Computational heterogeneity in the human mesencephalic dopamine system

**Kimberlee D'Ardenne · Terry Lohrenz ·
Krystle A. Bartley · P. Read Montague**

Abstract Recent evidence in animals has indicated that the mesencephalic dopamine system is heterogeneous anatomically, molecularly, and functionally, and it has been suggested that the dopamine system comprises distinct functional systems. Identifying and characterizing these systems in humans will have widespread ramifications for understanding drug addiction and mental health disorders. Model-based studies in humans have suggested an analogous computational heterogeneity, in which dopaminergic targets in striatum encode both experience-based learning signals and counterfactual learning signals that are based on hypothetical information. We used brainstem-tailored fMRI to identify mesencephalic sources of experiential and counterfactual learning signals. Participants completed a decision-making task based on investing in markets. This sequential investment task generated experience-based learning signals, in the form of temporal difference (TD) reward prediction errors, and counterfactual learning signals, in the form of "fictive errors." Fictive errors are reinforcement learning signals based on hypothetical information about "what could have been." An additional learning signal was constructed to be relatable to a motivational salience signal. Blood oxygenation level dependent responses in regions of substantia nigra (SN) and ventral tegmental area (VTA), where dopamine neurons are located, coded for TD and fictive errors, and additionally were related to the motivational salience signal. These results are highly consistent with animal electrophysiology and provide direct evidence that human SN and VTA heterogeneously handle important reward-harvesting computations.

K. D'Ardenne · T. Lohrenz · P. R. Montague
Virginia Tech Carilion Research Institute, Roanoke, VA, USA

K. A. Bartley
Pediatric Emergency Medicine, Baylor College of Medicine,
Houston, TX, USA

P. R. Montague
Department of Physics, Virginia Tech, Blacksburg, VA, USA

P. R. Montague
The Wellcome Centre for Neuroimaging,
University College London,
12 Queen Square, London, UK WC1N 3BG

P. R. Montague (✉)
Human Neuroimaging Laboratory, Virginia Tech Carilion Research
Institute, 2 Riverside Circle, Roanoke, VA 24016, USA
e-mail: read@vtc.vt.edu

Current theories of dopamine function posit a single computational role for dopamine neurons in communicating errors in reward prediction to target brain regions (Bayer & Glimcher, 2005; Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997). Recent empirical work in animals has shown that dopamine neurons in the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc) are heterogeneous, with anatomical, molecular, and—importantly—functional differences (Brischoux, Chakraborty, Brierley, & Ungless, 2009; Joshua, Adler, Rosin, Vaadia, & Bergman, 2008; Lammel et al., 2008; Matsumoto & Hikosaka, 2009; for a review, see Roeper, 2013).

Functional differences have been examined in both rodents and nonhuman primates. In rodents, some dopamine neurons were shown to phasically respond to aversive events (Brischoux et al., 2009), which dopamine neurons were explicitly thought not to do (Ungless, Magill, & Bolam, 2004). In nonhuman primates, as well, dopamine neurons have been identified that respond phasically to aversive events but also to cues predicting those aversive events. These dopamine

neurons were mainly located dorsolaterally within the SNc, relative to dopamine neurons that fired according to reward prediction error theory (Matsumoto & Hikosaka, 2009). This distinction has been interpreted as a functional gradient spanning the VTA and SNc, as opposed to anatomically discrete groups of dopamine neurons with different functional properties. The dopamine neurons responding to aversive events have been posited to encode a motivational salience signal, which can be used for more efficient reward harvesting (Kakade & Dayan, 2002), instead of just coding for reward prediction errors (Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Matsumoto & Hikosaka, 2009).

On the basis of this emerging animal literature illuminating heterogeneity in the dopamine system, it has been hypothesized that multiple dopamine systems, with different firing properties and different efferent projections (Lammel et al., 2008), exist in the brain (Bromberg-Martin et al., 2010). Identifying and characterizing putative functional subgroups of the human brainstem dopamine system will have widespread ramifications for understanding addiction and other brain disorders.

We investigated the complexity of the human mesencephalic dopamine system using brainstem-tailored functional magnetic resonance imaging (fMRI; D'Ardenne, McClure, Nystrom, & Cohen, 2008). Because the BOLD response measured in fMRI reflects composite neuronal activity, we focused on functional heterogeneity. We aimed to answer the question of whether distinct computational learning signals—some (like reward prediction errors) known to be computed by the dopamine neurons, and other signals hypothesized to be encoded by dopamine—had discriminable sources within the SN and VTA that were consistent with a functional gradient seen in animals.

To do this, we examined two classes of computational learning signals: experience-based signals and counterfactual signals based on hypothetical outcomes. The experiential learning signals that we examined were signed and unsigned temporal difference (TD) reward prediction errors. Signed TD errors are calculated as the ongoing difference between rewards (in this experiment, money) received and expected. The unsigned TD error—for simplicity, operationalized by taking the absolute value of the signed TD error—was used as a proxy signal that is relatable to a motivational salience signal (Bromberg-Martin et al., 2010; Matsumoto & Hikosaka, 2009). We sought to determine whether noninvasive measures of neuronal activity related to signed and unsigned TD errors would show patterns of activity in humans that were similar to those reported in the animal literature: namely, that unsigned TD error sources would primarily be localized dorsolaterally to signed TD errors.

The second type of computational learning signal that we studied was a counterfactual learning signal. Although it is well-established that dopamine neurons compute signed TD errors (Bayer & Glimcher, 2005; Montague et al., 1996;

Schultz et al., 1997; Tobler, Fiorillo, & Schultz, 2005), the role of the midbrain dopamine system in computing counterfactual computational signals is less clear. Work in animals and humans has implicated the anterior cingulate cortex (Hayden, Pearson, & Platt, 2009), the prefrontal cortex (Abe & Lee, 2011), and the striatum (Chiu, Lohrenz, & Montague, 2008; Lohrenz, McCabe, Camerer, & Montague, 2007) in the computation of counterfactual-information learning signals. Because the dopamine system is known to innervate these brain regions, and because counterfactual learning signals are easily integrated into TD reinforcement learning algorithms (Lohrenz et al., 2007), we hypothesized that the mesencephalic dopamine system was a candidate source for the computation of counterfactual learning signals.

