



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2013

Predictive value and reward in implicit classification learning

Lam, J M ; Wächter, T ; Globas, C ; Karnath, H O ; Luft, A R

DOI: <https://doi.org/10.1002/hbm.21431>

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-66874>
Journal Article

Originally published at:

Lam, J M; Wächter, T; Globas, C; Karnath, H O; Luft, A R (2013). Predictive value and reward in implicit classification learning. *Human Brain Mapping*, 34(1):176-185.

DOI: <https://doi.org/10.1002/hbm.21431>

Predictive value and reward in implicit classification learning

Judith M. Lam ^{a,b,*}, Tobias Wächter ^{b,c,*}, Christoph Globas ^{a,d}, Hans-Otto Karnath ^e, Andreas R. Luft ^a

^a Clinical Neurorehabilitation and Rehabilitation Initiative/Technology Center of Zurich, Department of Neurology, University of Zürich, Zürich, Switzerland

^b Center of Neurology and Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

^c German Center of Neurodegenerative Diseases (DZNE), Tübingen, Germany

^d Robert-Bosch-Krankenhaus, Klinik für geriatrische Rehabilitation, Stuttgart, Germany

^e Section of Neuropsychology, Center of Neurology, Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

*Authors contributed equally.

Abbreviated title: Predictive value and reward in classification learning

Correspondence:

Andreas R. Luft, MD, Department of Neurology, University of Zürich,

Frauenklinikstrasse 26, 8091 Zürich, Switzerland, Tel: +41 44 255 5400, Fax: +41 44

255 4649, andreas.luft@usz.ch

Abstract

Learning efficacy depends on its emotional context. The contents learned and the feedback received during training tinges this context. The objective here was to investigate the influence of content and feedback on the efficacy of implicit learning and to explore using functional imaging how these factors are processed in the brain. Twenty-one participants completed 150 trials of a probabilistic classification task (predicting sun or rain based on combinations of playing cards). Smileys or frowneys were presented as feedback. In 10 of these subjects, the task was performed during functional magnetic resonance imaging. Card combinations predicting sun were remembered better than those predicting rain. Similarly, positive feedback fortified learning more than negative feedback. The presentation of smileys recruited bilateral nucleus accumbens, sensorimotor cortex and posterior cingulum more than negative feedback did. The higher the predictive value of a card combination, the more activation was found in the lateral cerebellum. Both context and feedback influence implicit classification learning. Similar to motor skill acquisition, positive feedback during classification learning is processed in part within the sensorimotor cortex, potentially reflecting the activation of a dopaminergic projection to motor cortex (Hosp et al. J Neuroscience, 2011). Activation of the lateral cerebellum during learning of combinations with high predictive value may reflect the formation of an internal model.

Keywords: Sensorimotor cortex, nucleus accumbens, cerebellum, reward, classification learning, implicit learning, fMRI

Introduction

Learning depends on the emotional context in which it occurs (Cahill & McGaugh 1996). Pleasant cues with positive valence are better remembered than unpleasant ones with negative valence (Ali & Cimino 1997; Mneimne et al 2010) and the opposite may be true for directed forgetting (Minnema & Knowlton 2008). Feedback in the form of reward or punishment provided during trial-and-error learning also tinges the emotional context. Wächter and coworkers (2009) showed that implicit motor skill learning is more effective with positive than negative feedback. Brain networks involved in feedback processing include striatum, midbrain, amygdala, frontal and cingulate cortices; positive and negative feedback are handled by different circuits (Liu et al 2011).

The hypothesis here was that positive feedback, content of positive valence and high predictive value improve implicit learning and that improved learning is associated with stronger recruitment of brain networks encoding rewards. Because implicit learning mechanisms form the basis of many therapeutic interventions in rehabilitation, it is important to know the effect of these modifiable factors. We recruited subjects from middle to retirement age for a later comparison with individuals after stroke.

