

# Reward-Dependent Modulation of Movement Variability

Sarah E. Pekny,<sup>1\*</sup> Jun Izawa,<sup>1,2\*</sup> and Reza Shadmehr<sup>1</sup>

<sup>1</sup>Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, and

<sup>2</sup>Faculty of Engineering, Information and Systems, University of Tsukuba, Ibaraki 305-8573, Japan

Movement variability is often considered an unwanted byproduct of a noisy nervous system. However, variability can signal a form of implicit exploration, indicating that the nervous system is intentionally varying the motor commands in search of actions that yield the greatest success. Here, we investigated the role of the human basal ganglia in controlling reward-dependent motor variability as measured by trial-to-trial changes in performance during a reaching task. We designed an experiment in which the only performance feedback was success or failure and quantified how reach variability was modulated as a function of the probability of reward. In healthy controls, reach variability increased as the probability of reward decreased. Control of variability depended on the history of past rewards, with the largest trial-to-trial changes occurring immediately after an unrewarded trial. In contrast, in participants with Parkinson's disease, a known example of basal ganglia dysfunction, reward was a poor modulator of variability; that is, the patients showed an impaired ability to increase variability in response to decreases in the probability of reward. This was despite the fact that, after rewarded trials, reach variability in the patients was comparable to healthy controls. In summary, we found that movement variability is partially a form of exploration driven by the recent history of rewards. When the function of the human basal ganglia is compromised, the reward-dependent control of movement variability is impaired, particularly affecting the ability to increase variability after unsuccessful outcomes.

**Key words:** computational models; motor control; Parkinson's disease; reaching

## Introduction

*How shall I know, unless I go to Cairo and Cathay, whether or not this blessed spot is blest in every way?*

Edna St. Vincent Millay, 1922

Movement variability is often considered an unwanted and unavoidable byproduct of noise in the nervous system. However, behavioral evidence suggests that variability serves a critical role in motor learning (Tumer and Brainard, 2007; Wu et al., 2014). Variability may benefit learning because carefully controlled fluctuations in motor output can serve as a form of exploration, allowing the animal to find a better solution for achieving a goal. Indeed, variability fluctuates in response to changes in probability of success and failure. For example, humans increase their movement variability during periods of low success or minimal feedback, which is thought to reflect a search for a rewarding outcome (Izawa and Shadmehr, 2011; Galea et al., 2013). Similarly, monkeys increase the variability of their saccadic eye movements, altering peak velocity, latency, and amplitude, when their

movement is not paired with reward (Takikawa et al., 2002). When variability in a lever-pressing task is rewarded instead of repetition, pigeons can produce highly variable lever sequences similar to those produced by a random number generator (Page and Neuringer, 1985).

Deciding whether to repeat a movement or vary one's actions depends on the ability to predict future occurrences of punishment or reward. The difference between the actual and expected outcome is reward prediction error, which relies on dopamine-dependent processes (Schultz et al., 1997). It is therefore not surprising that variability, especially in terms of goal-directed exploration, has been linked to dopamine and the basal ganglia. In songbirds, the source of variability in song production is believed to be in brain structures homologous to the mammalian basal ganglia (Kao et al., 2005; Olveczky et al., 2005). Activating striatal D1 and D2 receptors in mice alters the decision process to stay with or switch from the current behavior to obtain reward (Tai et al., 2012). During periods of low variability, administration of a D2 agonist increases variability in rats (Pesek-Cotton et al., 2011). In humans, a D2 antagonist abolishes the increase in variability observed during periods of low reward (Galea et al., 2013).

Given this potential link between control of movement variability and the basal ganglia, we hypothesized that patients with basal ganglia dysfunction would have difficulty controlling their motor variability in response to reward prediction errors. Indeed, patients diagnosed with Parkinson's disease (PD) are known to have difficulties in certain cognitive learning tasks that depend on trial and error feedback (Knowlton et al., 1996), with some evidence suggesting a specific learning deficit based on negative re-

Received Aug. 5, 2014; revised Jan. 21, 2015; accepted Jan. 27, 2015.

Author contributions: J.I. and R.S. designed research; J.I. performed research; S.E.P. and J.I. analyzed data; S.E.P. wrote the paper.

This work was supported by the National Institutes of Health (Grant NS078311).

The authors declare no competing financial interests.

\*S.E.P. and J.I. contributed equally to this work.

Correspondence should be addressed to Sarah Pekny, Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins University School of Medicine, 416 Traylor Building, 720 Rutland Avenue, Baltimore, MD 21205. E-mail: sep205@gmail.com.

DOI:10.1523/JNEUROSCI.3244-14.2015

Copyright © 2015 the authors 0270-6474/15/354015-10\$15.00/0

**Table 1. Clinical characteristics of the PD group**

Subject ID	Age	Handedness	Sex	Disease duration (y)	Motor UPDRS	Total UPDRS
PD1	77	R	F	4	10	18
PD2	51	R	F	0.25	7	12
PD3	65	R	M	4	8	23
PD4	61	R	F	5	9	14
PD5	65	R	M	1	5	23
PD6	68	R	M	2	6	6
PD7	61	R	M	1	4	6
PD8	61	R	M	4	4	7
PD9	62	R	F	0.5	9	13

ward prediction errors (Frank et al., 2004; Frank et al., 2007; Bódi et al., 2009). Here, we considered a reaching task and provided subjects with binary feedback about the success of the reach. We manipulated the probability of reward and quantified the resulting changes in variability in healthy and PD populations.

## Materials and Methods

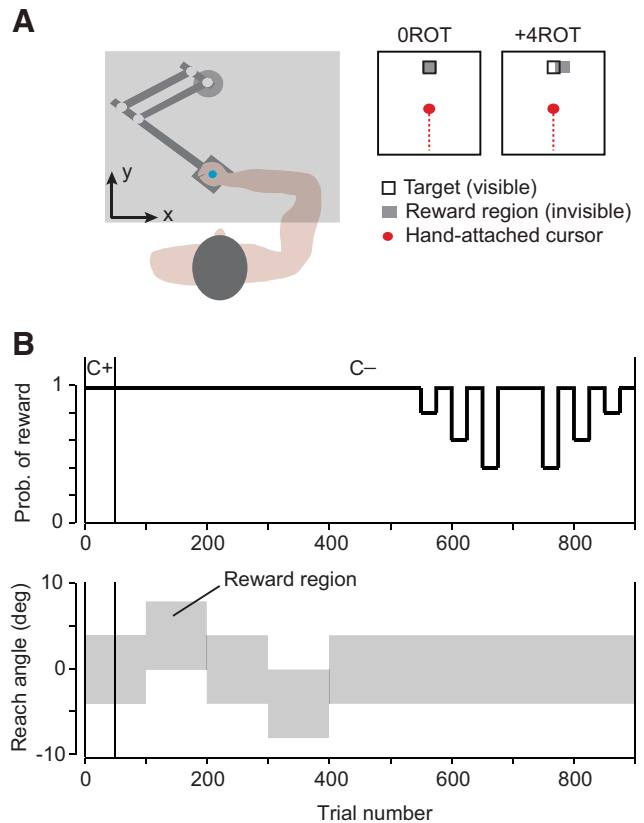
**Subjects.** A total of  $n = 26$  subjects participated in our study. Among them were  $n = 9$  mildly affected patients diagnosed with PD ( $63 \pm 6.9$  years old, including 4 females and 5 males) and  $n = 8$  healthy age-matched controls ( $65 \pm 8.1$  years old, including 4 females and 4 males). Because the dopaminergic system naturally undergoes degeneration with aging (Fearnley and Lees, 1991; Vaillancourt et al., 2012), we also included in our study a group of  $n = 9$  healthy young controls ( $25 \pm 5.6$  years old, mean  $\pm$  SD, including 7 females and 2 males) for comparison. All participants provided consent by signing a form approved by the Johns Hopkins University School of Medicine Institutional Review Board.

**PD patients.** All PD patients were free of dementia as assessed by a Mini-Mental Status Examination (Folstein et al., 1975) on which all subjects scored better than 28. Clinical severity was measured by using the Unified Parkinson's Disease Rating Scale (Movement Disorder Society Task Force on Rating Scales for Parkinson's disease, 2003), the results of which are provided in Table 1. All subjects were free of musculoskeletal disease and had no neurological disease other than PD, as confirmed by a neurologist. All subjects were taking dopamine agonist medications at the time of testing.

