Cogn Affect Behav Neurosci (2012) 12:513–526
DOI 10.3758/s13415-012-0097-7

Interactions between deliberation and delay-discounting in rats

Andrew E. Papale · Jeffrey J. Stott · Nathaniel J. Powell · Paul S. Regier · A. David Redish

Published online: 16 May 2012

Abstract When faced with decisions, rats sometimes pause and look back and forth between possible alternatives, a phenomenon termed *vicarious trial and error* (VTE). When it was first observed in the 1930s, VTE was theorized to be a mechanism for exploration. Later theories suggested that VTE aided the resolution of sensory or neuroeconomic conflict. In contrast, recent neurophysiological data suggest that VTE reflects a dynamic search and evaluation process. These theories make unique predictions about the timing of VTE on behavioral tasks. We tested these theories of VTE on a T-maze with return rails, where rats were given a choice between a smaller reward available after one delay or a larger reward available after an adjustable delay. Rats showed three clear phases of behavior on this task: *investigation*, characterized by discovery of task parameters; *titration*, characterized by iterative adjustment of the delay to a preferred interval; and *exploitation*, characterized by alternation to hold the delay at the preferred interval. We found that VTE events occurred during adjustment laps more often than during alternation laps. Results were incompatible with theories of VTE as an exploratory behavior, as reflecting sensory conflict, or as a simple neuroeconomic valuation process. Instead, our results were most consistent with VTE as reflecting a search process during deliberative decision making. This pattern of VTE that we observed is reminiscent of current navigational theories proposing a transition from a deliberative to a habitual decision-making mechanism.

A. E. Papale · J. J. Stott · N. J. Powell · P. S. Regier
Graduate Program in Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA

A. D. Redish (✉)
Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA
e-mail: redish@umn.edu

Keywords Decision-making · Delay-discounting · Impulsivity · Vicarious trial and error · VTE · Reinforcement learning

Introduction

When rats are faced with difficult choices, they sometimes pause and look back and forth down the possible paths, a behavioral process identified in the 1930s as *vicarious trial and error* (VTE; Muenzinger, 1938; Muenzinger & Gentry, 1931; Tolman, 1938, 1948). The terminology adopted by Muenzinger, Gentry, and Tolman implies that this pause-and-look behavior entails an imagination process—specifically, representing and evaluating future possibilities. While it was impossible in the 1930s to directly test this, recent neurophysiological experiments have determined that during these pause-and-look VTE events, place cell representations in the hippocampus sweep forward ahead of the rat down the potential future paths (Johnson & Redish, 2007). Reward-related cells in ventral striatal areas receiving hippocampal input show covert representations of reward (van der Meer & Redish, 2009, 2010), and reward-related cells in the orbitofrontal cortex reflect the expected outcomes (Steiner & Redish, 2010). These neurophysiological data suggest a strong relationship between VTE and model-based reinforcement learning algorithms (Daw, Niv, & Dayan, 2005; Johnson, van der Meer, & Redish, 2007; Niv, Joel, & Dayan, 2006; van der Meer, Kurth-Nelson, & Redish, in press). In humans, the hippocampus is critical for the imagination of future possibilities during deliberation (Buckner & Carroll, 2007; Hassabis, Kumaran, Vann, & Maguire, 2007) and, perhaps, during evaluation of discounted value (Peters & Büchel, 2010).

Early behavioral theories of VTE suggested that VTE occurred during investigation of alternatives (Tolman,

1948), while later theories suggested that VTE occurred as a result of conditioned orienting (Bower, 1959; Spence, 1960). Recently, Krajbich, Armel, and Rangel (2010) observed saccade–fixate–saccade (SFS) sequences in humans making decisions between snack foods; these human sequences share similar properties to VTE in rats. Subjects showed more SFS when the value between the choices was equal, suggesting an explanation for VTE based on an underlying neuroeconomic valuation process (Glimcher, Camerer, & Poldrack, 2008; Krajbich et al., 2010).

Here, we examine the behavioral timing of vicarious trial and error on a spatial delay-discounting task that dissociates exploratory, conditioned orienting, and value equalization explanations. We find all three hypotheses incompatible with the timing of VTE behaviors. Instead, we find that VTE on this task occurs during transient changes in choice behavior and is most consistent with a theory based on gathering information using a search-through-possibilities value calculation algorithm (Johnson et al., 2007; Johnson, Varberg, Benhardus, Maahs, & Schrater, 2012; Niv, Joel, & Dayan, 2006).

The spatial delay-discounting task

Delay-discounting experiments measure choices made between taking a smaller reward sooner versus waiting for a larger one later. In humans, the ability to wait for a larger reward is related to IQ (Burks, Carpenter, Goette, & Rustichini, 2009) and college SAT scores (Mischel & Underwood, 1974) and is diminished in addiction (Giordano et al., 2002; Madden, Petry, Badger, & Bickford, 1997; Mitchell, 2004; Odum, Madden, & Bickel, 2002; Petry, Bickel, & Arnett, 1998). Similarly, rats exposed to drug self-administration paradigms discount at higher rates than do unexposed rats (Paine, Dringenberg, & Olmstead, 2003), and rats who discount faster are more susceptible to drug acquisition and reinstatement in self-administration paradigms (Perry, Larson, German, Madden, & Carroll, 2005; Perry, Nelson, & Carroll, 2008).

Delay discounting in nonhuman animals has been primarily studied through the adjusting delay procedure (Madden & Johnson, 2010; Mazur, 1997, 2001). In these tasks, animals are given two choices (usually levers to press or holes to nose-poke into). Selecting one choice provides a small reward immediately, while selecting the other choice provides a large reward after a delay. In this task, the delay to the larger option is increased when the delayed option is selected and decreased when the nondelayed option is selected. Conceptually, selecting the larger-later option implies that its discounted value is larger than the value of the smaller-sooner option. Increasing the delay to the larger option decreases the discounted value of the larger option. Conversely, selecting the smaller-sooner option implies that the discounted value of the larger-later option is less than that of the smaller-sooner option. Decreasing the delay to the larger-later option increases the discounted value of the larger-later option, by shortening the time to reward (Mazur, 2001).

These animal delay-discounting experiments are usually presented in compound blocks of four leverpress choices (Cardinal, Daw, Robbins, & Everitt, 2002; Mazur, 1997; Simon et al., 2010): First, the animal is given two forced choice trials informing the animal of the two delays, and then the animal is given two free choice trials that drive changes in the delay to the larger-later option. Choosing the smaller-sooner option in both free choice trials within a block decreases the delay to the larger-later option, while choosing the larger-later option in both free choices increases the delay to the larger-later option. Theoretically, this should produce titration to a delay in which the discounted values are matched. However, rats do not actually titrate to a consistent delay on these tasks but show large swings in delay (Cardinal et al., 2002; Valencia Torres, da costa Araujo, Sanchez, Body, Bradshaw & Szabadi, 2011). In this article, we present a spatial T-maze version of the adjusting delay-discounting task. Because rats naturally prefer to alternate between options (spontaneous spatial alternation; Dember & Richman, 1989), the spatial delay-discounting task does not need forced choice trials to ensure that rats try both options, nor does it require complex leverpress or nose-poke pretraining to get them to perform the behavior.

Expected phases of behavior on the spatial delay-discounting task

A decision-making agent should begin a session by sampling the unknown parameters on that particular day. Because the initial delays on our task are random and the delayed side within a session is randomly chosen, the agent needs to determine which side will be delayed and how long the initial delay is on each day in order to make an informed choice. This initial phase would involve a few laps to fill in these pieces of information (*investigation*). Then, an agent choosing on the basis of the relative discounted value of each side would temporarily bias its choices to one side or the other, the delayed side to increase the delay or the nondelayed side to decrease the delay, until the discounted values of the two sides matched. This *titration* phase should last until the difference in discounted value approaches zero. For the remainder of the session, the favorable trade-off is maintained (*exploitation*; see Fig. 1c).

