

# 1 **Intrinsic rewards modulate sensorimotor adaptation**

2 Hyosub E. Kim<sup>1,2\*</sup>, Darius E. Parvin<sup>1,2</sup>, Richard B. Ivry<sup>1,2</sup>

3

4 <sup>1</sup>Department of Psychology and <sup>2</sup>Helen Wills Neuroscience Institute, University of California, Berkeley

5

6

## 7 **ABSTRACT**

8 Recent studies have demonstrated that task success signals can modulate behavioral changes during  
9 sensorimotor adaptation tasks, primarily through the engagement of explicit processes. In a series of  
10 reaching experiments with human participants, we explore a potential interaction between reward-  
11 based learning and implicit adaptation, using a method in which feedback is not contingent on task  
12 performance. We varied the size of the target to compare conditions in which visual feedback indicated  
13 an invariant angular error that either hit or missed the target. Hitting the target attenuated the behavioral  
14 changes from adaptation, an effect we attribute to the generation of an intrinsic reward signal. We  
15 evaluated two models, one in which reward and adaptation systems operate in parallel, and a second in  
16 which reward acts directly on the adaptation system. The results favor the latter, consistent with  
17 evidence showing communication, and possible overlap, between neural substrates underlying reward-  
18 based learning and sensorimotor adaptation.

19

20

## 21 **INTRODUCTION**

22 Multiple learning systems contribute to successful goal-directed actions in the face of changing  
23 physiological states, body structures, and environments (Huberdeau, Krakauer, & Haith, 2015;  
24 McDougle, Ivry, & Taylor, 2016; Jordan A. Taylor & Ivry, 2014). Among these different learning  
25 processes, implicit sensorimotor adaptation is of primary importance for maintaining appropriate  
26 calibration of sensorimotor maps over both short and long timescales. A large body of work has  
27 focused on how sensory prediction error (SPE), the difference between predicted and actual sensory  
28 feedback, drives sensorimotor adaptation (Shadmehr, Smith, & Krakauer, 2010). In addition to

29 sensorimotor adaptation, there is growing awareness of how reward-based learning contributes to  
30 motor control. While several recent studies have shown that rewarding successful actions alone is  
31 sufficient for the learning of perturbations (Izawa & Shadmehr, 2011; Therrien, Wolpert, & Bastian,  
32 2016, 2018), little is known about how rewards impact implicit adaptation. Thus, a central question  
33 remains as to how learning systems tuned to SPE versus those tuned to rewards interact during motor  
34 tasks.

35

36 Despite utilizing very similar task paradigms, initial studies have led to an inconsistent picture of how  
37 reward impacts performance in sensorimotor adaptation tasks. For example, in two separate  
38 visuomotor rotation studies using similar task paradigms and reward structures, the first study reported  
39 no effect of reward on adaptation rates but an enhancement of motor memory due to rewards (Galea,  
40 Mallia, Rothwell, & Diedrichsen, 2015), while the second reported a beneficial effect of rewards  
41 specifically on adaptation rate (Nikooyan & Ahmed, 2015). In a more recent study, however,  
42 manipulation of reward attenuated overall learning (Leow, Marinovic, & Carroll, 2018).

43

44 One factor that may contribute to these inconsistencies is highlighted by recent work showing that,  
45 even in relatively simple sensorimotor adaptation tasks, overall behavior reflects a combination of  
46 explicit and implicit processes (Jordan A. Taylor & Ivry, 2011; Jordan A. Taylor, Krakauer, & Ivry, 2014).  
47 Unless the explicit component is directly assayed (Jordan A. Taylor et al., 2014), measures of  
48 adaptation can be confounded by explicit aiming. That is, while the SPE is thought to drive adaptation  
49 (Tseng, Diedrichsen, Krakauer, Shadmehr, & Bastian, 2007), participants are often also consciously  
50 aware of the perturbation and decide to aim in order to compensate for it, thereby improving task  
51 performance. It may be that reward promotes the activation of explicit processes, which can be more  
52 flexibly implemented depending on the task demands (Bond & Taylor, 2015). A recent study provides  
53 evidence for this hypothesis (Codol, Holland, & Galea, 2017), showing that at least one of the putative  
54 effects of reward, the strengthening of motor memories (Shmuelof et al., 2012), is primarily the result of

55 the re-instantiation of explicit aiming strategies as opposed to a direct modulation of adaptation. As  
56 explicit learning is much more flexibly implemented and sensitive to task demands than implicit  
57 adaptation (Bond & Taylor, 2015), differential demands on strategies are likely to contribute toward the  
58 inconsistent effects reported across studies manipulating reward (Holland, Codol, & Galea, 2018).

59

60 A recently developed method, referred to as clamped visual feedback, isolates implicit adaptation from  
61 an invariant visual error signal (Morehead, Taylor, Parvin, & Ivry, 2017). During the clamp, the angular  
62 trajectory of a feedback cursor is invariant with respect to the target location and thus spatially  
63 independent of hand position (Kim, Morehead, Parvin, Moazzezi, & Ivry, 2018; Morehead et al., 2017;  
64 Shmuelof et al., 2012; Vandevoorde & Orban de Xivry, 2018; Vaswani et al., 2015). Participants'  
65 knowledge of the visual perturbation and instructions to ignore it are intended to prevent any explicit  
66 aiming, thus allowing a clean probe of implicit adaptation (Morehead et al., 2017).

67

68 Here, we employ the clamp method to better understand how rewards may affect implicit adaptation  
69 from SPE, without interference from explicit aiming strategies. In a series of three experiments, the  
70 clamp offset was held constant and only the size of the target was manipulated, affecting whether the  
71 cursor would hit or miss the target. Thus, we were able to experimentally manipulate both the putative  
72 SPE (angular offset of clamp) and the reward (hitting versus not hitting the target). We assume that  
73 hitting the target would be intrinsically rewarding (Leow et al., 2018; Xu-Wilson, Zee, & Shadmehr,  
74 2009), even though participants are explicitly aware that hitting the target is independent of their actual  
75 performance. Given this assumption, we ask how reward impacts adaptation from a constant SPE.

76

77 The results of the first two experiments revealed a strong attenuation of adaptation when the cursor hit  
78 the target. Based on these results, in Experiment 3 we evaluated two hypotheses regarding the  
79 mechanism by which intrinsic rewards affect adaptation. We hypothesized that either intrinsic reward  
80 activates reward-based reinforcement in parallel to SPE-driven adaptation, with movement being the

81 net result of these two independent processes (Movement Reinforcement model), or intrinsic reward  
82 directly modulates adaptation (Adaptation Modulation model). Our results provide support for the latter,  
83 although our model-based analyses suggest there may be a mixture of both mechanisms.

84

## 85 **RESULTS**

86 In all experiments we used clamped visual feedback, in which the angular trajectory of a feedback  
87 cursor is invariant with respect to the target location and thus spatially independent of hand position  
88 (Morehead et al., 2017; Fig. 1a). The instructions emphasized that the participant's behavior would not  
89 influence the cursor trajectory: They were to ignore this stimulus and always aim directly for the target.  
90 This method allows us to isolate implicit learning from an invariant SPE, eliminating potential  
91 contributions from strategic changes that might be used to reduce task performance error.

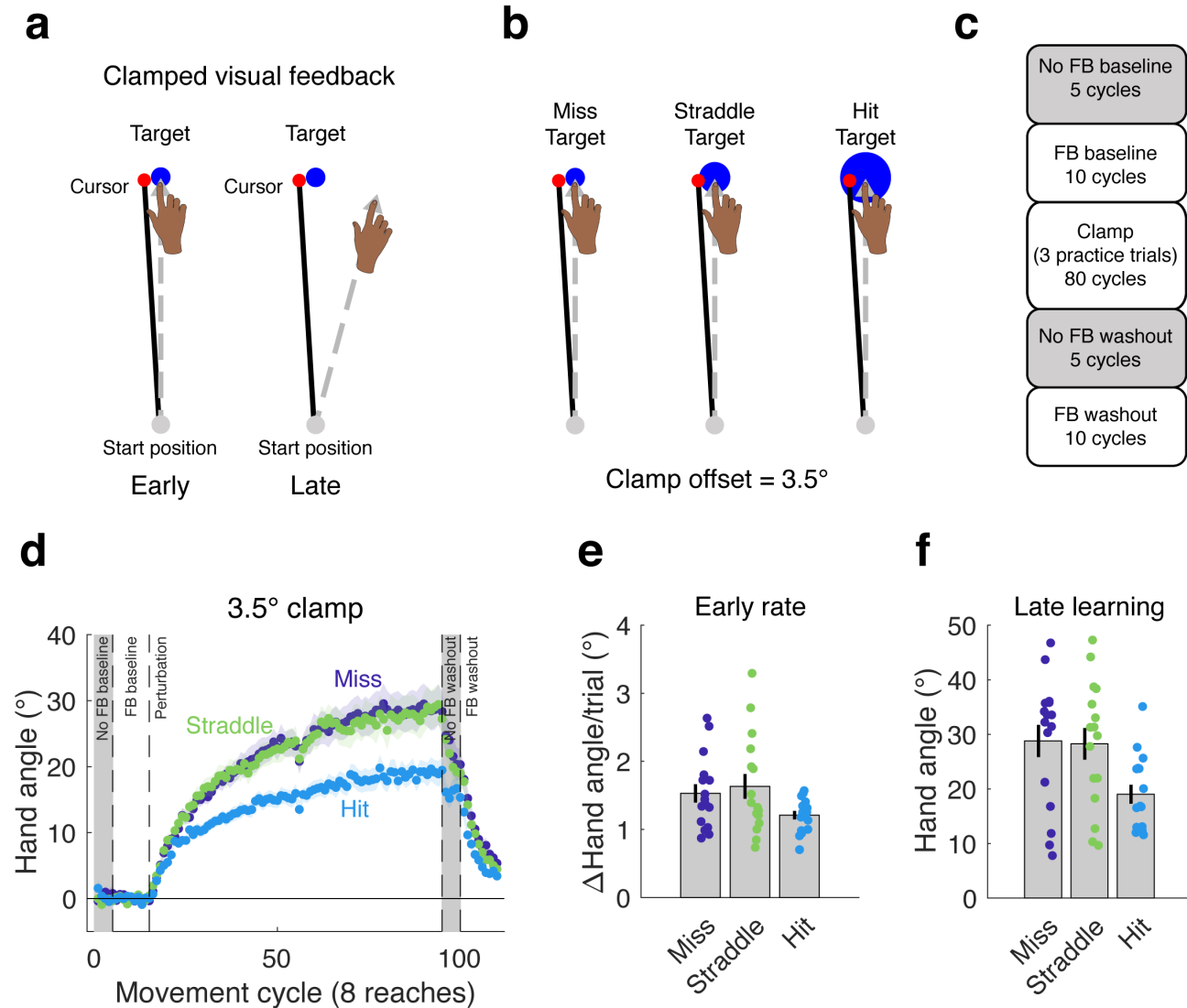
92

93 In Experiment 1, we asked whether hitting the target under conditions in which the feedback is not  
94 contingent on behavior would modulate adaptation, based on the assumption that this would be  
95 intrinsically rewarding. We tested three groups of participants ( $n=16/\text{group}$ ) with a  $3.5^\circ$  clamp offset for  
96 80 cycles (8 targets per cycle). The purpose of this experiment was to examine the effects of three  
97 different relationships between the clamp and target while holding the visual error (defined as the  
98 center-to-center distance between the cursor and target) constant (Fig. 1b): Hit Target (when the  
99 terminal position of the clamped cursor is fully embedded within a 16 mm diameter target), Straddle  
100 Target (when roughly half of the cursor falls within a 9.8 mm target, with the remaining part outside the  
101 target), Miss Target (when the cursor is fully outside a 6 mm target). As seen in Fig. 1d, hitting the  
102 target reduced the overall change in behavior. Statistically, there was a marginal difference on the rate  
103 of initial adaptation (one-way ANOVA:  $F(2,45)=2.67$ ,  $p=.08$ ,  $\eta^2=.11$ ; Fig. 1e) and a significant effect on  
104 late learning ( $F(2,45)=4.44$ ,  $p=.016$ ,  $\eta^2=.17$ ; Fig. 1f). For the latter measure, the value for the Hit Target  
105 group was approximately 35% lower than for the Straddle and Miss Target groups, with post-hoc

