



Published in final edited form as:

Curr Biol. 2011 April 12; 21(7): 557–562. doi:10.1016/j.cub.2011.02.030.

Reward improves long-term retention of a motor memory through induction of offline memory gains

Mitsunari Abe^a, Heidi Schambra^a, Eric M Wassermann^b, Dave Luckenbaugh^c, Nicolas Schweighofer^d, and Leonardo G Cohen^a

^a Human Cortical Physiology Section and Stroke Neurorehabilitation Section, National Institutes of Health, Bethesda, MD 20892

^b Brain Stimulation Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892

^c Mood and Anxiety Disorders Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892

^d Department of Biokinesiology and Physical Therapy, University of Southern California, CHP 155 1540 Alcazar Street, Los Angeles, CA 90089-9006, USA

Summary

In humans, training in which good performance is rewarded or bad performance punished results in transient behavioral improvements [1–3]. Their relative effects on consolidation and long-term retention, critical behavioral stages for successful learning [4,5], are not known. Here, we investigated the effects of reward and punishment on these different stages of human motor skill learning. We studied healthy subjects who trained on a motor task under rewarded, punished, or neutral control conditions. Performance was tested before, and immediately, 6 hs, 24 hs and 30 days after training in the absence of reward or punishment. Performance improvements immediately after training were comparable in the three groups. At 6 hs, the rewarded group maintained performance gains while the other two groups experienced significant forgetting. At 24 hs, the reward group showed significant offline (posttraining) improvements while the other two groups did not. At 30 days, the rewarded group retained the gains identified at 24 hs, while the other two groups experienced significant forgetting. We conclude that training under rewarded conditions is more effective than training under punished or neutral conditions in eliciting lasting motor learning, an advantage driven by offline memory gains that persist over time.

Results

Previous studies have shown that formation and retention of motor memories are dynamic processes that evolve over multiple behavioral stages: online learning, consolidation and long-term retention [4–6]. Consolidation has been defined as reduced fragility of fresh memories during the initial hours after the training period or as spontaneous memory improvements [4,5,7], measured at 24 hs post practice [8]. Long-term retention of newly acquired memories allows to recall them without further practice after longer delays [4,5].

Correspondence should be addressed to Leonardo G Cohen (CohenL@ninds.nih.gov).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Reward and punishment have been investigated in relation to their influence on short-term learning in conditioning tasks by different authors [1,2,9–11]. It has been demonstrated that learning under conditions in which good performance is rewarded or bad performance punished can transiently improve formation of new associations between events in animal models [12,13]. In humans, their relative effectiveness in inducing consolidation and long-term retention of memories is not known.

Activity in dopaminergic neurons [14], fundamental for formation [15–17] and retention [18] of new motor memories is differentially modulated by reward and punishment. Neuronal excitability increases with reward and decreases with punishment [19]. Reward's strong reliance on dopaminergic neurotransmission [20,21] makes it a reasonable candidate to influence long-term retention of newly acquired motor memories [15,22]. Here, we hypothesized that learning under rewarding conditions would result in better long-term retention of a newly acquired memory than learning under punished or neutral conditions and that this advantage would be driven through improved consolidation.

38 right-handed healthy subjects learned a tracking isometric pinch force task (Fig 1) under the influence of monetary reward, punishment, or a neutral control condition (different groups, Fig 2). Subjects were instructed to pinch a force transducer between the right thumb and index finger in order to maintain a red cursor within a moving blue target on a computer screen (Fig 1). At the end of each training trial, subjects were given feedback according to their group: the rewarded group earned money based on the amount of time the red cursor stayed within the blue target, the punished group lost money based on the amount of time the cursor stayed outside the target, and the neutral group received neutral monetary information irrespective of performance. Subjects were told that they would start with \$0 and earn money for time on target (rewarded group), would start at \$72 and lose money for time off target (punished group), or simply receive \$40 at the end of the training session (neutral group). The monetary values were based on preliminary data, so that all groups would have a comparable amount of money at the end of training (actual amounts were $\$40.4 \pm 1.2$, 38.6 ± 2.2 and 40.0 ± 0.0 in the rewarded, punished and neutral groups respectively, see Supporting information 1, S1 for training results). Mean error (Fig 1c) was evaluated at test blocks (Fig 2). Thus, all test measurements (baseline, Immediate, 6 hs, 24 hs and 30 days) were done in the absence of any reward or punishment while actual training trials were carried out under the influence of reward, punishment or neutral information.

