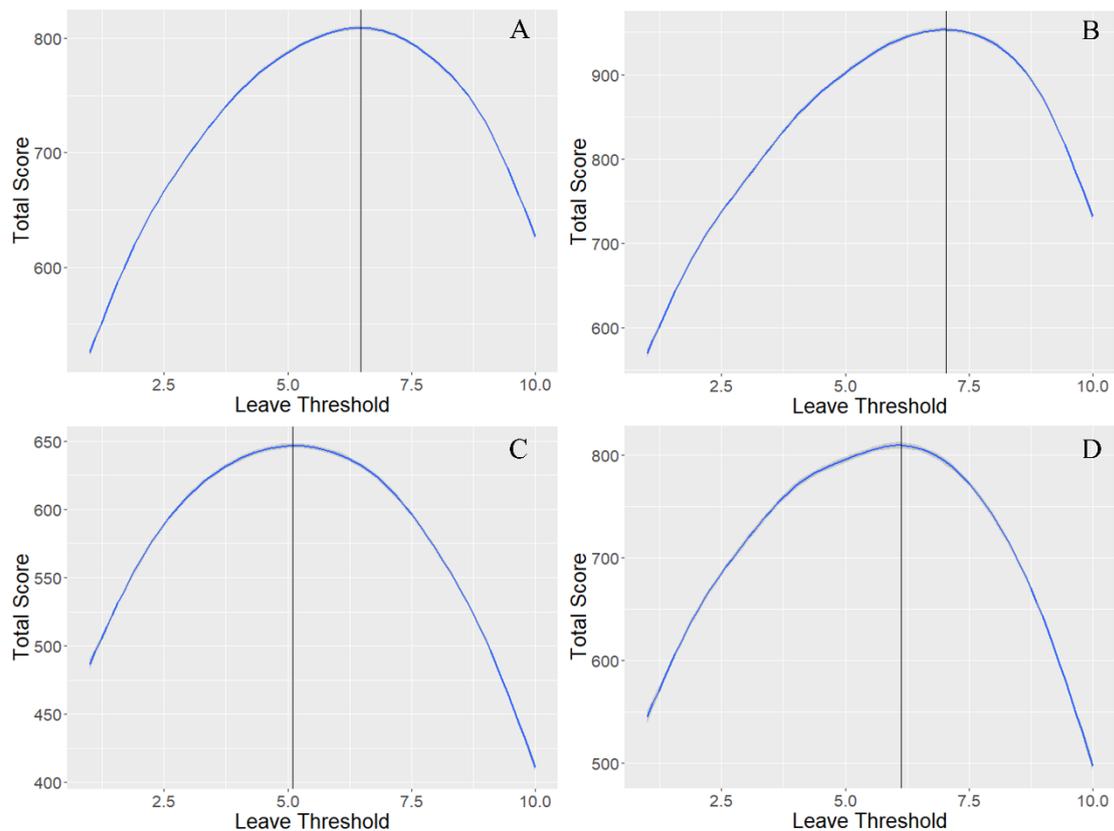
This study has implications for our understanding of the role of adolescence within the lifespan. Our findings indicate that exploration plays a key role at this developmental stage and can support individuals to benefit from novel opportunities available to them. As such, priority should be given to facilitating adolescents' exploration and providing them with the opportunity to learn in unfamiliar scenarios. The importance of allowing developmentally-appropriate levels of exploration and learning are currently highlighted in safeguarding documents for young people (NHS England, 2015). Findings of the present study support these guidelines as adolescents require exposure to novel experiences in order to develop knowledge of their surroundings.

Overall, findings of the present study suggest that compared to adults, adolescents exhibit heightened rates of exploration in novel, unknown environments. Contributing to the patch foraging literature, we replicate findings that adults overexploit patches compared to the optimal foraging strategy. Our findings suggest that when exploration is associated with reward, adolescents' heightened propensity for exploration and reward-driven behaviour can be beneficial, leading this age group to exhibit more optimal behaviour relative to adults. This strategy of novelty-seeking behaviour can allow this age group to benefit from novel opportunities and gain experience at the developmental stage where the individual transitions from childhood dependence to adult independence.

## 3.6. Supplementary Materials

### 3.6.1. Simulations

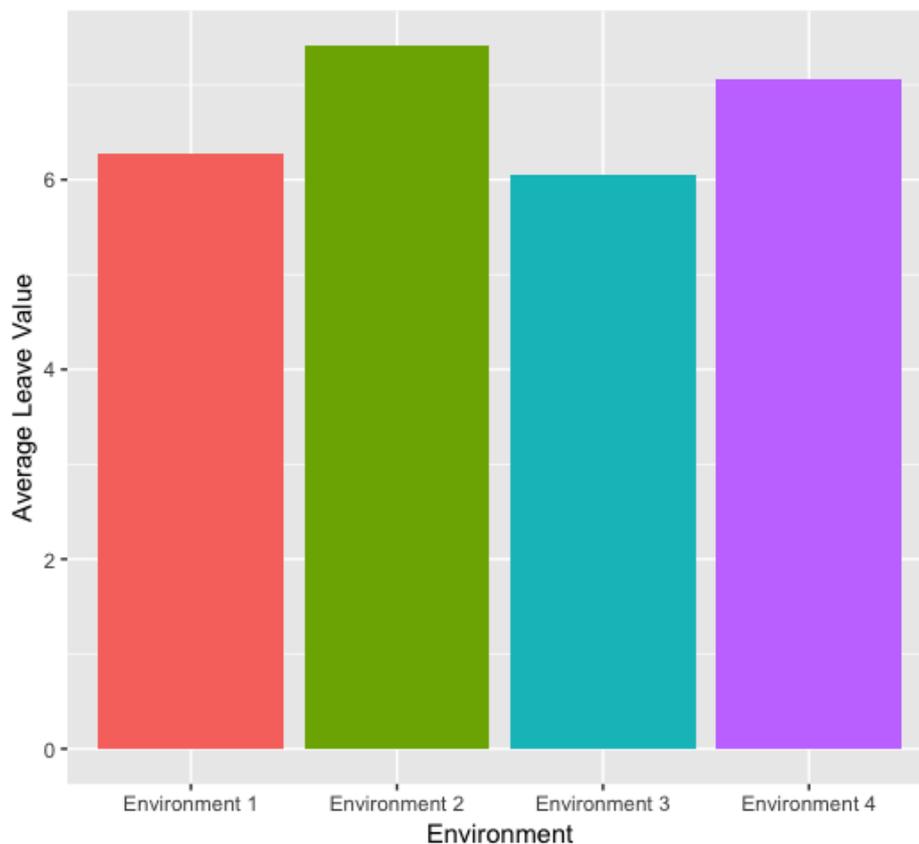
The following Figure (Supplementary Figure 3.1.) details the simulation results for our four patch foraging environments. The peak of the curve indicates the leaving threshold that maximises the average reward rate ( $p$ ), providing an empirical demonstration of the optimal leaving threshold for that environment.



*Supplementary Figure 3.1: Plots demonstrating the leaving threshold that led the simulated agent to accumulate the greatest number of rewards in each environment. Plot A corresponds to Environment 1 which had a fast depletion rate and short travel time. The optimal leaving threshold in this environment was 6.46. Plot B corresponds to Environment 2 which had a slow depletion rate and short travel time. The optimal leaving threshold for this environment was 7.04. Plot C corresponds to Environment 3 which had a fast depletion rate and long travel time. The optimal leaving threshold for this environment was 5.07. Plot D corresponds to Environment 3 which had a slow depletion rate and long travel time. The optimal leaving threshold for this environment was 5.90.*

### 3.6.2. Pilot Results

We collected pilot data from 7 participants (three adolescents and four adults) and plotted their leaving thresholds across the four environments. Consistent with the predictions of Marginal Value Theorem (MVT), participants had a lower leaving threshold in poorer quality environments (e.g. Environment 3 which had a fast depletion rate and long travel time) compared to richer environments (e.g. Environment 2, which had a slow depletion rate and short travel time). As participants behaved in a manner qualitatively consistent with MVT, we considered this evidence that adolescent and adult participants could understand the task sufficiently to run the main study.



*Supplementary Figure 3.2: Mean leaving threshold across the four environments for adult and adolescent participants*

### 3.6.3. Additional Exploratory Analyses

As the gender split between the adult and adolescent sample was imbalanced, we included gender added as a covariate in a mixed 2x2x2 ANCOVA. Results replicated a main

effect of age on deviation from the optimum  $F_{(1, 118)} = 14.52, p < .001$ , with adolescents' behaviour closer to the optimum leaving value compared to adults'. There was a main effect of travel time on deviation from the optimum  $F_{(1,118)} = 4.50, p = .036$ , with participants being closer to the optimum leaving threshold in the long travel time environments. In addition, there was a main effect of depletion rate on deviation from the optimum  $F_{(1,118)} = 11.02, p < .001$ . Participants were closer to the optimum leaving threshold in environments with a fast depletion rate compared to environments with a slow depletion rate. There was also a significant interaction effect between depletion and age group  $F_{(1,118)} = 7.74, p = .006$ . However, there were no interactions between travel time and age group  $F_{(1,122)} = 0.20, p = .655$ , nor travel time, depletion rate and age group  $F_{(1,122)} = 1.18, p = .279$ . These findings indicate that the gender imbalance between the adult and adolescent samples did not appreciably affect the results.

In addition, we ran an  $2 \times 2 \times 2 \times 2$  ANOVA with factors for gender (male/female), age group (adolescent/adult), depletion rate (fast/slow) and travel time (long/short) to determine whether there was a main effect of gender on foraging behaviour. Results of this analysis demonstrated that there was not a main effect of gender on leaving threshold  $F_{(1,115)} = 1.05, p = .307$ . However, the main effects of travel time  $F_{(1,115)} = 13.39, p < .001$ , depletion rate  $F_{(1,115)} = 28.70, p < .001$  and age group  $F_{(1,115)} = 15.29, p < .001$  were preserved. In addition, there was still an interaction effect between depletion rate and age group  $F_{(1,115)} = 5.78, p = .018$ . Interaction effects were not found for: travel time and gender  $F_{(1,115)} = 3.12, p = .08$ , travel time and age group  $F_{(1,115)} = 0.08, p = .78$ , travel time, gender and age group  $F_{(1,115)} = .27, p = .607$ , depletion rate and gender  $F_{(1,115)} = 1.50, p = .223$ , depletion rate, gender and age group  $F_{(1,115)} = 2.22, p = .139$ , travel time and depletion rate  $F_{(1,115)} = .96, p = .330$ , travel time, depletion rate and gender  $F_{(1,115)} = .002, p = .961$ , travel time, depletion rate and age group  $F_{(1,115)} = .87, p = .354$  nor travel time, depletion rate, gender and age group  $F_{(1,115)} = .16, p = .687$ .

To examine whether reaction time (RT) differed between age groups of environments, which could indicate waning interest, we conducted a mixed ANOVA with RT entered as the dependent variable. Results demonstrated that there was not a main effect of age  $F_{(1,130)} = 2.01$ ,  $p = .159$ , depletion rate  $F_{(1,130)} = 0.46$ ,  $p = .500$ , travel time on RT  $F_{(1,130)} = 0.11$ ,  $p = .744$ . In addition, there were no interactions between depletion rate and age group  $F_{(1,130)} = 1.32$ ,  $p = .253$ , travel time and group  $F_{(1,130)} = 0.95$ ,  $p = .332$ , depletion rate and travel time  $F_{(1,130)} = 0.74$ ,  $p = .392$  nor depletion rate, travel time and age group on RT  $F_{(1,130)} = 0.74$ ,  $p = .391$ . These results suggest that levels of interest or vigor did not systematically differ between age groups or environments.

#### **3.6.4. Post-hoc Power Analyses**

To test whether our sample had sufficient power to test our hypotheses, we conducted post-hoc power analyses. For the results of our ANOVA analyses, effect sizes given in  $\eta^2$  were first transformed into Cohen's  $f$  (Cohen, 1992). The effect size ( $f = 0.32$ ), sample size (138), number of groups (2) and measurements (4) were entered into G\*Power, which produced a power estimate of 99.7%. Following this, we conducted post-hoc power analyses for our main effect of travel time and depletion rate. The effect size ( $f = 0.1$ ), sample size (138), number of groups (2) and number of measurements (2) were entered which resulted in a power estimate of 64.21%. We therefore advise caution in interpreting the main effect of depletion rate and travel time on participants' behaviour. However, we note that these findings are consistent with a significant body of previous research (Constantino & Daw, 2015; Le Heron et al., in press; Lenow et al., 2017).

## **Chapter 4: Adolescents Utilise a Faster Learning Rate Whilst Foraging Compared to Adults**

### **4.1. Abstract**

Adolescence is a period of increasing independence, during which the individual must learn to negotiate their surroundings. The ability to learn about one's environment can be explained by reinforcement learning (RL) frameworks, where the individual uses reward feedback to update their expectation of outcomes associated with decisions. There are mixed findings with regards to how much emphasis adolescents place on recent versus more historic feedback relative to adults, measured in RL models as the learning rate. These inconsistencies may be reconciled by examining whether adolescents have a single learning rate for all outcomes or show asymmetry in their learning rate for better-than-expected and poorer-than-expected outcomes, which we compare in the present study. Our findings demonstrate that the single learning rate model provided a more parsimonious explanation of participants' exploration behaviour. Moreover, we find that adolescents are faster to integrate feedback into their decision-making while foraging, as they have a higher learning rate compared to adults. We consider these findings in the context of adolescence as a period of transition and how these learning strategies may support this age group to gain experiential knowledge that can support them to achieve independence.

## 4.2. Introduction

Adolescence is a developmental period during which the individual acquires the skills and experience necessary for adult independence (Ellis et al., 2012). Recently, it has been proposed that adolescents explore their environment more compared to adults and children as this provides them with opportunities to learn about their surroundings and gain experiential knowledge of their environment (Do et al., 2020; Romer et al., 2017). How individuals learn from exploring their surroundings can be explained by reinforcement learning (RL) frameworks, where the individual updates their expectation about the relative value of different actions based on the reward feedback from trialling these actions (Sutton & Barto, 2018). In the present study, we examined age-related differences in how adolescents and adults learn from reward feedback and how this ability may support adolescents when foraging from their environment. These findings have implications for our understanding of the role of exploration and learning in adolescent development.

To examine the development of reinforcement learning strategies across the lifespan, several studies have utilised tasks that manipulate the probability that stimuli (e.g., different coloured boxes or flowers; Davidow et al., 2016; Xia et al., 2021), if selected by the participant, will yield a reward. Successful participants must learn the reward probabilities associated with each stimulus through trialling each option and integrating the reward feedback (i.e., whether they received a reward or not) into their expected value of an option. The difference between the expected reward and the one received is known as ‘prediction error’, which can be both positive, when the outcome is better than expected, or negative, when the outcome is poorer than expected (Niv & Schoenbaum, 2008). The degree to which prediction errors are integrated into one’s expectation that a stimulus will produce a reward is controlled by one’s learning rate (typically denoted as  $\alpha$ ; Sutton & Barto, 2018). Higher learning rates suggest the individual relies on more recent feedback when estimating the reward value of available options whereas

lower values indicate the individual utilises more historic feedback in their decision-making. Individuals' learning rate can be measured with separate values for positive and negative prediction errors or measured with a single value for all prediction errors (e.g., Garrett & Daw, 2020). However, there is conflicting evidence regarding whether adolescents and adults utilise a single learning rate for all outcomes or have asymmetric learning rates for positive and negative prediction errors (Gershman, 2015; Jepma et al., 2020).

Moreover, there are gaps in our understanding of changes to RL parameters across the lifespan and the task conditions under which these age-related differences emerge. Previous studies have found that adolescents outperformed adults when learning the association between stimuli and outcomes, which allowed them to collect more rewards than adults (Davidow et al., 2016). Upon the receipt of a positive prediction error, adolescents exhibited heightened activation of neural regions responsible for encoding episodic memories (e.g., the hippocampus; Gabrieli, 1998) and regions responsible for processing reward (e.g., the ventral striatum; van den Bos et al., 2012) relative to adults. Moreover, adolescents' heightened neural activation, relative to adults, corresponded to lower learning rates for reward feedback. In contrast, adolescents exhibited similar levels of activation in these regions to adults for negative prediction errors (Davidow et al., 2016). Yet, several other studies have found that learning rates for positive prediction errors increase from adolescence to adulthood (e.g., Eckstein et al., 2020; Masters et al., 2020; Xia et al., 2020). Recently, it has been suggested that inconsistencies in findings about the development of the learning rate from adolescence to adulthood may be attributable to task differences, such as the degree of uncertainty in the structure of the task which can affect whether the use of a higher or lower learning rate is more optimal (Nussenbaum & Hartley, 2019). As such, the use of different tasks to measure computational parameters means that variables, such as the learning rate, do not generalise across RL tasks (Eckstein et al., 2021).

Given that adolescence is a time of heightened exploration and novelty-seeking (Romer et al., 2017; Steinberg, 2008), a task that directly measures exploration behaviours may be better suited to studying this developmental period. Here, we used a patch foraging task, which requires the organism to evaluate the trade-off between exploiting a patch they are located at, which gradually returns fewer rewards over time, and exploring a novel patch with a fresh distribution of rewards. It has been suggested that brain structures involved in processing reward feedback have, in part, evolved to optimise foraging behaviour (Hunt et al., 2021; Kolling & Akam, 2017; Niv et al., 2002; Passingham & Wise, 2012). During adolescence, there are significant developments in neural regions responsible for reward processing (e.g., the ventral striatum; Shulman et al., 2016; Van Leijenhorst et al., 2009). These developments may affect foraging strategies and it has been proposed that it is adaptive for younger organisms to utilise a highly exploratory foraging strategy to learn from their surroundings, whereas older organisms should utilise their experience to exploit bounteous patches more extensively (Mata et al., 2013). This proposal has empirical support, as one study found adolescents to explore more than adults while foraging (Lloyd et al., 2021b). Therefore, examining age-related changes in reward learning, and how these link to foraging strategies, may provide novel insight into the normative role of reward learning during development.

The organism's ability to forage optimally is important as the resources collected through this behaviour are directly related to the organism's adaptive needs, such as the collection of food for sustenance (Bettinger & Grote, 2016). According to a prominent model of patch foraging known as 'Marginal Value Theorem' (MVT), the forager should leave to explore a new patch when the rewards expected from exploiting their current patch fall below the average reward rate for the environment (Charnov, 1976). MVT has provided a robust account of optimal foraging theory and has been found to outperform other models to explain

foraging in both human and non-human animals (Bettinger & Grote, 2016; Marshall et al., 2013; Venkataraman et al., 2017).

However, normative models of patch foraging such as MVT rely on the assumption that the organism has perfect knowledge of the average reward rate for the environment. This assumption does not hold for foragers who encounter novel environments, or juvenile organisms whose relative inexperience means their knowledge of the environment is incomplete (Johnson & Wilbrecht, 2011). Under these conditions, the forager must learn the reward statistics of the foraging environment. To explain how reward feedback is utilised to learn the properties of foraging environments, Constantino and Daw (2015) proposed a model whereby reward feedback is integrated into the forager's estimate of the average reward rate for the environment. Testing their model on an adult sample, Constantino and Daw (2015) found that a simple RL-informed MVT model with a single learning rate outperformed other, more complex RL-informed models in explaining adults' foraging behaviour, such as models that included parameters measuring the temporal discounting of delayed rewards associated with travelling. Yet, across several environments which differed in the richness of rewards available to forage, adult participants exploited patches for longer than the optimal foraging policy would dictate (Constantino & Daw, 2015).

Building on this, it has recently been suggested that foraging behaviour might be better explained by a model that separates learning rates for positive and negative prediction errors, rather than a model that has a single learning rate for all prediction errors (Garrett & Daw, 2020). In this study, the authors found that adult participants were better able to learn the true reward value of stimuli, and adjust their behaviour accordingly, when they received positive prediction errors relative to negative prediction errors. This bias for positive prediction errors led participants to select highly rewarding options more frequently when the quality of the environment improved. Yet, participants took longer to learn the reward value of stimuli when

they moved onto an environment that was poorer in quality compared to the one they had previously experienced and as such, participants selected the more rewarding stimuli less often (though we found no such order effects in our previous study; see Supplementary Materials). The order effects observed by Garrett and Daw (2020) were only explained by a model which included separate learning rates for positive and negative prediction errors. These findings suggest that adults utilise reward feedback while foraging and that their behaviour may be explained by modelling learning rates for positive and negative prediction errors separately. However, whether adolescents utilise asymmetric learning rates for positive and negative prediction errors while foraging has not been tested. Age-related differences in strategies used to learn from reward feedback may be able to explain why adolescents exhibit greater rates of exploration, and hence more optimal foraging, compared to adults (Lloyd et al., 2021b).

In the present study, we examine the computational mechanisms used by human participants on a patch foraging task, and whether these mechanisms differ between adults and adolescents. Our first prediction was that participants' foraging behaviour would be better explained either by a model with a single learning rate for prediction errors (Constantino & Daw, 2015), or a dual learning rate model which separates this parameter for positive and negative prediction errors (Garrett & Daw, 2020). Our second prediction was that there would be age-related differences in the degree to which adolescents and adults weighted recent feedback, measured by their learning rate(s). However, as there is conflicting research about whether these parameters increase or decrease from adolescence to adulthood (Eckstein et al., 2020; Davidow et al., 2016; Jepma et al., 2020; Xia et al., 2020), we did not specify a direction for this hypothesis. These findings have implications for understanding how humans forage for resources and developmental changes in how humans learn the reward statistics of their surroundings.

### 4.3. Methods

#### 4.3.1. Participants

The present study is a secondary analysis of our previous work (Lloyd et al., 2021b). In this lab-based study, 68 adolescents were recruited ( $M_{\text{age}} = 16.57$ ,  $SD_{\text{age}} = 0.53$ , 77% female) and 69 adults were recruited ( $M_{\text{age}} = 30.77$ ,  $SD_{\text{age}} = 7.96$ , 46% female). Socioeconomic status, as measured by highest level of education (or parental education for adolescents) and ethnicity are reported in Table 4.1. Ethical approval for this study was received from Royal Holloway, University of London.

*Table 4.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents' highest level of education.*

	<b>Adolescent</b>	<b>Adult</b>
<i>Participant race/ethnicity</i>		
White (UK)	76.47%	29.41%
White (other)	4.41%	22.06%
Mixed race	2.94%	11.76%
Indian	1.47%	7.35%
Pakistani	0%	7.35%
Bangladeshi	4.41%	1.47%
Chinese	0%	10.29%
Asian (other)	16.18%	1.47%
African	2.94%	1.47%
Caribbean	0%	4.41%
Black (other)	0%	0%
Arab	0%	2.94%

#### 4.3.2. Measures

*Patch Foraging Task:* Participants completed a patch foraging task which measures explore/exploit decision-making. In the task, participants were instructed to collect as many rewards as possible within a limited period of time. We utilised an apple picking paradigm based on a previous study by Constantino and Daw (2015). In this foraging paradigm, patches are operationalised as trees and rewards are operationalised as apples. Participants could choose between exploiting their current patch for an immediate reward and exploring a novel patch, which would have a fresh distribution of rewards after incurring a brief time delay for “travel” (during which no rewards can be collected). When participants choose to exploit, the rewards gradually deplete from patches, meaning each successive exploit decision would yield fewer rewards. As such, participants must evaluate when the potential for reward on the new patch outweighs the opportunity cost associated with exploring.

In our original paper (Lloyd et al., 2021b) we had four foraging ‘environments’ which differed in the richness of rewards that were available to collect, with richer environments permitting participants to collect more rewards within the time limit relative to poorer environments. The quality of the environments was manipulated through changing the rate at which apples depleted within patches and the time cost associated with exploration. Four environments were derived by factorially crossing two travel times with two depletion rates (see Table 4.1; Lloyd et al., 2021b). According to MVT, in richer environments, participants should explore more relative to poorer environments, where participants should explore less. In the current study, we do not analyse these four environments separately, as this gave us greater statistical power to reliably estimate the free parameters in our computational models (including the learning rate parameter, see Computational Modelling below; Collins & Wilson,

2019; Daw, 2011). Doing so allowed us to address the first aim of the present study, which was to compare how a model with a single learning rate fits participants' behaviour compared to a model which separates learning rates for positive and negative prediction errors.

*Table 4.2: Features of the four foraging environments from Lloyd et al. (2021b).*

	Environment 1	Environment 2	Environment 3	Environment 4
Depletion Rate	0.88 (0.07)	0.94 (0.07)	0.88 (0.07)	0.94 (0.07)
Mean (SD)				
Travel Time	6	6	12	12
(seconds)				

### 4.3.3. Procedure

Participants were invited to an IT suite where they were given full information about the task and the time required to complete the study. After providing informed consent, participants completed a demographic questionnaire and a practice version of the task. The statistics of the practice environment were different to those of the main task; the travel time was set at 9 sec and the depletion rate was drawn from a Gaussian distribution with a mean of 0.90 (SD = 0.07). Practice trials were subsequently excluded from further analyses. Participants then completed the four foraging environments in counterbalanced order. Each environment lasted for 7-minutes, after which participants were informed they would be moving to a new environment. In this transition screen, participants were informed that this environment may be different to the one they had just previously experienced. They were further instructed that in some environments it may be better to stay with trees for longer, whereas in others it may be better to stay with tree for less time. The task took approximately 28 minutes to complete, after which participants were debriefed as to the aims of the study. Participants were remunerated £4 for their time and were awarded a performance-dependent bonus of up to £3,

which was calculated by dividing their total score earned during the task (i.e., the number of apples they collected) by 330.

#### 4.3.4. Computational Modelling

##### *Single Learning Rate Model*

To explain how reward feedback is utilised to guide participants' foraging behaviour, Constantino & Daw (2015) developed a model based on RL, where participants' estimate of the average reward rate is updated on each trial according to the reward feedback they receive. This equation is formalised as:

$$p_i = (1 - \alpha)^{T_i} \frac{s_i}{\tau_i} + (1 - (1 - \alpha)^{T_i}) p_{i-1} \quad (1)$$

In this equation,  $p_i$  is participants' running estimate of the average reward rate in the current environment, which is informed by the reward received ( $s_i$ ) on the  $i$ th trial along with the time cost ( $\tau_i$ ) associated with their decision to exploit or explore, the latter of which varied between environments. The equation also contains the free parameter  $\alpha$ , which, in the current formulation, measures the complement of traditional learning rates (Sutton & Barto, 2018), meaning that higher values on this parameter indicate a lower learning rate whereas lower values indicate a higher learning rate. Participants' estimate of the average reward rate is then entered into a SoftMax function:

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

Equation 2 calculates the probability that on any given trial participants will take a particular action ( $a_i$ ), which could be to exploit or explore. This equation contains a further two free parameters:  $c$ , which measures participants' overexploitation bias (Constantino & Daw, 2015) and  $\beta$ , which measures the amount of stochasticity in participants' decision-making. Higher values on  $c$  indicate participants have a greater bias to overexploit patches and lower

values on  $\beta$  indicate participants' behaviour is more stochastic. Finally, the term  $\kappa s_i - p_i h$  captures the difference between the rewards expected on the next trial  $\kappa s_i$  and the average reward rate for the environment ( $p_i h$ ). The term  $\kappa_i$  is the rate at which rewards deplete from patches, which varied between foraging environments. In the present study, we assumed that participants estimated a running average of the rate at which apples depleted in each environment. As such, the average depletion rate was calculated on each trial through averaging the true depletion rate participants had experienced in previous trials of that environment. The estimate of the depletion rate was reset to a null value at the beginning of each foraging environment, meaning the model would estimate the depletion rate for each environment separately. Finally,  $h$  refers to the time associated with harvesting patches, which was fixed at three seconds in all environments. Parameter recovery was successful for this model (see Supplementary Materials).

### *Dual Learning Rate Model*

Recent evidence has suggested that foraging behaviour may be better explained by a dual learning rate model containing separate learning rates for positive and negative prediction errors (Garret & Daw, 2020). To examine whether the dual learning rate model explained participants' behaviour in the present study, we separated the learning rate parameter into  $\alpha^+$ , which measured learning rate for positive prediction errors, and  $\alpha^-$ , which measured learning rate for negative prediction errors. On each trial, we calculated whether participants experienced a positive or negative prediction error, based on their running estimate of the average reward rate:

$$\delta_i = s_i / \tau_i - \rho_i \quad (3)$$

When  $\delta_i$  is greater than 0, participants experience a positive prediction error, and their estimate of the average reward rate is updated according to Equation (1), with  $\alpha^+$  replacing  $\alpha$

in this equation. When  $\delta$  is less than 0, participants experience a negative prediction error, and their estimate of the average reward rate is updated according to Equation (1), with  $\alpha$ - replacing  $\alpha$  in this equation. Participants' estimate of the average reward rate is then entered into the SoftMax function described in Equation (2). Parameter recovery for the dual learning rate model was also successful (see Supplementary Materials). However, it is important to note that during parameter recovery we observed significant positive correlations between estimated values of  $\alpha$  and  $\beta$ . We also found separate significant positive correlations between  $c$  and  $\beta$  during parameter recovery of the dual learning rate model. Therefore, when we estimate the parameters for the dual learning rate model from participants' data, we do not draw inferences about correlations we observe between parameter estimates in our behavioural data (Wilson & Collins, 2019).

#### *Model comparison*

To compare whether the single or dual learning rate model best explained participants' foraging behaviour, we calculated two commonly used indices of model fit: the Akaike Information Criterion (Akaike, 1998) and the Bayesian Information Criterion (BIC; Schwartz, 1978). Lower values on these indices suggest the model is a better fit to the data and comparing values on the AIC and BIC between the single and dual learning rate models indicates which of the models provides a more parsimonious explanation of the data (Daw, 2011).

## **4.4. Results**

### **4.4.1. Model Comparison**

To address our first aim, which was to compare whether participants' explore/exploit decisions were better explained by a single or dual learning rate model, we compared the AIC and BIC between these two models. We found lower average values on the AIC for the single

learning rate model compared to the dual learning rate model ( $AIC_{\text{single}} = -139.45$ ,  $AIC_{\text{dual}} = -131.23$ ;  $t(132) = 7.58$ ,  $p > .001$ ), as well as lower average values on the BIC for the single learning rate model ( $BIC_{\text{single}} = -127.45$ ,  $BIC_{\text{dual}} = -115.26$ ;  $t(132) = 10.08$ ,  $p > .001$ ). These findings suggest that the single learning rate was a better fit to both adolescents' and adults' foraging behaviour (see Figure 4.1).

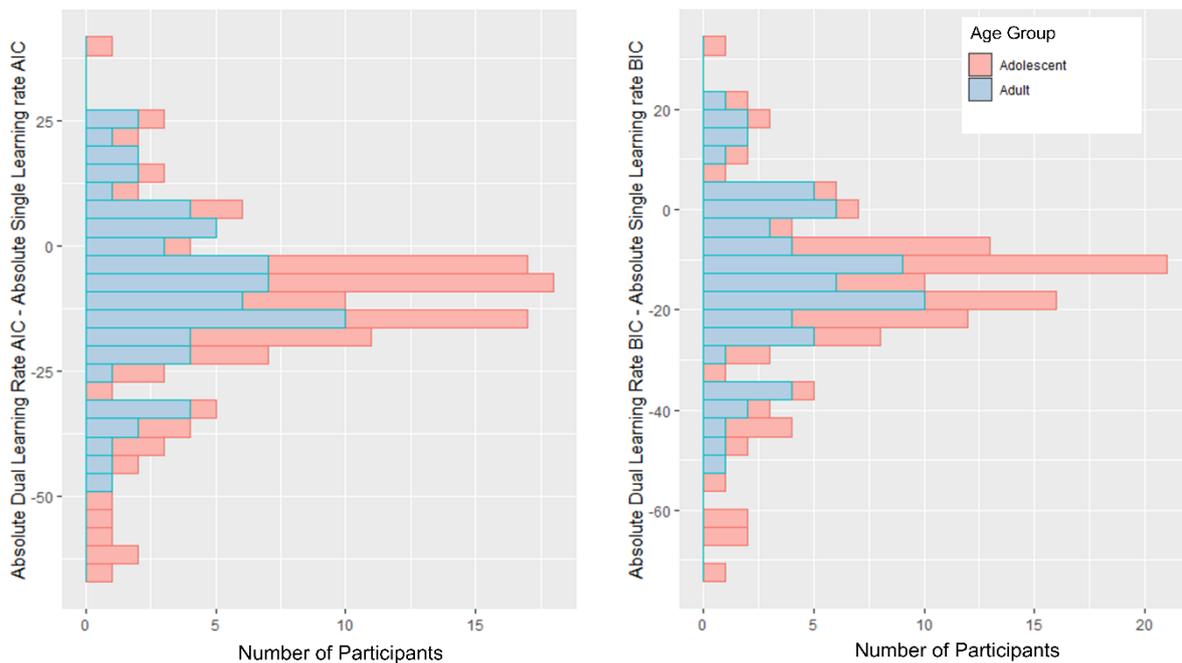


Figure 4.1: Figure demonstrating the difference between absolute values of the AIC and BIC for the dual learning rate model and the single learning rate model. The left plot demonstrates the difference between values of the AIC and the right plot demonstrates the difference between values of the BIC. Lower values on the y-axis indicate the single learning rate model is a better fit to the data

#### 4.4.2. Parameter Comparisons Between Groups

The second aim of the current study was to examine age-related differences on RL parameters. We analysed the three free parameters from the winning computational model:  $\alpha$ , measuring the complement of participants' learning rate;  $c$ , measuring participants' overexploitation bias; and  $\beta$ , measuring stochasticity in participants' decision-making. Levene's test on the variances in the two age groups was significant for the learning rate and

overexploitation parameters and as such, we conducted two non-parametric Mann-Whitney U tests to compare these parameters between adolescents and adults but a conventional  $t$ -test for the stochasticity parameter.

We found a significant difference between adolescents' and adults' learning rate while foraging ( $U = 425.00, p < .001$ ); adolescents utilised higher learning rate (i.e., lower values on  $\alpha$ ) compared to adults (See Table 4.2). However we did not find significant differences between adolescents and adults on the  $c$  parameter measuring overexploitation ( $U = 2386.00, p = .430$ ), nor the  $\beta$  parameter measuring decision stochasticity ( $t(131) = 0.315, p = .753$ ). Together, these findings suggest that the differences in participants' leaving thresholds observed in our previous study (Lloyd et al., 2021b) may be attributable to age-related differences in participants' learning rate. Supporting this interpretation, we found a significant negative correlation between the  $\alpha$  parameter and participants' average leaving threshold ( $r_s = -0.558, p < .001$ ), suggesting higher learning rates were associated with greater rates of exploration. See Table 4.2 for mean parameter estimates for each group.

Table 4.3: Mean (SD) estimates for the free parameters in the single learning rate model

Age Group	Parameter		
	$\alpha^*$	$c$	$\beta$
Adolescent	0.66 (0.24)	0.30 (1.63)	1.15 (0.73)
Adult	0.95 (0.18)	-0.02 (1.30)	1.20 (0.97)

\* Note, as the  $\alpha$  parameter we utilise is the complement of the learning rate in traditional RL models, lower values indicate a higher learning rate.

#### 4.4.3. Model Recovery

We next examined whether we were able to recover the winning model, which involved simulating data according to the single learning rate model and assessing the model fit for the single and dual learning rate models using the AIC and BIC. Successful model recovery is

demonstrated through the winning model (in our case the single learning rate model), which was used to simulate the data, providing a better fit to the data, as indicated by the AIC and BIC (Wilson & Collins, 2019). Foraging data were simulated using parameter estimates from the single learning rate model for all 137 participants in the present study. Once these data were simulated, both the single and dual learning rate models were fit to these data. Results demonstrated that the winning model could be successfully recovered, as we found the single learning rate model had lower values on the AIC and BIC compared to the dual learning rate model (see Table 4.3).

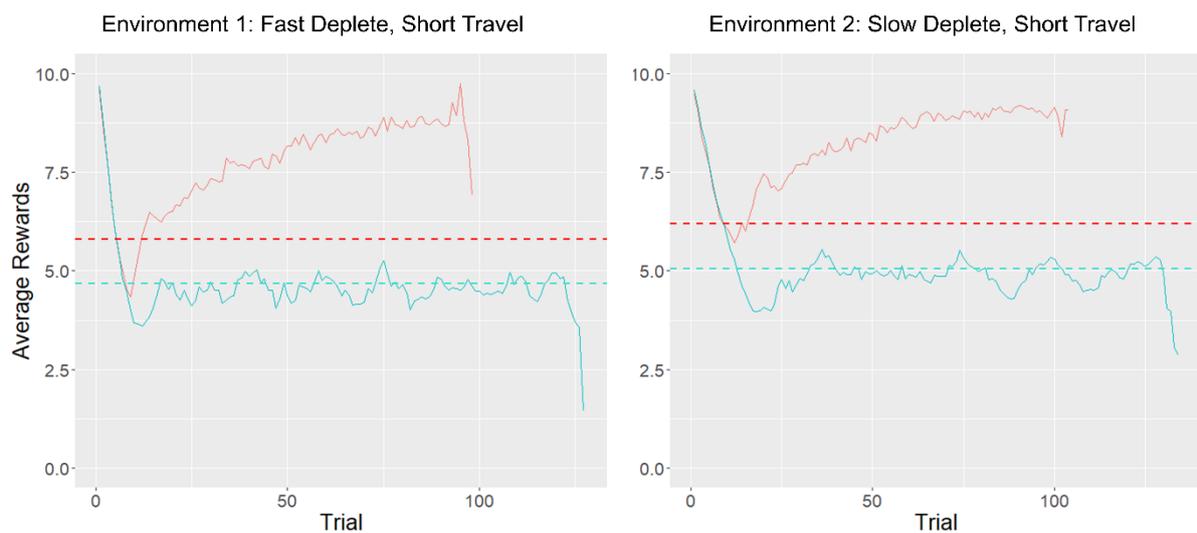
*Table 4.4: Average of the model fit indices used to assess which of the models provided a better explanation of participants' foraging behaviour.*

	<b>AIC</b>	<b>BIC</b>
<b>Model</b>		
Dual Learning Rate	-337.24	-320.55
Single Learning Rate	-349.35	-336.83

#### **4.4.4. Posterior Simulation**

In our previous study (Lloyd et al., 2021b) we observed that adolescents had significantly higher leaving thresholds and stayed with patches for less time (in seconds) compared to adults. To examine whether our model was able to validly capture these features of participants' foraging behaviour, we examined whether we could reproduce the age effects we observed through simulating our computational model. We took the mean parameter estimates for each age group on the three free parameters in the single learning rate model ( $\alpha$ ,  $c$  and  $\beta$ ) and used these to simulate explore/exploit decisions while foraging. We used these mean parameter estimates to simulate agents' explore/exploit choices on 200 probabilistic variations of foraging task, including 100 simulations using the mean parameter estimates from

adolescents' data and 100 simulations using the mean parameter estimates from adults' data. We used these mean parameter estimates to simulate data in all four foraging environments from our original study (see Table 4.1), where the depletion rate and travel time varied between environments. Plotting the explore/exploit decisions of these simulated agents suggested that these parameters were able to recapitulate the difference in adolescents' and adults' foraging behaviour; specifically, adults exploited patches until they were more extensively depleted compared to adolescents (see Figure 4.2). However, we note these simulations overestimate the average number of rewards that adolescents encountered, demonstrated by the discrepancy between the simulated adolescents' average rewards and the horizontal dashed line derived from our behavioural data (see Figure 4.2).



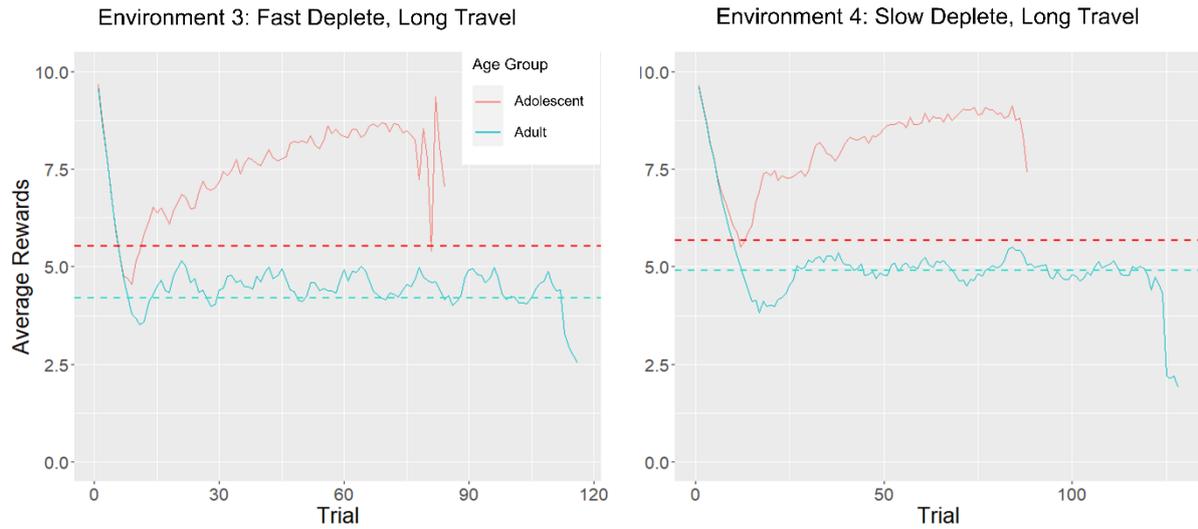


Figure 4.2: Simulation results demonstrating foraging behaviour for simulated adolescents and simulated adults. The y-axis denotes the average number of rewards experienced on that trial and the x-axis is the trial number, which corresponds to each explore/exploit decision. The red dashed line is the leaving threshold adolescent participants utilised in each environment whereas the dashed turquoise line is the leaving threshold adult participants used in each environment.

Analyses of these simulations produced similar effects of age group and environment that we observed in our original paper (Lloyd et al., 2021b). We utilised a  $2 \times 2 \times 2$  mixed ANOVA as used in Lloyd et al. (2021b), which included factors for: age group (adolescent or adult), depletion rate (fast or slow) and travel time (short or long). We were able to replicate a main effect of age group ( $F_{(1,198)} = 21491.83$ ),  $p < .001$ ,  $\eta^2 = .94$ ), with simulated adolescents utilising a higher leaving threshold relative to simulated adults. We also replicated a main effect of depletion rate ( $F_{(1,198)} = 656.44$ ),  $p < .001$ ,  $\eta^2 = .03$ ), with simulated participants exhibiting a higher leaving threshold in environments in environments with a slow depletion rate. Moreover, we also replicated a significant interaction between depletion rate and age group on simulated participants' leaving thresholds ( $F_{(1,198)} = 7.64$ ),  $p = .006$ ,  $\eta^2 < .001$ ). However, in contrast to our previous study where we did not find evidence that adolescents adjusted their leaving thresholds between environments with fast and slow depletion rates (Lloyd et al., 2021b), we found a significant difference between simulated adolescents' leaving thresholds

in the slow and fast depletion environments ( $p_{bonf} < .001$ ). In addition, we replicated our previous finding that there was a main effect of travel time on simulated participants' leaving thresholds ( $F_{(1,198)} = 18.77$ ),  $p < .001$ ,  $\eta^2 < .001$ ), with simulated participants exhibiting higher leaving thresholds in environments with shorter travel times relative to environments with longer times. However, unlike our previous study, we found an interaction between travel time and age group ( $F_{(1,198)} = 28.61$ ),  $p < .001$ ,  $\eta^2 = .001$ ). Post-hoc tests using the Bonferroni correction for six comparisons demonstrated that simulated adults did not use higher leaving thresholds in environments with shorter travel times compared to environments with longer travel times ( $p_{bonf} = 1$ ), whereas simulated adolescents did use higher leaving thresholds in environments with shorter travel times relative to environments with longer travel times ( $p_{bonf} < .001$ ). Finally, replicating our previous study (Lloyd et al., 2021b), there was not an interaction between depletion rate, travel time and age group ( $F_{(1,198)} = 1.61$ ),  $p = .237$ ,  $\eta^2 < .001$ ). Together, these findings suggest that the winning single learning rate model was able to replicate our finding that adolescents had higher leaving thresholds, and hence explored more, compared to adults. Moreover, this model was able to replicate most of the between environment effects we observed in our original study.

#### 4.5. Discussion

The present study examined the computational mechanisms involved in adolescents' and adults' patch foraging behaviour. In contrast to previous research (Garrett & Daw, 2020), we found that both age groups' explore/exploit decisions on a foraging task were best explained by a model with a single learning rate rather than a dual learning rate. Furthermore, examining developmental differences on the free parameters in this model, we found that adolescents had a higher learning rate compared to adults, but we did not find age-related changes on the parameters measuring overexploitation bias or decision stochasticity. Together, these findings

contribute to our understanding of how adolescents navigate novel foraging environments and how their learning strategies can facilitate the development of their experiential knowledge.

We found that adolescents' and adults' foraging behaviour was best explained by a model that included a single learning rate that applies to both positive and negative prediction errors. These findings are contrary to previous work that used a foraging task in which participants decided whether to accept or reject an option before being forcibly moved to the next 'patch' (Garrett & Daw, 2020). As such, this previous study did not contain patches with depleting rewards as we did here. This methodological difference may explain why we found the single learning rate model provided a better fit to our participants' data. In the current task, each successive exploit decision is processed as a negative prediction error, as the depleting rewards incrementally decreases the average reward rate. Therefore, a fast learning rate for negative prediction errors on each exploit trial would lead to lower estimates of the average reward rate (thereby increasing the likelihood the participant will explore) relative to using a slow learning rate where more historic feedback is integrated into the estimate of the average reward rate. Further, when arriving at a new patch, the participant receives a large positive prediction error having recently experienced no rewards while travelling. As such, a higher learning rate for positive prediction errors also promotes increased exploration. Therefore, in the current task, higher learning rates for positive and negative prediction errors both increase participants' exploratory behaviours, whereas in Garrett & Daw (2020), learning rates for positive and negative prediction errors produce different behavioural effects depending on whether the environment was improving or diminishing. The symmetrical association between learning rates for positive and negative prediction errors, and participants' exploration behaviour in our task may explain why participants utilise a single learning rate for all outcomes, as the separation of these learning rates does not produce any behavioural differences. As such, a simpler learning strategy using a single learning rate may be more

computationally efficient. Indeed, this interpretation is consistent with our finding that higher learning rates were associated with greater exploration. As such, in comparison to Garrett and Daw (2020) the present task may be better suited to the use of a single learning rate for all prediction errors rather than an asymmetric learning rate for positive and negative prediction errors.

Foraging problems are thought to have contributed to the evolution of brain structures involved in reward processing (e.g., Kolling & Akam, 2017; Niv et al., 2002) and therefore may provide insight into the normative development of the ability to learn from reward feedback across the lifespan. Our findings are consistent with Decker and colleagues (2015), who also found that estimates on the learning rate decreased from childhood to adulthood on a multi-armed bandit task. However, our findings are inconsistent with some studies that have found learning rates to increase as individuals mature (e.g., Davidow et al., 2016; Master et al., 2020). Methodological differences between our task and previous research may contribute to understanding conflicting findings in the literature, as we note that there are some aspects of patch foraging problems that may be better suited to a strategy of weighting recent information highly (i.e., having a fast learning rate). For example, while the statistics of the present task remained stable over the course of each environment, both the initial richness of patches and depletion rate were drawn from probabilistic distributions, which can make the environment more uncertain than environments where these values are fixed (e.g., Constantino & Daw, 2015). Environments that are more uncertain favour a learning strategy that relies on more recent experience (i.e., a higher learning rate) relative to paradigms where reward features of the task are fixed (Nussenbaum & Hartley, 2019). Studies that have found the learning rate to increase with age (e.g., Master et al., 2020) utilise tasks where feedback is deterministic, rather than probabilistic, which favours a slow learning rate (Nussenbaum & Hartley, 2019). These findings could indicate that adolescents are more adept than adults at calibrating their learning

rate to the statistics of different environments, whereas adults are more invariable in their learning rate. However, future research would need to test this empirically, as current studies have not compared how adolescents and adults adjust their learning rate between environments with different reward contingencies within a single task (see Chapter 5).

Our findings suggest that adolescents' use of a higher learning rate, compared to adults, supported them to engage in more optimal exploration, which may indicate adolescents are better at navigating probabilistic environments compared to adults (DePasque & Galvan, 2017). Adolescents' reliance on a higher learning rate compared to other age groups when learning to explore while foraging may be adaptive at this point in the lifespan, as this is a period of significant transition (e.g., starting a new school; Somerville et al., 2010). In such changeable environments, it is beneficial to have a higher learning rate (Behrens et al., 2007). Notably, unlike previous research that required participants to learn the reward values of specific stimuli (e.g., associative learning tasks; Master et al., 2020), we examined how participants update their reward estimates for the average reward rate of the environment. This modelling approach provides novel insight into how adolescents assess the relative bounty in their environment, which informs the value of exploring novel opportunities rather than exploiting a known option. As such, higher learning rates during adolescence, relative to adulthood, may support this age group to engage in exploratory behaviours and seek out novel opportunities in their real-world environments.

In addition, the use of a higher learning rate relative to adults could suggest that adolescents utilise recent experiences to guide future decisions about actions to approach and those to avoid (Do et al., 2020). This account would be consistent with recent theories of development, which propose that adolescence is a period of relative naiveté about one's surroundings, yet one where the individual develops the experiential knowledge which can support them in gaining adult independence (Romer et al., 2017). In contrast, relative to

adolescents, adults have a wider range of experiences spanning years to draw upon during real-world decision-making. Adults' use of previous experience to guide decision-making may be consistent with the use of a lower learning rate, which incorporates more historic knowledge into one's estimate of the value of available options (Behrens et al., 2007; Nussenbaum & Hartley, 2019). However, it remains unknown whether age-related differences in learning rates, measured on relatively brief experimental tasks, generalise to the use of feedback from real-world decisions that may have occurred over longer event horizons (i.e., years rather than minutes). As such, this possibility should be tested empirically in future research.

This study has some limitations, which are important to acknowledge. While the SoftMax function can measure features of decision-making, it does not provide a mechanistic explanation of how participants decide between two options (Fontanesi et al., 2019). Future research could consider more complex models that combine the reward valuation explained by RL models, alongside evidence accumulation processes, such as those explained by Drift Diffusion Models (e.g., Fontanesi et al., 2019). This is important as recent accounts suggest foraging decisions are made through a complex evidence accumulation process (Davidson & El Hady, 2019), which may be neglected by this relatively simple explanation of choice behaviour. Further, there are limitations with the foraging paradigm we utilised as participants deterministically moved onto patches, rather than having the option to select between a variety of patches. Recent work has demonstrated that allowing foragers autonomy over which patches to select within an environment can produce more accurate model fits of foraging behaviour in both simulation and behavioural work (Hall-McMaster et al., 2021; Hall-McMaster & Luyckx, 2019). In this work, participants' behaviour is best explained by a model that tracks within patch reward rates, rather than environmental reward rates, which could favour a slower learning rate to avoid overweighting within-patch fluctuations. As such, comparing adolescents' and adults' behaviour on this task would be an important avenue for future

research to examine whether adolescents' higher learning rate is detrimental in this foraging context.

In sum, the present study demonstrated that adolescents' and adults' patch foraging behaviour is best explained by a model with a single learning rate for positive and negative prediction errors rather than a dual learning rate model with separate parameters for these types of feedback. Examining developmental differences on free parameters within these models, we demonstrate that adolescents have a higher learning rate compared to adults and therefore integrate feedback into their decision-making more rapidly than their older counterparts. This may be an adaptive decision-making strategy at this point in the lifespan as the individual experiences significant periods of transition. Indeed, through rapidly integrating novel information into their expectation of outcomes associated with decisions, adolescents may be able to cultivate the experiential knowledge necessary for adult independence.

## 4.6. Supplementary Materials

### 4.6.1. Parameter Recovery

#### *Single Learning Rate Model*

Parameter recovery for the single learning rate model was successful. We hardcoded parameters across a range of values ( $\alpha$  {0,1} in increments of 0.1;  $c$  {-1, 3} in increments of 0.5 and  $\beta$  {0, 5} in increments of 0.5). We observed significant positive correlations between hardcoded and estimated values for the learning rate parameter ( $r(538) = 0.47, p < .001$ ),  $\beta$  parameter ( $r(538) = 0.27, p < .001$ ) and  $c$  parameter ( $r(538) = 0.33, p < .001$ ). We did observe a significant correlation between estimates for the learning rate parameter and  $\beta$  parameter ( $r(538) = 0.53, p < .001$ ). However, we did not find significant correlations between other parameters in this model.

#### *Dual Learning Rate*

Parameter recovery for the dual learning rate model was also successful. We observed significant correlations between hardcoded and estimated values for the  $\alpha_+$  parameter ( $r(1618) = 0.87, p < .001$ ),  $\alpha_-$  parameter ( $r(1618) = 0.52, p < .001$ ),  $c$  parameter ( $r(1618) = 0.27, p < .001$ ) and  $\beta$  parameter ( $r(1618) = 0.17, p < .001$ ). However, in this model we also observed significant correlations between parameter estimates for  $\alpha_+$  and  $\beta$  ( $r(1618) = 0.65, p < .001$ ) and  $c$  and  $\beta$  parameters ( $r(1618) = 0.48, p < .001$ ).

### 4.6.2. Exploratory Analysis of Dual Learning Rate Model Parameters

Although the dual learning rate model was not the best fit to the data, we conducted exploratory analyses on the free parameters in this model to examine whether these were consistent with previous research. As Levene's tests were significant for the age group factor, we conducted a Mann-Whitney U for the positive and negative learning rate estimates, and the overexploitation parameter estimate  $c$ . We found a significant difference between adolescents

and adults on the learning rate parameter for positive prediction errors ( $U = 1226, p < .001$ ). Consistent with our analyses of the single learning rate model, adolescents had a faster learning rate ( $M = 0.76, SD = 0.24$ ) for positive prediction errors relative to adults ( $M = 0.95, SD = 0.10$ ). Moreover, we found a significant difference between adolescents' and adults' learning rate for negative prediction errors ( $U = 3619, p < .001$ ). In contrast to the learning rate for positive prediction errors, adults ( $M = 0.13, SD = 0.18$ ) had a higher learning rate for negative prediction errors relative to adolescents ( $M = 0.40, SD = 0.18$ ). We note this finding is consistent with some previous research that has used n-armed bandit tasks (e.g., Christakou et al., 2012; Palminteri et al., 2016; however, see Burrítica et al., 2019; Hauser et al., 2015; van den Bos et al., 2012). In our analysis of this model, we also found significant differences between adolescents and adults on the overexploitation parameter ( $U = 1227.50, p < .001$ ), with adults having a higher value on this parameter ( $M = 2.48, SD = 0.91$ ) relative to adolescents ( $M = 1.33, SD = 1.37$ ). This would be consistent with findings from Lloyd et al. (2021b) that adolescents explore more than adults while foraging. Finally, using an independent samples t-test, we also observed a significant difference between age groups on the parameter measuring stochasticity ( $t_{(136)} = 3.18, p = .002, \text{Cohen's } d = 0.54$ ), with adolescents exhibiting more stochastic behaviour ( $M = 1.48, SD = 0.90$ ) relative to adults ( $M = 2.02, SD = 1.10$ ). The finding for the  $\beta$  parameter differed from our analysis of the  $\beta$  parameter in the single learning rate model, where we did not find differences between adolescents and adults on this parameter.

#### **4.6.3. Exploratory Analysis of Counterbalance Order**

Previous research has found that a model that separates learning rates for positive and negative errors was able to explain order effects whereby participants exhibited greater adjustment of the length of time they were willing to exploit a patch (i.e., their patch residency time) and the frequency with which they selected stimuli when the quality between two

environments improved compared to when it deteriorated (Garrett & Daw, 2020). We tested whether there were order effects on the extent to which participants exploited patches in the present study (i.e., their leaving threshold). Though we analysed leaving thresholds rather than the patch residency times, which differed to Garrett & Daw's (2020) analysis, we note in our original paper that we found similar patterns between participants' leaving thresholds and their patch residency times (Lloyd et al., 2021b). The absence of an order effect could explain why we found a model with a single learning rate to explain our data better than a model with separate learning rates for positive and negative prediction errors. We used a mixed ANCOVA with environment and age group (adolescent and adult) as fixed factors and counterbalance order for all four levels entered as a covariate. A Greenhouse-Geisser correction was applied to all within-subject effects as Mauchly's test of sphericity indicated that the assumption of sphericity was violated ( $p = .033$ ). These analyses suggested there was no evidence of an order effect as there was not a main effect of counterbalance order on participants leaving thresholds ( $F_{(1,121)} = 0.49, p = .484$ ), nor did we find an interaction between environment and counterbalance order ( $F_{(2.81,229.51)} = 0.96, p = .296$ ). Consistent with our original analyses (Lloyd et al., 2021b), we found a main effect of environment ( $F_{(2.81,229.51)} = 4.79, p = .003$ ), with participants exploring more in the richer environments relative to the poorer quality environments (see Lloyd et al., 2021b, Figure 2). Further, we found a main effect of age group ( $F_{(2.81,229.51)} = 12.99, p < .001$ ), with adolescents exploring more than adults (see Lloyd et al., 2021b, Figure 2). We also found no evidence of an interaction between environment and age group ( $F_{(2.81,229.51)} = 2.14, p = .100$ ). These findings suggest that there were no order effects in our data.