In this article, we report adjustment of a delay to a consistent indifference point on the spatial delay-discounting task. Our data show that the indifference point is a function of the

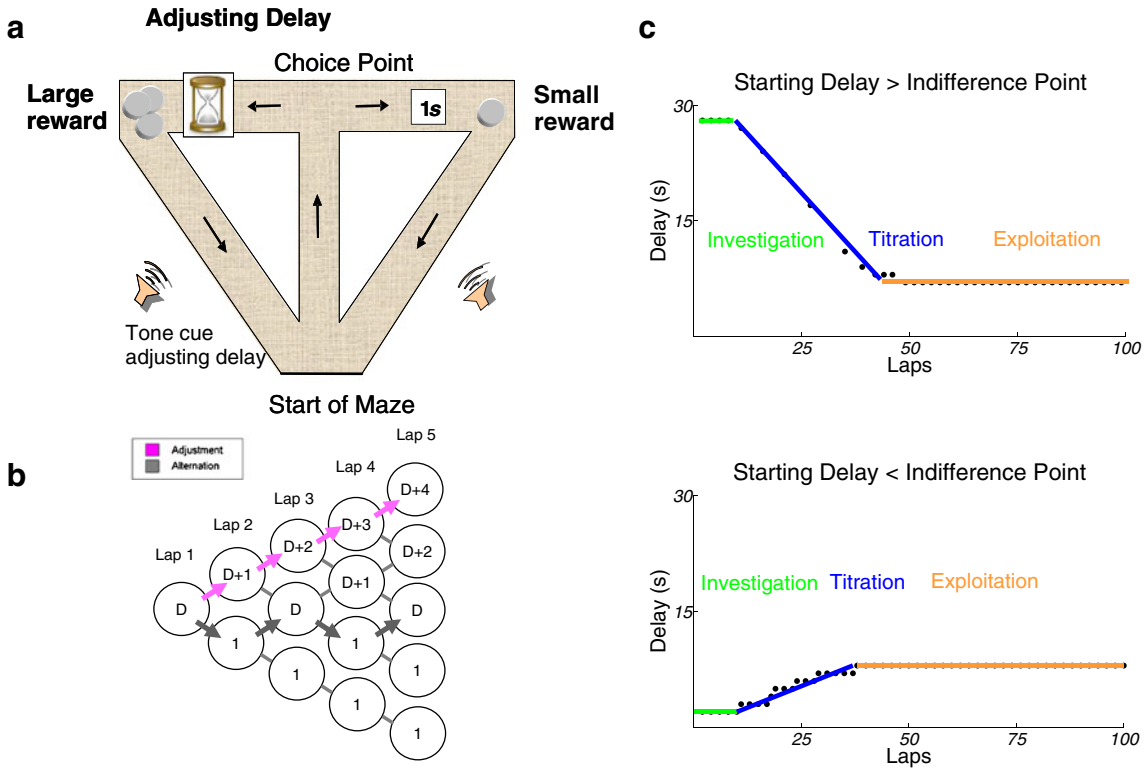

Fig. 1 Task diagram and predicted results a The spatial delay-discounting task. A lap (arrows) comprised a journey from the start of the maze to a choice point at the T-junction, to reward receipt at either the left feeder or the right feeder, and then back to the start of the maze. The delay before delivery of the larger of these two rewards varied on the basis of the choices of the rat. Sampling the larger reward caused the delay to increase; sampling the smaller reward caused it to decrease; thus, alternation maintained the same delay. An audible tone cue sequence accompanied the delay to food delivery. b Tree diagram illustrating changes in the adjusting delay as a function of lap. For a starting delay *D* on lap 1, the diagram illustrates two possible delay sequences for lap 2 through lap 5: (1) an upward titration composed of four adjustment laps to the large reward (magenta arrows) and (2) a sequence composed of four alternation laps from large-to-small reward and back (gray arrows). In the upward titration example, the large reward is sampled repeatedly, and the delay before reward delivery increases by 1 on each lap. In the alternation example, the large reward is sampled on odd laps, and the delay preceding the large reward is fixed at *D* seconds on each lap. c Sample sessions illustrating predicted behavior. First, information is acquired about the specific task parameters of the day (investigation, green), then the delay is adjusted to a preferred value (titration, blue), and lastly, this preference is maintained through alternation, maintaining the delay at the indifference point (exploitation, orange)

number of pellets of the large reward. We find that VTE occurs during adjustment laps, but not alternation laps. The variables of lap number, behavioral phase, and adjusting delay account for little variability in the occurrence of VTE. These results are interpreted in terms of theories of VTE and suggest that VTE occurs during flexible decision making.

Experimental procedures

Subjects

Fourteen adult male Fisher 344 Brown Norway rats (Harlan, Indianapolis, IN) were used in this experiment, 8–12 months of age at the start of behavioral training. Animals were food restricted to no less than 80 % of their free-feeding body weight, and water was available ad lib throughout the experiment. Animals were individually housed on a 12:12-h light:dark schedule. All procedures were conducted in full compliance with National Institute of Health guidelines for animal care and approved by the Institutional Animal Care and Use Committee at the University of Minnesota.

Task

The spatial delay-discounting task was run on a T-maze with return rails (Fig. 1a). Rewards were unflavored 45 mg pellets (Research Diets, New Brunswick, NJ) delivered by automated feeders (Med-Associates, St. Albans, VT) at the end of each T-arm. To receive a reward, rats traversed a navigational circuit from starting position to choice point to reward site, and then back to the starting position. Once rats entered the reward site, a tone sounded, beginning a countdown to

the reward. For each second of the countdown, a 100ms pure tone was played, indicating the time until reward delivery. Reward delivery always occurred simultaneously with a 1 kHz tone, and higher delays were accompanied by higher-frequency tones in steps of 175 Hz increase in frequency per 1 increase in delay. Thus, at least two tones (1.175 kHz at 1s to reward and 1 kHz at reward delivery) sounded on every lap, with longer time-to-reward accompanied by higher-frequency tones. Once the countdown began, if the rat left the reward site, the countdown stopped, and reward was not delivered. In practice, rats reliably waited out the delay, even on very long delays.

On each day, one reward site provided a "smaller" reward after 1 s (one 45 mg food pellet), while the other reward site provided a "larger" reward after a delay of D seconds. This reward was either three 45 mg food pellets (Experiment 2, N = 11 rats) or, one to five pellets, constant within a session but variable between sessions (Experiment 1, N = 4 rats). The notation R:1 is used throughout the article to indicate the larger-to-smaller reward ratio, with R being the magnitude of the larger reward in number of pellets. For both experiments, if the rat chose the larger option, the delay D increased by 1 s; if the rat chose the smaller option, the delay D decreased by 1 s. All reward sites had a minimum delay of 1 s.

Training

Sessions were stopped after 100 laps or a time-out period of 45 min (Experiment 1) or 60 min (Experiment 2). During initial training, the larger reward site was blocked off with wooden blocks to prevent the rat from going to the larger reward site. The blocked side alternated each day during training. During this initial training, the unblocked side provided one pellet after 1. After the rat reliably had run 100 laps (typically after 1–2 weeks), the rat then proceeded to the next stage of the experiment.

Experiment 1

Four rats ran 30 sessions over a pseudorandom distribution of five possible larger-to-smaller reward ratios (1:1, 2:1, 3:1, 4:1, 5:1) and six possible delays (1 s, 2 s, 5 s, 10 s, 20 s, 30 s). The delayed side was pseudorandomly counterbalanced from session to session.

Two weeks of training with a 3:1 reward ratio and with initial delays D selected pseudorandomly from the set of (1 s, 2 s, 5 s, 10 s, 20 s, 30 s) were given prior to data collection. This experiment was conducted in a separate room on an enclosed T-maze with 6-in.-high walls composed of multicolored Duplo bricks (LEGO Group). Rat position was tracked by use of an illuminated light emitting diode (LED) strapped to the back of the rat. One of these rats had previously completed 37 days of the variable-ratio protocol on the open T-maze (data not reported here). A second rat had previously completed the full 60-day experiment described below (data included in Experiment 2). Two of the rats were naive to the procedure at the start and received their initial training with blocked sides in the Duplo maze. No significant differences were seen between the 4 rats, so the 4 rats were pooled for analysis.

Experiment 2

Eleven rats received 30 days of training on an open T-maze elevated 6 in. above the floor with a 3:1 reward ratio (Fig. 1a). Initial delays were pseudorandomly selected for each session, without replacement, from 1 to 30 seconds. During the first 30 days, rat position was tracked by use of an LED strapped to the back of the rat, as in Experiment 1. After 30 days, rats were implanted with multielectrode hyperdrives targeting a variety of brain structures: 4× hippocampus, 4× dual-structure ventral striatum and orbitofrontal cortex, and 3× prefrontal cortex. Neurophysiological data from these rats are not reported here. After recovery from surgery, rats received 30 additional testing sessions with a 3:1 reward ratio and initial delays pseudorandomly uniformly distributed between 1 and 30 seconds. During these sessions, the position of the animal's head was tracked from LEDs on the headstage, so that head position and orientation were available, allowing the analysis of VTE behavior.

