



Published in final edited form as:

Behav Neurosci. 2014 October ; 128(5): 556–566. doi:10.1037/a0037128.

Dopamine Modulates Novelty Seeking Behavior During Decision Making

Vincent D. Costa, Valery L. Tran, Janita Turchi, and Bruno B. Averbeck

National Institute of Mental Health

Abstract

Novelty seeking refers to the tendency of humans and animals to explore novel and unfamiliar stimuli and environments. The idea that dopamine modulates novelty seeking is supported by evidence that novel stimuli excite dopamine neurons and activate brain regions receiving dopaminergic input. In addition, dopamine is shown to drive exploratory behavior in novel environments. It is not clear whether dopamine promotes novelty seeking when it is framed as the decision to explore novel options vs. the exploitation of familiar options. To test this hypothesis, we administered systemic injections of saline or GBR-12909, a selective dopamine transporter (DAT) inhibitor, to monkeys and assessed their novelty seeking behavior during a probabilistic decision making task. The task involved pseudorandom introductions of novel choice options. This allowed monkeys the opportunity to explore novel options or to exploit familiar options that they had already sampled. We found that DAT blockade increased the monkeys' preference for novel options. A reinforcement learning (RL) model fit to the monkeys' choice data showed that increased novelty seeking following DAT blockade was driven by an increase in the initial value the monkeys assigned to novel options. However, blocking DAT did not modulate the rate at which the monkeys learned which cues were most predictive of reward or their tendency to exploit that knowledge. These data demonstrate that dopamine enhances novelty-driven value and imply that excessive novelty seeking—characteristic of impulsivity and behavioral addictions—might be caused by increases in dopamine, stemming from less reuptake.

Keywords

dopamine; novelty seeking; reinforcement learning; exploration; exploitation; impulsivity; foraging; curiosity; reuptake

Novelty seeking refers to the tendency of humans and animals to explore unfamiliar stimuli and environments (Reed, Mitchell, & Nokes, 1996; Wilson & Goldman-Rakic, 1994). Identifying the neural mechanisms mediating novelty seeking is an important question, as deviations in novelty seeking characterize various psychiatric and neurological disorders (Averbeck et al., 2013; Dalley, Everitt, & Robbins, 2011).

Correspondence concerning this article should be sent to Vincent Costa or Bruno Averbeck, Laboratory of Neuropsychology, NIMH/NIH, Building 49 Room 1B80, 49 Convent Drive MSC 4415, Bethesda, MD 20892-4415. Contact: vincent.costa@nih.gov or bruno.averbeck@nih.gov.

All authors, Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health.

Exploring novel stimuli reduces uncertainty about the environment, and is necessary for acquiring information to optimize choice behavior. The trade-off between exploring novel options and exploiting known options is important in environments where choice options are dynamically changing (Cohen, McClure, & Yu, 2007). By learning the statistics of the environment, one can strategically manage this trade-off and maximize reward totals (Averbeck et al., 2013). While optimal solutions exist (Gittins, 1979), a useful heuristic is to assign novel stimuli a high enough initial value to forego the exploitation of familiar reward options. This ensures exploration, but not necessarily optimal behavior, and has led to speculation that novel stimuli are intrinsically processed as if they were themselves rewarding (Hazy, Frank, & O'Reilly, 2010; Kakade & Dayan, 2002).

Dopamine signaling seemingly contributes to the optimistic valuation of novel stimuli. Novel stimuli are known to excite dopamine neurons (Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Horvitz, 2000) as well as heighten hemodynamic signals in brain regions receiving dopaminergic input (Bunzeck, Dayan, Dolan, & Duzel, 2010; Bunzeck & Duzel, 2006). Also, in humans, electroencephalographic markers of novelty detection are enhanced after receipt of the dopamine agonist apomorphine (Rangel-Gomez, Hickey, van Amelsvoort, Bet, & Meeter, 2013). While these studies show that dopamine regulates attentional orienting, they fail to link novelty detection to novelty seeking. Direct manipulations of the dopamine system in rodents are shown to alter the exploration of novel objects (Dulawa, Grandy, Low, Paulus, & Geyer, 1999; Zhuang et al., 2001). A caveat to these findings is that gross changes in dopaminergic status alter locomotor activity, which might bias novelty seeking behavior. So it remains unclear if dopamine mediates novelty processing beyond attentional orienting or increased motor output. One exception is a fMRI study (Wittmann, Daw, Seymour, & Dolan, 2008), which reported novelty-induced increases in reward prediction error signals in ventral striatum. If dopamine does control novelty seeking, enhancing dopaminergic transmission should bias choice preferences in favor of exploring novel choice opportunities. This bias should persist when motor activity is equated across decisions to explore vs. exploit. It should also persist when novel stimuli are salient due to their potential reward value, not simply because they are alerting (Bromberg-Martin et al., 2010).

To determine if dopamine enhances novelty seeking, we systemically inhibited the dopamine transporter (DAT) in monkeys and examined how this pharmacological manipulation affected novelty seeking during decision making. We manipulated dopamine by injecting the dopamine transporter (DAT) inhibitor, GBR-12909, or saline, before the animals performed a probabilistic three-arm bandit task where novel options were randomly introduced. DAT is specific to the cell membrane of dopamine neurons and controls reuptake of extracellular dopamine. Blocking DAT therefore increases extracellular dopamine levels (Zhang, Doyon, Clark, Phillips, & Dani, 2009). To test if increasing extracellular dopamine heightened novelty seeking, we compared how often the monkeys chose novel vs. familiar options when dopamine levels were enhanced or not. Using a reinforcement learning (RL) model, we also examined if increased novelty seeking following DAT blockade was attributable to an increase in the value of novel options.

Methods

Subjects

Three male rhesus monkeys (*Macaca mulatta*) aged 5–6 years with weights ranging from 6.5–9.3 kg, were studied. Each animal was pair housed and had access to food *ad libitum*. On testing days the monkeys earned their fluid through performance on the task, while on non-testing days they were given free access to water. Experimental procedures for all monkeys were performed in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the Animal Care and Use Committee of the National Institute of Mental Health.

Pharmacology

Prior to the start of the drug testing schedule, the monkeys were first habituated for several days to intramuscular injections of sterile saline (pH 7.4, 0.1 ml/kg) into their hind limb. After this habituation period, the monkeys calmly extended their leg to receive an injection. During experimental sessions, the monkeys received intramuscular injections of the dopamine transporter blocker, GBR-12909 (Sigma-Aldrich, St. Louis, MO) or saline. Following the injection, the animals waited 30 minutes inside the testing box before starting the task. The length of the post-injection waiting period was based on pharmacokinetic evidence in monkeys that intramuscular injections of GBR-12909 first elevate extracellular dopamine levels significantly above baseline after 30 minutes.

