Journal of Experimental Psychology: General

Manuscript version of

Charting the Expansion of Strategic Exploratory Behavior During Adolescence

Leah H. Somerville, Stephanie F. Sasse, Megan C. Garrad, Andrew T. Drysdale, Nadine Abi Akar, Catherine Insel, Robert C. Wilson

Funded by:
- American Psychological Association
- National Science Foundation

Charting the expansion of strategic exploratory behavior during adolescence

Leah H. Somerville[1], Stephanie F. Sasse[1], Megan C. Garrad[1], Andrew T. Drysdale[2], Nadine Abi Akar[1], Catherine Insel[1] & Robert C. Wilson[3]

[1]Department of Psychology & Center for Brain Science, Harvard University, Northwest Building, 52 Oxford Street, Cambridge MA, USA

[2]Department of Psychiatry, Weill Cornell Medical College, 1300 York Avenue, New York NY USA

[3]Department of Psychology & Cognitive Science Program, University of Arizona, 503 E University Blvd., Tuscon AZ USA

Word count: 5,926

Corresponding Author:
Leah Somerville
Harvard University
52 Oxford Street, Room 290
Cambridge MA 02138
somerville@fas.harvard.edu
Phone: 617-495-7513

Abstract

While models of exploratory decision-making implicate a suite of strategies that guide the pursuit of information, the developmental emergence of these strategies remains poorly understood. This study takes an interdisciplinary perspective merging computational decision making and developmental approaches to characterize age-related shifts in exploratory strategy from adolescence to young adulthood. 149 participants aged 12-28 completed a computational explore-exploit paradigm that manipulated reward value, information value, and decision horizon (the utility information holds for future choices). Strategic directed exploration, defined as information seeking selective for long time horizons, emerged during adolescence and maintained its level through early adulthood. This age difference was partially driven by adolescents valuing immediate reward over new information. Strategic random exploration, defined as stochastic choice behavior selective for long time horizons, was invoked at comparable levels over the age range, and predicted individual differences in propensity for risk taking in daily life within the adolescent portion of the sample. Collectively, these findings reveal an expansion of the diversity of strategic exploration over development, implicate distinct mechanisms for directed and random exploratory strategies, and suggest novel mechanisms for adolescent-typical shifts in decision-making.

Keywords: adolescence; decision making; exploration; reward; development

During adolescence, there is a mounting demand to make self-directed decisions in an increasingly complex and uncertain environment. Some of the proximal decision dilemmas adolescents face require weighing the value of new options against those that are better known– inviting a new classmate or an old friend as a date to prom; trying a new drug or sipping a beer at a party; attending college in an unfamiliar city or in one's hometown. These dilemmas represent exemplars of the classic explore-exploit problem, deciding between an unknown option that could be better or worse than a known option. Solving exploration-exploitation dilemmas thus requires weighing the value of novel options that bring new information (i.e., exploration) relative to options with known value (i.e., exploitation) (Sutton & Barto, 1998). The present study characterizes the developmental emergence of strategies used to solve the explore-exploit dilemma from adolescence to adulthood. Moreover, because exploratory behavior has been conceptually linked to risk taking in adolescence (e.g., (Crone & Dahl, 2012)), we evaluate whether the use of particular exploratory strategies relate to adolescents' propensity to endorse risky behaviors in daily life.

Recent work suggests that people make explore-exploit decisions using at least two distinct strategies (Wilson et al. 2014). One strategy, *directed exploration*, is guided by assessment of the value of information that would be gained from choosing a particular option (Auer et al., 2002; Gittins, 1979). Directed exploration would guide an individual to select options presently lacking in information, for example choosing to take an alternative route to work over a very familiar one, even if the familiar route is quite fast. When implemented correctly, i.e. when information is valued appropriately, directed exploration can be optimal in the sense of maximizing reward over time. However, computing the correct value of information can be difficult, and performance can suffer when the value of options is over- or underestimated.

A second strategy, *random exploration*, represents a stochastic decision process that entails randomly selecting among options varying in value and information (Bridle, 1990; Thompson, 1933). Random exploration would guide an individual to select between options without reference to the information or value to be gained, for example flipping a coin to decide between two commute options despite being more familiar with one route, and despite one route being faster based on past experience. Although potentially less optimal than directed exploration, random exploration works well in practice and requires less precise tuning than directed exploration (Sutton & Barto, 1998; Watkins, 1989).

When used strategically, directed and random exploration should be modulated by the amount of time one has to utilize information in the future, the *horizon*. When the horizon is long, information has value because there is time to use it in the future, making exploration advantageous. When the horizon is short, information has no value and it is advantageous to exploit known options instead based on their value. Following this intuition, we have recently shown that adults increase directed and random exploration in long time horizons (Wilson et al., 2014).

Although work on non-human animals suggests strong exploratory motivations during adolescence (Adriani et al., 1998; Macri et al., 2002; Spear, 2000), work to date has not assessed developmental changes in the strategies of exploration. On one hand, the capacity for complex cognitive operations continues to refine during adolescence (Casey et al., 2005; Petersen, 1988), which could constrain the implementation of more complex forms of decision-making (Hartley & Somerville, 2015) such as directed exploration. Further, the neural systems critical to exploratory behavior, including dopaminergic, noradrenergic and frontal systems (Aston-Jones et al., 2000; Costa et al., 2014; Frank et al., 2009), show anatomical (Andersen et al., 2000; Brain_Development_Cooperative_Group, 2012) and functional (Braams et al., in press; Galvan et al., 2006) perturbations during adolescence. Nascent work in other domains of decision-making indicates that adolescents exhibit shifts in decision computations such as ambiguity tolerance (Blankenstein et al., 2016; Tymula et al., 2012), risk tolerance (Defoe et al., 2015), intertemporal choice (Steinberg et al., 2009; van den Bos et al., 2015), and reward sensitivity (Braams et al., in press; Steinberg, 2004). We predicted that these shifting features of adolescent motivation and decision processes would shape strategic exploration.