#### **4.6.4. Association Between Model Parameters and Disposition Measures**

We conducted exploratory analysis to examine whether the model parameters from the winning single learning rate model were associated with the disposition measures we collected

during this study. Notably, in contrast to previous research that has proposed a link between choice stochasticity and impulsivity (e.g., Hauser et al., 2016), we did not find an association between the  $\beta$  parameter, which measures stochasticity and our self-report measure of impulsivity (Barrett-11;  $r(137) = -0.06$ ,  $p = .482$ ). We did not find any other associations between model parameters and disposition measures ( $ps > .076$ ).

## Chapter 5: Optimal Exploration in Stable and Volatile Foraging

### Environments Depends on Response Stochasticity and Anxiety<sup>4</sup>

#### 5.1. Abstract

Sensitivity to the (in)stability of one's environment is important for survival. Greater instability should direct the organism to weigh recent feedback higher compared to more historic feedback, a calibration captured in reinforcement learning models by the individual's learning rate. As adolescence is a period of rapid developmental changes, we first hypothesised that adolescents would be more optimal than adults (who experience relative stability) when increasing their learning rate from stable to volatile environments. Second, we predicted that across age groups the ability to increase one's learning rate from stable to volatile environments would be impaired by anxiety. We thirdly predicted that we would replicate previous work that in stable environments, adolescents would explore more than adults. Adolescents aged 16-17 ( $N=88$ ) and adults aged 24+ ( $N=88$ ) completed a foraging paradigm, where they decided to exploit familiar options or explore novel ones. Our first and second hypotheses were not confirmed: Neither age group exhibited higher learning rates in the volatile relative to the stable condition. Instead, participants calibrated their stochasticity to the prevailing environmental volatility and anxiety impaired this calibration of stochasticity, rather than that of the learning rate as we originally hypothesised. Moreover, adolescents exhibited higher levels of stochasticity than adults in both foraging environments. Our third hypothesis, however, was confirmed: Adolescents explored more than adults in the stable foraging environment. Taken together, these findings characterise stochasticity as a key mechanism that drives the heightened novelty-seeking in adolescence relative to adulthood.

---

<sup>4</sup> OSF project page: <https://osf.io/dvqy6/>

## 5.2. Introduction

During adolescence (ages 10-24; Sawyer et al., 2018), there are significant changes to the individual's social and physical environment. This period has been associated with changes to social behaviour and peer relationships as the individual becomes independent of caregivers (e.g., Blakemore & Mills, 2014; Ellis et al., 2012), but adolescence also coincides with lifestyle changes, such as educational milestones (Eccles et al., 1993). The ability to adapt to such changes is necessary to ensure these transitions are successful and experimental evidence has demonstrated that adolescents are particularly adept at adjusting their motivations and goal priorities to negotiate changeable environments compared to other age groups (Crone & Dahl, 2012). Indeed, this ability may be important to allow adolescents to take advantage of opportunities to explore their environment, which can provide adolescents with experiences necessary for adult independence (Romer et al., 2017). However, this adaptive ability might be threatened by anxiety (Browning et al., 2015) and the onset of clinical disorders associated with anxiety are heightened during adolescence relative to childhood and adulthood (Rapee et al., 2019). The present study aims to examine how adolescents and adults adapt their behaviour to changes in their surroundings and the effects of anxiety on these cognitive abilities.

One paradigm that might be fruitful to address whether adolescents are more adept than adults at negotiating changeable environments is patch foraging. During patch foraging the individual chooses between collecting rewards from discrete patches that deplete the longer the forager remains with the patch and exploring a novel patch with a fresh distribution of rewards (Charnov, 1976). Recent research has demonstrated that adolescents explore more optimally than adults while foraging and are able to adjust their behaviour to the reward statistics of stable foraging environments that differ in the relative richness of rewards available (Lloyd et al., 2021b). According to a recent theory, adolescents' cognitive flexibility may support this age group to learn the properties of their surroundings and this experience can be used to guide

future decision-making (Romer et al., 2017). As such, examining how adolescents learn from exploring their surroundings is important to understand this developmental period and patch foraging tasks may be particularly well suited to testing hypotheses about the adaptive features of adolescents' exploration (Duell & Steinberg, 2021).

Adolescents' potentially marked ability to adapt their foraging behaviour between stable and volatile environments can be studied using reinforcement learning (RL) models adapted from a formal model of foraging known as Marginal Value Theorem (MVT; Constantino & Daw, 2015; Kolling & Akam, 2017). Broadly, RL models contain a parameter measuring how much emphasis the individual places on recent versus more historic feedback to guide decision-making, known as the 'learning rate', with higher values denoting that the individual places greater emphasis on more recent feedback (Sutton & Barto, 2018). In stable environments, stimulus-outcome contingencies are consistent over time (Behrens et al., 2007). Under these stable conditions, one should not place too much emphasis on recent feedback (i.e., utilise a lower learning rate), as this can lead the individual to base their expectations about stimulus-outcome contingencies on single random events that are unlikely to reflect the state of the environment (Lawson et al., 2017). Indeed, overweighting recent feedback in stable environments is thought to be an important mechanism in the development and maintenance of psychological symptoms such as delusions in schizophrenia (Deserno et al., 2020) and one recent study has demonstrated how the learning rate can be impacted by experiences associated with psychopathology, such as childhood trauma (Lloyd et al., 2022). Conversely, if the stimulus-outcome contingencies of the environment are known to be rapidly changing (i.e., are volatile), unexpected occurrences may signal meaningful changes to the state of the environment and therefore using more recent occurrences as a basis for one's decisions (i.e., utilising a higher learning rate) is beneficial (Manning et al., 2017). Adjusting one's behaviour

to this contrast between stability and volatility in the environment is important to successfully reap rewards from one's surroundings (Behrens et al., 2007).

As adolescence is associated with an increased ability to flexibly adapt behaviour to achieve goals (Crone and Dahl, 2012), adolescents may be better cognitively equipped than other age groups to negotiate changeable environments, though previous research has not examined this in the context of patch foraging. It has been demonstrated that adolescents more optimally weight recent evidence and accrue more rewards compared to other age groups in a reversal learning task examining children's, adolescents', and adults' learning strategies in volatile conditions (Eckstein et al., 2020). Adolescents also perform more optimally than adults in stable environments. Davidow and colleagues (2016) found that adolescents utilised a lower learning rate than adults in a stable environment, better learned the probabilities of reward associated with each stimulus and therefore more often selected the stimulus that led to a reward and subsequently collected more points. Together, these studies raise the hypothesis that adolescents are adept at calibrating their learning rates to both stable and volatile conditions. However, to our knowledge, no studies to date have experimentally manipulated stability/volatility within a single study to examine whether adolescents increase their learning rate as the environment becomes more volatile. Moreover, adolescents' and adults' ability to adjust their learning rate between conditions of stability and volatility has not been examined in context of patch foraging, which may be particularly well suited to testing hypotheses about exploration and cognitive flexibility in adolescence (Duell & Steinberg, 2021).

To empirically test whether adolescents demonstrate greater cognitive flexibility than adults between conditions of stability and volatility, we implemented a foraging paradigm where the (in)stability of the environment is manipulated. We also simulated RL models to estimate the optimal weight participants should place on recent versus historic feedback (i.e., the optimal learning rate) in these environments, which confirmed that it was optimal to utilise

a higher learning rate in the volatile environment relative to the stable environment. We use these methods to test our first pre-registered hypothesis (<https://osf.io/p3md9/>), which is that participants will utilise a higher learning rate in the volatile environment relative to the stable environment, though relative to adults, adolescents will exhibit a greater difference in their learning rate between the two environments, consistent with the simulated optimal RL agent.

However, the ability to adjust behaviour to one's surroundings might be disrupted by psychopathology, such as anxiety. Anxiety disorder can lead to poorer instrumental learning (Alvares et al., 2014) and a poorer ability for adults to adapt their learning rate between stable and volatile environments (Browning et al., 2015). Indeed, a recent study has demonstrated that self-reported anxiety predicts less directed exploration in an adult sample (Smith et al., 2021). As such, we expect anxiety to disrupt adolescents' and adults' ability to adjust their learning rate between stable and volatile environments. Examining anxiety in the context of adolescence is important to identify difficulties this age group may have in adjusting to novel environments after significant periods of transition, such as beginning a university programme. Our second hypothesis was that across age groups, higher rates of anxiety will be associated with a less optimal adjustment of participants' learning rate between the stable and volatile environment. Finally, our third hypothesis was that we would replicate previous research (Lloyd et al., 2021b), demonstrating that adolescents explore more than adults in stable foraging environments, which we measure in the current study as the point at which participants leave their current patch to explore a novel one (i.e., their leaving threshold).

## **5.3. Method**

### **5.3.1. Participants**

We recruited 176 participants including 88 adolescents ( $M_{\text{age}} = 16.32$ ,  $SD_{\text{age}} = 0.49$ , 70.33% female) and 88 adults ( $M_{\text{age}} = 33.54$ ,  $SD_{\text{age}} = 7.12$ , 68.09% female). The sample size was based on *a priori* power calculations using G\*Power (Faul et al., 2007) for the pre-

registered hypotheses detailed in the Introduction. For our first hypothesis, we conducted an *a priori* power analysis for an interaction effect with one between- and one within-participant factor in an ANOVA ( $\alpha = 0.05$ , power = 0.8, Cohen’s  $f = 0.25$ ), with the effect size based on previous research using this paradigm (Lloyd et al., 2021b). For our second hypothesis, we conducted an *a priori* power analysis for the covariate effect within a GLM ( $\alpha = 0.05$ , power = 0.8, effect size  $f^2 = 0.15$ ). For our third hypothesis, we conducted an *a priori* power analysis for an independent *t*-test ( $\alpha = 0.05$ , power = 0.80, effect size  $d = 0.5$ ), with the effect size determined based on our previous research (Lloyd et al., 2021b). Adolescent participants were recruited through a secondary school based in Surrey, UK. Adult participants were recruited through Prolific ([www.prolific.co](http://www.prolific.co)) and we restricted the location of participants recruited through this platform to the UK to match the location of the adolescent sample.

*Table 5.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents’ highest level of education.*

	<b>Adolescent</b>	<b>Adult</b>
<i>Participant race/ethnicity</i>		
White (UK)	69.32%	96.59%
White (other)	6.82%	0%
Mixed race	11.36%	1.14%
Indian	2.27%	0%
Pakistani	3.41%	3.41%
Bangladeshi	1.14%	0%
Chinese	0%	0%
Asian (other)	6.82%	3.41%
African	0%	0%
Caribbean	1.14%	1.14%

Black (other)	1.14%	1.14%
Arab	1.14%	0%
Education	2.66	3.24

---

### 5.3.2. Measures

*Patch Foraging:* The present study utilised an apple patch foraging paradigm (Constantino & Daw, 2015) to examine explore/exploit decision-making in conditions of stability and volatility. Participants attempted to collect as many rewards as possible within a fixed time limit. Rewards were operationalised as apples, patches were operationalised as trees that were available to forage, and environments were described as ‘orchards’ that participants collected apples from. If participants chose to stay in the current patch and exploit it, they would be presented with the number of apples collected on that trial alongside their cumulative number of apples. Reward would deplete on each trial until zero apples remained in the patch. Alternatively, participants could choose to leave and explore a novel tree, which would have a fresh distribution of rewards. Explore decisions incurred a time cost during which no new rewards could be harvested, which was fixed at six seconds in both stable and volatile environments. As participants had a fixed amount of time to collect rewards, exploring new patches limited the time available for participants to harvest apples. Participants had up to two seconds to decide whether to explore or exploit on each trial. If they did not make a response within this window, they would be presented with a timeout screen after which they arrived at a new patch. Timeout trials incurred a longer wait time than exploring to ensure that there was no benefit from not responding within the allocated time. Trials where participants timed out were excluded from analysis. In the stable condition, the statistics of the foraging environment remained consistent. The distribution of rewards on new patches (i.e., their initial richness) was set at a mean of 10 ( $SD = 1$ ). Rewards in patches underwent exponential depletion, in which

the current harvest was multiplied by a value drawn from a Gaussian distribution with a mean of 0.88 ( $SD = 0.07$ ).

The volatile environment was separated into rich and poor portions. In the rich portion of the volatile environment, the mean initial richness of patches was 13 ( $SD = 1$ ) with an exponential depletion rate drawn from a Gaussian distribution with a mean of 0.94 ( $SD = 0.07$ ). In the poor portion of the volatile environment, the mean initial richness of patches was seven ( $SD = 1$ ), and the exponential depletion rate was drawn from a Gaussian with a mean of 0.76 ( $SD = 0.07$ ). We did not manipulate the time cost of exploration as in previous studies (e.g., Constantino & Daw, 2015; Lloyd et al., 2021b) to avoid participants remaining in the poor-quality portion for substantially longer than the rich quality portion. In the volatile condition, the initial richness of rewards on new patches and depletion rate within patches switched after every 30 trials (i.e., after participants had made 30 choices to exploit or explore). Based on pilot data we found that using fewer than 30 trials could have allowed participants to experience only a single patch in each portion if they chose to exploit the patch until it was fully depleted.

For our behavioural analyses, we were interested in participants' exploration behaviour, operationalised in the present study as their leaving thresholds. We assume as in previous research (Constantino & Daw, 2015) that participants select a value of rewards as their 'threshold' and that when they expect the number of rewards from the next exploit decision to fall below this threshold value, they leave to explore a new patch. Higher values on the leaving threshold indicate participants explore more frequently. We calculated participants' leaving threshold as the average number of rewards harvested from the previous two exploit decisions before they chose to explore (Constantino & Daw, 2015). We compared participants' leaving thresholds to the optimal leaving threshold for the stable environment, which was calculated by conducting a grid search across leaving thresholds ranging from 1-10 in increments of 0.001 and identifying which yielded the highest number of rewards (see Lloyd et al., 2021b).

*Anxiety:* To measure anxiety, we utilised the General Anxiety Disorder-7 scale (GAD-7; Spitzer et al., 2006). The questionnaire asks participants to report the frequency that they have been bothered by a series of problems over the past two weeks. For example, participants are asked how often they have been bothered by “Worrying too much about different things?” on a four-point Likert scale, ranging from ‘Not at all’ to ‘Nearly every day’. Total scores were used, with scores of 5, 10 and 15 reflecting mild, moderate and severe anxiety respectively. This scale demonstrated good reliability in the sample recruited for this study  $\alpha = 0.90$ .

*Demographic Information:* Participants were asked to report their age, gender, and highest level of education. For adolescent participants, we asked them to report their parents’ highest level of education, which can be used to approximate socioeconomic status for developmental samples (Steinberg et al., 2018).

### **5.3.3. Procedure**

Participants completed the study online as recruitment took place during the COVID-19 pandemic. Both adolescent and adult samples were recruited within a 16-day window to avoid systematic differences between the samples that could arise from population level changes in anxiety associated with the pandemic (Fancourt et al., 2021; Shevlin et al., 2020). The study was hosted on Gorilla.sc, an online platform for behavioural research (Anwyl-Irvine et al., 2019). Adolescent participants were provided a link to complete the task in a school session supervised by their class teacher and adults were directed automatically to the task through Prolific.

Once participants began the study, they were provided with a consent form detailing the study protocol and the contact details of the researchers. After providing consent, participants completed a demographic questionnaire and a short practice of the foraging task. The practice environment lasted for two minutes, and participants were instructed that rewards collected during this practice would not count towards their overall score. The structure of this

practice task resembled the stable environment; the initial distribution of rewards on patches were drawn from a Gaussian distribution with a mean of 10 (SD = 1) and the depletion rate was drawn from a Gaussian distribution with a mean of 0.90 (SD = 0.07). Once the practice task was complete, participants were provided with a reminder of key information about the task, including the time they would have in each environment and that the environments might differ from one another. Participants then completed the stable and volatile foraging environments. The sequence in which the stable and volatile environments were presented was counterbalanced across participants, as was the presentation of the rich and poor portions in the volatile environment. Each environment was seven minutes in length and the foraging task took approximately 16 minutes to complete in total. After completing the main task, participants completed the GAD-7 questionnaire. Finally, participants were told their final score and debriefed as to the aims of the study. All participants were compensated a minimum of £3 for their participation. In addition to this base payment, participants were awarded a performance-dependent bonus which was worth up to £2 and was calculated as a proportion of the total number of rewards they accumulated in both environments.

#### 5.3.4. Computational Modelling

According to a prominent theory of optimal foraging known as Marginal Value Theorem (MVT), the forager should leave their current patch to explore a novel one when the rewards expected from exploiting the current patch fall below the average reward rate for the environment (Charnov, 1976). Drawing on the reinforcement learning framework, Constantino and Daw (2015) developed a model that describes how participants learn the average reward of the environment by integrating reward-feedback into their estimation of this value:

$$p_i = (1 - \alpha)^{T_i} \frac{S_i}{T_i} + (1 - (1 - \alpha)^{T_i}) p_{i-1} \quad (1)$$

The equation describes how the participant updates their estimate of the average reward rate ( $p$ ) on each trial ( $i$ ). This value was calculated from the time ( $T_i$ ) associated with either an explore or exploit decision ( $a_i$ ) and the reward received on each trial ( $s_i$ ). The equation additionally contains the free parameter  $\alpha$ , which measures the complement of participants' learning rate (i.e., the learning rate is  $1-\alpha$ ). This value captures the degree to which participants weight recent information in their decision-making (Sutton & Barto, 2018). Participants' estimation of the average reward rate is then entered into a SoftMax function, which estimates the probability that participants will choose to exploit on each trial:

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

In Equation (2), the parameter  $c$  measures participants' exploitation bias, with higher values indicating a greater overexploitation bias. The  $\beta$  parameter measures how reliably the participant selects the option that they believe to have the highest expected value. Higher values on  $\beta$  indicate the participant more reliably selects the option which they believe to have the highest expected value, whereas lower values on  $\beta$  indicate the participant more often diverts from the option which they believe to have the highest expected value (i.e., are more stochastic). The term  $k_{i-1}s_{i-1}$  refers to the last known reward obtained from exploiting the patch ( $s_i$ ), multiplied by the depletion rate  $k_i$ . In the present study, the value of  $k$  was updated on each trial ( $i$ ) and was calculated through averaging the true depletion rate experienced on all previous exploit decisions (as in previous research; Constantino & Daw, 2015). Finally, the term  $p_{i-1}h$  refers to participants' current estimate of the average reward rate ( $p_{i-1}$ ), multiplied by the harvest time ( $h$ ), which was fixed at three seconds in both environments. Further details about model validation and comparison are reported in the Supplementary Materials.

As we were interested in how participants are able to adjust their learning rates between stable and foraging environments, we calculated the optimal learning rate for each environment

using simulation (see Supplementary Materials for full simulation methodology). To provide a measure of how much participants should adjust their learning rate between the foraging environments, we calculated the difference between the optimal values of the learning rate in the stable and volatile environment. This value was subsequently subtracted from the difference in participants' own learning rate between the stable and volatile foraging environments to provide a measure how closely participants' behaviour was to the optimal RL agent. Values closer to zero on this variable indicate the participant's adjustment of their learning rate was closer to the optimal RL agent.

### **5.3.5. Design**

The present study was a 2×2 mixed factorial design with two independent variables: age (adolescent or adult) and foraging environment (stable or volatile). Our first hypothesis was that participants would increase their learning rate in the volatile condition, relative to the stable condition. Moreover, we predicted that the degree to which adolescents adjusted their learning rate between the stable and volatile environments would be greater than adults. As such, the dependent variable of interest for this first hypothesis was participants' learning rate. As we were also interested in the effect of anxiety on participants' ability to adjust their learning rates between environments (Hypothesis 2), self-reported anxiety and age group were entered as predictors of the difference between how much participants adjusted their learning rate and how much the optimal RL agent adjusted its learning rate between the environments. Our third hypothesis was that adolescents would explore more than adults in the stable environment. As such, the dependent variable of interest for the third hypothesis was participants' leaving thresholds, which measured participants' rates of exploration. We compared participants' leaving thresholds in the stable environment between age groups using an independent samples *t*-test. In addition, exploratory analyses were conducted on the other free parameters in this

model, which were the parameter measuring decision stochasticity ( $\beta$ ) and a parameter measuring participants' overexploitation bias ( $c$ ).

## 5.4. Results

### 5.4.1. Confirmatory Analyses

We simulated an RL foraging agent to determine the optimal learning rate values in each of our foraging environments, which demonstrated that it was optimal to adopt a higher learning rate in the volatile environment relative to the stable environment (see green lines, Figure 5.1; for a full description of the simulation methods, see Supplementary Materials). Our first pre-registered hypothesis was that, consistent with an optimal RL agent, adolescents would show greater adjustment of their learning rate between the stable and volatile foraging environments compared to adults. To test this, we conducted a mixed  $2 \times 2$  ANOVA with age (adolescent or adult) and environment (stable or volatile) as fixed factors. Participants' learning rate in each environment was the dependent variable. Inconsistent with our first hypothesis, there was no interaction between environment and age group ( $F_{(1,172)} = 0.01, p = .929$ ). Unexpectedly, we also found no main effect of environment on participants' learning rate ( $F_{(1,172)} = 0.81, p = .369$ ), meaning we did not find evidence that participants adjusted their learning rates between the stable and volatile environments as the optimal RL agent did. We did, however, find a main effect of age group ( $F_{(1,172)} = 5.26, p = .023, \eta^2 = 0.02$ ; see Figure 5.1), with adolescents using a higher learning rate compared to adults. These findings are inconsistent with our first hypothesis as we did not find evidence that participants adjusted their learning rate between the two foraging environments, nor did adolescents exhibit a greater adjustment of this parameter compared to adults.

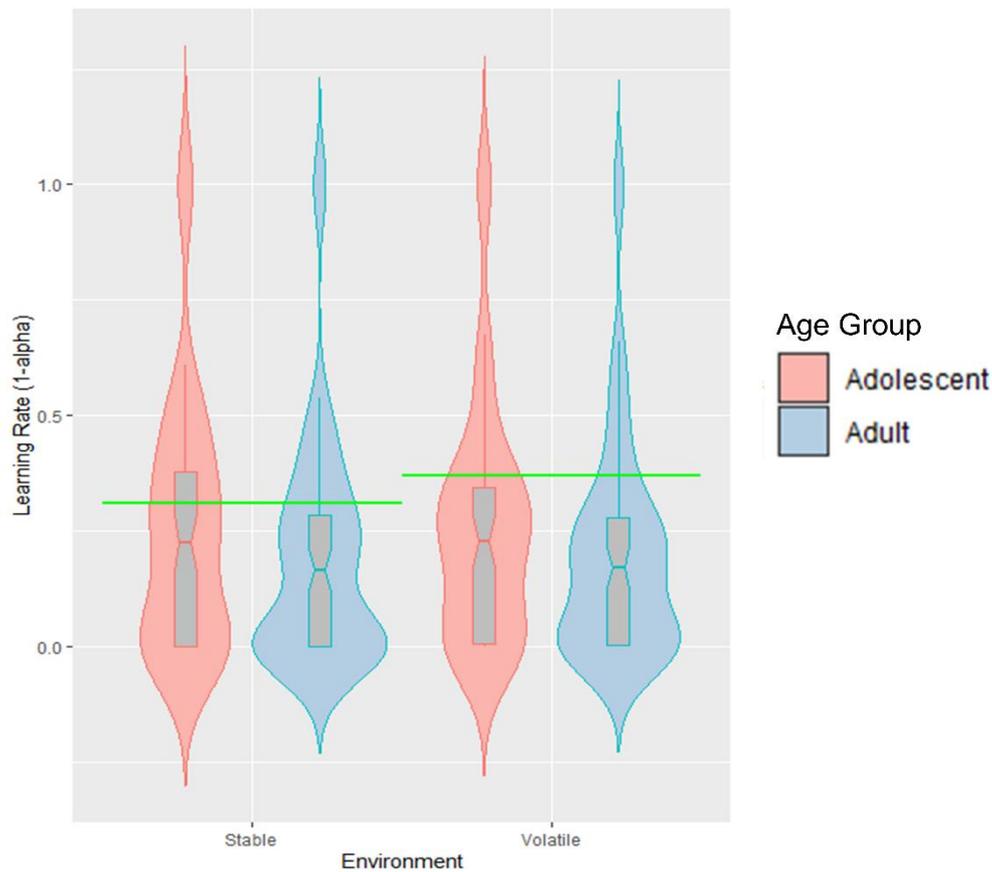


Figure 5.1: Violin plot demonstrating participants' learning rate in each environment. The middle bar of each boxplot indicates the mean for each age group and the upper and lower bars indicate the upper and lower quartiles, respectively. Green bars indicate the optimal learning rate.

Although we pre-registered a hypothesis that greater anxiety would be associated with less optimal adjustment of their learning rate to environmental volatility, our previous analysis provided minimal evidence that participants calibrate their learning rate to environmental volatility. Nevertheless, we implemented our pre-registered analysis to examine whether anxiety was associated with the extent to which participants adjusted their learning rates between environments. We first calculated the optimal amount that the learning rate should be adjusted between the two environments, which was the difference between the optimal RL agent's learning rate in the stable and volatile environments (see Figure 5.1). We then calculated how much participants adjusted their own learning rate between these environments (i.e., the difference in their learning rates between the stable and volatile environment). The difference between how much participants adjusted their learning rate and how much the

optimal RL agent adjusted its learning rate was entered into an ordinary least squares (OLS) regression to test whether the degree to which participants optimally adjusted their learning rate was associated with anxiety (Hypothesis 2). This analysis did not demonstrate evidence that the degree to which participants optimally adjusted their learning rate between environments was predicted by anxiety ( $\beta = -0.10$ ,  $t = 1.28$ ,  $p = 0.201$ ) or age group ( $\beta = -0.01$ ,  $t = 0.09$ ,  $p = 0.926$ ). The overall model was not significant ( $F_{(2,174)} = 0.83$ ,  $p = 0.439$ ,  $R^2 = 0.01$ ).

For our third hypothesis, we replicated previous results (Lloyd et al., 2021b) demonstrating that adolescents have higher leaving thresholds (i.e., explore more) compared to adults ( $t(179) = 3.07$ ,  $p = .002$ , Cohen's  $d = 0.46$ ; see Figure 5.2). Adolescents' heightened exploration, relative to adults, also translated to more optimal foraging behaviour (as indicated by the green line on Figure 5.2). However, we cannot attribute adolescents' heightened exploration to their learning rate, which was also heightened compared to adults (see analyses of hypothesis 1), as we did not find an association between participants' leaving threshold and their learning rate in the stable environment ( $r(176) = -0.12$ ,  $p = .109$ ).

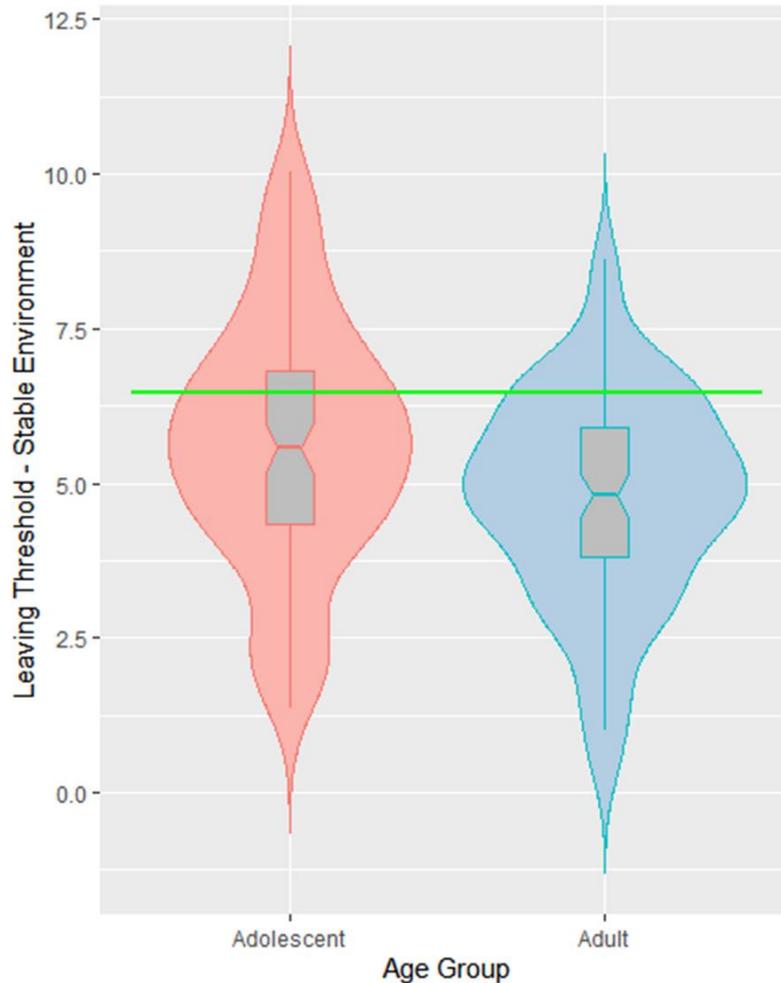


Figure 5.2: Violin plot demonstrating that adolescents had a significantly higher leaving threshold (i.e., explored more) compared to adults. Boxplots indicate the mean leaving threshold along with upper and lower quartiles. Horizontal green bar indicates the optimal leaving threshold in the stable environment.

#### 5.4.2. Exploratory Analyses

To understand whether participants' exploration behaviour in stable and volatile foraging environments could be explained by other computational parameters included in our model (i.e.,  $c$ , measuring participants' overexploitation bias and  $\beta$ , measuring how stochastic participants' choices were), we inspected our simulations of the optimal RL agent and the combination of parameter values that yielded the highest number of rewards (see Supplementary Materials). These simulations demonstrated that it was optimal to have more stochastic behaviour (i.e., lower values on the  $\beta$  parameter; see green line in Figure 5.3) in the volatile environment relative to stable environment. We subsequently conducted analyses on

estimates of participants'  $\beta$ , where lower values indicate greater decisional stochasticity, to examine whether participants showed a similar adjustment of stochasticity between foraging environments as shown by our simulations. Consistent with an optimal RL agent, we found a significant main effect of environment ( $F_{(1,172)} = 20.04, p < .001, \eta^2 = 0.03$ ), as participants exhibited more stochastic behaviour in the volatile condition relative to the stable condition (see Figure 5.3). This result was confirmed using simple main effects tests which demonstrated that adolescents had exhibited more stochastic behaviour in the volatile environment compared to the stable environment ( $F_{(1,172)} = 6.76, p = .011$ ), as did adults ( $F_{(1,172)} = 8.56, p = .004$ ). Further, there was a main effect of age group ( $F_{(1,172)} = 6.76, p = .010, \eta^2 = 0.03$ ), with adolescents exhibiting more stochastic behaviour relative to adults in both environments. This was confirmed using a simple main effects test which demonstrated that adolescents utilised more stochastic decision-making (i.e., lower values on  $\beta$ ) compared to adults in the stable environment ( $F_{(1,172)} = 7.49, p = .007$ ) and the volatile environment ( $F_{(1,172)} = 5.72, p = .018$ ). We did not find an interaction between foraging environment and age group ( $F_{(1,172)} = 0.47, p = .495$ ). Exploratory analyses of parameter estimates for  $c$ , which measured participants' overexploitation bias, did not reveal any between environment differences in our behavioural data (see Supplementary Material).

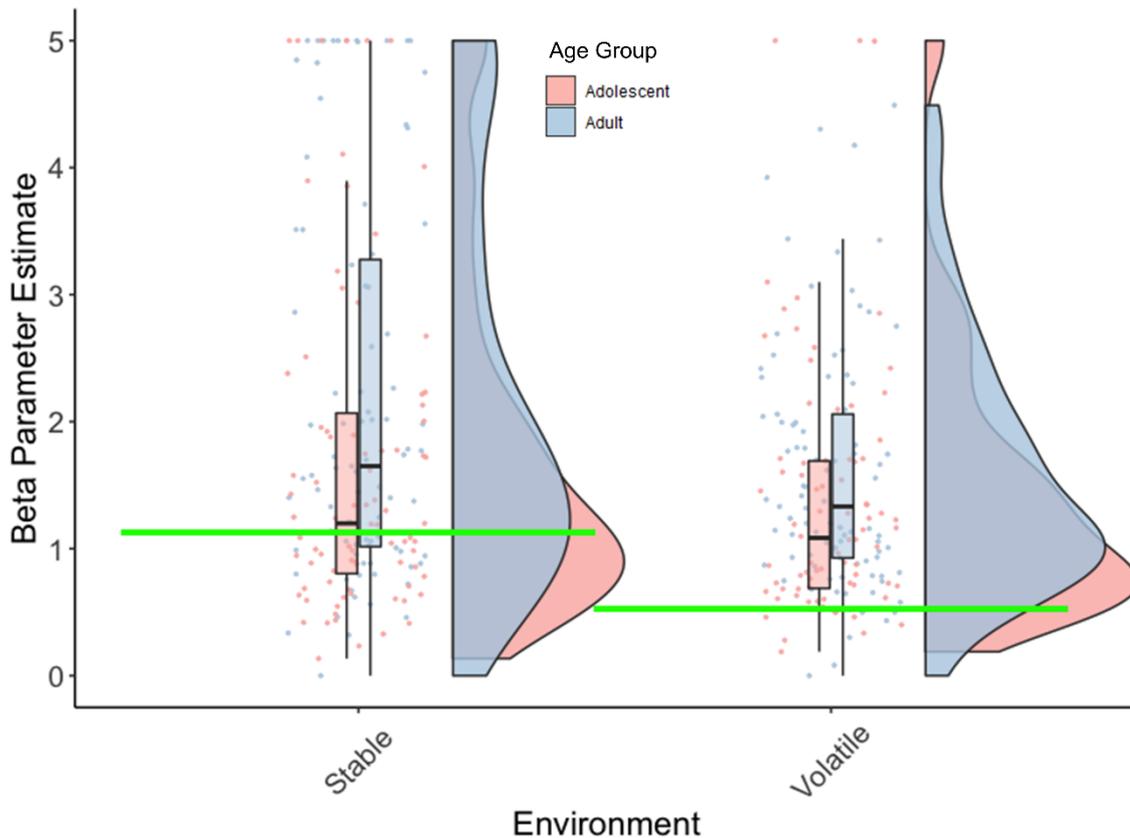


Figure 5.3: Raincloud plot of estimates for the  $\beta$  parameter, which measures stochasticity in participants' decision-making, with lower values on the y-axis denoting more stochastic choices and higher values denote more deterministic choices. Horizontal green lines indicate the optimal value of beta for each environment.

To examine whether decision stochasticity could account for participants' exploration in the stable and volatile environments, we conducted bivariate correlations on the  $\beta$  parameter and participants' leaving thresholds in both environments. We did not find evidence that leaving thresholds were associated with the degree of stochasticity in participants' choices (i.e., estimates of  $\beta$ ) in the stable environment ( $r(176) = 0.14, p = .065$ ), nor the rich portion of the volatile environment ( $r(176) = -0.02, p = .818$ ). We did, however, find that higher leaving thresholds were associated with less stochastic choices in the poor portion of the volatile condition ( $r(176) = 0.30, p < .001$ ).

Mirroring our analyses of participants' learning rate, we conducted exploratory analyses to examine whether the degree to which participants optimally adjusted their decision stochasticity (i.e., their estimates of  $\beta$ ) between environments was predicted by anxiety and age

group. We calculated the difference between the value of  $\beta$  used by the optimal RL agent in the stable and volatile environments (i.e., the optimal level of adjustment). We then calculated the difference between the estimates of  $\beta$  for participants in the stable and volatile environments (i.e., participants' adjustment). The difference between these two quantities measures the deviation from optimality of the participant's adjustments in  $\beta$  between one environment and the other. Using this value as an outcome variable and anxiety and age group as predictors in an OLS regression, we found that overall model was significant ( $F_{(2,174)} = 3.51, p = .032, R^2 = 0.04$ ). The degree to which participants optimally adjusted their decision stochasticity was predicted by anxiety ( $\beta = -0.18, t = 2.43, p = .016$ ; see Figure 5.4), but not age group ( $\beta = -0.08, t = 1.10, p = 0.271$ ). These findings suggest that rather than adjusting their learning rate between stable and volatile conditions, participants adjusted the level of stochasticity in their decision-making, and that the ability to optimally adjust decision stochasticity was modulated to some extent by anxiety.

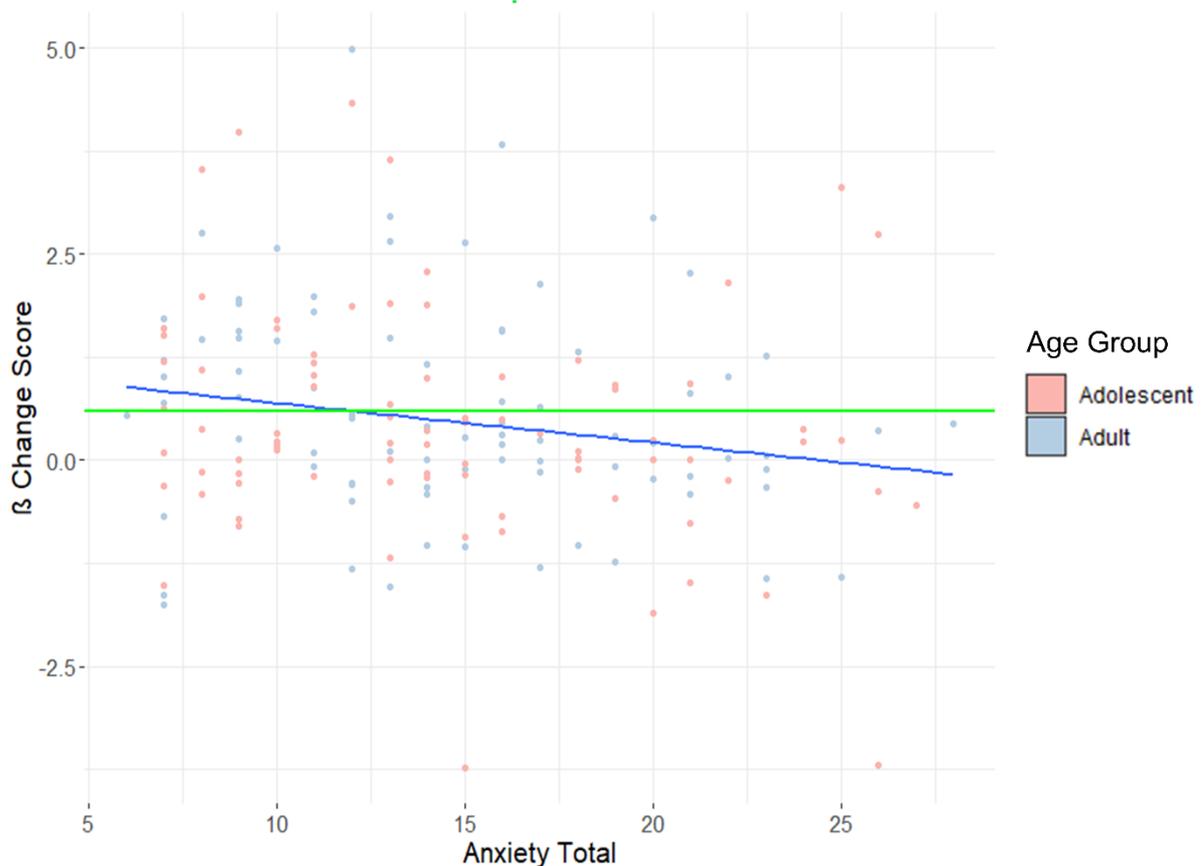


Figure 5.4: Plot demonstrating the association between anxiety (x-axis) and the degree to which participants adjusted their decision stochasticity (y-axis) between the two environments (calculated through subtracting stochasticity in the stable environment from stochasticity in the volatile environment). Values above 0 on the y-axis indicate participants had lower values on  $\beta$  in the volatile environment relative to the stable environment whereas values below 0 indicate participants who had higher values on  $\beta$  in the stable environment. The horizontal green line indicates the optimal amount of adjustment in stochasticity between stable and volatile environments, determined by our simulations.

We examined whether there were differences in the number of rewards participants collected throughout the task between age groups and environments, and whether participants' ability to collect rewards was affected by anxiety using an ANCOVA, with anxiety entered as a covariate. We found that anxiety impaired participants' ability to accrue rewards while foraging, as there was a main effect of anxiety on number of rewards collected ( $F_{(1,173)} = 4.27$ ,  $p = .040$ ,  $\eta^2 = 0.02$ ; see Figure 5.5), though we did not find an interaction between anxiety and environment ( $F_{(1,173)} = 0.80$ ,  $p = .373$ ). While adolescents had significantly higher leaving thresholds in the stable environment relative to adults (see Figure 5.2), this did not translate to

an effect on the number of rewards collected as we did not find a main effect of age ( $F_{(1,173)} = .57, p = .450$ ). This null finding may be attributed to the noisy reward schedule utilised in the foraging task, which introduced a degree of chance into the number of reward participants accumulated during the task. We did not find evidence of an interaction between age and environment ( $F_{(1,173)} = 0.02, p = .890$ ). There was, however, a main effect of environment, as participants collected a higher number of rewards in the volatile foraging environment compared to the stable environment ( $F_{(1,173)} = 12.51, p < .001, \eta^2 = 0.02$ ; see Figure 5.6).

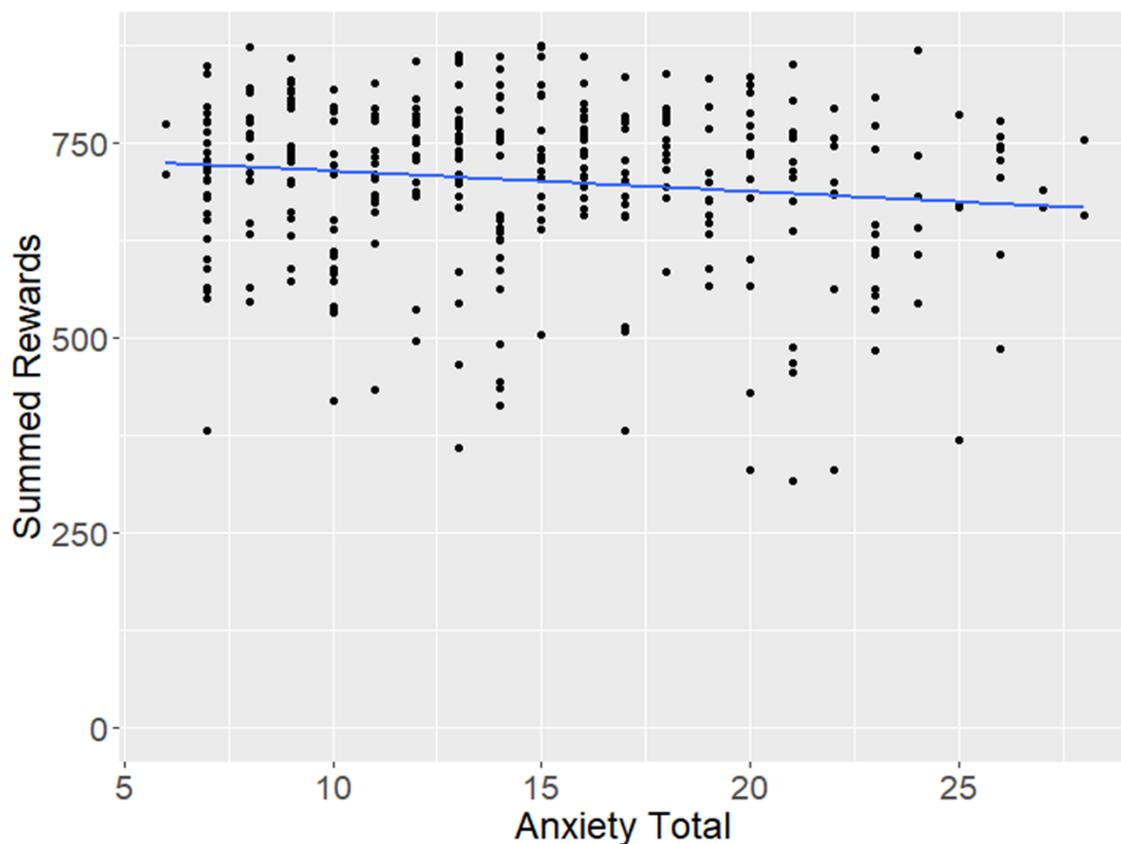


Figure 5.5: Plot demonstrating the association between anxiety and summed rewards collected in both environments.

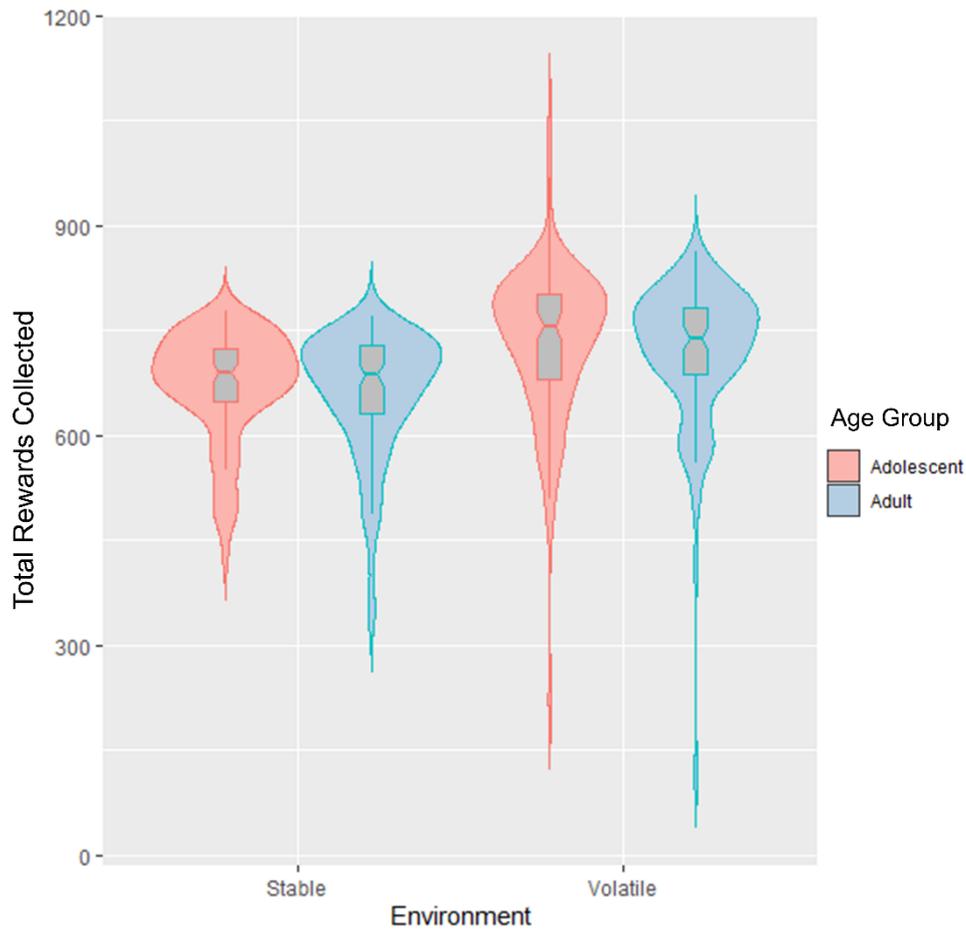


Figure 5.6: Violin plot demonstrating the total number of rewards accrued in each environment. The boxplot demonstrates the upper and lower quartile.

We next examined whether the degree to which participants adjusted the amount of stochasticity in their decision-making was associated with the absolute difference in the number of rewards they collected in the stable and volatile foraging environments. An optimal forager who adjusts their decision-making between the stable and volatile environments should minimise the difference in the number of rewards collected between environments (which afforded equal opportunity to collect rewards, see Supplementary Materials), and therefore values closer to zero on the absolute difference in rewards collected between environments reflects better performance. We found a negative association between the absolute difference in rewards collected between the environments and the degree to which participants adjusted their stochasticity between environments ( $r(180) = -0.24, p = .001$ ). Greater values on the variable measuring the degree to which participants adjust their stochasticity means

participants had lower values on  $\beta$  (i.e., exhibited more stochastic choices) in the volatile environment relative to the stable environment (see Figure 5.4). Therefore, participants who utilised more stochastic decision-making in the volatile environment relative to the stable environment minimised the disparity in the number of rewards collected between the two environments. This interpretation was further supported by analysis of the *total* number of rewards collected across both environments, as the degree to which participants adjusted the amount of stochasticity in their decision-making was positively associated with the total number of rewards collected ( $r(180) = 0.23, p = .002$ ). Therefore, participants who utilised more stochastic decision-making in the volatile environment relative to the stable environment collected more rewards throughout the task.

## 5.5. Discussion

The present study examined age-related differences in adolescents' and adults' ability to flexibly adjust their foraging behaviour between conditions of stability and volatility. We designed our volatile and stable environments such that the optimal strategy was to use a higher learning rate in the volatile environment. However, contrary to our hypotheses, we found no evidence that either adolescents or adults exhibited this calibration of learning rate to environmental volatility, nor that this ability was impaired by anxiety. We did, however, find that adolescents had a higher learning rate compared to adults in both environments. Moreover, we replicated findings that adolescents explore more than adults in stable foraging environments (Lloyd et al., 2021b). Using simulation, we demonstrated that there is an optimal strategy to utilise a more stochastic decision-making policy in volatile relative to stable environments. In contrast to our analyses of participants' learning rate, we found that participants calibrated their decisional stochasticity to the prevailing volatility of the environment, consistent with an optimal RL agent. Further, compared to adults, adolescents

utilised a more stochastic decision-making policy compared to their older counterparts in both stable and volatile environments. Together, these findings identify stochastic responding as a key computational mechanism that optimises human foraging in volatile environments, along with demonstrating the reduction of decisional stochasticity from adolescence to adulthood.

We originally predicted that participants would adopt a higher learning rate in the volatile environment relative to the stable environment, although this prediction was not supported by our data. This prediction derived from research that has attributed adults' successful decision-making in conditions of volatility to how much weight they give to recent information relative to more historic feedback (i.e., adjust their learning rate; Behrens et al., 2007; Browning et al., 2015), though these studies have not examined this ability in context of patch foraging. An alternative account has suggested that optimal exploration in conditions of volatility, where there is greater uncertainty about reward contingencies (Bland & Schaefer, 2012), is associated with introducing stochasticity into decision-making (Gershman, 2018). The use of a stochastic decision-making policy, known as 'random exploration', leads the individual to occasionally divert from the option that, to their knowledge, should have the highest expected value (Wilson et al., 2021). In conditions of volatility, the reward probabilities associated with actions change over the duration of the environment. Such changes affect which action objectively has the highest expected value. Using a decision-making strategy where the individual occasionally diverts from selecting the option they perceive to have the highest expected value (i.e., exhibiting more stochastic behaviour), the individual can identify changing stimulus-outcome contingencies, and therefore which action has the highest objective expected value, faster than if they were to deterministically select the option that they perceive to have the highest expected value (Gershman, 2018). Our simulations provided formal support for this account, as we demonstrated that an optimal RL agent utilised greater stochasticity in the volatile foraging environment relative to the stable one. Consistent with this account of

optimal exploration, we found that participants relied on a more stochastic decision-making policy in conditions of volatility relative to stability.

Previous research has demonstrated that adolescents and adults can utilise stochastic exploration strategically. For example, when participants have longer to explore their surroundings and therefore the feedback from exploring can be informative for later decisions, participants utilise stochastic exploration strategies to a greater degree compared to conditions where they have less time and exploration has less utility (Somerville et al., 2017; Wilson et al., 2014). Our findings contribute to research on decision-making during the explore/exploit trade-off by demonstrating that adolescents' and adults' use of stochastic exploration is strategically adjusted to the prevailing volatility of the environment.

Importantly, however, we found that adolescents' and adults' ability to increase the amount of stochasticity from stable to volatile environments is modulated by anxiety. It has been demonstrated that anxiety can also impair other features of decision-making, such as individuals' ability to increase their learning rate from stable to volatile conditions (Browning et al., 2015). One possible reason why we did not find that anxiety affected the learning rate in the current study is that Browning and colleagues (2015) utilised an instrumental learning task with aversive feedback, specifically an electric shock, whereas we utilised a foraging task with reward feedback. Previous research has causally implicated dopamine in controlling the parameter measuring stochasticity in conditions of positive reward feedback, but not for aversive feedback (Eisenegger et al., 2014). As anxiety is associated with impairments to the dopaminergic reward system (Lee et al., 2015), this could explain why we observed the effects of volatility on participants' stochasticity rather than their learning rate. Nevertheless, we identify a novel component of decision-making that is impacted by anxiety: decisional stochasticity. Notably, we found participants who calibrated their stochasticity to the prevailing environmental volatility reported some (i.e.,  $> 0$ ) anxiety, which is consistent with accounts

that propose psychopathologies arise from adaptive processes (Del Giudice & Haltigan, 2021). On these accounts, moderate levels of anxiety may support more optimal decision-making than chronic levels or an absence of these traits (Paulus & Yu, 2012); an effect we found in the current study. These findings may have clinical implications for understanding how anxiety impairs individuals' ability to adjust their behaviour during times of change, and may be targeted for intervention in populations more likely to present with symptoms of anxiety, such as individuals with experience of childhood adversity (Elmore & Crouch, 2020). Indeed, one recent study has demonstrated that childhood adversity is associated with reduced exploration and underweighting reward feedback (Lloyd et al., 2022). Future research could examine whether anxiety moderates the association between childhood adversity and suboptimal foraging strategies, such as the inability to adjust decision stochasticity between stable and volatile foraging environments we observed in the current study.

During adolescence, the individual undergoes significant physical, psychological, and neurobiological changes (Sawyer et al., 2018; Somerville et al., 2010). These developments are complemented by periods of external transition, such as beginning a new school and developing new friendships (Eccles et al., 1993; Blakemore, 2018). Consistent with previous research (e.g., Decker et al., 2015; Javadi et al., 2014; Palminteri et al., 2016), we found that adolescents exhibited a more stochastic decision-making policy relative to adults. The increased use of a stochastic decision-making policy at this point in the lifespan could be adaptive, as it could support adolescents to successfully negotiate the volatility associated with this developmental period compared to adults, who experience relative stability in their social and physical environments (Somerville et al., 2010).

Adolescents' reliance on more stochastic exploration strategies relative to adults could also account for the increase in novelty-seeking behaviours during this developmental period (Romer et al., 2017; Steinberg 2008), though we did not find an association between greater

choice stochasticity (i.e., lower values on  $\beta$ ) and heightened exploration in the current study. Nevertheless, stochastic exploration strategies lead the individual to trial a greater range of behaviours (Gopnick, 2020). As such, adolescents' increased novelty-seeking behaviours compared to other age groups may rather be associated with the exploration strategies they utilise to negotiate their changeable environment. Adolescents' novelty-seeking may be further heightened by the hyperactivity of the dopaminergic reward system during this developmental period (van Duijvenvoorde et al., 2017), which is positively associated with random exploratory behaviours (but not learning rate; Cinotti et al., 2019; Gershman & Tzovaras, 2018). Compared to accounts that attribute adolescents' novelty-seeking to an increase in risk-taking (Steinberg, 2008), our findings suggest that the rise in some novel and potentially risky behaviours may be associated with the strategies adolescents use to explore their changeable, volatile environment.