GBR-12909 was dissolved by sonication in sterile water and stored at –4C for use within the week. On drug injection days, aliquots were thawed and re-sonicated. Aliquot concentrations were determined to keep the total injection volume under 1.5 mL when administering a 1.3 mg/kg dose. Pharmacokinetic studies of GBR-12909 in monkeys have tested doses between 0.5–5.0 mg/kg and measured via microdialysis, significant elevations in striatal dopamine levels (160–630%) as early as 30–45 minutes post injection (Czoty, Justice, & Howell, 2000; Tsukada, Harada, Nishiyama, Ohba, & Kakiuchi, 2000). Within this dose range, the effects of GBR-12909 typically last for 2–5 hours depending on the route of administration, with DAT occupancies ranging between 23–76% (1–10 mg/kg; Eriksson, Langstrom, & Josephsson, 2011; Villemagne et al., 1999). We chose to use doses at the lower end of this range since a similar dose of GBR-12909 (1.7 mg/kg) selectively blocked cocaine-maintained responding while having little to no effect on food maintained-responding (Glowa, Wojnicki, Matecka, Rice, & Rothman, 1995). The chosen dose also approximated a well-tolerated single dose of GBR-12909 (~1.4 mg/kg) in healthy human subjects (Sogaard et al., 1990). In humans, higher doses of GBR-12909 (~ 2.8 or 4.3 mg/kg) are shown to have dose-related side effects including heart palpitations, asthenia, difficulties concentrating, and sedation (Sogaard et al., 1990). Injection volumes ranged between 0.7 mL to 1.5 mL, depending on the weight of the monkey. Saline injection volumes were matched to the volume of the most recent drug injection. For monkey M an escalating dose schedule was used after the initial session, as the initial dose caused undesirable side effects related to drug tolerance (i.e. increased agitation followed by drowsiness).

The monkeys completed a minimum of 6 sessions on GBR-12909. Monkey E completed 6 sessions for a total of 25 blocks, monkey G completed 6 sessions for a total of 30 blocks, and monkey M completed 7 sessions (1.3, 0.65, 0.95, 1.0, 1.0, 1.05, 1.05 mg/kg) comprising 35 total blocks. GBR-12909 sessions were spaced a minimum of 7 days apart. Saline sessions were run interleaved with drug sessions and occurred at a minimum 2 days after each drug session, to account for turnover in the dopamine transporter protein (Kahlig & Galli, 2003).

Experimental Setup

Stimulus presentation and behavioral monitoring were controlled by a PC computer running the MonkeyLogic (version 1.1) MATLAB-based behavioral control program (Asaad & Eskandar, 2008). Eye movements were monitored using an Arrington Viewpoint eye tracking system (Arrington Research, Scottsdale, AZ), low pass filtered at 350 Hz and sampled at 1 kHz. Stimuli were displayed on a LCD monitor situated 40 cm from the monkey's eyes. On rewarded trials 0.17 mL of apple juice was delivered through a pressurized plastic tube gated by a computer controlled solenoid valve (Mitz, 2005).

Task Design and Stimuli

Monkeys completed up to 6 blocks per session of a three-arm bandit choice task (Figure 1). The task was based on a similar multi-arm bandit task previously used in human studies of novelty seeking (Averbeck et al., 2013; Djamshidian, O'Sullivan, Wittmann, Lees, & Averbeck, 2011; Wittmann et al., 2008). Each block consisted of 650 trials where the monkey had to choose among three images that were probabilistically associated with juice reward. On each trial, the monkey had to first acquire and hold a central fixation point for a variable length of time (250–750 ms). After holding fixation, three peripheral choice targets were presented at the vertices of a triangle. The main vertex of the triangle could either point up or down on each trial (i.e. either an upright or inverted triangle), and the locations of the three stimuli were randomized from trial-to-trial. The animal was required to saccade to and maintain fixation on one of the peripheral choice targets for 500 ms. We excluded trials on which the monkey attempted to saccade to more than one choice target (< 1% of all trials). After this response, a juice reward was delivered probabilistically. Within a single block of 650 trials, 32 novel stimuli were introduced. When a novel stimulus was introduced it replaced one of the existing choice options. The interval between the introduction of two novel stimuli was restricted to be greater than or equal to 10 trials and less than or equal to 30 trials, but was otherwise random.

The stimuli were naturalistic scenes downloaded from the website Flickr. Each day, we downloaded the 200–500 Creative Commons licensed photos from the website Flickr using the web application Bulkr (<http://clipyourphotos.com/bulkr>). The downloaded images were the top images on Flickr that day, ranked in terms of the website's interestingness metric. Offline, the downloaded image sets were screened for image quality, discriminability, uniqueness, size, and color to obtain a final daily set of 210 images (35 images per block). To avoid novelty driven choices due to perceptual pop out, choice options were normalized in terms of spatial frequency and luminance using functions adapted from the SHINE toolbox for MATLAB (Willenbockel et al., 2010). To control spatial frequency, each image

was converted to a gray scale image and subjected to a 2-D FFT. The amplitude at each spatial frequency was summed across the two picture dimensions and then averaged across pictures to obtain a target amplitude spectrum. All images were then normalized to have this amplitude spectrum. Luminance histogram matching was used to normalize the luminance histogram of each color channel in each image, so that it matched the mean luminance histogram of the corresponding color channel, averaged across all 186 images. Luminance histogram matching always followed spatial frequency normalization. Images were manually screened each day before and after preprocessing to verify image integrity. Images deemed unrecognizable after post-processing were replaced with images that held up better to image processing (e.g. typically higher resolution images).

At the start of a block the three initial choice options were randomly assigned a reward probability. Novel choice options were pseudorandomly assigned a reward probability when they were introduced. Reward probabilities always remained fixed for each image. The only constraint was that all three options could not have the same reward probability. In each case the assigned reward probabilities were always drawn from a symmetric reward schedule, with low and high reward probabilities centered around an intermediate reward probability of $p_m = 0.5$. The reward schedule was occasionally adjusted to maintain motivation and task performance of the monkeys. However, the schedule was never varied within a daily session. High probability options varied from $p_h = 0.8$ to $p_h = 0.7$ and low probability options were always $1 - p_h$ (i.e. 0.2 to 0.3). The reward schedule was adjusted once for Monkey E from $p_h = 0.8$ to 0.75, once for Monkey G from $p_h = 0.7$ to 0.75, and three times for Monkey M from $p_h = 0.75$ to 0.8, then to 0.78 and then back to 0.8. Across monkeys, 51.12% of the data was collected with $p_h = 0.75$, 40.09% was collected with $p_h = 0.8$, 7.77% was collected with $p_h = 0.7$, and 1% was collected $p_h = 0.78$. Thus, the most common schedule was 0.75/0.5/0.25. Schedule use did not differ by drug conditions ($F(1, 75) = 1.75, p = 0.19$). Analyses that included the reward schedule p_h as a session level covariate yielded highly similar results.

Data analysis

Choice behavior—We quantified choice behavior as the fraction of times the monkey chose either the novel choice option or the best alternative option on the first trial in which the novel option was presented. The best alternative option was defined as the remaining option with the highest assigned reward probability. In cases where the remaining options were assigned equal probabilities (these occurred < 20% of the time), the best alternative was defined as the option reinforced more often over the previous ten trials. All choice estimates were arcsine transformed prior to their use in ANOVA analyses.

Saccade reaction times—Saccadic reaction times were measured from the presentation onset of the choice options to the onset of the saccade that targeted a choice option. Before statistical analysis, reaction times were log transformed to correct for positive skew.

Reinforcement learning model—We fit a RL model to the choice behavior of the monkeys to estimate learning rates, inverse temperature and the value of novel stimuli. The model was fit separately to the choice behavior from each block of trials within a session.

