Behavioral/Cognitive

Distinct Contributions of the Cerebellum and Basal Ganglia to Arithmetic Procedures

William Saban,[1,2] Pedro Pinheiro-Chagas,[3] Sravya Borra,[1,2] and Richard B. Ivry[4]

[1]Center for Accessible Neuropsychology, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel, [2]Department of Occupational Therapy, Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel, [3]UCSF Memory and Aging Center, Department of Neurology, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, California 94158, and [4]Department of Psychology, University of California, Berkeley, California 94720-1650

Humans exhibit complex mathematical skills attributed to the exceptional enlargement of neocortical regions throughout evolution. In the current work, we initiated a novel exploration of the ancient subcortical neural network essential for mathematical cognition. Using a neuropsychological approach, we report that degeneration of two subcortical structures, the cerebellum and basal ganglia, impairs performance in symbolic arithmetic. We identify distinct computational impairments in male and female participants with cerebellar degeneration (CD) or Parkinson's disease (PD). The CD group exhibited a disproportionate cost when the arithmetic sum increased, suggesting that the cerebellum is critical for iterative procedures required for calculations. The PD group showed a disproportionate cost for equations with increasing addends, suggesting that the basal ganglia are critical for chaining multiple operations. In Experiment 2, the two patient groups exhibited intact practice gains for repeated equations at odds with an alternative hypothesis that these impairments were related to memory retrieval. Notably, we discuss how the counting and chaining operations relate to cerebellar and basal ganglia function in other task domains (e.g., motor processes). Overall, we provide a novel perspective on how the cerebellum and basal ganglia contribute to symbolic arithmetic. Our studies demonstrate the constraints on the computational role of two subcortical regions in higher cognition.

Key words: arithmetic; ataxia; Parkinson's disease; subcortex

Significance Statement

Research on the neurobiology of mathematics has focused on the cerebral cortex, particularly the frontoparietal regions. In the present study, we asked how disorders primarily affecting subcortical structures impact performance on symbolic arithmetic operations. Participants with Parkinson's disease showed a greater impairment as the number of operations increased, and participants with cerebellar degeneration showed a greater impairment as the magnitude of the operations increased. This selective impairment points to the distinctive roles of the cerebellum and basal ganglia in symbolic arithmetic. These results suggest that two major subcortical structures can support symbolic complex cognition.

Introduction

A hallmark of human cognition is our ability to engage in complex reasoning that requires the understanding and utilizing of abstract concepts. One powerful example of this is mathematical cognition. While a sense of quantity and simple processes may be observed in many species (Agrillo et al., 2008, 2010; Dadda et al.,

2009; Leibovich-Raveh et al., 2021), humans are unique in their ability to engage in symbolic arithmetic reasoning. Even solving a relatively simple addition problem (e.g., $5 + 7 = 12$) is a complex cognitive process, requiring several numerical (e.g., counting) and more generic (e.g., memory retrieval) mental processes.

Our understanding of the neural network essential for mathematical cognition has benefitted from the use of the many tools of cognitive neuroscience. One prominent theme is the importance of neocortical regions, particularly the central role of frontoparietal neocortical areas (Gruber et al., 2001; Dehaene et al., 2004; Ansari and Dhital, 2006; Grabner et al., 2009; Zamarian et al., 2009; Arsalidou and Taylor, 2011; Andres et al., 2011; Cohen-Kadosh and Dowker, 2015). Meta-analyses of functional magnetic resonance imaging (fMRI) data point to the consistent engagement of the inferior parietal lobule and subregions of the

prefrontal cortex during arithmetic calculations (Arsalidou and Taylor, 2011). Developmental imaging studies have shown that the intraparietal sulcus (IPS) is a biomarker of arithmetic skills (Isaacs et al., 2001; Emerson and Cantlon, 2015). Correspondingly, neurological patients with damage to these regions exhibit impairments on tests of mathematical cognition (Van Harskamp and Cipolotti, 2001).

Notably, there has been little discussion of the possible contribution of subcortical regions to mathematical cognition (Saban and Gabay, 2023) and most of the previous research has consisted of case studies (Ojemann, 1974; Corbett et al., 1986; Hittmair-Delazer et al., 1994, 1995; Dehaene and Cohen, 1997). This is surprising given the expanding appreciation of the contribution of subcortical regions to higher-level cognition (Buckner, 2013; Saban et al., 2017, 2018c, 2019, 2021). Indeed, the functional domain of two prominent subcortical structures, the cerebellum and basal ganglia has been recognized to extend beyond the motor domain (Owen et al., 1992; Middleton and Strick, 1994; Balsters et al., 2013; Sokolov et al., 2017; Bostan and Strick, 2018; Schmahmann, 2019). Both regions have reciprocal connectivity with many of the cortical areas associated with mathematical cognition (Ide et al., 2011; Bostan and Strick, 2018; Milardi et al., 2019) and, although outside the focus of neuroimaging studies, activation changes in both the cerebellum and basal ganglia have been consistently observed even in contrasts that control for overt motor responses (Ischebeck et al., 2007; Zago et al., 2008).

In terms of neuropsychological research, studies involving individuals with Parkinson's Disease (PD) have generally been descriptive, involving assessments with a range of instruments used to evaluate mathematical abilities (Tamura et al., 2003; Delazer et al., 2004; Zamarian et al., 2006). In general, basic mathematical processes such as counting, magnitude comparison, and basic arithmetic procedures (e.g., addition, subtraction) appear to be unaffected in PD. The few reports of impairment tend to be on tasks that require solving relatively complex equations, where complexity might be number of single-digit numbers that can be mentally added or the time required to solve an equation (e.g., $2 + 5 + 3 + 7 + 6$). However, these studies have not directly manipulated or even operationalized complexity; as such, they do not provide clear insight into the specific computations affected by PD.

Even less research has examined how cerebellar dysfunction impacts mathematical cognition. We recently tested participants with cerebellar degeneration (CD) on a verification task in which they had to add or multiply two single-digit numbers. The CD group demonstrated an intriguing dissociation: Whereas they exhibited the typical increase in response time (RT) as the calculated sum becomes larger for addition and multiplication problems [i.e., the problem size effect (Ashcraft and Guillaume, 2009)], the slope of this function was selectively elevated relative to controls in the addition condition only (McDougle et al., 2022). We hypothesize that this impairment reflects a slowed rate of spatial movement along a "mental number line."

