By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism Supporting Online Material

Michael J. Frank[1†], Lauren C. Seeberger[2], Randall C. O'Reilly[1†]

[1] Dept of Psychology and Center for Neuroscience, University of Colorado Boulder
[2] Colorado Neurological Institute Movement Disorders Center

[†]To whom correspondence should be addressed; E-mail: frankmj/oreilly@psych.colorado.edu

1 Methods

Procedures were approved by the HealthONE Institutional Review Board and the University of Colorado Human Research Committee. We tested 30 PD patients and 19 healthy seniors matched for age, education, and scores on the North American Adult Reading Test (NAART), an estimate of premorbid verbal IQ (*1*). The demographics of seniors and PD patients are shown in Table 1. PD patients were recruited from the Colorado Neurological Institute. All patients were receiving daily L-Dopa preparations, with some supplemented with D2 receptor agonists and/or selegiline (monoamine activity enhancer), and were stable on their medication dose for at least 2 months. Participating senior controls were either the spouses of PD patients or were recruited from the Boulder Senior Center. Exclusionary criteria were as follows:

- significant medical history not related directly to PD (e.g. stroke, head injury, clinical dementia, epilepsy);
- concurrent illness such as schizophrenia and manic depression;
- documented or suspected history of drug abuse and/or alcoholism;
- PD patients with advanced symptoms (stage IV or V in the Hoehn and Yahr rating scale);

- PD patients with Mini Mental State Examination (MMSE) ratings of less than 24 to screen for dementia;
- patients and control subjects taking additional medication likely to confound interpretation of the findings were excluded to the best of our ability.

Note that we did not specifically attempt to exclude patients based on depressive symptoms. Although none of our patients were diagnosed with depression, there is some evidence for a co-morbidity of Parkinson's and depression. Although this may seem like a potential confound, in fact we think it may be in part due to the same mechanism as captured in our model. That is, depression may be associated with a propensity toward NoGo responding resulting from decreased dopamine levels. Furthermore, the NoGo responding in our study was reversed with dopamine medication, so it is unlikely to be due to a global depressive state.

Participants were tested in two separate experimental sessions, separated by a minimum of 7 days. To minimize potential learning effects between sessions, we employed two different procedural learning tasks in the two sessions. The order of these tasks is randomized, so that each patient is equally likely to perform either of the tasks ON or OFF medication, and each healthy senior participant is equally likely to perform each task in the first or second session. Both tasks require trial and error learning, in which some stimuli have a net positive reinforcement value (and should be chosen), whereas others have a negative reinforcement value (and should be avoided). Both tasks use two-alternative forced choice, in which participants press one of two keys on a keyboard to "choose" one of two stimuli presented on a computer screen.

Participants sit in front of a computer screen in a lighted room and view pairs of visual stimuli that are not easily verbalized (Japanese Hiragana characters. These stimuli are presented in black on a white background, in 72 pt font. They press keys on the left or right side of the keyboard depending on which stimulus they choose to be "correct". Note that precise motor control is not necessary because any of 12 keys on the appropriate half of the keyboard counts as a response, allowing us to control for motor deficits associated with PD. Furthermore, the forced-choice nature of the task controls for any differences in overall motor responding. Visual feedback is provided following each choice (the word "Correct!" printed in blue or "Incorrect" printed in red). If no response is made within four seconds, the words "no response detected" are printed in red.

1.1 Task I: Probabilistic Selection

1.1.1 Rationale

In previous work, we explored the involvement of BG/DA interactions in probabilistic classification via explicit computational modeling, which (a) accounted for PD deficits and (b) made novel predictions (*2*). Specifically, the model predicted different deficits in medicated and non-medicated PD patients, depending on whether the learning was from positive or negative feedback. Due to depleted DA, those OFF medication should have deficits in learning from positive feedback. DA medication should alleviate these deficits, but may impair learning from negative feedback as the medication blocks the effects of DA dips that are required for this learning (*2*). [1] We therefore designed a novel "Probabilistic Selection" (PS) task to specifically test this hypothesis.