**Behavioral task.** The experimental task was similar to those described in a previous study (Izawa and Shadmehr, 2011). Participants made shooting movements toward a single target in the horizontal plane holding onto the handle of a two-joint robotic manipulandum (Fig. 1A). An opaque screen was placed above the subject's arm, upon which a video projector painted the scene. At the start of each trial, a target of  $6^\circ$  width in reach space located 10 cm from start was displayed at  $90^\circ$  from horizontal. This single target was used for all trials through the experiment. Participants were instructed to make quick, shooting movements so that the robotic handle passed through this target. Once the participant finished a movement, the robot again guided the hand back to the start position.

Success was indicated after every reach via an animated target explosion when the participant's hand passed through an experimentally controlled rewarding target region. Movements were also required to have a reaction time (RT) of  $<0.6$  s and a movement time (MT) of  $<1$  s to be successful. After a successful reach, a point was added to the participant's score, which was displayed throughout the experiment. This target explosion and point were the reward given in our task. Participants were compensated for their time and the total payment was not based on task score.

All participants first performed a familiarization block of 50 trials in which full visual feedback of the movement was provided via a projected cursor ( $5 \times 5$  mm) representing hand position. These movements were performed in a 0ROT condition in which participants were rewarded if they passed within a region of  $\pm 4^\circ$  (in reach space) centered at  $0^\circ$ , the target center (rewarding target region is highlighted in gray in Fig. 1A,B). A clockwise rotation is defined as positive. After this initial block of

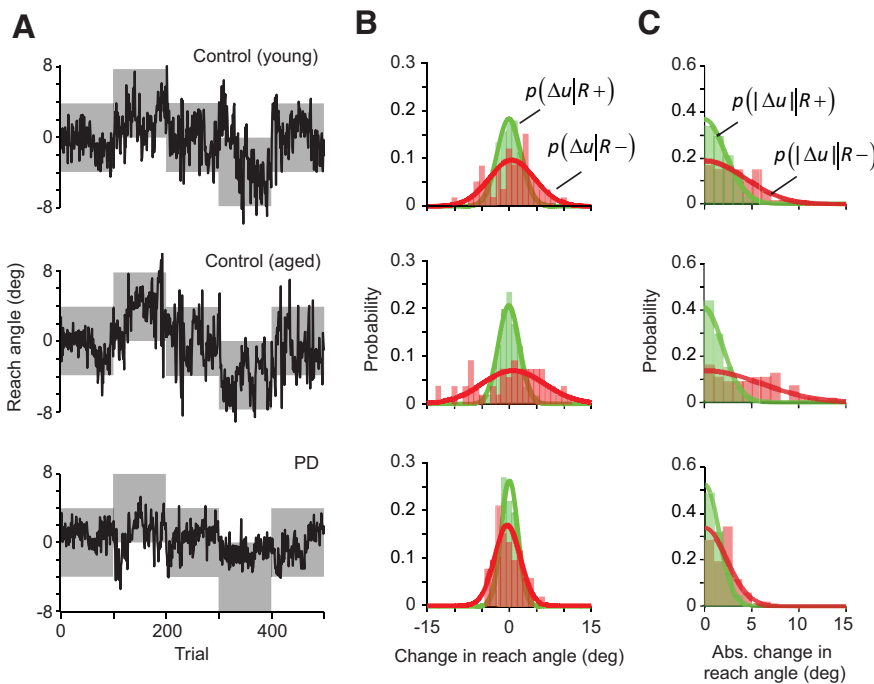


**Figure 1.** Experimental setup. **A**, Participants held the handle of a robotic manipulandum and made “shooting” movements through a target located 10 cm away from start. A screen was placed over the subject's hand to obscure the view of their arm and the robotic handle. **B**, During the first 50 trials of the experiment, a cursor was projected onto the screen to provide the subject with visual feedback of their movement (C+). After these 50 trials, the cursor was shut off (C−). For the remaining trials, the only feedback the participant received was a brief target explosion and a point added to their score after a successful trial. The gray area indicates the region through which the participant needed to reach to receive successful feedback. In the second half of the experiment, this region was held constant and the probability of receiving a reward (if the hand crossed through this area) was manipulated.

training, cursor feedback was shut off and participants did not receive visual information about the handle position for the remainder of the experiment. The only performance feedback that participants continued to receive was regarding success or failure of the trial. After the visual cursor feedback was removed, participants performed another block of 50 trials in the 0ROT condition. Participants next experienced a block of 100 trials in which, unbeknownst to the subject, the rewarding target region was shifted and centered at  $+4^\circ$  from midline, referred to as the +4ROT condition (reaches were now rewarded if they fell between  $0^\circ$  and  $+8^\circ$ , as illustrated in Fig. 1A,B). This block of trials was followed by two blocks of 50 trials in which the rewarding target region was adjusted again and then returned back to the 0ROT condition. Subjects then performed a block of 100 trials in a  $-4$ ROT condition (reaches now rewarded if they fell between  $-8^\circ$  and  $0^\circ$ ). Two blocks of 50 trials in the 0ROT condition followed this perturbation.

For the remainder of the experiment, the participants performed two blocks of 200 trials in the 0ROT condition, but the probability of reward was now controlled. For example, in the 40% reward condition, if the movement placed the cursor in the rewarding target area, the probability of reward was 40%. This reward probability was altered and left constant for 25 sequential trials. Participants experienced each reward condition of 40%, 60%, 80%, and 100% over two separate 25 trial intervals, as shown in Figure 1B.

**Data analysis.** Hand position and velocity were recorded at the robotic handle at 100 Hz and analyzed offline with MATLAB R2009b. The main



**Figure 2.** Participants responded differently to positive versus negative task outcomes. **A**, Reach angles of representative subjects. The gray area indicates the region where the reach would be rewarded. **B**, Signed change in reach angle for representative subjects. Plotted in green are the signed changes in reach angle between trial  $n$  and  $n+1$  for each subject after a successful trial ( $R+$ ). The changes in reach angle after an unsuccessful trial ( $R-$ ) are plotted in red. The fit of these changes to a normal distribution are plotted over the histogram of each individual’s data. **C**, Absolute change in reach angle for representative subjects. The fit of these absolute changes to a folded normal distribution are plotted over the histogram of each individual’s data.

variable for performance was the reach angle of the participant’s movement. First, a reach end point was defined as the point at which the participant’s hand crossed a circle with radius 10 cm centered at the start position. A reach angle was calculated for each movement as the angle between the hand path from start to reach end point and the line connecting start to target center.

For each movement, we also calculated the participant’s RT as the duration of time between target appearance and the hand velocity crossing a threshold of 0.03 m/s. MT was measured from the moment the hand crossed this initial velocity threshold until movement termination, when the hand passed a circle with radius 10 cm centered at the start point. Finally, an intertrial interval (ITIs) was calculated as the time between movement termination and the appearance of the target for the next trial.

Statistical analysis was performed using IBM SPSS version 22. All one-way ANOVA were tested for the assumption of homogeneity of variance using the Levene’s F test for equality of variance. For those one-way ANOVA tests in which this assumption is violated, the Brown–Forsythe statistic is reported. In these cases, the Games–Howell *post hoc* test was then used. For cases in which the assumption of homogeneity of variance has been met, the Tukey (HSD) test was used for *post hoc* analysis.

## Results

### Baseline reach variability was comparable between groups

Participants began the experiment with a familiarization block (50 trials) in which visual feedback was provided via a cursor ( $C+$  trials, Fig. 1B). These reaches were performed with veridical visual feedback (termed 0ROT condition) in which the invisible reach reward region ( $\pm 4^\circ$  in reach space, gray region in Fig. 1A) was centered on the visible target (black box in Fig. 1A). We observed no statistically significant differences among groups in the number of successful trials ( $F_{(2,13.211)} = 2.203, p = 0.149$ , one-way ANOVA for total reward in last 25 trials), reach variability (trial-to-trial change in reach direction,  $F_{(2,25)} = 0.300, p = 0.743$ , one-way ANOVA for average absolute difference in reach

angle in last 25 trials), or reach peak velocity ( $F_{(2,25)} = 0.877, p = 0.430$ , average maximum velocity in last 25 trials).