Data analysis

Tracking

Rats were tracked by an overhead camera sampling at 60 Hz. Pixels above a user-defined luminance threshold were digitized and time-stamped by a Cheetah data acquisition system (Neuralynx, Tucson, AZ). Position samples were deinterlaced by linear interpolation between even and odd sample frames to give a stable position measurement. Analysis was done by in-house programs written in MATLAB (Mathworks, Natick, MA).

In Experiment 1, 5/120 sessions were excluded from analysis because rats did not sample the adjusting delay. In Experiment 2, 14/348 backpack-tracked and 4/283 headstage-tracked sessions were excluded (all were from sessions 1–5) because rats ran fewer than 50 laps or did not sample the adjusting delay. In addition, 6/279 headstage-tracked sessions were excluded from the zIdPhi analysis (see below, in the VTE quantification with zIdPhi section of methods) because of technical errors with tracking. Of the remaining 273 sessions, 181/26,896 laps were omitted from analysis due to tracking errors on individual laps. Four additional laps where the adjusting delay sampled was greater than 30 s were excluded from analyses.

VTE quantification with zIdPhi

Headstage tracking from 273 sessions of Experiment 2 allowed measurement of VTE. Position samples starting halfway up the central stem of the T-maze and ending before entry into a reward site defined the *choice point* window. Note that the tone cue occurred only after the rat had exited the choice point window and entered a reward site. If the rat turned back into the choice point after triggering the countdown or after receiving reward, these samples were excluded from VTE analyses. VTE was quantified by the *z*-scored integrated absolute change in angular velocity of the head. Change in *x* and *y* position (*dx*, *dy*) through the pass was computed using an adaptive windowing of best-fit velocity vectors (Janabi-Sharifi, Hayward, & Chen, 2000). Orientation of motion (*Phi*) was then calculated from *dx* and *dy* using the arctangent. Orientation was unwrapped to prevent circular transitions. Change in orientation (*dPhi*) was then calculated using the same Janabi-Sharifi algorithm (Janabi-Sharifi et al., 2000). Absolute value of the change in orientation (|*dPhi*|) at each image sample (i.e., at 60 Hz) was integrated across the entire pass. This integrated score (*IdPhi*) was z-scored within session to produce a *zIdPhi* measure for each lap. zIdPhi reliably detected orient-reorient behaviors and was well-correlated with experimenter-scored VTE events and choice point pause time (Fig. 3).

Indifference point

The *indifference point* was quantified as the mean adjusted delay over the final 20 laps of a session.

Laps and phases

Adjustment laps were defined as consecutive laps in the same direction (i.e., LL or RR). The term *adjustment* was used because repeated laps to the same side cause changes in the adjusting delay. *Alternation* laps were defined as consecutive laps in opposite directions (i.e., LR or RL). Lap 1 was excluded from analysis.

We divided each session into behavioral phases by computing the percentage of adjustment-to-alternation laps within a sliding window of five laps. If two or more of the laps in the window were adjustment laps, it was classified as part of a titration phase. A window was classified as part of an investigation phase if it had zero or one adjustment laps and occurred before either the first titration phase or lap 30. Windows with one or no adjustment laps occurring after the first titration phase or after lap 30 were classified as part of an exploitation phase. Thus, the titration phase could include both adjustment and alternation laps but had a sufficiently high percentage of adjustment laps to change the adjusting delay.

Results

Experiment 1: Sensitivity to value

If the rats were truly comparing the discounted values of the two options, the adjusting delay at which the values of the two sides balance, the indifference point, should depend on the larger-to-smaller reward ratio. We tested this in Experiment 1 by varying the larger reward magnitude between sessions. The smaller reward site always provided one pellet, but the larger reward site delivered one, two, three, four, or five 45 mg pellets.

There was a significant effect of larger reward magnitude on the indifference point (Fig. 2) (Kruskal–Wallis; $p < 10^{-10}$; $\chi^2(114) = 55.03$). Hyperbolic discounting of delayed reward (Ainslie, 1975; Madden & Bickel, 2010; Mazur, 1987) predicts a linear relationship between the indifference point and the larger reward magnitude (Bradshaw & Szabadi, 1992; Ho, Woga, Bradshaw, & Szabadi, 1997). The relation between the variables in our data was well-described by a linear equation ($R^2 = .42$; $\beta = 2$; *t*-stat = 9; $p < 10^{-10}$), suggesting that the indifference point is a linear function of larger reward magnitude. Our data are, therefore, consistent with hyperbolic discounting of value.

Experiment 2: VTE

Different theories about VTE predict its occurrence during different phases of the spatial delay-discounting task. In Experiment 2, head position was obtained for 273 testing sessions over 11 rats. For these sessions, we were able to directly quantify VTE behavior with the zIdPhi measure, the *z*-scored integrated absolute change in angular displacement of the rat's head (see the Experimental Procedures section). Small zIdPhi indicated a "ballistic" trajectory, while large zIdPhi indicated a variability in orientation that characterizes VTE (Fig. 3).

Rats choose a consistent range of adjusting delays

In Experiment 2, when faced with a constant larger-to-smaller reward ratio of 3:1, rats consistently adjusted the delay toward a preferred range between 3 and 9 seconds. A histogram of the initial delay for all rats ($N = 11$) across all sessions ($N = 613$) shows a uniform distribution from 1 to 30 seconds. This was transformed into a consistent distribution of delays over the final 20 laps by the rats' adjustments (two-sample Kolmogorov–Smirnov test; $p < 10^{-10}$) (Fig. 4). The broad distribution of delays chosen during the first one third of a session reflected the uniform distribution of initial delays. During the middle one third of the session, the delays tended to converge onto a narrower range from 3 to 9 seconds and remained there for the remainder of the

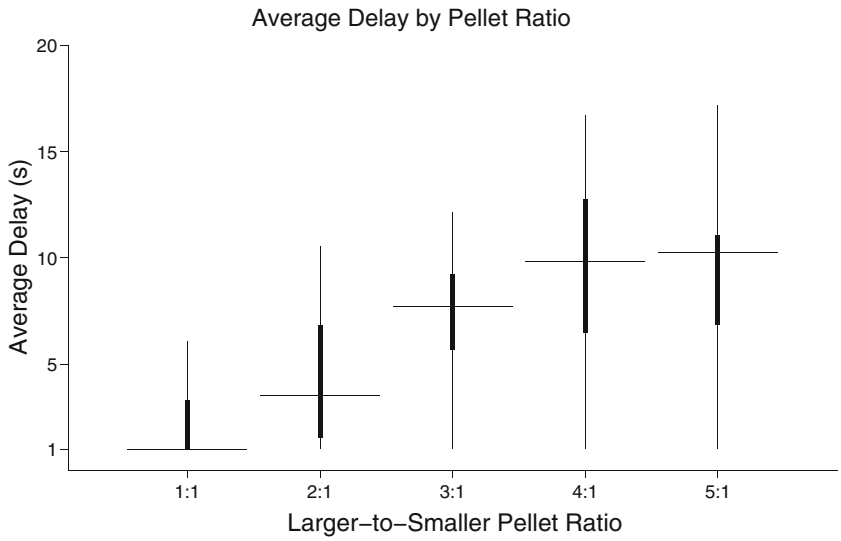

Fig. 2 Rats titrate to different delays as a function of different reward ratios. In Experiment 1, the indifference point was a function of the magnitude of the larger reward. Four rats ran 115 sessions with pseudorandomly selected larger rewards of *R* pellets (where *R* = 1, 2, 3, 4, 5, and *R*:1 indicates the larger-to-smaller reward ratio) and starting delays selected from a subset of 1 s to 30 s intervals. A boxplot of the indifference point versus the larger-to-smaller reward ratio is displayed where the horizontal line is the group median, the shaded box covers the 25th–75th percentile, and the whiskers are the 99th percentile of the data for each group. This relationship was well-described by a linear equation

Rats	p (slope)	F	beta	t-stat	df
R218	9.51E-03	7.89	1.24	2.81	26.00
R219	8.51E-06	30.50	2.51	5.52	27.00
R228	1.02E-05	28.78	2.14	5.37	29.00
R229	1.22E-05	28.10	1.86	5.30	29.00

session. These observations suggest that the rats titrated the adjusting delay to a preferred target on each session. We identify this target as the indifference point predicted by delay-discounting theories.

