High Cognitive Violation of Expectations is Compromised in Cerebellar Ataxia

Leonardo Daniel A[1,2], Eli Vakil[3] & William Saban[1,2]

[1]Center for Accessible Neuropsychology, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel.
[2]Department of Occupational Therapy, Faculty of Medical & Health Sciences, Tel Aviv University, Tel Aviv 69978, Israel.
[3]Department of Psychology and Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel.

Correspondence to: williamsaban@gmail.com

Keywords: Cerebellum; Arithmetic; Language; Cerebellar Ataxia; expectations.

Abstract

While traditionally considered a motor structure, the cerebellum is also involved in cognition. However, the underlying cognitive mechanisms through which the cerebellum contributes to evolutionarily novel cognitive abilities remain poorly understood. Another open question is how this structure contributes to a core unifying mechanism across domains. Motivated by the evolutionary principle of neural reuse, we suggest that a successful account of cerebellar contributions to higher cognitive domains will build on the structure's established role in motor behaviors. We conducted a series of neuropsychological experiments, assessing selective impairments in participants with cerebellar ataxia (CA) compared to neurotypicals in solving sequential discrete problems. In three experiments, participants were asked to solve symbolic subtraction, alphabet letter transformation, and novel artificial grammar problems, which were expected or unexpected. The CA group exhibited a disproportionate cost when comparing expected problems to unexpected problems, suggesting that the cerebellum is critical for violation of expectations (VE) across tasks. The CA group impairment was not found either when the complexity of the problem increased or in conditions of uncertainty. Together, these results demonstrate a possible causal role for the human cerebellum in higher cognitive abilities. VE might be a unifying cerebellar-dependent mechanism across motor and cognitive domains.

Significance Statement

While the cerebellum, a phylogenetically ancient brain region, is traditionally viewed as a motor structure, evidence suggests its involvement in cognition. However, the mechanisms by which the cerebellum supports evolutionarily novel cognitive abilities remain poorly understood. In addition, despite theoretical proposals, direct evidence for the cerebellum's contribution to a core unifying mechanism across non-motor domains is lacking. Drawing on the principle of neural reuse, we present neuropsychological evidence highlighting the cerebellum's causal role in symbolic arithmetic reasoning, alphabet transformation, and grammar problems via violation of expectations processes. The results offer a new perspective on how, rather than merely if, the cerebellum contributes to higher cognition, suggesting a constraint on its role in cognitive domains.

Introduction

Traditionally, the cerebellum has been primarily recognized as a motor structure. However, evidence indicates its involvement in various cognitive processes[1–6]. Yet, the underlying cognitive mechanisms through which the cerebellum contributes to higher cognitive abilities remain unknown. In addition, since the anatomy of the cerebellum is relatively uniform throughout its structure, it was suggested that its function may be consistent[4,7] (i.e., “universal cerebellar transform”). This led to the hypothesis that a cerebellar lesion will result in a core ubiquitous impairment across domains[4,7–9]. Despite these theoretical proposals, direct evidence for the cerebellum's contribution to a core unifying mechanism across non-motor domains is lacking.

A substantial body of neuropsychological, modeling, and imaging evidence supports the notion that the cerebellum is involved in motor control[6,7,10–14]. According to the forward model framework[13,14], the cerebellum encodes a predictive model [15](i.e., “internal model”). Specifically, the cerebellum is involved in prediction error[15], such that processing of the deviation between the predicted and the perceived information (i.e., error signal) leads to an internal model. Research shows that individuals with cerebellar pathology exhibit impairments in a range of sensorimotor tasks[16,17]. For example, in visuomotor rotation tasks, people with cerebellar ataxia (CA) exhibit a reduced ability to process experimental perturbations, where the participant’s predicted outcome is different from the perceived stimuli. This is consistent with the idea[6,18–20] that an intact cerebellum is required to process violation of expectations (VE). However, it is not fully understood how the concept of VE should be applied in cognitive domains[13,21].

Converging evidence indicates that the cerebellum’s function extends beyond mere motor control. The cerebellum is involved in many cognitive abilities, such as sequence learning, executive function, and even math and language processes[1,2,5,9,20,22–29]. This is evidenced by cognitive deficits in individuals with cerebellar pathology[2,16,21], bidirectional connectivity with the neocortex[26,27,30], developmental studies[31], and neuroimaging studies showing cerebellar activation in cognitive tasks[7,23,32]. While existing studies support a cerebellar role in nonmotor functions[1,5,18,31,33–35], direct evidence, and stronger theoretical constraints on the region’s specific function in higher cognition are needed[25]. In addition, despite the broad evidence in many cognitive domains, how this structure contributes to a fundamental, unifying cognitive mechanism remains an open question[8,36].

Furthermore, we observed that imaging and neuropsychological literature do not always converge, with the latter mainly showing mixed results. For instance, one language paradigm that activates the cerebellum is semantic processing, focusing on right cerebellar activation[8,19,24,37,38]. However, while some studies did not find that patients with cerebellar pathology have impairments in semantic processing tasks[39–41], others found significant impairment[19], specifically related to the processing of errors[18] or sequential processing[40]. Another example is working memory. The neuroimaging literature demonstrated the association of cerebellar activity with working memory tasks[7,42]. However, again, the results are not consistent in the neuropsychology literature. While some studies report impairments in individuals with cerebellar pathology on working memory tasks[43], others report null effects[21,44]. These mixed results may reflect the heterogeneity in the cerebellar patient samples, or it may be that the chosen cognitive tasks are not consistently sensitive to cerebellar pathology. To address these two challenges, we proposed to examine a homogenous sample of individuals with a single type of spinocerebellar ataxia (SCA6). Additionally, we suggested utilizing tasks designed to be more reliably sensitive in capturing the specific cognitive impairments associated with cerebellar dysfunction.

Several hypotheses have been proposed based on the idea that cerebellar contributions to motor control may extend to the cognitive domain[13,18,21]. For example, in one recent paper[21] it was hypothesized that the cerebellum supports dynamic continuous transformations of mental representations. However, the concrete implementation of 'continuity' in terms of higher mental representations remains an open question[45]. Currently, there is no established direct evidence supporting the existence of such 'continuous' and 'dynamic' higher cognitive processes, and, indeed, a recent later study by the same authors found no support for this hypothesis within the language domain (i.e., semantic processing task)[41]. It is unclear whether the utilized problems, such as simple addition or mental rotation, are entirely (or even partially) solved using “continuous transformation” or through previously learned procedural sequential knowledge of the required mental steps. The constraints on the cerebellum's role remain an open question, particularly whether its role is limited solely to tasks requiring "continuous mental transformation," as previously suggested[21].

Therefore, more consistent with well-established cerebellar models and parsimonious theories are still needed. As mentioned, in the motor domain, models of cerebellar function emphasize the importance of this structure in VE[46,47]. Expectations, which are based on top-down knowledge, are compared with visual feedback, with the difference between the two serving as an error signal[15]. However, given that violation of expectations is a general feature of brain function[48], a key challenge is to specify constraints on the cerebellum’s specific role.

When we examined previously used tasks, we noticed that many of them required procedure-based sequential processing[1,2,18,19,40,49–52]. People with CA show selective impairments in procedure-based tasks[2,16,49,53,54], requiring an algorithm of a sequence of discrete steps to solve[16,40,49,50]. While many studies explored whether the cerebellum contributes to language processes[1,22,38,55], the literature on the math domain is sparse. Recently, we revealed evidence for a distinct arithmetic impairment in a CA group[2,21]. The cerebellar group, compared to both neurotypical (NT) and Parkinson’s disease (PD) groups, exhibited a selective impairment in an addition procedure (i.e., counting, but not memory), requiring a sequence of discrete mental steps in order to solve each arithmetic problem[2] (e.g., 2 + 5 + 7 = 14 ?).

Herein, we proposed to investigate the cognitive mechanisms by which the cerebellum contributes to tasks requiring sequential processing: Symbolic subtraction, alphabet transformation, and grammar learning processes. Although defining the distinct cerebellum’s role in cognition is challenging[25,56], we anticipate that a successful theory will likely build on its well-established role in motor control[6,13–15,24].

Our logic is straightforward. VE is involved in almost every human function, from basic ones, such as motor control, to more complex ones, such as language and math. Evidence suggests that the cerebellum contributes to motor and cognitive tasks through the processing of VE[6,18,49,50,52,57]. Motivated by the biological and evolutionary principle of neural reuse[58,59], we propose and test the hypothesis that the cerebellum also contributes to higher cognition through cognitive VE. Evolutionarily novel cognitive abilities, such as arithmetic and language, also rely on the processing of VE[18,19,24,60,61]. For instance, based on known rules, a person predicts the correct answer to an equation, the next word in a sentence, or the next letter in a word. The expectations in these sequential processes are sometimes accurate and sometimes not. We suggest that through a VE mechanism, the cerebellum contributes not only to motor abilities but also to the cognitive procedures necessary for solving symbolic arithmetic and alphabet transformation problems[2,18,24].

Notably, these higher cognitive sequential problems are probably not solved by one continuous mental computation, and multi-step discrete computations are needed[45,62–64]. It is more likely that mental arithmetic and alphabet transformation problems are solved using discrete mental procedures by breaking down these complex cognitive computations into smaller steps[45,64–71]. For instance, when solving a simple subtraction problem (e.g., 9 - 5 = 3), a person typically follows a series of mental steps. For instance, first, they represent the stimuli mentally, understanding the need to subtract 5 from 9. They then retrieve relevant arithmetic facts from long-term memory or apply subtraction rules[65,72]. After arriving at a potential answer, they compare their expected computation (e.g., 4) with the actual visual information presented on the screen (e.g., 3). This algorithmic procedure highlights how sequential cognitive steps are used to solve even a simple arithmetic problem and how VE is a process needed for higher cognition as well.