The model updates the value, v_i , of a chosen option, i , based on reward feedback, r in trial t as:

$$v_i(t) = v_i(t-1) + \alpha (r(t) - v_i(t-1))$$

Thus, the updated value of an option is given by its old value, $v_i(t-1)$ plus a change based on the reward prediction error ($r(t) - v_i(t-1)$), multiplied by the learning rate parameter, α . When a novel stimulus is introduced in trial t' , there is no reward history. The initial values of novel options, when they were introduced were fit as 3 free parameters in the model, v_k^0 . The subscript k indexes the value of the option that was replaced i.e. $k \in \{p_H, p_M, p_L\}$. Therefore, we parameterized separate initial values for novel options that replaced high, medium or low valued options. Thus, whenever we introduced a novel option (independent of its value), we substituted the appropriate v_k^0 into the model, depending on the reward value of the option that was replaced, and updated this value on subsequent trials following feedback. The relative propensity of the monkeys to pick the novel option when it was introduced allowed us to estimate the value of that option relative to the other available options. The more often they picked the novel option when it was introduced, the higher the value of novel options. This is particularly true if the novel option is chosen when the other available options are of high value. The free parameters (the initial value of novel options, v_k^0 , the learning rate parameter, α and the inverse temperature, β , which estimates how consistently animals choose the highest valued option), were fit by maximizing the likelihood of the choice behavior of the participants, given the model parameters. Specifically, we calculated the choice probability $d_i(t)$ using:

$$d_i(t) = \frac{\exp(\beta v_i(t))}{\sum_{k=1}^3 \exp(\beta v_k(t))}$$

And then calculated the log-likelihood as:

$$ll = \sum_{t=1}^T \log \sum_{k=1}^3 c_k(t) d_k(t)$$

Where $c_k(t)=1$ when the subject chooses option k in trial t and $c_k(t)=0$ for all unchosen options. In other words, the model maximizes the choice probability ($d_k(t)$) of the actual choices the participants made. T is the total number of trials in the block for each monkey, usually 650. Parameters were maximized using standard techniques (Djamshidian et al., 2010). To avoid local minima, initial value and learning rate parameters were drawn from a normal distribution with a mean of 0.5 and a standard deviation of 3. The inverse temperature parameter was drawn from a normal distribution with a mean of 1 and a standard deviation of 5. Model fits were repeated 100 times to avoid local minima and the fit with the minimum log-likelihood was selected as the best fit (Table 1). No constraints were placed on the estimated parameters. Likelihood ratio tests were used to compare the fit between the estimated model and a null model that assigned novel options a fixed value of

0.5 (e.g. their empirical average reward expectation). Model fit was significantly improved relative to the null model in 92% of cases ($\chi^2(1) = 78.86, p < .001$).

Statistical analyses—Each dependent variable was entered into a full factorial, mixed effects ANOVA model implemented in MATLAB. For analyses involving choice behavior or reaction times, the model included the following fixed effects: drug (GBR-12909/ saline), choice option (novel/ best alternative), paired reward probability of the best alternative and worst alternative ($p_{LL}/p_{LM}/p_{MM}/p_{LH}/p_{MH}/p_{HH}$), reward probability of the cue replaced by the novel option ($p_L/p_M/p_H$) and reward receipt on the previous trial (reward/non-reward). Monkey identity and testing session were specified as random factors, with testing session hierarchically nested under monkey identity and drug condition. The Welch-Satterthwaite equation was used to calculate the effective denominator degrees of freedom in determining statistical significance (Neter, Wasserman, & Kutner, 1990).

ANOVA analyses of RL model parameters utilized a similar mixed effects framework but specified fewer fixed effects. This is because several choice-related factors are subsumed in fitting the RL model. Analyses of the learning rate and inverse temperature parameters only specified drug condition as a fixed effect. When comparing initial value estimates for novel options to the learned value of the best alternative option, we first averaged the model derived value estimates for the novel and best alternative options conditioned on the assigned reward probability of the cue replaced by the novel option. We then analyzed option value by specifying fixed effects of drug, choice option, and reward probability of the replaced cue. Post-hoc analyses of significant main effects used Fisher's least significant difference test to correct for multiple comparisons (Levin, Serlin, & Seaman, 1994). Post-hoc tests of significant interactions consisted of computing univariate ANOVAs for component effects and similarly correcting for multiple comparisons.

Results

Dopamine Modulates Novelty Driven Choice Behavior

We first examined how often the monkeys chose the novel option when it was introduced, relative to how often they chose the better of the two remaining alternative options (Figure 2A). The monkeys had an overall novelty preference, selecting the novel option more often than the best alternative option (Option, $F(1, 97) = 487.87, p < .001, d = 1.06$). This novelty preference was evident during both drug ($F(1, 20) = 29.82, p < .001, d = 2.44$) and saline sessions ($F(1, 72) = 66.23, p < .001, d = 1.91$). Furthermore, blocking dopamine reuptake increased the monkeys' novelty preference (Drug x Option, $F(1, 94) = 6.47, p = .012, d = .52$). Following GBR-12909 administration, the monkeys chose a greater proportion of the novel options than they did in saline sessions ($F(1, 90) = 4.02, p = .048, d = .42$). In addition, the monkeys were marginally less prone to choose the best alternative option during drug versus saline sessions ($F(1, 90) = 3.54, p = .063, d = .39$). Thus, the mean difference in how often the monkeys chose the novel or best alternative option was heightened by blocking DAT.

Since novelty seeking necessarily constitutes a switch in behavior, we also contrasted the likelihood of choosing the novel option with the likelihood of a switch in the animals' choice

when deciding between a stable set of options. On average and when choosing between options they had already seen, the monkeys switched their choice on 49.6% of trials. The likelihood of switching increased when the prior choice was unrewarded. However, animals on GBR-12909 showed no increased likelihood to switch after either reward or non-reward (Drug x Juice, $F(1, 79) < 1$, $p = .871$). We also tested if the likelihood that the monkeys would choose a novel option differed from the likelihood that they would switch their choice. The fraction of novel options chosen exceeded the fraction of trials on which the monkeys switched their choices during both drug ($M = 55.8\%$; $F(1, 19) = 5.05$, $p = .035$, $d = 1.03$) and saline ($M = 60.1\%$; $F(1, 72) = 9.04$, $p = .003$, $d = .72$) sessions. Moreover, heightened novelty seeking on GBR-12909 was evident even when the fraction of novel options selected was normalized by the fraction of switch trials ($F(1, 87) = 3.99$, $p = .048$, $d = .42$).

Next, we fit a RL model to the monkeys' choice behavior. We used this model to infer the initial value of novel options. This was done by parameterizing the initial value of novel options and optimizing these parameters to predict choice behavior. If the animals valued novel options more than the two alternative options they had already observed, the average initial value of novel options should exceed that of the best alternative option. Consistent with their choice behavior, the algorithm showed that, on average, monkeys did value novel options more than the best alternative option (Figure 2B; Option, $F(1, 85) = 4.88$, $p = .029$, $d = .15$). This effect was evident during GBR-12909 sessions ($F(1, 20) = 12.7$, $p < .001$, $d = .29$) and was marginally present in saline sessions ($F(1, 79) = 3.39$, $p = .069$, $d = .01$). Blocking dopamine reuptake potentiated the initial value assigned to novel stimuli (Drug x Option, $F(1, 375) = 5.92$, $p = .015$, $d = .25$). The monkeys assigned a higher initial value to novel cues on GBR-12909 compared to saline ($F(1, 94) = 5.33$, $p = .023$, $d = .47$). DAT blockade did not significantly decrease the mean value of the best remaining option, compared to saline ($F(1, 81) < 1$, $p = .81$). Although we conditioned the estimated initial value estimates on the reward probability of the cue replaced by the novel option (Table 1), the reward rate of the replaced cue did not significantly influence the initial value assigned to novel options (Replaced Cue Probability x Option, $F(2, 375) < 1$, $p = .477$). There were no other significant main effects or interactions on value estimates.

