Journal of Experimental Psychology: General
2017, Vol. 146, No. 2, 155–164

0096-3445/17/$12.00 http://dx.doi.org/10.1037/xge0000250

Charting the Expansion of Strategic Exploratory Behavior During Adolescence

Leah H. Somerville, Stephanie F. Sasse, and Megan C. Garrad
Harvard University

Andrew T. Drysdale
Weill Cornell Medical College

Nadine Abi Akar and Catherine Insel
Harvard University

Robert C. Wilson
University of Arizona

Although 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. Participants were 149 12–28-year-olds who completed a computational explore–exploit paradigm that manipulated reward value, information value, and decision horizon (i.e., the utility that 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 attitudes toward 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

Supplemental materials: http://dx.doi.org/10.1037/xge0000250.supp

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—e.g., inviting a new classmate versus an old friend as a date to prom; trying a new drug versus sipping a beer at a party; attending college in an unfamiliar city versus in one's hometown. These dilemmas are 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). For the present study, we characterized the developmental emergence of strategies used to solve the explore–exploit dilemma from adolescence to adulthood. Moreover, because exploratory behavior has been conceptually linked with risk taking in adolescence (e.g., Crone and Dahl, 2012), we evaluated whether the use

This article was published Online First December 15, 2016.

Leah H. Somerville, Stephanie F. Sasse, and Megan C. Garrad, Department of Psychology and Center for Brain Science, Harvard University; Andrew T. Drysdale, Department of Psychiatry, Weill Cornell Medical College; Nadine Abi Akar and Catherine Insel, Department of Psychology and Center for Brain Science, Harvard University; Robert C. Wilson, Department of Psychology and Cognitive Science Program, University of Arizona.

Leah H. Somerville and Robert C. Wilson conceived of the research; Megan C. Garrad, Nadine Abi Akar, and Stephanie F. Sasse performed research; Andrew T. Drysdale, Catherine Insel, and Robert C. Wilson contributed to task programming and analytic tools; Andrew T. Drysdale, Catherine Insel, Leah H. Somerville, Megan C. Garrad, Robert C. Wilson, and Stephanie F. Sasse compiled and analyzed data; Leah H. Somerville and Robert C. Wilson wrote the paper with substantive input from Catherine Insel, Megan C. Garrad, and Stephanie F. Sasse. We thank Erik Kastman for technical assistance and members of the Affective Neuroscience and Development Lab for valuable discussion. This material is based upon work supported by the National Science Foundation under Grant CAREER-BCS-1452530 to Leah H. Somerville and the American Psychological Association F. J. McGuigan Early Career Investigator Research Prize for Understanding the Human Mind to Leah H. Somerville.

Correspondence concerning this article should be addressed to Leah H. Somerville, Department of Psychology and Center for Brain Science, Harvard University, 52 Oxford Street, Room 290, Cambridge MA 02138. E-mail: somerville@fas.harvard.edu

of particular exploratory strategies related to adolescents' propensity to endorse risky behaviors in daily life.

Recent findings suggest that people make explore–exploit decisions using at least two distinct strategies (Wilson, Geana, White, Ludvig, & Cohen, 2014). One strategy, *directed exploration*, is guided by assessment of the value of information that would be gained from choosing a particular option (Auer, Cesa-Bianchi, & Fischer, 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 fast. When implemented correctly, that is, 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 commuting 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 use 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 nonhuman animals has suggested strong exploratory motivations during adolescence (Adriani, Chiarotti, & Laviola, 1998; Macrì, Adriani, Chiarotti, & Laviola 2002; Spear, 2000), researchers to date have not assessed developmental changes in the strategies of exploration. On one hand, the capacity for complex cognitive operations continues to refine during adolescence (Casey, Tottenham, Liston, & Durston, 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, Rajkowski, & Cohen, 2000; Costa, Tran, Turchi, & Averbeck, 2014; Frank, Doll, Oas-Terpstra, & Moreno, 2009), have shown anatomical (Andersen, Thompson, Rutstein, Hostetter, & Teicher, 2000; Brain Development Cooperative Group, 2012) and functional (Braams van Duijvenvoorde, Peper, & Crone, in press; Galvan et al., 2006) perturbations during adolescence. Nascent work in other domains of decision making has indicated that adolescents exhibit shifts in decision computations, such as ambiguity tolerance (Blankenstein, Crone, van den Bos, & van Duijvenvoorde, 2016; Tymula et al., 2012), risk tolerance (Defoe, Dubas, Figner, & van Aken, 2015), intertemporal choice (Steinberg et al., 2009; van den Bos, Rodriguez, Schweitzer, & McClure, 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 the use of each form of strategic exploration and evaluated whether such shifts related to a tendency to endorse real-life risky behavior.