In the current study, we take a neuropsychological approach to further explore the role of subcortical regions in mathematical cognition. We compared the performance of individuals with CD, PD and neurotypical control participants on an arithmetic verification task (AVT). We used manipulations that allowed us to probe two core procedures for addition, counting, and recursion, with the former manipulated by varying the sum of the equation and the latter manipulated by the number of addends. Although the sparsity of prior work on this problem

precludes strong predictions, we expected the CD group would exhibit a larger effect on sum manipulation, consistent with our earlier findings. Given the role of the basal ganglia in a variety of tasks that require chaining together multiple steps (Pascual-Leone et al., 1993; Meiran et al., 2004; Shohamy et al., 2005; Muslimović et al., 2007), we predicted that the PD group would show a disproportionately larger cost on problems involving more addends. In a second experiment, we measured the short-term practice benefits exhibited by the three groups on the addition problems, asking if either patient group exhibited a learning impairment in the use of arithmetic algorithms or memory retrieval (Logan, 1988; Rickard, 1997; Tenison et al., 2016).

Materials and Methods

Participants. An online platform, PONT (Saban and Ivry, 2021; Binoy et al., 2023), was used to recruit and test the participants. PONT entails five primary steps: (1) contact support group leaders/web-based platforms to advertise the project; (2) provide means for interested individuals to initiate contact, a requirement set by our IRB protocol; (3) conduct interactive, remote neuropsychological assessments; (4) automated administration of the experimental tasks; and (5) provide payment and obtain user feedback.

At the time of this project, there were approximately 183 individuals in the PONT database (Binoy et al., 2023). For Experiment 1, invitations were sent to 47, 72, and 64 individuals in the Control, CD, and PD groups, respectively. The overall response rate to the first email was approximately 15%, and after a few follow-up rounds of emails, we reached our goal of at least 20 participants per group. Two participants were excluded based on a failure to respond correctly on the attention probes (1 Control and 1 CD) and three participants reported connection issues and terminated the program (1 Control, 1 CD, and 1 PD). Of those who completed the study, we excluded the data from two participants who had accuracy scores at chance level (1 CD and 1 PD). Thus, the final sample consisted of 20 Control, 17 CD, and 18 PD in male and female participants. The CD group was composed of 10 participants with a known genetic subtype and 7 participants with an unknown etiology (idiopathic ataxia). The mean duration since diagnosis for the CD group was 8.2 years (SE = 2.6). The mean duration since diagnosis for the PD group was 6.1 years (SE = 1.0). None of the participants in the PD group had undergone surgical intervention as part of their treatment (e.g., DBS) and all were tested while on their current medication regimen.

The same recruitment procedure was followed in Experiment 2 with a goal of enlisting a minimum of 20 participants per group. Thirty-one participants (12 CD, 7 PD, 12 Control) who had participated in Experiment 1 also completed Experiment 2. One PD participant was excluded based on a failure to respond correctly on the attention check, three participants did not complete the study due to connectivity issues (1 from each group), and two completed the study but were not included in the analyses due to chance level of performance (1 Control and 1 PD). Thus, the final sample consisted of 20 Control, 20 CD, and 17 PD. The CD group included 12 participants with a known genetic subtype and 8 participants with idiopathic ataxia. The mean duration since diagnosis for the CD group was 4.2 years (SE = 1.0). The mean duration since diagnosis for the PD group was 8.1 years (SE = 1.3) and, as in Exp 1, none had undergone surgical intervention as part of their treatment, and all were tested while on their current medication regimen.

All participants provided informed consent under a protocol approved by the institutional review board at the University of California, Berkeley.

Neurological and neuropsychological assessment. Individuals were invited by email to participate in an online, live interview with an experimenter. After providing informed consent, the participant completed a demographic questionnaire. The experimenter then administered a modified version of the Montreal Cognitive Assessment test [MoCA (Nasreddine et al., 2005)] as a brief evaluation of cognitive status. For

Control participants, the session ended with the completion of the MoCA.

The PD and CD participants continued on to the medical evaluation phase. First, the experimenter obtained the participant's medical history, asking questions about age at diagnosis, medication and other relevant information (e.g., DBS for PD; genetic subtype if known for CD), and a screening for other neurological or psychiatric conditions. Second, the experimenter administered a modified version of the motor section of the Unified Parkinson's Disease Rating Scale, [UPDRS (Martínez-Martín et al., 1994)] to the PD participants and the Scale for Assessment and Rating of Ataxia (SARA (Schmitz-Hübsch et al., 2006) to the CD participants).

Modifications were made to these assessment instruments to make them better suited for online testing. For the MoCA test, we eliminated the "Alternating Trail Making" item since this requires providing the participant with a paper copy of the task. For the UPDRS and SARA, we modified items that require the presence of a trained individual to ensure safe administration. We eliminated the "Postural Stability task" from the UPDRS since it requires an experimenter to abruptly pull on the shoulders of the participant. We modified three items on the UPDRS ("Arising from Chair", "Posture", and "Gait"), obtaining self-reports from the participant rather than the standard evaluation by the experimenter. Similarly, we obtained self-reports of stance and gait for the SARA rather than observing the participant on these items. For the self-reports, we provided the scale options to the participant (e.g., on the SARA item for gait, 0 = normal/no difficulty and 8 = unable to walk even supported). The scores for the MoCA and UPDRS batteries were adjusted to reflect these modifications. For the online MoCA, the observed score was divided by 29 (the maximum online score), and then multiplied by 30 (the maximum score on the standard test). Hence, if a participant obtained a score of 26, the adjusted score will be (26/29) * 30, or 26.9. The same adjustment procedure was performed for the UPDRS. No adjustment was required for the SARA.

The interview took around 30 min for the control participants and 40–60 min for the PD and SCA participants. Table 1 provides demographic information for the three groups, as well as the adjusted MoCA, SARA (CD), and UPDRS (PD) scores.

Procedure. The experiments were programmed in Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) and designed to be compatible with any personal computer. Stimuli were presented as black characters on a white screen. The actual size in terms of visual angle varied given that participants used their own computer system, but we chose a font (7 HTML) that is clearly legible on all screens (as determined by testing when developing the system).

Participants were invited by email to participate in an experiment. The email provided an overview of the experimental task and included a link that could be clicked to initiate the experimental session. The instructions emphasized that the participant could start the experiment whenever he or she was ready but should only do so when they could complete the 30–45 min session. The link was associated with a unique participant ID, providing a means to ensure that the data were stored in an anonymized and confidential manner. Once activated, the link connected to the Gorilla platform is used to run the experimental session. The instructions were provided on the monitor in an automated manner, with the program advancing under the participant's control.

Each experiment involved an arithmetic verification task (AVT). The participant was asked to determine whether the mathematical equation (e.g., $3 + 2 + 6 = 11$) presented on the center of the screen was true (press the "M" key) or false (press the "Z" key). At the start of each trial in the AVT, a black fixation cross appeared in the middle of a white background. After 1,000 ms, the fixation cross was replaced by a stimulus display that consisted of the equation. The equation remained on the screen until a response was recorded or until 5,000 ms, whichever came first. The instructions emphasized that the response should be performed as quickly and accurately as possible. Visual feedback was presented for 500 ms above the equation, with a green checkmark ($\sqrt{}$) or red X indicating if the response was incorrect or correct. If a keypress was not detected after 3,500 ms, the participant was presented with the phrase "Respond faster" as feedback.