We tested our hypothesis using a classification learning paradigm, the weather prediction task (Knowlton et al 1996), in which subjects had to learn associations between a certain combination of four different playing cards and a dichotomous weather outcome, sun or rain. The associations were stochastic, that is, each combination of cards predicted sun or rain with a certain probability. The subject was supposed to learn which combination predicted which weather. Feedback was given in form of smiley or frowney faces. The stochastic nature limited the subject's

awareness of the association. Hence, learning this task was considered to be mainly implicit.

Materials and Methods

Subjects and Task

Twenty-one subjects were recruited via advertisements. Participants (14 females, 7 males) were between 43 and 85 years old (mean \pm SEM: 64.6 ± 2.1). Inclusion criteria were a Mini Mental State (MMS) ≥ 27 points and Beck's Depression Index (BDI) ≤ 11 points. Mean MMS was 29.5 ± 0.2 , SEM, and BDI was 5.1 ± 0.8 . Education quantified by the number of years spent in primary and secondary school was 11.7 ± 0.3 years. Ten of the 21 subjects qualified for (absence of claustrophobia and metal implants, 6 females, 4 males, age 64.5 ± 3.1 years, mean \pm SEM) and agreed to undergo fMRI testing. The sample was recruited as a control group for a later comparison to stroke survivors. The study was approved by the Ethics Committee of the University of Tübingen, Germany. All participants provided written informed consent.

The weather prediction task (WPT) was performed as described by Knowlton et al. (1996). The task is a forced-choice classification task with two alternative responses in which participants learn probabilistic associations between 14 different combinations of four playing cards (**Figure 1**) and two weather outcomes, sun and rain. Each card was linked to an outcome with a pre-specified probability (for sun: card 1 - 80%, card 2 - 54%, card 3 - 43%, card 4 - 20%). For each trial, either one, two or three cards were shown composing 14 combinations that predicted the weather each with a certain combined probability. **Table 1** shows for each combination of cards the probability and how often the combination was shown (as a fraction of 150 trials). Presentation of combinations of cards, detection of button-press responses and feedback were computer-controlled using Matlab (Mathworks Inc, Natick, MA, USA) and Psychtoolbox (www.psychtoolbox.org). The WPT was

verbally explained and demonstrated before the experiment. Participants were instructed not to talk with the investigator during the experiment. After presentation of a card combination, the subject had to respond within 4 seconds or the trial was scored as “incorrect”. After 3 seconds, a prompt („Please press a button“) appeared on the screen. After pressing either the „sun“ or the „rain“ button, feedback was shown for 2 sec in form of a smiley or a frowney face. Every 50 trials, a one-minute break was allowed. The experiment continued until 150 trials were completed.

[Figure 1, Table 1 approximately here]

Analysis of behavioral data

Trials were considered as “correct” when subjects chose the more probable weather (sun or rain) for the card combination presented. Trials in which subjects did not respond were scored as “incorrect”. Missed responses were few and an alternative analysis that excluded those responses yielded results similar to the ones reported below.

To investigate the time course of learning, a learning curve (performance over time) was constructed for each subject using a two-step procedure. First, a cumulative performance curve was computed by adding 1 for each „correct“ and subtracting 1 for each „incorrect“ trial. This curve was then smoothed using spline interpolation (Matlab’s `spapi` function, 2 knots). Second, to convert the cumulative into a performance-over-time curve the first derivative was computed (`fnder` function). The resulting curve showed that performance increased in a non-linear fashion over the course of training. Non-linear fitting of an exponential function [Boltzmann function, $p/(1+\exp(k*(a-x)))$] was used to derive parameters of learning: the plateau p , the turning point a of the sigmoid Boltzmann curve and the steepness in the turning point k . The plateau values were estimated for each card combination. A general linear

model were used to explore whether the variability of the plateau was explained by sun versus rain, the combinations' predictive probability and its presentation frequency. Predictive probability was classified as high, medium or low (**Table 1**). This stratification was done because some combinations were presented less frequently than others and by grouping we obtained prediction classes of approximately equal frequency. Whether or not a subject belonged to the subgroup receiving fMRI or not was included as an additional independent dichotomous variable to test for systematic differences between the samples.