All three groups had similar mean errors at baseline (rewarded vs. neutral, $p = 0.86$ and rewarded vs. punished, $p = 0.91$, multiple pairwise comparisons with Bonferroni adjustments) and immediately after training (rewarded vs. neutral, $p = 0.77$ and rewarded vs. punished, $p = 0.23$, Fig 3). Learning, measured as mean error change between the baseline and the immediate post-training time points ($\Delta_{\text{immediate} - \text{baseline}}$), was similar across groups (10.3 ± 0.5 , 10.6 ± 0.5 and 9.8 ± 0.6 for neutral, rewarded and punished groups respectively; rewarded vs. neutral, $p = 0.78$, rewarded vs. punished, $p = 0.45$ and neutral vs. punished, $p = 0.56$), although training performance while feedback information was provided differed between groups (S1).

Retention at 6 hours post-training ($\Delta_{\text{6 hours} - \text{immediate}}$) was significantly larger in the rewarded group than in the neutral ($p = 0.02$) and punished ($p = 0.04$) groups (Fig 3). Within-group comparisons between the immediate and 6 hours post-training time points showed mean errors that remained stable in the rewarded group ($p = 0.87$) but increased in the neutral and punished groups (worsened performance, $p < 0.001$ and $p = 0.01$ respectively, Fig 3 and S2).

By 24 hours post-training $\Delta_{24 \text{ hours} - \text{immediate}}$, a common measure of overnight consolidation [8], was significantly larger in the rewarded than in the neutral ($p = 0.02$) and the punished ($p = 0.04$) groups (Fig 3). Within-group comparisons between immediate and 24 hours post-training time points showed decreased mean error in the rewarded group ($p < 0.001$) in the absence of differences in the neutral and punished groups ($p = 0.39$ and $p = 0.19$, respectively, S2) indicating successful overnight offline consolidation in the rewarded group. Whereas the punished and neutral groups showed decreased retention at 6 hours relative to the rewarded group, all groups decreased mean errors (improved performance) to similar extents between 6 and 24 hours (1.24 ± 0.4 , 1.00 ± 0.2 and 1.35 ± 0.4 for neutral, rewarded and punished conditions respectively; rewarded vs. neutral: $p = 0.35$ and rewarded vs. punished: $p = 0.22$).

Most importantly, by 30 days post-training $\Delta_{30 \text{ days} - \text{immediate}}$ remained larger in the rewarded group than in the neutral ($p = 0.01$) and punished ($p = 0.001$) groups (Fig 3). Within-group comparison showed decreased mean error (improved performance) in the rewarded group ($p = 0.02$), in contrast to the increased mean error (worsened performance) in the punished and neutral groups ($p < 0.001$ and $p = 0.003$ respectively, S2). This difference was better accounted for by a relatively stable error between 24 hours and 1 month time points in the rewarded group ($p = 0.31$), whereas errors increased in the punished ($p < 0.001$) and the neutral ($p < 0.001$) groups (S2). These results clearly documented better long-lasting retention of post-training gains in the rewarded group relative to the other two. As a result, the rewarded group had a significantly smaller mean error than the neutral ($p = 0.03$) and punished ($p = 0.002$) groups at 30 days (Fig 3 and S2). Mean error at each time point and delta mean error between time points were not significantly different between the punished and neutral groups ($p > 0.4$ for all comparisons).

Finally, time on target during testing showed results comparable to distance error (Methods). A repeated mixed model ANOVA with factors GROUP (neutral/rewarded/punished) and TIME (baseline/immediate/6hs/24hs/30days) showed a significant GROUP \times TIME interaction on mean time on target ($F = 3.59$, $p = 0.008$) with all three groups having comparable values at baseline (3.25 ± 0.18 , 3.47 ± 0.17 and 3.03 ± 0.20 for neutral, rewarded and punished, respectively) and immediately after training (5.79 ± 0.17 , 5.98 ± 0.19 and 5.53 ± 0.30 for control, rewarded and punished, respectively). Consistent with the measurement of error as cumulative distance away from the target described above, the mean time on target at 30 days was better in the rewarded group (6.33 ± 0.15) than in the neutral (5.22 ± 0.24 , $p = 0.022$) or the punished (5.17 ± 0.20 , $p = 0.009$) groups.

Discussion

In summary, we found that training under rewarded conditions elicited substantial long-term retention of a newly acquired memory while training under punished or neutral conditions did not and that this advantage developed through stabilization of offline memory gains in subsequent days.