The goal of the present study was to characterize shifts in strategic exploratory behavior from early adolescence to early adulthood. In addition to examining developmental differences in these exploratory strategies, we tested whether the use of directed and random exploration strategies related to adolescents' endorsement of risk taking behaviors. Adolescent risk taking is thought to emerge, in part, from a desire to explore novel and arousing experiences (Steinberg, 2004). Thus, we reasoned that differential utilization of exploratory strategies might relate to greater willingness to endorse risk in service of novelty and/or potential reward. To address these questions, we took advantage of advances in computational approaches to explore-exploit decision making which are optimized for quantification of directed and random exploration strategies absent of reward confounds (Wilson et al., 2014). Adolescent participants also completed a self-report measure that assessed their endorsement of age-specific risky behaviors. Analyses quantified developmental shifts in utilization of each form strategic exploration and evaluated whether such shifts related to a tendency to endorse real-life risky behavior.

Materials and Methods

Sample.

N=149 participants aged 12.08 to 28.0 took part in this study. The sample size was selected based on Wilson et al. (2014), which documented evidence of strategic directed and random exploration in a sample of adults. The sample size was then doubled to allow sufficient power for developmental comparison, and the developmental subgroups were further bolstered in size by an additional ~third to accommodate the inherent rise in behavioral variability in developmental data. Children were not included in this study because of concerns with task comprehension.

Two participants were subsequently excluded from analysis: N=1 adolescent for withdrawing participation early, and N=1 adult for not following instructions. The final sample of N=147 usable participants consisted of N=65 males and N=82 females, distributed equivalently across the age range (see Figure S1 and Supplemental Online Materials).

All participants provided informed written consent and all minor participants received written permission for their participation from a parent or legal guardian. Research procedures were approved by the Committee on the Use of Human Subjects at Harvard University and by the Princeton University Institutional Review Board. The data from a subset of adult participants have been published previously (Wilson et al., 2014) and were repurposed for novel comparisons of age-related changes in task performance. These participants completed a longer version of the task and their data files were truncated to match the 160 game version that the developmental sample completed.

Horizon Task.

<u>Conceptual Overview.</u> During the Horizon Task (Wilson et al., 2014), participants made a series of choices between two one-armed bandits (i.e., slot machines) that paid out variable point values (see Figure 1). By selecting a bandit, participants saw only the points awarded by that bandit and not the other. In each game, lasting 5 or 10 choices, the computer determined the first four selections. These fixed choices manipulated quantity of information participants had about each bandit and the differential number of points the two bandits paid out. The dependent variable was the participant's first free choice following the four fixed choices.

Three task manipulations were imposed across games to reveal different strategies for exploratory behavior. The differential amount of *reward* (i.e., points) paid by each bandit manipulated the advantageousness of exploiting one bandit over the other. The differential amount of *information* available about the bandits' reward history manipulated the value of exploration. Selecting a bandit with fewer previous payouts displayed (i.e., "high information choices") would result in a proportionally greater boost in information gained (i.e., the exploratory choice), compared to selecting a bandit with more payouts displayed. The differential decision *horizon* (i.e. the number of choices in each game) manipulated the advantageousness of exploration, with exploration being more advantageous in long decision horizons because there is an opportunity to exploit that information in subsequent choices.

Using these manipulations, the following summary measures of exploratory strategy could be computed:

a) Directed Exploration was calculated using the unequal information games for which one option showed one previous trial payout and the other bandit showed three previous trial payouts (i.e., [1 3]; Figure 1A). Directed exploration was defined as selecting the option with one previous payout displayed, i.e., [1] the *high information choice*, even if it has the lower mean payout history. *Strategic* directed exploration was defined as selecting the high information choice option more frequently for horizon_6 games relative to horizon_1 games. A high value indicated a strategy of exploration for which participants were selectively motivated to accrue information about the option with less available information, moreso when subsequent exploration was possible.

b) Random Exploration was calculated using games for which both options showed two previous trial payouts (i.e, [2 2]; Figure 1B). As there is no difference in information between the two options in the equal condition, the optimal choice is simply to choose the option with highest mean payout history. Choosing the low mean option was therefore assumed to be caused by stochasticity in the decision process and was used as a measure of random exploration. *Strategic* random exploration was defined as selecting the low mean option more frequently for horizon_6 games relative to horizon_1 games. A high value would indicate more behavioral variability when subsequent exploration was possible.

Task instructions. Participants received animated, computerized instructions for the task (see Supplemental Material available online). For adults, the task was described as a bandit/slot machine task similar to what is in casinos. Because youths might not be familiar with slot machines, the stimuli were instead described to minors as stacks of boxes. Instructions for minors were also more elaborate than for adults, but were otherwise highly similar.

Participants were instructed that on a given trial they would see two slot machines/stacks of boxes. On each turn, participants selected between the two options (left or right), and then received the number of points displayed in that machine round/box. Participants were shown a series of fixed choices first, which revealed payout history information about the two options before the participant's turn to pick. When it was the participant's turn to pick, they selected the left or right option with a key press. The goal was to win as many points as possible, which would be translated into a cash bonus at the conclusion of the study. Participants could use any strategy they wished to select between the machines/stacks, including paying attention to the payout history visible for each bandit. While the payouts varied somewhat within a bandit, the payout history served as a reasonable proxy for a subsequent payout. The number of boxes on the screen indicated whether the participant would have one turn or six turns to select between the two sides (i.e., the horizon manipulation). Minor participants completed a comprehension test to ensure that the premise and parameters of the task were clear.

Game structure. In each game, the two bandits' payouts hovered around a different mean value, such that one was more advantageous on average, while the relative advantage varied across games (i.e., *reward differential*). The mean of one bandit was set to either 40 or 60 points and the mean of the other was set relative to the mean of the first, such that the difference between the two was sampled from 4, 8, 12, 20 and 30. Payouts were sampled (rounded to the nearest integer) from a Gaussian distribution with a fixed standard deviation of 8 points. Both the identity and the difference in means were counterbalanced over the entire experiment.

Games proceeded in two phases: fixed choice and free choice (see Figure 1). Each game began with four fixed choices that manipulated the reward differential and quantity of information about the bandits before the first free choice. For the fixed choices, a green square illuminated inside of one bandit, indicating participants were required to select that bandit by clicking on it. Allowing participants to click the fixed-choice selections (rather than simply displaying them on the screen) was intended to heighten engagement in the task and induce a sense of learning about the bandits from the participant's own experience (Hertwig et al., 2004).