106 comparisons confirming the substantial differences in late learning between the Hit Target and both the  
107 Straddle Target (95% CI [-16.13°, -2.34°],  $t(30)=-2.73$ ,  $p=.010$ ,  $d=.97$ ) and Miss Target (95% CI [-  
108 16.76°, -2.79°],  $t(30)=-2.86$ ,  $p=.008$ ,  $d=1.01$ ) groups. The learning functions for the Straddle and Miss  
109 Target groups were remarkably similar throughout the entire clamp block and reached similar  
110 magnitudes of late learning (95% CI [-7.90°, 8.97°],  $t(30)=.13$ ,  $p=.898$ ,  $d=.05$ ).

111

112 Interestingly, these results appear qualitatively different to those observed when manipulating the  
113 clamp offset. Our previous study using clamped visual feedback demonstrated that varying clamp offset  
114 alone results in different early learning rates, but produces the same magnitude of late learning (Kim et  
115 al., 2018). The results here in Experiment 1 however, suggest that the intrinsically rewarding feedback  
116 associated with hitting the target results in small differences in early learning that are amplified in late  
117 learning. Furthermore, the effect of intrinsic reward appears to be categorical, as it was only observed  
118 for the condition in which the cursor was fully embedded within the target (Hit Target), and not when the  
119 terminal position of the cursor fell partially outside the target (Straddle Target).



120

121

122

123 **Figure 1** Hitting the target attenuates adaptation from SPE. **(a)** During clamped visual feedback, the  
 124 angular offset of the cursor feedback is held constant throughout the perturbation block. Despite  
 125 participants' knowledge of the clamp, this manipulation elicits robust changes in hand angle. **(b)** The  
 126 clamp offset was equal across all three conditions tested in Experiment 1, with only the target size  
 127 varying between conditions. **(c)** Block design for experiment. **(d)** There was attenuation of adaptation in  
 128 the Hit Target condition, observed in the **(e)** early adaptation rate (the per cycle rate of change in hand  
 129 angle during clamp cycles 3-7), and more dramatically in **(f)** late learning (mean hand angle over last 10  
 130 clamp cycles). Dots represent individuals; shading and error bars denote SEM.

131

132 *Experiment 2*

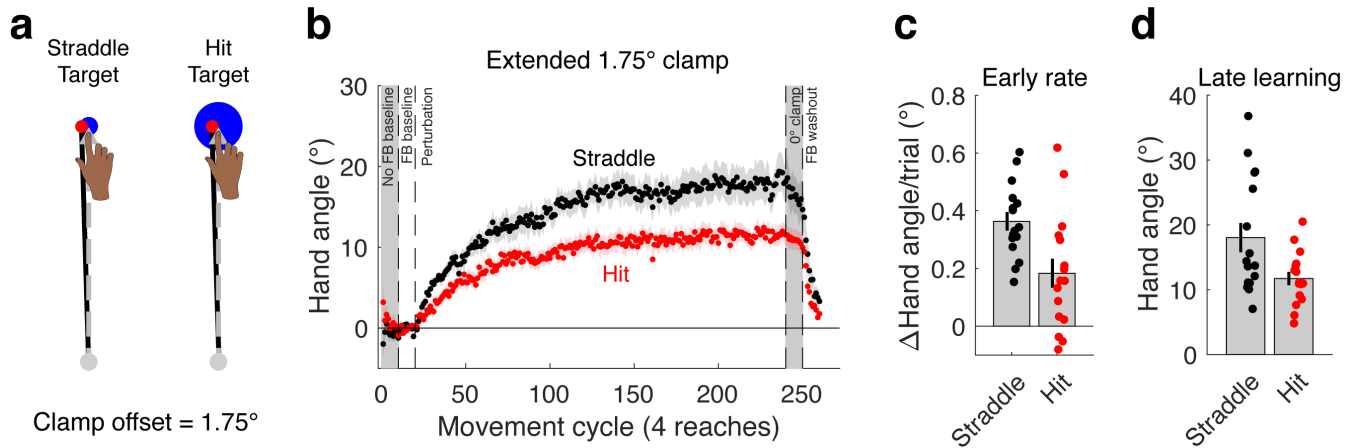
133 Experiment 2 was designed to extend the results of Experiment 1 in two ways: First, to verify that the  
 134 effect of hitting a target generalized to other contexts, we changed the size of the clamp offset. We

135 tested two groups of participants (n=16/group) with a 1.75° clamp offset. For the Hit Target group (Fig.  
136 2a), we used the large 16 mm target, and thus, the cursor was again fully embedded. For the Straddle  
137 Target group, we used the small 6 mm diameter target, resulting in an endpoint configuration in which  
138 the cursor was approximately half within the target and half outside the target. We did not test a Miss  
139 Target condition because having the clamped cursor land fully outside the target would have  
140 necessitated an impractically small target (~1.4 mm). Moreover, the results of Experiment 1 indicate  
141 that this condition is functionally equivalent to the Straddle Target group. The second methodological  
142 change was made to better assess asymptotic adaptation. We increased the number of clamped  
143 reaches to each location to 220 (reducing the number of target locations to four to keep the experiment  
144 within a 1.5 hour session). This resulted in a nearly three-fold increase in the number of clamped  
145 reaches per location.

146

147 Consistent with the results of Experiment 1, the Hit Target group showed an attenuated learning  
148 function compared to the Straddle Target group (Fig. 2b). Statistically, there was again only a marginal  
149 difference in the rate of early adaptation (95% CI [-.52°/cycle, .01°/cycle],  $t(30)=-1.96$ ,  $p=.06$ ,  $d=.69$ ; Fig.  
150 2c), whereas the difference in late learning was quite pronounced (95% CI [-11.38°, -1.25°],  $t(30)=-$   
151 2.54,  $p=.016$ ,  $d=.90$ ; Fig. 2d). Indeed, the 35% attenuation in asymptote for the Hit Target group  
152 compared to the Straddle Target group is approximately equal to that observed in Experiment 1.

153



154

155 **Figure 2** The attenuation of adaptation caused by hitting the target (a) generalizes to a different clamp  
156 offset and is stable over an extended clamp block (b). There were marginal differences in (c) early  
157 adaptation rates and (d) a significant difference in the magnitude of late learning. Dots represent  
158 individuals; shading and error bars denote SEM.

159

160

161 The results of these first two experiments converge in showing that adaptation from SPE is attenuated  
162 when the cursor hits the target, relative to conditions in which at least part of the cursor falls outside the  
163 target. This effect replicated across two experiments that used different clamp offsets.

164

165 *Attenuated behavioral changes are not due to differences in motor planning*

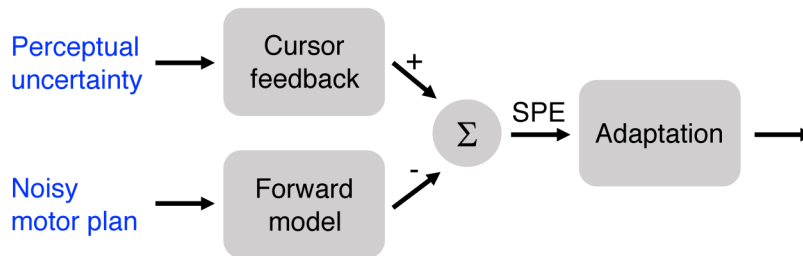
166 Although we interpret the attenuation of behavioral change as the effect of an intrinsic reward signal,  
167 generated in the Hit Target conditions, there are some alternative explanations for the effect of target  
168 size on adaptation. We aim to address these alternatives by analyzing the kinematic data in  
169 Experiments 1 and 2.

170

171 One alternative is that participants in the Hit Target groups had reduced accuracy demands relative to  
172 the other groups, given that they were reaching to a larger target (Soechting, 1984). If the accuracy  
173 demands were reduced for these large targets, then the motor command could be more variable,  
174 resulting in more variable sensory predictions from a forward model, and thus a weaker SPE (Körding &  
175 Wolpert, 2004; see Fig. 3). While we do not have direct measures of planning noise, a reasonable



176 proxy can be obtained by examining movement variability during unperturbed baseline trials (data from  
177 clamped trials would be problematic given the induced change in behavior). If there is substantially  
178 more noise in the plan for the larger target, then the variability of hand angles should be higher in this  
179 group (Churchland, Afshar, & Shenoy, 2006). In addition, one may expect faster movement times (or  
180 peak velocities) and/or reaction times for reaches to the larger target, assuming a speed-accuracy  
181 tradeoff (Fitts' law; Fitts, 1992).



182

183 **Figure 3** Adaptation could be affected by different factors, such as perceptual uncertainty or greater  
184 variability in motor planning. In the case of perceptual uncertainty, the feedback signal is weakened,  
185 thus leading to a weaker SPE signal. In the case of noisy motor planning, the forward model prediction  
186 would also be more variable and effectively weaken the SPE.

187

188 Examination of kinematic and temporal variables did not support this noisy motor plan hypothesis.

189 During baseline trials with veridical feedback, mean spatial variability, measured in terms of hand angle,

190 was actually lower for the group reaching to the larger target (Hit Target group:  $3.09^\circ \pm .18^\circ$ ; Straddle

191 Target group:  $3.56^\circ \pm .16^\circ$ ;  $t(30)=-1.99$   $p=.056$ ,  $d=0.70$ ). Further supporting the argument that planning

192 was no different across conditions, neither reaction times (Hit Target:  $378 \pm 22$  ms; Straddle Target:

193  $373 \pm 12$  ms) nor movement times (Hit Target:  $149 \pm 8$  ms; Straddle Target:  $157 \pm 8$  ms) differed

194 between the groups ( $t(30)=-0.183$ ,  $p=.856$ ,  $d=.06$  and  $t(30)=0.71$ ,  $p=.484$ ,  $d=.25$ , respectively).

