Implicit rewards modulate sensorimotor adaptation

Hyosub E. Kim, Darius E. Parvin, Matthew A. Hernandez, Richard B. Ivry

Recent work has begun to shed light on interactions between task performance and sensory-based feedback in motor learning. In sensorimotor adaptation, task performance error (TPE) can engage explicit learning processes, whereas sensory prediction error (SPE), the difference between predicted and actual sensory feedback drives implicit adaptation\(^1\).\(^2\). Despite various demonstrations of the obligatory nature of learning from SPE, an open question is whether, and to what degree, the adaptation system is sensitive to TPE. The manipulation of success rate by varying target size has been shown to affect generalization following visuomotor adaptation\(^3\), and reward may increase learning rate\(^4\) or enhance retention of recently formed motor memories\(^5\). However, since TPE and SPE are both contingent on behavior in these studies, it is difficult to disentangle the effects of reward on different learning processes. Here we divorce task performance feedback from behavior in order to better understand how TPE signals may influence implicit adaptation from SPE.

We used clamped visual feedback, in which the angular trajectory of a feedback cursor is invariant with respect to the target location and thus spatially independent of hand position (Fig.1b; groups always counterbalanced CW and CCW). The instructions emphasized that the participants’ behavior would not influence the cursor trajectory; they were to ignore this stimulus and always aim directly for the target. This method allows us to measure learning from an invariant SPE, eliminating potential contributions from strategic changes that might be used to reduce TPE. The clamp offset was held constant with only the size of the target manipulated. The goal of this manipulation was to ask if an “implicit” task success signal (e.g., target hits or misses) modulated the response to a constant SPE. We use the phrase “implicit” task success because participants are told to ignore the feedback as it is noncontingent on their behavior.

In Experiment 1, two groups of participants (n=16/group) trained with a 1.75° clamp for 220 cycles (1 reach to each of 4 targets). For the small target group, the clamped cursor feedback straddled the 6 mm diameter target at reach endpoint. For the big target group, the feedback cursor ended within the 16 mm target. Consistent with the hypothesis that the adaptation system is sensitive to implicit TPE, the big target group had a slower early adaptation rate (p=.039) and lower asymptote (p=.025) compared to the small target group (Fig.1c-e). Despite adapting less, assessment of retention during 10 cycles of a 0° clamp revealed the participants in the big target group retained a higher proportion of what was learned (p=.046; Fig.1f).

In Experiment 2, three groups of participants (n=16/group) were trained with a 3.5° clamp for 80 cycles (8 targets per cycle). Between groups, the target size was either 6 mm, 9.8 mm, or 16 mm. Thus, with the larger clamp, the final position of the cursor was either outside (small), straddled (medium), or inside (big) the target. (Fig.2a) Consistent with the results of Experiment 1, we observed a significant attenuation of adaptation for the big target group (Fig.2b). Interestingly, there were no reliable differences between the small and medium groups, suggesting that partial overlap of clamp and target does not act as an effective implicit reward signal for modulating adaptation (Fig.2c,d).

We hypothesize that with the big targets, an implicit reward is generated when the cursor lands within the target. This could serve as a positive reinforcement signal, strengthening the representation of rewarded movements\(^6\), and operating as a counterweight to the learning drive associated with an SPE. An action-reinforcement hypothesis would be consistent with the stronger retention observed in the big target group during the 0° clamp phase of Experiment 1. As a stronger test of this hypothesis, we tested two groups (n=12/group) in Experiment 3, using a 1.75° clamp in a transfer design. In the first 120 cycles, participants trained with either a small or big target (Fig.3a). Following the first 120 cycles, the target sizes were reversed for the next 80 (small-to-big or big-to-small). Our main predictions focused on the transfer phase, evaluating the hypothesis that clamped feedback embedded within the target reinforces actions. The big-to-small group should show increased adaptation at transfer since the movements no longer receive implicit reward. In contrast, the small-to-big group should remain around the level of adaptation observed at the end of the first training phase, since the SPE remains the same and their actions now become rewarded. While the big-to-small group did show an increase in adaptation (p<.001), the small-to-big group exhibited a reduction in adaptation at transfer (p<.001; Fig.3c). The marked decay in adaptation suggests that implicit reward may modulate adaptation more directly, such as through attenuation of the SPE.

A third group (n=12) was added to test whether the attenuation of adaptation in the big target condition was due to perceptual uncertainty. Here, the clamped cursor landed within the target, but a bright bisecting line indicated that the cursor was off-center (Fig.3b). Performance for this group was similar to the big target group.

In summary, the results of these studies demonstrate that implicit reward modulates adaptation from SPE. As emphasized in earlier studies, these rewards may reinforce associated actions. In addition, implicit reward may also modulate the efficacy of learning from SPE.
Figure 1. Adaptation demonstrates sensitivity to implicit task performance error signals. (a) Illustration of experimental apparatus and basic task structure. Note that the cursor feedback is invariant and spatially independent of hand angle (i.e., hand angle changes, but cursor feedback stays constant). Clamp offsets are equal and only target sizes are different between groups. For all experiments, groups were counterbalanced with half of the participants training with a clockwise clamp and the other half with a counterclockwise clamp. (b) Scaled view of visual stimuli for clamped feedback. Learning functions for small and big target groups. (c) Learning functions for small and big target groups. (d) The big target group demonstrated both an attenuated adaptation rate (mean change of hand angle over cycles 2-11 of clamp) and (e) a lower asymptote (mean hand angle over last 10 clamp cycles). (f) Bootstrapped 95% confidence intervals are shown for the retention parameter estimates used to fit the 10 cycles of reaching with a 0° clamp following the perturbation (p=.046 for difference between groups). Retention was modeled during this phase using the following equation: \( x(n+1) = Ax(n) \), where \( A \) represents the proportion of the motor output, \( x \), retained on trial \( n \). Dots are individuals; shading and error bars represent SEM, except in figure f. Gray shading denotes trials with no visual feedback.

Figure 2. Partial overlap between clamped cursor and target does not affect normal adaptation. (a) In this experiment, all three groups trained with a 3.5° clamp, with the only difference between groups being the size of the target. (b) The learning functions for the small and medium size target groups were similar. Only the big target group (i.e., group with fully embedded clamp) demonstrated attenuation of adaptation. (c) Although there were no reliable differences in early adaptation rates between groups (\( F_{3,38}=1.98, p=.15 \)), there were differences in early adaptation rates between groups (\( F_{3,38}=3.88, p=.028 \)), with the big target group showing less adaptation than either the small (\( p=.011 \)) or medium (\( p=.016 \)) target groups. n.s. means non-significant.

Figure 3. The effects of implicit reward cannot be explained by reinforcement alone. (a) During the first phase of the perturbation, the big-to-small target group adapted slower (\( p=.002 \)) and reached a lower asymptote than the small-to-big target group (\( p<.001 \)). A third group trained with a big target with a bright, bisecting line and adapted in a comparable manner as the big-to-small target group did in the first phase. In terms of both early adaptation rate (\( p=.13 \)) and asymptote (\( p=.59 \)), (b) following the transfer phase to different target sizes, the small-to-big group demonstrated significant decay in adaptation. The bisected target group showed no change in adaptation after going from the bisected big target to a non-bisected big target (\( p=.22 \)). **p<.001.