Choice Consistency and Exploration

Use of the RL model allowed us to additionally estimate an inverse temperature parameter, β . This parameter quantifies how consistently the animals chose the highest value option for a given distribution of values across the three choices. As β increases, choice behavior is increasingly invariant. Therefore β is often described as representing the balance between exploration and exploitation. However, exploration suggests an intentional process, and an unmotivated animal, because it is not optimizing its decisions, may also make choices inconsistent with value estimates. Furthermore, if average total reward is being maximized, greedy policies that have minimal exploration are optimal when only a small number of choices are available and the animal is familiar with the distribution of reward probabilities from which the stimuli are drawn (Averbeck et al., 2013). Despite the effects of DAT blockade on novelty seeking behavior, there were no drug related differences on the inverse temperature parameter (Figure 2C; Drug, $F(1, 134) = 1.15$, $p = .285$, $d = .18$).

We also calculated the entropy in the monkeys' choice behavior. Entropy is mathematically related to the inverse temperature and measures variability in the probability distribution describing how often the monkey chose the best, intermediate, or worst available option. Low entropy implies a peak in the probability distribution and the frequent selection of one of the options relative to the others. High entropy implies that choice behavior is more uniformly distributed. This can correspond to increased exploration or decreased learning. Choice entropy was highest on trials proximal to experiencing a novel option, and decreased as the monkeys became familiar with a set of options (Trials, $F(1,94) = 8.39$, $p < .001$, $d = .59$; Figure 2D). This is consistent with the animals learning to select the best available option in the array. Choice entropy did not differ between drug and saline sessions, overall (Drug, $F(1,87) = 1.24$, $p = .267$, $d = 0.23$), or in time (Drug x Trials, $F(1,94) < 1$, $p = .678$). Thus, based on the inverse temperature parameter and choice entropy, animals tended to sample sub-optimal choice options equally often, both on and off drug.

Additional Reward Related Effects on Novelty Seeking

Returning to the analysis of novelty driven choice behavior, we found that the monkeys' overall novelty preference scaled with three additional factors: the opportunity costs associated with declining the two remaining options, the reward probability of the cue replaced by the novel option, and the choice outcome on the previous trial (Figure 3). None of these reward related effects were altered by inhibiting DAT (i.e. all interactions involving drug condition were non-significant at $p > .43$).

Opportunity costs—As the reward probability of the best alternative option increased, there was a progressive decline in the monkeys' novelty preference (Reward Environment x Option, $F(5, 419) = 10.14$, $p < .001$, $d = .31$). There was a greater likelihood that the monkeys would select the novel option when the two remaining alternatives were rewarded at a low rate, compared to when they were both rewarded at a high rate (Figure 3A). Selection of the best alternative option mirrored this trend. The animals showed a consistent novelty preference when we analyzed the difference in the likelihood of choosing either the novel or best alternative option (all $p < .002$). The animals were biased to select the novel option even when both alternative options were rewarded at the maximal rate and the novel option, due to task constraints, had to be of lesser value ($F(1, 81) = 8.19$, $p = .002$, $d = .63$). In other words novelty seeking scaled with the cost of declining the best alternative option even when monkeys became more novelty prone as a result of DAT inhibition.

Replaced cue probability—The monkeys were more likely to select the novel option when it replaced a high or medium probability reward cue compared to when it replaced a low probability cue (Replaced Cue Probability x Option, $F(2, 177) = 8.77$, $p < .001$, $d = .44$; Figure 3B). Likewise the monkeys chose the best alternative option more often when it replaced a low versus a high or medium probability cue. Again, this effect was consistent across drug and saline sessions.

Prior choice outcome—The monkeys were more likely to select the novel option if their choice on the previous trial went unrewarded (Juice x Option, $F(1, 101) = 6.03$, $p = .015$, d

= .48; Figure 3C), whereas reward receipt on the previous trials increased the likelihood that the monkeys chose the best alternative option.

Learning and Value Updating

We compared model derived learning rate parameters across drug conditions, to test whether blocking dopamine reuptake altered the rate at which the monkeys updated the value of each choice option. Overall the learning rate parameter was positive ($M = .157$, $SE = 0.018$) and differed significantly from zero ($t(81) = 8.78$, $p < .001$, $d = 1.95$). This indicated the monkeys had learned to discriminate between high and low value options. There were no differences in the learning rate parameter between the two drug conditions (Figure 4A; Drug, $F(1,115) < 1$, $p = .577$). Thus, DAT blockade with GBR-12909 did not affect the rate at which the animals learned to select the best valued option, compared to saline injections.

To further examine learning, we also compared the fraction of times the monkeys continued to choose a novel option based on its assigned reward probability, up to 20 trials from its initial presentation (Figure 4B). As the monkeys gained experience with a novel option their continued selection of that option reflected its assigned reward probability (Reward Probability x Trial, $F(2, 210) = 40.42$, $p < .001$, $d = .88$). On average across drug and saline conditions, the monkeys learned to select high probability options more often than they chose medium probability cues ($t(442) = 3.96$, $p < .001$, $d = .37$), and to choose medium probability cues more often than low probability cues ($t(185) = 2.75$, $p = .007$, $d = .41$). There were no drug related differences in the ability of the monkeys to discriminate between the high, medium or low probability options (Drug x Reward Probability, $F(2,69) < 1$, $p = .986$; Drug x Reward Probability x Trial, $F(2,69) < 1$, $p = .995$).

We also compared the average amount of reward earned per block for each drug conditions. Equivalent amounts of juice reward were earned on GBR-12909 ($M = 68.68$ mL $SE = 1.65$) and saline ($M = 69.83$ mL $SE = 1.62$; Drug, $F(1,97) = 2.38$, $p = .134$).

Psychostimulant and Reaction Time Effects

We did not find drug related differences in task performance in terms of the number of errors made initiating a trial, holding baseline fixation, or holding choice selection (all $p > .45$). Overall the monkeys' saccadic choice reactions were faster on GBR-12909 versus saline (Drug, $F(1,29) = 12.86$, $p = .001$, $d = 1.33$). When we separately analyzed trials on which a novel option was introduced, choice reactions times were similarly faster on GBR-12909 ($M_{RT} = 165.68$ ms, $SE = 3.76$) than on saline ($M_{RT} = 176.26$ ms, $SE = 3.92$; $F(1,29) = 8.57$, $p = .006$, $d = 1.09$). Otherwise there were no additional effects on the monkeys' reaction times (all $p > .15$).

Discussion

To determine if dopamine enhances novelty seeking, we systemically inhibited the dopamine transporter (DAT) in monkeys and examined how this pharmacological manipulation affected the monkeys' novelty preference. We found that DAT blockade increased the initial value monkeys assigned to novel options, biasing them to select novel over familiar choice options. This dopamine dependent increase in novelty seeking behavior did not reflect

differences in the rate at which monkeys learned to select cues predictive of reward or their tendency to exploit that knowledge, measured by the inverse temperature and choice entropy.