In a second analysis, we measured how well single card combinations were remembered. We counted the number of trials in which the subject responded identical to a preceding trial with the same card combination and a smiley reward. The two trials could have been subsequent or several trials apart. Trials with the same response after a rewarded (smiley) trial will be referred to as „same-after-smiley“ trials (SAS), otherwise they will be termed „opposite-after-smiley“ (OAS). Conversely to examine, if subjects remembered to change their response behavior after seeing a frowney, we counted „opposite-after-frowney“ (OAF) and „same-after-frowney“ trials. The ratio of (SAS+OAF) / all trials was then used as an index of memory. Because memory improved during training, only the last 60 trials of 150 were considered to compute this index. Using the index as a dependent variable, we tested for effects of combination (sun versus rain), predictive value (high-medium-low, **Table 1**) and feedback. The independent variable feedback was defined as the number of smileys - frowneys that a subject saw during the initial 30 trials of training. It was assumed that no relevant memory was formed during these initial 30 trials, hence, there was no bias towards smileys because some combinations had already been memorized. In fact frowneys were slightly but significantly more frequent during

the initial 30 trials (1.9 versus 2.8, $p=0.045$) excluding a bias towards smileys. Whether or not a subject belonged to the subgroup receiving fMRI or not was included as an additional independent dichotomous variable to test for systematic differences between the samples.

JMP (version 8, SAS Institute Inc, Cary, NC, USA) was used for statistical calculations.

Functional magnetic resonance imaging (fMRI)

Using a 3 Tesla scanner (Trio-Tim with 8-channel phased-array head coil, Siemens, Erlangen, Germany) fMRI was performed in subjects without metal implants or claustrophobia who agreed to participate. Visual cues were presented via a projection system installed in the scanner room; responses were collected using an MRI-compatible button-box. All participants responded using their right hand.

The WPT task was performed in participants naïve to this task as described above, except that the intertrial interval was 5 seconds, subjects had to respond within 4 seconds and did not receive the second prompt („Please press a button“). Additionally, a control task was performed before and after the WPT to record brain activity related to visual processing and movement similar to the WPT. In the control task one, two or three cards were shown and subjects were asked to respond with the right button when two cards were presented and the left button when one or three cards were shown. Thirty training trials of the WPT were performed outside the scanner without feedback stimuli to avoid learning before the actual experiment was started. Brain activity during WPT was measured in three blocks of 50 trials each separated by 30 seconds of fixation. Fifty trials of the control task were performed before the WPT.

A high-resolution T1-weighted scan was acquired for anatomical localization. Functional imaging used gradient-echo planar T2*-weighted images (EPI) with blood oxygenation level (BOLD)-contrast (TR=2.4 s, TE=30 ms, flip angle=90°). Thirty-eight slices (slice thickness 3 mm) were acquired to cover the entire brain.

fMRI Analysis

fMRI data were processed using Brainvoyager QX (version 2.2, Brain Innovation BV, Maastricht, The Netherlands). BOLD-weighted EPI datasets were corrected for slice acquisition timing and head motion. Motion correction parameters were used as confound predictors in first-level GLM analyses. Datasets were registered to Talairach space in correspondence to the anatomical dataset. Images were spatially (Gaussian kernel, full-width at half-maximum of 8 mm) and temporally (3 cycles, GLM-Fourier-high-pass-filter) smoothed.

The statistical analysis modeled each trial as two events, one before (*presentation period*) and one after the button press (*feedback period*). Five general linear models were computed:

1. In the first, the hemodynamic response was estimated for each of the following conditions, control trial presentation, control trial feedback, WPT trial presentation, WPT trial feedback. Random effects second-level analysis of variance (ANOVA) was used to construct WPT versus control activation maps for the presentation and feedback periods.
2. The second model was constructed analogous to model 1 except for replacing the control trials with chance trials (combination with no predictive value).
3. In the third model, the WPT trials were separated according to whether the card combination predicted sun or rain, and according to the predictive value of the combination (prediction class, **Table 1**). Random effects ANOVA was used to

extract activation maps for the effects of sun/rain, predictive value and their interaction (F-tests). Contrasts (t-tests) were computed for high > low predictive value and sun > rain. This second model was computed for presentation and feedback periods separately.