195 Qualitatively similar results for baseline behavior were observed in Experiment 1 (see Supplement).

196

197 One reason for not observing an effect of target size on accuracy or temporal measures could be due

198 to the constraints of the task. Studies which observe effects of target size on motor planning typically

199 utilize point-to-point movements (Knill, Bondada, & Chhabra, 2011; Soechting, 1984) in which accuracy

200 requires planning of both movement direction and extent. In our experiments, we utilized shooting  
201 movements, thus minimizing demands on the control of movement extent. Endpoint variability is  
202 generally larger for movement extent compared to movement direction (Gordon, Ghilardi, & Ghez,  
203 1994). It is possible that participants are near ceiling-level performance in terms of hand angle  
204 variability. Another reason for the absence of a speed accuracy trade-off in the current experiment  
205 could be that with the clamp method, participants do not receive task performance feedback throughout  
206 the experiment.

207

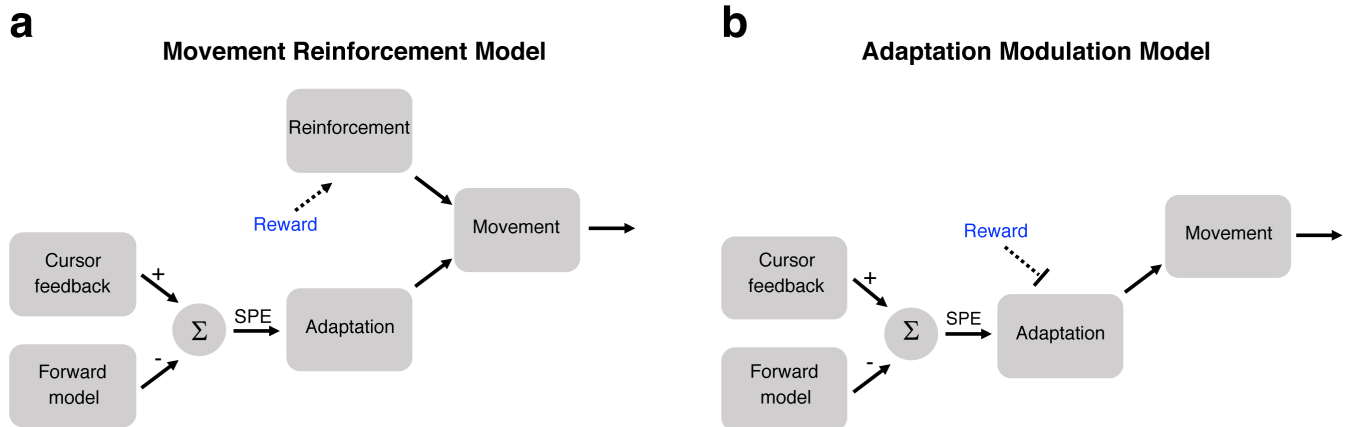
208 A second alternative to the intrinsic reward hypothesis is that participants adapted less during Hit  
209 Target conditions due to perceptual uncertainty, an idea we test in a control condition in Experiment 3.

210

### 211 *Experiment 3*

212 Based on the results of Experiments 1 and 2, we considered two ways in which an intrinsic reward,  
213 generated by hitting the target, could attenuate the rate and asymptotic level of learning. First, intrinsic  
214 reward could act as a positive reinforcement signal, strengthening the representation of rewarded  
215 movements (Shmuelof et al., 2012) (Fig. 4a, Movement Reinforcement model). This would effectively  
216 operate as a resistance to the directional change in hand angle induced by SPEs, since the reward  
217 would reinforce the executed motor command. By this model, intrinsic reward has no direct effect on  
218 the adaptation process, in that reward and error-based learning are operating in parallel, with the final  
219 movement being a composite of two different processes. Alternatively, intrinsic reward might directly  
220 modulate adaptation, attenuating the trial-to-trial change induced by the SPE (Fig. 4b, Adaptation  
221 Modulation model). For example, the reward signal might serve as a gain controller, reducing the rate  
222 at which an internal model is updated, attenuating the learning drive.

223



224

225 **Figure 4** Two models of how intrinsic reward could affect overall learning. **(a)** In the Movement  
226 Reinforcement model, reward signals cause reinforcement learning processes to bias future  
227 movements towards previously rewarded movements. The adaptation process is sensitive only to SPE  
228 and not reward. The overall movement reflects a composite of the two processes. **(b)** In the Adaptation  
229 Modulation model, reward directly attenuates adaptation to SPE.  
230  
231

232 The experimental design employed in Experiments 1 and 2 cannot distinguish between these two  
233 hypotheses because both make similar predictions when the clamp is introduced. In the Movement  
234 Reinforcement model, the attenuated asymptote arises because movements are rewarded throughout,  
235 including during early learning, biasing future movements towards baseline. The Adaptation Modulation  
236 model makes a similar prediction, but here the effect arises because the adaptation system is directly  
237 attenuated.

238

239 However, a transfer design in which the target size changes after an initial adaptation phase affords an  
240 opportunity to contrast the two models. In Experiment 3, we tested a group of participants (n=12) with a  
241 1.75° clamp, using the design depicted in Fig. 5a (Straddle-to-Hit group). In an initial acquisition phase  
242 (first 120 clamp cycles), the target was small, such that the clamp always straddled the target. Based  
243 on the results of Experiments 1 and 2, we expect to observe a relatively large change in hand angle at  
244 the end of this phase. The key test comes during the transfer phase (final 80 clamp cycles), in which  
245 the target size is increased such that the invariant clamp now results in a target hit. By the Movement

246 Reinforcement model, hitting the target will produce an intrinsic reward signal, reinforcing the  
247 associated movement. Therefore, there should be no change in performance (hand angle) following  
248 transfer: The SPE remains the same, and with the introduction of a reward signal, the executed  
249 movements would now be reinforced (Fig. 5b). In contrast, the Adaptation Modulation model assumes  
250 that the introduction of the reward signal will directly attenuate the output of the adaptation system. As  
251 such, this model predicts a marked decay in hand angles following transfer, relative to the initial  
252 asymptote.

253

254 In addition to the Straddle-to-Hit group described above, we also tested a second group (n=12) in which  
255 the large target (reward) was used in the acquisition phase and the small target (no reward) in the  
256 transfer phase (Hit-to-Straddle group). Both models make the same predictions for the Hit-to-Straddle  
257 group. At the end of the acquisition phase, there should be a relatively small change in hand angle due  
258 to the presence of an intrinsic reward signal. Following transfer, the Movement Reinforcement model  
259 predicts that, with the switch to the small target, the intrinsic reward signal will now be absent,  
260 weakening the contribution of the reward-based system to the motor output. As such, there should be  
261 an increase in hand angle following transfer. The Adaptation Modulation model predicts a similar  
262 change in behavior due to the removal of the direct inhibitory effect of the reward system on adaptation  
263 following transfer. Although this group in isolation does not discriminate between the models, it does  
264 provide a second test of each model, as well as an opportunity to rule out alternative hypotheses for the  
265 behavioral effects at transfer. For example, the absence of a change at transfer might be due to  
266 reduced sensitivity to the clamp following a long initial acquisition phase. With the Hit-to-Straddle group,  
267 both models predict a marked increase in hand angle.

268

269 For our analyses, we first examined performance during the acquisition phase (Fig. 5c). Consistent with  
270 the results from Experiments 1 and 2, the Hit-to-Straddle Target group adapted slower than the  
271 Straddle-to-Hit group (95% CI [-.17°/cycle, -.83°/cycle],  $t(22)=-3.15$ ,  $p=.005$ ,  $d=1.29$ ; Fig. 5d) and

272 reached a lower asymptote (95% CI [-5.25°, -15.29°],  $t(22)=-4.24$ ,  $p=.0003$ ,  $d=1.73$ ). The reduction at  
273 asymptote was approximately 45%.

274

275 We next examined performance during the transfer phase where the target size reversed for the two  
276 groups. Our primary measure of behavioral change for each subject was the difference in late learning  
277 (average hand angle over last 10 cycles) between the end of the acquisition phase and the end of the  
278 transfer phase. As seen in Fig. 5c, the two groups showed opposite changes in behavior in the transfer  
279 phase, evident by the strong (group x phase) interaction ( $F(2,33)=43.1$ ,  $p<10^{-7}$ , partial  $\eta^2=.72$ ). The  
280 results of a within-subjects t-test showed that the Hit-to-Straddle group showed a marked increase in  
281 hand angle following the decrease in target size (95% CI [4.9°, 9.1°],  $t(11)=7.42$ ,  $p<.0001$ ,  $d_z=2.14$ ; Fig.  
282 5e), consistent with the predictions for both the Movement Reinforcement and Adaptation Modulation  
283 models, assuming that transfer resulted in the removal of the intrinsic reward signal.

284

285 The Straddle-to-Hit group's transfer performance provides the critical test of the two hypotheses.  
286 Following the switch to the large target, there was a decrease in hand angle. Applying the same  
287 statistical test, the mean decrement in hand angle was 5.7° from the final cycles of the training phase to  
288 the final cycles of the transfer phase (95% CI [-3.1°, -8.2°],  $t(11)=-4.84$ ,  $p=.0005$ ,  $d_z=1.40$ ; Fig. 5e). This  
289 result is consistent with the prediction of the Adaptation Modulation model, namely that the introduction  
290 of an intrinsic reward signal attenuated the output of the adaptation system. The reduction in hand  
291 angle cannot be accounted for by the Movement Reinforcement model. For all participants in both  
292 groups, the directional changes in hand angle following transfer were consistent with the predictions of  
293 the Adaptation Modulation model (Fig. 5e).

294

295 To quantitatively evaluate the Adaptation Modulation model, we simulated the results of the transfer  
296 phase of Experiment 3 based on parameters estimated from the acquisition phase of both groups. We  
297 fit the data using a single rate state-space model of the following form:

$$298 \quad x_{n+1} = Ax_n + U(e_n) \quad \text{[Equation 1]}$$

299 where  $x$  represents the motor output on trial  $n$ ,  $A$  is a retention factor, and  $U$  represents the  
300 update/correction size (or, learning rate) as a function of the error size,  $e$ . This model is mathematically  
301 equivalent to a standard single rate state-space model (Thoroughman & Shadmehr, 2000), with the  
302 only modification being the replacement of the error sensitivity term,  $B$ , with a correction size function.  
303 Unlike standard adaptation studies where error size changes over the course of learning, however,  $e$  is  
304 a constant with clamped visual feedback, and therefore  $U(e)$  can be fit as a single parameter (for further  
305 details, see Kim et al. 2018). We refer to this model as the motor correction variant of the standard  
306 state space model.