The indifference point did not vary significantly across session (Kruskal–Wallis; $p = .5$; $\chi^2(612) = 58.45$), suggesting that the delay target remained stable over the course of the 60-day experiment. The group average indifference point of 7.4 ± 3.0 s across all sessions for Experiment 2 was comparable to the average indifference point across the 3:1 sessions of Experiment 1.

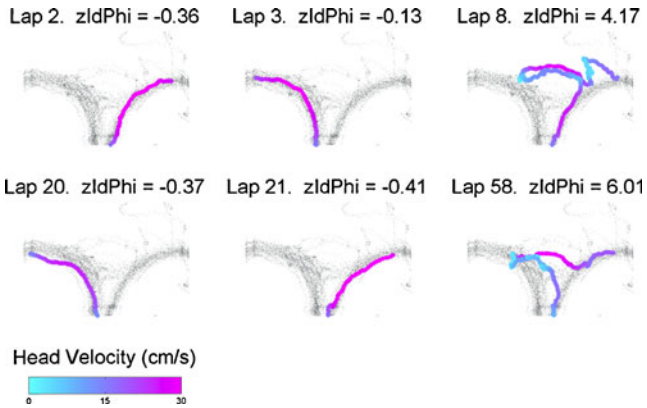

Fig. 3 Vicarious trial and error (VTE) quantification with zIdPhi. VTE was quantified at the choice point with a *z*-scored integrated angular velocity measure (zIdPhi). The position of the rat's head was measured from an overhead camera as it passed through a choice point on the T-maze. Six laps are displayed, with the rat entering the choice point at the bottom and walking upward and to either the left or the right arm of the T-maze. The small gray dots are position samples for all laps of this session, and the larger colored dots are the position samples for the indicated lap. The blue-to-magenta shading indicates head velocity from 0 cm/s (light blue) to 30 cm/s (magenta). On laps 2, 3, 20, and 21, the head of the rat passes through the choice point quickly with a smooth trajectory. These laps have low zIdPhi scores and would not be classified as VTE by manual observation. On laps 8 and 58, the rat pauses and reorients right-to-left and then left-to-right before passing through the choice point. These laps have high zIdPhi scores and are qualitatively recognizable as VTE events

Predictable phases of behavior were observed on each session

Plotting the percentage of adjustment laps by lap number reveals that the number of adjustment laps increased from laps 1 to 10, peaked on laps 10 to 30, and decreased through the remainder of the session, reaching an asymptotic low of 10 % around lap 75. Alternation laps dominated for the final two thirds of a session (Fig. 5a). The pattern of adjustment laps suggests the existence of three phases: one phase composed mostly of adjustment laps, usually occurring during laps 10–30, and two phases composed mostly of alternation laps, one before and one after the adjustment lap peak.

A five-lap sliding window was used to classify the three behavioral phases suggested by the distribution of adjustment laps. Labels were assigned to the phases on the basis of the predicted behavior of the idealized decision-making agent described in the introduction. There was a clear progression through investigation, titration, and exploitation

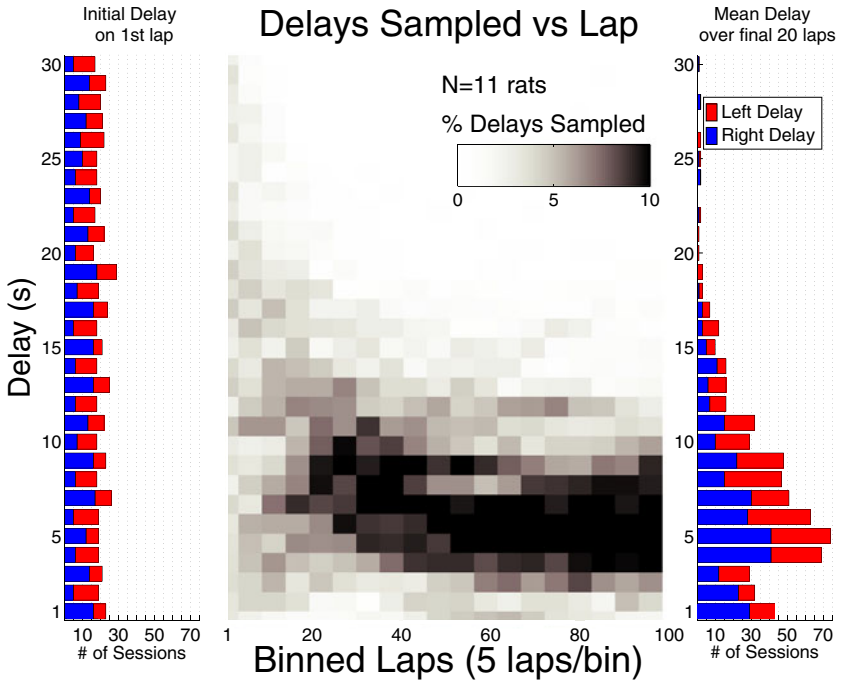

Fig. 4 Rats choose a consistent range of adjusting delays. For the 3:1 reward ratio in Experiment 2, rats ($N = 11$) titrated an adjusting delay to a consistent indifference point across all sessions ($N = 613$). A histogram of the initial delays (left panel) for right-side delay (red) and left-side delay (blue) shows a uniform distribution of initial delays. The probability of choosing a given delay on a given lap is displayed in white-to-black shading, with darker shading indicating a higher probability (center panel). The proportion of delays chosen in the 3 s to 9 s range increased with lap. A histogram of the indifference point for all sessions (right panel) shows that the initial delays were transformed into a consistent distribution in this lower third of the delay range

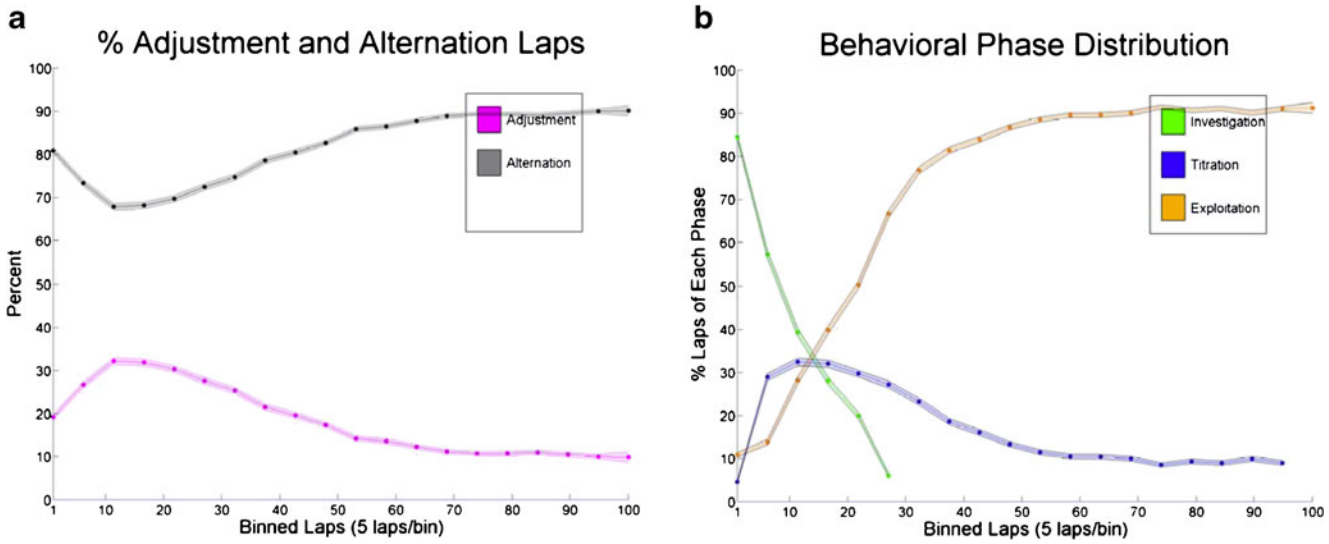

Fig. 5 Categorizing three phases of behavior. a The distribution of adjustment and alternation laps. Adjustment laps (magenta) were defined as consecutive laps to the same side, while alternation laps (gray) were defined as laps to opposite sides. The mean (shaded circles) and standard error (shading) are displayed. The distribution of adjustment laps peaks around laps 10–20, with alternation laps dominating both before and after this peak. This profile suggests three dissociable phases of behavior on the task. b Individual sessions were divided into three phases of behavior using a five-lap sliding window. The mean percentage of laps falling into the definition of each behavioral phase is plotted by lap number (shaded circles) with the standard error (shading). The profiles of the behavioral phases followed the distribution of adjustment and alternation laps. The phases were labeled investigation, titration, and exploitation according to the predictions made on the basis of the ideal decision-making agent

phases across lap. Over all 11 rats, 85 % of laps 1–5 were classified as investigation, but by lap 10 this proportion had dropped to 40 %. The peak of the titration phase occurred between laps 5 and 25. The percentage of laps in the titration phase decreased uniformly across lap, and by laps 70–75, 90 % of laps were classified as exploitation phase (Fig. 5b). Together, these observations suggest that each session began with alternation laps (investigation), followed by adjustment laps toward the indifference point (titration), and ended with alternation to maintain the delay at the indifference point once it was reached (exploitation).