Notably, this does not necessarily mean that the cerebellum involvement will increase when the need for a sequential procedure increases. Rather, we propose that the cerebellum role primarily emerges under the constraints of cognitive tasks that require sequential processing. In a sequential task requiring a series of discrete steps to solve, expectations are formed within each trial rather than between trials. During the operation of these mental steps, participants develop a prediction model (i.e., ‘internal model’), predicting the correct outcome. When an incorrect answer is presented, it leads to a violation of the participant's expectation within that specific trial. This process allows for the investigation of expectation formation, and its violation, within the context of a single trial, providing valuable insights into the underlying cognitive mechanism for each problem. Taken together, rather than being restricted to continuous transformations as suggested before[21], we propose that the cerebellum is necessary for solving sequential higher cognitive problems, which involve discrete multi-step procedures, via VE.

Hypothesis and goals

We hypothesized that the cerebellum is a generative structure – it generates predictions, and the deviation between the expected and perceived information could be a cerebellar-dependent problem-solving general mechanism. Rather than merely implementing previously learned cognitive processes (e.g., memory retrieval[2,49]), the cerebellum generates or adapts procedures (i.e., “internal model”) by processing VE in higher cognitive domains. The universal cerebellar transform theory[73] suggests that since the cerebellum is a relatively uniform structure, its function may also be uniform across domains[6,9,24,36]. Thus, we hypothesized that cerebellar degeneration would result in a core ubiquitous impairment in processing VE across complex cognitive tasks.

Given the scarcity of literature in this field, which remains inconclusive[2,18,24,39,41,44,61], we first aim to examine whether individuals with cerebellar degeneration are impaired in three high cognitive tasks, which require sequential processing. Second, we aim to assess the cerebellum's contribution, specifically in processing VE, compared to other cognitive processes (e.g., task complexity). Third, we asked whether the processing of VE could be a core, domain-general cerebellar-dependent mechanism across these cognitive tasks.

To investigate the cerebellum's role in processing error signals in high cognitive processes, specifically in arithmetic reasoning and alphabet transformation, we employed a neuropsychological approach, comparing the performance of individuals with CA to NT. Our study consisted of three experiments designed to measure participants' responses to correct and incorrect stimuli, focusing on the effect of errors on participant’s behavior.

In typical paradigms assessing cerebellar contributions to motor processes[17,74,75], deliberate incorrect information is displayed on a computer screen using visual perturbation. The incorrect information contradicts the expected (correct) visual input. This experimental manipulation aligns with the notion that an intact cerebellum is crucial in processing error signals. Similarly, an event-related potentials study[60] by Posner and others, utilized correct equations (e.g., 1 + 1 = 2), in which the presented solution (e.g., 2) is in accordance with the participant's expectation. However, in incorrect equations (e.g., 1 + 1 = 1), the presented solution (e.g., 1) violates the expected results and creates a cognitive VE. Thus, to maintain a comparable experimental paradigm in the cognitive domain, while having a point of comparison, we compared correct and incorrect cognitive problems. Comparing incorrect to correct problems allowed us to control for potential confounds, such as memory and perceptual processes[19,24,65,76].

In Experiments 1 and 2, we utilized a symbolic arithmetic verification task and an alphabet transformation task, respectively. Our goal was to assess whether CA modulates the processing of VE in two different tasks that require sequential processing: Subtraction reasoning and transformation of alphabet letters. Participants were required to identify whether a given subtraction equation or alphabet transformation problem was correct or incorrect. In each task, we manipulated two processes: the effect of errors on participants’ responses and, for comparison, the effect of task complexity (e.g., the number of mental steps). Comparing these two effects allowed us to assess potential confounds, such as task difficulty and perceptual processes[2]. To achieve this, we also varied the number of steps required to solve the problems (e.g., two-digit vs. three-digit problems). Given our hypothesis and previous findings[19], we predicted that the CA group would show a distinct disproportionate cost for VE (unexpected minus expected problems) but not for the complexity effect (high complexity minus low complexity problems).

While in Experiment 1 and 2 participants identify if a given problem is correct or not based on their previous knowledge, in Experiment 3, we probed novel cognitive expectations under uncertainty. We formed a new scenario where lifetime learned expectations are absent (no prior knowledge), and participants are required to learn a new artificial grammatical rule (i.e., organizing a sequence of letters in a specific order) within the task. The participants' expectations were novel and not based on previous top-down processes, allowing for minimizing memory effects. To do so, we utilized an artificial grammar learning (AGL) task[77,78]. During the training phase, participants were exposed to a sequence of letters that followed a predefined novel complex artificial grammar. In the subsequent test phase, participants needed to identify whether novel strings of letters were grammatical or nongrammatical (following the rule or not). Given previous studies demonstrating the cerebellum's role specifically in procedure-based learning[2,13,49,50,79], we predicted that the cerebellum would contribute to procedure-based learning[18,40] of an artificial grammar.

To assess the effect of uncertainty on participants' accuracy, we manipulated the level of similarity between grammatical and nongrammatical problems (i.e., sensitivity). We predicted that the CA group would exhibit selective impairment only in the low similarity condition, where there is greater certainty (higher probability), making it easier to differentiate between grammatical and nongrammatical problems.

Notably, given that processing VE is highly dependent on previous knowledge and top-down memory-related processes[19,23], it is important to consider these factors. Thus, across the three experiments, we examined VE in both established and newly learned cognitive procedures, considering varying levels of potential top-down effects. Furthermore, we aimed to determine whether the between-group differences in the cost of processing VE are domain-general in three cognitive tasks.

Results

Experiment 1 – Cerebellar contribution to symbolic subtraction.

In Experiment 1, we used manipulation to probe two major subtraction processes. We utilized a subtraction task because, compared to addition or multiplication that were used in previous studies[2,21], subtraction necessitates more spatial procedural processes[80–82], which are directly related to known cerebellar functions[83–85].

First, we manipulated the expectancy effect (expected vs. unexpected) of the equation. While at correct equations (e.g., 9 – 5 – 1 = 3), the presented solution (e.g., 3) is in accordance with the participant's expectation, at incorrect equation (e.g., 9 – 5 – 1 = 4), the presented solution (e.g., 4) violates the participant's expectations (see also[60]). Accordingly, response times (RT) are typically higher for incorrect problems compared to correct problems[63,65,86]. We computed the “expectancy effect” score by subtracting the mean RT of expected equations from the mean RT of unexpected equations. Second, we manipulated the complexity of the equation by probing the number of steps required to solve it. We employed problems that involved either one or two operators. Accordingly, subtracting one single-digit number (e.g., 9 - 5 = 4) takes less time than subtracting two single-digit numbers (9 - 3 - 2 = 4) because the latter requires more cognitive steps to solve the problem[2]. We computed the “complexity effect” score by subtracting the mean RT of low-complexity trials from that of high-complexity trials. RT was calculated on the participants' correct responses (>90% for both groups).

Participants completed a subtraction verification task (see Fig. 1). At the beginning of each trial, a fixation cross appeared in the center of the screen for 1000 ms. Then, the fixation cross was replaced by a display of an equation. The participant was required to determine whether a given subtraction equation was correct (by pressing the ‘Z’ key) or incorrect (by pressing the ‘M’ key). Three practice equations were presented before the experimental trials to familiarize participants with the procedure. The equation remained on the screen until a response was recorded or until 5 s’ had elapsed, whichever occurred first. Participants were instructed to respond as quickly and accurately as possible. Visual feedback was provided for 500 ms above the equation, with a green checkmark (√) or a red X indicating the accuracy detected.

In Experiment 1, we created 64 subtraction equations. To minimize the effect of memory, each equation only appeared once. To manipulate the expectancy effect, half (32) of the equations were correct, and half were incorrect. To manipulate complexity, half of the equations were of low complexity level, and the other half were of high complexity level. All the experimental conditions were counterbalanced and presented in a random order. In the middle of the task, the participant had a ten-second break. The experiment lasted approximately 15 minutes.

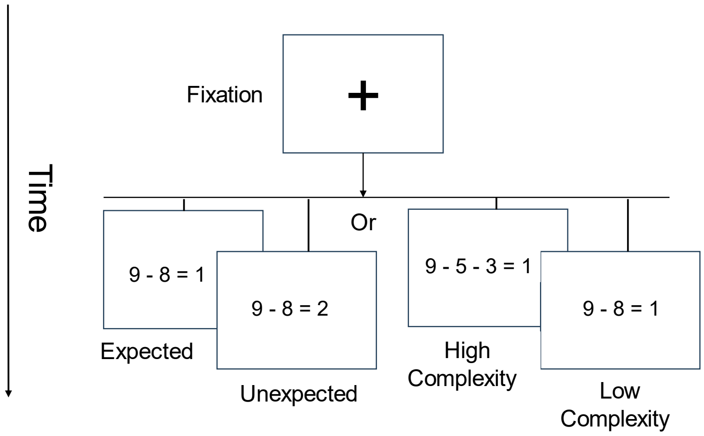

Fig. 1. Experiment 1 – Subtraction verification task.

Figure 2 shows RTs as a function of group (NT/CA), expectancy (expected/unexpected), and complexity (high/low). We utilized a linear mixed-effects (LME[87,88]) model with group, expectancy, and complexity as fixed effects, and participant ID as a random factor.

