Supporting Online Material for

Human Substantia Nigra Neurons Encode Unexpected Financial Rewards

Kareem A. Zaghloul,* Justin A. Blanco, Christoph T. Weidemann, Kathryn McGill, Jurg L. Jaggi, Gordon H. Baltuch, Michael J. Kahana*

*To whom correspondence should be addressed. E-mail: zaghlouk@uphs.upenn.edu (K.A.Z.); kahana@psych.upenn.edu (M.J.K.)

Published 13 March 2009, *Science* **323**, 1496 (2009)
DOI: 10.1126/science.1167342

This PDF file includes:

Supporting Online Material

Materials and Methods

We captured intraoperative activity of SN in ten Parkinson's patients (6 men, 4 women, mean age 60.75 years). Seven subjects had completed college level education or higher. Three subjects did not attend college. Subjects participated voluntarily in the study after informed consent was obtained during pre-operative consultation for DBS surgery. Per routine presurgical protocol, Parkinson's medications were stopped on the night prior to surgery (12 hours preoperatively); hence, subjects engaged in the study while in an OFF state. In routine DBS procedures, microelectrodes are used to localize STN and advanced into SN to identify the inferior border of STN. We used this opportunity to capture microelectrode recordings from SN while subjects were awake and engaged in a probabilistic learning task (see (*1*) for recording details), in accordance with a University of Pennsylvania IRB approved protocol. Ethical constraints limited recordings to only subjects with clinical indications for DBS surgery.

Because DBS electrodes are often implanted bilaterally, we captured activity from 17 microelectrode SN recordings. We converted stereotactic coordinates for each recording site to Schaltenbrand-Wahren coordinates, referenced to the mid-commisural point (*2*). Mean electrode coordinates were $x = 12.21 \pm 0.28$ mm (mean $\pm$ SEM), $y = -1.28 \pm 0.42$ mm, and $z = -7.02 \pm 0.51$ mm for left electrode recordings, and $x = -11.93 \pm 0.23$ mm, $y = -1.45 \pm 0.29$ mm, and $z = -7.65 \pm 0.47$ mm for right electrode recordings. These coordinates correspond to left and right SN on the Schaltenbrand-Wahren brain atlas.

Intraoperative microelectrode recordings were performed with 1μm diameter tungsten tip electrodes advanced with a power-assisted microdrive (*3*). Microelectrode recordings were captured using a StimPilot® recording system and Spike2® data acquisition software (*1*). Signals are sampled at 24 kHz (16-bit A/D converter) for data analysis.

The probability learning task we use to examine reinforcement learning involves repeatedly drawing cards from two decks presented on a computer screen. Probabilities of financial reward are randomly assigned prior to the experiment – the good deck carries

a 65% chance of yielding a reward and the bad deck carries a 35% chance of yielding a reward. Subjects are instructed to try to win as much virtual money as they can in the five minutes allocated for the experiment. They have up to four seconds to make each choice, and feedback is presented for two seconds. Following feedback, the two decks are immediately presented on the screen for the subject to make the next choice. Limitations of the intraoperative recording environment on human subjects constrained the cognitive task to five minutes. Accumulated financial rewards were virtual in nature and were not paid out to participants.

We quantified learning rates by dividing the number of total trials (draws) in a session of the experiment into ten blocks and calculating how often the subjects chose the higher payoff deck within each block. Because every subject completed a different number of trials per experiment (depending on how quickly they made their selections) dividing the trials in this manner allowed us to compare learning rates across subjects.

For our model of expectation (see text, Equation 1), we set initial $E_d[1] = 0.5$ for each deck. We set α such that the weights of the power function approximate one over infinite trials for a given τ. This ensures an unbiased estimate of the effect of prior outcomes on expectation and limits expectation to the range between zero and one. We assumed subjects would make a choice on each trial based on this estimate of expectation. Thus, we fit τ for each subject to the sequence of choices and rewards observed in each experimental session. We performed a grid search optimization, varying τ and computing how often each subject chose the deck with the higher expected reward, as determined by Equation 1, on each trial. For each subject, we chose the τ that corresponded to the highest rate of optimal choices (i.e. choices that corresponded to the deck with the higher E_d).

To restrict our analysis to dopaminergic cells, we extracted and sorted single-unit activity using the WaveClus spike sorting package (*4*). Previous work on non-human primates relied on published examples of rat SN dopaminergic neurons to establish criteria for postulating the dopaminergic nature of the neurons studied (*5, 6*). We followed a similar approach and applied the following criteria to each recorded spike cluster. We considered only spike clusters with firing rates between 1.5 and 12 Hz, corresponding to findings for typical dopaminergic cells in animals (*6–8*). We then compared the average waveform of each such cluster with published examples of extracellular action potential morphologies for mammalian SN neurons recorded in a similar configuration to ours (*5–7, 9*). Typical extracellular recordings of dopaminergic SN neurons exhibit broader bi- and tri-phasic waveforms ($>$ 2 msec) than GABAergic neurons ($<$ 1.5 msec) (*5, 6, 8, 9*). To estimate the width of recorded waveforms, we calculated the average time from the beginning of an individual spike's waveform to its return to baseline ("baseline width") and the average time between the two positive peaks of the waveform ("peak-to-peak width"). We required that clusters have an average baseline width exceeding 2 msec and a peak-to-peak width