307

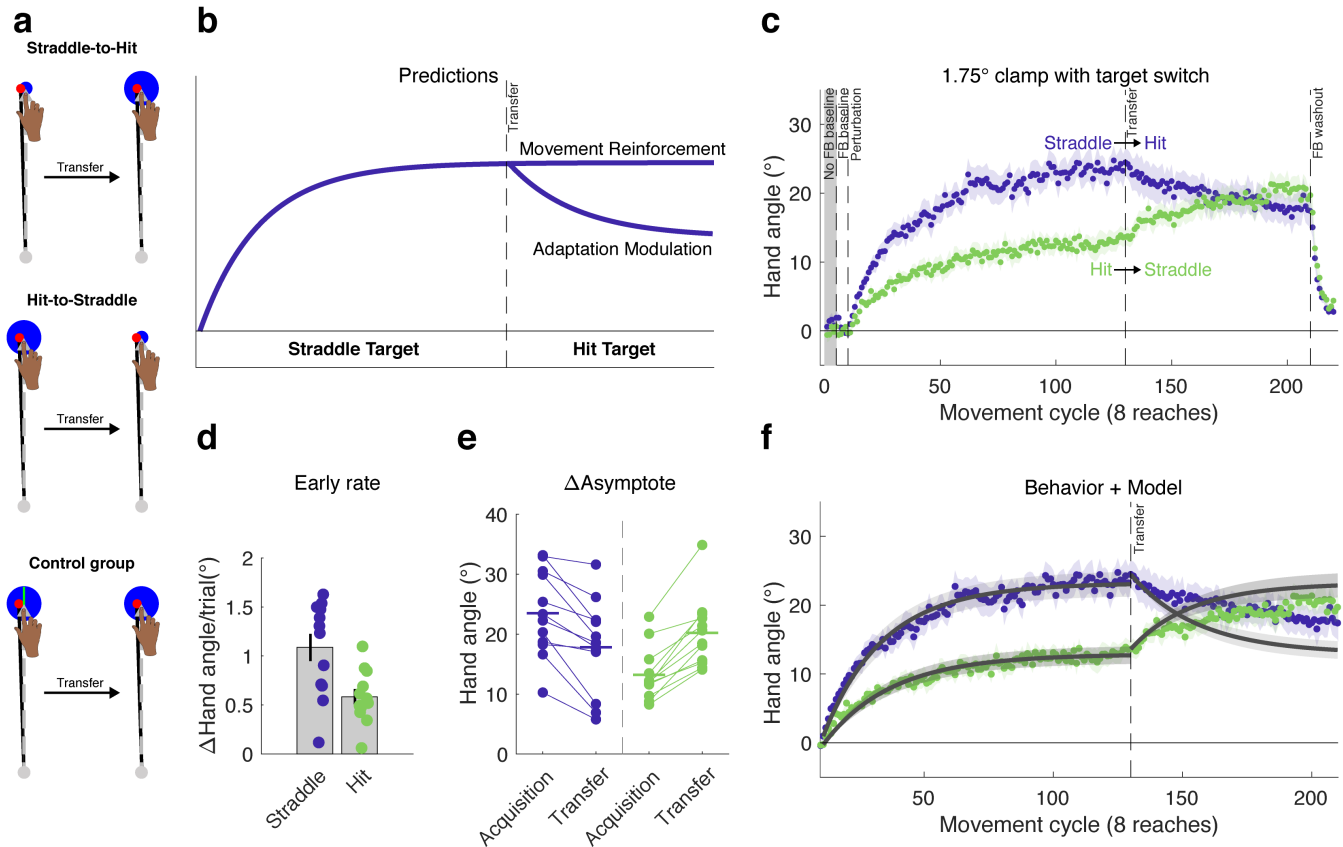
308 To estimate  $A$  and  $U(e)$ , we fit the bootstrapped samples of mean behavior, using only the data from  
309 the acquisition phase. The model provided good fits of behavioral change during the acquisition phase  
310 (Fig. 5f), with a median r-squared value of .94 (95% CI: [.86, .96]). Parameter estimates for the Hit-to-  
311 Straddle group were [.952, .973] for  $A$  and [.33°, .62°] for  $U(e)$  (values represent bootstrapped 95%  
312 CIs). For the Straddle-to-Hit group, estimated  $A$  and  $U(e)$  values during acquisition were [.939 .970] and  
313 [.69° 1.37°], respectively. Using non-parametric permutation tests on the parameter estimates for  
314 individual data, reliable differences between groups were observed for  $U(e)$  ( $p = .012$ ), but not  $A$  ( $p =$   
315  $.802$ ). Thus, the analysis of the parameter estimates indicates that reward modulates the error size-  
316 dependent motor correction within the adaptation system, effectively reducing the size of the trial-to-trial  
317 correction.

318

319 To further test whether the effect of intrinsic reward was better explained by a reduction in learning rate  
320 rather than a change in retention, we compared models in which only  $U$  or  $A$  were free to vary, asking  
321 how well these models fit the bootstrapped samples of the acquisition phase data. For the model in  
322 which  $U$  was a free parameter, we fixed  $A$  to its median value from the original fits (.96); for the model  
323 in which  $A$  was a free parameter, we fixed  $U$  to its median value from the original fits (.64). The free  $U$   
324 model explained, on average, ~8% more of the variance in the data than the free  $A$  model, and also  
325 provided excellent fits of the data (95% CIs for r-squared values: free  $U$  model [.85, .96]; free  $A$  model  
326 [.61, .94]).

327

328 The estimated parameters for each group's acquisition phase data were then used to predict the  
329 transfer performance for the other group. That is, parameter estimates from the Hit-to-Straddle group  
330 were used to predict the transfer performance of the Straddle-to-Hit group. In a complementary  
331 manner, parameter estimates from the Straddle-to-Hit group were used to predict the transfer  
332 performance of the Hit-to-Straddle group. We used all 1000 sets of parameter estimates from each  
333 group to generate the mean and variance of the predicted behavior (Fig. 5f). During transfer, the model  
334 captures the qualitative change in performance for both groups, with an increase in hand angle for the  
335 Hit-to-Straddle group and decrease in hand angle for the Straddle-to-Hit group. However, the  
336 predictions of the model slightly underestimate the observed rates of change for both groups. We return  
337 to this issue in the Discussion; for now, we note that modeling results are consistent with the hypothesis  
338 that intrinsic reward directly modulates the adaptation system.



339

340 **Figure 5** Intrinsic rewards directly modulate the output of the adaptation system. Using a transfer  
 341 design (a), the two models diverge in their behavioral predictions for the Straddle-to-Hit group following  
 342 transfer (b). The Movement Reinforcement model predicts a consistent asymptote following transfer, as  
 343 the learning drive from the SPE is the same and the reward system now reinforces the movement. In  
 344 contrast, the Adaptation Modulation model predicts an immediate decay in hand angle, given that the  
 345 intrinsic reward attenuates the output of the adaptation system after transfer. (c) The learning functions  
 346 were consistent with the predictions of the Adaptation Modulation model. The rise in hand angle for the  
 347 Hit-to-Straddle group is consistent with both models. Note that during the acquisition phase, we again  
 348 observed differences in the early adaptation rate (d) as well as late learning. (e) All participants in both  
 349 groups demonstrated changes in late learning from acquisition to transfer phases that were consistent  
 350 with Adaptation Modulation model predictions. (f) A two-parameter state-space model was able to  
 351 characterize the learning functions during the acquisition phase. However, the changes in behavior  
 352 during the transfer phase were somewhat slower than predicted, based on predictions made by  
 353 switching parameters for both groups. Dots represent individuals; shading and error bars denote SEM.

354

355

356 *Control group for testing perceptual uncertainty hypothesis*

357 Across the three experiments, the amount of adaptation induced by clamped visual feedback was  
 358 attenuated when participants reached to the large target. We considered if this effect could be due, in  
 359 part, to the differences between the Hit and Straddle/Miss conditions in terms of perceptual uncertainty.



360 For example, the reliability of the visual error signal might be weaker if the cursor is fully embedded  
361 within the target; in the extreme, failure to detect the angular offset might lead to the absence of an  
362 SPE on some percentage of the trials.

363

364 To evaluate this perceptual uncertainty hypothesis, we tested an additional group in Experiment 3 with  
365 a large target, but modified the display such that a bright line, aligned with the target direction, bisected  
366 the target (Fig. 5a). With this display, the feedback cursor remained fully embedded in the target, but  
367 was clearly off-center. If the attenuation associated with the large target is due to perceptual  
368 uncertainty, then the inclusion of the bisecting line should produce an adaptation effect similar to that  
369 observed with small targets. Alternatively, if perceptual uncertainty does not play a prominent role in the  
370 target size effect, then the adaptation effects would be similar to that observed with large targets.

371

372 Consistent with the second hypothesis, performance during the acquisition phase for the group  
373 reaching to a bisected target was similar to that of the group reaching to the standard large target (Hit-  
374 to-Straddle, see Supplement). Planned pair-wise comparisons showed no significant differences  
375 between the two groups (early adapt: 95% CI [-.34°/cycle, .22°/cycle],  $t(22)=-.47$ ;  $p=.64$ ;  $d=.19$ ; late  
376 learning: 95% CI [-7.80° 1.19°],  $t(22)=-1.52$ ;  $p=.14$ ;  $d=.62$ ). In contrast, the group reaching to bisected  
377 targets showed slower early adaptation rates (95% CI [-.81°/cycle, -.07°/cycle],  $t(22)=-2.49$ ,  $p=.02$ ,  
378  $d=1.02$ ) and lower magnitudes of late learning (95% CI [-12.58°, -1.35°],  $t=-2.57$ ,  $p=0.017$ ,  $d=1.05$ )  
379 when compared with the group reaching to small targets (Straddle-to-Hit). Given our analysis plan  
380 entailed multiple comparisons, we also performed an omnibus one-way ANOVA on the late learning  
381 data at the end of the acquisition phase. The effect of group was significant ( $F(2,33)=9.33$ ,  $p=.0006$ ,  
382  $\eta^2=.36$ ).

383

384 During the transfer phase, the target size for the perceptual uncertainty group remained large, but the  
385 bisection line was removed. If perceptual uncertainty contributed to the Hit Target effect, we would  
386 expect to observe a decrease in hand angle (since uncertainty would increase following transfer).  
387 However, following transfer to the non-bisected large target, there was no change in asymptote (95%  
388 CI [-.87°, 2.32°],  $t(11)=1.0$ ,  $p=.341$ ,  $d_z=.29$ ). In sum, the results from this control group indicate that the  
389 attenuated adaptation observed when the cursor is fully embedded within the target is not due to  
390 perceptual uncertainty,

391

392

393

## 394 **DISCUSSION**

395 The impact of reward on sensorimotor adaptation has been the focus of recent debate and  
396 investigation. A number of studies have demonstrated, either through the direct manipulation of reward  
397 (Galea et al., 2015; Nikooyan & Ahmed, 2015), or indirectly by varying task outcomes (Leow et al.,  
398 2018; Reichenthal, Avraham, Karniel, & Shmuelof, 2016; Jordan A. Taylor & Ivry, 2011), that task  
399 success signals can modulate performance changes in sensorimotor adaptation tasks. What remains  
400 unclear, however, is how to characterize the interaction of reward-based and error-based learning  
401 systems. Based on previous results and modeling work, reward signals have been hypothesized to  
402 operate on certain aspects of learning such as consolidation (e.g., Shmuelof et al., 2012 and Galea et  
403 al., 2015). Other studies suggest rewards are exploited by learning systems distinct from SPE-driven  
404 implicit adaptation (Codol, Holland, & Galea, 2018), with the resulting performance a composite of  
405 changes resulting from the independent operation of these different systems (Jordan A. Taylor & Ivry,  
406 2011; Jordan A. Taylor et al., 2014). The interpretation of the results from these studies is complicated  
407 by the fact that the experimental tasks conflate different learning processes. In the present study, we  
408 sought to avoid this complication by using a new method to study adaptation, one in which performance  
409 changes arise implicitly in response to an invariant visual error signal.