To establish a baseline, we assessed the NT group's performance. As predicted, the NT groups showed a significant expectancy effect (beta for simple effect estimator (est.) = 238 ms, p < .0001) and a significant complexity effect (est. = 831 ms, p < .0001). The CA group exhibited a quantitatively slower response than the NT group across all conditions, but this difference did not reach statistical significance (main effect est. = 304 ms, p = .120).

Our focus in Experiment 1 is comparing the effects of expectancy and complexity between the CA and NT groups (Fig. 2B). Notably, the expectancy effect was significantly larger for the CA group compared to the NT group (two-way interaction: est. = 160 ms, p = .013, cohen's d = 0.897; large effect size), but there was no significant difference between the groups in the complexity effect (two-way interaction: est. = 79 ms, p = .229). In terms of covariates, there were no significant differences between the groups in years of education and age (p > .05). In addition, accuracy rates were not significantly different between the groups (NT = 91%, CA = 89%, p > .05). Additionally, there was no significant interaction between Group and Expectancy when accuracy was the dependent variable (est. = 3.5, p = 0.61). This similar accuracy rates indicate that accuracy is not a sensitive measure for detecting group differences in this experiment.

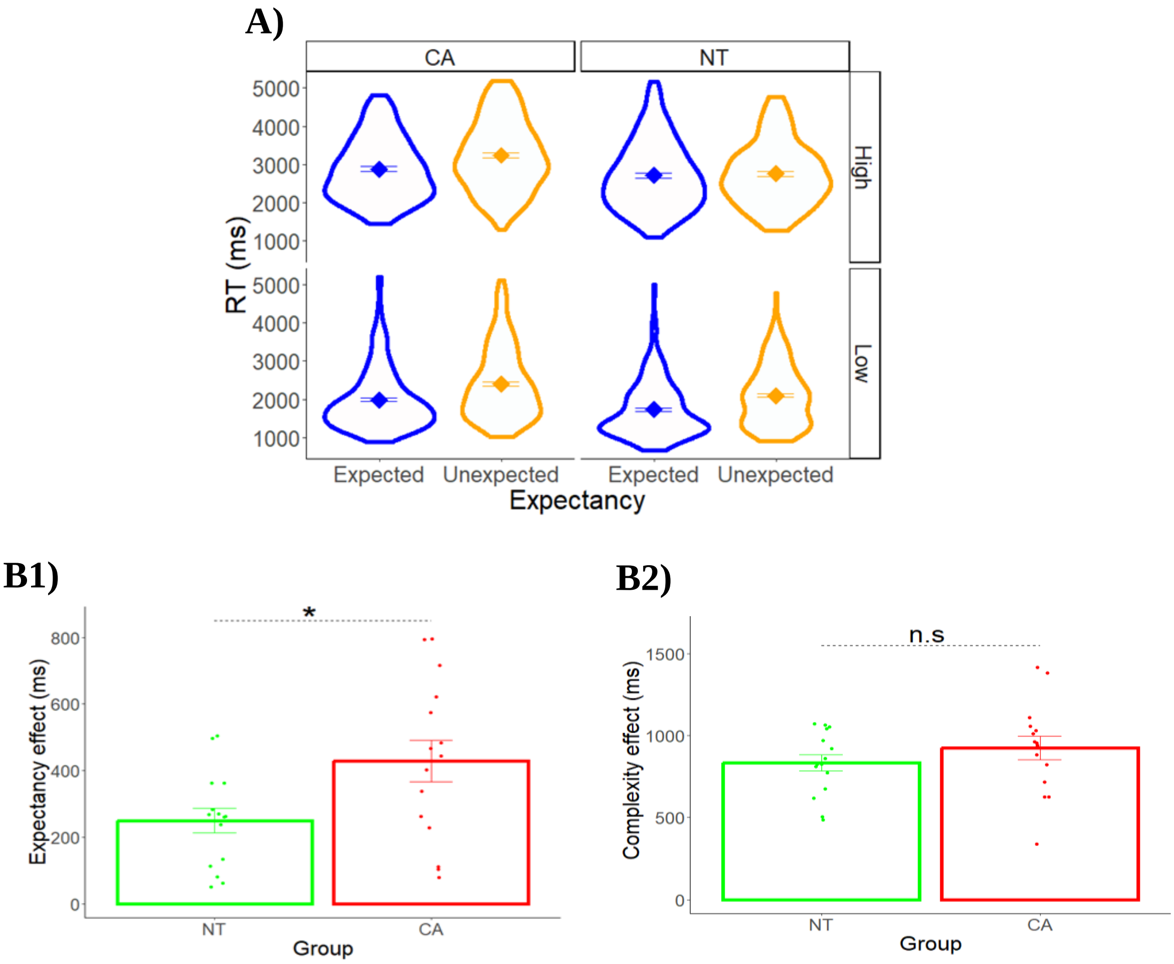

Fig. 2. A) RTs as a function of group (NT/CA), expectancy effect (unexpected minus expected), and complexity effect (high minus low). **B1)** The expectancy effect for each group. **B2)** The complexity effect for each group. Each datapoint is a participant. Error bars = SEM. * indicates p = .009. n.s = not significant.

To conclude, two primary insights can be derived from the results of Experiment 1. Firstly, the CA group demonstrated a deficiency in a symbolic subtraction task, providing novel evidence of the cerebellum's contribution to higher cognitive functions. Secondly, the findings indicate a selective impairment of the CA group compared to the NT group. The CA group exhibited a disproportionate expectancy effect but not a complexity effect. In line with the established role of the cerebellum in the motor domain, we propose that the selective impairment reflects the cerebellum's role in processing arithmetic VE.

Experiment 2 – Cerebellar contribution to alphabet transformation.

Participants completed an alphabet letter sequential transformation task. The design of this task was specifically tailored to achieve the following two objectives. First, we wanted to generalize the findings from arithmetic digits to alphabet letters. Participants were required to determine the correctness of an alphabet letter rule-based transformation. Similar to Experiment 1, the transformation rule was an arithmetic operator (e.g., +1), but the stimuli were alphabetic letters (e.g., A, B, C). This required participants to utilize their understanding of alphabetical sequences and relationships for each step in a discrete controlled manner. Second, we aimed to reduce top-down potential effects. While most educated participants have experience solving arithmetic equations, they probably have less experience transforming alphabetic letters using an arithmetic operator. Consequently, the alphabet transformation task reduces reliance on prior numerical proficiency. This allows for assessing VE in the context of relatively newly formed procedures.

In the alphabet transformation task, participants were presented with alphabet letters transformation based on an arithmetic operator (see Fig. 3). On the upper row, we presented letters (e.g., "C M") that needed to be transformed by a given rule in a discrete manner for each letter. In the second row, the transformation rule appeared (i.e., +1 or +2). In the third lower row, the transformation results appeared (e.g., "D N"). Participants were required to determine the correctness of the transformation displayed on the screen. Six practice problems were administered before the task to familiarize participants with the necessary procedure. At the onset of each experimental trial, a fixation cross appeared in the center of the screen for 1 s'. Then, the fixation cross was replaced by an alphabet transformation problem. The stimulus remained on the screen until a response was recorded or until 10 s' had elapsed, whichever occurred first. Participants were instructed to respond as quickly and accurately as possible. Visual feedback was presented for 1 s' above the equation, with a green checkmark (√) or a red X indicating the accuracy of the response. If a response was not detected within 9 s', participants received the feedback message "Respond faster." RT was calculated on the participants correct responses.

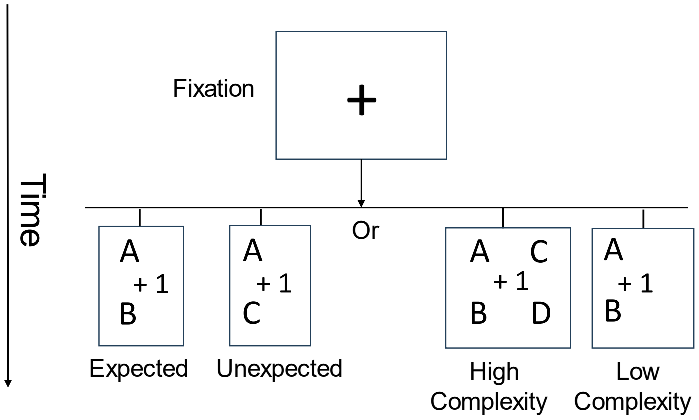

Fig. 3. Experiment 2 – Alphabet transformation task.

In Experiment 2, we created 96 unique letter transformation problems. The letters were chosen randomly, but each letter appeared only once in each problem. Similarly to Experiment 1, to minimize the effect of memory, each problem appeared only once. Again, we manipulated the expectancy effect by presenting expected or unexpected alphabet transformation problems (50% of the trials). We also manipulated the level of complexity: The problems either required transforming one letter (low complexity condition) or transforming two letters (high complexity condition). We calculated the complexity effect (complex minus simple problems). All the experimental conditions were counterbalanced and presented in a random order. The experiment lasted approximately 25 minutes.

Figure 4 shows RTs as a function of group (NT/CA), expectancy (expected/unexpected), and complexity (high/low). We utilized a linear mixed-effects (LME[87,88]) model with group, expectancy, and complexity as fixed effects, and participant ID as a random factor.

As predicted, the NT group responded slower to the unexpected than the expected alphabet transformation problems (the expectancy effect, est. = 478 ms, p < .0001). In addition, this group was slower to respond to the high complexity condition (two-letter problems) compared to the low complexity condition (the complexity effect, est. = 1173 ms, p < .0001). Across all trials, we observed that the CA group was significantly slower than the NT group (est. = 1470 ms, p < .0001).