VTE occurred on adjustment laps

We first examined the distribution of VTE events across lap for adjustment laps versus alternation laps. Average zIdPhi was higher for adjustment laps than for alternation laps (Kruskal–Wallis; $p < 10^{-10}$; $\chi^2(1) = 2{,}519$; $\eta^2 = .31$). The average zIdPhi was above the within-session average during adjustment laps, independent of the overall lap number. In contrast, average zIdPhi on alternation laps was near to or below the within-session average, uniformly so during the last two thirds of the session (Fig. 6a).

VTE occurred on early laps

As is shown in Fig. 5a, most adjustment laps occurred during the first one third of the session. Therefore, it is possible that the observed relationship between adjustment laps and VTE was a result of the increased likelihood of adjustment laps occurring early in the session.

zIdPhi was significantly impacted by lap (Kruskal–Wallis; $p < 10^{-10}$; $\chi^2(99) = 583.6$; $\eta^2 = .15$). Over the first few laps, zIdPhi was higher than the session average, and it decreased steadily to the session average by lap 30. After dipping below the session average, zIdPhi remained low for the remainder of the session (Fig. 6b).

In order to determine the relative contributions for the factors of lap number and adjustment versus alternation, a two-way ANOVA was performed. The main significant effect on zIdPhi was between adjustment and alternation laps ($p < 10^{-10}$; $F = 1{,}578$; $df = 1$; $\eta^2 = .24$). Although there was a significant effect of lap number ($p < 10^{-10}$; $F = 2.28$; $df = 98$; $\eta^2 = .095$), the effect of adjustment versus alternation explained much more of the total variance. This analysis suggests that the primary explanation for the increased zIdPhi on early laps was the occurrence of adjustment laps, rather than the lap number itself.

VTE occurred during high delays

As can be inferred from Fig. 5a, most adjustment laps occurred on delays that were above the indifference point, tending to drive the delay into the lower one third of the displayed range. Therefore, it is possible that the observed relationship between adjustment laps and VTE was a result

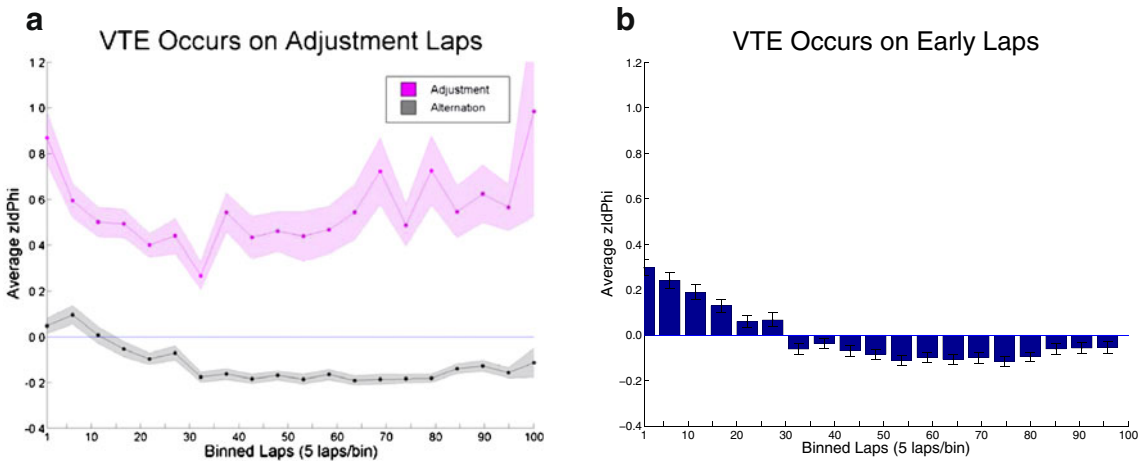

Fig. 6 Vicarious trial and error (VTE) occurred on early laps. VTE was quantified at the choice point from headstage-tracked sessions of Experiment 2 by integrating the change in angular position of the head and *z*-scoring within session (zIdPhi; see the Experimental Procedures section). Larger zIdPhi indicated high variability in head trajectory as a rat passed through the choice point, while smaller zIdPhi indicated that the trajectory was smooth and stereotyped. a zIdPhi for adjustment versus alternation lap types versus lap. Adjustment laps were pairs of laps that proceeded in the same direction (magenta), and alternation laps were pairs of laps to opposite directions (gray). zIdPhi was averaged across all sessions for all rats in 5-lap bins within each group. The mean (filled circles) and standard error (shading) are displayed. The average zIdPhi on adjustment laps is uniformly higher than on alternation laps: VTE is more likely to occur on adjustment laps. b zIdPhi versus lap. Average zIdPhi decreased with lap and was above within-session average for the first 30 laps and below session average for the final 60 laps. VTE tends to occur on early laps

of the increased likelihood of adjustment laps occurring during high delays.

Average zIdPhi increased with delay (Kruskal–Wallis; $p < 10^{-10}$; $\chi^2(29) = 266$; $\eta^2 = .09$), tending to remain below the within-session average for delays lower than the group indifference point but above it for delays greater than the group indifference point (Fig. 7a). Parsing out adjustment and alternation laps, zIdPhi remained high for adjustment laps regardless of delay and low for alternation laps on all delays. However, alternation laps during high delays did tend to show higher zIdPhi than did alternation laps at low delays (Fig. 7b).

In order to determine the relative contributions for the factors of adjustment versus alternation and delay, a two-way ANOVA was performed. The main significant effect on zIdPhi was between adjustment and alternation laps ($p < 10^{-10}$; $F = 1{,}501$; $df = 1$; $\eta^2 = .25$). Although there was a significant effect of delay ($p < 10^{-10}$; $F = 4.57$; $df = 29$; $\eta^2 = .069$), the effect size was small, and adjustment versus alternation explained much more of the total variance.

VTE occurred during the titration phase

During the titration and investigation phases, zIdPhi was higher, as compared with the exploitation phase (Kruskal–Wallis; $p < 10^{-10}$; $\chi^2(2) = 348.36$; $\eta^2 = .013$). zIdPhi was above the session average during the investigation phase and the titration phase. During exploitation, a decrease in zIdPhi was observed through the first one third of the session, with the remainder of the session having below average zIdPhi (Fig. 8a). These results indicate that VTE occurred during all behavioral phases in the first one third of a session but increased during titration phases, while diminishing during exploitation phases throughout the remaining two thirds of a session.

Because the adjustment laps occurred predominantly during the titration phase, the effect of adjustment versus alternation on VTE might be explained better as the occurrence of a titration phase. To test this, zIdPhi was averaged separately for adjustment and alternation laps within each phase. We found that zIdPhi on adjustment laps was higher than on alternation laps, regardless of phase (Fig. 8b).

In order to determine the relative contributions for the factors of adjustment versus alternation and behavioral phase, a two-way ANOVA was performed. The main significant effect on zIdPhi was between adjustment and alternation laps ($p < 10^{-10}$; $F = 1{,}626$; $df = 1$; $\eta^2 = .059$). Although there was a significant effect of behavioral phase ($p < 10^{-10}$; $F = 29.11$; $df = 2$; $\eta^2 = .002$), the effect of adjustment versus alternation explained much more of the total variance.