Method and Materials

Sample

One hundred forty-nine participants between the ages of 12.08 and 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 approximately one 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: one adolescent for withdrawing from participation early, and one 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 in the 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 of the task that the developmental sample completed.

Horizon Task

Conceptual overview. During the horizon task (Wilson et al., 2014), participants made a series of choices between two onearmed 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 five 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 advantage 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 with selecting a bandit with more payouts displayed. The differential decision horizon (i.e., the number of choices in each game) manipulated the advantage of exploration, with exploration being more advantageous in long decision horizons because there was an opportunity to exploit that information in subsequent choices.

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

Directed exploration. Directed exploration was calculated using the unequal-information games for which one option showed one previous trial payout and the other option showed three previous trial payouts (i.e., [1, 3]; see Figure 1A). Directed exploration was defined as selecting the option with one previous payout displayed, that is, the high information choice [1], even if it had the lower mean payout history. Strategic directed exploration was defined as selecting the high information choice option more frequently for Horizon-6 games than for 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, more so when subsequent exploration was possible.

Random exploration. Random exploration was calculated using games for which both options showed two previous trial payouts (i.e., [2, 2]; see 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

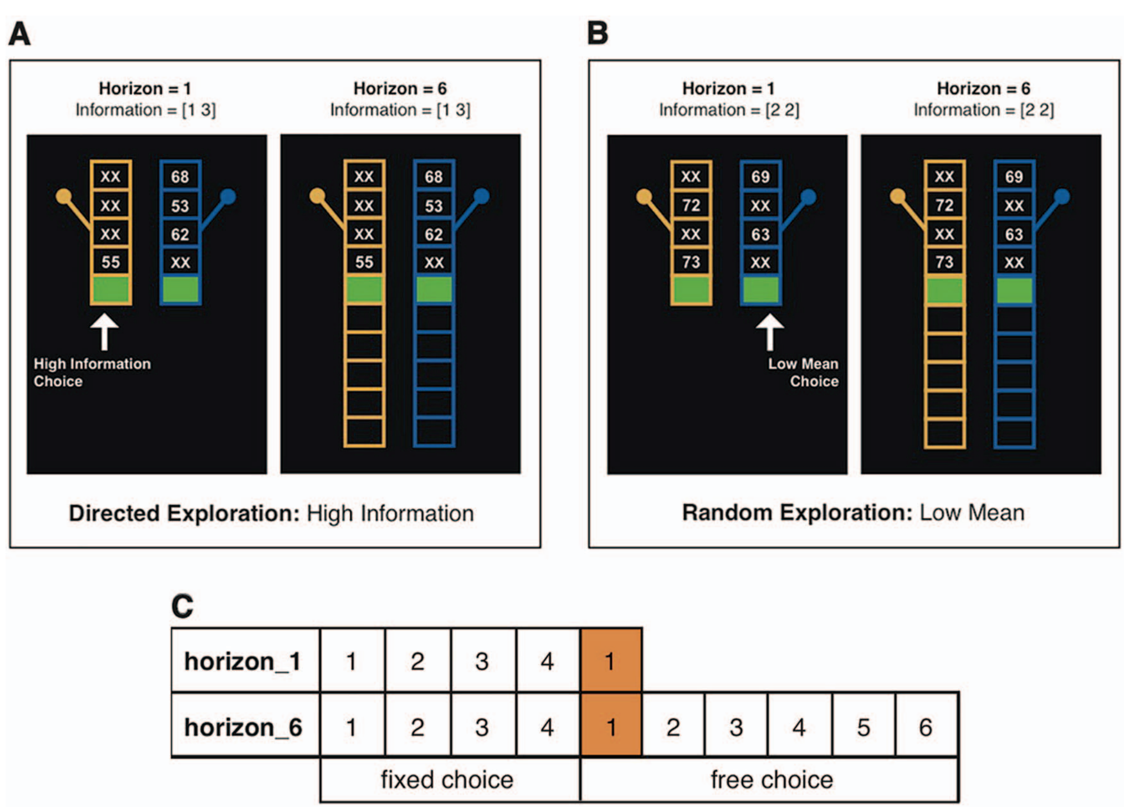

Figure 1. Task design. In the horizon task, participants selected between two colored bandits with varying amounts of information and reward. Directed exploration was measured in unequal-information games (A) where exploration was defined as selecting the bandit with less information available (left). 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).

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 online supplemental materials). For adults, the task was described as a bandit/slot-machine task similar to the kind seen 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. Participants were shown a series of fixed choices first, which revealed payout history information about the two options before it was the participant's turn to pick. When it was the participant's turn to pick, he or she 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. Although 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; 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 one bandit, indicating that 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, Barron, Weber, & Erev, 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]; see 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]; see 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, and the corresponding space for the unplayed option was filled with "XX."