Under our experimental conditions, all three groups improved significantly although to different extents during rewarded, punished or neutral training. Immediately after training, when testing was carried out in the absence of any reward or punishment, all groups showed comparable and marked learning. Memory changes after completion of training could involve stabilization (consistent performance over time), offline gains (performance improvements beyond stabilization) [4,5,7], often referred to as consolidation [7], or offline forgetting (performance worsening over time)[6]. We found that at 6 hours, mean errors increased in the punished and neutral groups, reflecting offline forgetting of memory but

remained stable in the rewarded group. These findings indicate substantive differences in the strength of the motor memory during the initial hours of the consolidation period depending on training type with stabilization of memory gains present only in the rewarded group.

Consolidation at 24 hours [8], measured as the difference in performance between the immediate and 24-hour post-training time points, was larger in the rewarded than in the punished or neutral groups. Within group analysis demonstrated offline improvements only in the reward group, which remained present 30 days later. In contrast, the punished and neutral groups did not have significant offline gains and by 30 days showed substantial performance loss (Fig 3).

These results represent to our knowledge the first demonstration of a benefit of reward on long-term retention of a motor memory in animals or humans. Long-term retention is important because it impacts our ability to maintain an acquired memory over time without the need to relearn it each time memory retrieval is required [4]. Learning under the rewarded condition induced significant offline memory gains, while learning under the punished and the neutral conditions resulted in the opposite effect: offline memory losses (forgetting) (Fig 3). Therefore, training under reward not only compensated for the offline forgetting seen in the punished and neutral groups by 30 days, but also resulted in a reversal from offline forgetting to lasting offline learning (Fig 4).

To explore the possibility that performance during training influenced long-term retention we calculated single exponential fits of the training data on the decay parameter and then looked for correlations with retention at 30 days. We found no significant correlations between individual subjects' decay parameters and retention at 30 days for any of the groups ($p = 0.27$, $p = 0.23$, $p = 0.35$ for neutral, rewarded and punished groups respectively) or for all subjects together ($p = 0.52$). Across groups, retention at 30 days was not predicted by mean error in the last 10 training trials either (S3). Thus, we found no evidence that the decay parameter during training or training performance in the last 10 trials predicted long-term retention, an issue plausible to experimental testing in the future.

Reward is associated with increased dopaminergic function [23] in the midbrain [11] and striatum [17], which is influential on memory retention in humans [18], possibly through D1/D5-dopamine-dependent long-term potentiation (LTP) [17,24,25]. It is conceivable that dopaminergic neurotransmission could represent a common mechanistic link underlying the synergistic effects of training and reward on long-term retention of motor memories [15]. Dopamine-dependent LTP develops gradually over hours [24] and persists for days to weeks [26], a time course similar to that of the developing reward benefits in our study. It operates in cortico-striatal loops [17,25], which are engaged in motor memory formation [15] and is activated by motor training [4,5] and reward protocols [23,27]. It is possible that the facilitatory effect of reward on long-term motor memory retention reveals an underpinning of D1/D5-dopamine-dependent LTP-like mechanisms [15,17,25] at the intersection of networks that mediate motor learning and reward processing [15], as proposed recently in relation to episodic memory [28] and habit formation [29]. Training under punishment did not significantly modify memory formation stages relative to the neutral group, an effect that could be accounted for by its depressing influence on dopamine-dependent LTP-like mechanisms [24] or/and its predominant reliance on activity in serotonergic pathways [30,31] that are not part of the network mediating consolidation and long-term retention of motor memories, consistent with results reported in a procedural learning paradigm [32].

We conclude that training under rewarded conditions is more effective than training under punished or neutral conditions in inducing long-term retention of newly learned memories. Understanding the learning stages influenced by reward, driven through reduction in

degradation of the fresh memory and induction of persistent offline memory gains, may influence the design of practice protocols in education as well as the treatment of memory disorders and rehabilitation of function after brain lesions.

Experimental procedures

Subjects—Forty-one young adults (24.3 ± 5.2 years, mean \pm S.D., 18 females) were enrolled in this study. All subjects were recruited at the laboratory of the Human Cortical Physiology and Stroke Neurorehabilitation Section, NINDS, National Institutes of Health. All participants were right-handed as assessed by the Edinburgh Handedness Inventory, had no abnormal physical or neurological findings, had no past history of neurological or psychiatric diseases, and did not take chronic medications. All subjects gave written informed consent to participate in the study before the experiment. The study was approved by the Ethics Committee of the National Institute of Neurological Disorders and Stroke. We excluded three subjects from the analysis, because their baseline performances were two standard deviations beyond the mean baseline performance of all subjects. Thus, 38 subject's data were used for data analysis (rewarded training, $n = 13$; punished training, $n = 12$; control training, $n = 13$).