Dopaminergic and Novelty Driven Decision Making

The present study was motivated by prior work showing that Parkinson's disease patients that develop behavioral addictions following treatment with dopamine agonists, are more prone to select novel choice options in a closely matched decision making task (Averbeck et al., 2013; Djamshidian et al., 2011). Pharmacological imaging data in these patients suggests that they have low levels of DAT in their ventral striatum (Cilia et al., 2010), compared to Parkinson's patients without behavioral addictions. Related studies with non-Parkinsonian patients diagnosed with behavioral addictions similarly report lower levels of DAT in the striatum (Chang, Alicata, Ernst, & Volkow, 2007; Leroy et al., 2012), and elevated questionnaire based reports of novelty and sensation seeking (Kreek, Nielsen, Butelman, & LaForge, 2005; Wills, Vaccaro, & McNamara, 1994). Here we explicitly demonstrate that reduced dopamine reuptake enhances novelty-based choice behavior during reinforcement learning in monkeys. This result, viewed together with prior clinical evidence, suggests dopamine reuptake capacity may contribute to excessive novelty seeking.

Evidence that blocking DAT heightens the reward value of novel images and therefore promotes exploration also provides direct support for the theory that novelty is intrinsically valued despite choice uncertainty (Hazy et al., 2010; Kakade & Dayan, 2002; Wittmann et al., 2008). Until now, this view has only received indirect support from studies that operationalized novelty seeking as a non-decision process. Instead, past studies generally investigated novelty processing in terms of heightened attention to passive encounters with rare or deviant stimuli (Bromberg-Martin et al., 2010), or the active exploration of singular novel objects placed in an open field (Dulawa et al., 1999; Zhuang et al., 2001). In an fMRI study that did examine novelty seeking as a decision process (Wittmann et al., 2008)—using a version of the current task—activation in ventral striatum encoded both standard reward prediction errors and enhanced reward prediction errors during novelty driven exploration. However, the assumption that such signals are dopamine related relies on the inference that striatal BOLD activity has a dopaminergic basis.

We manipulated dopamine signaling by injecting GBR-12909, a selective inhibitor of DAT. DAT inhibition with GBR-12909 is known to increase tonic dopamine levels in the striatum by slowing reuptake (Zhang et al., 2009). Since we chose to block systemically we cannot say where critical dopaminergic changes occurred in the brain. A likely candidate is the ventral striatum, considering prior evidence of reduced DAT in the striatum of novelty prone patients (Cilia et al., 2010), and enhanced encoding of novelty driven prediction errors in the ventral striatum (Wittmann et al., 2008). Yet a single dose of cocaine—which similarly blocks the action of the dopamine transporter—is shown to alter glucose metabolism in the ventral striatum, orbitofrontal cortex, entorhinal cortex, and hippocampus (Lyons, Friedman, Nader, & Porrino, 1996). DAT is also present in most axons containing tyrosine hydroxylase in the cortex of primates (Lewis et al., 2001). Future studies involving microdialysis, voltammetry, or single unit recordings paired with systemic DAT blockade may aid in

clarifying exactly which features of dopamine transmission contribute to increases in novelty seeking.

In conjunction with DAT, catechol-o-methyltransferase (COMT) also regulates dopamine metabolism. It is plausible that reducing COMT might similarly heighten novelty-based choice behavior, as seen here. For example, infusion of the COMT inhibitor tolcapone into the dorsal hippocampus of rats increases their exploration of novel arms in a spatial novelty preference task (Laatikainen, Sharp, Bannerman, Harrison, & Tunbridge, 2012). Also, novelty seeking personality traits are associated with genetic polymorphisms of both DAT and COMT (Golimbet, Alfimova, Gritsenko, & Ebstein, 2007).

Dopamine and Behavioral Exploration

Besides examining how dopamine biased novelty driven exploration, we also examined if blocking DAT caused animals to sample sub-optimal stimuli more often, quantified by the inverse temperature parameter, β in the RL algorithm. In a previous experiment comparing transgenic and wild type mice, DAT knockdown mice showed increased exploration of an infrequently rewarded option. This behavior caused a decrease in β when RL models were fit to the animals' choice behavior (Beeler, Daw, Frazier, & Zhuang, 2010). We did not find a significant reduction in the inverse temperature parameter during drug sessions. While there was an average decrease in the inverse temperature in all three monkeys, this result should not be over interpreted since related measures of exploration were not modulated by DAT blockade. Monkeys were equally likely to switch their choice from the previous trial on GBR-12909 or saline. Choice entropy was also equivalent on and off drug. Together these results suggest blocking DAT had no effect on how consistently the monkeys chose to exploit learned values. Instead, greater exploration following DAT blockade was restricted to the monkeys' initial encounters with novelty.

Dopamine, Novelty Seeking and Foraging

Novelty seeking relies on predictions about the relative value of available resources in order to select among novel and familiar items. This closely resembles descriptions of foraging behavior. Foraging relies on the optimization of exploratory behavior in order to maximize the long-term rate of return, taking into account the cost of foraging and resource availability (Stephens & Krebs, 1986). Previous studies of foraging behavior in humans and monkeys have identified several brain regions, including anterior cingulate cortex (ACC), that encode the average value of the reward environment as well as the cost of foraging (Hayden, Pearson, & Platt, 2011; Kolling, Behrens, Mars, & Rushworth, 2012; Sugrue, Corrado, & Newsome, 2004). Dopamine is thought to contribute to foraging behavior based on its role in feeding and goal-directed behavior (Hills, 2006). Computational models of instrumental behavior following dopaminergic manipulations also posit that tonic dopamine in the striatum encodes the average reward rate to regulate opportunity costs and response vigor (Niv, Daw, Joel, & Dayan, 2007)

This could lead to the prediction that heightened novelty seeking on GBR-12909 results from altered sensitivity to task parameters relevant to novelty seeking. However, opportunity costs, and priors for the reward history of the replaced cue and immediately preceding

choice, all influenced novelty seeking in the same way, on and off drug. So DAT blockade apparently increases novelty seeking independent of contextual factors. For example, in the present study, one of our animals (Monkey G), on average, avoided choosing the novel option unless the opportunity costs were low (e.g. the best alternative was rewarded at a low or medium rate). Even in this animal GBR-12909 generally enhanced the selection and valuation of novel options.

One preliminary interpretation is that dopamine modulates neuronal thresholds in circuits regulating value based decision making. For example, in a patch foraging task neurons in ACC gradually increase their firing rate over repeated decisions to harvest a patch with diminishing returns (Hayden et al., 2011). Decisions to seek out a new patch are predicted by a threshold crossing in the firing rate of these cells, and this threshold is influenced by foraging costs. Dopamine might similarly influence threshold computations related to novelty seeking, possibly by modulating the excitability of neurons (Seamans & Yang, 2004) forming circuits that regulate exploration based on current biological needs (Murray & Rudebeck, 2013). Excessive novelty seeking may then occur when dopamine overwhelms the ability of relevant circuits to optimize exploratory behavior.

Dopamine and Learning

Considering that dopamine manipulations can modulate reinforcement learning (Djamshidian et al., 2010; Frank, Seeberger, & O'Reilly R, 2004; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006), it is important to point out that DAT blockade did not affect the ability of the monkeys to discriminate high, medium, and low value options, their learning rates, or the total amount of reward they earned per block. Thus, dopamine related increases in novelty preference could not be attributed to differences in the value updating processes that drive learning. Modulation of dopaminergic tone through genetic knockdown of DAT (Beeler et al., 2010) or tyrosine hydroxylase similarly have no impact on learning ability (Robinson, Sandstrom, Denenberg, & Palmiter, 2005). Also, Parkinson's disease patients that were novelty prone in a version of the current task, learn at the same rate as healthy controls, both on and off their dopamine medication (Averbeck et al., 2013; Djamshidian et al., 2011; Housden, O'Sullivan, Joyce, Lees, & Roiser, 2010). These results suggest that equivalent learning rates on and off GBR-12909 might be attributable to the manipulation of tonic rather than phasic dopamine signaling.