Experimental design and statistical analysis. In Experiment 1, we created ten different sets of 288 equations. Each individual was randomly presented with one of the sets. To minimize the effect of memory, each equation only appeared twice (50% of which were true). For any given equation, the same digit did not appear twice and only the digits 1–9 were used. In Experiment 1, we examined two procedures involved in solving arithmetic problems: Complexity and counting. As a manipulation of complexity, the equations could either require adding two addends (Simple) or three addends (Complex). The sum of the equations varied from 3 to 17 (e.g., $2 + 1 = 3$ to $9 + 8 = 17$) and served as our proxy for variation in counting. For each level of complexity, 144 unique equations were presented. Each participant had three breaks of 1 min each. The equations were presented in a random order. The experiment was completed within approximately 25 min.

In Experiment 2, we created ten different sets of 296 equations. Each participant was randomly presented with one of the sets. The stimulus set was limited to equations with three addends. For any given equation, the same digit did not appear twice. Eight of these equations constituted the Repeated stimuli (the same equation); the others constituted the No Repetition stimuli. The experiment consisted of a total of 36 blocks of 16 trials each (576 total trials; 50% of which were true). Each block was composed of one trial with each of the eight Repeat equations and eight trials with No Repetition (unique) equations. The two types of equations were included to provide measures of two forms of learning, memory-based learning (Repetition equations) and algorithm-based learning (No Repetition equations). We averaged across every three blocks to collapse the 36 blocks into 12 cycles. The experiment was completed within approximately 45 min.

To ensure that participants remained attentive, we included five "attention probes" that appeared in the start phase of the experiment or during the experimental block. For example, an attention probe might instruct the participant to press a specific key rather than selecting the "next" button on the screen to advance the experiment (e.g., "Do not press the 'next' button. Press the letter 'A' to continue"). If the participant failed to respond as instructed on these probes, the experiment continued, but the participant's results were not included in the analysis. After completing the experimental task, each participant answered a

Table 1. Demographic and neuropsychological summary of all groups

	Years of education	No. of females	Age	MoCA	Motor assessment
Experiment 1					
Control	17.8 ± 0.5 (14–22)	11	64.2 ± 2.0 (40–78)	27.9 ± 0.4 (25–30)	
CD	15.9 ± 0.9 (12–28)	14	61.7 ± 2.6 (40–81)	26.6 ± 0.4 (25–30)	10.1 ± 0.8 (2.0–18.5)
PD	16.5 ± 0.6 (12–20)	9	64.0 ± 1.0 (47–75)	26.5 ± 0.7 (25–30)	19.4 ± 1.6 (11.3–36.6)
Experiment 2					
Control	18.3 ± 0.5 (13–22)	7	66.9 ± 1.5 (51–78)	27.7 ± 0.4 (25–30)	
CD	16.1 ± 0.9 (12–25)	9	63.0 ± 3.2 (34–81)	26.7 ± 0.4 (25–30)	10.4 ± 1.1 (2.0–19.5)
PD	18.3 ± 0.8 (12–25)	9	67.6 ± 1.9 (49–80)	27.4 ± 0.7 (25–30)	17.1 ± 1 (10.7–26.9)

Mean ± SE and (range) for each demographic and neuropsychological variable. The scores of the MoCA and UPDRS are adjusted for online administration.

feedback questionnaire about their experience of the study (e.g., "How well was the study instructions explained." or "Did all the images display correctly?").

Results

Experiment 1

We used manipulations in Experiment 1 to probe two core procedures for addition. First, under the assumption that performing addition problems with single digit numbers involves a counting procedure that references a mental number line (Groen and Parkman, 1972), the time required to solve the problem will be related to the spatial distance that has to be traversed on this linear number line. Second, we manipulated the number of required steps, employing problems that involved either two or three operands. Adding up three single-digit numbers will take more time than adding up two single-digit numbers, presumably because the former requires chaining together procedures required for addition.

Figure 1A shows RT as a function of sum and group, with separate figures provided for the two-addend (simple) and three-addend (complex) problems. We used a linear mixed-effects [LME (Bates et al., 2015)] model with the factors Group (Control/PD/CD), Complexity (complex/simple), and Sum (17-3), with the participant as a random factor. Years of education, age, and MoCA score were included as covariates. Collapsing

across conditions, mean RTs were 1,482 ms, 1,585 ms, and 1,829 ms for the Control, PD, and CD groups, respectively. We log-transformed the RTs to fit the assumptions of the LME model in both experiments (e.g., the independent variables are related linearly to the RT and the errors are normally distributed). We observed that the CD group was significantly slower than the Control group (estimator (est) = 0.209, SE = 0.055, $p < 0.001$) and the PD group (est. = 0.143, SE = 0.056, $p = 0.014$). The PD group was numerically slower than the Control group, but this difference was not significant (est. = 0.066, SE = 0.054, $p = 0.229$).

Turning to our mathematical variables of interest, RT increased as a function of both sum and complexity, a pattern that was observed in all three groups. Considering the Control group as establishing baseline performance, this group was, on average, slower to respond to the three-addend problems compared to the two-addend problems (the "complexity effect"; est. = 0.381, SE = 0.005, $p < 0.001$). In terms of the sum effect, this group showed an increase in RT as the sum increased (est. = 0.0314, SE = 0.0008, $p < 0.001$).