Our focus in Experiment 2 is on the comparison of the effects of expectancy and complexity between groups (Fig. 4B). Notably, only the expectancy effect was significantly larger for the CA group compared to the NT group (expectancy effect: est. = 372 ms, p < .0001, cohen's d = 1.388 (large); complexity effect: est. = 108 ms, p = .270). In terms of covariates, there were no significant differences between the groups in years of education, MoCA, and age (p > .05). In addition, accuracy rates were not significantly different between the groups (NT = 75%, CA = 76%, p > .05). Additionally, there was no significant interaction between group and expectancy when accuracy was the dependent variable (est. = 4.0, p = 0.277). This similar accuracy rates indicate that accuracy is not a sensitive measure for detecting group differences in this experiment.

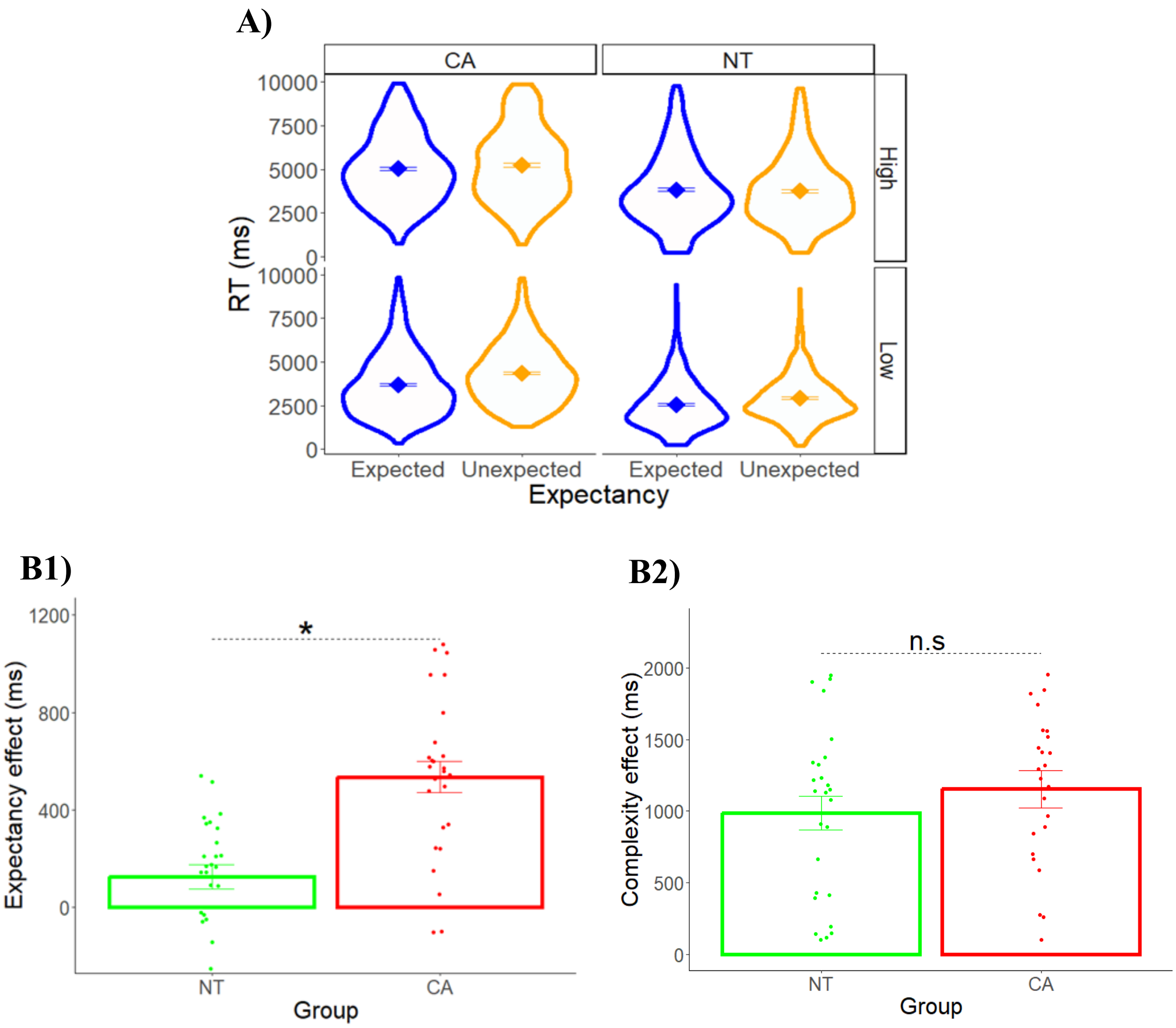

Fig. 4. **A)** RTs as a function of group (NT/CA), expectancy (unexpected/expected), and complexity (high/low). **B1)** The expectancy effect (unexpected minus expected) for each group. **B2)** The complexity effect (high minus low) for each group. Each datapoint is a participant. Error bars = SEM. * indicates p < .0001. n.s = not significant.

Consistent with Experiment 1, in Experiment 2, using the alphabet transformation task, the CA group showed a selective disproportionate expectancy effect, but not complexity effect. In Experiment 2, the required transformation was implemented on alphabetic letters, which allowed for the generalization of the results from arithmetic digits to language-related stimuli. Furthermore, transforming alphabetic letters by an arithmetic operator is a less commonly used cognitive operation than subtracting digits, as evidenced by the general increase in RT (~5,000ms vs. ~3,000ms). This allowed us to reduce the potential effects of top-down processes and show that CA distinct impairment also appears in these conditions.

Experiment 3 – Cerebellar contribution to cognitive expectations under uncertainty.

In Experiment 3, we aimed to test our hypothesis in novel scenarios involving cognitive expectations under uncertainty. We manipulated the participants' expectations by forming new predictions – learning new rules (i.e., grammar) without prior knowledge of the required procedures. Participants were first exposed to strings of letters generated according to a specific novel grammatical rule. During this training phase, they were not informed about the rules regarding how to organize a sequence of letters in a specific order. Next, they were shown new strings and asked to judge whether these strings followed the grammatical rules or not. We assessed learning by measuring how well participants could distinguish between grammatical and nongrammatical strings.

During the test stage, we assessed the participant's response to unexpected (nongrammatical) versus expected (grammatical) problems. As there are no expectations based on previous lifetime knowledge, the participant's expectations are formed during the task rather than before it. This approach allowed us to minimize the influence of previously learned top-down processes on newly learned grammar regarding a specific order of a sequence of letters.

We achieved this by using an artificial grammar learning (AGL) task[77]. In this task, during the training phase, we exposed participants to 23 strings of 2 to 6 letters (e.g., 'XVJ') four times each (92 training trials). In each trial, a fixation cross appeared for 500 ms, and participants were requested to type in the string (each string appeared for 3 s'). Next, in the test phase, participants were informed that the order of the letters in the previous strings was determined by a complex set of grammatical rules, without explicitly stating the rules. See Fig. 5 for the Markovian grammar chain used to produce the training and test strings. We then presented novel strings of letters (32 strings, each three times), and the participant was required to decide whether each new string was formed according to the grammatical rule or not. If the new sequence of letters followed the rule, it was considered "grammatical" (e.g., 'XVXJ'). Otherwise, it was considered "nongrammatical" (e.g., 'XVXT'). Given the results of the previous two experiments, complexity was not a factor of interest in this experimental design.

Notably, utilizing the AGL task, we also probed novel cognitive expectations under uncertainty. In the test phase, to assess the effect of uncertainty, we manipulated the level of similarity between grammatical and nongrammatical strings. In the low similarity condition, it was relatively easy for participants to distinguish between grammatical and nongrammatical strings. Participants were able, with a higher level of sensitivity, to differentiate which problems adhere to the expected grammar and which deviate from it (i.e., incorrect). In contrast, in the high similarity condition, participants' sensitivity was decreased. The discrimination between grammatical vs. nongrammatical strings was more challenging given the elevated degree of similarity between them. This manipulation also allowed us to disentangle between the difficulty in learning the rule during training and the difficulty in responding to the new strings during testing.

We presented strings that were utilized in previous studies[77]. For each test string, we calculated a similarity value, which is sometimes called a ‘chunk strength’ value. Similarity refers to the frequency of specific letter combinations (bigrams and trigrams) that participants have been exposed to during the training phase. Higher similarity indicates that certain letter combinations have been repeated more frequently during the training phase. This similarity value served as a measure of uncertainty since, in higher similarity levels, the sensitivity to grammatical versus nongrammatical is lower. Similarly to previous studies[77], we then categorized the test strings into four groups based on grammaticality and low- or high-similarity values: Grammatical high, grammatical low, nongrammatical high, and nongrammatical low.

Of the grammatical and nongrammatical test strings[77], there were an equal number of high-similarity (i.e., chunk strengths) and low-similarity items. The similarity was calculated as the average number of times each of the bigrams and trigrams in the string had been presented in the training set. The average similarity of high-similarity items was 8.5; the average of low-similarity items was 5.6.

To describe the magnitude of the sensitivity to grammatical status, we calculated a percent correct score for each participant in each condition. The analysis of accuracy not only provided insights into cognitive performance but also served a secondary benefit by reducing potential motor effects on RT. This approach helps to disentangle cognitive processes from motor execution, which is impacted by CA.