VTE was driven by behavioral flexibility

A high percentage of adjustment laps occurred during the titration phase; however, VTE was best explained as occurring on adjustment laps rather than during the titration phase. To investigate this discrepancy, we looked at the average zIdPhi with respect to the percentage of alternation laps in a 10-lap sliding window. Results were robust to

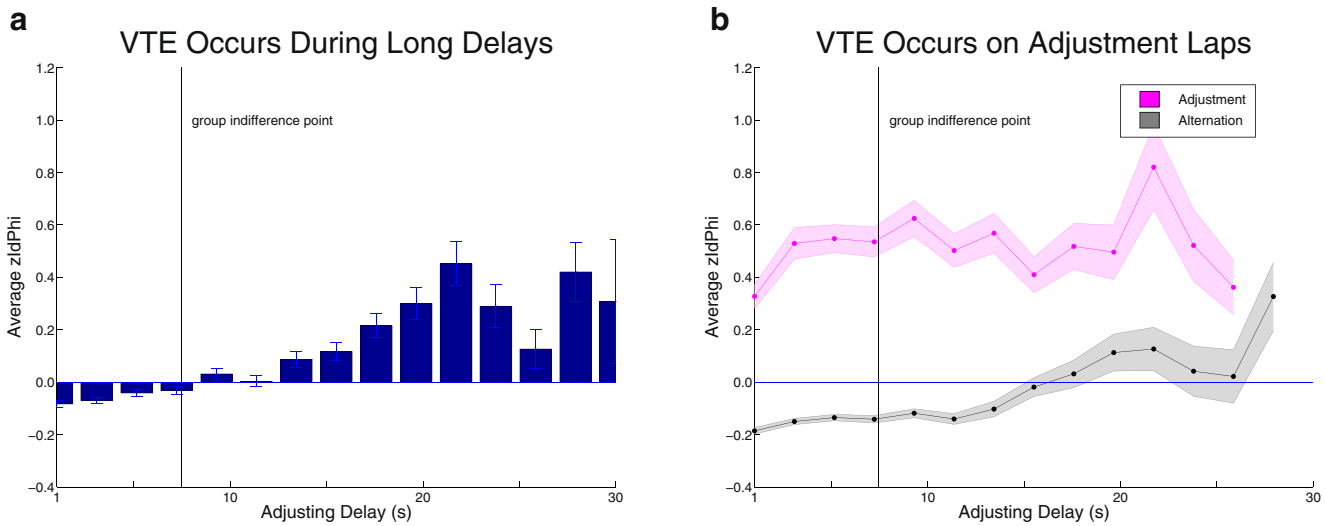

Fig. 7 Vicarious trial and error (VTE) occurred during high delays. a zIdPhi versus delay. zIdPhi was averaged over all laps for each adjusting delay from 1 s to 30 s, with delays greater than 30 s excluded from analysis. There was a small but significant increase in zIdPhi with delay. The shading indicates the standard error for each datapoint. b zIdPhi by lap category versus delay. zIdPhi was averaged within group for adjustment laps and alternation laps in five-lap bins. The mean (filled circles) and standard error (shading) are displayed. While VTE occurred on titration laps, it also occurred more frequently during alternation laps with high delays

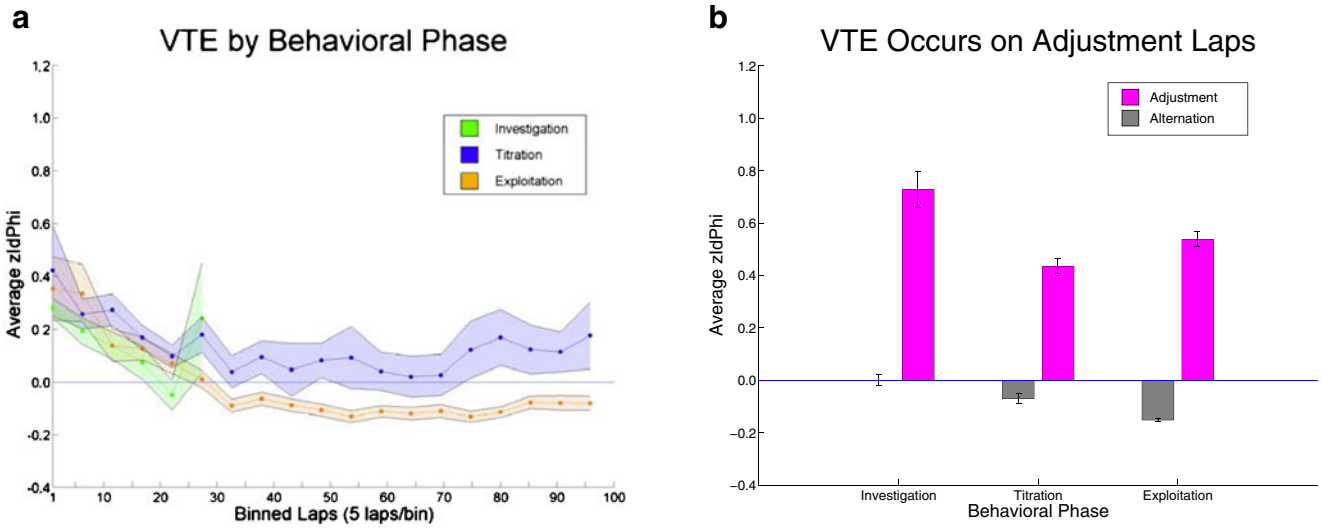

Fig. 8 Vicarious trial and error (VTE) occurred during the titration phase. a Behavioral phase classifications of investigation, titration, and exploitation were assigned to segments of each session on the basis of the distribution of alternation and adjustment laps in a five-lap sliding window. VTE tended to occur on early laps for all phases and during the titration phases for the remainder of the session. b zIdPhi was averaged within phase for adjustment and alternation lap types. Average zIdPhi was significantly above session average on adjustment laps, higher than zIdPhi on alternation laps regardless of phase categorization

different size windows (5–15 laps; data not shown). Average zIdPhi was low during alternation laps regardless of the percentage of alternation laps in the window. In contrast, zIdPhi on adjustment laps increased with the percentage of alternation laps in the window, indicating that the fewer adjustment laps there were in a set of 10 laps, the more likely it was that VTE would be observed when the animal did perform an adjustment lap (Fig. 9). VTE tended to occur most frequently on adjustment laps that occurred amid groups of alternation laps.

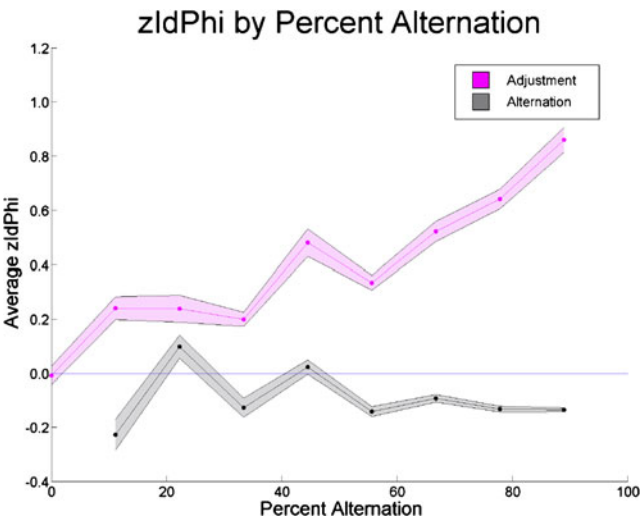

Fig. 9 Vicarious trial and error (VTE) was driven by behavioral flexibility. zIdPhi was uniformly low for alternation laps, no matter how frequently they occurred within a 10-lap sliding window. In contrast, zIdPhi was higher during adjustment laps, particularly when they occurred in the midst of alternation laps, indicating that VTE is more likely to occur during isolated adjustment laps, rather than during groups of consecutive adjustment laps

Discussion

Our results show that rats tracked the discounted economic value of reward on a spatial version of the adjusting delay task. Each day, rats adjusted an initially random delay to a consistent final delay; choosing either a larger-later option to increase it or a smaller-sooner option to decrease it. After titrating to their preferred waiting period, rats alternated for the remainder of a session. This behavior is compatible with titration to an indifference point where the subjective value between the two options is equivalent. VTE, a pause-and-look behavior in rats, occurred primarily during adjustment laps. VTE occurred most frequently on isolated adjustment laps but also occurred on adjustment laps that were grouped together into a titration phase.