Conclusion

We found that monkeys show a preference for novel vs. familiar choice options during a probabilistic decision making task. Systemic blockade of the dopamine transporter with GBR-12909 caused monkeys to become more novelty prone. When novel options were first encountered, DAT blockade led monkeys to optimistically value and over select novel options relative to the best alternative and familiar option. These findings demonstrate that increases in extracellular dopamine levels underlie the positive valuation of novel stimuli in order to promote exploratory behavior. They also suggest that alterations in dopamine reuptake may contribute to excessive novelty seeking and impulsivity.

Acknowledgments

We thank O. Dal Monte, A. Mitz, and S. Gallagher for technical assistance. This research was supported by the Intramural Research Program of the NIH, National Institute of Mental Health (NIMH).

References

- Asaad WF, Eskandar EN. A flexible software tool for temporally-precise behavioral control in Matlab. *J Neurosci Methods*. 2008; 174(2):245–258. DOI: 10.1016/j.jneumeth.2008.07.014 [PubMed: 18706928]
- Averbeck BB, Djamshidian A, O’Sullivan SS, Housden CR, Roiser JP, Lees AJ. Uncertainty about mapping future actions into rewards may underlie performance on multiple measures of impulsivity in behavioral addiction: Evidence from Parkinson’s disease. *Behavioral neuroscience*. 2013; 127(2): 245–255. DOI: 10.1037/a0032079 [PubMed: 23565936]
- Beeler JA, Daw N, Frazier CRM, Zhuang XX. Tonic dopamine modulates exploitation of reward learning. *Frontiers in Behavioral Neuroscience*. 2010; 4doi: 10.3389/Fnbeh.2010.00170
- Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*. 2010; 68(5):815–834. DOI: 10.1016/j.neuron.2010.11.022 [PubMed: 21144997]
- Bunzeck N, Dayan P, Dolan RJ, Duzel E. A common mechanism for adaptive scaling of reward and novelty. *Hum Brain Mapp*. 2010; 31(9):1380–1394. DOI: 10.1002/hbm.20939 [PubMed: 20091793]
- Bunzeck N, Duzel E. Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron*. 2006; 51(3):369–379. DOI: 10.1016/J.Neuron.2006.06.021 [PubMed: 16880131]
- Chang L, Alicata D, Ernst T, Volkow N. Structural and metabolic brain changes in the striatum associated with methamphetamine abuse. *Addiction*. 2007; 102(Suppl 1):16–32. DOI: 10.1111/j.1360-0443.2006.01782.x [PubMed: 17493050]
- Cilia R, Ko JH, Cho SS, van Eimeren T, Marotta G, Pellecchia G, ... Strafella AP. Reduced dopamine transporter density in the ventral striatum of patients with Parkinson’s disease and pathological gambling. *Neurobiology of disease*. 2010; 39(1):98–104. DOI: 10.1016/j.nbd.2010.03.013 [PubMed: 20338240]
- Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Phil Trans R Soc*. 2007; 362(1481):933–942.
- Czoty PW, Justice JB Jr, Howell LL. Cocaine-induced changes in extracellular dopamine determined by microdialysis in awake squirrel monkeys. *Psychopharmacology (Berl)*. 2000; 148(3):299–306. [PubMed: 10755743]
- Dalley JW, Everitt BJ, Robbins TW. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron*. 2011; 69(4):680–694. DOI: 10.1016/J.Neuron.2011.01.020 [PubMed: 21338879]
- Djamshidian A, Jha A, O’Sullivan SS, Silveira-Moriyama L, Jacobson C, Brown P, ... Averbeck BB. Risk and Learning in Impulsive and Nonimpulsive Patients With Parkinson’s Disease. *Movement Disorders*. 2010; 25(13):2203–2210. DOI: 10.1002/Mds.23247 [PubMed: 20721918]
- Djamshidian A, O’Sullivan SS, Wittmann BC, Lees AJ, Averbeck BB. Novelty seeking behaviour in Parkinson’s disease. *Neuropsychologia*. 2011; 49(9):2483–2488. DOI: 10.1016/J.Neuropsychologia.2011.04.026 [PubMed: 21565210]
- Dulawa SC, Grandy DK, Low MJ, Paulus MP, Geyer MA. Dopamine D4 receptor-knock-out mice exhibit reduced exploration of novel stimuli. *J Neurosci*. 1999; 19(21):9550–9556. [PubMed: 10531457]
- Eriksson O, Langstrom B, Josephsson R. Assessment of receptor occupancy-over-time of two dopamine transporter inhibitors by [C-11]CIT and target controlled infusion. *Upsala Journal of Medical Sciences*. 2011; 116(2):100–106. DOI: 10.3109/03009734.2011.563878 [PubMed: 21443419]
- Frank MJ, Seeberger LC, O’Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*. 2004; 306(5703):1940–1943. DOI: 10.1126/science.1102941 [PubMed: 15528409]

