

## REWARD LEARNING IN THE CEREBELLUM

### 1      The role of the cerebellum in learning to predict reward: evidence from 2      cerebellar ataxia

3      Jonathan Nicholas<sup>1,2</sup>, Christian Amlang<sup>3,4</sup>, Chi-Ying R. Lin<sup>5</sup>, Leila Montaser-Kouhsari<sup>6</sup>, Natasha  
4      Desai<sup>3,4</sup>, Ming-Kai Pan<sup>7,8,9</sup>, Sheng-Han Kuo<sup>3,4,\*</sup>, Daphna Shohamy<sup>1,2,10,\*</sup>

5      1. Department of Psychology, Columbia University

6      2. Zuckerman Mind Brain Behavior Institute, Columbia University

7      3. Department of Neurology, Columbia University Medical Center

8      4. Initiative for Columbia Ataxia and Tremor, Columbia University Medical Center

9      5. Department of Neurology, Baylor College of Medicine

10     6. Department of Neurology, Stanford University School of Medicine

11     7. Department of Medical Research, National Taiwan University Hospital

12     8. Department and Graduate Institute of Pharmacology, National Taiwan University College of  
13     Medicine

14     9. Cerebellar research center, National Taiwan University Hospital, Yun-Lin Branch

15     10. Kavli Institute for Brain Science, Columbia University, New York, NY, USA

16     Corresponding Authors:

17     Daphna Shohamy

18     Columbia University Zuckerman Institute Quad 3D, 3227 Broadway, New York, NY 10027

19     ds2619@columbia.edu

21     Sheng-Han Kuo

22     650 W. 168<sup>th</sup> St, Rm 305, New York, NY, 10032

23     sk3295@cumc.columbia.edu

25     Acknowledgements:

26     J.N. was supported by the NSF Graduate Research Fellowship (1644869). S.H.K. was supported  
27     by NINDS R01NS104423, NINDS R01 NS118179, NINDS R01 NS124854, and National Ataxia  
28     Foundation. D.S. was supported by an NSF CRCNS award (1822619), NIMH R01 MH121093  
29     and the Kavli Foundation.

30     The authors declare no conflicts of interest.

31

## REWARD LEARNING IN THE CEREBELLUM

### 32 Abstract

33 Recent findings in animals have challenged the traditional view of the cerebellum solely as the site  
34 of motor control, suggesting that the cerebellum may also be important for learning to predict  
35 reward from trial-and-error feedback. Yet, evidence for the role of the cerebellum in reward  
36 learning in humans is lacking. Moreover, open questions remain about which specific aspects of  
37 reward learning the cerebellum may contribute to. Here we address this gap through an  
38 investigation of multiple forms of reward learning in individuals with cerebellum dysfunction,  
39 represented by cerebellar ataxia cases. Nineteen participants with cerebellar ataxia and 57 age- and  
40 sex-matched healthy controls completed two separate tasks that required learning about reward  
41 contingencies from trial-and-error. To probe the selectivity of reward learning processes, the tasks  
42 differed in their underlying structure: while one task measured incremental reward learning ability  
43 alone, the other allowed participants to use an alternative learning strategy based on episodic  
44 memory alongside incremental reward learning. We found that individuals with cerebellar ataxia  
45 were profoundly impaired at reward learning from trial-and-error feedback on both tasks, but  
46 retained the ability to learn to predict reward based on episodic memory. These findings provide  
47 evidence from humans for a specific and necessary role for the cerebellum in incremental learning  
48 of reward associations based on reinforcement. More broadly, the findings suggest that alongside  
49 its role in motor learning, the cerebellum likely operates in concert with the basal ganglia to support  
50 reinforcement learning from reward.

51 **Keywords:** reward; reinforcement learning; cerebellum; ataxia

## REWARD LEARNING IN THE CEREBELLUM

### 52 Introduction

53 It is well established that the cerebellum is required for refining movement through supervised  
54 motor learning.<sup>1–4</sup> The cerebellum receives error signals from climbing fiber input which then  
55 alters Purkinje cell plasticity to adapt motor behavior in service of minimizing future error.<sup>5–7</sup>  
56 However, recent findings have challenged the notion that the cerebellum is solely responsible for  
57 supervised learning of motor behavior and instead suggest that the cerebellum may also be  
58 involved in the processing of reward more generally<sup>8–19</sup>. In particular, climbing fiber inputs to the  
59 cerebellum encode expected reward<sup>13,15,17,19</sup>, and cerebellar Purkinje cells have been found to  
60 report reward-based prediction errors<sup>11,12,18</sup>. These signals are essential ingredients for  
61 reinforcement learning, or learning that allows an organism to determine from trial-and-error  
62 feedback which actions should be taken in order to maximize future expected reward. The presence  
63 of reward-related processing in the cerebellum suggests that it may play a role in reinforcement  
64 learning alongside its capacity for supervised motor learning<sup>20</sup>. This proposal challenges not only  
65 our current understanding of cerebellar function, but also our understanding of how the brain learns  
66 from reward more broadly<sup>5,21</sup>.

67 Although research on the cerebellum’s function in reward learning is growing, the vast majority  
68 of work has been done in animal models<sup>10–19</sup>, and evidence in humans remains limited. Human  
69 neuroimaging studies have revealed correlational evidence that the cerebellum is involved in tasks  
70 unrelated to movement<sup>22</sup>, however, despite some reports of BOLD activity in the cerebellum in  
71 response to reward across several early imaging studies<sup>23–25</sup>, more direct investigations of the role  
72 of cerebellum in reward-related behaviors in humans are lacking. The aim of the present study was  
73 to fill this gap by testing whether individuals with damage to the cerebellum, as occurs in cerebellar  
74 ataxia (CA), are impaired in their ability to acquire stimulus-reward associations.

75 Our study builds upon a rich literature focused on learning about reward from trial-and-error  
76 feedback. This process has been studied extensively using models of incremental learning, which  
77 rely on error-driven rules that summarize experiences with a running average<sup>26–28</sup>. During reward  
78 learning of this type, an agent uses the outcome of a recent decision to associate some stimulus  
79 with an action. Following successful learning, actions that are more likely to be rewarded are more

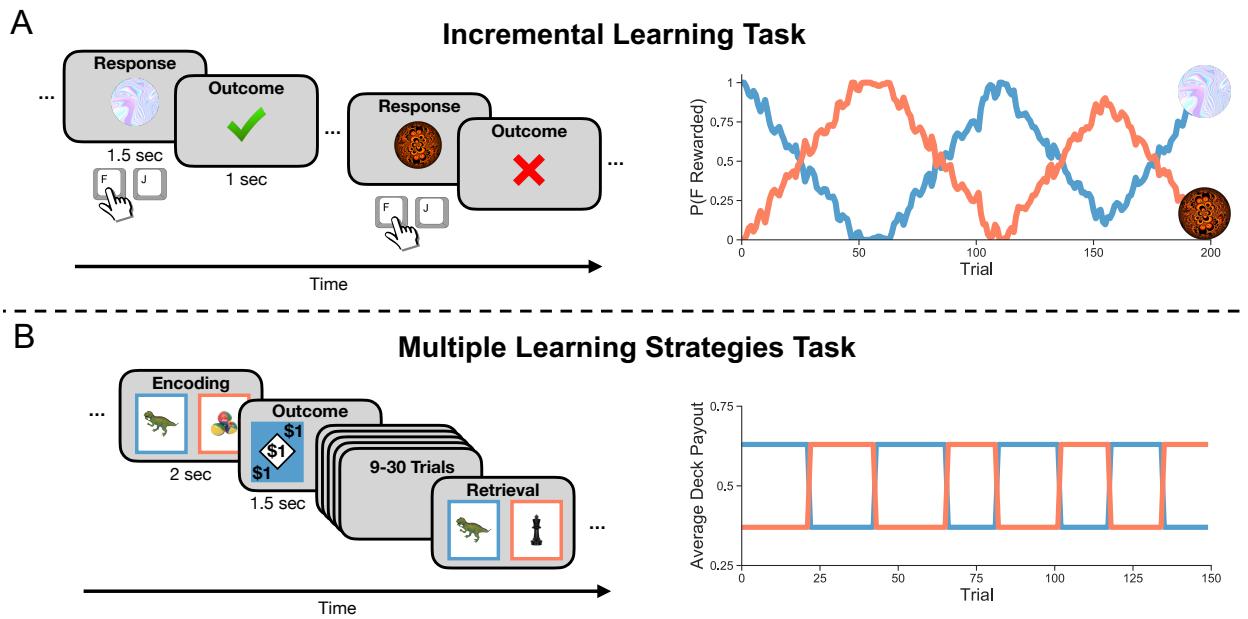
## REWARD LEARNING IN THE CEREBELLUM

likely to be repeated. This simple mechanism has been evoked to explain conditioning behavior and is well-captured by reward prediction error signals in midbrain dopamine neurons that project to the striatum<sup>27,29</sup>. This error signal is also precisely what has been implicated in recent animal models of cerebellar contributions to reward learning<sup>9</sup>, suggesting an additional, albeit unclear, role for the cerebellum in this process. Whether these cerebellar contributions are actually needed for successful incremental reward learning in humans is at present unknown.

To answer this question, we asked individuals with CA to complete a series of tasks that required them to learn associations between stimuli from trial-and-error feedback in order to maximize expected reward. CA is defined as a lack of coordination caused by disorders that affect cerebellar function<sup>30</sup>. A large variety of conditions can cause CA, ranging from immune-mediated disease to genetic and neurodegenerative disorders. Given the presence of cerebellar dysfunction in CA cases, studying individuals with CA is a common method used to investigate the necessary physiological functions of the cerebellum in humans.

Nineteen individuals with CA and 57 age- and sex-matched healthy controls (HC) completed two tasks (**Figure 1**). The first, referred to throughout as the *incremental learning* task, allowed us to measure each participants' ability to learn about reward incrementally. This task was motivated by recent work using a similar simplified paradigm to investigate cerebellar-based incremental learning in non-human primates<sup>10,11</sup>. The second task, referred to throughout as the *multiple learning strategies* task, allowed us to measure whether any impairments were specific to incremental learning alone. In the multiple learning strategies task, learning about reward can be supported by an alternative strategy based on episodic memory for trial-unique past outcomes. Healthy adults readily use of both of these strategies in this task<sup>31,32</sup>. We hypothesized that cerebellar dysfunction would lead specifically to impaired incremental reward learning relative to healthy controls.

## REWARD LEARNING IN THE CEREBELLUM



104

105 **Figure 1. Design of the incremental learning and multiple learning strategies tasks.** (A) Left: Trial  
106 design for the incremental learning task. Participants saw one of two fractal cues on the screen and were  
107 required to press either the F key with their left hand or the J key with their right hand. Following their  
108 choice, they received binary probabilistic feedback about whether they were correct or not. Right: Drifting  
109 cue-response-reward contingencies over the course of the incremental learning task. The probability that  
110 the F key is rewarded is shown for each cue in blue and orange. (B) Left: Trial design for the multiple  
111 learning strategies task. Participants chose between two decks of cards (one blue and one orange) and  
112 received an outcome between \$0-\$1 in intervals of 20 cents. Each card featured a trial-unique object that  
113 could repeat once every 9-30 trials. Participants were told that if they saw the same card again, it would be  
114 worth the same amount as the first time that it appeared. Right: An example of how average deck value  
115 reversed throughout the course of the multiple learning strategies task.