Tracking pinch force task—Seated subjects pinched a force transducer between the right thumb pad and lateral middle phalanx of the index finger (Fig. 1a), which controlled the vertical movements of a red cursor (0.6 cm^2). Subjects were asked to modulate their pinch force to keep the red cursor in the blue target (1.5 cm^2). The blue target moved in a sequential pattern along a single vertical axis for 9 s during each trial (Fig. 1b). The force required to reach the target increased logarithmically with the vertical displacement. Error was defined as the vertical distance between the edges of the blue target and the red cursor at each sampled time point, as shown in Fig 1c.

Testing paradigm—Subjects were randomly allocated to the rewarded, punished, or neutral control training groups (Fig. 2). Each group practiced the same task over one session (80 trials total). After the end of every trial, the red cursor and the blue box disappeared for 0.5 s. All subjects received visual monetary reward / monetary punishment / neutral information for 1 s specific to their training group (Fig. 2). The range of the positive or negative monetary outcomes was + \$ 0.00 to 0.80 or - \$ 0.80 to 0.00 \$ per trial, respectively. The neutral information in the control group consisted of a sequence of characters (“#####”). Unbeknownst to the subjects, all groups ultimately earned a comparable amount of money (see main text). Monetary reward or punishment depended on the amount of time during which the subjects kept the red cursor in the blue target per trial, a measurement tightly correlated with mean error (S4).

Data Analysis—For all outcome measures, assumption of a normal distribution (Shapiro-Wilk test of normality) and homogeneity of variance (Mauchly's Sphericity Test) were verified. Multiple pairwise comparisons with Bonferroni adjustments were performed to compare delta (mean error change) across groups at each delta time point. A repeated measures mixed-model ANOVA with factors GROUP (rewarded/punished/neutral) and TIME (baseline/immediate/6 hours/24 hours/30 days) on mean error was performed. Mean error was also compared with Bonferroni adjustments across groups at each level of time (baseline, or the immediate or the 30 days post-training) and compared between the immediate and the other post-training time points within each group. Time-series of single trial error during training was mathematically modeled by a single exponential decay function: $\text{Error}(t) = a \cdot \exp(-b \cdot t) + c$, where t indicates the number of trials [33]. The decay parameter (term b) for each individual was calculated and then possible correlations between these decay parameters and retention at 30 days were tested for each group and for all

subjects lumped together. Secondly, time on target (see *Testing paradigm*) was also computed during testing and then analyzed with a repeated measures mixed model ANOVA, followed by multiple pairwise comparisons with Bonferroni adjustments, with the same factorial design and post-hoc comparisons across groups as the ones when mean error was tested. All the analyses were done in SPSS 17.0 (SPSS Inc., Chicago, IL). Significance level was set to $p < 0.05$. All data are reported as mean \pm SEM.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke (NINDS, NIH). We thank Drs Toyomi Abe, Nitzan Censor, Mark Hallett, Eran Dayan, Ethan Buch, John Krakauer and Pablo Celnik for critical comments on earlier versions of this manuscript.

References

- O'Doherty JP. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol.* 2004; 14:769–776. [PubMed: 15582382]
- Kim H, Shimojo S, O'Doherty JP. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* 2006; 4:e233. [PubMed: 16802856]
- Seymour B, Singer T, Dolan R. The neurobiology of punishment. *Nat Rev Neurosci.* 2007; 8:300–311. [PubMed: 17375042]
- Doyon J, Penhune V, Ungerleider LG. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia.* 2003; 41:252–262. [PubMed: 12457751]
- Doyon J, Benali H. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol.* 2005; 15:161–167. [PubMed: 15831397]
- Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci U S A.* 2009; 106:1590–1595. [PubMed: 19164589]
- Robertson EM, Pascual-Leone A, Miall RC. Current concepts in procedural consolidation. *Nat Rev Neurosci.* 2004; 5:576–582. [PubMed: 15208699]
- Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG. Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proc Natl Acad Sci U S A.* 2002; 99:1017–1022. [PubMed: 11805340]
- Hull, CL. *Principles of behaviors: An introduction to behavior theory.* New York: Appleton-Century; 1943.
- Ferster, CB.; Skinner, BF. *Schedules of reinforcement.* New York: Appleton Century Croft; 1957.
- Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol.* 1994; 72:1024–1027. [PubMed: 7983508]
- Tempel BL, Bonini N, Dawson DR, Quinn WG. Reward learning in normal and mutant *Drosophila*. *Proc Natl Acad Sci U S A.* 1983; 80:1482–1486. [PubMed: 6572401]
- Nakatani Y, Matsumoto Y, Mori Y, Hirashima D, Nishino H, Arikawa K, Mizunami M. Why the carrot is more effective than the stick: different dynamics of punishment memory and reward memory and its possible biological basis. *Neurobiol Learn Mem.* 2009; 92:370–380. [PubMed: 19435611]
- Schultz W. Getting formal with dopamine and reward. *Neuron.* 2002; 36:241–263. [PubMed: 12383780]
- Wickens JR, Reynolds JN, Hyland BI. Neural mechanisms of reward-related motor learning. *Curr Opin Neurobiol.* 2003; 13:685–690. [PubMed: 14662369]