Following the fixed choices, participants engaged in the free-choice phase in which they freely selected between the two bandits (see 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 would 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 Foundation, Vienna, Austria), MATLAB (MathWorks, 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 (after 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; computational modeling data are presented in the online supplemental materials.

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 regardless of the reward differential between bandits. Taking the reward differential into account quantitatively using a computational model (see online supplemental materials) and qualitatively by inspecting choice curves leads to convergent conclusions and thus are presented descriptively and used in targeted post hoc 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 semiarbitrary 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 concerned whether age differences minimized into early adulthood, presumably because the underlying processes had asymptoted developmentally (e.g., emergent predictor), rendering adolescence a period of rapid change, or whether age differences continued 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 change rapidly during adolescence and stabilize 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 online supplemental materials 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 (i.e., How likely you would be to engage in that activity?), Risk Perception (i.e., How risky do you feel the activity is?), and Expected Benefits (i.e., How much would you benefit from engaging in that activity?). Within this sample, Cronbach’s α reliability indices indicated strong reliability (all scales > .8; see supplemental materials, 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, thereby solely maximizing reward and never engaging in exploration. Although participants could have adopted this strategy, their choice behavior suggested that they did not. For equal-information [2, 2] games, participants exhibited a significant nonzero rate of selecting the bandit with the lower mean payout history (Horizon 1: 9.94% ± 10.92%; $t_{146} = 11.04$, $p < .001$; Horizon 6: 20.90% ± 13.35; $t_{146} = 18.98$, $p < .001$). 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, $t_{146} = 9.01$, $p < .001$; Horizon 6: 27.29% ± 21.34, $t_{146} = 15.51$, $p < .001$). 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 ($F_{1,\ 145} = 24.60$, $p < .001$, $\eta^2 = 0.145$; high-information choices Horizon 1 = 53% ± 9.84%; Horizon 6 = 58.81% ± 11.15%). This main effect was qualified by a significant Age × Horizon interaction ($F_{1,\ 145} = 18.85$, $p < .001$, $\eta^2 = 0.115$; see Figure 2C). Post hoc 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, $t_{146} = 2.68$, $p = .008$, $R^2 = 0.047$; see Figure 2B, and an overall reduction of information seeking on games for which information could not be used to inform subsequent choices, $t_{146} = -2.55$, $p = .012$, $R^2 = 0.043$; see Figure 2A. Overall selection of the high-information option did not differ by age (main effect of age: $F_{1,\ 145} = 0.08$, $p > .250$), suggesting that overall levels of directed exploration were age-invariant.