4. A fourth model was computed separating trials in which a smiley and a frowney feedback was received. Random effects ANOVA was then used to extract the activation map for smiley > frowney. Only the feedback period of the trial was included in this model.
5. In a fifth random effects ANOVA model we investigated a potential interaction between sun/rain and smiley/frowney as independent variables. This model did not yield any results and is therefore not further mentioned.

To test for effects of age, the subject sample was split according to the median age and the age group was included as a between-subject predictor in all models. For all random effects models the statistical threshold was set $p < 0.05$ corrected for multiple comparisons using a false discovery rate (FDR) method. Talairach coordinates and average p-values were measured for each activation cluster equal or larger than 10 voxels (10 x 3mm x 3mm x 3mm).

Results

Learning depends on repetition and predictive value of the combination and on the content of the association being learned (sun versus rain)

During the task, performance (correct responses over time) increased in a sigmoid fashion reaching a plateau between trials 70 and 80 (**Fig. 2A**). For card combinations predicting sun plateau performance was higher than for card combinations predicting rain (general linear model, effect of the dichotomous variable sun/rain: $p=0.039$, **Fig. 2B**) indicating that combinations predicting sun were better learned than those predicting rain. Plateau performance was also higher for combinations occurring more frequently (general linear model, effect of frequency: $p<0.0001$, $r=0.97$, **Fig. 2C**) and for those with higher predictive value, i.e., a larger difference from chance probability to predict either sun or rain (general linear model, effect of probability: $r=0.84$, $p<0.0001$, **Fig. 2D**). The latter effect, however, did not remain significant if *frequency* was included in the model. Whether or not the subject was in the fMRI group had no significant effect.

[Figure 2 approximately here]

Learning depends on feedback

To evaluate learning from positive feedback, we counted how often a subject responded identical to a card combination that had occurred before and was rewarded with a smiley („same-after-smiley“, SAS). Vice versa, to evaluate learning from negative feedback we counted „opposite-after-frowney“ (OAF) trials. SAS trials were significantly more frequent than OAF trials (paired t-test, $p<0.001$, **Fig. 3a**) indicating that card combinations leading to positive feedback were remembered better than those leading to negative feedback.

[Figure 3 approximately here]

Using the SAS-CAF index (see Methods) as a measure of how well card combinations were remembered, learning was better for sun than rain combinations ($p=0.036$) and for combinations for which more smileys than frowneys were received during the initial 30 trials of WPT training ($p=0.001$, **Fig. 3b**). The interaction between the two independent variables was not significant. Likewise the predictive value of the combination (high-medium-low) had no significant effect if the variables sun/rain and smiley-frowneys were included in the model. Whether or not the subject was in the fMRI group had no significant effect.

Brain activation

During the presentation period, the control task was associated with more activation in several brain areas as compared with the WPT (**Table 2a**). In contrast, during the feedback period, WPT lead to stronger activation in the left inferior frontal gyrus (**Table 2b**). In comparison with WPT card combinations without predictive value (chance trials), the combinations with a predictive value lead to stronger activation in the left posterior cingulum, Brodmann's area 31, during presentation (**Table 2c**). No differences between chance and predictive combinations were observed during feedback.

[Table 2 approximately here]

The third random effects ANOVA model tested the effect of combinations predicting sun- versus rain (content) and the combinations' predictive value (combinations grouped in high - 92%, medium - 89% and low - 83%, 78%, 55% to account for uneven frequency of presentation, **Table 1**) on activation. For the presentation period of the trial, no significant voxels were found related to either the interaction or to the individual effects of content (sun/rain) or predictive value. For the feedback period, estimating the effect of the within-subject variable „predictive value“

(high/medium/low) using an F-test as well as the contrast (t-test) high > low predictive value showed significant voxels in the right lateral cerebellum (**Figure 4, Table 3a**).