16. Rossato JI, Bevilacqua LR, Izquierdo I, Medina JH, Cammarota M. Dopamine controls persistence of long-term memory storage. *Science*. 2009; 325:1017–1020. [PubMed: 19696353]
17. Calabresi P, Picconi B, Tozzi A, Di Filippo M. Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci*. 2007; 30:211–219. [PubMed: 17367873]
18. Floel A, Garraux G, Xu B, Breitenstein C, Knecht S, Herscovitch P, Cohen LG. Levodopa increases memory encoding and dopamine release in the striatum in the elderly. *Neurobiol Aging*. 2008; 29:267–279. [PubMed: 17098331]
19. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
20. Zald DH, Boileau I, El-Dearedy W, Gunn R, McGlone F, Dichter GS, Dagher A. Dopamine transmission in the human striatum during monetary reward tasks. *J Neurosci*. 2004; 24:4105–4112. [PubMed: 15115805]
21. Hakyemez HS, Dagher A, Smith SD, Zald DH. Striatal dopamine transmission in healthy humans during a passive monetary reward task. *Neuroimage*. 2008; 39:2058–2065. [PubMed: 18063390]
22. Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. *Nature*. 2001; 413:67–70. [PubMed: 11544526]
23. Gaspar P, Stepniewska I, Kaas JH. Topography and collateralization of the dopaminergic projections to motor and lateral prefrontal cortex in owl monkeys. *J Comp Neurol*. 1992; 325:1–21. [PubMed: 1362430]
24. Huang YY, Kandel ER. D1/D5 receptor agonists induce a protein synthesis-dependent late potentiation in the CA1 region of the hippocampus. *Proc Natl Acad Sci U S A*. 1995; 92:2446–2450. [PubMed: 7708662]
25. Jay TM. Dopamine: a potential substrate for synaptic plasticity and memory mechanisms. *Prog Neurobiol*. 2003; 69:375–390. [PubMed: 12880632]
26. Abraham WC. How long will long-term potentiation last? *Philos Trans R Soc Lond B Biol Sci*. 2003; 358:735–744. [PubMed: 12740120]
27. Kapogiannis D, Champion P, Grafman J, Wassermann EM. Reward-related activity in the human motor cortex. *Eur J Neurosci*. 2008; 27:1836–1842. [PubMed: 18371077]
28. Shohamy D, Adcock RA. Dopamine and adaptive memory. *Trends Cogn Sci*. 2010; 14:464–472. [PubMed: 20829095]
29. Turchi J, Devan B, Yin P, Sigrist E, Mishkin M. Pharmacological evidence that both cognitive memory and habit formation contribute to within-session learning of concurrent visual discriminations. *Neuropsychologia*. 2010; 48:2245–2250. [PubMed: 20144631]
30. Ogren SO, Eriksson TM, Elvander-Tottie E, D’Addario C, Ekstrom JC, Svenningsson P, Meister B, Kehr J, Stiedl O. The role of 5-HT(1A) receptors in learning and memory. *Behav Brain Res*. 2008; 195:54–77. [PubMed: 18394726]
31. Dai JX, Han HL, Tian M, Cao J, Xiu JB, Song NN, Huang Y, Xu TL, Ding YQ, Xu L. Enhanced contextual fear memory in central serotonin-deficient mice. *Proc Natl Acad Sci U S A*. 2008; 105:11981–11986. [PubMed: 18695238]
32. Wachter T, Lungu OV, Liu T, Willingham DT, Ashe J. Differential effect of reward and punishment on procedural learning. *J Neurosci*. 2009; 29:436–443. [PubMed: 19144843]
33. Lang CE, Bastian AJ. Cerebellar subjects show impaired adaptation of anticipatory EMG during catching. *J Neurophysiol*. 1999; 82:2108–2119. [PubMed: 10561391]
34. Floyer-Lea A, Matthews PM. Changing brain networks for visuomotor control with increased movement automaticity. *J Neurophysiol*. 2004; 92:2405–2412. [PubMed: 15381748]
35. Ramnani N, Elliott R, Athwal BS, Passingham RE. Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage*. 2004; 23:777–786. [PubMed: 15528079]
36. Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci*. 2004; 7:887–893. [PubMed: 15235607]
37. Izawa J, Rane T, Donchin O, Shadmehr R. Motor adaptation as a process of reoptimization. *J Neurosci*. 2008; 28:2883–2891. [PubMed: 18337419]