1.1.2 Procedures

In the PS task, three different stimulus pairs (AB, CD, EF) are presented in random order. Feedback follows the choice to indicate whether it was correct or incorrect, but this feedback is probabilistic. Choosing stimulus A leads to correct (positive) feedback in 80% of AB trials, whereas choosing stimulus B leads to incorrect (negative) feedback in these trials. CD and EF pairs are less reliable: stimulus C is correct in 70% of CD trials, while E is correct in 60% of EF trials. Over the course of training participants learn to choose stimuli A, C and E more often than B, D, or F.

We enforced a performance criterion (evaluated after each training block of 60 trials) to ensure that all participants were at the same performance level before advancing to test. Because of the different probabilistic structure of each stimulus pair, we used a different criterion for each (65% A in AB, 60% C in CD, 50% E in EF)[2].

After reaching this criterion, participants were subsequently tested with the same training pairs, in addition to all novel combinations of stimuli, in random sequence. They were instructed (prior to the test phase) to use "gut instinct" if they did not know how to respond to these novel pairs. Each test pair was presented 6 times.

[1] Note that although aversive events have been reported to increase firing in a small proportion of cells in DA-producing brain areas (*3*), these neurons are in fact non-dopaminergic; DA cells in this region are indeed inhibited during aversive events (*4*). Further, we argue that "incorrect" feedback signals in trial-and error tasks are not actively aversive, and may be more analogous to a lack of positive reinforcement, widely accepted to induce DA dips in animal studies (*3,5*).

[2] In the EF pair, stimulus E is correct 60% of the time, but this is particularly difficult to learn. We therefore used a 50% criterion for this pair simply to ensure that if participants happened to "like" stimulus F at the outset, they nevertheless had to learn that this bias was not going to consistently work.

Group/Task	n	n filt	Sex ratio (m:f)	Age	Years Education	NAART IQ	Hoehn & Yahr ratio (1:2:2.5)
Seniors	19		10:9	64.8 (1.3)	17.1 (0.6)	118.5 (1.6)	N/A
PS	15	11	5:6	64.9 (1.1)	16.7 (0.9)	119.0 (1.9)	N/A
TI	18	17	9:8	64.1 (1.4)	16.7 (0.6)	118.4 (1.7)	N/A
PD patients	30		19:11	63.8 (2.2)	16.9 (0.4)	114.5 (1.9)	1:22:7
ON	29*	26	17:9	62.6 (2.6)	16.9 (0.5)	112.6 (2.0)	1:18:7
PS	12	9	7:2	69.1 (2.6)	16.9 (0.8)	112.3 (3.4)	0:6:3
TI	17	17	10:7	59.2 (3.5)	16.8 (0.7)	112.7 (2.6)	1:12:4
OFF	17	17	10:7	65.6 (3.3)	17.3 (0.7)	115.9 (2.5)	0:14:3
PS	9	9	5:4	62.3 (5.2)	16.8 (0.9)	113.2 (3.9)	0:8:1
TI	8	8	6:2	70.5 (1.5)	17.4 (0.9)	119.7 (1.9)	0:6:2

Table 1: Demographic variables for seniors and PD patients, with no significant differences between groups in any of the demographic variables. Imbalance in N's resulted from random assignment to task in session 1, together with differences in filtering. Because only a subset of participants returned for the second session, variables are broken down to show the number of participants in each condition that actually performed each task (PS = probabilistic selection; TI = transitive inference). The "n filt" column shows the number of remaining participants who were not filtered out for data analysis (see Data Filtering sections); participants who were filtered out were not included in the demographic means displayed for that row (as they were not used in the statistical comparisons). NAART IQ = premorbid IQ as estimated with the North American Adult Reading Test. For PD patients, disease severity is indicated in terms of the number of patients classified in Hoehn and Yahr stages 1, 2, and 2.5. Values represent mean (standard error). * Although other patients participated ON medication in the first session, one patient was OFF medication on that day (and did not return ON medication), explaining the discrepancy between the N for total PD patients and the N for patients ON medication. Because of a lack of session effect in either controls or patients, this does not confound any of our reported results.

1.2 Task II: Transitive Inference

1.2.1 Rationale

Our framework holds that PD patients are not impaired at learning probabilistic tasks *per se*. Rather, we argue that these tasks are simply examples of situations in which choices for particular stimuli are only differentiated by subtle reinforcement histories. We therefore predicted that other non-probabilistic tasks that engender subtle differences in associative strength would also be learned in a qualitatively different way by PD patients.

