The influence of task outcome on implicit motor learning

Hyosub E. Kim[1,2,3]*, Darius E. Parvin[1,2], Richard B. Ivry[1,2]

[1]Department of Psychology and [2]Helen Wills Neuroscience Institute, University of California, Berkeley
[3]Department of Physical Therapy, University of Delaware

ABSTRACT

Recent studies have demonstrated that task success signals can modulate learning during sensorimotor adaptation tasks, primarily through engaging explicit processes. Here we examine the influence of task outcome on implicit adaptation, using a reaching task in which adaptation is induced by feedback that is not contingent on actual performance. We imposed an invariant perturbation (rotation) on the feedback cursor while varying the target size. In this way, the cursor either hit or missed the target, with the former producing a marked attenuation of implicit motor learning. We explored different computational architectures that might account for how task outcome information interacts with implicit adaptation. The results fail to support an architecture in which adaptation operates in parallel with a model-free operant reinforcement process. Rather, task outcome may serve as a gain on implicit adaptation or provide a distinct error signal for a second model-based process, in addition to implicit adaptation.

INTRODUCTION

Multiple learning processes contribute to successful goal-directed actions in the face of changing physiological states, body structures, and environments (Taylor et al., 2014; Huberdeau et al., 2015; McDougle et al., 2016). Among these processes, implicit sensorimotor adaptation is of primary importance for maintaining appropriate calibration of sensorimotor maps over both short and long timescales. A large body of work has focused on how sensory prediction error (SPE), the difference between predicted and actual sensory feedback, drives sensorimotor adaptation (Shadmehr et al., 2010). In addition, there is growing appreciation of the contribution of other processes to sensorimotor learning, including strategic aiming and reward-based learning (Taylor et al., 2014; Wu et al., 2014; Bond and Taylor, 2015; Galea et al., 2015; Nikooyan and Ahmed, 2015; Summerside et al., 2018). In terms of the latter, several recent studies have shown that rewarding successful actions alone is sufficient to learn a new sensorimotor mapping (Izawa and Shadmehr, 2011; Therrien et al., 2016, 2018).

Little is known about how feedback about task outcome impacts adaptation from SPE; indeed, the literature presents an inconsistent picture of how reward impacts performance in sensorimotor adaptation tasks. For example, two recent visuomotor rotation studies using similar tasks and reward structures led to divergent conclusions: One reported that reward enhanced retention of the adapted state, but had no effect on the rate of adaptation (Galea et al., 2015), whereas the other reported a beneficial effect of rewards specifically on adaptation rate (Nikooyan and Ahmed, 2015). More recently, Leow and colleagues (Leow et al., 2018) created a situation in which task outcome was experimentally manipulated by shifting the target on-line to either intersect a rotated cursor or move away from the cursor. Task success, artificially imposed by allowing the displaced cursor to intersect the target, led to attenuated adaptation.

46 One factor that may contribute to these inconsistencies is highlighted by studies showing that, even in

47 relatively simple sensorimotor adaptation tasks, overall behavior reflects a combination of explicit and

48 implicit processes (Taylor and Ivry, 2011; Taylor et al., 2014). That is, while SPE is thought to drive

49 adaptation (Tseng et al., 2007), participants are often consciously aware of the perturbation and

50 strategically aim as one means to counteract the perturbation. It may be that reward promotes the

51 activation of such explicit processes (Bond and Taylor, 2015). Consistent with this hypothesis, Codol

52 and colleagues (Codol et al., 2017), showed that at least one of the putative effects of reward, the

53 strengthening of motor memories (Shmuelof et al., 2012), is primarily the result of re-instantiating an

54 explicit aiming strategy rather than via the direct modulation of adaptation. As explicit processes are

55 more flexible than implicit processes (Bond and Taylor, 2015), differential demands on strategies may

56 contribute toward the inconsistent effects reported across previous studies manipulating reward

57 (Holland et al., 2018).

58

59 We recently introduced a new method, referred to as clamped visual feedback, designed to isolate

60 learning from implicit adaptation (Morehead, Taylor, Parvin, & Ivry, 2017; Kim, Morehead, Parvin,

61 Moazzezi, & Ivry, 2018). During the clamp, the angular trajectory of the feedback cursor is invariant with

62 respect to the target location and thus spatially independent of hand position (Shmuelof et al., 2012;

63 Vaswani et al., 2015; Morehead et al., 2017; Kim et al., 2018; Vandevoorde and Orban de Xivry, 2018).

64 Participants are informed of the invariant nature of the visual feedback and instructed to ignore it. In this

65 way, explicit aiming should be eliminated and, thus, allow for a clean probe of implicit learning

66 (Morehead et al., 2017).

67

68 Here, we employ the clamp method to revisit how task outcome, even when divorced from actual

69 performance, influences implicit adaptation. In a series of three experiments, the clamp angle was held

70 constant and only the target size was manipulated. We assume that the clamp angle, defined with

71 respect to the centers of the target and feedback cursor, specifies the SPE. In contrast, by varying the

72   target size, we independently manipulate the information regarding task outcome, comparing conditions

73   in which the feedback cursor signals the presence or absence of a target error (TE), defined by whether

74   the cursor misses or hits the target. Given that the participants have no control over the feedback

75   cursor, the effect of this task outcome information would presumably operate in an implicit, automatic

76   manner, similar to how we assume the clamped feedback provides an invariant SPE signal.

77

78   Our experiments show that hitting the target has a strong effect on performance, attenuating the rate

79   and magnitude of learning. Through computational modeling, we explore a series of hypotheses that

80   might account for this effect. In particular, we consider models in which implicit learning is driven by

81   both SPE and TE, or where hitting the target serves as an intrinsic reward signal, one that can reinforce

82   associated movements or modulate adaptation.

83

84   **RESULTS**

85   In all experiments we used clamped visual feedback, in which the angular trajectory of a feedback

86   cursor is invariant with respect to the target location and thus spatially independent of hand position

87   (Morehead et al., 2017). The instructions (see Supplement) emphasized that the participant's behavior

88   would not influence the cursor trajectory: They were to ignore this stimulus and always aim directly for

89   the target. This method allows us to isolate implicit learning from an invariant error, eliminating potential

90   contributions from explicit aiming that might be used to reduce task performance error.

91

92   In Experiment 1, we asked if the task outcome, defined in terms of whether or not the cursor hit the

93   target, would modulate learning under conditions in which the feedback is not contingent on behavior.

94   We tested three groups of participants (n=16/group) with a 3.5° clamp for 80 cycles (8 targets per

95   cycle). The purpose of this experiment was to examine the effects of three different relationships

96   between the clamp and target while holding the visual error (defined as the center-to-center distance

97   between the cursor and target) constant (Fig. 1b): Hit Target (when the terminal position of the clamped

4

98   cursor is fully embedded within a 16 mm diameter target), Straddle Target (when roughly half of the

99   cursor falls within a 9.8 mm target, with the remaining part outside the target), Miss Target (when the

100   cursor is fully outside a 6 mm target).

101

102   Hitting the target reduced the overall change in behavior (Fig. 1d). Statistically, there was a marginal

103   difference on the rate of initial adaptation (one-way ANOVA: $F(2,45)=2.67$, $p=.08$, $\eta^2=.11$; permutation

104   test: $p=.08$; Fig. 1e) and a significant effect on late learning ($F(2,45)=4.44$, $p=.016$, $\eta^2=.17$; Fig. 1f). For

105   the latter measure, the value for the Hit Target group was approximately 35% lower than for the

106   Straddle and Miss Target groups, with post-hoc comparisons confirming the substantial differences in

107   late learning between the Hit Target and both the Straddle Target (95% CI [-16.13°, -2.34°], $t(30)=-2.73$,

108    $p=.010$, $d=.97$) and Miss Target (95% CI [-16.76°, -2.79°], $t(30)=-2.86$, $p=.008$, $d=1.01$) groups. These

109   differences were also evident in the aftereffect measure, taken from the first cycle of the no feedback

110   block (see Methods). The learning functions for the Straddle and Miss Target groups were remarkably

111   similar throughout the entire clamp block and reached similar magnitudes of late learning (95% CI [-

112   7.90°, 8.97°], $t(30)=.13$, $p=.898$, $d=.05$).

113

114   As seen in Fig. 1d, the change in hand angle from the final cycle of the clamp block to the final cycle of

115   the no feedback block was less for the Hit than the Straddle and Miss groups (one-way ANOVA:

116    $F(2,45)=4.42$, $p=.018$, $\eta^2=.16$; Hit vs Miss: 95% CI [1.47°, 8.00°], $t(30)=2.96$, $p=.006$, $d=1.05$; Hit vs

117   Straddle: 95% CI [1.06°, 8.74°], $t(30)=2.61$, $p=.014$, $d=.92$). This result indicates that retention was

118   strongest in the Hit group. However, retention is generally analyzed as a relative, rather than absolute

119   measure, especially when the amount of learning differs between groups. We thus re-analyzed the

120   change in hand angle across the no feedback block, but now as the ratio of the last no-feedback cycle

121   relative to the last clamp cycle. In this analysis, there was no difference between the three groups (Fig.

122   1g; $F(2,45)=2.06$, $p=.139$, $\eta^2=.08$; permutation test: $p=.138$).

123

124   Interestingly, the results from this experiment are qualitatively different to those observed when

125   manipulating the angular deviation of the clamp. Our previous study using clamped visual feedback

126   demonstrated that adaptation in response to errors of varying size, which was assessed by

127   manipulating the clamp angle, results in different early learning rates, but produces the same

128   magnitude of late learning (Kim et al., 2018). In contrast, the results in Experiment 1 show that the

129   hitting the target attenuates learning, with the effect becoming pronounced after prolonged exposure to

130   the perturbation. Furthermore, the effect of task outcome appears to be categorical, as it was only

131   observed for the condition in which the cursor was fully embedded within the target (Hit Target), and not

132   when the terminal position of the cursor fell partially outside the target (Straddle Target).

133
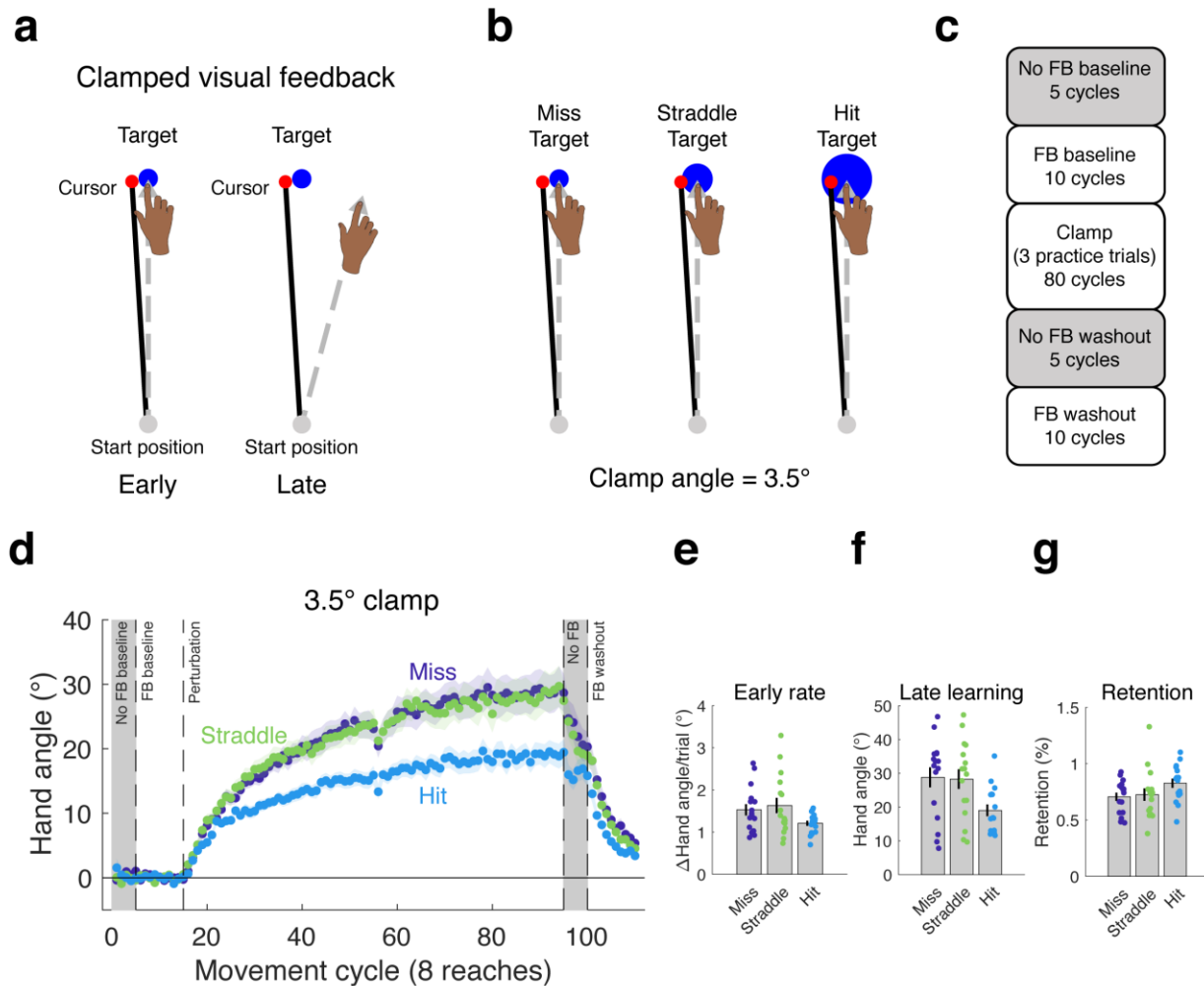
Figure 1 Hitting the target attenuates the behavioral change from clamped feedback. (a) During clamped visual feedback, the angular deviation of the cursor feedback is held constant throughout the perturbation block, and participants are fully informed of the manipulation. **(b)** The clamp angle was equal across all three conditions tested in Experiment 1, with only the target size varying between conditions. **(c)** Block design for experiment. **(d)** As in previous studies with clamped feedback, the manipulation elicits robust changes in hand angle. However, the effect was attenuated in the Hit Target condition, observed in the **(e)** rate of early adaptation, and, more dramatically, in **(f)** late learning. **(g)** The percentage of learning retained over the no feedback block following the clamp did not differ between groups. Dots represent individuals; shading and error bars denote SEM.

