Categorization system–switching deficits in typical aging and Parkinson’s disease

Sébastien Hélie
Department of Psychological Sciences, Purdue University

Madison Fansher
Department of Psychological Sciences, Purdue University

Objective: Numerous studies documenting cognitive deficits in Parkinson’s disease (PD) revealed impairment in a variety of tasks related to memory, learning, and attention. One ubiquitous task that has not received much attention, is categorization system–switching. Categorization system–switching is a form of task–switching requiring participants to switch between different categorization systems. In this article, we explore whether older adults and people with PD show deficits in categorization system–switching.

Method: Twenty older adults diagnosed with PD, 20 neurologically intact older adults, and 67 young adults participated in this study. Participants were first trained in rule–based (RB) and later information–integration (II) categorization separately. After training on the tasks, participants performed a block of trial–by–trial switching where the RB and II trials randomly intermixed. Finally, the last block of trials also intermixed RB and II trials randomly but additionally changed the location of the response buttons.

Results: Contrary to our hypothesis, the results show no difference in accuracy between older adults and people with PD during the intermixed trial block, as well as no difference in response time (RT) switch cost. However, both groups were less accurate during intermixed trial blocks and had a higher RT switch cost when compared to young adults. In addition, the proportion of participants able to switch systems was smaller in people with PD than in young adults.

Conclusions: The results suggest that older adults and people with PD have impaired categorization system–switching ability, and that this ability may be related to a decrease in tonic dopamine (DA) levels associated with normal aging and PD.

Public significance statement: From categorizing objects as edible or inedible, to categorizing people as friends or enemies, everyday life is filled with thousands of category decisions. Hence, understanding how categorization system–switching is affected by aging and PD is critical and may point to intervention targets to improve quality of life.

Keywords: Parkinson’s disease, aging, system–switching, perceptual categorization

More than 10,000,000 people are affected by PD worldwide and 60,000 more Americans are diagnosed with PD each year.¹ Parkinsonian motor symptoms include tremor, rigidity, bradykinesia, and akinesia. In addition to motor deficits, non–demented PD patients present cognitive symptoms that resemble those observed in patients with frontal damage (Zgaljardic, Foldi, & Borod, 2004). Specifically, numerous studies documenting the cognitive deficits of PD patients have revealed impairment in a variety of tasks related to memory, learning, visuospatial skills, and attention (Gotham, Brown, & Marsden, 1988; Zgaljardic et al., 2006). Many of these cognitive functions are critical for assigning everyday objects to categories (Price, Filoteo, & Maddox, 2009).

Everyday life is filled with thousands of category decisions. Over the past 20 years, mounting evidence has been gathered that category learning is achieved using a number of different psychological and biological systems (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Valentin, 2017; Erickson & Kruschke, 1998; Hélie & Cousineau, 2015; Hélie, Roeder, & Ashby, 2010; Nosofsky, Palmeri, & McKinley, 1994; Waldschmidt & Ashby, 2011). While existing multiple–systems theories and models of categorization sometimes disagree about the number or nature of the different systems, all assume that people can switch between systems seamlessly depending on the task at hand. However, one ubiquitous task that has not received much attention is categorization system–switching (Ashby & Crossley, 2010; Crossley, Roeder, Hélie, & Ashby, in press; Erickson, 2008): Can people flexibly switch between the different categorization systems on a trial–by–trial basis? For example, imagine you are in the basement of a building completing your laundry and a fire alarm sounds. How difficult is it to switch from a declarative task (e.g., sorting clothes by color), to a procedural task (e.g., climbing stairs)? Past research has shown that tonic DA levels are associ-

¹Parkinson’s Disease Foundation, retrieved on 07/24/2017: http://www.pdf.org/parkinson_statistics
ated with cognitive flexibility (Price et al., 2009), and DA–producing neurons die at a rate of 5%–10% per decade of life with normal aging (Karrer, Josef, Mata, Morris, & Samanez-Larkin, 2017). Furthermore, PD is caused by the accelerated death of DA–producing neurons, and PD motor symptoms appear after DA–levels have been reduced by 70%–80% in the striatum (Hélie, Paul, & Ashby, 2012a). Hence, it is critical to understand how categorization system–switching is affected by aging and PD. In this article, we explore whether older adults and people with PD show deficits in categorization system–switching.