Typically, in experiments using the AGL task[77], the NT group is more accurate in grammatical strings compared to nongrammatical strings. In addition, this group demonstrates higher accuracy in low similarity conditions, where grammatical strings are likely more distinguishable from non-grammatical strings, compared to high similarity conditions. Accordingly, and to control potential response bias/motor-related abilities, a common dependent measure is *d’* (sensitivity) in discriminating between grammatical vs. nongrammatical strings. To investigate the expectancy effect under uncertainty conditions, our focus in Experiment 3 was on the comparison of the *d’* between groups in each level of similarity. Notably, when looking at *d’*, we predicted an interaction between group and similarity, such that the *d’* will be larger for the NT group compared to the CA group in the low similarity condition only.

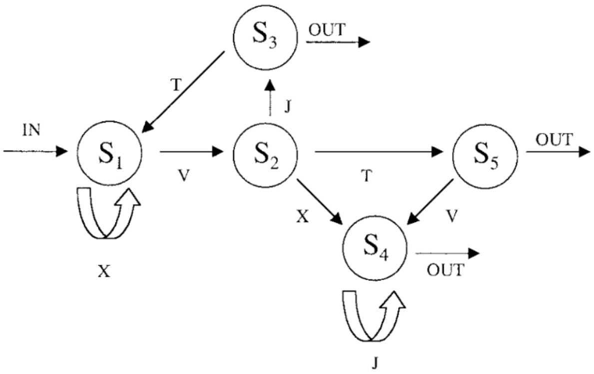

Fig. 5. The finite-state Markovian rule system used in Experiment 3. S1-S5 indicate the states that occur during the generation of grammatical letter strings. Letter strings are formed by starting at one entry and then by navigating from one transition to another, with each transition being able to generate a letter. A letter string terminates when an exit is reached.

Figure 6A shows accuracy as a function of group (NT/CA), expectancy (expected/unexpected), and similarity (high/low). We utilized an LME model with group, expectancy, and similarity as fixed effects, and participant ID as a random factor.

We observed no differences between the groups in mean accuracy across conditions (57.4%, est. = 4.02%, p = .310), indicating that potential specific group differences in this task cannot be fully explained by motor abilities. Then, turning to our main variables of interest, as predicted[61], the NT group was more accurate in grammatical strings compared to nongrammatical strings (error effect est. = 28.20%, p < 0.0001). In addition, the NT group was more accurate in the low similarity condition compared to the high similarity condition (est. = 11.33%, p = 0.024). Accordingly, and to control for potential response bias, we also found higher *d'* (sensitivity) in discriminating between grammatical vs. nongrammatical strings for the low similarity condition compared to the high similarity condition (est. = 0.635, p < 0.0001).

Our focus in Experiment 3 is on the comparison of the *d'* between groups in each level of similarity (Fig. 6B). Importantly, when looking at *d'*, we found an interaction between group and similarity (est. = 0.459, p = 0.022). As predicted, the *d'* was larger for the NT group compared to the CA group in the low similarity condition only (est. = 0.464, p < .0001, effect size = 1.106 (large); High similarity: est. = 0.004, p = .978).

In terms of covariates, there were no significant differences between the groups in years of education, MoCA, and age (p > .05). As expected, no significant group differences in RT were observed for this task (p = 0.549, NT = 2,568 ms, CA = 2,815 ms). Additionally, the interaction between group and expectancy was insignificant when RT was the depended vaibale (Low similarity: est. = 255, p = 0.310; High similarity: est. = 352, p = 0.180). Therefore, as commonly used in AGL tasks, we focused on accuracy measures.

We also conducted a criterion (c') analysis to examine potential effects of response bias. First, we did not find group differences in response bias (est. = 0.291, p = 0.210). Second, while we observed a two-way interaction between Similarity and Group when d' was the dependent variable, we did not find a significant interaction when we examined response bias (est. = 0.114, p = 0.262). This pattern of results supports the interpretation of reduced sensitivity in the CA group rather than the presence of a systematic response bias. Furthermore, we are unaware of any robust prior studies demonstrating a consistent response bias in the CA group.

Additionally, if the CA group had a bias to respond “no” or exhibited any other bias, such a bias should have been evident across different conditions. However, our inclusion of a control condition allowed us to examine this possibility. When we examined the High similarity condition, we did not find significant differences between the groups in d'. If the CA group exhibited a bias, one would expect significant performance differences from the control NT group across conditions and not only in the Low similarity condition. This alternative explanation, however, is not supported by the current empirical data.

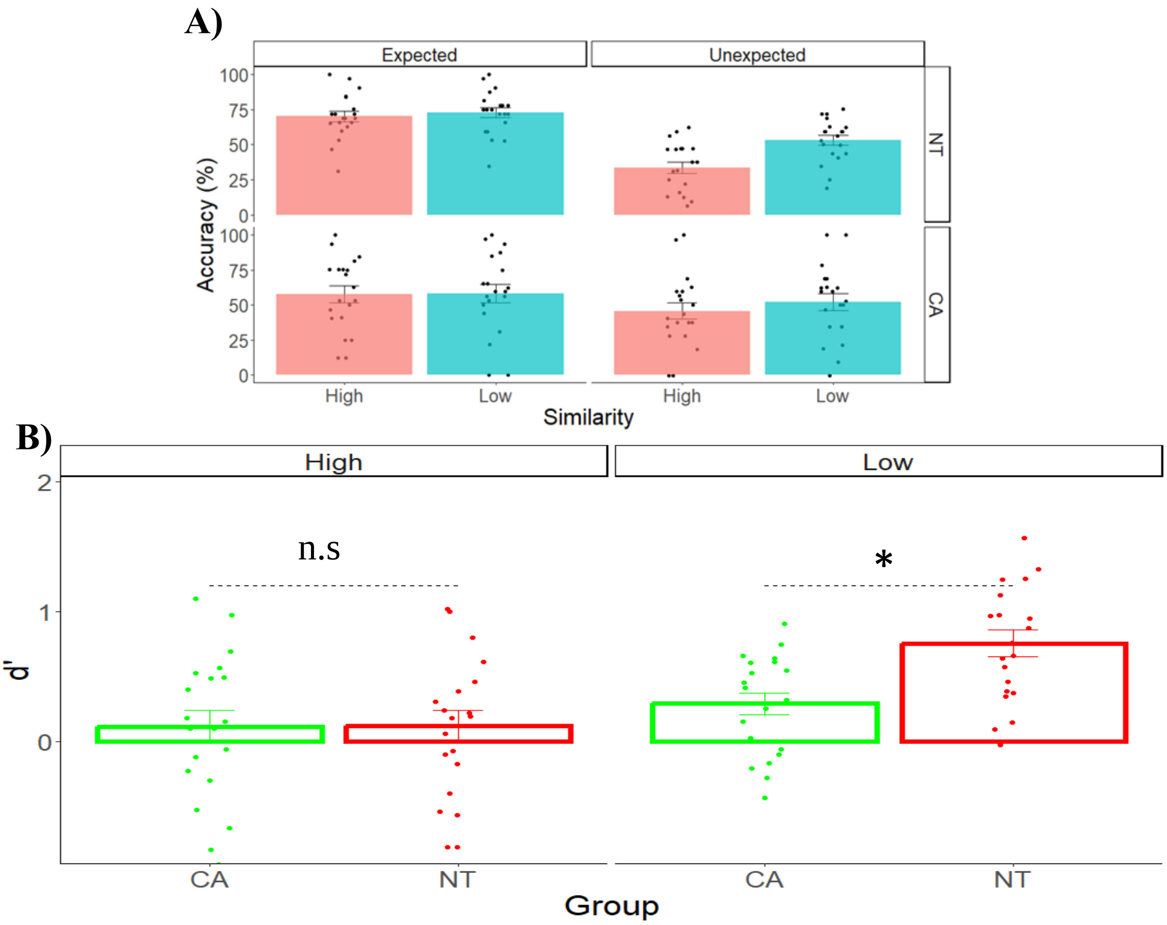

Fig. 6. A) Accuracy as a function of group (NT/CA), expectancy (expected/unexpected), and similarity (high/low). **B)** Sensitivity (d') as a function of group for each level of similarity (high/low). Each datapoint is a participant. Error bars = SEM. * indicates p < .0001. n.s = not significant.

Discussion

In three experiments, patients with cerebellar ataxia (CA) and neurotypical (NT) healthy participants solved symbolic arithmetic, alphabet transformation, and grammar sequential problems. These sequential problems required a series of discrete steps to be completed in a specific order to reach a solution. In Experiment 1, the CA group demonstrated impairment in a subtraction task, providing novel evidence of the cerebellum's contribution to symbolic arithmetic reasoning. In addition, we found a selective impairment of the CA group compared to the NT group. The CA group exhibited a disproportionate expectancy effect only (no difference in the complexity effect). Consistent with Experiment 1, in Experiment 2, using the alphabet transformation task, we found a distinct disproportionate expectancy effect (no difference in the complexity effect). Transforming alphabetic letters by an arithmetic operator is a less commonly used cognitive operation than subtracting digits. This allowed us to reduce the potential effects of top-down processes and show that CA's impairment also appears when there are fewer top-down effects. Rather, this probably increased the need to develop algorithmic procedures during the task.

While in Experiments 1 and 2, participants identify if a given problem is correct or not based on their previous knowledge, in Experiment 3, we probed novel cognitive VE under uncertainty. We formed a new scenario where prior lifetime knowledge is absent, and participants are required to learn a new grammatical rule within the task. The participants' expectations were novel and were not based on previous top-down processes. Utilizing an artificial grammar learning (AGL) task[77], participants needed to learn a new Markovian grammar, regarding how to organize a sequence of letters in a specific order. To assess the effect of the participant's uncertainty, we manipulated the level of similarity between grammatical and nongrammatical problems. We found that the CA group showed selective impairment only in the low similarity condition (higher sensitivity), probably where expectations are higher. Although it was not a causal finding and no behavioral evidence was found, an fMRI study[24] revealed that activity in the right posterolateral cerebellum correlated with the predictability of the upcoming target word. Together with our current findings and others[18,19], this pattern of results might indicate that the cerebellum is necessary for higher cognition through a VE mechanism.