- Gittins JC. Bandit Processes and Dynamic Allocation Indexes. *Journal of the Royal Statistical Society Series B-Methodological*. 1979; 41(2):148–177.
- Glowa JR, Wojnicki FHE, Matecka D, Rice KC, Rothman RB. Effects of dopamine reuptake inhibitors on food- and cocaine-maintained responding: II. Comparisons with other drugs and repeated administrations. *Experimental and Clinical Psychopharmacology*. 1995; 3(3):232–239. DOI: 10.1037/1064-1297.3.3.232
- Golimbet VE, Alifimova MV, Gritsenko IK, Ebstein RP. Relationship between dopamine system genes and extraversion and novelty seeking. *Neurosci Behav Physiol*. 2007; 37(6):601–606. DOI: 10.1007/s11055-007-0058-8 [PubMed: 17657431]
- Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci*. 2011; 14(7):933–939. DOI: 10.1038/nn.2856 [PubMed: 21642973]
- Hazy TE, Frank MJ, O'Reilly RC. Neural mechanisms of acquired phasic dopamine responses in learning. *Neurosci Biobehav Rev*. 2010; 34(5):701–720. DOI: 10.1016/j.neubiorev.2009.11.019 [PubMed: 19944716]
- Hills TT. Animal foraging and the evolution of goal-directed cognition. *Cognitive Science*. 2006; 30(1):3–41. DOI: 10.1207/S15516709cog0000_50 [PubMed: 21702807]
- Horvitz JC. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*. 2000; 96(4):651–656. DOI: 10.1016/S0306-4522(00)00019-1 [PubMed: 10727783]
- Housden CR, O'Sullivan SS, Joyce EM, Lees AJ, Roiser JP. Intact Reward Learning but Elevated Delay Discounting in Parkinson's Disease Patients With Impulsive-Compulsive Spectrum Behaviors. *Neuropsychopharmacology*. 2010; 35(11):2155–2164. DOI: 10.1038/Npp.2010.84 [PubMed: 20631686]
- Kahlig KM, Galli A. Regulation of dopamine transporter function and plasma membrane expression by dopamine, amphetamine, and cocaine. *Eur J Pharmacol*. 2003; 479(1–3):153–158. [PubMed: 14612146]
- Kakade S, Dayan P. Dopamine: generalization and bonuses. *Neural Networks*. 2002; 15(4–6):549–559. Pii S0893-6080(02)00048-5. DOI: 10.1016/S0893-6080(02)00048-5 [PubMed: 12371511]
- Kolling N, Behrens TE, Mars RB, Rushworth MF. Neural mechanisms of foraging. *Science*. 2012; 336(6077):95–98. DOI: 10.1126/science.1216930 [PubMed: 22491854]
- Kreek MJ, Nielsen DA, Butelman ER, LaForge KS. Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nature Neuroscience*. 2005; 8(11):1450–1457. DOI: 10.1038/nn1583 [PubMed: 16251987]
- Laatikainen LM, Sharp T, Bannerman DM, Harrison PJ, Tunbridge EM. Modulation of hippocampal dopamine metabolism and hippocampal-dependent cognitive function by catechol-O-methyltransferase inhibition. *Journal of Psychopharmacology*. 2012; 26(12):1561–1568. DOI: 10.1177/0269881112454228 [PubMed: 22815336]
- Leroy C, Karila L, Martinot JL, Lukasiewicz M, Duchesnay E, Comtat C, ... Trichard C. Striatal and extrastriatal dopamine transporter in cannabis and tobacco addiction: a high-resolution PET study. *Addiction biology*. 2012; 17(6):981–990. DOI: 10.1111/j.1369-1600.2011.00356.x [PubMed: 21812871]
- Levin JR, Serlin RC, Seaman MA. A Controlled, Powerful Multiple-Comparison Strategy for Several Situations. *Psychological Bulletin*. 1994; 115(1):153–159. DOI: 10.1037//0033-2909.115.1.153
- Lewis DA, Melchitzky DS, Sesack SR, Whitehead RE, Auh S, Sampson A. Dopamine transporter immunoreactivity in monkey cerebral cortex: regional, laminar, and ultrastructural localization. *J Comp Neurol*. 2001; 432(1):119–136. [PubMed: 11241381]
- Lyons D, Friedman DP, Nader MA, Porrino LJ. Cocaine alters cerebral metabolism within the ventral striatum and limbic cortex of monkeys. *J Neurosci*. 1996; 16(3):1230–1238. [PubMed: 8558251]
- Mitz AR. A liquid-delivery device that provides precise reward control for neurophysiological and behavioral experiments. *J Neurosci Methods*. 2005; 148(1):19–25. doi: <http://dx.doi.org/10.1016/j.jneumeth.2005.07.012>. [PubMed: 16168492]
- Murray EA, Rudebeck PH. The drive to strive: goal generation based on current needs. *Front Neurosci*. 2013; 7:112.doi: 10.3389/fnins.2013.00112 [PubMed: 23818876]
- Neter, John, Wasserman, William, Kutner, Michael H. *Applied linear statistical models: regression, analysis of variance, and experimental designs*. 3. Homewood, IL: Irwin; 1990.

- Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 2007; 191(3):507–520. DOI: 10.1007/s00213-006-0502-4 [PubMed: 17031711]
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 2006; 442(7106):1042–1045. [PubMed: 16929307]
- Rangel-Gomez M, Hickey C, van Amelsvoort T, Bet P, Meeter M. The Detection of Novelty Relies on Dopaminergic Signaling: Evidence from Apomorphine's Impact on the Novelty N2. *Plos One*. 2013; 8(6) ARTN e66469. doi: 10.1371/journal.pone.0066469
- Reed P, Mitchell C, Nokes T. Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Animal Learning & Behavior*. 1996; 24(1):38–45. DOI: 10.3758/Bf03198952
- Robinson S, Sandstrom SM, Denenberg VH, Palmiter RD. Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav Neurosci*. 2005; 119(1):5–15. DOI: 10.1037/0735-7044.119.1.5 [PubMed: 15727507]
- Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in Neurobiology*. 2004; 74(1):1–57. DOI: 10.1016/J.Pneurobio.2004.05.006 [PubMed: 15381316]
- Sogaard U, Michalow J, Butler B, Laursen AL, Ingersen SH, Skrumsager BK, Rafaelsen OJ. A Tolerance Study of Single and Multiple Dosing of the Selective Dopamine Uptake Inhibitor Gbr-12909 in Healthy-Subjects. *International Clinical Psychopharmacology*. 1990; 5(4):237–251. [PubMed: 2150527]
- Stephens, David W., Krebs, JR. Foraging theory. Princeton, N.J: Princeton University Press; 1986.
- Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science*. 2004; 304(5678):1782–1787. DOI: 10.1126/science.1094765 [PubMed: 15205529]
- Tsukada H, Harada N, Nishiyama S, Ohba H, Kakiuchi T. Dose-response and duration effects of acute administrations of cocaine and GBR12909 on dopamine synthesis and transporter in the conscious monkey brain: PET studies combined with microdialysis. *Brain Res*. 2000; 860(1–2):141–148. [PubMed: 10727633]
- Villemagne VL, Rothman RB, Yokoi F, Rice KC, Matecka D, Dannals RF, Wong DF. Doses of GBR12909 that suppress cocaine self-administration in non-human primates substantially occupy dopamine transporters as measured by [C-11] WIN35,428 PET scans. *Synapse*. 1999; 32(1):44–50. DOI: 10.1002/(Sici)1098-2396(199904)32:1<44::Aid-Syn6>3.0.Co;2-9 [PubMed: 10188637]
- Willenbockel V, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW. Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*. 2010; 42(3):671–684. DOI: 10.3758/Brm.42.3.671 [PubMed: 20805589]
- Wills TA, Vaccaro D, McNamara G. Novelty seeking, risk taking, and related constructs as predictors of adolescent substance use: an application of Cloninger's theory. *Journal of substance abuse*. 1994; 6(1):1–20. [PubMed: 8081104]
- Wilson FA, Goldman-Rakic PS. Viewing preferences of rhesus monkeys related to memory for complex pictures, colours and faces. *Behav Brain Res*. 1994; 60(1):79–89. [PubMed: 8185855]
- Wittmann BC, Daw ND, Seymour B, Dolan RJ. Striatal activity underlies novelty-based choice in humans. *Neuron*. 2008; 58(6):967–973. DOI: 10.1016/j.neuron.2008.04.027 [PubMed: 18579085]
- Zhang L, Doyon WM, Clark JJ, Phillips PE, Dani JA. Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. *Mol Pharmacol*. 2009; 76(2):396–404. DOI: 10.1124/mol.109.056317 [PubMed: 19460877]
- Zhuang X, Oosting RS, Jones SR, Gainetdinov RR, Miller GW, Caron MG, Hen R. Hyperactivity and impaired response habituation in hyperdopaminergic mice. *Proc Natl Acad Sci U S A*. 2001; 98(4):1982–1987. DOI: 10.1073/pnas.98.4.1982 [PubMed: 11172062]