Current theories of decision-making suggest that there are at least three dissociable action selection systems in the mammalian brain: a Pavlovian system that releases actions (unconditioned responses) based on associations between stimuli and outcomes, a deliberative system that considers future possibilities, and a habit system that learns to associate actions with stimuli (Daw et al., 2005; Montague, Dolan, Friston, & Dayan, 2012; Redish, Jensen, & Johnson, 2008; van der Meer et al., in press). We postulate that during the titration phase, the deliberative system may dominate as rats make online evaluations of action–outcome relationships to adjust to the changing conditions of the world, while during

the exploitation phase, the habit system dominates as rats alternate during a constant adjusting delay.

Different theories about VTE predict its occurrence at different times during the spatial delay-discounting task. The three theories are (1) investigation of alternatives through exploration, (2) mediation of sensory conflict through conditioned orienting, and (3) discrimination between fixed-value alternatives. Finally, we turn to the multiple decision-making system theory and discuss the relation of our VTE data to this theory.

Is VTE an exploratory behavior?

Tolman's (1948) original explanation for VTE was that it facilitated exploration of the structure of the task—for example, which of two stimuli in a visual discrimination task led to reward on a given day. Tolman found that VTE tracked performance, rising with the percentage of correct choices, falling off with asymptotic performance, and reemerging when the discrimination was made more difficult. While free choice tasks such as the spatial delay-discounting task do not have a percent correct measurement for assessing learning across sessions, within-session, rats must determine the location of the larger-later and smaller-sooner options and the unknown initial delay. If VTE was facilitating learning of these parameters, we would expect a particularly sharp decrease following the investigation phase of the session, a prediction that was not supported by our data. An exploration-associated behavior would also be expected on early laps, as compared with late laps. While we did find that VTE decreased with lap number, the effect of adjustment versus alternation was much larger in both cases. An interpretation of VTE in terms of exploration or learning would need also to account for its reemergence during isolated adjustment laps throughout the session.

Johnson et al. (2012) proposed that VTE mediates investigation of alternatives based on previously learned expectations about the environment. The theory contrasts undirected exploration of novel environments and directed investigation of familiar environments, noting that search can be carried out efficiently in familiar environments when unexpected outcomes violate expectations. Johnson et al. argued that VTE and its associated neural activity simulate potential future outcomes of decisions in directed investigation. This argument follows directly from the idea of the rat "searching for rules" to maximize reward in the Y-maze discrimination task (Hu, Xu, & Gonzalez-Lima, 2006). In terms of the spatial delay-discounting task, VTE occurring during adjustment laps could be interpreted as simulation of the possible temporal estimates of the delay (Buhusi & Meck, 2005), allowing covert evaluation of the option (Steiner & Redish, 2010), potentially through formation of a somatic marker for the upcoming reward (Damasio, 1996).

Is VTE discrimination between fixed-value alternatives?

Schrier and Povar (1979) hypothesized that SFS sequences in monkeys served the same fundamental operation as VTE orienting behavior in rats. Their results showed that SFS sequences in monkeys were present during early trials in a sensory discrimination task and decreased following asymptotic levels of performance. However, they found no relationship between SFS and learning rate in their study, leading them to the conclusion that the two processes were not causally related. They concluded that SFS sequences were most consistent with the search for more efficient visual scanning methods, consistent with the directed search theory of Johnson et al. (2012).

Krajbich et al. (2010) measured SFS sequences in humans during choice between two food items from a set that had been preference-ranked by subjects before testing. They found that SFS was higher between two similarly valued items, interpreting this as evidence for a neurally based value-integration-to-threshold mechanism (Gold & Shadlen, 2002; Ratcliff & McKoon, 2008). According to this interpretation, choice between similarly valued items requires longer integration times because the weighting of competing outcome representations is more or less equal. If VTE represented an underlying process of comparing fixed-value alternatives in a race-to-threshold manner, we would expect it to occur while the adjusting delay was close to the indifference point on the spatial delay-discounting task. However, VTE tended to occur less frequently during later laps and when the adjusting delay was close to the indifference point. This discrepancy may be explainable in terms of methodological differences between the tasks. In the Krajbich et al. study, there was no way for the subject to predict the value of the upcoming choice, preventing subjects from switching to a habitual action selection mechanism. In contrast, in the spatial delay-discounting task, once the adjusting delay becomes predictable and the exploitation phase begins, the rat can switch to a procedural or habitual action selection system without negative consequences.

Is VTE conditioned orienting?

In discrimination tasks, sensory cues are used to explicitly define correct from incorrect behavioral responses. VTE is a prominent behavior on these tasks, suggesting that it may elicit Pavlovian approach through contact with different sets of stimuli (Bower, 1959; Spence, 1960). While sensory cues are almost certainly used for navigation within the spatial delay-discounting maze, they are fixed with respect to the counterbalanced reward options and, therefore, do not provide consistent landmarks for action selection across sessions. Because the effectiveness of conditioned stimuli generally diminishes with a longer CS–US interval (Holland,

1980), if VTE were mediating Pavlovian sensory conflicts, we would expect different rates of VTE for different delays, particularly for shorter delays; however, this hypothesis was not supported by our data, because VTE increased with delay.

VTE and win-shift versus win-stay

The pattern of adjustment and alternation was not as simple as predicted by our theoretical analysis of the spatial delay-discounting task. While predictions from Fig. 1c provide a useful first-order account to categorize behavioral phase, titration was better characterized as occurring in multiple discrete fragments, rather than a single unified phase. This pattern suggests that rats were shifting strategies between alternation and adjustment. In the language of machine learning and game theory, adjustment laps are analogous to a "win-stay" response while alternation laps are a "win-shift" response. Titration to an indifference point may also be thought of as an iterative process of approaching a strict win-shift strategy during the exploitation phase. It can be inferred that rats prefer a win-shift strategy from the well-characterized phenomena of spontaneous spatial alternation (Dember & Fowler, 1958). In their review of spatial alternation in the rat, Dember and Fowler concluded that the probability of win-shift trials increased with the length of time at one location. VTE was strong during downward titration from high delays on the spatial delay-discounting task, requiring many win-stay trials, in conflict with the preferred win-shift strategy. This suggests that VTE coincides with times where the immediate preference to win-shift contradicts the long-term goal to reach an indifference point.

VTE and deliberative decision making

Gray and McNaughton (2000) hypothesize a behavioral inhibition system that coordinates suppression of motor output and increases in arousal and attention while information is accumulated toward the resolution of a conflict between approach and avoidance behaviors. Suggested to be part of the behavioral inhibition system, the hippocampus is thought to represent available goals on the basis of previous experiences, and this information can be evaluated online to help make a decision. We note the similarity of the behavioral inhibition theory with the search and evaluation processes described in neural ensembles (van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010).

Deliberation is hypothesized to be a two-step process: (1) predicting and, then, (2) evaluating a future outcome (van der Meer & Redish, 2010). Model-based theories of decision making, such as this deliberation theory, argue that neural representations of action–outcome relationships can be used to guide behavior (Daw et al., 2005; Johnson et al., 2007; Niv, Joel, & Dayan, 2006). While time-consuming and computationally expensive, deliberation is thought to be flexible to changing demands faced in real-world situations (Daw et al., 2005; Keramati, Dezfouli, & Piray, 2011; Niv, Daw, & Dayan, 2006; van der Meer, et al., in press). In contrast, model-free decision-making systems are fast but inflexible, subserving repetitive or habitual behaviors (Daw et al., 2005; Yin & Knowlton, 2006). The pattern of alternation seen during the exploitation phase is suggestive of habitual action selection, while the goal-directed titration to an indifference point suggests the operation of a deliberative system. During titration, responses shifted between win-stay and win-shift, suggesting online reevaluation of actions and outcomes.

VTE may be a behavioral correlate of a deliberative action selection system. The evolutionary advantage of vicarious simulation is that potential actions can be evaluated without an organism facing the potentially deadly consequences of learning action–outcome relationships by trial and error (Campbell, 1956). On the spatial delay-discounting task, where choices change future outcomes, vicarious estimation is necessary; a rat that must sample an actual outcome to evaluate it would be unable to titrate the delay effectively, and his adjustment laps would likely appear stochastically throughout the session, instead of being grouped together within a titration phase. In contrast, we found that VTE occurred during the titration phase and during adjustment laps that required inhibition of a preferred alternation response. VTE and adjustment co-occurred even in well-trained animals that had extensive knowledge of task parameters, suggesting that VTE occurs even in familiar environments. These observations suggest that VTE reflects the use of a deliberative action selection system, a pathway employed to resolve conflict between approach/avoid or win-stay/win-shift responses. In other words, that VTE really is *vicarious trial and error.*

Author note This work was supported by HFSP 2010-RGP/0039 and an equipment grant from the Minnesota Medical Foundation (MMF/2005). N.J.P. was supported by a graduate research fellowship from 5T32HD007151.