After this baseline block, cursor feedback was removed ( $C-$  trials, Fig. 1B) and participants performed another block of 50 trials in the 0ROT condition. We again found no statistically significant differences across groups in terms of the number of successful trials ( $F_{(2,25)} = 0.967, p = 0.395$ , one-way ANOVA for reward in last 25 trials), reach variability (trial-to-trial change in reach direction,  $F_{(2,25)} = 0.677, p = 0.517$ , one-way ANOVA for average absolute difference in reach angle in last 25 trials), or reach peak velocity ( $F_{(2,25)} = 1.578, p = 0.228$ , average maximum velocity in last 25 trials).

Therefore, the patients were able to perform the task successfully even with the absence of visual feedback. In addition, there was no evidence of baseline differences in trial-to-trial reach variability or success rate across the groups.

### Reach variability increased after an unrewarded trial

In trials 100–500 (Fig. 1B), we shifted the reward region covertly with respect to the target, requiring participants to alter their reach direction to continue receiving reward.

Because no cursor feedback was available in these and all subsequent trials, the only information provided at the end of each trial was the successful acquisition of reward ( $R+$ ) or failure ( $R-$ ).

During trials 100–200, the reward region was shifted by  $+4^\circ$  (termed  $+4$ ROT condition). That is, the reaches were rewarded only if the hand crossed between  $0^\circ$  and  $+8^\circ$  in reach space, as illustrated in Figure 1B. This block of training was followed by 100 trials of washout in which the reward region was returned to the 0ROT condition. Participants then experienced 100 trials in the  $-4$ ROT condition, followed by another 100 trials of washout in the 0ROT condition.

Reach angles are plotted in Figure 2A for a typical subject from each group. (These three participants were selected for display because they achieved similar scores during this block of trials, receiving reward on 88.0%, 88.8%, and 89.4% of the 500 trials for the young control, aged control, and PD patient, respectively.) The data in Figure 2A suggest that the subjects varied their reach to find the reward zone. To analyze the data, we quantified how much the reach angle changed from one trial to the next as a function of whether the initial trial was rewarded ( $R+$ ) or not ( $R-$ ). In this analysis, we measured change in reach angle  $u$  from trial  $n$  to trial  $n+1$ , and represented this change as follows:

$$\Delta u^{(n)} = u^{(n+1)} - u^{(n)} \quad (1)$$

We quantified the change in reach angles after each  $R+$  trial, resulting in the conditional probability distribution  $p(\Delta u|R+)$  for each subject (green colored distribution, Fig. 2B). Similarly, we quantified the change in reach angles after each  $R-$  trial, resulting in the distribution  $p(\Delta u|R-)$  (red colored distribution, Fig. 2B). As a proxy for variability, we also computed the quantity  $|\Delta u^{(n)}|$ ,

which provided a measure of the unsigned “motor exploration” that followed a rewarded or unrewarded trial. The conditional probability distributions  $p(|\Delta u||R+)$  and  $p(|\Delta u||R-)$  for each subject are plotted in Figure 2C.

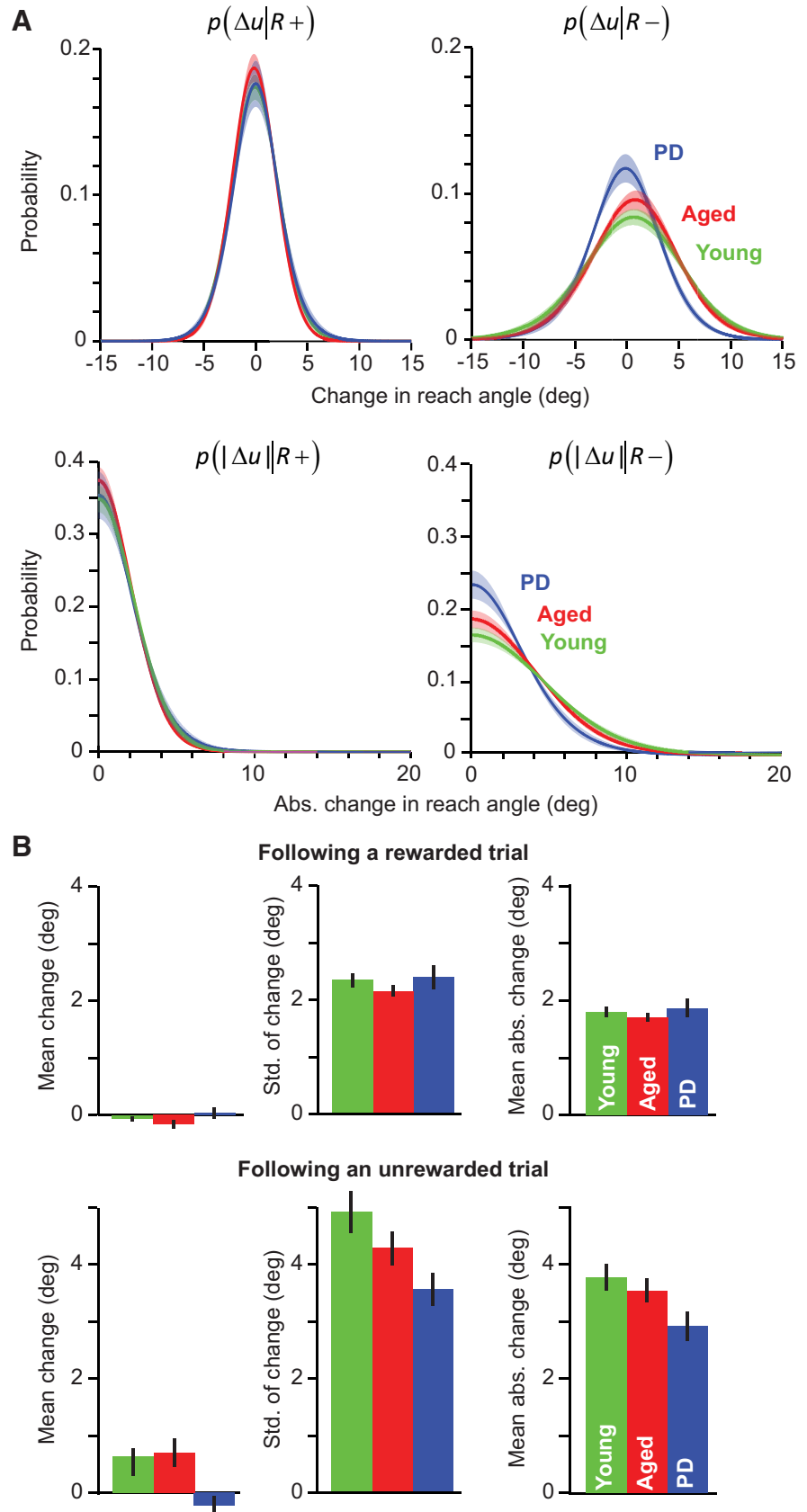
We found that, in general, the change in motor commands was greater after an R- trial compared with an R+ trial, as indicated by the fact that the red-colored probability distributions in Figure 2, B and C, were broader than the green-colored distributions. This indicated that, after an unrewarded trial, the subjects changed their reach angle by a larger amount than after a rewarded trial. Importantly, in the representative PD patient, the distribution after an R+ trial appeared similar to the two healthy controls (green distributions, Fig. 2B,C). However, the distribution for  $p(\Delta u|R-)$  and  $p(|\Delta u||R-)$  appeared narrower than normal. This suggested that, for the PD subject, there was less change in the reach angles after an unrewarded trial than in the healthy controls.