116

## Results

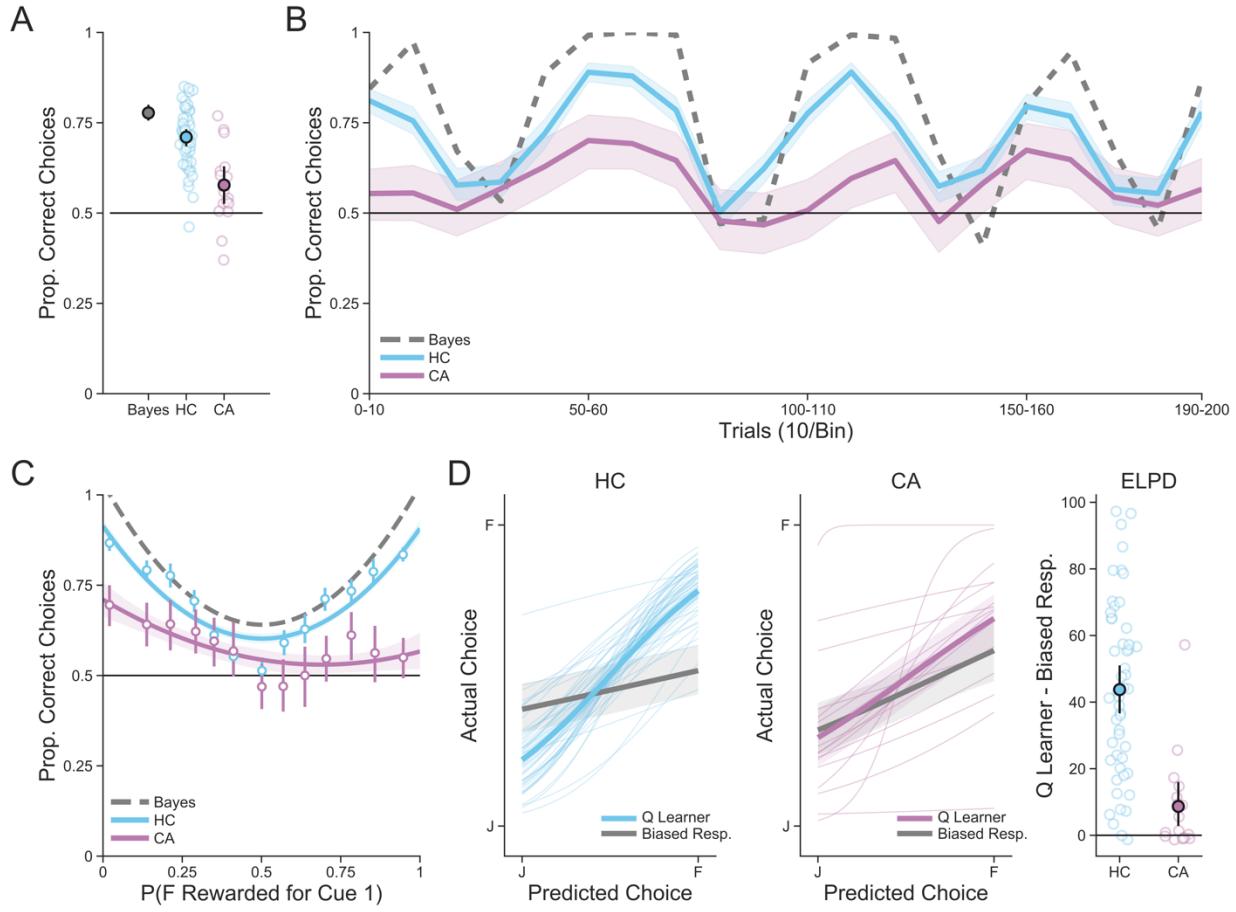
117

### Impaired reward learning in the incremental learning task

118 Our first goal was to assess CA participants' baseline ability to learn incrementally from reward  
119 using the incremental learning task. On this task, CA participants made overall fewer correct  
120 choices compared to healthy controls ( $\beta_{Group} = -0.88$ , 95% CI = [-1.55, -0.144]; **Figure**  
121 **2A**). CA participants' choices were less correct throughout the entirety of the task, even during  
122 periods of learning where action-outcome contingencies were more deterministic (e.g. close to  
123 100%) compared to more difficult periods of learning ( $\beta_{Group \times pFReward_1^2} = -5.49$ , 95% CI =  
124 [-7.57, -3.52]; **Figure 2B-C**). Overall, this difference in performance indicates that CA

## REWARD LEARNING IN THE CEREBELLUM

participants did not learn from reward feedback. Although CA participants responded slightly more slowly than healthy controls on this task ( $\beta_{Group} = -115.81$ , 95% CI = [-201.26, -33.55]), we included reaction times as a covariate in the above regression analysis to ensure that differences in choice accuracy were not attributed to motor slowing in CA participants.



**Figure 2. Performance on the incremental learning task.** (A) Performance on the incremental learning task averaged across all trials for healthy controls (HC) and CA participants compared to a Bayesian observer in gray, which represents normative performance on the task. Individual points are averages for each subject and filled in points represent group-level averages. Error bars are 95% confidence intervals. (B) Performance on the incremental learning task over time. Each timepoint represents ten trials. Lines are group averages and bands are 95% confidence intervals. For normative comparison, the performance of the Bayesian observer is shown as a dotted gray line. (C) Performance on the incremental learning task as a function of task difficulty, which is indexed by the true underlying probability that pressing the F key was the correct response (>50%) on each trial. Points represent group level averages from 13 bins with an equal number of trials, lines represent the fit of a second-order linear model, and error bars and bands represent 95% confidence intervals. (D) Model performance of the Q Learner and baseline Biased Responder models. Left: Posterior predictive performance. Individual lines represent Q learner fits for each individual, whereas thick lines represent the group-level average fit (with the Q Learner in color and Biased Responder in gray). Bands represent 95% confidence intervals. Right: The difference in estimated out-of-sample predictive

## REWARD LEARNING IN THE CEREBELLUM

144 performance (as measured by expected log pointwise predictive density; ELPD) between the Q Learner and  
145 the Biased Responder model for each group. Individual points are the ELPD difference for each subject and  
146 filled in points represent group-level averages. Error bars are 95% confidence intervals.

147

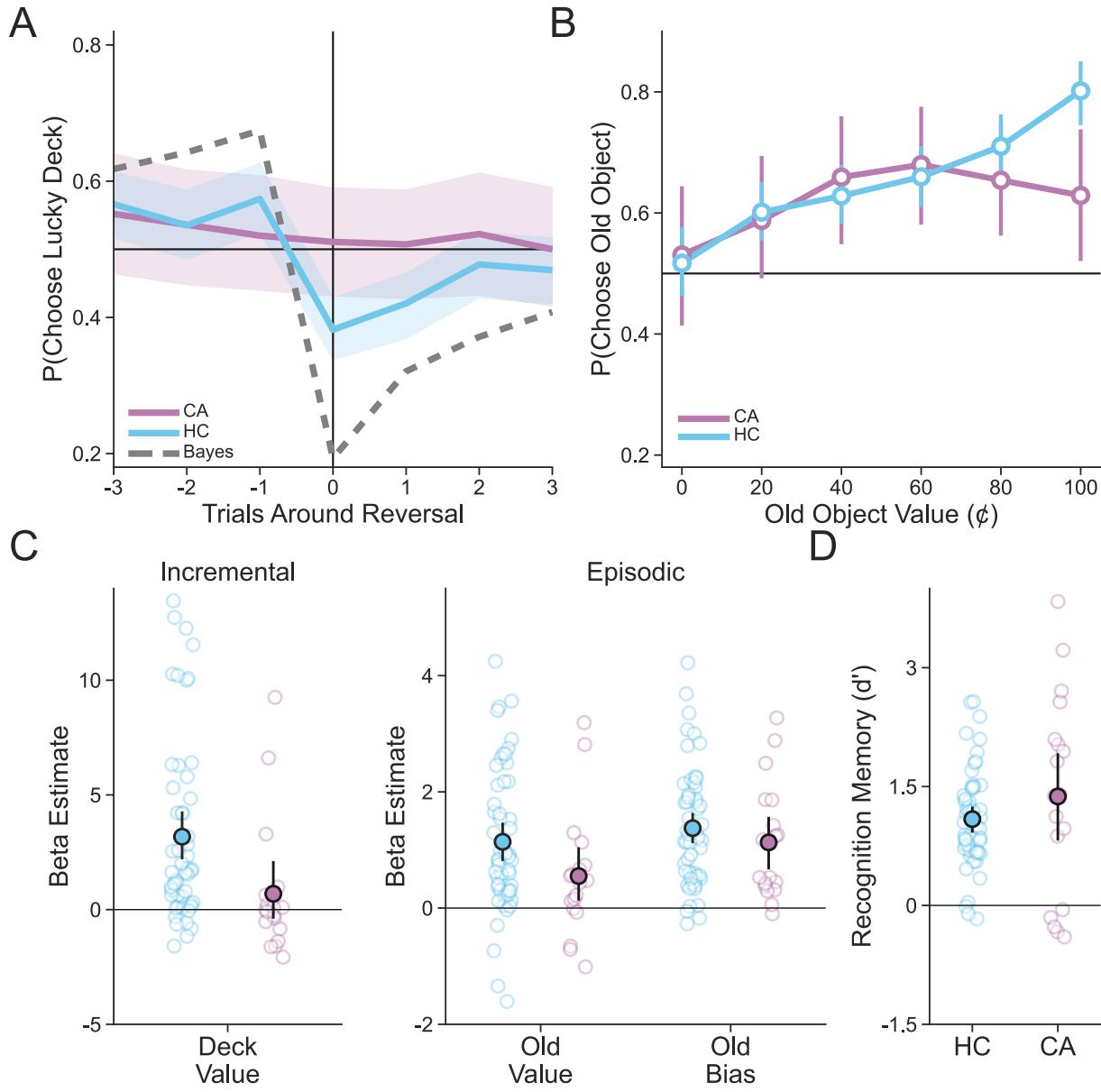
148 Next, to more formally assess participants' performance on this task, we fit a standard Q learning  
149 model to participants' responses. This model captures the extent to which each participant  
150 incorporated trial-by-trial outcomes into running estimates of the value of pressing each button in  
151 response to each cue, as well as whether choices are based on these estimates. As a baseline, we  
152 compared the performance of this model to a biased responder that merely estimated the extent to  
153 which each participant pressed one button over the other, regardless of outcome, in response to  
154 each cue. While healthy controls' responses were well described by the Q learning model, this  
155 model did no better than the biased responder at predicting CA participants' decisions, thus  
156 demonstrating that CA participants engaged in little-to-no incremental learning (**Figure 2D**). On  
157 a measure of estimated out-of-sample predictive performance, controls were substantially better  
158 fit by the Q learner compared to the biased responder, while this improvement in fit was largely  
159 absent for CA participants ( $\beta_{Group} = 30.94$ , 95% CI = [16.465, 46.0]). Thus, while healthy  
160 controls incorporated feedback into their estimates about the relationship between cue and action  
161 at each timepoint, CA participants generally did not.

162 Together, these results indicate that individuals with CA are impaired at reward learning from trial-  
163 and-error.

### 164 **Impaired incremental reward learning but intact episodic memory 165 in the multiple learning strategies task**

166 After establishing that CA participants were impaired in a task that measured solely incremental  
167 reward learning, we wanted to examine both the specificity and generalizability of this impairment  
168 by i) providing an alternative means of reward-based decision making alongside incremental  
169 learning and ii) altering the incremental learning task structure to measure responses to reversal  
170 events rather than drifting probabilities. The multiple learning strategies task was thus used to  
171 accomplish both of these goals.

REWARD LEARNING IN THE CEREBELLUM



172  
 173 **Figure 3. Performance on the multiple learning strategies task.** (A) Deck learning performance on the  
 174 multiple learning strategies task as indicated by the proportion of trials on which the currently lucky deck  
 175 was chosen as a function of how distant those trials were from a reversal in deck value. Performance for  
 176 both healthy controls (HC) and CA participants is shown alongside a Bayesian observer with perfect  
 177 episodic memory for visual comparison. Lines represent group averages and bands represent 95%  
 178 confidence intervals. (B) Object value usage on trials in which a previously seen object appeared. Points  
 179 represent group averages and error bars represent 95% confidence intervals. (C) Inverse temperature  
 180 estimates from the Hybrid model. Individual points represent estimates for each subject, group-level  
 181 averages are shown as filled in points and error bars represent 95% confidence intervals. (D) Recognition  
 182 memory performance on the subsequent memory task. Individual points represent each participant's dprime  
 183 score, filled in points are group-level averages and error bars are 95% confidence intervals.

## REWARD LEARNING IN THE CEREBELLUM

184 Consistent with the results of the incremental learning task, CA participants in the multiple  
185 learning strategies task were less responsive to reward outcomes compared to controls (**Figure**  
186 **3A**). Specifically, controls tended to choose the lucky deck more than CA participants immediately  
187 prior to a reversal ( $\beta_{Group \times t=t-1} = 0.397$ , 95% CI = [0.002, 0.807]), and this tendency was  
188 disrupted by reversals; CA participants did not show this pattern ( $\beta_{Group \times t=0} =$   
189  $-0.897$ , 95% CI = [-1.28, -0.535]), and remained below chance performance after a reversal  
190 occurred ( $\beta_{Group \times t=t+1} = -0.595$ , 95% CI = [-0.984, -0.21]). This indicates that CA  
191 participants were unable to learn which deck had the higher expected value at any given time  
192 throughout the task.