Source data 1 This file contains hand angle data for each trial and participant in Experiment 1, and was used to generate Figure 1d-g. Reaction times (RTs) and movement times (MTs) are also included. Note that hand angles were flipped for participants who experienced a counter-clockwise clamp.

7

150   *Experiment 2*

151   Experiment 2 was designed to extend the results of Experiment 1 in two ways: First, to verify that the

152   effect of hitting a target generalized to other contexts, we changed the size of the clamp angle. We

153   tested two groups of participants (n=16/group) with a small 1.75° clamp. For the Hit Target group (Fig.

154   2a), we used the large 16 mm target, and thus, the cursor was fully embedded. For the Straddle Target

155   group, we used the small 6 mm diameter target, resulting in an endpoint configuration in which the

156   cursor was approximately half within the target and half outside the target. We did not test a Miss

157   Target condition because having the clamped cursor land fully outside the target would have

158   necessitated an impractically small target (~1.4 mm). Moreover, the results of Experiment 1 indicate

159   that this condition is functionally equivalent to the Straddle Target group. The second methodological

160   change was made to better assess asymptotic learning. We increased the number of clamped reaches

161   to each location to 220 (reducing the number of target locations to four to keep the experiment within a

162   1.5-hour session). This resulted in a nearly three-fold increase in the number of clamped reaches per

163   location.

164

165   Consistent with the results of Experiment 1, the Hit Target group showed an attenuated learning

166   function compared to the Straddle Target group (Fig. 2b). Statistically, there was again only a marginal

167   difference in the rate of early adaptation (95% CI [-.52°/cycle, .01°/cycle], t(30)=-1.96, p=.06, d=.69; Fig.

168   2c), whereas the difference in late learning was more pronounced (95% CI [-11.38°, -1.25°], t(30)=-

169   2.54, p=.016, d=.90; permutation test: p=.007; Fig. 2d). Indeed, the 35% attenuation in asymptote for

170   the Hit Target group compared to the Straddle Target group is approximately equal to that observed in

171   Experiment 1.

172

173   We used a different approach to examine retention in Experiment 2, having participants complete 10

174   cycles with a 0° clamp following the extended 1.75° clamp block (Shmuelof et al., 2012). We opted to

175   use this alternative method since the presence of the 0° clamp would create less contextual change

176   when switching from the clamp to the retention block, compared to the no feedback block of Experiment

177   1. In terms of absolute change across the 0° clamp block, there was a trend for greater retention in the

178   Hit group compared to the Straddle group (95% CI [-.27°, 3.53°], t(30)=1.75, p=.090, d=.62). However,

179   when analyzed as a proportional change, the difference was not reliable (95% CI [-.06, .27], t(30)=1.27,

180   p=.21, d=.45).

181

182
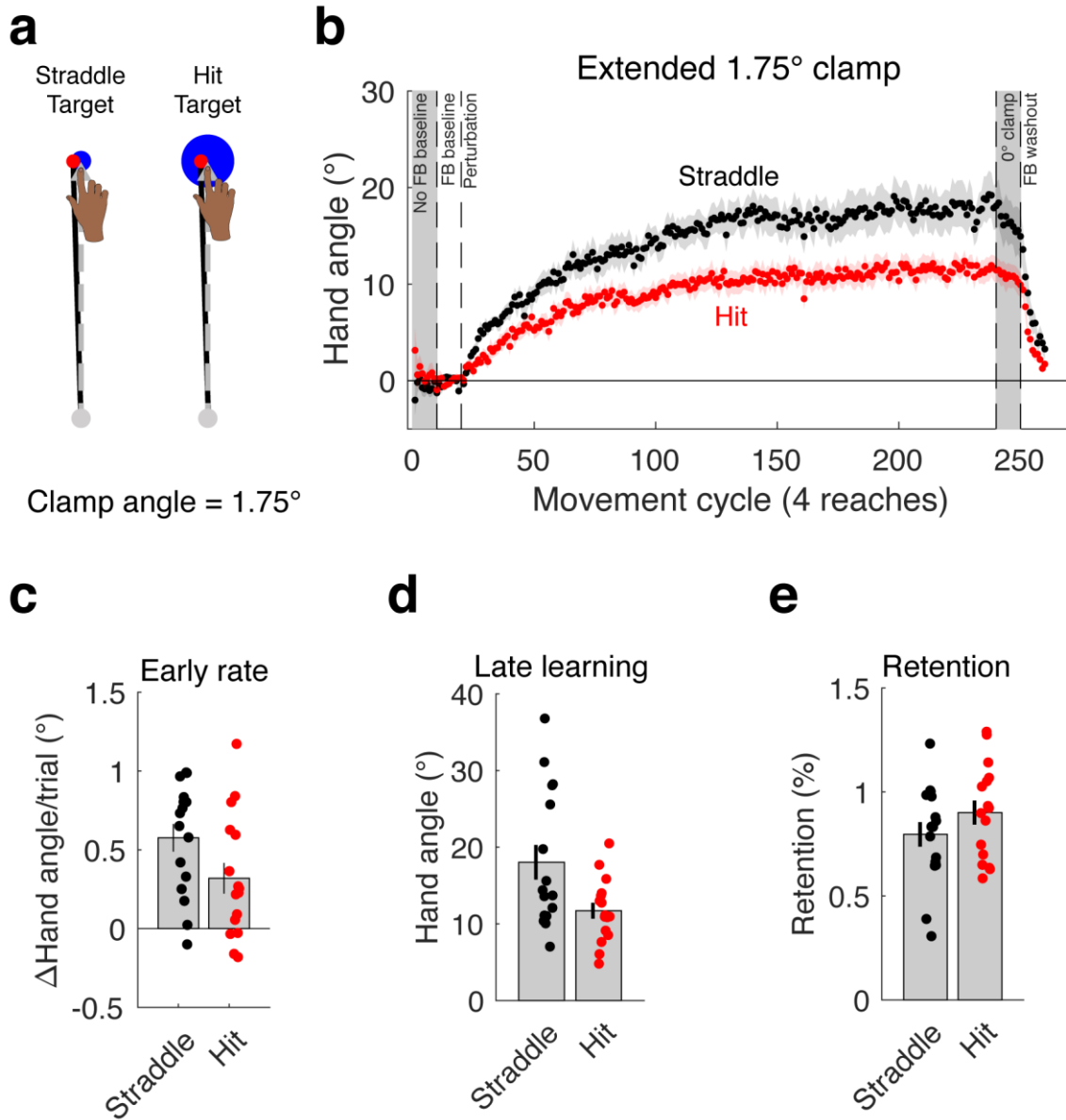
Figure 2 The attenuation of adaptation caused by hitting the target **(a)** generalizes to a different clamp angle and is stable over an extended clamp block **(b)**. As in Experiment 1, there was **(c)** a marginal difference in early adaptation rate that became **(d)** a more dramatic difference in late learning. **(e)** Again, there was no difference in the percentage of retention, this time during a 0° clamp block. Dots represent individuals; shading and error bars denote SEM.

Source data 2 This file contains hand angle data for each trial and participant in Experiment 2, and was used to generate Figure 2b-e. Reaction times (RTs) and movement times (MTs) are also included. Note that hand angles were flipped for participants who experienced a counter-clockwise clamp.

The results of these first two experiments converge in showing that learning from an invariant error is

attenuated when the cursor hits the target, relative to conditions in which at least part of the cursor falls

outside the target. This effect replicated across two experiments that used different clamp sizes.

Attenuated behavioral changes are not due to differences in motor planning

Although we hypothesized that manipulating target size in Experiments 1 and 2 would influence

learning mechanisms that respond to the differential task outcomes (i.e., hit or miss), it is also important

to consider alternative explanations for the effect of target size on learning. Figure 3 provides a

schematic of the core components of sensorimotor adaptation. The figure highlights that changes in

adaptation might arise because target size alters the inputs on which learning operates, rather than

from a change in the operation of the learning process itself. For example, increasing the target size

may increase perceptual uncertainty, creating a weaker error signal. We test this hypothesis in a control

condition in Experiment 3.

Another hypothesis centers on how variation in target size might alter motor planning. Assuming target

size influences response preparation, participants in the Hit Target groups had reduced accuracy

demands relative to the other groups, given that they were reaching to a larger target (Soechting,

1984). If the accuracy demands were reduced for these large targets, then the motor command could

be more variable, resulting in more variable sensory predictions from a forward model, and thus a

213   weaker SPE (Körding and Wolpert, 2004). While we do not have direct measures of planning noise, a

214   reasonable proxy can be obtained by examining movement variability during the unperturbed baseline

215   trials (data from clamped trials would be problematic given the induced change in behavior). If there is

216   substantially more noise in the plan for the larger target, then the variability of hand angles should be

217   higher in this group (Churchland et al., 2006). In addition, one may expect faster movement times (or

218   peak velocities) and/or reaction times for reaches to the larger target, assuming a speed-accuracy

219   tradeoff (Fitts, 1992).
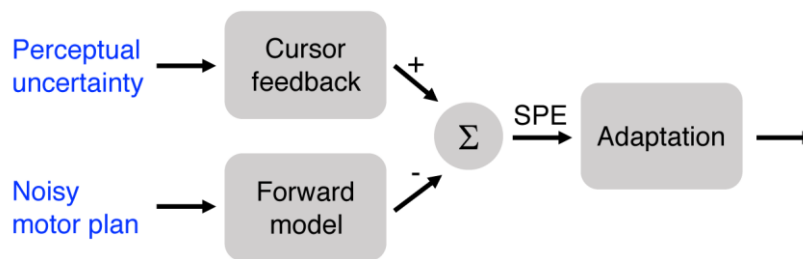
220

221   **Figure 3** Target size could affect adaptation due to increased perceptual uncertainty or greater
222   variability in motor planning. In the case of perceptual uncertainty, the feedback signal is weakened,
223   thus leading to a weaker SPE signal. In the case of noisy motor planning, the forward model prediction
224   would also be more variable and effectively weaken the SPE.
225

226   Examination of kinematic and temporal variables (see Supplement) did not support the noisy motor

227   plan hypothesis. In Experiment 1, average movement variability across the eight targets during cycles

228   2-10 of the veridical feedback baseline block were not reliably different between groups (variability:

229    $F(2,45)=2.32$, $p=.110$, $\eta^2=.093$). Movement times across groups were not different ($F(2,45)=2.19$,

230    $p=.123$, $\eta^2=.089$). However, we did observe a difference in baseline RTs ($F(2,45)=4.48$, $p=.017$,

231    $\eta^2=.166$), with post hoc t-tests confirming that the large target (Hit) group had faster RTs than the small

232   target (Miss) group (95%CI [-108ms, -16ms], $t(30)=-2.74$, $p=.010$, $d=.97$) and medium target (Straddle)

233   group (95%CI [-66ms, -10ms], $t(30)=-2.76$, $p=.010$, $d=.97$). The medium target (Straddle) and small

234   target groups' RTs were not reliably different (95% CI [-74ms, 26ms], $t(30)=-.984$, $p=.333$, $d=.348$). This

235   baseline difference in RTs was only observed in this experiment (see Supplement), and there was no

236   correlation between baseline RT and late learning for the large target group (r = .09, p = .73),

237   suggesting that RTs are not associated with the magnitude of learning.

238

239   During baseline trials with veridical feedback in Experiment 2, mean spatial variability, measured in

240   terms of hand angle, was actually lower for the group reaching to the larger target (Hit Target group:

241   3.09° ± .18°; Straddle Target group: 3.56° ± .16°; t(30)=-1.99 p=.056, d=0.70). Further supporting the

242   argument that planning was no different across conditions, neither reaction times (Hit Target: 378 ± 22

243   ms; Straddle Target: 373 ± 12 ms) nor movement times (Hit Target: 149 ± 8 ms; Straddle Target: 157 ±

244   8 ms) differed between the groups (t(30)=-0.183, p=.856, d=.06 and t(30)=0.71, p=.484, d=.25,

245   respectively).

246

247   One reason for not observing consistent effects of target size on accuracy or temporal measures could

248   be due to the constraints of the task. Studies showing an effect of target size on motor planning

249   typically utilize point-to-point movements (Soechting, 1984; Knill et al., 2011) in which accuracy

250   requires planning of both movement direction and extent. In our experiments, we utilized shooting

251   movements, thus minimizing demands on the control of movement extent. Endpoint variability is

252   generally larger for movement extent compared to movement direction (Gordon et al., 1994). It is also

253   possible that participants are near ceiling-level performance in terms of hand angle variability.

254

255   *Theoretical analysis of the effect of task outcome on implicit learning.*

256   Having ruled out a motor planning account of the differences in performance in Experiments 1 and 2,

257   we next considered different ways in which target error could affect the rate and asymptotic level of

258   learning. Adaptation from SPE can be thought of as recalibrating an internal model that learns to predict

259   the sensory outcome of a motor command (Figure 3). Here, we model adaptation with a single rate

260   state-space equation of the of the following form:

12

261 $$x(n+1) = A*x(n) + U(e) \qquad\qquad \text{[Equation 1]}$$

262 where x represents the motor output on trial n, A is a retention factor, and U represents the

263 update/correction size (or, learning rate) as a function of the error (clamp) size, e. This model is

264 mathematically equivalent to a standard single rate state-space model (Thoroughman and Shadmehr,

265 2000), with the only modification being the replacement of the error sensitivity term, B, with a correction

266 size function, U (Kim et al. 2018). Unlike standard adaptation studies where error size changes over the

267 course of learning, e is a constant with clamped visual feedback and thus, $U(e)$ can be estimated as a

268 single parameter. We refer to this model as the motor correction variant of the standard state space

269 model. The first two experiments make clear that a successful model must account for the differences

270 between hitting and missing the target, even while holding the error term in Eqn. 1 (clamp angle)

271 constant.

272

273 We consider three variants to the basic model that might account for how task outcome influences

274 learning. The first model is motivated by previous studies that have considered how reinforcement

275 processes might operate in sensorimotor adaptation tasks, and in particular, the idea that task outcome

276 information impacts a model-free operant reinforcement process (Huang et al., 2011; Shmuelof et al.,

277 2012). We can extend this idea to the clamp paradigm, considering how the manipulation of target size

278 affects reward signals: When the clamp hits the target, the feedback generates a positive reinforcement

279 signal; when the clamp misses (or straddles) the target, this reinforcement signal is absent. We refer to

280 the positive outcome as an intrinsic reward given that it is not contingent on the participant's behavior.