Choice curves in Figure 2D–G show the tendency to select the information-maximizing option as a function of reward differential between the two bandits and the horizon for four age subgroupings. Participants of all ages exhibited a tendency to exploit the high-reward option when it was also the high-information option (the right half of each choice curve), regardless of horizon. The primary age differences in directed exploration occurred during reward–information conflict games, that is, when one bandit contained higher reward value and the other contained higher information value (the left half of each choice curve). A post hoc analysis of age differences in conflict games yielded a significant Age × Horizon interaction ($F_{1,\ 145} = 5.40$, $p = .022$, $\eta^2 = 0.036$). In Horizon 1, participants across ages were uniformly likely to exploit the high reward option. However, in Horizon 6, increasing age related to an increasing tendency to forego the high-reward option in favor of the lower value but 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 in Horizon 1 ($F_{1,\ 145} = 170.56$, $p < .001$, $\eta^2 = 0.541$). Selection of the low-mean option did not differ by age (main effect of age: $F_{1,\ 145} = 0.976$, $p > .250$), suggesting that overall levels of random exploration were age-invariant. The Age × Horizon interaction was not significant ($F_{1,\ 145} = 0.976$, $p > .250$; see Figure 4A–C), suggesting that strategic random exploration was also age-invariant.

Choice-curve data (see Figure 4D–G) reveal the consistency of random exploration across four age subgroups. The choice curves

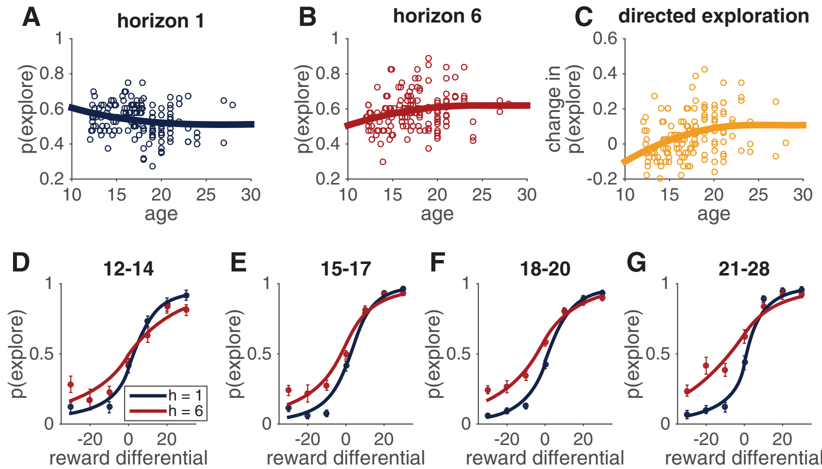

Figure 2. Strategic use of directed exploration increases from adolescence to young adulthood. Age differences in probability of selecting the bandit with less information available onscreen (*y* axis; $p_{explore}$) for Horizon-1 (A) and Horizon-6 (B) games. Age differences in strategic directed exploration (*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 (C). 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 which the bandit with the higher mean payout was also the high-information choice; negative values denote games in which 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; D–G). Error bars represent standard error of the mean.

show the tendency to select the reward-maximizing option as a function of differential reward value of the two bandits and the horizon. Overall, participants of all ages exploited reward maximization in Horizon-1 choices (blue lines). However, in Horizon-6 conditions (red lines), participants demonstrated less reward maximization (flatter curves). These patterns were consistent across the four age subgroups.

