# Classical conditioning drives learned reward prediction signals in climbing fibers across the lateral cerebellum

William Heffley and Court Hull*

*Department of Neurobiology, Duke University Medical School, Durham NC, United States*

*Correspondence: Court Hull

Email: Hull@neuro.duke.edu

## Abstract

Classical models of cerebellar learning posit that climbing fibers operate according to a supervised learning rule to instruct changes in motor output by signaling the occurrence of movement errors. However, cerebellar output is also associated with non-motor behaviors, and recently with modulating reward association pathways in the VTA. To test how the cerebellum processes reward related signals in the same type of classical conditioning behavior typically studied to evaluate reward processing in the VTA and striatum, we have used calcium imaging to visualize instructional signals carried by climbing fibers across the lateral cerebellum before and after learning. We find distinct climbing fiber responses in three lateral cerebellar regions that can each signal reward prediction, but not reward prediction errors per se. These instructional signals are well suited to guide cerebellar learning based on reward expectation and enable a cerebellar contribution to reward driven behaviors.

## Introduction

The cerebellum plays a key role in motor control and motor learning. However, it is becoming widely appreciated that the cerebellum also contributes significantly to many non-motor processes (Schmahmann, 1991) such as cognition (Kim et al., 1994), social processing (Van Overwalle et al., 2014), aggression (Reis et al., 1973) and emotion (Schmahmann and Caplan, 2006). Moreover, cerebellar deficits are associated with non-motor conditions including autism spectrum disorders (Wang et al., 2014), deficits of language processing and vocal learning (Ackermann, 2008), schizophrenia (Mothersill et al., 2016) and other impairments (Ivry and Spencer, 2004). Together, these observations strongly suggest that the cerebellum, and cerebellar learning, must operate in a manner compatible with adjusting motor as well as non-motor processes.

Classical models posit that cerebellar learning is instructed by signals carried by an afferent fiber projection from the inferior olive called climbing fibers (Albus, 1971; Ito, 1972; Marr, 1969). Such models suggest that climbing fibers operate according to a supervised learning rule to instruct changes in motor output by signaling the occurrence of movement errors. However, it is unclear how such a supervised learning rule could contribute to modification of many non-motor behaviors. In line with this view, accumulating evidence has demonstrated a wide range of climbing fiber responses that are distinct from motor error signals (Kitazawa et al., 1998; Streng et al., 2017), including prediction error and reward related signals (Heffley et al., 2018; Larry et al., 2019; Ohmae and Medina, 2015). Together, such data provide evidence that the cerebellum

may utilize alternative learning rules to support a wider range of behaviors than can be driven by error-based supervised learning alone.

Under what conditions might the climbing fiber system instruct alternative learning rules? Rabies tracing experiments have identified a disynaptic pathway from the lateral cerebellum to the striatum (Hoshi et al., 2005), and recent functional connectivity experiments have revealed a direct, monosynaptic connection from the lateral cerebellum to the ventral tegmental area (VTA) (Carta et al., 2019). These experiments also showed that lateral cerebellar output can modulate reward-driven behaviors by the same pathway. If the cerebellum acts to modulate striatal and VTA pathways involved in reward-driven behaviors, it must contain information related such behaviors, and likely mediate cerebellar learning in manner consistent with reward-based associations. Indeed, there is evidence that cerebellar granule cells can display reward related signals (Wagner et al., 2017), and climbing fibers can display signals consistent with reward prediction errors in certain behaviors (Heffley et al., 2018; Ohmae and Medina, 2015). However, it remains unclear how cerebellar learning, and in particular, how learning signals provided by cerebellar climbing fibers, operate during the types of classical conditioning behaviors classically studied to evaluate reward processing and learning in the VTA and striatum.

A key first step in understanding how cerebellar learning can contribute to reward processing and reward driven behaviors is to test what instructional signals are carried by cerebellar climbing fibers in behavioral paradigms known to involve striatal computations. To this end, we have utilized an appetitive classical conditioning task that mirrors well-studied reward prediction paradigms (Schultz et al., 1997), and performed *in vivo* calcium imaging to measure cerebellar climbing fiber driven complex spike (CpS) activity across the lateral cerebellum in awake mice. These experiments reveal distinct, functionally compartmentalized, reward-related CpS signals in each of the lateral cerebellar lobules. Specifically, in Lobule Simplex (LS), Crus I and Crus II, robust CpS responses emerge with learning in response to a visual cue that predicts reward. In LS and Crus II, unexpected reward produces CpS responses, both before and after learning. In contrast, Crus I shows no CpS response to unexpected reward in naïve animals, but a robust CpS response to unexpected reward once animals have associated a visual cue with reward. Imaging CpS activity across learning suggests that the same climbing fibers that respond to reward in naïve animals can instead respond to a reward predictive cue after learning. Finally, we observe no evidence of negative prediction error following reward omission, suggesting that in this paradigm, CpS responses may not instruct a strict temporal difference (TD) learning model. The learned, reward predictive CpS signals observed here are, however, consistent with a model

that would allow learned cerebellar output to provide information about rewarding stimuli to downstream structures such as the VTA and striatum. Hence, these data provide a key link between the instructional signals known to be involved in cerebellar learning and the types of reward based processing observed in downstream areas.

## Results

### An appetitive classical conditioning paradigm to investigate climbing fiber responses in the lateral cerebellum

To investigate climbing fiber driven CpS activity during an appetitive classical conditioning paradigm that parallels tasks typically used to study reward processing in the VTA and striatum, we first established an appropriate behavior for head-fixed mice that is compatible with calcium imaging in the lateral cerebellum (**Fig 1A-D**). In this task, mice learn a simple association between a visual cue and the timing of subsequent reward delivery. Specifically, mice were trained by viewing a continuously present, full screen vertical grating that was transiently (100 ms) replaced with a horizontal grating to signal the delivery of a saccharine reward at a 600 ms delay (**Fig 1D, top**). After multiple training sessions (mean = 5.9 ±0.5 days), mice learned the association between the visual cue and the time of reward delivery, significantly reducing both their reaction times (Day 1, 730.7 ±26.9 ms; Day N+1, 545.8 ±14.3 ms; $p=1.64\times10^{-6}$, n = 24, paired t-test, **Fig. 1 E**) and miss rates (Day 1, 0.40±0.04; Day N+1, 0.03±0.01; $p=6.25\times10^{-9}$, paired t-test, **Fig. 1 F**). In addition to developing well-timed licking responses that approximated the time of reward delivery on rewarded trials, mice also developed robust licking responses following the visual cue on omission trials after learning that were absent in naïve animals (**Fig. 1G**). Together, these data indicate that mice successfully learn the association between visual cue and the timing of reward delivery in this paradigm.

### CpS activity shifts earlier in time after learning

We began by measuring CpS activity before and after learning in lobule simplex (LS), where previous work has identified reward related activity in granule cells and climbing fibers after learning in different behavioral paradigms (Heffley et al., 2018; Wagner et al., 2017). To obtain a broad view of CpS activity across LS and assess whether any reward-related signals are carried by climbing fibers in this paradigm, we first performed single-photon, mesoscale calcium imaging of Purkinje cell (PC) dendrites expressing GCaMP6f. This method has previously been demonstrated to accurately report CpS activity (Heffley et al., 2018). Using this approach, we

observed a significant increase in the CpS activity following reward delivery in naïve animals (n = 308 trials, 5 animals, p =5x$10^{-19}$, paired t-test), but no significant response on reward omission trials (n =74 trials, 5 animals, p =0.595, paired t-test) (**Fig 2A, top**). After learning, CpS activity still maintained a significant response (n =749 trials, 6 animals, p = 3.3 x$10^{-100}$, paired t-test), but was shifted in time, and preceded reward delivery on rewarded trials (**Fig. 2A, bottom, Fig 2B**). Moreover, after learning there was also a significant response on omission trials preceding reward delivery (n =156 trials, 6 animals, p = 5.4 x $10^{-16}$, paired t-test) (**Fig 2A, bottom, Fig. 2C**). Across experiments, the latency of CpS activity decreased significantly (n = 6 animals, p=1.04×$10^{-6}$, paired t-test) (**Fig 2B**). Notably, this shift in CpS response timing ($\Delta_{CpS}$ = 441 ms) was greater than the change in lick timing ($\Delta_{lick}$ = 185 ms), indicating that the CpS response did not simply follow the timing of motor output. In addition, we find that CpS responses after learning had the same timing for rewarded and omission trials (Rewarded trial latency = 383 ±15 ms, Omission trial latency = 378 ±19ms, n =6 animals, p=0.78×$10^{-6}$, paired t test) (**Fig 2C**). In contrast, when reward was delivered unexpectedly after learning (i.e. in absence of the visual cue), CpS activity followed reward delivery, and was significantly delayed compared with rewarded trials (Day N+2; rewarded trial latency = 366.1 ±31 ms, unexpected trial latency 793 ± 31ms, n = 5 animals, p=2.40×$10^{-4}$, paired t-test). The latency of CpS responses on unexpected reward trials was, however, equivalent to the timing of responses to rewarded trials in naïve animals (p=0.38, two-sample t-test) (**Fig 2D**). Together, these data reveal that complex spiking shifts in time as a result of learning, initially following reward delivery in naïve animals and later preceding reward delivery in trained animals, occurring with the same timing on both rewarded an omission trials after learning.