The fixed choices also manipulated how much information was available about each bandit. For half of the games, the four fixed choices were split as two payouts per bandit (i.e., *equal information*; [2 2; Figure 1B) and for the other half of games, the participant viewed one payout for one bandit and three payouts for the other (i.e., *unequal information;* [1 3]; Figure 1A). Throughout a given game, choice and outcome history remained onscreen inside each of the bandits. After a particular option was played, the points added to the visual display, while the corresponding space for the unplayed option was filled with an 'XX'.

Following the fixed choices, participants engaged in the free choice phase where they freely selected between the two bandits (Figure 1C). For half of the games, participants made one free choice (i.e., horizon_1), and for the other half participants made six free choices (i.e., horizon_6). Participants were aware of the number of free choices at the outset of the game based on the visual display. An ideal decision maker should explore more on the first free choice in horizon_6 than horizon_1 because participants have the opportunity to use the information gained by exploring on later choices. Thus, the horizon manipulation was implemented to compare strategic exploration as a function of utility using a horizon_6 > horizon_1 comparison.

Participants completed a total of 160 self-paced games in randomized order. Games were counterbalanced on information, reward amount, and whether the left or right side of the screen depicted the higher points-mean bandit.

Data analysis.

Data analyses were carried out in R (R Core Team, 2014), Matlab 2014b (Natick, MA), and IBM SPSS Statistics for Macintosh, Version 22.0 (Armonk, NY). To evaluate whether participants' tendency to use directed and random exploration strategies varied as a function of age, we conducted two streams of analyses: one that analyzed choice in its native form, and one that submitted each participant's choice data to a logistic computational model (as in (Wilson et al., 2014)) that summarized information seeking and behavioral variability in participants' choice data. Results of both sets of analyses were highly convergent. Results from the native data are reported in the main manuscript while computational modeling data are presented in the Supplemental Material available online.

The dependent measure in every game was the participant's bandit selection on the first free choice. Additional choices in the horizon_6 games were examined as manipulation checks, but they were not used in primary analyses because reward and information rapidly become correlated over subsequent choices (Wilson et al., 2014). By contrast, reward and information manipulations were designed to be orthogonal on the first free choice.

Primary analyses examined rates of strategic directed and strategic random exploration irrespective of the reward differential between bandits. Taking the reward differential into account quantitatively using a computational model (see Supplemental Material available online) and qualitatively by inspecting choice curves leads to convergent conclusions and thus are presented descriptively and used in targeted posthoc analyses. The computational model allows more precise quantification of strategic directed and strategic random exploration fully dissociated from reward differential and spatial bias (e.g. a preference for choosing left) that are not taken into account with the model-free metrics.

Analysis of age differences.

As in our prior work, age was invoked as a continuous predictor of developmental differences to maximize statistical power and to mitigate the need to create semi-arbitrary boundaries between age groups (Somerville et al., 2013). Analyses of age differences tested two *a priori* models of potential age differences: linear (monotonic change with age) and adolescent-emergent (rapid change through adolescence that asymptotes into adulthood; see Figure S2). The difference between models concerns whether age differences minimized into early adulthood presumably because the underlying processes have asymptoted developmentally (e.g., emergent predictor) rendering adolescence a period of rapid change, or whether age differences continue steadily throughout the entire age range (e.g., linear predictor). To adjudicate between the age models, model comparison was conducted on the age effect on directed exploration by comparing Bayesian Information Criterion (BIC) values between the models. The emergent age model had lower BIC values (Linear: BIC: 2232.4; Emergent: BIC: 2230.0) indicating a better model fit, suggesting that the processes of interest are changing rapidly during adolescence and stabilizing into early adulthood. Therefore, the emergent model was used for all age-related analyses.

Additional analyses evaluated choice markers of task comprehension and data comparability across the age range. Results indicated that participants across the full age range displayed behavior consistent with comprehension of the task, and that task data were not subject to problematic confounds with age. See Supplemental Material available online for the methods and results of these analyses.

Measures of risk taking.

Participants younger than 18 and one 18 year old completed the Child or Adolescent version of the Domain Specific Risk Taking questionnaire (DOSPERT) questionnaire (Blais & Weber, 2006). Participants younger than 14 completed the Child version, and participants 14-18 years old completed the Adolescent version. The DOSPERT is a three-part questionnaire containing scenarios describing a range of age-appropriate risks. Example items from the Child version include, "*Walking home alone after dark,"* and *"Climbing up a very high tree".* Example items from the Adolescent scale include, "*Skateboarding down a steep hill*," and "*Speaking out against an unpopular opinion at school*." Multivariate responses to each scenario were obtained to characterize three related features of participants' risk attitudes: Risk Taking (*how likely you would be to engage in that activity*), Risk Perception (*how risky you feel the activity is)*, and Expected Benefits *(how much you would benefit from engaging in that activity).* Within this sample, Cronbach's alpha reliability indices indicate strong reliability (all scales >0.8; see Supplemental Material available online).

Results

Baseline evidence of exploration. One possible strategy for performing the Horizon Task would be to select the higher payout bandit for every choice – solely reward-maximizing and never engaging in exploration. Although participants could have adopted this strategy, their choice behavior suggests that they did not. For equal information [2 2] games, participants exhibited a significant non-zero rate of selecting the bandit with the lower mean payout history (horizon_1: 9.94% +/- 10.92%; $t(146)=11.04$, $p<0.001$; horizon_6: 20.90% +/- 13.35; $t(146)=18.98$, $p<0.001$). For unequal information [1 3] games, participants selected the high information option even when it also had a lower mean payout history (horizon_1: 9.77% +/- 13.15; $t(146)=9.01$, $p<0.001$; horizon_6: 27.29% +/- 21.34; $t(146)=15.51$; $p<0.001$). These findings indicate that the task successfully evoked exploratory motivations for all trial types.

Strategic use of directed exploration emerges in adolescence. As in prior work (Wilson et al., 2014), participants invoked a behavioral pattern consistent with directed exploration in [1 3] unequal information games, selecting the more informative option more often in horizon_6 than horizon_1 ($F(1,145)=24.60$, $p<0.001$; $\eta^2=0.145$; high information choices horizon_1=53% +/- 9.84%; horizon_6=58.81% +/- 11.15%). This main effect was qualified by a significant age x horizon interaction ($F(1,145)=18.85$, $p<0.001$; $\eta^2=0.115$; Figure 2C). Posthoc contrasts indicated that with increasing age there was both an emerging tendency to seek information selectively during games in which it could be used to inform subsequent choices ($t(146)=2.68$, $p=0.008$; $R^2=0.047$; Figure 2B), and an overall reduction of information seeking on games for which information cannot be used to inform subsequent choices ($t(146)=-2.55$, $p=.012$; $R^2=0.043$; Figure 2A). Overall selection of the high-information option did not differ by age (main effect of age: $F(1,145)=0.08$, $p>0.250$)), suggesting that overall levels of directed exploration were age-invariant.