We focused on a specific counterfactual learning signal, the so-called "fictive error," that is based on hypothetical outcomes that could have happened (Lohrenz et al., 2007; Montague, King-Casas, & Cohen, 2006). Fictive errors are the ongoing difference between rewards that could have been gained and actual rewards. The present study extends previous work on fictive errors in humans (Chiu et al., 2008; Lohrenz et al., 2007) to a detailed examination of these signals in the mesencephalon.[1]

Method

Participants

The Institutional Review Panel for Baylor College of Medicine (BCM) approved this experiment for human participation. Informed written consent was obtained from all participants, who were recruited from within the BCM community, as well as from the surrounding Houston, Texas, area. Of the 90 participants imaged, 23 were excluded due to excessive head motion (motion greater than 1.5 mm in any direction). The remaining 67 participants (35 males, 32 females) varied in age from 19 to 53 years, and all but three were right-handed.

Sequential investment task

To examine experience-based and counterfactual learning signals, we had participants complete the "sequential decision-making task," which on each trial required them to place money into a market (Fig. 1a; Chiu et al., 2008; Lohrenz et al., 2007).

---

[1] As has been done previously, we restricted our definition of fictive errors to the difference between the obtained and maximum outcomes in a sequential decision-making task (Figs. 1a and b), instead of focusing on relationships to models of regret-based decision making (Bell, 1982; Loomes & Sugden, 1982). Additionally, these kinds of signals are often termed "off-policy," since they utilize information generated by behavioral options not taken—that is, off the behavioral policy (Sutton & Barto, 1998).

The markets were based on real historical markets, such as stock markets. The task consisted of 100 total trials divided among five markets, similar to previous versions of the sequential investment task (cf. Lohrenz et al., 2007). Thus, 20 trials were presented per market. The five markets used in our task were:

1. the Dow Jones Industrial Average (7/26/1927–11/12/1929)
2. the Nikkei 225 Index (01/05/2000–04/24/2002)
3. the Deutsche Mark/US Dollar Exchange Rate (8/12/1983–11/29/1985)
4. the NASDAQ Composite Index (09/11/1998–12/29/2000), and
5. the Hang Seng Index (09/11/1992–12/30/1994).

During the sequential investment task, participants viewed a screen that contained a trace of the market activity at the top and information about the participant's earnings at the bottom (Fig. 1a). A slider bar was used to indicate the bet on each trial. At the beginning of each trial, the slider bar turned red; the participant then used the slider bar to indicate what percentage of the current endowment should be invested in the market. Participants had an unlimited amount of time to decide how much to invest. The average bet across participants on a given trial was 48.95 ± 24.24 (SD), and the bets ranged from 0 % to 100 %. After the investment decision was made, the slider bar turned gray, and a variable delay of 4–10 s (in 2-s increments) was imposed before the outcome was displayed. After the bet was placed, the market value fluctuated up or down, and participants consequently won or lost money. Each trial outcome provided experiential information (how much money was won/lost, Fig. 1b black pathway) and counterfactual information (i.e., how much money could have been won/lost; gray pathways in Fig. 1b). At the time of the outcome, the change in the market value appeared, the percentage change in the endowment was also displayed in red, and the total amount of money that the participant had earned was updated. A variable intertrial interval (4 – 10 s, in 2-s increments) occurred before the start of subsequent trials.

TD errors are experience-based learning signals and were calculated as the difference between market value (r_t) and the expected earnings (taken as the bet, b_t): $TD_t = r_t - b_t$. Market value, r_t, was a positive number, whereas the bet, b_t, was a percentage. Similarly, fictive errors are counterfactual learning signals and were calculated as the difference between the maximum reward that could have been earned (the optimal bet, b_{opt}, multiplied by the market value, r_t) and the actual earnings (bet b_t multiplied by the market value r_t): $f_t = b_{opt} \cdot r_t - b_t r_t$. To enable comparisons with previous work (Chiu et al., 2008; Lohrenz et al., 2007), we focused on "fictive
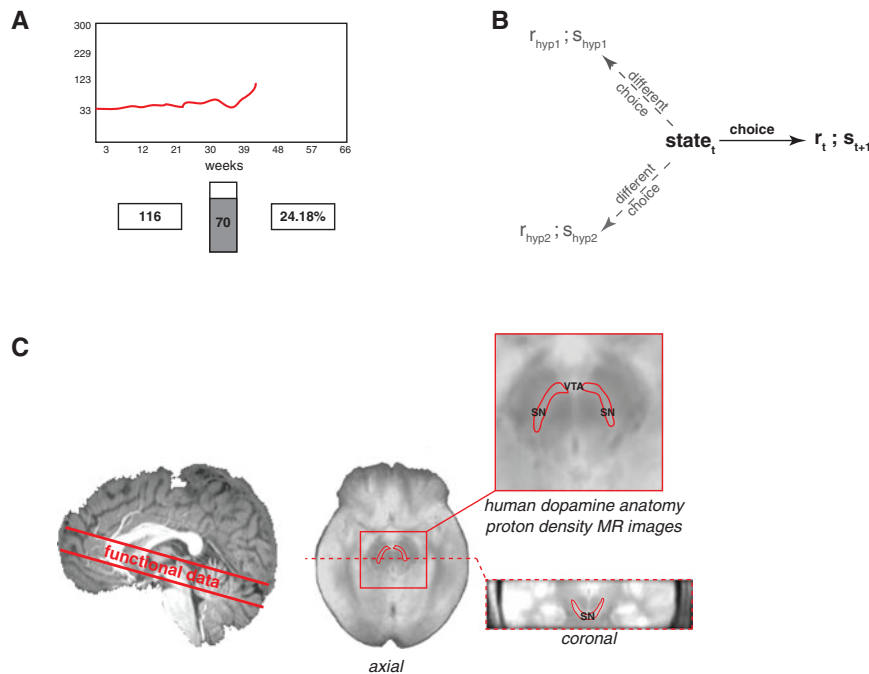
Fig. 1 a Participants began the task with $100 (running total, left box). On each trial, they had unlimited time to place a bet (slider bar) into a market (graph). After a variable delay of 4–10 s, the change in market value (trace), the amount of money gained or lost (right box), and the overall earnings (left box) were updated. Trials were separated by an intertrial interval of 4–10 s. b Information from the trial outcomes enabled computation of experiential and counterfactual learning signals. A learner's actual experience is indicated in black. The actual choice transitions the learner from a given state, s_t, to a new state, s_{t+1}, while the participant receives a reward, r_t. Hypothetical experience is indicated in gray. Information can be gained from this hypothetical experience, from what would have happened if different choices had been made. Experiential learning signals (black) correspond to TD errors, which reflect the difference between rewards received and expected. Counterfactual learning signals (gray) correspond to fictive errors, operationalized as the difference between the best possible outcome and the actual outcome. c Functional data were restricted to a slab centered on the mesencephalon and tilted to cover the substantia nigra and ventral tegmental area