**CpS responses in LS signal reward predictions that emerge with learning**

Our mesoscale imaging experiments suggest that CpS activity in LS exhibits reward related activity that changes as a function of learning in manner that is consistent with reward expectation. However, to further disambiguate CpS activity related to reward, reward expectation and licking, and to test whether CpS activity in individual PC dendrites changes across learning, we next used two-photon imaging to visualize CpS responses in individual PC dendrites. As in our mesoscale imaging experiments, these data revealed significantly elevated CpS rates following reward delivery in naïve animals (n = 1040 dendrites, p = 5.5 x $10^{-48}$, one-tailed t-test) (**Fig. 3A**) but no significant response during this window on omission trials (p = 0.98). After learning, we observed significantly elevated CpS rates prior to reward delivery on both rewarded (n = 906 dendrites, p = 2.3 x $10^{-60}$, one-tailed t-test) and omission trials (p = 1.4 x $10^{-33}$) that were comparable across trial

159 types ($p = 0.01$, paired t-test; **Fig. 3B**). In addition, post-reward CpS rates were significantly
160 reduced after learning ($p = 2.0 \times 10^{-19}$, unpaired t-test), with small elevations in spiking (reward
161 trials: $p = 3.3 \times 10^{-4}$; omission trials: $p = 6.7 \times 10^{-6}$; one-tailed t-test) that were equivalent on both
162 rewarded and omission trials ($p = 0.46$; paired t-test). Notably, there were no CpS responses that
163 were reliably suppressed below baseline after the time of expected reward on omission trials (see
164 Methods). These data indicate that CpS activity shifts from an association with reward prior to
165 learning to the cue that predicts reward after learning. To test whether unique populations of PC
166 dendrites exhibit CpS responses to reward and following the visual cue after learning, we
167 identified the subset of PC dendrites with significant post-reward responses (n = 204/906; **Fig. 3**
168 **– Figure Supplement 1**). This small subset of PC dendrites also responded following the visual
169 cue ($p = 5.9 \times 10^{-36}$; one-tailed t-test), indicating that a minority of PC dendrites can exhibit CpS
170 responses to both reward and the visual cue after learning. These data suggest the possibility
171 that some CpS responses had not fully transitioned between encoding the reward and the reward-
172 predicting cue. However, the majority of PC dendrites (n = 702/906) did not respond to the reward
173 after learning, and instead only responded to the reward-predicting cue ($p = 3.9 \times 10^{-36}$; one-tailed
174 t-test; **Fig. 3– Figure Supplement 1**).
175
176 To disambiguate CpS responses related to reward versus reward expectation, we measured
177 responses to unexpected reward after learning. On these trials, CpS rates increased significantly
178 after the time of unexpected reward delivery (n = 243, $p = 1.6 \times 10^{-8}$, one-tailed t-test; **Fig. 3C**).
179 Moreover, after an additional day of training, there was no longer a significant post-reward
180 response on trials where reward was expected ($p=0.8$). These data indicate that after learning,
181 CpS responses in area LS are only reward responsive if the reward is unexpected. In the same
182 experiments, CpS rates increased significantly prior to reward delivery when reward was expected
183 ($p = 1.2 \times 10^{-12}$, one-tailed t-test; **Fig. 3 C**). Notably, the PC dendrites that were significantly
184 responsive following the reward-predicting cue (n = 84/243), were also significantly responsive
185 after the unexpected reward ($p = 2.3 \times 10^{12}$; one-tailed t-test; **Fig. 3D**). This suggests that the
186 same PC dendrites exhibit CpS responses to the reward-predictive cue and unexpected reward,
187 thus signaling predictions under two different conditions. However, lick timing also differs between
188 these two trial types, and changes across learning. Thus, we also sought to further disambiguate
189 reward related CpS responses from possible CpS responses driven by motor output during
190 licking.
191

192 To directly assess the role of licking in modulating CpS rates, we first generated a lick-triggered
193 average of CpS activity from licks generated during the inter-trial interval (ITI). These data
194 revealed no significant CpS response to licking ($p = 0.6$; $n = 768$; one-tailed t-test; **Fig. 3E**),
195 consistent with previous reports showing a lack of lick-related CpS activity in area LS (Bryant et
196 al., 2010). To further address the contribution of licking to the emergent CpS responses following
197 the reward-predicting cue after learning, we segregated rewarded trials into subsets according to
198 lick timing. Specifically, we averaged trials with the earliest and latest quartiles of licking. This
199 analysis revealed that on trials with early compared to late lick onsets, the peak of the pre-reward
200 CpS occurred at nearly the same time relative to the cue (latency on early trials: 300 ms; late
201 trials: 330 ms), and with similar, but slightly smaller, amplitude ($p = 4.4 \times 10^{-4}$; $n = 906$ dendrites;
202 paired t-test; **Fig 3F**). There was, however, a difference in CpS rate following reward delivery ($p$
203 $= 8.8 \times 10^{-9}$), likely due to the increased variability in the timing of reward consumption for trials
204 with late licks. These data indicate that pre-reward CpS responses are related to the cue driven
205 expectation of reward, and not motor output due to licking.
206
207 **The same climbing fibers can respond to reward and a reward-predictive visual cue**
208 While our data indicate CpS activity within individual dendrites can represent different reward-
209 related events (i.e. the reward-predicting cue and unexpected reward), these data do not address
210 whether it is the same climbing fibers that shift their responses, or whether separate populations
211 of climbing fibers generate different responses before and after learning. To address this
212 question, we measured CpS activity from PC dendrites in the same coordinates before and after
213 learning in a subset of experiments. To evaluate responses, we took two approaches. First, we
214 manually identified individual PC dendrites across imaging days and measured their CpS
215 responses before and after learning (**Fig 4A,B**). This approach revealed clear examples of single
216 PC dendrites with reward responsive CpS activity in naïve animals that became cue responsive
217 after learning. However, because this approach relies on user identification of the same PC
218 dendrites, we also sought to test an automated approach. To do so, we first independently
219 generated pixel masks for PC dendrites from our pre and post-learning fields of view (**Fig 4C-D**).
220 We then motion registered the mean fluorescence image from the post-learning session to the
221 pre-learning session (**Fig 4E**), and applied the same pixel shifts from this registration to the
222 segmented pixel masks from the post-learning condition. By overlaying the shifted post-learning
223 pixel masks with the pre-learning masks, we assessed the overlap between imaging days (**Fig**
224 **4F-H**). Despite closely matched fields of view, this process lead to a majority of PC dendrite masks
225 with little overlap. We thus employed a stringent criteria of greater than 50% overlap to increase

226 the probability of accurately identifying the same PC dendrites across imaging sessions. With
227 these criteria, we identified a group of PC dendrites likely corresponding to the same PC dendrites
228 before and after learning. In these PC dendrites, as in our other experiments, we find that CpS
229 activity was initially absent following the visual cue in naïve animals ($p = 0.59$; $n = 61$; one-tailed
230 t-test; **Fig. 4I**), and that cue driven responses emerged with learning ($p = 1.2 \times 10^{-8}$). We note that
231 because of the tight packing and similar size and orientation of PC dendrites, such approaches
232 cannot provide definitive evidence of tracking the same PC dendrites across learning. However,
233 these data support the conclusion that individual CFs can develop a learned response to the
234 reward predictive cue as a result of learning.
235

236 **<u>Crus I and Crus II exhibit reward-related CpS responses that are distinct from area LS</u>**

237 Anatomical and physiological evidence suggests that CpS responses are functionally
238 compartmentalized across the medio-lateral axis of the cerebellum (Apps and Garwicz, 2005),
239 raising the possibility of different computations and diverse reward-related CpS activity patterns
240 in lobules other than LS. Thus, we next measured CpS responses in Crus I and Crus II .
241

242 In Crus I, CpS activity differed substantially from LS. Specifically, in naïve animals, CpS activity
243 was sensory-related, and followed the visual cue at a time well before reward delivery or licking
244 on both rewarded ($p = 9.5 \times 10^{-17}$; $n = 105$; one-tailed t-test; **Fig. 5A**) and omission trials ($p = 7.9$
245 $\times 10^{-14}$). After learning, however, once the cue had been associated with upcoming reward, CpS
246 responses to the visual cue were significantly enhanced ($p = 1.4 \times 10^{-6}$; unpaired t-test; $n = 105$
247 and 151; **Fig. 5B**). In contrast, there were only weak or non-significant reward responses on
248 rewarded or omission trials, either before (rewarded: $p = 0.005$; omission: $p = 0.23$; $n = 105$; one-
249 tailed t-test; **Fig. 5A,B**) or after learning (rewarded: $p = 1.0$; omission: $p = 0.99$; $n = 151$; one-
250 tailed t-test). Again, however, we did not identify any PC dendrites that were reliably suppressed
251 at the time of expected reward after learning. Notably, despite the lack of a CpS response to
252 unexpected reward in naïve animals, there was a large CpS response to unexpected reward in
253 the same animals after learning ($p = 1.5 \times 10^{-12}$; $n = 120$; one-tailed t-test; **Fig. 5C**). These data
254 suggest that in Crus I, where CpS responses are sensory related in naïve animals, activity can
255 also be generated by expectations that are specifically linked to a learned sensory cue after
256 learning.
257

258 To test whether the same PC dendrites that exhibit CpS responses following the visual cue also
259 respond to unexpected reward, we identified those PC dendrites that were significantly cue

responsive. Indeed, these PC dendrites (n = 83/110) also showed a robust response to unexpected reward ($p = 4.7 \times 10^{-11}$; one-tailed t-test; **Fig. 5D**), indicating that PC dendrites can signal expectation under distinct conditions. As further evidence that CpS responses in Crus I are sensory related in naïve animals, there was no significant CpS response to licks during the ITI (p = 0.05; n = 71; one-tailed t-test; **Fig 5E**). To test whether licking contributes to the learned increase in response amplitude following the cue, we segregated trials according to the timing of the first lick. This revealed that the timing (early and late lick latency: 230 ms) and amplitude (p = 0.02, n = 151; paired t-test; **Fig 5F**) of CpS activity following the visual cue was largely independent of lick timing. Together, these data indicate that CpS activity in Crus I is sensory related in naïve animals, but can also be modulated by reward expectation because it is enhanced by associating the visual cue with reward, and driven by delivering a reward unexpectedly.

