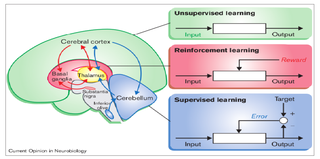
The role of error-based and reward-based mechanisms in a motor learning task

Sarah Hillenbrand[1], Katherine Okpara[2], & Richard Ivry[1,2]
1 Helen Wills Neuroscience Institute, 2 Department of Psychology, University of California--Berkeley
UU11
hillenbrands@berkeley.edu

Introduction

Learning a new motor task involves making corrections based on a combination of sensory and reward prediction errors.



Doya 2010

Reward-based learning:
* Reward prediction error
* Action selection
* Goal changes
* Basal ganglia

Error-based learning:
* Sensory prediction error
* Action specification
* Goal unchanged
* Cerebellum

Corrections may occur in response to a **change in goal** or in response to an **error**.

Changes in goal may arise from **large errors** (subject-produced action selection) or from a **change in target location** (world-produced).

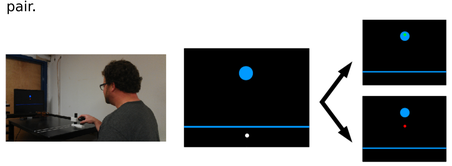
Errors may arise from **motor noise** (subject-produced action specification) or from **perturbations** (world-produced).

GOAL: Develop a task to dissociate the neural correlates of:

* **Changes in goal vs. errors**
* **Repeated successes vs. error corrections**
* **Action selection vs. action specification**

Task

* We developed a novel "**shuffleboard**" task.
* Participants learn **target-appropriate movement speeds/ forces.**
* Target appears at one of 3-4 locations. Participant moves robot arm to "release" puck, with distance traversed by puck determined by speed of arm when crossing start line.
* **Target location is fixed for two successive trials** and then changes position.
* **Endpoint-only feedback**, sometimes restricted to first trial of pair.

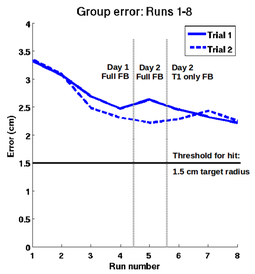


fMRI Pilot:

We plan to search for differences in brain activation across the following dimensions of interest:

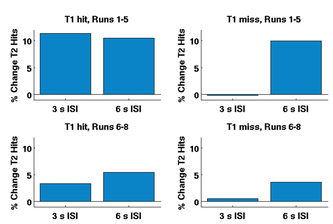
Pair success: Repeated success vs. error correction vs. uncorrected error
Error magnitude: Large vs. small error
Error sign: Too fast vs. too slow
Correction magnitude: Large vs. small correction
Correction sign: Speed up vs. slow down
Correction quality: Well-corrected vs. poorly corrected
Action selection: New target vs. repeated target

Behavioral pilot: Establish learning profile for paired trial shuffleboard task

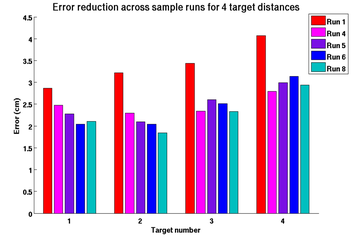


Effect of training:
* Error reduction in runs 1-4
* Improvement continues even when feedback is withheld for second trial of each pair.
* Consistent with improved goal selection over training.

Effect of withholding T2 feedback:
* Error reduction from Trial 1 to Trial 2 disappears.
* Could be due to performance asymptote, overcompensation, or attenuation of error-based learning, fatigue, or decay of T1-T2 correction model



Effect of ISI jitter:
* Examined two most frequent inter-stimulus intervals
  * Longer ISI confers an advantage.
* This advantage is most pronounced following T1 misses.
* Provides further evidence against strong contribution of error-based learning since error signals decay with time.



Effect of target distance:
* Learning is observed for all four distances.
* Continuous improvement across session is only seen for closest targets.

Experiment 1 parameters

2 sessions:
Day 1: 4 runs full feedback
Day 2: 1 run full feedback, 3 runs trial 1 only feedback

* 8 runs total, 112 trials per run, 3 second ISI
* 4 target distances, order randomized, no repeats
* Each target presented for 14 trial pairs per run
* 54 null trials randomly inserted in each run, allowed within or between trial pairs

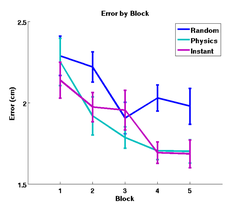
fMRI goal: Examine activation to trial 1 and trial 2 (within a pair) separately

SUMMARY

* While participants improved on the shuffleboard task, the evidence suggested only weak error-based learning.
* To amplify error-based learning, we will introduce perturbations of the feedback signal.