281 This signal could strengthen the representation of its associated movement (Castro et al., 2011;

282 Shmuelof et al., 2012).

283

284 We combine this idea with the state space model to create a Movement Reinforcement model (Fig. 4a).

285 Intuitively, this model accounts for the attenuated learning functions for the Hit conditions in

286 Experiments 1 and 2 because the effect of movement reinforcement resists the directional change in

13

287   hand angle induced by SPEs. In this model, intrinsic reward has no direct effect on SPE-driven

288   adaptation. That is, reward and error-based learning are assumed to operate independently of each

289   other, with the final movement being a composite of these two processes.

290

291   To formalize the Movement Reinforcement model, the motor output, *y*, is a weighted sum of a model-

292   free reinforcement process and an adaptation process, *x*:

293    $$y(n) = (1-V_l(n))*x(n) + V_l(n)*V_d(n) \qquad [Equation\ 2]$$

294   where a population vector (Georgopoulos et al., 1986), *V*, indicates the current bias of motor

295   representations within the reinforcement system (see Methods). The direction of this vector (V_d)

296   corresponds to the mean preferred direction resulting from the reinforcement history, with the length

297   (V_l) corresponding to the strength of this biasing signal. This vector can be viewed as a weight on the

298   movement reinforcement process (0=no weight, 1=full weight), relative to the adaptation process.

299

300   In this framework, the vector is composed of directionally-tuned units, with the strength of each unit

301   reflective of its reward history. The vector representing the weights on every unit, **r**, is updated on each

302   trial based on the task outcome:

303    $$\mathbf{r}_\theta(n+1) = A'*\mathbf{r}_\theta(n) + s \quad [Equation\ 3]$$

304    $$\mathbf{r}_{\sim\theta}(n+1) = A'*\mathbf{r}_{\sim\theta}(n) \qquad [Equation\ 4]$$

305   where θ indexes the unit corresponding to the direction of the movement, *y(n)*, on hit trials, and $\sim\theta$

306   indexes all of the other units on hit trials and all units on miss trials. In this simplified reward scheme,

307   the weight to the unit corresponding to the rewarded movement direction is increased by magnitude *s*

308   on a trial-by-trial basis, and all weights are decremented due to a retention factor, *A'*, on every trial. The

309   latter ensures that these reward-dependent weights revert back to zero in the absence of reward.

310

14

311   In sum, the Movement Reinforcement model entails four parameters, composed of separate update

312   and retention parameters for the reinforcement learning process and the adaptation process. The

313   former is model-free, dependent on an operant conditioning process by which a task outcome signal

314   modifies movement biases, whereas the latter is model-based, using SPE to recalibrate an internal

315   model of the sensorimotor map.

316

317   The second model entails a single process whereby the task outcome directly modulates the adaptation

318   process. For example, an intrinsic reward signal associated with hitting the target could modulate

319   adaptation, attenuating the trial-to-trial change induced by the SPE (Fig. 4b). In this Adaptation

320   Modulation model, the reward signal can be interpreted as a gain controller, similar to previous efforts

321   to model the effect of explicit rewards and punishments on adaptation (Galea et al., 2015). In

322   Experiments 1 and 2, hitting the target presumably reduces the gain on adaptation, thus leading to

323   attenuated learning.

324

325   We formalize the Adaptation Modulation model as follows:

326    $$x(n+1) = \gamma_A * A * x(n) + \gamma_u * U(e) \quad \text{[Equation 5]}$$

327   where γ_A and γ_u are gains on the retention and update parameters, respectively. In the current

328   implementation, we set γ_A and γ_u to 1 on miss trials and estimate the values of γ_A and γ_u for the hit trials.

329   Although this could be reversed (e.g., set gains to 1 on hit trials and estimate values on miss trials), our

330   convention seems more consistent with previous modeling studies of adaption where the movements

331   generally miss the target. We impose no additional constraint on the gain parameters; the effect of

332   retention or updating can be larger or smaller on hit trials compared to miss trials. As with the

333   Movement Reinforcement model, the Adaptation Modulation model has four free parameters.

334

335 The third model we consider here, the Dual Error model, postulates that learning is the composite of

336 two model-based learning processes that operate on different error signals. The first is an adaptation

337 process driven by SPE (as in Eqn. 1). The second process operates in the same manner as adaptation,

338 but here the error signal is sensitive to the task outcome. This idea of a TE-sensitive process stems

339 from previous studies in which an error is produced, not by perturbing the visual feedback of hand

340 position, but rather by displacing the visual feedback of the target position (Magescas and Prablanc,

341 2006; Cameron et al., 2010, 2010; Schmitz et al., 2010). The resulting mismatch between the hand

342 position and displaced target position can be viewed as a TE rather than SPE, under the assumption

343 that the veridical feedback of hand position roughly matches the predicted hand position (see

344 Discussion). When this error signal is consistent (e.g., target is displaced in the same direction on every

345 trial), a gradual change in heading angle is observed, similar to that seen in studies of visuomotor

346 adaptation. Moreover, this form of learning is implicit: By shifting the target position during a saccade,

347 just prior to the reach, the participants are unaware of the target displacement.

348

349 In the Dual Error model, the motor output is the sum of two processes:

350 $$x_{total}(n) = x_{spe}(n) + x_{te}(n), \qquad \text{[Equation 6]}$$

351 where

352 $$x_{spe}(n+1) = A_{spe}*x_{spe}(n) + U_{spe}*(SPE) \quad \text{[Equation 7]}$$

353 $$x_{te}(n+1) = A_{te}*x_{te}(n) + U_{te}*(TE), \qquad \text{[Equation 8]}$$

354 Equation 7 is the same as in the other two models, describing adaptation from a sensory prediction

355 error, but with the notation modified here to explicitly contrast with the second process. Eqn. 8

356 describes a second model-based process, but one that is driven by the target error.

357

358 The SPE-sensitive process updates from the error term on every trial given that the SPE is always

359 present, even on hit trials. In contrast, the TE-sensitive process only updates from the error term on

360 miss trials. The error component of Eqn. 8 is absent on hit trials. This would account for the attenuated

16

361 learning observed in the large target (Hit) conditions in Experiments 1 and 2. In the context of our clamp

362 experiments, TE is modeled as a step function (Fig. 4c), set to 0 when the cursor hits the target and 1

363 when the cursor misses or straddles the target. However, if the cursor position varied (as in studies with

364 contingent feedback), TE might take on continuous, signed values, similar to SPE.

365

366 We note that the Dual Error model is similar to the influential two-process state space model of

367 adaption introduced by Smith and colleagues (Smith et al., 2006). In their model, dual-adaptation

368 processes have different learning rates and retention factors, resulting in changes that occur over

369 different time scales. Here the different learning rates and retention factors are related to the different

370 error signals, TE and SPE. Whereas the dual-rate model imposes a constraint on the parameters (i.e.,

371 process with faster learning must also have faster forgetting), the four parameters in the Dual Error

372 model are unconstrained relative to each other.

373

374
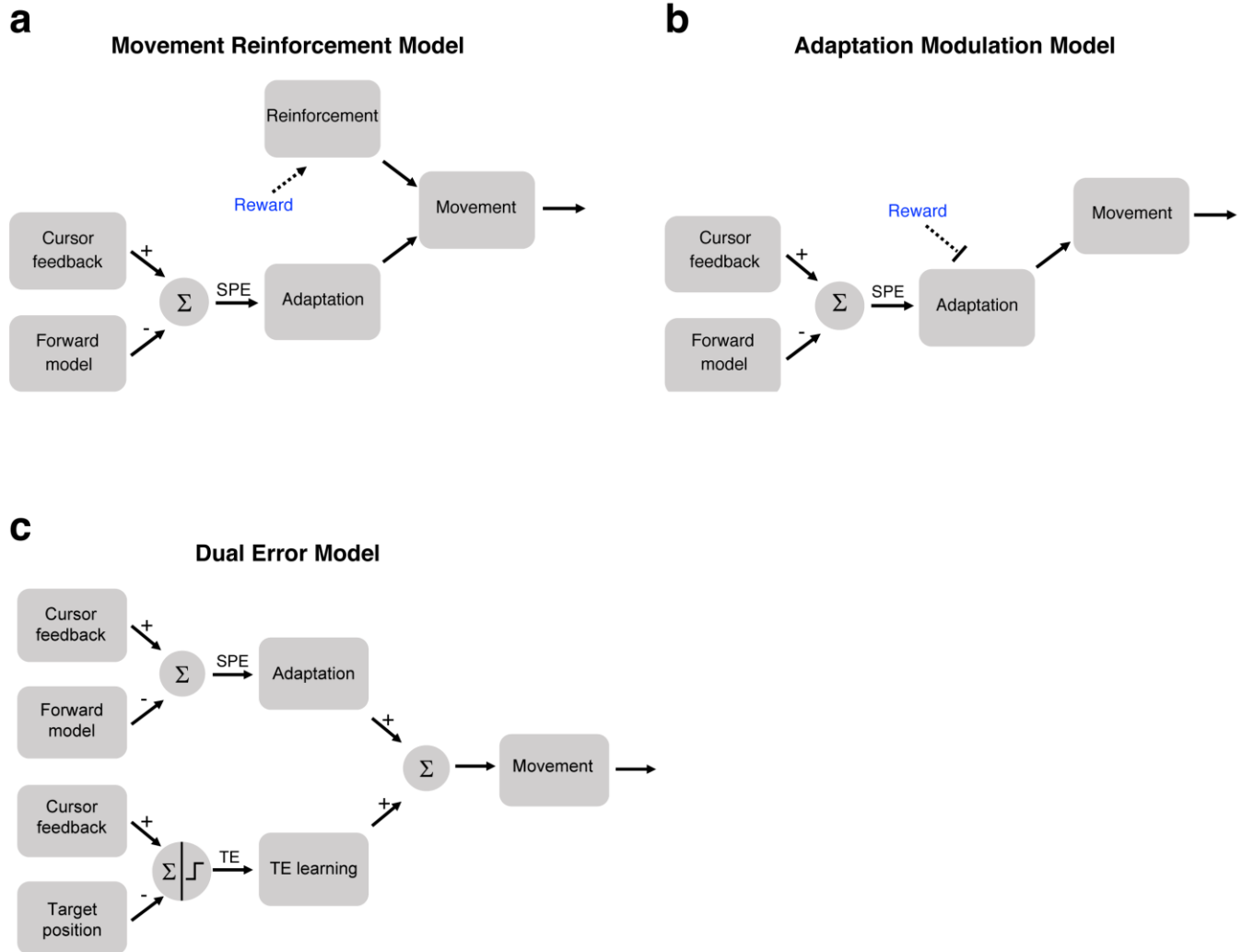
Figure 4 Three models of how intrinsic reward or target error could affect learning. (a) In the Movement Reinforcement model, reward signals cause reinforcement learning processes to bias future movements towards previously rewarded movements. The adaptation process is sensitive only to SPE and not reward. The overall movement reflects a composite of the two processes. **(b)** In the Adaptation Modulation model, reward directly attenuates adaptation to SPE. **(c)** In the Dual Error model, a second, independent model-based process, one driven by TE, combines with SPE-based adaptation to modify performance.

Experiment 3

The experimental design employed in Experiments 1 and 2 cannot distinguish between these three models because all make qualitatively similar predictions. In the Movement Reinforcement model, the attenuated asymptote in response to Hit conditions arises because movements are rewarded throughout, including during early learning, biasing future movements towards baseline. The Adaptation

18

388   Modulation model predicts a lower asymptote during the Hit condition because the adaptation system is

389   directly attenuated by reward. The Dual Error model similarly predicts a lower asymptote because only

390   one of two learning processes is active when there is no target error.

391

392   In contrast to the single perturbation blocks used in Experiments 1 and 2, a transfer design in which the

393   target size changes after an initial adaptation phase affords an opportunity to contrast the three models.

394   In Experiment 3, we tested two groups of participants (n=12/group) with a 1.75° clamp, varying the

395   target size between the first and second halves of the experiment (Fig. 5a). The key manipulation

396   centered on the order of when the target was large (hit condition) or small (straddle condition).

397

398   For the Straddle-to-Hit group, a small target was used in an initial acquisition phase (first 120 clamp

399   cycles). Based on the results of Experiments 1 and 2, we expect to observe a relatively large change in

400   hand angle at the end of this phase since the outcome is always an effective "miss". The key test

401   comes during the transfer phase (final 80 clamp cycles), in which the target size is increased such that

402   the invariant clamp now results in a target hit. For the Movement Reinforcement model, hitting the

403   target will produce an intrinsic reward signal, reinforcing the associated movement. Therefore, there

404   should be no change in performance (hand angle) following transfer since the SPE remains the same

405   and the current movements are now reinforced (Fig. 5b). In contrast, both the Adaptation Modulation

406   and Dual Error models predict that, following transfer to the large target, there will be a drop in hand

407   angle, relative to the initial asymptote. For the former, hitting the target will attenuate the adaptation

408   system; for the latter, hitting the target will shut down learning from the process that is sensitive to

409   target error.

410

411   We also tested a second group in which the large target (hit) was used in the acquisition phase and the

412   small target (effective "miss") in the transfer phase (Hit-to-Straddle group). All three models make the

413   same qualitative predictions for this group. At the end of the acquisition phase, there should be a

414   smaller change in hand angle compared to the Straddle-to-Hit group, due to the persistent target hits.

415   Following transfer, all three models predict an increase in hand angle, relative to the initial asymptote.

416   For the Movement Reinforcement model, the reduction in target size removes the intrinsic reward

417   signal, which over time, lessens the contribution of the reinforcement process as the learned movement

418   biases decay in strength. The Adaptation Modulation model predicts that hand angle will increase due

419   to the removal of the attenuating effect on adaptation following transfer. The Dual Error model also

420   predicts an increase in hand angle, but here the effect occurs because the introduction of a target error

421   activates the second implicit learning process. Although the Hit-to-Straddle group does not provide a

422   discriminative test between the three models, the inclusion of this group does provide a second test of

423   each model, as well as an opportunity to rule out alternative hypotheses for the behavioral effects at

424   transfer. For example, the absence of a change at transfer might be due to reduced sensitivity to the

425   clamp following a long initial acquisition phase.