In the *Implicit Transitive Inference* (TI) task (*6*), the reinforcement for each stimulus pair is deterministic, but stimulus pairs are partially overlapping. Four pairs of stimuli are presented: A+B-, B+C-, C+D- and D+E-, where + and – indicate positive and negative feedback. During a subsequent testing phase, the novel combination BD is presented, and participants tend to choose stimulus B, despite having no explicit awareness of any hierarchical structure among the items (*6, 7*).

Although this behavior is consistent with explicit logical inference about the transitive relationship between B and D, there is considerable evidence that, in this implicit form, it is instead based on greater associative strengths for stimuli at the top of the hierarchy versus those at the bottom (*6*). Indeed, the same associative principles may give rise to transitive responding in animals trained in similar paradigms, and computational models (*8–14*).

Specifically, the associative strength hypothesis holds that the top and bottom pairs (AB, DE) "anchor" the development of associative weights: A obtains a strong positive association, while E obtains a strong negative one. This then carries over to adjacent pairs, such that B in BC has a stronger positive association, while D in CD has a stronger negative one (for details of how this happens, see (*9*)).

1.2.2 Procedures

Training consisted of four phases of blocked trials, followed by a fifth phase of randomly interleaved trials. Each phase was terminated after criterion performance of at least 75% correct across all pairs was achieved. In the first phase, the stimulus pairs were presented in blocks of 6 trials, such that the first block consisted of AB trials, the second block consisted of BC trials, and so on. In phase 2, blocks were shortened to 4 trials per block. Phase 3 consisted of 3 trials per block, and phase 4 consisted of 2 trials per block. In phase 5, all pairs were randomly interleaved for a total of 25 trials before criterion performance was evaluated. If criterion was not met, the random sequence was repeated. The test phase was similar to the training phase 5 in that all pairs were randomly interleaved. However, no feedback was provided and the two transitive pairs BD and AE were added to the mix of randomly ordered pairs. All pairs were presented 6 times each. Following the experiment, all participants were given a questionnaire to assess their awareness of the logical hierarchy of the stimuli, and to determine what strategies, if any, were used to respond to the novel test pairs (*6*). Out of 17 seniors, only one became explicitly aware of the hierarchical structure. No PD patients became explicitly aware.

2 Detailed Results and Analysis

2.1 Probabilistic Selection Results

2.2 Training

During the training phase of the PS task, all PD patients and seniors successfully learned to choose positively associated stimuli more often than negatively associated stimuli. Groups did not significantly differ in the number of training trials needed to reach criterion [OFF: 200(40); ON: 273(33); SEN: 250(48), $F(2,26) = 0.75$, n.s.]. In this and all subsequent cases, numbers in parentheses are standard errors (SEM). There was also no between-groups differences in accuracy, either across all training conditions [$F(2,26) = 0.55$, n.s.], or when considering just the critical AB pair alone [$F(2,26) = 0.74$, n.s.].

2.3 Test

2.3.1 Data Filtering

Because we were interested in the extent to which participants learned about the positive versus negative outcomes of their choices, we had to first ensure that they learned the basic task. While the training criteria were meant to address this issue, some participants were globally confused by the lack of feedback and addition of novel pairs during test and therefore performed poorly all around, including in pairs which were easiest for them during training. To reduce the amount of noise caused by this confound, we eliminated participants from the analysis who did not perform better than chance during test in the easiest training pair conditions. In the PS task, we eliminated three patients ON medication and four seniors who did not choose A over B more than 50% of the time when the AB pair was presented at test, reasoning that if they could not reliably choose A/avoid B in this pair, then the results in novel pairs were meaningless. The results described below apply to the remainder of participants, amounting to 11 seniors, 9 patients ON medication and 9 patients OFF medication.