Finally, we measured CpS activity in Crus II. Similar to area LS, CpS rates were elevated in response to reward delivery in naïve animals ($p = 1.3 \times 10^{-6}$; n = 58; one-tailed t-test; **Fig. 6A**), but not on reward omission trials (p = 0.98). After learning, a robust cue response was observed in Crus II on both rewarded ($p = 9.5 \times 10^{-13}$; n = 148; one-tailed t-test; **Fig. 6B**) and omission trials ($p = 2.9 \times 10^{-14}$). However, distinct from area LS, the reward response persisted after learning ($p = 3.3 \times 10^{-12}$; **Fig. 6B**). Again, however, we observed no significant increased response on omission trials at the time of expected reward (p = 0.99); nor did we observe any significant suppression. When reward was delivered unexpectedly after learning, we observed enhanced CpS rates as compared to the CpS responses to expected reward ($p = 4.5 \times 10^{-5}$, n = 115; **Fig. 6C**), suggesting a component of the learned response that scales with expectation. Notably, the same PC dendrites that respond to the cue after learning (n = 41/115) continue to respond to both the expected ($p = 1.4 \times 10^{-5}$; one-tailed t-test; **Fig. 6D**) and unexpected reward ($p = 3.0 \times 10^{-11}$) after learning.

In contrast with area LS and Crus I, in naïve mice we observed a significant lick related response in Crus II as measured by the lick triggered average of CpS activity during the ITI (p = 0.003, n = 40; one-tailed t-test; **Fig. 6E**). This result is consistent with previous reports suggesting that CpS activity can be triggered by licking in this area (Gaffield et al., 2016; Gaffield et al., 2018; Welsh et al., 1995). Interestingly, however, these responses disappeared after learning (p = 1; n = 78). Moreover, after learning, neither the timing (latency on early trials: 300 ms; late trials: 330 ms) nor the amplitude (p = 0.32; n = 148 dendrites; paired t-test; **Fig 6F**) of the CpS following the cue

293 depended on the timing of licking. These results suggest that, while some activity can be driven
294 by licking, motor output alone is not the central determinant of CpS activity in Crus II.
295

296 **Discussion**

297 We have used a classical conditioning paradigm of the kind typically studied to evaluate reward
298 signaling in the VTA and striatum to reveal that climbing fibers provide a diverse set of learned
299 reward prediction signals across the lateral cerebellum. Specifically, in LS and Crus II, CpS
300 responses emerged as a function of learning to a cue that predicted upcoming reward. In Crus I,
301 while a sensory-evoked CpS response to the visual cue was already evident in naïve animals,
302 this response was enhanced once the cue had been associated with upcoming reward. Hence,
303 in all three areas, there was a learned CpS response related to reward prediction. CpS responses
304 also exhibited significant differences across areas. In LS, the response to expected reward was
305 largely lost after learning, but was pronounced when reward was delivered unexpectedly. In Crus
306 I, while there was no reward response in naïve animals, a CpS response to unexpected reward
307 emerged only after the animals established an expectation of reward delivery when in response
308 to the visual cue. Finally, in Crus II, reward responses persisted across learning, even after a
309 separate CpS response emerged to the reward-predictive visual cue. The post-learning CpS
310 response to unexpected reward was enhanced, however, as compared to trials when reward was
311 expected.
312

313 Together, these common, learned responses to a reward-predictive cue and to unexpected
314 reward indicate that all three areas exhibit CpS responses related to reward prediction. Such
315 commonalities suggest a broad role for the lateral cerebellum in reward-based learning. The
316 differences in responses, however, also suggest that learned changes in each area are overlaid
317 on the unique sensory and motor CpS signals pre-existing in each region. Thus, these data also
318 indicate that the lateral cerebellum can harness other distinct sensory and motor signals across
319 each area for different forms of sensorimotor learning, and perhaps to utilize alternate learning
320 strategies that depend on task requirements.
321

322 In all three areas, cue and reward-associated CpS responses could be clearly distinguished from
323 motor responses due to licking, as the timing of licking did not correlate with the timing CpS
324 responses. In a previous study, we identified a subset of lick-responsive PC dendrites after
325 learning in area LS, and we concluded that such CpS responses were learned and related to
326 reward expectation rather than motor output per se (Heffley et al., 2018). Here, by imaging naïve

327 animals, we were able to demonstrate no significant lick response in absence of learning in area
328 LS and Crus I, consistent with previous reports indicating CpS responses to licking are localized
329 primarily to Crus II. Also in agreement with these data, we did observe a significant CpS response
330 to licking in Crus II in naïve animals. Interestingly, however, once the association between visual
331 cue and reward had been established, licking during the ITI when no reward was possible failed
332 to produce CpS activity. These data suggest that even in areas with motor related signals, CpS
333 responses are more complex than was previously appreciated. While earlier studies investigating
334 CpS responses to licking did not attempt to disambiguate between motor output, reward, and
335 reward expectation (Bryant et al., 2010; Gaffield et al., 2016; Gaffield et al., 2018; Welsh et al.,
336 1995), we suggest that such efforts are crucial to understanding the role CpS signals in cerebellar
337 learning.
338
339 This study is part of a growing literature that suggests new roles for climbing fiber signals distinct
340 from classical error-based supervised learning (Heffley et al., 2018; Kitazawa et al., 1998; Ohmae
341 and Medina, 2015; Streng et al., 2017). In a previous study, we identified CpS activity consistent
342 with an unsigned reinforcement learning rule in an operant conditioning task after learning (Heffley
343 et al., 2018). In this study, we measured CpS activity both before and after learning in a task with
344 different requirements, and found responses that showed many similarities but also significant
345 differences in CpS activity. In both tasks, we observed CpS responses to an event that predicted
346 upcoming reward. However, in the current study, CpS activity was linked to a visual cue that
347 predicted reward delivery, whereas previously we observed CpS activity linked to an arm
348 movement that was required to produce reward delivery. Indeed, a key difference between our
349 previous operant task and the classical conditioning task studied here is that animals do not affect
350 trial outcome with their motor output in this task. Specifically, while animals learn to lick earlier
351 after learning in this task, a change in lick timing is not required for reward delivery, and there is
352 no consequence for licks occurring with 'correct' or 'incorrect' timing (i.e. correctly or incorrectly
353 matching the time of reward delivery). Hence, we speculate that observed differences in CpS
354 activity across tasks are related to the distinct sensory input and motor output that predicts reward
355 delivery. In this task, the reward-predictive event was a visual cue alone. Previously, however,
356 despite a visual cue that instructed arm movement, correct execution of the movement was
357 required for reward delivery, and this action was therefore the necessary predictor of task
358 outcome. In support of a role for task-specific movement requirements is shaping instructional
359 signals, in the mesolimbic dopamine system it has been shown with a similar operant task that

360 dopamine release is governed by correct movement initiation, and not simply reward prediction
361 (Syed et al., 2016).
362
363 Another notable difference from our previous work is that we do not observe clear reward omission
364 responses in this task. While it is unclear why such responses are absent in the current task, we
365 again speculate that such responses may only be generated when movement is required to
366 determine task outcome. In neither task, however, did we observe decreases in CpS activity at
367 the time when reward was expected on reward omission trials. Thus, we have not observed any
368 evidence of negative prediction errors in CpS activity. These results contrast with a recent
369 eyeblink conditioning study that used an aversive stimulus to drive learning (Ohmae and Medina,
370 2015), and further suggests a key role for task parameters in establishing the specific features of
371 CpS responses. Another possibility is suggested by the mesolimbic dopamine system, where
372 neurons have different response profiles depending on their striatal projection target (Menegas
373 et al.; Parker et al., 2016). It is thus remains possible that climbing fibers also exhibit an analogous
374 target-specific response variability, and signal diverse reward-related information such as
375 negative prediction errors in other regions of the cerebellum. Lastly, recent measurements of VTA
376 dopamine neuron responses have revealed that increased firing to reward can persist after
377 learning (Coddington and Dudman, 2018), similar to what we observe in Crus II. In those
378 experiments, while negative prediction errors did emerge eventually, they occurred only in the
379 form of a reduced peak firing very late in training, indicating that over-training can be required for
380 negative prediction error to develop.
381
382 Finally, in this study we have measured CpS responses across learning, thus revealing how CpS
383 activity shifts from reward-related responses to responses driven by reward predictive cues in the
384 same animals. Our data suggest that the same climbing fibers can adjust their responses as a
385 function of learning, and can signal diverse reward expectation information (i.e. both reward-
386 predicting cues and unexpected reward in the same climbing fibers). These learned, reward-
387 predictive climbing fiber signals are ideally suited to instruct both cerebellar learning and
388 modulation of reward processing in downstream brain regions. In particular, a recent study has
389 demonstrated that output from the lateral cerebellum via the deep cerebellar nuclei (DCN) forms
390 a monosynaptic, excitatory connection onto dopamine neurons in the VTA (Carta et al., 2019).
391 Activating these DCN projections is sufficient to positively modulate reward-driven behaviors, and
392 these same projections are endogenously active under social conditions. These data strongly
393 suggest that the cerebellum must contain, or have the ability to learn, information about rewarding