Choice curves in Figure 2D-2G show the tendency to select the information-maximizing option as a function of reward differential between the two bandits and horizon for four age subgroupings. Participants of all ages exhibit a tendency to exploit the high reward option when it is also the high information option (right half of each choice curve) regardless of horizon. The primary age differences in directed exploration occurred during *reward-information conflict* games – when one bandit contained higher reward value and the other contained higher information value (left half of each choice curve). A posthoc analysis of age differences in conflict games yielded a significant age by horizon interaction ($F(1,145)=5.40$, $p=0.022$; $\eta^2=0.036$). In horizon_1, participants across ages were uniformly likely to exploit the high reward option. However, in horizon_6, increasing age relates to an increasing tendency to forego the high reward option in favor of the lower value yet higher information option (see Figure 3).

Strategic use of random exploration is age-invariant from adolescence to adulthood. Analyses of the [2 2] equal information games indicated that participants, on average, invoked a behavioral pattern consistent with random exploration, selecting the lower mean option more in horizon_6 than horizon_1 ($F(1,145)=170.56$, $p<0.001$; $\eta^2=0.541$). Selection of the low-mean option did not differ by age (main effect of age: $F(1,145)=0.976$, $p>0.250$), suggesting that overall levels of random exploration were age-invariant. The age by horizon interaction was not significant ($F(1,145)=0.976$, $p>0.250$; Figure 3A-C), suggesting that strategic random exploration was also age-invariant.

Choice curve data (Figure 4D-G) reveal the consistency of random exploration across four age subgroups. The choice curves show the tendency to select the reward-maximizing option as a function of differential reward value of the two bandits and horizon. Overall, participants of all ages exploit reward maximization in horizon_1 choices (blue lines). However, in horizon_6 conditions (red lines), participants demonstrate less reward maximization (i.e., flatter curves). These patterns were consistent across the four age subgroups.

Relationship between exploration strategy and risky behavior. N=86 participants younger than 18 years and one 18 year old completed measures of daily life risk taking and risk attitudes using the DOSPERT questionnaire (Blais & Weber, 2006). Participants in the sample endorsed a wide range of risky behavior and risk attitudes (see Supplemental Material available online for descriptive statistics). A multiple linear regression treated risk taking scores as a dependent measure, inputting directed and random exploration as independent variables. This analysis yielded a significant relationship between tendency to use strategic random exploration and risk taking ($\beta=0.168$; $F(1,85)=4.287$, $p=0.0415$; 95% confidence interval (CI) [0.007 0.330]; same analysis statistically controlling for age $p=0.0422$). Directed exploration and random exploration were not significantly related to Risk Perceptions or Expected Benefits scores. This finding suggests that adolescents who tend to utilize random exploration strategically in the Horizon

Task are more likely to endorse greater propensity for risk taking in daily life. Computational based indices of random exploration yield highly convergent findings (see Supplemental Material available online).

Follow-up analyses tested for relationships between risk taking and exploration in horizons 1 and 6 separately. Results indicated that risk taking is significantly related to a tendency to engage in random exploration less in horizon_1 ($t(85)=-2.193$, $p=0.031$; CI=[-0.040, -0.002), and a trend-level tendency to engage in random exploration more in horizon_6 ($t(85)=1.696$, $p=0.094$; CI=[-0.002 0.030]); see Figure 5.

Analysis of sex differences in exploratory behavior and risk taking. Additional analyses tested whether participant sex or age by sex interactions explained additional variance in strategic exploratory behavior. Results indicated that participant sex did not explain variance in strategic exploratory behavior or risk measures, nor did the primary findings change with inclusion of sex in statistical models. Thus, the findings reported here are equivalently applicable to males and females. The full analyses are presented in Supplemental Online Materials.

Discussion

A burgeoning challenge of adolescence is to engage in self-guided decisions that involve weighing the value of known and unknown options. Here we applied a decision-theoretic approach founded on the explore-exploit dilemma to chart the development of strategic exploratory behaviors. While overall levels of exploration did not vary with age, the *strategic* use of directed exploration showed robust changes from adolescence into adulthood. These findings reveal mechanisms that contribute to unique facets of adolescents' decision-making.

Previous research has suggested that adolescents engage in more exploration than adults (Nunnally & Lemond, 1973; Spear, 2000). Adolescent-aged rodents spend more time in novel sections of a physical environment (Adriani et al., 1998; Macri et al., 2002), which has been described in terms of exploratory and novelty seeking motivations, but also reduced anxiety. Existing work on human adolescents cannot decouple exploratory motivation from risk or reward-related motivations (Humphreys et al., 2015) and/or had sample characteristics that did not permit developmental comparison (Humphreys et al., 2015; Kayser et al., 2015), but see (Christakou et al., 2013). The Horizon Task enabled *a)* isolation of exploratory tendencies that are orthogonal to risk and reward related factors, and *b)* quantification of the degree to which choices to explore or exploit are contingent on the utility of exploration for future choice (i.e., *strategic* exploration). Our findings indicate that baseline levels of exploration (measured by tendency to engage in directed and random exploration irrespective of horizon) did not change developmentally, whereas *strategic* use of exploration changed robustly from adolescence to adulthood.

Directed exploration.

Directed exploration is defined as a tendency to make decisions that favor choice options that lack information. We observed a rapid emergence of strategic directed exploration during adolescence that stabilized into early adulthood. Dissecting adolescents' use of directed exploration revealed underlying processes that are 'tuned' differently across adolescence to young adulthood. Adolescents' tendency to select more informative choices in short time horizons, and to select less informative choices in long time horizons, indicates lessened reliance on horizon to guide exploration. In other words, adolescents' use of directed exploration differs from adults in that it is more indifferent to its future utility.