Across the three experiments, we examined the effect of VE in both established and newly learned cognitive procedures, considering varying levels of potential top-down effects. The CA group showed a disproportionate expectancy effect compared to the NT group. We found that the between-group differences in the expectancy effect are consistent across tasks. Notably, the results indicate that CA patients had both intact processing of the problems' complexity (i.e., number of steps) and intact ability to discriminate between correct and incorrect problems when certainty decreased (i.e., sensitivity was lower). Thus, the results indicate a distinct role of the cerebellum in processing VE across these sequential cognitive tasks.

Several theories are in line with the principle of neural reuse and our hypothesis that the cerebellum contributes to many domains using the same core cognitive mechanism[25]. Paul Rozin[58] proposed that computations that initially evolved to solve specific problems become accessible to other systems through evolution as well as within the individual lifetime of an organism. Change or expansion of a function, because it is more generally available or accessible, "*would have adaptive value when an area of behavioral function could profit from programs initially developed for another purpose.*" This idea has been reframed and elaborated upon in Gallese's "neural exploitation" hypothesis[89] and Anderson's "massive redeployment" hypothesis[59,90]. The core idea is that neural networks can acquire new uses after establishing an initial function. Decades of empirical research from human and animal experiments, including our own, support this framework[6,25,91,92]. The broad involvement of the cerebellum in motor and nonmotor functions supports the idea of neural reuse, indicating this specific structure's potential ability to reuse its core function.

Importantly, while our findings point to a cerebellar unique and consistent role in VE processes in sequential tasks, we do not aim to generalize this role to all forms of expectations[18–20]. Notably, previous experimental paradigms aiming to assess violation of expectations utilized tasks that manipulated errors and/or probability-related processes, but did not fully dissociate these constructs. We propose that these previous experimental operational manipulations[17,19,40,93–95], applied in both motor and non-motor domains, may have inadvertently manipulated different cognitive constructs, such as statistical probability, error identification, or error correction. In our Experiments 1 and 2, we specifically manipulated error signals derived from previous top-down effects, while in Experiment 3, the participant's VE was derived from within-task probability processes. If feasible, we propose that future studies will disentangle different forms of expectations by operationalizing them in experimental tasks in an orthogonal manner. This will allow us to achieve a more well-defined cerebellar cognitive and motor mechanistic account.

Additionally, mathematical and linguistic tasks involve intricate cognitive processes. However, our study's scope is limited, and we do not propose that the cerebellum contributes to all mathematical or linguistic functions. These tasks depend on many neural networks beyond the cerebellum, which fall outside our current investigation. Our research aimed to better define a constraint on specific aspects of VE and the selective contributions of the cerebellum to higher cognition.

Our study is subject to two main limitations. First, information about anatomical-behavioral relationships can provide valuable data and will certainly be important in the long run for understanding how the cerebellum contributes to cognition. For example, one can ask if performance is related to gross measures such as total cerebellar volume or finer measures such as whether the observed deficits are associated with atrophy in particular regions. These analyses typically require large sample sizes, especially when dealing with atrophic processes (where the pathology tends to be relatively diffuse). Relatedly, we could not perform brain connectivity analysis, which limits our ability to examine the interactions between the cerebellum and other brain regions, such as the basal ganglia (BG) and the frontal lobe. This lack of brain connectivity data also restricts our ability to compare the cerebellum's functional contributions to those of other regions. For example, in a recently published paper (2024), we found that the BG plays a distinct role in mathematical complexity processes[2]. Unfortunately, we do not have imaging data for many of the patients.

Yet, there is added value in studies that include behavioral results from neurological groups defined based on clinical diagnosis. This is common in the literature, as seen in other recently (2019, 2022, 2024) published studies[2,21,96] in high-impact journals (PNAS, Journal of Neuroscience, Brain). In addition, the current experimental design is already complex, with two groups and three experiments, including control conditions. However, future work should use lesion analysis or connectivity methods to identify the specific anatomical-behavioral relationships critical for arithmetic and language operations.

Conclusion

Our findings support neural reuse hypotheses[58,90] in that the cerebellum contributes to both motor and non-motor functions using a similar mechanism[6,19,97]. The cerebellum not only contributes to motor control but also to cognitive procedures necessary for processing sequential arithmetic reasoning, alphabet transformation, and grammar problems using a similar mechanism – processing VE.

To conclude, theories of universal cerebellar transform propose that the cerebellum plays an essential role in modulating not only motor functions but also cognitive and affective processes[4,9]. The dysmetria of thought theory posits that cognitive and affective symptoms observed in cerebellar patients arise from the same dysfunction that affects motor control. This concept highlights the interconnectedness of motor, emotional, and cognitive domains, suggesting that impairments in one domain can reflect core problems in others. Our study provides convergent empirical evidence for the potential core role of the cerebellum and aligns with these previous theoretical frameworks.

Methods

Participants

The sample size was determined based on previous comparable work in CA patients. To calculate the required sample size, we conducted a power analysis (alpha = .05; power = .95) using effect sizes (Cohen's d = 1.36) derived from previous studies[2,21]. This analysis suggested a minimal sample size of 13 participants for each group. As such, the sample sizes of our groups (>14) had sufficient power to detect group differences.

For Experiment 1, 15 individuals with CA were tested along with 15 neurotypical (NT) healthy participants. For Experiment 2, 27 individuals with CA and 27 NT participants were tested. Three of these CA had participated in Experiment 1. Two participants (1/group) were excluded because they had accuracy scores below chance level (50%). For Experiment 3, 22 individuals with CA were tested along with 22 NT participants. Two of these CA had participated in Experiment 2. Four participants were excluded (2/group) based on a failure to respond correctly to the attention probes. Thus, in Experiments 1–3, 63 individuals with CA and 63 healthy age- and education-matched NT participants were included in the final analyses: 15/group in Experiment 1, 26/group in Experiment 2, and 20/group in Experiment 3.

The participants' demographic and medical information are presented in Table 1. Individuals with a previously established diagnosis of CA were recruited through our clinical Center for Accessible Neuropsychology (CAN) database. The CA groups included participants with a genetically confirmed SCA6, in which the pathology is relatively limited to the cerebellum[98]. Based on a prescreening interview, we included participants who were diagnosed with ataxia, had MRI evidence of cerebellar degeneration, and had genetic confirmation of SCA6. Individuals with other neurological conditions (not CA), psychiatric conditions, learning disabilities, and severe visual or auditory impairments were excluded from the study. The NT participants were recruited through the Prolific platform[99] for Experiment 1, and through our CAN database for Experiments 2 and 3, targeting participants that met the same demographic criteria as the CA participants in terms of age, years of education, gender distribution, and no general cognitive impairment. As previously mentioned, except in Experiment 1 (where MoCA was not collected), MoCA, age, and years of education were included as covariates in the primary analyses of all the experiments. All participants were above the age of 18 and were required to be able to understand and provide informed consent. The Tel Aviv University Institutional Review Board approved the protocol.

Neurological and neuropsychological assessment

We followed the online neuropsychological testing published protocol[16,41,100–102]. Individuals were invited by email to participate in an online, live interview with an experimenter. After providing informed consent, the participant completed a demographic and medical questionnaire. The trained experimenter then administered the Montreal Cognitive Assessment test (MoCA[100,101,103]) as a brief evaluation of cognitive status. The CA participants continued to the medical evaluation phase. First, the experimenter obtained the participant's medical history, collecting information about age at diagnosis, medication, primary symptoms, genetic subtype (based on molecular genetic testing), diet, other neurological or psychiatric conditions, and more relevant information. Second, the experimenter administered the Scale for Assessment and Rating of Ataxia (SARA[16,101,104]) as an evaluation of disease severity. The mean duration since diagnosis was 7 years in Experiment 1, 6.4 years in Experiment 2, and 6.8 years in Experiment 3. To avoid participants with general cognitive impairments, all CA participants have early to moderate ataxia severity (see SARA scores), providing further evidence for the absence of general cognitive impairments. This session took 40-60 minutes to complete.

Table 1. Demographic and medical summary table of the CA and NT groups in each Experiment. Mean ± SEM.

Group	Years of Education	# Of Females	Age	MoCA	SARA
Experiment 1					
NT	16.6 ± 0.8	10	58.0 ± 1.70		
CA	16 ± 0.7	13	60.6 ± 3.12	27.9 ± 0.47	13.6 ± 1.4
Experiment 2					
NT	16.7 ± 0.8	11	57.4 ± 2.62	27.1 ± 0.78	NA
CA	16.8 ± 0.7	12	57.1 ± 2.31	26.4 ± 0.81	12.2 ± 1.4
Experiment 3					
NT	16.1 ± 0.7	11	57.9 ± 2.43	27.3 ± 0.84	NA
CA	15.8 ± 0.6	10	58.1 ± 2.10	26.2 ± 1.36	14.2 ± 1.3

Procedure

The experiments were programmed in Gorilla Experiment Builder[105] and designed to be compatible with personal computers. Stimuli were presented at the center of the screen as black characters on a white screen. The actual size in terms of visual angle varied given that participants used their computer system, but we chose a font (7 HTML) that is clearly readable on all screens (as determined by pilots when developing the tasks). 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 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 was 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.