Our findings also contribute to the growing literature examining the development of foraging strategies across the lifespan (Bach et al., 2020; Lloyd et al., 2021b). We replicated findings (Lloyd et al., 2021b), which show that adolescents explore more than adults and that this behaviour is closer to the optimal leaving threshold prescribed by Marginal Value Theorem (Charnov, 1976). Further we demonstrate adolescents' more optimal behaviour is reflected in a computational parameter that measures this overexploitation bias, which is lower in this age group relative to adults (see Supplementary Materials). Moreover, we found that adolescents had a higher learning rate relative to adults, consistent with some previous studies (e.g., Davidow et al., 2016). These findings indicate that adolescents' heightened novelty-seeking compared to other age groups is not exclusively associated with negative outcomes (such as dangerous driving or substance misuse; Chein et al., 2011; Willoughby et al., 2013), as adolescents were better than other age groups at determining the appropriate amount of exploration in the explore/exploit trade-off. Indeed, adolescents' more optimal foraging

compared to adults is consistent with previous work that has demonstrated adolescents are better at harvesting rewards from stable environments than their older counterparts (Davidow et al., 2016). As such, the rapid development of the dopaminergic reward system at the onset of puberty (Shulman et al., 2016), which promotes exploratory behaviour (Costa et al., 2014; Gershman & Tzovaras, 2018), may confer benefits in this context of decision-making.

There are some important limitations to consider with this study. One design feature of foraging tasks is that the trial structure is not fixed as in other decision-making paradigms, as participants selected how long they remained in each patch (Gabay & Apps, 2021). As such, participants may have differed the number of patches in the rich and poor portions of the volatile environment, which we found to be the case as participants visited a greater number of patches in the poor portion relative to the rich portion of the volatile environment (see Supplementary Materials). This variability in the number of patches visited in the rich and poor portions of the volatile environment could have affected participants' estimation of the average reward rate in this condition, as it may have biased participants to believe there were a greater number of poor patches in the volatile environment. Yet, despite this variability in number of patches visited, we observed significant differences between participants' leaving thresholds in the poor and rich portions of the volatile environment, indicating that participants did adjust their behaviour between different portions of the volatile environment (see Supplementary Materials).

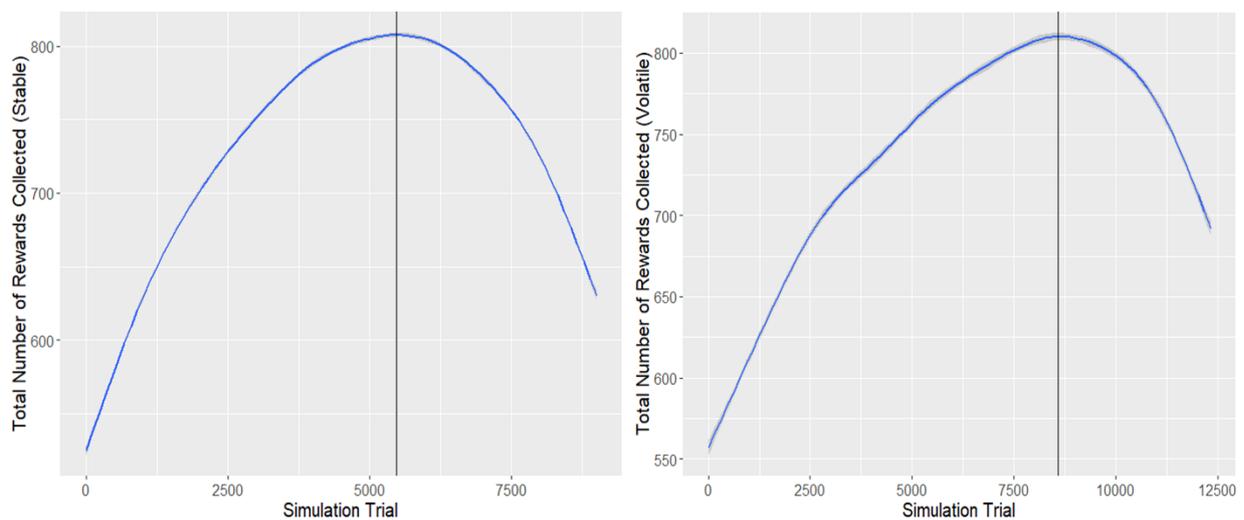
In sum, the present study has demonstrated that adolescents and adults are able to flexibly adjust their exploration strategies between stable and volatile foraging environments. Specifically, both age groups calibrated the amount of stochasticity in their decision-making to the volatility of the environment, consistent with recent accounts of optimal exploration under volatile conditions (e.g., Gershman 2018; Wilson et al., 2021). However, we found that adolescents utilised a more stochastic decision-making strategy in both environments relative

to adults, which may be adaptive during this developmental period as the individual experiences significant social and psychological changes. Importantly, the adaptive ability to adjust stochasticity in decision-making was modulated by anxiety, highlighting features of decision-making that can be negatively affected by disorders associated with anxiety. Consistent with our pre-registered predictions, we also replicated findings that adolescents explore more than adults while foraging. However, inconsistent with our predictions, participants deviated from an optimal RL model insofar as they did not adjust their learning rate between stable and volatile environments. Further, we did not find evidence that the ability to adjust one's learning rate while foraging was modulated by anxiety. Nevertheless, our findings contribute to our understanding of how exploration is employed strategically by adolescents and adults to negotiate their surroundings, along with the computational mechanisms associated with foraging in volatile environments.

## 5.6. Supplementary Material

### 5.6.1. Simulating Stable and Volatile Foraging Environments

We designed the foraging environments such that participants could collect a similar number of rewards in both the stable and volatile environments, which allowed us to compare performance between the two environments. The simulation method to determine the maximum number of rewards that could be collected in the stable environment was similar to that reported in Lloyd et al. (2021b). For the volatile environment, we systematically varied combinations of leaving thresholds for the rich and poor portions of the patch (which ranged from 1-12). Through adopting this approach, we could identify the combination of leaving thresholds that allowed the simulated agent to collect the maximum number of rewards in that environment. These simulations demonstrated that the stable and volatile environments afforded participants the opportunity to collect a similar number of rewards if they adopted an optimal foraging strategy (see Supplementary Figure 5.1).



*Supplementary Figure 5.1. Plots demonstrating the total number of points available to collect in the stable environment (left) and volatile environment (right). Each environment affords the participant the opportunity to collect a maximum of approximately 800 rewards (i.e., apples).*

### 5.6.2. Parameter Recovery and Model Comparison

*Parameter Recovery and Model Comparison for the Single Learning Rate Model*

We conducted parameter recovery to ensure that the model could provide accurate estimates of each parameter. To evidence this, we hardcoded parameter values into simulations of foraging behaviour in the stable and volatile foraging environments. Parameter estimation was then conducted on the simulated data, and we examined the correlations between these hardcoded and estimated parameters. We ran parameter recovery testing combinations of values that ranged across the parameter space (alpha {0,1} in increments of 0.1; c {-1, 3} in increments of 0.5 and beta {0, 5} in increments of 0.5). We found significant positive correlations between hardcoded and parameter estimates for alpha ( $r(538) = 0.47, p < .001$ ),  $c$  ( $r(538) = 0.25, p < .001$ ) and  $\beta$  ( $r(538) = 0.27, p < .001$ ). However, during this process we observed a significant correlation between parameter estimates for alpha and beta ( $r(538) = 0.53, p < .001$ ). While correlations between parameters are not ideal, these are sometimes unavoidable in formal models (Wilson & Collins, 2019). As such, we were careful not to interpret correlations between these parameters in the parameter estimates of participants' behaviour (Wilson & Collins, 2019).

As we found correlations between parameter estimates for  $\alpha$  and  $\beta$ , one possibility is that the model mistook the effects of one parameter on the data for the effects of another parameter. For example, it could be the case that lower values on  $\alpha$  produced patterns in the data that had similar expected effects to lower values of  $\beta$ , which could incorrectly bias the parameter estimates. One way to exclude the possibility that parameter values are biasing one another is to fit models to the behavioural data where one of the correlated parameters is fixed between participants, which can be used to examine whether the same effects can be replicated in the other, unfixed parameter (Daw, 2011). As such, we fitted two additional models to the behavioural data: one where  $\alpha$  was fixed across participants and another where  $\beta$  was fixed across participants. The fixed values of  $\alpha$  and  $\beta$  in their respective models were taken from the population estimates of these parameters in the model where neither parameter is fixed. We

then fitted these two additional models to the behavioural data to estimate the remaining free parameter (i.e., in the model where  $\beta$  was fixed we estimated values of  $\alpha$  and  $c$  and in the model where  $\alpha$  was fixed we estimated values of  $\beta$  and  $c$ ).

Notably, even when all participants' data was fit with one shared  $\beta$ , which was unable to vary across participants, we replicated the pattern of results for the learning rate parameter. Specifically, we found in the model where we assumed  $\beta$  was shared that there was no main effect of environment ( $F_{(179)} = 0.04, p = .847$ ), consistent with the analysis reported in the main text. Replicating the other findings reported in the main text, we found a main effect of age group on  $\alpha$  ( $F_{(179)} = 4.13, p = .044, \eta^2 = .02$ ), with adolescents having a higher learning rate relative to adults and no interaction between age group and environment ( $F_{(179)} = 0.04, p = .847$ ). Moreover, even when all participants' data were fit with one shared  $\alpha$ , we replicated the pattern of results for the  $\beta$  parameter measuring decisional stochasticity. Results from an ANCOVA, with anxiety entered as a covariate, demonstrated a main effect of age ( $F_{(179)} = 7.26, p = .008, \eta^2 = .03$ ) and environment ( $F_{(179)} = 15.70, p < .001, \eta^2 = .02$ ) on estimates of the  $\beta$  parameter. We also found that anxiety impaired participants' ability to adjust their behaviour as there was an interaction between anxiety and environment ( $F_{(179)} = 5.92, p = .016, \eta^2 = .01$ ), along with a main effect of anxiety ( $F_{(179)} = 5.09, p = .025, \eta^2 = .02$ ) though there was no interaction between age group and environment ( $F_{(179)} = 1.91, p = .169$ ). Together, these findings suggest that although parameter estimates of  $\beta$  and  $\alpha$  were correlated, they did not obscure the main effects of environment on estimates of  $\beta$  or age group on estimates of  $\beta$  and  $\alpha$ . As such, we considered it valid to interpret the results of the model where no parameters were shared (Daw, 2011; Wilson & Collins, 2019).

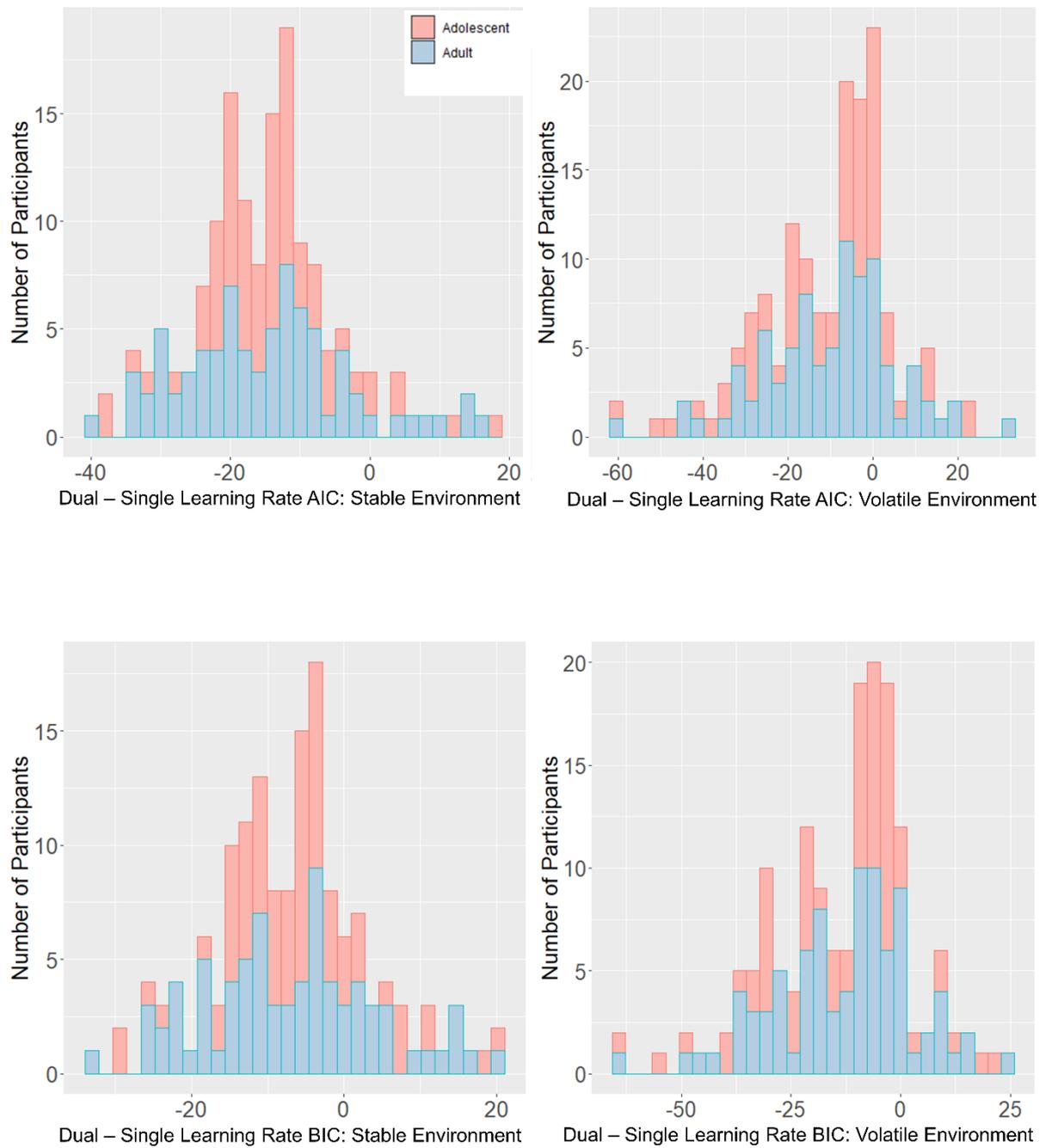
#### *Model Comparison Between Single and Dual Learning Rate Models*

We also compared the formal model described in the main text to a model which separated learning rates for positive and negative prediction errors (Garrett & Daw, 2020). In this model, we calculated whether the reward feedback experienced by participants is better than expected (i.e., is a positive prediction error), or poorer than expected (i.e., a negative prediction error) using the following formula:

$$\delta_i \leftarrow s_i / \tau_i - \rho_i$$

When delta ( $\delta$ ) is greater than 0, participants update their estimation of the average reward rate according to Equation (1) described in the main manuscript and the weight placed on this feedback is controlled by the free parameter  $\alpha_{\text{pos}}$ . When delta is less than 0, this feedback is integrated into the participants' estimate of the average reward rate, controlled by the free parameter  $\alpha_{\text{neg}}$ . The participants' estimate of the average reward rate is then entered into the SoftMax function described in Equation (2) in the main manuscript.

We conducted model comparison for each of these models, calculating the AIC and BIC of each model (Wilson & Collins, 2019). Lower values on these criteria indicate the model is a better fit to the data (Daw, 2011). Results of model comparison of the AIC and BIC in each foraging environment demonstrated overwhelming support for the simpler model that included a single learning rate for all prediction errors (see Supplementary Figure 5.2). As such, we proceeded with our analyses using this winning model.

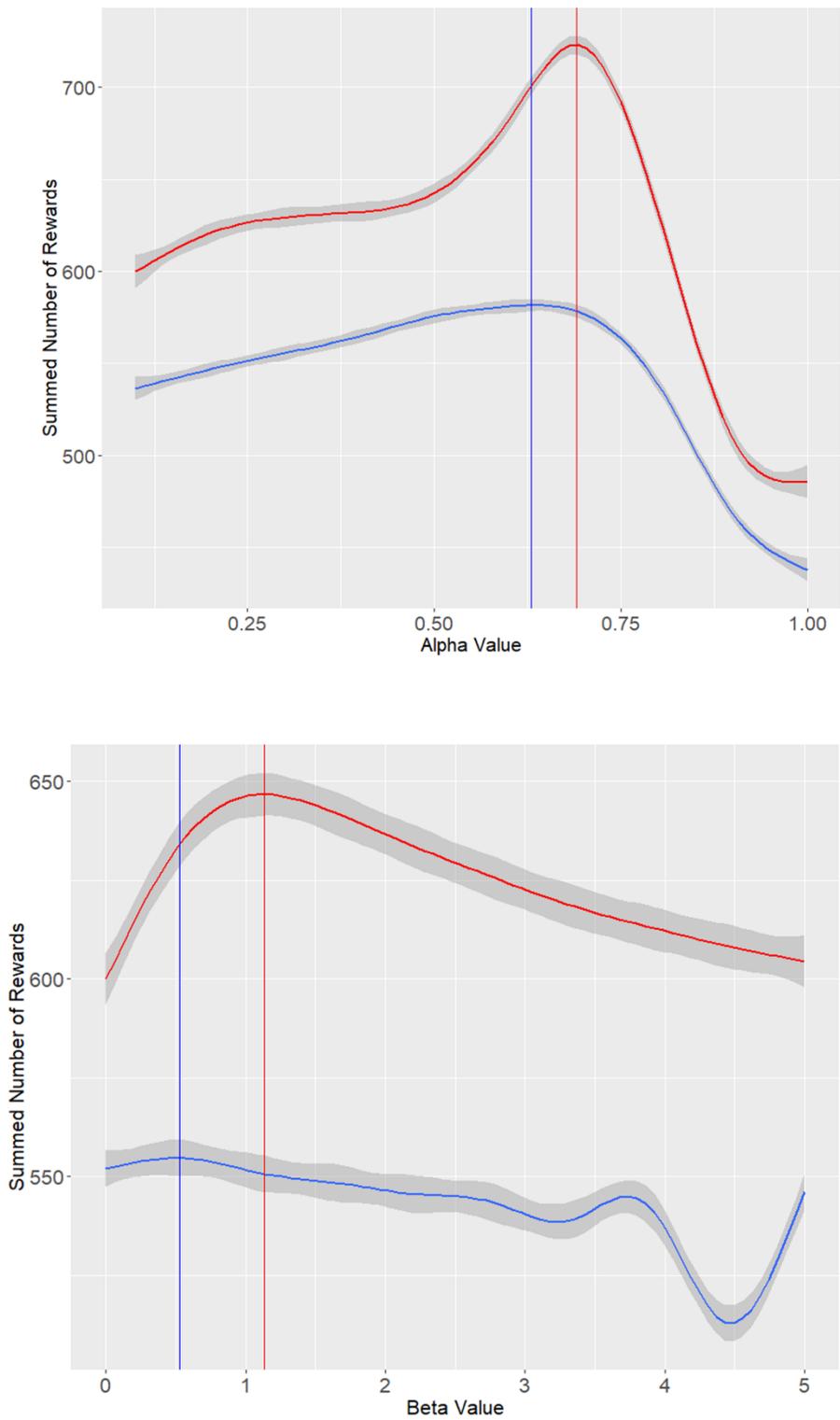


Supplementary Figure 5.2: Results of model comparison for each participant. Values below zero indicate the model with a single learning rate was a better fit to the data than the model with separate learning rates for positive and negative feedback. The y-axis denotes the number of participants who are represented by each value on the x-axis.

### 5.6.3. Simulating Optimal Parameter Values

As we were interested in how participants' behaviour compares to an optimal reinforcement learning (RL) agent, we simulated task behaviour systematically varying the combination of parameter values to identify which parameters yielded the greatest number of rewards. We simulated an RL-informed forager's behaviour in the stable and volatile foraging environments, conducting a grid search across different combinations of free parameters within the ranges:  $\alpha \{0,1\}$  in increments of 0.1,  $c \{-1,3\}$  in increments of 0.5 and  $\beta \{0,5\}$  in increments of 0.5. Here, we considered a parameter value optimal only if it maximised the reward value in a way that was as robust as possible to variations in the other two parameters. Therefore, the values we selected as our optimal parameter values were those that maximised the summed rewards across the values of the other two parameters (for example, which value of  $\alpha$  yielded the highest number of rewards across all possible values of  $\beta$  and  $c$ ). These simulations indicated that the optimal alpha was 0.69 for the stable environment and 0.63 for the volatile environment (see Supplementary Figure 5.3). Note, as we operationalised alpha as the complement of the learning rate in traditional learning rate models (i.e., our learning rate is equal to  $1-\alpha$ ), higher values of alpha in our model indicate the participant utilises a lower learning rate and vice versa.

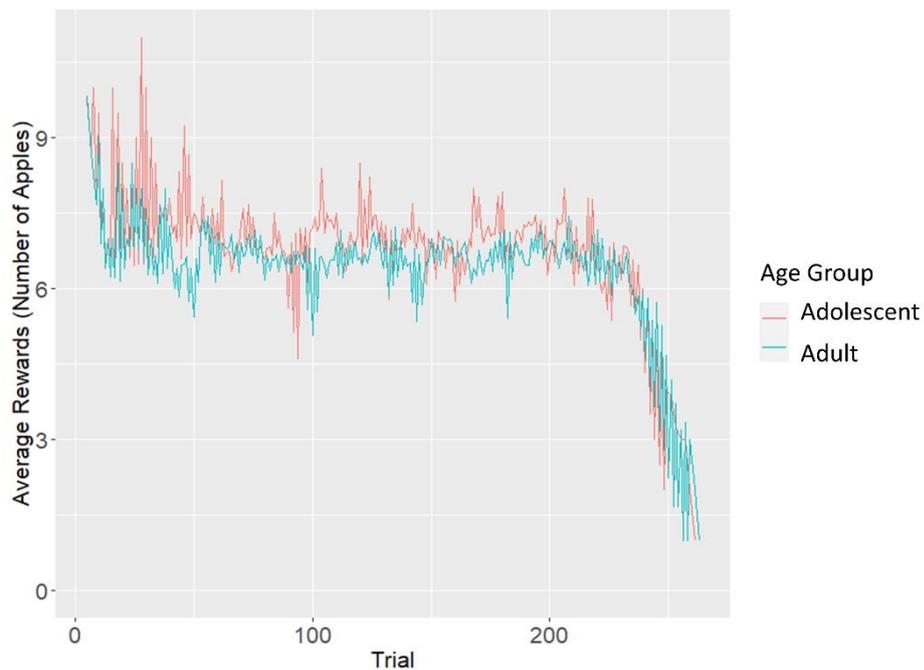
As we did not find any difference between participants' learning rate between the stable and volatile environments, we inspected the values on the  $\beta$  parameter that produced optimal behaviour. Lower values on  $\beta$  indicate more stochastic behaviour, whereas higher values indicate more deterministic behaviour. Our simulations demonstrated that the optimal values on the beta parameter were 1.13 for the stable environment and 0.53 for the volatile environment (see Supplementary Figure 5.3). As such, it was optimal to utilise more stochastic behaviour in the volatile environment relative to the stable environment.

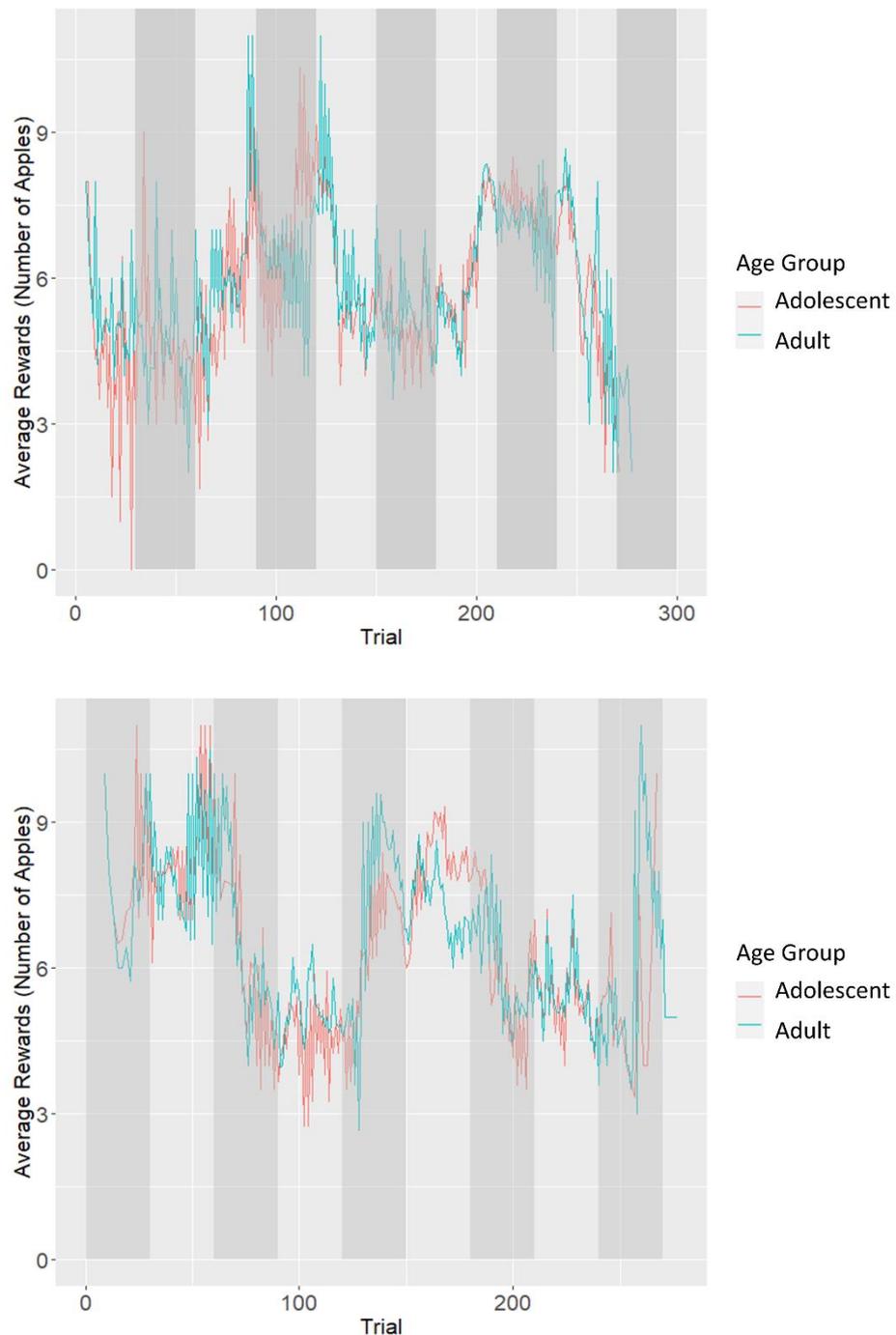


Supplementary Figure 5.3: Plots of the optimal values on alpha (top) and beta (bottom) parameters in the stable environment (red lines) and volatile environment (blue lines). Vertical lines indicate the value of the parameter that yielded the greatest number of points (in their respective colours). Gray shaded areas indicate the confidence interval around the number of rewards collected for each parameter value.

#### 5.6.4. Analyses of the Average Reward Experienced & Posterior Predictive Checks

We conducted exploratory analyses to examine whether our manipulation of the reward statistics in the volatile environment had been successful in producing differences in the rewards participants experienced during the rich and poor portion of this environment. We inspected the average number of rewards that participants experienced on each trial, which provided an indication of whether the changes to the reward statistics in the volatile environment impacted the rate at which participants accumulated rewards during the task. These plots indicate that the rate at which rewards were acquired differed during the rich and poor portions of the volatile environment, whereas the rate at which rewards were accumulated remained relatively constant in the stable environment (see Supplementary Figure 5.4). As such, these plots indicate that our manipulation of the reward statistics between the stable and volatile environments was successful.

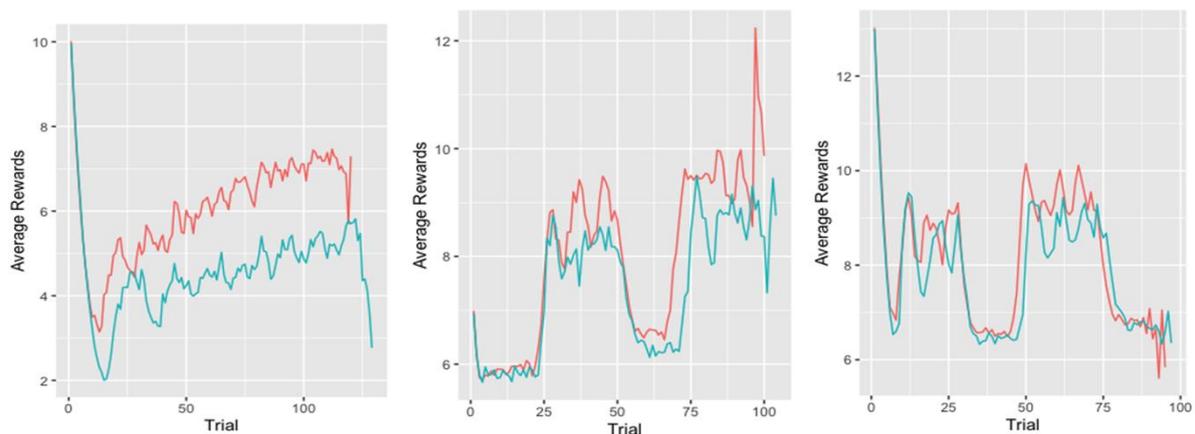




*Supplementary Figure 5.4: Participants' experience of rewards in each environment, averaged separately for each age-group. The y-axis is the number of rewards (apples) experienced on each trial. Local peaks indicate when participants travelled to a new patch and subsequently experienced a large number of rewards. The top plot is the average reward experienced in the stable environment, the middle plot is the average reward experienced in the volatile environment where the poor portion was presented first, and the bottom plot is the average reward experienced in the volatile environment where the rich portion was presented first. In the volatile environments,*

dark grey blocks indicate the rich portions of the condition and light grey blocks indicate the poor portions of the environment.

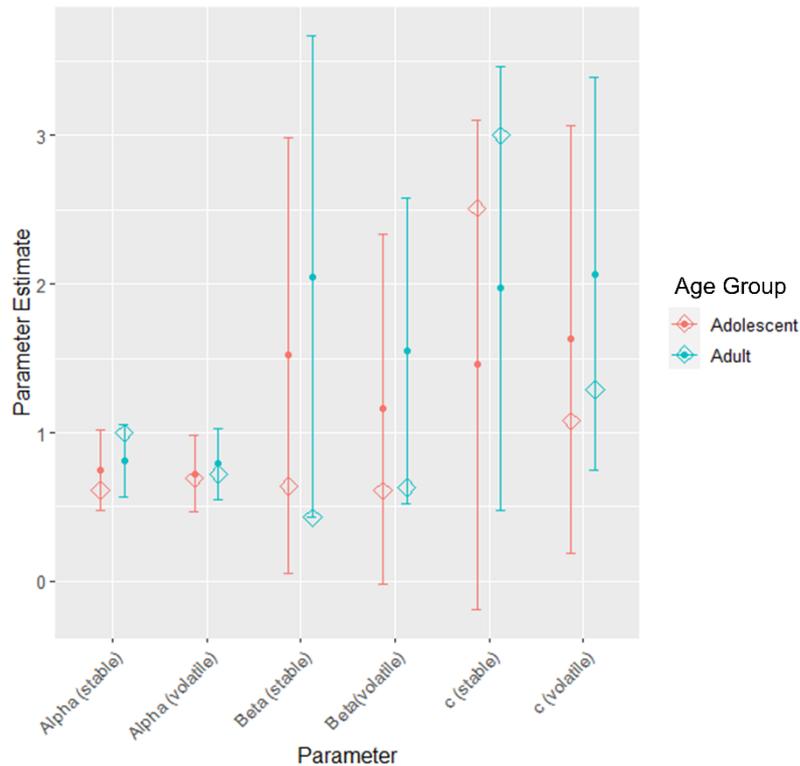
We also conducted posterior predictive checks which demonstrated that the configuration of fitted parameters reported in the main text were able to replicate the higher leaving thresholds seen in adolescents relative to adults, as well as the changes to the number of rewards accrued in rich versus poor portions of the volatile environment (see Supplementary Figure 5.5). To achieve this, we simulated the foraging behaviour of 100 agents using the mean parameter estimates from each age group and plotted this behaviour. Comparing Supplementary Figure 5.3 to Supplementary Figure 5.4, we see the same pattern of peaks and troughs in the volatile environments.



*Supplementary Figure 5.5: Simulation of participants' behaviour using mean parameter estimates for adolescents (red) and adults (blue). The y-axis indicates the average number of rewards collected on each trial (x-axis). The leftmost panel is the simulated stable environment, the middle panel is the simulated volatile environment with the poor portion encountered first and the rightmost panel is the simulated volatile environment with the rich portion encountered first.*

In a final step, we took the data generated through these posterior checks (i.e., the data plotted in Supplementary Figure 5.5) and compared the estimated parameters from the model fitting to the parameter values that were used to generate the synthetic data. This demonstrated

that median parameter estimates were within 1 SD of the values used to generate the data (see Supplementary Figure 5.6). These results suggest that simulations used to generate the plots in Supplementary Figure 5.5 (i.e., the mean values that described our behavioural data for each age group) were adequately able to describe the behavioural trends we observed in the data.



Supplementary Figure 5.6: Plot demonstrating the fixed parameter values used to simulate the synthetic data presented in Supplementary Figure 5.4 (circular points). Vertical lines indicate  $\pm 1$  SD from the participants' mean parameter estimate. Diamond points indicate the mean parameters estimated by the model when fitted to these simulated data.

### 5.6.5. Exploratory Covariate, Parameter and Leaving Threshold Analyses

We conducted a mixed ANCOVA on the parameter  $c$ , where higher values indicate more of an overexploitation bias, with anxiety entered as a covariate. We found a main effect of age ( $F_{(1,172)} = 6.94, p = .009, \eta^2 = 0.01$ ), with adolescents exhibiting lower values on this parameter ( $M_{\text{stable}} = 1.46, SD_{\text{stable}} = 1.65, M_{\text{volatile}} = 1.62, SD_{\text{volatile}} = 1.44$ ) compared to adults ( $M_{\text{stable}} = 1.97, SD_{\text{stable}} = 1.49, M_{\text{volatile}} = 2.07, SD_{\text{volatile}} = 1.32$ ). This finding is consistent with

our analyses of participants' leaving thresholds in the stable environment, where we found that adolescents explored more than adults (see Figure 5.2). However, we found no differences on this parameter between the two foraging environments ( $F_{(1,178)} = 1.19, p = .276$ ), nor did we find an interaction between age group and environment ( $F_{(1,178)} = 0.08, p = .785$ ). We found no effect of anxiety  $F_{(1,172)} = 0.01, p = .930$ ).

To examine whether participants visited a significantly different number of patches in the rich and poor portions of the volatile block, we ran a paired  $t$ -test. Results demonstrated that participants visited significantly more patches in the poor portion of this condition relative to the rich portion ( $t(174) = 8.32, p < .001$ , Cohen's  $d = 0.63$ ). It is therefore possible that participants believed the environment was primarily populated by poor quality patches and therefore used a single, lower leaving threshold, which would suggest participants did not adjust their behaviour between the rich and poor portions of the volatile environment. To exclude this possibility, we ran analyses comparing leaving thresholds between the rich and poor portions of the volatile environment. We found that participants did adjust their leaving thresholds within the volatile environment, as there was a main effect of the portion of the environment (rich or poor;  $F_{(1,173)} = 74.56, p > .001, \eta^2 = 0.05$ ). Consistent with our analysis of leaving thresholds in the stable environment, we found adolescents had a significantly higher leaving threshold in both rich and poor portions of the volatile condition relative to adults, demonstrated by a main effect of age group ( $F_{(1,173)} = 5.68, p = .018, \eta^2 = 0.03$ ). We also found a significant interaction between age group and portion of the volatile block ( $F_{(1,173)} = 6.81, p = .010, \eta^2 = 0.01$ ). Post hoc tests using the Bonferroni correction demonstrated that adolescents had a significantly higher leaving threshold in the rich portion of the volatile block compared to adults ( $p_{bonf} < .001$ ) whereas there was no significant difference between adolescents' and adults' leaving thresholds in the poor portion of the volatile block ( $p_{bonf} = 1$ ). Tentatively, this

interaction effect could indicate that adolescents exhibited greater flexibility in adjusting their behaviour between the rich and poor portions of the volatile block.

*Supplementary Table 5.1: Mean (SD) of participants' leaving thresholds in the rich portion of the volatile condition.*

	Rich	Poor
Adolescent	5.96 (2.55)	4.74 (1.60)
Adult	5.03 (1.92)	4.38 (1.51)

## **Chapter 6: Adolescents' and Adults' Exploration Behaviour is Susceptible to Social Influence While Foraging<sup>5</sup>**

### **6.1. Abstract**

Adolescence is a period of heightened susceptibility to social influence relative to childhood and adulthood. Previous research has focussed on the effects of social influence on adolescents' risky decisions. Yet, there is limited evidence regarding how peers influence exploration, which is also heightened in adolescence relative to other age groups. To address this, 116 younger adolescents (aged 13-15), 116 older adolescents (aged 16-17) and 107 adults (aged 21+) completed a patch foraging task, where they chose between exploiting an option with a known but gradually diminishing number of rewards and exploring an unknown option with a fresh distribution of rewards. During the task, participants received advice from a fictional peer, who either endorsed optimal or suboptimal rates of exploration. Our findings demonstrated that both adolescents' and adults' exploration strategies were susceptible to social influence while foraging. Using computational modelling, we demonstrate that younger adolescents had higher values on a parameter measuring the weight they placed on social influence, whereas older adolescents and adults had lower values on this parameter. These findings demonstrate how social influence affects adolescents' and adults' explore/exploit decision-making and provide insight into the risk and protective factors associated with susceptibility to social influence in adolescence.

---

<sup>5</sup> OSF project page: <https://osf.io/3t2wa/>

## 6.2. Introduction

Foraging for resources is a near-ubiquitous behaviour across the animal kingdom (Mobbs et al., 2018). The ability to forage successfully has been directly linked with humans' adaptive needs, such as acquiring food for sustenance, and has therefore been an important context of decision-making in humans' evolutionary history (Passingham & Wise, 2012). Previous research has considered how humans use reward-learning to optimise decision-making while patch foraging (e.g., Constantino & Daw, 2015; Kolling & Akam, 2017). However, there is a paucity of research on how humans utilise social information while foraging, despite evidence that human foraging typically occurs within groups (Bird et al., 2019; Mithen, 1989). As such, the present study examined how social influence affects foraging behaviour. Specifically, we examined whether social information is utilised to a greater extent by adolescents compared to adults, as adolescence is a developmental period where individuals exhibit heightened susceptibility to social influence during risk taking (see Blakemore, 2018, for a review). Yet, there is still a limited understanding about the extent to which adolescents' exploration behaviour is also susceptible to social influence. These findings have implications for our understanding of the mechanisms involved in adolescents' ability to negotiate their surroundings during a period where the individual acquires the experience necessary for adult independence (Ellis et al., 2012; Romer et al., 2017).

Examining the development of exploration strategies is important as this behaviour has been used across human beings' evolutionary history to forage for resources necessary for survival, such as food and mates (Smith, 1983; Venkataraman et al., 2017). Patch foraging requires the individual to evaluate the trade-off between exploiting a known patch that gradually depletes of resources or incurring a time and energy cost to explore a new patch that has a fresh distribution of resources (otherwise known as the explore/exploit trade-off; Bettinger & Grote, 2016). A theorem of optimal foraging, known as Marginal Value Theorem

(MVT), prescribes that the organism should leave their current patch when the rewards expected from exploiting the patch fall below the average reward rate for the environment (Charnov, 1976). Yet, human adults are typically suboptimal foragers, exploiting patches for longer than the optimal policy would dictate (Constantino and Daw, 2015; Le Heron et al., 2020). Recent research has demonstrated that adolescents do not exhibit this bias to the same extent as adults (Lloyd et al., 2021b), suggesting heightened exploration in adolescence is associated with more optimal navigation of foraging environments. In addition to our other hypotheses, which are described below, our first hypothesis (H1) is that we should replicate previous research that shows that in the absence of social influence (i.e., an asocial foraging task), adolescents will forage more optimally than adults because they explore more.

Upon encountering a novel environment, the naïve forager does not know the availability of rewards in the environment (Kolling & Akam, 2017). As such, the forager must learn key properties of the environment, such as the average reward rate for the environment (Gabay & Apps, 2021). One route to learn these properties is to trial explore/exploit actions and update one's expectations based on the reward outcomes of those decisions (Kolling & Akam, 2017). However, social information can also be utilised to determine the optimal leaving policy (Dunlap et al., 2016; Grüter et al., 2011). For example, foraging bees (*Bombus terrestris*) utilise social information from more experienced bees to determine how much time to remain with individual patches before exploring novel ones (Baude et al., 2008). These findings suggest that social information from conspecifics can influence non-human animals' estimate of the availability of rewards in the environment (i.e., the average reward rate in the environment), which is a key quantity in the MVT equation (Charnov, 1976). Yet, to our knowledge, no research to date has tested whether human foraging is similarly susceptible to social influence. As such, we also examined how social influence affects adolescents' and adults' foraging decisions.

Much of the literature examining age-related trends in susceptibility to social influence has utilised risk taking paradigms (see Defoe et al., 2015). In these tasks, feedback does not have utility for future decision-making, as the probabilities involved in choices are made explicit (e.g., sampling the risk option in the Wheel of Fortune Task; Smith et al., 2014). These studies have demonstrated that adolescents emulate the level of risk taking (i.e., how frequently they select the option with a higher probability of loss compared to the option with little or no probability of loss) they observe in other individuals – an effect known as ‘risk contagion’. Though, there is contrasting evidence about whether adolescents are more susceptible to social influence that endorses or discourages risk taking (Braams et al., 2019; Ciranka & van den Bos, 2019; Reiter et al., 2019). Yet, many adolescent behaviours characterised in the social influence literature as ‘risk taking’ (e.g., trying an illegal drug for the first time; Smith et al., 2014), may instead be attributed to adolescents’ disposition for novelty-seeking, which motivates them to explore new experiences (Do et al., 2020). Unlike risk taking behaviour, exploration can be utilised to gather information about one’s surroundings and can guide future decisions, as the reward statistics of the environment are initially uncertain (such as while foraging; Garrett & Daw, 2020). Indeed, it has been proposed that feedback from exploration is important to allow adolescents to gain the experiential knowledge necessary for adulthood (Romer et al., 2017). However, relative to what is understood about risk-taking behaviour, there is limited understanding about the extent to which adolescents’ exploration behaviour is susceptible to social influence.

The limited studies that have examined whether adolescents’ exploration behaviour is susceptible to social influence have demonstrated that peers can increase adolescents’ exploratory behaviours, though these studies have not manipulated whether peers endorsed exploration or exploitation. For example, Burritica and colleagues (2019) demonstrated that on the Iowa Gambling Task, adolescents initially followed the recommendation of a peer, but in

later trials explored other options, which allowed them to identify other advantageous stimuli. In contrast, adults did not deviate from the advice of their peer and failed to identify other advantageous options available to them. These findings suggest that adolescents utilise social information when they have less knowledge about which option yields the highest rewards. Supporting this account, recent simulation-based work has demonstrated that social influence has a greater effect on adolescents when they perceive the outcomes of their actions as more uncertain, and peer effects diminish once the properties of the environment are learned (Ciranka & van den Bos, 2021). While previous research suggests that social influence can increase adolescents' exploration behaviour (Burrítica et al., 2019; Silva et al., 2015), it remains to be tested to what extent adolescents will follow advice that reduces their predisposition for novelty-seeking (Romer et al., 2017).

In addition to our first pre-registered hypothesis that we would replicate findings that adolescents explore more than adults in an asocial version of the task (H1), the present study examined a novel hypothesis about whether adolescents and adults demonstrate susceptibility to social influence while foraging. Specifically, our second set of hypotheses (H2a-c) addressed whether participants would adjust their rates of exploration based on the advice of a virtual, similar-aged peer and, mirroring the risk-taking literature (e.g., Braams et al., 2019; Reiter et al., 2019), whether adolescents and adults would be differentially susceptible to advice that promoted exploitation or exploration. Relative to adults, adolescents have a heightened predisposition for exploration (Romer et al., 2017). As such we predicted that when participants received advice that endorsed suboptimally high rates of exploration, adolescents would be more susceptible to this advice than adults, and therefore adolescents would exhibit less optimal foraging behaviour than adults by over-exploring (H2a). We also introduce a novel condition where participants receive advice to lower their rates of exploration. As adults typically overexploit while foraging (Constantino & Daw, 2015), we predicted that when

participants received advice that endorsed lower rates of exploration, adults would be more susceptible to this advice than adolescents, and therefore adults would exhibit less optimal foraging behaviour than adolescents by overexploiting (H2b). In previous work, we have found that adolescents are more optimal foragers relative to adults (Lloyd et al., 2021b). As such we predicted that when participants received advice that endorsed the optimal amount of exploration, adolescents would be more susceptible to this advice than adults, and therefore adolescents would exhibit more optimal foraging behaviour than adults (H2c). If these results hold, they will demonstrate that adolescents' propensity for exploration can either be enhanced or diminished by the views of their peers, which has implications for how adolescents learn about their surroundings during this developmental period.

### 6.3. Methods

#### 6.3.1. Participants

A total of 339 participants were recruited across the three age groups based on an *a priori* power analysis for a small-medium effect size (alpha = 0.05, power = 0.8, Cohen's *f* effect size = 0.15). Adolescents were recruited from local schools in Surrey, UK and adults were recruited from a participant recruitment site ([www.sona-systems.com](http://www.sona-systems.com)) hosted by a UK University. There were 116 younger adolescents aged between 13-15 ( $M_{\text{age}} = 14.25$ ,  $SD_{\text{age}} = 0.60$ , 48.74% female), 116 older adolescents aged between 16-17 ( $M_{\text{age}} = 16.33$ ,  $SD_{\text{age}} = 0.51$ , 75.63% female) and 107 adults aged 21 and above ( $M_{\text{age}} = 30.21$ ,  $SD_{\text{age}} = 10.64$ , 61.4% female). For participants under 16, parental consent was obtained. All participants provided consent prior to completing the study. Ethical approval was received from Royal Holloway, University of London.

*Table 6.1. Descriptive statistics for the adolescent and adult samples. Adults were asked to report their own highest level of education whereas adolescents were asked to report their parents' highest level of education.*

	Young Adolescent	Old Adolescent	Adult
<i>Participant</i>			
<i>race/ethnicity</i>			
White (UK)	86.21%	82.76%	57.01%
White (other)	0%	8.62%	3.74%
Mixed race	7.76%	6.90%	8.41%
Indian	0.86%	1.72%	7.48%
Pakistani	0.86%	0%	1.87%
Bangladeshi	0.86%	0%	3.74%
Chinese	2.59%	0%	5.61%
Asian (other)	1.72%	3.45%	2.80%
African	2.59%	0.86%	5.61%
Arab	1.72%	0.86%	0.93%
Other	2.59%	1.72%	1.87%
Education	2.32	2.45	3.28

### 6.3.2. Materials

*Demographic Questionnaire:* Participants were asked to report their age, gender (male or female), ethnicity and socioeconomic status. Gender was restricted to a binary choice as this determined the fictional peer they were matched with for the main task (see Social Patch Foraging Task below), though participants were free to select the gender they identified with, rather than the sex they were assigned at birth. For adolescent participants, socioeconomic status was measured by proxy through a self-report measure of their parents' highest level of education, which has been used in previous research (Steinberg et al., 2018). Adults were asked

to report their own highest level of education. Ethnicity options were selected based on the options listed by the Office of National Statistics (Office for National Statistics, 2009). However, participants who did not fall into one of these categories could enter their ethnicity in a text-entry box.

*Social Patch Foraging Task:* The present study utilised an adapted version of an apple picking foraging paradigm, where the aim was to collect as many apples as possible within the allocated time (Constantino & Daw, 2015). Participants could decide whether to stay with their current patch to exploit rewards that depleted with each subsequent exploit decision or to explore a novel patch with an unknown distribution of rewards. The task consisted of two foraging environments, one that was rich in quality and one that was poor in quality. Environment quality was manipulated by the amount of time it took to travel between patches. In the rich quality environment, the travel time was six seconds between patches. In the poor-quality environment, the travel time was 12 seconds between patches. The environment with the short travel time allowed participants to collect more reward in the time limit and was therefore richer in quality relative to the environment with the longer travel time. The manipulation of travel time was utilised to check whether participants adjusted their leaving thresholds according to the qualities of the environment, consistent with optimal foraging theory (Charnov, 1976). Each block ran for 7-minutes, and the two environments were presented in counterbalanced order.

There were four social influence conditions in the present task, to which participants were randomly assigned. The first condition was one with no social influence, which we term the asocial condition. This asocial condition was a traditional patch foraging task, in which participants attempted to estimate the optimal leaving threshold based only on the reward-feedback from their decisions. In the three other conditions, in which social influence was introduced, participants were told that a peer that was a similar age and gender to them had

previously completed the study and had provided their suggestions of the best points to leave patches. However, this advice was in fact programmed by the researchers. When participants chose to leave each patch to explore a novel one, participants were presented with an image of the peer and text stating whether they recommended staying for longer, staying for less time, or staying for the same amount of time with the next patch (see Figure 5.1).

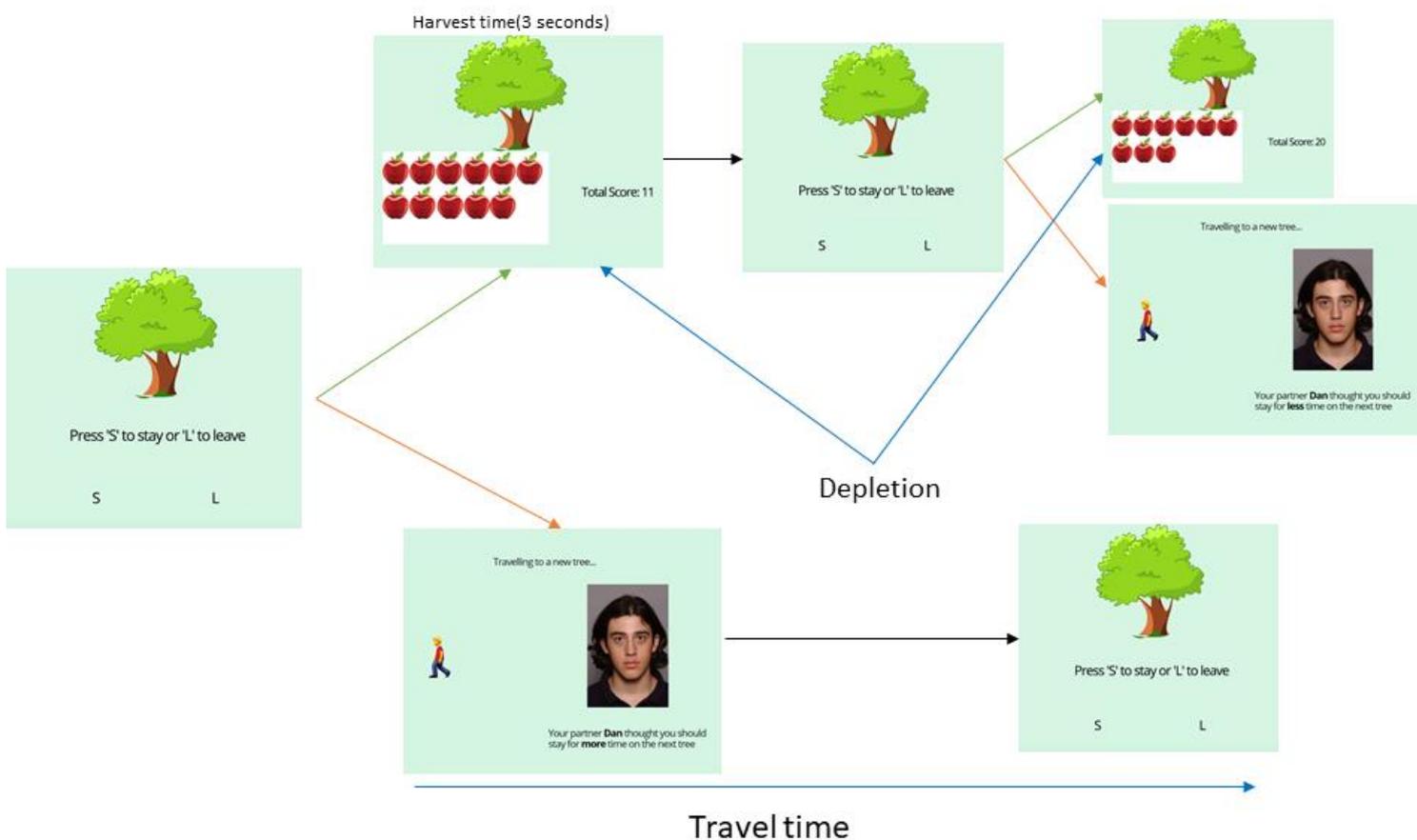


Figure 6.1: Schematic of the Social Patch Foraging task. When participants chose to stay (top panel), they were presented with the number of apples they accrued along with their total score. When participants chose to leave (bottom panel), they were presented with a travel screen which included an image of the peer and text that indicated whether they recommended staying for more, less, or the same amount of time on the next tree (patch).

The second condition was the good advice condition, in which the fictional peer’s feedback directed the participant to the optimal leaving threshold (determined by simulation,

see below). For example, if participants stayed with a patch longer than the optimal leaving threshold, the participant would receive advice to ‘stay for less time’ on subsequent trees. The third and fourth conditions were those in which the peer provided suboptimal advice, either directing participants to stay for too long (over-exploit) or too little time (over-explore). Suboptimal advice was coded as  $\pm 2$  units (i.e., apples) from the optimal leaving threshold. The peer images presented to participants were drawn from the NIMH-ChEFS database (Coffman et al., 2015) for adolescent faces and the FACES database (Ebner, Riediger, Lindenberger, 2010) for adult faces. Only neutral faces were utilised to avoid emotional valence confounding participants’ behaviour, as adolescents have heightened sensitivity to emotive stimuli compared to adults (Silvers et al., 2012). Similar operationalisations of social influence have been utilised in previous studies (e.g., Braams et al., 2019).

Our pre-registered dependent variable, which is the primary measure we consider here, was how far participants’ leaving threshold deviated from the optimal leaving threshold in each of the two qualities of environment. To calculate how far participants deviated from the optimal leaving threshold, we first calculated participants’ own leaving threshold. Consistent with previous research (Constantino & Daw, 2015), we assume participants select a value of apples as their leaving threshold and only explore when the number of apples expected from the next exploit decision falls below this threshold. Participants’ leaving threshold was calculated from the average of the last two harvests from the patch, which provided an upper and lower bound of participants’ leaving threshold (Constantino & Daw, 2015).

We next estimated the optimal leaving threshold, which was determined by simulating the foraging environments and running a grid search across leaving thresholds (ranging from 1-10) in increments of 0.001. The leaving threshold that yielded the highest number of points was considered the optimal leaving threshold for that environment (Lloyd et al., 2021b). Consistent with MVT, these simulations demonstrated that participants should have a higher

leaving threshold (i.e., explore more) in the rich quality environment, where the average reward rate is higher. Further, the simulations demonstrated that in the poor-quality environment, where the average reward rate is lower, participants should stay until the patch is more extensively depleted (i.e., should explore less). To calculate how much participants deviated from the optimal leaving threshold, we subtracted the optimal leaving threshold from participants' average leaving threshold. Values closer to zero on the deviation from the optimal leaving threshold variable indicate the participant exhibited more optimal behaviour. Positive values on the deviation from the optimal leaving threshold variable indicate that participants exploited for longer than was optimal and negative values indicate the participants explored more than was optimal.

In addition to pre-registering hypotheses about deviation from leaving threshold, we also preregistered hypotheses about the parameters from our computational modelling, described below. Moreover, we calculated a further two dependent variables, which were not preregistered: participants' patch residency time and the degree to which participants stayed in patches for the length of time advised by their peer (i.e., how closely participants converged on the value of apples that we coded the fictional peer to recommend in each environment). Descriptions and analyses involving these variables can be found in the Supplementary Materials.

*Manipulation Check:* We included a series of questions that probed whether participants believed they had received advice from a real peer, which was comprised of Likert scale questions, such as "How sure were you during the game that the other player was real". The degree to which participants believed they were receiving advice from a fictional peer was calculated as the average of their ratings on these questions. See Supplementary Materials for analyses involving this measure.