**Aging, Parkinson’s disease, and categorization**

Category learning tasks can generally be separated into declarative (e.g., RB) and procedural (e.g., II) learning tasks (Ashby et al., 1998). Declarative category–learning tasks are those in which the category structures can be learned via some explicit reasoning process. Frequently, the rule that maximizes accuracy (i.e., the optimal rule) is easy to describe verbally. In the most common applications, only one stimulus dimension is relevant, and the participant’s task is to discover this relevant dimension and then map the different dimensional values to the relevant categories (Hélie, Ell, Filoteo, & Maddox, 2015). For example, sorting the laundry into piles of different color would be a declarative categorization task. Peoples with PD display many of the same deficits in rule–learning tasks as patients with frontal lobe damage (Owen et al., 1993). These tasks demand attention, working memory, and logical reasoning to maximize performance. Ashby, Noble, Filoteo, Waldron, and Ell (2003) tested PD patients, aged–matched controls (OC), and young controls (YC) in a RB categorization task similar to the Wisconsin Card Sorting Test (WCST) (Heaton, Chelune, Talley, Kay, & Curtiss, 1993), except that the stimuli varied on four dimensions instead of three. Like the WCST, however, a simple one–dimensional rule could be used to categorize the stimuli perfectly. Each participant was classified as a learner if s/he could produce 10 consecutively correct responses within 200 trials. Significantly more PD patients failed to learn in this task than both the YC and OC, showing an impairment in categorization rule–learning for people with PD.

The other type of categorization tasks involve procedural categorization. These tasks are those in which accuracy is maximized only if information from two or more stimulus components (or dimensions) is integrated at some pre–decisional stage. Perceptual integration can take many forms – from treating the stimulus as a Gestalt to computing a weighted linear combination of the dimensional values. Typically, the optimal strategy in procedural tasks is difficult or impossible to describe verbally, and the categorical knowledge is specific to the motor response (Ashby, Ell, & Waldron, 2003). For example, selecting the appropriate golf swing as a function of all the available environmental information (e.g., roughness of the area where the ball lie, wind speed, terrain incline, etc.) is a procedural categorization task. Declarative strategies can be applied in procedural categorization tasks, but they generally lead to sub–optimal levels of performance because declarative strategies make separate decisions about each stimulus component, rather than integrating this information. Ashby, Noble, et al. (2003) tested PD patients with an II category–learning task that used the same stimuli as in the RB task described above. Towards this end, the stimuli were separated into two categories in such a way that no easily verbalized rule would yield optimal performance (an example procedural categorization task). Interestingly, PD patients were unimpaired in this task compared to OC, although both groups were massively impaired relative to YC. Similarly, PD patients showed no deficits compared to aged–matched controls in two other II category–learning tasks that used two–dimensional continuous–valued stimuli when the categories were linearly separable, although they were impaired relative to aged–matched controls when the categories were nonlinearly separable (Filoteo, Maddox, Salmon, & Song, 2005; Maddox & Filoteo, 2001). These results suggest that PD patients are impaired relative to OC in tasks that rely on procedural learning, but only when the task is sufficiently complex.

**Categorization system–switching**

Many theories and models of category learning assume that declarative and procedural categorization are learned by functionally (Erickson & Kruschke, 1998; Nosofsky et al., 1994) and anatomically (Ashby et al., 1998) distinct systems. However, it took over a decade after the proposals of multiple–systems theories of categorization for the first empirical investigation of system–switching to be performed (Erickson, 2008). Erickson asked undergraduate students to categorize “space shuttle” schematics into one of four categories. Two of the categories could be distinguished using a simple verbal rule (i.e., RB categories) while the other two categories could not (i.e., II categories). Decision–bound models (Ashby, 1992; Maddox & Ashby, 1993; Hélie, Turner, Crossley, Ell, & Ashby, 2003) were individually fit to the RB and II data to identify ‘switchers’ and ‘non–switchers’. Switchers were participants whose RB and II data were best fit by optimal models (i.e., a decision bound on one of the stimulus dimensions for RB trials and the general linear classifier for the II trials). All other participants were labeled as non–switchers. Only 37% of the participants were able to switch categorization systems on a trial–by–trial basis (Erickson, 2008). With continuous–dimension stimuli, Crossley et al. (in press) obtained a proportion of undergraduate student switchers of about 40%. The highest proportion of undergraduate student switchers observed so far is 65.7% (Hélie, 2017). This higher proportion of
switchers required 2 sessions of training and preparation time for system–switching before the categorization stimulus is presented. In all cases, system–switching appears to be extremely difficult for healthy young adults.

While categorization system–switching has yet to be studied with older adults and people with PD, older adults and people with PD have been tested in regular task–switching paradigms. Task–switching is a more general construct than system–switching in that the participants are asked to switch between tasks that may or may not rely on separate systems (Vandierendonck, Liefooghe, & Verbruggen, 2010). Crossley et al. (in press) showed that switch cost, a measure quantifying difficulty of trial–by–trial switching, is higher when switching between systems than when switching within a system. Older adults typically have a higher switch cost, meaning that they tend to be slower and less accurate on trials in which the task differs from the previous trial (Kray & Lindenberger, 2000). In addition to higher switch costs, people with PD struggle with self–initiated task–switching (e.g., when the task sequence is predictable of the task but there is no environmental cue) but there is no evidence of impairment when the task–switch is triggered by an external cue (Werheid, Koch, Reichert, & Brass, 2007), as in typical system–switching experiments.