Our main focus in Experiment 1 is on the comparison of the effect of complexity and sum between groups (Fig. 1B). The complexity effect (three-addends minus two-addends) was larger for the PD group compared to both the Control group (est. = −0.0447, SE = 0.0074, $p < 0.001$) and the CD group (est. = −0.0542, SE = 0.0078, $p < 0.001$). The Control and CD groups did not differ on this variable (est. = −0.0095, SE = 0.0076, $p = 0.213$).
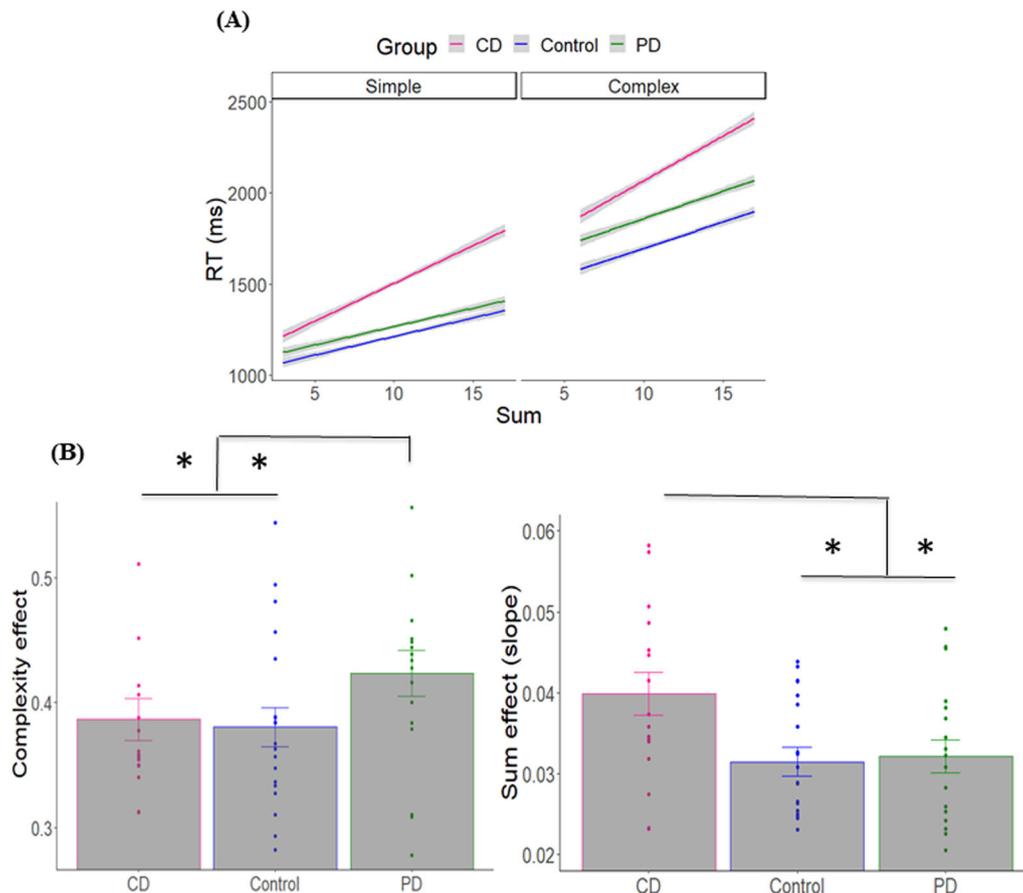
Figure 1. *A*, RT as a function of sum for two-addend problems (left) and three-addend problems (right). Error bars = 95% confidence interval. *B*, The left panel shows the effect of complexity, defined as the difference in mean RT for the three- and two-addend problems. The right panel shows the effect of sum, defined as the slope of the function relating RT to the sum. Note that the data were log-transformed to fit the assumptions of the LME model. Thus, the values in this figure represent the difference between the conditions in log(RT). Dots indicate the performance of each individual participant. Error bars = SEM; *$p < 0.001$.

In contrast, the sum effect was larger for the CD group compared to both the Control group (est. = 0.0297; SE = 0.0038, $p < 0.001$) and the PD group (est. = 0.0314, SE = 0.0039, $p < 0.001$). The Control and PD groups did not differ on this variable (est. = −0.0017, SE = 0.0037, $p = 0.638$). In terms of the covariates, there were no significant effects of education (est. = 0.005, SE = 0.025, $p = 0.821$), age (est. = 0.020, SE = 0.025, $p = 0.440$), or MoCA score (est. = −0.007, SE = 0.026, $p = 0.779$).

We performed a secondary analysis to determine the number of individuals within each group that showed a selective impairment in line with the group data. Since our sample size is lower than 50, a one-sided t-test (Type-I error rate of 5%) was used to compare each patient's test score against norms derived from the controls (Crawford and Howell, 1998). Note that for this type of analysis, using a test statistic with a t-distribution is more conservative compared to a test statistic based on a standard normal distribution. When looking at the complexity effect, 72.2% of the PD group (13) and 29.4% of the CD group (5) exhibited an impairment relative to the controls. When looking at the Sum effect, 76.4% CD (13) and 38.8% PD (7) were impaired compared to the controls. Thus, while around three-quarters of the participants in each patient group exhibit the impairment associated with their group, there are also cases in which a participant showed an impairment associated with the other group (e.g., CD on the complexity effect) or impairments on both tasks (3/group).

To examine the distributional data more formally, we used a chi-squared test, asking if the proportion of patients identified as impaired on the non-associated task (e.g., CD on Complexity) exceeded the expected proportion based on the Control data. For the Complexity effect, the Control and the CD proportions were not different ($p = 0.51$). Similarly, for the Sum effect, the Control and the PD proportions were not different ($p = 0.09$). Regarding a dual deficit, there was no difference between the Control and each patient group (CD: $p = 0.50$; PD: $p = 0.54$). Thus, these analyses indicate that the proportion of impairment observed within each group on the non-associated task was not different from that expected by chance based on the variability observed within the control participants.

To conclude, two main points can be taken from these results. First, both patient groups exhibited an impairment in the arithmetic verification task, providing novel evidence of the contribution of the cerebellum and basal ganglia to higher cognition. Second, the results point to a selective impairment of each patient group in terms of *how* the cerebellum and basal ganglia contribute to algorithmic procedures required for symbolic arithmetic. The CD group showed an impairment in the magnitude manipulation, exhibiting a larger slope for the function relating RT to the sum of the digits. In contrast, the PD group showed a bigger cost compared to the other two groups when adding three digits relative to two digits.

We hypothesize that the selective Sum effect reflects the involvement of the cerebellum in counting, a procedure that has been hypothesized to entail mental movement along a number line (Dehaene, 2003). The added cost incurred by the PD group for complex problems may reflect the involvement of the basal ganglia in chaining together a series of operations (Shohamy et al., 2005) or in facilitating transitions between successive operations [i.e., set switching (Meiran et al., 2004)].

Experiment 2

Arithmetic is a practiced skill in most literate adults. Nonetheless, we expect performance will improve over the course of the experiment, reflecting the benefits of short-term practice. An important distinction in the numerical learning literature has been made between benefits that accrue from improved efficiency in algorithmic procedures and benefits that accrue from memory processes (Logan, 1988; Rickard, 1997; Tenison et al., 2016). Prior work has shown that short-term benefits can arise from improved efficiency in counting (algorithmic learning) and enhanced memory retrieval (Ashcraft, 1982; Baroody, 1994; LeFevre et al., 1996; Campbell and Xue, 2001). Although the basal ganglia and the cerebellum have been associated with procedural learning in a variety of task domains (47, 65), we are unaware of any work looking at the contribution of these subcortical structures to learning in the math domain.

We compared practice benefits for the PD, CD, and controls on the AVT in Experiment 2. To measure short-term practice benefits associated with algorithmic and memory-based learning, we assessed problems that appeared only a single time (No-Repetition condition) and those that were repeated multiple times (Repetition condition). An algorithmic-based impairment should be manifest as a selective reduction of practice gains for non-repeated items. In contrast, a memory-based impairment would be evident as reduced practice gains for both non-repeated and repeated items.