394 stimuli. While speculative, such a model fits well with the observations made here. Specifically,
395 we have shown CpS instructional signals can become associated with stimuli that predict reward.
396 If combined with contextual information from the mossy fiber pathway, such CpS signals would
397 be effective in instructing plasticity to depress Purkinje cell output and thereby enable DCN output
398 to the VTA in response to reward-associated stimuli or actions. Hence, a key next step will be to
399 evaluate the intersection of cerebellar learning with cerebellar output to the mesolimbic dopamine
400 system during behavior. For the present, the patterns of CpS activity demonstrated here provide
401 key information about how instructional signals carried by cerebellar climbing fibers can change
402 as a function of learning to signal reward expectation across the lateral cerebellum.
403
404
405 **Methods**
406
407 Mice
408 All experimental procedures using animals were carried out with the approval of the Duke
409 University Animal Care and Use Committee. All experiments were performed during light cycle
410 using adult mice (>p60) of both sexes, randomly selected from breeding litters. All mice were
411 housed in a vivarium with normal light/dark cycles in cages with 1-5 mice. Imaging experiments
412 were performed using Tg(PCP2-Cre)3555Jdhu mice (Jackson Labs, 010536; n=24). A subset of
413 these animals (n=4) were the result of a genetic cross with Ai148(TIT2L-GC6f-ICL-tTA2)-D mice
414 (Jackson Labs, 030328). We used two exclusion criteria for animals in this study: (1) poor
415 recovery or other health concerns following surgical intervention or (2) missed virus injection, as
416 determined by *in vivo* imaging and post-hoc histological analysis.
417
418 Surgical Procedures
419 3-10 hours prior to surgery, animals received dexamethasone (3mg/kg) and ketoprofen (5mg/kg).
420 Surgical procedures were performed under anesthesia, using an initial dose of ketamine/xylazine
421 (50mg/kg and 5mg/kg) 5 minutes prior to surgery and sustained during surgery with 1.0-2.0%
422 isoflurane. Toe pinches and breathing were used to monitor anesthesia levels throughout
423 surgeries. Body temperature was maintained using a heating pad (TC-1000 CWE Inc.). Custom-
424 made titanium headplates (HE Parmer) were secured to the skull using Metabond (Parkell). For
425 imaging experiments, a 3mm diameter craniotomy was made over the lobule simplex at
426 approximately 1.4mm lateral and 2.8mm posterior to lambda or crus I and II at approximately
427 3.0mm lateral and 4.3mm posterior to lambda. Glass cover slips consisting of two 3mm bonded

428 to a 5mm coverslip (Warner Instruments No. 1) with index matched adhesive (Norland No. 1)
429 were secured in the craniotomy using Metabond. Buprenex (0.05mg/kg) and cefazolin (50mg/kg)
430 were administered following surgery twice a day for two days. Following a minimum of 4 recovery
431 days, animals were water deprived for 3 days, or until body weight stabilized at 85% of initial
432 weight, and were habituated to head restraint (3-5 days) prior to behavioral training.
433
434 For imaging experiments, the glass cover slip was removed following behavioral training, and
435 mice were injected (WPI UMP3) with AAV1.CAG.Flex.GCaMP6f.WPRE.SV40 (UPenn vector
436 core, titer = $1.48 \times 10^{13}$ or $7.60 \times 10^{12}$). 150 nL virus diluted 1:1-1:5 in ACSF was injected at a rate of
437 30nl/min and a depth of 150 μm at 1-3 sites in dorsal lobule simplex. Imaging was performed
438 beginning 10-14 days following injection. Transgenic mice crossed with the Ai148 mouse line
439 received no injection and were given 10 days to recover from surgery before starting training.
440
441 Behavior
442 During behavioral training, animals were head-fixed and placed in front of a computer monitor
443 and reward delivery tube. Animals were trained to associate a visual cue with a saccharine
444 reward. A high contrast vertical black and white grating was present at all times, including the
445 inter-trial interval (ITI; 24-30 s), except when the reward cue was presented. The reward cue
446 consisted of a horizontal grating, presented for 100ms, and followed reward delivery after a 500ms
447 delay (600 ms from cue onset). Training sessions lasted for 62.1 ±0.9 minutes and consisted of
448 131.8 ±1.8 trials (n = 124 sessions). Imaging sessions lasted an average of 64.1 ±1.3 minutes
449 and consisted of 130.9 ±2.8 trials (n = 62 sessions). For reward omission sessions, 20% of
450 randomly determined trials were unrewarded. For unexpected reward sessions, no reward cue
451 was presented on 20% of randomly determined trials. Behavioral parameters including cue
452 presentation, reward delivery, and licking were monitored using Mworks (http://mworks-
453 project.org) and custom software written in MATLAB (Mathworks). Learning was defined by
454 stabilization of reaction times and miss rates, with required reaction times less than 650 ms and
455 a miss rates below 15%. Across animals, 5.9 ± 0.5 training sessions (n = 24 mice) was required
456 to meet these criteria.
457
458 Calcium Imaging
459 *Mesoscale imaging*: Single photon imaging was performed using a customized microscope
460 (Sutter SOM) affixed with a 5x objective (Mitutoyo, 0.14NA) and CMOS camera (Qimaging, Rolera
461 em-$c^2$). Excitation (470 nm) was provided by an LED (ThorLabs, M470L3), and data were

462 collected through a green filter (520-536 nm band pass, Edmund Optics) at a frame rate of 10Hz,
463 with a field of view of 3.5x3.5mm at 1002x1004 pixels.
464
465 *Two-Photon Imaging*: Two-photon imaging was performed with a resonant scanning microscope
466 (Neurolabware) using a 16x water immersion objective (Nikon CFI75 LWD 16xW 0.80NA).
467 Imaging was performed using a polymer to stabilize the immersion solution (MakingCosmetics,
468 0.4% Carbomer 940). A Ti:Sapphire laser tuned to 920nm (Spectra Physics, Mai Tai eHP
469 DeepSee) was raster scanned via a resonant galvanometer (8 kHz, Cambridge Technology) at a
470 frame rate of 30 Hz with a field of view of 1030 μm x 581 μm (796 x 264 pixels). Data were
471 collected through a green filter (510 ± 42 nm band filter (Semrock)) onto GaAsP photomultipliers
472 (H10770B-40, Hamamatsu) using Scanbox software (Neurolabware). A total of 24 mice were
473 used for imaging experiments (6 mice for mesoscale imaging in LS; 15 mice for two-photon
474 imaging in LS; 4 mice for two-photon imaging in Crus I; 2 mice for two-photon imaging in Crus II).
475
476 Data Analysis and Statistics
477 *Behavior:* Analysis was performed using custom MATLAB code. Reaction times were defined as
478 the first lick occurring 200-1000ms after reward cue onset. Miss rate was defined as the percent
479 of trials without a lick in the 1s following cue presentation.
480
481 *Mesoscale imaging:* Imaging analysis was performed using custom MATLAB code. Regions of
482 interest (ROIs) were selected within LS according to the pattern of GCaMP expression. Window
483 location and the lobule identity were identified according to folia patterns visible through the
484 cranial window, landmarks recorded during surgery, and post-hoc histology. Baseline
485 fluorescence (F) was measured on a trial-by-trial basis during the ITI as the mean F between 900
486 and 200ms before trial initiation. Normalized fluorescence ($\Delta F/F$) was calculated according to the
487 cumulative activity within an ROI. $\Delta F/F$ latency was measured by first finding the peak $\Delta F/F$
488 response following cue onset averaged across trials. Next, the latency was measured as the time
489 when mean $\Delta F/F$ trace achieved a maximum rate of change (the peak of the first derivative) in a
490 500ms window preceding the identified peak $\Delta F/F$ response.
491
492 *Two-photon imaging*: Motion in the X and Y planes was corrected by sub-pixel image registration.
493 To isolate signals from individual PC dendrites, we utilized principal component analysis (PCA)
494 followed by independent component analysis (ICA). Final dendrite segmentation was achieved
495 by thresholding the smoothed spatial filters from ICA. A binary mask was created by combining

highly correlated pixels (correlation coefficient > 0.8) and removing any overlapped regions between segmented dendrites. Notably, image segmentation using these criteria did not extract PC soma, which were clearly visible in some two-photon imaging experiments. Fluorescence changes ($\Delta F$) were normalized to a window of baseline fluorescence (F) between 500ms and 100ms preceding trial initiation.

To extract events corresponding to CpSs, the first derivative of the raw fluorescence trace was thresholded at 1.7 standard deviations from baseline. All events separated by at least one frame were included. The pre-reward window was defined as the 600 ms between the cue and the reward delivery; the post-reward window was defined as the 600 ms following reward (or the time of expected reward). Peak spike rates for individual PC dendrites were defined by first identifying the peak spike rate in the mean peri-stimulus histogram (PSTH) across all PC dendrites within either the Pre- or Post-reward window. Next, this peak was used to define a 100 ms analysis window centered on the peak, to calculate spike rate. To identify PC dendrites responsive at the time of this peak, we selected only those PC dendrites with a peak response at least one standard deviation above their pre-cue baseline activity. To determine whether responses were significant across the population, we compared the responses of all PC dendrites at the time of the peak in the PSTH to a peak found in a 1200 ms window during the baseline and performed a one-tailed t-test across PC dendrites.

To identify PC dendrites that were responsive to task events (i.e. visual cue or reward), we selected PC dendrites with spike rates that were elevated 1 standard deviation above the pre-cue period for two consecutive frames in the pre- or post-reward windows. To identify PC dendrites that were suppressed following expected reward, we selected PC dendrites with spike rates that were suppressed 1 standard deviation below both their pre-cue and pre-reward firing rates for two consecutive frames during the post-reward window. These criteria selected a very small number of dendrites, which upon inspection, revealed no clear suppression associated with the time of expected reward (**Fig. 3 – Figure Supplement 2**).

Single licks during the ITI (separated by at least 250 ms from other licks) were used to generate lick triggered averages; only experiments with at least 4 licks were included for this analysis. Peaks in the lick triggered average were identified within a window spanning 5 frames centered on the lick. Significance was calculated by comparing the amplitude of this peak with two frames 250 ms before the lick and two frames 250 ms after the lick. To segregate trials according to lick

times, on each trial we identified the first lick bout (at least 3 licks within 300 ms) following the cue, and separated trials into quartiles according to lick latency.