To compare reach variability across groups after R+ and R- trials, we estimated  $p(\Delta u|R+)$  and  $p(\Delta u|R-)$  via a normal distribution for each subject first and then computed the group mean  $\mu$  and SD  $\sigma$  from the resulting distribution of means. The results are shown in the top row of Figure 3A. Similarly, we fitted a folded normal distribution to the measured  $|\Delta u|$  data for each subject after an R+ and an R- trial, using the following equation:

$$f(|\Delta u|; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(-|\Delta u| - \mu)^2}{2\sigma^2}\right) + \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(|\Delta u| - \mu)^2}{2\sigma^2}\right) \quad (2)$$

From the estimate of each subject’s mean, we then computed the probability distribution for each group, as shown in the bottom row of Figure 3A. The results suggested three ideas: (1) the variability (as indicated by the width of the distributions) after a rewarded trial was comparable in the three groups; (2) in all groups, the variability increased after an unrewarded trial; and (3) in PD, the variability after an unrewarded trial was smaller than normal.

Using repeated-measures ANOVA with reward condition as the within-subject measure and groups as the between-subject measure, we found both a significant effect



**Figure 3.** PD participants are less sensitive to negative outcomes. **A**, Group average for the individual subject fit to a normal distribution for change in reach angle, and fit to a folded normal distribution for the absolute change in reach angle after a successful (R+) or unsuccessful trial (R-). Shaded area indicates between-subject SEM. **B**, The mean and SD of the signed change in reach angle and mean absolute change in reach angle for each group after rewarded or unrewarded trials (R+ or R-). Error bars indicate SEM.

of condition (R+ vs R−) and a significant interaction between condition and group for all measures of reach variability. The SD of  $\Delta u$  was significantly affected by reward condition (effect of condition,  $F_{(1,23)} = 159.339$ ,  $p < 0.001$ ) and displayed a significant condition by group interaction ( $F_{(2,23)} = 7.442$ ,  $p = 0.003$ ). Similarly,  $|\Delta u|$  was significantly affected by reward condition ( $F_{(1,23)} = 193.806$ ,  $p < 0.001$ ) and displayed a significant condition by group interaction ( $F_{(2,23)} = 6.231$ ,  $p = 0.007$ ). We observed that, across all groups, these measures were greater after an R− trial than an R+ trial (*post hoc* on condition;  $p < 0.001$  for SD of change;  $p < 0.001$  for absolute mean change).

A *post hoc* test in which we analyzed each condition individually revealed that, after an R− trial, variability was significantly different between the PD and the young control groups. After an R− trial, the SD of  $\Delta u$  was significantly smaller in the PD group ( $p = 0.015$ , ne-way ANOVA,  $F_{(2,25)} = 4.641$ ,  $p = 0.030$ ). Similarly, after an R− trial, the measure  $|\Delta u|$  was significantly smaller in the PD group ( $p = 0.040$ , after one-way ANOVA on mean absolute change in reach angle,  $F_{(2,25)} = 3.640$ ,  $p = 0.042$ ). However, there were no statistically significant differences between the young and age-matched controls for either measure ( $p = 0.367$  for SD of signed change,  $p = 0.775$  for mean absolute change) nor between the PD and age-matched controls ( $p = 0.277$  for SD of signed change,  $p = 0.175$  for mean absolute change), indicating that, in this part of the experiment, behavior of the age-matched subjects fell somewhere between the young controls and PD patients.

In addition to the above changes in reward-dependent variability of motor output, there were also small but consistent changes in the biases of the motor commands, particularly after an R− trial in the two control groups, as shown in the right column of Figure 3A. We found that, after an R− trial, the control groups preferred a change in reach angle that was slightly clockwise to the previous movement (an effect of condition on the mean change in reach angle,  $F_{(1,23)} = 6.293$ ,  $p = 0.020$ ; and condition by group interaction  $F_{(2,23)} = 1.512$ ,  $p = 0.021$ ). This bias was not significantly different between the two control groups ( $p = 0.775$ , after a one-way ANOVA on mean change in reach angle after an R+ trials,  $F_{(2,25)} = 3.640$ ,  $p = 0.042$ ) and the PD patients did not show this bias in behavior ( $p = 0.047$  PD vs young control,  $p = 0.017$  PD vs age-matched controls). Presence of this bias after an R− trial is interesting because it is in the direction for which the effective mass of the arm becomes smaller (Gordon et al., 1994), resulting in movements that have reduced effort.

In contrast to the between group differences after R− trials, the various groups behaved similarly after an R+ trial. A *post hoc* test in which we analyzed each condition individually revealed that, after an R+ trial, the mean change in reach angle did not differ significantly between the three groups (one-way ANOVA on mean change in reach angle after an R+ trial,  $F_{(2,25)} = 0.512$ ,  $p = 0.606$ ). After an R+ trial, variability was also not different among the PD, aged, and young groups (one-way ANOVA on SD of signed change in reach angle after an R+ trial,  $F_{(2,25)} = 0.590$ ,  $p = 0.563$ ; on mean absolute change in reach angle after an R+ trial,  $F_{(2,25)} = 0.512$ ,  $p = 0.606$ ).

Was this policy of changing the reach angle after an unsuccessful trial useful in acquisition of reward? We found that the PD patients had a lower average score (number of rewarded trials) at the end of the +4ROT and −4ROT conditions (*post hoc* comparisons,  $p = 0.038$  and  $p = 0.007$  against age-matched and young controls, respectively, after a significant one-way ANOVA

on total reward in last 25 trials from both blocks,  $F_{(2,16,098)} = 6.638$ ,  $p = 0.008$ ). Therefore, after an unsuccessful trial, increasing the reach direction variability was a good strategy for acquiring more reward.

In summary, when a motor command was rewarded (R+ trials), the trial-to-trial change in the command was similar across all three groups. When a motor command was not rewarded (R− trials), there was a larger trial-to-trial change. However, after an R− trial, the PD patients changed their movements less than the control groups in terms of both a bias in behavior and the variability around that bias. This hinted that sensitivity to an unrewarded trial was lower in PD than in controls. To test systematically for the relationship between reward and change in motor commands, we performed an additional experiment.

### Relationship between probability of reward and reach variability

To test directly whether the absence of reward resulted in increased trial-to-trial change in the reach angles, we controlled the probability of reward on each trial (final 400 trials of Fig. 1B). In this part of the experiment, all trials were in the 0ROT condition but we regulated the probability of reward: if the subject's reach placed the unseen cursor in the reward region, reward was provided at a probability of 40%, 60%, 80%, or 100% for bins of 25 trials, as shown in Figure 1B.

For each subject, we computed the mean and SD of  $\Delta u$  in each probability condition, as well at the mean of  $|\Delta u|$ . As illustrated in the left column of Figure 4A, we observed no effect of reward condition on the mean  $\Delta u$ , nor an interaction between reward condition and group (effect of reward condition,  $F_{(2,359,54,258)} = 1.582$ ,  $p = 0.212$  and condition by group interaction  $F_{(4,718,54,258)} = 2.249$ ,  $p = 0.066$ ). Therefore, during this second portion of the experiment, there were no biases in the trial-to-trial changes in reach angle. Importantly, in the healthy groups both measures of motor variability (left column of Fig. 4B, C) were largest when the probability of reward was lowest and then gradually declined as the probability of reward increased. Therefore, in healthy controls, we found that a lower probability of reward coincided with larger trial-to-trial change in reach angle.