Figure 2 shows RT as a function of the learning cycle and group, with separate functions for the Repetition and No Repetition conditions. In Experiment 2, the LME model included the factors Group (Control/PD/CD), Repetition Condition (Repetition/No repetition), and Cycle (1–12), with participant as a random factor. Years of education, age, and MoCA score were included as covariates. Collapsing across all conditions, the mean RTs were 1,953 ms, 2,148 ms, and 2,281 ms for the Control, PD, and CD groups, respectively, a pattern similar to that observed in Exp 1. Both patient groups were slower than the control group although this effect was only significant for the CD group (CD vs Control: est. = 0.177, SE = 0.059, $p = 0.004$; PD vs Control: est. = 0.110, SE = 0.062, $p = 0.080$). The mean RTs for CD and PD groups were not significantly different (est. = 0.066, SE = 0.063, $p = 0.299$).

As expected, the participants got faster at the task over the course of the experiment, and this improvement was especially marked in the Repetition condition. Using the control group to establish baseline performance, this group got faster for repeated items (est. = −41.4, SE = 2.2, $p < 0.0001$) and for non-repeated items (est. = −26.4, SE = 2.2, $p < 0.0001$). The difference between these two slopes was significant (est. difference in slope = −14.9, SE = 3.1, $p < 0.0001$), consistent with the hypothesis that performance in the Repetition condition benefits from memory retrieval.

Turning to our main analysis (Fig. 2B), there was a significant three-way interaction between Group × Repetition × Cycle (CD vs Control: est. = −0.022, SE = 0.022, $p = 0.046$; PD vs Control: est. = −0.003, SE = 0.011, $p = 0.766$). We compared the learning effect (slope of the function relating RT to Cycle) between the groups separately for the Repetition and No Repetition conditions. For the Repetition condition, the three groups showed a similar change in performance across cycles (CD vs Control: est. = 0.004, SE = 0.007, $p = 0.531$; PD vs Control: est. = 0.006, SE = 0.007, $p = 0.419$; CD vs PD: est. = −0.001, SE = 0.008, $p = 0.839$). However, for the No Repetition condition, the CD group showed less improvement over cycles than the Control group (est. = 0.023, SE = 0.007, $p = 0.002$) and the PD group (est. = 0.016, SE = 0.008, $p = 0.047$). The comparison between the Control and PD groups was not significant (est. = 0.006, SE = 0.008, $p = 0.425$). Neither education (est. = 0.009, SE =
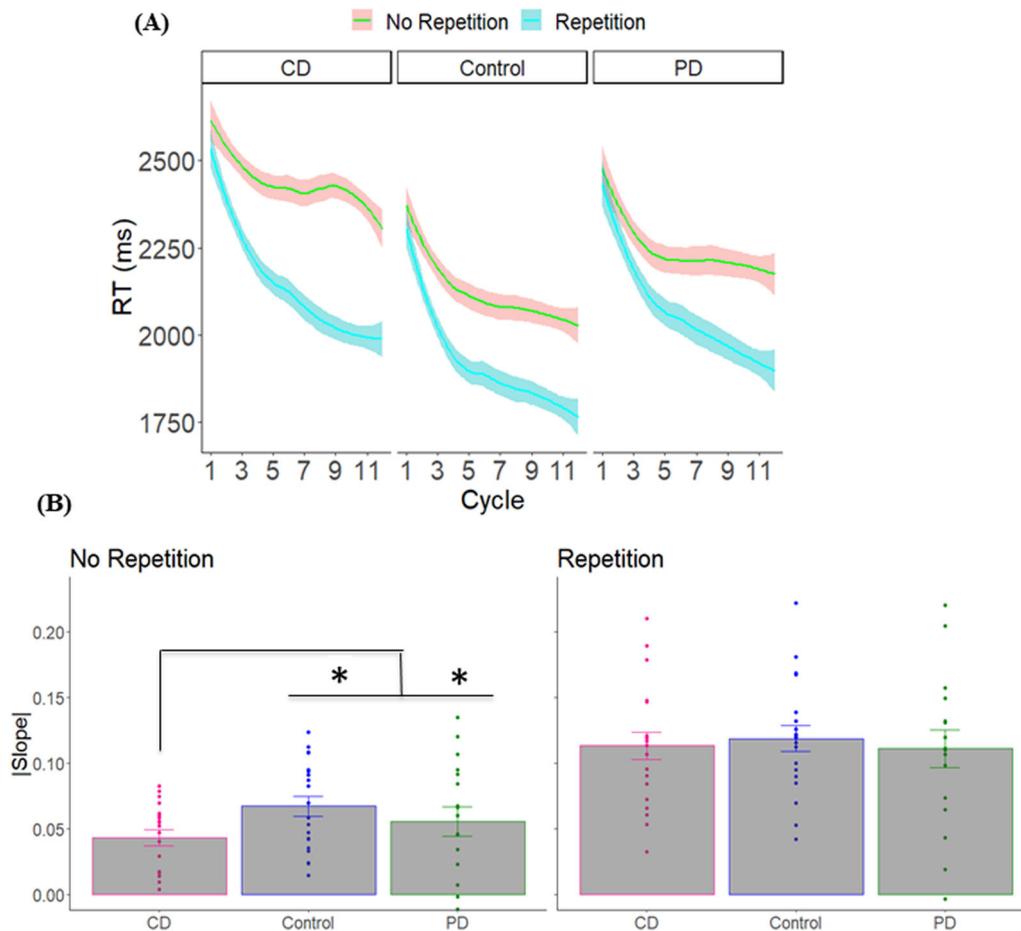
Figure 2. **A**, RT as a function of learning cycle for the Repetition and No Repetition conditions. Error bars = 95% confidence interval. **B**, Rate of improvement across cycles for each group for the No Repetition (left) and Repetition (right) conditions. Dots indicate performance of each individual participant. Error bars = SEM; *$p < 0.001$.

0.023, $p = 0.670$), age (est. = 0.005, SE = 0.023, $p = 0.815$), or MoCA score (est. = −0.015, SE = 0.024, $p = 0.521$) were significant when entered as covariates in the model.

In summary, the results of Experiment 2 show dissociable benefits from different forms of learning. Participants exhibited a large decrease in RT on the equations that repeated multiple times, an effect we assume is driven by the availability of item-specific memory-based retrieval. Given that the magnitude of this effect was similar in all three groups, we assume that the basal ganglia and cerebellum are not essential for the processes required to encode and retrieve the equations. In contrast, the CD group showed a reduced practice benefit in evaluating the novel equations, a signature of what we interpret as an indication of an impairment in algorithm-based learning. Given the results of Experiment 1, this selective impairment may be related to the counting procedure, here manifest as a problem in becoming more facile in counting over the course of the experimental session.