The current experiment

The current experiment aimed at testing the effects of aging and PD on categorization system–switching. Young adults, older adults, and people with PD were first trained in RB categorization and II categorization separately to learn the categories (blocked trials). In order to ensure that any observed deficits in system–switching were caused by switch difficulty, and not by impaired learning in the individual categorization tasks, the RB and II categorization tasks were designed to be easy enough so that group differences in blocked trials were not expected. After training on the tasks separately, participants then performed a block of trial–by–trial switching where the RB and II trials were randomly intermixed. This block was used to compute the switch cost. Finally, the last block of trials also intermixed RB and II trials randomly, but in addition the location of the response buttons were changed. This manipulation has been shown to affect II trials more than RB trials (Ashby, Ell, & Waldron, 2003; Crossley et al., in press), but it is unclear whether all groups would be equally affected. This last manipulation was more exploratory, as we did not have any strong hypothesis before running the experiment.

Given the observed deficits associated with aging and PD in task–switching (i.e., higher switch costs), we hypothesized that older adults and people with PD would be impaired in categorization system–switching, and that this impairment may be worse for people with PD. Specifically, a higher switch cost should translate into a smaller proportion of participants that can switch on a trial–by–trial basis, which should reduce the overall accuracy in intermixed trial blocks. To anticipate, the results show that, as predicted, older adults and people with PD were less accurate during intermixed trial blocks and had a higher RT switch cost when compared to young adults (but did not differ from each other). In addition, the proportion of switchers was smaller in people with PD than in young adults. However, there was no difference between the groups in accuracy switch costs and in button–switch interference.

Method

Participants

Twenty older adults diagnosed with PD and 20 neurologically intact older adults were recruited from the Lafayette (IN) community and the surrounding area to participate in this experiment. In addition, 67 young adults were recruited from the Purdue University undergraduate population. OC and people with PD were given a $25 monetary compensation for their time, while YC received credits for participation as partial fulfillment of a course requirement. Participant demographic information is illustrated in Table 1. All procedures were approved by the Purdue University Human Research Program Institutional Review Board.

Before the beginning of the experiment, people with PD were asked if they had a diagnosis of Parkinson’s disease and no issues with their thinking in a pre–screening phase (self–report). The pre–screening form also asked that participants were not diagnosed with any additional neurological disease (including dementia), did not have significant cognitive impairment, and did not have any visual problems. Only participants who answered in the affirmative for all questions were included in the experiment. After completing the experiment, permission was obtained to contact each participant’s neurologist to confirm the presence of a diagnosis of Parkinson’s disease and the absence of dementia. Medical records were obtained from their treating neurologist to confirm the absence of other neurological conditions (including dementia), their medication, and record their Hoehn and Yahr (HY) scores. A HY score of I indicates only unilateral involvement while a HY score of V indicates confinement to a bed or wheelchair (Hoehn & Yahr, 1967). HY scores could not be collected for three participants due to various complications with medical record data. The HY distribution of the remaining 17 participants was as follows: I = 6, II = 8, and III = 3. No participant scored higher than III. All participants were tested “ON” medication and a list of medication is shown in Table 2.

Older adults were eligible to participate in the study if they had not been diagnosed with any neurological disease, did not report significant cognitive impairment, did not report any visual problem, and were not diagnosed with depression.
Table 1

Demographic information for the participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Male</th>
<th>Female</th>
<th>Age</th>
<th>White</th>
<th>Black/ African American</th>
<th>Hispanic/ Latino</th>
<th>Not Hispanic/ Latino</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>13</td>
<td>7</td>
<td>67.2 (9.5)</td>
<td>52.3%</td>
<td>47.7%</td>
<td>19.5 (1.5)</td>
<td>70.3%</td>
</tr>
<tr>
<td>OC</td>
<td>8</td>
<td>12</td>
<td>64.5 (9.8)</td>
<td>41.7%</td>
<td>58.3%</td>
<td>18 (1.5)</td>
<td>80%</td>
</tr>
<tr>
<td>YC</td>
<td>52.3%</td>
<td>47.7%</td>
<td>19.5 (1.5)</td>
<td>70.3%</td>
<td>4.3%</td>
<td>8.8%</td>
<td>Note. For PD and OC, entries represent the sample demographic data. The YC were randomly recruited from the Purdue University Department of Psychological Sciences participant pool. However, the sample demographic information was not collected for this group, so entries reflect the participant pool demographic information. Mixed race or did not provide the information. For Ethnicity, the data missing from the Table was Unknown.</td>
</tr>
</tbody>
</table>

Note. For PD and OC, entries represent the sample demographic data. The YC were randomly recruited from the Purdue University Department of Psychological Sciences participant pool. However, the sample demographic information was not collected for this group, so entries reflect the participant pool demographic information. Mixed race or did not provide the information. For Ethnicity, the data missing from the Table was Unknown.