### 6.3.3. Procedure

Participants were invited to an IT suite either in their school classroom or at Birkbeck, University of London, where they were informed about the task instructions. They were informed that a base rate of £3 would be awarded for taking part and that there was the potential to earn an additional £2 based on their performance on the task. After being briefed and provided with the opportunity to ask questions, participants completed a consent form, demographic questionnaire and practice task which introduced them to the foraging paradigm. This practice task did not include advice from a fictional peer and had a travel time of nine seconds between patches, which differed from either environment in the main task.

After completing the practice, participants were randomly allocated to one of the experimental conditions where social influence (or the lack thereof) was manipulated. Each participant completed only one of the four social influence conditions. A total of 85 (64.29% female) participants completed the asocial condition, 86 (61.63% female) participants completed the good advice condition, 84 (61.18% female) completed the overexploit condition and 84 (62.36% female) completed the over-explore condition. Each social condition was matched in terms of age, as demonstrated by the absence of a significant difference in age between social influence conditions for younger adolescents ( $F_{(113)} = 0.08, p = .973$ ), older adolescents ( $F_{(113)} = 0.58, p = .630$ ) and adults ( $F_{(113)} = 0.12, p = .947$ ). For participants in each of the conditions where they received advice from a fictional peer, participants were told that an individual of a similar age and gender completed the task previously and had provided advice to the next participant about the best point to leave patches in that environment, which would be presented each time they travelled to a new patch. Participants were informed that this advice may not be correct, but it was up to them to decide whether to follow their peer's advice. In the condition with no social feedback (the asocial condition), participants were told they should try to collect as many apples as possible within the time limit and had no further

instruction. The reward statistics of rich and poor environments were the same for all participants with the only difference being the type of social influence they received during the task. Every participant completed the rich and poor-quality environments of the patch foraging task in counterbalanced order. Upon completion of the task, participants completed a questionnaire comprised of the Barratt-11, BIS/BAS and for adolescents only, the Pubertal Scale (see Supplementary Measures for descriptions of the Barratt-11, BIS/BAS and Pubertal Scale). Participants were debriefed as to the aims of the experiment, including the peer manipulation, at the end of the study.

#### **6.3.4. Computational Modelling**

In the present study, we compare two computational models that explain how participants determine when to explore new patches, which are both derived from MVT (Charnov, 1976). In the first, asocial model, we assume that participants do not integrate social information into their decisions and instead exclusively use the reward feedback from explore or exploit decisions to determine when to leave patches. We compare this asocial model to a novel, social learning model where the advice received from the fictional peer is integrated into the estimated reward value of exploration and exploitation, which determines when participants leave patches. We only compare the asocial and social models in conditions where potentially socially influencing stimuli were present: when participants receive advice from the fictional peer (i.e., the good advice condition, the overexploit condition and the over-explore condition).

According to MVT, when deciding whether to leave a patch, the forager must compare two values. The first is the reward value expected from exploiting the current patch, which is given by multiplying the most recently observed reward from exploiting ( $s$ ) on each trial ( $i$ ) by the rate at which apples deplete from patches ( $\kappa_i$ ) within the environment. This depletion rate was drawn on each trial from a Gaussian distribution with a mean of 0.88 (SD = 0.07). As in previous research (Constantino & Daw, 2015), we assume that participants have a running

estimate of the depletion rate, which we calculate through averaging across all true depletion rates experienced by the participant in the previous trials for each environment. Under MVT, agents compare this expected reward rate to a second reward value, which is the participants' estimate of the average reward rate for the environment on that trial ( $p_i$ ) multiplied by the time associated with exploiting the current patch ( $h$ ), which in the current task was fixed at three seconds. When the rewards expected from exploiting the current patch fall below the average reward rate in the environment, the participant should leave to explore a novel patch (i.e., when  $\kappa_i s_i < p_i h$  is true; Constantino & Daw, 2015).

In both models, we assume that the forager does not know the structure of the foraging environments *a priori* and must therefore learn the average reward rate. For the asocial model, we use an equation developed by Constantino and Daw (2015), which explains how the participant estimates the average reward rate of the foraging environment ( $p_i$ ), which is expressed as:

$$p_i = (1 - \alpha)^{T_i} \frac{s_i}{T_i} + (1 - (1 - \alpha)^{T_i}) p_{i-1} \quad (1)$$

In Equation (1), the free parameter alpha ( $\alpha$ ) represents the degree to which the participant weighs recent feedback. As we operationalised alpha as the complement of the traditional learning rate parameter (i.e.,  $1 - \alpha$ ), lower values on this free parameter indicate that the participant weighs recent feedback highly in their estimates of the average reward rate, whilst higher values indicate the participant weighs recent feedback lower, instead integrating a wider range of trials into their estimate. Included in Equation 1 is the time cost associated with the participant's decision ( $T_i$ ), which was coded as the true number of seconds they experienced with each stay or leave choice ( $a_i$ ). The time cost associated with exploitation was three seconds and the time cost associated with exploration was manipulated between

environments. The value of  $p_i$  is subsequently entered into a SoftMax equation, which determines the probability that the participant will stay or leave on each trial:

$$P(a_i = harvest) = 1/(1 + \exp(-(c + \beta(\kappa_{i-1}s_{i-1} - p_{i-1}h)))) \quad (2)$$

In Equation (2),  $c$  is a free parameter measuring participants' overexploitation bias and  $\beta$  is a free parameter measuring stochasticity in participants' decisions. Higher values on  $\beta$  indicate participants consistently select the option with the highest expected reward value, whereas lower  $\beta$  values indicate choices are more stochastic and occasionally deviate from the option the participant perceives to have the highest expected value. Finally,  $\kappa_i s_i - p_i h$  represents the difference between the number of rewards expected from the next exploit decision and the participants' estimate of the average reward rate.

To model the role of social influence on foraging decisions, we developed a variation of the learning model described in Equations (1) and (2). This social variation of the learning model included an additional free parameter which measured the degree to which participants integrated social influence into their foraging choices, based on similar work (Burrutca et al., 2019; Ciranka & van den Bos, 2019). In this model, we assumed that peers influence participants' estimate of the average reward rate for the environment (see Amano et al., 2006; Gabay & Apps, 2021), which signals when participants should leave their current patch. We formalise this MVT Social Learning model as:

$$p_i = (1 - \alpha)^{T_i} \frac{s_i}{T_i} + (1 - (1 - \alpha)^{T_i}) q_{i-1} \quad (3)$$

Equation (3) closely resembles Equation (1), with the only difference being the quantity  $q_i$ , which is the estimate of the average reward rate subject to the degree to which participants are susceptible to social influence, which was expressed as:

$$q_i = p_i + (v_i - p_i) \cdot \psi \mid \{\psi \in \mathbb{R} : 0 < \psi < 1\} \quad (4)$$

In Equation (4), the quantity  $v_i$  represents the estimated value of the advice that the fictional peer provides, which is expressed in units of reward per second to ensure that it is scaled to the average reward rate. The value of  $v_i$  is updated on each leave trial when participants receive social feedback, increasing by one apple (which, when scaled to the average reward rate, was equal to 0.33) when the peer advised exploring more and decreasing by one apple (or 0.33 when scaled to the average reward rate) when the peer advised exploring less. In cases where the participant uses the leaving threshold consistent with the advice of the fictional peer, the participant would receive advice that would state they should “stay for the same amount of time” in future patches. However, as  $v_i$  was an estimate that varied across the course of the environment, this value did not necessarily equal the leaving threshold that was advised by the peer (i.e., the number of apples that the fictional peer would direct the participant towards on explore trials, see Figure 6.1), which was fixed in each environment. As such, in cases when the participants’ leaving threshold was consistent with the advice of the fictional peer *and*  $v_i h$  was lower than this leaving threshold, the value of  $v_i$  would increase by one apple (i.e., 0.33). In cases when the participants’ leaving threshold was consistent with the advice of the fictional peer *and*  $v_i h$  was higher than this leaving threshold, the value of  $v_i$  would decrease by one apple (i.e., 0.33). The free parameter  $\psi$  measures the extent to which participants update their estimate of the average reward rate based on advice from their peer, with higher values indicating participants’ estimate of the average reward rate was more susceptible to social influence. The updated value of  $q_i$  is subsequently entered into the SoftMax equation to calculate the probability participants will stay or leave on each trial (Equation 5). Parameter and model recovery for both the asocial and social patch foraging models was successful, indicating the models had acceptable parameter identifiability (see Supplementary Materials).

$$P(a_i = harvest) = 1/(1 + \exp(-(c + \beta(\kappa_{i-1}s_{i-1} - q_{i-1}h)))) \quad (5)$$

To examine whether participants integrated social information into their foraging behaviour, we compared the ability of the asocial learning and social learning models to explain participants' explore/exploit behaviour in the three social influence conditions in which peers provided advice. When estimating the values of the free parameters ( $\alpha$ ,  $c$ ,  $\beta$ , and for the social model only,  $\psi$ ), we fitted the models to the explore/exploit choices from participants' data in both environments rather than running the models on each environment separately. This decision was made to increase the number of trials that were used to estimate parameters, thereby increasing the reliability of these parameter estimates (Daw, 2011). To assess whether the asocial or social model provided a better fit to participants' data, we utilised the Akaike Information Criterion (AIC; Akaike, 1998) and Bayesian Information Criterion (BIC; Schwarz, 1978) as indices of model fit. Lower values on these criteria indicate the model is a better fit to the data (Daw, 2011). We then tested for the effects of age and social influence condition on the winning model parameters.

### **6.3.5. Design**

The design of the study was a  $2 \times 3 \times 4$  mixed factorial design. The dependent variables in our analyses were how much participants' leaving threshold deviated from the optimal leaving strategy and the parameter estimates from our computational models ( $\alpha$ ,  $c$ ,  $\beta$ , and for the social model only,  $\psi$ ). The within-participant factor was the quality of the foraging environment participants completed, which had two levels of quality (rich and poor). The second factor, age, had three levels, which were: younger adolescent, older adolescent, and adult. These groups were selected to examine how social influence changes over development. Noting the positive association between exploration and risk taking (Addicott et al., 2017), previous research has demonstrated that younger adolescents' risk seeking behaviours are more susceptible to social influence compared to older adolescents and adults (e.g., Knoll et al., 2015; 2017). Moreover, there is conflicting research regarding whether older adolescents' risk

seeking behaviours are more susceptible to social influence that endorses a safe option than that which endorses a risky one (Braams et al., 2019; Reiter et al., 2019). The third factor was the social influence condition participants were assigned to during the task, which was a between subjects factor. This factor had four levels: no social feedback, good advice that would direct participants to leave patches at the optimal point, advice to over-exploit patches and advice to over-explore patches. Our pre-registered analysis plan can be found at: <https://osf.io/3t2wa/>.

#### 6.4. Results

Our first hypothesis (H1) was that adolescent participants would exhibit more optimal foraging relative to adults in the asocial condition of the foraging paradigm, replicating previous work (Lloyd et al., 2021b). To test this, we used the data from participants that completed the asocial condition to conduct a 3×2 mixed ANOVA with the degree to which participants deviated from the optimal leaving threshold entered as the dependent variable. The model had two independent variables: age group (whether participants were younger adolescents, older adolescents, or adults) and environment quality (rich or poor). Inconsistent with our first hypothesis, there was not a main effect of age group ( $F_{(2, 81)} = 0.33, p = 0.720$ ), demonstrating there is no evidence that younger adolescents, older adolescents or adults differed in how optimally they performed in the asocial condition (see Figure 6.2). Other findings from this ANOVA include that participants were more optimal in the poor quality environment ( $F_{(1,81)} = 6.59, p = .012, \eta^2 = 0.01$ ), which replicated our previous findings (Lloyd et al., 2021b). There was no interaction between age group and environment ( $F_{(1,81)} = 0.93, p = .400, \eta^2 = 0.002$ ). These effects remained after controlling for gender to account for the

overrepresentation of female adolescents in the older adolescent and adult groups (see Supplementary Materials).

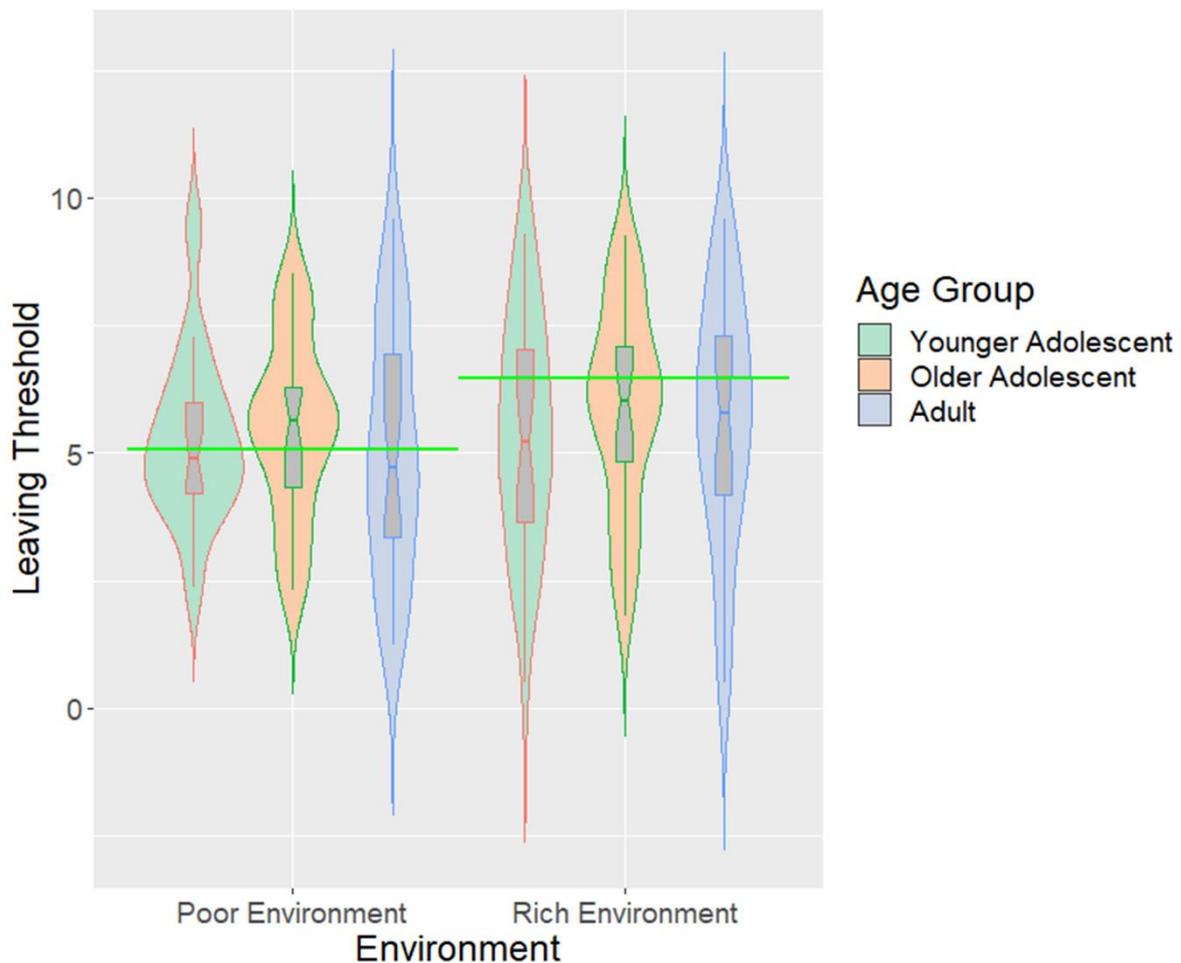


Figure 6.2: Violin plots demonstrating participants' leaving thresholds in the asocial condition of the task. Middle bars indicate the mean leaving threshold for that age group with upper and lower bars reflecting the upper and lower quartiles, respectively. Horizontal green bars indicate the optimal leaving threshold for that environment.

Our second set of hypotheses (H2a-c) predicted that, adolescents would perform less optimally than adults in the over-explore condition (H2a), whereas adults would perform less optimally than adolescents in the overexploit condition (H2b). Further, we predicted that compared to other adults, adolescents would perform more optimally when they received advice that promoted optimal foraging (H2c). To test these predictions, we examined how far participants deviated from the optimal leaving threshold in each foraging environment using a

3×2×4 ANOVA. The factors in this analysis were: age group (younger adolescent, older adolescent, or adult), environment quality (rich or poor) and social influence condition (asocial, good advice, overexploit advice, or over-explore advice). Inconsistent with hypotheses H2a, H2b and H2c, the interaction between social influence condition and age group was not significant ( $F_{(6,326)} = 0.37, p = 0.896$ ), which indicates that the degree to which participants' leaving threshold optimality was susceptible to social influence that endorsed exploration or exploitation did not differ by age group. The absence of an interaction effect was also found when the deviation from the optimal leaving threshold was treated as an absolute, rather than relative value (see Supplementary Materials), further suggesting adolescents and adults were not differentially susceptible to social influence that promoted exploration or exploitation.

Additional findings from this 3×2×4 ANOVA were that there was a main effect of social influence condition on deviation from the optimal leaving threshold ( $F_{(3,326)} = 8.96, p < .001, \eta^2 = .06$ ; see Figure 6.3), indicating that our social influence manipulation was successful. Additional findings from this 3×2×4 ANOVA were that there was a main effect of social influence condition on deviation from the optimal leaving threshold ( $F_{(3,326)} = 8.96, p < .001, \eta^2 = .06$ ; see Figure 6.3), indicating that our social influence manipulation was successful. Using a one sample t-test, we found that participants' deviation from the optimal leaving threshold in the asocial condition did not differ from zero ( $t(83) = 1.30, p = .199$ ) suggesting that across age groups, participants' exploration was close to optimal foraging behaviour in the absence of social stimuli. In addition, participants' deviation from the optimal leaving threshold in the good advice condition did not differ from zero ( $t(83) = 1.80, p = .076$ ). Further, in the over-explore condition, participants' leaving threshold was higher and they explored more than was optimal (statistically differing from zero  $t(84) = 3.39, p = .001$ ), whereas in the overexploit condition participants' leaving thresholds were lower, exploiting patches more than was optimal (which differed statistically from zero  $t(84) = 4.21, p < .001$ ). The effect of the social

influence manipulation was supported by post hoc tests using the Bonferroni correction adjusted for six comparisons (see Figure 6.3). Moreover, there was not a main effect of age group ( $F_{(2,326)} = 1.17, p = 0.312$ ), suggesting there was no evidence that the degree to which participants deviated from the optimal leaving threshold varied by age group. There was, however, a main effect of environment quality ( $F_{(1,326)} = 158.02, p < .001, \eta^2 = 0.05$ ), with participants performing more optimally in the rich quality environment relative to the poor quality environment. Further, we did not find significant interactions between: environment and social influence condition ( $F_{(3,326)} = 2.45, p = 0.064$ ) environment and age group ( $F_{(2,326)} = 0.12, p = 0.884$ ), or environment, social influence condition and age group ( $F_{(6,326)} = 1.26, p = 0.276$ ). Analysis of the degree to which participants stayed in patches for the length of time advised by their peer were consistent with the analysis of participants' deviation from the optimal leaving thresholds, as we did not find a main effect of age on how closely participants converged on the leaving threshold advised by the fictional peer nor an interaction between age group and social influence condition (see Supplementary Materials).

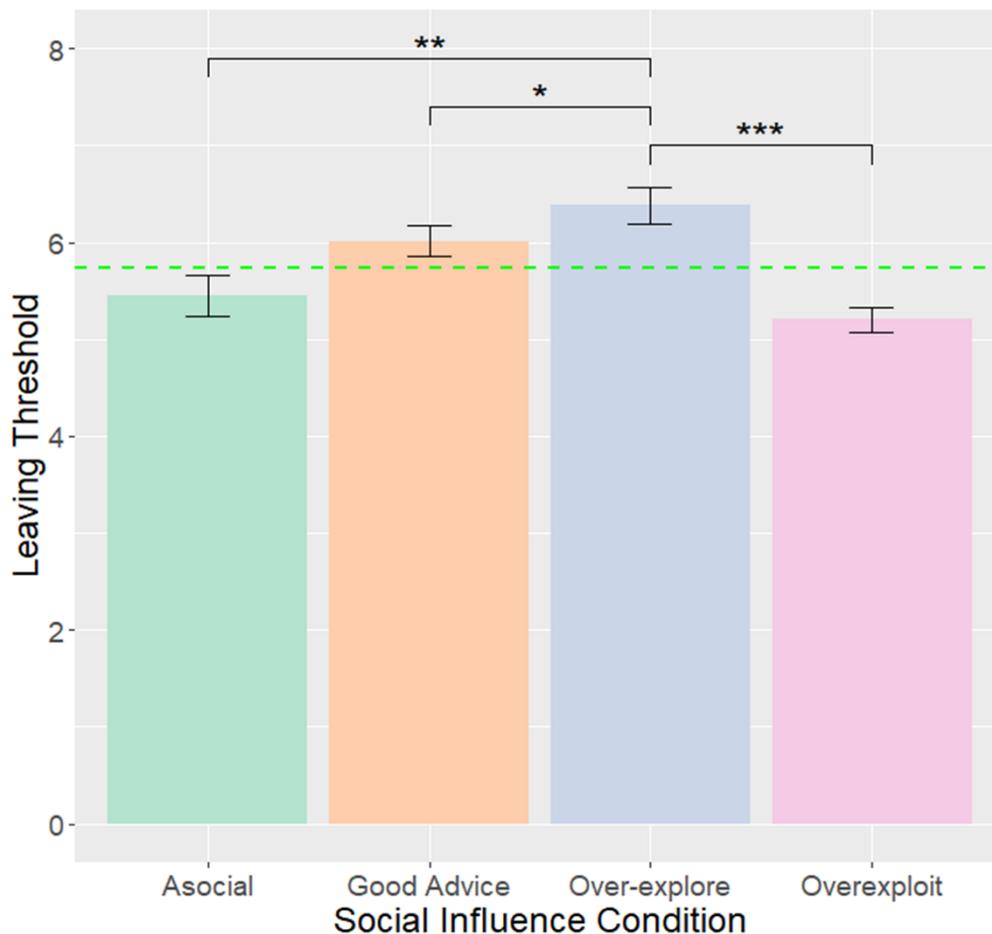


Figure 6.3: Plot demonstrating participants' leaving thresholds in each of the social influence conditions for both foraging environments. Error bars indicate standard deviations for each condition. The green dashed line indicates the optimal leaving threshold averaged across both foraging environments. \*  $p < .050$ , \*\*  $p < .010$ , \*\*\*  $p < .001$

We conducted exploratory analyses on data from each of the social influence conditions to examine whether age-related differences on susceptibility to social influence were restricted to peer advice that either endorsed overexploitation, over-exploration, or optimal levels of exploration. Hypothesis 2a predicted that adolescents would exhibit less optimal behaviour than adults when participants received advice to engage in suboptimal levels of exploration. Using data from the condition where participants received advice to over-explore, we did not find evidence that adolescents were more susceptible to this type of social influence relative to adults, as we did not find a main effect of age group on the degree to which participants

deviated from the optimal leaving threshold ( $F_{(2,81)} = 1.44, p = .242$ ). Hypothesis 2b predicted that adults would exhibit less optimal behaviour than adolescents when participants received advice to overexploit patches. Using data from the condition where participants received advice to overexploit, we did not find evidence that adults were more susceptible to this type of social influence relative to adolescents, as we did not find a main effect of age group on the degree to which participants deviated from the optimal leaving threshold ( $F_{(2,81)} = 0.05, p = .948$ ). Hypothesis 2c predicted that adults would exhibit less optimal behaviour than adolescents when participants received advice that endorsed optimal rates of exploration. Using data from the condition where participants received advice for optimal levels of exploration, we did not find evidence that adolescents were more susceptible to this type of social influence, as we did not find a main effect of age group on the degree to which participants deviated from the optimal leaving threshold ( $F_{(2,81)} = 0.06, p = .939$ ). Our finding that there was no main effect of age group in either the overexploit and over-explore social influence conditions is consistent with the lack of a significant interaction between age group and social influence condition reported in our  $2 \times 3 \times 4$  ANOVA.

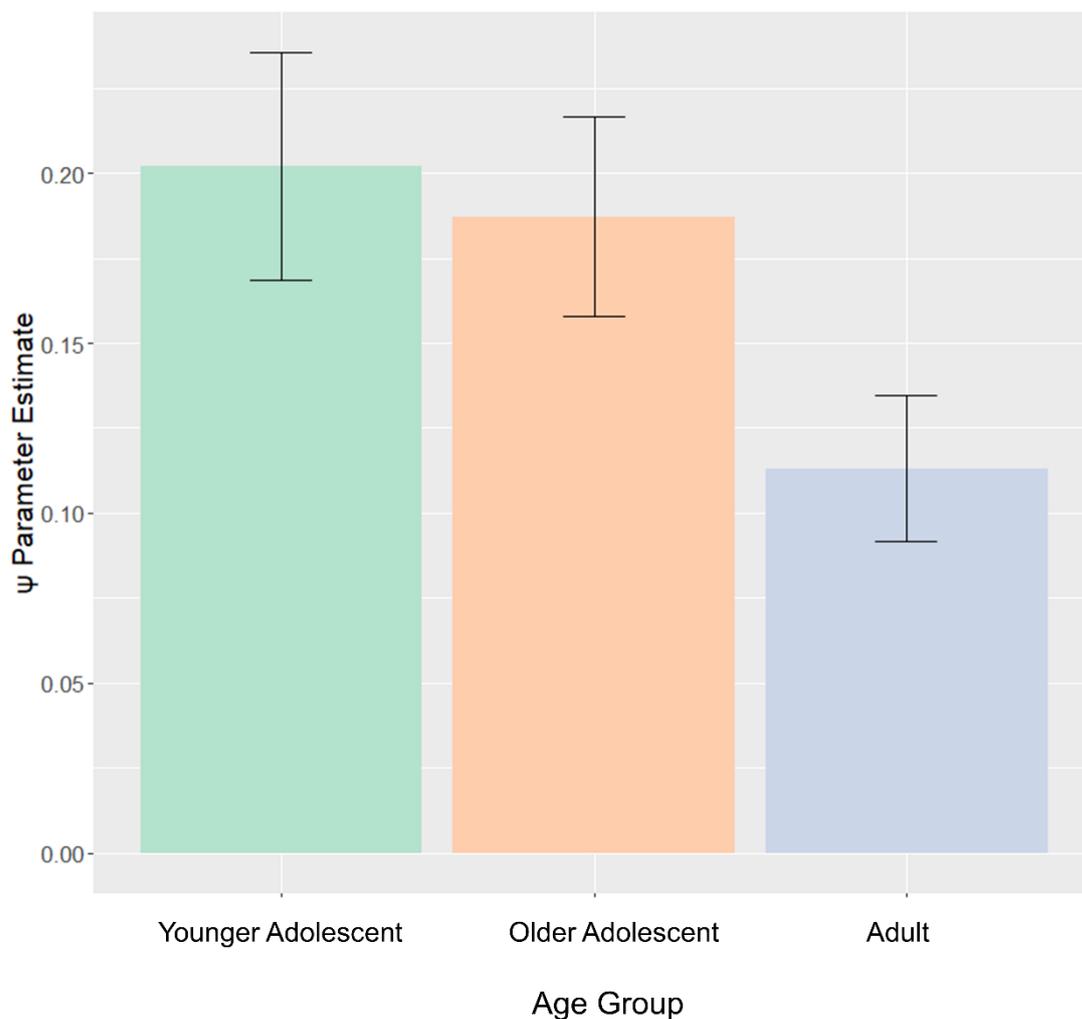
We developed a novel computational model that attempted to explain how participants integrated social information into their foraging choices. In pre-registered analyses, we first compared the indices of model fit from our novel social learning model with an asocial learning model that has previously been utilised to explain foraging behaviour (Constantino & Daw, 2015). Comparison of the AICs indicated that the social model was a better fit to the data compared to the asocial model ( $AIC_{\text{social}} = -81.59, AIC_{\text{asocial}} = -78.97$ ). However, BIC values indicated that the asocial model was a better fit to the data compared to the social MVT model ( $BIC_{\text{social}} = -68.44, BIC_{\text{asocial}} = -69.06$ ). We note that the difference between the BIC for the social and asocial models was less than one, which provides weak evidence that the asocial model was better able to explain participants' explore/exploit decisions. These patterns in the

model fit metrics were observed separately in both adolescents and adults (see Supplementary Material). Relative to the AIC, the BIC has a harsher penalty for more complex models with a higher number of free parameters, which may account for the inconsistency between the two indices of model fit (Daw, 2011; Lockwood & Klein-Flugge, 2021). Moreover, only the social model was able to replicate the pattern of participants' leaving thresholds, and how far these deviated from the optimal leaving thresholds, seen in Figure 6.3 (see Supplementary Figure 6.2). Together, we suggest that these findings indicate the social learning model provided a better explanation of participants' exploration behaviour in the conditions of our paradigm where participants received advice from a fictional peer.

Examining mechanistic features underlying participants' behaviour in the social influence conditions, we conducted a 3×3 ANOVA with the free parameter  $\psi$ , measuring susceptibility to social influence, as the dependent variable. We compared estimates of this free parameter between age groups and between the social influence conditions. The first factor included as its levels the three social influence conditions, excluding the asocial condition with no social feedback, and the second factor was participants' age group. Results of the ANOVA demonstrated a main effect of social influence condition ( $F_{(2,226)} = 4.87, p = .008, \eta^2 = .04$ ). Specifically, participants had higher values on the social influence parameter in the over-explore condition ( $M = 0.24, SD = 0.29$ ) relative to the overexploit condition ( $M = 0.14, SD = 0.25; p_{bonf} = .043$ ) and the good advice condition ( $M=0.12, SD=0.22, p_{bonf} = .015$ ; see Supplementary Figure 6.3) which did not differ from each other. However, the main effect of age group was not significant ( $F_{(2,226)} = 2.99, p = .052, \eta^2 = .02$ ), nor was the interaction between social influence condition and age group ( $F_{(4,226)} = 2.07, p = .085, \eta^2 = .03$ ).

To account for the overrepresentation of female participants in the older adolescent and adult groups, we conducted an exploratory ANCOVA on the parameter estimate  $\psi$ , controlling for gender. In contrast to the analyses without controlling for gender, we found a main effect

of age group ( $F_{(2,225)} = 3.18$ ,  $p = .044$ ,  $\eta^2 = .03$ ; see Figure 6.4). Post-hoc tests using the Bonferroni correction for three pairs of analyses demonstrated that younger adolescents had significantly higher values on the  $\psi$  parameter compared to adults ( $p_{bonf} = .045$ ), though there was no significant difference between younger adolescents and older adolescents ( $p_{bonf} = 1.00$ ), nor older adolescents and adults ( $p_{bonf} = 0.249$ ). Replicating the analysis where gender was not controlled for, we also found a main effect of social influence condition ( $F_{(2,225)} = 4.85$ ,  $p = .009$ ,  $\eta^2 = .04$ ) and no interaction between social influence condition and age group ( $F_{(4,225)} = 2.11$ ,  $p = .080$ ,  $\eta^2 = .03$ ). We did not find any differences between younger adolescents, older adolescents, or adults on the other free parameters before or after controlling for gender (see Supplementary Materials).



*Figure 6.4: Estimates of the social influence parameter ( $\psi$ ) for each age group. Error bars indicate standard deviations in each age group. Larger values indicate that participants are more susceptible to social influence.*

In summary, our analyses provide partial support for our hypotheses (H2a-c). Specifically, we found that the degree to which participants deviated from the optimal leaving threshold could be influenced by the advice they received from a fictional peer. However, these effects were not restricted to a particular age group, nor were adolescents and adults differentially susceptible to social influence that endorsed exploration or exploitation. As demonstrated through our computational modelling, younger adolescents' estimate of the average reward rate, and hence their exploration behaviour, was more susceptible to social influence compared to adults. However, this finding was inconsistent with our prediction that adolescents' susceptibility to social influence would differ depending on whether the peer endorsed exploration or exploitation.

## **6.5. Discussion**

The first hypothesis (H1) of the present study was that adolescents would be more optimal foragers than adults in the asocial condition of the task, replicating previous findings (Lloyd et al., 2021b). Contrary to our predictions, we did not find that adolescents exhibited more optimal exploration than adults. Our second hypothesis (H2) was that optimality of adolescents' leaving thresholds would be more susceptible to social influence compared to adults. We introduced a novel paradigm and associated computational model that incorporates social influence into human foraging. Using these new methods, we demonstrated that adolescents' and adults' exploration behaviour is susceptible to social influence. These findings provide insight into the social factors that affect humans' explore/exploit strategies and how these factors can promote both optimal and suboptimal levels of foraging. Analysis of our

social learning model suggested that younger adolescents' exploration behaviour was more susceptible to social influence compared to adults, as they placed greater weight on this advice when estimating the average reward rate of the environment. However, inconsistent with our hypotheses (H2a-c), we did not find evidence that adolescents' and adults' leaving thresholds were differentially susceptible to social influence that endorsed exploration or exploitation, as both age groups exhibited similar levels of susceptibility to social influence in these conditions. Rather, analysis of our computational model suggested that all participants were susceptible to social influence that endorsed suboptimal rates of exploration. Nevertheless, these findings can inform our understanding of how adolescents and adults utilise social information to explore novel scenarios and provide insight on how social feedback influences decision-making across the lifespan.

Our findings contribute to the foraging literature through the innovation of a model that incorporates social influence into patch foraging behaviour. We demonstrate that participants' exploration behaviour is susceptible to social influence, evidenced through the main effect of social influence on the degree to which participant deviated from the optimal leaving threshold. To explain these effects, we have proposed a formal model whereby social information is incorporated into the forager's estimate of the richness of the environment, which according to MVT determines how often the individual should explore new patches. Similar models have been proposed in ecology, whereby animals signal to other conspecifics about the richness of the environment (Amano et al., 2006). This social learning model may have utility for understanding how humans use social information to evaluate the trade-off between local (i.e., within patch) and global (i.e., average environmental) rewards (Charnov, 1976). Recently, there has been increased interest in foraging problems within social neuroscience (Gabay & Apps, 2021); our study can contribute to understanding the psychological processes involved in this complex behaviour and can be adapted to understand phenomena such as prosocial

decision-making (Gabay & Apps, 2021; Zacharopoulos et al., 2018). Furthermore, understanding how social cues influence the perception of global reward rates could be useful to applied foraging contexts, such as understanding how the perception of the availability of food items in supermarkets is influenced by social cues in times of crises (Bentall et al., 2021; Dickins & Schalz, 2020).

Our findings also demonstrate how adolescents' and adults' explore/exploit decision-making is susceptible to social influence. The susceptibility to social influence during adolescence has typically been linked with negative outcomes such as increased risk-taking behaviours (Chein et al., 2011). However, recent work has demonstrated that social influence can be a positive influence on adolescents' behaviour, encouraging risk-aversion and prosocial tendencies (Braams et al., 2019; Foulkes et al., 2018). Our findings further bolster the evidence showing that adolescents adjust their behaviour to conform to the views of their peers and that this behaviour is more prevalent in younger adolescents relative to adults (Knoll et al., 2017). In contrast to previous literature, which has focussed on how social influence affects adolescents' decisions to take risks, our findings provide novel evidence that adolescents are susceptible to feedback from peers when evaluating whether to exploit a known option or explore a novel one.

Explore/exploit decision-making is notably different from risk taking. In tasks measuring risk-taking, such as the Wheel of Fortune, there is no information that can be learned through sampling the risky option where the full information about the probabilities of a positive and negative outcome occurring is explicit and are known to the decision-maker. Feedback from sampling an option where probabilities are known (e.g., sampling the risk option in the Wheel of Fortune task) therefore does not have utility for future decision-making. In contrast, decision-making in the explore/exploit trade-off involves estimating the reward value of each option and feedback on each trial can inform future decisions (Garrett & Daw,

2020). We demonstrate that participants' propensity for exploration – and therefore their opportunities to learn the reward value of different actions - can be influenced by the views of their peers. We did not find evidence that adolescents are differentially susceptible to social influence that endorses exploration or exploitation. This finding is in contrast to studies examining risk taking decisions, where there is conflicting evidence about whether adolescents are more susceptible to social influence that endorses risk seeking or risk averse behaviour (Braams et al., 2019; Ciranka & van den Bos, 2019). As such, our findings contribute to understanding that adolescents' propensity for exploration in novel and uncertain environments can either be enhanced or diminished by the views of their peers.

When encountering novel environments where the best course of action is not immediately known, following the advice of another individual may be an adaptive decision-making strategy (Winet et al., 2020). Previous research has demonstrated that adolescents are more likely to follow peer advice when they are more uncertain about their own preferences (Reiter et al., 2021) or when there is greater ambiguity associated with the available choices (van Hoorn et al., 2017). In the present foraging paradigm, participants are naïve to the optimal strategy when encountering each environment and as such, social cues may provide a valuable information source (Dunlap et al., 2016; Grüter et al., 2011). Indeed, the utility of relying on social information has been observed in multiple species including human beings (Baude et al., 2008; King et al., 2010), and it has been suggested that social information was important in successful foraging in hunter-gatherer societies (King et al., 2010). Our findings are consistent with this account, as participants were closer to the optimal leaving threshold in the social influence condition where participants received good advice (i.e., that which recommended leaving at the optimal leaving threshold; see Figure 6.3). However, we note that participants' performance in the asocial condition was already close to the optimal leaving threshold (i.e., did not differ statistically from zero) and therefore there was little room for performance to be

improved in the good advice condition. Nevertheless, in addition to presenting evidence that peers can reinforce optimal rates of exploration while foraging, we present novel findings that susceptibility to social influence could lead to suboptimal outcomes, as participants followed the advice of a similar aged peer in conditions where advice promoted suboptimal levels of exploration.

These findings have several implications for understanding the risk and protective factors associated with susceptibility to social influence in adolescence. Our results further substantiate evidence that adolescents' decision-making is susceptible to social influence and that this can promote or discourage novelty-seeking behaviours (Foulkes et al., 2018; Knoll et al., 2015). The importance placed on the views of same-aged peers can be used to encourage positive or prosocial behaviours. For example, one recent paper has suggested adolescents' susceptibility to social information could be harnessed to promote adherence to novel health guidance during the COVID-19 pandemic (Andrews et al., 2020). Exploration is important during adolescence as it can facilitate the development of experiential knowledge, which can be utilised in later life (Romer et al., 2017). As such, social influence could be used to promote novel experiences that support adolescent development, such as engaging in extracurricular activities, which are associated with positive developmental outcomes (Farb & Matjasko, 2012). Conversely, heightened rates of exploitation have been associated with compulsive, habitual behaviours (Addicott et al., 2017), which highlights the potential role for peers in promoting behaviours associated with compulsion (e.g., gambling, or excessive substance use; Blom et al., 2011; el-Guebaly et al., 2011). Indeed, adults may also be susceptible to such peer influences, as we found adults' exploration behaviours were also susceptible to social influence. Together, our findings contribute to the debate around the protective and risk factors associated with susceptibility to social influence in adolescence.

In contrast to previous research (Lloyd et al., 2021b), we did not find evidence that adolescents exhibited more optimal exploration relative to adults in the asocial condition of our foraging task. One possible reason for these findings is that we did not manipulate features of the task related to reward contingencies as we did in our previous studies (e.g., initial richness or the rate at which rewards depleted from patches; Lloyd et al., 2021b). Previous research has found that adolescents exhibit heightened sensitivity to changes to probabilistic reward contingencies relative to adults (DePasque & Galvan, 2017; Hauser et al., 2015), which may explain why we previously found age-related differences in optimal exploration, as in previous studies we also manipulated the rate at which rewards depleted from patches (Lloyd et al., 2021b). In contrast, in the current study, we manipulated the quality of the foraging environments only through altering the length of time associated with travelling to new patches. Previous research has demonstrated that adolescents and adults exhibit similar evaluations of time cost when determining how much cognitive control to exert during decision-making (Devine et al., 2021). As such, our manipulation of the travel time, but not the depletion rate or initial richness of patches, may explain the similar levels of exploration we observed between adolescents and adults. However, this possibility should be tested empirically in future research.

There are some notable limitations to consider with the present study. It is important to note that the age effect on the  $\psi$  parameter was only significant after an exploratory covariate (i.e., gender) was added to this analysis, which may be consistent with some research that has found male adolescents' risk taking is more susceptible to social influence relative to female adolescents (McCoy et al., 2019). As such, these findings should be considered preliminary evidence for age effects on the social influence parameter  $\psi$ . To make this finding more robust, future research should seek to replicate these findings in a larger sample with an equal ratio of male and female participants, which would improve the ability to detect the small effect sizes

that may characterise age-related differences on the social influence parameter. In addition, we selected the images used to represent peers with the aim of matching these fictional peers to the participant's own age. As the adolescent cohorts were recruited from local schools, we were able to match the fictional peer to these participants' ages perfectly, as there was little variance in age within the classes recruited. In contrast, adult participants were recruited from the community and consequently there was a larger distribution of ages in this sample. As such, adults may not have believed they were receiving advice from an individual that was similar in age to them. However, we selected images of adults who were aged 30, based on the mean age of the adult sample from our previous study (Lloyd et al., 2021b), which closely matched the mean age of adult participants in the present study ( $M = 30.21$ ). Another limitation of our social influence manipulation was that the faces used to present peer advice were not randomised across social influence conditions, which means we cannot exclude the possibility that visual biases towards certain avatars contributed to our findings. Yet, our main findings were robust to the inclusion of our manipulation check (see Supplementary Material), suggesting participants' susceptibility to social influence was not affected by the extent to which they believed the manipulation.

In conclusion, the present study demonstrated that adolescents' and adults' exploration behaviours are susceptible to social influence while foraging. These findings are complemented by our computational modelling, where we debut a novel model that incorporates social influence during foraging. These analyses demonstrated that younger adolescents had higher values on a parameter measuring how much weight was placed on the advice from a fictional peer relative to adults. Notably, this model demonstrated the role that peers can have in signalling information relevant for computing the average reward rate of the environment, which can direct the individual's explore/exploit strategies. While we did not find evidence that participants' susceptibility to social influence is modulated by the type of advice they

receive, our findings can help to understand the mechanisms through which peers affect adolescents' decision-making (Ciranka & van den Bos, 2019). Overall, these findings contribute to understanding how peers influence exploration strategies during adolescence and how this can lead to (sub)optimal outcomes during the explore/exploit trade-off.

## 6.6. Supplementary Material

### 6.6.1. Supplementary Measures

Though we did not use these measures in our analyses, we also recorded the following measures:

*Pubertal Development:* The Pubertal Status scale (Carskadon & Acebo, 1993) is a measure of adolescents' pubertal development. As development is subject to individual differences, this measure provides information about an individual's development that is independent from chronological age. The scale is comprised of five questions that ask whether developmental milestones, such as physical growth, is underway or complete. The individual's pubertal development is calculated as a mean of their scores on the scale.

*Behavioural Inhibition and Behavioural Activation:* The BIS/BAS (Carver & White, 1994) is a combination of 20 items capturing behavioural inhibition scores (BIS) and behavioural approach scores (BAS) on a Likert scale ranging from 1-4. Higher scores on the BIS subscale denote higher sensitivity towards non-rewarding or unpleasant stimuli. The BAS is divided into three subscales measuring drive, fun seeking and reward responsiveness. The BAS scale therefore captures participants' responsiveness to rewarding stimuli. Together, the BIS/BAS questionnaire is a measure of participants' responsiveness to rewarding and aversive stimuli. In the current sample, mean scores on each subscale ranged from 1-4 and subscales ranged from acceptable to good reliability ( $\alpha = 0.64-0.79$ ). The BIS/BAS was collected as an exploratory measure, and we did not use this data in any further analyses.

*Impulsivity:* The Barratt-11 (Patton, Stanford & Barratt, 1995) measures the degree to which an individual acts impulsively. The scale comprises 30 items on a Likert scale ranging from 1-4. The items capture the degree to which individuals act without thinking, or do not engage in self-regulation. The individual's impulsivity is calculated as their mean score on the items in

the scale, which in the current sample ranged from 2-4. This scale demonstrated good reliability ( $\alpha = .80$ ). Self-reported impulsivity was also collected as an exploratory measure and not utilised in further analyses.

### **6.6.2. Exploratory Variables and Analyses**

Two dependent variables were calculated, which were not preregistered. The first was patch residency time, which was the average amount of time (in seconds) participants remained in patches before exploring. This variable was examined as we noted after data collection that participants in the social influence conditions were instructed by the fictional peer to stay for longer, less time, or the same amount of time in subsequent patches, rather than the instruction referring to the number of apples participants had seen prior to exploring (i.e., their leaving threshold). As such, we wanted to examine whether the effect of the social influence differed when analysing participants' patch residency times compared to participants' leaving thresholds.

Analyses using participants' patch residency time demonstrated that participants adjusted the length of time they spent in patches in a similar manner to their leaving thresholds. These findings suggest that the phrasing of the instructions did not alter our main findings, as we found similar trends between participants' patch residency time and the degree to which they deviated from the optimal leaving threshold, which we report in the main text. Specifically, a mixed  $2 \times 3 \times 4$  ANOVA demonstrated that participants remained significantly longer with patches in the poor quality environment compared to the rich environment ( $F_{(1,319)} = 62.78, p < .001, \eta^2 = 0.02$ ), consistent with optimal foraging theory (Charnov, 1976). Further, there was a main effect of social influence condition, such that participants adjusted the length of time they remained with patches according to the feedback they received from the peer ( $F_{(3,319)} = 5.11, p = .002, \eta^2 = 0.04$ ). Participants stayed for longer in patches in the overexploit social influence condition, which meant that participants would reach a lower leaving threshold

compared to the good advice condition and over-explore condition (see Supplementary Table 6.1). However, we did not observe a main effect of age group ( $F_{(2,319)} = 1.51, p = .222, \eta^2 = 0.006$ ), nor interactions between environment and social influence condition ( $F_{(3,319)} = 2.47, p = .062, \eta^2 = 0.003$ ), environment and age group ( $F_{(2,319)} = 0.89, p = .893, \eta^2 < .001$ ), social influence condition and age group ( $F_{(6,319)} = 0.41, p = .873, \eta^2 = 0.006$ ) nor environment, social influence condition and age group ( $F_{(6,319)} = 1.07, p = .381, \eta^2 = 0.002$ ). These results are consistent with our analyses of participants' deviation from the leaving thresholds that are reported in the main text.

*Supplementary Table 6.1: Mean (SD) of participants' average patch residency time in seconds.*

	No Feedback		Good Advice		Overexploit		Over-explore	
	Rich	Poor	Rich	Poor	Rich	Poor	Rich	Poor
Adult	21.59 (15.76)	24.66 (14.11)	17.22 (11.46)	19.59 (10.30)	19.01 (4.65)	23.87 (9.61)	17.54 (7.98)	20.02 (11.52)
Old	18.95 (9.89)	19.94 (8.14)	16.36 (7.18)	19.26 (5.02)	20.12 (5.02)	22.35 (7.01)	13.24 (5.55)	16.30 (7.07)
Young adolescent	22.69 (13.12)	22.82 (10.05)	15.54 (6.61)	20.00 (11.90)	18.15 (5.92)	23.93 (6.58)	16.15 (12.47)	19.69 (15.90)

The second variable calculated for the social influence conditions was the degree to which participants deviated from the peer's advice. To calculate this variable, participants' leaving threshold was subtracted from the value of the advice that we coded for the fictional peer. In the good advice condition, this was the optimal leaving threshold in each environment. In the overexploit condition, this value was two points (i.e., apples) less than the optimal leaving threshold and in the over-explore condition this was two points (i.e., apples) more than

the optimal leaving threshold. Positive values indicate participants remained with patches for less time than the peer recommended, and negative values indicate participants remained with patches for longer than the peer recommended.

We conducted a  $2 \times 3 \times 3$  ANOVA on the deviations from advice we coded for the peer in all the conditions where participants received advice from the fictional peer (i.e., the good advice conditions, overexploit condition and over-explore condition). Results of this ANOVA demonstrated a main effect of environment ( $F_{(1,245)} = 11.48, p < .001, \eta^2 = 0.004$ ), suggesting that participants deviated less from the leaving threshold recommended by the fictional peer in the rich quality environment relative to the poor quality environment. Further, there was a main effect of social influence condition on the how much participants deviated from the leaving threshold recommended by the peer ( $F_{(2,245)} = 78.55, p < .001, \eta^2 = .35$ ). The pattern of means on Supplementary Table 6.2 suggest that participants exhibited closer leaving thresholds to those recommended by the peer when the peer's advice was good, compared to when the advice was to either over-exploit or over-explore. Further, the descriptive statistics suggest that participants' behaviour trended towards the optimal foraging strategy, as participants in the overexploit condition explored more than their peer advised, while those in the over-explore condition exploited more than their peer advised (see Supplementary Table 6.2). However, there was not a main effect of age group ( $F_{(2,245)} = 1.05, p = 0.353$ ), indicating that adherence to the peer's advice did not vary across age groups. The interaction effect between environment and social influence condition was not significant ( $F_{(2,245)} = 0.53, p = 0.592$ ). Moreover, there were no significant interactions between environment and age group: ( $F_{(2,245)} = 1.03, p = 0.360$ ), social influence condition and age group ( $F_{(4,245)} = 0.54, p = 0.705$ ), or environment, social influence condition and age group ( $F_{(4,245)} = 0.99, p = .414$ ).

Supplementary Table 6.2: Mean (SD) of the degree to which participants deviated from their peer's advice in each environment. Negative values indicate the participant explored more than the peer advised whereas positive values indicate the participant exploited patches for longer than the peer advised.

	Good Advice		Over-exploit		Over-explore	
	Rich	Poor	Rich	Poor	Rich	Poor
Adult	-0.29 (1.89)	-0.56 (1.61)	-1.52 (1.05)	-1.82 (1.62)	1.81 (1.80)	1.30 (1.86)
Old	-0.46 (1.47)	-0.55 (1.05)	-1.53 (1.01)	-2.00 (1.42)	0.83 (1.52)	0.71 (1.52)
Adolescent	-0.55 (1.45)	-0.61 (1.70)	-1.79 (1.47)	-1.72 (1.40)	1.20 (2.13)	0.83 (2.07)

Notably, we observed a significant relationship between parameter estimates for  $\psi$ , which was the free parameter measuring sensitivity to social feedback, and our behavioural measure of how much participants deviated from the feedback provided by the peer ( $r(1, 226) = -0.14, p = .018$ ). This finding suggests that our formal model was able to capture a distinct mechanism relating to how participants utilised social feedback in the foraging task.

### 6.6.3. Parameter recovery

#### *Parameter recovery results*

To test whether the parameters in our models were identifiable we conducted parameter recovery on the social and asocial models we compared in the present study. Parameter recovery involves simulating task data with the values of  $\alpha$ ,  $c$ ,  $\beta$  and (for the social model only)  $\psi$  hard coded into the simulation and running the models over these synthetic datasets. Parameter recovery was run over the following values for each parameter:  $\alpha$  {0.1, 1} in increments of 0.1,  $\beta$  {1, 5} in increments of 0.5,  $c$  {-1, 3} in increments of 0.5 and  $\psi$  {0.1, 1}

in increments of 0.1. Significant positive correlations between hardcoded parameters and those that are estimated by the model indicate parameter recovery is successful.

Parameter recovery for the asocial learning model was successful, as there were significant positive correlations between simulated and estimated values for  $\alpha$  ( $r(658) = .88, p < .001$ ),  $\beta$  ( $r(658) = .25, p < .001$ ) and  $c$  ( $r(658) = .33, p < .001$ ). However, we did find a significant correlation between estimates for  $\alpha$  and  $\beta$  ( $r(658) = 0.55, p < .001$ ).

Parameter recovery for the social learning model was also successful as we observed significant positive correlations between simulated and estimated values for each of the free parameters. Correlation statistics were as follows:  $\alpha$  ( $r(748) = 0.28, p < .001$ ),  $\beta$  ( $r(748) = 0.51, p < .001$ ),  $c$  ( $r(748) = 0.54, p < .001$ ),  $\psi$  ( $r(748) = 0.32, p < .001$ ). We did not observe any correlations above  $r = 0.16$  between the parameters in this model.

#### **6.6.4. Pilot results**

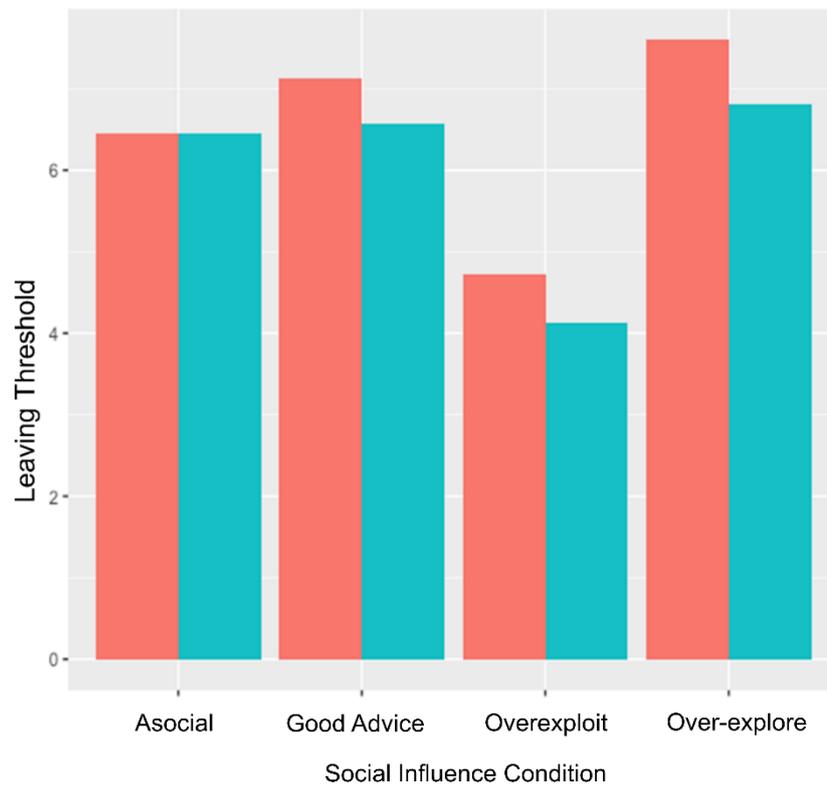
##### *Participants*

Sixteen adolescent participants aged between 16-19 years ( $M = 16.75, SD = 0.08, 87.5\%$  female) were recruited from a convenience sample on an open day hosted at the authors' institution.