We performed an exploratory analysis limited to the individuals who participated in both experiments. Here we asked if there was a correlation between the Sum effect in Experiment 1 and the practice benefit in the No Repetition condition in Experiment 2, the two measures we have used to assess counting processes. None of the correlations were significantly different from zero (CD: $r = 0.29$, $p = 0.18$, Control: $r = −0.10$, $p = 0.37$, and PD: $r = −0.19$, $p = 0.34$). This may indicate that the two measures probe different aspects of counting; for example, the sum measure might

reflect the application of an algorithm whereas the practice measure might reflect the short term benefit in deploying the algorithm. However, we also recognize that this analysis is limited given that the sample size for each group is low (12 CD, 7 PD, 12 Controls) and the absence of test-retest measures to establish the reliability of the two measures, a prerequisite for establishing an upper bound on the correlations.

Discussion

Converging evidence has highlighted the involvement of the basal ganglia and cerebellum in a broad range of cognitive domains including cognitive control, decision making, and language (Middleton and Strick, 2000; Buckner, 2013; Anon, 2016; Bostan and Strick, 2018; King et al., 2019; Schmahmann, 2019). In the present study, we examined the involvement of these subcortical structures in mathematical cognition, a domain in which current models focus on a frontal-parietal cortical network (Arsalidou and Taylor, 2011; Arsalidou et al., 2018). Taking a neuropsychological approach, we tested patients with Parkinson's disease (PD) and cerebellar degeneration (CD) on mental addition tasks, using these groups as models to evaluate the role of the basal ganglia and cerebellum, respectively.

The results of Experiment 1 revealed a selective impairment for each patient group. Relative to both the Control and PD groups, the CD group showed a larger slope for the function relating RT to the sum of the digits (i.e., problem size effect).

In contrast, the PD group showed a larger cost relative to the other two groups when evaluating problems with three addends relative to two addends. These results not only implicate the basal ganglia and cerebellum in mathematical cognition but also point to distinct computational contributions. In Experiment 2, we used a learning design that allowed us to examine item-specific and item-general practice benefits. Both patient groups showed a similar practice benefit as the Control group on repeated equations. Coupled with the results of Experiment 1, this null finding argues against the hypothesis that the patients' impairments might be fully explained by a deficit in memory processes. Interestingly, the CD group showed an attenuated practice effect when tested with non-repeated novel equations. This pattern of results underscores a selective contribution of the cerebellum to algorithm-based learning.

Addition problems involving single digits can be solved by memory retrieval or/and by counting procedures (Groen and Parkman, 1972; Ashcraft, 1982; Butterworth, 2005; Ashcraft and Guillaume, 2009; Barrouillet and Thevenot, 2013; Cohen-Kadosh and Dowker, 2015; Chen and Campbell, 2018; Grotheer et al., 2018). From a memory perspective, the problem size effect may appear because we have more experience with small-size problems (e.g., $3 + 2$) compared to large-size problems (e.g., $3 + 8$). From a counting perspective, the problem size effect arises from a process that references a mental number line; thus, adding eight will take more time than adding two because of the additional iterations required by the former. In one variant of this model, these iterations are conceptualized as mental movement along a spatialized number line (Dehaene, 2003). Recently, it was found that participants with CD show a larger problem size effect on addition problems, but not multiplication problems (McDougle et al., 2022). This previous result points to a role of the cerebellum in a counting-like procedure, given that a retrieval problem should impact both types of procedures. The results from Experiment 1 are consistent with this hypothesis and points to some degree of neural specificity in that the PD group did not show a similar problem size effect.

The results from Experiment 2 can also be interpreted as evidence of a counting deficit associated with cerebellar degeneration. With repeated items, the demands on counting can be reduced since the answer can be obtained by reference to recently activated memories (Logan, 1988). The fact that the CD group showed comparable practice gains for repeated items as the Control and PD groups suggests that this type of memory is intact. However, as suggested by the results of Experiment 1, non-repeated items require more specific mathematical procedures including counting. As such, the CD impairment would continue to be manifest. Moreover, the reduced rate of learning found in Experiment 2 suggests that some of the improvement exhibited by the Control and PD groups for non-repeated items reflects short-term benefits in the rate of counting. We recognize the inferential reasoning underlying this argument. Future experiments could directly examine how different arithmetic procedures (i.e., counting, carrying, number of steps) change over the course of learning.

To this point, we have focused on the hypothesis that the problem size effect may be reflective of a counting-like procedure. However, one should note that the problem size effect may also arise from a different mathematical procedure, carrying: Problems with larger sums are more likely to require a carry procedure (e.g., $6 + 7$) than problems with smaller sums (e.g., $3 + 4$). One way to assess the carry procedure effect is to compare equations that either require carrying or not, that are matched in terms of the lower addend (e.g., $5 + 3$ vs $8 + 3$) given that people tend to increment from the larger digit (Barrouillet and Thevenot, 2013). Although the sample is limited, for equations matched in this manner, we did observe longer RTs for equations in which the sum was greater than 10. However, we also observed a significant positive slope in RT if we only consider the set of equations that summed to higher than 10. The latter result underscores that the problem size effect is not solely due to carrying. It would be interesting in future experiments to create balanced stimulus sets to have sufficient power to independently examine the effect of cerebellar damage on carrying and counting.

As noted above, the PD group showed a different pattern of performance than the CD group. Their problem size effect was comparable to that observed for the Control Group. However, relative to both of the other groups, the PD group exhibited a larger cost in adding three single-digit numbers compared to adding two single-digit numbers. While we referred to this manipulation as one of complexity, the two- and three-addend problems involve the same procedures; where they differ is that the latter require additional steps (e.g., add two numbers and that sum becomes an addend to go with the third number). Models of basal ganglia function have highlighted the involvement of this structure in executive functions, including working memory (Zamarian et al., 2006), chaining a series of operations (Shohamy et al., 2005), or facilitating the transition between successive procedures [i.e., set switching (Meiran et al., 2004)]. The larger complexity effect observed in the PD group would be consistent with these computations: Three-addend problems are likely more taxing on working memory, as well as require performing a longer series of procedures. Future work can use tasks that hone in on different computations; at present, it is intriguing to consider that the impairment observed in the math domain may be indicative of a domain-independent impairment.

Despite the inflated complexity effect observed in the PD group, these participants showed similar practice benefits as the Controls for both repeated and non-repeated items in Experiment 2. This intact rate of learning suggests that the improvement exhibited by the Control group for non-repeated items is unlikely to be due to short-term benefits in procedures underlying the complexity effect (working memory, chaining, task switching). However, this hypothesis needs to be viewed with caution since it rests on a null result, the absence of a practice deficit in the PD group. Future experiments should directly examine how subcortical (and cortical) contributions to different arithmetic procedures (i.e., counting, carrying, complexity) change over the course of learning.