Table 2

Medication distribution in the sample of PD participants

<table>
<thead>
<tr>
<th>Medication</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>L–dopa</td>
<td>13</td>
</tr>
<tr>
<td>MAO inhibitor</td>
<td>8</td>
</tr>
<tr>
<td>Anticonvulsant</td>
<td>6</td>
</tr>
<tr>
<td>Dopamine agonist</td>
<td>5</td>
</tr>
<tr>
<td>Antidepressant/Anxiolytic</td>
<td>4</td>
</tr>
<tr>
<td>Anticholinergic</td>
<td>3</td>
</tr>
<tr>
<td>Anticholinesterase</td>
<td>2</td>
</tr>
<tr>
<td>Muscle relaxant</td>
<td>2</td>
</tr>
<tr>
<td>COMT inhibitor</td>
<td>1</td>
</tr>
<tr>
<td>Dopamine reuptake inhibitor (DRI)</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. Medication was only available for 17 participants (see main text).

Material

The stimuli used were circular sine–wave gratings generated with the Matlab Psychophysics toolbox (Brainard, 1997) that occupied approximately 5° of visual angle. Stimuli were of constant contrast and size and were presented on a 21-inch monitor (1,920 × 1,080 resolution). Each stimulus was defined by a set of points in 2D space (frequency, orientation) where frequency was indicative of bar width and was calculated in cycles per degree (cpd), and orientation was the angle of counterclockwise rotation from horizontal calculated in radians. During each trial a single stimulus was presented in the center of the screen. Figure 1a shows an example stimulus.

The stimuli were generated into an arbitrary 200 × 100 coordinate system using the randomization technique of Ashby and Gott (1988). Stimuli were separated into four categories and Figure 1b shows the stimuli grouped into categories arbitrarily labeled from left–to–right with letters A–D. Category A and B structures are II and were generated using bivariate normal distributions: \( \mu_A = (42, 80) \), \( \Sigma_A = (145 135) \), \( \mu_B = (58, 64) \), \( \Sigma_B = \Sigma_A \). The category C and D structures are RB and were generated using bivariate normal distributions: \( \mu_C = (138, 72) \), \( \Sigma_C = (10 0) \), \( \mu_D = (162, 72) \), and \( \Sigma_D = \Sigma_C \). The stimulus arbitrary coordinate system was then re–scaled into a frequency x orientation space using a non–linear transformation (Hélie, 2017). This yielded stimuli ranging in frequency from 0.29 to 8.6 cpd and from 34° to 95° in orientation (counterclockwise from horizontal). II stimuli were presented on a green background while RB stimuli were presented on a blue background. Perfect accuracy was possible and optimal performance would require responding to the A–B stimuli using a procedural system and responding (self–report). The only inclusion criterion for YC was that they had not participated in a categorization experiment with similar stimuli previously.
to the C–D stimuli with a declarative system. The optimal rule for C/D categorization was a rule on bar width, with wider bars belonging to the C category and narrower bars belonging to the D category. Orientation was irrelevant for C/D categorization. No such verbalizable rule existed for A/B categorization because measures of orientation and bar width are non-commensurable and had to be integrated at a pre-decisional stage.

Stimulus presentation, feedback, and response recording were displayed and acquired using Matlab. The screen background color during the task indicated possible category choices to participants. A blue background was used when presenting RB stimuli (C–D) and a green background was used when presenting II stimuli (A–B). Participants in the YC group gave responses on a standard keyboard: the ‘s’ key was used for category A, the ‘d’ key was used for category B, the ‘k’ key was used for category C, and the ‘l’ key was used for category D. Response keys A and B were covered with blank green stickers indicating their use for II trials and response keys C and D were covered with blank blue stickers indicating their use for RB trials. The category labels (A–D) were displayed at the bottom of the screen below the stimulus mapping the response buttons on the keyboard.

Participants in the PD group and OC group used a button box with large buttons to record their responses. Participants made responses using a button box instead of a keyboard because tremors associated with PD could potentially interfere with participant responses on a standard keyboard. The button box contained four buttons. The two leftmost buttons were for selecting categories A or B and had a green mark above them to indicate that the buttons were to be selected during II trials. The two rightmost buttons were for selecting categories C or D and had a blue mark above them to indicate that the buttons were to be selected during RB trials. Note that the button box configuration matched that of the keyboard for the YC group.

After each response, participants were presented with auditory feedback: a high pitch tone for a correct response, a buzz sound for an incorrect response, and a two note sound for an incorrect key selection.

Procedure

The procedure was identical to the 1-session preparation (1S/PREP) condition in Hélie (2017). The experiment lasted about 55 minutes, and the session was divided into 7 blocks of 100 trials (for a total of 700 trials). Participants were first trained only in RB categorization (categories C/D) in Block 1. Participants were then trained only in II categorization (categories A/B) in Blocks 2–5. Participants were trained longer in II categorization due to previous work showing that II category structures are more difficult to acquire than RB (Hélie & Ashby, 2012), and also because successful acquisition of II category knowledge was multidimensional (frequency, orientation), while successful RB category identification was dependent only on a single dimension (frequency).