193 We next assessed the extent to which both incrementally constructed value and episodic value  
194 contributed to choice in a combined Hybrid choice model. This model combined a standard Q  
195 learning model with three inverse temperature parameters that captured each participants'  
196 sensitivity to estimated deck value, the true value of previously seen objects, and a bias toward  
197 choosing previously seen objects regardless of their value (**Figure 3B-C**). This model of hybrid  
198 choice outperformed a biased responder, which again served as a baseline, for both CA participants  
199 and controls as there was no difference between groups in estimated out-of-sample predictive  
200 performance ( $\beta_{Group} = -1.631$ , 95% CI = [-11.425, 8.444]). Importantly, this indicates that  
201 the behavior of both CA participants and controls was well described by the hybrid choice model,  
202 which is expected if CA participants are unimpaired at episodic value learning. For each group,  
203 we then assessed whether sensitivity differed from zero and, if so, concluded that participants in  
204 that group made choices that were affected by each possible predictor. While healthy controls  
205 incorporated deck value into their decisions ( $\beta_{HC} = 3.173$ , 95% CI = [2.181, 4.189]), CA  
206 participants generally did not ( $\beta_{CA} = 0.681$ , 95% CI = [-0.668, 2.066]). This reward learning  
207 deficit was specific to value acquired incrementally, however, because CA participants and  
208 controls were both sensitive to episodic value ( $\beta_{HC} = 1.373$ , 95% CI = [1.095, 1.654];  $\beta_{CA} =$   
209  $1.13$ , 95% CI = [0.59, 1.643]) and were both similarly biased by previously seen objects  
210 regardless of their value ( $\beta_{HC} = 1.142$ , 95% CI = [0.798, 1.477];  $\beta_{CA} = 0.551$ , 95% CI =  
211 [0.028, 1.056]). Furthermore, while there were no differences between groups for the effects of  
212 either episodic value ( $\beta_{Group} = 0.244$ , 95% CI = [-0.311, 0.820]) or bias ( $\beta_{Group} =$

## REWARD LEARNING IN THE CEREBELLUM

213 0.586, 95% CI = [-0.051, 1.246]), healthy controls were indeed more sensitive to learned deck  
214 value than CA participants ( $\beta_{Group} = 2.47$ , 95% CI = [0.362, 4.572]).

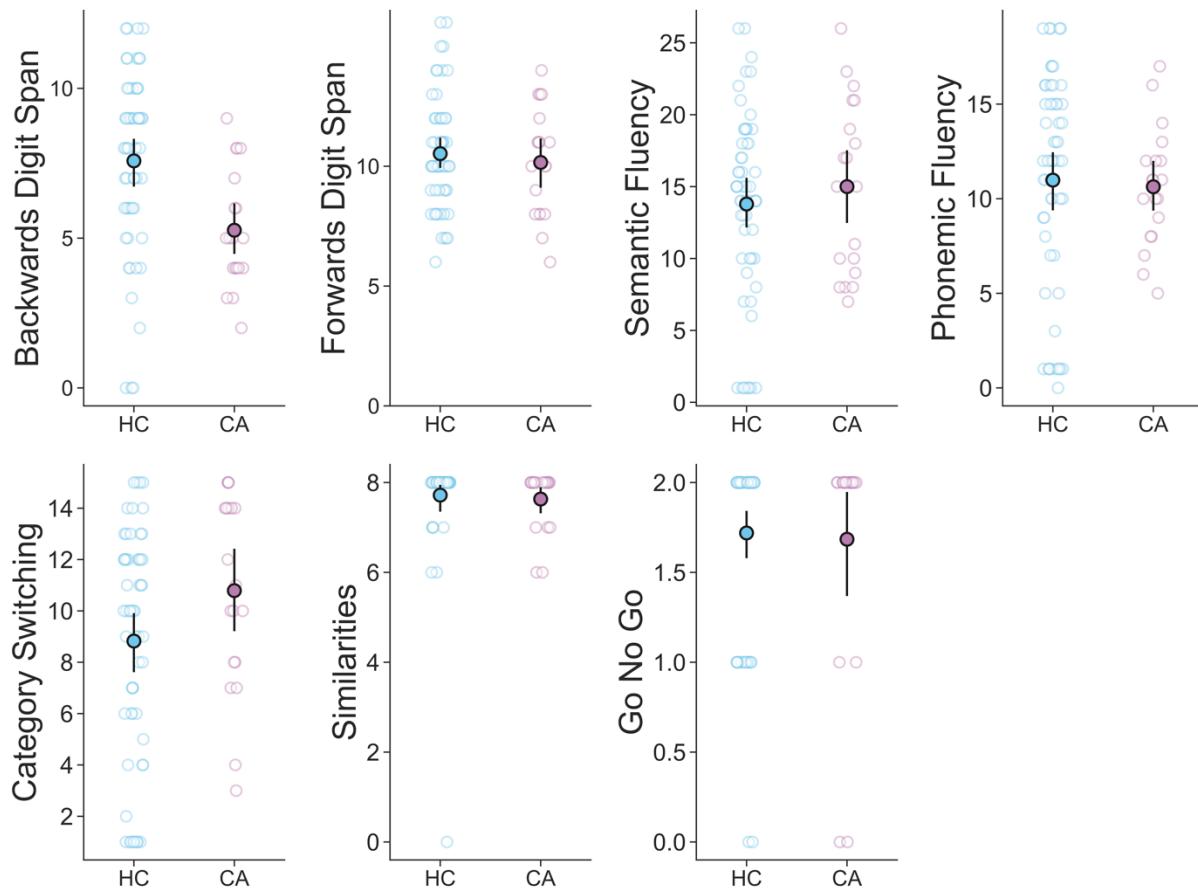
215 We additionally had each participant complete a subsequent memory test for a subset of objects  
216 shown during the multiple learning strategies task. There was no difference in recognition memory  
217 performance between groups ( $\beta_{Group} = -0.487$ , 95% CI = [-1.144, 0.157]). This result  
218 provides further evidence that CA participants were unimpaired at using episodic memory  
219 throughout the task relative to their stark impairments in incremental learning. Lastly, CA  
220 participants and healthy controls demonstrated no differences in reaction time on this task  
221 ( $\beta_{Group} = -86.40$ , 95% CI = [-222.28, 44.76]), suggesting that the behavioral differences  
222 reported here cannot be attributed to motor slowing in CA.

### 223 **Controlling for effects of non-motor deficits and disease subtype**

224 We next sought to ensure that the differences in the tasks reported here were specific to deficits in  
225 reward learning rather than general cognitive impairment. Controlling for cognitive impairment is  
226 particularly important because recent work<sup>33,34</sup> has suggested that incremental learning  
227 experiments tax higher level functions, like executive control and working memory, in addition to  
228 learning from reward prediction error. To address this issue and assess possible cognitive  
229 impairment, we conducted a battery of neuropsychological measures on CA participants (see  
230 **Methods**). Of these, a subset of measures were also completed by healthy controls (**Figure 4**). We  
231 found no differences in performance between groups on all measures except for the backwards  
232 digit span task, which indexes working memory ability, and on which healthy controls scored  
233 higher than CA participants (**Table 2**;  $\beta_{Group} = -2.57$ , 95% CI = [-4.18, -0.92]). Backwards  
234 digit span scores were thus included as covariates in regression analyses where possible (see  
235 **Methods**) in order to control for impacts of this performance difference on impairments in  
236 incremental learning. To further ensure that CA participants' deficient incremental learning was  
237 not due to broad cognitive impairment, we also repeated all analyses excluding seven CA  
238 participants (and their matched controls) with mild cognitive impairment (MCI), as indicated by  
239 scoring lower than 26 on the MoCA (**Table 1**). While CA participants with MCI consisted of some  
240 of the lowest performing participants in our sample (**Figure 2—Figure supplement 1, Figure 2—**

## REWARD LEARNING IN THE CEREBELLUM

241 **Figure supplement 2**), there were no differences in the results across both tasks when they were  
 242 excluded. It is therefore unlikely that CA participants' impaired reward learning ability is due to  
 243 either working memory deficits or cognitive decline more broadly. A full report of these analyses  
 244 can be found in **Appendix A**.



245  
 246 **Figure 4. Neuropsychological test performance for the subset of neuropsychological measures that**  
 247 **were completed by both CA participants and healthy controls (HC).** Scores from individual participants  
 248 are plotted as empty circles behind group-level means plotted as filled circles with uncertainty represented  
 249 by 95% confidence intervals.

250 We next sought to further characterize the nature of CA participants' reward learning impairment  
 251 by looking at the relationship between incremental learning sensitivity, as measured by the Q  
 252 learning models in each task, and performance on our neuropsychological battery. The extent to  
 253 which CA participants learned about cues in the incremental learning task related only to total  
 254 CCAS score ( $r = 0.84$ ,  $p < 0.001$ , Bonferroni corrected; **Table 3**), suggesting that the  
 255 specific contributions of the cerebellum to cognition may impact performance in this task. The

## REWARD LEARNING IN THE CEREBELLUM

CCAS scale was recently developed to measure the exact types of cognitive impairment that result from damage to the cerebellum<sup>35</sup>. Because more focal cerebellar lesions tend to lead to lower total CCAS scores<sup>36</sup>, this provides further evidence of the necessity for the cerebellum to successfully perform the incremental learning task. The relationship between total CCAS score and performance was driven by the timed portions of the CCAS scale (e.g. the Semantic Fluency and Category Switching measures; **Table 3—Table supplement 1**), suggesting a potential effect of slowed responses in the incremental learning task. While CA participants did indeed respond more slowly than healthy controls on this task (see above), we controlled for this difference in our behavioral analysis. Finally, there was no relationship between any measure and incremental learning ability in the multiple learning strategies task (**Table 3**).

Finally, we addressed the possibility that the subset of our sample of CA participants consisting of diagnoses that were less restricted to the cerebellum, namely the three individuals with multiple system atrophy (MSA) and the two individuals with Friedrich's ataxia (FA), could be responsible for the deficits reported here. We repeated all analyses with these five participants excluded and found no differences in the results (**Figure 2—Figure supplement 1**, **Figure 2—Figure supplement 2**). A full report of these analyses can be found in **Appendix B**.

## Discussion

The results of the present work demonstrate that individuals with cerebellar dysfunction, represented by CA cases in our cohort, are impaired at reward learning. While the cerebellum and basal ganglia have traditionally been treated as making separate contributions to learning<sup>5,21</sup>, recent findings have called this dichotomy into question<sup>8–19</sup>. This work has suggested that, alongside its role in motor learning, the cerebellum likely operates in concert with the basal ganglia to support reinforcement learning from reward. Our study corroborates these findings from animal models<sup>10–19</sup>, providing evidence that the human cerebellum is necessary for learning associations from reward. In comparison to age- and sex-matched healthy controls, CA participants were impaired at reward-based learning from trial-and-error. Further, CA participants retained the ability to employ an alternative strategy based in episodic memory to guide their decisions, demonstrating that this impairment is specific to incremental learning. These results challenge the idea that the

## REWARD LEARNING IN THE CEREBELLUM

284 cerebellum is used primarily for motor learning and shed light on how multiple neural systems  
285 may interact with one another to support learning in the non-motor domain.

286 Our findings join a litany of recent research suggesting that the cerebellum plays a broad role in  
287 human cognition<sup>22,37-39</sup>. Indeed, individuals with damage to the cerebellum demonstrate  
288 impairment in a wide range of cognitive functions including cognitive control<sup>40</sup> and impulsivity<sup>41</sup>.  
289 Human functional neuroimaging studies have also revealed cerebellar activity in a variety of  
290 different non-motor tasks<sup>22,38</sup>. Many of these functions are likely supported by the robust  
291 bidirectional connections the cerebellum shares with the prefrontal cortex<sup>42,43</sup>. In particular, recent  
292 findings have indicated that individuals with CA have heightened domain-specific impulsive and  
293 compulsive behaviors, which is a common symptom of underlying reward system dysfunction<sup>44,45</sup>.  
294 Our study adds to this work by suggesting that the cerebellum is additionally necessary for reward  
295 learning in humans.

296 While there is growing evidence validating the implication of the cerebellum in reward-based  
297 learning in animals, there is only limited work on this topic in humans. Early imaging studies, for  
298 example, demonstrated cerebellar BOLD activity in patients with substance use disorder who  
299 performed reward-based learning tasks<sup>23</sup> and experienced cravings<sup>24</sup>, and also in response to  
300 unexpected reward<sup>25</sup>. However, it remains unknown how cerebellar damage impacts reward  
301 learning, as investigations of reward learning in the cerebellum are rare. While two previous  
302 studies employed reward-based experimental tasks in individuals with isolated ischemic lesions of  
303 the cerebellum<sup>46,47</sup>, results until this point have remained far from conclusive. Thoma et al. (2008)  
304 used a reward-based learning task consisting of an initial acquisition phase in which eight  
305 participants with cerebellar damage were rewarded for learning associations between colors and  
306 symbols followed by a reversal portion in which they had to disremember previously acquired  
307 knowledge and learn new associations for each cue. While participants with cerebellar damage  
308 demonstrated no impairment at acquiring new, reward-based knowledge, they were selectively  
309 impaired at learning from a single reversal. While this study complements our findings, we found  
310 evidence for more global impairment: CA participants in both of our tasks were unable to learn  
311 associations from reward on a trial-by-trial basis. Rustemeier et al. (2016) took a different approach  
312 by asking twelve individuals with cerebellar damage to learn a simple acquisition task from

## REWARD LEARNING IN THE CEREBELLUM

313 probabilistic feedback and subsequently transfer this knowledge to re-arranged stimuli. While  
314 participants were unimpaired behaviorally at this task, electroencephalographic (EEG) results  
315 revealed that they may process reward-based feedback differently from controls. Our findings  
316 support this interpretation and further suggest that processing of trial-by-trial feedback is not just  
317 different, but impaired, in individuals with cerebellar damage. Finally, while other related studies  
318 showed impairment in learning from reinforcement in participants with cerebellar damage<sup>20,48</sup>, this  
319 work has focused primarily on movement-dependent deficiencies.