errors over gains" (f_t^+), defined thus: $f_t^+ = 1 \cdot r_t - b_t r_t$, where, if the change in market value is positive, the best investment is a 100 % bet. When the market value goes down, fictive errors are mathematically equal to the money lost, since in that case, the optimal bet is 0 %.

Because TD errors are experience-based learning signals, whereas fictive errors are constructed from hypothetical outcomes, these two signals could be encoded in different brain regions and could influence participants' decisions in distinct ways. Indeed, previous work has shown robustly separable striatal regions encoding TD and fictive errors (Chiu et al., 2008; Lohrenz et al., 2007), as well as dissociable behavioral effects for the different errors in nicotine addicts (Chiu et al., 2008).

Behavioral data analysis

The behavioral data were analyzed using MATLAB (The MathWorks, Natick, MA). We investigated the behavioral impacts of signed TD errors, unsigned TD errors, and fictive errors over gains on task behavior by relating each signal to the change in bet from the current to the next trial. For signed TD and fictive errors, we performed an additional behavioral analysis, on the basis of previous work using this task (Chiu et al., 2008; Lohrenz et al., 2007). We regressed the change in bet from the current to the next trial against the current bet and each signal. Regression coefficients for both TD and fictive errors were correlated with the BOLD response from mesencephalic regions encoding TD and fictive errors, respectively (Figs. 2a and b in the Results).

MR image acquisition

All images were acquired using a 3-T Siemens Trio MRI system in the Human Neuroimaging Laboratory at Baylor College of Medicine. The visual stimuli were displayed on a rear projection screen and viewed by the participants through a mirror attached to the head coil. High-density foam padding was used to stabilize each participant's head to minimize head motion during the experiment.

High-resolution (0.25-mm^3 voxels) T1-weighted structural images were acquired with an MP-RAGE pulse sequence at the beginning of the scanning session.

All functional data were acquired using methods tailored for imaging from the human brainstem (D'Ardenne et al., 2008). A high-resolution echoplanar-imaging pulse sequence (128 × 128 matrix, 1.5 × 1.5 mm^2 in-plane voxels, 1.9-mm-thick slices, TE 41 ms) that was cardiac gated was used to collect functional images. Cardiac gating has been shown to reduce the physiological noise associated with fMRI of brainstem regions (Baria, Baliki, Parrish, & Apkarian, 2011; Guimaraes et al., 1998). A finger pulse-oximeter that interfaces with the scanner was used to monitor the participant's

pulse and to trigger the scanner during functional imaging. The pulse-oximeter was placed on the middle finger of the nondominant hand of each participant. The number of slices was determined by the participant's heart rate and remained constant throughout the entire experiment. To determine the number of slices, the experimenter observed the participant's heart rate during acquisition of the T1-weighted image and then selected a conservative number of slices. Additionally, the volume acquisition time (in non-cardiac-gated imaging, this corresponds to the repetition time, TR) and the maximum length of the acquisition window were determined on the basis of the participant's heart rate. The acquisition window corresponds to the amount of time during which the scanner will wait for a heartbeat to trigger the next image acquisition. The scanner was set to acquire an image every second heartbeat; a two-heartbeat interval was selected to balance the amount of data that we could acquire with the increase in signal-to-noise ratio that accompanies cardiac-gated brainstem fMRI data acquisition (Zhang et al., 2006). For a participant with a heart rate of 60 beats per minute (or 1,000 ms per beat), 18 slices would be used, with a volume acquisition time of 1,800 ms and an acquisition window of 1,900 ms. The volume acquisition time was always set to be as fast as possible, depending on the number of slices used. The flip angle (FA) was determined according to the Ernst angle:

$$\cos(\alpha_E) = e^{\left(-\frac{TR}{T1}\right)}.$$

The flip angle is α_E, TR is the repetition time (image acquisition time, in this experiment), and T1 is the T1 value for gray matter at 3 T. For 8 – 9 slices, FA = 60 deg; for 10 – 13 slices, FA = 70 deg; for 14 – 15 slices, FA = 75 deg; and for 16 – 20 slices, FA = 80 deg.

The midbrain was identified on the central sagittal slice of the high-resolution structural, and a slab comprising axial/coronal slices (each 1.9 mm thick) was centered on the midbrain and tilted to include as much of the SN, VTA, and striatum as possible (Fig. 1c). The mean number of slices was 14.1 ± 2.3 (SD), with a maximum of 19 slices and a minimum of eight slices across participants. The mean volume acquisition time was 1,128.2 ms ± 577.1, and the mean acquisition window was 1,457.5.5 ms ± 227.8. All scanner trigger times during functional imaging were recorded and used in data analysis.