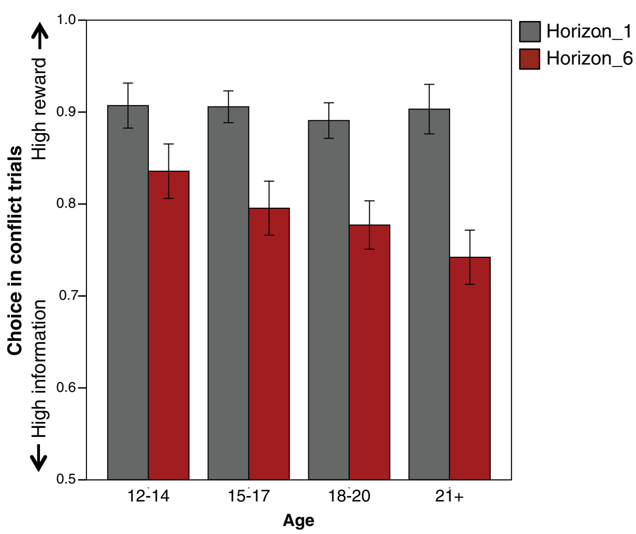

Figure 3. Age shifts the strategies used to adjudicate choice when reward and information conflict. In Horizon-1 games, participants consistently selected the high-reward option. In Horizon-6 games, younger participants were more likely to select the high-reward option and with increasing age, there is an increasing tendency to select the high-information option instead. Error bars represent standard error of the mean.

Relationship Between Exploration Strategy and Risky Behavior

Measures were completed by 86 participants younger than 18 years and one 18-year-old of daily life risk taking and risk attitudes using the DOSPERT questionnaire (Blais & Weber, 2006). Participants in the sample endorsed a wide range of risk endorsement and risk attitudes (see online supplemental materials 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 ($\beta = 0.168$, $F_{1, 85} = 4.287$, $p = .0415$; 95% CI [0.007, 0.330]; the same analysis statistically controlled for age was also significant, $p = .0422$). Directed exploration and random exploration were not significantly related to Risk Perceptions or Expected Benefits scores. This finding suggests that adolescents who tend to use random exploration strategically in the horizon task are more likely to endorse greater propensity for risk taking in daily life. Computation-based indices of random exploration yielded highly convergent findings (see supplemental materials).

Follow-up analyses tested for relationships between risk taking and exploration in Horizons 1 and 6 separately. Results indicate that risk taking is significantly related to a lesser tendency to engage in random exploration in Horizon 1, $t_{85} = -2.193$, $p =$

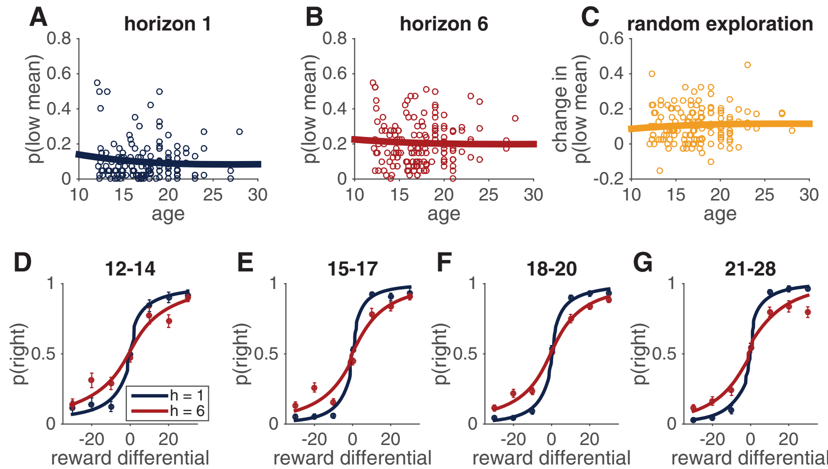

Figure 4. Random exploration strategies do not differ from adolescence to early adulthood. Age differences in probability of selecting the bandit with the lower mean payout history (*y* axis; $p_{\text{low mean}}$). Age does not modulate low mean choices in Horizon-1 (A) or Horizon-6 games (B). Age differences in strategic random exploration, defined as $p_{\text{low mean}}$ for Horizon 6 > Horizon 1. Plotting by age (*x* axis) indicates stability of strategic random exploration through adolescence and young adulthood (C). 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 (D–G). Error bars represent standard error of the mean.