2.3.2 Session Effects

To minimize practice and learning effects across session, each participant performed one task in the first session and the other task in the second. However, it is nevertheless possible that there were non-specific transfer effects across sessions that are unrelated to the particular details of each

Group	Session 1	Session 2
Seniors	0.021 (0.21)	-0.29 (0.2)
PD	0.026 (0.11)	0.15 (0.12)

Table 2: Session effects on test phase accuracy, converted to Z-scores to permit comparison across tasks. Values represent mean (standard error).

task. To examine this possibility, we performed an ANOVA across all participants to test for a main effect of session on overall test performance. We found no such effect [F(1,27) = 0.10, n.s.]. As shown in Table 2, the general trend for better performance of OFF medication patients relative to ON patients is evident by greater Z-scores in the second session. However, senior participants showed the reverse (non-significant) trend, supporting our claim that performance differences in the patient group were related to medication effects, rather than generalized transfer effect that theoretically could have carried over from the first session. Because patients in the second session were always OFF medication, we did not include session number in subsequent analyses, as it would likely mask any legitimate medication effects.

2.3.3 Test Pair Analysis

Next, we performed a general linear model (GLM) regression on positive/negative test pair accuracy, using between-subjects contrast-coded factors of patients versus seniors and patients ON versus OFF medication (forming a full set of orthogonal contrast codes) (*15*). There were no overall differences in performance levels between patients and seniors [F(1,26) = 1.0, n.s.], or between patients ON and OFF medication [F(1,26) = 0.57, n.s.]. There was no within-subjects main effect of positive/negative (A/B) test condition [F(1,26) = 0.14, n.s.]. Critically, within-subjects differences in positive/negative test condition interacted significantly with the ON/OFF medication status [F(1,26) = 4.33, p = .04] (see Figure 1b in the main paper). There was a trend for patients ON medication to perform better than those OFF medication at choosing positive stimuli [F(1,26) = 2.3, p = 0.14], while OFF patients performed better than ON patients when having to avoid negative stimuli [F(1,26] = 3.5, p = .07]. Finally, to explicitly compare performance of each of the medication conditions with the senior control participants, we performed two additional GLM's to test for the specific contrasts of patients ON medication versus seniors and patients OFF medication versus seniors (each of these tests was also accompanied by the orthogonal contrast code). Patients ON medication were non-significantly better than healthy seniors at choosing positive stimuli [F(1,26) = 1.6, n.s.] but worse than seniors at avoiding negative stimuli [F[1,26] =

0.2]. In contrast, patients OFF medication were numerically worse than seniors at choosing positive stimuli [$F(1,26) = 0.11$], but better than seniors at avoiding negative stimuli [$F(1,26) = 2.1$, $p = 0.16$].

2.3.4 Trial-to-trial Go/NoGo Learning from Positive/Negative Feedback

If our hypothesis is correct, we could potentially find evidence of it on a trial-to-trial basis during training, in terms of the effectiveness of positive/negative feedback in one trial to modify behavior in the subsequent trial of the same type. In other words, how much of participants' choice behavior in a given trial could be predicted by the type of feedback in the previous trial of the same type?

We restricted our analysis to the first 100 trials of training, because participants are less sensitive to feedback from individual trials in the latter parts of training as choice behavior stabilizes. We analyzed the conditional probabilities that (a) feedback on the previous trial was positive given that participants chose the same stimulus in the current trial; (b) feedback on the previous trial was negative given that participants switched to the alternative stimulus on the current trial. For the positive feedback case (a), the conditional probabilities were SEN: 55.4 (4.36); PD OFF 57.7 (3.7); PD ON 54.1 (2.3). There were no significant differences between PD groups in the efficacies of positive feedback in modifying behavior [$F(1,26) = 0.42$, n.s.]. For the negative feedback case (b), the conditional probabilities were SEN: 58.2 (2.44); PD OFF 62.0 (4.15); PD ON 49.4 (4.41). This difference between patients OFF and ON medication was significant [$F(1,26) = 5.62$, $p = .025$].

The enhancement in learning from positive feedback for the PD ON group was not apparent on a trial-to-trial basis during training, but was evident during test choice behavior. This suggests that greater levels of DA in medicated patients were effective when positive feedback had to be integrated over several trials, but were not sufficient to modify behavior after a single trial. In contrast, non-medicated patients were significantly more likely to avoid a negative stimulus, and this was evident both during test and in the effectiveness of a single negative feedback signal to modify behavior during training. That trial-to-trial effects were found for negative, but not positive, feedback is consistent with the fact that the tendency to avoid negative stimuli by OFF patients was generally greater than the tendency to choose positive stimuli by ON patients.