To track single PC dendrites across imaging sessions, we first independently generated pixel masks for PC dendrites from each field of view. We then motion registered the mean fluorescence image from the post-learning session to the pre-learning session, and applied the same pixel shifts from this registration to the segmented pixel masks from the post-learning condition. By overlaying the shifted post-learning pixel masks with the pre-learning masks, we assessed the overlap between imaging days. We employed a stringent criteria of greater than 50% overlap to increase the probability of accurately identifying the same PC dendrites across imaging sessions. With these criteria, we identified a group of PC dendrites likely corresponding to the same neurons before and after learning.

*Additional statistics*: Data are presented as mean ± S.E.M., unless stated otherwise. Statistical tests were two-sided, except as specifically noted. Differences were considered statistically significant when $P < 0.05$. No correction for multiple comparisons was applied. No statistical methods were used to predetermine sample sizes. Data distribution was assumed to be normal but this was not formally tested. Data collection and analysis were not performed blind to the conditions of the experiments, but data collection relied on automatized measurements and subsequent analysis was based on code uniformly applied across experimental conditions.

**Acknowledgements**

This work was supported by grants from the NIH NINDS (5R01NS096289-02) (CH) and (F31NS103425) (WH), the Sloan Foundation (CH), and the Whitehall Foundation (CH). We thank L. Glickfeld for helpful discussions and input on calcium imaging approaches and analyses, and members of the Hull and Glickfeld labs for input and technical assistance throughout the project.

**Author Contributions**

WH and CH designed experiments, WH conducted experiments, WH and CH analyzed data, and WH and CH wrote the manuscript.

**Competing Interests**

The authors declare no competing interests.

564

565 **Figure Legends**

566

567 **Figure 1**. Appetitive classical conditioning regime for head-fixed mice. **a)** Diagram of headplate
568 and 3mm cranial window. **b)** Scale representation of the 3mm cranial windows over either Lobule
569 Simplex (LS) or Crus I and II. **c)** Mice viewed a static full screen vertical grating that was transiently
570 replaced by a horizontal grating for 100ms to cue reward delivery following a 600ms delay. **d)**
571 Trial structure and progression of learning sessions. **e)** Left, mean reaction time, defined as the
572 first lick following the visual cue, on rewarded trials across training days. Right, mean reaction
573 time on rewarded trials before learning and after learning. **f)** Left, mean miss rate, defined as the
574 fraction of trials with no licks within 1 s after the cue, on rewarded trials across training days.
575 Right, mean miss rate on rewarded trials before and after learning **g)** Mean lick rates aligned to
576 visual cue onset, before (top) and after (bottom) learning for rewarded trials (RW, left) and
577 omission trials (right, dotted line indicates time when reward would have been delivered).

578

579 **Figure 2**. Single photon, mesoscale imaging of CpS activity during behavior. **a)** Average $\Delta F/F$
580 timecourse (black, collected at 10Hz) and lick rate (blue bars) for an example animal before (top)
581 and after (bottom) learning for rewarded (left) and reward omission (right) trials. **b,c,d)** Summary
582 of bulk CpS response latencies relative to the visual cue (b,c) or the reward (d) for all trial types.
583 Response latency is defined by the maximum positive change in $\Delta F/F$ rate within 500ms before
584 the peak response.

585

586 **Figure 3**. CpS modulation in LS across learning. **a)** Top, mean cue-aligned peri-stimulus
587 histograms (PSTHs) of CpS rate on reward (black) and omission (red) trials for all PC dendrites,
588 measured via two-photon calcium imaging for the first day of training in naive mice. Bottom,
589 summary scatterplot comparing the CpS rate for individual PCs on rewarded vs omission trials in
590 naïve mice. Spike rates were measured in a window preceding reward delivery (left) or
591 immediately after reward delivery (right) (methods). **b)** Same as A, but for rewarded and reward
592 omission trials on post-learning day N +1 **c)** Same as a and b, but for rewarded and unexpected
593 reward trials on post-learning day N+2. **d)** Mean cue-aligned PSTH for PC dendrites that exhibited
594 CpS responses to the visual cue. **e)** Mean lick-triggered PSTH for licks during the inter-trial interval
595 (ITI) in naïve animals. **f)** Mean cue-aligned PSTHs for trained animals with rewarded trials
596 segregated according trials with the earlies 1/4 of licks (black) and latest 1/4 of licks (blue). Data

points with horizontal error bars represent the mean lick timing ± SEM. For all PSTHs, shaded area represents ± SEM across dendrites.

**Figure 4.** CpS responses measured in the same fields of view across learning. **a)** Example of manual identification from a field of view containing several of the same PC dendrites on day 1 and day N+1. Left, fields of view for day 1 (top) and day N+1 (bottom). Right, magnified view from the white box at left. Red arrows indicates an example PC dendrite identified both pre and post learning. **b)** Cue aligned calcium transients averaged across trials extracted from the PC dendrite identified in **a** for day 1 (top) and day N+1 (bottom). **c)** Mean images taken from day 1 (left) and day N+1 (right) after independent rigid motion registration on each dataset. **d)** Pixel masks of PC dendrites independently extracted from the datasets in **c** (methods). **d)** The post-learning data set from c, right, was motion registered to the pre-learning dataset from c, left. The resulting x-y pixel shifts were applied to the post-learning dendrite pixel masks from d, right. **f)** Following registration, the overlap between pre and post learning pixel masks was compared quantitatively (f) and graphically (g), with pre-learning masks labeled red, post-learning masks labeled green, and overlap labeled yellow. **h)** Dendritic masks from **g** which had >50% overlap. **i)** Summary of CpS firing rates for individual PC dendrites in the pre (left) and post (right) reward window for all dendrites which had a >50% overlap in their dendritic masks across learning (n=61).

**Figure 5**. CpS modulation in Crus I across learning. **a)** Top, mean cue-aligned PSTHs of CpS rate on reward (black) and omission (red) trials for all PC dendrites, measured via two-photon calcium imaging for the first day of training in naive mice. Bottom, summary scatterplot comparing the CpS rate for individual PCs on rewarded vs omission trials in naïve mice. Spike rates were measured in a window preceding reward delivery (left) or immediately after reward delivery (right) (methods). **b)** Same as A, but for rewarded and reward omission trials on post-learning day N +1 **c)** Same as a and b, but for rewarded and unexpected reward trials on post-learning day N+2. **d)** Mean cue-aligned PSTH for PC dendrites that exhibited CpS responses to the visual cue. **e)** Mean lick-triggered PSTH for licks during the inter-trial interval (ITI) in naïve animals. **f)** Mean cue-aligned PSTHs for trained animals with rewarded trials segregated according trials with the earlies 1/4 of licks (black) and latest 1/4 of licks (blue). Data points with horizontal error bars represent the mean lick timing ± SEM. For all PSTHs, shaded area represents ± SEM across dendrites.

**Figure 6**. CpS modulation in Crus II across learning. **a)** Top, mean cue-aligned PSTHs of CpS rate on reward (black) and omission (red) trials, measured via two-photon calcium imaging for the

first day of training in naive mice. Bottom, summary scatterplot comparing the CpS rate for individual PCs on rewarded vs omission trials in naïve mice. Spike rates were measured in a window preceding reward delivery (left) or immediately after reward delivery (right) (methods). **b)** Same as A, but for rewarded and reward omission trials on post-learning day N +1 **c)** Same as a and b, but for rewarded and unexpected reward trials on post-learning day N+2. **d)** Mean cue-aligned PSTH for PC dendrites that exhibited CpS responses to the visual cue. **e)** Mean lick-triggered PSTH for licks during the inter-trial interval (ITI) in naïve animals. **f)** Mean cue-aligned PSTHs for trained animals with rewarded trials segregated according trials with the earlies 1/4 of licks (black) and latest 1/4 of licks (blue). Data points with horizontal error bars represent the mean lick timing ± SEM. For all PSTHs, shaded area represents ± SEM across dendrites.

**Fig. 3 – Figure Supplement 1.** A minority of PC dendrites can respond to reward and the reward-predictive cue after learning. **a)** Mean cue-aligned PSTH for PC dendrites that exhibited CpS responses to reward. **b)** Mean cue-aligned PSTH for PC dendrites that did not exhibit CpS responses to reward.

**Fig. 3 – Figure Supplement 2.** No evidence of CpS suppression following reward omission. **a-c)** CpS responses identified as suppressed below baseline according to statistical criteria (methods) in LS, Crus I and Crus II show no suppression upon visual inspection. These small numbers of PC dendrites were likely identified based on noise in the PSTHs that crossed the threshold for significance.

## References


Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. Trends Neurosci 31, 265-272.

Albus, J.S. (1971). A theory of cerebellar function. Mathematical Biosciences 10, 25-61.

Apps, R., and Garwicz, M. (2005). Anatomical and physiological foundations of cerebellar information processing. Nat Rev Neurosci 6, 297-311.

Bryant, J.L., Boughter, J.D., Gong, S., LeDoux, M.S., and Heck, D.H. (2010). Cerebellar cortical output encodes temporal aspects of rhythmic licking movements and is necessary for normal licking frequency. Eur J Neurosci 32, 41-52.

Carta, I., Chen, C.H., Schott, A.L., Dorizan, S., and Khodakhah, K. (2019). Cerebellar modulation of the reward circuitry and social behavior. Science (New York, NY) 363, eaav0581.

Coddington, L.T., and Dudman, J.T. (2018). The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nature Neuroscience 21, 1563-1573.