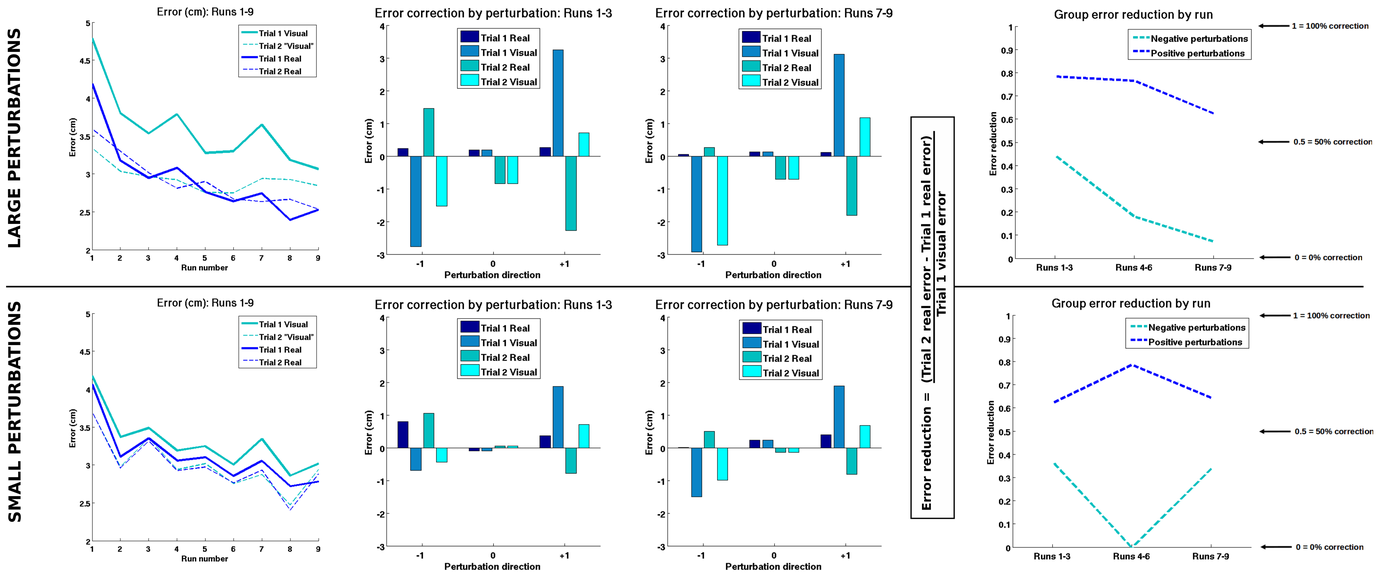
A note on feedback timing

We will use instantaneous feedback timing, which we previously found to produce equivalent performance to an ecologically realistic timing scheme ("Physics" timing).



fMRI pilot: Establish error-based learning in response to feedback perturbations

* Feedback was perturbed, randomly appearing at a location that was farther (**+**) or closer (**-**) than the actual produced distance, or was veridical (**0**).
* Two sizes of perturbations were used: **Large =** 1 sd of performance noise on Expt. 1 task. **Small =** 0.5 sd of performance noise on Expt. 1 task.



Effect of learning:
* Participants learn mean target speed over the course of the session, as evidenced by reduction in real error.

Evidence of error-based learning:
* Sign change from T1 Visual to T2 Real indicates responsiveness to perturbations.

Reduction in error correction:
* Both groups responded less to negative perturbations.
* The large perturbation group responded less to the perturbations over the course of the session, while the small perturbation group maintained their responsiveness.

Experiment 2 parameters

2 groups:
Group 1: 1 standard deviation perturbations
Group 2: 0.5 standard deviation perturbations

* Single session
* 9 runs total, 54 trials per run, 3 second ISI
* 3 target distances, order randomized, repeats allowed
* Transitions between targets balanced within each run
* Each target presented for 9 trial pairs per run
* 3 perturbations: -1, 0, +1
* Perturbation magnitude was target-specific
* Perturbation order was randomized
* Each perturbation was shown 3 times per target per run
* Perturbation size was determined by group average standard deviation of run 4, 1st trial of pairs only, from Experiment 1

fMRI goal: Examine signatures of error-based learning across trial pairs with high- and low-quality corrections

SUMMARY

* Participants learn the mean desired speed for each of three targets even as their feedback is perturbed.
* Participants demonstrate evidence of error-based learning: they produce directionally appropriate corrections in response to perturbations.
* Across the session, participants' response to large perturbations was reduced. This may indicate greater reliance on a learned (goal-based) action for the different targets.
* Across the session, participants' response to small perturbations fluctuated but was, on average, maintained. This may indicate that participants relied more on error-based learning in the face of perceived fluctuations.

Conclusions

Trained participants can produce meaningful error corrections even when they know feedback will be withheld.

Introducing perturbations in untrained participants yields similar signatures of error-based learning. This allows us to quantify error-based learning over the course of a session.

However, learning the mean target speed across the session may interfere with error-based learning within trial pairs, and vice versa.

Acknowledgment:
This work was supported by the National Science Foundation Graduate Research Fellowship Program and NIH R01 HD060306 (National Institute of Child Health and Human Development).