2.4 Transitive Inference Results

2.5 Training

During the training phase of the TI task, both Parkinson patients and aged-matched controls successfully learned to choose correctly across all stimulus pairs. Again, groups did not significantly differ in the number of training trials needed to reach the performance criteria required to advance to test [OFF: 280 (44); ON: 359 (35); SEN: 302 (39), F(2,39) = 1.03, n.s.]. There was also no between-groups differences in training phase accuracy [F(2,39) = 0.55, n.s.].

2.6 Test

2.6.1 Data Filtering

As in the PS task, we filtered out participants who were globally confused by the test phase and did not perform better than chance even at the easiest training pair conditions during test. This amounted to filtering out only one senior who did not perform better than 50% measured across anchor pairs AB and DE.[3] Results described below apply to the remainder of participants, amounting to 17 seniors, 17 patients ON medication and 8 patients OFF medication.

2.6.2 Session Effects

As in the PS task, an initial simple ANOVA revealed no main effects of session number [F(1,40) = 0.02, n.s.] on overall test pair performance.

2.6.3 Test Pair Analysis

As in the PS task, we performed a GLM regression on positive/negative test pair accuracy, testing the specific contrasts of patients versus seniors, and patients ON versus OFF medication. There were no overall differences in performance levels between patients and seniors [F(1,39) = 0.62, n.s.], or between patients ON and OFF medication [F[1,39] = 0.13, n.s.]. There was no within-subjects main effect of positive/negative (A/B) test condition [F(1,39) = 1.51, n.s.]. Critically, within-subjects differences in positive/negative test condition interacted significantly with

[3]The reason only one participant failed to transfer training knowledge to test performance in the TI task, while a total of seven failed in the PS task, may relate to the addition of novel test pairs. In the PS task, eight of the test pairs were novel with only three pairs that repeated from the training phase. In the TI task, there were only six test pairs, of which four of them were the original training pairs.

the ON/OFF medication status [F(1,39) = 5.54, p = .023] (see Figure 1c in the main paper). Patients ON medication performed non-significantly better than those OFF medication at choosing positive stimuli [F(1,39) = 1.6, p = 0.2], while OFF patients performed significantly better than ON patients when having to avoid negative stimuli [F(1,39] = 4.2, p = .047]. Finally, in comparison with healthy seniors, patients ON medication were non-significantly better than healthy seniors at choosing positive stimuli [F(1,39) = 1.6, n.s.] but worse than seniors at avoiding negative stimuli [F[1,39] = 0.5]. In contrast, patients OFF medication were numerically worse than seniors at choosing positive stimuli [F(1,39) = 0.1], but better than seniors at avoiding negative stimuli [F(1,26) = 2.3, p = 0.14].

Further insight comes from analysis of the more detailed pattern of results for each training pair (Figure 1), which was predicted by our computational models. The intact model learned "Go" to the stimulus at the top of the hierarchy (A) and NoGo to the one at the bottom (E). This enabled the model to allow B to take on a net positive value to facilitate performance in the BC pair, because this positive B value would not interfere with the very strong positive strength associated with A. A similar process held for the D item in the CD case, which took on a net negative value that did not compete with the very strong negative value for E. The resulting positive B and negative D values explains why B is chosen over D in implicit versions of the TI task in humans (*6*) and prior studies with animals (*8–10, 13*).

In contrast to this intact case, models with simulated PD (OFF medication) were biased to learn with a "NoGo strategy", such that they chose A in AB indirectly by learning NoGo to B. This led to worse performance on the BC pair, in which stimulus B should be chosen. In contrast, this NoGo bias helped performance on the CD pair by increasing the model's tendency to avoid D. This is exactly the pattern seen in Figure 1, where the OFF medication PD patients perform worse at BC, and better at CD.

Conversely, models with simulated DA medication were biased to learn with a "Go strategy", such that they learned Go to D in the DE pair instead of NoGo to E. This led to worse performance on the CD pair. However, the Go bias improved performance in the BC case, leading to a stronger tendency to select B. Again, this is the pattern seen in the ON medication PD patients in Figure 1, where they perform better at BC than CD, reversing the pattern seen in the OFF medication group.