To maintain and check attentiveness, in each of the three experiments, we included three attention probes that appeared before, during, and after the experimental block (e.g., 'Do not press the Z key to continue; press the M key to continue'). At the end of the task, the participants were asked to provide feedback on their experience (e.g., 'How well were the study instructions explained?').

Acknowledgments

References

1. Strick, P. L., Dum, R. P. & Fiez, J. A. Cerebellum and Nonmotor Function. *Annu. Rev. Neurosci.* (2009). doi:10.1146/annurev.neuro.31.060407.125606
2. Saban, W., Pinheiro-Chagas, P., Borra, S. & Ivry, R. B. Distinct contributions of the cerebellum and basal ganglia to arithmetic procedures. *J. Neurosci.* (2024). doi:10.1523/JNEUROSCI.1482-22.2023
3. Middleton, F. A. & Strick, P. L. Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res. Rev.* **31**, 236–250 (2000).
4. Guell, X., Gabrieli, J. D. E. & Schmahmann, J. D. Embodied cognition and the cerebellum: Perspectives from the Dysmetria of Thought and the Universal Cerebellar Transform theories. *Cortex* **100**, 140–148 (2018).
5. Ravizza, S. M. *et al.* Cerebellar damage produces selective deficits in verbal working memory. *Brain* **129**, 306–320 (2006).
6. Hull, C. Prediction signals in the cerebellum: Beyond supervised motor learning. *Elife* **9**, 1–22 (2020).
7. Stoodley, C. J., Valera, E. M. & Schmahmann, J. D. Functional topography of the cerebellum for motor and cognitive tasks: An fMRI study. *Neuroimage* **59**, 1560–1570 (2012).
8. Stoodley, C. J. The cerebellum and cognition: evidence from functional imaging studies. *Cerebellum* **11**, 352–365 (2012).
9. Schmahmann, J. D. Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *J. Neuropsychiatry Clin. Neurosci.* **16**, 367–378 (2004).
10. Heffley, W. *et al.* Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nat. Neurosci.* **21**, 1431–1441 (2018).
11. Debas, K. *et al.* Brain plasticity related to the consolidation of motor sequence learning and motor adaptation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 17839–17844 (2010).
12. Imamizu, H. *et al.* Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* **403**, 192–195 (2000).
13. Ito, M. Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience* **9**, 304–313 (2008).
14. Miall, R. C. & Wolpert, D. M. Forward models for physiological motor control. *Neural Networks* **9**, 1265–1279 (1996).
15. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347 (1998).
16. Saban, W. & Ivry, R. B. Pont: A protocol for online neuropsychological testing. *J. Cogn. Neurosci.* **33**, 2413–2425 (2021).
17. Taylor, J. A., Krakauer, J. W. & Ivry, R. B. Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *J. Neurosci.* **34**, 3023–3032 (2014).
18. Fiez, J. A., Petersen, S. E., Cheney, M. K. & Raichle, M. E. IMPAIRED NON-MOTOR LEARNING AND ERROR DETECTION ASSOCIATED WITH CEREBELLAR DAMAGEA SINGLE CASE STUDY. *Brain* **115**, 155–178 (1992).
19. Moberget, T., Gullesen, E. H., Andersson, S., Ivry, R. B. & Endestad, T. Generalized role for the cerebellum in encoding internal models: Evidence from semantic processing. *J. Neurosci.* **34**, 2871–2878 (2014).

20. Sokolov, A. A., Miall, R. C. & Ivry, R. B. The Cerebellum: Adaptive Prediction for Movement and Cognition. *Trends Cogn. Sci.* **21**, 313–332 (2017).
21. McDougle, S. D. *et al.* Continuous manipulation of mental representations is compromised in cerebellar degeneration. *Brain* **145**, 4246–4263 (2022).
22. Argyropoulos, G. P. D. The cerebellum, internal models and prediction in ‘non-motor’ aspects of language: A critical review. *Brain Lang.* **161**, 4–17 (2016).
23. Lesage, E., Nailer, E. L. & Miall, R. C. Cerebellar BOLD signal during the acquisition of a new lexicon predicts its early consolidation. *Brain Lang.* **161**, 33–44 (2016).
24. Lesage, E., Hansen, P. C. & Miall, R. C. Right Lateral Cerebellum Represents Linguistic Predictability. *J. Neurosci.* **37**, 6231–6241 (2017).
25. Saban, W. & Gabay, S. Contributions of Lower Structures to Higher Cognition : Towards a Dynamic Network Model. *J. Intell.* **11**, (2023).
26. Caligiore, D. *et al.* Consensus Paper: Towards a Systems-Level View of Cerebellar Function: the Interplay Between Cerebellum, Basal Ganglia, and Cortex. *Cerebellum* **16**, 203–229 (2017).
27. Bostan, A. C. & Strick, P. L. The basal ganglia and the cerebellum: Nodes in an integrated network. *Nat. Rev. Neurosci.* **19**, 338–350 (2018).
28. BARTH, A. E. *et al.* A test of the cerebellar hypothesis of dyslexia in adequate and inadequate responders to reading intervention. *J. Int. Neuropsychol. Soc.* **16**, 526–536 (2010).
29. Nicolson, R. I., Fawcett, A. J., Brookes, R. L. & Needle, J. Procedural learning and dyslexia. *Dyslexia* **16**, 194–212 (2010).
30. Watson, T. C., Becker, N., Apps, R. & Jones, M. W. Back to front: Cerebellar connections and interactions with the prefrontal cortex. *Front. Syst. Neurosci.* **8**, 1–11 (2014).
31. Riva, D. & Giorgi, C. The cerebellum contributes to higher functions during development: evidence from a series of children surgically treated for posterior fossa tumours. *Brain* **123 (Pt 5)**, 1051–1061 (2000).
32. Walz, A. D. *et al.* Changes in cortical, cerebellar and basal ganglia representation after comprehensive long term unilateral hand motor training. *Behav. Brain Res.* **278**, 393–403 (2015).
33. Schmahmann, J. D. The cerebellum and cognition. *Neurosci. Lett.* **688**, 62–75 (2019).
34. Buckner, R. L. The Cerebellum and Cognitive Function: 25 Years of Insight from Anatomy and Neuroimaging. *Neuron* **80**, 807–815 (2013).
35. Nicholas, J. *et al.* The Role of the Cerebellum in Learning to Predict Reward: Evidence from Cerebellar Ataxia. *Cerebellum* **23**, 1355–1368 (2023).
36. Guell, X. *et al.* Dysmetria of Thought and the Universal Cerebellar Transform: empirical evidence, future approaches, and relevance for embodied cognition. (S18.007). *Neurology* **88**, (2017).
37. Fiez, J. A. *et al.* A positron emission tomography study of the short-term maintenance of verbal information. *J. Neurosci.* **16**, 808–822 (1996).
38. Murdoch, B. E. The cerebellum and language: historical perspective and review. *Cortex.* **46**, 858–868 (2010).
39. Gasparini, M. *et al.* Linguistic impairment after right cerebellar stroke: a case report. *Eur. J. Neurol.* **6**, 353–356 (1999).
40. Riva, D. The cerebellar contribution to language and sequential functions: evidence from a child with cerebellitis. *Cortex.* **34**, 279–287 (1998).

41. King, M., Bruinsma, S. & Ivry, R. B. No Evidence for Semantic Prediction Deficits in Individuals With Cerebellar Degeneration. *Neurobiol. Lang.* **5**, 635–651 (2024).
42. E, K. H., Chen, S. H. A., Ho, M. H. R. & Desmond, J. E. A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Hum. Brain Mapp.* **35**, 593–615 (2014).
43. Ziemus, B. *et al.* Impaired working-memory after cerebellar infarcts paralleled by changes in BOLD signal of a cortico-cerebellar circuit. *Neuropsychologia* **45**, 2016–2024 (2007).
44. Appollonio, I. M., Grafman, J., Schwartz, V., Massaquoi, S. & Hallett, M. Memory in patients with cerebellar degeneration. *Neurology* **43**, 1536–1544 (1993).
45. Dietrich, E. & Markman, A. B. Discrete Thoughts: Why Cognition Must Use Discrete Representations. *Mind Lang.* **18**, 95–119 (2003).
46. Miall, R. C., Christensen, L. O. D., Cain, O. & Stanley, J. Disruption of State Estimation in the Human Lateral Cerebellum. *PLOS Biol.* **5**, e316 (2007).
47. Wolpert, D. & Flanagan, J. Motor prediction. *Curr Biol* **11**, (2001).
48. Friston, K. Does predictive coding have a future? *Nat. Neurosci. 2018 218* **21**, 1019–1021 (2018).
49. Elyoseph, Z., Mintz, M., Vakil, E., Zaltzman, R. & Gordon, C. R. Selective Procedural Memory Impairment but Preserved Declarative Memory in Spinocerebellar Ataxia Type 3. *Cerebellum* **19**, 226–234 (2020).
50. Morgan, O. P. *et al.* The Cerebellum and Implicit Sequencing: Evidence from Cerebellar Ataxia. *Cerebellum* **20**, 222–245 (2021).
51. Pascual-Leone, A. *et al.* Procedural learning in Parkinson's disease and cerebellar degeneration. *Ann. Neurol.* **34**, 594–602 (1993).
52. Molinari, M. *et al.* Cerebellum and procedural learning: Evidence from focal cerebellar lesions. *Brain* **120**, 1753–1762 (1997).
53. Knowlton, B. J., Siegel, A. L. M. & Moody, T. D. 3.17 - Procedural Learning in Humans☆. in (ed. Byrne, J. H. B. T.-L. and M. A. C. R. (Second E.) 295–312 (Academic Press, 2017). doi:https://doi.org/10.1016/B978-0-12-809324-5.21085-7
54. Knowlton, B. J., Siegel, A. L. M. & Moody, T. D. *Procedural learning in humans*. *The Curated Reference Collection in Neuroscience and Biobehavioral Psychology* **3**, (Elsevier, 2016).
55. Adamaszek, M. & Kirkby, K. C. Cerebellum and Grammar Processing. in *The Linguistic Cerebellum* 81–105 (Elsevier Inc., 2016). doi:10.1016/B978-0-12-801608-4.00005-0
56. Parvizi, J. Corticocentric myopia: old bias in new cognitive sciences. *Trends Cogn. Sci.* **13**, 354–359 (2009).
57. Saywell, N. & Taylor, D. The role of the cerebellum in procedural learning - Are there implications for physiotherapists' clinical practice? *Physiother. Theory Pract.* **24**, 321–328 (2008).
58. Rozin, P. The Selection of Foods by Rats, Humans, and Other Animals. *Adv. Study Behav.* (1976). doi:10.1016/S0065-3454(08)60081-9
59. Anderson, M. L. Neural reuse: A fundamental organizational principle of the brain. *Behav. Brain Sci.* **33**, 245–266 (2010).
60. Berger, A., Tzur, G. & Posner, M. I. Infant brains detect arithmetic errors. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12649–12653 (2006).
61. Argyropoulos, G. P. D. The cerebellum, internal models and prediction in 'non-