The fourth random effects ANOVA model tested the effects of smiley versus frowney feedback during the feedback period of the trial. Smiley rewards were related to stronger activation in bilateral nucleus accumbens, bilateral posterior cingulum, left primary motor cortex, right postcentral gyrus, and right premotor cortex (**Figure 5, Table 3b**). No brain region was identified in which frowneys produced stronger activation than smileys.

[Table 3, Figures 4 and 5 approximately here]

Including age group (dichotomous variable by median split according to age) into the smiley-frowney model as a between-subject variable revealed significant voxels in left cingulate gyrus, medial frontal gyrus and putamen (**Table 4**) for the interaction young>old x smiley>frowney. Including age group as a between-subject variable in the other statistical models did not reveal any significant results.

[Table 4 approximately here]

Discussion

These results demonstrate that learning efficacy in a probabilistic classification task depends on predictive value of the cue, the content (sun/rain) to be learned and the feedback (smiley/frowney) provided. Content and feedback seem to be independent factors. Memory of card combinations was better for combinations of cards with higher predictive value, for those combinations that predicted sun instead of rain, and if smiley feedback was provided. Higher predictive value lead to stronger activation of the lateral cerebellar hemisphere. Smiley feedback was associated with stronger activation of Nucleus accumbens (NAcc), sensorimotor cortex, and cingulum.

Subjects remembered card-weather associations markedly better when positive instead of negative feedback was provided. Wächter and colleagues (2009) reported a similar finding during implicit motor learning in healthy individuals. It may be that as individuals age negative feedback becomes more important as shown for avoidance learning in subjects with higher age (77 years) than our sample (Frank & Kong 2008). In fMRI smiley rewards were associated with stronger activation in bilateral NAcc (left>right; left dominance may be a consequence of right-hand button presses, Haruno et al 2004). NAcc activation is frequently observed during reward processing (Bischoff-Grethe et al 2009; Jensen et al 2007; Linke et al 2010; Seger & Cincotta 2005; Wächter et al 2009; Ullsperger & von Cramon 2003; Aron et al 2006; Poldrack et al 2001). During an over-learned cue-response task NAcc activation correlated with the amount of anticipated monetary reward (Knutson et al 2001). A metaanalysis of functional imaging studies on reward confirms bilateral activation of the NAcc and the posterior cingulum during positive versus negative feedback (Liu et al 2011). Also, the orbitofrontal cortex was reported to be over-active during presentation of reward versus punishment (Liu et al 2011; Jensen et al 2007) and reward prediction

error (O'Doherty et al 2003), but was not activated here. It is possible that this finding is related to the age of our subject sample; the frontal cortex is specifically vulnerable to age-related metabolic dysfunction (Curiati et al 2011) and structural atrophy (Raz et al 2004). In support of this interpretation frontal cortex activation was stronger in the younger half of our subjects.

We found stronger activation of sensorimotor cortex (primary motor, premotor and somatosensory cortex) with smiley than frowney feedback which has not been reported by reward studies using explicit learning paradigms. Paradigm differences may explain this discrepancy. Most reward studies use tasks that are over-learned, involve minimal learning or are not designed to learn associations based on feedback (e.g. Elliott et al 2003, Knutson et al 2001). Here, rewards were provided to learn associations between card combinations and outcomes. These associations were stochastic, hence, learning was largely implicit. Reward processing for implicit learning may involve the motor cortices like it does for motor skill learning (Wächter et al 2009). Whether this activation reflects a dopaminergic reward signal that is routed directly to motor cortex remains speculative (Molina-Luna et al 2009; Hosp et al 2011).

The posterior cingulate cortex (PCC) was an additional brain region activated by positive > negative feedback confirming previous reports (Liu et al 2011). In primates, PCC activity corresponds to decisions deviating from a standard, i.e. decision salience (Heilbronner et al 2011). It has been proposed that the PCC detected change relative to a standard (expected) signal in general (Pearson et al 2011). In our context, positive rewards may have had greater impact on changing behavior than negative ones.