During Block 6, II and RB trials were randomly intermixed requiring participants to switch between procedural and declarative systems on a trial–by–trial basis. Block 7 was similar to Block 6 as it required participants to switch between II and RB categorization on a trial–by–trial basis except these trials included a button–switch component. The button or key associated with category A was now the button or key previously associated with category B and vice–versa. Likewise, the button or key associated with category C was now the button or key previously associated with category D and vice-versa. During Block 7 the labels at the bottom of the computer screen were changed to match this alteration (so they now read B, A, D, C).

Figure 2 shows an example of a trial. Each trial went as follows: a fixation point (crosshair) was presented on the screen for 1,500 ms. The crosshair background was either
green (signaling an II trial) or blue (signaling a RB trial). The 1,500 ms cue allowed for preparation for the participants. Previous work has shown that preparation time reduces the accuracy switch cost for participants who can successfully switch system on a trial–by–trial basis (Hélie, 2017). After 1,500 ms the crosshair disappeared and was replaced by a stimulus. Stimulus background color was indicative of possible response choices for participants (green for A–B, blue for C–D). The stimulus disappeared after the participant pressed a response button and auditory feedback was provided. Participants were given the opportunity to take breaks between blocks.

Participants were told that they would be completing a categorization task and that they would need to learn the category membership of stimuli through trial–and–error learning. Participants were told prior to Block 1 that for trials with a blue background bar width would determine category membership and that stimuli would belong to either category C or D. Participants were told prior to Block 2 that for trials with a green background bar width and orientation would both determine category membership and that stimuli would belong to either category A or B. Participants were informed about the button–switch component of Block 7 prior to completing the block. Due to the complexity of the task instructions, the experimenter stayed with the participants for the first few trials to ensure that they correctly understood the task.

Results

The mean accuracy in each block for each group is shown in Figure 3. As can be seen, participants in all groups performed well in the initial RB training block and were also able to learn the II stimuli. However, group differences emerged when the trial intermixing began (Block 6) and continued into the button–switch block (Block 7). These observations were supported by a Group (YC, OC, PD) × Block (1–7, repeated) ANOVA. Both the effects of Group \((F(2, 104) = 3.40, p = .04, \eta^2 = 0.06)\) and Block \((F(6, 624) = 36.66, p < .001, \eta^2 = 0.67)\) reached statistical significance. However, these main effects must be interpreted within the context of a statistically significant Group × Block interaction \((F(12, 624) = 3.19, p < .001, \eta^2 = 0.14)\). Based on the hypothesis that trial intermixing would affect the groups differently, we decomposed the interaction by looking at the effect of Group within each level of Block. There were no group differences in any of the single–task blocks (Blocks 1–5; all \(F(2, 104) < 2.20, p > .11, \eta^2 < 0.05\)). For the RB task (Block 1), the mean accuracy was 90.5%. For the II task, initial accuracy was 73.3% (Block 2) and improved to 82.1% (Block 5).

As suggested by Figure 3, the effect of Group was statistically significant in Block 6 \((F(2, 104) = 8.63, p < .001, \eta^2 = 0.14)\). Bonferroni–corrected pairwise t–tests show that YC participants were more accurate than PD \((t(85) = 4.28, p < .001, \eta^2 = 0.18)\) and OC \((t(85) = 2.91, p = .01, \eta^2 = 0.09)\) participants. OC and PD participants did not differ from each other \((t(38) = 0.58, p = 1, \eta^2 = 0.01)\). Accuracy during Block 6 was: YC = 83.2%, OC = 72.2%, and PD = 68.2%. The effect of Group was also statistically significant in Block 7 \((F(2, 104) = 3.50, p = .03, \eta^2 = 0.06)\), showing that YC (79.0%) were more accurate than OC (70.3%) during the button–switch block. However, none of the Bonferroni–corrected pairwise t–tests survived multiple testing correction in Block 7 (all \(t < 2.20, p > .09, \eta^2 < 0.06\)). The mean accuracy of the PD group in Block 7 was 71.0%.

Model–based analyses. Participants who can or cannot switch systems on a trial–by–trial basis are typically identified using model–based analysis (Ashby & Crossley, 2010; Crossley et al., in press; Erickson, 2008; Hélie, 2017). Specifically, the data from the intermixed trials block (Block 6 identified with a ‘M’ in Figure 3) were selected and RB trials were separated from II trials. Decision–bound models (Ashby, 1992; Maddox & Ashby, 1993) were then fit separately to the RB and II data. There are three general classes of decision–bound models, namely guessing models, explicit–reasoning models, and procedural–learning models (Hélie et al., 2017). For each data set, the best model was selected using the Bayes information criterion (Hélie, 2006). Participants whose data were best–fit by the optimal models were labeled as ‘switchers’. All other participants were labeled as ‘non–switchers’. In the current experiment, the optimal decision–bound model for the RB data was an explicit–reasoning model (i.e., a unidimensional rule on the \(x\)–axis of Figure 1b) and the optimal decision bound model for the II data was a procedural–learning model (i.e., the general linear classifier). More details on decision–bound models and fitting procedures can be found in Maddox and Ashby (1993) or Hélie et al. (2017).