320 While our findings suggest that the cerebellum is necessary for incremental reward learning, they  
321 cannot speak to the neural circuitry underlying this role. One intriguing possibility is that the  
322 cerebellum may operate in tandem with the basal ganglia—canonically seen as the seat of  
323 reinforcement learning in the brain<sup>5,21</sup>—to learn about reward incrementally. Reward prediction  
324 error signals in midbrain dopamine neurons that provide input to the basal ganglia<sup>27,29</sup> have also  
325 been found to be encoded by cerebellar neurons<sup>9,15,17,19</sup>. Further, through excitatory projections to  
326 the ventral tegmental area, the cerebellum has widespread reciprocal connections with the basal  
327 ganglia and has recently been shown to influence reward-driven behavior through these  
328 projections<sup>8,49</sup>. While reinforcement learning via the basal ganglia and supervised learning via the  
329 cerebellum have typically been treated as fulfilling entirely separate roles<sup>5,21</sup>, these systems appear  
330 to be more interdependent than previously thought. Future investigations of the relationship  
331 between the basal ganglia and cerebellum are needed to clarify the exact mechanisms underlying  
332 reinforcement learning in the brain.

333 Lastly, there are several potential limitations related to the nature of our sample that should be  
334 considered when interpreting these findings. First, cerebellar dysfunction in our sample of CA  
335 participants was caused by several different conditions. While most of these pathologies are  
336 predominantly restricted to the cerebellum, non-cerebellar brain areas and circuits could also be  
337 affected, particularly in participants diagnosed with either MSA or FA. There was, however, no  
338 change in the reported reward-based learning deficits when these participants were excluded.  
339 Second, while cognitive impairment due to neurodegenerative disease could potentially contribute  
340 to some of the deficits measured here, we accounted for this possibility by establishing that the  
341 incremental reward learning deficits reported here persist regardless of MCI status. We also

## REWARD LEARNING IN THE CEREBELLUM

342 collected basic neuropsychological measures from all participants, and CA participants were not  
343 different from controls on the vast majority of measures. We focused particularly on possible  
344 contributions of working memory given recent work suggesting that working memory plays an  
345 important role in incremental reward learning<sup>33,34</sup>. While CA participants and controls performed  
346 similarly on the forward digit span task, CA participants were somewhat impaired at backwards  
347 digit span. We controlled for this difference by including backwards digit span scores as covariates  
348 in our analyses. Finally, while our control participants completed the study online, we accounted  
349 for potential variability caused by this difference in setting by collecting three matched controls  
350 for each CA participant in our sample.

351 Taken together, our findings suggest that the human cerebellum is necessary for reward learning.  
352 These results provide new constraints on models of non-motor learning and suggest that the  
353 cerebellum and basal ganglia work in tandem to support learning from reinforcement.

## 354 Materials and Methods

### 355 Cerebellar Ataxia Participants

356 Nineteen individuals with cerebellar ataxia were recruited from the Ataxia Clinic, Columbia  
357 University Medical Center and completed both tasks (see **Table 1** for information about basic CA  
358 participant demographics and diagnoses). Due to hardware issues, data from one participant on  
359 each task was not saved. The first CA participant also completed a shorter pilot version of the  
360 incremental learning task, and several changes were made before running this task on the other 18  
361 CA participants. Thus, the final sample for the incremental learning task was 17 CA participants,  
362 and the final sample for the multiple learning strategies task was 18 CA participants. Task order  
363 was counterbalanced. A neuropsychological battery comprised of the Montreal Cognitive  
364 Assessment (MOCA), Beck's Depression Inventory (BDI), MESA digit forward and backward  
365 span, trail making test A and B, and the cerebellar cognitive affective syndrome scale (CCAS) was  
366 conducted between tasks for each participant. This battery was specifically selected based on the  
367 current understanding of the cerebellum's role and association with non-motor symptoms, such as  
368 depression<sup>50</sup>, executive function<sup>51,52</sup>, and attention<sup>53</sup>.

**369 Healthy Controls**

370 Age- and sex-matched participants were recruited through Amazon Mechanical Turk using the  
371 Cloud Research Approved Participants feature<sup>54</sup>. To account for potential variability due to online  
372 data collection, three matched controls were collected for each CA participant, bringing the total  
373 number of controls to 57 (3:1 match). Data from one control was excluded for the multiple learning  
374 strategies task due to random responding. Task order was counterbalanced such that the tasks were  
375 completed in the identical order to each control's matched CA participant. A modified online  
376 neuropsychological battery consisting of 7 measures was completed in between each task for  
377 comparison to individuals with CA. Five of these measures (Semantic Fluency, Phonemic Fluency,  
378 Category Switching, Similarities and Go No Go) were directly taken from the CCAS, and two  
379 others were comprised of the MESA digit forward back backward span. Participant recruitment  
380 was restricted to the United States. Before starting each task, all participants were required to score  
381 100% on a quiz that tested their comprehension of the instructions. Informed consent was obtained  
382 with approval form the Columbia University Institutional Review Board.

**383 Experimental Design****384 Incremental Learning Task**

385 In the incremental learning task (**Figure 1A**), participants were told that they would be playing a  
386 game where they were required to press a key, either F or J, whenever one of two symbols was  
387 seen, and that they would receive feedback about whether they had pressed correctly following  
388 each trial. They were then informed that it was their job to determine which key they should press  
389 for each symbol, and that what key is best will change throughout the experiment. Outcomes were  
390 determined by a drifting probability such that each button was correct for each image 50% of the  
391 time. Critically, these probabilities differed over time, thus encouraging constant learning  
392 throughout the task. Participants were told to press the F key with their left index finger and the J  
393 key with their right index finger. The response period during which the symbol remained on the  
394 screen lasted 1.5 seconds, with feedback displayed for 1 second immediately following the  
395 response period. An intertrial interval featuring a fixation cross was shown for an average of 1  
396 second, but varied between 0.5 and 1.5 seconds. Lastly, to provide a rewarding outcome for correct

## REWARD LEARNING IN THE CEREBELLUM

397 responses, participants were informed that they could earn bonus money based on their  
398 performance. Correct responses were worth an additional cent each.

### 399 **Multiple Learning Strategies Task**

400 The other task completed by participants was previously developed by our lab<sup>31,32</sup> to measure the  
401 relative contribution of incremental learning and episodic memory to decisions (**Figure 1B**).  
402 Participants were told that they would be playing a card game where their goal was to win as much  
403 money as possible. Each trial consisted of a choice between two decks of cards that differed based  
404 on their color (red or blue). Participants had two seconds to decide between the decks. The outcome  
405 of each decision was then immediately displayed for 1.5 seconds. Following each decision,  
406 participants were shown a fixation cross during the intertrial interval period which varied in length  
407 (mean = 1.5 seconds, min = 1 seconds, max = 2 seconds). Decks were equally likely to appear on  
408 either side of the screen on each trial. Participants completed a total of 150 trials.

409 Participants were made aware that there were two ways they could earn bonus money throughout  
410 the task, which allowed for the use of incremental learning and episodic memory respectively.  
411 First, at any point in the experiment one of the two decks was “lucky”, meaning that the expected  
412 value ( $V$ ) of one deck color was higher than the other ( $V_{lucky}=63\text{¢}$ ,  $V_{unlucky}=37\text{¢}$ ). Outcomes  
413 ranged from \$0 to \$1 in increments of 20¢. Critically, the mapping from  $V$  to deck color reversed  
414 periodically throughout the experiment, which incentivized participants to utilize each deck’s  
415 recent reward history to determine the identity of the currently lucky deck. Second, to assess the  
416 use of episodic memory throughout the task, each card within a deck featured an image of a trial-  
417 unique object that could re-appear once throughout the experiment after initially being chosen.  
418 Participants were told that if they encountered a card a second time it would be worth the same  
419 amount as when it was first chosen, regardless of whether its deck color was currently lucky or  
420 not. On a given trial  $t$ , cards chosen once from trials  $t - 9$  through  $t - 30$  had a 60% chance of  
421 reappearing following a sampling procedure designed to prevent each deck’s expected value from  
422 becoming skewed by choice, minimize the correlation between the expected value of previously  
423 seen cards and deck expected value, and ensure that choosing a previously selected card remained  
424 close to 50¢.

## REWARD LEARNING IN THE CEREBELLUM

Following completion of the multiple learning strategies task, we tested participants' memory for trial-unique objects. Participants completed up to 54 three-part memory trials. An object was first displayed on the screen and participants were asked whether or not they had previously seen the object and were given five response options: Definitely New, Probably New, Don't Know, Probably Old, Definitely Old. If the participant indicated that they had not seen the object before or did not know, they moved on to the next trial. If, however, they indicated that they had seen the object before they were then asked if they had chosen the object or not. Lastly, if they responded that they had chosen the object, they were asked what the value of that object was.

## Computational Models

In order to capture subjective estimates of incrementally constructed value on each task, we fit computational models to participants' choices. Below we describe each of these models in detail.

### Q Learning Models

We modeled incremental reward learning using a Q Learning model, which is a standard model-free reinforcement learner that assumes a stored value ( $Q$ ) for each deck is updated over time<sup>26,28</sup>.  $Q$  is then referenced on each decision in order to guide choices. After each outcome,  $r_t$ , the value for an option  $Q_1$  is updated according to the following rule if that option is chosen:

$$Q_{1,t+1} = Q_{1,t} + \alpha(r_t - Q_{1,t})$$

And is not updated if a different option is chosen:

$$Q_{1,t+1} = Q_{1,t}$$

Likewise, if a different option is chosen, its value is updated equivalently. Large differences between estimated value and outcomes therefore have a larger impact on updates, but the overall degree of updating is controlled by the learning rate,  $\alpha$ , which is a free parameter constrained to lie between 0 and 1.

For the incremental learning task, the model learned separate Q values for each cue and button combination, such that four Q values were estimated in total. Decisions were then modeled using the following rule:

## REWARD LEARNING IN THE CEREBELLUM

451  $P(\text{Choose}F) = \sigma(\beta_{0,1} + \beta_{0,2} + \beta_{1,1}(Q_{F,1} - Q_{J,1}) + \beta_{1,2}(Q_{F,2} - Q_{J,2}))$

452  $\sigma(x) = \frac{1}{1+e^{-x}}$

453 such that four inverse temperatures  $\beta$  were estimated to capture a bias toward choosing a key for  
454 each cue ( $\beta_{0,1}$  and  $\beta_{0,2}$ ) and sensitivity to incrementally learned value for each cue ( $\beta_{1,1}$  and  $\beta_{1,2}$ ).  
455 This model is referred to as the “Q Learner” model throughout the text.

456 For the multiple learning strategies task, the model learned separate Q values for each deck color,  
457 such that two Q values were estimated in total. Decisions were then modeled using the following  
458 rule:

459  $P(\text{Choose}Red) = \sigma(\beta_1(Q_R - Q_B) + \beta_2(OldValue) + \beta_3(Old))$

460 such that three inverse temperatures  $\beta$  were estimated to capture sensitivity to incrementally  
461 learned value ( $\beta_1$ ), sensitivity to the value of previously seen objects ( $\beta_2$ ), and a bias toward  
462 choosing the deck featuring a previously seen object regardless of its value ( $\beta_3$ ). The predictor  
463 *OldValue* was the coded true value of a previously seen object (ranging from 0.5 if the value was  
464 \$1 on the red deck or \$0 on the blue deck to -0.5 if the value was \$0 on the red deck and \$1 on the  
465 blue deck) and the predictor *Old* was coded as 0.5 if the red deck featured a previously seen object  
466 and -0.5 if the blue deck did instead. For both of these predictors, trials that did not feature a  
467 previously seen object were coded as 0. This model is referred to as the “Hybrid” model throughout  
468 the text.

### 469 Biased Responder Model

470 For both tasks, we compared the performance of the Q Learning models to a model which made  
471 choices that were completely independent of reward information. For the incremental learning  
472 task, this model was simply:

473  $P(\text{Choose}F) = \sigma(\beta_{0,1} + \beta_{0,2})$

474 such that choices depended only on choosing a button to press for each cue throughout the  
475 experiment. For the multiple learning strategies task, this model was:

476  $P(\text{ChooseRed}) = \sigma(\beta_0)$

477 such that choices depended only on preferring one deck over the other throughout the experiment.  
 478 Our logic in using this model as a baseline was that responses captured by the Q learning models  
 479 should, at a minimum, outperform a biased responder that did not consider reward in order for it  
 480 to make meaningful predictions about participants' behavior.