38. Shadmehr R, Orban de Xivry JJ, Xu-Wilson M, Shih TY. Temporal discounting of reward and the cost of time in motor control. *J Neurosci*. 2010; 30:10507–10516. [PubMed: 20685993]
39. Walker MP, Stickgold R, Alsop D, Gaab N, Schlaug G. Sleep-dependent motor memory plasticity in the human brain. *Neuroscience*. 2005; 133:911–917. [PubMed: 15964485]
40. Shank SS, Margoliash D. Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature*. 2009; 458:73–77. [PubMed: 19079238]
41. Penhune VB, Doyon J. Dynamic cortical and subcortical networks in learning and delayed recall of timed motor sequences. *J Neurosci*. 2002; 22:1397–1406. [PubMed: 11850466]

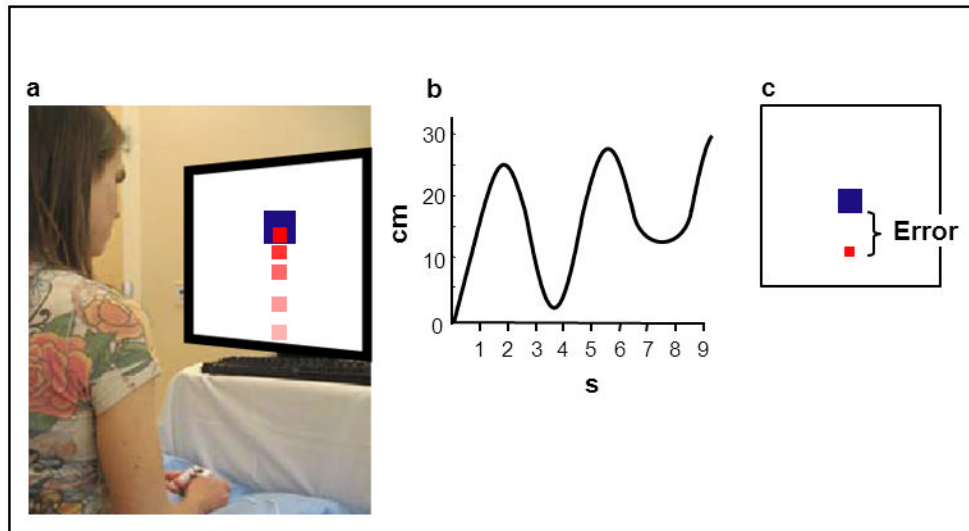


Figure 1. behavioral task. (a) Tracking isometric pinch force task. Subjects pinched a force transducer between the right thumb and index finger. Squeezing the force transducer resulted in the upward movement of a red cursor on the computer screen, while relaxing caused the cursor to move downward (cursor movements with force shown in shades of red on the monitor). The goal was to maintain the red cursor within the moving blue target by modulating pinch force. Each trial started with the red cursor and the blue target overlapping at the bottom of the screen. (b) Target path. During each trial, the blue target moved in a vertical line for 9 seconds along a consistently repeated trajectory to disappear for 0.5 s at the end of the trial. The y-axis displays the vertical distance (cm) from the lower edge of the blue box and the bottom of the screen. The x-axis shows the elapsed time (s). Intertrial intervals varied randomly between 1 and 2 seconds. (c) Mean error calculation. Error was defined as the vertical distance between the edges of the blue target and the red cursor at each sampled time point, as shown. Each single-trial error was calculated by averaging the errors along all sampled time points in the trajectory (sampled every 20 ms for 9 s). Mean error was calculated as the mean across all 10 trials within each test block.