Unique features of adolescent strategic exploration were revealed by choices on *reward-information conflict* games where one bandit yielded more points, and the other bandit yielded more information. Adolescents and adults similarly exploited the high reward option when time horizons were short. However, in long horizon games, adolescents were more likely than adults to forego the more informative option, favoring the high-reward option instead. This pattern suggests that adolescents place greater value on immediate rewards relative to the value of information that holds potential to boost long-term utility. This finding echoes prior work showing increasingly patient choices in delay discounting tasks (Christakou et al., 2011; van den Bos et al., 2015) and a rise in future-oriented cognition (Steinberg et al., 2009) from adolescence to adulthood. The present study extends these observations by demonstrating how the *utility of information* is subject to differential valuation during adolescence.

Random exploration.

In contrast, strategic random exploration was age-invariant in our sample of 12-28 year olds. Participants, regardless of age, demonstrated a more stochastic, variable pattern of exploratory behavior in long horizon games relative to short horizon games. Although random exploration is based on high-variability choices that sometimes entail selecting lower mean options, random exploration is in fact an efficacious exploration strategy (Sutton & Barto, 1998; Watkins, 1989) that successfully uncovers information and reward. That even early adolescents engage in more stochastic behavioral choices in long horizons than short horizons indicates that they exhibit a capability of using horizon to guide exploration. These findings constrain the interpretation of developmental shifts in directed exploration as unlikely to be due to baseline cognitive capability to manipulate information about horizon. It is more likely that changes in value assignment to reward and information underlie age-related changes in directed exploration, underscoring a theoretical viewpoint that developmental shifts in decision-making are rooted in valuation and cost-benefit processes (Hartley & Somerville, 2015).

Task performance.

It is important to contextualize the age differences in strategic exploration in task performance patterns: no age differences were observed in the overall success at earning points in the task. Thus, use of different strategies with age was equivalently well suited to the statistics of this task. In part, this is because the long horizon condition, with six choices, is still relatively short and so the benefits of exploration tend to be small (see (Wilson et al., 2014) for a discussion of the optimal model). In the real world, the decision horizon is often much longer, in some cases lifelong, and is almost always unspecified. As such, a generalized bias towards stochastic behavior may be advantageous for younger teens, although this may come at the cost of engaging in exploratory behavior in situations where there is less potential benefit, and failing to capitalize on the benefits of information gathering when it could be useful for subsequent choice.

Exploratory strategy and risk taking propensity.

Within the adolescent portion of the sample, we observed an association between greater strategic random exploration and propensity for risk taking in daily life. Although baseline risk taking did not vary with age within the constrained age range available for this analysis (12-17 years), there were robust individual differences both in exploration strategy and propensity to endorse risk taking. The limited age range may have hindered sensitivity to observe age differences in risky behavior using this self report measures (Figner et al., 2015).

Risk taking in daily life is shaped by a host of decisional subprocesses including valuation, risk assessment, availability of risk, consideration of long and short term consequences, and exploratory biases (Figner & Weber, 2011; Hartley & Somerville, 2015; Steinberg, 2004). Adolescents who reported a greater willingness to engage in risk taking in daily life were oriented toward selecting more low-mean options in horizon 6 and fewer low-mean options in horizon 1, consistent with a highly strategic stochastic behavioral pattern. Stochastic, random behavior could result in stumbling on risky acts during adolescence due to the burgeoning availability of risky situations. Although this initial finding awaits replication, it introduces the possibility of an additional, poorly understood “route to risk” during adolescence – behaving randomly when horizons are long (as they typically are in everyday life).

Our findings suggest a partial dissociation between the mechanisms underlying directed exploration, which correlates with age but not risk taking, and random exploration, which correlates with risk taking but not age. Given that the brain's frontal association areas undergo pronounced developmental changes throughout adolescence (Somerville & Casey, 2010), a hypothesis for future work is that changes in prefrontal cortex function and functional integration subserve the emergence of strategic directed exploration. This would be consistent with a number of studies that have implicated prefrontal cortex, and frontal pole in particular, in exploratory choice (Badre et al., 2012; Beharelle et al., 2015; Daw et al., 2006; Frank et al., 2009). For random exploration, the association with risk taking may be related to subcortical reward systems, and is consistent with work implicating dopamine and norepinephrine in exploration and behavioral variability (Aston-Jones, Gary & Cohen, 2005; Costa et al., 2014; Tervo et al., 2014). It is important to note that most previous studies of exploratory choice were not designed to dissociate directed and random exploration. More work will be required to understand the neural substrates of directed and random exploration and how these relate to the patterns observed in the present study.

Conclusion.

Although there was not a shift in overall exploratory motivation from adolescence to adulthood, there were clear age-related differences in the exploratory strategies used in decision-making. These findings offer a framework to study strategic exploratory behavior that could be expanded to reveal the developmental, hormonal, and experiential mechanisms that shape the unique features of complex decision-making through adolescence.

Acknowledgements

We thank Erik Kastman for technical assistance, and to members of the Affective Neuroscience & Development Lab for valuable discussion. This material is based upon work supported by the National Science Foundation under Grant No. (CAREER-BCS-1452530) to LHS and the APA F. J. McGuigan Early Career Investigator Research Prize for Understanding the Human Mind to LHS.

Author contributions

LHS and RCW conceived of the research; MCG, NAA, and SFS performed research; ATD, CI, and RCW contributed to task programming and analytic tools; ATD, CI, LHS, MCG, RCW, and SFS compiled and analyzed data; LHS and RCW wrote the paper with substantive input from CI, MCG, and SFS.

References

Adriani, W., Chiarotti, F., & Laviola, G. (1998). Elevated novelty seeking and peculiar d-amphetamine sensitization in periadolescent mice compared with adult mice. *Behavioral Neuroscience, 112*(5), 1152-1166.

Andersen, S. L., Thompson, A. T., Rutstein, M., Hostetter, J. C., & Teicher, M. H. (2000). Dopamine receptor pruning in prefrontal cortex during the periadolescent period in rats. *Synapse, 37*(2), 167-169.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403-450.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). Locus coeruleus and regulation of behavioral flexibility and attention. *Progress in Brain Research, 126*, 165-182.

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of multiarmed bandit problem. *Machine Learning, 47*, 235.

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron, 73*(3), 595-607.

Beharelle, A. R., Polanía, R., Hare, T. A., & Ruff, C. C. (2015). Transcranial stimulation over frontopolar cortex elucidates the choice attributes and neural mechanisms used to resolve exploration–exploitation trade-offs. *The Journal of Neuroscience, 35*(43), 14544-14556.