Models of the brain networks supporting mathematical cognition have focused mainly on the cerebral cortex, and in particular, the frontoparietal network. The present results point to the need for a broader conceptualization, one that incorporates the basal ganglia and cerebellum. We note that we have only looked on one elementary feature of numerical fluency, the addition of single-digit numbers. A more comprehensive picture will require experiments that examine subcortical contributions to a broad range of mathematical procedures (e.g., subtraction, multiplication, geometric relations). The selective impairment of each patient group in counting and complexity gives reason to expect that this line of research will help reveal how the basal ganglia and cerebellum work in concert with the cortex to support mathematical cognition.

We recognize that we have treated the basal ganglia and cerebellum as unitary structures, whereas there is evidence

demonstrating functional specialization within each of these subcortical structures. While a more detailed analysis of anatomical-behavioral relationships would be useful in understanding how these subcortical regions contribute to cognition, we were unable to acquire sufficient MRI records for the patient groups. Given that our sample was distributed over a large geographic area and the data were collected during the pandemic, we were not in a position to obtain MRIs for the participants. Nonetheless, future work should use lesion analysis or neuroimaging methods to examine regions within the cerebellum and basal ganglia that are critical for arithmetic operations.

Considered more broadly, the present results provide a novel causal demonstration of how subcortical systems contribute to higher-level cognition (Anon, 2016; Saban et al., 2017, 2018a,b, 2019, 2021). We assume that the computations provided by the cerebellum and basal ganglia for math initially evolved to support more elementary functions, functions that might be performed in the absence of a developed cortex (Güntürkün and Bugnyar, 2016; Saban and Gabay, 2023). Coupling these subcortical systems with the expanded representational capacity of the cortex allows for the emergence of complex cognitive representations such as arithmetic. As Paul Rozin has noted, a process that evolved to solve a specific problem may come to be exploited across a broad range of task domains (Rozin, 1976). In this manner, the functional domain of subcortical regions has expanded, evolving in parallel with the cortex to create novel cognitive competences.

Conclusion

Very little is known about the neuro-evolutionary development of numerical abilities. To date, the literature has emphasized the role of cortical regions in arithmetic abilities. A central question addressed by the present study is whether ancient subcortical neural mechanisms are involved in humans' arithmetic abilities. Divergent patterns of impairment were observed in participants with either degeneration of the basal ganglia or cerebellum on an arithmetic task. These results highlight that these subcortical structures make distinct computational contributions to symbolic arithmetic procedures. Taken together, these results provide compelling support for the constraints on the computational role of two major subcortical regions in higher cognition (Saban and Gabay, 2023).

References

Agrillo C, Dadda M, Serena G (2008) Choice of female groups by male mosquitofish (Gambusia holbrooki). Ethology 114:479–488.

Agrillo C, Piffer L, Bisazza A (2010) Large number discrimination by mosquitofish. PLoS One 5:e15232.

Anon (2016) Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function author(s): Frank A. Middleton and Peter L. Strick published by: American Association for the Advancement of Science Stable. Available at: .

Andres M, Pelgrims B, Michaux N, Olivier E, Pesenti M (2011) Role of distinct parietal areas in arithmetic: an fMRI-guided TMS study. Neuroimage 54:3048–3056.

Ansari D, Dhital B (2006) Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: an event-related functional magnetic resonance imaging study. J Cogn Neurosci 18:1820–1828.

Anwyl-Irvine AL, Massonnié J, Flitton A, Kirkham N, Evershed JK (2020) Gorilla in our midst: an online behavioral experiment builder. Behav Res Methods 52:388–407.

Arsalidou M, Pawliw-Levac M, Sadeghi M, Pascual-Leone J (2018) Brain areas associated with numbers and calculations in children: meta-analyses of fMRI studies. Dev Cogn Neurosci 30:239–250.

Arsalidou M, Taylor MJ (2011) Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. Neuroimage 54:2382–2393.

Ashcraft MH (1982) The development of mental arithmetic: a chronometric approach. Dev Rev 2:213–236.

Ashcraft MH, Guillaume MM (2009) Mathematical cognition and the problem size effect. Psychol Learn Motiv 51:121–151.

Balsters JH, Whelan CD, Robertson IH, Ramnani N (2013) Cerebellum and cognition: evidence for the encoding of higher order rules. Cereb Cortex 23:1433–1443.

Baroody AJ (1994) An evaluation of evidence supporting fact-retrieval models. Learn Individ Differ 6:1–36.

Barrouillet P, Thevenot C (2013) On the problem-size effect in small additions: can we really discard any counting-based account? Cognition 128:35–44.

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Binoy S, Woody R, Ivry RB, Saban W (2023) Feasibility and efficacy of online neuropsychological assessment. Sensors 23:5160.

Bostan AC, Strick PL (2018) The basal ganglia and the cerebellum: nodes in an integrated network. Nat Rev Neurosci 19:338–350.

Buckner RL (2013) The cerebellum and cognitive function: 25 years of insight from neuroimaging. Neuron 80:807–815.

Butterworth B (2005) The development of arithmetical abilities. J Child Psychol Psychiatry 46:3–18.

Campbell JID, Xue Q (2001) Cognitive arithmetic across cultures. J Exp Psychol Gen 130:299–315.

Chen Y, Campbell JID (2018) "Compacted" procedures for adults' simple addition: a review and critique of the evidence. Psychon Bull Rev 25: 739–753.

Cohen-Kadosh R, Dowker A (2015) The Oxford handbook of numerical cognition. Oxford, England: Oxford University Press.

Corbett AJ, Mccusker EA, Davidson OR (1986) Acalculia following a dominant-hemisphere subcortical infarct. Arch Neurol 43:964–966.

Crawford JR, Howell DC (1998) Comparing an individual's test score against norms derived from small samples. Clin Neuropsychol 12:482–486.

Dadda M, Piffer L, Agrillo C, Bisazza A (2009) Spontaneous number representation in mosquitofish. Cognition 112:343–348.

Dehaene S (2003) The neural basis of the Weber–Fechner law: a logarithmic mental number line. Trends Cogn Sci 7:145–147.

Dehaene S, Cohen L (1997) Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. Cortex 33:219–250.

Dehaene S, Molko N, Cohen L, Wilson AJ (2004) Arithmetic and the brain. Curr Opin Neurobiol 14:218–224.

Delazer M, Domahs F, Lochy A, Karner E, Benke T, Poewe W (2004) Number processing and basal ganglia dysfunction : a single case study. 42:1050–1062.

Emerson RW, Cantlon JF (2015) Continuity and change in children's longitudinal neural responses to numbers. Dev Sci 18:314–326.

Grabner RH, Ansari D, Koschutnig K, Reishofer G, Ebner F, Neuper C (2009) To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. Neuropsychologia 47:604–608.

Groen GJ, Parkman JM (1972) A chronometric analysis of simple addition. Psychol Rev 79:329–343.