This reward-dependent modulation of variability was missing in the PD group. Rather, the PD patients appeared to exhibit approximately the same level of  $\Delta u$  across all reward probabilities. A repeated-measures ANOVA with reward probability as the within-subject measure and groups as the between-subject measure found a significant group by reward interaction for both the SD of  $\Delta u$  ( $F_{(6,69)} = 3.096$ ,  $p < 0.001$ ) and mean of  $|\Delta u|$  ( $F_{(6,69)} = 4.699$ ,  $p < 0.001$ ). A *post hoc* test in which we analyzed each group individually revealed a significant effect of reward probability for both control groups (one-way ANOVA on reward probability for the SD of  $\Delta u$ ,  $F_{(3,31)} = 6.122$ ,  $p = 0.002$  for aged controls and  $F_{(3,35)} = 8.648$ ,  $p < 0.001$ , and for  $|\Delta u|$ ,  $F_{(3,31)} = 6.51$ ,  $p = 0.002$ , for aged controls and  $F_{(3,35)} = 9.01$ ,  $p < 0.001$  for young control). In contrast, in the patients, we found no significant effect of reward probability on the SD of  $\Delta u$  ( $F_{(3,35)} = 0.271$ ,  $p = 0.846$ , one-way ANOVA on reward condition) or on the mean of  $|\Delta u|$  ( $F_{(3,35)} = 0.281$ ,  $p = 0.839$ , one-way ANOVA on reward condition). Notably, the three groups had similar performance during the highly rewarded condition for both measures of variability (one-way ANOVA on 100% reward condition alone,  $F_{(2,27)} = 2.171$ ,  $p = 0.135$  for SD of signed change,  $F_{(2,27)} = 0.868$ ,  $p = 0.432$  for mean absolute change). These results indicated that the PD patients were not impaired in responding to successful trials,

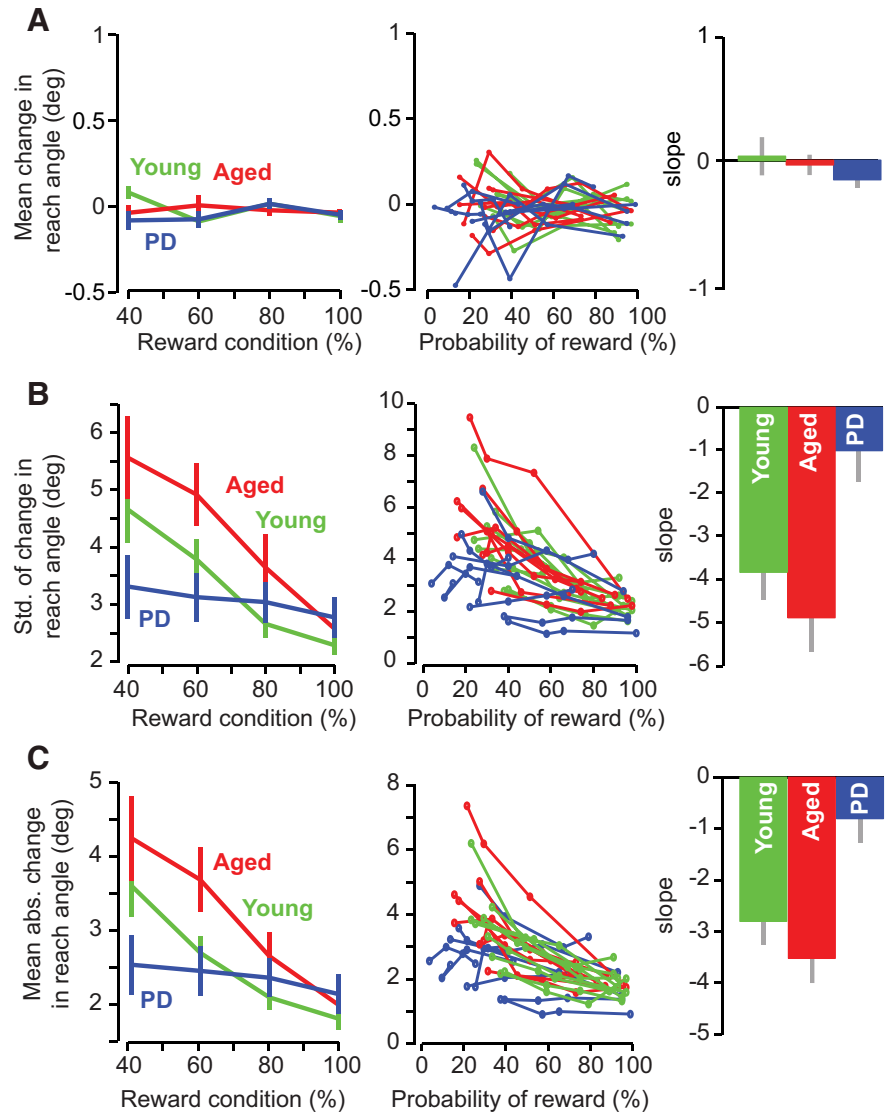
but instead did not adjust their level of variability in response to unsuccessful trials.

The left column in Figure 4 displays the reward probabilities that were experimentally imposed. However, this probability does not necessarily equal the probability of reward that the subjects actually experienced during the experiment. To examine this question, in the middle column of Figure 4, we plotted the changes in our behavioral measures over the actual reward probability achieved. To compare the relationship of these measures against the probability of reward, for each subject, we applied a linear regression and estimated the slope (right column, Fig. 4). We found that the PD group exhibited a significantly smaller slope than the two control groups for both measures of variability ( $\Delta u$ , *post hoc*  $p = 0.025$  vs young, and  $p = 0.003$  vs aged, after a significant effect of group, one-way ANOVA,  $p = 0.003$ ,  $F_{(2,25)} = 7.606$ ;  $|\Delta u|$  *post hoc*  $p = 0.031$  vs young, and  $p = 0.003$  vs aged, after a significant effect of group, one-way ANOVA,  $p = 0.003$ ,  $F_{(2,25)} = 7.686$ ). There was no difference in slope across the two control groups for either measure ( $p = 0.575$  age-matched vs young controls, SD of signed change;  $p = 0.502$ , mean absolute change). In addition, there was no differences across groups for the mean change in reach angle during this experimental session (one-way ANOVA,  $F_{(2,25)} = 2.062$ ,  $p = 0.150$ ).

In summary, in the two control groups, the probability of reward significantly modulated the change in motor output: as reward probability decreased, the trial-to-trial change in reach angles increased. In contrast, in PD patients, the probability of reward was not a significant modulator of variability.

**Sensitivity to history of reward**

To describe the relationship between changes in motor variability and reward mathematically, we extended our trial-to-trial analysis to include the history of reward. In Figure 5A, we have plotted  $|\Delta u|$  as a function of the history of reward (for this analysis, we have included data from all trials, 1–900, from both parts of the experiment). To describe the history of past rewards, we considered all possible combinations of successful and unsuccessful feedback for three consecutive trials. This history of reward for three consecutive trials was represented by variables  $R(n)$ ,  $R(n - 1)$ , and  $R(n - 2)$ , indicating whether the subject was successful in trials  $n$ ,  $n - 1$ , and  $n - 2$ , respectively. In Figure 5A, history of reward is plotted as a binary vector, ordered from left to right along the  $x$ -axis. All combinations are considered, starting from three consecutive successful trials [ $R(n) = 1$ ,  $R(n - 1) = 1$ , and  $R(n - 2) = 1$  on the left, to three consecutive unsuccessful trials ( $R(n) = 0$ ,  $R(n - 1) = 0$ ,  $R(n - 2) = 0$  on the right)]. We found that across groups,  $|\Delta u|$  was largest when the last three trials had