Gaffield, M.A., Amat, S.B., Bito, H., and Christie, J.M. (2016). Chronic imaging of movement-related Purkinje cell calcium activity in awake behaving mice. J Neurophysiol 115, 413-422.

Gaffield, M.A., Rowan, M.J.M., Amat, S.B., Hirai, H., and Christie, J.M. (2018). Inhibition gates supralinear Ca(2+) signaling in Purkinje cell dendrites during practiced movements. Elife 7.

Heffley, W., Song, E.Y., Xu, Z., Taylor, B.N., Hughes, M.A., McKinney, A., Joshua, M., and Hull, C. (2018). Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. Nature Neuroscience 21, 1431-1441.

Hoshi, E., Tremblay, L., Féger, J., Carras, P.L., and Strick, P.L. (2005). The cerebellum communicates with the basal ganglia. Nature neuroscience 8, 1491-1493.

Ito, M. (1972). Neural design of the cerebellar motor control system. Brain Res 40, 81-84.

Ivry, R.B., and Spencer, R.M. (2004). The neural representation of time. Curr Opin Neurobiol 14, 225-232.

Kim, S.G., Ugurbil, K., and Strick, P.L. (1994). Activation of a cerebellar output nucleus during cognitive processing. Science 265, 949-951.

Kitazawa, S., Kimura, T., and Yin, P.-B. (1998). Cerebellar complex spikes encode both destinations and errors in arm movements. Nature 392, 494-497.

Larry, N., Yarkoni, M., Lixenberg, A., and Joshua, M. (2019). Cerebellar climbing fibers encode expected reward size. bioRxiv, 533653.

Marr, D. (1969). A theory of cerebellar cortex. The Journal of Physiology 202, 437-470.

Menegas, W., Babayan, B.M., Uchida, N., and Watabe-Uchida, M. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice.

Mothersill, O., Knee-Zaska, C., and Donohoe, G. (2016). Emotion and Theory of Mind in Schizophrenia-Investigating the Role of the Cerebellum. Cerebellum 15, 357-368.

Ohmae, S., and Medina, J.F. (2015). Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nature Neuroscience 18, 1798-1803.

Parker, N.F., Cameron, C.M., Taliaferro, J.P., Lee, J., Choi, J.Y., Davidson, T.J., Daw, N.D., and Witten, I.B. (2016). Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat Neurosci 19, 845-854.

Reis, D.J., Doba, N., and Nathan, M.A. (1973). Predatory attack, grooming, and consummatory behaviors evoked by electrical stimulation of cat cerebellar nuclei. Science 182, 845-847.

Schmahmann, J.D. (1991). An emerging concept. The cerebellar contribution to higher function. Arch Neurol 48, 1178-1187.

699 Schmahmann, J.D., and Caplan, D. (2006). Cognition, emotion and the cerebellum. Brain 129,
700 290-292.
701 Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward.
702 Science (New York, NY) 275, 1593-1599.
703 Streng, M.L., Popa, L.S., and Ebner, T.J. (2017). Climbing fibers predict movement kinematics
704 and performance errors. J Neurophysiol 118, 1888-1902.
705 Syed, E.C.J., Grima, L.L., Magill, P.J., Bogacz, R., Brown, P., and Walton, M.E. (2016). Action
706 initiation shapes mesolimbic dopamine encoding of future rewards. Nature Neuroscience 19, 34-
707 36.
708 Van Overwalle, F., Baetens, K., Marien, P., and Vandekerckhove, M. (2014). Social cognition and
709 the cerebellum: a meta-analysis of over 350 fMRI studies. Neuroimage 86, 554-572.
710 Wagner, M.J., Kim, T.H., Savall, J., Schnitzer, M.J., and Luo, L. (2017). Cerebellar granule cells
711 encode the expectation of reward. Nature 544, 96-100.
712 Wang, S.S., Kloth, A.D., and Badura, A. (2014). The cerebellum, sensitive periods, and autism.
713 Neuron 83, 518-532.
714 Welsh, J.P., Lang, E.J., Suglhara, I., and Llinas, R. (1995). Dynamic organization of motor control
715 within the olivocerebellar system. Nature 374, 453-457.

716

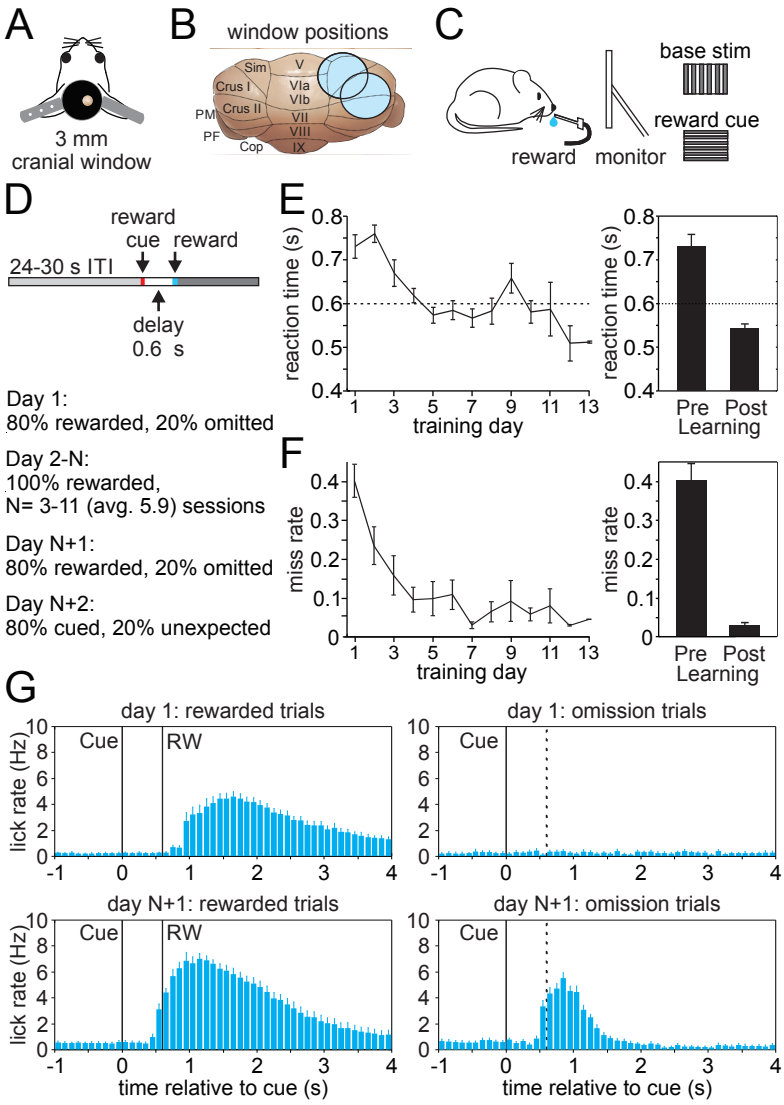
A
3 mm
cranial window
B
window positions
Sim
V
Crus I
VIa
VIb
Crus II
PM
VII
PF
VIII
Cop
IX
C
base stim
reward cue
reward
monitor
D
reward
cue
reward
24-30 s ITI
delay
0.6 s
Day 1:
80% rewarded, 20% omitted
Day 2-N:
100% rewarded,
N= 3-11 (avg. 5.9) sessions
Day N+1:
80% rewarded, 20% omitted
Day N+2:
80% cued, 20% unexpected
E
reaction time (s)
training day
Pre
Post
Learning
F
miss rate
training day
Pre
Post
Learning
G
day 1: rewarded trials
day 1: omission trials
day N+1: rewarded trials
day N+1: omission trials
Cue
RW
lick rate (Hz)
time relative to cue (s)

**Figure 1**

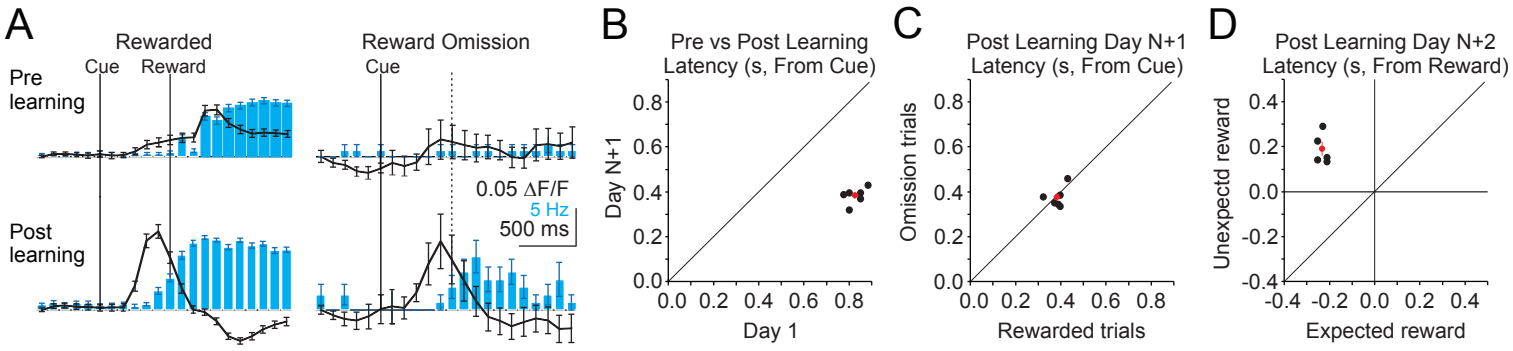
A
Rewarded
Cue
Reward
Pre learning
Post learning
Reward Omission
Cue
0.05 ΔF/F
5 Hz
500 ms
B
Pre vs Post Learning
Latency (s, From Cue)
Day N+1
Day 1
0.0
0.2
0.4
0.6
0.8
C
Post Learning Day N+1
Latency (s, From Cue)
Omission trials
Rewarded trials
D
Post Learning Day N+2
Latency (s, From Reward)
Unexpectd reward
Expected reward
-0.4
-0.2
0.0
0.2
0.4