As expected, learning efficacy correlated directly with the card combination's predictive value. As compared to low predictive value, high-value trials over-activated the right lateral cerebellum. During sensorimotor learning, the cerebellum encodes error signals. Additional activation is unrelated to error (Imamizu & Kawato 2009) and has been suggested to reflect the generation of an internal model that serves as a cognitive framework for task-related decisions (Ito 2005). One can speculate that cerebellar activation observed here for high>low predictive value trials reflects internal model formation for the classification task. High value trials are more informative than low-value trials for the formation of such a model. If cerebellar activity would represent an error signal, activation should have been stronger for low-value combinations, because for those subjects made more errors.

Combinations of cards predicting sun were better remembered than those predicting rain. Given everyday experience that positive events are better remembered than negative ones, this finding seems plausible in the context of explicit learning. That it transfers to implicit classification learning has – to our knowledge – not been reported before. That this finding is spurious and caused by receiving more smiley rewards for sun combinations was excluded by showing the smileys at the beginning of training were not more frequently presented for sun than rain trials. That this distribution became uneven later is expected because sun combinations were learned better and rewarded with more smileys. We also found statistically no interaction between the variables sun/rain and the number of smileys-frowney at the beginning of training in their effect on how well patterns were remembered. Nevertheless, was the behavioral difference between sun and rain trials small, which is probably why we did not observe a differential effect on brain activation.

The fMRI control task was chosen to subtract the activation related to visual presentation and motor response from performing the WPT. A comparison of the presentation phase of control and WPT trials showed stronger control-related activation in bilateral frontal, parietal and cerebellar areas. This likely reflects the fact that more intense processing was required for counting cards than for implicitly assessing their predictive value and deciding in favor of sun or rain during WPT. Counting is known to be associated with activation of fronto-temporal language areas, frontal and parietal cortices and cerebellum (Sveljo et al 2010; Hinton et al 2004; Kansaku et al 2006; Ardila 2010). During the feedback phase of a trial, findings were opposite in that the WPT was associated with stronger activation than control. This activation localized to the inferior frontal gyrus (Brodmann area 45). While this region is part of Broca's language area, it also is involved in risk assessment (d'Acremont et al 2009). Risk prediction error processing is likely more important during WPT than control.

A limitation of this study is that the individual valence of positive and negative stimuli was not assessed or controlled, neither for content nor feedback. Individuals may have found sun pleasant but rain rather neutral. This is a common criticism for many learning paradigms that focus on valence (Lang et al 1990; Mneimne et al 2010). We used stimuli of small valence (smiley/frowney as feedback, imagined sun/rain as a response) to minimize potential differences in salience thereby hoping to reduce this confound. Although we think the effect is small, we cannot rule out a possible influence. A difference to prior fMRI studies using the WPT is that our control task was not interleaved with WPT trials. We chose this design to render the WPT data comparable to WPT training outside the scanner performed by subjects that did not qualify or opted against MR scanning. To minimize sequencing effects the control

condition was performed before and after WPT training and for analysis the data of the two control periods were combined. A limitation to acknowledge is the small size of the fMRI sample. Nevertheless, random effects models that offer generalizability, yielded significant results. A limitation is also the advanced age of our subject sample. Subjects were collected as an age-matched control group for a comparison with individuals after a stroke to be reported elsewhere. Further studies are warranted to investigate the effects of feedback, content and predictive value in young healthy subjects.

In conclusion, our data show that pleasant content and feedback improve implicit classification learning. Positive feedback is associated with stronger activation of NAcc, sensorimotor cortex and posterior cingulum as compared with negative feedback. Learning also depends on the predictive value of the visual cues which is in part processed within the lateral cerebellum possibly reflecting the formation of an internal model.