After functional scanning, a non-cardiac-gated whole-brain functional image (TR/TE 2,500/42 ms, 25 slices, 6 mm thick, 30 % gap, FA 80 deg, four volumes) with the same center and orientation as the functional images was acquired solely to facilitate registration of the whole-brain structural image to the functional data (D'Ardenne et al., 2008; for a visualization of this method see Fig. 1 in Limbrick-Oldfield et al., 2012).
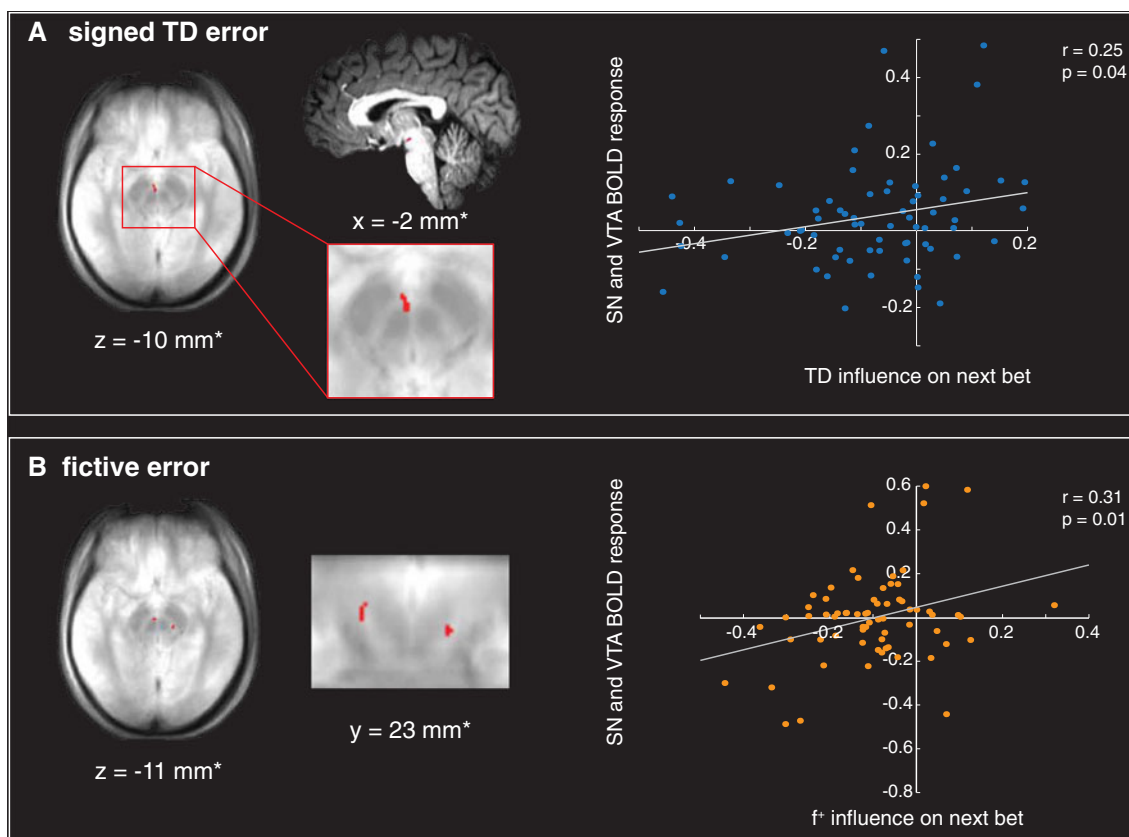
Fig. 2 a A random-effect GLM analysis revealed that a ventromedial region of substantia nigra (SN) and ventral tegmental area (VTA) (–2, 14, – 10)* encoded signed TD errors. Blood oxygenation level dependent responses in the indicated region were significantly correlated with the behavioral impact of TD errors (scatterplot). b The GLM analysis also showed that a ventromedial region of SN and VTA (-4, 16, –11)* and a dorsolateral region of SN (12, 21, –12)* encoded fictive errors over gains. BOLD responses in the indicated regions were significantly correlated with the behavioral impact of fictive errors (scatterplot). In both panels, highlighted regions indicate all voxels surviving thresholding ($n = 67$; p < .05, two-tailed t test, corrected for multiple comparisons). All MR images are shown according to radiological convention (i.e., left = right). Statistical maps were overlaid on a group-average proton-density weighted image (axial images in panels A and B, and expanded coronal image in panel B) or on a T1-weighted image (sagittal image in panel A) that have been brainstem-normalized (indicated by asterisks; Napadow et al., 2006). Scatterplots compare the results of regressions examining the influences of TD and fictive errors on future bets in terms of the SN and VTA BOLD responses to these signals

Finally, to enable localization of the SN and VTA in the midbrain, a proton-density weighted image (TR/TE 6,000/16 ms, FA 149 deg, echo spacing 15.6 ms, 0.75 × 0.75 mm² voxels in-plane) was acquired using the slice center and orientation information from the functional images (Oikawa, Sasaki, Tamakawa, Ehara, & Tohyama, 2002). On the proton-density image, the SN are visualized as a hyperintense (light) band between the cerebral peduncles and the circular red nucleus. Once the SN are located, the VTA can be readily identified, because it is medial to the SN in the rostral two-thirds of the midbrain (Naidich et al., 2009; Paxinos & Huang, 1995).

MR image analysis

First, functional data were corrected for the T1 variations inherent to cardiac-gated collection (Guimaraes et al., 1998)

using software written and implemented in MATLAB (The MathWorks, Natick, MA). The corrected data were then preprocessed and analyzed using AFNI (Cox, 1996). To account for the variable time between image acquisitions, regressors were calculated at higher temporal resolution and then resampled at image acquisition times. The functional images were first corrected for slice-timing offset and motion. The motion correction parameters were used to determine whether the participant moved the head more than 1.5 mm in any direction. Additionally, motion correction parameters were used as regressors of noninterest in multiple linear regression analyses. The data were then spatially smoothed with a 3-mm full-width-at-half-maximum Gaussian kernel and mean subtracted. For all participants, the most superior and inferior slices were excluded from the analysis, as a precaution against those slices shifting into previously unexcited regions.

The T1-weighted whole-brain structural image was aligned to the functional data and then transformed into Talairach and brainstem-normalized space (Napadow, Dhond, Kennedy, Hui, & Makris, 2006). The transform to Talairach space and brainstem-normalized space was then applied to the functional data. Coordinates from the group analysis are reported in brainstem-normalized space as Talairach coordinates with an asterisk.