The results for novel test pairs BD and AE are shown in Table 3. There was no significant effect of PD group on test pair performance [F(1,39) = 1.61, n.s.]. Note that we did not predict a difference in test pair performance, because participants can choose B over D or A over E *either*

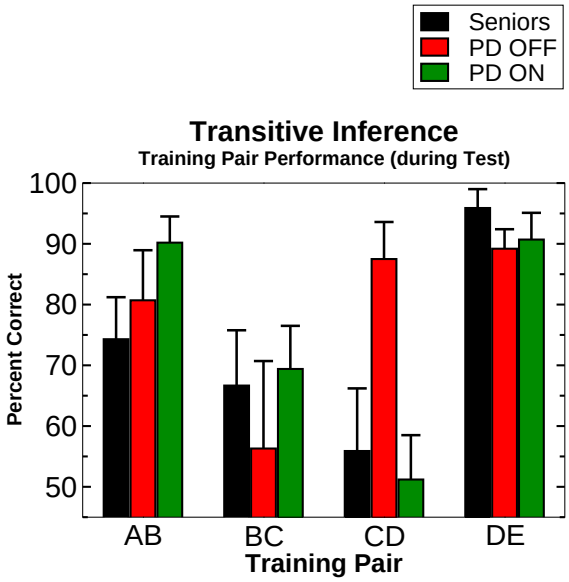

Figure 1: Training pair performance during the test phase of the transitive inference task. Patterns were consistent with model predictions: while all participants performed relatively well at the anchor pairs AB and DE, patients OFF medication performed poorly on the BC pair, whereas those ON medication had difficulty on the CD pair. See text for explanation.

Group	AE	BD
Seniors	98 (1.3)	61.8 (10.0)
PD OFF	95.8 (2.7)	66.2 (12.9)
PD ON	84.3 (5.8)	58.4 (9.5)

Table 3: Novel test pair results in the transitive inference task. Values represent mean (standard error). As predicted, groups did not differ in test pair generalization, since both AE and BD could be solved either by choosing the positive stimulus or by avoiding the negative stimulus.

by learning that A and B are good *or* by learning that D and E are bad.

2.7 Combined Z-score Analysis

Because we wanted to examine the extent to which participants learned from positive versus negative feedback across both tasks, we needed a way to equate performance on positively and negatively associated cues in the two tasks. To do this, we first computed Z-scores for performance in the two conditions for each task alone, by normalizing accuracy measures so that the distribution across all participants had a mean of 0 and a standard deviation of 1. In the PS task, we computed participants' Z-scores for test performance on A pairs and B pairs, while in the TI task we computed Z-scores for test performance on AB/BC pairs and CD/DE pairs. We then combined Z scores across the two tasks in positive and negative conditions, which were now in the same metric

(Figure 1d in the main paper).

We first performed a GLM regression to test for within-subject differences in positive versus negative test pair accuracy, and whether these differences interacted with group (patients versus seniors) or medication status (ON versus OFF). Across all participants, there was no within-subject main effect of positive/negative test condition [$F(1,68) = 0.3$, n.s.]. Critically, there was a highly significant interaction between Z-scores on positive/negative test conditions and PD medication status (ON/OFF) [$F(1,68) = 10.4$, $p = 0.0019$]. Each of the simple main effects were also significant when analyzed across both tasks: planned contrasts revealed that patients ON medication had significantly greater Z-scores for choosing positive stimuli than those OFF medication [$F(1,68) = 4.3$, $p = .04$]. Conversely, when avoiding negative stimuli, OFF patients had significantly greater Z-scores than ON patients [$F(1,68) = 8.0$, $p = .006$].

For further between-groups analysis, we performed two additional planned pair-wise comparisons across all participants—one for accuracy on positive test pairs, and one for accuracy on negative test pairs. For positive pairs, we tested whether Z-scores of patients ON medication were better than those of the other two groups. For negative pairs, we tested whether Z-scores of patients OFF medication were better than those of the other two groups. These analyses revealed that patients ON medication had significantly greater positive Z-scores than the other two groups combined [$F(1,69) = 4.8$, $p = .03$], while those OFF medications had significantly greater negative Z-scores than the other two groups [$F(1,69) = 7.6$, $p = .007$]. Compared with just the senior group, patients ON medication had marginally higher positive Z-scores [$F(1,69) = 3.3$, $p = .07$] and non-significantly lower negative Z-scores [$F(1,69) = 0.7$]. Conversely, when compared with seniors, those OFF medication had significantly greater Z-scores for avoiding negative stimuli [$F(1,69) = 4.6$, $p = .035$], with numerically but non-significantly lower positive Z-scores [$F(1,69) = 0.2$].