References

- Ali N, Cimino CR (1997) Hemispheric lateralization of perception and memory for emotional verbal stimuli in normal individuals. *Neuropsychology* 11: 114-125.
- Ardila A (2010) On the evolution of calculation abilities. *Front Evol Neurosci* 2.
- Aron AR, Gluck MA, Poldrack RA (2006) Long-term test-retest reliability of functional MRI in a classification learning task. *Neuroimage* 29: 1000-1006.
- Bischoff-Grethe A, Hazeltine E, Bergren L, Ivry RB, Grafton ST (2009) The influence of feedback valence in associative learning. *Neuroimage* 44: 243-251.
- Cahill L, McGaugh JL (1996) Modulation of memory storage. *Curr Opin Neurobiol* 6: 237-242.
- Curiati PK, Tamashiro-Duran JH, Duran FL, Buchpiguel CA, Squarzoni P, Romano DC, Vallada H, Menezes PR, Scazufca M, Busatto GF, Alves TC (2011) Age-Related Metabolic Profiles in Cognitively Healthy Elders: Results from a Voxel-Based [18F]Fluorodeoxyglucose-Positron-Emission Tomography Study with Partial Volume Effects Correction. *AJNR Am J Neuroradiol* .
- d'Acremont M, Lu ZL, Li X, Van der Linden M, Bechara A (2009) Neural correlates of risk prediction error during reinforcement learning in humans. *Neuroimage* 47: 1929-1939.
- Elliott R, Newman JL, Longe OA, Deakin JF (2003) Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23: 303-307.
- Frank MJ, Kong L (2008) Learning to avoid in older age. *Psychology and aging* 23: 392-398.
- Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M (2004) A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J Neurosci* 24: 1660-1665.
- Heilbronner SR, Hayden BY, Platt ML (2011) Decision salience signals in posterior cingulate cortex. *Front Neurosci* 5: 55.
- Hinton SC, Harrington DL, Binder JR, Durgerian S, Rao SM (2004) Neural systems supporting timing and chronometric counting: an FMRI study. *Brain Res Cogn Brain Res* 21: 183-192.
- Hosp JA, Pektanovic A, Rioult-Pedotti MS, Luft AR (2011) Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning. *J Neurosci* 31: 2481-2487.
- Imamizu H, Kawato M (2009) Brain mechanisms for predictive control by switching internal models: implications for higher-order cognitive functions. *Psychol Res* 73: 527-544.
- Ito M (2005) Bases and implications of learning in the cerebellum--adaptive control and internal model mechanism. *Prog Brain Res* 148: 95-109.
- Jensen J, Smith AJ, Willeit M, Crawley AP, Mikulis DJ, Vitcu I, Kapur S (2007) Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum Brain Mapp* 28: 294-302.
- Kansaku K, Johnson A, Grillon ML, Garraux G, Sadato N, Hallett M (2006) Neural correlates of counting of sequential sensory and motor events in the human brain. *Neuroimage* 31: 649-660.
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273: 1399-1402.

- Knutson B, Adams CM, Fong GW, Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21: RC159.
- Lang PJ, Bradley MM, Cuthbert BN (1990) Emotion, attention, and the startle reflex. *Psychological review* 97: 377-395.
- Linke J, Kirsch P, King AV, Gass A, Hennerici MG, Bongers A, Wessa M (2010) Motivational orientation modulates the neural response to reward. *Neuroimage* 49: 2618-2625.
- Liu X, Hairston J, Schrier M, Fan J (2011) Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neurosci Biobehav Rev* 35: 1219-1236.
- Minnema MT, Knowlton BJ (2008) Directed forgetting of emotional words. *Emotion* 8: 643-652.
- Mneimne M, Powers AS, Walton KE, Kosson DS, Fonda S, Simonetti J (2010) Emotional valence and arousal effects on memory and hemispheric asymmetries. *Brain and cognition* .
- Molina-Luna K, Pekanovic A, Rohrich S, Hertler B, Schubring-Giese M, Rioult-Pedotti MS, Luft AR (2009) Dopamine in motor cortex is necessary for skill learning and synaptic plasticity. *PLoS One* 4: e7082.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38: 329-337.
- Pearson JM, Heilbronner SR, Barack DL, Hayden BY, Platt ML (2011) Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci* 15: 143-151.
- Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA (2001) Interactive memory systems in the human brain. *Nature* 414: 546-550.
- Raz N, Gunning-Dixon F, Head D, Rodrigue KM, Williamson A, Acker JD (2004) Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol Aging* 25: 377-396.
- Seger CA, Cincotta CM (2005) The roles of the caudate nucleus in human classification learning. *J Neurosci* 25: 2941-2951.
- Sveljo OB, Koprivsek KM, Lucic MA, Prvulovic MB, Culic M (2010) Gender differences in brain areas involved in silent counting by means of fMRI. *Nonlinear Biomed Phys* 4 Suppl 1: S2.
- Ullsperger M, von Cramon DY (2003) Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J Neurosci* 23: 4308-4314.
- Wächter T, Lungu OV, Liu T, Willingham DT, Ashe J (2009) Differential effect of reward and punishment on procedural learning. *J Neurosci* 29: 436-443.