410

411 Using the visual clamp method (Morehead et al., 2017), we observed a striking difference between  
412 conditions in which the final position of the cursor was fully embedded in the target compared to  
413 conditions in which the cursor either terminated outside or straddled the target: When the cursor was  
414 fully embedded, the rate of learning was reduced and, more strikingly, the asymptotic level of learning  
415 was attenuated. Interestingly, the effect of varying the target size was qualitatively different than what  
416 we observed in previous studies in which we varied the angular direction of the clamp. In that work,  
417 small clamp angles reduced the rate of adaptation (Kim et al., 2018), but, over a large range of values,  
418 failed to produce reliable differences in asymptotic levels of learning (Kim et al., 2018; Morehead et al.,  
419 2017).

420

421 The difference in behavioral change as a function of relative target size was observed across different  
422 clamp sizes and did not appear to be because of differences in perceptual sensitivity or motor  
423 competence. This was supported by our control analyses, perceptual control experiment, and our  
424 finding that the Straddle group in Experiment 1 was similar to the Hit group, suggesting that the effect of  
425 target size was categorical. As such, we assume that the effect of target size on behavior arises from  
426 the generation of an intrinsic reward signal, one that is generated when the cursor lands fully within the  
427 target. In the final experiment, we explored two ways in which an intrinsic reward signal could impact  
428 performance. One hypothesis centered on the idea that reward modulates the strength of movement  
429 representations associated with task success, a variant of the idea that reward and SPE engage  
430 distinct representations and learning systems (Shmuelof et al., 2012). The other hypothesis considered  
431 a more direct modulatory impact on the adaptation process. The results showed that the differences in  
432 asymptote cannot be attributed solely to strengthening of rewarded movements. Rather, intrinsic  
433 reward directly attenuates the operation of the adaptation system.

434

435 We recognize that our interpretation of the results rests on the assumption that “hitting” the target with  
436 the cursor is intrinsically rewarding (Huang et al., 2011; Leow et al., 2018). If correct, this assumption

437 holds despite the participants' awareness that the angular motion of the cursor is causally unrelated to  
438 their behavior. Our earlier work with clamped feedback had shown that adaptation can be driven by a  
439 task-irrelevant error signal, the SPE defined by the difference between the cursor and target. Here we  
440 see the automatic operation of an intrinsic reward signal. Of course we do not have evidence,  
441 independent of the behavior, that hitting a target is rewarding; this might require using methods such as  
442 fMRI (Daw, Gershman, Seymour, Dayan, & Dolan, 2011) or pupillometry (Manohar, Finzi, Drew, &  
443 Husain, 2017) to assess the presence of well-established signatures of reward.

444

445 State space models have provided a concise computational account of sensorimotor adaptation  
446 (Huang, Haith, Mazzoni, & Krakauer, 2011; Smith, Ghazizadeh, & Shadmehr, 2006; Tanaka, Krakauer,  
447 & Sejnowski, 2012; Thoroughman & Shadmehr, 2000). In the simplest version, these models entail two  
448 parameters, a memory term,  $A$ , representing the retention of the current state from trial to trial, and a  
449 learning rate term,  $B$ , representing how the state is updated based on the error from the current trial  
450 (the  $A$  and  $U(e)$  terms in Eq. 1, respectively). Given this framework, we can consider how reward might  
451 modulate adaptation. One possibility is that reward modulates retention. This hypothesis is consistent  
452 with the results of a recent visuomotor adaptation study comparing groups that either received only  
453 cursor feedback or cursor feedback and a monetary reward, scaled to their accuracy. The latter showed  
454 greater retention during a washout block in which the feedback was removed (Galea et al., 2015).  
455 When the data were fit with the standard state space model, this effect was accounted for by an  
456 increase in the retention term,  $A$ , interpreted as indicating that rewarded movements are better  
457 consolidated.

458

459 A retention-based account, however, does not accord well with the current results. If the memory term  
460 was larger in conditions with intrinsic reward (i.e., Hit Target conditions), then we should have observed  
461 a higher asymptote when the cursor was embedded in the target compared to when it missed (or  
462 straddled) the target, since the SPE is invariant and more of the current state is retained from trial to

463 trial. The behavior went in the opposite direction of this prediction: The Hit Target conditions  
464 consistently resulted in lower asymptotic values. Thus, a retention-based account of the intrinsic reward  
465 effect would mandate lower values of  $A$ , a situation in which the memory term results in the adaptation  
466 system being resistant to learning from errors. We suspect that the washout results observed in Galea  
467 et al. (2015) are not due to a change in the adaptation process, but rather reflect the residual effects of  
468 an aiming strategy induced by the reward. That is, the monetary rewards might have reinforced a  
469 strategy during the rotation block, and this carried over into the washout block. Indeed, the idea that  
470 reward impacts strategic processes has been advanced in studies comparing conditions in which the  
471 performance could be enhanced by re-aiming (Codol et al., 2018; Holland et al., 2018).

472

473 Alternatively, intrinsic reward could influence overall learning by modulating the learning rate  
474 parameter. A priori, one might suppose that reward would enhance learning (Nikooyan & Ahmed,  
475 2015), either by increasing the sensitivity and responsiveness to error, or by promoting exploratory  
476 behavior to generate appropriate compensatory strategies. The latter would be a case where the  
477 learning rate parameter encompasses the effects of both implicit and explicit learning processes,  
478 especially relevant in standard adaptation studies where the task outcome is contingent on the  
479 participant's behavior and the perturbation is large (Bond & Taylor, 2015; Jordan A. Taylor & Ivry, 2011;  
480 Jordan A. Taylor et al., 2014).

481

482 In contrast, the clamp method, by eliminating the contribution of strategic processes, allows us to  
483 directly examine how reward might influence estimated rates of implicit learning. Here we see that the  
484 effect would suggest that reward reduces the learning rate, made salient by the parameter estimates  
485 from the acquisition phase of Experiment 3 (see also, Leow et al., 2018). A reduction in the learning  
486 rate can be conceptualized as a gain factor attenuating the system's response to error. In terms of the  
487 standard state space model, this would translate into reducing the system's sensitivity to error; in the  
488 motor correction variant of the state space model, this would translate into reducing the amount of

489 change induced by an error of a given size. In either conceptualization, the end result is that in the  
490 presence of an intrinsic reward signal, the error-dependent drive is reduced.

491

492 The hypothesis that reward attenuates the learning rate within the adaptation system provides a  
493 parsimonious account of the data from all three experiments. Following the introduction of the clamped  
494 feedback, a lower asymptote was observed in Experiments 1-3 when the cursor hit the target.  
495 Assuming the memory process is unaffected, the reduced error-dependent drive will result in a lower  
496 asymptote. Moreover, the rate of change in behavior, operationalized here by the early learning rate,  
497 should also be lower, a pattern evident in all three experiments, although only statistically significant in  
498 Experiment 3. Moreover, a change in learning drive can account for the behavioral effects observed in  
499 the transfer phase of Experiment 3. The loss of an intrinsic reward signal (Hit-to-Straddle group) would  
500 increase the error-dependent learning drive, resulting in an increase in hand angle. Conversely, the  
501 introduction of an intrinsic reward signal (Straddle-to-Hit group) would decrease the learning drive,  
502 resulting in a drop in hand angle.

503

504 Although the results indicate that reward directly modulates adaptation, the observed changes in  
505 behavior during the transfer phase of Experiment 3 were more gradual than predicted based on  
506 parameter estimates derived from the initial acquisition phase data. The quantitative predictions here  
507 assume that the adaptation system is time-invariant. This assumption may be too rigid; for example,  
508 learning parameters may change with increased exposure to a perturbation (Mawase, Shmuelof, Bar-  
509 Haim, & Karniel, 2014; Zarah, Weston, Liang, Mazzoni, & Krakauer, 2008) or change as a function of  
510 the context (Herman, Harwood, & Wallman, 2009).

511

512 We also recognize that behavioral changes here may reflect the operation of multiple processes  
513 (Krakauer & Mazzoni, 2011), and the composite effects of these processes might account for why the  
514 observed changes were more gradual than predicted. For example, intrinsic reward may not only

515 directly modulate adaptation, but may also reinforce an executed movement (Castro, Monsen, & Smith,  
516 2011), a combination of the Movement Reinforcement and Adaptation Modulation models sketched in  
517 Figure 4. For example, in the Straddle-to-Hit condition, the introduction of intrinsic reward at transfer  
518 would reinforce movements at the initial asymptote, resisting the effect of reduced learning drive.

519

520 Studies involving non-human primates and rodents have provided insights into possible neural  
521 substrates supporting the interaction of systems involved in reward- and error-based learning.  
522 Converging evidence points to a critical role for the cerebellum in sensorimotor adaptation (Butcher et  
523 al., 2017; Izawa, Criscimagna-Hemminger, & Shadmehr, 2012; J A Taylor, Klemfuss, & Ivry, 2010;  
524 Tseng et al., 2007), including the observation that patients with cerebellar degeneration show a  
525 reduced response to visual error clamps (Morehead et al., 2017). Reward-based learning is associated  
526 with a more distributed network of cortical and subcortical areas, including a prominent role for  
527 dopaminergic signals in the basal ganglia (Schultz, 2015). Neuroanatomical studies have identified di-  
528 synaptic reciprocal connections between the basal ganglia and cerebellum (Bostan, Dum, & Strick,  
529 2010), as well as direct connections between the cerebellum and dopaminergic nuclei in the brainstem  
530 (Perciavalle, Berretta, & Raffaele, 1989; Watabe-Uchida, Zhu, Ogawa, Vamanrao, & Uchida, 2012).  
531 These connections might provide a relatively direct pathway for reward signals to modulate cerebellar  
532 activity. Alternatively, or perhaps complementary, recent work has indicated that both simple (Wagner,  
533 Kim, Savall, Schnitzer, & Luo, 2017) and complex (Ohmae & Medina, 2015) spike activity in the  
534 cerebellum may signal information about rewards or anticipated rewards. This work suggests a more  
535 expansive view may be required to understand cerebellar function, one in which error-based learning is  
536 modulated by contextual factors. The current study provides a striking example of how intrinsic reward  
537 signals, in the form of persistent target hits, may serve as one such contextual factor that can modulate  
538 cerebellar-dependent sensorimotor adaptation.

539

540

541 **ACKNOWLEDGMENTS**

542 We thank Matthew Hernandez and Wendy Shwe for assistance with data collection. We are also  
543 grateful to Maurice Smith, Ryan Morehead, Guy Avraham, and Ian Greenhouse for helpful discussions  
544 regarding this work.