481 **Posterior Inference and Model Comparison**

482 Model parameters for each participant were estimated using Bayesian inference. The joint  
 483 posterior was approximated using No-U-Turn Sampling<sup>55</sup> as implemented in stan<sup>56</sup>. Four chains  
 484 with 2000 samples (1000 discarded as burn-in) were run for a total of 4000 posterior samples per  
 485 model per subject. Chain convergence was determined by ensuring that the Gelman-Rubin statistic  
 486  $\hat{R}$  was close to 1 for all parameters. For the incremental learning task, the Q learner did not  
 487 converge for one CA participant, and so that individual and their matched controls were removed  
 488 from further model-based analyses. For the multiple learning strategies task, all models for all  
 489 participants converged.

490 Under this approach, the likelihood function for all models can be written as:

491  $c_t \sim \text{Bernoulli}(\theta_t)$

492 where  $c_t$  is 1 if the subject chose F (in the response mapping task) or red (in the multiple learning  
 493 strategies task). Here,  $\theta_t$  is the linear combination of inverse temperature parameters and predictors  
 494 explained above for each model. For the Q learning models, the learning rate,  $\alpha$ , had the following  
 495 weakly informative prior:

496  $\alpha \sim \beta(0,1)$

497 For all models, every inverse temperature parameter had the following weakly informative prior:

498  $\beta \sim \mathcal{N}(0,5)$

## REWARD LEARNING IN THE CEREBELLUM

499 Model fit was assessed using approximate leave-one-out cross validation estimated using Pareto-  
500 smoothed importance sampling<sup>57</sup>. The expected log pointwise predictive density (ELPD) was  
501 computed and used as a measure of out-of-sample predictive fit for each model.

### 502 Bayesian Observers

503 In order to provide a normative performance benchmark, we simulated beliefs about incremental  
504 value as estimated by Bayesian observers for each task. For the incremental learning task, this  
505 learner was a Kalman Filter<sup>58</sup> and for the multiple learning strategies task this learner was a reduced  
506 Bayesian change-point detection model<sup>59</sup>. Choices in the incremental learning task were made  
507 according to which button the observer believed was the most likely to be rewarded for each cue  
508 at each time point. Choices in the multiple learning strategies task were made differently depending  
509 on whether a previously seen object was present. For trials in which no previously seen object was  
510 shown, the observer responded according to its beliefs about deck value. For trials in which a  
511 previously seen object was present, however, the observer compared the value of that object to its  
512 belief about deck value for the opposing deck and chose accordingly. In this way, the observer was  
513 augmented with “perfect” episodic memory.

### 514 Regression Models

515 Mixed effects Bayesian regressions were used to test effects of group (CA participant or control).  
516 Group membership was allowed to vary randomly by CA participant identifier, *pid*, such that CA  
517 participants and matched controls were assigned the same ID. In these models, *GroupID* was  
518 coded as -0.5 for CA participants and 0.5 for controls. We additionally controlled for working  
519 memory ability by including backwards digit span scores, *dsBwd*, as a standardized covariate in  
520 these analyses.

521 For the incremental learning task, we assessed behavioral performance using the following logistic  
522 regression:

$$p(\text{Correct}) = \sigma(\beta_0 + b_{0,pid[t]} + \text{GroupID}_t(\beta_1 + b_{1,pid[t]}) + \\ pFReward1_t(\beta_2 + b_{2,pid[t]}) + \text{GroupID}_t \times pFReward1_t(\beta_3 + b_{3,pid[t]}) + \\ pFReward1_t^2(\beta_4 + b_{4,pid[t]}) + \text{GroupID}_t \times pFReward1_t^2(\beta_5 + b_{5,pid[t]}) + \\ dsBwd\beta_6 + RT\beta_7)$$

## REWARD LEARNING IN THE CEREBELLUM

524 Here, and in all regressions described in this section,  $\beta$  stands for fixed effects and  $b$  stands for  
525 random effects of CA participant ID. The predictor  $pFReward1$  indicates the true underlying  
526 difficulty of the task and is the probability that the F key was rewarding for cue one. A second-  
527 order polynomial was included for this predictor as extreme values indicate portions of the task  
528 that are easier and middling values indicate portions of the task that were more difficult. Interaction  
529 effects of this predictor and group were included to capture differences in sensitivity to the  
530 underlying task difficulty between the groups. Lastly, the reaction time,  $RT$ , on each decision was  
531 included as a standardized covariate in this analysis to account for any differences that may be due  
532 to slowed responding by individuals with CA on this task.

533 For both the incremental learning and multiple learning strategies tasks, we assessed whether there  
534 were differences between the groups on Q learning model performance compared to the baseline  
535 biased responder model with the following linear regression:

536 
$$ELPDDifference = \beta_0 + b_{0,pid[t]} + GroupID_t(\beta_1 + b_{1,pid[t]}) + dsBwd\beta_2$$

537 where  $ELPDDifference$  was the difference in model performance (Q Learning model ELPD -  
538 Biased Responder ELPD; see above) for each subject.

539 For the multiple learning strategies task, we assessed behavioral incremental learning performance  
540 using the following logistic regression:

541 
$$p(ChooseLucky) = \sigma(\beta_0 + b_{0,pid[t]} + T_{-3:3} \times GroupID_t(\beta_{1:7} + b_{1:7,pid[t]}) + dsBwd\beta_8)$$

542 In this regression, we grouped trials according to their distance from a reversal, up to three trials  
543 prior to ( $t = -3:-1$ ), during ( $t = 0$ ), and after ( $t = 1:3$ ) a reversal occurred. We then dummy  
544 coded them to measure their effects on the degree to which the lucky deck was chosen and  
545 interacted each dummy coded regressor with group to measure how this was affected by group  
546 membership.

547 We then assessed the degree to which each group used either incrementally learned deck value,  
548 the value of previously seen objects, or a bias toward previously seen objects regardless of their  
549 value as estimated by the Hybrid Q learning model using a simple linear regression of the following  
550 form for each of these inverse temperature parameters and groups:

## REWARD LEARNING IN THE CEREBELLUM

551  $InvTemp_s = \beta_0 + dsBwd\beta_1$

552 Here we interested primarily in the intercept,  $\beta_0$ , as this determined the degree to which each  
553 group's inverse temperatures were above zero. We additionally assessed differences between  
554 groups on each of these measures by including fixed and random effects for group that varied by  
555 matched participant ID, as in previously described regression analyses.

556 We also assessed the impact of group on subsequent memory performance following the multiple  
557 learning strategies task using the following linear regression:

558  $Dprime = \beta_0 + b_{0,pid[t]} + GroupID_t(\beta_1 + b_{1,pid[t]}) + dsBwd\beta_2$

559 where  $Dprime$  is the signal detection measure  $d'$ , which is the difference in z scored hit rate and  
560 false alarm rate for each participant.

561 We were also interested in determining whether there were any differences in reaction times  
562 between individuals with CA and matched controls due to motor impairment. For both tasks, we  
563 did this by assessing whether there were any differences in reaction time between groups:

564  $RT = \beta_0 + b_{0,pid[t]} + GroupID_t(\beta_1 + b_{1,pid[t]})$

565 where  $RT$  was the median reaction time across trials in either task. A separate regression of this  
566 form was used for each of the two tasks. We also assessed whether there were differences on each  
567 neuropsychological measure using a similar regression.

568 For all regression analyses, fixed effects are reported in the text as the mean of each parameter's  
569 marginal posterior distribution alongside 95% credible intervals, which indicate where 95% of the  
570 posterior density falls. Parameter values outside of this range are unlikely given the model, data,  
571 and priors. Thus, if the range of likely values does not include zero, we conclude that a meaningful  
572 effect was observed.

573 **Code Accessibility**

574 All code used to analyze the data in this study may be found here:

575 <https://github.com/boomsbloom/ataxia-rl>

576 **References**

- 577 1. Raymond, J. L., Lisberger, S. G. & Mauk, M. D. The Cerebellum: A Neuronal Learning  
578 Machine? *Science* **272**, 1126–1131 (1996).
- 579 2. Llinás, R. & Welsh, J. P. On the cerebellum and motor learning. *Curr. Opin. Neurobiol.* **3**,  
580 958–965 (1993).
- 581 3. Ito, M. & Itō, M. *The Cerebellum and Neural Control*. (Raven Press, 1984).
- 582 4. Marr, D. A theory of cerebellar cortex. *J. Physiol.* **202**, 437–470 (1969).
- 583 5. Doya, K. What are the computations of the cerebellum, the basal ganglia and the cerebral  
584 cortex? *Neural Netw.* **12**, 961–974 (1999).
- 585 6. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn.*  
586 *Sci.* **2**, 338–347 (1998).
- 587 7. Raymond, J. L. & Medina, J. F. Computational Principles of Supervised Learning in the  
588 Cerebellum. *Annu. Rev. Neurosci.* **41**, 233–253 (2018).
- 589 8. Caligiore, D., Arbib, M. A., Miall, R. C. & Baldassarre, G. The super-learning hypothesis:  
590 Integrating learning processes across cortex, cerebellum and basal ganglia. *Neurosci.*  
591 *Biobehav. Rev.* **100**, 19–34 (2019).
- 592 9. Hull, C. Prediction signals in the cerebellum: Beyond supervised motor learning. *eLife* **9**,  
593 e54073 (2020).
- 594 10. Sendhilnathan, N. & Goldberg, M. E. *The mid-lateral cerebellum is necessary for*  
595 *reinforcement learning*. <http://biorxiv.org/lookup/doi/10.1101/2020.03.20.000190> (2020)  
596 doi:10.1101/2020.03.20.000190.

## REWARD LEARNING IN THE CEREBELLUM

- 597 11. Sendhilnathan, N., Semework, M., Goldberg, M. E. & Ipata, A. E. Neural Correlates of  
598 Reinforcement Learning in Mid-lateral Cerebellum. *Neuron* **106**, 188–198.e5 (2020).
- 599 12. Sendhilnathan, N., Ipata, A. & Goldberg, M. E. Mid-lateral cerebellar complex spikes encode  
600 multiple independent reward-related signals during reinforcement learning. *Nat. Commun.*  
601 **12**, 6475 (2021).
- 602 13. Larry, N., Yarkoni, M., Lixenberg, A. & Joshua, M. Cerebellar climbing fibers encode  
603 expected reward size. *eLife* **8**, e46870 (2019).
- 604 14. Carta, I., Chen, C. H., Schott, A. L., Dorizan, S. & Khodakhah, K. Cerebellar modulation of  
605 the reward circuitry and social behavior. *Science* **363**, eaav0581 (2019).
- 606 15. Heffley, W. & Hull, C. Classical conditioning drives learned reward prediction signals in  
607 climbing fibers across the lateral cerebellum. *eLife* **8**, e46764 (2019).
- 608 16. Wagner, M. J., Kim, T. H., Savall, J., Schnitzer, M. J. & Luo, L. Cerebellar granule cells  
609 encode the expectation of reward. *Nature* **544**, 96–100 (2017).
- 610 17. Heffley, W. *et al.* Coordinated cerebellar climbing fiber activity signals learned sensorimotor  
611 predictions. *Nat. Neurosci.* **21**, 1431–1441 (2018).
- 612 18. Kostadinov, D., Beau, M., Blanco-Pozo, M. & Häusser, M. Predictive and reactive reward  
613 signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat. Neurosci.* **22**,  
614 950–962 (2019).
- 615 19. Ohmae, S. & Medina, J. F. Climbing fibers encode a temporal-difference prediction error  
616 during cerebellar learning in mice. *Nat. Neurosci.* **18**, 1798–1803 (2015).
- 617 20. Therrien, A. S., Wolpert, D. M. & Bastian, A. J. Effective reinforcement learning following  
618 cerebellar damage requires a balance between exploration and motor noise. *Brain J.*  
619 *Neurol.* **139**, 101–114 (2016).
- 620 21. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor  
621 control. *Curr. Opin. Neurobiol.* **10**, 732–739 (2000).

## REWARD LEARNING IN THE CEREBELLUM

- 622 22. King, M., Hernandez-Castillo, C. R., Poldrack, R. A., Ivry, R. B. & Diedrichsen, J. Functional  
623 boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat. Neurosci.*  
624 **22**, 1371–1378 (2019).
- 625 23. Volkow, N. D. *et al.* Expectation enhances the regional brain metabolic and the reinforcing  
626 effects of stimulants in cocaine abusers. *J. Neurosci. Off. J. Soc. Neurosci.* **23**, 11461–  
627 11468 (2003).
- 628 24. Grant, S. *et al.* Activation of memory circuits during cue-elicited cocaine craving. *Proc. Natl.  
629 Acad. Sci. U. S. A.* **93**, 12040–12045 (1996).
- 630 25. Ramnani, N., Elliott, R., Athwal, B. S. & Passingham, R. E. Prediction error for free  
631 monetary reward in the human prefrontal cortex. *NeuroImage* **23**, 777–786 (2004).
- 632 26. Sutton, R. S. & Barto, A. G. Reinforcement Learning: An Introduction. 352.
- 633 27. Houk, J. C., Adams, J. L. & Barto, A. G. A model of how the basal ganglia generate and use  
634 neural signals that predict reinforcement. in *Models of information processing in the basal  
635 ganglia* 249–270 (The MIT Press, 1995).
- 636 28. Rescorla, R. A. & Wagner, A. R. 3 A Theory of Pavlovian Conditioning : Variations in the  
637 Effectiveness of Reinforcement and Nonreinforcement. in (1972).
- 638 29. Schultz, W., Dayan, P. & Montague, P. R. A Neural Substrate of Prediction and Reward.  
639 *Science* **275**, 1593–1599 (1997).
- 640 30. Kuo, S.-H. Ataxia. *Contin. Minneap. Minn* **25**, 1036–1054 (2019).
- 641 31. Duncan, K., Semmler, A. & Shohamy, D. Modulating the Use of Multiple Memory Systems in  
642 Value-based Decisions with Contextual Novelty. *J. Cogn. Neurosci.* 1–13 (2019)  
643 doi:10.1162/jocn\_a\_01447.
- 644 32. Nicholas, J., Daw, N. D. & Shohamy, D. Uncertainty alters the balance between incremental  
645 learning and episodic memory. *eLife* **11**, e81679 (2022).