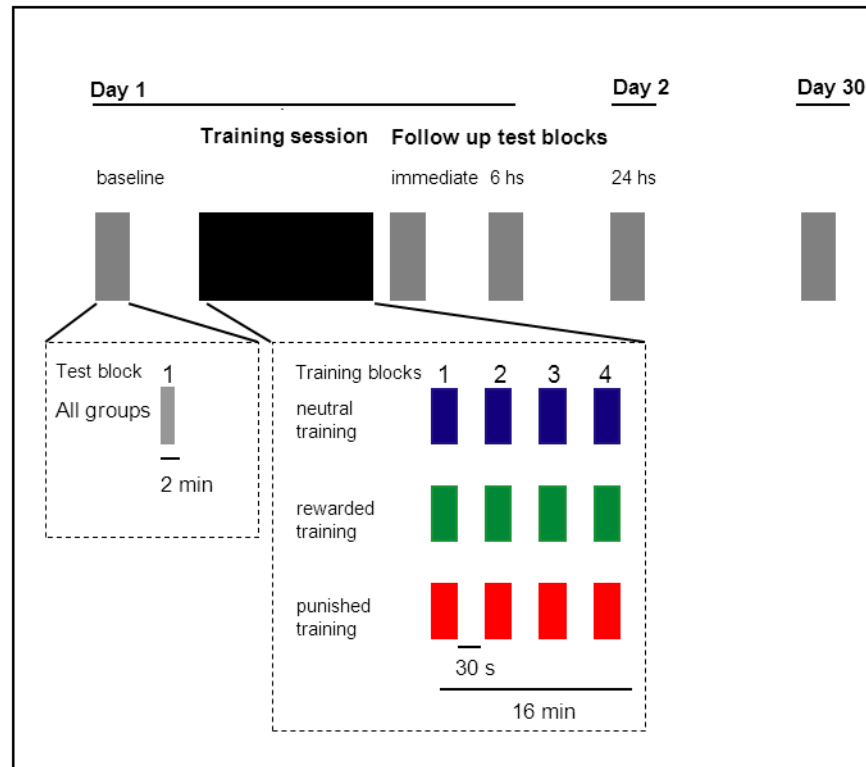


Figure 2.

Experimental design. Subjects participated in 3 different sessions (days 1, 2 and 30) separated into 3 training groups who practiced the task over 4 blocks (20 trials each, black rectangle) under the influence of monetary reward (green, $n = 13$), monetary punishment (red, $n = 12$) or neutral conditions (blue, $n = 12$) in a factorial design. During training blocks, monetary reward, monetary punishment, or neutral visual information was presented for 1 s at the end of each trial depending on the group. Test blocks were evaluated in all subjects in the absence of reward/punishment/neutral information before training (baseline), and immediately, 6, 24 hs and 30 days after training.