545

546

547 **Competing Interests**

548 No competing interests, financial or otherwise, are declared by the authors.



549 **REFERENCES**

- 550 Bond, K. M., & Taylor, J. A. (2015). Flexible explicit but rigid implicit learning in a visuomotor adaptation  
551 task. *Journal of Neurophysiology*, *113*(10), 3836–3849. <https://doi.org/10.1152/jn.00009.2015>  
552 Bostan, A. C., Dum, R. P., & Strick, P. L. (2010). The basal ganglia communicate with the cerebellum.  
553 *Proceedings of the National Academy of Sciences*, *107*(18), 8452–8456.  
554 <https://doi.org/10.1073/pnas.1000496107>  
555 Butcher, P. A., Ivry, R. B., Kuo, S.-H., Rydz, D., Krakauer, J. W., & Taylor, J. A. (2017). The cerebellum  
556 does more than sensory prediction error-based learning in sensorimotor adaptation tasks.  
557 *Journal of Neurophysiology*, *118*(3), 1622–1636. <https://doi.org/10.1152/jn.00451.2017>  
558 Castro, L. N. G., Monsen, C. B., & Smith, M. A. (2011). The Binding of Learning to Action in Motor  
559 Adaptation. *PLOS Computational Biology*, *7*(6), e1002052.  
560 <https://doi.org/10.1371/journal.pcbi.1002052>  
561 Churchland, M. M., Afshar, A., & Shenoy, K. V. (2006). A Central Source of Movement Variability.  
562 *Neuron*, *52*(6), 1085–1096. <https://doi.org/10.1016/j.neuron.2006.10.034>  
563 Codol, O., Holland, P. J., & Galea, J. M. (2017). The relationship between reinforcement and explicit  
564 strategies during visuomotor adaptation. *BioRxiv*, 206284.  
565 Codol, O., Holland, P. J., & Galea, J. M. (2018). The relationship between reinforcement and explicit  
566 control during visuomotor adaptation. *Scientific Reports*, *8*(1), 9121.  
567 <https://doi.org/10.1038/s41598-018-27378-1>  
568 Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences  
569 on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215.  
570 <https://doi.org/10.1016/j.neuron.2011.02.027>  
571 Fitts, P. M. (1992). The information capacity of the human motor system in controlling the amplitude of  
572 movement. *Journal of Experimental Psychology: General*, *121*(3), 262–269.  
573 <https://doi.org/10.1037/0096-3445.121.3.262>  
574 Galea, J. M., Mallia, E., Rothwell, J., & Diedrichsen, J. (2015). The dissociable effects of punishment  
575 and reward on motor learning. *Nature Neuroscience*, *18*(4), 597–602.  
576 <https://doi.org/10.1038/nn.3956>  
577 Gallivan, J. P., Logan, L., Wolpert, D. M., & Flanagan, J. R. (2016). Parallel specification of competing  
578 sensorimotor control policies for alternative action options. *Nature Neuroscience*, *19*(2), 320–  
579 326. <https://doi.org/10.1038/nn.4214>  
580 Gordon, J., Ghilardi, M. F., & Ghez, C. (1994). Accuracy of planar reaching movements - I.  
581 Independence of direction and extent variability. *Experimental Brain Research*, *99*(1), 97–111.  
582 <https://doi.org/10.1007/BF00241415>  
583 Hadjiosif, A., & Smith, M. (2013). Savings is restricted to the temporally labile component of motor  
584 adaptation. In *Translational and Computational Motor Control (TCMC)*. San Diego, CA.  
585 Holland, P. J., Codol, O., & Galea, J. M. (2018). The contribution of explicit processes to reinforcement-  
586 based motor learning. *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00901.2017>  
587 Huang, V. S., Haith, A., Mazzoni, P., & Krakauer, J. W. (2011). Rethinking motor learning and savings  
588 in adaptation paradigms: model-free memory for successful actions combines with internal  
589 models. *Neuron*, *70*(4), 787–801. <https://doi.org/10.1016/j.neuron.2011.04.012>  
590 Huberdeau, D. M., Krakauer, J. W., & Haith, A. M. (2015). Dual-process decomposition in human  
591 sensorimotor adaptation. *Curr Opin Neurobiol*, *33*, 71–77.  
592 <https://doi.org/10.1016/j.conb.2015.03.003>  
593 Izawa, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2012). Cerebellar contributions to reach  
594 adaptation and learning sensory consequences of action. *J Neurosci*, *32*(12), 4230–4239.  
595 <https://doi.org/10.1523/jneurosci.6353-11.2012>  
596 Izawa, J., & Shadmehr, R. (2011). Learning from sensory and reward prediction errors during motor  
597 adaptation. *PLoS Comput Biol*, *7*(3), e1002012–e1002012.  
598 <https://doi.org/10.1371/journal.pcbi.1002012>

- 599 Kim, H. E., Morehead, J. R., Parvin, D. E., Moazzezi, R., & Ivry, R. B. (2018). Invariant errors reveal  
600 limitations in motor correction rather than constraints on error sensitivity. *Communications*  
601 *Biology*, 1(1), 19. <https://doi.org/10.1038/s42003-018-0021-y>
- 602 Knill, D. C., Bondada, A., & Chhabra, M. (2011). Flexible, Task-Dependent Use of Sensory Feedback to  
603 Control Hand Movements. *Journal of Neuroscience*, 31(4), 1219–1237.  
604 <https://doi.org/10.1523/JNEUROSCI.3522-09.2011>
- 605 Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*,  
606 427(6971), 244–247. <https://doi.org/10.1038/nature02169>
- 607 Krakauer, J. W. (2005). Adaptation to Visuomotor Transformations: Consolidation, Interference, and  
608 Forgetting. *Journal of Neuroscience*, 25(2), 473–478.  
609 <https://doi.org/10.1523/JNEUROSCI.4218-04.2005>
- 610 Krakauer, J. W., & Mazzoni, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond.  
611 *Curr Opin Neurobiol*, 21(4), 636–644. <https://doi.org/10.1016/j.conb.2011.06.012>
- 612 Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical  
613 primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4.  
614 <https://doi.org/10.3389/fpsyg.2013.00863>
- 615 Leow, L.-A., Marinovic, W., & Carroll, T. (2018). Task errors contribute to implicit remapping in  
616 sensorimotor adaptation. *BioRxiv*, 263988.
- 617 Manohar, S. G., Finzi, R. D., Drew, D., & Husain, M. (2017). Distinct Motivational Effects of Contingent  
618 and Noncontingent Rewards. *Psychological Science*, 28(7), 1016–1026.  
619 <https://doi.org/10.1177/0956797617693326>
- 620 Mawase, F., Shmuelof, L., Bar-Haim, S., & Karniel, A. (2014). Savings in locomotor adaptation  
621 explained by changes in learning parameters following initial adaptation. *Journal of*  
622 *Neurophysiology*, 111(7), 1444–1454. <https://doi.org/10.1152/jn.00734.2013>
- 623 McDougle, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking Aim at the Cognitive Side of Learning in  
624 Sensorimotor Adaptation Tasks. *Trends in Cognitive Sciences*, 20(7), 535–544.  
625 <https://doi.org/10.1016/j.tics.2016.05.002>
- 626 Morehead, J. R., Taylor, J. A., Parvin, D. E., & Ivry, R. B. (2017). Characteristics of Implicit  
627 Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. *Journal of Cognitive*  
628 *Neuroscience*, 29(6), 1061–1074. [https://doi.org/10.1162/jocn\\_a\\_01108](https://doi.org/10.1162/jocn_a_01108)
- 629 Nikooyan, A. A., & Ahmed, A. A. (2015). Reward feedback accelerates motor learning, 633–646.  
630 <https://doi.org/10.1152/jn.00032.2014>
- 631 Ohmae, S., & Medina, J. F. (2015). Climbing fibers encode a temporal-difference prediction error during  
632 cerebellar learning in mice. *Nature Neuroscience*, 18(12), 1798–1803.  
633 <https://doi.org/10.1038/nn.4167>
- 634 Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory.  
635 *Neuropsychologia*, 9(1), 97–113.
- 636 Perciavalle, V., Berretta, S., & Raffaele, R. (1989). Projections from the intracerebellar nuclei to the  
637 ventral midbrain tegmentum in the rat. *Neuroscience*, 29(1), 109–119.  
638 [https://doi.org/10.1016/0306-4522\(89\)90336-9](https://doi.org/10.1016/0306-4522(89)90336-9)
- 639 Reichenthal, M., Avraham, G., Karniel, A., & Shmuelof, L. (2016). Target size matters: target errors  
640 contribute to the generalization of implicit visuomotor learning. *Journal of Neurophysiology*,  
641 116(2), 411–424. <https://doi.org/10.1152/jn.00830.2015>
- 642 Schultz, W. (2015). Neuronal Reward and Decision Signals: From Theories to Data. *Physiol Rev*, 95(3),  
643 853–951. <https://doi.org/10.1152/physrev.00023.2014>
- 644 Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and  
645 adaptation in motor control. *Annu Rev Neurosci*, 33, 89–108. <https://doi.org/10.1146/annurev-neuro-060909-153135>
- 647 Shmuelof, L., Huang, V. S., Haith, A. M., Delnicki, R. J., Mazzoni, P., & Krakauer, J. W. (2012).  
648 Overcoming Motor “Forgetting” Through Reinforcement Of Learned Actions. *Journal of*  
649 *Neuroscience*, 32(42), 14617–14621a. <https://doi.org/10.1523/JNEUROSCI.2184-12.2012>