## REWARD LEARNING IN THE CEREBELLUM

- 646 33. Collins, A. G. E. & Frank, M. J. How much of reinforcement learning is working memory, not  
647 reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur. J.*  
648 *Neurosci.* **35**, 1024–1035 (2012).
- 649 34. Yoo, A. H. & Collins, A. G. E. How Working Memory and Reinforcement Learning Are  
650 Intertwined: A Cognitive, Neural, and Computational Perspective. *J. Cogn. Neurosci.* **34**,  
651 551–568 (2022).
- 652 35. Hoche, F., Guell, X., Vangel, M. G., Sherman, J. C. & Schmahmann, J. D. The cerebellar  
653 cognitive affective/Schmahmann syndrome scale. *Brain* **141**, 248–270 (2018).
- 654 36. Chirino-Pérez, A. *et al.* Mapping the Cerebellar Cognitive Affective Syndrome in Patients  
655 with Chronic Cerebellar Strokes. *The Cerebellum* **21**, 208–218 (2022).
- 656 37. McDougle, S. D. *et al.* Continuous manipulation of mental representations is compromised  
657 in cerebellar degeneration. *Brain J. Neurol.* awac072 (2022) doi:10.1093/brain/awac072.
- 658 38. Buckner, R. L. The cerebellum and cognitive function: 25 years of insight from anatomy and  
659 neuroimaging. *Neuron* **80**, 807–815 (2013).
- 660 39. Koziol, L. F. *et al.* Consensus Paper: The Cerebellum's Role in Movement and Cognition.  
661 *Cerebellum Lond. Engl.* **13**, 151–177 (2014).
- 662 40. Alexander, M. P., Gillingham, S., Schweizer, T. & Stuss, D. T. Cognitive impairments due to  
663 focal cerebellar injuries in adults. *Cortex J. Devoted Study Nerv. Syst. Behav.* **48**, 980–990  
664 (2012).
- 665 41. Amokrane, N., Lin, C.-Y. R., Desai, N. A. & Kuo, S.-H. The Impact of Compulsivity and  
666 Impulsivity in Cerebellar Ataxia: A Case Series. *Tremor Hyperkinetic Mov.* **10**, 43.
- 667 42. Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C. & Yeo, B. T. T. The organization  
668 of the human cerebellum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**,  
669 2322–2345 (2011).

## REWARD LEARNING IN THE CEREBELLUM

- 670 43. Middleton, F. A. & Strick, P. L. Cerebellar Projections to the Prefrontal Cortex of the  
671 Primate. *J. Neurosci.* **21**, 700–712 (2001).
- 672 44. Amokrane, N. et al. Impulsivity in Cerebellar Ataxias: Testing the Cerebellar Reward  
673 Hypothesis in Humans. *Mov. Disord.* **35**, 1491–1493 (2020).
- 674 45. Chen, T. X. et al. Impulsivity Trait Profiles in Patients With Cerebellar Ataxia and Parkinson  
675 Disease. *Neurology* **99**, e176–e186 (2022).
- 676 46. Thoma, P., Bellebaum, C., Koch, B., Schwarz, M. & Daum, I. The Cerebellum Is Involved in  
677 Reward-based Reversal Learning. *The Cerebellum* **7**, 433 (2008).
- 678 47. Rustemeier, M., Koch, B., Schwarz, M. & Bellebaum, C. Processing of Positive and  
679 Negative Feedback in Patients with Cerebellar Lesions. *Cerebellum Lond. Engl.* **15**, 425–  
680 438 (2016).
- 681 48. McDougle, S. D. et al. Credit assignment in movement-dependent reinforcement learning.  
682 *Proc. Natl. Acad. Sci.* **113**, 6797–6802 (2016).
- 683 49. Caligiore, D. et al. Consensus Paper: Towards a Systems-Level View of Cerebellar  
684 Function: the Interplay Between Cerebellum, Basal Ganglia, and Cortex. *The Cerebellum*  
685 **16**, 203–229 (2017).
- 686 50. Hariri, A. R. The Emerging Importance of the Cerebellum in Broad Risk for  
687 Psychopathology. *Neuron* **102**, 17–20 (2019).
- 688 51. Bellebaum, C. & Daum, I. Cerebellar involvement in executive control. *The Cerebellum* **6**,  
689 184–192 (2007).
- 690 52. Beuriat, P.-A. et al. A New Insight on the Role of the Cerebellum for Executive Functions  
691 and Emotion Processing in Adults. *Front. Neurol.* **11**, (2020).
- 692 53. Mannarelli, D. et al. The Cerebellum Modulates Attention Network Functioning: Evidence  
693 from a Cerebellar Transcranial Direct Current Stimulation and Attention Network Test Study.  
694 *The Cerebellum* **18**, 457–468 (2019).

## REWARD LEARNING IN THE CEREBELLUM

- 695 54. Litman, L., Robinson, J. & Abberbock, T. TurkPrime.com: A versatile crowdsourcing data  
696 acquisition platform for the behavioral sciences. *Behav. Res. Methods* **49**, 433–442 (2017).
- 697 55. Hoffman, M. D. & Gelman, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in  
698 Hamiltonian Monte Carlo. 31.
- 699 56. Team, S. D. *Stan Reference Manual*.
- 700 57. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-  
701 out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017).
- 702 58. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.*  
703 **82**, 35–45 (1960).
- 704 59. Nassar, M. R., Wilson, R. C., Heasly, B. & Gold, J. I. An Approximately Bayesian Delta-Rule  
705 Model Explains the Dynamics of Belief Updating in a Changing Environment. *J. Neurosci.*  
706 **30**, 12366–12378 (2010).

707

708 **Tables**709 **Table 1. Basic CA participant demographics and neuropsychiatric measures**

Participant	Age (years)	Sex	Diagnosis	CCAS	MoCA	BDI	QUIP	FDS	BDS	TMTA (sec)	TMTB (sec)
Participant 1	40	M	SCA3	59	21	15	40	8	5	50	150
Participant 2	33	M	SCA3	77	21	9	12	11	5	21	261
Participant 3	20	F	SCA2	92	27	23	10	13	4	41	71
Participant 4	52	F	MSA-C	85	27	4	0	8	3	23	158
Participant 5	61	F	SCA2	92	23	15	38	8	4	49	169
Participant 6	56	M	MSA-C	100	26	7	8	13	4	66	127
Participant 7	52	M	SCA2	62	26	13	55	11	4	105	287
Participant 8	41	F	SCA2	95	29	18	2	10	8	45	97
Participant 9	43	M	SCA1	87	27	0	6	10	6	52	104
Participant 10	62	F	MSA-C	70	21	4	25	6	6	45	121
Participant 11	54	M	SCA2	72	21	0	5	11	5	32	100
Participant 12	67	F	ILOCA	101	29	12	16	12	5	51	88
Participant 13	60	F	SCA3	104	28	1	0	14	9	42	113
Participant 14	51	F	SCA10	74	25	13	8	8	2	35	83
Participant 15	66	M	SCA1	60	23	4	20	7	3	66	127
Participant 16	49	M	IMCA	86	28	17	6	13	7	81	225
Participant 17	54	M	FA	98	26	24	8	10	8	50	80
Participant 18	33	F	FA	84	27	3	26	9	4	38	84
Participant 19	54	F	IMCA	113	28	18	4	11	8	33	62

710 SCA – Spinocerebellar ataxias, MSA-C – Multiple system atrophy, cerebellar type, ILOCA – Idiopathic late onset cerebellar ataxia, IMCA –  
 711 Immune-mediated cerebellar ataxia, FA – Friedreich's ataxia, CCAS – Cerebellar cognitive affective/Schmahmann syndrome scale, MoCA –  
 712 Montreal cognitive assessment, BDI – Beck's depression inventory, QUIP – Questionnaire for impulsive-compulsive disorders in Parkinson's  
 713 disease, FDS – Forward digit span, BDS – Backward digit span, TMTA – Trail making test part A, TMTB – Trail making test part B

## REWARD LEARNING IN THE CEREBELLUM

715 **Table 2. Neuropsychological test regression analysis results**  
716

Measure	$\beta$ Estimate	95% Credible Interval
Backwards	-2.57	[-4.18, -0.92]
Forwards	-0.09	[-1.35, 1.15]
Semantic	1.54	[-2.12, 5.14]
Phonemic	0.04	[-2.75, 2.88]
Category	1.68	[-0.76, 4.10]
Similarities	-0.20	[-0.50, 0.12]
Go No Go	0.0001	[-0.33, 0.33]

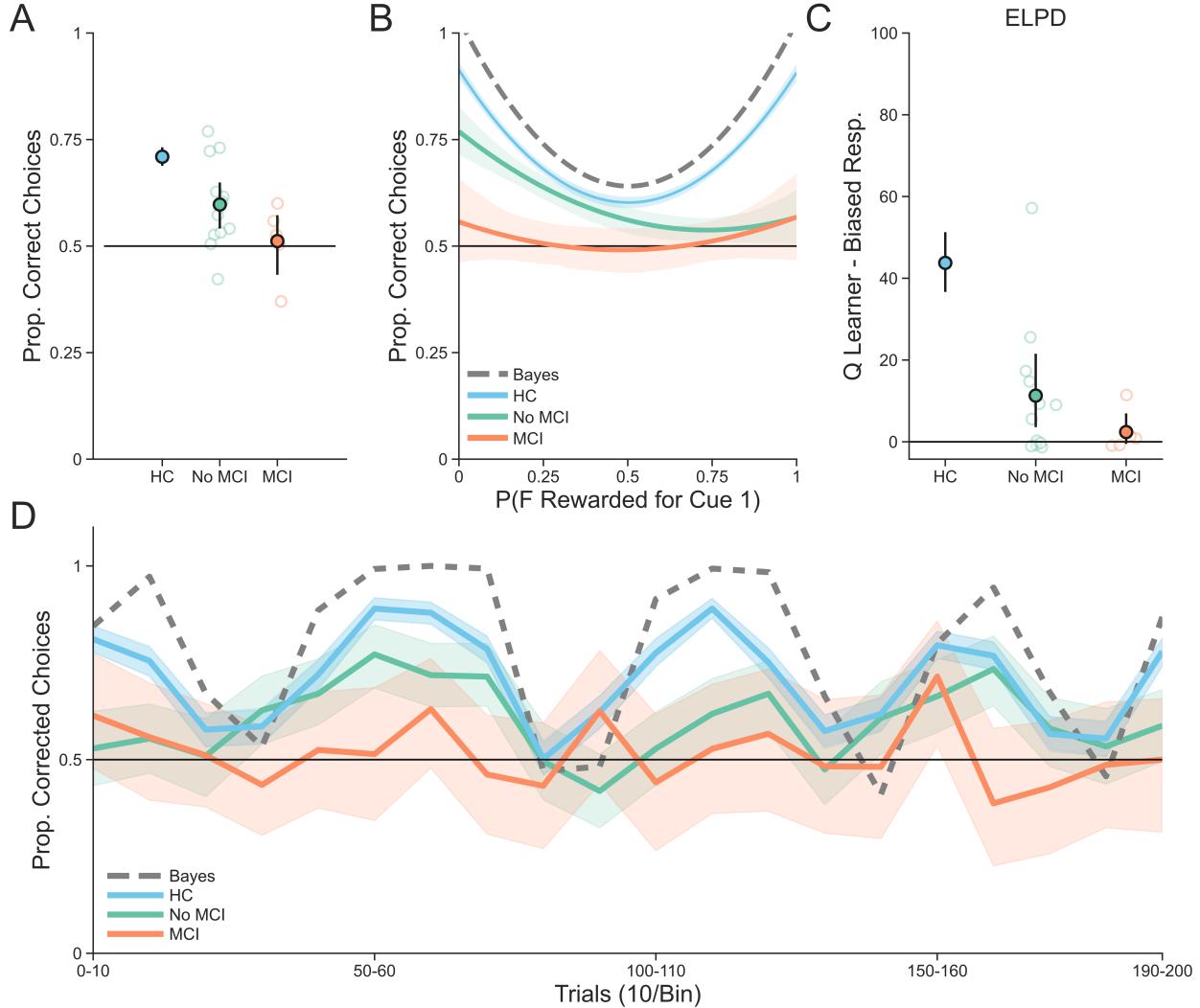
717 Results of regression analyses assessing differences in neuropsychological test performance between CA participants and healthy controls.