exceeding 0.8 msec. Finally, we assumed that dopaminergic cells would exhibit physiologic responses to feedback (*5, 6*), whereas GABAergic neurons would not. To determine which clusters exhibited feedback responses, we tabulated the total number of spikes in response to both positive and negative feedback for the first 500 msec after feedback onset; we then performed a Wilcoxon rank-sum test on the difference between total spike count in trials corresponding to positive and negative feedback during this period. We used a liberal threshold ($p < 0.2$) to retain for analysis only those clusters that demonstrated a difference in spike count between the two conditions in either direction.

To quantify spike activity, we generated peri-event spike histograms (75 msec bins; 1250 msec time window: 250 msec pre-feedback, 1000 msec post-feedback) for each trial, and averaged over all positive and all negative feedback trials to create a single mean histogram for each case. We calculated a continuous time firing rate by smoothing the spike train from a given trial with a Gaussian kernel with standard deviation of 25 msec, and normalizing by sampling rate. We compared continuous time firing rates and spike histograms to baseline spiking activity, defined as activity occurring in the interval 250 msec prior to card presentation on every trial. Spiking activity on every trial is z-scored by subtracting the mean and dividing by the standard deviation of baseline spiking activity in order to compare across cells and across subjects. We found significant effects only for the three continuous 75 msec intervals between 150 msec and 375 msec. We used non-overlapping 225 msec intervals (temporal epochs: -75 to 150; 150 to 375; 375 to 600; 600 to 825; 825 to 1050 msec relative to feedback onset) for all statistical analyses.

For statistical analysis of pooled data, we used a three-way ANOVA to examine differences in response to positive and negative feedback. We set feedback (positive vs negative) and expectation (unexpected vs expected) as fixed variables and cell as a random variable. To insure statistical significance of independent temporal events, we used a Bonferroni-corrected significance threshold of 0.009, taking into account six 225 msec epochs. For two-way ANOVAs, we set feedback as a fixed variable and cell as a random variable. In total there were 17.7 ± 1.5 (mean $\pm$ SEM) trials per experiment corresponding to unexpected gains, 20.6 ± 1.8 corresponding to unexpected losses, 30.1 ± 2.6 corresponding to expected gains, and 21.5 ± 1.9 corresponding to expected losses.

To examine the correlation between spike activity and prediction error, we a defined prediction error surrogate as the change in expected value for a given deck from one draw to the next. For each experimental session, we identified the median positive and negative differences in expected value and divided prediction errors associated with each trial into those positive and negative differences larger and smaller than the positive and negative median, respectively. This allowed us to normalize prediction error across subjects, who may have different peak and trough levels of expectation. We grouped spike activity into one of four categories — large positive, small positive, large negative, and small negative

prediction errors — and examined the mean z-scored spike rates associated with each category in the interval between 150 and 375 msec after feedback onset. We used a two-way ANOVA to determine whether activity associated with large positive prediction errors differed from activity associated with small positive prediction errors. We set prediction error (large positive vs small positive) as a fixed variable and cell as a random variable. We also computed a linear regression of differences in expected reward to mean z-scored spike rate during this interval for each subject. For each regression, a positive slope suggests that spike activity increases with prediction error. We compared the distribution of these slopes ($\beta = 0.95 \pm 0.53$, mean$\pm$SEM) to the null hypothesis using a two-tailed t-test ($p = 0.09$).

Supporting Online Text

In addition to z-scored continuous time firing rates and spike counts, we also compared raw, unnormalized spike counts in response to unexpected and expected positive and negative feedback using a three-way ANOVA for the interval between 150 and 375 msec after feedback onset. We found a significant difference between responses to positive and negative feedback [$F(1, 14) = 12.6$, $MSE = 20$, $p < 0.005$]. We found that this main effect of feedback was modulated by a significant interaction with expectation [$F(1, 14) = 12.1$, $MSE = 21.6$, $p < 0.005$]. Examining only unexpected trials during the same post-feedback interval, raw spike counts in response to unexpected gains were significantly greater than spike rates in response to unexpected losses [$F(1, 14) = 13.8$, $MSE = 30.5$, $p < 0.005$]. Expected trials, however, demonstrated no significant difference in raw spike counts during this interval between expected gains and losses [$F(1, 14) < 1$, n.s.]. We used z-scored firing rates and spike counts for our analyses in the main text to insure unbiased comparisons across different cells.