426

427   *Experiment 3 – Behavioral Analyses*

428   For our analyses, we first examined performance during the acquisition phase. Consistent with the

429   results from Experiments 1 and 2, the Hit-to-Straddle Target group adapted slower than the Straddle-to-

430   Hit group (95% CI [-.17°/cycle, -.83°/cycle], $t(22)=-3.15$, $p=.005$, $d=1.29$; Fig. 5c) and reached a lower

431   asymptote (95% CI [-5.25°, -15.29°], $t(22)=-4.24$, $p=.0003$, $d=1.73$; permutation test: $p=.0003$; Fig. 5d).

432   The reduction at asymptote was approximately 45%.

433

434   We next examined performance during the transfer phase where the target size reversed for the two

435   groups. Our primary measure of behavioral change for each subject was the difference in late learning

436   (average hand angle over last 10 cycles) between the end of the acquisition phase and the end of the

437   transfer phase. As seen in Fig. 5d, the two groups showed opposite changes in behavior in the transfer

438   phase, evident by the strong (group x phase) interaction ($F(2,33)=43.1$, $p<10^{-7}$, partial $\eta^2=.72$). The

20

439   results of a within-subjects t-test showed that the Hit-to-Straddle group showed a marked increase in

440   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.

441   5e), consistent with the predictions for all three models.

442

443   The Straddle-to-Hit group's transfer performance provides an opportunity to compare differential

444   predictions, and in particular, to pit the Movement reinforcement model against the other two models.

445   Following the switch to the large target, there was a decrease in hand angle. Applying the same

446   statistical test, the mean decrement in hand angle was 5.7° from the final cycles of the training phase to

447   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

448   result is consistent with the prediction of the Adaptation Modulation and Dual Error models. In contrast,

449   the reduction in hand angle cannot be accounted for by the Movement Reinforcement model.
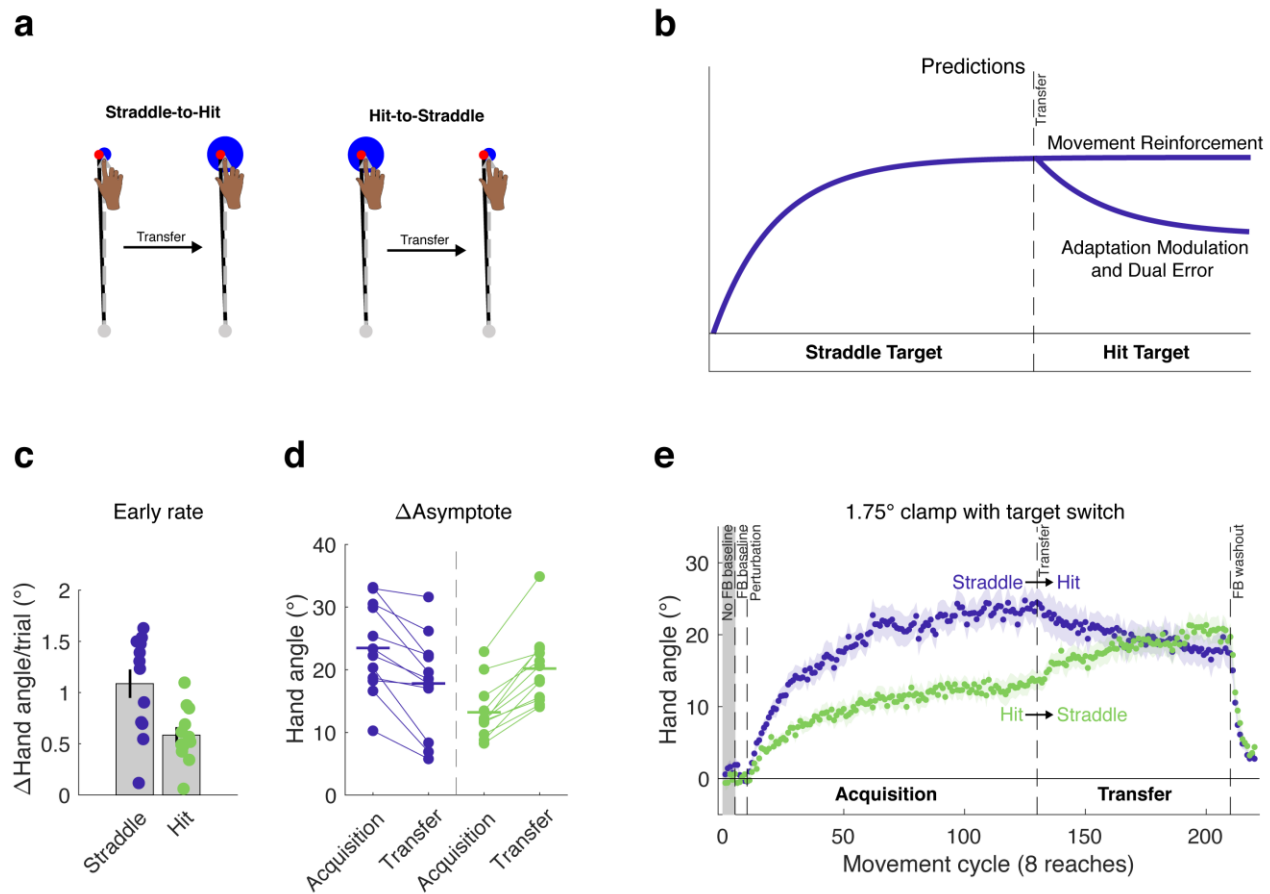
Figure 5 Within-subject transfer design to evaluate models of the impact of task outcome on implicit motor learning. (a) Using a transfer design, **(b)** the models diverge in their behavioral predictions for the Straddle-to-Hit group following transfer. The Movement Reinforcement model predicts a persistent asymptote following transfer, whereas the Adaptation Modulation and Dual Error models predict a decay in hand angle. During the acquisition phase, we again observed differences between the Hit and Straddle groups in the **(c)** early adaptation rate as well as **(d)** late learning**.** All participants in both groups demonstrated changes in reach angle consistent with the Adaptation Modulation and Dual Error models. **(e)** The learning functions were inconsistent with the Movement Reinforcement model. Note that the rise in hand angle for the Hit-to-Straddle group is consistent with all three models. Dots represent individuals; shading and error bars denote SEM.

Source data 3 This file contains hand angle data for each trial and participant in Experiment 3, and was used to generate Figure 5c-e and Figure 5-figure supplement 1. Reaction times (RTs) and movement times (MTs) are also included. Note that hand angles were flipped for participants who experienced a counter-clockwise clamp.

469   *Experiment 3 -- Modeling Results*

470   We evaluated the three models by simultaneously fitting group-averaged data for both groups. As

471   depicted in Fig. 6, all three models capture the initial plateau followed by increased learning of the Hit-

472   to-Straddle group. However, the quality of the fits diverges for the Straddle-to-Hit group, where the

473   Movement Reinforcement model cannot produce a decrease in hand angle once the large target is

474   introduced. Instead, the best-fit parameters for this model result in an asymptote that falls between the

475   hand angle values observed during the latter part of each phase. In contrast, the Adaptation Modulation

476   and Dual Error models both predict the drop in hand angle during the second phase of the experiment

477   for the Straddle-to-Hit group.

478

479   Consistent with the preceding qualitative observations, the Movement Reinforcement model yielded a

480   lower R^2 value and higher Akaike Information Criterion (AIC) score (higher AIC indicates relatively

481   worse fit) than the Adaptation Modulation and Dual Error models (Table 1). A comparison of the latter

482   two shows that the Dual Error model provides the best account of the results. This model yielded a

483   lower AIC score and accounted for 90% of the variance in the group-averaged data compared to 86%

484   for the Adaptation Modulation model.

a

b

c
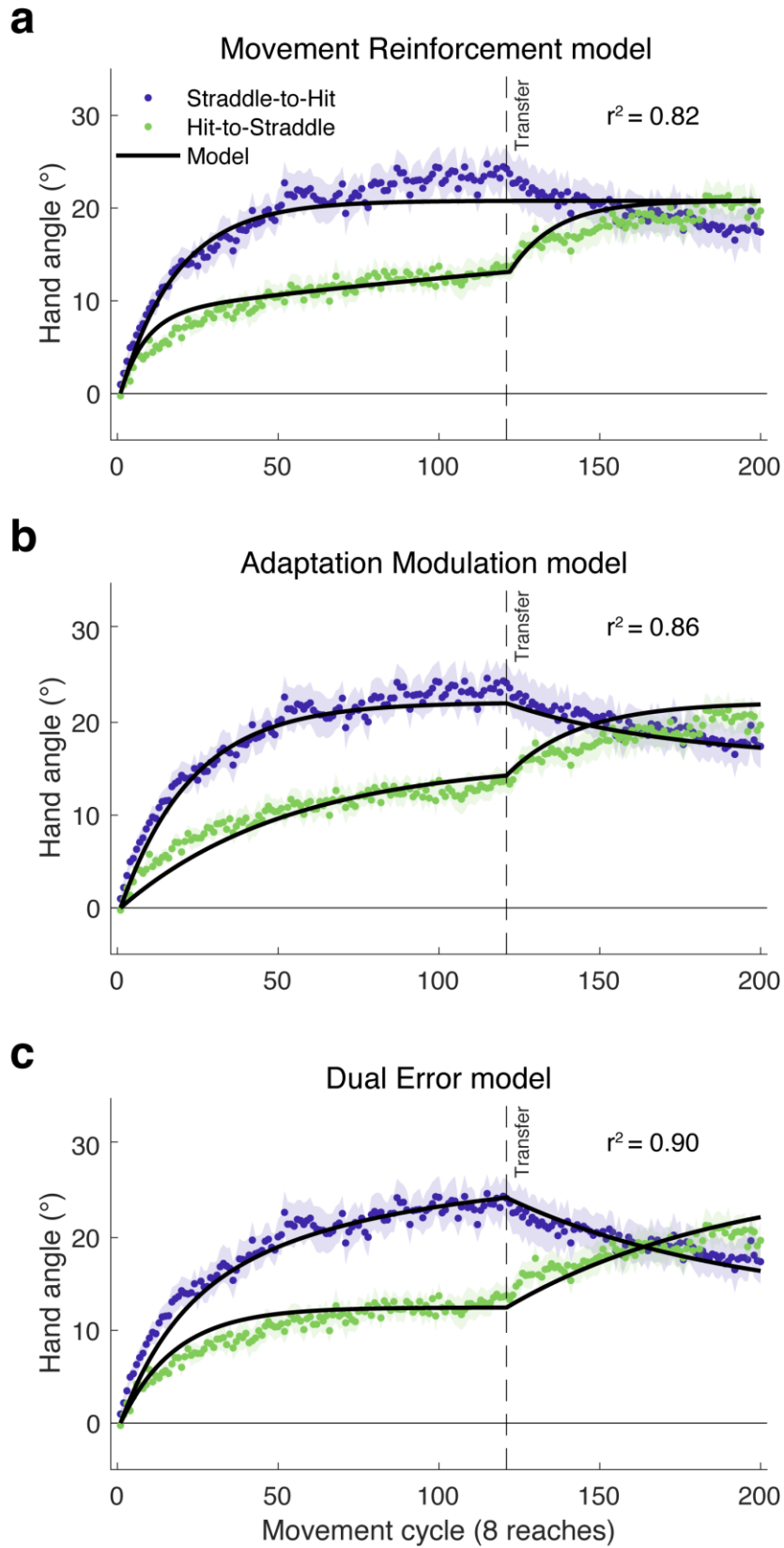
486 **Figure 6 Model fits of the learning functions from Experiment 3.** The failure of the **(a)** Movement
487 Reinforcement model to qualitatively capture the decay in hand angle following transfer in the Straddle-
488 to-Hit condition argues against the idea that the effect of task outcome arises solely from a model-free
489 learning process that operates independent of model-based adaptation. In contrast, both the **(b)**
490 Adaptation Modulation and **(c)** Dual Error models accurately predict the changes in hand angle
491 following transfer in both the Hit-to-Straddle and Straddle-to-Hit conditions.
492

493 To understand whether the effects of target size were due to changes in learning or retention, we

494 examined the parameter estimates for the Adaptation Modulation and Dual Error models. We first

495 generated 1000 bootstrapped samples of group-averaged behavior by resampling with replacement

496 from each group. We then fit each of the bootstrapped samples simultaneously and report the results

497 here in terms of 95% confidence intervals. The parameter estimates indicate that the learning rate

498 parameter was very sensitive to the two task outcome conditions. For the Adaptation Modulation model,

499 the estimates of γ_u*U were larger during miss than hit conditions, with no overlap of the confidence

500 intervals ([.693, 1.302] vs [.182, .573], respectively); thus, the error-driven adjustment in the state of the

501 internal model was much larger after a miss than a hit. For the Dual Error model, the estimates of U_{spe}

502 were larger than for U_{te}, again with no overlap of the confidence intervals [(.414, 1.08), vs [.157, .398]),

503 indicating that the state change was more strongly driven by SPE than TE. For each model, the

504 process that produced a larger error-based update in the internal model also had the lower retention

505 factor, although here there was overlap in the 95% confidence intervals for the latter (γ_r*A for Miss:

506 [.939,.969] vs Hit: [.961,.989]; A_{spe}: [.900,.972] vs A_{te}: [.938,.993]). In sum, the impact of task outcome

507 (hit or miss) was primarily manifest in the estimates of the learning rate parameters.

508

509 The behavioral pattern observed in Experiment 3, complemented by the modeling results, are

510 problematic for the Movement Reinforcement model, challenging the idea that the effect of task

511 outcome arises solely from a model-free learning process that operates independent of model-based

512 adaptation. However, this does not exclude the possibility that task outcome information influences both

513 model-free and model-based processes. For example, hitting the target might not only reinforce an

514   executed movement, but might also modulate adaptation. Formally, this hypothesis would correspond

515   to a hybrid model that combines the Adaptation Modulation and Movement Reinforcement models.

516   Indeed, hybrids that combine the Movement Reinforcement model with either the Adaptation

517   Modulation or Dual Error models (see Methods) yield improved model fits and lower AIC values, with

518   the two hybrids producing comparable values (see Table 1).