.031 CI [−0.040, −0.002], and greater tendency (at a trend level) to engage in random exploration in Horizon 6, $t_{85} = 1.696$, $p = .094$, 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 indicate 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 the supplemental online materials.

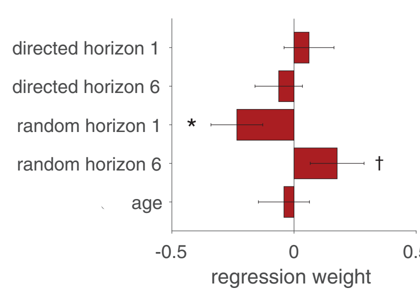

Figure 5. Task based exploration predicts endorsement of risk taking. Greater tendency to endorse risk taking was associated with less random exploration in short horizons (Horizon 1) and more random exploration in long horizons (Horizon 6). Directed exploration and age did not predict endorsement of risk-taking behavior. * $p < .05$; † $p < .1$. Error bars represent standard error of the mean.

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, 1974; Spear, 2000). Adolescent 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 have had sample characteristics that did not permit developmental comparison (Humphreys et al., 2015; Kayser, Op de Macks, Dahl, & Frank, 2016), 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 choices (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 toward 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 in which 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, Brammer, & Rubia, 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 extended 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 during 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 toward stochastic behavior may be advantageous for younger teens, although this may come at the cost of engaging in exploratory behavior in situations in which there is less potential benefit, and failing to capitalize on the benefits of information gathering when it could be useful for subsequent choices.

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 self-report measures (Figner, van Duijvenvoorde, Blankenstein, & Weber, 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, that is, 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, Doll, Long, & Frank, 2012; Raja Beharelle, Polanía, Hare, & Ruff, 2015; Daw, O'Doherty, Dayan, Seymour, & Dolan, 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 & 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.

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,* 1152–1166. http://dx.doi.org/10.1037/0735-7044.112.5.1152

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,* 167–169. http://dx.doi.org/10.1002/1098-2396(200008)37:2<167::AID-SYN11>3.0.CO;2-B

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. http://dx.doi.org/10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). Locus coeruleus and regulation of behavioral flexibility and attention. *Progress in Brain Research, 126,* 165–182. http://dx.doi.org/10.1016/S0079-6123(00)26013-5

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of multiarmed bandit problem. *Machine Learning, 47,* 235–256. http://dx.doi.org/10.1023/A:1013689704352

Badre, D., Doll, B. B., Long, N. M., & Frank, M. J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron, 73,* 595–607. http://dx.doi.org/10.1016/j.neuron.2011.12.025

Blais, A.-R., & Weber, E. U. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. *Judgment and Decision Making, 1,* 33–47. Retrieved from http://journal.sjdm.org/jdm06005.pdf

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, 41,* 77–92. http://dx.doi.org/10.1080/87565641.2016.1158265

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,* 1–12. http://dx.doi.org/10.1093/cercor/bhr018

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). San Francisco, CA: Morgan Kaufmann.