motor' aspects of language: A critical review. *Brain Lang.* **161**, 4–17 (2016).
62. Logan, G. D. & Klapp, S. T. Automatizing Alphabet Arithmetic: I. Is Extended Practice Necessary to Produce Automaticity? *J. Exp. Psychol. Learn. Mem. Cogn.* **17**, 179–195 (1991).
63. Ashcraft, M. H. Cognitive arithmetic: A review of data and theory. *Cognition* **44**, 75–106 (1992).
64. Verhaeghen, P., Kliegl, R. & Mayr, U. Sequential and coordinative complexity in time-accuracy functions for mental arithmetic. *Psychol. Aging* **12**, 555–564 (1997).
65. Cohen-Kadosh, R. & Dowker, A. The Oxford handbook of numerical cognition. in *The Oxford handbook of numerical cognition* (Oxford University Press., 2015).
66. LeFevre, J. A., Sadesky, G. S. & Bisanz, J. Selection of procedures in mental addition: Reassessing the problem size effect in adults. *J. Exp. Psychol. Learn. Mem. Cogn.* **22**, 216–230 (1996).
67. DeStefano, D. & LeFevre, J. A. The role of working memory in mental arithmetic. *Eur. J. Cogn. Psychol.* **16**, 353–386 (2004).
68. Grabner, R. H. & De Smedt, B. Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biol. Psychol.* **87**, 128–136 (2011).
69. Taatgen, N. a. The nature and transfer of cognitive skills. *Psychol. Rev.* **120**, 439–71 (2013).
70. Fias, W., Sahan, M. I., Ansari, D. & Lyons, I. M. From counting to retrieving: Neural networks underlying alphabet arithmetic learning. *J. Cogn. Neurosci.* **34**, 16–33 (2021).
71. tiberghien, K., De Smedt, B., Fias, W. & Lyons, I. M. Distinguishing between cognitive explanations of the problem size effect in mental arithmetic via representational similarity analysis of fMRI data. *Neuropsychologia* **132**, 107120 (2019).
72. Grabner, R. H. *et al.* To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia* **47**, 604–608 (2009).
73. Schmahmann, J. D., Guell, X., Stoodley, C. J. & Halko, M. A. The Theory and Neuroscience of Cerebellar Cognition. *Annu. Rev. Neurosci.* **42**, 337–364 (2019).
74. Tzvi, E. *et al.* Cerebellar degeneration affects cortico-cortical connectivity in motor learning networks. *NeuroImage Clin.* **16**, 66–78 (2017).
75. Bares, M. *et al.* Impaired predictive motor timing in patients with cerebellar disorders. *Exp. Brain Res.* **180**, 355–365 (2007).
76. Ashcraft, M. H. & Battaglia, J. Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *J. Exp. Psychol. Hum. Learn. Mem.* **4**, 527–538 (1978).
77. Chang, G. Y. & Knowlton, B. J. Visual feature learning in artificial grammar classification. *J. Exp. Psychol. Learn. Mem. Cogn.* **30**, 714–722 (2004).
78. Pothos, E. M. Theories of artificial grammar learning. *Psychol. Bull.* **133**, 227–244 (2007).
79. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* **10**, 732–739 (2000).
80. Ward, J., Sagiv, N. & Butterworth, B. The impact of visuo-spatial number forms on simple arithmetic. *Cortex* **45**, 1261–1265 (2009).
81. Dormal, V., Schuller, A. M., Nihoul, J., Pesenti, M. & Andres, M. Causal role of

spatial attention in arithmetic problem solving: Evidence from left unilateral neglect. *Neuropsychologia* **60**, 1–9 (2014).
82. Díaz-Barriga Yáñez, A. *et al.* Learning to run the number line: the development of attentional shifts during single-digit arithmetic. *Ann. N. Y. Acad. Sci.* **1477**, 79–90 (2020).
83. Mandolesi, L., Leggio, M. G., Spirito, F. & Petrosini, L. Cerebellar contribution to spatial event processing: do spatial procedures contribute to formation of spatial declarative knowledge? *Eur. J. Neurosci.* **18**, 2618–2626 (2003).
84. Leggio, M. G. *et al.* Representation of actions in rats: the role of cerebellum in learning spatial performances by observation. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 2320–2325 (2000).
85. Mandolesi, L., Leggio, M. G., Graziano, A., Neri, P. & Petrosini, L. Cerebellar contribution to spatial event processing: involvement in procedural and working memory components. *Eur. J. Neurosci.* **14**, 2011–2022 (2001).
86. Saban, W., Sklar, A. Y., Hassin, R. R. & Gabay, S. Ancient visual channels have a causal role in arithmetic calculations. *Sci. Rep.* **11**, (2021).
87. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, (2015).
88. Figner, B. *et al.* Standard Operating Procedures For Using Mixed-Effects Models. 1–24 (2020).
89. Gallese, V. & Cuccio, V. The neural exploitation hypothesis and its implications for an embodied approach to language and cognition: Insights from the study of action verbs processing and motor disorders in Parkinson's disease. *Cortex* (2018). doi:10.1016/j.cortex.2018.01.010
90. Anderson, M. L. The massive redeployment hypothesis and the functional topography of the brain. *Philos. Psychol.* (2007). doi:10.1080/09515080701197163
91. Saban, W., Raz, G., Grabner, R. H., Gabay, S. & Kadosh, R. C. Primitive visual channels have a causal role in cognitive transfer. *Sci. Rep.* **11**, 8759 (2021).
92. Balsters, J. H., Whelan, C. D., Robertson, I. H. & Ramnani, N. Cerebellum and Cognition: Evidence for the Encoding of Higher Order Rules. *Cereb. Cortex* **23**, 1433–1443 (2013).
93. Butcher, P. A. *et al.* The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. *J. Neurophysiol.* **118**, 1622–1636 (2017).
94. Sokolov, A. A., Miall, R. C. & Ivry, R. B. The Cerebellum: Adaptive Prediction for Movement and Cognition. *Trends Cogn. Sci.* **21**, 313–332 (2017).
95. Fiez, J. A., Petersen, S. E., Cheney, M. K. & Raichle, M. E. Impaired non-motor learning and error detection associated with cerebellar damage. A single case study. *Brain* **115 Pt 1**, 155–178 (1992).
96. Breska, A. & Ivry, R. B. Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. (2019). doi:10.1073/pnas.1810596115
97. Diedrichsen, J., King, M., Hernandez-Castillo, C., Sereno, M. & Ivry, R. B. Universal Transform or Multiple Functionality? Understanding the Contribution of the Human Cerebellum across Task Domains. *Neuron* **102**, 918–928 (2019).
98. Klockgether, T., Mariotti, C. & Paulson, H. L. Spinocerebellar ataxia. *Nat. Rev. Dis. Prim.* **5**, 1–21 (2019).
99. Palan, S. & Schitter, C. Prolific.ac—A subject pool for online experiments. *J.*

Behav. Exp. Financ. **17**, 22–27 (2018).

100. Binoy, S., Monstaser-Kouhsari, L., Ponger, P. & Saban, W. Remote Assessment of Cognition in Parkinsons Disease and Cerebellar Ataxia: The MoCA Test in English and Hebrew. *Front. Hum. Neurosci.* **17**, (2023).
101. Picciotto, Y. De, Algon, A. L., Amit, I., Vakil, E. & Saban, W. Large-scale evidence for the validity of remote MoCA administration among people with cerebellar ataxia administration among people with cerebellar ataxia. *Clin. Neuropsychol.* **0**, 1–17 (2024).
102. Tsay, J. S. *et al.* Large-scale citizen science reveals predictors of sensorimotor adaptation. *Nat. Hum. Behav.* (2024). doi:10.1038/s41562-023-01798-0
103. Nasreddine, Z. S. *et al.* The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).
104. Schmitz-Hübsch, T. *et al.* Scale for the assessment and rating of ataxia: Development of a new clinical scale. *Neurology* **66**, 1717–1720 (2006).
105. Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. & Evershed, J. K. Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Methods* **52**, 388–407 (2020).