519

520   **Table 1: Model evaluations.**

Basic Models	# of free parameters	R-squared	AIC
Movement Reinforcement	4	.824	363
Adaptation Modulation	4	.861	269
Dual Error	4	.895	156
Hybrid Models			
Movement Reinforcement + Adaptation Modulation	6	.945	-100
Movement Reinforcement + Dual Error	6	.945	-97

521

522

523   *Control group for testing perceptual uncertainty hypothesis*

524   Across the three experiments, the amount of learning induced by clamped visual feedback was

525   attenuated when participants reached to the large target. We considered if this effect could be due, in

526   part, to the differences between the Hit and Straddle/Miss conditions in terms of perceptual uncertainty.

527   For example, the reliability of the visual error signal might be weaker if the cursor is fully embedded

528   within the target; in the extreme, failure to detect the angular offset might lead to the absence of an

529   error signal on some percentage of the trials.

530

531   To evaluate this perceptual uncertainty hypothesis, we tested an additional group in Experiment 3 with

532   a large target, but modified the display such that a bright line, aligned with the target direction, bisected

533   the target (Figure 5-figure supplement 1). With this display, the feedback cursor remained fully

534   embedded in the target, but was clearly off-center. If the attenuation associated with the large target is

26

535   due to perceptual uncertainty, then the inclusion of the bisecting line should produce an adaptation

536   effect similar to that observed with small targets. Alternatively, if perceptual uncertainty does not play a

537   prominent role in the target size effect, then the adaptation effects would be similar to that observed

538   with large targets.

539

540   Consistent with the second hypothesis, performance during the acquisition phase for the group

541   reaching to a bisected target was similar to that of the group reaching to the standard large target (Hit-

542   to-Straddle, see Supplement). To provide support for this observation, we first performed an omnibus

543   one-way ANOVA on the late learning data at the end of the acquisition phase, given our analysis plan

544   entailed multiple planned pair-wise comparisons. There was a significant effect of group ($F_{(2,33)}=9.33$,

545    $p=.0006$, $\eta^2=.36$). Subsequent planned pair-wise comparisons showed no significant differences

546   between the bisected target and standard large target (Hit-to-Straddle) groups (early adapt: 95% CI [-

547   .34°/cycle, .22°/cycle], $t(22)=-.47$; $p=.64$,; $d=.19$; late learning: 95% CI [-7.80° 1.19°], $t(22)=-1.52$; $p=.14$;

548    $d=.62$). In contrast, the group reaching to bisected targets showed slower early adaptation rates (95%

549   CI [-.81°/cycle, -.07°/cycle], $t(22)=-2.49$, $p=.02$, $d=1.02$) and lower magnitudes of late learning (95% CI

550   [-12.58°, -1.35°], $t=-2.57$, $p=0.017$, $d=1.05$) when compared with the group reaching to small targets

551   (Straddle-to-Hit).

552

553   During the transfer phase, the target size for the perceptual uncertainty group remained large, but the

554   bisection line was removed. If perceptual uncertainty underlies the effect we have attributed to hitting

555   the target, we would expect to observe a decrease in hand angle following transfer, since uncertainty

556   would increase. However, following transfer to the non-bisected large target, there was no change in

557   asymptote (95% CI [-.87°, 2.32°], $t(11)=1.0$, $p=.341$, $d_z=.29$). In sum, the results from this control group

558   indicate that the attenuated adaptation observed when the cursor is fully embedded within the target is

559   not due to perceptual uncertainty,

27

560

DISCUSSION

562 Models of sensorimotor adaptation have emphasized that this form of learning is driven by sensory

563 prediction errors, the difference between the observed and predicted sensory consequences of a

564 movement. In this formulation, task outcome, defined as hitting or missing the target, is not part of the

565 equation (although in most adaptation tasks, the sensory prediction is at the target, thus conflating SPE

566 and task outcome). While a number of recent studies have demonstrated that task outcome signals can

567 influence overall performance in these tasks (Galea et al., 2015; Reichenthal et al., 2016; Leow et al.,

568 2018), it is unclear whether these reinforcement signals impact sensorimotor adaptation (Shmuelof et

569 al., 2012; Galea et al., 2015), or whether they are exploited by other learning systems, distinct from

570 SPE-driven implicit adaptation (Codol et al., 2018; Holland et al., 2018).

571

572 The interpretation of the results from these studies is complicated by the fact that the experimental

573 tasks may conflate different learning processes. In the present study, we sought to avoid this

574 complication by employing a new method to study implicit learning, one in which participants are

575 specifically instructed to ignore an invariant visual error signal, thus eliminating explicit processes

576 (Morehead et al., 2017). Using this clamp method, we observed a striking difference between

577 conditions in which the final position of the cursor was fully embedded in the target compared to

578 conditions in which the cursor either terminated outside or straddled the target: When the cursor was

579 fully embedded, the rate of learning was reduced and the asymptotic level of learning was markedly

580 attenuated.

581

582 *Characterizing the Information Associated with Task Outcome*

583 We manipulated task outcome by varying the size of the target, and, across experiments, manipulated

584 SPE by varying the clamp size. Although the experimental instructions remained unchanged, these

585 stimulus changes might be expected to also influence the perception of the error or motor planning

28

586   processes. However, the behavioral differences arising from the manipulation of task outcome did not

587   appear to arise from these factors. Movement kinematics were essentially the same when reaching to

588   the different sized targets, and the perceptual control condition showed that reducing perceptual

589   uncertainty did not influence performance. Moreover, the finding in Experiment 1 that the Straddle

590   group performed similar to the Miss group, suggests that the effect of target size is, to some degree,

591   categorical rather than continuous.

592

593   With clamped visual feedback, participants have no control over the invariant task outcome. In our

594   earlier work with this method, we hypothesized that the cursor feedback is interpreted by the adaptation

595   system as an error signal. We assume the adaptation system is "fooled" by the temporal correlation

596   between the motion of the hand and feedback signal, even though the participants are fully aware that

597   the angular position of the cursor is causally unrelated to their behavior (Morehead et al., 2017). This

598   hypothesis is consistent with earlier work showing that SPEs will drive implicit adaptation, even at the

599   cost of reduced task success (Mazzoni and Krakauer, 2006; Taylor and Ivry, 2011).

600

601   One interpretation of the effect of task outcome is that an automatic signal is generated when the

602   cursor hits the target; that is, this outcome is intrinsically rewarding (Huang et al., 2011; Leow et al.,

603   2018), even though the participant is aware that the outcome does not depend on the accuracy of their

604   movements. In two of our proposed models, we assume that hitting the target leads to the automatic

605   generation of a positive reinforcement signal. In the Movement Reinforcement model, this signal

606   strengthens associated movement representations, producing a bias on behavior. In the Adaptation

607   Modulation model, this signal directly attenuates adaptation. Alternatively, one could emphasize the

608   other side of the coin, namely, that the absence of reward (i.e., missing the target) results in a negative

609   reinforcement signal, or what we refer to here as target error. Consideration of two types of error

610   signals is, of course, central to the Dual Error model. We could also reframe the Adaptation Modulation

611   model: Rather than view adaptation as being attenuated following a positive task outcome, it may be

612   that adaptation is enhanced following a negative task outcome.

613

614   With the current procedure, we do not have evidence, independent of the behavior, that the task

615   outcome with non-contingent feedback results in a reinforcement signal (either positive or negative).

616   Methods such as fMRI (Daw et al., 2011) or pupillometry (Manohar et al., 2017) could provide an

617   independent means to assess the presence of well-established signatures of reward. Nonetheless, our

618   results indicate, more generally, that task outcome is an important factor mediating the rate and

619   magnitude of implicit motor learning.

620

621   *Modeling the Influence of Task Outcome on Implicit Changes in Performance*

622   Our modeling analysis makes clear that parallel, independent activity of purely SPE-driven adaptation

623   and task outcome-driven operant reinforcement processes cannot account for the behavioral changes

624   observed in the present set of experiments. In particular, the Movement Reinforcement Model fails to

625   predict the change in reach direction observed when the target size was decreased in the Straddle-to-

626   Hit condition of Experiment 3. In this model, the Straddle-to-Hit group's asymptotic learning during the

627   acquisition phase is due to the isolated operation of the adaptation system, given that none of the

628   reaches are rewarded. The SPE signal would be expected to persist following transfer, maintaining this

629   asymptote. Moreover, movements in this direction would be further strengthened given that, with the

630   introduction of the large target, they would be reinforced by an intrinsic reward signal. Importantly, the

631   predicted absence of behavioral change following transfer should hold for all models in which a model-

632   free reinforcement-based process is combined with a task outcome-insensitive model-based adaptation

633   process. For example, the prediction is independent of whether the reinforcement process follows a

634   different time course than adaptation (e.g., faster or slower), or if we model the effect of reinforcement

635   as basis functions (Donchin et al., 2003; Tanaka et al., 2012; Taylor et al., 2013) rather than discrete

636   units. Thus, we propose that any model in which adaptation and reinforcement processes act

637   independently will fail to show the observed decrease in hand angle following transfer from a miss

638   condition to a hit condition.

639

640   The failure of the Movement Reinforcement model requires that we consider alternatives in which

641   information about the task outcome interacts with model-based processes. The Adaptation Modulation

642   model postulates that a signal associated with the task outcome directly modulates the adaptation

643   process. In the current instantiation, we propose that hitting the target results in an intrinsic reward

644   signal that reduces the gain on adaptation (Leow et al., 2018), although an alternative interpretation

645   would be that missing the target results in an error signal that amplifies the gain. This model was able

646   to account for the reduced asymptote observed in the Hit-to-Straddle condition of Experiment 3,

647   outperforming the Movement Reinforcement model.

648

649   The Adaptation Modulation model makes explicit assumptions of previous work in which reward was

650   proposed to act as a gain controller on the adaptation process (Galea et al., 2015; Nikooyan and

651   Ahmed, 2015). In terms of the standard state space model, the results indicate that the main effect of

652   task outcome was on the learning rate parameter. Hitting the target reduced the learning rate by

653   approximately 40%, consistent with other studies showing reduced behavioral changes when hitting the

654   target (Reichenthal et al., 2016; Leow et al., 2018).

655

656   Galea et al. (2015) also used a model-based approach to examine the influence of reinforcement on

657   adaptation, comparing conditions in which participants received or lost money during a standard

658   visuomotor rotation task. Their results indicated that reward had a selective effect on the retention

659   parameter in the state space model, suggesting the effect was on memory rather than learning. We

660   also observed higher retention parameters when the cursor hit the target, although the effect size here

661   was a relatively smaller ~3% increase and not reliably different from the miss/straddle condition, based

662   on bootstrapped parameter estimates. We suspect that the effect on retention in Galea et al. (2015)

663   was, in large part, not due to a change in the adaptation process itself, but rather the residual effects of

664   an aiming strategy induced by the reward. That is, the monetary rewards might have reinforced a

665   strategy during the rotation block, and this carried over into the washout block. Indeed, the idea that

666   reward impacts strategic processes has been advanced in studies comparing conditions in which the

667   performance could be enhanced by re-aiming (Codol et al., 2018; Holland et al., 2018). By using non-

668   contingent clamped feedback, we eliminate strategy use and thus provide a purer assessment of how

669   reward influences adaptation.

670

671   We recognize that the hypothesized modulation of sensorimotor adaptation by task outcome is, at least

672   superficially, contrary to previous conjectures concerning the independent effects of SPE and TE

673   (Mazzoni and Krakauer, 2006; Taylor and Ivry, 2011; Taylor et al., 2014; Morehead et al., 2017; Kim et

674   al., 2018). One argument for independence comes from a visuomotor adaptation task in which

675   participants are instructed to use an aiming strategy to compensate for a large visuomotor rotation

676   (Mazzoni and Krakauer, 2006; Taylor and Ivry, 2011). By using the instructed strategy, the cursor

677   immediately intersects the target, eliminating the target error. However, over the course of subsequent

678   reaches, the participants' performance deteriorates, an effect attributed to the persistence of an SPE,

679   the difference between the aiming location and cursor position. Taylor and Ivry (2011) modeled this

680   behavior by assuming the operation of two independent learning processes, adaptation driven by SPE

681   and strategy adjustment driven by TE. In light of the present results, it is important to note that there

682   were actually very few trials in which target hits actually occurred, given that the large SPE on the initial

683   reaches resulted in target misses on almost all trials. In addition, the strength of a task success signal

684   may fall off with larger SPEs (Cashaback et al., 2017). As such, the current study, in which SPE and

685   task outcome are held constant throughout learning, provides a much stronger assessment on the

686   effect of task outcome on sensorimotor adaptation.

687

688   The Dual Error model suggests an alternative, model-based account of the effect of task outcome on

689   performance. This model assumes that performance is the composite of two independent error-based

690   processes, an adaptation system that is sensitive to SPE, and a second implicit process that is

691   sensitive to target error. Of the three models tested here, the Dual Error model provided the best

692   account of the behavior in Experiment 3, accounting for 90% of the variance when the group-averaged

693   data from both the Straddle-to-Hit and Hit-to-Straddle conditions of Experiment 3 were fit

694   simultaneously.

695

696   Interestingly, in previous work, TE was thought to be a driving signal for explicit learning, and in

697   particular, for adjusting a strategic aiming process that can lead to rapid improvements in performance

698   (Taylor and Ivry, 2011; Taylor et al., 2014; McDougle et al., 2015; Day et al., 2016). Conceptualizing

699   TE-based learning as supporting an explicit process does not appear warranted here. We have no

700   evidence, either based on performance or verbal reports obtained during post-experiment debriefing

701   sessions (Kim et al., 2018), that participants employ a strategy to counteract the clamp. Rather, all of

702   the observed changes in behavior are implicit.

703

704   Alternatively, we can consider whether the TE-based process constitutes a form of implicit aiming. The

705   notion of implicit aiming has previously been suggested in work showing that, with extended practice,

706   strategic aiming may become automatized (Huberdeau et al., 2017). One interpretation of this effect is

707   that aiming strategies eventually become "cached" and are automatically retrieved during response

708   preparation (Haith and Krakauer, 2018). While the idea of a cached strategy may be reasonable in the

709   context of traditional sensorimotor perturbation studies, it does not seem to offer a reasonable

710   psychological account of the effect of task outcome in the current context. Given that participants do not

711   employ a strategy to counteract the clamp, there is no strategy to cache. Furthermore, parameter

712   estimates for the Dual Error model indicate that the TE-sensitive process learned at a slower rate and

713   retained more than the SPE-sensitive process. Were implicit aiming to share core features of explicit

33

714   aiming, the modeling results would be inconsistent with previous work indicating that explicit aiming

715   from TE is faster (McDougle et al., 2015) and more flexible (Bond and Taylor, 2015; Hutter and Taylor,

716   2018) than adaptation from SPE.