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,* 104–110. http://dx.doi.org/10.1016/j.tics.2005.01.011

Christakou, A., Brammer, M., & Rubia, K. (2011). Maturation of limbic corticostriatal activation and connectivity associated with developmental changes in temporal discounting. *NeuroImage, 54,* 1344–1354. http://dx.doi.org/10.1016/j.neuroimage.2010.08.067

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,* 1807–1823. http://dx.doi.org/10.1162/jocn_a_00447

Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2014). Dopamine modulates novelty seeking behavior during decision making. *Behavioral Neuroscience, 128,* 556–566. http://dx.doi.org/10.1037/a0037128

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nature Reviews Neuroscience, 13,* 636–650. http://dx.doi.org/10.1038/nrn3313

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature, 441,* 876–879. http://dx.doi.org/10.1038/nature04766

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. http://dx.doi.org/10.1037/a0038088

Figner, B., van Duijvenvoorde, A. C., Blankenstein, N., & Weber, E. U. (2015, September). *Risk-taking, perceived risks, and perceived benefits across adolescence: A domain-specific risk-return approach*. Paper presented at the 3rd Annual Flux Congress: The Society for Developmental Cognitive Neuroscience, Leiden, the Netherlands.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? Determinants of risk taking. *Current Directions in Psychological Science, 20,* 211–216. http://dx.doi.org/10.1177/0963721411415790

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,* 1062–1068. http://dx.doi.org/10.1038/nn.2342

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. *The Journal of Neuroscience, 26,* 6885–6892. http://dx.doi.org/10.1523/JNEUROSCI.1062-06.2006

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. http://dx.doi.org/10.1016/j.cobeha.2015.09.004

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15,* 534–539. http://dx.doi.org/10.1111/j.0956-7976.2004.00715.x

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,* 313–321. http://dx.doi.org/10.1002/dev.21293

Kayser, A. S., Op de Macks, Z., Dahl, R. E., & Frank, M. J. (2016). A neural correlate of strategic exploration at the onset of adolescence. *Journal of Cognitive Neuroscience, 28,* 199–209. http://dx.doi.org/10.1162/jocn_a_00896

Macrì, 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,* 541–546. http://dx.doi.org/10.1006/anbe.2002.4004

Nunnally, J. C., & Lemond, L. C. (1974). Exploratory behavior and human development. *Advances in Child Development and Behavior, 8,* 59–109. http://dx.doi.org/10.1016/S0065-2407(08)60493-0

Petersen, A. C. (1988). Adolescent development. *Annual Review of Psychology, 39,* 583–607. http://dx.doi.org/10.1146/annurev.ps.39.020188.003055

Raja Beharelle, A., 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,* 14544–14556. http://dx.doi.org/10.1523/JNEUROSCI.2322-15.2015

Somerville, L. H., & Casey, B. J. (2010). Developmental neurobiology of cognitive control and motivational systems. *Current Opinion in Neurobiology, 20,* 236–241. http://dx.doi.org/10.1016/j.conb.2010.01.006

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science, 24,* 1554–1562. http://dx.doi.org/10.1177/0956797613475633

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24,* 417–463. http://dx.doi.org/10.1016/S0149-7634(00)00014-2

Steinberg, L. (2004). Risk taking in adolescence: What changes, and why? *Annals of the New York Academy of Sciences, 1021,* 51–58. http://dx.doi.org/10.1196/annals.1308.005

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,* 28–44. http://dx.doi.org/10.1111/j.1467-8624.2008.01244.x

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,* 21–32. http://dx.doi.org/10.1016/j.cell.2014.08.037

Thompson, W. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika, 25,* 285–294. http://dx.doi.org/10.1093/biomet/25.3-4.285

Tymula, A., Rosenberg Belmaker, L. A., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 109,* 17135–17140. http://dx.doi.org/10.1073/pnas.1207144109

van den Bos, W., Rodriguez, C. A., Schweitzer, J. B., & McClure, S. M. (2015). Adolescent impatience decreases with increased frontostriatal connectivity. *PNAS: Proceedings of the National Academy of Sciences of the United States of America, 112,* E3765–E3774. http://dx.doi.org/10.1073/pnas.1423095112

Watkins, C. J. C. H. (1989). *Learning from delayed rewards (Unpublished doctoral dissertation)*. King's College, 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,* 2074–2081. http://dx.doi.org/10.1037/a0038199

Received June 14, 2016
Revision received September 18, 2016
Accepted October 14, 2016 ■