References

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82*, 463–496.

Bower, G. H. (1959). Choice-point behavior. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory* (pp. 109–124). Stanford, CA: Stanford University Press.

Bradshaw, C. M., & Szabadi, E. (1992). Choice between delayed reinforcers in a discrete-trials schedule: The effect of deprivation level. *Quarterly Journal of Experimental Psychology*, *44B*, 1–16.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49–57.

Buhusi, C., & Meck, W. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*, 755–765.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, *106*, 7745–7750.

Campbell, D. (1956). Perception as substitute trial and error. *Psychological Review*, *63*, 330–342.

Cardinal, R. N., Daw, N., Robbins, T., & Everitt, B. J. (2002). Local analysis of behaviour in the adjusting-delay task for choice of delayed reinforcement. *Neural Networks*, *15*, 617–634.

Damasio, A. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society B*, *351*, 1413–1420.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704–1711.

Dember, W.N., & Fowler, H. (1958). Spontaneous alternation behavior. *Psychological Bulletin*, *55*, 412–428.

Dember, W. N., & Richman, C. L. (Eds.). (1989). *Spontaneous alternation behavior*. New York: Springer.

Giordano, L. A., Bickel, W. K., Loewenstein, G., Jacobs, E. A., Marsch, L., & Badger, G. J. (2002). Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money. *Psychopharmacology*, *163*, 174–182.

Glimcher, P. W., Camerer, C., & Poldrack, R. A. (Eds.). (2008). *Neuroeconomics: Decision making and the brain*. San Diego, CA: Academic Press.

Gold, J.I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, *36*, 299–308.

Gray, J., & McNaughton, N. (2000). *The neuropsychology of anxiety.* Oxford: Oxford University Press.

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *PNAS*, *104*, 1726–1731.

Ho, M.-Y., Woga, M. A., Bradshaw, C., & Szabadi, E. (1997). Choice between delayed reinforcers: Interaction between delay and deprivation level. *Quarterly Journal of Experimental Psychology*, *50B*, 193–202.

Holland, P. (1980). CS–US interval as a determinant of the form of Pavlovian appetitive conditioned responses. *Journal of Experimental Psychology*, *6*, 155–174.

Hu, D., Xu, X., & Gonzalez-Lima, F. (2006). Vicarious trial-and-error behavior and hippocampal cytochrome oxidase activity during Y-maze discrimination learning in the rat. *International Journal of Neuroscience*, *116*, 265–280.

Janabi-Sharifi, F., Hayward, V., & Chen, C. S. J. (2000). Discrete-time adaptive windowing for velocity estimation. *IEEE Transactions on Control Systems Technology*, *8*, 1003–1009.

Johnson, A., & Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, *27*, 12176–12189.

Johnson, A., van der Meer, M. A. A., & Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology*, *17*, 692–697.

Johnson, A., Varberg, Z., Benhardus, J., Maahs, A., & Schrater, P. (2012). The hippocampus and exploration: Dynamically evolving behavior and neural representations. *Frontiers in Human Neuroscience* (in press).

Keramati, M., Dezfouli, A., & Piray, P. (2011). Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Computational Biology*, *7*, e1002055.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature*, *13*, 1292–1298.

Madden, G. J., & Bickel, W. K. (Eds.). (2010). *Impulsivity: The behavioral and neurological science of discounting*. Washington, DC: American Psychological Association.

Madden, G. J., & Johnson, P. S. (2010). A delay-discounting primer. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 11–37). Washington, DC: American Psychological Association.

Madden, G. J., Petry, N. M., Badger, G. J., & Bickford, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control patients: Drug and monetary rewards. *Experimental and Clinical Psychopharmacology*, *5*, 256–262.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: The effect of delay and of intervening events* (Vol. 5, pp. 55–73). Hillsdale, NJ: Erlbaum.

Mazur, J. E. (1997). Choice, delay, probability and conditioned reinforcement. *Animal Learning and Behavior*, *25*, 131–147.

Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, *108*, 96–112.

Mischel, W., & Underwood, B. (1974). Instrumental ideation in delay of gratification. *Child Development*, *45*, 1083–1088.

Mitchell, S. H. (2004). Measuring impulsivity and modeling its association with cigarette smoking. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 261–275.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*, 72–80.

Muenzinger, K. F. (1938). Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *Journal of Genetic Psychology*, *53*, 75–86.

Muenzinger, K. F., & Gentry, E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology*, *12*, 195–206.

Niv, Y., Daw, N. D., & Dayan, P. (2006). Choice values. *Nature Neuroscience*, *9*, 987–988.

Niv, Y., Joel, D., & Dayan, P. (2006). A normative perspective on motivation. *Trends in Cognitive Sciences*, *10*, 375–381.

Odum, A. L., Madden, G. J., & Bickel, W. K. (2002). Discounting of delayed health gains and losses by current, never- and ex-smokers of cigarettes. *Nicotine and Tobacco Research*, *4*, 295–303.

Paine, T. A., Dringenberg, H. C., & Olmstead, M. C. (2003). Effects of chronic cocaine on impulsivity: Relation to cortical serotonin mechanisms. *Behavioural Brain Research*, *147*, 135–147.

Perry, J. L., Larson, E. B., German, J. P., Madden, G.J., & Carroll, M. E. (2005). Impulsivity (delay discounting) as a predictor of acquisition of IV cocaine self-administration in female rats. *Psychopharmacology*, *178*, 193–201.

Perry, J. L., Nelson, S., & Carroll, M. E. (2008). Impulsive choice as a predictor of acquisition of IV cocaine self-administration and reinstatement of cocaine-seeking behavior in male and female rats. *Behavioral and Clinical Psychopharmacology*, *16*, 165–177.

Peters, J., & Büchel, C. (2010). Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal-mediotemporal interactions. *Neuron*, *66*, 138–148.

Petry, N. M., Bickel, W. K., & Arnett, M. (1998). Shortened time horizons and insensitivity to future consequences in heroin addicts. *Addiction*, *93*, 729–738.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.

Redish, A. D., Jensen, S., & Johnson, A. (2008), A unified framework for addiction: Vulnerabilities in the decision process. *Behavioral and Brain Sciences*, *31*, 415–487.

Schrier, A., & Povar, M. (1979). Eye movements of stumptailed monkeys during discrimination learning: VTE revisited. *Animal Learning and Behavior*, *7*, 239–245.

Simon, N. W., LaSarge, C. L., Montgomery, K. S., Williams, M. T., Mendez, I.A., Setlow, B., & Bizon, J. L. (2010). Good things come to those who wait: Attenuated discoutning of delayed rewards in aged Fischer-344 rats. *Neurobiology of Aging*, *31*, 853–862.

Spence, K. (1960). Conceptional models of spatial and non-spatial selective learning. In K. Spence (Ed.), *Behavior theory and learning: Selected papers* (pp. 366–392). Englewood Cliffs, NJ: Prentice-Hall.

Steiner, A., & Redish, A. D. (2010). Orbitofrontal cortical ensembles during deliberation and learning on a spatial decision-making task [Abstract]. *Society for Neuroscience Abstracts*.

Tolman, E. C. (1938). The determiners of behavior at a choice point. *Psychological Review*, *45*, 1–41.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*, 189–208.

Valencia Torres, L., da costa Araujo, S., Sanchez, C. O., Body, S., Bradshaw, C., & Szabadi, E. (2011). Transitional and steady-state choice behavior under and adjusting-delay schedule. *Journal of the Experimental Analysis of Behavior*, *95*, 57–74.

van der Meer, M. [A. A.], Kurth-Nelson, Z., & Redish, A. D. (in press). Information processing in decision-making systems. *The Neuroscientist*.

van der Meer, M. A. A., & Redish, A. D. (2009). Covert expectation-of-reward in rat ventral striatum at decision points. *Frontiers in Integrative Neuroscience*, *3*, 1–15.

van der Meer, M. A. A., & Redish, A. D. (2010). Expectancies in decision making, reinforcement learning, and ventral striatum. *Frontiers in Neuroscience*, *4*:6.

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, *67*, 25–32.

Yin, H., & Knowlton, B. (2006). The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, *7*, 464–476.