Blais, A.-R., & Weber, E. U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1*(1).

Blankenstein, N. E., Crone, E. A., van den Bos, W., & van Duijvenvoorde, A. C. (2016). Dealing with uncertainty: testing risk-and ambiguity-attitude across adolescence. *Developmental Neuropsychology*, 1-16.

Braams, B. R., van Duijvenvoorde, A. C. K., Peper, J. S., & Crone, E. A. (in press). Longitudinal changes in adolescent risk-taking: A comprehensive study of neural respones to rewards, pubertal development, and risk-taking behavior. *Journal of Neuroscience*.

Brain_Development_Cooperative_Group. (2012). Total and regional brain volumes in a population-based normative sample from 4 to 18 years: the NIH MRI study of normal brain development. *Cerebral Cortex, 22*(1), 1-12.

Bridle, J. S. (1990). Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimates of parameters. In D. S. Touretzky (Ed.), *Advances in Neural Information Processing Systems* (Vol. 2, pp. 211-217).

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences, 9*(3), 104-110.

Christakou, A., Brammer, M., & Rubia, K. (2011). Maturation of limbic corticostriatal activation and connectivity associated with developmental changes in temporal discounting. *Neuroimage, 54*(2), 1344-1354.

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

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.

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.

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.

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, 1*(1), 48-84.

Figner, B., van Duijvenvoorde, A. C., Blankenstein, N., & Weber, E. U. (2015). *Risk-taking, perceived risks, and perceived benefits across adolescence: A domain-specific risk-return approach.* Paper presented at the Flux: The Society for Developmental Cognitive Neuroscience, Leiden, Netherlands.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science, 20*(4), 211-216.

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience, 12*(8), 1062-1068.

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.

Gittins, J. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B (Methodological), 41*, 148.

Hartley, C. A., & Somerville, L. H. (2015). The neuroscience of adolescent decision-making. *Current Opinion in Behavioral Sciences, 5*, 108-115.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events on risky choice. *Psychological Science, 15*, 534-539.

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. *Developmental Psychobiology, 57*(3), 313-321.

Kayser, A. S., Op de Macks, Z., Dahl, R. E., & Frank, M. J. (2015). A neural correlate of strategic exploration at the onset of adolescence. *Journal of Cognitive Neuroscience, 28*(2), 199-209.

Macri, S., Adriani, W., Chiarotti, F., & Laviola, G. (2002). Risk taking during exploration of a plus-maze is greater in adolescent than in juvenile or adult mice. *Animal Behaviour, 64*(4), 541-546.

Nunnally, J. C., & Lemond, L. C. (1973). Exploratory behavior and human development. *Advances in Child Development and Behavior, 8*, 59-109.

Petersen, A. C. (1988). Adolescent development. *Annual review of psychology, 39*(1), 583-607.

Somerville, L. H., & Casey, B. J. (2010). Developmental neurobiology of cognitive control and motivational systems. *Current Opinion in Neurobiology, 20*(2), 236-241.

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science, 24*(8), 1554-1562.

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24*(4), 417-463.

Steinberg, L. (2004). Risk taking in adolescence: what changes, and why? *Ann N Y Acad Sci, 1021*, 51-58.

Steinberg, L., Graham, S. J., O'Brien, L., Woolard, J., Cauffman, E., & Banich, M. (2009). Age differences in future orientation and delay discounting. *Child Development*.

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

Tervo, D. G. R., Proskurin, M., Manakov, M., Kabra, M., Vollmer, A., Branson, K., & Karpova, A. Y. (2014). Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. *Cell, 159*(1), 21-32.

Thompson, W. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika, 25*, 285-294.

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.

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.

Watkins, C. J. C. H. (1989). *Learning from delayed rewards.* (PhD), Cambridge University, Cambridge, England.

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), 2074-2082.