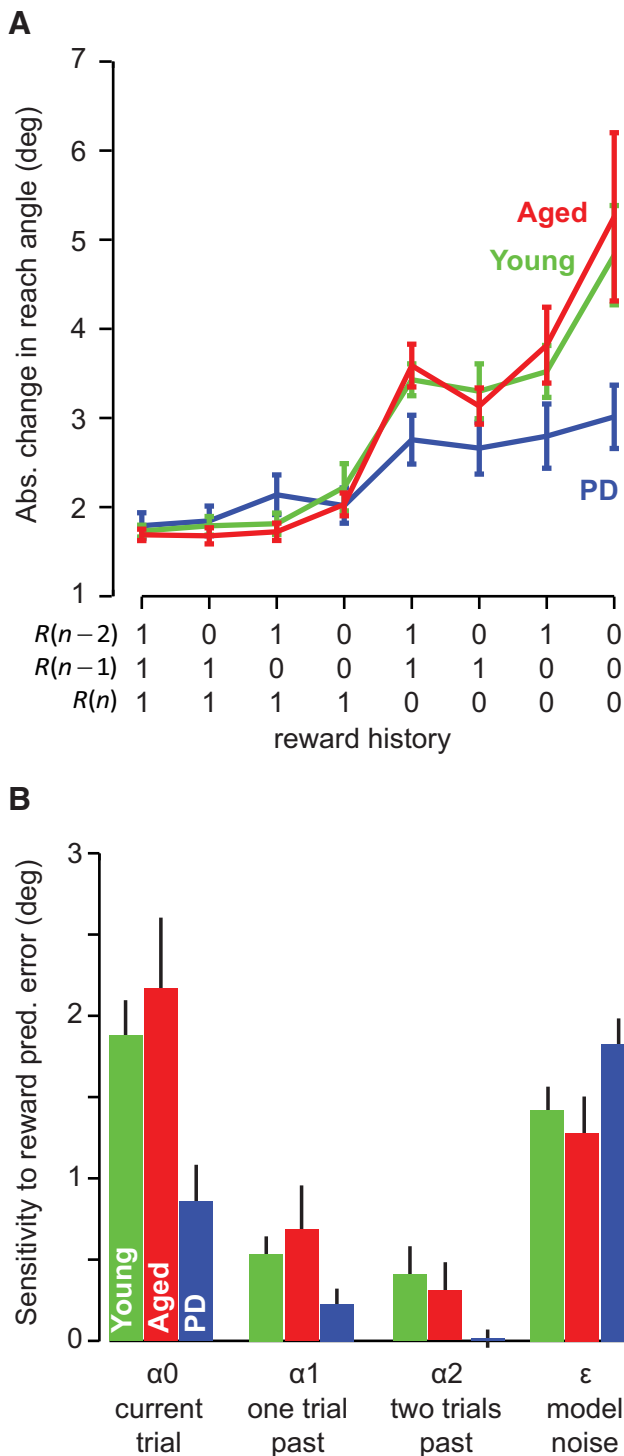
**Figure 4.** Reward-based modulation of reach variability is impaired in PD. The first column of each row displays group data indicating the average of each behavioral measure within each reward condition. Error bars indicate SEM. The second column of each row plots the individual subject data indicating the average of each performance measure against the actual probability of reward experienced during each condition. The third column of each row plots slope of the relationship between performance versus probability of reward fitted to each individual subject. Error bars indicate SEM. **A**, Mean of signed change in reach angle. **B**, SD of signed change in reach angle. **C**, Mean absolute change in reach angle.

been unsuccessful. A repeated-measures ANOVA with reward as the within-subject measure and group as the between-subject measure produced a significant effect of reward history on  $|\Delta u|$  ( $p < 0.001$ ,  $F_{(7,161)} = 45.073$ ), as well as an interaction between group and reward ( $p < 0.001$ ,  $F_{(14,161)} = 3.453$ ), suggesting that sensitivity to unsuccessful trials was smaller in PD patients compared with the other two groups.

To quantify this relationship across each participant, we formulated a state-space model to relate  $|\Delta u|$  to the history of the past rewards as follows:

$$|\Delta u(n)| = \alpha_0(1 - R(n)) + \alpha_1(1 - R(n - 1)) + \alpha_2(1 - R(n - 2)) + \varepsilon \quad (3)$$

In the above equation, the change in reach angle on trial  $n$  is written as a function of reward history in the last three trials. The term  $\alpha_0$  represents sensitivity to failure in the current trial and the



**Figure 5.** Sensitivity to positive and negative outcomes. **A**, Group data for the absolute change in reach angle between trials  $n$  and  $n + 1$  as a function of the reward history for the three most recent trials [ $R(n)$ ,  $R(n - 1)$ ,  $R(n - 2)$ ]. Error bars indicate SEM. **B**, A state-space model was used to determine each subject's change in reach angle as a function of the past history of reward. Here,  $\alpha_0$  represents sensitivity to reward prediction error in the current trial,  $\alpha_1$  represents sensitivity to reward prediction error in trial  $n - 1$ , and  $\epsilon$  is the variability in the movements that is not accounted for by reward prediction errors. Data indicate group average. Error bars indicate SEM.

term  $\alpha_1$  represents sensitivity to failure in the previous trial. The term  $\epsilon$  is the variability that cannot be explained by the recent history of failures. A large  $\alpha_0$  indicates that, after an unsuccessful reach on the current trial, there is a large change in reach angle.

We fitted the above equation to the data from each participant and plotted the parameter values in Figure 5B. We found that, across all groups, sensitivity was largest to failure in the current trial and then this sensitivity declined with trial history, as follows:  $\alpha_0 > \alpha_1 > \alpha_2$ . A repeated-measures ANOVA with sensitivity as the within-subject measure and group as the between-subject measure found a significant effect of sensitivity ( $F_{(2,46)} = 90.222$ ,  $p < 0.001$ ) and a significant group by sensitivity interaction ( $F_{(4,46)} = 4.363$ ,  $p = 0.004$ ). Sensitivity to failure in the current trial  $\alpha_0$  was significantly smaller in PD patients than in the two control groups ( $p = 0.032$  for young,  $p = 0.014$  for aged). There were no differences between this sensitivity value for the control groups ( $p = 0.657$ ).

Although the PD patients did respond to negative reward prediction error, as evidenced by a nonzero  $\alpha_0$  value ( $z$ -test,  $p < 0.001$ ), they appeared to be less affected by this feedback and did not adjust their motor commands to the same extent as the controls. PD patients were also less affected by reward prediction errors from trials further in the past. The  $\alpha$  values for these participants quickly decreased to the point where the  $\alpha_2$  was not significantly different from zero ( $z$ -test,  $p = 0.407$ ).

Once exploration successfully leads to reward, the best strategy is to maintain this performance. This behavior is captured by the  $\epsilon$  value, which determines the change in reach angle after a series of rewarded trials. The PD patients had a similar  $\epsilon$  value as the control groups (no effect of group on a one-way ANOVA,  $F_{(2,25)} = 2.642$ ,  $p = 0.093$ ), indicating that these participants had the same amount of variability after a rewarded trial. This fact is further supported by the similar trial-to-trial changes that were observed across groups during the initial baseline blocks and in the 100% reward condition, in which many trials were rewarded and motor exploration was unnecessary.

In summary, trial-to-trial change in motor output was partially driven by the history of reward. These changes were most sensitive to reward prediction error in the current trial and had smaller sensitivity to prediction errors in the previous trials. In PD patients, reward induced variability that was similar to the control group. However, if the trial was not rewarded, the resulting variability was abnormally small.

### Time between movements and other factors

Learning from error is affected by the time that passes between the trial in which the error is experienced and the after trial in which the change in motor output is assayed. For example, differences in intertrial intervals (ITIs) have been shown to alter adaptation rates (Bock et al., 2005; Francis, 2005; Huang and Shadmehr, 2007) as well as movement speed (Haith et al., 2012). Indeed, with passage of time between movements, the error-dependent change in the motor memory may decay (Yang and Lisberger, 2014). It is possible that a similar time-dependent decay process may modulate the influence of reward in one trial on the variability of movements in the subsequent trial. We therefore tested whether the patient and control populations experienced similar ITIs.

For each trial, we calculated the ITI between consecutive trials (duration of time between the end of trial  $n$ , and the start of trial  $n + 1$ ), the RT, and the MT. Across the experiment, both the age-matched controls and the patients exhibited occasional trials with a very long RT or ITI, which produced asymmetric distributions. We computed a robust mean of the data for each subject by fitting the data of each participant

with an exponentially modified Gaussian function (termed ex-Gaussian; Izawa et al., 2012b) as follows:

$$f(x | \mu, \sigma, \lambda) = \frac{\lambda}{2} \exp\left(\frac{\lambda}{2}(2\mu + \lambda\sigma^2 - 2x)\right) \times \left(1 - \operatorname{erf}\left(\frac{\mu + \lambda\sigma^2 - x}{\sqrt{2}\sigma}\right)\right) \quad (4)$$

In the above expression,  $\mu$  is the robust mean and  $\sigma$  is the robust SD of the normally distributed component of the data, and the parameter  $\lambda$  describes the exponentially distributed component, accounting for the positive tail that arises from the few, uncharacteristically long durations. To compare various groups, we represented the data for each subject via the robust mean  $\mu$  of their data.

We observed an ITI of  $\sim 2.2$  s for each group ( $2.18 \pm 0.10$  s for young controls,  $2.19 \pm 0.16$  s for aged, and  $2.22 \pm 0.20$  s for PD, mean  $\pm$  SD) and noted no statistically significant differences among the groups ( $F_{(2,25)} = 0.182$ ,  $p = 0.835$ , one-way ANOVA for average ITI across all 900 trials). This demonstrated that the time between trials was not different between groups.