##### *Results*

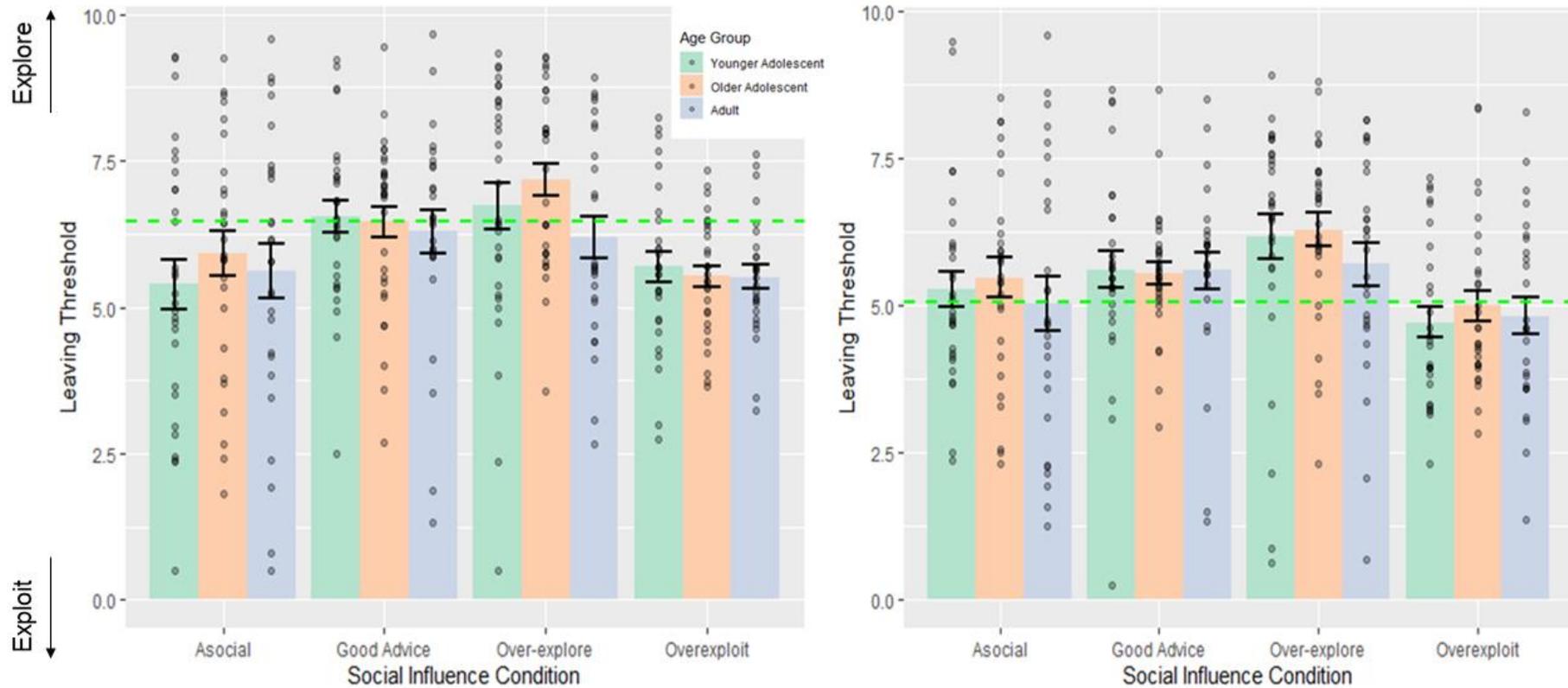
We examined the pilot data of 16 adolescent participants to determine whether our social influence manipulation was successful and could be utilised for the main study. These findings indicated that participants had adjusted their leaving thresholds in a manner that was consistent with the social influence condition they were assigned to. For example, in the over-explore condition, participants were advised to utilise a higher leaving threshold. Consistent with this feedback, participants utilised a higher leaving threshold in this condition relative to the control condition where they received no feedback (see Supplementary Figure 6.1). In contrast, in the

overexploit condition participants were advised to utilise a lower leaving threshold. Consistent with this feedback, participants utilised a lower leaving threshold in this condition relative to the control condition where participants did not receive feedback. This plot suggests there were behavioural differences between our social influence conditions, and we subsequently considered it appropriate to run this with our main sample.

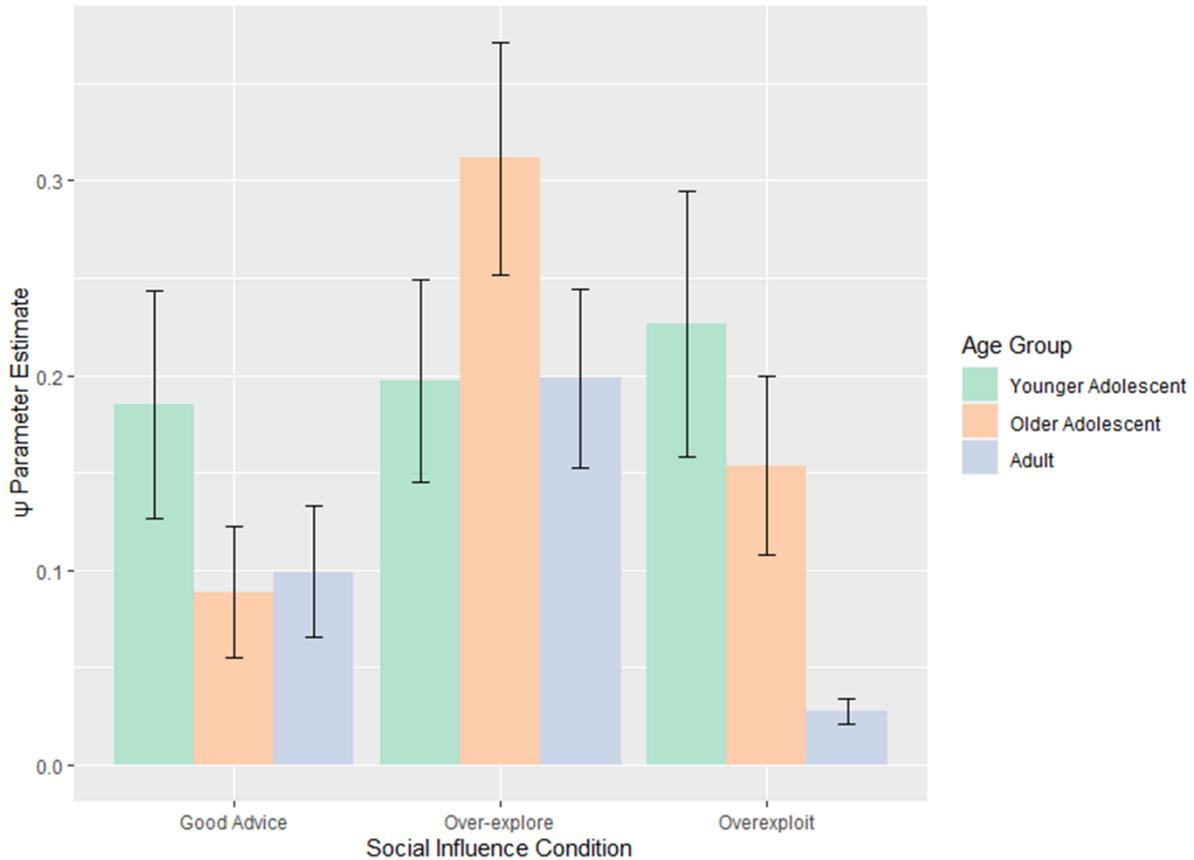


*Supplementary Figure 6.1: Bar plot demonstrating the leaving threshold (y-axis) for each of the social influence conditions (x-axis). Higher values denote more exploration. Red bars are the rich quality environment and turquoise bars indicate the poor-quality environment.*

### 6.6.5. Supplementary Visualisations and Discussion



Supplementary Figure 6.2: Plot demonstrating the mean and standard deviation of participants' leaving thresholds (y-axis) for each age group across the social influence conditions (x-axis), with individual datapoints in gray. The left plot is for the rich environment and the right plot is for the poor environment. The dashed green line indicates the optimal leaving threshold for that environment.



Supplementary Figure 6.3: Bar plot demonstrating the mean and standard error of parameter estimates for  $\psi$  (y-axis) across the three social influence conditions (x-axis). Error bars indicate standard deviations in each condition. Higher values on the y-axis indicate the participant was more susceptible to social influence. The asocial condition did not have a parameter estimate for susceptibility to social feedback and was omitted from this analysis.

While we did not find a statistically significant interaction between participants' age group and the social feedback they received on the computational parameter measuring susceptibility to social influence ( $\psi$ ), we did observe nonsignificant trends in estimates of this parameter (see Supplementary Table 6.3). As previous research has demonstrated that adults have a greater overexploitation bias relative to adolescents (Lloyd et al., 2021b), we predicted that adults would be more susceptible to social influence that endorsed overexploitation. However, we instead found a trend that compared to both younger and older adolescents, adults had lower values on the social influence parameter, suggesting they were less susceptible to

social information when the peer advised the participant to exploit patches for longer than optimal. One interpretation of this finding is that adults utilised social information less in the over-exploit condition relative to other social influence conditions, as this age group typically exploits patches for longer than optimal (Constantino & Daw, 2015). Due to this overexploitation bias, adults therefore required less social influence to engage in suboptimal exploitation. Moreover, as adolescence is a period of heightened exploration, we also predicted that adolescents would be more susceptible to advice that promoted over-exploration. Consistent with this hypothesis, we found a trend that older adolescents' exploration behaviour was more susceptible to social influence when the peer endorsed exploration (see Supplementary Figure 6.3), which could indicate a heightened susceptibility to information that promotes novelty-seeking. This finding is consistent with evidence that adolescents are more likely to adopt the behaviour of peers when this behaviour models risky behaviour rather than risk-averse behaviour (Reiter et al., 2019), as exploration and risk-taking are positively associated (Addicott et al., 2017). However, we recommend caution in interpreting these findings as differences on the parameter measuring social susceptibility did not translate to behavioural differences on the degree to which participants' leaving thresholds deviated from the optimal leaving threshold, though group-level differences in computational parameters do not always correspond to group-level differences in behaviour (e.g., Hauser et al., 2015).

*Supplementary Table 6.3: Mean (SD) estimates of the  $\psi$  parameter for younger adolescents, older adolescents and adults in each of the social influence conditions (excluding the asocial condition).*

<b>Condition</b>	<b>Age Group</b>	<b>Mean</b>	<b>SD</b>
Exploit	Adult	0.028	0.031
	Older Adolescent	0.153	0.229

	Young Adolescent	0.226	0.333
Explore	Adult	0.199	0.235
	Older Adolescent	0.311	0.321
	Young Adolescent	0.197	0.285
Good	Adult	0.099	0.168
	Older Adolescent	0.089	0.179
	Young Adolescent	0.185	0.299

---

#### 6.6.6. Separate ANOVAs for Each Age Group

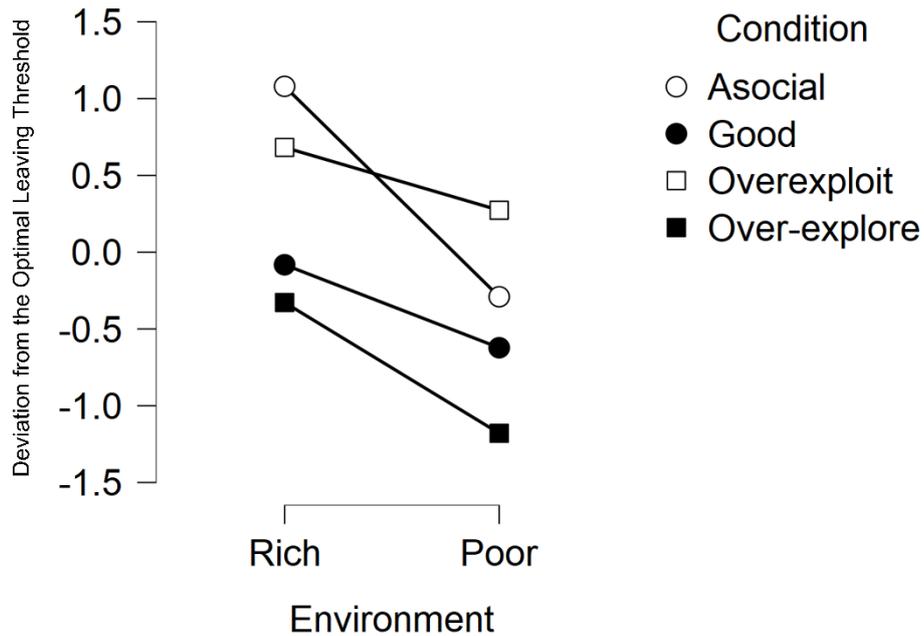
In our pre-registration document, we planned to conduct separate 2×4 ANOVAs for each age group to examine whether younger adolescents, older adolescents and adults exhibited susceptibility to social influence while foraging. However, we deviated from this pre-registered plan as we determined that it would be more appropriate to run separate ANOVAs for each social influence condition, rather than each age group, to test whether adolescents and adults were equally susceptible to social influence that endorsed exploration or exploitation. Our pre-registered analyses are reported below.

Examining young adolescents' data, we found a main effect of social influence condition on the degree to which younger adolescents deviated from the optimal leaving threshold ( $F_{(1,111)} = 3.61, p = .016$ ; see Supplementary Table 6.4). Moreover we found a main effect of foraging environment on the degree to which these participants deviated from the optimal leaving threshold ( $F_{(3,111)} = 49.32, p < .001$ ), with younger adolescent exhibiting more optimal behaviour (i.e., values closer to zero) in the rich quality environment relative to the

poor quality environment. In addition, we found an interaction between foraging environment and social influence condition ( $F_{(3,111)} = 3.61, p = .016$ ). Post-hoc tests using the Bonferroni correction demonstrated that the degree to which younger adolescents deviated from the optimal leaving threshold was significantly different between the rich and poor environments in the asocial condition ( $p_{bonf} < .001$ ), with participants exhibiting more optimal exploration in the poor-quality environment (see Supplementary Table 6.4). Moreover, the degree to which younger adolescents deviated from the optimal leaving threshold was significantly different between the rich and poor environments in the over-explore condition ( $p_{bonf} = .007$ ), with participants exhibiting more optimal behaviour in the rich quality environment relative to the poor-quality environment (see Supplementary Table 6.4); there were no between environment effects for the other social influence conditions ( $p_{bonf} > .050$ ). Inspection of this disordinal interaction revealed that this finding was driven by a cross over between the asocial condition and the overexploit condition (see Supplementary Figure 6.4). Specifically, young adolescents in the asocial condition overexploited in the rich environment and over-explored in the poor environment, whereas young adolescents in the overexploit condition exploited for longer than was optimal in the rich and poor environments. As we did not find a disordinal interaction between social influence conditions where younger adolescents received advice from a fictional peer (i.e., the good advice, overexploit and over-explore conditions), we suggest it is valid to interpret the main effect of social influence condition on the degree to which younger adolescents deviated from the optimal leaving threshold. Indeed, analysis of young adolescents' data excluding the asocial condition preserved the main effect of social influence condition ( $F_{(2,83)} = 4.16, p = .019$ ) without an interaction between environment and social influence condition ( $F_{(2,83)} = 1.47, p = .235$ ).

*Supplementary Table 6.4: Descriptive statistics demonstrating the degree to which younger adolescents deviated from the optimal leaving threshold in each foraging environment.*

<b>Environment</b>	<b>Condition</b>	<b>Mean</b>	<b>SD</b>
Poor	Asocial	-0.290	1.663
	Good	-0.622	1.704
	Over-explore	-1.179	2.066
	Overexploit	0.273	1.398
Rich	Asocial	1.080	2.263
	Good	-0.082	1.453
	Over-explore	-0.328	2.127
	Overexploit	0.682	1.323



Supplementary Figure 6.4: Plot demonstrating the interaction between social influence condition and environment in the younger adolescent group.

Examining older adolescents' data, we found a main effect of social influence condition on the degree to which older adolescents deviated from the optimal leaving threshold ( $F_{(1,113)} = 5.91, p < .001$ ; see Supplementary Table 6.5). Moreover we found a main effect of foraging environment on the degree to which these participants deviated from the optimal leaving threshold ( $F_{(3,113)} = 52.35, p < .001$ ), with older adolescents exhibiting more optimal leaving thresholds in the rich quality environment relative to the poor quality environment. We did not find an interaction between environment and social influence condition on the degree to which participants deviated from the optimal leaving threshold ( $F_{(3,113)} = 1.20, p = .314$ ).

Supplementary Table 6.5: Descriptive statistics demonstrating the degree to which older adolescents deviated from the optimal leaving threshold in each foraging environment.

Environment	Condition	Mean	SD
Poor	Asocial	-0.489	1.741

<b>Environment</b>	<b>Condition</b>	<b>Mean</b>	<b>SD</b>
	Good	-0.564	1.051
	Over-explore	-1.299	1.521
	Overexploit	-0.009	1.416
Rich	Asocial	0.547	2.006
	Good	0.012	1.472
	Over-explore	-0.705	1.517
	Overexploit	0.945	1.005

In contrast to the analyses of adolescents' data, we did not find a main effect of social influence condition on the degree to which adults deviated from the optimal leaving threshold ( $F_{(3,102)} = 1.39, p = .250$ ; see Supplementary Table 6.6). There was, however, a main effect of environment ( $F_{(1,102)} = 56.63, p < .001$ ), with adults exhibiting more optimal behaviour in the poor environment relative to the rich environment, replicating our previous work (Lloyd et al., 2021b). In addition, there was not an interaction between environment and social influence condition ( $F_{(3,102)} = 0.19, p = .907$ ). Tentatively, these findings could indicate that younger adolescents' and older adolescents' exploration strategies were susceptible to social influence, whereas there is not evidence for this effect in adults. However, we recognise the absence of an interaction effect reported with regards to hypotheses 2a-c in the main text.

Supplementary Table 6.6: Descriptive statistics demonstrating the degree to which adults deviated from the optimal leaving threshold in each foraging environment.

Environment	Condition	Mean	SD
Poor	Asocial	-0.035	2.375
	Good	-0.605	1.610
	Over-explore	-0.712	1.864
	Overexploit	0.169	1.623
Rich	Asocial	0.853	2.423
	Good	0.181	1.894
	Over-explore	0.277	1.799
	Overexploit	0.949	1.046

### 6.6.7. Analyses Using Absolute Deviation from the Optimal Leaving Threshold

In the main text we conducted analyses on the relative degree to which participants deviated from the optimal leaving threshold, with positive values indicating participants overexploited and negative values indicating participants explored more than was optimal. Using these relative values allowed us to test our hypotheses related to whether adolescents and adults explored or exploited more than was optimal in each of the social influence conditions. However, using relative values might have obscured our comparison of the overexploit and over-explore social influence conditions, as each of these conditions were differently signed (i.e., exploit behaviour was positively signed and explore behaviour was negatively signed). As such, we conducted analyses on the *absolute* deviation from the optimal

leaving threshold in the overexploit and over-explore conditions to examine whether adolescents and adults were more susceptible to social influence that endorsed suboptimal rates of exploration or exploitation. These findings replicated the results reported in the main text; namely, adolescents and adults were not differentially susceptible to social influence that endorsed exploration or exploitation. Moreover, replicating the trends observed in the main text, there was a significant difference between participants' susceptibility to social influence in the over-explore and overexploit conditions (see Figure 6.3).

We conducted a  $2 \times 3 \times 2$  ANOVA with factors for: environment (rich or poor), age group (younger adolescent, older adolescent or adult) and social influence condition (overexploit or over-explore). Notably, there was not an interaction between age group and social influence condition ( $F_{(1,162)} = 0.66, p = .517$ ), suggesting the susceptibility to social influence that endorsed suboptimal levels of exploration and exploitation did not differ between adolescents and adults. There was a main effect of condition ( $F_{(1,162)} = 15.39, p < .001$ ), with participants exhibiting less optimal exploration in the over-explore condition ( $M_{\text{rich}} = 1.55, SD_{\text{rich}} = 1.04, M_{\text{poor}} = 1.82, SD_{\text{poor}} = 1.07$ ) relative to the overexploit condition ( $M_{\text{rich}} = 1.20, SD_{\text{rich}} = 0.82, M_{\text{poor}} = 1.24, SD_{\text{poor}} = 0.79$ ). Further, there was a main effect of environment ( $F_{(1,162)} = 4.29, p = .040$ ), with participants exhibiting less optimal behaviour in the poor quality environment ( $M = 1.53, SD = 0.99$ ) relative to the rich environment ( $M = 1.36, SD = 0.95$ ). We did not find a main effect of age group ( $F_{(2,162)} = 1.22, p = .297$ ), nor an interaction between environment and social influence condition ( $F_{(1,162)} = 1.56, p = .214$ ), environment and age group ( $F_{(2,162)} = 0.18, p = .835$ ) or environment, age group and social influence condition ( $F_{(2,162)} = 1.16, p = .317$ ).

#### **6.6.8. Analyses Controlling for Gender**

We controlled for gender in our analyses of how much participants deviated from the optimal leaving threshold. Replicating the findings reported in the main text, the interaction between social influence condition and age group was not significant ( $F_{(6,325)} = 0.37, p = 0.897$ ).

There was a main effect of social influence condition on deviation from the optimal leaving threshold ( $F_{(3,325)} = 8.98, p < .001, \eta^2 = .07$ ); consistent with analyses reported in the main text, participants exhibited more optimal exploration in the overexploit condition relative to the good advice condition. Further, participants exhibited less optimal exploration in the overexploit condition relative to the over-explore condition. In addition, there was not a main effect of age group ( $F_{(2,325)} = 1.00, p = 0.370$ ), indicating that the degree to which participants deviated from the optimal leaving threshold did not vary by age group. We found a main effect of environment ( $F_{(1,325)} = 8.447, p = .004, \eta^2 = 0.003$ ), with participants performing more optimally in the poor quality environment. Further, we did not find significant interactions between: environment and social influence condition ( $F_{(3,325)} = 2.41, p = 0.067$ ), environment and age group ( $F_{(2,325)} = 0.14, p = 0.869$ ), or environment, social influence condition and age group ( $F_{(6,325)} = 1.26, p = 0.278$ ), environment and sex ( $F_{(6,325)} = 0.38, p = 0.537$ ), nor a main effect of gender ( $F_{(5,325)} = 1.31, p = 0.253$ ).

#### **6.6.9. Analyses Controlling for Manipulation Check**

We examined whether participants' belief that they were receiving feedback from a real peer altered whether they followed the advice of the peer. To test this, we ran analyses on the degree to which Participants deviated from the optimal leaving threshold in the conditions in which participants received advice from a fictional peer (i.e., the good advice condition, the overexploit condition and the over-explore condition). Our manipulation check, which was a continuous measure, was entered as a covariate in this model. These results duplicated the findings where this manipulation check was not included, reported in the main text, as we found a main effect of environment quality ( $F_{(1,243)} = 9.13, p = .003, \eta^2 = 0.004$ ), suggesting participants were more optimal in the poorer quality environments (see Supplementary Figure 6.2). Further, there was a main effect of social influence condition on how much participants deviated from the optimal leaving threshold ( $F_{(2,243)} = 12.96, p < .001, \eta^2 = .08$ ), with

participants exhibiting more optimal behaviour in the good advice condition relative to the overexploit condition ( $p_{bonf} = .002$ ). Consistent with our analyses reported in the main text, post hoc tests also demonstrated that participant exhibited less optimal leaving thresholds in the overexploit condition relative to the over-explore condition ( $p_{bonf} < .001$ ), though we did not find evidence that there was a difference between how much participants deviated from the optimal leaving threshold between the good advice condition and the over-explore condition ( $p_{bonf} = .405$ ). In addition, there was not an effect of the manipulation check ( $F_{(2,243)} = 0.16, p = .687$ ), nor age group ( $F_{(2,243)} = 1.05, p = .351$ ), indicating there were no age-related differences in participants' deviation from the optimal leaving threshold. The interaction effect between environment and social influence condition was not significant ( $F_{(3,243)} = 0.74, p = .481$ ), nor were the interactions between environment and age group: ( $F_{(2,243)} = 0.67, p = .513$ ), social influence condition and age group ( $F_{(4,243)} = 0.53, p = .712$ ), or environment, social influence condition and age group ( $F_{(4,243)} = 1.27, p = .281$ ). Importantly, these results suggest that our results were not affected by the degree to which participants believed the peer was real.

#### **6.6.10. Group comparisons of model fit indices**

Previous research has demonstrated that the best fitting model can differ between populations, yielding insights into mechanistic differences in cognition between groups (e.g., Collins et al., 2014). To test whether different models explained adolescents' and adults' foraging behaviour, we compared the AIC and BIC for the social and asocial models between age groups. We conducted two mixed ANOVAs with the model (social or asocial) as the within subject factors and age group as the between subject factors. The AIC and BIC of each model were included as the DVs in each separate analysis. Results of the ANOVAs demonstrated that the model fits did not vary systematically by age group as indicated by the model AICs ( $F_{(2,211)} = 1.49, p = .227$ ) or BIC ( $F_{(2,211)} = 1.52, p = .222$ ). Importantly, there was no interaction between model and age group for AIC ( $F_{(2,211)} = 0.36, p = .700$ ) or BIC ( $F_{(2,211)} = 0.34, p = .711$ ).

### 6.6.11. Group comparisons of additional free parameters

We conducted ANCOVAs on each of the additional free parameters in our social RL model, controlling for gender. These results demonstrated there was not a main effect of age on participants' learning rate (i.e.,  $1-\alpha$ ;  $F_{(2,225)} = 0.97, p = .380$ ). There was a significant main effect of social influence condition on learning rate ( $F_{(2,225)} = 18.32, p < .001$ ), with participants utilising a lower learning rate (i.e., higher values on  $\alpha$ ) in the overexploit social influence condition relative to the good advice social influence condition ( $p_{bonf} < .001$ ) and the over-explore condition ( $p_{bonf} < .001$ ; see Supplementary Table 6.6). However, there was not a significant interaction between social influence condition and age group ( $F_{(2,225)} = 0.86, p = .491$ ).

Examining the parameter estimates for  $\beta$ , which measures the degree of stochasticity in participants' decision making, we did not find a main effect of age group on this parameter ( $F_{(2,225)} = 1.10, p = .335$ ). There was a significant main effect of social influence condition on  $\beta$  ( $F_{(2,225)} = 15.53, p < .001$ ), with participants exhibiting more stochastic behaviour (i.e., lower values on  $\beta$ ) in the overexploit social influence condition relative to the over-explore ( $p_{bonf} < .001$ ) and good ( $p_{bonf} < .001$ ) social influence conditions (see Supplementary Table 6.7). However, there was not a significant interaction between social influence condition and age group ( $F_{(2,225)} = 1.04, p = .387$ ).

Examining the parameter estimates for participants' overexploitation bias ( $c$ ), we also found no main effect of age group ( $F_{(2,225)} = 0.70, p = .500$ ). There was a significant main effect of social influence condition on overexploitation bias ( $F_{(2,225)} = 28.58, p < .001$ ). Participants had a higher overexploitation bias in the social influence condition where they received advice to overexploit, relative to the good advice ( $p_{bonf} < .001$ ) and the over-explore ( $p_{bonf} < .001$ ) social influence conditions (see Supplementary Table 6.6). However, there was not a

significant interaction between social influence condition and age group ( $F_{(2,225)} = 0.59, p = .667$ ).

*Supplementary Table 6.7: Mean (SD) parameter estimates for the free parameters  $\alpha$ ,  $\beta$  and  $c$  in the winning social learning model.*

	Younger Adolescent			Older Adolescent			Adult		
	Over-exploit	Good	Over-explore	Over-exploit	Good	Over-explore	Over-exploit	Good	Over-explore
$\alpha^*$	0.73 (0.32)	0.53 (0.43)	0.44 (0.42)	0.67 (0.33)	0.35 (0.38)	0.37 (0.41)	0.77 (0.33)	0.48 (0.41)	0.27 (0.38)
$\beta$	0.69 (1.06)	0.12 (0.69)	0.18 (0.78)	0.72 (0.89)	0.12 (0.78)	0.01 (0.71)	1.00 (1,17)	0.50 (1.02)	0.05 (0.68)
$c$	2.62 (0.96)	1.54 (1.08)	1.65 (1.15)	2.76 (0.57)	1.73 (1.13)	1.21 (1.43)	2.56 (0.63)	1.53 (1.17)	1.31 (1.15)

\* Note, in our implementation, the learning rate is equal to  $1-\alpha$  and therefore higher values on this parameter denote a lower learning rate.

Notably, in contrast to the correlations reported in the parameter recovery above (see 6.6.3. Parameter Recovery), we observed significant positive correlations in the parameters estimated from participants' data. Specifically, we observed significant correlations between the learning rate parameter ( $\alpha$ ), the overexploitation parameter ( $c$ ), and the  $\beta$  parameter (see Supplementary Table 6.8). In addition, we found a significant negative correlation between the social influence parameter ( $\psi$ ) and the parameter measuring stochasticity in participants' decision making ( $\beta$ ). Lower values on the  $\beta$  parameter indicate that participants were less consistent in their decision making. As such, one interpretation of this finding is that the use of social information increased participants' cognitive load, which can increase stochasticity in decision-making (Tymula, 2019).

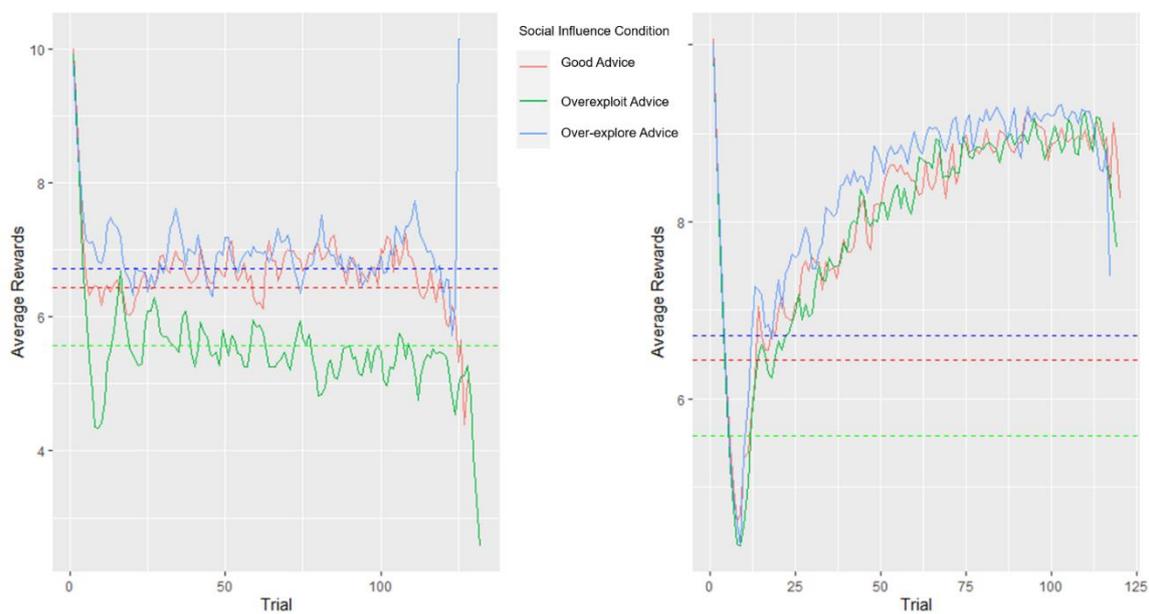
Supplementary Table 6.8: Correlations between free parameter estimates for the social learning model.

Variable		$\alpha$	$c$	$\beta$	$\psi$
1. $\alpha$	Pearson's r	—			
	p-value	—			
2. $c$	Pearson's r	0.277	—		
	p-value	< .001	—		
3. $\beta$	Pearson's r	0.616	0.232	—	
	p-value	< .001	< .001	—	
4. $\psi$	Pearson's r	-0.126	-0.114	-0.225	—
	p-value	0.054	0.082	< .001	—

### 6.6.12. Post-Hoc Simulations and Model Recovery

We conducted post-hoc simulations to examine whether we could replicate the main effect of the social influence condition on participants' leaving thresholds, specifically the effects reported in Figure 6.3. Further, we aimed to examine whether the pattern of the social influence conditions was better reproduced by the social model compared to the asocial model. We took the mean parameter estimates for each model in each condition and simulated 100 agents' behaviour according to these models. Plots of these simulated agents demonstrated that the social model was a better reflection of participants' actual performance in the foraging task relative to the asocial model (see Supplementary Figure 6.5). Specifically, the lines in Supplementary Figure 6.5 represent the average rewards experienced by the simulated agents. Local troughs represent the value of rewards the agent received before leaving for a new patch and local peaks represent new patches. The plots of the simulated social model (left, Supplementary Figure 6.5) demonstrate that the local troughs were higher in the condition

where the agent received advice to over-explore, indicating a higher leaving threshold, whereas the local troughs are lower in the overexploit condition, indicating a lower leaving threshold. These patterns replicate the post-hoc tests reported in the main text that demonstrate the greatest difference between the social influence conditions was between the overexploit and over-explore conditions (see Figure 6.3). These simulations provide additional evidence that the winning social model provided a better explanation of our data compared to the asocial model.



*Supplementary Figure 6.5: Plots of the social model (left) and asocial model (right) using means of participants' parameter estimates for each social influence condition. Horizontal dashed lines indicate the average leaving threshold observed in our behavioural data for each condition.*

### *Model recovery*

We used the social learning model to simulate the data of 20 participants and then verified that our model comparison on the simulated data could correctly recover the social learning model instead of the asocial model. These findings indicated that the model was recoverable when using the AIC as the index of fit, whereas model recovery was not successful when using

the BIC as the index of model fit (see Supplementary Table 6.9). However, as we note in the main manuscript, the penalty for the inclusion of additional free parameters is harsher for the BIC relative to the AIC. As the difference between the asocial and social model BICs are below one, we suggest that our findings support the view that the social learning model was recoverable.

*Supplementary Table 6.9: Model fit indices for the model recovery analyses.*

	<b>AIC</b>	<b>BIC</b>
Asocial Model	-67.20	-58.68
Social Model	-69.22	-57.85

## **Chapter 7: Individuals with Adverse Childhood Experiences Explore Less and Underweight Reward Feedback<sup>6,7</sup>**

### **7.1. Abstract**

Adverse Childhood Experiences (ACEs) are extreme stressors that lead to negative psychosocial outcomes in adulthood. Non-human animals explore less after exposure to early stress. Therefore, in this pre-registered study we hypothesised that reduced exploration following ACEs would also be evident in human adults. Further, we predicted that adults with ACEs, in a foraging task, would adopt a decision-making policy that relies on the most recent reward feedback, a rational strategy for unstable environments. We analysed data from 145 adult participants, 47 with four or more ACEs and 98 with fewer than four ACEs. In the foraging task, participants evaluated the trade-off between exploiting a known patch with diminishing rewards and exploring a novel one with a fresh distribution of rewards. Using computational modelling, we quantified the degree to which participants' decisions weighted recent feedback. As predicted, participants with ACEs explored less. However, contrary to our hypothesis, they underweighted recent feedback. These unexpected findings indicate that early adversity may dampen reward sensitivity. Our results may help to identify cognitive mechanisms that link childhood trauma to the onset of psychopathology.

---

<sup>6</sup> This Chapter is now in press at *Proceedings of the National Academy of Sciences of the United States of America* as: Lloyd, A., McKay, R., & Furl, N. (in press). Individuals with Adverse Childhood Experiences Explore Less and Underweight Reward Feedback. *Proceedings of the National Academy of Sciences of the United States of America*.

<sup>7</sup> OSF page: <https://osf.io/8znyx/>

## **Significance Statement**

Adverse Childhood Experiences (ACEs) are extreme stressors that have a profound impact on cognitive development. Using an explore/exploit foraging paradigm, we demonstrate that ACEs are associated with reduced exploration, leading these individuals to accumulate fewer rewards from their environment. Using computational modelling, we identify that reduced exploration is associated with ACE-exposed individuals underweighting reward feedback, which highlights a cognitive mechanism that may link childhood trauma to the onset and maintenance of psychopathology.

## 7.2. Introduction

Across animal species, an organism's survival depends on its ability to adapt to the conditions of its environment (McNamara & Houston, 1985). 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 (Boyce & Ellis, 2005). Early experiences of adversity have been associated with significant negative psychosocial outcomes (Danese & McEwen, 2012). However, these outcomes may result from strategies of decision-making which individuals adopt to cope with their early caregiving environments (Humphreys et al., 2015). 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 0-18 years old; Danese & McEwen, 2012). 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; Hughes et al., 2017). Notably, higher rates of ACEs are associated with poorer health and social outcomes in adulthood, such as substance misuse and antisocial behaviour (Baglivio et al., 2015; Mersky et al., 2013).

A theoretical account known as life history theory proposes that adverse rearing conditions direct individuals to adopt strategies that maximise short-term benefits (Ellis & Del Giudice, 2019). This behaviour 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. Behaviourally, greater delay discounting is negatively correlated with exploration, suggesting these preferences both reflect cognitive

processes related to the individual's temporal horizon (Sadeghiyeh et al., 2020). Consistent with life history theory, adolescent rats exposed to early stress explore their environment less compared to controls (Spivey et al., 2008). Moreover, in human adolescents, adversity related to childhood poverty is associated with a preference for immediate rather than delayed rewards in delay discounting paradigms (Oshri et al., 2019). These findings are notable as decision-making that prioritises short term rewards can lead to poorer socioeconomic outcomes (Haushofer & Fehr, 2014) and has also been linked to problematic health behaviours (e.g., substance misuse; Kollins, 2003). 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 focussed 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 (Belsky, 2021). This benefit has been demonstrated in adolescents with experience of early life stress (specifically events leading to institutionalisation) using a variant of the Balloon Analogue Risk Task (BART; Lejuez et al., 2002). In a study by Humphreys and colleagues, adolescents who had experienced institutionalisation explored less when the environment favoured less exploration, meaning they collected more rewards in this condition compared to adolescents with more stable upbringings (Humphreys et al., 2015). Furthermore, in this same study, adolescents with experience of institutionalisation explored less when the optimal decision-making strategy was to explore more, meaning 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 (2015) 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 towards negative feedback (Miu et al., 2017) and therefore previous research demonstrating reduced exploration following early stress may be confounded by heightened loss aversion in this population (Huh et al., 2016). 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 institutionalisation led to a reduction in exploration in adulthood. Humphreys and colleagues' (2015) focus on institutionalisation 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 (Office of National Statistics, 2020; Bellis et al., 2015). Identifying how ACEs affect exploration can contribute towards 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 favour 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 (Charnov, 1976; Stephens & Krebs, 1976.). 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 (Bettinger & Grote, 2016). This is formalised in a computational account of foraging known as Marginal Value Theorem (MVT). The theorem proves that to maximise 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 (Charnov, 1976). Previous research has found that human adults adjust their

foraging strategies according to qualities of the environment so as to maximise rewards (Constantino & Daw, 2015). Nevertheless, adults explore less than an optimal foraging strategy would dictate (Lloyd et al., 2021b). 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 behaviour and the individual learns that positive and negative behaviours occur at random (Novick et al., 2018). Indeed, a recent paper has characterised early adversity as a violation of environmental predictability, which has profound consequences for sensitive periods of development (Nelson III & Gabard-Durnam, 2020). 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 (Danese & McEwen, 2012). It has been argued that such conditions of adversity can lead to schemas of unpredictability, where the world is perceived as unstable (Ross & Hill, 2002). Consistent with this view, individuals exposed to childhood adversity develop neurobiological and behavioural adaptations to navigate changeable environments, such as rapidly shifting attention (Jensen et al., 1997). Together, this evidence suggests that individuals who have been exposed to ACEs perceive the environment as relatively volatile, which can impact neurobiological and behavioural outcomes (Danese & McEwen, 2012).

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 (Behrens et al., 2007). As such, decisions in volatile environments should be based on more recent feedback about whether actions will be rewarded (Browning et al., 2015). As a real-world example, during the COVID-19 pandemic, information regarding health behaviours rapidly changed in response to the spread of the virus and, as such, adaptive decision-making involved utilising the most recent information to make health-related choices (Lloyd et al., 2021a). The utility of relying on recent feedback in volatile environments has been demonstrated in instrumental learning paradigms, where 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 (Behrens et al., 2007). 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 (Sutton & Barto, 2018). By contrast, in environments where the probability of receiving rewards from each action remains stable, the optimal learning strategy is to utilise a wider range of historic experience during decision-making to avoid overweighting rare events (i.e., use a lower learning rate; Behrens et al., 2007).

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; Lawson et al., 2017). 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 (Pulcu & Browning, 2019). Neuroimaging research has demonstrated that ACEs lead to reduced volume of the anterior cingulate cortex (ACC; Cohen et al., 2006), a region that has been implicated in tracking environmental volatility (Behrens et al., 2007) and the value of exploring new patches while foraging (Kolling et al., 2012; Blanchard & Hayden, 2014; however, see Shenhav et al., 2014). As such, we hypothesised 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 (Danese & McEwen, 2012), 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 behaviour (leaving thresholds) on a patch foraging task in individuals with more or fewer ACEs and fit a reinforcement learning model (Constantino & Daw, 2015) to their behaviour 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 (see Hypothesis 1a), they should be closer to optimal in conditions where exploitation garners greater rewards (Humphreys et al., 2015). 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 behaviours.

### 7.3. 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 (Constantino & Daw, 2015; Lloyd et al., 2021b) that in the rich environment, participants had a higher leaving threshold than in the poor quality environment ( $F_{(1,137)} = 28.26, p < .001, \eta^2 = .03$ ). Furthermore, participants in the high ACE group remained in patches significantly longer (i.e., explored less) than participants with less exposure to ACEs ( $F_{(1,137)} = 4.46, p = .037, \eta^2 = .03$ ) (see Figure 7.1). There was no interaction between environment type and ACE exposure ( $F_{(1,137)} = 0.63, p = .429, \eta^2 < .001$ ). 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 (see Supplementary Materials).

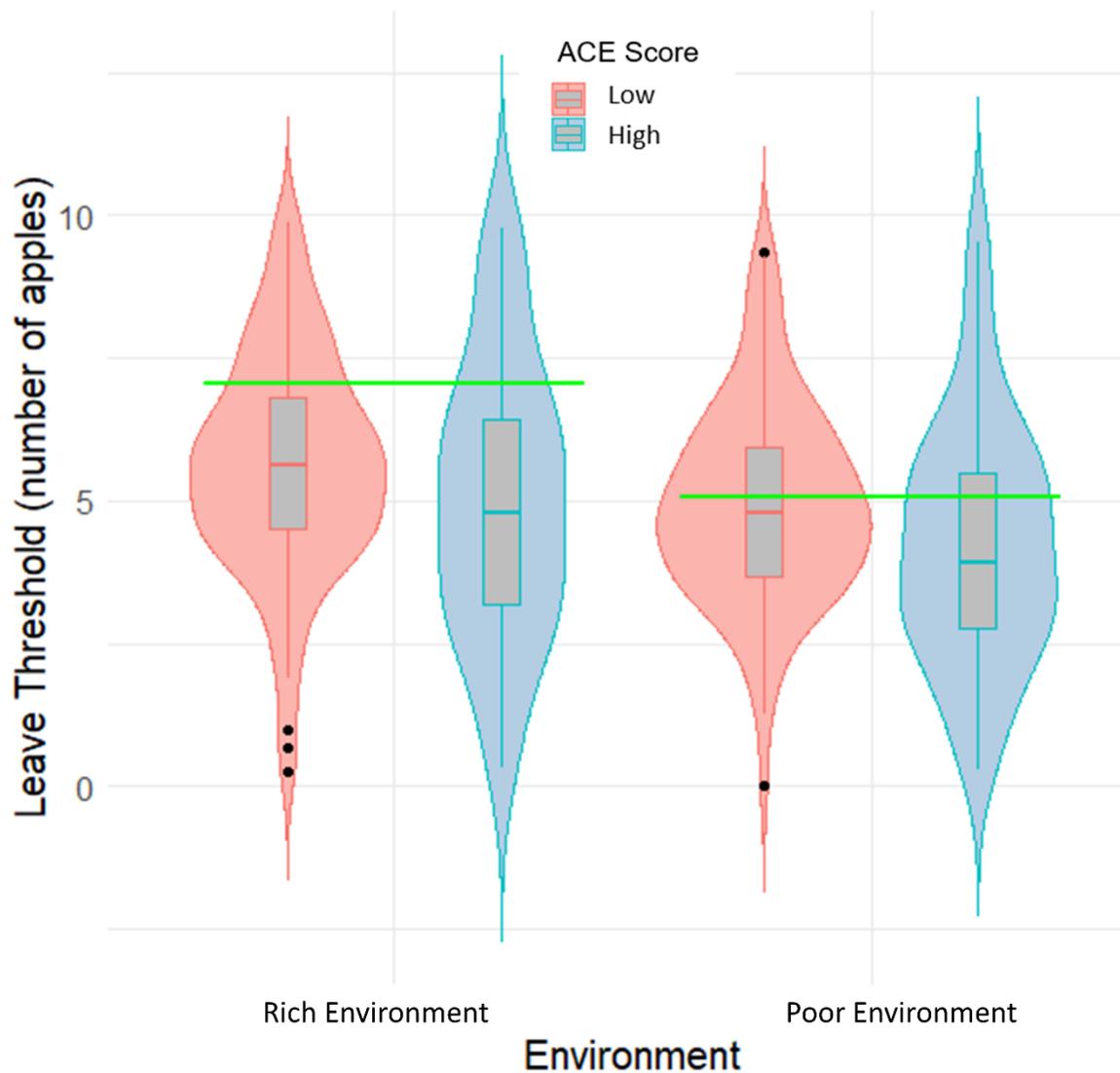
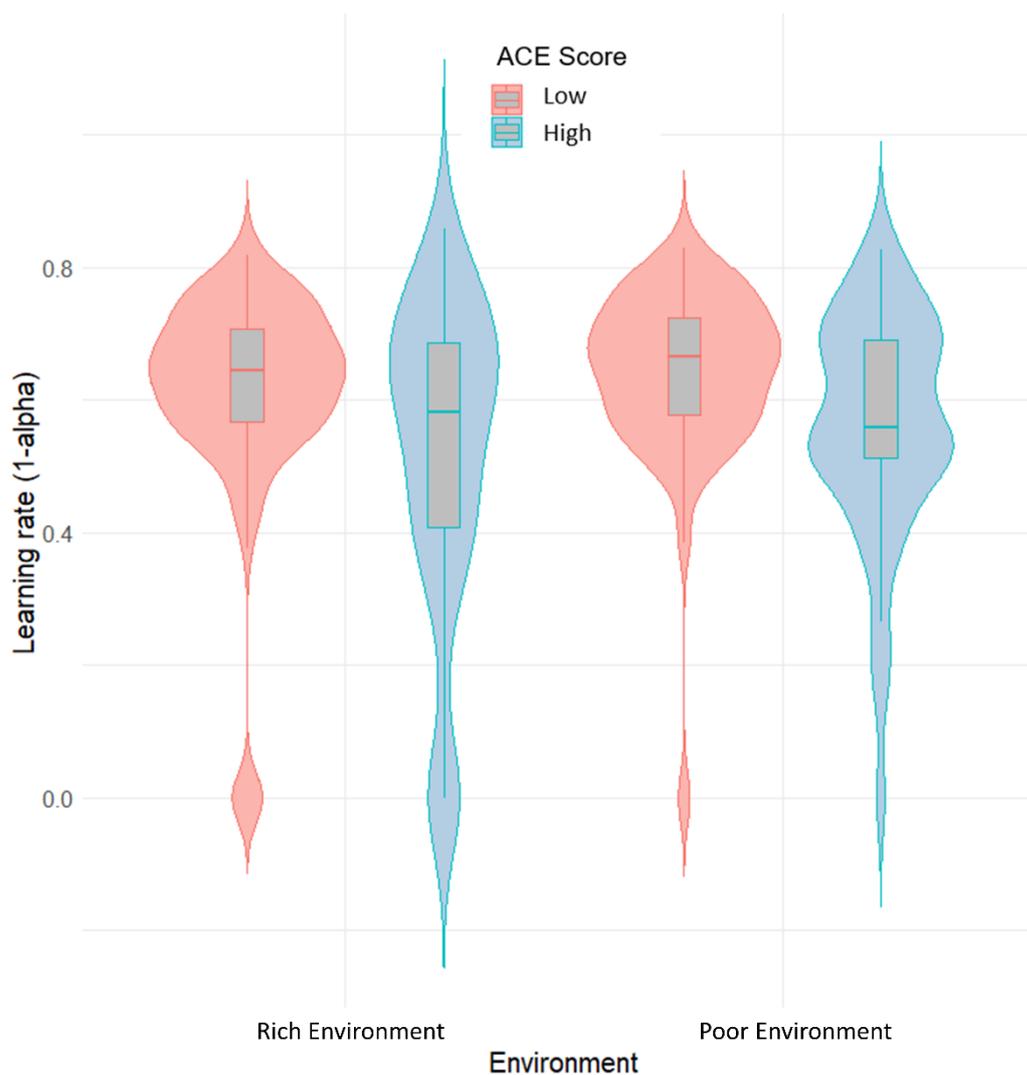


Figure 7.1: Violin and boxplots demonstrating the leaving thresholds for each environment. The central line in each boxplot refers to the mean. Upper and lower lines reflect the upper and lower quartiles, respectively. Black dots identify outliers that are greater than 1.5 times the interquartile range from the upper and lower quartile. Horizontal green bars indicate the optimal leaving thresholds for each environment. Higher leaving threshold indicates more exploration and less exploitation.

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 group ( $F_{(1,137)} = 8.92, p = .003, \eta^2 = 0.05$ ). This finding was in the opposite direction to our hypothesis and suggested those in the high ACE group weighted recent feedback lower

than those in the low ACE group (see Figure 7.2). This analysis also revealed that participants adjusted their learning rate between the two environments ( $F_{(1,137)} = 9.63, p = .002, \eta^2 = .01$ ), as participants utilised 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 ( $F_{(1,137)} = 0.40, p = .527, \eta^2 < .001$ ). 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 ( $F_{(1,137)} = 0.25, p = .621, \eta^2 < .001$ ) and the intercept parameter  $c$  ( $F_{(1,137)} = 0.76, p = .377, \eta^2 = .003$ ).



*Figure 7.2: 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.*

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 ( $r(149) = -.91, p < .001$ ) and the poor-quality environment ( $r(149) = -.71, p < .001$ ). 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' behaviour were able to reproduce the differences in leaving thresholds we observed between the high and low ACE groups as well as between environments (see Supplementary Materials).

To examine whether ACEs are associated with heightened risk taking (Hypothesis 2), we conducted several regression analyses. Each subscale of the 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 (see Supplementary Materials for full model statistics).

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 ( $M = 2.04$ ,  $SD = 2.46$ ) and poor-quality environments ( $M = 0.79$ ,  $SD = 2.16$ ) compared to individuals with low ACE scores ( $M_{rich} = 1.43$ ,  $SD_{rich} = 1.94$ ,  $M_{poor} = 0.03$ ,  $SD_{poor} = 1.68$ ,  $F_{(1,137)} = 4.46$ ,  $p = .037$ ,  $\eta^2 = .03$ ). Participants across the board exhibited more optimal behaviour in the poor quality environment than in the rich quality one ( $F_{(1,137)} = 107.31$ ,  $p < .001$ ,  $\eta^2 = .09$ ). There was no interaction between ACE score and environment ( $F_{(1,137)} = 0.63$ ,  $p = .429$ ,  $\eta^2 < .001$ ).

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 ( $F_{(1,137)} = 24.39$ ,  $p < .001$ ,  $\eta^2 = .09$ ; see Figure 7.3). Further, participants accumulated more points in the rich environment compared to the poor environment, as demonstrated by a main effect of environment type ( $F_{(1,137)} = 187.14$ ,  $p < .001$ ,  $\eta^2 = 0.24$ ). There was not a significant interaction between environment type and ACE group ( $F_{(1,137)} = 3.79$ ,  $p = .053$ ,  $\eta^2 = .01$ ).

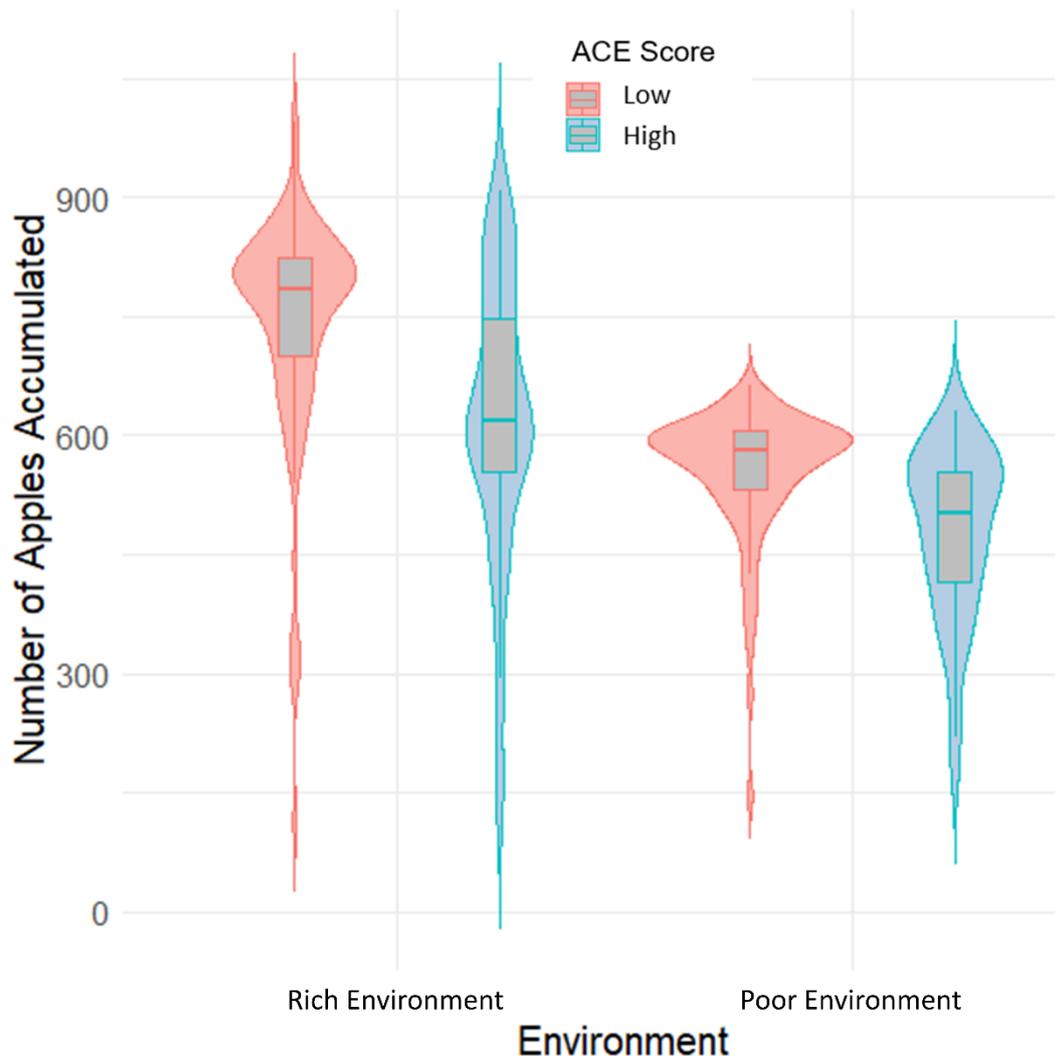


Figure 7.3: Apples accumulated in each environment by ACE group. The central line in each boxplot refers to the mean. Lines above and below the mean reflect the upper and lower quartile, respectively.

#### 7.4. Discussion

The present study tested whether ACEs are associated with reduced exploration and with the degree to which individuals weight recent feedback, as measured by participants' learning rate. Consistent with our pre-registered predictions, individuals with more ACEs explored their environment significantly less (had lower leaving thresholds) than individuals who reported fewer ACEs. However, contrary to our hypotheses, ACEs were associated with a *lower* learning rate, meaning participants who reported these experiences integrated recent

feedback less in their decision-making. We also found no associations between ACEs and self-reported risk taking. While these results were not entirely in line with our predictions, they are consistent with evidence that for individuals who are exposed to them, ACEs introduce patterns into decision-making, which have deleterious outcomes that last into adulthood.

Using an explore/exploit foraging paradigm, our findings demonstrate that early adversity is associated with reduced exploration of one's surroundings. Our findings are consistent with a previous study, which found that rats exposed to early stress demonstrated reduced exploration (Spivey et al., 2008). In addition, these findings build on important work with human participants that has investigated the impact of early stress in adolescence (e.g., Humphreys et al., 2015) by demonstrating how a novel set of stressors, ACEs, impact behavioural outcomes in adulthood.

While foraging, the decision-maker must compare the value of the current patch to the average reward rate for the environment when evaluating whether to explore or exploit (Charnov, 1976; Kolling et al., 2012). Participants exposed to ACEs appeared poorer at evaluating this trade-off and were more likely to select the option that yielded an immediate reward (i.e., exploitation). These findings are consistent with empirical and theoretical work which suggests that experience of instability during childhood directs the individual later in life towards decision-making strategies that maximise short-term rewards (Ellis & Del Giudice, 2019; Oshri et al., 2019). Results of our computational modelling indicated that the preference for immediate rewards in ACE-exposed individuals was driven by a reduced sensitivity to reward feedback. Specifically, upon arriving at a new patch with a large initial harvest, participants who have a lower learning rate integrate this reward feedback less into their estimate of the average reward rate compared to participants who have a higher learning rate. As such, participants with a lower learning rate may underweight the larger bounty of rewards that can be gained through exploration relative to participants with a higher learning rate,

leading to less frequent exploration. Furthermore, our findings demonstrate that reduced exploration exhibited by ACE-exposed individuals led them to accumulate fewer rewards from the environment. These findings highlight how a preference for immediate reward can prevent individuals from taking advantage of the full panoply of rewards available in an environment. This is important as the preference for immediate rewards has been causally linked to poorer socioeconomic outcomes (Haushofer & Fehr, 2014), so our results inform our understanding of the link between childhood trauma and adult poverty (Metzler et al., 2017).