The proportion of participants identified as switchers in each condition is shown in Figure 4. As can be seen, 52.2% of YC were identified as switchers. This is in line with previous results with undergraduate students (Hélie, 2017). However, only 30.0% of the participants in the OC group and 15% of the participants in the PD group were identified as switchers. A \(\chi^2\)–test shows that the proportion of participants identified as switchers was affected by group membership \((\chi^2(2) = 10.08, p = .007, \phi = 0.31)\). Bonferroni–corrected pairwise comparisons show that a smaller proportion of people with PD was identified as ‘switchers’ when compared to YC \((\chi^2(1) = 7.24, p = .02, \phi = 0.29)\). Other pairwise comparisons did not reach statistical significance \((\chi^2(1) < 2.23, p > .13, \phi < 0.17)\).

Trial–by–trial switch cost (accuracy). One well–known effect in the task–switching literature is switch cost (Kiesel et al., 2010). Essentially, when tasks are alternating, there is a cost associated with the switch that is visible in accuracy and RT. Crossley et al. (in press) argued that system–
switching could be interpreted as a form of task-switching, and Hélie (2017) extended this work by testing the effect of factors known to affect task switching in a system-switching paradigm. Switch cost was calculated as follows: Using the data from each participant’s Block 6, trials were labelled as “stay” or “switch” depending on whether the previous trial required using the same system (or not) for optimal performance. As is customary, the first trial of the block was discarded since there was no previous trial and could not be labeled. Next, for each participant, the mean accuracy for Switch trials was subtracted from the mean accuracy for Stay trials.

Figure 5 shows the accuracy switch cost for switchers and non-switchers in each group. A Switch (Switchers, Non-switchers) × Group ANOVA was performed. The ANOVA showed no statistically significant effect of Group

\( (F(2, 101) = 0.81, p .45, \eta^2 = 0.01) \), Switch \( (F(1, 101) = 0.30, p .59, \eta^2 = 0.00) \), or interaction \( (F(2, 101) = 0.96, p .39, \eta^2 = 0.02) \). Hence, there was no evidence of differential accuracy switch costs for each group or switch status. The mean accuracy switch cost was 4.5%.

Another informative way to analyze the data is to compute whether the switch cost was statistically different from 0. One-sample t-tests were performed for each group, and the results show that the switch cost was statistically significant for YC switchers \( (t(34) = 3.42, p .002, \eta^2 = 0.26) \) and non-switchers \( (t(31) = 3.27, p .003, \eta^2 = 0.26) \). For older adults, a statistically significant accuracy switch cost was found for PD non-switchers \( (t(16) = 3.32, p .004, \eta^2 = 0.41) \) as well as OC switchers \( (t(5) = 4.06, p .01, \eta^2 = 0.77) \). The effect was trending for OC non-switchers \( (t(13) = 1.87, p .08, \eta^2 = 0.21) \) and non-significant for PD switch-
ners ($t(2) = 0.16, p = .89, \eta^2 = 0.01$).

**Trial–by–trial switch cost (response time).** The RT switch cost was calculated as follows: First, the data in each participant’s Block 6 were labeled as *Switch* or *Stay* using the same method as when calculating the accuracy switch cost (see previous subsection). Second, error trials were removed. Finally, the mean *Stay* RT was subtracted from the mean *Switch* RT individually for each participant. Figure 6 shows the RT switch cost for switchers and non–switchers in each group. Similar to accuracy, a *Switch* × Group ANOVA was performed. The effect of Group reached statistical significance ($F(2,101) = 8.68, p < .001, \eta^2 = 0.18$), but the effect of *Switch* ($F(1,101) = 0.31, p = .58, \eta^2 = 0.01$) and the interaction between the factors ($F(2,101) = 0.27, p = .76, \eta^2 = 0.01$) both failed to reach statistical significance. Bonferroni–corrected pairwise $t$–tests show that RT switch cost was smaller for YC than for OC ($t(85) = 4.51, p < .001, \eta^2 = 0.19$) and PD ($t(85) = 4.20, p < .001, \eta^2 = 0.17$) participants. The OC and PD conditions did not differ from each other ($t(38) = 0.32, p = 1, \eta^2 = 0.00$). The mean RT switch cost was: YC = 111 ms, OC = 357 ms, and PD = 401 ms.