General linear model (GLM) analysis

For each participant, design matrices were created in which each experimental event was considered an impulse stimulus that generated a hemodynamic response function of unknown amplitude. In addition to regressors for each experimental event, the design matrix contained regressors of noninterest that modeled baseline drift (scanning run mean, linear, and quadratic trends) and head motion. The experimental events modeled included the trial start, keypresses, indications of when a new market started, and the trial outcome.

The equations for TD errors and fictive errors are described above. As had been done previously (Chiu et al., 2008; Lohrenz et al., 2007), TD error regressors were computed using normalized values for the changes in market value and bet on the basis of what the participant had already experienced. Because the sequential investment task did not include any reward omissions, we computed a regressor relatable to a motivational salience signal by taking the absolute value of the signed TD error regressor.

The fictive-error-over-gains regressor and the loss regressor were orthogonalized with respect to the signed TD error regressor by subtracting the orthogonal projection of the fictive error onto the TD error from the fictive-error regressor. It is important to note that we also ran the analysis with the TD error regressor orthogonalized with respect to the fictive-error regressor and the loss regressor, and the results were unchanged.

The GLM analysis determined how mesencephalic dopamine regions encoded experiential signed TD errors, fictive error learning signals (Fig. 2 below), and unsigned TD errors (Fig. 3 below). Statistical maps of all events of interest were generated for each participant and then thresholded to identify brain regions where the regression coefficients (beta values) for modeled events were significantly different from zero (two-tailed *t* test).

The statistical significance of the results was determined using a small-volume correction that was constrained on the basis of the physical shape and size of the SN and VTA complex. We defined an anatomical mask comprising the SN and VTA using the hypointense regions on a group-average proton-density weighted anatomical image. The SN and VTA together are approximately 10–12 mm along the ventrodorsal axis and 4–5 mm at the widest location along the
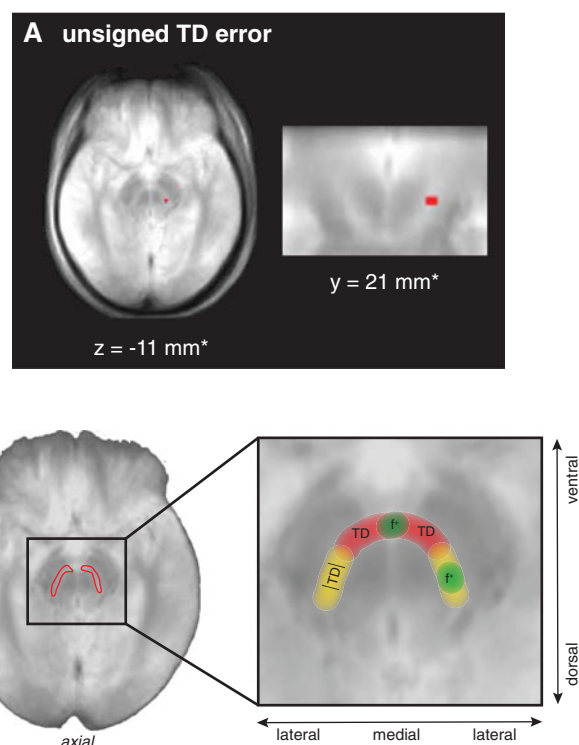
Fig. 3 a A random effects GLM analysis revealed that a dorsolateral region of human SN [10, 21, –11]* encoded the unsigned TD error. The regions indicated show all voxels surviving thresholding (*n* = 67; *p* < .05, two-tailed *t* test, corrected for multiple comparisons). MR images are shown according to radiological convention (i.e., left = right). Statistical maps are overlaid on a group-average proton-density weighted image, and asterisks denote coordinates in brainstem-normalized space (Napadow et al., 2006). **b** Schematic of a functional gradient in the human mesencephalic dopamine system. Ventromedial regions of SN and VTA are primarily coded for signed TD errors, and dorsolateral regions of SN and VTA are primarily coded for unsigned TD errors, which are relatable to salience signals (Bromberg-Martin et al., 2010). The regions within SN and VTA labeled f⁺ indicate areas found to code for fictive errors over gains

mediolateral axis (Naidich et al., 2009; Paxinos & Huang, 1995). We used the AFNI program 3dClustSim, which implements the cluster-size threshold procedure as a protection against Type I error (Forman et al., 1995), to define a corrected *p* value that fell within the physical size of the SN and VTA. We determined that a corrected *p* value of .05 was achieved with a minimum cluster size of six contiguous voxels, each significant at *p* < .01. All group results were calculated in brainstem-normalized space, and active regions within the brainstem were visualized on brainstem-normalized T1-weighted and proton-density anatomical images.

Correlation of mesencephalic and striatal BOLD activity

We performed a region-of-interest (ROI) analysis in the striatum to determine the relationship between striatal and mesencephalic BOLD activity to signed TD and fictive errors.

Striatal ROIs were based on regions previously identified as coding for signed TD errors (centroids [± 9, 14, − 2], on the basis of Lohrenz et al., 2007) and for fictive errors (centroids [± 8, 14, 4]). Because the number of slices included in the functional MR data set was restricted on the basis of participant heart rate, we did not image from the entire striatum. Striatum coverage spanned from $z = 3$ to − 15 mm (Talairach coordinates). All 67 participants had data in the region where TD errors were previously shown to be encoded, but just over half had coverage in the region where fictive errors were encoded ($n = 40$). For signed TD errors, we correlated regression coefficients, which are proportional to the magnitude of the BOLD response, from the regions of the SN and VTA shown to code for signed TD errors (Fig. 2a) with regression coefficients from the signed TD error striatal ROI. Likewise for fictive errors over gains, we correlated the regression coefficients from the SN and VTA regions shown to encode fictive errors (Fig. 2b) with regression coefficients from the fictive-error striatal ROI. Scatterplots are included in the supplemental materials, Fig. S1.