- 650 Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting Adaptive Processes with Different  
651 Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, 4(6), e179.  
652 <https://doi.org/10.1371/journal.pbio.0040179>
- 653 Soechting, J. F. (1984). Effect of target size on spatial and temporal characteristics of a pointing  
654 movement in man. *Experimental Brain Research*, 54(1), 121–132.  
655 <https://doi.org/10.1007/BF00235824>
- 656 Tanaka, H., Krakauer, J. W., & Sejnowski, T. J. (2012). Generalization and Multirate Models of Motor  
657 Adaptation. *Neural Computation*, 24(4), 939–966. [https://doi.org/10.1162/NECO\\_a\\_00262](https://doi.org/10.1162/NECO_a_00262)
- 658 Taylor, J. A., Klemfuss, N. M., & Ivry, R. B. (2010). An explicit strategy prevails when the cerebellum fails  
659 to compute movement errors. *Cerebellum*, 9(4), 580–586. <https://doi.org/10.1007/s12311-010-0201-x>
- 660 Taylor, Jordan A., & Ivry, R. B. (2011). Flexible Cognitive Strategies during Motor Learning. *PLoS*  
661 *Computational Biology*, 7(3), e1001096. <https://doi.org/10.1371/journal.pcbi.1001096>
- 662 Taylor, Jordan A., & Ivry, R. B. (2014). Cerebellar and Prefrontal Cortex Contributions to Adaptation,  
663 Strategies, and Reinforcement Learning. In *Progress in Brain Research* (Vol. 210, pp. 217–  
664 253). Elsevier. <https://doi.org/10.1016/B978-0-444-63356-9.00009-1>
- 665 Taylor, Jordan A., Krakauer, J. W., & Ivry, R. B. (2014). Explicit and Implicit Contributions to Learning in  
666 a Sensorimotor Adaptation Task. *Journal of Neuroscience*, 34(8), 3023–3032.  
667 <https://doi.org/10.1523/JNEUROSCI.3619-13.2014>
- 668 Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2016). Effective reinforcement learning following  
669 cerebellar damage requires a balance between exploration and motor noise. *Brain*, 139(1),  
670 101–114. <https://doi.org/10.1093/brain/awv329>
- 671 Therrien, A. S., Wolpert, D. M., & Bastian, A. J. (2018). Increasing Motor Noise Impairs Reinforcement  
672 Learning in Healthy Individuals. *Eneuro*, ENEURO.0050-18.2018.  
673 <https://doi.org/10.1523/ENEURO.0050-18.2018>
- 674 Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor  
675 primitives. *Nature*, 407(6805), 742–747. <https://doi.org/10.1038/35037588>
- 676 Tseng, Y. W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory  
677 prediction errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol*, 98(1), 54–  
678 62. <https://doi.org/10.1152/jn.00266.2007>
- 679 Vandevorde, K., & Orban de Xivry, J.-J. (2018). Motor adaptation but not internal model recalibration  
680 declines with aging. *BioRxiv*. <https://doi.org/10.1101/292250>
- 681 Vaswani, P. A., Shmuelof, L., Haith, A. M., Delnicki, R. J., Huang, V. S., Mazzoni, P., ... Krakauer, J.  
682 W. (2015). Persistent Residual Errors in Motor Adaptation Tasks: Reversion to Baseline and  
683 Exploratory Escape. *Journal of Neuroscience*, 35(17), 6969–6977.  
684 <https://doi.org/10.1523/JNEUROSCI.2656-14.2015>
- 685 Wagner, M. J., Kim, T. H., Savall, J., Schnitzer, M. J., & Luo, L. (2017). Cerebellar granule cells encode  
686 the expectation of reward. *Nature*. <https://doi.org/10.1038/nature21726>
- 687 Watabe-Uchida, M., Zhu, L., Ogawa, S. K., Vamanrao, A., & Uchida, N. (2012). Whole-Brain Mapping  
688 of Direct Inputs to Midbrain Dopamine Neurons. *Neuron*, 74(5), 858–873.  
689 <https://doi.org/10.1016/j.neuron.2012.03.017>
- 690 Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects  
691 saccade velocities. *Experimental Brain Research*, 196(4), 475–481.  
692 <https://doi.org/10.1007/s00221-009-1879-1>
- 693 Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining Savings for  
694 Visuomotor Adaptation: Linear Time-Invariant State-Space Models Are Not Sufficient. *Journal of*  
695 *Neurophysiology*, 100(5), 2537–2548. <https://doi.org/10.1152/jn.90529.2008>
- 696  
697  
698  
699  
700

701 **METHODS**

702 *Participants:* Healthy, young adults (N=116, 69 females, age =  $20.9 \pm 2.1$  years old) were recruited from  
703 the University of California, Berkeley, community. Each participant was tested in only one experiment.  
704 All participants were right-handed, as verified with the Edinburgh Handedness Inventory (Oldfield,  
705 1971). Participants provided informed consent and received financial compensation for their  
706 participation. The Institutional Review Board at UC Berkeley approved all experimental procedures.

707

708 *Experimental Apparatus:* The participant was seated at a custom-made tabletop housing an LCD  
709 screen (53.2 cm by 30 cm, ASUS), mounted 27 cm above a digitizing tablet (49.3 cm by 32.7 cm,  
710 Intuos 4XL; Wacom, Vancouver, WA). The participant made reaching movements by sliding a modified  
711 air hockey "paddle" containing an embedded stylus. The position of the stylus was recorded by the  
712 tablet at 200 Hz. The experimental software was custom written in Matlab, using the Psychtoolbox  
713 extensions<sup>26</sup>.

714

715 *Reaching Task:* Center-out planar reaching movements were performed from the center of the  
716 workspace to targets positioned at a radial distance of 8 cm. Direct vision of the hand was occluded by  
717 the monitor, and the lights were extinguished in the room to minimize peripheral vision of the arm. The  
718 start location and target location were indicated by white and blue circles, respectively (start circle: 6  
719 mm in diameter; target: either 6, 9.8 or 16 mm depending on condition).

720

721 To initiate each trial, the participant moved the digitizing stylus into the start location. The position of the  
722 stylus was indicated by a white feedback cursor (3.5 mm diameter). Once the start location was  
723 maintained for 500 ms, the target appeared. For Experiments 1 and 3, the target could appear at one of  
724 8 locations, placed in 45° increments around a virtual circle (0°, 45°, 95°, 135°, 180°, 225°, 270°, 315°).  
725 For Experiment 2, the target could appear at one of four locations placed in 90° increments around a  
726 virtual circle (45°, 135°, 225°, 315°). We reduced the number of targets from 8 to 4 in this experiment in

727 order to increase the overall number of training cycles with the clamp, while keeping the experiment  
728 under 1.5 hours, and so that participants would reach a stable asymptote. Participants were instructed  
729 to accurately and rapidly "slice" through the target, without needing to stop at the target location. Visual  
730 feedback, when presented, was provided during the reach until the movement amplitude exceeded 8  
731 cm. As described below, the feedback either matched the position of the stylus (veridical) or followed a  
732 fixed path (clamped). If the movement was not completed within 300 ms, the words "too slow" were  
733 generated by the sound system of the computer.

734

735 After the hand crossed the target ring, endpoint cursor feedback was provided for 50 ms either at the  
736 position in which the hand crossed the virtual target ring (veridical feedback) or at a fixed distance  
737 determined by the size of the clamp. During the return movement, the feedback cursor reappeared  
738 when the participant's hand was within 1 cm of the start.

739

740 *Experimental Feedback Conditions:* Across the experimental session, there were three types of visual  
741 feedback. On no-feedback trials, the cursor disappeared when the participant's hand left the start circle  
742 and only reappeared at the end of the return movement. On veridical feedback trials, the cursor  
743 matched the position of the stylus during the 8 cm outbound segment of the reach. On clamped  
744 feedback trials, the feedback followed a path that was fixed along a specific hand angle. The radial  
745 distance of the cursor from the start location was still based on the radial extent of the participant's  
746 hand during the 8 cm outbound segment, but the angular position was fixed relative to the target (i.e.,  
747 independent of the angular position of the hand).

748

749 The primary instructions to the participant remained the same across the experimental session:  
750 Specifically, that they were to reach directly towards the visual target. Prior to the introduction of the  
751 clamped feedback trials, participants were briefed about the feedback manipulation. They were  
752 informed that the position of the cursor would now follow a fixed trajectory and that the angular position

753 would be independent of their movement. They were explicitly instructed to ignore the cursor and  
754 continue to reach directly to the target. Participants also performed three instructed trials with the clamp  
755 perturbation on. During these practice trials, a target appeared at the 90 deg location (straight ahead),  
756 and the experimenter instructed the participant to first “reach straight to the left” (ie, 180 deg). For the  
757 second practice trial, the participant was instructed to “reach straight to the right” (0 deg). For the last  
758 trial, the participant was instructed to “reach straight down (towards your torso)” (ie, 270 deg). The  
759 purpose of these trials was to familiarize the participant with the exact clamp condition they were about  
760 to experience. Following these three practice trials, the experimenter confirmed with the participant they  
761 understood now what was meant by clamped visual feedback. These practice trials were removed from  
762 future analyses.

763

764 The same instructions in abbreviated form (“Ignore the cursor and move your hand directly to the target  
765 location”) were repeated verbally and with onscreen text at every block break during the clamp  
766 perturbation. Participants were debriefed at the end of the experiment and asked whether they ever  
767 intentionally tried to reach to locations other than the target. All subjects reported aiming to the target  
768 throughout the experiment.

769

770 We counterbalanced clockwise and counterclockwise clamp offsets within each group for all three  
771 experiments.

772

### 773 *Experiment 1*

774 Participants (n=48, 16/group) were randomly assigned to one of three groups, each training with a 3.5°  
775 clamp but differing only in terms of the size of the target: 6mm, 9.8, or 16 mm diameter. These sizes  
776 were chosen so that at an 8 cm radial distance the clamped cursor would be adjacent to the target  
777 without making any contact (Target Miss group), straddling the target by being roughly half inside and  
778 half outside the target (Straddle Target group), or fully embedded within the target (Hit Target group).

779 The Euclidean distance for this clamp offset, measured from the centers of cursor and target, was 4.9  
780 mm.

781

782 The session began with two baseline blocks, the first comprised of 5 movement cycles (40 total  
783 reaches to 8 targets) without visual feedback and the second comprised of 10 cycles with a veridical  
784 cursor displaying hand position. The experimenter then informed the participant that the visual  
785 feedback would no longer be veridical and would now be clamped at a fixed angle from the target  
786 location. Immediately following these general instructions, the experimenter continued providing  
787 instructions for the three practice trials which immediately followed (see Experimental Feedback  
788 Conditions). After the practice trials and confirming the participant's understanding of the task, the  
789 clamp block ensued for a total of 80 cycles. A short break (<1 min), as well as a reminder of the task  
790 instructions, was provided after 40 cycles (i.e., at the halfway point of this block). Immediately following  
791 the perturbation block, there were two washout blocks, first a 5 cycle block in which there was no visual  
792 feedback, followed by 10 cycles with veridical visual feedback. These blocks were preceded by  
793 instructions regarding the change in experimental condition and participants were reminded to always  
794 aim for the target and to attempt to slice through it with their hand.

795

## 796 *Experiment 2*

797 In Experiment 2 we assessed adaptation over an extended number of clamped visual feedback trials.  
798 The purpose of extending the perturbation block was to ensure that participants reached asymptotic  
799 levels of learning. In order to achieve a greater number of training cycles, we reduced the number of  
800 target locations within the set from 8 to 4.

801

802 Participants (n=32, 16/group) trained with a 1.75° clamp (2.4 mm distance between target and cursor  
803 centers) and were assigned to either a small (Straddle) or large (Hit) target condition. The session  
804 started with two baseline blocks, 10 cycles (40 reaches) without visual feedback and then 10 cycles

805 with veridical feedback. Following 3 practice trials with the clamp, the number of cycles in the clamped  
806 visual feedback block was nearly tripled from that of Experiment 1 to 220 cycles, with breaks provided  
807 after every 70 cycles. Following 220 cycles of training with a 1.75° clamp, there were two washout  
808 blocks, first a 10 cycle block in which there was a 0° clamp, followed by 10 cycles with veridical visual  
809 feedback. Prior to washout, participants were again instructed to always aim directly to the target.