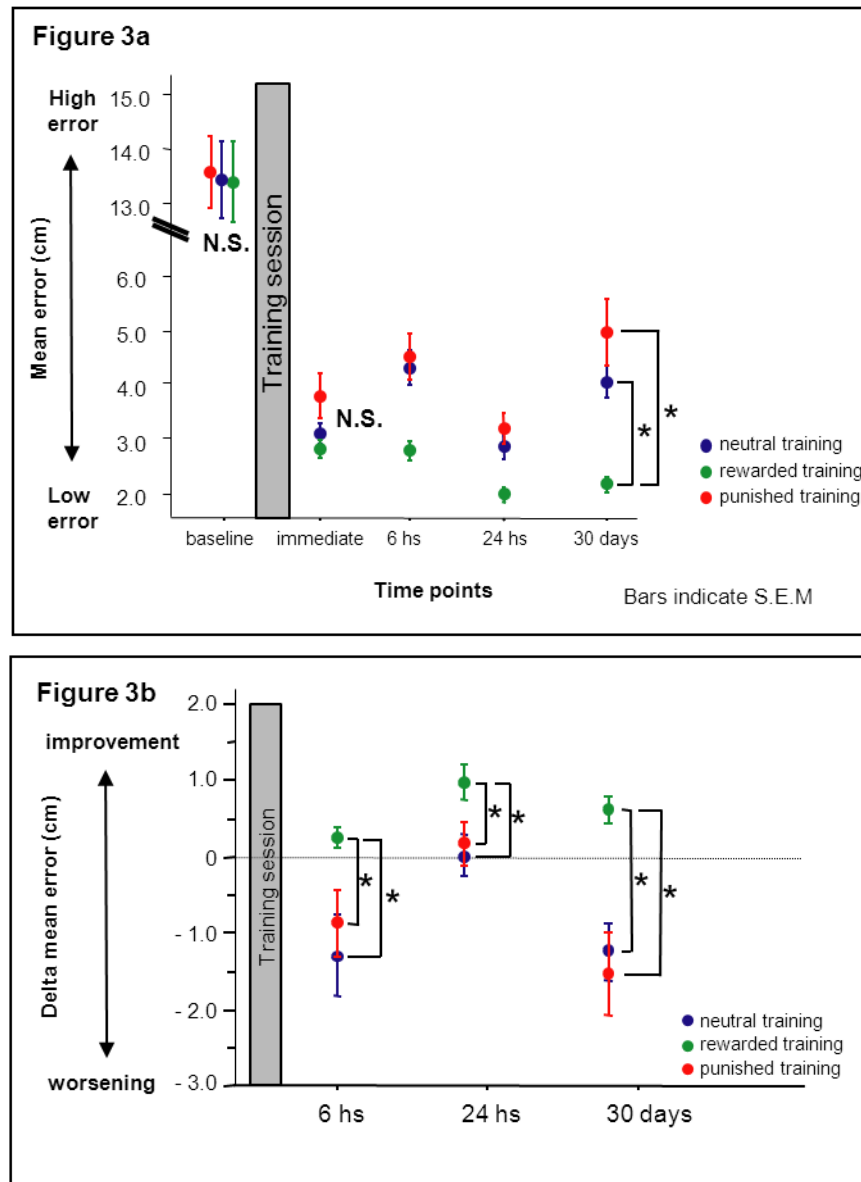


Figure 3.

(a) Effect of reward and punishment on motor skill. Mean errors in the rewarded, punished and neutral groups as a function of time. A repeated measures mixed-model ANOVA with factors GROUP (rewarded/punished/neutral) and TIME (baseline/immediate/6 hours/24 hours/30 days) on mean error showed a significant effect of TIME ($F(4, 140) = 403.1, p < 0.001$) and a trend for GROUP ($F(2, 35) = 2.72, p = 0.08$). Most importantly, there was a significant GROUP x TIME interaction ($F(8, 140) = 4.41, p < 0.001$), indicating a different time course of performance changes across groups. At baseline and immediately after training, all groups had similar mean errors. Of note are the comparable mean errors immediately after training and the significantly smaller mean error at 30 days post-training in the rewarded relative to the neutral and punished groups, a difference that started to develop as early as 6 hours post-training.

(b) Effect of reward and punishment on motor skill retention after training. Changes in mean error (delta) between the immediate and the other (6 hs, 24 hs or 30 days) post-training time

points. Values above 0 indicate decrease in mean error (improved performance), while those below 0 indicate increase in mean error (worsened performance). Of note is the significantly greater delta at 30 days post-training in the rewarded relative to the neutral and punished groups, a difference that started to develop as early as 6 hours post training. Data shown as $\text{delta} \pm \text{SEM}$. Asterisks indicate corrected $p < 0.05$.

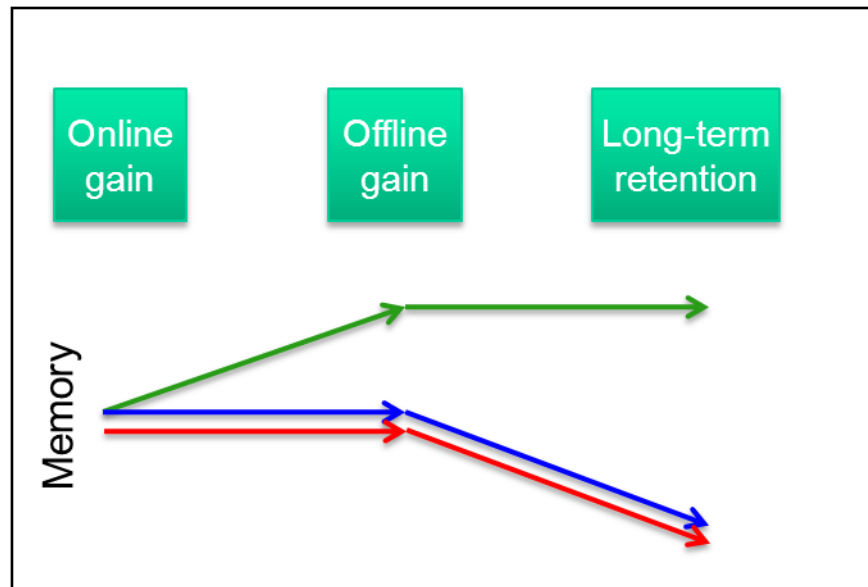


Figure 4.

Time course of memory changes. Online gains were comparable in the three groups. While the reward group (green) experienced substantial offline memory gains, the other two groups did not. By 30 days memory in the rewarded group stabilized offline gains while in the other two groups deteriorated. The neural structures mediating different stages of rewarded learning remain to be determined firmly but may include the cerebellum, neocortex and striatum in the intersection of networks that provide substrate for learning of this task [34] and processing of reward information [32,35,36]. It remains to be determined to which extent these observations apply to motor adaptation paradigms [37,38]. It is conceivable that the cerebellum could contribute to error-based learning while the striatum and neocortex may become engaged in later stages [39,40] and long-term retention [41] under rewarded conditions.