We did not find evidence that individuals with higher levels of ACE exposure were more optimal in a foraging environment that favoured less exploration (the poor-quality environment). However, several studies have found that adults typically exploit patches for longer than is optimal (Constantino & Daw, 2015; Le Heron et al., 2019), suggesting even typical adults' foraging behaviour is already somewhat suited to environments that are poorer in quality. As such, future research should seek to recruit populations that do not demonstrate this bias to overexploit, such as adolescents (Lloyd et al., 2021b). Examining how ACEs affect exploration in adolescence would also be important for theoretical reasons, as this is a period where exploration serves a developmental purpose, providing this age group with the experiential knowledge necessary for adult independence (Romer et al., 2017). As the present foraging task lends itself to formal modelling techniques, this paradigm can be used to identify mechanistic explanations for reduced exploration in adolescents who experience early adversity (Humphreys et al., 2015). A failure to engage in typical levels of exploration during adolescence can have long-term psychosocial impacts (Larsen & Luna, 2018), highlighting the need to understand environmental factors that lead to reduced rates of exploration at this point in the lifespan.

Unexpectedly, we found that ACEs were associated with lower learning rates, meaning individuals with high numbers of ACEs weighted recent feedback lower than individuals with

fewer of these experiences. Our original prediction was based on evidence that ACEs dispose the individual to perceive the environment as unstable (Danese & McEwen, 2012) and on the predictions of theories such as life history theory that these early experiences lead the individual to adopt decision-making strategies to suit their environment (Ellis et al., 2011). We therefore hypothesised that ACE-exposed individuals would utilise decision-making strategies that were adapted to the unstable reward availability in their formative environments, leading them to adopt a strategy that prioritises recent feedback.

While we did not find support for this hypothesis, this inconsistency might be reconciled if one considers the evidence that individuals with high numbers of ACEs underweight recent feedback about stimulus-outcome contingencies due to their difficulty utilising positive feedback (Hanson et al., 2015). For example, women who had experienced childhood sexual abuse demonstrated a poorer ability to utilise positive reward feedback to guide future decisions compared to participants without these experiences (Pechtel & Pizzagalli, 2013). This insensitivity to reward feedback might arise from hypoactive neural responses to rewards. Indeed, individuals with experience of trauma exhibit less activation in the ventral striatum upon receiving reward feedback compared to controls (Boecker et al., 2014; Hanson et al., 2015). As the striatum encodes stimulus-outcome contingencies for gains (but not losses; Taswell et al., 2018), it is a prime candidate for the region where recent reward feedback might be underweighted in individuals with high numbers of ACEs. However, this hypothesis would need to be tested empirically (Kolling et al., 2012) in future research.

Along with being associated with hypoactive striatal responses to rewards (Hanson et al., 2017), ACEs have also been associated with hyperactive responses to punishment (Miu et al., 2017). This poses a further testable hypothesis, which could explain why our findings with respect to the learning rate were in the opposite direction to our predictions, as the reward feedback used in the current paradigm was positively valenced, with participants making

choices to accumulate rewards rather than avoid punishment. Future research should compare how participants with ACEs weight feedback in response to both rewards and punishments. Based on our novel findings, we predict that in adulthood, ACEs will lead to overweighting feedback to avoid punishment and underweighting feedback to accumulate rewards compared to individuals without these experiences. This asymmetry in learning from reward and punishment could yield important insights into how childhood trauma is associated with the development and maintenance of psychopathology across the lifespan (Novick et al., 2018).

Our study has several limitations that are important to consider. We did not control for rates of stress, which mediate the association between ACEs and adult psychopathology (McLaughlin et al., 2010). State and trait stress have been associated with decreased exploration in a foraging paradigm (Lenow et al., 2017), which could explain some of the differences observed in foraging behaviour between our high and low ACE groups. Indeed, it could be interesting for future research to consider whether stress mediates the relationship between ACE exposure and foraging behaviour. Moreover, the ACE measure utilised in the present study includes a wide range of experiences, which may provide a less specific measure of participants' exposure to adverse events compared to previous operationalisations of early stress (Humphreys et al., 2015). For example, the ACE questionnaire does not ask about the frequency of each experience. A final limitation is that we manipulated both the travel time and depletion rate between foraging environments, meaning we cannot separately examine whether participants have greater sensitivity to changes in the depletion rate or to travel time, and whether these sensitivities differed by ACE-exposure. Future research could address this limitation by comparing environments with long and short travel times, while independently manipulating fast and slow depletion rates (e.g., Lloyd et al., 2021b). Administering environments more than once (e.g., Lenow et al., 2017) might further enhance the effect of environment quality on foraging behaviour that we observed in the current study.

In sum, this study has demonstrated that ACEs are associated with reduced exploration and with underweighting positive reward feedback in a patch foraging paradigm. These findings demonstrate the negative impacts on reward processing that are associated with adversity in childhood, further highlighting the need for children to be protected from these experiences. Our findings identify a novel computational component of decision-making that is impacted by ACEs: learning rate. This can provide direction for future work examining how reward-based decision-making is affected by ACEs and how this contributes to the heightened rates of psychopathology observed in this population.

## **7.5. Methods**

### **7.5.1. Participants**

To selectively recruit participants who had been exposed to ACEs, we advertised amongst four international charities and support groups for adult survivors of childhood trauma. These were: 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; [www.sona-systems.com](http://www.sona-systems.com)) hosted by a UK-based university and through Prolific ([www.prolific.co](http://www.prolific.co)). The Prolific sample was recruited from the same regions that the charities were based in (the UK and Europe). We recruited a total of 151 participants ( $M_{\text{age}} = 38.91$ ,  $SD_{\text{age}} = 11.09$ ), with a mean of 2.66 ACEs ( $SD = 2.76$ ). For group level analyses, we categorised 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., Bellis et al., 2015; Hughes et al., 2017). 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

$t(143) = 1.28, p = .202$ , nor did level of education  $t(141) = 0.01, p = .991$ . Because there were more females in the high ACE group compared to the control group  $\chi^2(1, N = 143) = 13.10, p < .001$ , 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).

### 7.5.2. Materials

*Adverse Childhood Experiences.* The Adverse Childhood Events Scale (Felitti et al., 1998) is a self-report measure of the number of ACEs 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 0-18), which is coded as a binary option (yes/no). Total scores on the measure range from 0–10, with higher values denoting the individual has been exposed to more ACEs (see Table 7.1).

Table 7.1: Prevalence of each Adverse Childhood Experience in the current sample.

<b>Adverse Childhood Experience</b>	<b>Percentage of the total sample</b>
<b><i>Threatening Events</i></b>	
Emotional Abuse	44.08%
Physical Abuse	30.26%
Sexual Abuse	36.25%
<b><i>Neglect</i></b>	
Emotional Neglect	44.08%
Physical Neglect	14.47%
<b><i>Family Adversity</i></b>	

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%

---

*Risk-taking:* The Domain-Specific Risk-Taking scale (DOSPERT; Blais & Weber, 2006) is a 30-item self-report measure assessing an individual's risk-taking propensity. The scale measures five domains of risk taking: financial, health/safety, recreational, ethical and social risks. Each domain is measured by six items ranging from 1-7, where 1 denotes that the individual would be highly unlikely to engage in that behaviour and 7 denotes that the individual is highly likely to engage in that behaviour. In the current sample, mean scores for each domain of risk-taking ranged from 1-6. 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 (Figure 7.4) in which participants harvested apples (rewards) from trees (patches; Constantino & Daw, 2015). 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 three seconds, regardless of participants' reaction times. The three seconds included participants' reaction time and the presentation of rewards on that trial. This design feature was implemented to ensure that faster reaction times did not impact how quickly participants could accumulate rewards. For example, a participant who responded in 0.6 seconds would be presented with

their score for 2.4 seconds. Participants had up to two seconds to make a decision before they were presented with a timeout screen. Timeout trials were subsequently excluded from further analyses.

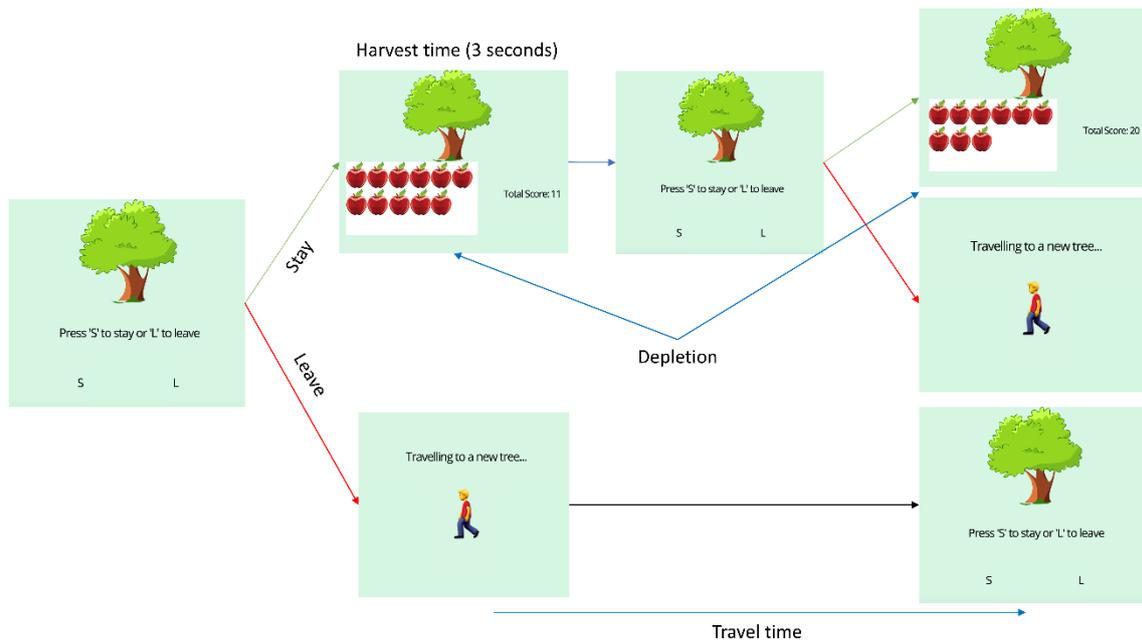


Figure 7.4: Outline of the patch foraging paradigm. The leftmost screen is presented to participants when they first enter the environment. They decide between staying or leaving the current patch. If participants choose to stay (top screen), they are presented with the number of apples they have harvested 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.

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 maximise 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 (Constantino & Daw, 2015; Lloyd et al., 2021b). The depletion rate was applied using the following formula:  $s_{i+1} = s_i * N_i$ , 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 six seconds 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 second 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 seven minutes in a counterbalanced order.

We utilised two behavioural variables from this task, in addition to the learning rate parameter derived from our computational model (described below). The first was the participants' leaving thresholds for each environment. As in previous research (Constantino & Daw, 2015), 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 (Constantino & Daw, 2015). 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 (Lenow et al., 2017; Lloyd et al., 2021b), we determined

the optimal leave threshold by running a grid search across leaving thresholds between 1-10 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 (see green horizontal lines in Figure 7.2).

### **7.5.3. 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 behavioural study platform (Anwyl-Irvine et al., 2019). Participants completed a consent form which informed them that they would be asked to complete a computerised 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 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 where they filled out demographic information. Following this, participants were provided with the task instructions and completed a two-minute 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 seconds. 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.

#### 7.5.4. Computational Modelling

According to a prominent theory of foraging behaviour (MVT; Charnov, 1976), the rational agent aiming to maximise 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 (2015). 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 (Sutton & Barto, 2018). This parameter captures the degree to which participants weight recent feedback to guide their decision-making:

$$p_i = (1 - \alpha)^{T_i} \frac{S_i}{T_i} + (1 - (1 - \alpha)^{T_i}) p_{i-1} \quad (1)$$

Although traditional reinforcement learning models directly equate the alpha parameter with the learning rate, it can be seen from Equation (1) that the model developed by Constantino and Daw (2015) parameterises 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:

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

This SoftMax function 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 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 (Constantino & Daw, 2015), 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 three seconds in both environments. As such, the term  $\kappa_i s_i - p_i h$  captures the difference between the reward participants expect from the next exploit decision and their current estimate of the average reward rate (Constantino & Daw, 2015). Parameter recovery indicated all parameters could be identified uniquely without parameter correlations, though we found the parameter  $c$  was less well recovered compared to the other parameters in the model (see Supplementary Materials). We compared this model, which uses only a single learning rate for all outcomes, to a model where the learning rate was split for better-than-expected and poorer-

than-expected outcomes (Garrett & Daw, 2020). Details about this additional model can be found in the Supplementary Materials.

## 7.6. Supplementary Materials

### 7.6.1. Instructions to participants (note, bold text was also highlighted when instructions were provided to participants)

*Instructions presented prior to the practice task*

Please read these instructions carefully before you begin.

Your aim is to collect as many apples as possible within the **time limit**. The more apples you collect, the larger your score at the end of this experiment and the bigger your prize will be.

- You can either **stay** to continue picking apples from the current tree or **leave** and find a new tree. If you leave and travel to a new tree, you have to wait a fixed amount of time. This time is fixed and has nothing to do with your internet connection or page loading.
- You will only need your keyboard for this task. You can either press 'S' to stay with the tree or 'L' to leave the tree and find a new one.
- You will not know how many apples are on a new tree until you stay and pick them, so it is a good idea to stay with each tree at least once before moving on.
- The number of apples left on a tree will decrease with time, meaning there will be fewer apples left on the tree to collect the longer you stay there. Apples do not grow back on each tree, so your job is to decide how long you want to spend at each tree.

**After seven minutes you will move into a completely new environment (think of it as a new orchard). This environment may be richer or poorer than the others. In some environments it may be better to stay with a tree for longer and in others it may be better to stay with a tree for less time.**

You will now begin a quick practice run of the study. Your practice score will not count towards your score in the main task. The task should take 15 minutes altogether (including the practice).

*Second set of instructions presented after practice*

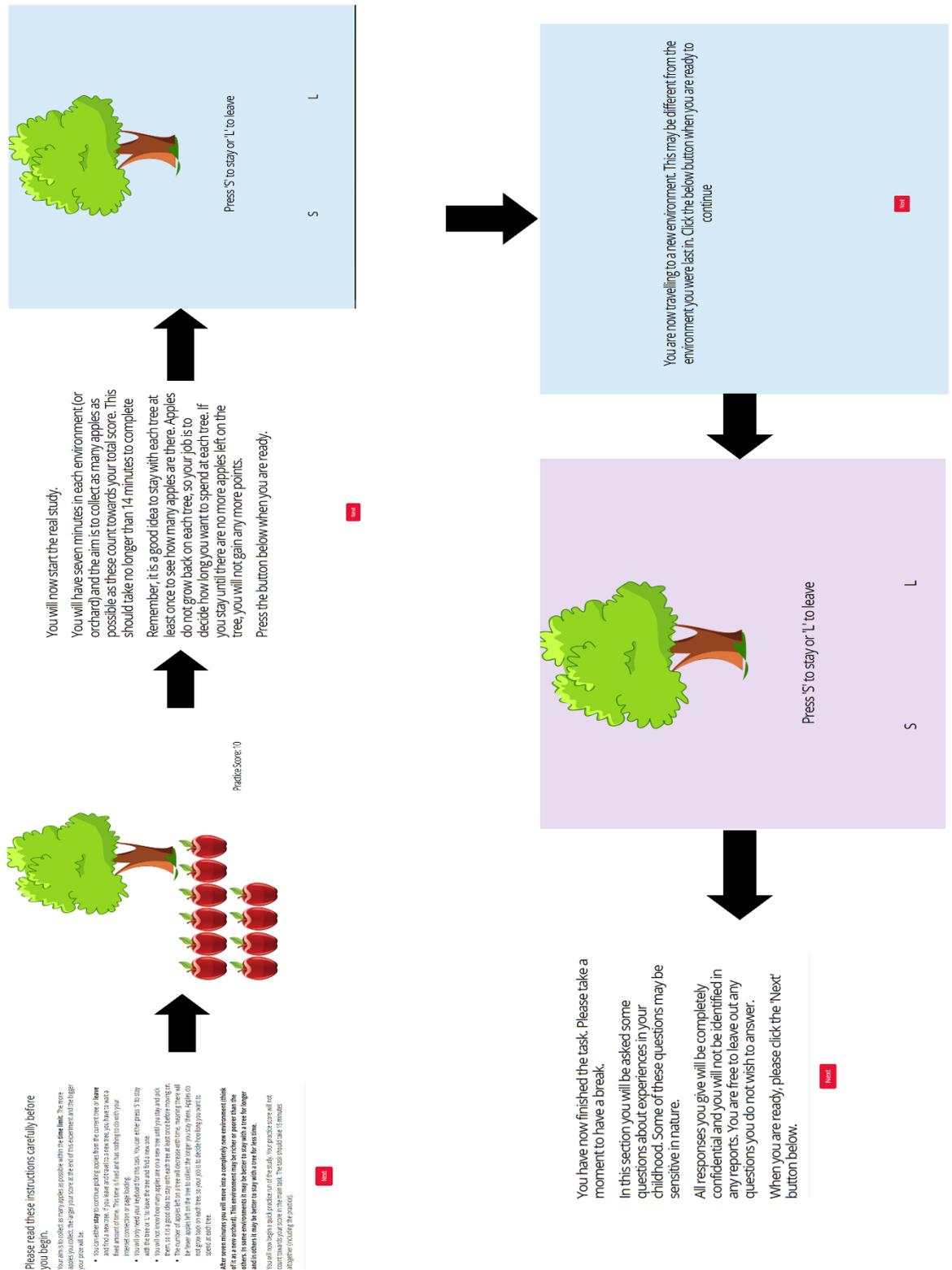
You will now start the real study.

You will have seven minutes in each environment (or orchard) and the aim is to collect as many apples as possible as these count towards your total score. This should take no longer than 14 minutes to complete

Remember, it is a good idea to stay with each tree at least once to see how many apples are there. Apples do not grow back on each tree, so your job is to decide how long you want to spend at each tree. If you stay until there are no more apples left on the tree, you will not gain any more points.

Press the button below when you are ready.

## 7.6.2. Full Outline of the Foraging Task



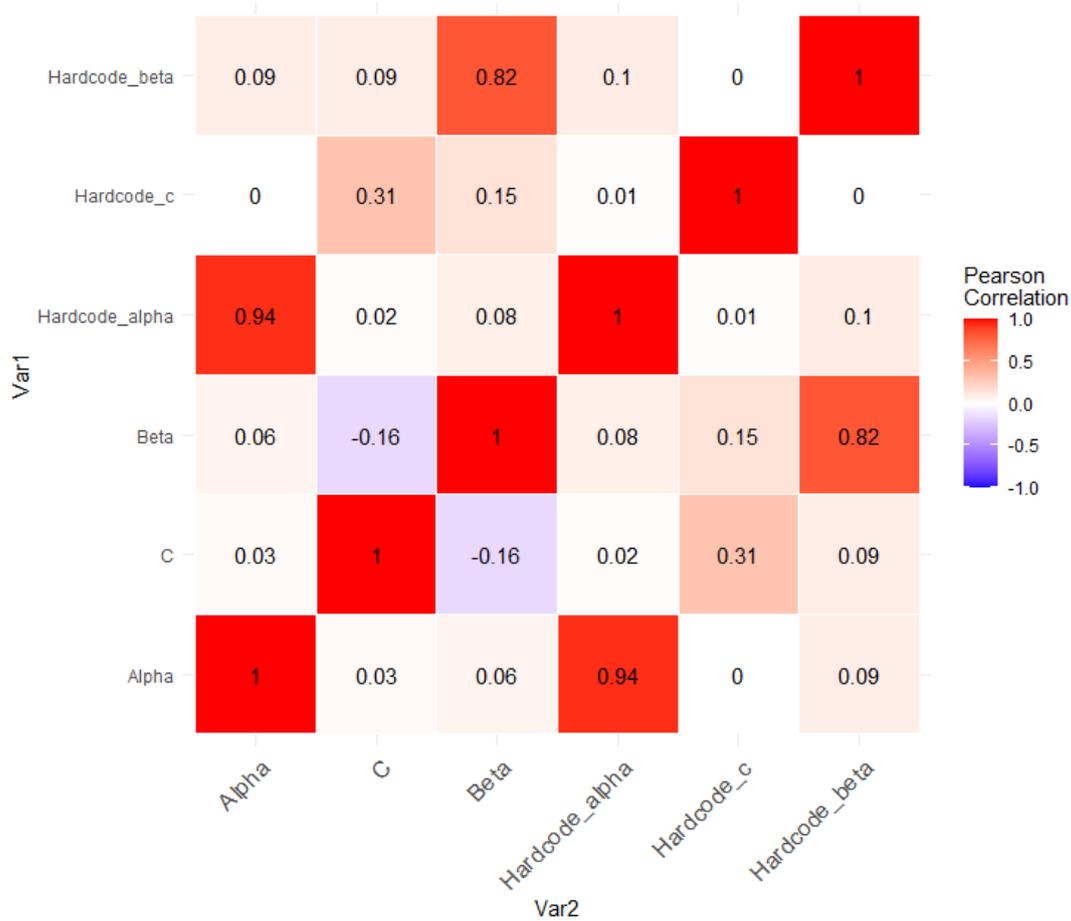
Supplementary Figure 7.1: Order with which participants completed the foraging task. The screens detail (in order): the initial instructions, the practice environment, final instructions, first environment (note, the order with

which the colours were presented was counterbalanced), interim screen signalling the change in environment, the second environment and finally a break screen notifying participants about the upcoming ACE questionnaire.

### **7.6.3. Computational Modelling**

#### *Parameter Recovery*

To ensure the model could estimate the correct value of the three free parameters, we simulated data for the rich and poor environments used in the behavioural study. For each free parameter, we simulated data across a range of values ( $\alpha$  {0,1} in increments of 0.1;  $c$  {-1, 3} in increments of 0.2 and  $\beta$  {1, 5} in increments of 0.4) and ran the computational model over these data. For our simulated and behavioural data, we estimated the parameters for each individual using the `optim` function (Nash, 2014) for R v.3.5.1. This function uses Maximum Likelihood Estimation to optimise the negative log likelihood of participants' stay/leave decisions produced by the SoftMax function (Equation 2). We found that the parameters were recoverable, meaning that there was a strong correlation between the parameters hard coded into the simulation and those estimated by the model (alpha:  $r(999) = .94, p < .001, 95\% \text{ CI } [0.93, 0.94]$ , beta:  $r(999) = .82, p < .001, 95\% \text{ CI } [0.80, 0.84]$ ,  $c: r(999) = .31, p < .001, 95\% \text{ CI } [0.25, 0.36]$ ). This suggests that the model was able to estimate the correct value of the free parameters, though we found that  $c$  was less well recovered. Notably, we did not observe correlations between the parameters (Daw, 2011). For correlations between the simulated and recovered parameters, see Supplementary Figure 7.2.



Supplementary Figure 7.2: Heatmap demonstrating the correlations between simulated (i.e., variables with the ‘hardcode’ prefix) and recovered parameters. Figures in each grid square indicate the Pearson’s  $r$  correlation coefficient.

#### 7.6.4. Model Comparison Analysis

We compared two models based on reinforcement learning to examine which of these presented a better explanation to participants’ behaviour. The first model had a single learning rate parameter, which measured how much participants weighted reward feedback during their decision-making. The formalisation for this model is detailed in the main manuscript (see Equations 1 and 2). We compared this single learning rate model to a model which had separate learning rates for positive ( $\alpha_+$ ) and negative ( $\alpha_-$ ) prediction errors. In this model, we calculate whether reward feedback on each trial is better-than-expected (i.e., a positive prediction error) or poorer-than-expected (i.e., a negative prediction error), which is expressed as:

$$\delta_i \leftarrow s_i/\tau_i - \rho_i$$

When  $\delta > 0$ , participants' expectation of the average reward rate ( $\rho_i$ ), as formulated in Equation 1, is updated according to the learning rate for positive prediction errors ( $\alpha_+$ ). When  $\delta < 0$ , participants' estimate of the average reward rate, as formulated in Equation 1, is updated according to the learning rate for negative prediction errors ( $\alpha_-$ ). Participants' estimate of the average reward rate is then entered into the SoftMax equation detailed in Equation 2 in the main manuscript.

We compared the models using two indices of model fit: the Bayesian Information Criterion (BIC) and Akaike's Information Criterion (AIC), where lower values indicate the model provides a better explanation of participants' data (Daw, 2011). Comparison of the AIC and BIC for both models demonstrated that the single learning rate model had lower AIC and BIC values than the dual learning rate model in both the rich and poor foraging environments (see Supplementary Table 7.1), which was true for participants regardless of ACE exposure (see Supplementary Figure 7.3). These results suggest that a model with a single learning rate for both positive and negative prediction errors was a better fit to the data across both environments. As such, analyses in the main text were conducted using this model.

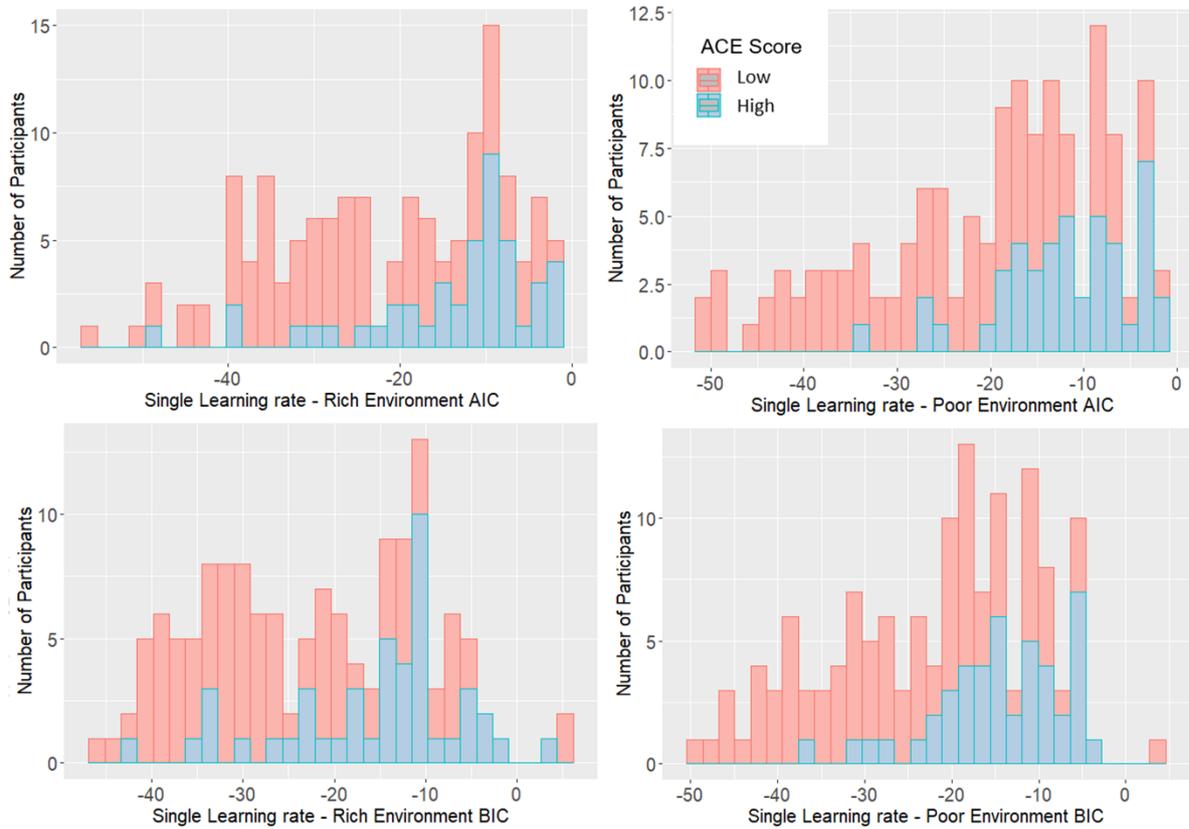
*Supplementary Table 7.1: Average of the indices of model fits from participants' data for the single and dual learning rate models, in each environment.*

		Rich Environment		Poor Environment	
Model		AIC	BIC	AIC	BIC
Dual	learning rate model	-26.35	-14.61	-27.82	-17.97

Single learning rate model

-47.47                      -39.49                      -48.39                      -40.86

---

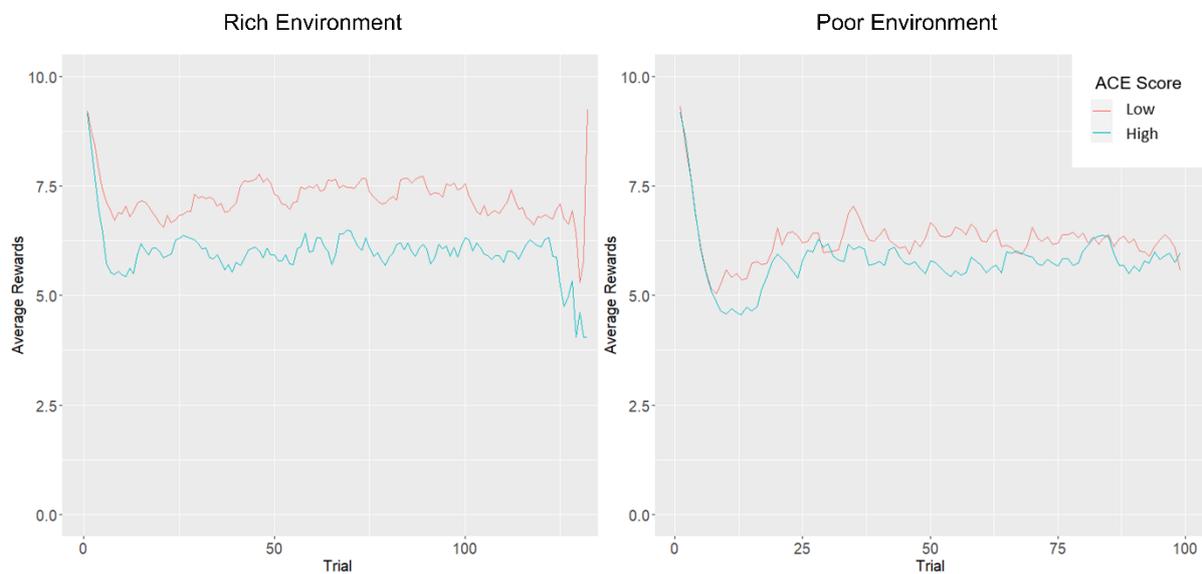


Supplementary Figure 7.3: Plots demonstrating the relative fit of the single learning rate model compared to the dual learning rate model. Values below zero indicate that the single learning rate model was a better fit to participants' data whereas values above zero indicate the dual learning rate model was a better fit to the data.

### 7.6.5. Post-Hoc Simulation

We used post-hoc simulation to demonstrate that we could recapitulate the main effect of ACE exposure reported in the Results section on participants' leaving threshold. We conducted simulations using the mean parameter from the single learning rate model for each ACE group in each foraging environment. A total of 200 agents were simulated, 100 using the mean parameter estimates from the high ACE group and 100 using the mean parameter estimates from the low ACE group. The plots in Supplementary Figure 7.4 descriptively

indicate that our model was able to replicate the effect of ACE exposure on participants' leaving threshold (plotted in Figure 7.1 in the main manuscript), as we found the simulated 'high ACE' group to utilise a lower leaving threshold than the simulated 'low ACE group'. This group difference was confirmed statistically ( $F_{(1,198)} = 297.86, p < .001$ ). We also found that simulated agents utilised a higher leaving threshold in the rich environment relative to the poor environment ( $F_{(1,198)} = 52.57, p < .001$ ), consistent with optimal foraging theory and participants' leaving threshold as reported in the Results section.



*Supplementary Figure 7.4: Plots demonstrating the average reward experienced on each trial for the simulated foraging agents in the rich environment (left) and poor environment (right). Local troughs indicate the simulated agents' leaving thresholds.*

### 7.6.6. Analyses Controlling for Gender

When controlling for gender in an ANCOVA of participants' leaving threshold, we preserved all the effects reported in the Results of the main manuscript, including the two main effects and null interaction effect. We found a main effect of ACE exposure ( $F_{(1,134)} = 4.84, p = .030, \eta^2 = .03$ ), and a main effect of environment type ( $F_{(1,134)} = 20.96, p < .001, \eta^2 = .002$ ).

In addition, there was no interaction between ACE exposure and environment ( $F_{(1,134)} = 0.57$ ,  $p = .452$ ,  $\eta^2 < .001$ ).

When including gender as a covariate in an analysis of participants' learning rate, we also preserved the main effects and null interaction effect observed in the main manuscript. We found a main effect of ACE exposure ( $F_{(1,134)} = 9.26$ ,  $p = .003$ ,  $\eta^2 = .05$ ) and environment type ( $F_{(1,134)} = 5.13$ ,  $p = .025$ ,  $\eta^2 = .006$ ). Moreover, we did not observe an interaction between ACE exposure and environment ( $F_{(1,134)} = 0.36$ ,  $p = .546$ ,  $\eta^2 < .001$ ).

### 7.6.7. Regression Model Statistics

When each of the DOSEPERT subscales were entered into separate regression models, neither ACE score nor age predicted subscales of the DOSPERT. With the exception of the social risk subscale, we found that being male predicted the subscales of the DOSPERT (see Supplementary Table 7.2 for full model statistics).

*Supplementary Table 5.2: Table demonstrating regression model statistics for the subscales of the DOSPERT. Bold cells indicate results that were statistically significant at the .050 level.*

<i>Subscale</i>	<i>ACE Score</i>		<i>Age</i>		<i>Gender</i>		<i>Model Statistics</i>		
	$\beta$	$p$	$\beta$	$p$	$\beta$	$p$	$F$	$R^2$	$p$
<i>Social</i>	-0.09	.299	-0.12	.158	0.04	.631	1.17	0.03	.325
<i>Ethical</i>	0.10	.247	-0.06	.515	<b>0.19</b>	<b>.033</b>	1.73	0.04	.163
<i>Financial</i>	-0.13	.120	-0.11	.183	<b>0.25</b>	<b>.004</b>	<b>5.84</b>	<b>.001</b>	<b>0.11</b>
<i>Health and Safety</i>	0.15	.093	-0.10	.251	<b>0.23</b>	<b>.009</b>	<b>2.98</b>	<b>.033</b>	<b>0.06</b>
<i>Recreational</i>	-0.05	.588	-0.09	.272	<b>0.18</b>	<b>.044</b>	2.30	0.05	.080

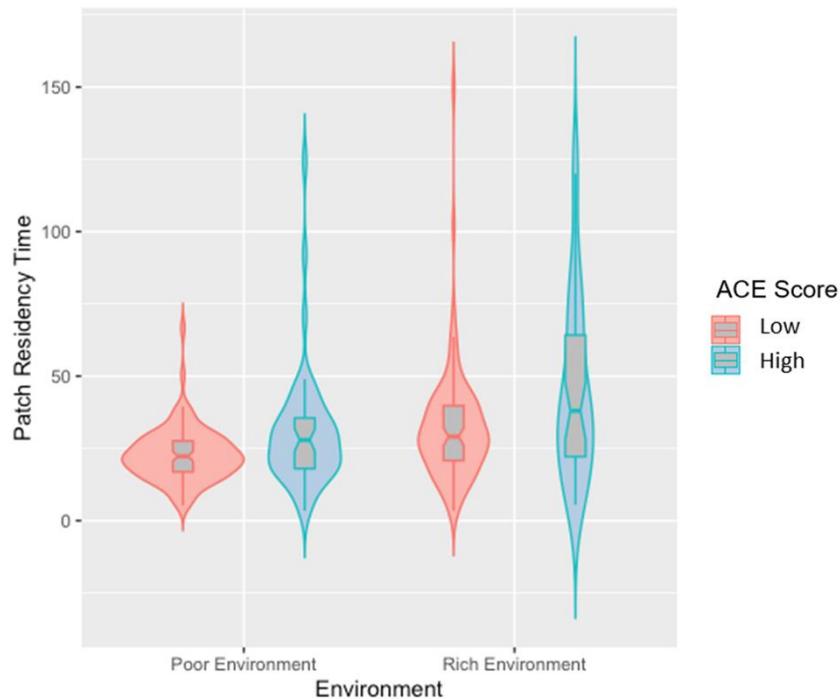
### 7.6.8. Analyses of ACE Score as a Continuous Measure

We found similar patterns in our data when nonparametric correlations were run using ACE scale as a continuous measure. Specifically, we found that there were significant negative associations between ACE scores and learning rate in the rich environment  $r(144) = -0.15, p = .033$  and poor environment  $r(144) = -0.20, p = .007$ . With regard to leaving thresholds, we found a significant negative association between ACE score and leaving threshold in the poor quality environment  $r(144) = -0.15, p = .033$ . However, the relationship between ACE score and leaving threshold in the rich environment was not significant  $r(144) = -0.09, p = .144$ .

### 7.6.9. Analyses of Participants' Patch Residency Times

One possibility is that the main effects of ACE group on participants' leaving threshold reported in the Results section occurred because participants in the high ACE group remained for the same or similar amount of time in patches in both environments, producing an illusory effect that they were adjusting their leaving thresholds between these two environments. To exclude this interpretation, we conducted analyses on participants' patch residency time measured in seconds (i.e., how long they remained in patches before exploring). Analysis of this variable replicated the findings we observed with participants' leaving thresholds. Specifically, participants adjusted their patch residency time as a function of the environment ( $F_{(1,136)} = 55.24, p < .001, \eta^2 = 0.07$ ), with participants remaining for longer in the richer environment relative to the poorer environment, which is consistent with optimal foraging theory (Charnov, 1976). We also found a main effect of ACE exposure ( $F_{(1,136)} = 11.78, p < .001, \eta^2 = 0.06$ ; see Supplementary Figure 7.5), consistent with our analyses of participants' leaving thresholds (see Figure 7.1 in the main manuscript). However, we did not find an interaction between environment and ACE exposure ( $F_{(1,136)} = 1.31, p = .255, \eta^2 = 0.002$ ). Together, these findings suggest that participants adjusted the time they remained with patches

in a similar manner to their leaving thresholds between the two foraging environments. Moreover, ACE exposed individuals stayed with patches for longer than individuals without these experiences.



*Supplementary Figure 7.5: Plot demonstrating the average patch residency time (y-axis) in the two foraging environments (x-axis). These results are comparable to our analyses of participants' leaving thresholds detailed in Figure 7.1 in the main manuscript.*

#### **7.6.10 Analyses Controlling for Counterbalance Order Effects**

We conducted additional analyses which suggested that our primary findings were not affected by order effects related to counterbalancing of the environments, as the effects reported in the Results section were preserved after accounting for the order with which participants experienced the two foraging environments. In these analyses we included the order with which participants experienced the rich and poor environments as an additional factor in the analyses of participants' leaving threshold and learning rates.

Examining participants' leaving thresholds, we did find a significant order\*environment interaction ( $F_{(1,136)} = 5.31, p = .023$ ). Despite this interaction, our results

were consistent with those reported in the main manuscript as we still find a main effect of ACE exposure ( $F_{(1,135)} = 4.39, p = .038, \eta^2 = 0.03$ ) and environment ( $F_{(1,135)} = 28.65, p < .001, \eta^2 = 0.03$ ), with the effect sizes for these differences being exactly the same as reported in our original analyses. We did not find a significant order\*environment\*ACE group interaction ( $F_{(1,135)} = 0.44, p = .510$ ), indicating these order effects did not account for the different leaving thresholds we observed between groups. As we had a priori predictions that leaving thresholds would be higher in the rich environment relative to the poor environment, we ran a one-tailed planned comparison between leaving thresholds in the rich and poor environments, accounting for counterbalance order. Results of this planned comparison demonstrated that leaving thresholds were significantly higher in the rich environment relative to the poor environment in the counterbalance condition where the rich environment was presented first ( $t(135) = 2.08, p = .039$ ) and in the counterbalance condition where the poor environment was presented first ( $t(135) = 5.51, p < .001$ ). Examining participants' learning rates, an ANOVA demonstrated that there was no interaction between counterbalance order\*environment ( $F_{(1,142)} = 0.01, p = .915$ ), though consistent with the results in the main manuscript we find a main effect of ACE exposure ( $F_{(1,142)} = 8.65, p = .004$ ) and environment ( $F_{(1,142)} = 0.01, p = .915$ ). These analyses demonstrate that the results reported in the original manuscript are not affected by order effects.

#### **7.6.11. Analyses of Participants' Reaction Times**

Analysing participants' reaction times, we found that participants in the high ACE group had slower reaction times ( $M_{\text{rich}} = 0.53, SD_{\text{rich}} = 0.24, M_{\text{poor}} = 0.51, SD_{\text{poor}} = 0.23$ ) compared to participants in the low ACE group ( $M_{\text{rich}} = 0.43, SD_{\text{rich}} = 0.17, M_{\text{poor}} = 0.43, SD_{\text{poor}} = 0.16; F_{(1,137)} = 8.53, p = .004, \eta^2 = 0.05$ ). However, we found no main effect of environment on RT ( $F_{(1,137)} = 0.39, p = .535$ ) nor an interaction between ACE group and environment ( $F_{(1,137)} = 0.84, p = .361$ ).

### **7.6.12. Exploration Behaviour in the High ACE Group Only**

To exclude the possibility that participants in the high ACE group simply did not change their leaving thresholds much between the two environments, we ran independent *t*-tests on the leaving threshold and patch residency time in the high ACE group. These analyses demonstrated that participants in the high ACE group did modulate their leaving thresholds as a function of the environment, as these participants explored more in the rich environment ( $M = 5.00$ ,  $SD = 2.46$ ) than the poor environment ( $M = 4.26$ ,  $SD = 2.16$ ;  $t(44) = 3.35$ ,  $p = .002$ , Cohen's  $d = 0.50$ ). In addition, we also found that the high ACE group had significantly longer patch residency times in the rich environment relative to the poorer environment ( $t(44) = 4.59$ ,  $p < .001$ , Cohen's  $d = 0.68$ ), which is consistent with optimal foraging theory (Gabay & Apps, 2021).

### **7.6.13. Analyses by Type of Adversity**

It has been suggested by some researchers (e.g., McLaughlin & Sheridan, 2016) that behavioural differences following adversity are driven by the type of experience. These experiences can be separated into the following categories: threatening events (including physical and sexual abuse), neglect, and family adversity (e.g., parental imprisonment; see Table 7.1 in the main manuscript; McLaughlin et al., 2014; McLaughlin & Sheridan, 2016; Sheridan & McLaughlin, 2014). As some participants experienced more than one type of adversity, we constructed three factors: One that dummy coded for the presence versus absence of threatening events, one that coded for the presence versus absence of neglect, and one that coded for the presence versus absence of family adversity. For each of our dependent variables (leaving thresholds and learning rate), we ran separate ANOVAs for each of the three adversity factors to examine whether behavioural differences were driven by a particular type of adverse experience. We also included environment type (rich or poor) as a second factor in these ANOVA analyses.

In our analysis of participants' leaving thresholds, we did not find evidence that a particular category of adversity explains the differences we observe in participants' leaving thresholds. We ran three ANOVAs which each had only one of the types of adversity (threatening events, neglect, or family adversity) entered as the independent variable. We applied the Bonferroni correction to these three exploratory analyses. When experience of threatening events was entered as the independent variable, we did not find a main effect of threatening events on participants' leaving thresholds ( $F_{(1,137)} = 2.17, p_{bonf} = .429$ ), nor an interaction between threatening events and environment ( $F_{(1,137)} = 0.03, p_{bonf} = 1$ ) though in this model we did find a main effect of environment, consistent with the analyses presented in the main manuscript ( $F_{(1,137)} = 28.83, p_{bonf} = .003$ ). When experience of neglect was entered as the independent variable, we did not find a main effect of neglect ( $F_{(1,137)} = 0.21, p_{bonf} = 1$ ) nor an interaction between environment and neglect ( $F_{(1,137)} = 2.61, p_{bonf} = .327$ ), though we did find a main effect of environment, consistent with our original analyses ( $F_{(1,137)} = 31.38, p_{bonf} = .003$ ). Finally, when experience of family adversity was entered as the independent variable, we did not find a main effect of family adversity ( $F_{(1,137)} = 3.73, p_{bonf} = .168$ ), nor an interaction between family adversity and environment ( $F_{(1,137)} = 1.82, p_{bonf} = .540$ ), though we did find a main effect of environment ( $F_{(1,137)} = 28.15, p_{bonf} = .003$ ).

Although there were no effects of adversity type for leaving threshold, we found that group level differences in learning rate were driven by experiences of threatening events. We ran three ANOVAs examining the effect of each type of adversity on participants' learning rate. We Bonferroni-corrected for these three exploratory analyses. When experience of threatening events was entered as an independent variable, we found a main effect of threatening events on participants learning rate ( $F_{(1,137)} = 7.02, p_{bonf} = .027, \eta^2 = 0.05$ ) with participants exposed to these events having a lower learning rate ( $M_{rich} = 0.53, SD_{rich} = 0.24, M_{poor} = 0.59, SD_{poor} = 0.17$ ) compared to participants who were not exposed to these

experiences ( $M_{\text{rich}} = 0.62$ ,  $SD_{\text{rich}} = 0.14$ ,  $M_{\text{poor}} = 0.64$ ,  $SD_{\text{poor}} = 0.17$ ). We also found a main effect of environment ( $F_{(1,137)} = 9.31$ ,  $p_{\text{bonf}} = .009$ ,  $\eta^2 = 0.01$ ), though we did not find an interaction between environment and threatening events ( $F_{(1,137)} = 2.27$ ,  $p_{\text{bonf}} = .405$ ). When experience of neglect was entered as an independent variable, we did not find a main effect of neglect ( $F_{(1,137)} = 1.53$ ,  $p_{\text{bonf}} = .654$ ) nor an interaction between environment and neglect ( $F_{(1,137)} = .327$ ), though we did find a main effect of environment ( $F_{(1,137)} = 31.38$ ,  $p_{\text{bonf}} = .003$ ). Finally, when experience of family adversity was entered as an independent variable, we did not find a main effect of family adversity ( $F_{(1,137)} = 4.29$ ,  $p_{\text{bonf}} = .120$ ). We found a main effect of environment ( $F_{(1,137)} = 9.38$ ,  $p = .003$ ,  $p_{\text{bonf}} = .009$ ), though we did not find an interaction between environment and family adversity ( $F_{(1,137)} = 0.03$ ,  $p_{\text{bonf}} = 1$ ). These findings may suggest that differences in learning rate were driven by threatening experiences.

## Chapter 8: General Discussion

### 8.1. Summary of Findings

The present thesis investigated adolescents' and adults' patch foraging behaviour to address three objectives, all of which were to test the predictions of theories that explain decision-making across the lifespan. The first objective was to test the predictions of theories that account for the rise in novelty-seeking and risky behaviours in adolescence, which I addressed in Chapters 3, 4 and 5. In Chapter 3, I found that adolescents explored more than adults and that this was associated with more optimal foraging behaviour, which I replicated in Chapter 5 but failed to replicate in Chapter 6. In Chapter 4, I conducted a computational analysis of data from Chapter 3, demonstrating that adolescents' heightened exploration was associated with a higher learning rate relative to adults. I next examined how adolescents adapted their behaviour between conditions of stability and volatility in Chapter 5 demonstrating that adolescents exhibited greater stochasticity in their behaviour compared to adults, though both age groups were able to adjust this feature of their decision making between these environments. These findings are consistent with theories of adolescent development that propose novelty-seeking is heightened during this developmental period as it allows adolescents to trial experiences which can support the development of experiential knowledge, supporting them in the transition to adulthood (Romer et al., 2017).

The second objective was to test the predictions of theories that explain how social influence affects adolescents' decision-making, which I addressed in Chapter 6. Using a social manipulation of the foraging task, I present evidence that adolescents and adults were susceptible to social influence while foraging. Results of my computational model indicated that compared to adults, younger adolescents had higher values on a parameter measuring susceptibility to social feedback. These findings are consistent with theories of social influence

in adolescence that suggest adolescents are motivated by the views of their peers (Ciranka & van den Bos, 2019), rather than being hypersensitive to reward in the presence of same-aged peers (Chein et al., 2011).

The third and final objective was to test the predictions of a theory that purports to explain how experiences of adversity during childhood and adolescence can impact decision-making in adulthood, known as Life History Theory (Ellis et al., 2012). I found that individuals who were exposed to ACEs explored less than individuals without these experiences. Moreover, in contrast to my predictions, ACE-exposed individuals exhibited lower learning rates for reward feedback, suggesting these individuals underweighted recent information compared to individuals without these experiences. That individuals with ACEs explored less than individuals without these experiences is consistent with Life History Theory, though my finding that ACE-exposed individuals underweight reward feedback may not be consistent with the predictions of Life History Theory. Nevertheless, these findings contribute to our understanding of how early adversity affects features of decision-making and the deleterious outcomes associated with childhood trauma.

Together, these studies provide empirical evidence that highlights the importance of adolescence within lifespan development. I show that adolescence is a period of heightened exploration relative to adulthood and that heightened exploration can support adolescents in learning about their surroundings. I also present a computational account of the rise in novelty-seeking in adolescence and suggest these exploratory behaviours are related to mechanistic features of decision-making, such as the learning rate and the level of stochasticity in adolescents' choices. In examining the role of social influence, I demonstrate how this can lead to both positive, optimal outcomes for adolescents, as well as the potential for social influence to promote suboptimal outcomes. Finally, I highlight the implications of adverse experiences

during adolescence for longer term outcomes on decision-making, demonstrating how trauma during development is associated with a poorer ability to navigate one's environment.

## **8.2. Adolescents Explore Their Environment More Than Adults**

Across three empirical chapters (Chapters 3, 4 and 5) I have examined whether the rise in novelty-seeking observed during adolescence can be better explained by imbalance models or the Lifespan Wisdom Model. Imbalance models propose that adolescents are more prone to risky behaviour due to the rapid development of the dopaminergic reward system at the onset of puberty and the relative immaturity of neural regions responsible for complex cognitive control (Casey et al., 2008; Steinberg, 2005). In contrast, the Lifespan Wisdom Model suggests that the rapid development of the dopaminergic reward system promotes novelty-seeking behaviours, which in turn promotes exploration of one's surroundings (Romer et al., 2017). The experience gained through trialling one's surroundings can then be used to guide future decision-making in adulthood (Romer et al., 2017).

The patch foraging paradigm I utilised was able to arbitrate between these models and findings using this task favour the Lifespan Wisdom Model. Imbalance models propose that the hyperactivity of the reward system and relative immaturity of cognitive control capabilities mean that adolescents are likely to seek rewards without considering the potential for negative outcomes that could result from their behaviour (Steinberg, 2008). As such, I would expect according to imbalance models that adolescents would have been motivated to seek immediate rewards gained through exploiting patches, forgoing delayed rewards that can be gained through exploring fresh patches, whereas adults would be use their mature cognitive control abilities to explore more often. In contrast, the Lifespan Wisdom Model predicts that adolescents are more motivated to explore compared to adults, as this supports them to learn

the structure of their surroundings. I demonstrated empirically that adolescents exhibited heightened rates of exploration compared to adults, supporting predictions of the Lifespan Wisdom Model. In Chapters 3 and 5, I found that adolescents explored more than adults (though noting I failed to replicate this finding in Chapter 6). That adolescents exhibit heightened rates of exploration suggests they were able to exercise cognitive control to explore new patches with potential for longer term rewards, despite this option incurring a time cost associated with travelling, during which rewards could not be collected (Otto et al., 2015). These findings are consistent with evidence that adolescents explore more in behavioural tasks such as the Iowa Gambling Task (e.g., Almy et al., 2018; Hooper et al., 2004; Smith et al., 2012). However, adolescents' heightened exploration in the Iowa Gambling Task relative to adults has typically been associated with adolescents accruing fewer points in the Iowa Gambling Task and therefore demonstrating less optimal behaviour relative to adults. In contrast, heightened exploration in my patch foraging paradigm allowed adolescents to collect more rewards compared to adults. These findings contribute to our understanding of the conditions under which adolescents' heightened novelty seeking, relative to their older counterparts, may confer benefits to this age group.

My findings may appear to be inconsistent with some findings from previous literature. For example, the finding that adolescents explore more than adults while foraging is in contrast to one study that did not find age related differences in a risky foraging task from adolescence to early adulthood (Bach et al., 2020). However, I note that this task did not measure *patch* foraging and the behaviour of interest in this study was how participants were able to avoid a virtual predator to collect rewards, whereas I was interested in adolescents' rates of exploration relative to adults. Indeed, the absence of an age-related difference in the risky foraging task utilised by Bach and colleagues (2020), combined with my findings of heightened exploration in adolescence relative to adulthood, could lend further support to the argument that

adolescents' behaviour is better characterised by exploration rather than risk taking (Romer et al., 2017).

My findings are also inconsistent with research that has demonstrated that, relative to adults, adolescents prefer the immediate rather than delayed option in delay discounting paradigms (e.g., de Water et al., 2014; van den Bos et al., 2015). One possible reason for this inconsistency is that the delays presented in discounting tasks refer to periods of up to months, whereas exploration only incurred a delay of seconds in the current foraging paradigm (Odum, 2011). It may be that, relative to adults, adolescents exhibit a preference for immediate rewards when the delays associated with a larger reward exceed more than a few seconds, though this would need to be tested empirically in future research. Yet, another possible explanation is that the foraging paradigm required participants to estimate the future reward values associated with exploration and exploitation, whereas delay discounting paradigms explicitly present reward values to participants. This difference between estimated versus explicit rewards may parallel developmental differences observed between conditions of risk and ambiguity (e.g., van den Bos & Hertwig, 2017). In these studies, relative to adults, adolescents are more likely to select the ambiguous option, whereas adolescents are less likely to select the risk option compared to adults (Tymula et al., 2012). As such, adolescents may be more willing to accept the delayed reward when the available options are not explicit, as in my foraging tasks, compared to when they are explicit such as in delay discounting paradigms. Indeed, one recent study did not observe associations between foraging leaving thresholds and delay discounting preferences in an adult sample, suggesting these tasks measure distinguishable cognitive processes (Lloyd et al., 2021a).

Building on the finding that adolescents explore more than adults, in Chapter 4 I demonstrate that age-related differences in foraging behaviour were driven by differences in participants' learning rates, with adolescents having a faster learning rate relative to adults

(though I failed to replicate this finding in Chapter 6). While one study using a bandit task has found that adolescents have a higher learning rate compared to adults (Decker et al., 2015), a larger body of research using similar bandit paradigms have found that learning rates increase from adolescence to adulthood (Eckstein et al., 2020; Davidow et al., 2016; Xia et al., 2020). One possible reason for this inconsistency is that the degree to which the individual should utilise a low or high learning rate is determined by the structure of the task and estimates for this parameter do not necessarily generalise across paradigms (Eckstein et al., 2021). During the bandit task, the use of a lower learning rate allows the individual to approximate the probabilities associated with each bandit, should these remain stable during the course of the task. Moreover, when participants receive rewards, these are of the same magnitude (i.e., participants receive one point every time they are rewarded on the task). However, in my foraging task, the rewards gained through exploiting patches vary depending on the length of time participants remain in patches, as these rewards deplete on each successive exploit decision. As such, using a higher learning rate may lead participants to place greater emphasis on the relative bounty of rewards when arriving at a new patch compared to participants with a lower learning rate. Moreover, participants with a higher learning rate may also place more emphasis on the relatively fewer rewards earned from each successive exploit decision compared to participants with a lower learning rate. The changing reward feedback encountered in the foraging task may mean that higher learning rates bias individuals towards greater exploration. In comparison, in bandit tasks, the rewards received from each stimulus remain consistent, meaning that lower learning rates are associated with less exploration (Eckstein et al., 2021). Consistent with this interpretation, I found a higher learning rate was correlated with greater exploration in Chapters 4 and 7, and that adolescents utilised a higher learning rate relative to adults in Chapter 4.

The patch foraging paradigm and computational models I utilise throughout this thesis have several additional key differences to studies that have utilised bandit tasks and their associated RL models to measure changes to cognition across the lifespan. For example, RL models that explain performance on bandit tasks assume that participants track reward values for each available stimulus and these reward values are updated according to the feedback from sampling each of these stimuli (e.g., Davidow et al., 2016). In contrast, the learning rate parameter in the models I utilise contributes to participants' estimation of the average reward rate for the environment. Recent work has considered models where participants estimate reward values for individual patches (Hall-McMaster et al., 2021), which may be a closer approximation of the role that the learning rate has in other RL studies. As such, future research could compare whether adolescents' and adults' learning rate change according to whether the reward value being estimated is for within patch rewards or the average reward rate for the environment. Nevertheless, the findings of Chapter 4 have implications for our understanding of how adolescents and adults update their reward estimates of the relative opportunities afforded by their environment, which is important for a number of real world contexts, such as deciding whether to accept a job offer or continue searching for another opportunity.