## REWARD LEARNING IN THE CEREBELLUM

718 **Table 3. Correlations between CA participant-level measures and incremental value**  
 719 **sensitivity**

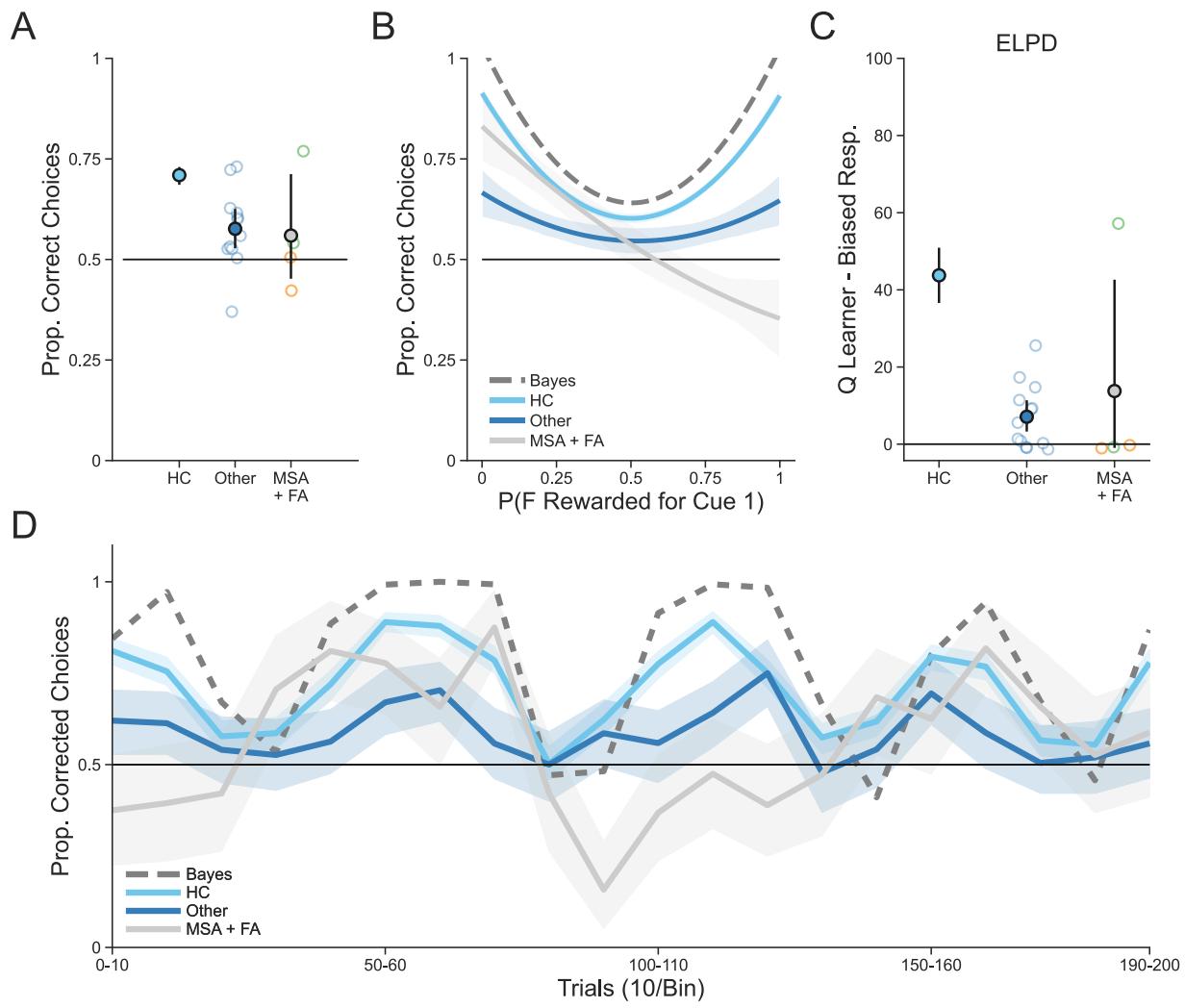
Measure	Pearson's R	P Value (Bonferroni Corrected)	Task
Symptom Duration	-0.0661	1	Incremental Learning
SARA (Total)	0.1067	1	Incremental Learning
MoCA (Total)	0.5225	0.1885	Incremental Learning
<b>CCAS (Total)</b>	<b>0.8419</b>	<b>0.0001*</b>	<b>Incremental Learning</b>
BDI (Total)	0.2977	1	Incremental Learning
QUIP (Total)	-0.3892	0.7356	Incremental Learning
Symptom Duration	-0.3699	0.7848	Multiple Learning Strategies
SARA (Total)	-0.2141	1	Multiple Learning Strategies
MOCA (Total)	0.1438	1	Multiple Learning Strategies
CCAS (Total)	0.3849	0.6887	Multiple Learning Strategies
BDI (Total)	0.4641	0.3141	Multiple Learning Strategies
QUIP (Total)	-0.074	1	Multiple Learning Strategies

720 CA participant-level measures consist of total neuropsychological scores and symptom duration and incremental value sensitivity consists of  
 721 estimates by the Q Learning model in the incremental learning task and as by the Hybrid Q Learning model in the multiple learning strategies task.  
 722 CCAS – Cerebellar cognitive affective/Schmahmann syndrome scale, SARA – Scale for the Assessment and Rating of Ataxia, MoCA – Montreal  
 723 cognitive assessment, BDI – Beck's depression inventory, QUIP – Questionnaire for impulsive-compulsive disorders in Parkinson's disease. \*\*\*p  
 724 < 0.001

725 **Supplementary Information**

726  
 727 **Figure 2—Figure supplement 1. CA participant performance on the incremental learning task**  
 728 **separated by mild cognitive impairment (MCI versus all others) and compared to healthy controls**  
 729 **(HC).** (A) Performance on the incremental learning task averaged across all trials. Individual points are  
 730 averages for each subject and filled in points represent group-level averages. Error bars are 95% confidence  
 731 intervals. (B) Performance on the incremental learning task as a function of task difficulty, which is indexed  
 732 by the true underlying probability that pressing the F key was the correct response (>50%) on each trial.  
 733 Points represent group level averages from 13 bins with an equal number of trials, lines represent the fit of  
 734 a second-order linear model, and error bars and bands represent 95% confidence intervals. (C) The  
 735 difference in estimated out-of-sample predictive performance (as measured by expected log pointwise  
 736 predictive density; ELPD) between the Q Learner and the Biased Responder model for each group.  
 737 Individual points are the ELPD difference for each subject and filled in points represent group-level  
 738 averages. Error bars are 95% confidence intervals. (D) Performance on the incremental learning task over  
 739 time. Each timepoint represents ten trials. Lines are group averages and bands are 95% confidence intervals.  
 740 For normative comparison, the performance of the Bayesian observer is shown as a dotted gray line.

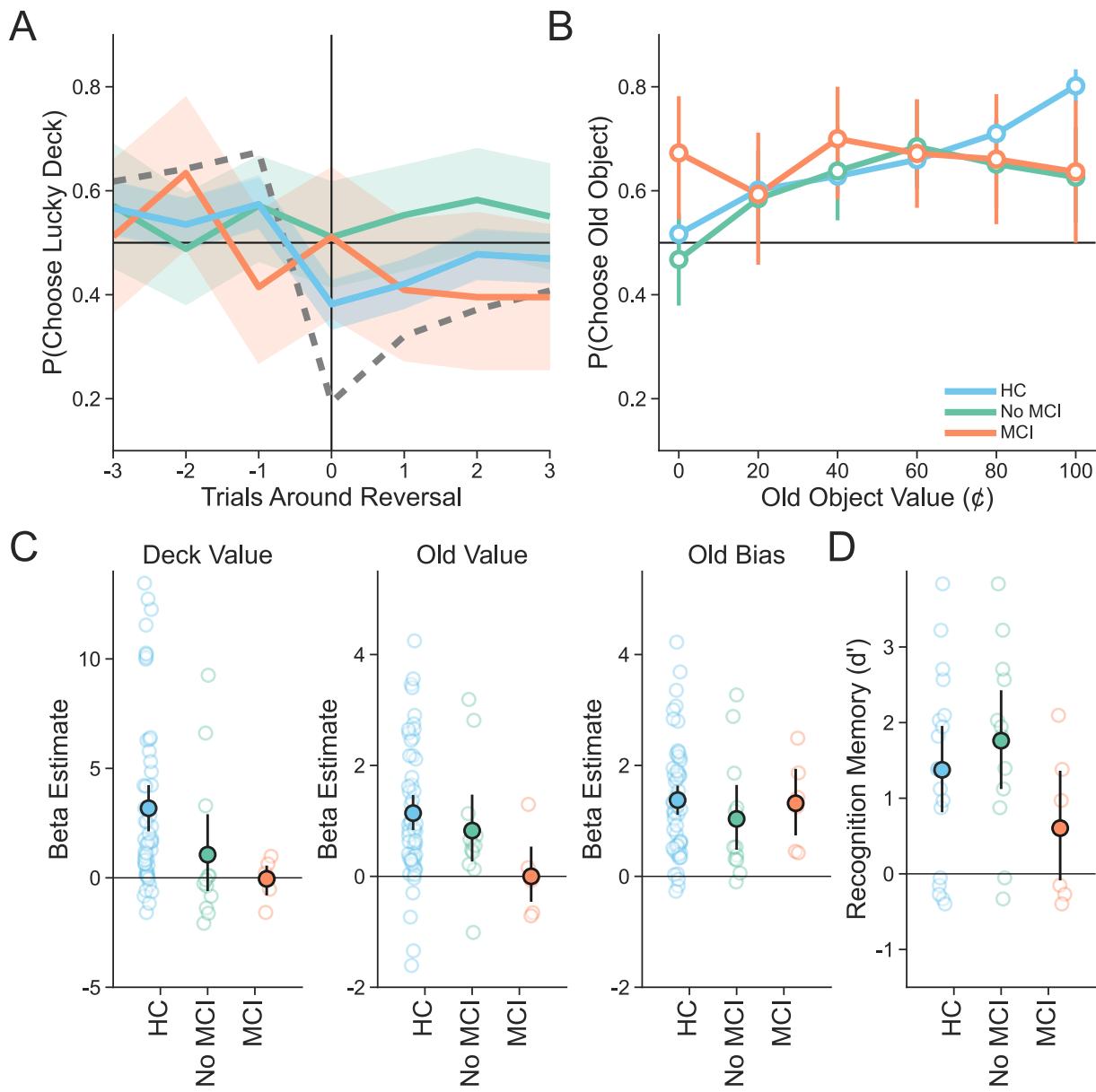
## REWARD LEARNING IN THE CEREBELLUM



741

742 **Figure 2—Figure supplement 2. CA participant performance on the incremental learning task**  
 743 **separated by diagnosis (MSA + FA individuals versus all others) and compared to healthy controls**  
 744 **(HC).** (A) Performance on the incremental learning task averaged across all trials. Individual points are  
 745 averages for each subject and filled in points represent group-level averages. Error bars are 95% confidence  
 746 intervals. MSA participants are shown in green and FA participants are shown in orange. (B) Performance  
 747 on the incremental learning task as a function of task difficulty, which is indexed by the true underlying  
 748 probability that pressing the F key was the correct response (>50%) on each trial. Points represent group  
 749 level averages from 13 bins with an equal number of trials, lines represent the fit of a second-order linear  
 750 model, and error bars and bands represent 95% confidence intervals. (C) The difference in estimated out-  
 751 of-sample predictive performance (as measured by expected log pointwise predictive density; ELPD)  
 752 between the Q Learner and the Biased Responder model for each group. Individual points are the ELPD  
 753 difference for each subject and filled in points represent group-level averages. Error bars are 95%  
 754 confidence intervals. (D) Performance on the incremental learning task over time. Each timepoint represents  
 755 ten trials. Lines are group averages and bands are 95% confidence intervals. For normative comparison, the  
 756 performance of the Bayesian observer is shown as a dotted gray line.

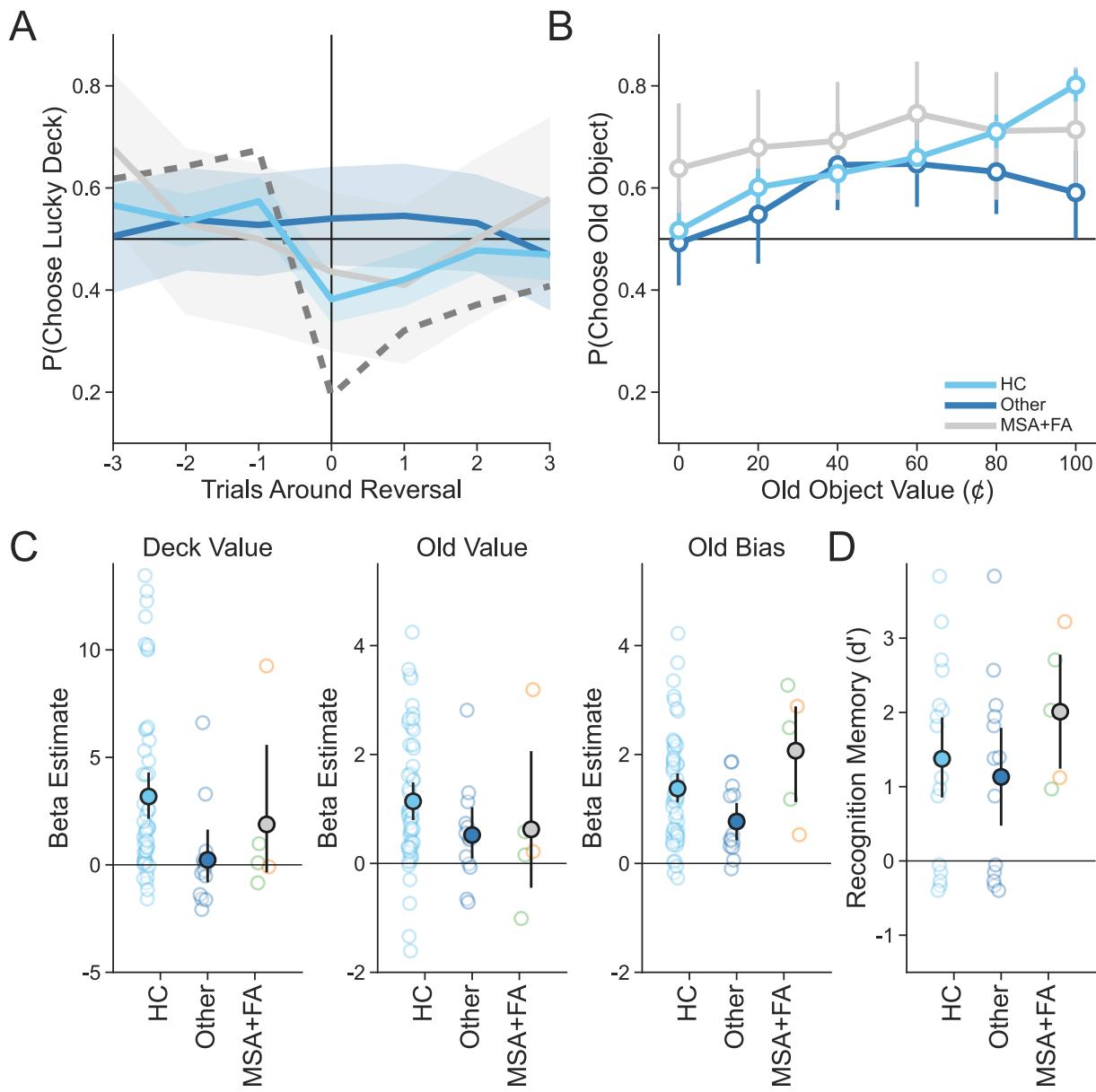
# REWARD LEARNING IN THE CEREBELLUM



757

758 **Figure 3—Figure supplement 1. CA participant performance on the multiple learning strategies task**  
 759 **separated by mild cognitive impairment (MCI versus all others) and compared to healthy controls**  
 760 **(HC).** (A) Deck learning performance on the multiple learning strategies task as indicated by the proportion  
 761 of trials on which the currently lucky deck was chosen as a function of how distant those trials were from  
 762 a reversal in deck value. Lines represent group averages and bands represent 95% confidence intervals. (B)  
 763 Object value usage on trials in which a previously seen object appeared. Points represent group averages  
 764 and error bars represent 95% confidence intervals. (C) Inverse temperature estimates from the Hybrid  
 765 model. Individual points represent estimates for each subject, group-level averages are shown as filled in  
 766 points and error bars represent 95% confidence intervals. (D) Recognition memory performance on the  
 767 subsequent memory task. Individual points represent each participant's  $d'$  prime score, filled in points are  
 768 group-level averages and error bars are 95% confidence intervals.

REWARD LEARNING IN THE CEREBELLUM



769  
 770 **Figure 3—Figure supplement 2. CA participant performance on the multiple learning strategies task**  
 771 **separated by diagnosis (MSA + FA individuals versus all others) and compared to healthy controls**  
 772 **(HC)** (A) Deck learning performance on the multiple learning strategies task as indicated by the proportion  
 773 of trials on which the currently lucky deck was chosen as a function of how distant those trials were from  
 774 a reversal in deck value. Lines represent group averages and bands represent 95% confidence intervals.  
 775 MSA participants are shown in green and FA participants are shown in orange. (B) Object value usage on  
 776 trials in which a previously seen object appeared. Points represent group averages and error bars represent  
 777 95% confidence intervals. (C) Inverse temperature estimates from the Hybrid model. Individual points  
 778 represent estimates for each subject, group-level averages are shown as filled in points and error bars  
 779 represent 95% confidence intervals. (D) Recognition memory performance on the subsequent memory task.  
 780 Individual points represent each participant's  $d'$  prime score, filled in points are group-level averages and  
 781 error bars are 95% confidence intervals.

## REWARD LEARNING IN THE CEREBELLUM

782 **Table 3—Table supplement 1. Correlations between CA participant CCAS subscale**  
783 **measures and incremental value sensitivity**

CCAS Measure	Pearson's R	P Value (Bonferroni Corrected)
Semantic Fluency	0.6693	0.033*
Phonetic Fluency	0.6235	0.0749
Category Switching	0.7168	0.012*
Digit Span Fwd (CCAS)	0.295	1
Digit Span Bwd (CCAS)	0.3146	1
Cube Drawing	0.5009	0.4055
Verbal Recall	0.4224	0.9118
Similarities	0.5881	0.1302
Go No Go	-0.2654	1
Affect	0.2944	1

784 Incremental value sensitivity consists of estimates by the Q Learning model in the incremental learning task. \* $p < 0.05$

785

786 **Appendix A**

787 **Results excluding CA participants with MCI**

788 In order to ensure that our results were not biased by patients with mild cognitive impairment  
 789 (scoring <26 on the MoCA), here we report all primary analyses with these participants (1, 2, 5,  
 790 10, 11, 14, and 15) excluded. Further mention of CA participants in this section refers to all CA  
 791 participants excluding these seven participants and their matched controls.

792 In the incremental learning task, CA participants made overall fewer correct choices compared to  
 793 healthy controls ( $\beta_{Group} = -0.10$ , 95% CI = [-0.02, -0.18]; **Figure 2—Figure Supplement**  
 794 **1A**). CA participants' choices were also less correct during periods of learning where action-  
 795 outcome contingencies were more deterministic (e.g. close to 100%) compared to more difficult  
 796 periods of learning ( $\beta_{Group \times pFReward1^2} = -5.69$ , 95% CI = [-8.37, -3.20]; **Figure 2—Figure**  
 797 **Supplement 1B-C**). Overall, this difference in performance indicates that CA participants did not  
 798 learn from reward feedback. CA participants responded slightly more slowly than healthy controls  
 799 on this task ( $\beta_{Group} = -147.69$ , 95% CI = [-226.03, -72.97]). Controls were also much better  
 800 fit by the Q learner compared to the biased responder, while this improvement in fit was largely  
 801 absent for CA participants ( $\beta_{Group} = 25.54$ , 95% CI = [9.21, 41.42]).

802 In the multiple learning strategies task, CA participants were less responsive to reward outcomes  
 803 compared to controls (**Figure 3—Figure Supplement 1A**). Specifically, controls were disrupted  
 804 more than CA participants by reversals ( $\beta_{Group \times t=0} = -0.71$ , 95% CI = [-1.21, -0.22]) and  
 805 remained below chance performance after a reversal occurred ( $\beta_{Group \times t=t+1} = -0.51$ , 95% CI =  
 806 [-1.00, -0.03]). The model of hybrid choice outperformed the baseline biased responder for both  
 807 CA participants and controls as there was no difference between groups in estimated out-of-sample  
 808 predictive performance ( $\beta_{Group} = -7.01$ , 95% CI = [-20.17, 5.47]). While controls  
 809 incorporated deck value into their decisions ( $\beta_{HC} = 3.05$ , 95% CI = [1.87, 4.24]), CA  
 810 participants generally did not ( $\beta_{CA} = 1.04$ , 95% CI = [-1.20, 3.22]; **Figure 3—Figure**  
 811 **Supplement 1C**). CA participants and controls were both sensitive to episodic value ( $\beta_{HC} =$   
 812  $1.41$ , 95% CI = [1.69, 1.12];  $\beta_{CA} = 1.03$ , 95% CI = [0.31, 1.76]; **Figure 3—Figure**

813 **Supplement 1B-C)** and were both similarly biased by previously seen objects regardless of their  
 814 value ( $\beta_{HC} = 1.18$ , 95% CI = [0.77, 1.58];  $\beta_{CA} = 0.83$ , 95% CI = [0.14, 1.53]). CA  
 815 participants and healthy controls demonstrated no differences in reaction time on this task  
 816 ( $\beta_{Group} = -7.91$ , 95% CI = [-173.68, 157.65]). Lastly, there was no difference in recognition  
 817 memory performance between groups ( $\beta_{Group} = -0.59$ , 95% CI = [-1.34, 0.157]).

## 818 Appendix B

### 819 Results excluding CA participants with MSA + FA

820 In order to ensure that our results were not biased by patients with diagnoses that were not  
 821 restricted to the cerebellum (those with MSA or FA), here was report all primary analyses with  
 822 these participants (4, 6, 10, 17, 18) excluded. Further mention of CA participants in this section  
 823 refers to all CA participants excluding these five participants and their matched controls.

824 In the incremental learning task, CA participants made overall fewer correct choices compared to  
 825 healthy controls ( $\beta_{Group} = -1.27$ , 95% CI = [-0.74, -1.82]; **Figure 2—Figure Supplement**  
 826 **2A**). CA participants' choices were also less correct during periods of learning where action-  
 827 outcome contingencies were more deterministic (e.g. close to 100%) compared to more difficult  
 828 periods of learning ( $\beta_{Group \times pFReward1^2} = -4.76$ , 95% CI = [-6.32, -3.08]; **Figure 2—Figure**  
 829 **Supplement 2B-C**). Overall, this difference in performance indicates that CA participants did not  
 830 learn from reward feedback. CA participants responded slightly more slowly than healthy controls  
 831 on this task ( $\beta_{Group} = -112.40$ , 95% CI = [-217.35, -5.03]). Controls were also much better  
 832 fit by the Q learner compared to the biased responder, while this improvement in fit was largely  
 833 absent for CA participants ( $\beta_{Group} = 37.30$ , 95% CI = [20.41, 54.63]).

834 In the multiple learning strategies task, CA participants were less responsive to reward outcomes  
 835 compared to controls (**Figure 3—Figure Supplement 2A**). Specifically, controls were disrupted  
 836 more than CA participants by reversals ( $\beta_{Group \times t=0} = -1.15$ , 95% CI = [-1.66, -0.65]) and  
 837 remained below chance performance after a reversal occurred ( $\beta_{Group \times t=t+1} = -0.66$ , 95% CI =  
 838 [-1.12, -0.20]). The model of hybrid choice outperformed the baseline biased responder for both  
 839 CA participants and controls as there was no difference between groups in estimated out-of-sample

## REWARD LEARNING IN THE CEREBELLUM

predictive performance ( $\beta_{Group} = 3.96$ , 95% CI = [-7.29, 14.89]). While controls incorporated deck value into their decisions ( $\beta_{HC} = 2.70$ , 95% CI = [1.57, 3.92]), CA participants generally did not ( $\beta_{CA} = 0.22$ , 95% CI = [-1.36, 1.71]; **Figure 3—Figure Supplement 2C**). CA participants and controls were both sensitive to episodic value ( $\beta_{HC} = 1.31$ , 95% CI = [1.63, 0.98];  $\beta_{CA} = 0.77$ , 95% CI = [0.37, 1.15]; **Figure 3—Figure Supplement 2B-C**). Controls were biased by previously seen objects regardless of their value ( $\beta_{HC} = 1.35$ , 95% CI = [1.75, 0.95]), but CA participants demonstrated this effect less strongly ( $\beta_{CA} = 0.51$ , 95% CI = [1.09, -0.05]).

CA participants and healthy controls demonstrated no differences in reaction time on this task ( $\beta_{Group} = -39.00$ , 95% CI = [-192.12, 113.59]). Lastly, there was no difference in recognition memory performance between groups ( $\beta_{Group} = -0.27$ , 95% CI = [-0.97, 0.470]).