810

### 811 Experiment 3

812 We assume that with the large targets, an intrinsic reward is generated when the cursor lands within the  
813 target. This could serve as a positive reinforcement signal, strengthening the representation of  
814 rewarded movements, and operating as a resistance to the learning drive associated with an SPE  
815 (Movement Reinforcement model). Alternatively, intrinsic reward may directly modulate the output of  
816 the adaptation system (Adaptation Modulation model). As a test of these hypotheses, we tested two  
817 main groups (n=12/group) in Experiment 3, using a 1.75° clamp in a transfer design. The session  
818 started with two baseline blocks, 5 cycles (40 reaches) without visual feedback and then 5 cycles with  
819 veridical feedback. After the baseline blocks, clamp instructions and three practice trials were provided  
820 to all participants. The first clamp block lasted 120 cycles, with participants training with either a small  
821 or large target. Following the first 120 cycles, the target sizes were reversed for the next 80 (Straddle-  
822 to-Hit or Hit-to-Straddle conditions). Our main predictions focused on the transfer phase, comparing the  
823 behavior to the predictions of both Movement Reinforcement and Adaptation Modulation models.  
824 Breaks of < 1 min were provided after every 35 cycles of training. On the break preceding the transfer  
825 (15 cycles before target switch), participants were told that everything would continue on as before,  
826 except that the target size would change at some point during the block. The purpose of staggering the  
827 break with the transfer was to mitigate any change in adaptation due to temporal decay that could result  
828 from a break in training (Hadjiosif & Smith, 2013).

829

### 830 *Control group*



831 A third group (n=12) was added to test whether the attenuation of adaptation in the large target  
832 condition was due to perceptual uncertainty. Here, the block structure was identical to the first two  
833 groups. We used a modified large target (16mm), one which had a bright green bisecting line through  
834 the middle, aligned with the target direction. The clamped cursor always fell within one half of the target  
835 (either clockwise or counter-clockwise depending on the condition), thus providing a clear indication  
836 that the cursor was off center. At the transfer, the bisecting line was removed and participants trained  
837 for 80 cycles with the standard large target.

838

### 839 *Data Analysis*

840 All statistical analyses and modeling were performed using Matlab 2015b and the Statistics Toolbox.  
841 The primary dependent variable in all experiments was hand angle at peak radial velocity, defined by  
842 the angle of the hand relative to the target at the time of peak radial velocity (i.e., angle between lines  
843 connecting start position to target and start position to hand). Throughout the text, we refer to this  
844 variable as hand angle. Additional analyses were performed using hand angle at “endpoint” (angle of  
845 the hand as it crossed the invisible target ring) rather than peak radial velocity. The results were  
846 essentially identical for the two dependent variables; as such, we only report the results of the analyses  
847 using peak radial velocity.

848

849 Outlier responses were removed from the analyses. For the sole purpose of identifying outliers, the  
850 Matlab “smooth” function was used to calculate a moving average (using a 5-trial window) of the hand  
851 angle data for each target location. Outliers were trials in which the observed hand angle was greater  
852 than 90° or deviated by more than 3 standard deviations from the moving average. In total, less than  
853 0.8% of trials overall were removed, and the most trials removed for any individual across all three  
854 experiments was 2%.

855

856 Individual baseline biases for each target location were subtracted from all data. Biases were defined  
857 as the average hand angles across cycles 2-10 (Experiments 1 and 2) or 2-5 (Experiment 3) of the  
858 feedback baseline block. These same cycles were used to calculate mean RTs, MTs, and movement  
859 variability (SD). To calculate each participant's baseline RT or MT, we took the average of median  
860 values at each target location. To calculate each participant's movement variability, we took the  
861 average of the standard deviations of hand angles at each target location.

862

863 In order to pool all of the data and to aid visualization, we flipped the hand angles for all participants  
864 clamped in the counterclockwise direction.

865

866 For Experiments 1 and 3, movement cycles consisted of 8 consecutive reaches (1 reach/target); for  
867 Experiment 2, we only used four targets, thus a movement cycle consisted of 4 consecutive reaches (1  
868 reach/target). Early adaptation rate was quantified by averaging the hand angle values over cycles 3-7  
869 of the clamp, and dividing by the number of cycles (i.e., 5) to get an estimate of the per trial rate of  
870 change in hand angle. We opted to use this measure of early adaptation rather than obtain parameter  
871 estimates from exponential fits since the latter approach gives considerable weight to the asymptotic  
872 phase of performance and, therefore would be less sensitive to early differences in rate. This would be  
873 especially problematic in Experiment 2, which utilized 220 clamp cycles. Asymptotic adaptation was  
874 defined as the last 10 cycles within a clamp block. In Experiment 1, the aftereffect was quantified by  
875 using the data from the first no-feedback cycle following the last clamp cycle. We also performed a  
876 secondary analysis of early adaptation rates using cycles 2-11 (Krakauer, 2005), rather than 3-7.  
877 Results from using this alternate metric were consistent with the reported analyses (i.e., slower rates for  
878 Hit Target groups), only they resulted in larger effect sizes due to the gradually increasing divergence of  
879 learning functions.

880

881 All t-tests were two-tailed. Posthoc pairwise comparisons following significant ANOVAs were performed  
882 using two-tailed t-tests. Cohen's d, eta squared ( $\eta^2$ ), partial eta squared (for mixed model ANOVA), and  
883  $d_z$  (for within-subjects design) values are provided as standardized measures of effect size (Lakens,  
884 2013). Values in main text are reported as 95% CIs in brackets and mean  $\pm$  SEM.

885  
886

### 887 *Modeling*

888 For our model fitting and simulation procedures we applied standard bootstrapping techniques,  
889 constructing group-averaged hand angle data 1000 times by randomly resampling with replacement  
890 from the pool of participants within each group. Using Matlab's *fmincon* function, we started with ten  
891 different initial sets of parameter values and estimated the retention and learning parameters which  
892 minimized the least squared error between the bootstrapped data and model output ( $x_n$ ).

893

894 We also fit the model to the acquisition phase data of each participant in Experiment 3 in order to  
895 compare parameter values between groups using a non-parametric permutation test. We first  
896 calculated our two test statistics, the average difference in *A* values and the average difference in *U*  
897 values between groups. Then, we randomly shuffled the group assignments using 10000 Monte Carlo  
898 simulations in order to create the null distributions for mean *A* and *U* parameter values, separately. We  
899 then calculated exact p-values by summing the proportion of each respective null distribution that was  
900 at least as or more extreme than our test statistic values (i.e., using 2-sided tests).

901

902 No statistical methods were used to predetermine sample sizes. The chosen sample sizes were based  
903 on our previous study using the clamp method (Kim et al., 2018; Morehead et al., 2017), as well as  
904 prior psychophysical studies of human sensorimotor learning (Galea et al., 2015; Gallivan, Logan,  
905 Wolpert, & Flanagan, 2016; Huang et al., 2011; Vaswani et al., 2015).

906

907 **SUPPLEMENTAL INFORMATION**

908

909 *Kinematic variables*

910

911 *Experiment 1*

912

913 Average movement variability across the eight targets during cycles 2-10 of the veridical feedback

914 baseline block were not different between groups (variability:  $F(2,45)=0.232$ ,  $p=.110$ ,  $\eta^2=.093$ ).

915 Movement times across groups were not different ( $F(2,45)=2.19$ ,  $p=.123$ ,  $\eta^2=.089$ ). However, we did

916 observe a difference in baseline RTs ( $F(2,45)=4.48$ ,  $p=.017$ ,  $\eta^2=.166$ ), with post hoc Tukey-Kramer

917 tests confirming that the large target (Hit) group had faster RTs ( $325 \pm 7$  ms) than the small target

918 (Miss) group ( $387 \pm 22$  ms). The medium target (Straddle) group's RTs ( $362 \pm 12$  ms) were not reliably

919 different from either group. This baseline difference in RTs was only observed in this experiment (see

920 below), and there was no correlation between baseline RT and late learning for the large target group ( $r$

921  $= .09$ ,  $p = .73$ ), suggesting that RTs are not associated with the magnitude of adaptation.

922

923 *Experiment 2*

924 Baseline movement variability was marginally less for the group reaching to the larger target (Hit Target

925 group:  $3.09^\circ \pm .18^\circ$ ; Straddle Target group:  $3.56^\circ \pm .16^\circ$ ;  $t(30)=-1.99$   $p=.056$ ,  $d=0.70$ ). Further

926 supporting the argument that planning was no different across conditions, neither reaction times nor

927 movement times differed between the groups ( $t(30)=-0.183$ ,  $p=.856$ ,  $d=.06$  and  $t(30)=0.71$ ,  $p=.484$ ,

928  $d=.25$ , respectively).

929

930 *Experiment 3*

931 Baseline movement variability was not different across all three groups, including the control group

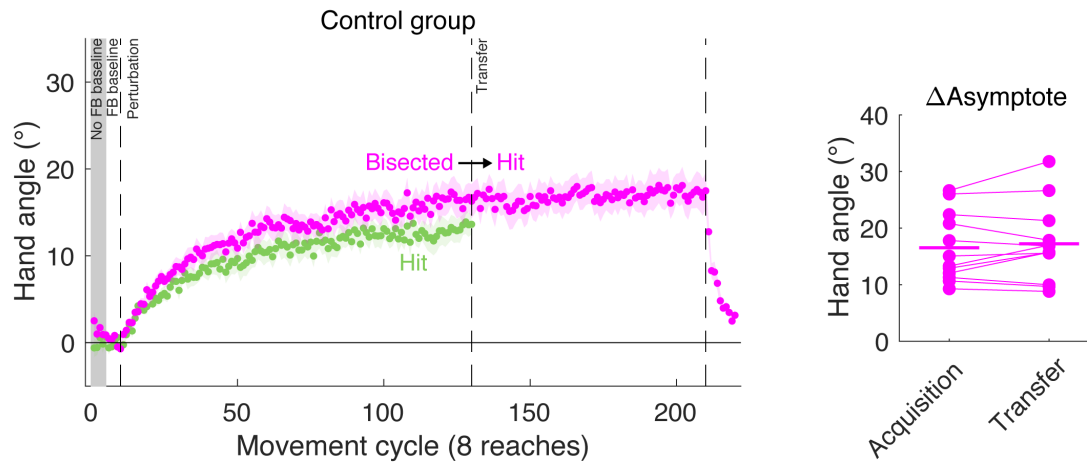
932 trained with the bisected target ( $F(2,33)=1.38$ ,  $p=.267$ ,  $\eta^2=.077$ ). Also, no differences across groups

933 were observed for either RTs ( $F(2,33)=1.51$ ,  $p=.236$ ,  $\eta^2=.0084$ ) or MTs ( $F(2,33)=.46$ ,  $p=.634$ ,  $\eta^2=.027$ ).

934

935 *Control group from Experiment 3*

936 The behavior of the control group reaching to a large target bisected by a line (aligned with the target  
937 direction) during the acquisition phase is shown in magenta. The change in hand angle was not  
938 significantly different than that observed for the group that was tested with the large target in the  
939 acquisition phase of Experiment 3 (replotted here in green), suggesting that perceptual uncertainty did  
940 not make a substantive contribution to the effects of hitting the target. We omitted the transfer behavior  
941 of the large target group as this was when the large target was replaced with the small target.  
942



943