The finding that adolescents utilise a higher learning rate relative to adults is consistent with the predictions of the Lifespan Wisdom Model. This theory predicts that adolescents' behaviour is driven by their heightened sensitivity to reward relative to older individuals, which can motivate adolescents to engage in exploration (Romer et al., 2017). The use of a higher learning rate in adolescence has also been associated with greater activation of the hippocampus during reward-learning compared to adults (Davidow et al., 2016). Taken in context of this previous finding, my finding that adolescents use a higher learning rate compared to adults may suggest that heightened reward sensitivity during adolescence supports the individual to better learn stimulus-outcome contingencies and use this information to guide

future decision-making. This interpretation would support predictions of the Lifespan Wisdom Model, which posits that adolescence is a period during which the individual gains experience that is used to guide future choices (Romer et al., 2017). Indeed, in the foraging tasks used in this thesis, adolescents performed more optimally than adults, which may have resulted from an improved ability to integrate feedback to learn stimulus-outcome contingencies, reflected computationally by their higher learning rate relative to adults.

Adolescents' faster learning rate compared to adults may be adaptive in scenarios requiring the individual to track the average reward rate in the environment. Utilising a high learning rate is beneficial in environments that are rapidly changing, as higher learning rates provide the individual with more accurate estimates of stimulus-outcome contingencies compared to a lower learning rate, where more historic feedback becomes rapidly outdated (Behrens et al., 2007). As adolescence is a stage where the individual undergoes significant periods of transition (for example, starting secondary school), utilising a higher learning rate may support adolescents to update their expectations of the affordances in their environment in an appropriate manner. Indeed, previous research has found that adolescents perform better than adults in a volatile reversal learning task (Eckstein et al., 2020), indicating they are cognitively equipped to navigate changeable environments. Together, these findings are consistent with the Lifespan Wisdom Model that adolescents utilise feedback to learn from their surroundings and the results of this thesis suggest that they place more emphasis on recent information compared to adults.

In Chapter 5 I replicated the findings from Chapters 3 and 4 that compared to adults, adolescents had a higher leaving threshold and higher learning rate (though these findings did not replicate in Chapter 6). In Chapter 5, I examined how adolescents and adults adjust their behaviour between stable and volatile foraging environments and found that this ability is driven by calibrating the parameter that measures stochasticity in decision-making. These

findings are consistent with theoretical accounts that explain optimal exploration in volatile conditions (Gershman, 2018; Wilson et al., 2021). Under conditions of volatility, occasionally diverting from the option with the perceived highest expected value can support the individual to identify changes to stimulus-outcome contingencies when the environment is changing (Gershman, 2018). In contrast, if the decision-maker only selects the option which they perceive to have the highest expected value, when the reward contingencies associated with this option change and become less favourable, the decision-maker will continue to select this option for longer despite other more rewarding stimuli being available. As such, stochasticity in decision-making can be adaptive under conditions of volatility (Wilson et al., 2021). Moreover, these findings are consistent with empirical work that has demonstrated adolescents consistently utilise a more stochastic decision-making strategy compared to adults (Eckstein et al., 2020; Decker et al., 2015; Palminteri et al., 2016; Xia et al., 2020). The results I present in Chapter 5 contribute to this literature by demonstrating how stochasticity is deployed strategically by adolescents and adults in conditions of volatility, and that this ability is impaired by anxiety.

A stochastic decision-making policy leads the individual to trial a greater range of actions (Gopnick, 2020) and occasionally divert from the option that to the decision-maker's knowledge has the highest expected value at that time (Gershman, 2018). As such, these findings may be considered more consistent with the Lifespan Wisdom Model compared to imbalance models. If adolescents' decision-making were motivated by maximising immediate reward, as predicted by imbalance models, using a stochastic decision-making strategy would occasionally divert them from the most rewarding option and lead them to forfeit these rewards. Instead, adolescents utilise a strategy of decision-making that increases the rate of exploratory decisions, relative to adults, which supports the predictions of the Lifespan Wisdom Model (Romer et al., 2017).

However, an alternative interpretation of the finding that adolescents exhibit more stochastic choice behaviour relative to adults is that this group lacks the cognitive capabilities to identify the option with the highest expected value while foraging, thereby demonstrating more variable behaviour. On this account, heightened stochasticity in adolescence relative to adulthood may reflect an immature decision-making strategy, potentially resulting from the protracted development of neural regions responsible for cognitive control during adolescence (e.g., the vIPFC and dlPFC; Hooper et al., 2004). This novel proposal of the association between immature cognitive control capacities and more stochastic behaviour would be consistent with imbalance models, which posit that adolescents' decision-making is driven by hyperactive responses to reward and underdeveloped cognitive control capacities (Steinberg 2008; Steinberg et al., 2018). Indeed, the increased variability in adolescents' decision-making that is associated with stochasticity may be a product of this age group's reduced working memory capacities compared to adults (Luciana et al., 2005). A poorer ability to retain information in working memory may prevent adolescents from identifying whether exploration or exploitation has the highest expected value. As such, adolescents may deviate from the option with the highest expected value at a greater frequency compared to adults due to inefficient decision-making strategies, rather than adaptive ones.

Furthermore, the stochasticity parameter in reinforcement learning models has also been associated with impulsivity, which may be consistent with imbalance models that predict adolescents' risk taking is driven by their preference for immediate reward without considering longer-term consequences. Several studies have identified a link between attention deficit hyperactivity disorder (ADHD), a disorder characterised by higher levels of impulsivity, and stochastic responding in multi-armed bandit tasks (Dubois et al., 2020; Hauser et al., 2016). As such, the finding that adolescents utilised more stochastic responding relative to adults could indicate that this group are more impulsive relative to adults, consistent with imbalance models

(Steinberg et al., 2008). However, in contrast to previous research (e.g., Dubois et al., 2020) I note that there was no association between self-reported impulsivity and the  $\beta$  parameter, which measures stochasticity, in Chapter 4.

Moreover, it is important to recognise that adolescents adjusted the amount of stochasticity in their decision-making between the stable and volatile environments, which was consistent with an optimal RL model. This effect could suggest that adolescents strategically adjust their amount of stochasticity in their decision-making according to the statistics of the environment. While adolescents may have heightened stochasticity in their decision-making due to structural immaturities in regions responsible for complex cognitive control (Shulman et al., 2016) and increased impulsivity (Steinberg et al., 2008), in showing that adolescents calibrate the amount of stochasticity in their decision-making based on the (in)stability of the environment, I suggest that there is greater evidence in favour of this ability being utilised to promote exploration during adolescence. This account would be consistent with the Lifespan Wisdom Model, which highlights the adaptive role of heightened exploration in adolescence relative to other age groups.

I also demonstrated that the adaptive ability to calibrate the amount of stochasticity in one's decision-making was impacted by anxiety. Previous research has demonstrated that anxiety can impact the ability to adjust one's learning rate between stable and volatile environments (Browning et al., 2015), whereas we highlight a novel computational mechanism that is impacted by anxiety: decisional stochasticity. These findings have implications for understanding how adolescents and adults adjust to changes in their environment. An inability to adapt behaviour to changes in the environment may be particularly detrimental during adolescence, as the individual experiences significant changes to their physical and social environment (Somerville et al., 2010). As such, these results may provide one avenue for

clinical intervention to reduce the negative impacts of disorders associated with anxiety in adolescence.

Although our stochasticity effects are consistent with the Lifespan Wisdom model, one prediction of the model that I did not find support for is the existence of distinct subgroups within adolescent populations. The Lifespan Wisdom Model suggests that only a small subgroup of adolescents engage in excessive risk taking, which is driven by impulsivity (Khurana et al., 2018; Romer et al., 2017). I failed to find an association between measures of impulsivity and our behavioural or computational measures in Chapters 3 and 4. However, one possible reason for this finding is that I did not have sufficient heterogeneity within the samples recruited to detect these different subgroups, as the 68 adolescent participants for Chapters 3 and 4 were recruited from a single UK school. Future research could recruit a more diverse sample, or selectively recruit adolescents high on impulsivity to examine individual differences that predict foraging strategies.

The data from Chapters 3, 4 and 5 collectively supports the predictions of the Lifespan Wisdom Model, which accounts for the real-world risk taking behaviours adolescents engage in, such as dangerous driving, risky sexual practices and substance misuse (Potard et al., 2008; Swedler et al., 2012; Willoughby et al., 2014). Rather than resulting from a tolerance for risk, without consideration of negative consequences as proposed by imbalance models (Steinberg, 2005; 2008), these findings suggest that adolescents' risky behaviour arises from their exploratory behaviours, consistent with the Lifespan Wisdom Model (Romer et al., 2017). This thesis contributes towards the development of the Lifespan Wisdom Model by identifying computational parameters that facilitate exploration in adolescence, providing a mechanistic account of heightened novelty-seeking during this developmental period. Future research could consider how adolescents' novelty-seeking relates to a wider range of behavioural outcomes,

such as the link between heightened exploration, foraging and positive risk taking as suggested in a recent review (Duell & Steinberg, 2021).

Overall, I have addressed my first objective through testing the predictions of models that explain the rise in novelty-seeking in adolescence. Rather than adolescents' behaviour being driven by the preference for immediate reward that may cause this age group to take risks, I find evidence that adolescents' behaviour is driven by heightened exploration relative to adults. These findings support the predictions of the Lifespan Wisdom Model in favour of predictions of imbalance models. Further, these findings contribute to our understanding of the role of exploration in adolescence, and how these exploratory behaviours can support this age group to learn about their surroundings. In doing so, these findings support the view that heightened novelty-seeking in adolescence plays a function role in development, allowing adolescents to acquire the independence that supports the transition from childhood to adulthood (Ellis et al., 2012).

### **8.3. Social Influence in Adolescence**

The second objective of this thesis was to test theories that explain how social influence affects adolescents' decision-making. Distraction theories propose that the presence of peers disrupts adolescents' ability to engage in goal-directed behaviour, making their actions more stochastic relative to when peers are absent (Ciranka & van den Bos, 2019). On this account, stochasticity results from distraction rather than the strategic implementation of this feature of decision-making discussed in relation to Chapter 5. In contrast, reward sensitivity theories suggest that the presence of same-aged peers makes adolescents hypersensitive to the potential for reward that can be acquired through risk taking. Finally, social motivation theories argue that adolescents are sensitive to the views that peers express and will adjust their behaviour to

conform to the preferences expressed by same-aged peers. My findings most closely support the predictions of social motivation models, which I demonstrate empirically in Chapter 6. In this study, I utilised a patch foraging paradigm where participants received feedback from a fictional peer, which I manipulated to endorse either optimal or suboptimal levels of exploration. Distraction theories predict that adolescents would be less precise in their behaviour than adults (Ciranka & van den Bos, 2019) and one way in which this may manifest is through adolescents exhibiting more stochastic behaviour in the presence of their peers relative to when their peers are absent. In contrast, reward sensitivity theories predict that peers sensitise adolescents to rewards (Chein et al., 2011), meaning the presence of peers should have a uniform effect across regardless of the type of feedback adolescents received from their fictional peer. Finally, social motivation theories would predict that adolescents will adjust their leaving thresholds according to the views their peers express (Knoll et al., 2015; 2017; Foulkes et al., 2018) and would therefore exhibit differences in their rates of exploration between social influence conditions.

The findings in Chapter 6 most closely resemble those predicted by social motivation theories, as I find evidence that adolescents adjust their leaving thresholds in the same direction of the feedback they receive from a similar-aged peer. Moreover, I find computational evidence that younger adolescents are more susceptible to adjusting their leaving thresholds in response to social influence compared to adults. However, inconsistent with social motivation theories that propose susceptibility to social influence during decision-making is heightened across adolescence, I also found that adults' exploration behaviour was susceptible to social influence. Nevertheless, these findings are consistent with research that has suggested that adolescents adjust their behaviour to conform to the views of similar aged peers, as the peers in our image set appeared to be the same age as our participants (Foulkes et al., 2018; Knoll et al., 2015; 2017). Notably, these findings highlight the potential for social influence to be both a positive

and negative influence during this developmental period, as adolescents were susceptible whether the peer advice promoted optimal or suboptimal levels of exploration.

The findings in Chapter 6 build on work that has examined how social influence affects adolescents' decisions to engage in risk taking (e.g., Braams et al., 2019; Reiter et al., 2019). Based on the findings in Chapters 3-5 I suggest that some behaviours that are typically considered reckless (e.g., trialling novel drugs) are more appropriately explained by adolescents' tendency for exploration. Examining how exploration behaviours are affected by social influence, I demonstrate that adolescents' leaving thresholds are affected by the views expressed by peers, similar to their risk choices in lab-based tasks. These findings are important as the feedback from exploration is utilised to inform future decision-making (Garrett & Daw, 2020). As such, the views and opinions of same aged peers may support adolescents to develop the 'wisdom' or experiential knowledge proposed by the Lifespan Wisdom Model (Romer et al., 2017). However, exploring new behaviours is not always positive and can be linked to harmful outcomes, such as trialling novel drugs. Indeed, this proposition would be consistent with evidence that socialising with peers who endorse substance use increases adolescents' likelihood of engaging in these harmful activities (Greenwood et al., 2021). Future research could expand on findings in this thesis by considering how adolescents adopt the exploration strategies of peers without explicit advice but instead through observation (i.e., demonstrate effects similar to risk contagion when making explore/exploit decisions; Reiter et al., 2019).

The findings from Chapter 6 further highlights the potential positive role for exploration in adolescent development that was also suggested in Chapters 3-5. That adolescents' explore/exploit decision-making is susceptible to social influence has important implications for our understanding of the positive role for peers during this developmental period. For example, Foulkes and colleagues (2018) demonstrated that adolescents rated positive, prosocial behaviour as more favourable if these activities were rated highly by peers. As the Lifespan

Wisdom model posits that exploration is associated with trialling new experiences (Romer et al., 2017), adolescents' susceptibility to social influence may be harnessed to guide interventions that can promote this age group to engage in positive activities (for example, using peers to encourage adolescents to trial novel educational activities). Conversely, these findings also highlight the risk factors associated with susceptibility to social influence during this developmental period. Overexploitation during explore/exploit dilemmas has been associated with compulsive traits (Addicott et al., 2017). As such, peers who endorse exploitation, rather than exploration, may also promote behaviours associated with compulsion (e.g., gambling; el-Guebaly et al., 2012). The findings of this thesis can inform interventions that seek to harness adolescents' susceptibility to social influence to promote positive behavioural outcomes for this age group.

One finding that I failed to replicate in Chapter 6 was that adolescents explored more than adults in the asocial condition of this task, which is an effect I found in Chapter 3, where we found that adolescents explored more than adults in four foraging environments and Chapter 5, where I found that adolescents explored more than adults in the stable condition of the stable and volatile environments. Moreover, I did not replicate the finding that adolescents had a higher learning rate than adults, which I observed in the computational analysis of participants' foraging behaviour in Chapter 4 and the modelling of participants behaviour in the stable and volatile environments in Chapter 5. One possible reason that I failed to replicate these findings is that the difference between foraging environments in Chapter 6 was operationalised through manipulating only the travel time between patches. In contrast, in Chapters 3, 4 and 5 I manipulated the within patch depletion rate and initial richness of patches in addition to travel time. It may be that adolescents' perception of the opportunity cost associated with travelling is similar to adults, meaning they would be less likely to explore if this time cost is perceived as outweighing the potential rewards available on new patches. Indeed, this interpretation

would be consistent with recent evidence that adolescents' and adults' exhibit similar perceptions of the opportunity cost of time, which impacts their willingness to engage cognitive control (Devine et al., 2021). Rather, the age-related effects we observed in Chapters 3, 4 and 5 may be driven instead by adolescents' perception of changing reward statistics (i.e., depletion rates and initial richness of patches) between foraging environments, which would be consistent with evidence that adolescents exhibit heightened sensitivity to changes in reward contingencies in paradigms which similarly manipulate probabilistic reward statistics (DePasque & Galvan, 2017; Hauser et al., 2015). The absence of changes to the reward statistics between environments would also impact the learning rate parameter, as participants would not be required to learn changes to the reward contingencies between environments, which may mean adolescents utilised a lower value on this parameter relative to the other empirical chapters in this thesis. Future research could consider modelling the perceived opportunity cost associated with changes to travel time to empirically test the interpretation that adolescents and adults do not differ in their sensitivity to time costs.

Future research could also expand on the findings of this thesis by considering how adolescents forage for social information. Research has demonstrated that non-human animals search for social information in a manner consistent with MVT (Turrin et al., 2017). Considering how adolescents and adults forage for social information could be applied to novel contexts, such as social media (Drias & Pasi, 2020). These sites contain discrete patches (i.e., social media pages) which contain social information (i.e., images or videos). One avenue for future research would be to examine whether adolescents exhibit social foraging behaviour that is consistent with MVT, and whether aberrant strategies in social foraging behaviour predict negative psychosocial outcomes offline, such as social anxiety (Shaw et al., 2015; though the relationship between social media use and psychopathology is strongly contested, see Orben & Przybylski, 2019). This field of research could be particularly timely as adolescents' social

interactions are increasingly moved online due to technological advances and factors such as the COVID-19 pandemic (Orben et al., 2020).

In summary, I have addressed my second objective, which was to test the predictions of theories of how social influence affects adolescents' decision-making. I find evidence to support social motivation theories over distraction or reward sensitivity theories. Indeed, the data are consistent with the view that adolescents place increased emphasis on the perspectives of their peers and will adjust their exploratory behaviour to conform to these views. I also highlight the potential for peers as both a protective and risk factor in adolescents' harmful or hazardous behaviours, which can inform interventions that seek to protect adolescents from negative outcomes whilst allowing this age group to explore their surroundings. Together, these findings contribute towards understanding the mechanisms through which social influence affects adolescents' behaviour.

#### **8.4. Adversity, Exploration and Reward Processing**

The third and final objective of this thesis was to test the predictions of Life History Theory, which explains how experiences of adversity during childhood and adolescence can impact decision-making in adulthood. Life History Theory proposes that experiences of adversity during development can direct the individual towards strategies of decision-making that maximise the individual's reproductive fitness. Individuals who experience unstable upbringings perceive resources as scarce, and this perception directs them towards a strategy of decision-making that prioritises short term rewards (i.e., a fast life history; Belsky et al., 1991). For example, the preference for immediate rewards associated with a fast life history strategy has been linked with poorer health outcomes, such as increased rates of obesity (Gunstad et al., 2006; Kaplan et al., 2000). These behaviours may be adaptive from an

evolutionary perspective, as they maximise calorific intake which can support short-term survival, but individuals who engage in these behaviours may not consider the longer-term health implications (Ellis et al., 2012). In contrast, individuals who experience relative stability in their upbringing are directed towards a strategy of decision-making that prioritises long term rewards and such a strategy is associated with positive health outcomes (i.e., a slow life history; Belsky et al., 1991). Due to their prioritisation of short-term rewards, as predicted by Life History Theory, I expected that individuals with Adverse Childhood Experiences (ACEs) would exhibit a preference of the immediate rewards that could be gained from exploiting patches rather than the delayed rewards gained through exploration relative to individuals without these experiences. Further, I examined whether ACE-exposed individuals' overexploitation bias was linked to computational mechanisms of reward learning, such as the learning rate.

The findings presented in Chapter 7 support the prediction of Life History Theory that individuals with experience of early adversity exhibit a greater overexploitation bias during patch foraging relative to individuals without these experiences. ACE-exposed individuals did, however, adjust their leaving thresholds and learning rates between foraging environments, suggesting they can adapt their behaviour to changes in the reward statistics of the environment. As exploiting patches is associated with an immediate reward at the expense of a larger, delayed reward gained through exploring, these findings further substantiate evidence that early life stress directs individuals towards a decision-making strategy that prioritises short term rewards. However, unlike previous research that has typically evidenced this preference through delay discounting paradigms (e.g., Lee et al., 2018), I find evidence for this preference in a sequential decision-making task, demonstrating these effects persist across a range of decision-making domains and to decisions with shorter temporal horizons (i.e., those with delays of seconds, as in the patch foraging task, rather than months, as in many delay discounting tasks; Odum,

2011). Though, I do recognise the conflicting findings between explore/exploit decisions and delay discounting preferences (Lloyd et al., 2021a; Sadeghiyeh et al., 2020). I also found that ACE-exposed individuals collected fewer rewards on the task, demonstrating how these experiences prevent the individual from taking advantage of the full bounty of rewards available in their environment. Together these findings can help to inform our understanding of the link between childhood adversity and adult poverty (Metzler et al., 2017).

In contrast to my predictions, I did not find evidence that individuals with ACEs had a higher learning rate compared to individuals without these experiences. Based on Life History Theory, I predicted that individuals' decision-making strategies would be adapted to their unstable and changeable caregiving environments, reflected computationally by a higher learning rate. Yet, I instead found that ACE-exposed individuals underweighted reward feedback relative to those without these experiences. ACE-exposed individuals' use of a lower learning rate may still have adaptive purposes, though not in the formulation specified by Life History Theory. Previous research has demonstrated that ACE-exposure is associated with diminished reward processing (Hanson et al., 2015) and a lower learning rate is one way in which diminished reward processing might manifest behaviourally (Halahakoon et al., 2020). Disordered reward processing is also a robust transdiagnostic risk factor for mental disorders, such as depression (Nusslock & Alloy, 2017; Qi et al., 2021). It has been suggested that mood and anxiety disorders form a cluster of psychopathologies that are an evolutionary adaptation to conditions of adversity (Boyce & Ellis, 2005). These disorders promote vigilant and avoidant behaviours (Ioannou et al., 2004; Ottenbreit & Dobson, 2004), which direct the individual away from hazards in their aversive environment (Del Giudice & Haltigan, 2021). Indeed, this account is consistent with neuroimaging research that has found regions such as the amygdala, which are responsible for threat detection, are hyperactive following ACEs (Cohen et al., 2013; Hanson et al., 2015; Tottenham et al., 2011; Van Tiegheem & Tottenham, 2017). As such, while

the finding that ACE-exposed individuals exhibit reduced reward processing may be inconsistent with Life History Theory, it may still be a product of adaptations made to avoid potentially harmful scenarios. This finding may be useful in contributing our understanding of the link between early adversity and psychopathology, though the specific link between ACEs, reward learning and psychopathology would need to be substantiated by future research.

The finding that ACE-exposed individuals have a lower learning rate compared to individuals without these experiences is consistent with the view that early adversity impairs functional connectivity between the amygdala and striatum (Fareri & Tottenham, 2016). Functional connectivity between the amygdala and striatum is necessary for the valuation and learning of reward contingencies (Costa et al., 2016; Fareri & Tottenham, 2016; Niv & Schoenbaum, 2008). The findings in Chapter 7 present preliminary behavioural evidence that the ability to learn stimulus-outcome contingencies is impaired following early adversity and manifests through reduced learning from reward feedback. The use of a lower learning rate impacts the ability of ACE-exposed individuals to evaluate the explore/exploit trade-off and assess global versus local reward rates while foraging.

In sum, my third objective was to test the predictions of Life History Theory, which purports to explain how experiences of adversity in childhood and adolescence impact decision-making in adulthood. Consistent with this theory, I find that ACEs are associated with the preference for immediate rewards gained through exploiting rather than accepting the cost of a delay to later achieve the rewards that can be gained from exploring. I also find that ACE-exposed individuals have a lower learning rate compared to individuals without these experiences and consider this evidence in context of the link between childhood trauma and psychopathology (Del Giudice & Haltigan, 2021). Together, these findings contribute to our understanding of the deleterious outcomes following childhood trauma and highlight direction for future work to understand the link between ACEs and psychopathology.

## **8.5. Limitations, Implications and Future Directions**

There are some limitations with this thesis that should be highlighted. One major limiting factor of the body of work presented in this thesis is the cross-sectional approach taken to studying the development of decision-making. Cross-sectional approaches preclude us from making causal inferences about how exploration behaviour develops from adolescence to adulthood, and how these behaviours are affected by social influence. Future research should consider measuring the development of explore/exploit decision-making over repeated time points to provide an explanation of the causal mechanisms underlying the age-related differences I observed in the foraging paradigms (Hamaker et al., 2020). An important design issue that future longitudinal foraging studies will need to consider is how participants learn the structure of foraging environments over repeated time points (i.e., those measured over months or years) and whether practice effects will confound any development trends observed.

A further limitation is that I did not measure the foraging behaviour of children aged below 10, which means I cannot make inferences about the development of foraging strategies across the lifespan. While ecological work has examined the development of foraging across the lifespan of non-human animals (e.g., Johnson & Wilbrecht, 2011), whether this translates to human development cannot be concluded from the present thesis. This limitation is important to consider, as I interpret the findings of Chapter 5 as suggesting that adolescents have more stochastic behaviour compared to other age groups due to this developmental stage being directed towards gaining diverse experience in the support of developing ‘wisdom’ (Romer et al., 2017). However, several studies have demonstrated that relative to adolescents, children exhibit more stochastic decision-making (e.g., Eckstein et al., 2020), which would challenge some arguments proposed in this thesis. Indeed, adopting the interpretation that

stochasticity can be adaptive (rather than indicating less precise decision-making), then the use of a stochastic decision-making in childhood may serve a similar purpose as it does during adolescence, allowing children to learn the structure of their environment (Gopnick, 2020; Pelz & Kidd, 2020). As such, it would be important for future research to compare children's, adolescents', and adults' foraging behaviour.

Finally, a methodological limitation of the present study is that I operationalised exploration as a fixed option, where participants do not have autonomy to decide which patches they visited. Recent work has demonstrated that adults track within patch reward rates when they have autonomy about which patches to visit and are able to return to patches they have previously exploited (Hall-McMaster et al., 2021). Moreover, formal models that accounted for tracking within patch reward rates provided a better explanation of participants' leaving thresholds than tracking the global reward rate (Hall-McMaster et al., 2021; Hall-McMaster & Luyckx, 2019). As such, the developmental trends I observed in reward-learning and exploration may be attributed to the deterministic exploration process I utilised in the design of the foraging tasks. Indeed, tracking reward rates for individual patches may require additional cognitive resources, such as working memory, to retain the expected value of each patch (Chatham & Badre, 2013). These additional requirements may favour adults' more mature cognitive capacities, as they have greater working memory capacities relative to adolescents (Luna et al., 2010; Steinberg et al., 2018). Yet, this task assumes that patches replenish within a timeframe where they could be utilised by the forager, which is not true of the ecological conditions under which hunter-gatherers forage (Venkataraman et al., 2017). As such, the foraging paradigms I utilise have the advantage of measuring how the forager estimates the background rate of environments where patches are encountered only once. Nevertheless, future research could compare adolescents' and adults' behaviour on these two paradigms and examine whether there are developmental differences in whether foraging

behaviour is best explained by tracking reward rates for individual patches or the average reward rate for the environment.

To build on the findings presented in the current thesis, future research could examine the neural correlates of foraging in adolescents and adults. There is evidence that the anterior cingulate cortex is responsible for monitoring the average reward rate in the environment, which is a quantity in the MVT equation (Kolling et al., 2012), and that the circuitry in the prefrontal cortex has an important role in evaluating foraging choices (Hunt et al., 2021; Passingham & Wise, 2012). Examining whether adolescents' developing prefrontal circuitry (Hooper et al., 2004) is associated with their heightened exploration behaviour would provide further insight into how neural adaptations support adolescents to explore their surroundings and learn from their environment. Indeed, this would also contribute evidence towards understanding how these neural regions develop across the lifespan, which can inform debates about the development of decision-making from childhood to adulthood (Ernst et al., 2006; Romer et al., 2017; Shulman et al., 2016; Steinberg, 2008).

An additional direction that future research could examine is how adolescents' heightened exploration, relative to adults, translates to information foraging. Information foraging describes how individuals search for information from online sources, where individual webpages represent patches (Pirolli & Card, 1999). In this context, the forager must evaluate the quantity of information being obtained from the information patch and whether the information gained from exploiting the current source outweighs the background rate of potential information that could be gathered from exploring the online environment (Pirolli & Fu, 2003). This context of decision-making may be timely as many education settings have moved online to adapt to the COVID-19 pandemic (Chakraborty et al., 2021). As such, examining how adolescents forage for information in an online setting may have implications for educational research.

Examining information foraging could also be of interest to when considering how online sources predict the formation and maintenance of individuals' attitudes. In the context of information foraging, the overexploitation bias that I and others observe in adults (e.g., Constantino & Daw, 2015) may prevent individuals from exploring the full range of information available in the online environment. The use of a limited number of information sources to gather information online predicts the formation of beliefs that have the potential for harm, such as conspiracy beliefs (Hartman et al., 2021) and extremist attitudes (Bright, 2017). Based on my developmental findings, adolescents may be less susceptible to the development of these views, as they explore more while foraging, which may translate to sampling a greater diversity of online sources. However, whether these exploratory behaviours translate to an information foraging context would need to be examined empirically in future research.

## **8.6. Conclusion**

In conclusion, this thesis has contributed to the literature on adolescent decision-making through testing the predictions of theories that explain how these cognitive abilities change across the lifespan. Specifically, I have addressed three key objectives, which were to test theories that attempt to explain how decision-making changes from adolescence to adulthood, to test theories that attempt to explain how social influence affects adolescents' decision-making, and finally to test theories about how experiences in childhood and adolescence affect decision-making in adulthood. Addressing these predictions, I have utilised several patch foraging paradigms, which has allowed me to measure decision-making in the explore/exploit trade-off. I have found that adolescents explore more than adults while foraging, and that this is associated with the degree to which they integrate reward feedback into their decision-making. Moreover, I have demonstrated that both adolescents and adults are adept at adjusting their decision-making between conditions of stability and volatility, though this ability can be impaired by anxiety. I have also demonstrated that adolescents' and adults' foraging behaviour

is susceptible to social influence, and that adverse childhood experiences are associated with reduced exploration and reward learning. Together, this thesis has demonstrated how heightened exploration in adolescence is associated with positive behavioural outcomes and that these behaviours may play a functional role in development. Importantly, these behaviours are susceptible to social influence and adverse experiences, highlighting potential avenues for intervention to protect adolescents from harmful outcomes and promote positive outcomes for this age group.

## References

- Addicott, M. A., Pearson, J. M., Schechter, J. C., Sapyta, J. J., Weiss, M. D., & Kollins, S. H. (2021). Attention-deficit/hyperactivity disorder and the explore/exploit trade-off. *Neuropsychopharmacology*, *46*(3), 614-621.
- 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.
- Afkinich, J. L., & Blachman-Demner, D. R. (2020). Providing incentives to youth participants in research: A literature review. *Journal of Empirical Research on Human Research Ethics*, *15*(3), 202-215.
- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike* (pp. 199-213). New York, US: Springer.
- Albert, D., & Steinberg, L. (2011). Judgment and decision making in adolescence. *Journal of Research on Adolescence*, *21*, 211-224.
- Albert, D., Chein, J., & Steinberg, L. (2013). The teenage brain: Peer influences on adolescent decision making. *Current Directions in Psychological Science*, *22*(2), 114-120.
- Alley, J., & Diamond, L. M. (2021). Early childhood adversity and Women's sexual behavior: The role of sensitivity to sexual reward. *Developmental Review*, *61*, 100982.
- Almy, B., Kuskowski, M., Malone, S. M., Myers, E., & Luciana, M. (2018). A longitudinal analysis of adolescent decision-making with the Iowa Gambling Task. *Developmental Psychology*, *54*(4), 689-702.
- Alonso, L. C., & Rosenfield, R. L. (2002). Oestrogens and puberty. *Best practice & research Clinical Endocrinology & Metabolism*, *16*, 13-30.

- Alvares, G. A., Balleine, B. W., & Guastella, A. J. (2014). Impairments in goal-directed actions predict treatment response to cognitive-behavioral therapy in social anxiety disorder. *PLoS One*, *9*(4), e94778.
- Amano, T., Ushiyama, K., Moriguchi, S., Fujita, G., & Higuchi, H. (2006). Decision-making in group foragers with incomplete information: test of individual-based model in geese. *Ecological Monographs*, *76*(4), 601-616.
- Andrews, J. L., Ahmed, S., & Blakemore, S. J. (2021). Navigating the social environment in adolescence: The role of social brain development. *Biological Psychiatry*, *89*(2), 109-118.
- Andrews, J. L., Foulkes, L., & Blakemore, S. J. (2020). Peer influence in adolescence: Public-health implications for COVID-19. *Trends in Cognitive Sciences*, *24*(8), 585-587.
- Andrews-Hanna, J. R., Seghete, K. L. M., Claus, E. D., Burgess, G. C., Ruzic, L., & Banich, M. T. (2011). Cognitive control in adolescence: neural underpinnings and relation to self-report behaviors. *PloS one*, *6*(6), e21598.
- Anokhin, A. P., Golosheykin, S., Grant, J., & Heath, A. C. (2009). Heritability of risk-taking in adolescence: a longitudinal twin study. *Twin Research and Human Genetics*, *12*(4), 366-371.
- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388-407.
- Aoki, C., Romeo, R. D., & Smith, S. S. (2017). Adolescence as a critical period for developmental plasticity. *Brain Res*, *1654*, 85-86.

- Arnett, J. J. (2002). Developmental sources of crash risk in young drivers. *Injury Prevention*, 8, ii17-ii23.
- Arnett, J. J., & Taber, S. (1994). Adolescence terminable and interminable: When does adolescence end?. *Journal of Youth and Adolescence*, 23(5), 517-537.
- Bach, D. R., Moutoussis, M., Bowler, A., Neuroscience in Psychiatry Network consortium, & Dolan, R. J. (2020). Predictors of risky foraging behaviour in healthy young people. *Nature Human Behaviour*, 4(8), 832-843.
- Baglivio, M. T., Wolff, K. T., Piquero, A. R., & Epps, N. (2015). The Relationship between Adverse Childhood Experiences (ACE) and Juvenile Offending Trajectories in a Juvenile Offender Sample. *Journal of Criminal Justice*, 43(3), 229–241. <https://doi.org/10.1016/j.jcrimjus.2015.04.012>
- Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., ... & Reiss, A. L. (2005). White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cerebral Cortex*, 15(12), 1848-1854.
- Baude, M., Dajoz, I., & Danchin, E. (2008). Inadvertent social information in foraging bumblebees: effects of flower distribution and implications for pollination. *Animal Behaviour*, 76(6), 1863-1873.
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214–1221. <https://doi.org/10.1038/nn1954>
- Bellis, M. A., Ashton, K., Hughes, K., Ford, K., Bishop, J., Paranjothy, S., & Public Health Wales. (2015). Adverse Childhood Experiences and their impact on health-harming behaviours in M.A Bellis, K. Ashton, K. Hughes, K. Ford, J. Bishop & S. Paranjothy

(Ed.) *Welsh adult population: Alcohol Use, Drug Use, Violence, Sexual Behaviour, Incarceration, Smoking and Poor Diet*. Cathays, Cardiff: Centre for Public Health.

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development*, 62(4), 647-670.

Bentall, R. P., Lloyd, A., Bennett, K., McKay, R., Mason, L., Murphy, J., ... & Shevlin, M. (2021). Pandemic buying: Testing a psychological model of over-purchasing and panic buying using data from the United Kingdom and the Republic of Ireland during the early phase of the COVID-19 pandemic. *PLOS ONE*, 16, e0246339.

Bethlehem, R., Seidlitz, J., White, S., Vogel, J., Anderson, K., Adamson, C., Adler, S., Alexopoulos, G., Anagnostou, E., Areces-Gonzalez, A., Astle, D., Auyeung, B., Ayub, M., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S., Benegal, V., Beyer, F., ... Alexander-Bloch, A. (2021). Brain charts for the human lifespan. *BioRxiv*, doi: <https://doi.org/10.1101/2021.06.08.447489>.

Bettinger, R. L., & Grote, M. N. (2016). Marginal value theorem, patch choice, and human foraging response in varying environments. *Journal of Anthropological Archaeology*, 42, 79–87. <https://doi.org/10.1016/j.jaa.2016.03.002>

Bird, D. W., Bird, R. B., Copping, B. F., & Zeanah, D. W. (2019). Variability in the organization and size of hunter-gatherer groups: Foragers do not live in small-scale societies. *Journal of Human Evolution*, 131, 96-108.

Bissonette, G. B., Burton, A. C., Gentry, R. N., Goldstein, B. L., Hearn, T. N., Barnett, B. R., ... & Roesch, M. R. (2013). Separate populations of neurons in ventral striatum encode value and motivation. *PLoS One*, 8(5), e64673.

- Bjork, J. M., & Pardini, D. A. (2015). Who are those “risk-taking adolescents”? Individual differences in developmental neuroimaging research. *Developmental Cognitive Neuroscience, 11*, 56-64.
- Blais, A.-R., & Weber, E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1*, 33-47.
- Blakemore, S. J. (2018). Avoiding social risk in adolescence. *Current Directions in Psychological Science, 27*(2), 116-122.
- Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47*(3-4), 296-312.
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing?. *Annual Review of Psychology, 65*, 187-207.
- Blakemore, S. J., Burnett, S., & Dahl, R. E. (2010). The role of puberty in the developing adolescent brain. *Human Brain Mapping, 31*(6), 926-933.
- Blakemore, S. J., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience, 15*(9), 1184-1191.
- Blanchard, T. C., & Gershman, S. J. (2018). Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective, & Behavioral Neuroscience, 18*, 117-126.
- Blanchard, T. C., & Hayden, B. Y. (2014). Neurons in Dorsal Anterior Cingulate Cortex Signal Postdecisional Variables in a Foraging Task. *Journal of Neuroscience, 34*(2), 646–655.  
<https://doi.org/10.1523/JNEUROSCI.3151-13.2014>
- Bland, A. R. & Schaefer, A. (2012). Different varieties of uncertainty in human decision-making. *Frontiers in Neuroscience, 6*, 1-12.

- Blankenstein, N. E., & van Duijvenvoorde, A. C. (2019). Neural tracking of subjective value under risk and ambiguity in adolescence. *Cognitive, Affective, & Behavioral Neuroscience, 19*(6), 1364-1378.
- Blankenstein, N. E., Crone, E. A., van den Bos, W., & van Duijvenvoorde, A. C. (2016). Dealing with uncertainty: Testing risk and ambiguity attitudes across adolescence. *Developmental Neuropsychology, 41*(1-2), 77-92.
- Blankenstein, N. E., Huettel, S. A., & Li, R. (2021). Resolving ambiguity: Broadening the consideration of risky decision making over adolescent development. *Developmental Review, 62*, 100987.
- Blankenstein, N. E., Schreuders, E., Peper, J. S., Crone, E. A., & van Duijvenvoorde, A. C. (2018). Individual differences in risk-taking tendencies modulate the neural processing of risky and ambiguous decision-making in adolescence. *NeuroImage, 172*, 663-673.
- Blom, R. M., Koeter, M., van den Brink, W., de Graaf, R., Ten Have, M., & Denys, D. (2011). Co-occurrence of obsessive-compulsive disorder and substance use disorder in the general population. *Addiction, 106*(12), 2178-2185.
- Boecker, R., Holz, N. E., Buchmann, A. F., Blomeyer, D., Plichta, M. M., Wolf, I., Baumeister, S., Meyer-Lindenberg, A., Banaschewski, T., Brandeis, D., & Laucht, M. (2014). Impact of early life adversity on reward processing in young adults: EEG-fMRI results from a prospective study over 25 years. *PLoS ONE, 9*(8), e104185. <https://doi.org/10.1371/journal.pone.0104185>
- Botvinick, M. M., Niv, Y., & Barto, A. G. (2009). Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition, 113*(3), 262-280.

- Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, *17*(2), 271-301. <https://doi.org/10.1017/S0954579405050145>
- Braams, B. R., Davidow, J. Y., & Somerville, L. (2020). Information about others' choices selectively alters risk tolerance and medial prefrontal cortex activation across adolescence and young adulthood, *Developmental Cognitive Neuroscience*, *52*, 101039-101051
- Braams, B. R., Davidow, J. Y., & Somerville, L. H. (2019). Developmental patterns of change in the influence of safe and risky peer choices on risky decision-making. *Developmental Science*, *22*, e12717.
- Braams, B. R., van Duijvenvoorde, A. C., Peper, J. S., & Crone, E. A. (2015). Longitudinal changes in adolescent risk-taking: a comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *Journal of Neuroscience*, *35*(18), 7226-7238.
- Brieant, A., Peviani, K. M., Lee, J. E., King-Casas, B., & Kim-Spoon, J. (2021). Socioeconomic Risk for Adolescent Cognitive Control and Emerging Risk-Taking Behaviors. *Journal of Research on Adolescence*, *31*, 71-84.
- Bright, J. (2017). Explaining the emergence of echo chambers on social media: the role of ideology and extremism. Retrieved from: [arXiv:1609.05003](https://arxiv.org/abs/1609.05003)
- British Psychological Society. (2019). Code of ethics and conduct. Retrieved from: <https://www.bps.org.uk/sites/bps.org.uk/files/Policy%20-%20Files/BPS%20Code%20of%20Ethics%20and%20Conduct%20%28Updated%20July%202018%29.pdf>

- Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience, 18*(4), 590–596. <https://doi.org/10.1038/nn.3961>
- Brumbach, B. H., Figueredo, A. J., & Ellis, B. J. (2009). Effects of harsh and unpredictable environments in adolescence on development of life history strategies. *Human Nature, 20*, 25-51.
- Buritica, J. M. R., Heekeren, H. R., & van den Bos, W. (2019). The computational basis of following advice in adolescents. *Journal of Experimental Child Psychology, 180*, 39-54.
- Burnett, S., Bault, N., Coricelli, G., & Blakemore, S. J. (2010). Adolescents' heightened risk-seeking in a probabilistic gambling task. *Cognitive Development, 25*(2), 183-196.
- Camara, E., Rodriguez-Fornells, A., & Münte, T. F. (2009). Functional connectivity of reward processing in the brain. *Frontiers in Human Neuroscience, 2*, 1-14.
- Carskadon, M. A., & Acebo, C. (1993). A self-administered rating scale for pubertal development. *Journal of Adolescent Health, 14*(3), 190-195.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319.
- Casey, B. J. (2015). Beyond simple models of self-control to circuit-based accounts of adolescent behavior. *Annual Review of Psychology, 66*, 295-319.
- Casey, B. J., Getz, S., & Galvan, A. (2008). The adolescent brain. *Developmental Review, 28*, 62-77.

- Casey, B. J., Jones, R. M., & Somerville, L. H. (2011). Braking and accelerating of the adolescent brain. *Journal of Research on Adolescence*, *21*, 21-33.
- Casey, B. J., Jones, R. M., Levita, L., Libby, V., Pattwell, S. S., Ruberry, E. J., ... & Somerville, L. H. (2010). The storm and stress of adolescence: insights from human imaging and mouse genetics. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, *52*(3), 225-235.
- Chaby, L. E., Sheriff, M. J., Cavigelli, S. A., Hirrlinger, A. M., Lim, J., & Braithwaite, V. A. (2016). Stress during adolescence shapes performance in adulthood: Context-dependent effects on foraging and vigilance. *Ethology*, *122*(4), 284-297.
- Chaby, L. E., Sheriff, M. J., Hirrlinger, A. M., & Braithwaite, V. A. (2015). Does early stress prepare individuals for a stressful future? Stress during adolescence improves foraging under threat. *Animal Behaviour*, *105*, 37-45.
- Chakraborty, P., Mittal, P., Gupta, M. S., Yadav, S., & Arora, A. (2021). Opinion of students on online education during the COVID-19 pandemic. *Human Behavior and Emerging Technologies*, *3*(3), 357-365.
- Chamorro, J., Bernardi, S., Potenza, M. N., Grant, J. E., Marsh, R., Wang, S., & Blanco, C. (2012). Impulsivity in the general population: a national study. *Journal of Psychiatric Research*, *46*(8), 994-1001.
- Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129-136. [https://doi.org/10.1016/0040-5809\(76\)90040-X](https://doi.org/10.1016/0040-5809(76)90040-X)
- Chatham, C. H., & Badre, D. (2013). Working memory management and predicted utility. *Frontiers in Behavioral Neuroscience*, *7*, 83.

- Chein, J., Albert, D., O'Brien, L., Uckert, K., & Steinberg, L. (2011). Peers increase adolescent risk taking by enhancing activity in the brain's reward circuitry. *Developmental Science, 14*(2), F1-F10.
- Cinotti, F., Fresno, V., Aklil, N., Coutureau, E., Girard, B., Marchand, A. R., & Khamassi, M. (2019). Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Scientific Reports, 9*, 1-14.
- Ciranka, S., & van den Bos, W. (2019). Social influence in adolescent decision-making: A formal framework. *Frontiers in Psychology, 10*, 1915.
- Ciranka, S., & van den Bos, W. (2021). Adolescent risk-taking in the context of exploration and social influence. *Developmental Review, 61*, Article 100979.
- Coffman, M. C., Trubanova, A., Richey, J. A., White, S. W., Kim-Spoon, J., Ollendick, T. H., & Pine, D. S. (2015). Validation of the NIMH-ChEFS adolescent face stimulus set in an adolescent, parent, and health professional sample. *International Journal of Methods in Psychiatric Research, 24*(4), 275-286.
- 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 Science of Learning, 5*, 1-9.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.
- Cohen, J. R., Asarnow, R. F., Sabb, F. W., Bilder, R. M., Bookheimer, S. Y., Knowlton, B. J., & Poldrack, R. A. (2010). A unique adolescent response to reward prediction errors. *Nature Neuroscience, 13*(6), 669-671.

- Cohen, M. M., Jing, D., Yang, R. R., Tottenham, N., Lee, F. S., & Casey, B. J. (2013). Early-life stress has persistent effects on amygdala function and development in mice and humans. *Proceedings of the National Academy of Sciences*, *110*(45), 18274-18278.
- Cohen, M. X., Elger, C. E., & Weber, B. (2008). Amygdala tractography predicts functional connectivity and learning during feedback-guided decision-making. *Neuroimage*, *39*(3), 1396-1407.
- Cohen, R. A., Grieve, S., Hoth, K. F., Paul, R. H., Sweet, L., Tate, D., Gunstad, J., Stroud, L., McCaffery, J., Hitsman, B., Niaura, R., Clark, C. R., MacFarlane, A., Bryant, R., Gordon, E., & Williams, L. M. (2006). Early Life Stress and Morphometry of the Adult Anterior Cingulate Cortex and Caudate Nuclei. *Biological Psychiatry*, *59*(10), 975–982. <https://doi.org/10.1016/j.biopsych.2005.12.016>
- Collins, A. G., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working memory contributions to reinforcement learning impairments in schizophrenia. *Journal of Neuroscience*, *34*(41), 13747-13756.
- Constantino, S. M., & Daw, N. D. (2015). Learning the opportunity cost of time in a patch-foraging task. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(4), 837–853. <https://doi.org/10.3758/s13415-015-0350-y>
- Costa, V. D., Dal Monte, O., Lucas, D. R., Murray, E. A., & Averbeck, B. B. (2016). Amygdala and ventral striatum make distinct contributions to reinforcement learning. *Neuron*, *92*(2), 505-517.
- Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2014). Dopamine modulates novelty seeking behavior during decision making. *Behavioral Neuroscience*, *128*(5), 556-566.

- Crone, E. A. (2009). Executive functions in adolescence: inferences from brain and behavior. *Developmental Science*, *12*(6), 825-830.
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, *13*(9), 636-650.
- Crone, E. A., van Duijvenvoorde, A. C., & Peper, J. S. (2016). Annual Research Review: Neural contributions to risk-taking in adolescence—developmental changes and individual differences. *Journal of Child Psychology and Psychiatry*, *57*(3), 353-368.
- da Silva, C. F., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour*, *4*(10), 1053-1066.
- Danese, A., & McEwen, B. S. (2012). Adverse childhood experiences, allostasis, allostatic load, and age-related disease. *Physiology & Behavior*, *106*, 29–39. <https://doi.org/10.1016/j.physbeh.2011.08.019>
- Davidow, J. Y., Foerde, K., Galván, A., & Shohamy, D. (2016). An upside to reward sensitivity: the hippocampus supports enhanced reinforcement learning in adolescence. *Neuron*, *92*, 93-99.
- Davidson, J. D., & El Hady, A. (2019). Foraging as an evidence accumulation process. *PLoS Computational Biology*, *15*(7), e1007060.
- Davis, M., and Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Mol. Psychiatry*, *6*, 13–34.
- Davis-Kean, P. E., & Ellis, A. (2019). An overview of issues in infant and developmental research for the creation of robust and replicable science. *Infant Behavior and Development*, *57*, 101339.

- Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision making, affect, and learning: Attention and performance XXIII*, 23(1), 1-26.
- Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12), 1704-1711.
- Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095), 876-879.
- Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: reevaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2), 473-492.
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4), 429-453.
- de Boer, A., & Harakeh, Z. (2017). The effect of active and passive peer discouragement on adolescent risk taking: An experimental study. *Journal of Research on Adolescence*, 27(4), 878-889.
- de Water, E., Cillessen, A. H., & Scheres, A. (2014). Distinct age-related differences in temporal discounting and risk taking in adolescents and young adults. *Child Development*, 85(5), 1881-1897.
- Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective, & Behavioral Neuroscience*, 15(2), 310-320.

- 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. *Psychological Science, 27*(6), 848-858.
- Defoe, I. N., Dubas, J. S., Figner, B., & Van Aken, M. A. (2015). A meta-analysis on age differences in risky decision making: adolescents versus children and adults. *Psychological Bulletin, 141*, 48-84.
- Del Giacco, A. C., Jones, S. A., Morales, A. M., Kliamovich, D., & Nagel, B. J. (2021). Adolescent novelty seeking is associated with greater ventral striatal and prefrontal brain response during evaluation of risk and reward. *Cognitive, Affective, & Behavioral Neuroscience, 1-11*.
- Del Giudice, M., & Haltigan, J. D. (2021). An integrative evolutionary framework for psychopathology. *Development and Psychopathology, 1-11*. doi: 10.1017/S0954579421000870
- DePasque, S., & Galván, A. (2017). Frontostriatal development and probabilistic reinforcement learning during adolescence. *Neurobiology of Learning and Memory, 143*, 1-7.
- Deserno, L., Boehme, R., Mathys, C., Katthagen, T., Kaminski, J., Stephan, K. E., ... & Schlagenhauf, F. (2020). Volatility estimates increase choice switching and relate to prefrontal activity in schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(2), 173-183.
- Devine, S., Neumann, C., Otto, A. R., Bolenz, F., Reiter, A., & Eppinger, B. (2021). Seizing the opportunity: Lifespan differences in the effects of the opportunity cost of time on cognitive control. *Cognition, 216*, 104863.

- Dickins, T. E., & Schalz, S. (2020). Food shopping under risk and uncertainty. *Learning and Motivation*, 72, 101681.
- Dillon, D. G., Holmes, A. J., Birk, J. L., Brooks, N., Lyons-Ruth, K., & Pizzagalli, D. A. (2009). Childhood adversity is associated with left basal ganglia dysfunction during reward anticipation in adulthood. *Biological Psychiatry*, 66(3), 206-213.
- Do, K. T., Moreira, J. F. G., & Telzer, E. H. (2017). But is helping you worth the risk? Defining prosocial risk taking in adolescence. *Developmental Cognitive Neuroscience*, 25, 260-271.
- Do, K. T., Sharp, P. B., & Telzer, E. H. (2020). Modernizing conceptions of valuation and cognitive-control deployment in adolescent risk taking. *Current Directions in Psychological Science*, 29, 102-109.
- Doll, B. B., Simon, D. A., & Daw, N. D. (2012). The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology*, 22(6), 1075-1081.
- Donati, M. A., Chiesi, F., & Primi, C. (2015). The role of probabilistic reasoning abilities on adolescent risk taking. *The Mathematics Enthusiast*, 12, 262-274.
- Drias, Y., & Pasi, G. (2020, April). Credible information foraging on social media. In *World Conference on Information Systems and Technologies* (pp. 415-425). Cham, CH: Springer.
- Dubois, M., Bowler, A., Moses-Payne, M. E., Habicht, J., Steinbeis, N., & Hauser, T. U. (2020). Tabula-rasa exploration decreases during youth and is linked to ADHD symptoms. *BioRxiv*.

- Dubois, M., Habicht, J., Michely, J., Moran, R., Dolan, R. J., & Hauser, T. U. (2021). Human complex exploration strategies are enriched by noradrenaline-modulated heuristics. *Elife*, *10*, e59907.
- Duell, N., & Steinberg, L. (2019). Positive risk taking in adolescence. *Child Development Perspectives*, *13*, 48-52.
- Duell, N., & Steinberg, L. (2020). Differential correlates of positive and negative risk taking in adolescence. *Journal of Youth and Adolescence*, *49*(6), 1162-1178.
- Duell, N., & Steinberg, L. (2021). Adolescents take positive risks, too. *Developmental Review*, *62*, 100984
- Duell, N., Steinberg, L., Chein, J., Al-Hassan, S. M., Bacchini, D., Lei, C., ... & Alampay, L. P. (2016). Interaction of reward seeking and self-regulation in the prediction of risk taking: A cross-national test of the dual systems model. *Developmental Psychology*, *52*(10), 1593-1605.
- Dumontheil, I. (2016). Adolescent brain development. *Current Opinion in Behavioral Sciences*, *10*, 39-44.
- Dunlap, A. S., Nielsen, M. E., Dornhaus, A., & Papaj, D. R. (2016). Foraging bumble bees weigh the reliability of personal and social information. *Current Biology*, *26*(9), 1195-1199.
- Eaton, D. K., Kann, L., Kinchen, S., Shanklin, S., Flint, K. H., Hawkins, J., ... Whittle, L. (2012). Youth risk behavior surveillance—United States, 2011. *Morbidity and Mortality Weekly Report: Surveillance Summaries*, *61*(4), 1–162.

- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*, 351-362.
- Eccles, J. S., Jacobs, J. E., Harold, R. D., Yoon, K. S., Abreton, A., & Freedman-Doan, C. (1993). *Parents and gender-role socialization during the middle childhood and adolescent years*. New York, US: Sage Publications, Inc.
- Eckstein, M. K., Master, S. L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2020). Understanding the Unique Advantage of Adolescents in Stochastic, Volatile Environments: Combining Reinforcement Learning and Bayesian Inference. *BioRxiv*.
- Eckstein, M. K., Master, S. L., Xia, L., Dahl, R. E., Wilbrecht, L., & Collins, A. G. E. (2021). Learning Rates Are Not All the Same: The Interpretation of Computational Model Parameters Depends on the Context. *BioRxiv*.
- Eckstein, M. K., Wilbrecht, L., & Collins, A. G. (2021). What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. *Current Opinion in Behavioral Sciences*, *41*, 128-137.
- Edwards, S., & Koob, G. F. (2010). Neurobiology of dysregulated motivational systems in drug addiction. *Future Neurology*, *5*(3), 393-410.
- el-Guebaly, N., Mudry, T., Zohar, J., Tavares, H., & Potenza, M. N. (2012). Compulsive features in behavioural addictions: the case of pathological gambling. *Addiction*, *107*(10), 1726-1734.
- Ellis, B. J., & Del Giudice, M. (2019). Developmental Adaptation to Stress: An Evolutionary Perspective. *Annual Review of Psychology*, *70*, 111–139.  
<https://doi.org/10.1146/annurev-psych-122216-011732>

- Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology*, *23*, 7–28. <https://doi.org/10.1017/S0954579410000611>
- Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicius, V., ... & Wilson, D. S. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology*, *48*(3), 598-623.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 643-669.
- Elmore, A. L., & Crouch, E. (2020). The association of adverse childhood experiences with anxiety and depression for children and youth, 8 to 17 years of age. *Academic Pediatrics*, *20*(5), 600-608.
- Ernst, M., Pine, D. S., & Hardin, M. (2006). Triadic model of the neurobiology of motivated behavior in adolescence. *Psychological Medicine*, *36*(3), 299-312.
- Fairchild, G., van Goozen, S. H., Stollery, S. J., Aitken, M. R., Savage, J., Moore, S. C., & Goodyer, I. M. (2009). Decision making and executive function in male adolescents with early-onset or adolescence-onset conduct disorder and control subjects. *Biological Psychiatry*, *66*(2), 162-168.
- Fancourt, D., Steptoe, A., & Bu, F. (2021). Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: a longitudinal observational study. *The Lancet Psychiatry*, *8*(2), 141-149.

- Farb, A. F., & Matjasko, J. L. (2012). Recent advances in research on school-based extracurricular activities and adolescent development. *Developmental Review, 32*, 1-48.
- Fareri, D. S., & Tottenham, N. (2016). Effects of early life stress on amygdala and striatal development. *Developmental Cognitive Neuroscience, 19*, 233-247.
- Fareri, D. S., Martin, L. N., & Delgado, M. R. (2008). Reward-related processing in the human brain: developmental considerations. *Development and Psychopathology, 20*(4), 1191-1211.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.
- Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults. *American Journal of Preventive Medicine, 14*(4), 245–258. [https://doi.org/10.1016/S0749-3797\(98\)00017-8](https://doi.org/10.1016/S0749-3797(98)00017-8)
- Felzmann, H. (2009). Ethical Issues in School-Based Research. *Research Ethics, 5*(3), 104-109. <https://doi.org/pdf/10.1177/174701610900500304>
- Ferguson, H. J., Brunson, V. E., & Bradford, E. E. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*, 1-17.
- Figuer, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science, 20*(4), 211-216.

- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M., Sefcek, J. A., Tal, I. R., ... & Jacobs, W. J. (2006). Consilience and Life History Theory: From genes to brain to reproductive strategy. *Developmental Review, 26*(2), 243-275.
- Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review, 26*(4), 1099-1121.
- Forbes, E. E., Ryan, N. D., Phillips, M. L., Manuck, S. B., Worthman, C. M., Moyles, D. L., ... & Dahl, R. E. (2010). Healthy adolescents' neural response to reward: associations with puberty, positive affect, and depressive symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry, 49*(2), 162-172.
- Foulkes, L., Leung, J. T., Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2018). Age differences in the prosocial influence effect. *Developmental Science, 21*(6), e12666.
- Fried, E. I. (2020). Lack of theory building and testing impedes progress in the factor and network literature. *Psychological Inquiry, 31*(4), 271-288.
- Fuhrmann, D., Knoll, L. J., & Blakemore, S. J. (2015). Adolescence as a sensitive period of brain development. *Trends in Cognitive Sciences, 19*(10), 558-566.
- Gabay, A. S., & Apps, M. A. (2021). Foraging optimally in social neuroscience: computations and methodological considerations. *Social Cognitive and Affective Neuroscience, 16*(8), 782-794.
- Gabrieli, J. D. (1998). Cognitive neuroscience of human memory. *Annual Review of Psychology, 49*, 87-115.

- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *Journal of Neuroscience*, *26*(25), 6885-6892.
- Gamble, J. (2017). Early starters. *Nature*, *550*(7674), S10-S11.
- Gardner, M., & Steinberg, L. (2005). Peer influence on risk taking, risk preference, and risky decision making in adolescence and adulthood: an experimental study. *Developmental Psychology*, *41*(4), 625.
- Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*, 3417. <https://doi.org/10.1038/s41467-020-16964-5>
- Gershman, S. J. (2015). Do learning rates adapt to the distribution of rewards?. *Psychonomic Bulletin & Review*, *22*(5), 1320-1327.
- Gershman, S. J. (2018). Deconstructing the human algorithms for exploration. *Cognition*, *173*, 34-42.
- Gershman, S. J. (2019). How to never be wrong. *Psychonomic Bulletin & Review*, *26*, 13-28.
- Gershman, S. J., & Tzovaras, B. G. (2018). Dopaminergic genes are associated with both directed and random exploration. *Neuropsychologia*, *120*, 97-104.
- Gibbons, F. X., Houlihan, A. E., & Gerrard, M. (2009). Reason and reaction: The utility of a dual-focus, dual-processing perspective on promotion and prevention of adolescent health risk behaviour. *British Journal of Health Psychology*, *14*(2), 231-248.
- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., ... & Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*(10), 861-863.

- Gigerenzer, G. (1998). Surrogates for theories. *Theory & Psychology*, 8(2), 195-204.
- Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585-595.
- Goosby, B. J., Bellatorre, A., Walsemann, K. M., & Cheadle, J. E. (2013). Adolescent loneliness and health in early adulthood. *Sociological Inquiry*, 83(4), 505-536.
- Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B*, 375(1803), 20190502.
- Greenwood, C. J., Youssef, G. J., Letcher, P., Spry, E. A., Thomson, K. C., Hagg, L. J., ... & Olsson, C. A. (2021). Exploring a causal model in observational cohort data: The role of parents and peers in shaping substance use trajectories. *Addictive Behaviors*, 112, 106597.
- Grüter, C., Czaczkes, T. J., & Ratnieks, F. L. (2011). Decision making in ant foragers (*Lasius niger*) facing conflicting private and social information. *Behavioral Ecology and Sociobiology*, 65(2), 141-148.
- Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science*, 16(4), 789-802.
- Gunstad, J., Paul, R. H., Spitznagel, M. B., Cohen, R. A., Williams, L. M., Kohn, M., & Gordon, E. (2006). Exposure to early life trauma is associated with adult obesity. *Psychiatry Research*, 142, 31-37.
- Gurven, M., Kaplan, H., & Gutierrez, M. (2006). How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *Journal of Human Evolution*, 51(5), 454-470.

- Haddad, A. D., Harrison, F., Norman, T., & Lau, J. Y. (2014). Adolescent and adult risk-taking in virtual social contexts. *Frontiers in Psychology, 5*, 1476.
- Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic review and meta-analysis. *JAMA psychiatry, 77*(12), 1286-1295.
- Hall, G. S. (1904). Adolescence: Its psychology and its relations to physiology, anthropology, sociology, sex, crime, religion, and education , Vols. 1 & 2 . New York, US: Appleton.
- Hall-McMaster, S., & Luyckx, F. (2019). Revisiting foraging approaches in neuroscience. *Cognitive, Affective, & Behavioral Neuroscience, 19*(2), 225-230.
- Hall-McMaster, S., Dayan, P., & Schuck, N. W. (2021). Control over patch encounters changes foraging behaviour. *Available at SSRN 3770102*.
- Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. *Developmental Cognitive Neuroscience, 100867*.
- Hanson, J. L., Hariri, A. R., & Williamson, D. E. (2015). Blunted Ventral Striatum Development in Adolescence Reflects Emotional Neglect and Predicts Depressive Symptoms. *Biological Psychiatry, 78*(9), 598–605.  
<https://doi.org/10.1016/j.biopsych.2015.05.010>
- Hanson, J. L., Nacewicz, B. M., Sutterer, M. J., Cayo, A. A., Schaefer, S. M., Rudolph, K. D., ... & Davidson, R. J. (2015). Behavioral problems after early life stress: contributions of the hippocampus and amygdala. *Biological Psychiatry, 77*(4), 314-323.
- Hanson, J. L., van den Bos, W., Roeber, B. J., Rudolph, K. D., Davidson, R. J., & Pollak, S. D. (2017). Early adversity and learning: Implications for typical and atypical behavioral

- development. *Journal of Child Psychology and Psychiatry*, 58(7), 770–778.  
<https://doi.org/10.1111/jcpp.12694>
- Harakeh, Z., & de Boer, A. (2019). The effect of active and passive peer encouragement on adolescent risk-taking. *Journal of Adolescence*, 71, 10-17.
- Hart, G., Leung, B. K., & Balleine, B. W. (2014). Dorsal and ventral streams: the distinct role of striatal subregions in the acquisition and performance of goal-directed actions. *Neurobiology of Learning and Memory*, 108, 104-118.
- Hartman, T. K., Marshall, M., Stocks, T. V., McKay, R., Bennett, K., Butter, S., ... & Bentall, R. P. (2021). Different conspiracy theories have different psychological and social determinants: Comparison of three theories about the origins of the COVID-19 virus in a representative sample of the UK population. *Frontiers in Political Science*, 3, 44.
- Hauser, T. U., Fiore, V. G., Moutoussis, M., & Dolan, R. J. (2016). Computational psychiatry of ADHD: neural gain impairments across Marrian levels of analysis. *Trends in Neurosciences*, 39(2), 63-73.
- Hauser, T. U., Iannaccone, R., Walitza, S., Brandeis, D., & Brem, S. (2015). Cognitive flexibility in adolescence: neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *Neuroimage*, 104, 347-354.
- Haushofer, J., & Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186), 862–867.  
<https://doi.org/10.1126/science.1232491>
- Hawkes, K., Hill, K., & O'Connell, J. F. (1982). Why hunters gather: optimal foraging and the Ache of eastern Paraguay. *American Ethnologist*, 9(2), 379-398.

- Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933.
- Holmes, C., Briant, A., Kahn, R., Deater-Deckard, K., & Kim-Spoon, J. (2019). Structural home environment effects on developmental trajectories of self-control and adolescent risk taking. *Journal of Youth and Adolescence*, *48*, 43-55.
- Hoogendam, J. M., Kahn, R. S., Hillegers, M. H., van Buuren, M., & Vink, M. (2013). Different developmental trajectories for anticipation and receipt of reward during adolescence. *Developmental Cognitive Neuroscience*, *6*, 113-124.
- Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents' performance on the Iowa Gambling Task: implications for the development of decision making and ventromedial prefrontal cortex. *Developmental Psychology*, *40*(6), 1148.
- Hopfer, C., Salomonsen-Sautel, S., Mikulich-Gilbertson, S., Min, S. J., McQueen, M., Crowley, T., ... & Hewitt, J. (2013). Conduct disorder and initiation of substance use: a prospective longitudinal study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(5), 511-518.
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, *49*(5), 765-775.
- Hughes, K., Bellis, M. A., Hardcastle, K. A., Sethi, D., Butchart, A., Mikton, C., Jones, L., & Dunne, M. P. (2017). The effect of multiple adverse childhood experiences on health: A systematic review and meta-analysis. *The Lancet Public Health*, *2*(8), e356–e366. [https://doi.org/10.1016/S2468-2667\(17\)30118-4](https://doi.org/10.1016/S2468-2667(17)30118-4)
- Huh, H. J., Baek, K., Kwon, J.-H., Jeong, J., & Chae, J.-H. (2016). Impact of childhood trauma and cognitive emotion regulation strategies on risk-averse and loss-averse patterns

- of decision-making in patients with depression. *Cognitive Neuropsychiatry*, *21*(6), 447–461. <https://doi.org/10.1080/13546805.2016.1230053>
- Humphreys, K. L., Lee, S. S., Telzer, E. H., Gabard-Durnam, L. J., Goff, B., Flannery, J., & Tottenham, N. (2015). Exploration-exploitation strategy is dependent on early experience: Exploration-Exploitation. *Developmental Psychobiology*, *57*(3), 313–321. <https://doi.org/10.1002/dev.21293>
- Hunt, L. T. (2021). Frontal circuit specialisations for decision making. *European Journal of Neuroscience*, *53*(11), 3654-3671
- Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Muga, U., Procyk, E., ... & Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, 1-14.
- Icenogle, G., & Cauffman, E. (2021). Adolescent decision making: A decade in review. *Journal of Research on Adolescence*, *31*(4), 1006-1022.
- Ioannou, M. C., Mogg, K., & Bradley, B. P. (2004). Vigilance for threat: Effects of anxiety and defensiveness. *Personality and Individual Differences*, *36*(8), 1879-1891.
- JASP Team (2019). JASP (Version 0.11.1) [Computer software].
- Javadi, A. H., Schmidt, D. H., & Smolka, M. N. (2014). Adolescents adapt more slowly than adults to varying reward contingencies. *Journal of Cognitive Neuroscience*, *26*(12), 2670-2681.
- Jaworska, N., & MacQueen, G. (2015). Adolescence as a unique developmental period. *Journal of Psychiatry & Neuroscience: JPN*, *40*(5), 291.

- Jensen, P. S., Mrazek, D., Knapp, P. K., Steinberg, L., Pfeffer, C., Schowalter, J., & Shapiro, T. (1997). *Evolution and Revolution in Child Psychiatry: ADHD as a Disorder of Adaptation*, 161, 160-168
- Jepma, M., Schaaf, J. V., Visser, I., & Huizenga, H. M. (2020). Uncertainty-driven regulation of learning and exploration in adolescents: A computational account. *PLoS Computational Biology*, 16(9), e1008276.
- Johnson, A. W., & Earle, T. K. (2000). *The evolution of human societies: from foraging group to agrarian state*. Stanford, US: Stanford University Press.
- Johnson, C., & Wilbrecht, L. (2011). Juvenile mice show greater flexibility in multiple choice reversal learning than adults. *Developmental Cognitive Neuroscience*, 1(4), 540-551.
- Jorm, A. F., Christensen, H., Rodgers, B., Jacomb, P. A., & Easteal, S. (2004). Association of adverse childhood experiences, age of menarche, and adult reproductive behavior: does the androgen receptor gene play a role?. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 125, 105-111.
- Kahneman, D. and Tversky, A., Choices, values, and frames. In Choices, Values, and Frames, edited by D. Kahneman and A. Tversky, Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: Diet, intelligence, and longevity. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 9(4), 156-185.
- Kasperek, S. W., Jenness, J. L., & McLaughlin, K. A. (2020). Reward Processing Modulates the Association Between Trauma Exposure and Externalizing Psychopathology. *Clinical Psychological Science*, 8(6), 989-1006.

- Katehakis, M. N., & Veinott Jr, A. F. (1987). The multi-armed bandit problem: decomposition and computation. *Mathematics of Operations Research*, *12*(2), 262-268.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196-217.
- Khurana, A., Romer, D., Betancourt, L. M., & Hurt, H. (2018). Modeling trajectories of sensation seeking and impulsivity dimensions from early to late adolescence: universal trends or distinct sub-groups?. *Journal of Youth and Adolescence*, *47*(9), 1992-2005.
- Kiai, A. (2019). To protect credibility in science, banish “publish or perish”. *Nature Human Behaviour*, *3*(10), 1017-1018.
- Kievit, R. A. (2020). Sensitive periods in cognitive development: a mutualistic perspective. *Current Opinion in Behavioral Sciences*, *36*, 144-149.
- Kim, S. H., Yoon, H., Kim, H., & Hamann, S. (2015). Individual differences in sensitivity to reward and punishment and neural activity during reward and avoidance learning. *Social Cognitive and Affective Neuroscience*, *10*(9), 1219-1227.
- King, A. J., Narraway, C., Hodgson, L., Weatherill, A., Sommer, V., & Sumner, S. (2011). Performance of human groups in social foraging: the role of communication in consensus decision making. *Biology Letters*, *7*(2), 237-240.
- King, J. A., Colla, M., Brass, M., Heuser, I., & von Cramon, D. Y. (2007). Inefficient cognitive control in adult ADHD: evidence from trial-by-trial Stroop test and cued task switching performance. *Behavioral and Brain Functions*, *3*, 1-19.
- Knoll, L. J., Leung, J. T., Foulkes, L., & Blakemore, S. J. (2017). Age-related differences in social influence on risk perception depend on the direction of influence. *Journal of Adolescence*, *60*, 53-63.

- Knoll, L. J., Magis-Weinberg, L., Speekenbrink, M., & Blakemore, S. J. (2015). Social influence on risk perception during adolescence. *Psychological Science, 26*(5), 583-592.
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience, 16*(8), 1412-1425.
- Kolling, N., & Akam, T. (2017). (Reinforcement?) Learning to forage optimally. *Current Opinion in Neurobiology, 46*, 162-169.
- Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science, 336*(6077), 95–98. <https://doi.org/10.1126/science.1216930>
- Kollins, S. H. (2003). Delay discounting is associated with substance use in college students. *Addictive Behaviors, 28*(6), 1167–1173. [https://doi.org/10.1016/S0306-4603\(02\)00220-4](https://doi.org/10.1016/S0306-4603(02)00220-4)
- Korn, C. W., & Bach, D. R. (2015). Maintaining homeostasis by decision-making. *PLoS Computational Biology, 11*(5), e1004301.
- Kraft, P. M., & Griffiths, T. (2018). Levels of Analysis in Computational Social Science. *CogSci*, 1-6.
- Ladouceur, C. D., Peper, J. S., Crone, E. A., & Dahl, R. E. (2012). White matter development in adolescence: the influence of puberty and implications for affective disorders. *Developmental Cognitive Neuroscience, 2*, 36-54.
- Larsen, B., & Luna, B. (2018). Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews, 94*, 179–195. <https://doi.org/10.1016/j.neubiorev.2018.09.005>

- Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299. <https://doi.org/10.1038/nn.4615>
- Le Heron, C., Kolling, N., Plant, O., Kienast, A., Janska, R., Ang, Y. S., ... & Apps, M. (2020). Dopamine modulates dynamic decision-making during foraging. *Journal of Neuroscience*, *40*(27), 5273-5282.
- Lee, A. J., DeBruine, L. M., & Jones, B. C. (2018). Individual-specific mortality is associated with how individuals evaluate future discounting decisions. *Proceedings of the Royal Society B*, *285*(1880), 20180304.
- Lee, L. T., Tsai, H. C., Chi, M. H., Chang, W. H., Chen, K. C., Lee, I. H., ... & Yang, Y. K. (2015). Lower availability of striatal dopamine transporter in generalized anxiety disorder: a preliminary two-ligand SPECT study. *International Clinical Psychopharmacology*, *30*(3), 175-178.
- Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4), 1-9.
- Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). *Journal of Experimental Psychology: Applied*, *75*–84.
- Lenow, J. K., Constantino, S. M., Daw, N. D., & Phelps, E. A. (2017). Chronic and Acute Stress Promote Overexploitation in Serial Decision Making. *The Journal of Neuroscience*, *37*(23), 5681–5689. <https://doi.org/10.1523/JNEUROSCI.3618-16.2017>

- Lenroot, R. K., & Giedd, J. N. (2006). Brain development in children and adolescents: insights from anatomical magnetic resonance imaging. *Neuroscience & Biobehavioral Reviews*, *30*(6), 718-729.
- Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. California, US: SAGE publications.
- Li, R. (2017). Flexing dual-systems models: How variable cognitive control in children informs our understanding of risk-taking across development. *Developmental Cognitive Neuroscience*, *27*, 91-98.
- Lindström, B., Bellander, M., Chang, A., Tobler, P. N., & Amodio, D. M. (2019). A computational reinforcement learning account of social media engagement. *Nature Communications*, *12*, 1-10.
- Lloyd, A., & Döring, A. K. (2019). When do peers influence adolescent males' risk taking? Examining decision making under conditions of risk and ambiguity. *Journal of Behavioral Decision Making*, *32*(5), 613-626.
- Lloyd, A., McKay, R., Hartman, T. K., Vincent, B. T., Murphy, J., Miller, J. G., ... & Mason, L. (2021a). Delay discounting and under-valuing of recent information predict lower adherence to social distancing measures during the COVID-19 pandemic. *Scientific Reports*, *11*, 19237. Doi: 10.1038/s41598-021-98772-5
- Lloyd, A., McKay, R., Sebastian, C. L., & Balsters, J. H. (2021b). Are adolescents more optimal decision-makers in novel environments? Examining the benefits of heightened exploration in a patch foraging paradigm. *Developmental Science*, e13075. <https://onlinelibrary.wiley.com/doi/full/10.1111/desc.13075>

- Lloyd, A., McKay, R., & Furl, N. (2022). Individuals with Adverse Childhood Experiences Explore Less and Underweight Reward Feedback. *Proceedings of the National Academy of Sciences* 119 (4), e2109373119. doi: 10.1073/pnas.2109373119
- Lockwood, P. L., & Klein-Flügge, M. C. (2021). Computational modelling of social cognition and behaviour—a reinforcement learning primer. *Social Cognitive and Affective Neuroscience*, 16(8), 761-771.
- Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a ‘social’ brain? Implementations and algorithms, *Trends in Cognitive Sciences*, 24(10), 802-813.
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, 76(3), 697-712.
- Luna, B., Padmanabhan, A., & O’Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence?. *Brain and Cognition*, 72, 101-113.
- Mallet, P., & Vignoli, E. (2007). Intensity seeking and novelty seeking: Their relationship to adolescent risk behavior and occupational interests. *Personality and Individual Differences*, 43(8), 2011-2021.
- Manning, C., Kilner, J., Neil, L., Karaminis, T., & Pellicano, E. (2017). Children on the autism spectrum update their behaviour in response to a volatile environment. *Developmental Science*, 20(5), e12435.
- Marques, J. C., Li, M., Schaak, D., Robson, D. N., & Li, J. M. (2019). Internal state dynamics shape brainwide activity and foraging behaviour. *Nature*, 1-5.
- Marr, D. (1977). Artificial intelligence—a personal view. *Artificial Intelligence*, 9, 37-48.

- Marr, D., & Poggio, T. (1976). From understanding computation to understanding neural circuitry. *Neurosci Res Program Bull.* 1-22
- Marsh, R., Zhu, H., Schultz, R. T., Quackenbush, G., Royal, J., Skudlarski, P., & Peterson, B. S. (2006). A developmental fMRI study of self-regulatory control. *Human Brain Mapping, 27*(11), 848-863.
- Marshall, A. T., Munson, C. N., Maidment, N. T., & Ostlund, S. B. (2020). Reward-predictive cues elicit excessive reward seeking in adolescent rats. *Developmental Cognitive Neuroscience, 45*, 100838.
- Marshall, H. H., Carter, A. J., Ashford, A., Rowcliffe, J. M., & Cowlshaw, G. (2013). How do foragers decide when to leave a patch? A test of alternative models under natural and experimental conditions. *Journal of Animal Ecology, 82*(4), 894-902.
- Martin, A. K., Dzafic, I., Ramdave, S., & Meinzer, M. (2017). Causal evidence for task-specific involvement of the dorsomedial prefrontal cortex in human social cognition. *Social Cognitive and Affective Neuroscience, 12*(8), 1209-1218.
- Maslowsky, J., Owotomo, O., Huntley, E. D., & Keating, D. (2019). Adolescent risk behavior: Differentiating reasoned and reactive risk-taking. *Journal of Youth and Adolescence, 48*(2), 243-255.
- Master, S. L., Eckstein, M. K., Gotlieb, N., Dahl, R., Wilbrecht, L., & Collins, A. G. (2020). Disentangling the systems contributing to changes in learning during adolescence. *Developmental Cognitive Neuroscience, 41*, 100732.
- Mata, R., Wilke, A., & Czienskowski, U. (2013). Foraging across the life span: is there a reduction in exploration with aging? *Frontiers in Neuroscience, 7*, 1-7.

- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean?. *American Psychologist*, *70*(6), 487-498.
- McCormick, E. M., & Telzer, E. H. (2017). Failure to retreat: Blunted sensitivity to negative feedback supports risky behavior in adolescents. *NeuroImage*, *147*, 381-389.
- McCormick, E. M., Peters, S., Crone, E. A., & Telzer, E. H. (2021). Longitudinal network re-organization across learning and development. *NeuroImage*, *229*, 117784.
- McCoy, S. S., Dimler, L. M., Samuels, D. V., & Natsuaki, M. N. (2019). Adolescent susceptibility to deviant peer pressure: Does gender matter?. *Adolescent Research Review*, *4*, 59-71.
- McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: a meta-analysis of functional neuroimaging data. *Frontiers in Human Neuroscience*, *11*, 154.
- McLaughlin, K. A., Kubzansky, L. D., Dunn, E. C., Waldinger, R., Vaillant, G., & Koenen, K. C. (2010). Childhood social environment, emotional reactivity to stress, and mood and anxiety disorders across the life course. *Depression and Anxiety*, *27*(12), 1087–1094. <https://doi.org/10.1002/da.20762>
- McLaughlin, K. A., & Sheridan, M. A. (2016). Beyond cumulative risk: A dimensional approach to childhood adversity. *Current Directions in Psychological Science*, *25*(4), 239-245.
- McLaughlin, K. A., Sheridan, M. A., & Lambert, H. K. (2014). Childhood adversity and neural development: deprivation and threat as distinct dimensions of early experience. *Neuroscience & Biobehavioral Reviews*, *47*, 578-591.

- McNamara, J. M., & Houston, A. I. (1985). Optimal foraging and learning. *Journal of Theoretical Biology*, *117*(2), 231–249. [https://doi.org/10.1016/S0022-5193\(85\)80219-8](https://doi.org/10.1016/S0022-5193(85)80219-8)
- Mehta, P. H., Welker, K. M., Zilioli, S., & Carré, J. M. (2015). Testosterone and cortisol jointly modulate risk-taking. *Psychoneuroendocrinology*, *56*, 88-99.
- Mersky, J. P., Topitzes, J., & Reynolds, A. J. (2013). 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 & Neglect*, *37*(11), 917–925. <https://doi.org/10.1016/j.chiabu.2013.07.011>
- Metzler, M., Merrick, M. T., Klevens, J., Ports, K. A., & Ford, D. C. (2017). Adverse childhood experiences and life opportunities: Shifting the narrative. *Children and Youth Services Review*, *72*, 141–149. <https://doi.org/10.1016/j.chilyouth.2016.10.021>
- Mills, K. L., Goddings, A. L., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). The developmental mismatch in structural brain maturation during adolescence. *Developmental Neuroscience*, *36*(3-4), 147-160.
- Mills, K. L., Goddings, A. L., Herting, M. M., Meuwese, R., Blakemore, S. J., Crone, E. A., ... & Tammes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *Neuroimage*, *141*, 273-281.
- Mills, K. L., Lalonde, F., Clasen, L. S., Giedd, J. N., & Blakemore, S. J. (2014). Developmental changes in the structure of the social brain in late childhood and adolescence. *Social Cognitive and Affective Neuroscience*, *9*, 123-131.

- Mills, K. L., Siegmund, K. D., Tamnes, C. K., Ferschmann, L., Bos, M. G., Wierenga, L. M., ... & Herting, M. L. (2021). Individual variability in structural brain development from late childhood to young adulthood. *BioRxiv*.
- Millstein, S. G., & Halpern–Felsher, B. L. (2002). Judgments about risk and perceived invulnerability in adolescents and young adults. *Journal of Research on Adolescence*, *12*(4), 399-422.
- Mithen, S. J. (1988). Modeling hunter-gatherer decision making: complementing optimal foraging theory. *Human Ecology*, *17*, 59-83.
- Miu, A. C., Bîlc, M. I., Bunea, I., & Szentágotai-Tătar, A. (2017). Childhood trauma and sensitivity to reward and punishment: Implications for depressive and anxiety symptoms. *Personality and Individual Differences*, *119*, 134–140.  
<https://doi.org/10.1016/j.paid.2017.07.015>
- Mobbs, D., Trimmer, P. C., Blumstein, D. T., & Dayan, P. (2018). Foraging for foundations in decision neuroscience: insights from ethology. *Nature Reviews Neuroscience*, *19*(7), 419-427.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review*, *100*(4), 674.
- Moffitt, T. E. (2018). Male antisocial behaviour in adolescence and beyond. *Nature Human Behaviour*, *2*(3), 177-186.
- Moreno, O. M., & Rosokha, Y. (2016). Learning under compound risk vs. learning under ambiguity—an experiment. *Journal of Risk and Uncertainty*, *53*(2-3), 137-162.

- Morris, L. S., Baek, K., Kundu, P., Harrison, N. A., Frank, M. J., Voon, V. (2016). Biases in the explore-exploit tradeoff in addictions: the role of avoidance of uncertainty. *Neuropsychopharmacology* **41**: 940–948.
- Moses-Payne, M. E., Habicht, J., Bowler, A., Steinbeis, N., & Hauser, T. U. (2021). I know better! Emerging metacognition allows adolescents to ignore false advice. *Developmental Science*, 1-13.
- Murty, V. P., Calabro, F., & Luna, B. (2016). The role of experience in adolescent cognitive development: Integration of executive, memory, and mesolimbic systems. *Neuroscience & Biobehavioral Reviews*, *70*, 46-58.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90-100.
- Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, *4*(1), 79-95.
- Nakagawa, S., Takeuchi, H., Taki, Y., Nouchi, R., Sekiguchi, A., Kotozaki, Y., ... & Kawashima, R. (2015). White matter structures associated with loneliness in young adults. *Scientific Reports*, *5*(1), 1-11.
- Nash, J. C. (1990). *Compact Numerical Methods for Computers. Linear Algebra and Function Minimisation*. Adam Hilger
- Nash, J. C. (2014). On Best Practice Optimization Methods in R. *Journal of Statistical Software*, *60*(2). <https://doi.org/10.18637/jss.v060.i02>
- Nash, J. C., & Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, *43*(9), 1-14.

- Nelder, J. A. and Mead, R. (1965). A simplex algorithm for function minimization. *Computer Journal*, *7*, 308–313. doi: 10.1093/comjnl/7.4.308.
- Nelson III, C. A., & Gabard-Durnam, L. J. (2020). Early Adversity and Critical Periods: Neurodevelopmental Consequences of Violating the Expectable Environment. *Trends in Neurosciences*, *43*(3), 133-143. <https://doi.org/10.1016/j.tins.2020.01.002>
- Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex*, *2*(3), 244-258.
- NHS England. (2015). *Safeguarding Policy*. Leeds: Nursing Directorate.
- Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, *12*(7), 265-272.
- Niv, Y., Joel, D., Meilijson, I., & Ruppin, E. (2002). Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422-1425.
- Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., ... & Vazire, S. (2019). Preregistration is hard, and worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815-818.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600-2606.
- Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*(6), 615-631.

- Novick, A. M., Levandowski, M. L., Laumann, L. E., Philip, N. S., Price, L. H., & Tyrka, A. R. (2018). The effects of early life stress on reward processing. *Journal of Psychiatric Research, 101*, 80–103. <https://doi.org/10.1016/j.jpsychires.2018.02.002>
- Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research?. *Developmental Cognitive Neuroscience, 40*, 100733.
- Nusslock, R., & Alloy, L. B. (2017). Reward processing and mood-related symptoms: An RDoC and translational neuroscience perspective. *Journal of Affective Disorders, 216*, 3-16.
- O'Brien, L., Albert, D., Chein, J., & Steinberg, L. (2011). Adolescents prefer more immediate rewards when in the presence of their peers. *Journal of Research on Adolescence, 21*(4), 747-753.
- Odum, A. L. (2011). Delay discounting: I'm a *k*, you're a *k*. *Journal of the Experimental Analysis of Behavior, 96*(3), 427-439.
- Office for National Statistics. (2009). Final recommended questions for the 2011 Census in England and Wales: Ethnic group
- Office of National Statistics. (2020). *Children Looked after in England (Including Adoptions and Care Leavers): 2019–2020*. <https://explore-education-statistics.service.gov.uk/find-statistics/children-looked-after-in-england-including-adoptions/2020>
- Olson, E. A., Hooper, C. J., Collins, P., & Luciana, M. (2007). Adolescents' performance on delay and probability discounting tasks: contributions of age, intelligence, executive

- functioning, and self-reported externalizing behavior. *Personality and Individual Differences*, 43(7), 1886-1897.
- Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, 3(2), 173-182.
- Orben, A., Tomova, L., & Blakemore, S. J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health*, 4(8), 634-640.
- Ordaz, S. J., Foran, W., Velanova, K., & Luna, B. (2013). Longitudinal growth curves of brain function underlying inhibitory control through adolescence. *Journal of Neuroscience*, 33(46), 18109-18124.
- Oshri, A., Hallowell, E., Liu, S., MacKillop, J., Galvan, A., Kogan, S. M., & Sweet, L. H. (2019). Socioeconomic hardship and delayed reward discounting: Associations with working memory and emotional reactivity. *Developmental Cognitive Neuroscience*, 37, 100642. <https://doi.org/10.1016/j.dcn.2019.100642>
- Osmont, A., Camarda, A., Habib, M., & Cassotti, M. (2021). Peers' Choices Influence Adolescent Risk-taking Especially When Explicit Risk Information is Lacking. *Journal of Research on Adolescence*, 31(2), 402-416.
- Osmont, A., Moutier, S., Simon, G., Bouhours, L., Houdé, O., & Cassotti, M. (2017). How does explicit versus implicit risk information influence adolescent risk-taking engagement?. *Journal of Behavioral Decision Making*, 30(5), 1093-1103.
- Ottenbreit, N. D., & Dobson, K. S. (2004). Avoidance and depression: the construction of the Cognitive–Behavioral Avoidance Scale. *Behaviour Research and Therapy*, 42(3), 293-313.

- Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *Journal of Cognitive Neuroscience*, *27*(2), 319-333.
- Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S. J. (2016). The computational development of reinforcement learning during adolescence. *PLoS Computational Biology*, *12*(6), e1004953.
- Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight* (No. 50). Oxford University Press.
- Patterson, M. W., Pivnick, L., Mann, F. D., Grotzinger, A. D., Monahan, K. C., Steinberg, L. D., ... & Harden, K. P. (2019). Positive risk-taking: Mixed-methods validation of a self-report scale and evidence for genetic links to personality and negative risk-taking. *PsyArXiv*, 1-45.
- Patton, G. C., Olsson, C. A., Skirbekk, V., Saffery, R., Wlodek, M. E., Azzopardi, P. S., ... & Sawyer, S. M. (2018). Adolescence and the next generation. *Nature*, *554*(7693), 458-466.
- Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768-774.
- Paulus, M. P., & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends in Cognitive Sciences*, *16*(9), 476-483.
- Pechtel, P., & Pizzagalli, D. A. (2013). 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*(5), 499. <https://doi.org/10.1001/jamapsychiatry.2013.728>

- Peeters, M., Oldehinkel, A., Veenstra, R., & Vollebergh, W. (2019). Unique developmental trajectories of risk behaviors in adolescence and associated outcomes in young adulthood. *PloS ONE*, *14*(11), e0225088.
- Pei, R., Lauharatanahirun, N., Cascio, C. N., O'Donnell, M. B., Shope, J. T., Simons-Morton, B. G., ... & Falk, E. B. (2020). Neural processes during adolescent risky decision making are associated with conformity to peer influence. *Developmental Cognitive Neuroscience*, *44*, 100794.
- Pelz, M., & Kidd, C. (2020). The elaboration of exploratory play. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190503.
- Peper, J. S., Braams, B. R., Blankenstein, N. E., Bos, M. G., & Crone, E. A. (2018). Development of multifaceted risk taking and the relations to sex steroid hormones: A longitudinal study. *Child Development*, *89*(5), 1887-1907.
- Pfeifer, J. H., & Allen, N. B. (2012). Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. *Trends in Cognitive Sciences*, *16*(6), 322-329.
- Phillips, J. G., Evans, M., Hughes, B., & Ogeil, R. P. (2020). Patterns of Cannabis Consumption, Social Networks, and Foraging. *Journal of Drug Issues*, *50*(1), 63-76.
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643.
- Pirolli, P., & Fu, W. T. (2003). SNIF-ACT: A model of information foraging on the World Wide Web. In *International Conference on User Modeling* (pp. 45-54). Springer, Berlin, Heidelberg.
- Popper, K. R. (1963). Science as falsification. *Conjectures and Refutations*, *1*(1963), 33-39.

- Potard, C., Courtois, R., & Rusch, E. (2008). The influence of peers on risky sexual behaviour during adolescence. *The European Journal of Contraception & Reproductive Health Care*, *13*(3), 264-270.
- Potter, M. H., & Font, S. A. (2019). Parenting influences on adolescent sexual risk-taking: Differences by child welfare placement status. *Children and Youth Services Review*, *96*, 134-144.
- Prokopy, R. J., & Roitberg, B. D. (1984). Foraging behavior of true fruit flies: concepts of foraging can be used to determine how tephritids search for food, mates, and egg-laying sites and to help control these pests. *American Scientist*, *72*(1), 41-49.
- Protzko, J., & Schooler, J. W. (2019). Kids these days: Why the youth of today seem lacking. *Science Advances*, *5*(10), eaav5916.
- Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*, *23*(10), 865–875. <https://doi.org/10.1016/j.tics.2019.07.007>
- Pyke, G. H. (1978). Optimal foraging in hummingbirds: testing the marginal value theorem. *American Zoologist*, *18*(4), 739-752.
- Qi, S., Schumann, G., Bustillo, J., Turner, J. A., Jiang, R., Zhi, D., ... & Sui, J. (2021). Reward Processing in Novelty Seekers: A Transdiagnostic Psychiatric Imaging Biomarker. *Biological Psychiatry*.
- Qu, Y., Galvan, A., Fuligni, A. J., Lieberman, M. D., & Telzer, E. H. (2015). Longitudinal changes in prefrontal cortex activation underlie declines in adolescent risk taking. *Journal of Neuroscience*, *35*(32), 11308-11314.

- R Core Team (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL <http://www.R-project.org/>
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raine, N. E., & Chittka, L. (2007). The adaptive significance of sensory bias in a foraging context: floral colour preferences in the bumblebee *Bombus terrestris*. *PLoS ONE*, 2(6), e556.
- Rapee, R. M., Oar, E. L., Johnco, C. J., Forbes, M. K., Fardouly, J., Magson, N. R., & Richardson, C. E. (2019). Adolescent development and risk for the onset of social-emotional disorders: A review and conceptual model. *Behaviour Research and Therapy*, 123, 103501.
- Reiter, A. M., Moutoussis, M., Vanes, L., Kievit, R., Bullmore, E. T., Goodyer, I. M., ... & Dolan, R. J. (2021). Preference uncertainty accounts for developmental effects on susceptibility to peer influence in adolescence. *Nature Communications*, 12(1), 1-13.
- Reiter, A. M., Suzuki, S., O'Doherty, J. P., Li, S. C., & Eppinger, B. (2019). Risk contagion by peers affects learning and decision-making in adolescents. *Journal of Experimental Psychology: General*, 148(9), 1494.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Proktsy (Eds.), *Classical conditioning 1/: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts

- Reynolds, E. K., MacPherson, L., Schwartz, S., Fox, N. A., & Lejuez, C. W. (2014). Analogue study of peer influence on risk-taking behavior in older adolescents. *Prevention Science, 15*(6), 842-849.
- Robb, C. E., de Jager, C. A., Ahmadi-Abhari, S., Giannakopoulou, P., Udeh-Momoh, C., McKeand, J., ... & Middleton, L. (2020). Associations of social isolation with anxiety and depression during the early COVID-19 pandemic: a survey of older adults in London, UK. *Frontiers in Psychiatry, 11*.
- Romer, D. (2010). Adolescent risk taking, impulsivity, and brain development: Implications for prevention. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 52*(3), 263-276.
- Romer, D., Betancourt, L. M., Brodsky, N. L., Giannetta, J. M., Yang, W., & Hurt, H. (2011). Does adolescent risk taking imply weak executive function? A prospective study of relations between working memory performance, impulsivity, and risk taking in early adolescence. *Developmental Science, 14*(5), 1119-1133.
- Romer, D., Reyna, V. F., & Satterthwaite, T. D. (2017). Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental Cognitive Neuroscience, 27*, 19-34.
- Rosenbaum, G. M., Botdorf, M. A., Patrianakos, J. L., Steinberg, L., & Chein, J. M. (2017). Working memory training in adolescents decreases laboratory risk taking in the presence of peers. *Journal of Cognitive Enhancement, 1*(4), 513-525.
- Ross, L., & Hill, E. (2002). Childhood Unpredictability, Schemas for Unpredictability, and Risk Taking. *Social Behavior and Personality: An International Journal, 30*, 453-473.  
<https://doi.org/10.2224/sbp.2002.30.5.453>

- Sadeghiyeh, H. (2019). Social influence, temporal discounting and active-passive gap in explore-exploit dilemma. *PhD Thesis: University of Arizona*.
- Sadeghiyeh, H., Wang, S., Alberhasky, M. R., Kylo, H. M., Shenhav, A., & Wilson, R. C. (2020). Temporal discounting correlates with directed exploration but not with random exploration. *Scientific Reports*, *10*(1), 4020. <https://doi.org/10.1038/s41598-020-60576-4>
- Said-Mohamed, R., Prioreshi, A., Nyati, L. H., van Heerden, A., Munthali, R. J., Kahn, K., ... & Norris, S. A. (2018). Rural–urban variations in age at menarche, adult height, leg-length and abdominal adiposity in black South African women in transitioning South Africa. *Annals of Human Biology*, *45*(2), 123-132.
- Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health*, *2*(3), 223-228.
- Schreuders, E., Braams, B. R., Blankenstein, N. E., Peper, J. S., Güroğlu, B., & Crone, E. A. (2018). Contributions of reward sensitivity to ventral striatum activity across adolescence and early adulthood. *Child Development*, *89*(3), 797-810.
- Schulz, E., Franklin, N. T., & Gershman, S. J. (2020). Finding structure in multi-armed bandits. *Cognitive Psychology*, *119*, 101261.
- Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H., Kronbichler, M., & Friston, K. J. (2019). Computational mechanisms of curiosity and goal-directed exploration. *Elife*, *8*, e41703.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461-464.

- Sebastian, C. L., Tan, G. C., Roiser, J. P., Viding, E., Dumontheil, I., & Blakemore, S. J. (2011). Developmental influences on the neural bases of responses to social rejection: implications of social neuroscience for education. *Neuroimage*, *57*(3), 686-694.
- Sebastian, C., Viding, E., Williams, K. D., & Blakemore, S. J. (2010). Social brain development and the affective consequences of ostracism in adolescence. *Brain and Cognition*, *72*(1), 134-145.
- Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R., ... & Gray, J. R. (2008). Individual differences in delay discounting: relation to intelligence, working memory, and anterior prefrontal cortex. *Psychological Science*, *19*(9), 904-911.
- Sharp, P. B., Do, K. T., Lindquist, K. A., Prinstein, M. J., & Telzer, E. H. (2021). Cognitive control deployment is flexibly modulated by social value in early adolescence. *Developmental Science*, e13140.
- Shaw, A. M., Timpano, K. R., Tran, T. B., & Joormann, J. (2015). Correlates of Facebook usage patterns: The relationship between passive Facebook use, social anxiety symptoms, and brooding. *Computers in Human Behavior*, *48*, 575-580.
- Shaw, P., Kabani, N. J., Lerch, J. P., Eckstrand, K., Lenroot, R., Gogtay, N., ... & Wise, S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *Journal of Neuroscience*, *28*(14), 3586-3594.
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240.

- Shenhav, A., Straccia, M. A., Cohen, J. D., & Botvinick, M. M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience*, *17*(9), 1249–1254. <https://doi.org/10.1038/nn.3771>
- Sheridan, M. A., & McLaughlin, K. A. (2014). Dimensions of early experience and neural development: Deprivation and threat. *Trends in Cognitive Sciences*, *18*(11), 580–585. <https://doi.org/10.1016/j.tics.2014.09.001>
- Shevlin, M., McBride, O., Murphy, J., Miller, J. G., Hartman, T. K., Levita, L., ... & Bentall, R. P. (2020). Anxiety, depression, traumatic stress and COVID-19-related anxiety in the UK general population during the COVID-19 pandemic. *BJPsych Open*, *6*(6).
- Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*, 487-510.
- Shulman, E. P., Smith, A. R., Silva, K., Icenogle, G., Duell, N., Chein, J., & Steinberg, L. (2016). The dual systems model: Review, reappraisal, and reaffirmation. *Developmental Cognitive Neuroscience*, *17*, 103-117.
- Siegel, S., & Allan, L. G. (1996). The widespread influence of the Rescorla-Wagner model. *Psychonomic Bulletin & Review*, *3*(3), 314-321.
- Silva, K., Shulman, E. P., Chein, J., & Steinberg, L. (2016). Peers increase late adolescents' exploratory behavior and sensitivity to positive and negative feedback. *Journal of Research on Adolescence*, *26*(4), 696-705.
- Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence*, 103535.

- Silvers, J. A., McRae, K., Gabrieli, J. D., Gross, J. J., Remy, K. A., & Ochsner, K. N. (2012). Age-related differences in emotional reactivity, regulation, and rejection sensitivity in adolescence. *Emotion, 12*(6), 1235.
- Simons-Morton, B., Lerner, N., & Singer, J. (2005). The observed effects of teenage passengers on the risky driving behavior of teenage drivers. *Accident Analysis & Prevention, 37*(6), 973-982.
- Smaldino, P. E. (2020). How to translate a verbal theory into a formal model. *Social Psychology, 51*(4), 207–218
- Smith, A. R., Chein, J., & Steinberg, L. (2014). Peers increase adolescent risk taking even when the probabilities of negative outcomes are known. *Developmental Psychology, 50*(5), 1564.
- Smith, D. G., Xiao, L., & Bechara, A. (2012). Decision making in children and adolescents: Impaired Iowa Gambling Task performance in early adolescence. *Developmental Psychology, 48*(4), 1180.
- Smith, E. A. (1983). Anthropological applications of optimal foraging theory: A critical review. *Current Anthropology, 24*(5), 625-651.
- Smith, E. A., Bettinger, R. L., Bishop, C. A., Blundell, V., Cashdan, E., Casimir, M. J., ... & Stini, W. A. (1983). Anthropological applications of optimal foraging theory: a critical review [and comments and reply]. *Current Anthropology, 24*(5), 625-651.
- Smith, R., Taylor, S., Wilson, R. C., Chuning, A. E., Persich, M., Wang, S., & Killgore, W. (2021). Lower levels of directed exploration and reflective thinking are associated with greater anxiety and depression. *PsyArXiv*. <https://doi.org/10.31234/osf.io/3w4je>

- Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, *20*(10), 635-644.
- Somerville, L. H., Hare, T., Casey, B. J. (2011). Frontostriatal maturation predicts cognitive control failure to appetitive cues in adolescents. *J. Cogn. Neurosci.* *23*(9):2123–34
- Somerville, L. H., Haddara, N., Sasse, S. F., Skwara, A. C., Moran, J. M., & Figner, B. (2019). Dissecting “peer presence” and “decisions” to deepen understanding of peer influence on adolescent risky choice. *Child Development*, *90*(6), 2086-2103.
- Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition*, *72*(1), 124-133.
- Somerville, L. H., Sasse, S. F., Garrad, M. C., Drysdale, A. T., Abi Akar, N., Insel, C., & Wilson, R. C. (2017). Charting the expansion of strategic exploratory behavior during adolescence. *Journal of Experimental Psychology: General*, *146*(2), 155-164.
- Song, M., Bnaya, Z., & Ma, W. J. (2019). Sources of suboptimality in a minimalistic explore–exploit task. *Nature Human Behaviour*, *3*(4), 361-368.
- Spear, L. P. (2000). Neurobehavioral changes in adolescence. *Current Directions in Psychological Science*, *9*(4), 111-114.
- Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience & Biobehavioral Reviews*, *24*(4), 417-463.
- Spear, L. P. (2011). Rewards, aversions and affect in adolescence: emerging convergences across laboratory animal and human data. *Developmental Cognitive Neuroscience*, *1*(4), 390-403.

- Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in Cognitive Science*, 7(2), 351-367.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092-1097.
- Spivey, J., Barrett, D., Padilla, E., & Gonzalez-Lima, F. (2008). Mother–infant separation leads to hypoactive behavior in adolescent Holtzman rats. *Behavioural Processes*, 79(1), 59–65. <https://doi.org/10.1016/j.beproc.2008.05.002>
- Spoormaker, V. I., Andrade, K. C., Schröter, M. S., Sturm, A., Goya-Maldonado, R., Sämann, P. G., & Czisch, M. (2011). The neural correlates of negative prediction error signaling in human fear conditioning. *Neuroimage*, 54(3), 2250-2256.
- Steenbergen, M. R. (2006). Maximum likelihood programming in R. *University of North Carolina, Chapel Hill*.
- Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9(2), 69-74.
- Steinberg, L. (2007). Risk taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science*, 16(2), 55-59.
- Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, 28(1), 78-106.
- Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 52(3), 216-224.

- Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age differences in sensation seeking and impulsivity as indexed by behavior and self-report: evidence for a dual systems model. *Developmental Psychology*, *44*(6), 1764.
- Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development*, *80*(1), 28-44.
- Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., ... & Fanti, K. A. (2018). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*, *21*(2), e12532.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton University Press: Princeton, NJ
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Swedler, D. I., Bowman, S. M., & Baker, S. P. (2012). Gender and age differences among teen drivers in fatal crashes. In *Annals of Advances in Automotive Medicine/Annual Scientific Conference* (Vol. 56, p. 97). Association for the Advancement of Automotive Medicine.
- Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(10), 1231-1238.
- Tammes, C. K., & Mills, K. L. (2020). Imaging structural brain development in childhood and adolescence. In D Poeppel, GR Mangun, MS Gazzaniga (eds.), *The Cognitive Neurosciences VI* (pp.17-25). MIT Press.

- Tamnes CK, Ostby Y, Fjell AM, Westlye LT, Due-Tønnessen P, Walhovd KB (2010): Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb Cortex* 20: 534–548.
- Taswell, C. A., Costa, V. D., Murray, E. A., & Averbeck, B. B. (2018). Ventral striatum's role in learning from gains and losses. *Proceedings of the National Academy of Sciences*, 115(52), E12398–E12406. <https://doi.org/10.1073/pnas.1809833115>
- Telzer, E. H. (2016). Dopaminergic reward sensitivity can promote adolescent health: A new perspective on the mechanism of ventral striatum activation. *Developmental Cognitive Neuroscience*, 17, 57-67.
- Timberlake, W., Gawley, D. J., & Lucas, G. A. (1987). Time horizons in rats foraging for food in temporally separated patches. *Journal of Experimental Psychology: Animal Behavior Processes*, 13(3), 302.
- Tottenham, N., & Sheridan, M. A. (2010). A review of adversity, the amygdala and the hippocampus: a consideration of developmental timing. *Frontiers in Human Neuroscience*, 3, 68.
- Tottenham, N., Hare, T. A., Millner, A., Gilhooly, T., Zevin, J. D., & Casey, B. J. (2011). Elevated amygdala response to faces following early deprivation. *Developmental Science*, 14(2), 190-204.
- Turrin, C., Fagan, N. A., Dal Monte, O., & Chang, S. W. (2017). Social resource foraging is guided by the principles of the Marginal Value Theorem. *Scientific Reports*, 7(1), 1-13.

- Tymula, A. (2019). An experimental study of adolescent behavior under peer observation: Adolescents are more impatient and inconsistent, not more risk-taking, when observed by peers. *Journal of Economic Behavior & Organization*, *166*, 735-750.
- Tymula, A., Belmaker, L. A. R., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences*, *109*(42), 17135-17140.
- Urošević, S., Collins, P., Muetzel, R., Lim, K., & Luciana, M. (2012). Longitudinal changes in behavioral approach system sensitivity and brain structures involved in reward processing during adolescence. *Developmental Psychology*, *48*(5), 1488.
- van den Bos, W., & Eppinger, B. (2016). Developing developmental cognitive neuroscience: From agenda setting to hypothesis testing. *Developmental Cognitive Neuroscience*, *17*, 138.
- van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports*, *7*, 40962, 1-11.
- van den Bos, W., Bruckner, R., Nassar, M. R., Mata, R., & Eppinger, B. (2018). Computational neuroscience across the lifespan: Promises and pitfalls. *Developmental Cognitive Neuroscience*, *33*, 42-53.
- van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum–medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral cortex*, *22*(6), 1247-1255.
- van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *Proceedings of the National Academy of Sciences*, *112*(29), E3765-E3774.

- van der Meer, M. A., & Redish, A. D. (2011). Ventral striatum: a critical look at models of learning and evaluation. *Current Opinion in Neurobiology*, *21*(3), 387-392.
- van Duijvenvoorde, A. C., Jansen, B. R., Bredman, J. C., & Huizenga, H. M. (2012). Age-related changes in decision making: Comparing informed and noninformed situations. *Developmental Psychology*, *48*(1), 192.
- van Duijvenvoorde, A. C., de Macks, Z. A. O., Overgaauw, S., Moor, B. G., Dahl, R. E., & Crone, E. A. (2014). A cross-sectional and longitudinal analysis of reward-related brain activation: effects of age, pubertal stage, and reward sensitivity. *Brain and cognition*, *89*, 3-14.
- van Duijvenvoorde, A. C., Peters, S., Braams, B. R., & Crone, E. A. (2016). What motivates adolescents? Neural responses to rewards and their influence on adolescents' risk taking, learning, and cognitive control. *Neuroscience & Biobehavioral Reviews*, *70*, 135-147.
- van Hoorn, J., Crone, E. A., & Van Leijenhorst, L. (2017). Hanging out with the right crowd: Peer influence on risk-taking behavior in adolescence. *Journal of Research on Adolescence*, *27*(1), 189-200.
- van Leijenhorst, L., & Crone, E. A. (2009). Paradoxes in adolescent risk taking. *Developmental Social Cognitive Neuroscience*, 209-225.
- van Leijenhorst, L., Moor, B. G., de Macks, Z. A. O., Rombouts, S. A., Westenberg, P. M., & Crone, E. A. (2010a). Adolescent risky decision-making: neurocognitive development of reward and control regions. *Neuroimage*, *51*(1), 345-355.

- Van Leijenhorst, L., Westenberg, P. M., & Crone, E. A. (2008). A developmental study of risky decisions on the cake gambling task: age and gender analyses of probability estimation and reward evaluation. *Developmental Neuropsychology*, *33*(2), 179-196.
- van Leijenhorst, L., Zanolie, K., Van Meel, C. S., Westenberg, P. M., Rombouts, S. A., & Crone, E. A. (2010b). What motivates the adolescent? Brain regions mediating reward sensitivity across adolescence. *Cerebral Cortex*, *20*(1), 61-69.
- van Rooij, I., & Baggio, G. (2020). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science*, 1745691620970604.
- van Tieghem, M. R., & Tottenham, N. (2017). Neurobiological programming of early life stress: functional development of amygdala-prefrontal circuitry and vulnerability for stress-related psychopathology. In *Behavioral Neurobiology of PTSD* (pp. 117-136). Springer, Cham.
- Venkataraman, V. V., Kraft, T. S., Dominy, N. J., & Endicott, K. M. (2017). Hunter-gatherer residential mobility and the marginal value of rainforest patches. *Proceedings of the National Academy of Sciences*, *114*(12), 3097-3102.
- Vidal-Ribas, P., Benson, B., Vitale, A. D., Keren, H., Harrewijn, A., Fox, N. A., ... & Stringaris, A. (2019). Bidirectional associations between stress and reward processing in children and adolescents: a longitudinal neuroimaging study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(10), 893-901.
- Wahlstrom, D., Collins, P., White, T., & Luciana, M. (2010). Developmental changes in dopamine neurotransmission in adolescence: behavioral implications and issues in assessment. *Brain and Cognition*, *72*(1), 146-159.

- Wajnberg, E., Fauvergue, X., & Pons, O. (2000). Patch leaving decision rules and the Marginal Value Theorem: an experimental analysis and a simulation model. *Behavioral Ecology*, *11*(6), 577-586.
- Walker, R., Hill, K., Kaplan, H., & McMillan, G. (2002). Age-dependency in hunting ability among the Ache of Eastern Paraguay. *Journal of Human Evolution*, *42*(6), 639-657.
- Weller, J. A., Parker, A. M., Reynolds, M., & Kirisci, L. (2021). Developmental trajectory classes in psychological dysregulation predict later decision-making competence. *Addictive Behaviors*, *112*, 106650.
- Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., Van Aert, R., & Van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*, 1832.
- Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *Neuroimage* *96*, 67–72
- Willoughby, T., Good, M., Adachi, P. J., Hamza, C., & Tavernier, R. (2014). Examining the link between adolescent brain development and risk taking from a social–developmental perspective. *Brain and Cognition*, *89*, 70-78.
- Willoughby, T., Heffer, T., Good, M., & Magnacca, C. (2021). Is adolescence a time of heightened risk taking? An overview of types of risk-taking behaviors across age groups. *Developmental Review*, *61*, 100980.
- Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *Elife*, *8*, e49547.

- Wilson, R. C., Bonawitz, E., Costa, V. D., & Ebitz, R. B. (2021). Balancing exploration and exploitation with information and randomization. *Current Opinion in Behavioral Sciences*, 38, 49-56.
- Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, 143(6), 2074.
- Winet, Y. K., Tu, Y., Choshen-Hillel, S., & Fishbach, A. (2020). Social exploration: When people deviate from options explored by others. *Journal of Personality and Social Psychology*.
- Wu, S., Sun, S., Camilleri, J. A., Eickhoff, S. B., & Yu, R. (2021). Better the devil you know than the devil you don't: Neural processing of risk and ambiguity. *NeuroImage*, 236, 118109.
- Xia, L., Master, S., Eckstein, M., Wilbrecht, L., & Collins, A. G. E. (2020). Learning under uncertainty changes during adolescence. In *Proceedings of the Cognitive Science Society*.
- Zacharopoulos, G., Shenhav, A., Constantino, S., Maio, G. R., & Linden, D. E. (2018). The effect of self-focus on personal and social foraging behaviour. *Social Cognitive and Affective Neuroscience*, 13(9), 967-975.