However, we did find a significant effect of group on RT and MT data ( $F_{(2,25)} = 4.718$ ,  $p = 0.019$  and  $F_{(2,25)} = 12.024$ ,  $p < 0.001$ , one-way ANOVA for average RT and MT across all 900 trials). Here, the PD group exhibited a slightly longer RT ( $0.28 \pm 0.02$  s, mean  $\pm$  SD) and MT ( $0.32 \pm 0.06$  s) than the young controls (RT:  $0.24 \pm 0.03$  s, MT:  $0.21 \pm 0.02$  s), which was a statistically significant difference ( $p = 0.015$  and  $p < 0.001$  for RT and MT, *post hoc* analysis). Crucially, we did not observe any significant differences between the age-matched controls (RT:  $0.25 \pm 0.02$  s, MT:  $0.27 \pm 0.05$  s) and the PD patients in measures of RT and MT ( $p = 0.212$  and  $p = 0.094$ , *post hoc* for RT and MT, respectively).

## Discussion

We examined the hypothesis that variability in motor commands is partly driven by the history of reward. Our hypothesis made two predictions. First, trial-to-trial change in motor commands should reflect the success or failure of each movement. Subjects should stay with their current motor command if the trial was successful, but change if it was not. As a result, the search for a rewarding outcome should lead to large performance variability during periods of low reward probability, but low variability during periods of high reward probability. Second, reward-dependent motor variability should be dependent on the integrity of the basal ganglia. As a result, if the ability of the brain to encode reward prediction error is compromised, then the trial-to-trial change in motor commands in response to failure will also be affected. We tested our ideas in three groups of people: young controls, age-matched controls, and PD patients. Because aging is associated with decreases in dopamine neurons (Fearnley and Lees, 1991; Vaillancourt et al., 2012) and impaired physiological representations of reward processing (Schott et al., 2007; Chowdhury et al., 2013), we suspected that the process of aging may contribute to deficits in control of reward-dependent motor variability.

We considered a reaching task in which the only performance feedback was success or failure of the trial. We observed that, during periods of low reward probability, both young and age-matched controls exhibited large intertrial changes in reach angle, but these changes were smaller in PD. To estimate the relationship between change in reach angle and history of success,

we considered a state-space model in which the change in motor command from one trial to the next was driven by the recent history of reward. (A similar model was used to represent trial-to-trial change in dopamine as a function of the history of reward; Bayer and Glimcher, 2005.) We found that the control participants were highly sensitive to failure in the most recent trial (indicated by a high  $\alpha_0$  value), changing their reach angles in response to an unrewarded outcome. Patients with PD also showed the greatest dependence on the outcome of the most recent trial, but exhibited smaller than normal sensitivity to failure. Compared with controls, the PD patients had similar levels of variability during periods of high reward probability, but a much smaller change in trial-to-trial reach angle after unrewarded trials. Our results consistently indicated that PD patients were impaired at increasing their variability, but only in response to unrewarding outcomes. However, we did not observe statistically significant differences in these measures between the aged and young controls, indicating no major effect of aging in this task.

By our measures, the impairment in the PD patients was a reduced ability to increase their variability to search the task space and achieve reward. It is thought that one source of motor variability arises during execution due to the noise in the peripheral motor organs, including motor neurons (Jones et al., 2002). However, variability exists in neural activity during motor planning in the premotor cortex, which also contributes to variability in movement execution (Churchland et al., 2006). We think it is reasonable that the motor variability observed in a typical motor control experiment is composed of motor noise arising from both the peripheral and CNS. Importantly, the baseline motor variability of PD was comparable to that of healthy controls, suggesting that the motor noise in the peripheral motor organs was not influenced by PD; rather, it appeared that PD specifically reduced the ability to modulate motor planning noise in response to negative reward prediction error.

In our task, the PD patients were less likely to increase their performance variability after a negative reward prediction error, indicating a reluctance to switch from their current action despite the lack of success. This result is similar to other studies finding that PD patients may settle on a less than optimal solution to complete a task (Shohamy et al., 2004; Vakil et al., 2014), which in this instance is persistence instead of exploration. A major limitation of our study however, was that the patients in our PD group were only tested while on their usual schedule of medication. Therefore, our results cannot be attributed to the disease state alone. Because PD patients experience significant discomfort during periods of withdrawal from their medication schedule, we chose to not collect data from an experimental group of PD patients in the off-medication state. In addition, PD is characterized by bradykinesia, a slowing of movements, making it possible that, in the off-medication state, they may not have been able to perform the motor components of this task adequately.

Due to this limitation in our study design, we cannot separate the contributions of disease state from the effects of medication on response to reward. Several studies suggest that, when the patients are on medication, they are specifically impaired at learning from negative reward prediction errors (Frank et al., 2004; Frank et al., 2007; Bódi et al., 2009). Although deficits have been reported in many associative learning tasks, this has not proven to be ubiquitous. The heterogeneity of results is believed to be due to differences in task demands, clinical severity of the disease, and, importantly, the presence or absence of medication (Cools et al., 2001; Frank et al., 2004; Shohamy et al., 2006; Frank



et al., 2007; Bódi et al., 2009; Rutledge et al., 2009). Even in healthy controls, dopaminergic agents have been found to alter the ability to learn during feedback-based tasks (Pessiglione et al., 2006; Pizzagalli et al., 2008). The administration of L-DOPA to healthy older adults has been found to improve performance in a reward-based task in some participants, but impair performance in others, depending on baseline levels of dopamine (Chowdhury et al., 2013). Even the selection of low-level parameters of movement, such as the velocity, acceleration, and latency of an action, are thought to be affected by dopamine (Mazzoni et al., 2007; Niv et al., 2007; Shadmehr et al., 2010; Galea et al., 2013). With such an interplay between behavior and dopamine levels, we cannot rule out the possibility that the deficits observed in the PD group after an unrewarded trial in our study are not simply due to the dopaminergic medication.

Despite impairments in learning from reward prediction errors, PD patients exhibit normal behavior in many motor learning tasks. PD patients are able to perform comparably to controls on motor skill and mirror inversion tasks while on medication (Agostino et al., 1996; Paquet et al., 2008). Several studies have shown that PD patients adapt to visuomotor rotations as well as control participants, although consolidation of this learning is impaired in those with the disease (Marinelli et al., 2009; Bédard and Sanes, 2011; Leow et al., 2012; Leow et al., 2013). This intact performance during motor adaptation in PD is presumably due to their ability to recruit learning processes that do not depend on reward prediction errors, such as learning from sensory prediction errors which may depend on the cerebellum (Izawa et al., 2012a). Indeed, we previously found distinct signatures of learning from sensory versus reward prediction errors and suggested that the basal ganglia structures were responsible for altering movements in response to reward (Izawa and Shadmehr, 2011). Motor adaptation may employ several learning processes, including control of error sensitivity (Herzfeld et al., 2014) and reinforcement of action through reward (Huang et al., 2011). Consolidation may depend on some or all of these processes. The question of motor memory consolidation and its possible dependence on function of the basal ganglia remains to be more fully explored.

In conclusion, we found that trial-to-trial change in motor commands was driven by the history of past rewards. Control of variability was most sensitive to a recent unrewarded trial, resulting in increased variability that explored the task space for a more rewarding solution. Therefore, during periods of low reward probability, the healthy brain increased motor variability. This control of variability appeared to depend on the integrity of the basal ganglia because PD patients on medication exhibited impaired sensitivity to a negative outcome, but normal sensitivity to a positive outcome.

Our results may be combined with recent ideas about computational processes in motor learning to suggest that two forms of prediction error drive trial-to-trial change in motor commands: sensory prediction errors alter the mean of the motor commands (Marko et al., 2012), whereas reward prediction errors alter the variance of the commands. The first form of learning appears to depend on the integrity of the cerebellum and the second appears to depend on the integrity of the basal ganglia.