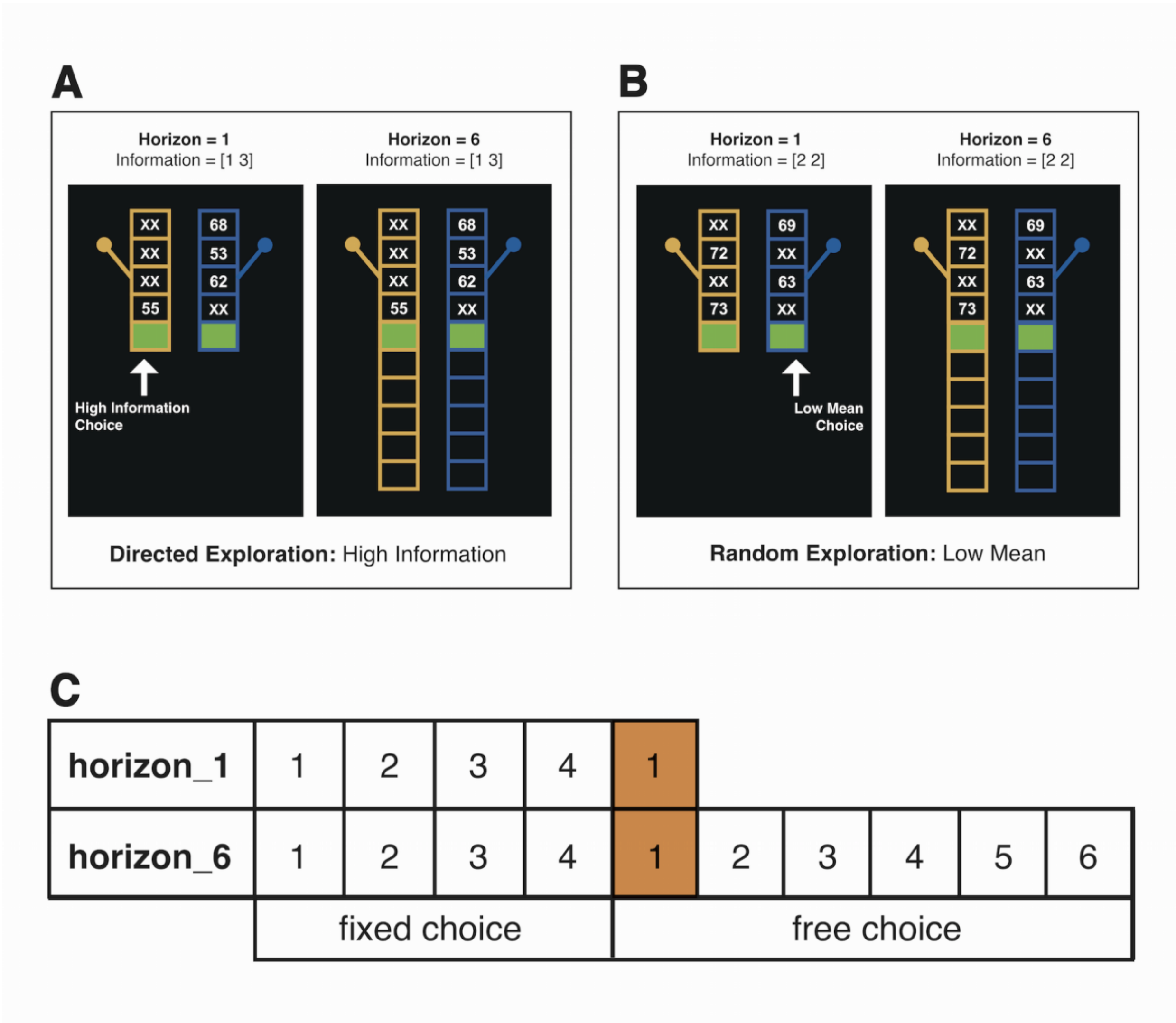

Figure 1. Task design. In the Horizon Task, participants selected between two colored bandits with varying amounts of information and reward. Directed exploration was measured in unequal information games (A) where exploration was defined as selecting the bandit with less information available (left). Random exploration was measured in equal information games (B) where exploration was defined as selecting the bandit with a lower mean payout history. The horizon manipulation was intended to render exploration differentially advantageous, because information gained by exploration could be used subsequently on horizon_6 but not horizon_1

games. (C) Games revealed four bandit payouts as fixed choices, followed by the dependent measure of first free choice decision (in orange).

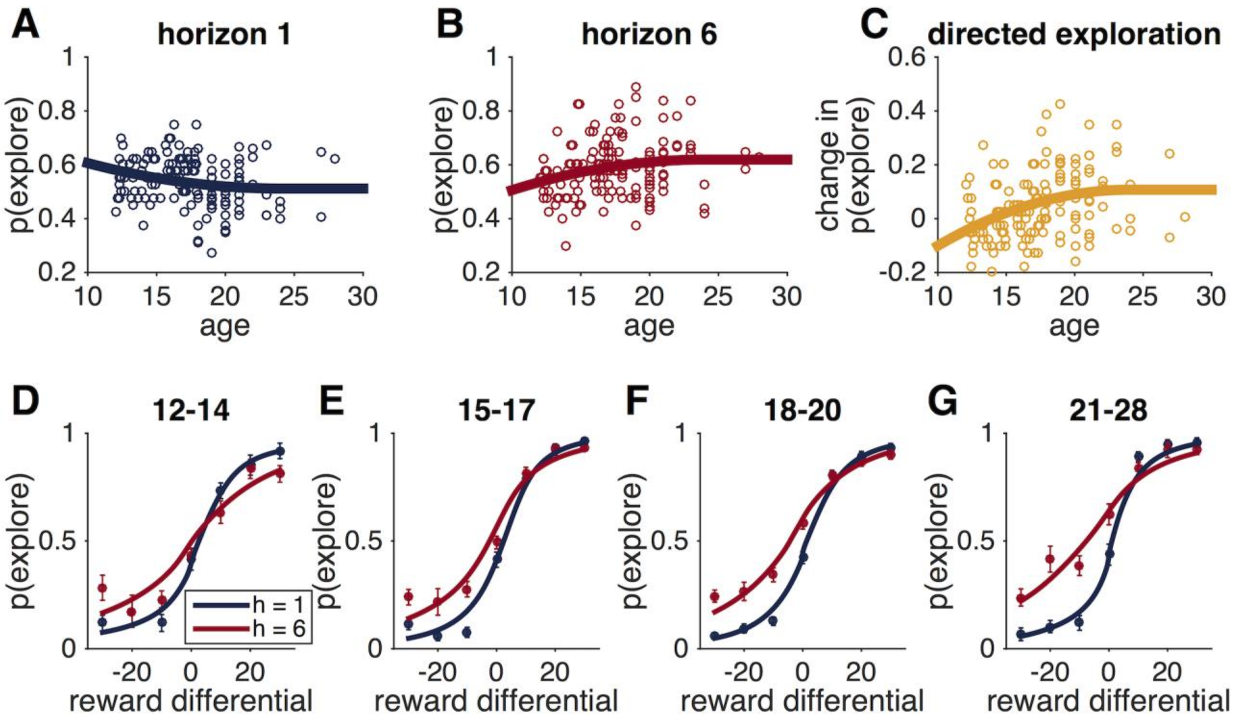

Figure 2. Strategic use of directed exploration increases from adolescence to young adulthood. A-B. Age differences in probability of selecting the bandit with less information available onscreen (y-axis; p(explore)) for horizon_1 (A) and horizon_6 (B) games. C: Age differences in strategic directed exploration (y-axis: differential exploration for horizon_6 > horizon_1 games). With increasing age (x-axis), there was a reduction in exploration in horizon_1 games and an increase in exploration in horizon_6 games. This indicates a rise in strategic directed exploration through adolescence that stabilizes in young adulthood. D-G: Choice curves for four age subgroups plotting the mean differential points between bandits on the x-axis and the probability of selecting the bandit with less information available onscreen (p(explore)) on the y-axis. Positive values on the x-axis denote games where the bandit with the higher mean payout was also the high information choice; negative values denote games where the bandit with the lower mean payout was also the high information choice (i.e., a conflict

between the more rewarding and more informative choice). Error bars represent standard error of the mean.