Similar to accuracy switch cost, we next computed whether RT switch costs were statistically greater than 0. The results show that the RT switch cost was larger than 0 for all groups (all $t > 2.88, p < .03, \eta^2 > 0.21$) except for PD switchers ($t(2) = 2.14, p = 0.17, \eta^2 = 0.70$). Note that PD switchers had the largest RT switch cost (498 ms), but only 3 participants in the PD group were identified as switchers, so this non–significant result was likely caused by a lack of statistical power.

**Button–switch interference.** Previous research has shown interference (i.e. lower accuracy) when changing the location of response buttons after learning II, but not RB, categories in a non–system–switching experiment (Ashby, Ell, & Waldron, 2003). Button–switch interference was calculated as follows: For each participant, select the data from Block 6 and separate the RB from the II trials. Separately compute the mean RB and II accuracy. Repeat the same procedure with the data from Block 7. Subtract the RB and II accuracy in Block 6 from the corresponding means in Block 7. The button–switch interference is shown in Figure 7. As can be seen, the interference was fairly small in all conditions. A Switch × Group × Categories (RB vs. II, repeated) ANOVA showed a statistically significant effect of Categories ($F(1,101) = 7.81, p = .006, \eta^2 = 0.07$). The main effect of Categories shows that, as predicted, the mean button–switch interference was larger for II trials (4.8%) than for RB trials (0.0%). All other main effects and interactions failed to reach statistical significance (all $F < 1.65, p > .19, \eta^2 < 0.04$).

Similar to switch cost, we next calculated for each condition whether the button–switch interference was statistically different from 0. Button–switch interference was significantly larger than 0 for YC/II switchers ($t(34) = 2.81, p = .008, \eta^2 = 0.19$) and non–switchers ($t(31) = 3.21, p = .003, \eta^2 = 0.25$). In contrast, there was facilitation for PD/II switchers ($t(2) = -5.20, p = .04, \eta^2 = 0.93$). All other button–switch interference did not differ from 0 (all $|t| < 1.50, p > .14, \eta^2 < 0.07$).

**General discussion**

Previous research has shown that trial–by–trial system–switching in categorization is extremely difficult for young adults (Ashby & Crossley, 2010; Crossley et al., in press; Erickson, 2008; Hélie, 2017). However, categorization system–switching had never been studied with older adults. The goal of this experiment was to test the hypothesis that older
adults as well as people with PD would be impaired in trial–by–trial system switching. To test for this hypothesis, we replicated the 1–session condition with preparation time of Hélie (2017) with older adults and people with PD. The results showed that older adults and people with PD can learn the separate (blocked) categorization tasks as well as young adults, but that trial–by–trial system–switching is impaired. This was supported by: (1) lower overall accuracy in an in–termixed trials block, (2) a smaller proportion of PD participants who can switch systems when compared to YC, and (3) a higher RT switch cost for both older adults and people with PD. Below we discuss some implications of these results in more detail.

### Category learning

Previous results showed that older adults and people with PD performed similarly in linear II categorization but were impaired compared to young adults (Ashby, Noble, et al., 2003). Ashby, Noble, et al. further observed that people with PD were impaired in RB categorization when compared to both older and younger adults. Because the ability to switch between categorization systems requires being able to learn the categories individually, the present experiment design was made to be easy to learn. This manipulation was effective, as no group differences were observed during the blocked trials (Blocks 1–5). Importantly, we did not observe an impairment in RB categorization for people with PD. There is evidence that inhibitory processes are particularly impaired in PD (Price et al., 2009), so this difference between the present results and Ashby, Noble, et al. could be accounted for by a difference in the number of stimulus dimensions. The stimuli used in the present experiment included only two dimensions whereas the stimuli in Ashby, Noble, et al. included four dimensions. Hence, fewer dimensions needed to be inhibited in this experiment. For II categorization, we did not observe any difference between older adults (both controls and peoples with PD) and younger adults, and this could also be a result of the number of stimulus dimensions. In Ashby, Noble, et al., there was one irrelevant stimulus dimension in the II task, whereas there was no irrelevant dimension in the II task included in the present experiment. Hence, the number of irrelevant stimulus dimensions may be key when studying category learning in older adults and people with PD, with fewer irrelevant dimensions being easier. More research is needed to characterize the interaction between the number of irrelevant dimensions and aging in categorization.

### Switch cost

Previous results with task–switching have shown higher switch costs for older adults (Kray & Lindenberger, 2000), and people with PD are not impaired when compared to aged–matched controls in environmentally cued task–switching (Werheid et al., 2007). The current categorization system–switching experiment used environmentally cued system–switching and reproduced the higher switch cost for older adults and people with PD, who did not differ from each other. This cost was most visible in RTs. Crossley et al. (in press) showed that categorization system–switching is generally more difficult than task–switching since many task–switching experiments only require within–system switching. This could explain why fewer people with PD were identified as ‘switchers’ when compared to YC, while the proportion of ‘switchers’ did not differ between YC and OC (or between PD and OC). The presence of environmental switching cues could mask a difference of magnitude in system–switching deficits between people with PD and aged–matched controls. Future work should test an internally–cued categorization system–switching paradigm, with system–switching being cued by, e.g., a regular trial ordering (e.g., RB → RB → II → II → RB → etc...), to see if the additional difficulty would allow for distinguishing between people with PD and aged–matched controls.