717

718   Although the Dual Error model provided a better fit of the behavioral results compared to the Adaptation

719   Modulation model, the challenge for future research is to design experiments that can evaluate their

720   unique predictions. The current experiments, by manipulating the size of the target to vary TE while

721   holding SPE constant, highlights the distinctive roles of these two outcome signals. We can envision

722   experiments in which SPE is eliminated (by using a $0°$ clamp), but the target jumps by a small amount,

723   with the size of the displaced target manipulated such that the cursor either misses or hits the target.

724   The Dual Error model, as presently formulated would predict learning in the former, but not in the latter.

725   The Adaptation Modulation model, on the other hand, would predict no learning in either case since

726   there is no SPE.

727

728   In terms of neural mechanisms, converging evidence points to a critical role for the cerebellum in SPE-

729   driven sensorimotor adaptation (Tseng et al., 2007; Taylor et al., 2010; Izawa et al., 2012; Schlerf et al.,

730   2012; Butcher et al., 2017), including the observation that patients with cerebellar degeneration show a

731   reduced response to visual error clamps (Morehead et al., 2017). An important question for future

732   research is whether the cerebellum is also essential for learning driven by information concerning task

733   outcome. A recent behavioral study showed that individuals with cerebellar degeneration were

734   unimpaired in learning from binary, reward-based feedback, once the motor variability associated with

735   their ataxia was taken into consideration (Therrien et al., 2016). This finding provides one instance in

736   which the cerebellum is not essential for learning from task outcome. However, the complete retention

737   observed in that study would indicate that learning was of a different form than adaptation, perhaps

738   related to the use of an explicit strategy (Holland et al., 2018). Evidence that the cerebellum may be

739   integral to processing task outcome signals that could support implicit processes comes from research

34

740   with animal models indicating that both simple (Wagner et al., 2017) and complex (Ohmae and Medina,

741   2015) spike activity in the cerebellum may signal information about task outcome and reward prediction

742   errors. By testing individuals with cerebellar impairment on a clamp design in which SPE is held

743   constant and TE is manipulated, one can simultaneously assess the role of the cerebellum in learning

744   from these two error signals.

745

746   *Conclusions*

747   By using non-contingent feedback, we were able to re-examine the effect of task outcome on

748   sensorimotor learning. The results clearly show that 1) implicit learning processes are influenced by

749   information concerning task outcome, either through the generation of an intrinsic reward or task error

750   signal and 2) that the effect cannot be accounted for by the engagement of a model-based SPE-driven

751   adaptation process operating in tandem with an independent model-free operant reinforcement

752   process. The behavioral results and our modeling work indicate the need for a more nuanced view of

753   sensorimotor adaptation. We outline two directions to consider. In the Adaptation Modulation model,

754   task outcome signals are proposed to serve as a gain on adaptation, contrary to previous views of a

755   modular system that is immune to information about task success. The Dual Error model suggests the

756   need for a more expansive definition of adaptation in which multiple model-based processes operate to

757   keep the sensorimotor system well-calibrated. These models can serve as a springboard for future

758   research designed to further delineate how information about motor execution and task outcome

759   influence implicit sensorimotor learning.

760

761

762

763

764

765

766 **ACKNOWLEDGMENTS**

767 We thank Matthew Hernandez and Wendy Shwe for assistance with data collection. We are also

768 grateful to Maurice Smith, Ryan Morehead, Guy Avraham, and Ian Greenhouse for helpful discussions

769 regarding this work.

770

771

772 **Competing Interests**

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

774

775

776 **Ethics**

777 Human subjects: All participants provided written informed consent to participate in the study and to

778 allow publication of their data, and received financial compensation for their participation. The

779 Institutional Review Board at UC Berkeley approved all experimental procedures under ID number

780 2016-02-8439.

REFERENCES

Bond KM, Taylor JA (2015) Flexible explicit but rigid implicit learning in a visuomotor adaptation task. J Neurophysiol 113:3836–3849.

Butcher PA, Ivry R, Kuo S-H, Rydz D, Krakauer JW, Taylor JA (2017) The Cerebellum Does More Than Sensory-Prediction-Error-Based Learning In Sensorimotor Adaptation Tasks. bioRxiv:139337.

Cameron BD, Franks IM, Inglis JT, Chua R (2010) Implicit motor learning from target error during explicit reach control. Exp Brain Res 206:99–104.

Cashaback JGA, McGregor HR, Mohatarem A, Gribble PL (2017) Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. PLOS Comput Biol 13:e1005623.

Castro LNG, Monsen CB, Smith MA (2011) The Binding of Learning to Action in Motor Adaptation. PLOS Comput Biol 7:e1002052.

Churchland MM, Afshar A, Shenoy KV (2006) A Central Source of Movement Variability. Neuron 52:1085–1096.

Codol O, Holland PJ, Galea JM (2017) The relationship between reinforcement and explicit strategies during visuomotor adaptation. bioRxiv:206284.

Codol O, Holland PJ, Galea JM (2018) The relationship between reinforcement and explicit control during visuomotor adaptation. Sci Rep 8:9121.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-Based Influences on Humans' Choices and Striatal Prediction Errors. Neuron 69:1204–1215.

Day KA, Roemmich RT, Taylor JA, Bastian AJ (2016) Visuomotor Learning Generalizes Around the Intended Movement. eNeuro 3 Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4894913/ [Accessed September 21, 2017].

Donchin O, Francis JT, Shadmehr R (2003) Quantifying Generalization from Trial-by-Trial Behavior of Adaptive Systems that Learn with Basis Functions: Theory and Experiments in Human Motor Control. J Neurosci 23:9032–9045.

Fitts PM (1992) The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychol Gen 121:262–269.

Galea JM, Mallia E, Rothwell J, Diedrichsen J (2015) The dissociable effects of punishment and reward on motor learning. Nat Neurosci 18:597–602.

Gallivan JP, Logan L, Wolpert DM, Flanagan JR (2016) Parallel specification of competing sensorimotor control policies for alternative action options. Nat Neurosci 19:320–326.

Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416–1419.

Gordon J, Ghilardi MF, Ghez C (1994) Accuracy of planar reaching movements - I. Independence of direction and extent variability. Exp Brain Res 99:97–111.

817 Hadjiosif A, Smith M (2013) Savings is restricted to the temporally labile component of motor
818   adaptation. In: Translational and Computational Motor Control (TCMC). San Diego, CA.

819 Haith AM, Krakauer JW (2018) The multiple effects of practice: skill, habit and reduced cognitive load.
820   Curr Opin Behav Sci 20:196–201.

821 Holland PJ, Codol O, Galea JM (2018) The contribution of explicit processes to reinforcement-based
822   motor learning. J Neurophysiol Available at: https://www-physiology-
823   org.libproxy.berkeley.edu/doi/abs/10.1152/jn.00901.2017 [Accessed May 17, 2018].

824 Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking motor learning and savings in
825   adaptation paradigms: model-free memory for successful actions combines with internal
826   models. Neuron 70:787–801.

827 Huberdeau DM, Krakauer JW, Haith AM (2015) Dual-process decomposition in human sensorimotor
828   adaptation. Curr Opin Neurobiol 33:71–77.

829 Huberdeau DM, Krakauer JW, Haith AM (2017) Practice induces a qualitative change in the memory
830   representation for visuomotor learning. bioRxiv:226415.

831 Hutter SA, Taylor JA (2018) Relative sensitivity of explicit re-aiming and implicit motor adaptation.
832   Available at: http://biorxiv.org/lookup/doi/10.1101/308510 [Accessed September 14, 2018].

833 Izawa J, Criscimagna-Hemminger SE, Shadmehr R (2012) Cerebellar contributions to reach adaptation
834   and learning sensory consequences of action. J Neurosci 32:4230–4239.

835 Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor
836   adaptation. PLoS Comput Biol 7:e1002012–e1002012.

837 Kim HE, Morehead JR, Parvin DE, Moazzezi R, Ivry RB (2018) Invariant errors reveal limitations in
838   motor correction rather than constraints on error sensitivity. Commun Biol 1:19.

839 Kim HE. 2019. Influence-of-task-outcome-on-implicit-motor-learning. Github.
840 https://github.com/hyosubkim/Influence-of-task-outcome-on-implicit-motor-learning. 0583d26.
841

842 Knill DC, Bondada A, Chhabra M (2011) Flexible, Task-Dependent Use of Sensory Feedback to
843   Control Hand Movements. J Neurosci 31:1219–1237.

844 Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427:244–247.

845 Krakauer JW, Ghez C, Ghilardi MF (2005) Adaptation to Visuomotor Transformations: Consolidation,
846   Interference, and Forgetting. J Neurosci 25:473–478.

847 Lakens D (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical
848   primer for t-tests and ANOVAs. Front Psychol 4 Available at:
849   https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00863/full [Accessed June 29, 2018].

850 Leow L-A, Marinovic W, de Rugy A, Carroll TJ (2018) Task errors contribute to implicit aftereffects in
851   sensorimotor adaptation. Eur J Neurosci Available at: http://doi.wiley.com/10.1111/ejn.14213
852   [Accessed November 24, 2018].

853   Magescas F, Prablanc C (2006) Automatic Drive of Limb Motor Plasticity. J Cogn Neurosci 18:75–83.

854   Manohar SG, Finzi RD, Drew D, Husain M (2017) Distinct Motivational Effects of Contingent and
855   Noncontingent Rewards. Psychol Sci 28:1016–1026.

856   Mazzoni P, Krakauer JW (2006) An implicit plan overrides an explicit strategy during visuomotor
857   adaptation. J Neurosci 26:3642–3645.

858   McDougle SD, Bond KM, Taylor JA (2015) Explicit and Implicit Processes Constitute the Fast and Slow
859   Processes of Sensorimotor Learning. J Neurosci 35:9568–9579.

860   McDougle SD, Ivry RB, Taylor JA (2016) Taking Aim at the Cognitive Side of Learning in Sensorimotor
861   Adaptation Tasks. Trends Cogn Sci 20:535–544.

862   Morehead JR, Taylor JA, Parvin DE, Ivry RB (2017) Characteristics of Implicit Sensorimotor Adaptation
863   Revealed by Task-irrelevant Clamped Feedback. J Cogn Neurosci 29:1061–1074.

864   Nikooyan AA, Ahmed AA (2015) Reward feedback accelerates motor learning. :633–646.

865   Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error during
866   cerebellar learning in mice. Nat Neurosci 18:1798–1803.

867   Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into
868   movies. Spat Vis 10:437–442.

869   Reichenthal M, Avraham G, Karniel A, Shmuelof L (2016) Target size matters: target errors contribute
870   to the generalization of implicit visuomotor learning. J Neurophysiol 116:411–424.

871   Schlerf J, Ivry RB, Diedrichsen J (2012) Encoding of Sensory Prediction Errors in the Human
872   Cerebellum. J Neurosci 32:4913–4922.

873   Schmitz G, Bock O, Grigorova V, Ilieva M (2010) Adaptation of eye and hand movements to target
874   displacements of different size. Exp Brain Res 203:479–484.

875   Shadmehr R, Smith MA, Krakauer JW (2010) Error correction, sensory prediction, and adaptation in
876   motor control. Annu Rev Neurosci 33:89–108.

877   Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW (2012) Overcoming Motor
878   "Forgetting" Through Reinforcement Of Learned Actions. J Neurosci 32:14617-14621a.

879   Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting Adaptive Processes with Different
880   Timescales Underlie Short-Term Motor Learning. PLOS Biol 4:e179.

881   Soechting JF (1984) Effect of target size on spatial and temporal characteristics of a pointing
882   movement in man. Exp Brain Res 54:121–132.

883   Summerside E, Shadmehr R, Ahmed AA (2018) Vigor of reaching movements: reward discounts the
884   cost of effort. J Neurophysiol Available at:
885   https://www.physiology.org/doi/abs/10.1152/jn.00872.2017 [Accessed March 16, 2018].

886   Tanaka H, Krakauer JW, Sejnowski TJ (2012) Generalization and Multirate Models of Motor
887   Adaptation. Neural Comput 24:939–966.

888 Taylor JA, Hieber LL, Ivry RB (2013) Feedback-dependent generalization. J Neurophysiol 109:202–
889   215.

890 Taylor JA, Ivry RB (2011) Flexible Cognitive Strategies during Motor Learning Diedrichsen J, ed. PLoS
891   Comput Biol 7:e1001096.

892 Taylor JA, Klemfuss NM, Ivry RB (2010) An explicit strategy prevails when the cerebellum fails to
893   compute movement errors. Cerebellum 9:580–586.

894 Taylor JA, Krakauer JW, Ivry RB (2014) Explicit and Implicit Contributions to Learning in a
895   Sensorimotor Adaptation Task. J Neurosci 34:3023–3032.

896 Therrien AS, Wolpert DM, Bastian AJ (2016) Effective reinforcement learning following cerebellar
897   damage requires a balance between exploration and motor noise. Brain 139:101–114.

898 Therrien AS, Wolpert DM, Bastian AJ (2018) Increasing Motor Noise Impairs Reinforcement Learning in
899   Healthy Individuals. eneuro:ENEURO.0050-18.2018.

900 Thoroughman KA, Shadmehr R (2000) Learning of action through adaptive combination of motor
901   primitives. Nature 407:742–747.

902 Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ (2007) Sensory prediction errors
903   drive cerebellum-dependent adaptation of reaching. J Neurophysiol 98:54–62.

904 Vandevoorde K, Orban de Xivry J-J (2018) Motor adaptation but not internal model recalibration
905   declines with aging. bioRxiv Available at: http://biorxiv.org/lookup/doi/10.1101/292250
906   [Accessed June 12, 2018].

907 Vaswani PA, Shmuelof L, Haith AM, Delnicki RJ, Huang VS, Mazzoni P, Shadmehr R, Krakauer JW
908   (2015) Persistent Residual Errors in Motor Adaptation Tasks: Reversion to Baseline and
909   Exploratory Escape. J Neurosci 35:6969–6977.

910 Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the
911   expectation of reward. Nature Available at:
912   http://www.nature.com/doifinder/10.1038/nature21726 [Accessed April 2, 2017].