**Figure 2**

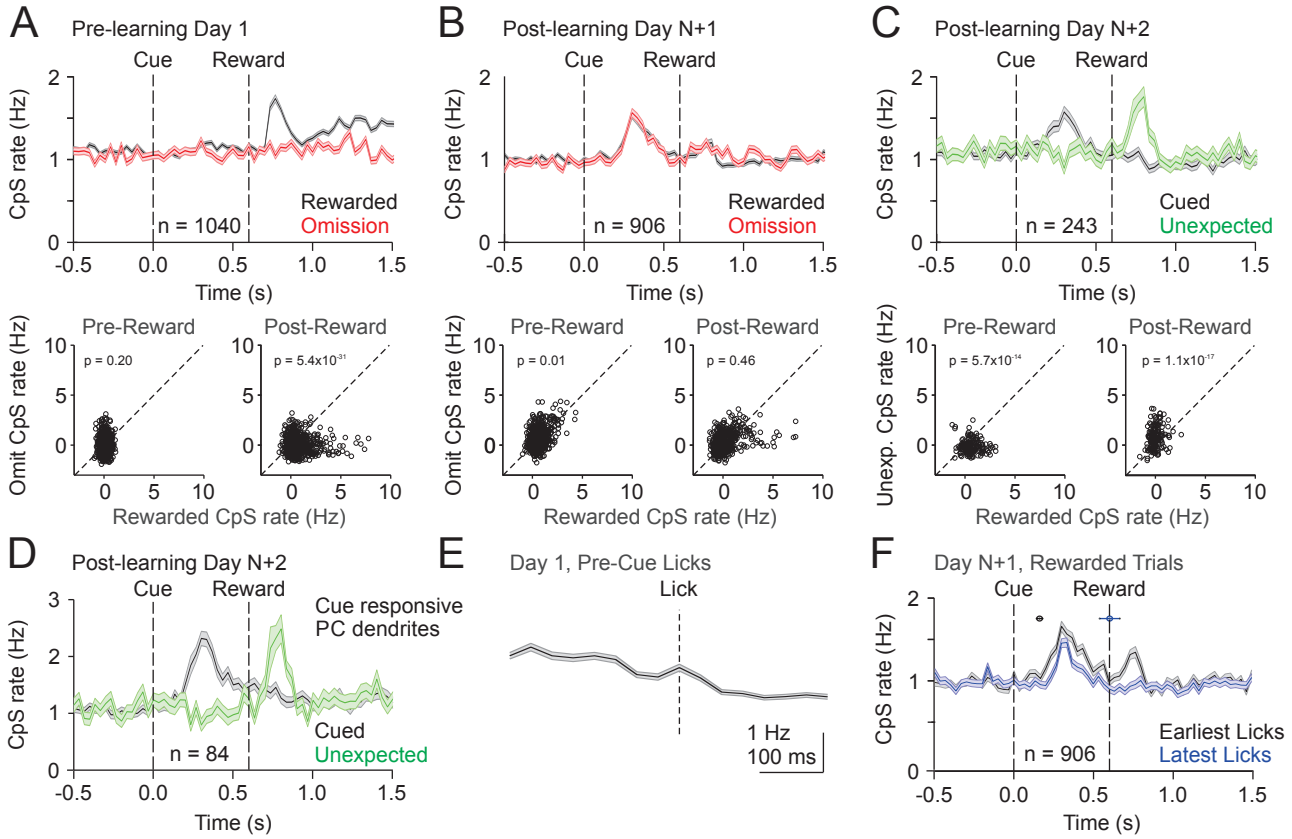
A
Pre-learning Day 1
Cue
Reward
CpS rate (Hz)
n = 1040
Rewarded
Omission
Time (s)
Pre-Reward
Post-Reward
p = 0.20
p = 5.4x10-31
Omit CpS rate (Hz)
Rewarded CpS rate (Hz)
B
Post-learning Day N+1
n = 906
p = 0.01
p = 0.46
C
Post-learning Day N+2
n = 243
Cued
Unexpected
p = 5.7x10-14
p = 1.1x10-17
Unexp. CpS rate (Hz)
D
Post-learning Day N+2
Cue responsive
PC dendrites
n = 84
E
Day 1, Pre-Cue Licks
Lick
1 Hz
100 ms
F
Day N+1, Rewarded Trials
n = 906
Earliest Licks
Latest Licks

**Figure 3**

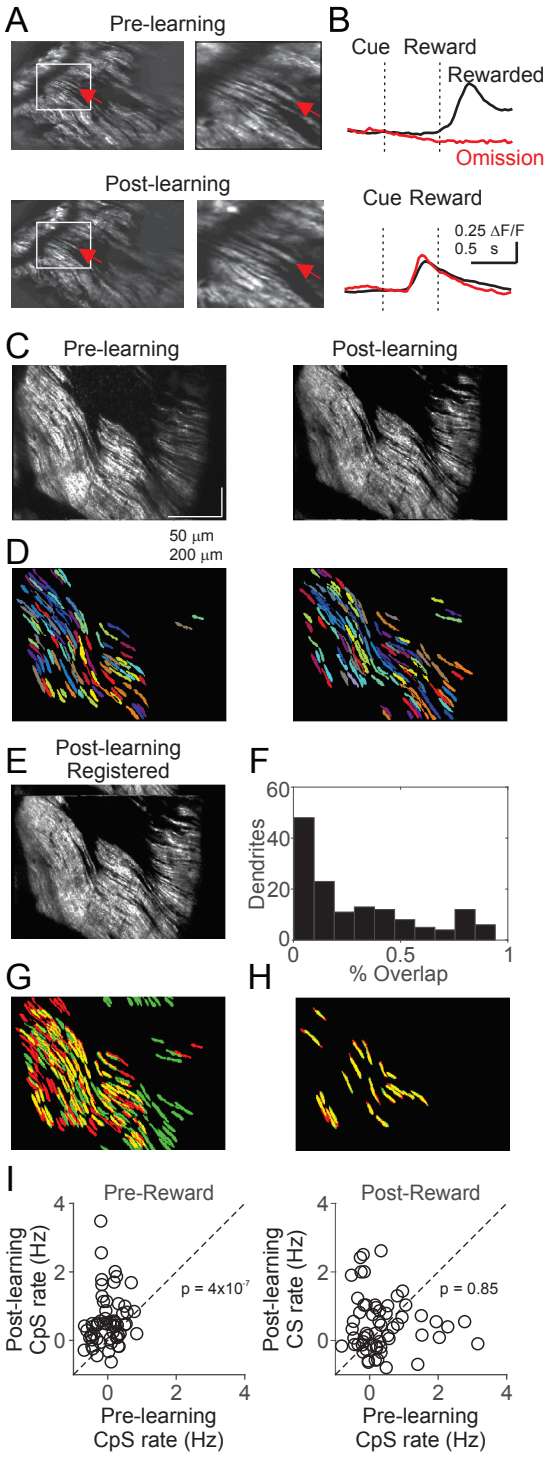
A
Pre-learning
Post-learning
B
Cue
Reward
Rewarded
Omission
Cue Reward
0.25 ΔF/F
0.5 s
C
Pre-learning
Post-learning
50 μm
200 μm
D
E
Post-learning
Registered
F
Dendrites
% Overlap
G
H
I
Pre-Reward
Post-Reward
Post-learning
CpS rate (Hz)
Post-learning
CS rate (Hz)
Pre-learning
CpS rate (Hz)
p = 4x10^-7
p = 0.85