### Button–switch interference

The last block of the experiment included a button–switch manipulation. This manipulation was exploratory, as button–switch interference had never been explored in categorization with older adults and people with PD. With young adults, previous results had shown a selective impairment for II categories but not for RB categories (Ashby, Ell, & Waldron, 2003). The results with young adults were reproduced, but there was no button–switch interference for older adults or...
people with PD for either RB or II categories.\textsuperscript{2} Ashby, Ell, and Waldron (2003) interpreted the button–switch interference with II stimuli as evidence of procedural learning in II categorization, and there is evidence of reduced procedural learning in older adults (Janacek, Fiser, & Nemeth, 2012) and people with PD (Nagy et al., 2007). Hence, this result may not be surprising and suggests that older adults may not learn II categories in the same way as young adults. Future research should explore this possibility for both functional and biological differences.

\section*{Limitations}

The present research had a number of limitations, chief among them was the absence of educational data. Earlier research has shown that education is usually not associated with performance in simple perceptual tasks such as perceptual categorization. For example, Reimers and Maylor (2005) did not find an association between task switching and education in a sample of 12,103 data sets using a facial categorization task. Furthermore, Possin, Cagigas, Strayer, and Filoteo (2006) found no correlation between level of education and performance in an object–based attention task. For this reason, educational data were not collected in the present study. In hindsight, it is clear that collecting this information would have assisted with interpretation of the results. Any future work on system–switching with older adults should collect this information.

Another limitation of this study is the absence of standard measures to evaluate general cognitive, affective, and visuo–perceptual functioning. Unfortunately, these measures are not typically used in the perceptual categorization literature so this data was not collected. Instead, self–report was used for inclusion in the study in a pre–screening interview for all of the older participants. For participants with Parkinson’s disease, the self–reported information was later confirmed by contacting their neurologists. Additionally, we observed that categorization performance did not differ between groups during blocked trials (Figure 3, Blocks 1–5), confirming the absence of general cognitive impairment. Still, the task instructions were complex, and having the experimenter stay with the participants for the first few trials may not have fully addressed the possibility that participants may not have fully understood the instructions. Future work on system–switching with older adults should collect this information to ensure that participants are not cognitively impaired and fully understand the instructions.

\section*{Conclusion}

So why are older adults and people with PD impaired in categorization system–switching? DA–producing neurons die at a rate of 5\%–10\% per decade of life with normal aging (Karrer et al., 2017), and PD is caused by the accelerated death of DA–producing neurons (Hélie et al., 2012a). Reduced DA levels have been associated with less flexible cognition (Price et al., 2009), which could explain difficulty with categorization system–switching. In line with these findings, Hélie et al. (2012a) reduced DA levels in COVIS (Ashby et al., 1998) to reproduce the effects of aging and PD in a large number of procedural learning and rule–based tasks (see also Hélie, Paul, \& Ashby, 2012b). It is thus possible that the ability to switch between categorization systems is related to the levels of tonic DA, especially in the dorsolateral prefrontal cortex (Zgaljardic et al., 2006). The present experiment showed similar deficits in system–switching for older adults and people with PD. This could have been caused by the task not being sufficiently difficult, or by a ceiling effect of the role of DA in categorization system–switching. Future research should use positron emission tomography (PET) or animal models to directly measure DA levels and how it relates to categorization system–switching. Differences in tonic DA levels may also account for individual differences between ‘switchers’ and ‘non–switchers’ within each age group. These could be easily measured using resting–state eye–blink rates (Jongkees \& Colzato, 2016). Finally, the finding that environmental cueing with preparation time allows for people with PD to flexibly switch between categorization systems without additional disease–related impairment suggests that interventions attempting to improve flexible cognition in people with PD may benefit from preparatory environmental cues. Clearly, much work is still needed to understand categorization system–switching and how cognitive flexibility can be improved in healthy aging and PD.

\section*{Acknowledgments}

This research was funded in part by the National Institute of Mental Health, award #2R01MH063760. The authors would like to thank Dr. Jessica Huber, Dr. S. Elizabeth Zaubuer, and Sandy Snyder for help with recruiting participants. Correspondence concerning this article should be addressed to Sébastien Hélie, Department of Psychological Sciences, Purdue University (e-mail: shelie@purdue.edu).

\section*{References}


\textsuperscript{2}There was a facilitation effect with RB categories for ‘switchers’ with PD, but only 3 PD participants were identified as switchers so this result should be interpreted with caution.


doi: 10.1037/xhp0000063


doi: 10.1016/j.neubiorev.2016.08.020

doi: 10.1016/j.neurobiolaging.2017.05.006