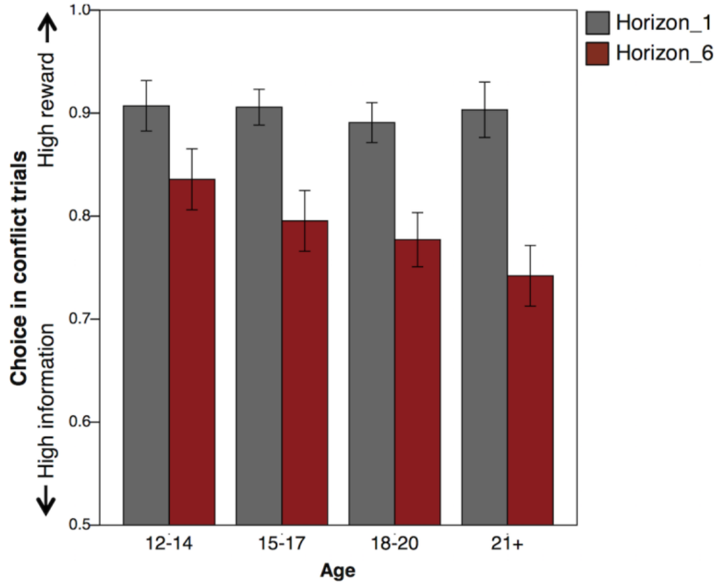

Figure 3. Age shifts the strategies used to adjudicate choice when reward and information conflict. In horizon_1 games, participants consistently select the high reward option. In horizon_6 games, younger participants were more likely to select the high reward option and with increasing age there is an increasing tendency to select the high information option instead. Error bars represent standard error of the mean.

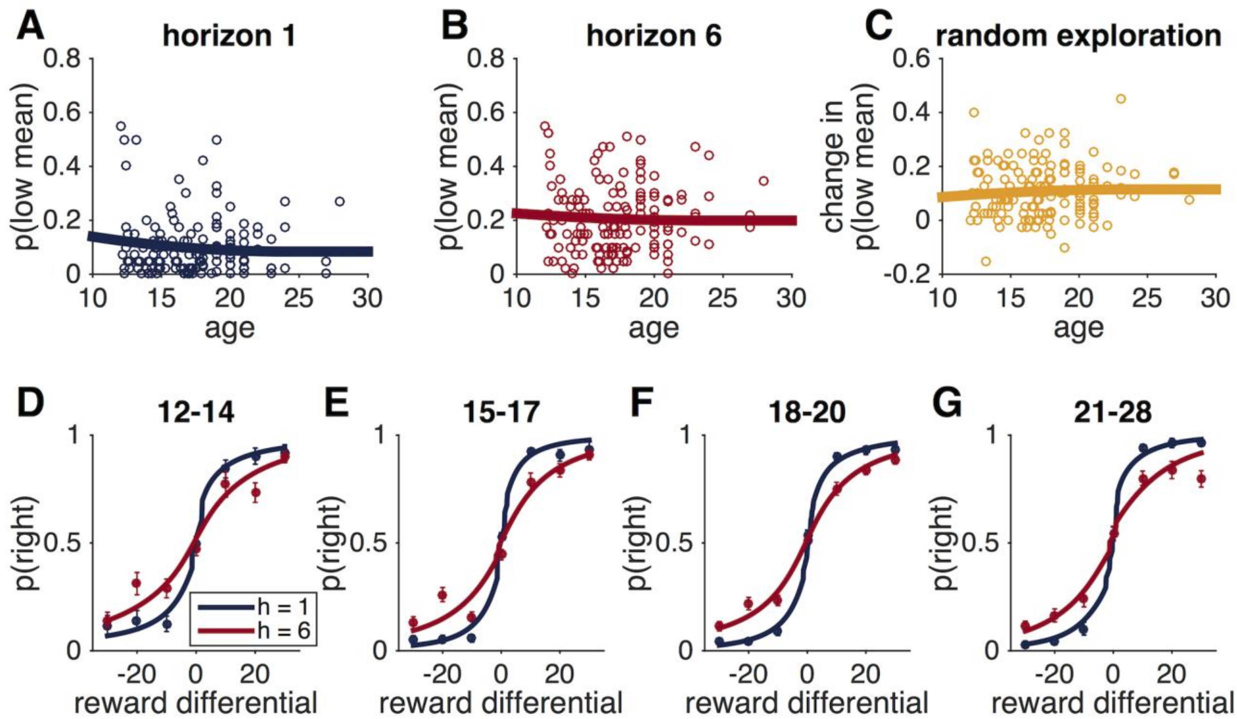

Figure 4. Random exploration strategies do not differ from adolescence to early adulthood. A-B. Age differences in probability of selecting the bandit with the lower mean payout history (y-axis; p(low mean)). Age does not modulate low mean choices in horizon_1 (A) or horizon_6 games (B). C. Age differences in strategic random exploration, defined as p(low mean) for horizon_6 > horizon_1. Plotting by age (x-axis) indicates stability of strategic random exploration through adolescence and young adulthood. D-G: Choice curves for four age subgroups plotting the mean differential points between bandits on the x-axis and probability of selecting the right-side option on the y-axis. Positive values on the x-axis indicate that the right side option has a greater point mean than the left side option, and negative values indicate that left side option has a greater point mean than the right side option. Error bars represent standard error of the mean.

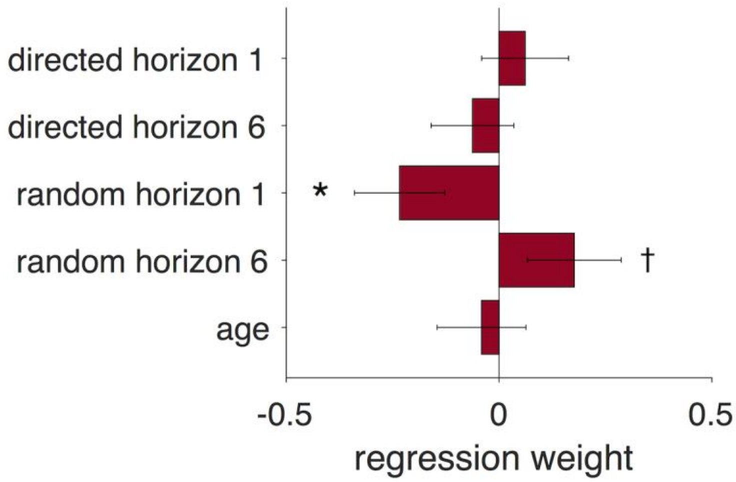

Figure 5. Task based exploration predicts endorsement of risk taking. Greater tendency to endorse risk taking was associated with less random exploration in short horizons (horizon_1) and more random exploration in long horizons (horizon_6). Directed exploration and age did not predict endorsement of risk taking behavior. * $p<0.05$; † $p<0.1$. Error bars represent standard error of the mean.