913 Wu HG, Miyamoto YR, Castro LNG, Ölveczky BP, Smith MA (2014) Temporal Structure of Motor
914   Variability is Dynamically Regulated and Predicts Motor Learning Ability. Nat Neurosci 17:312–
915   321.

916
917
918
919
920

921 **METHODS**

922 *Participants:* Healthy, young adults (N=116, 69 females; average age = 20.9 years old, range: 18.2 –

923 27.8) were recruited from the University of California, Berkeley, community. Each participant was tested

924 in only one experiment and was right-handed, as verified with the Edinburgh Handedness Inventory

925 (Oldfield, 1971). All participants provided written informed consent to participate in the study and to

926 allow publication of their data, and received financial compensation for their participation. The

927 Institutional Review Board at UC Berkeley approved all experimental procedures under ID number

928 2016-02-8439.

929

930 *Experimental Apparatus:* The participant was seated at a custom-made tabletop housing an LCD

931 screen (53.2 cm by 30 cm, ASUS), mounted 27 cm above a digitizing tablet (49.3 cm by 32.7 cm,

932 Intuos 4XL; Wacom, Vancouver, WA). The participant made reaching movements by sliding a modified

933 air hockey "paddle" containing an embedded stylus. The position of the stylus was recorded by the

934 tablet at 200 Hz. The experimental software was custom written in Matlab, using the Psychtoolbox

935 extensions (Pelli, 1997)**.**

936

937 *Reaching Task:* Center-out planar reaching movements were performed from the center of the

938 workspace to targets positioned at a radial distance of 8 cm. Direct vision of the hand was occluded by

939 the monitor, and the lights were extinguished in the room to minimize peripheral vision of the arm. The

940 starting and target locations were indicated by white and blue circles, respectively (start circle: 6 mm in

941 diameter; target: either 6, 9.8 or 16 mm depending on condition).

942

943 To initiate each trial, the participant moved the digitizing stylus into the start location. The position of the

944 stylus was indicated by a white feedback cursor (3.5 mm diameter). Once the start location was

945 maintained for 500 ms, the target appeared. For Experiments 1 and 3, the target could appear at one of

946 8 locations, placed in 45° increments around a virtual circle (0°, 45°, 95°, 135°, 180°, 225°, 270°, 315°).

41

947 For Experiment 2, the target could appear at one of four locations placed in 90° increments around a

948 virtual circle (45°, 135°, 225°, 315°). We reduced the number of targets from 8 to 4 in Experiment 2 in

949 order to increase the overall number of training cycles with the clamp to ensure that participants reach

950 a stable asymptote, while keeping the experiment under 1.5 hours. Participants were instructed to

951 accurately and rapidly "slice" through the target, without needing to stop at the target location. Visual

952 feedback, when presented, was provided during the reach until the movement amplitude exceeded 8

953 cm. As described below, the feedback either matched the position of the stylus (veridical) or followed a

954 fixed path (clamped). If the movement duration (excluding RT) was not completed within 300 ms, the

955 words "too slow" were generated by the sound system of the computer.

956

957 After the hand crossed the target ring, endpoint cursor feedback was provided for 50 ms either at the

958 position in which the hand crossed the virtual target ring (veridical feedback) or at a fixed distance

959 determined by the size of the clamp. During the return movement, the feedback cursor reappeared

960 when the participant's hand was within 1 cm of the start position.

961

962 *Experimental Feedback Conditions:* Across the experimental session, there were three types of visual

963 feedback. On no-feedback trials, the cursor disappeared when the participant's hand left the start circle

964 and only reappeared at the end of the return movement. On veridical feedback trials, the cursor

965 matched the position of the stylus during the 8 cm outbound segment of the reach. On clamped

966 feedback trials, the feedback followed a path that was fixed along a specific hand angle. The radial

967 distance of the cursor from the start location was still based on the radial extent of the participant's

968 hand during the 8 cm outbound segment, but the angular position was fixed relative to the target (i.e.,

969 independent of the angular position of the hand).

970

971　The primary instructions to the participant (experiment script included) remained the same across the

972　experimental session: Specifically, that they were to reach directly towards the visual target. Prior to the

973　introduction of the clamped feedback trials, participants were briefed about the feedback manipulation.

974　They were informed that the position of the cursor would now follow a fixed trajectory and that the

975　angular position would be independent of their movement. They were explicitly instructed to ignore the

976　cursor and continue to reach directly to the target. Participants also performed three instructed trials

977　with the clamp perturbation on. During these practice trials, a target appeared at the 90° location

978　(straight ahead), and the experimenter instructed the participant to first "reach straight to the left" (ie,

979　180°). For the second practice trial, the participant was instructed to "reach straight to the right" (0°).

980　For the last trial, the participant was instructed to "reach straight down (towards your torso)" (ie, 270°).

981　The purpose of these trials was to familiarize the participant with the exact clamp condition they were

982　about to experience. Following these three practice trials, the experimenter confirmed with the

983　participant they understood now what was meant by clamped visual feedback. These practice trials

984　were removed from future analyses.

985

986　The same instructions in abbreviated form ("Ignore the cursor and move your hand directly to the target

987　location") were repeated verbally and with onscreen text at every block break during the clamp

988　perturbation. Participants were debriefed at the end of the experiment and asked whether they ever

989　intentionally tried to reach to locations other than the target. All subjects reported aiming to the target

990　throughout the experiment.

991

992　We counterbalanced clockwise and counterclockwise clamps within each group for all three

993　experiments.

994

995　*Experiment 1*

43

996   Participants (n=48, 16/group) were randomly assigned to one of three groups, each training with a 3.5°

997   clamp but differing only in terms of the size of the target: 6mm, 9.8, or 16 mm diameter. These sizes

998   were chosen so that at an 8 cm radial distance the clamped cursor would be adjacent to the target

999   without making any contact (Target Miss group), straddling the target by being roughly half inside and

1000   half outside the target (Straddle Target group), or fully embedded within the target (Hit Target group).

1001   The Euclidean distance for this clamp size, measured from the centers of cursor and target, was 4.9

1002   mm.

1003

1004   The session began with two baseline blocks, the first comprised of 5 movement cycles (40 total

1005   reaches to 8 targets) without visual feedback and the second comprised of 10 cycles with a veridical

1006   cursor displaying hand position. The experimenter then informed the participant that the visual

1007   feedback would no longer be veridical and would now be clamped at a fixed angle from the target

1008   location. Immediately following these general instructions, the experimenter continued providing

1009   instructions for the three practice trials which immediately followed (see Experimental Feedback

1010   Conditions). After the practice trials and confirming the participant's understanding of the task, the

1011   clamp block ensued for a total of 80 cycles. A short break (<1 min), as well as a reminder of the task

1012   instructions, was provided after 40 cycles (i.e., at the halfway point of this block). Immediately following

1013   the perturbation block, there were two washout blocks, first a 5 cycle block in which there was no visual

1014   feedback, followed by 10 cycles with veridical visual feedback. These blocks were preceded by

1015   instructions regarding the change in experimental condition and participants were reminded to always

1016   aim for the target and to attempt to slice through it with their hand.

1017

1018   *Experiment 2*

1019   In Experiment 2 we assessed adaptation over an extended number of clamped visual feedback trials.

1020   The purpose of extending the perturbation block was to ensure that participants reached asymptotic

1021   levels of learning. In order to achieve a greater number of training cycles, we reduced the number of

1022   target locations within the set from 8 to 4.

1023

1024   Participants (n=32, 16/group) trained with a 1.75° clamp (2.4 mm distance between target and cursor

1025   centers) and were assigned to either a small (Straddle) or large (Hit) target condition. The session

1026   started with two baseline blocks, 10 cycles (40 reaches) without visual feedback and then 10 cycles

1027   with veridical feedback. Following 3 practice trials with the clamp, the number of cycles in the clamped

1028   visual feedback block was nearly tripled from that of Experiment 1 to 220 cycles, with breaks provided

1029   after every 70 cycles. Following 220 cycles of training with a 1.75° clamp, there were two washout

1030   blocks, first a 10 cycle block in which there was a 0° clamp, followed by 10 cycles with veridical visual

1031   feedback. Prior to washout, participants were again instructed to always aim directly to the target.

1032

1033   *Experiment 3*

1034   Experiment 3 used a transfer design to evaluate different hypotheses concerning the role of task

1035   outcome on implicit sensorimotor learning. Our main predictions focused on the transfer phase,

1036   comparing the participants' behavior to the predictions of three models (see section, *Theoretical*

1037   *analysis of the effect of task outcome on implicit learning).* We tested two main groups (n=12/group) in

1038   Experiment 3, using a 1.75° clamp in both the acquisition and transfer phases. The session started with

1039   two baseline blocks, 5 cycles (40 reaches) without visual feedback and then 5 cycles with veridical

1040   feedback. After the baseline blocks, clamp instructions and three practice trials were provided to all

1041   participants. The first clamp block (acquisition phase) lasted 120 cycles, with participants training with

1042   either a small or large target. Following the first 120 cycles, the target sizes were reversed for the next

1043   80 cycles (transfer phase: Straddle-to-Hit or Hit-to-Straddle conditions). Breaks of < 1 min were

1044   provided after every 35 cycles of training. On the break preceding the transfer (15 cycles before target

1045   switch), participants were told that everything would continue on as before, except that the target size

1046   would change at some point during the block. The purpose of staggering the break with the transfer

1047   was to mitigate any change in adaptation due to temporal decay that could result from a break in

1048   training (Hadjiosif and Smith, 2013).

1049

1050   *Control group*

1051   A third group (n=12) was added to test whether the attenuation of adaptation in the large target

1052   condition was due to perceptual uncertainty. Here, the block structure was identical to the first two

1053   groups. We used a modified large target (16mm), one which had a bright green bisecting line through

1054   the middle, aligned with the target direction. The clamped cursor always fell within one half of the target

1055   (either clockwise or counter-clockwise depending on the condition), thus providing a clear indication

1056   that the cursor was off center. At the transfer, the bisecting line was removed and participants trained

1057   for 80 cycles with the standard large target.

1058

1059   *Data Analysis*

1060   All statistical analyses and modeling were performed using Matlab 2015b and the Statistics Toolbox.

1061   Data and code are available on GitHub at: https://github.com/hyosubkim/Influence-of-task-outcome-on-

1062   implicit-motor-learning (Kim, 2019). The primary dependent variable in all experiments was hand angle

1063   at peak radial velocity, defined by the angle of the hand relative to the target at the time of peak radial

1064   velocity (i.e., angle between lines connecting start position to target and start position to hand).

1065   Throughout the text, we refer to this variable as hand angle. Additional analyses were performed using

1066   hand angle at "endpoint" (angle of the hand as it crossed the invisible target ring) rather than peak

1067   radial velocity. The results were essentially identical for the two dependent variables; as such, we only

1068   report the results of the analyses using peak radial velocity.

1069

1070   Data used in statistical analyses were tested for normality and homogeneity of variance using Shapiro-

1071   Wilks and Levene's tests, respectively. When normality or homogeneity of variance was violated, we

1072   performed non-parametric permutation tests in addition to standard parametric tests (i.e., t-tests and

1073   ANOVAs) and report results from both. For comparisons between two groups, we used the difference

1074   between group means as our test statistic. This value was compared to a null distribution, created by

1075   random shuffling of group assignment in 10,000 Monte Carlo simulations (resampling with

1076   replacement), to obtain an exact p-value. When a comparison involved more than two groups, we used

1077   a similar approach, but used the F-value obtained from a one-way ANOVA as our test statistic.

1078

1079   Outlier responses were removed from the analyses. For the sole purpose of identifying outliers, the

1080   Matlab "smooth" function was used to calculate a moving average (using a 5-trial window) of the hand

1081   angle data for each target location. Outliers were trials in which the observed hand angle was greater

1082   than 90° or deviated by more than 3 standard deviations from the moving average. In total, less than

1083   0.8% of trials overall were removed, and the most trials removed for any individual across all three

1084   experiments was 2%.

1085

1086   Individual baseline biases for each target location were subtracted from all data. Biases were defined

1087   as the average hand angles across cycles 2-10 (Experiments 1 and 2) or 2-5 (Experiment 3) of the

1088   feedback baseline block. These same cycles were used to calculate mean baseline RTs, MTs, and

1089   movement variability (SD). To calculate each participant's baseline RT or MT, we took the average of

1090   median values at each target location. To calculate each participant's movement variability, we took the

1091   average of the standard deviations of hand angles at each target location.

1092

1093   In order to pool all of the data and to aid visualization, we flipped the hand angles for all participants

1094   clamped in the counterclockwise direction.

1095

1096   For Experiments 1 and 3, movement cycles consisted of 8 consecutive reaches (1 reach/target); for

1097   Experiment 2, we only used four targets, thus a movement cycle consisted of 4 consecutive reaches (1

1098   reach/target). To estimate the rate of early adaptation, we calculated the mean change in hand angle

1099 per cycle over the first five cycles. To provide a more stable estimate of hand angle at cycle 5, we

1100 averaged over cycles 3-7 of the clamp block. We opted to use this measure of early adaptation rather

1101 than obtain parameter estimates from exponential fits since the latter approach gives considerable

1102 weight to the asymptotic phase of performance, and, therefore, would be less sensitive to early

1103 differences in rate. This would be especially problematic in Experiment 2, which utilized 220 clamp

1104 cycles. We also performed a secondary analysis of early adaptation rates using a larger window, cycles

1105 2-11 (Krakauer et al., 2005). Results from using this alternate metric were consistent with the reported

1106 analyses (i.e., slower rates for Hit Target groups), only they resulted in larger effect sizes due to the

1107 gradually increasing divergence of learning functions. Asymptotic adaptation (i.e., late learning) was

1108 defined as the average hand angle over the last 10 cycles within a clamp block. In Experiment 1, the

1109 aftereffect was quantified by using the data from the first no-feedback cycle following the last clamp

1110 cycle. This measure yielded similar statistical results as that based on the analysis of asymptotic

1111 adaptation.

1112

1113 All t-tests were two-tailed. Posthoc pairwise comparisons following significant ANOVAs were performed

1114 using two-tailed t-tests, with a corrected α of .017 due to multiple comparisons. Cohen's d, eta squared

1115 (η^2), partial eta squared (for mixed model ANOVA), and d_z (for within-subjects design) values are

1116 provided as standardized measures of effect size (Lakens, 2013). Values in main text are reported as

1117 95% CIs in brackets and mean $\pm$ SEM.