**Figure 4**

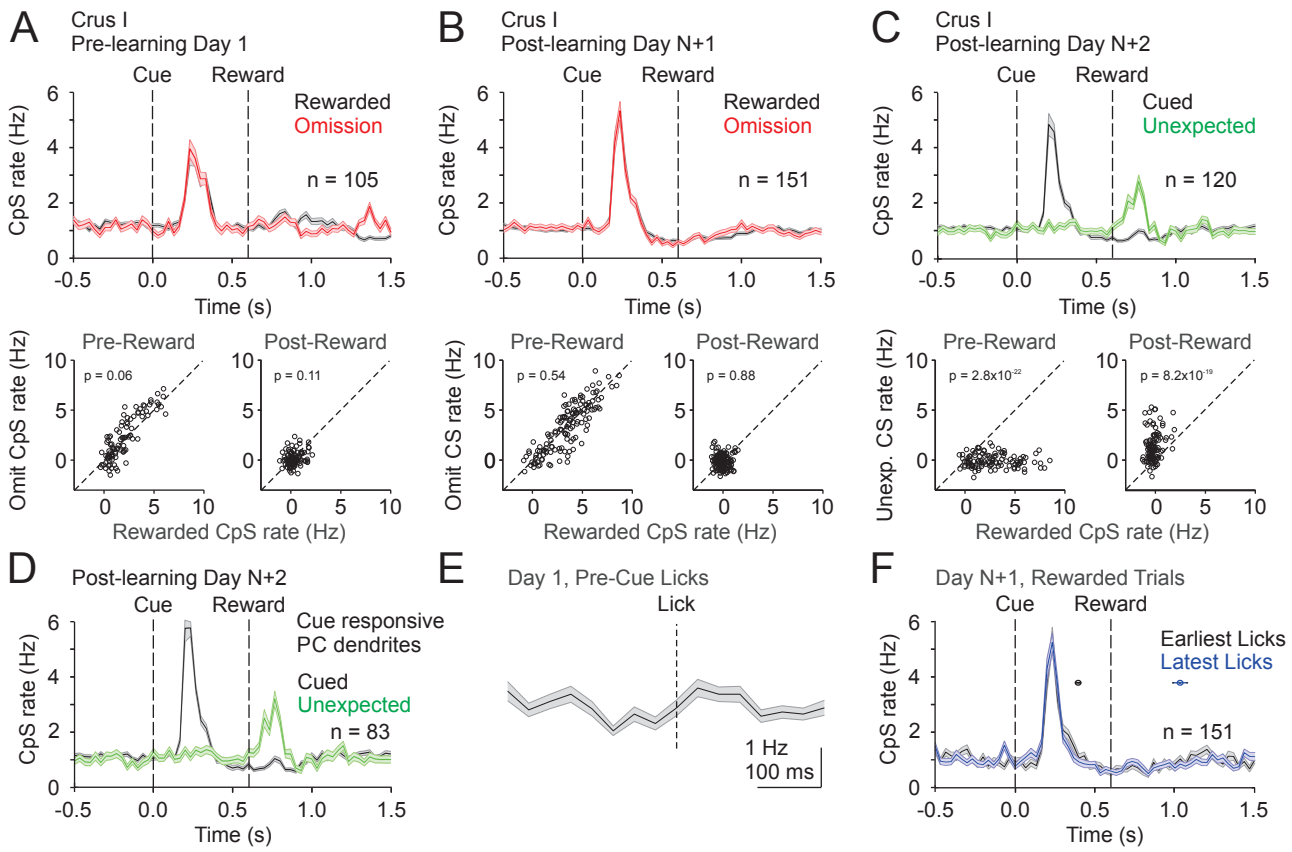
A
Crus I
Pre-learning Day 1
Cue
Reward
Rewarded
Omission
n = 105
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 0.06
p = 0.11
Omit CpS rate (Hz)
Rewarded CpS rate (Hz)
B
Crus I
Post-learning Day N+1
Cue
Reward
Rewarded
Omission
n = 151
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 0.54
p = 0.88
Omit CS rate (Hz)
Rewarded CpS rate (Hz)
C
Crus I
Post-learning Day N+2
Cue
Reward
Cued
Unexpected
n = 120
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 2.8x10-22
p = 8.2x10-19
Unexp. CS rate (Hz)
Rewarded CpS rate (Hz)
D
Post-learning Day N+2
Cue
Reward
Cue responsive
PC dendrites
Cued
Unexpected
n = 83
CpS rate (Hz)
Time (s)
E
Day 1, Pre-Cue Licks
Lick
1 Hz
100 ms
F
Day N+1, Rewarded Trials
Cue
Reward
Earliest Licks
Latest Licks
n = 151
CpS rate (Hz)
Time (s)

**Figure 5**

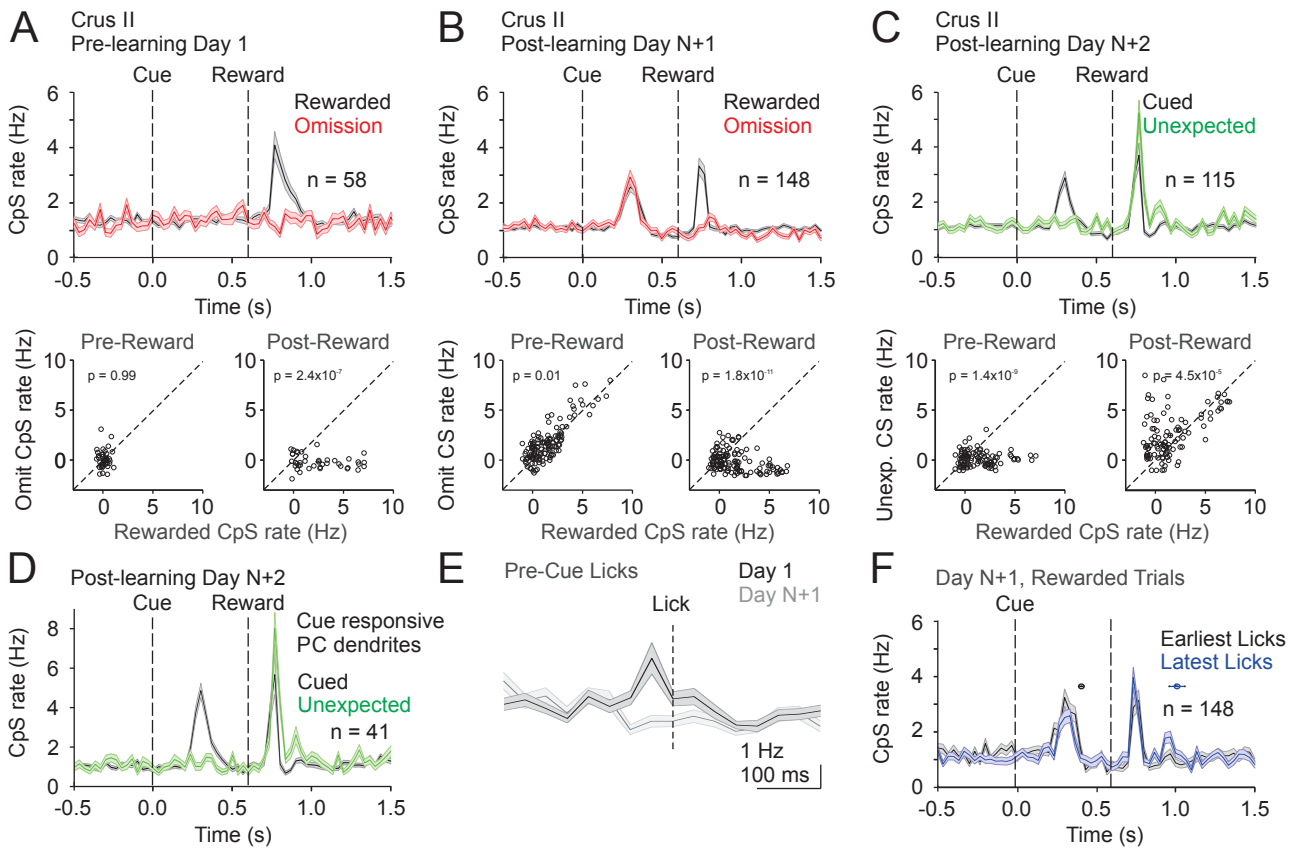
A
Crus II
Pre-learning Day 1
Cue
Reward
Rewarded
Omission
n = 58
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 0.99
p = 2.4x10-7
Omit CpS rate (Hz)
Rewarded CpS rate (Hz)
B
Crus II
Post-learning Day N+1
Cue
Reward
Rewarded
Omission
n = 148
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 0.01
p = 1.8x10-11
Omit CS rate (Hz)
Rewarded CpS rate (Hz)
C
Crus II
Post-learning Day N+2
Cue
Reward
Cued
Unexpected
n = 115
CpS rate (Hz)
Time (s)
Pre-Reward
Post-Reward
p = 1.4x10-9
pc = 4.5x10-5
Unexp. CS rate (Hz)
Rewarded CpS rate (Hz)
D
Post-learning Day N+2
Cue
Reward
Cue responsive
PC dendrites
Cued
Unexpected
n = 41
CpS rate (Hz)
Time (s)
E
Pre-Cue Licks
Day 1
Day N+1
Lick
1 Hz
100 ms
F
Day N+1, Rewarded Trials
Cue
Earliest Licks
Latest Licks
n = 148
CpS rate (Hz)
Time (s)

**Figure 6**

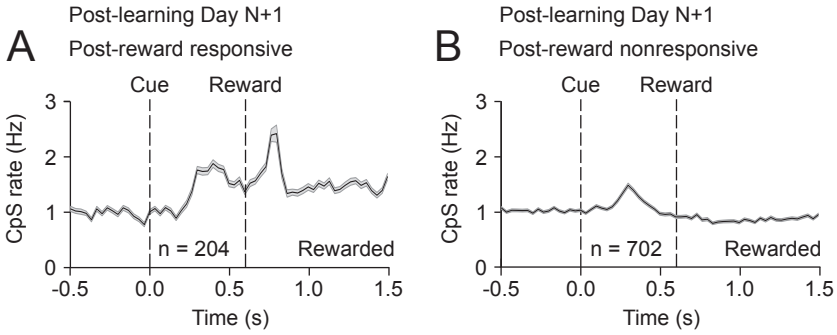


Fig. 3 - Figure Supplement 1

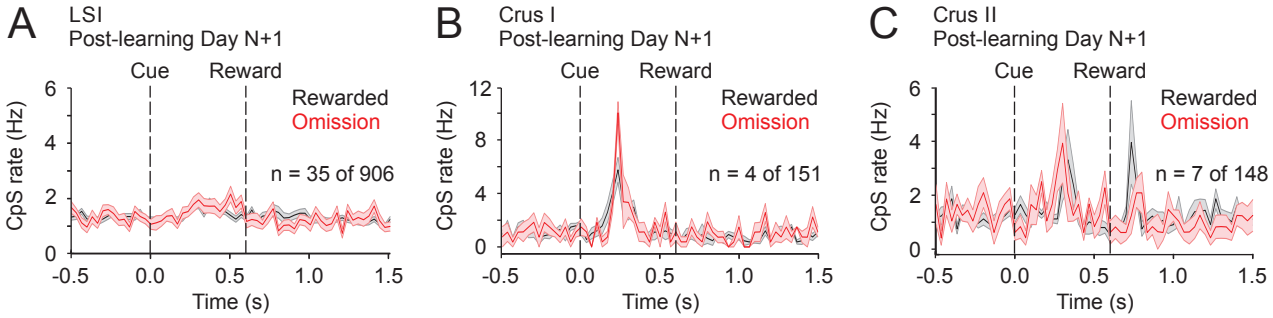


Fig. 3 - Figure Supplement 2