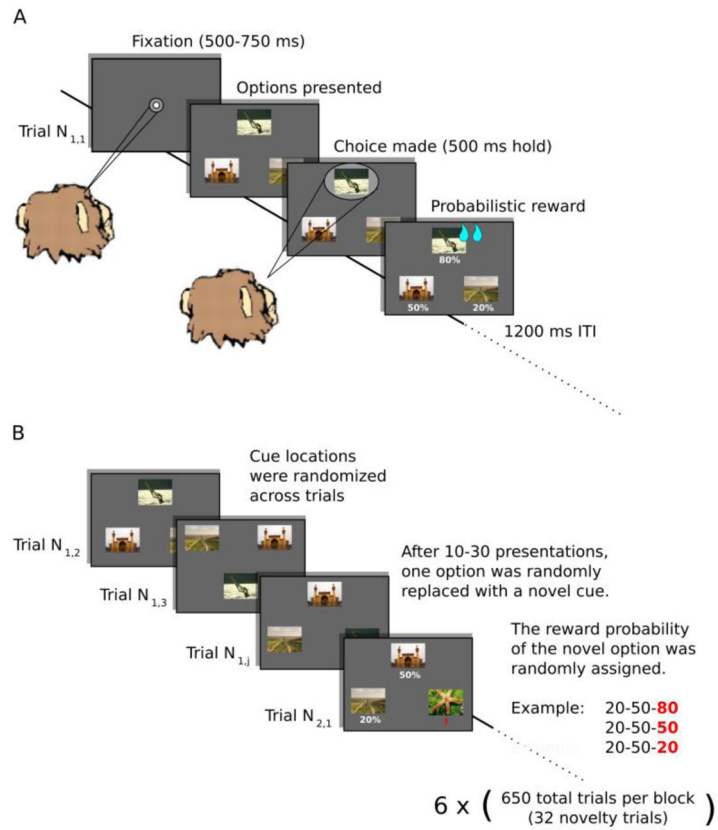


Figure 1. Task design

A) Sequence of events on a single trial. Following an initial fixation period, the choice options were presented and the animals made a saccade to one of the options to indicate their choice. A juice reward was then delivered probabilistically based on the assigned reward probability of the chosen cue. B) Novelty manipulation. A set of options was repeatedly presented for 10–30 trials, after which one of the existing options was replaced with a novel option. The novel option was randomly assigned its own reward probability. Novel options were introduced 32 times in a block of 650 trials. Trial $N_{i,j}$ refers to the number of times, j , a particular set of choice options, i , was seen.

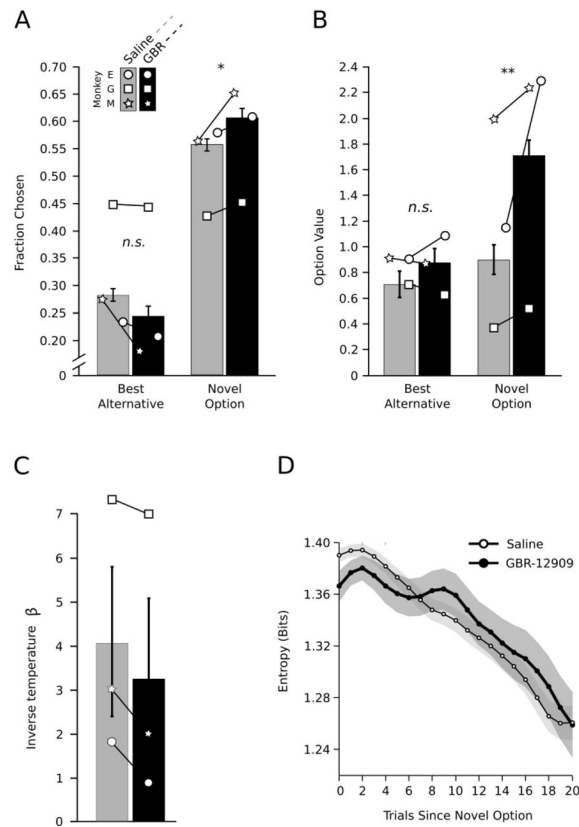


Figure 2. DAT blockade increases novelty seeking behavior and the initial value of novel options
 A) Mean fraction of times each option was chosen. B) Mean initial value (v_i^0) of novel options and the mean value of the best alternative when the novel option was introduced, as estimated by the RL model for drug and saline sessions. C and D) The inverse temperature parameter β and choice entropy H —which summarize the overall balance in explore/exploit behavior—did not differ between the drug and saline conditions.

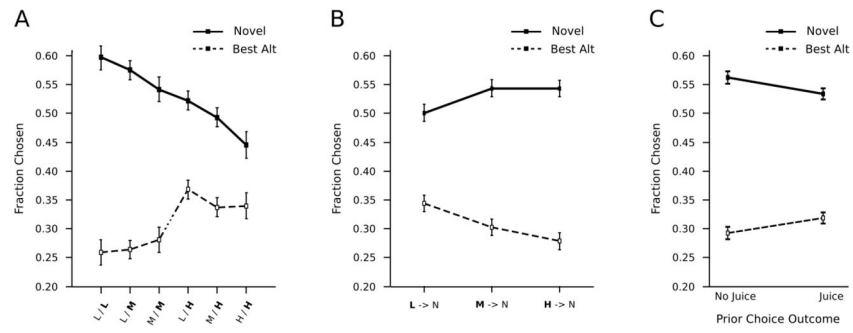


Figure 3. Reward related effects on novelty seeking

A) Fraction of times the novel option or best alternative option was selected broken out by A) the combined reward probabilities of the two alternative options, B) the reward probability of the cue replaced by the novel option, and C) reward receipt on the immediately preceding trial.

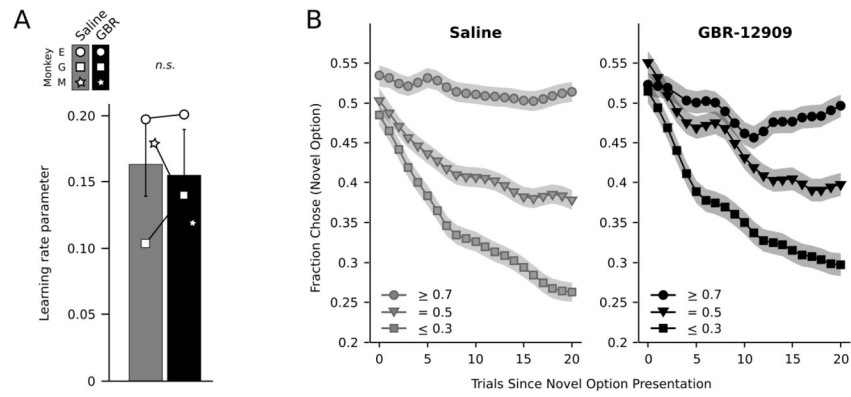


Figure 4. Dopamine does not affect learning rates

A) Learning rate parameter fit to blocks completed on GBR-12909 or saline. B) Fraction of times a novel option, with an assigned reward probability, was selected on GBR-12909 or saline up to 20 trials after its initial presentation. Choice data were smoothed with a 6 trial moving average kernel.

Table 1

Mean parameter estimates for reinforcement learning models fit to choice behavior for drug and saline sessions

	GBR-12909		Saline	
v_i^{ρ} pL	1.41	(0.19)	0.75	(0.19)
v_i^{ρ} pM	1.78	(0.25)	0.82	(0.18)
v_i^{ρ} pH	1.86	(0.44)	0.91	(0.18)
α	0.159	(0.06)	0.165	(0.05)
β	3.34	(0.36)	4.03	(0.38)
H	1.33	(0.10)	1.36	(0.10)
$\chi^2(3)^a$	63.73	(6.43)	58.86	(5.61)

^aCandidate model was tested for improved fit against a nested null model that fixed v_i^{ρ} at 0.5 using a critical value $\chi^2(3) = 7.81$, $p < .05$. The number of blocks showing a significant improvement in fit (85.3%) did not differ by drug ($\chi^2(1) = 0.68$, $p = .406$).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript