Supplementary Online Materials

To accompany

Charting the expansion of strategic exploratory behavior during adolescence

Somerville, Sasse, Garrad, Drysdale, Abi Akar, Insel, & Wilson

Contents

I. Sample characteristics and age predictors
II. Analyses to validate developmental comparison
III. Descriptive statistics for DOSPERT questionnaire
IV. Relationship between directed and random exploration
V. Computational modeling results
VI. Sex differences in exploration
VII. Task instructions and comprehension questions
I. Sample characteristics and age predictors

The usable sample consisted of $N = 43$ 12-14 year olds, $N = 46$ 15-17 year olds, $N = 34$ 18-20 year olds, and $N = 24$ 21-28 year olds, with male and female participants equally distributed across the age span (Chi-squared test of equal age distributions: $\chi^2 = 3.84$, $p = 0.28$; see Figure S1).

![Figure S1. Sample characterized by age and sex.](image)

As in our prior work (Somerville et al., 2013; van Duijvenvoorde et al., 2015), linear and nonlinear differences were considered. Based on the age range of data, two age patterns were interrogated: linear (monotonic change with age, created by mean-centering raw age), and adolescent-emergent, modeled as a cubic function centered at 24 years of age and retaining the maximum value for further ages that was subsequently mean-centered (see Figure S2). For most adult participants, age was available only as an integer value. These age predictors are collinear, and thus were submitted to model comparison rather than being tested head-to-head.
in the same statistical models. As reported in the main text, the adolescent emergent function was the best-fitting function to the key test of age differences in directed exploration. Thus, the emergent function was used for all analyses that test for age differences across the full sample.

Figure S2: Age functions used in statistical analysis. (Left) mean-centered linear age, (right) transformed function reflecting rapid change through adolescence that asymptotes in early adulthood. Y-axes are arbitrary units. Each dot represents a participant (with some participants overlapping); red: fit line.

II. Analyses to validate developmental comparison

Several analyses were conducted on data orthogonal to the primary variables reported in the main manuscript. The goal was to evaluate potential developmental confounds that could influence task performance through mechanisms other than exploration strategy. Indices of points earned, reward maximization, and the quality of computational model fits were evaluated across the age group to address whether any of these factors co-varied with primary analyses.

A. Overall success in earning points. Analyses were conducted to determine whether individuals across the age span exhibited differential overall success toward the goal of winning points in the task. The mean number of points earned per choice (mean = 47.83 +/- sd = 1.26) did not
vary by age \((F(1,145) = 1.45, p = 0.149)\). This finding suggests that participants regardless of age comprehended the superordinate goal of the task. More importantly, this finding indicates that despite differences in strategic exploration with age, these differences yielded similar levels of success in the overall goal of the task. Thus, the strategic differences observed across age should not be interpreted as differentially optimal.

**B. Response time.** Participants on average took 1.562 seconds (+/- sd 0.380) to respond during the first free choice. Response times were marginally longer for the first horizon_6 free choice (0.596s +/- 0.461) compared to horizon_1 (0.529 +/- 0.406; \(F(1,145) = 3.712, p = 0.056\)). This lengthened decision time is consistent with the intention that horizon_6 provoke more elaborate decisional strategies. However, the greater visual complexity inherent to the horizon_6 display (which always displayed more boxes on the screen) could have also contributed to slower processing times.

Consistent with a host of developmental studies (Kail, 2007; Somerville et al., 2011), younger aged individuals took significantly longer to make choices than older individuals (main effect of age on grand mean reaction time: \(F(1,145) = 9.68, p = 0.002\)). However, there was no interaction between horizon and age on response time (\(F(1,145) = 0.001, p = 0.973\)). Thus, although younger participants demonstrated a baseline shift toward longer responses, this tendency did not manifest differentially depending on horizon. This finding increases confidence that the key variables of directed and random exploration, which are calculated as a difference between horizon_6 > horizon_1, are not contaminated by developmental “time on task” confounds.

**C. Reward maximization behavior when exploration is mitigated.** Age differences in choice data were examined for the sixth free choice of horizon_6 games. This choice is orthogonal to the
key dependent variables reported in the main manuscript, which only reflect the first free choice. By the sixth choice of horizon_6 games, the higher mean bandit is likely to be obvious because the trial history contains 9 values, and motivation to explore should be minimized because there are no further choices on which to utilize the information gained by exploration. Participants generally opted for the high mean option on these choices (average high mean choices = 83.6% +/- 0.70). This final choice was analyzed to identify whether participants of all ages exhibited comparable reward maximization behavior.

A significant positive relationship was observed between age and the propensity to select the high mean option on the sixth horizon_6 choice (F(1,145) = 14.95, p<0.001). This large effect was driven by N = 9 statistical outliers 12-19 years of age who exhibited much lower reward maximization tendencies (47.5%-63% compared to group mean of 83.6%). If these individuals are excluded from the analysis, there remains a reduced but nonetheless significant effect of age on reward maximization behavior F(1,136) = 5.05, p = 0.026. Thus, there is a small but significant tendency to reward-maximize more with greater age in the late stages of horizon_6 games.

The significance of the main findings in the manuscript remained unchanged if the N = 9 participants were excluded, and when statistically controlling for the tendency to reward-maximize. Although these findings might suggest that the N = 9 participants’ choices did not fully reflect the goals of the task, they could have been operating under a different exploratory strategy that, while different from the other participants, is not inherently invalid. Thus, to remain conservative with regard to the range of acceptable strategies, we have opted to retain these participants in all analyses reported in the main manuscript.

D. Computational modeling fit quality. Computational modeling analyses (presented in the next section) were tested for age confounds in overall quality of fit, which could preclude
interpretation of developmental effects (Hartley & Somerville, 2015; van den Bos et al., 2012).
The overall quality of model fit, quantified with Bayesian Information Criterion (BIC; mean =
185.17 +/- sd = 49.95), did not vary by age (r(146) = -0.093, p = 0.262). The consistency of fit
quality over age validates the integrity of direct comparison of age effects on model-derived
exploratory strategies.

III. Descriptive statistics for DOSPERT questionnaire.
Developmental variants of the Domain Specific Risk Taking questionnaire (Blais & Weber, 2006)
were used to assess risk attitudes. Participants aged 12-13 completed the Child version and
participants 14-17 completed the Adolescent version. One 18 year old also completed the
Adolescent version; omitting the 18 year old did not affect the significance of the results and
thus their data were included in all analyses.

Participants completed the questionnaires on a digital notepad which encouraged open
disclosure of tendency to take risks as experimenters could not view their responses until they
were downloaded at a later date. Minor participants and their parents were informed that study
responses would not be shared with parents. All participants completed the Risk Taking
subscale, N = 1 participant did not complete the Expected Benefits subscale, and N = 2
participants did not complete the Risk Perceptions subscale. Complete subscale data were
retained in analyses when possible.

Psychometric validations of the Child and Adolescent versions of the DOSPERT scale
have not been published yet, although preliminary analyses have been presented which support
the validity of the questionnaires with an alpha = 0.88 in a sample of N = 448 (Figner et al., 2015
and laboratory of Elke Weber, personal communication). Data from the present sample indicate
good reliability (see Table S1) and a sufficient range of scores for each of the three subscales to
assess individual differences.
As predicted by the DOSPERT framework, Risk Taking was associated with reduced Risk Perceptions ($r(85) = -0.412$, $p<0.000080$), and Risk Taking was associated with greater Expected Benefits ($r(86) = 0.474$, $p<0.000004$). The relationship between Risk Perceptions and Expected Benefits was not significant ($r(85) = -0.182$, $p = 0.096$). Within this constrained age range (from 12-17 years), age and Risk Taking were not correlated ($p = 0.657$), and age and Expected Benefits were not correlated ($p = 0.70$). Greater age was associated with greater Risk Perception ($r(84) = 0.217$, $p = 0.045$).

### IV. Relationship between directed and random exploration.

Because direct and random exploration measures were derived from distinct sets of trials, their scores are mathematically independent and thus can be standardized and directly compared. To determine whether participants of different ages relied more strongly on one exploration strategy over another, an ANOVA was conducted with exploration strategy [z-scored directed, z-scored random] as a within subjects variable, and age as a continuous between subjects covariate. This analysis yielded a significant age by exploration strategy interaction ($F(1,145) = 6.46$, $p = 0.012$) such that greater age was positively associated with greater utilization of directed over random exploration. While the younger participants (<18 years) invoked an

<table>
<thead>
<tr>
<th></th>
<th>Cronbach α</th>
<th>Descriptive statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risk taking</td>
<td>Risk perception</td>
</tr>
<tr>
<td><strong>Child version</strong></td>
<td>0.836</td>
<td>0.879</td>
</tr>
<tr>
<td><strong>Adolescent version</strong></td>
<td>0.860</td>
<td>0.885</td>
</tr>
</tbody>
</table>

Table S1: Reliability and descriptive data on DOSPERT questionnaires. Left columns: Cronbach α reliability for the three subscales of the Child and Adolescent versions of the DOSPERT. Right columns: Descriptive statistics for the three subscales of the Child and Adolescent versions of the DOSPERT. Row 1: Mean (Standard deviation); Row 2: Minimum-Maximum.
equivalent utilization of directed and random exploration ($z_{direct} - z_{random} = -0.230$), older participants (>18) showed a greater utilization from directed over random exploration ($z_{direct} - z_{random} = 0.352$). There was also a positive correlation between tendency to use directed and random exploration strategies ($r(146) = 0.223$, $p = 0.007$) that persisted when controlling for age ($r(144) = 0.209$, $p = 0.012$), which can be thought of as reflecting global attunement to decision horizon.

V. Computational Modeling Results.

To complement our model-free analysis in the main paper, we conducted a separate model-based analysis using a simple logistic model previously described in Wilson et al. (2014). The results of this analysis are in line with those reported in the main text.

A. Model. As with the model-free analysis, the focus of the model is on the first free-choice trial in each game. In particular, we assume that the probability of choosing the left bandit is a function of the difference in the observed means for the two options, $\Delta \mu = \mu_{left} - \mu_{right}$ and the difference in information available for each option $\Delta I = I_{left} - I_{right}$ (which, for simplicity is defined as +1 when playing the left bandit is more informative and -1 when the right bandit is more informative). Thus,

$$p_1 = \frac{1}{1 + \exp \left( -\frac{\Delta \mu + A \Delta I + B}{\sqrt{2} \sigma} \right)}$$

Where $A$ is the information bonus that quantifies directed exploration, $B$ is the spatial bias and $\sigma$ is the standard deviation of the decision noise that quantifies random exploration.
Using this equation, we fit the parameters $A$, $B$ and $\sigma$ for each participant in each of the four horizon x information conditions using a maximum \textit{a posteriori} approach. The solution sought parameters that maximized the posterior probability of the parameters given the choices, $c_{1:t}$, i.e.

$$p(A, B, \sigma | c_{1:t}) \propto p(c_{1:t} | A, B, \sigma)p(A)p(B)p(\sigma)$$

where $p(A)$, $p(B)$ and $p(\sigma)$ are priors over the information bonus, spatial bias and decision noise. In particular, we assumed a Gaussian prior with mean zero and standard deviation of 20 for the information bonus, $A$, a uniform prior for $B$, and an exponential prior with scale parameter of 20 for the decision noise, $\sigma$. These priors help to keep the parameters in a reasonable range, but have little bearing on the main result.

**B. Strategic directed exploration increases with age.** After fitting the information bonus, $A$, spatial bias, $B$, and decision noise, $\sigma$, for each condition and each subject, we performed a similar analysis to that in the main text to determine whether directed (quantified by $A$) or random (quantified by $\sigma$) varied with age. For directed exploration, a repeated measures ANOVA with emerging age and horizon as factors found main effects of horizon ($F(1, 145) = 92.4, p < 2 \times 10^{-16}$) and emerging age ($F(1, 145) = 7.87, p = 0.0057$). In line with our model-free analysis we also found a strong interaction between emerging age and horizon ($F(1, 145) = 7.34, p = 0.0076$). Post hoc analysis revealed that this interaction was primary driven by change in horizon 6 ($t(146) = 3.25, p = 0.0014$) with age, while the information bonus in horizon 1 was independent of age ($t(146) = 0.90, p = 0.37$). These findings are summarized in Figure S3.
**Figure S3. Strategic use of directed exploration increases from adolescence to young adulthood (N=147).** A-B. Age differences in probability of selecting the bandit with less information available onscreen (y-axis; information bonus) for horizon_1 (A) and horizon_6 (B) trials. C. Age differences in strategic directed exploration, defined as information bonus for horizon_6 > horizon_1. With increasing age (x-axis) there are no differences in information bonus for horizon_1 games (A) and an increase in information bonus in horizon_6 games (B) leading to a rise in strategic directed exploration through adolescence that stabilizes in young adulthood (C). D-G: Average information bonus (y-axis) for the different age groups as a function of horizon (x-axis). Age shifts are primarily evident for horizon_6 decisions. Error bars represent standard error of the mean.

For random exploration, as with the model-free analysis we found no relationship between emerging age and decision noise. Because the distribution of decision noise across participants was non-Gaussian (for the [2 2] condition, skew = 3.0 for horizon_1 and 1.7 for horizon_6, kurtosis = 12.5 for horizon_1 and 5.6 for horizon_6; for the [1 3] condition, skew = 3.0 for horizon_1 and 1.9 for horizon_6, kurtosis = 13.7 for horizon_1 and 8.2 for horizon_6) we focused our analysis on the square root transform of decision noise (for the [2 2] condition, skew = 1.1 for horizon_1 and 0.75 for horizon_6, kurtosis = 4.9 for horizon_1 and 3.7 for horizon_6;
for the [1 3] condition, skew = 1.1 for horizon_1 and 0.67 for horizon_6, kurtosis = 5.7 for horizon_1 and 3.7 for horizon_6). Using this square root transform, for the [2 2] condition we found only a main effect of horizon on decision noise (F(1,145) = 173, p < 2×10^{-16}) and neither a main effect of emerging age (F(1,145) = 1.6, p = 0.21) nor an interaction between emerging age and horizon (F(1,145) = 0.007, p = 0.935).

In addition to replicating our results from the main paper for the [2 2] condition, the model allows us to quantify decision noise in the [1 3] condition, allowing us an independent check of the findings. This analysis yielded almost identical results: main effect of horizon (F(1,145) = 80.5, p = 1.4×10^{-15}), no main effect of emerging age (F(1,145) = 1.74, p = 0.19) and no interaction (F(1,145) = 0.57, p = 0.45). These findings are summarized in Figure S4.

Figure S4. Random exploration strategies do not differ from adolescence to early adulthood. A-B. Age differences in decision noise (y-axis; square root of decision noise) in the equal information condition. Age does not modulate decision noise in horizon_1 (A) or horizon_6 games (B). C. Age differences in random exploration, defined as decision noise for horizon_6 > horizon_1. Plotting by age (x-axis) indicates stability of strategic random exploration through
adolescence and young adulthood. D-F Similar results also hold for random exploration in the unequal information condition. Age does not modulate decision noise in the unequal condition in horizon_1 (D) or horizon 6 (E) and there is no difference in strategic random exploration with age (F).

**C. Random exploration correlates with risk taking.** As in the main paper, we used linear regression to quantify the relationship between risk attitudes (as measured by the DOSPERT questionnaire) and exploration. In particular, we assumed that each of the three risk measurements (risk taking, risk perception and expected benefit) was linearly related to the information bonus in both horizons, square root decision noise in both horizons and emerging age. Using this analysis we found significant associations between risk seeking and expected benefit measures and random exploration, such that there was a negative association in horizon_1 and a positive association in horizon_6 (for risk taking: $\beta$(decision noise, horizon_1) = -0.37, $p = 0.0013$; $\beta$(decision noise, horizon_6) = 0.34, $p = 0.005$; for expected benefit: $\beta$(decision noise, horizon_1) = -0.33, $p = 0.002$; $\beta$(decision noise, horizon_6) = 0.25, $p = 0.02$). There was no significant association with risk perception ($\beta$(decision noise, horizon_1) = -0.07, $p = 0.55$; $\beta$(decision noise, horizon_6) = 0.12, $p = 0.37$). This suggests that random exploration, defined as the change in decision noise between horizons correlates with risk attitudes (for risk taking: Spearman’s $\rho$(87) = 0.31, $p = 0.0036$; risk perception: Spearman’s $\rho$(87) = 0.03, $p = 0.80$; expected benefit: Spearman’s $\rho$(87) = 0.19, $p = 0.086$). However, the correlations for risk perception and expected benefit are driven almost entirely by a single subject who had especially low scores for risk perception and expected benefit. When this subject was removed from the analysis, only the risk taking result remained significant (for risk taking: Spearman’s $\rho$(86) = 0.29, $p = 0.0065$; risk perception: Spearman’s $\rho$(86) = -0.008, $p = 0.94$; expected benefit: Spearman’s $\rho$(86) = 0.16, $p = 0.15$). Likewise, when this subject was removed from the regression analysis, only risk taking had significant associations with decision noise in horizons
1 and 6 (for risk taking: $\beta$(decision noise, horizon_1) = -0.36, $p = 0.0035$; $\beta$(decision noise, horizon_6) = 0.34, $p = 0.0055$).

**Figure S5.** Regression analysis showing the relationship between the DOSPERT measures of risk attitude and the different types of exploration and age. Only random exploration correlates significantly with risk taking. * $p<0.05$. Error bars represent standard error of the mean.

**VI. Sex differences in strategic exploration.**

Because males and females mature along different trajectories, we evaluated whether participant sex moderated strategic exploration or interacted with age to predict strategic exploration. To do so, all critical statistical tests from the main manuscript were recomputed with sex as a between subjects variable included as a moderator of the observed effects. As reported below, sex did not explain variance in our exploratory measures nor did it moderate the reported findings.

There were no overall differences between males and females on measures of strategic exploration (male vs female strategic directed exploration: $t(145) = 0.167, p = 0.867$; male vs female strategic random exploration: $t(145) = 0.124, p = 0.902$). Age x sex interactions were also not significant for strategic directed exploration ($F(1,143) = 1.24, p = 0.267$) or random
exploration (F(1,143) = 0.590, p = 0.444). Analysis of conflict trials (see main manuscript; Figure 3) did not yield a significant effect of sex (main effect of sex: F(1,143) = 1.427, p = 0.234); or an age x sex interaction (F(1,143) = 1.015, p = 0.315).

For analyses associating directed and random exploration with risk taking and risk attitudes (as measured by the DOPSRERT), sex did not interact with directed or random exploration to predict risk taking or risk attitudes (all p’s >0.3). Further, male and female participants endorsed equivalent levels of risk taking (sex differences on risk taking subscale t(86) = 0.124, p = 0.901), risk perception (sex differences on risk perceptions subscale t(84) = 0.08, p = 0.937), and expected benefits of risk taking (sex differences on expected benefits subscale t(85) = 0.482, p = 0.631).

To summarize, participant sex did not modulate any of the findings reported in the main manuscript. Furthermore, all of the findings reported retained their level of significance when participant sex was accounted for in statistical tests.

VII. Task instructions and comprehension questions.

Instructions were delivered by computer, with matching visuals. Text is reproduced below.

Please contact the authors for the fully programmed version.

INSTRUCTIONS FOR ADULTS 18+ YEARS OF AGE

Welcome! Thank you for volunteering for this experiment.

In this experiment we would like you to choose between two one-armed bandits of the sort you might find in a casino. The one-armed bandits will be represented like this.

Every time you choose to play a particular bandit, the lever will be pulled like this ...
... and the payoff will be shown like this. For example, in this case, the left bandit has been played and is paying out 77 points.

Each bandit tends to pay out about the same amount of reward on average, but there is variability in the reward on any given play.

For example, the average reward for the bandit on the right might be 50 points, but on the first play we might see a reward of 52 points because of the variability ...

... on the second play we might get 56 points ...

... if we open a third box on the right we might get 45 points this time ...

... and so on, such that if we were to play the right bandit 10 times in a row we might see these rewards ...

Both bandits will have the same kind of variability and this variability will stay constant throughout the experiment.

One of the bandits will always have a higher average reward and hence is the better option to choose on average.

To make your choice: Press < to play the left bandit. Press > play the right bandit

On any trial you can only play one bandit and the number of trials in each game is determined by the height of the bandits. For example, when the bandits are 10 boxes high, there are 10 trials in each game ...

... when the bandits are 5 boxes high there are only 5 trials in the game.

Finally, the first 4 choices in each game are instructed trials where we will tell you which option to play. This will give you some experience with each option before you make your first choice.
These instructed trials will be indicated by a green square inside the box we want you to open and you must press the button to choose this option in order to move on to see the reward and move on the next trial. For example, if you are instructed to choose the left box on the first trial, you will see this: If you are instructed to choose the right box on the second trial, you will see this:

Once these instructed trials are complete you will have a free choice between the two bandits that is indicated by two green squares inside the two boxes you are choosing between.

Press space when you are ready to begin. Good luck!

**INSTRUCTIONS FOR MINORS 17- YEARS OF AGE**

*Note: For minors, the visual display was described as stacks of boxes rather than bandits/slot machines due to likely unfamiliarity with the operation of slot machines.*

Welcome! Thank you for volunteering for this experiment.

In this experiment you will see stimuli like these:

Each of these represents a stack of boxes, one stack on the left and one stack on the right. Inside each of the boxes is a reward of between 1 and 100 points.

During this experiment you will be able to open some of the boxes and receive the reward inside. This reward is translated into real money at the end of the experiment so to maximize your earnings you need to get the most points possible.

For example, in this case, a box from the left stack has been opened and is paying out 77 points. The XX denotes that the corresponding box in the right stack was not opened.

The boxes in each stack tend to pay out about the same amount of reward on average, but there is variability in the reward of an individual box.

For example, the average reward for the stack on the right might be 50 points, but if we
were to open the first box we might see a reward of 52 points, if we open a second box on the right we might get 56 points, if we open a third box on the right we might get 45 points this time and so on, if we open all the boxes on the right we might see these rewards. One of the stacks will always have a higher average reward and is the better option to choose on average.

To make your choice: Press < to open a box from the left stack. Press > to open a box from the right stack.

On any trial you can only open one box and the number of trials in each game is determined by the height of the stack. For example, when the stacks are 10 boxes high, there are 10 trials in each game when the stacks are 5 boxes high there are only 5 trials in the game.

Finally, for the first 4 choices in each game we will tell you which option to play. This will give you some experience with each option before you make your first choice.

These instructed trials will be shown by a green square inside the box we want you to open and you must press the button to choose this option in order to move on to see the reward and move on the next trial. For example, if you are instructed to choose the left box on the first trial, you will see this:

If you are instructed to choose the right box on the first trial, you will see this:

Once these instructed trials are complete you will have a free choice between the two stacks that is indicated by two green squares inside the two boxes you are choosing between.

Press space when you are ready to begin.  Good luck!

Beginning block
Press space to begin.

Well done! You averaged " points!

Press space to continue

You earned $

**COMPREHENSION QUESTIONS**

These questions were administered verbally to participants. Experimenters corrected and re-instructed participants who did not provide a reasonable answer to an item.

1. When it’s your turn to choose, what choice are you making?
   - *Left or right sets of boxes/bandits*

2. What keys are you going to use?
   - *Left arrow for the left side, right arrow for the right side*

3. If you see a box with the number 40 in it what does that tell you about what numbers are in the other boxes in that stack?
   - *They will be around 40.*

4. If you see a green box only on the left side, that means you have to choose that side. But what does it mean when the green is on both sides?
   - *My turn to pick between left side and right side.*

5. How are you going to decide what side to pick?
   - *Reference to goals of task: to get as many points as possible, and/or decide based on the information on each side’s boxes.*
References