Grotheer M, Jeska B, Grill-Spector K (2018) A preference for mathematical processing outweighs the selectivity for Arabic numbers in the inferior temporal gyrus. Neuroimage 175:188–200.

Gruber O, Indefrey P, Steinmetz H, Kleinschmidt A (2001) Dissociating neural correlates of cognitive components in mental calculation. Cereb Cortex 11:350–359.

Güntürkün O, Bugnyar T (2016) Cognition without cortex. Trends Cogn Sci 20:291–303.

Hittmair-Delazer M, Sailer U, Benke T (1995) Impaired arithmetic facts but intact conceptual knowledge a single—case study of dyscalculia. Cortex 31:139–147.

Hittmair-Delazer M, Semenza C, Denes G (1994) Concepts and facts in calculation. Brain 117:715–728.

Ide JS, Li C, Shan R (2011) A cerebellar thalamic cortical circuit for error-related cognitive control. Neuroimage 54:455–64.

Isaacs EB, Edmonds CJ, Lucas A, Gadian DG (2001) Calculation difficulties in children of very low birthweight: a neural correlate. Brain 124:1701–1707.

Ischebeck A, Zamarian L, Egger K, Schocke M, Delazer M (2007) Imaging early practice effects in arithmetic. Neuroimage 36:993–1003.

King M, Hernandez-Castillo CR, Poldrack RA, Ivry RB, Diedrichsen J (2019) Functional boundaries in the human cerebellum revealed by a multi-domain task battery. Nat Neurosci 22:1371–1378.

LeFevre JA, Sadesky GS, Bisanz J (1996) Selection of procedures in mental addition: reassessing the problem size effect in adults. J Exp Psychol Learn Mem Cogn 22:216–230.

Leibovich-Raveh T, Raveh A, Vilker D, Gabay S (2021) Magnitude integration in the archerfish. Sci Rep 11:1–11.

Logan GD (1988) Toward an instance theory of automatization. Psychol Rev 95:492–527.

Martínez-Martín P, Gil-Nagel A, Gracia LM, Gómez JB, Martinez-Sarries J, Bermejo F, Cooperative Multicentric Group (1994) Unified Parkinson's disease rating scale characteristics and structure. Mov Disord 9:76–83.

Mcdougle SD, Tsay JS, Pitt B, King M, Saban W, Taylor JA, Ivry RB (2022) Continuous manipulation of mental representations is compromised in cerebellar degeneration. Brain 145:4246–4263.

Meiran N, Friedman G, Yehene E (2004) Parkinson's disease is associated with goal setting deficits during task switching. Brain Cogn 54:260–262.

Middleton FA, Strick PL (1994) Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. Science 266:458–461.

Middleton FA, Strick PL (2000) Basal ganglia and cerebellar loops: motor and cognitive circuits. Brain Res Rev 31:236–250.

Milardi D, et al. (2019) The cortico-basal ganglia-cerebellar network: past, present and future perspectives. Front Syst Neurosci 13:61.

Muslimović D, Post B, Speelman JD, Schmand B (2007) Motor procedural learning in Parkinson's disease. Brain 130:2887–2897.

Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 53:695–699.

Ojemann GA (1974) Mental arithmetic during human thalamic stimulation. Neuropsychologia 12:1–10.

Owen AM, James M, Leigh PN, Summers BA, Marsden CD, Quinn NP, Lange KW, Robbins TW (1992) Fronto-striatal cognitive deficits at different stages of Parkinson's disease. Brain 115:1727–1751.

Pascual-Leone A, Grafman J, Clark K, Stewart M, Massaquoi S, Lou J-S, Hallett M (1993) Procedural learning in Parkinson's disease and cerebellar degeneration. Ann Neurol 34:594–602.

Rickard TC (1997) Bending the power law: a CMPL theory of strategy shifts and the automatization of cognitive skills. J Exp Psychol Gen 126:288–311.

Rozin P (1976) The selection of foods by rats, humans, and other animals. Adv Study Behav 6:21–76.

Saban W, Gabay S (2023) Contributions of lower structures to higher cognition: towards a dynamic network model. J Intell 11:121.

Saban W, Gabay S, Kalanthroff E (2018a) More than just channeling: the role of subcortical mechanisms in executive functions—evidence from the stroop task. Acta Psychol 189:36–42.

Saban W, Ivry RB (2021) Pont: a protocol for online neuropsychological testing. J Cogn Neurosci 33:2413–2425.

Saban W, Klein RM, Gabay S (2018b) Probabilistic versus "pure" volitional orienting: a monocular difference. Atten Percept Psychophys 80:669–676.

Saban W, Raz G, Grabner RH, Gabay S, Kadosh RC (2021) Primitive visual channels have a causal role in cognitive transfer. Sci Rep 11:8759.

Saban W, Sekely L, Klein RM, Gabay S (2017) Endogenous orienting in the archer fish. Proc Natl Acad Sci 114:7577–7581.

Saban W, Sekely L, Klein RM, Gabay S (2018c) Monocular channels have a functional role in endogenous orienting. Neuropsychologia 111:1–7.

Saban W, Weinbach N, Gabay S (2019) Monocular channels have a functional role in phasic alertness and temporal expectancy. Atten Percept Psychophys 81:752–763.

Schmahmann JD, et al. (2019) The cerebellum and cognition. Neurosci Lett 688:62–75.

Schmitz-Hübsch T, et al. (2006) Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology 66:1717–1720.

Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA (2005) The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. Behav Brain Res 156:191–199.

Sokolov AA, Miall RC, Ivry RB (2017) The cerebellum: adaptive prediction for movement and cognition. Trends Cogn Sci 21:313–332.

Tamura I, Kikuchi S, Otsuki M, Kitagawa M, Tashiro K (2003) Deficits of working memory during mental calculation in patients with Parkinson's disease. J Neurol Sci 209:19–23.

Tenison C, Fincham JM, Anderson JR (2016) Phases of learning: how skill acquisition impacts cognitive processing. Cogn Psychol 87:1–28.

Van Harskamp NJ, Cipolotti L (2001) Selective impairments for addition, subtraction and multiplication. Implications for the organisation of arithmetical facts. Cortex 37:363–388.

Zago L, Petit L, Turbelin MR, Andersson F, Vigneau M, Tzourio-Mazoyer N (2008) How verbal and spatial manipulation networks contribute to calculation: an fMRI study. Neuropsychologia 46:2403–2414.

Zamarian L, Ischebeck A, Delazer M (2009) Neuroscience of learning arithmetic—evidence from brain imaging studies. Neurosci Biobehav Rev 33:909–925.

Zamarian L, Visani P, Delazer M, Seppi K, Mair KJ, Diem A, Poewe W, Benke T (2006) Parkinson's disease and arithmetics: the role of executive functions. J Neurol Sci 248:124–130.