Finally, we also performed separate repeated measures ANOVA's for Z-scores of each medication condition, to evaluate whether patients ON medication were better at choosing positive stimuli than they were at avoiding negative stimuli (and vice versa for patients OFF medication). Indeed, across both tasks, patients ON medication were better at choosing positive stimuli than these same patients were at avoiding negative stimuli [$F(1,25)= 4.98$, $p = .03$]. Conversely, those OFF medication better avoided negative stimuli than they chose positive stimuli [$F(1,15) = 5.42$, $p=.03$]. Healthy seniors did not differ in their performance on positive versus negative stimuli [$F[1,18] = 0.02$, n.s.].

3 Model Methods & Statistics

As discussed in the main paper, the observed behavioral pattern of results in medicated and non-medicated PD patients was predicted by our computational model (*2*). To explicitly compare these results to the model, we simulated the probabilistic selection task in with three model manipulations (intact, simulated PD, and simulated DA medication, as described in the main paper). 25 networks with different sets of random initial synaptic weights were run with each manipulation. Following training, Go/NoGo associations were recorded from the model's striatum in response to input stimuli A and B (Figure 2c in the main paper). For positive Go learning, we computed Go - NoGo associations for stimulus A. For negative NoGo learning, we computed NoGo - Go associations for stimulus B.

We then performed a GLM exactly as in the behavioral analyses. This analysis revealed a highly significant interaction between stimulus condition and the contrast between simulated medication and simulated PD ($F[1,72] = 20.1$, $p < .0001$). This is similar to the behavioral interaction observed between choosing A and avoiding B and the ON/OFF medication contrast in PD patients. Networks with simulated DA medication had significantly higher overall Go associations for stimulus A then did networks with simulated PD [$F(1,72) = 13.6$, $p = .0004$]. Conversely, PD networks had greater NoGo associations for stimulus B than did those with simulated medication [$F(1,72) = 9.2$, $p = .003$].

The observed significant differences between simulated PD and medication networks in Go/NoGo learning could result either from enhancements or impairments in the respective conditions (or both). To disentangle these possibilities, we performed separate planned comparisons that compare Go/NoGo associations of each of the DA manipulations networks to those of the intact networks. Compared with intact networks, simulated PD resulted in significantly lower Go associations to stimulus A [$F(1,48) = 7.8$, $p = .0075$], with non-significantly greater NoGo associations to stimulus B [$F(1,48) = 0.93$, n.s.]. Conversely, networks with simulated DA medication had significantly weaker NoGo associations to stimulus B [$F(1,48) = 4.33$, $p = .04$], and numerically but not significantly stronger Go associations to stimulus A [$F(1,48) = 0.75$, n.s.].

3.1 Model Interpretation Issues

There might appear to be some conflict between the way the model is described in the Frank (in press) paper (*2*), and the way it is characterized here. Specifically, whereas in this paper we

emphasize the NoGo bias effects of low tonic dopamine levels in Parkinson networks, Frank (in press) also discusses the decreased "dynamic range" effects of decreased dopamine levels on phasic dips. In fact, both effects are present in the model. Thus, the model has an overall NoGo bias, which means that if it is going to learn anything at all, it will be more able to learn to avoid negative stimuli than to choose positive stimuli. However, because the amount of tonic dopamine is already low, phasic dips during negative feedback actually have less of a differential effect on NoGo learning than in intact networks. This dynamic range issue may result in decreased ability to resolve fine differences in Go and NoGo associations across multiple cue configurations (as in more complex probabilistic categorization tasks (*16*)). However, this issue is not relevant in tasks that allow one to adopt either a simple Go or NoGo strategy, without having to resolve fine differences. For instance, in our probabilistic selection task, adopting a global NoGo strategy can lead to preserved training performance because each trial always involves one stimulus that has an overall negative association and should be avoided.

4 Appendix

4.1 Post-Experiment Questionnaire Analysis for the Transitive Inference Task

Seven questions were asked, as follows.