Results

To examine the complexity of the human brainstem dopamine system, we identified behavioral correlates and mesencephalic dopaminergic sources of three computational learning signals: signed TD errors, unsigned TD errors, and fictive errors. Participants completed the sequential investment task (Chiu et al., 2008; Lohrenz et al., 2007) that generated these signals while we measured BOLD responses from SN and VTA (D'Ardenne et al., 2008).

We first related signed TD errors, unsigned TD errors, and fictive errors over gains to task behavior. Signed TD errors ($p = 10^{-5}$, two-tailed t test) and fictive errors over gains ($p = 10^{-8}$, two-tailed t test) were found to have positive linear relationships, indicated by the positive slopes of the best-fitting lines, with the change in bets between the current and the next trial. Unsigned TD errors showed an inverse relationship to subsequent bets ($p = 10^{-14}$, two-tailed t test). We replicated previous behavioral regression analyses for signed TD errors and fictive errors over gains (cf. Chiu et al., 2008; Lohrenz et al., 2007; and Fig. 2).

The functional data were analyzed using a GLM analysis, and the resulting statistical parametric maps were thresholded ($p < .05$, corrected for multiple comparisons). We additionally performed a native-space within-subjects analysis to examine the discriminability of sources for signed TD errors, unsigned TD errors, and fictive errors, and also to test for directionality differences between the signed and unsigned TD errors.

Signed TD errors were localized to ventromedial regions of dopaminergic nuclei in the midbrain (Fig. 2a), whereas fictive errors were localized in both ventromedial and dorsolateral

regions of the SN and VTA (Fig. 2b). Regression coefficients indicating the behavioral influence of signed TD errors were positively correlated with the BOLD response in the SN and VTA regions that encoded TD errors ($r = .25$, $p = .04$; Fig. 2a). The fictive errors over gains were also positively correlated with the BOLD response in both mesencephalic dopaminergic regions encoding fictive errors ($r = .31$, $p = .01$; Fig. 2b).

The relationship between BOLD activity in mesencephalic dopaminergic regions and in the striatal regions previously identified as coding for signed TD errors and fictive errors over gains (Lohrenz et al., 2007) was also examined. We performed an ROI analysis on the regions of the striatum previously shown to code for signed TD errors and fictive errors and correlated the regression coefficients from mesencephalic and striatal regions for each signal separately (Fig. S1). Mesencephalic BOLD responses to signed TD errors were positively correlated with BOLD responses in the caudate (Fig. S1A; $r = .43$, $p = .0003$). In the subset of participants ($n = 40$; Fig. 1c) who had data in the region of the putamen previously shown to code for fictive errors, mesencephalic BOLD responses to fictive errors over gains were positively related to BOLD responses in the putamen, but this relationship was likely underpowered and was not statistically significant (Fig. S1B).

Because we hypothesized that BOLD responses measured from human SN would code for a signal relatable to the motivational salience signal recently identified in nonhuman primates (cf. Bromberg-Martin et al., 2010; Matsumoto & Hikosaka, 2009), we examined the neural correlates of the unsigned TD error. The sequential investment task does not have reward omissions, and our experiment constituted a special case in which a motivational salience signal could be computed by taking the absolute value of the signed TD error generated from this task. We found that BOLD responses in a dorsolateral region of SN were significantly related to the unsigned TD error signal (Fig. 3a). At the group level, sources for signed TD errors and motivational salience signals within human SN and VTA are consistent with electrophysiological results in nonhuman primates (Matsumoto & Hikosaka, 2009).

To examine the topography and discriminability of mesencephalic sources of the signed TD, unsigned TD, and fictive error signals, we carried out a native-space within-subjects analysis. We performed a sign test on the coordinates of the peak voxel for each computational signal in order to test whether they were overlapping within our test criterion of 4 mm (which corresponded to more than 2 voxels within-plane in our data set). For signed TD errors and fictive errors over gains, 58 of 67 participants (87 %) had nonoverlapping sources, whereas peak voxels for unsigned TD errors and fictive errors were nonoverlapping in 51 participants (76 %).

In comparing signed and unsigned TD errors, we found that 55 of our 67 participants (82 %) had nonoverlapping peak

voxels. We additionally tested for directionality differences between the sources for signed and unsigned TD errors along a ventromedial–dorsolateral gradient, with signed TD errors being more ventromedial and unsigned TD errors being more dorsolateral. Of the 55 participants with nonoverlapping sources for signed and unsigned TD errors, 45 of them (82 %) had signed TD sources that were more ventromedial than the unsigned TD-error sources. Overall, the percentage of participants showing a ventromedial/dorsolateral separation pattern for signed and unsigned TD errors was 67 %.

Discussion

The sequential decision-making task provides an ecologically valid framework for investigating computational-learning signals used in reward harvesting. We studied the behavioral correlates of TD reward prediction errors, a signal relatable to motivational salience, and of fictive errors over gains, while using fMRI methods tailored to the human brainstem (D'Ardenne et al., 2008) to identify their mesencephalic sources.

Our results identifying a region encoding fictive errors over gains (Fig. 2b) agree with the known anatomy of the dopamine system (Björklund & Dunnett, 2007) and with previous work examining fictive errors (Abe & Lee, 2011; Chiu et al., 2008; Hayden et al., 2009; Lohrenz et al., 2007), but they also suggest a role for the dopamine system in the computation of counterfactual information. In this task, fictive errors over gains quantify a specific kind of counterfactual information—namely, how much money the participant could have gained if the bet had been different. This learning signal has previously been integrated into TD reward prediction error learning algorithms and has been shown to have an impact on behavior in humans (Chiu et al., 2008; Lohrenz et al., 2007). Similar counterfactual learning signals have been shown to drive brain activity and behavior in animals (Abe & Lee, 2011; Hayden et al., 2009).

People with altered dopamine systems, because of brain diseases like drug addiction, mental health disorders, and degenerative pathologies like Parkinson's disease, are known to have deficits in decision making (Antonelli, Ray, & Strafella, 2011; Bach & Dolan, 2012; Bickel, Jarmolowicz, Mueller, Koffarnus, & Gatchalian, 2012; Hamilton & Potenza, 2012; Montague & Berns, 2002; Montague et al., 2006). In these populations, tracking fictive errors could provide both behavioral and neurobiological markers of the said deficits (see Chiu et al., 2008, for examination of the behavioral and neural correlates of fictive errors in nicotine addicts) and also could identify possible targets for therapeutic intervention.