## References

- Agostino R, Sanes JN, Hallett M (1996) Motor skill learning in Parkinson's disease. *J Neurol Sci* 139:218–226. [CrossRef Medline](#)
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141. [CrossRef Medline](#)
- Bédard P, Sanes JN (2011) Basal ganglia-dependent processes in recalling learned visual-motor adaptations. *Exp Brain Res* 209:385–393. [CrossRef Medline](#)
- Bock O, Thomas M, Grigorova V (2005) The effect of rest breaks on human sensorimotor adaptation. *Exp Brain Res* 163:258–260. [CrossRef Medline](#)
- Bódi N, Kéri S, Nagy H, Moustafa A, Myers CE, Daw N, Dibó G, Takáts A, Bereczki D, Gluck MA (2009) Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain* 132:2385–2395. [CrossRef Medline](#)
- Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ (2013) Dopamine restores reward prediction errors in old age. *Nat Neurosci* 16:648–653. [CrossRef Medline](#)
- Churchland MM, Afshar A, Shenoy KV (2006) A central source of movement variability. *Neuron* 52:1085–1096. [CrossRef Medline](#)
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11:1136–1143. [CrossRef Medline](#)
- Fearnley JM, Lees AJ (1991) Ageing and Parkinson's disease: substantia nigra regional selectivity. *Brain* 114:2283–2301. [CrossRef Medline](#)
- Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 12:189–198. [CrossRef Medline](#)
- Francis JT (2005) Influence of the inter-reach-interval on motor learning. *Exp Brain Res* 167:128–131. [CrossRef Medline](#)
- Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943. [CrossRef Medline](#)
- Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* 318:1309–1312. [CrossRef Medline](#)
- Galea JM, Ruge D, Buijink A, Bestmann S, Rothwell JC (2013) Punishment-induced behavioral and neurophysiological variability reveals dopamine-dependent selection of kinematic movement parameters. *J Neurosci* 33:3981–3988. [CrossRef Medline](#)
- Gordon J, Ghilardi MF, Cooper SE, Ghez C (1994) Accuracy of planar reaching movements. II. Systematic extent errors resulting from inertial anisotropy. *Exp Brain Res* 99:112–130. [CrossRef Medline](#)
- Haith AM, Reppert TR, Shadmehr R (2012) Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32:11727–11736. [CrossRef Medline](#)
- Herzfeld DJ, Vaswani PA, Marko MK, Shadmehr R (2014) A memory of errors in sensorimotor learning. *Science* 345:1349–1353. [CrossRef Medline](#)
- Huang VS, Shadmehr R (2007) Evolution of motor memory during the seconds after observation of motor error. *J Neurophysiol* 97:3976–3985. [CrossRef Medline](#)
- Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* 70:787–801. [CrossRef Medline](#)
- Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol* 7:e1002012. [CrossRef Medline](#)
- Izawa J, Criscimagna-Hemminger SE, Shadmehr R (2012a) Cerebellar contributions to reach adaptation and learning sensory consequences of action. *J Neurosci* 32:4230–4239. [CrossRef Medline](#)
- Izawa J, Pekny SE, Marko MK, Haswell CC, Shadmehr R, Mostofsky SH (2012b) Motor learning relies on integrated sensory inputs in ADHD, but over-selectively on proprioception in autism spectrum conditions. *Autism Res* 5:124–136. [CrossRef Medline](#)
- Jones KE, Hamilton AF, Wolpert DM (2002) Sources of signal-dependent noise during isometric force production. *J Neurophysiol* 88:1533–1544. [CrossRef Medline](#)
- Kao MH, Doupe AJ, Brainard MS (2005) Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* 433:638–643. [CrossRef Medline](#)
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273:1399–1402. [CrossRef Medline](#)
- Leow LA, Loftus AM, Hammond GR (2012) Impaired savings despite intact

- initial learning of motor adaptation in Parkinson's disease. *Exp Brain Res* 218:295–304. [CrossRef Medline](#)
- Leow LA, de Rugy A, Loftus AM, Hammond G (2013) Different mechanisms contributing to savings and anterograde interference are impaired in Parkinson's disease. *Front Hum Neurosci* 7:55. [Medline](#)
- Marinelli L, Crupi D, Di Rocco A, Bove M, Eidelberg D, Abbruzzese G, Ghilardi MF (2009) Learning and consolidation of visuo-motor adaptation in Parkinson's disease. *Parkinsonism Relat Disord* 15:6–11. [CrossRef Medline](#)
- Marko MK, Haith AM, Harran MD, Shadmehr R (2012) Sensitivity to prediction error in reach adaptation. *J Neurophysiol* 108:1752–1763. [CrossRef Medline](#)
- Mazzoni P, Hristova A, Krakauer JW (2007) Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J Neurosci* 27:7105–7116. [CrossRef Medline](#)
- Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease (2003) The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Movement Disorders* 18:738–750. [CrossRef Medline](#)
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520. [CrossRef Medline](#)
- Olveczky BP, Andalman AS, Fee MS (2005) Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol* 3:e153. [CrossRef Medline](#)
- Page S, Neuringer A (1985) Variability is an operant. *J Exp Psych Anim Behav Proc* 11:429–452.
- Paquet F, Bédard MA, Levesque M, Tremblay PL, Lemay M, Blanchet PJ, Scherzer P, Chouinard S, Filion J (2008) Sensorimotor adaptation in Parkinson's disease: evidence for a dopamine dependent remapping disturbance. *Exp Brain Res* 185:227–236. [CrossRef Medline](#)
- Pesek-Cotton EF, Johnson JE, Newland MC (2011) Reinforcing behavioral variability: an analysis of dopamine-receptor subtypes and intermittent reinforcement. *Pharmacol Biochem Behav* 97:551–559. [CrossRef Medline](#)
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045. [CrossRef Medline](#)
- Pizzagalli DA, Evins AE, Schetter EC, Frank MJ, Pajtas PE, Santesso DL, Culhane M (2008) Single dose of a dopamine agonist impairs reinforcement learning in humans: behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology (Berl)* 196:221–232. [CrossRef Medline](#)
- Rutledge RB, Lazzaro SC, Lau B, Myers CE, Gluck MA, Glimcher PW (2009) Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J Neurosci* 29:15104–15114. [CrossRef Medline](#)
- Schott BH, Niehaus L, Wittmann BC, Schütze H, Seidenbecher CI, Heinz HJ, Düzél E (2007) Ageing and early-stage Parkinson's disease affect separable neural mechanisms of mesolimbic reward processing. *Brain* 130:2412–2424. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY (2010) Temporal discounting of reward and the cost of time in motor control. *J Neurosci* 30:10507–10516. [CrossRef Medline](#)
- Shohamy D, Myers CE, Onlaor S, Gluck MA (2004) Role of the basal ganglia in category learning: how do patients with Parkinson's disease learn? *Behav Neurosci* 118:676–686. [CrossRef Medline](#)
- Shohamy D, Myers CE, Gekhman KD, Sage J, Gluck MA (2006) L-dopa impairs learning, but spares generalization, in Parkinson's disease. *Neuropsychologia* 44:774–784. [CrossRef Medline](#)
- Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat Neurosci* 15:1281–1289. [CrossRef Medline](#)
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O (2002) Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* 142:284–291. [CrossRef Medline](#)
- Tumer EC, Brainard MS (2007) Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. *Nature* 450:1240–1244. [CrossRef Medline](#)
- Vaillancourt DE, Spraker MB, Prodoehl J, Zhou XJ, Little DM (2012) Effects of aging on the ventral and dorsal substantia nigra using diffusion tensor imaging. *Neurobiol Aging* 33:35–42. [CrossRef Medline](#)
- Vakil E, Hassin-Baer S, Karni A (2014) A deficit in optimizing task solution but robust and well-retained speed and accuracy gains in complex skill acquisition in Parkinson's disease: multi-session training on the Tower of Hanoi puzzle. *Neuropsychologia* 57:12–19. [CrossRef Medline](#)
- Wu HG, Miyamoto YR, Gonzalez Castro LN, Ölveczky BP, Smith MA (2014) Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat Neurosci* 17:312–321. [CrossRef Medline](#)
- Yang Y, Lisberger SG (2014) Role of plasticity at different sites across the time course of cerebellar motor learning. *J Neurosci* 34:7077–7090. [CrossRef Medline](#)