1. Do you have any prior knowledge of the symbols used in the experiment?
2. If you answered "Yes" to question 1, please indicate to what extent you are familiar with these characters.
3. Did you have the impression that some of the pairs were easier to choose from than others?
4. Did you think any of the symbols were ALWAYS correct (no matter what the other symbol was)?
5. Did you think any of the symbols were ALWAYS incorrect (no matter what the other symbol was)?
6. Did you have the impression that there was some kind of logical rule, order, or hierarchy of symbols in the experiment? If so, please explain briefly.

7. In the test phase, were there any new symbols or new combinations of symbols?

8. If you answered "Yes" to question 7, how did you make your choice in these cases? (e.g., guessed, went with instinct, used some sort of rule - explain)

Awareness judgments were made by assessing the above written questionnaires, and asking participants to clarify some responses, while being completely blind to their performance (*6*).

Only one senior was judged to be explicitly aware of the logical hierarchy ordering. The remaining 16 seniors and 17 PD patients were judged to be unaware of any notion of logical order or hierarchy among premise pairs. Eight out of 16 seniors, 8 out of 17 patients ON medication, and 4 out of 8 patients OFF medication did in fact notice that there was a stimulus that was always correct (stimulus A). Nine seniors, 6 patients ON medication, and 4 patients OFF medication noticed that there was a stimulus that was always incorrect (stimulus E). There were no differences between groups in these explicit awareness measures. When asked to describe the "rule", some of these participants stated that they memorized specific pairs, but could not describe any notion of logical order, and didn't explicitly know how to respond to the novel test pairs because they had not memorized the correct response to them during training. Moreover, the "unaware" participants that advanced to the test phase did not use any logical rule or order to determine their choices during test. Many didn't notice there were novel test pairs that differed from the training pairs, and those that did simply "guessed", or went with "instinct." A few participants employed an explicit rule that was incorrect (e.g., "I chose the symbol that was widest").

References

1. J. Blair, O. Spreen, *Clinical Neuropsychologist* **3**, 129 (1989).

2. M. Frank, *Journal of Cognitive Neuroscience* (in press).

3. W. Schultz, *Neuron* **36**, 241 (2002).

4. M. Ungless, P. Magill, J. Bolam, *Science* **303**, 2040 (2004).

5. T. Satoh, S. Nakai, T. Sato, M. Kimura, *Journal of Neuroscience* **23**, 9913 (2003).

6. M. Frank, J. Rudy, W. Levy, R. O'Reilly, *Memory and Cognition* (in press).

7. A. J. Greene, B. A. Spellman, J. A. Dusek, H. B. Eichenbaum, W. B. Levy, *Memory and Cognition* **29**, 893 (2001).

8. M. VanElzakker, R. C. O'Reilly, J. W. Rudy, *Hippocampus* **13**, 334 (2003).

9. M. Frank, J. Rudy, R. O'Reilly, *Hippocampus* **13**, 341 (2003).

10. L. von Fersen, C. D. L. Wynne, J. D. Delius, J. E. R. Staddon, *Journal of Experimental Psychology: Animal Behavior Processes* **17**, 334 (1991).

11. C. Wynne, *Animal Learning & Behavior* **23**, 207 (1995).

12. C. Wynne, *Models of Action*, C. Wynne, J. Staddon, eds. (Lawrence Erlbaum Associates, New Jersey, 1998), pp. 269–307.

13. J. Delius, M. Siemann, *Behavioural Processes* **42**, 107 (1998).

14. M. Siemann, J. Delius, *European Journal of Cognitive Psychology* **10**, 307 (1998).

15. C. M. Judd, G. H. McClelland, *Data Analysis, A Model-Comparison Approach* (Harcourt Brace Jovanovich, Orlando, FL, 1989).

16. B. J. Knowlton, J. A. Mangels, L. R. Squire, *Science* **273**, 1399 (1996).

17. Supported by ONR grants N00014-00-1-0246 and N00014-03-1-0428, and NIH grant MH069597-01. We thank the staff of the Colorado Neurological Institute for their assistance, and Karen Richardson for help in administering cognitive tasks to patients and senior participants. We also thank Tim Curran, Gary McClelland, and Yuko Munakata for helpful comments.