To assess whether recorded neurons encode the presentation of the stimulus cue, we measured changes in firing rate at the time of deck presentation. We found no significant changes in firing rate, compared to baseline, during any interval after the time of deck presentation [$F(1, 14) < 1$, n.s]. This did not change over the course of the experiment, and suggests that a visual stimulus alone is not enough to elicit responses in our SN neurons. We also compared firing rates at the time of deck presentation during trials when subjects chose the better deck to trials associated with the worse deck. We found no significant difference in continuous time firing rates in the interval between 150 and 375 msec after deck presentation between the two conditions [$F(1, 14) < 1$, n.s.]. The remaining intervals after deck presentation exhibited no significant differences between the two conditions as well. This suggests that the response to cue presentation could not predict the choice made for each trial in our study. Recent data conflict about whether firing rates at the time of cue presentation correlate with choice (*10, 11*). In these studies,

multiple cues are learned and presented. It is possible that after learning, the novel presentation of the rewarding cue acts as a surrogate for the reward itself. In our study, on the other hand, the same cues are presented on every trial. Since cue presentation rapidly loses its novelty in our paradigm, we would not expect a reward response until a choice is actually made and a reward presented.

The differences we observe in spike activity following unexpected gains versus losses could be driven by an increase in response to unexpected gains, a decrease in response to unexpected losses, or both. To probe this issue, we compared responses to baseline firing rates. Continuous time spike rates in response to unexpected gains were significantly greater than baseline firing rates during the 150 to 375 msec post-feedback interval [$F(1, 14) = 8.9$, $MSE = 53$, $p < 0.005$]. Spike rates in response to unexpected losses demonstrated a relative decrease in activity, but this response did not differ significantly from baseline firing rates during any interval [$F(1, 14) = 1.0$, $MSE = 2.7$, $p = 0.32$]. We hypothesize that low baseline firing rates prevented a significant decrease below baseline in response to unexpected losses. We found similar results when examining z-scored spike counts.

We found significant differences in neuronal responses in the 150 to 375 msec post-feedback interval. Direct measurements of non-human primate dopaminergic neurons in response to reward feedback demonstrate similar response onset time, albeit with longer depressions in activity in response to negative feedback (*12, 13*). This correspondence is further evidence that we are indeed studying the human analog of non-human primate reinforcement learning. That the response duration we measured is shorter than that observed in non-human primates, however, may be explained by differences in the saliency of the reward signals used by human and non-human primates, or by neuroanatomic differences between species that may lend these neurons different electrophysiologic profiles (*14, 15*).

In our study, although a relatively small number of cells (22%) exhibited responses to feedback, the responses within these cells were quite robust. Because of the degenerative nature of Parkinson's disease, the number of functioning dopaminergic neurons is depleted, thus compromising the ability to respond to behavioral feedback (*16–18*). This could explain the small number of behaviorally responsive cells we recorded. It seems plausible that healthy individuals who have larger populations of viable dopaminergic cells could mount more significant SN responses to unexpected feedback.

References and Notes

1. J. Moyer, *et al.*, *Operative Neurosurgery Supplement 1* **60**, E177 (2007).
2. G. Shaltenbrand, W. Wahren, *Atlas for Stereotaxy of the Human Brain* (Thieme Medical Publishers, Stuttgart, Germany, 1977).
3. J. Jaggi, *et al.*, *Stereotactic and Functional Neurosurgery* **82**, 104 (2004).
4. R. Quian Quiroga, Z. Nadasdy, Y. Ben-Shaul, *Neural Computation* **16**, 1661 (2004).
5. W. Schultz, A. Ruffieux, P. Aebischer, *Experimental Brain Research* **51**, 377 (1983).
6. W. Schultz, *Journal of Neurophysiology* **56**, 1439 (1986).
7. R. White, F.J. Wang, *Science* **221**, 1054 (1983).
8. A. Grace, B. Bunney, *Neuroscience* **10**, 301 (1983).
9. T. Korotkova, L. Haas, R. Brown, *Neuroscience Letters* **320**, 133 (2002).
10. G. Morris, A. Nevet, D. Arkadir, E. Vaadia, H. Bergman, *Nature Neuroscience* **9**, 1057 (2006).
11. M. Roesch, D. Calu, G. Schoenbaum, *Nature Neuroscience* **10**, 1615 (2007).
12. W. Schultz, *Neuron* **36**, 241 (2002).
13. J. Hollerman, W. Schultz, *Nature Neuroscience* **1**, 304 (1998).
14. R. Kotter, M. Feizelmeier, *Progress in Neurobiology* **54**, 619 (1998).
15. C. Hardman, *et al.*, *Journal of Comparative Neurology* **445**, 238 (2002).
16. R. Cools, R. Barker, B. J. Sahakian, T. W. Robbins, *Cerebral Cortex* **11**, 1136 (2001).
17. M. J. Frank, L. Seeberger, R. C. O'Reilly, *Science* **306**, 1940 (2004).
18. R. Cools, R. Barker, B. J. Sahakian, T. W. Robbins, *Neuropsychologia* **41**, 1431 (2003).