1118

1119 No statistical methods were used to predetermine sample sizes. The chosen sample sizes were based

1120 on our previous studies using the clamp method (Morehead et al., 2017; Kim et al., 2018), as well as

1121 prior psychophysical studies of human sensorimotor learning (Huang et al., 2011; Galea et al., 2015;

1122 Vaswani et al., 2015; Gallivan et al., 2016).

1123

1124

1125

1126   *Modeling*

1127   For the Movement Reinforcement model, a population vector (Georgopoulos et al., 1986), V, indicates

1128   the current bias of motor representations within the reinforcement system. The direction of this vector

1129   (V_d) was calculated for each trial in the following manner:

1130 $$V_x(n) = \mathbf{r}(n) \bullet \mathbf{u_x}$$

1131 $$V_y(n) = \mathbf{r}(n) \bullet \mathbf{u_y}$$

1132 $$V_d(n) = \tan^{-1}(V_y(n) / V_x(n))$$

1133   Here, $\mathbf{r}$ represents the weights on every unit (36,000 total units representing a resolution of .01°); $\mathbf{u}$ is a

1134   vector filled with unit vectors pointing in all 36,000 directions; and x and y subscripts represent the x-

1135   and y-components for both V and $\mathbf{u}$, respectively. The mean preferred direction, V_d, was converted

1136   from radians into degrees. The strength of the biasing signal, V_l, is equal to the population vector

1137   length: $\sqrt{V_x^2 + V_y^2}$, with the constraint that $0 \leq V_l \leq 1$.

1138

1139   In order to calculate confidence intervals for the parameter estimates, we applied standard

1140   bootstrapping techniques, constructing group-averaged hand angle data 1000 times by randomly

1141   resampling with replacement from the pool of participants within each group. Using Matlab's *fmincon*

1142   function, we started with ten different initial sets of parameter values and estimated the retention and

1143   learning parameters that minimized the least squared error between the bootstrapped data and model

1144   output (x_n). Parameter estimates were bounded such that *0 < A < 1* and *0 < U(e) < e*, where *e* is equal

1145   to the clamp size in degrees.

1146

1147   The hybrid models combined the Movement Reinforcement with either the Adaptation Modulation or

1148   Dual Error model. Each hybrid incorporated the equations for the Movement Reinforcement model

49

1149 (Eqns. 2-4). However, when movement reinforcement was combined with the Adaptation Modulation

1150 model, the contribution of the adaptation system, x, to the motor output, y, was derived from the gain

1151 modulation equation (Eqn. 5). When movement reinforcement was combined with the Dual Error

1152 model, equations 6-8 were used, with x_{total} now substituting for x in Eqn. 2.

1153 **SUPPLEMENTAL INFORMATION**

1154 *Experiment 3: Kinematic variables*

1155 Baseline movement variability was not different across all three groups, including the control group

1156 trained with the bisected target ($F(2,33)=1.38$, $p=.267$, $\eta^2=.077$). Similarly, no differences across groups

1157 were observed for either RTs ($F(2,33)=1.51$, $p=.236$, $\eta^2=.0084$) or MTs ($F(2,33)=.46$, $p=.634$, $\eta^2=.027$).

1158
1159
1160 **Table 2: Average Reaction Times (RTs).** Values represent mean $\pm$ SEM.

Experiment 1	Baseline	Early clamp	Late clamp	No feedback
Hit	325 $\pm$ 7	327 $\pm$ 7	347 $\pm$ 11	344 $\pm$ 12
Straddle	362 $\pm$ 12	359 $\pm$ 14	397 $\pm$ 32	407 $\pm$ 33
Miss	386 $\pm$ 22	383 $\pm$ 19	378 $\pm$ 15	385 $\pm$ 15
Experiment 2				**0° clamp**
Hit	378 $\pm$ 22	376 $\pm$ 27	354 $\pm$ 9	351 $\pm$ 9
Straddle	373 $\pm$ 12	366 $\pm$ 13	368 $\pm$ 15	373 $\pm$ 16
Experiment 3				
Hit-to-Straddle	356 $\pm$ 19	350 $\pm$ 15	326 $\pm$ 9	N/A
Straddle-to-Hit	360 $\pm$ 8	360 $\pm$ 7	355 $\pm$ 7	N/A
Bisected-to-Normal	400 $\pm$ 28	395 $\pm$ 27	400 $\pm$ 25	N/A

1161
1162
1163 **Table 3: Average Movement Times (MTs).** Values represent mean $\pm$ SEM.

| Experiment 1	Baseline	Early clamp	Late clamp	No feedback
Hit	153 $\pm$ 11	150 $\pm$ 10	137 $\pm$ 8	133 $\pm$ 9
Straddle	162 $\pm$ 8	149 $\pm$ 8	139 $\pm$ 7	131 $\pm$ 7
Miss	137 $\pm$ 7	134 $\pm$ 7	124 $\pm$ 6	118 $\pm$ 6
Experiment 2				**0° clamp**
Hit	149 $\pm$ 8	159 $\pm$ 20	155 $\pm$ 11	127 $\pm$ 7
Straddle	157 $\pm$ 8	161 $\pm$ 15	170 $\pm$ 18	130 $\pm$ 8
Experiment3				
Hit-to-Straddle	158 $\pm$ 7	189 $\pm$ 12	168 $\pm$ 12	N/A
Straddle-to-Hit	164 $\pm$ 11	207 $\pm$ 28	169 $\pm$ 13	N/A
Bisected-to-Normal	151 $\pm$ 11	165 $\pm$ 14	166 $\pm$ 15	N/A

1164
1165
1166
1167
1168
1169
1170
1171

51

1172   **Table 4: Movement variability during baseline block.** Values represent mean $\pm$ SEM.

Experiment 1	Baseline SD
Hit	$4.19 \pm .26°$
Straddle	$3.61 \pm .16°$
Miss	$3.80 \pm .15°$
Experiment 2	
Hit	$3.09 \pm .18°$
Straddle	$3.57 \pm .16°$
Experiment 3	
Hit-to-Straddle	$3.30 \pm .22°$
Straddle-to-Hit	$3.85 \pm .37°$
Bisected-to-Normal	$3.97 \pm .31°$

1173
1174
1175
1176
1177
1178

1179
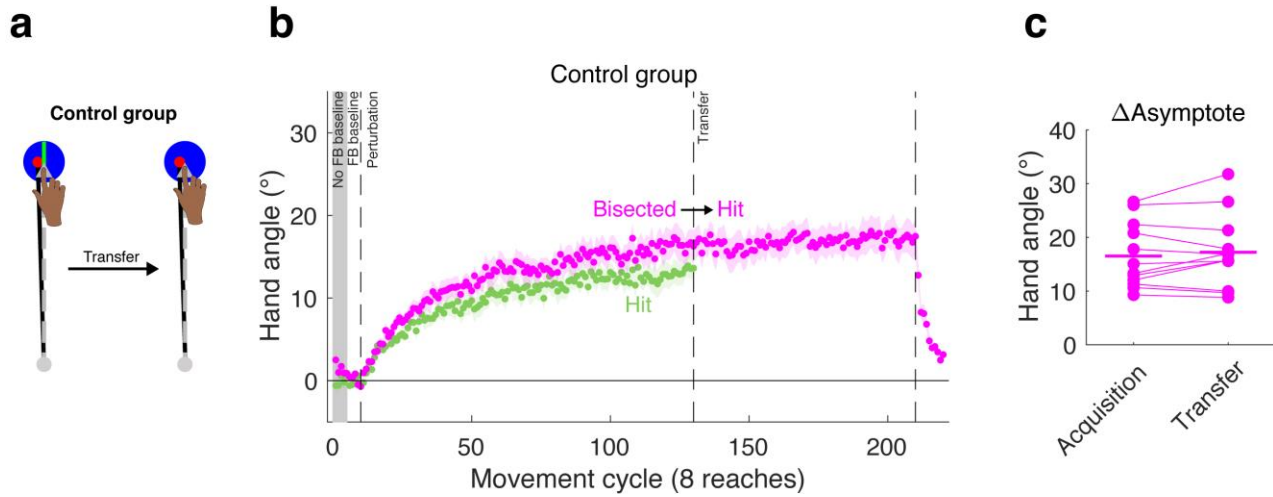
1180
1181
1182 **Figure 5-figure supplement 1 Control group from Experiment 3. (a)** Using a transfer design, **(b)** the
1183 behavior of the control group reaching to a large target bisected by a line (aligned with the target
1184 direction) during the acquisition phase is shown in magenta. The change in hand angle was not
1185 significantly different than that observed for the group that was tested with the large target in the
1186 acquisition phase of Experiment 3 (re-plotted here in green), suggesting that perceptual uncertainty did
1187 not make a substantive contribution to the effects of hitting the target. We omitted the transfer behavior
1188 of the large target group as this was when the large target was replaced with the small target. **(c)** There
1189 was no change in asymptote when going from the bisected target to the standard large target.
1190

1191

Target Size Experiment Instructions

Pre-Experiment Instructions (verbal):

- "The point of this experiment is to better understand how the brain controls reaching movements. We are especially interested in how motor control is impacted by various neurological disorders, like Parkinson's disease and cerebellar degeneration. As a healthy student, your data will be used as a normative baseline for future comparisons with these neurological patients. The information we gain from this experiment could help improve rehabilitation for these patients, so please try your best to pay attention and follow all instructions."
- "You will be holding this 'pen' at the red base and should maintain the same grip throughout the experiment. (Demonstrate how to hold pen) You will be making fast center-out reaching movements towards a blue target that will appear in one of several different locations." (Physically demonstrate the reaching movements)
- "Try your best to reach quickly and accurately in a straight line, and try to slice through the blue targets rather than stopping at the target. So, slice and come back to the home position."
- "These are center-out reaches, which means that every trial starts from the exact same home location."
- "During the experiment, you will see three different scenarios: you either won't be able to see the cursor, your cursor will reflect your hand position, or you won't be controlling the cursor at all. Regardless of what phase of the experiment you are in, your goal should always be to bring your hand directly to the blue target. And, of course, I will always tell you ahead of time which condition you are in."
- "You will go through several different blocks of trials, some lasting longer than others, but there will be several breaks at different points of the experiment."
- "I will give you time to adjust the seat height and scoot in close to the work station; you will be making many reaches towards the end of the tablet, and I want you to be able to do that without moving any parts of your body other than your arm." (demonstrate)
- "Try to keep the same posture throughout the experiment."
- "Rest your left hand in your lap."
- "I will talk you through the first several trials."
- "Do you have any questions?" (Run game code and switch seats.)
- "I'll give you a minute to get comfortable before turning out the lights."

**Start program, move mouse and keypad
**Participant gets situated before turn off the light
**Turn off lights and close door

No feedback baseline block:
"In the center of the screen you can see a white circle that indicates the start position for each trial. When your hand is close to the start position, a cursor will appear indicating your actual hand position. For this first block of trials you won't get to see the cursor when you reach to the target, but remember to still move your hand quickly to the target."

(Explain "too slow" message after it happens. Remind them to make slicing movements.)

1240 (After several trials): "That knocking sound you hear means it was a valid trial and that you moved far
1241 enough and fast enough. It does not mean you hit the target. There is no special sound for hitting the
1242 target."

1243

1244 **Veridical feedback:** *After first pause (screen: "Good job!")*
1245 "Great! The only difference for the next block of trials is that now you can see your cursor as you move
1246 to the target. Continue to move your hand directly towards the target."

1247

1248 *Hit space bar.*

1249

1250 **Error clamp block:**

1251

1252 **Practice trials**
1253 *(screen: "You will now get to familiarize yourself with the next condition.")*
1254 - "Great! During the next long block of trials you will continue to see the cursor, but its movement
1255   direction will **not** be under your control. We will want you to ignore the cursor as you **continue**
1256   **to focus on slicing through the targets with your hand."**
1257 - "Before we start the next block, though, we'll do three practice trials, just so that you will be
1258   familiar with exactly what will happen in the next part of the experiment."
1259 - "As a demonstration of what we mean by the cursor not being under your control, in the very
1260   next trial, after the blue target appears, move your hand quickly to the left." (*hit space bar*)
1261 - "Did you see how the cursor moved when you did, but it followed a fixed path independent of
1262   where your hand moved?"
1263 - "I want you to now move quickly to the right after you find home and the target appears." (*hit
1264   space bar*)
1265 - "For the last practice trial, move in the opposite direction of the target."

1266

1267 **Perturbation**
1268 *(screen: "Ignore the cursor and move your hand directly to the target location.")*
1269 - "Is it clear now that the cursor's direction will not be under your control?"
1270 - "Those practice trials were so that you could see and experience what's going to happen in the
1271   next block. The cursor will move like that towards all the targets now. You'll be controlling the
1272   distance the cursor travels, but not its direction. We want to see how well you can reach to the
1273   target while ignoring the cursor."
1274 - "Your job is again to always reach directly for the target. But, remember, for the next block of
1275   trials you do not have control over the cursor's direction, so try to ignore the cursor and always
1276   aim directly for the targets."
1277 - "Is this clear?"

1278

1279 After first couple of trials "Remember that you are controlling how far that cursor moves, but not its
1280 direction, even though it may look like it. No matter where you move your hand, the cursor will still
1281 travel along the same path relative to the target. Continue to ignore the cursor and move your hand
1282 directly to the target."

1283

1284 *If slow, and trying to gauge cursor direction:* "Please make quick, straight reaches to the target."

1285

1286 **Short break – occurs after X trials (screen: "Ignore the cursor and move your hand directly to**
1287 **the target location.")**

1288   •   "Feel free to take a quick break if you'd like, but please return back to the same grip and
1289   posture when you're ready. If you're okay, we can just keep going. Like before, you still won't be
1290   controlling the cursor, but we still want you to move your hand directly to the blue target."
1291
1292
1293
1294 **No feedback washout:**
1295   •   "Great. Keep bringing your hand directly to the target. The only change is you won't see the
1296   cursor."
1297
1298
1299 **Veridical feedback washout:**
1300   •   "Nice job. You are back in control of the cursor now. Keep bringing your hand to the target.
1301   Remember to make quick, straight reaches. Once you begin your reach, follow through with it."
1302   (You are making sure they are not trying to make online feedback corrections)
1303