The behavioral correlates of signed TD errors and fictive errors showed a positive linear relationship with BOLD responses in regions of the SN and VTA identified as encoding these signals (Fig. 2). It is intriguing to note that some participants' behavioral regression coefficients for signed TD errors

and/or fictive errors were negative. A possible explanation for these behavioral regression coefficients is erroneous perceptions of the outcome probabilities, or the gambler's fallacy (Tversky & Kahneman, 1971). Interestingly, the mesencephalic sources of fictive errors over gains are located in regions of the SN and VTA known to target prefrontal regions. Recent whole-brain fMRI studies examining the gambler's fallacy in similar decision-making tasks have shown that, relative to the striatum, prefrontal regions selectively code for responses relevant to the gambler's fallacy (Jessup & O'Doherty, 2011; Xue, Lu, Levin, & Bechara, 2011).

The negative relationship of the unsigned TD error signal to future bets also supports the gambler's fallacy as an interpretation of task behavior. We additionally explored the relationship between the unsigned TD error signal and the associability term, as defined by the Pearce–Hall theory (Pearce & Hall, 1980). The Pearce–Hall associability term assesses how surprising an event is and aids in learning relationships between cues and reinforcement. When reinforcement is fully predicted, learning from the associability term is slow, but when reinforcement is not fully predicted, learning is faster.

Because the Pearce–Hall associability term has been shown to be encoded by the human amygdala (Li, Schiller, Schoenbaum, Phelps, & Daw, 2011), we determined whether the regions within the SN and VTA that we identified as encoding unsigned TD errors were consistent with the known origins of dopaminergic projections to the amygdala. In rodents and nonhuman primates, A8 dopamine cells project to the amygdala; the A8 cells are located dorsal to the lemniscus (Dahlström & Fuxe, 1964). On an axial slice of a proton-density image, A8 cells would be located dorsal (down) and medial to the bright regions corresponding to the SN (Naidich et al., 2009; Paxinos & Huang, 1995). The mesencephalic region that we identified in the group analysis as encoding unsigned TD errors (Fig. 3a) is indeed near the putative location of A8 cells in humans, but because it does not extend outside the SN, we cannot attribute it to an A8 source.

We also examined the behavioral impact of unsigned TD errors for each participant, to see whether the effect agreed with the Pearce–Hall model. To do this, we plotted the value of the unsigned TD error on the current trial against the change in bet from the current to the next trial. We computed the slope of the best-fitting line for each participant and then determined whether all slopes were different from zero. If the unsigned TD error is relatable to the Pearce–Hall associability term, one would expect to see a positive relationship between the unsigned TD error and future changes in bets. We found the opposite pattern: The current value of the unsigned TD error was anticorrelated with upcoming bets ($p = 10^{-14}$).

For signed TD errors, it is also interesting to consider participants with both negative behavioral regression coefficients and negative BOLD responses (Fig. 2a). We found

that most of these participants bet more aggressively after negative TD errors generated from losses. This suggests that the BOLD response that we measured was not signaling only reward prediction errors per se, but could also be signaling how behavior should adapt on the basis of that signal.

Mesencephalic sources of signed TD errors, unsigned TD errors, and fictive errors over gains were separable at the group level, and within subjects, the majority of participants had separable sources for the two signals. When we tested for directionality differences between signed and unsigned TD errors, we found that the peak voxels coding for unsigned TD errors were primarily located dorsolaterally to those coding for signed TD errors. These results lend support to our hypothesis that the human brainstem dopamine system is computationally heterogeneous and organized as a functional gradient (Fig. 3b).

In subcortical structures such as the brainstem, BOLD responses are thought to predominantly reflect the summation of afferent synaptic inputs, as opposed to neuronal spiking (Logothetis, 2008). It is important to note that the BOLD response is a composite signal, reflecting many contributions from neuronal populations. Also, mesencephalic dopaminergic regions are composed of other neurons in addition to dopamine. Our measurements included contributions from these other neuronal populations and, additionally, contributions from substantia nigra *pars reticulata*. Previous work has shown that midbrain BOLD responses encode computations that dopamine neurons are known to carry out, suggesting a dominant contribution of the dopamine system (D'Ardenne et al., 2008). On the basis of synaptic input alone, we would be unable to distinguish between many of the potential computations subsumed by the SN and VTA. This issue is highlighted with the comparison of the motivational salience signal and TD error. Both of these signals derive from expected and actual reward values, and differ only in their input–output relations. Our human fMRI data agree well with animal electrophysiology studies that have differentiated dopaminergic regions on the basis of motivational salience versus TD error output, suggesting that at least part of our measured BOLD signal reflects regionally specific output, perhaps due to recurrent collateral synaptic activity.

Although functional differentiation in the dopamine system has been observed in animal electrophysiology (Brischoux et al., 2009; Matsumoto & Hikosaka, 2009), such topography has not been anticipated previously in humans. Our results highlight the utility of high-resolution, cardiac-gated fMRI methods when combined with precise hypotheses.

References

Abe, H., & Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron, 70,* 731–741. doi:10.1016/j.neuron.2011.03.026

Antonelli, F., Ray, N., & Strafella, A. P. (2011). Impulsivity and Parkinson's disease: More than just disinhibition. *Journal of Neurological Sciences, 310,* 202–207. doi:10.1016/j.jns.2011.06.006

Bach, D. R., & Dolan, R. J. (2012). Knowing how much you don't know: A neural organization of uncertainty estimates. *Nature Reviews Neuroscience, 13,* 572–586. doi:10.1038/nrn3289

Baria, A. T., Baliki, M. N., Parrish, T., & Apkarian, A. V. (2011). Anatomical and functional assemblies of brain BOLD oscillations. *Journal of Neuroscience, 31,* 7910–7919. doi:10.1523/JNEUROSCI.1296-11.2011

Bayer, H., & Glimcher, P. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47,* 129–141. doi:10.1016/j.neuron.2005.05.020

Bell, D. (1982). Regret in decision making under uncertainty. *Operations Research, 30,* 961–981.

Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Koffarnus, M. N., & Gatchalian, K. M. (2012). Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: Emerging evidence. *Pharmacology and Therapeutics, 134,* 287–297. doi:10.1016/j.pharmthera.2012.02.004

Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: An update. *Trends in Neurosciences, 30,* 194–202. doi:10.1016/j.tins.2007.03.006

Brischoux, F., Chakraborty, S., Brierley, D. I., & Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences, 106,* 4894–4899. doi:10.1073/pnas.0811507106

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron, 68,* 815–834. doi:10.1016/j.neuron.2010.11.022

Chiu, P. H., Lohrenz, T. M., & Montague, P. R. (2008). Smokers' brains compute, but ignore, a fictive error signal in a sequential investment task. *Nature Neuroscience, 11,* 514–520. doi:10.1038/nn2067

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29,* 162–173.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science, 319,* 1264–1267. doi:10.1126/science.1150605

Dahlström, A., & Fuxe, K. (1964). Evidence for the existence of monoamine-containing neurons in the central nervous system: I. Demonstration of monoamines in the cell bodies of brain stem neurons. *Acta Physiologica Scandinavica. Supplementum, 232,* 1–55.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647. doi:10.1002/mrm.1910330508

Guimaraes, A. R., Melcher, J. R., Talavage, T. M., Baker, J. R., Ledden, P., Rosen, B. R., … Weisskoff, R. M. (1998). Imaging subcortical auditory activity in humans. *Human Brain Mapping, 6,* 33–41. doi:10.1002/(SICI)1097-0193(1998)6:1<33::AID-HBM3>3.0.CO;2-M

Hamilton, K. R., & Potenza, M. N. (2012). Relations among delay discounting, addiction, and money mismanagement: Implications and future directions. *The American Journal of Drug and Alcohol Abuse, 38,* 30–42. doi:10.3109/00952990.2011.643978

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science, 324,* 948–950. doi:10.1126/science.1168488

Jessup, R. K., & O'Doherty, J. P. (2011). Human dorsal striatal activity during choice discriminates reinforcement learning behavior from the gambler's fallacy. *Journal of Neuroscience, 31,* 6296–6304. doi:10.1523/JNEUROSCI.6421-10.2011

Joshua, M., Adler, A., Rosin, B., Vaadia, E., & Bergman, H. (2008). Encoding of probabilistic rewarding and aversive events by pallidal and nigral neurons. *Journal of Neurophysiology, 101,* 758–772. doi:10.1152/jn.90764.2008

Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks, 15,* 549–559.

Lammel, S., Hetzel, A., Häckel, O., Jones, I., Liss, B., & Roeper, J. (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron, 57,* 760–773. doi:10.1016/j.neuron.2008.01.022

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience, 14,* 1250–1252. doi:10.1038/nn.2904

Limbrick-Oldfield, E. H., Brooks, J. C. W., Wise, R. J. S., Padormo, F., Hajnal, J. V., Beckmann, C. F., & Ungless, M. A. (2012). Identification and characterisation of midbrain nuclei using optimised functional magnetic resonance imaging. *NeuroImage, 59,* 1230–1238. doi:10.1016/j.neuroimage.2011.08.016

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453,* 869–878. doi:10.1038/nature06976

Lohrenz, T., McCabe, K., Camerer, C. F., & Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences, 104,* 9493–9498. doi:10.1073/pnas.0608842104

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal, 92,* 805–825.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature, 459,* 837–841. doi:10.1038/nature08028

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36,* 265–284.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience, 16,* 1936–1947.

Montague, P. R., King-Casas, B., & Cohen, J. D. (2006). Imaging valuation models in human choice. *Annual Review of Neuroscience, 29,* 417–448. doi:10.1146/annurev.neuro.29.051605.112903

Naidich, T. P., Duvernoy, H. P., Delman, B. N., Sorensen, A. G., Kollias, S. S., & Haacke, E. M. (2009). *Duvernoy's atlas of the human brain stem and cerebellum*. New York, NY: Springer.

Napadow, V., Dhond, R., Kennedy, D., Hui, K. K. S., & Makris, N. (2006). Automated brainstem co-registration (ABC) for MRI. *NeuroImage, 32,* 1113–1119. doi:10.1016/j.neuroimage.2006.05.050

Oikawa, H., Sasaki, M., Tamakawa, Y., Ehara, S., & Tohyama, K. (2002). The substantia nigra in Parkinson disease: Proton density-weighted spin-echo and fast short inversion time inversion-recovery MR findings. *American Journal of Neuroradiology, 23,* 1747–1756.

Paxinos, G., & Huang, X. (1995). *Atlas of the human brainstem*. San Diego, CA: Academic Press.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. doi:10.1037/0033-295X.87.6.532

Roeper, J. (2013). Dissecting the diversity of midbrain dopamine neurons. *Trends in Neurosciences, 36,* 336–342. doi:10.1016/j.tins.2013.03.003

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599.

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

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science, 307,* 1642–1645. doi:10.1126/science.1105370

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin, 76,* 105–110. doi:10.1037/h0031322

Ungless, M. A., Magill, P. J., & Bolam, J. P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science, 303,* 2040–2042. doi:10.1126/science.1093360

Xue, G., Lu, Z., Levin, I. P., & Bechara, A. (2011). An fMRI study of risk-taking following wins and losses: Implications for the gambler's fallacy. *Human Brain Mapping, 32,* 271–281. doi:10.1002/hbm.21015

Zhang, W. T., Mainero, C., Kumar, A., Wiggins, C. J., Benner, T., Purdon, P. L., & Sorensen, A. G. (2006). Strategies for improving the detection of fMRI activation in trigeminal pathways with cardiac gating. *NeuroImage, 31,* 1506–1512. doi:10.1016/j.neuroimage.2006.02.033