Table 1

Combination	Combination class according to predictive value	Card				Percent of trials with combination	Probability for predicting sun
		1	2	3	4		
5	high	0	0	1	1	13 %	0.08
1	medium	0	0	0	1	9.5 %	0.11
7	medium	0	1	1	1	9.5 %	0.11
3	low	0	1	0	1	6 %	0.17
11	low	1	1	0	1	4.5 %	0.22
6	low	0	1	1	0	4.5 %	0.44
2	none	0	1	0	0	3 %	0.5
13	none	1	0	1	1	3 %	0.5
4	low	0	0	1	0	4.5 %	0.55
9	low	1	0	0	1	4.5 %	0.78
12	low	1	0	1	0	6 %	0.83
8	medium	1	0	0	0	9.5 %	0.89
14	medium	1	1	1	0	9.5 %	0.89
10	high	1	1	0	0	13 %	0.92

Table 2a

Region of interest	Brodmann	Side	x	y	z	t	p
Somatosensory cortex	3	left	-34	-23	42	-4.93	0.000109
Premotor cortex	6	left	-10	-2	60	-5.79	0.000017
Premotor cortex	6	left	-28	-2	45	-4.54	0.000254
Middle Temporal Gyrus	22	left	-55	-35	6	-9.37	0.000000
Orbitofrontal cortex	47	left	-49	31	-3	-6.94	0.000002
Thalamus		left	-1	-11	12	-5.97	0.000012
Putamen		left	-16	10	-6	-4.81	0.000141
Cerebellar hemisphere		left	-43	-47	-33	-4.91	0.000113
Premotor cortex	6	right	23	-5	45	-5.40	0.000040
Fusiform Gyrus	20	right	50	-38	-24	-4.35	0.000385
Posterior Cingulate	31	right	11	-32	33	-5.18	0.000064
Cerebellar hemisphere		right	14	-56	-33	-4.74	0.000165

Table 2b

Region of interest	Brodmann	Side	x	y	z	t	p
Inferior frontal gyrus	45	left	-49	28	9	6.48	0.000004

Table 2c

Region of interest	Brodmann	Side	x	y	z	t	p
Posterior cingulum	31	left	-19	-26	36	4.20	0.000539

Table 3a

Region of interest	Brodmann	Side	x	y	z	t	p
Cerebellar hemisphere		right	47	-47	-36	6.97	0.000002

Table 3b

Region of interest	Brodmann	Side	x	y	z	t	p
Primary motor cortex	4	left	-28	-29	45	8.50	0.000014
Primary motor cortex	4	left	-37	-14	54	7.69	0.000030
Posterior Cingulate	31	left	-16	-29	36	8.93	0.000009
Nucleus accumbens		left	-7	10	0	8.37	0.000015
Somatosensory cortex	3	right	44	-20	51	8.68	0.000012
Premotor cortex	6	right	11	-11	57	9.05	0.000008
Premotor cortex	6	right	8	-26	57	6.95	0.000067
Posterior Cingulate	23	right	11	-35	27	9.70	0.000005
Posterior Cingulate	31	right	23	-20	33	7.74	0.000029
Nucleus accumbens		right	10	7	-2	5.23	0.000585

Table 4

Region of interest	Brodmann	Side	x	y	z	t	p
Posterior Cingulate	23	left	-4	-32	27	26.6	0.000871
Anterior Cingulate		left	-10	37	-3	30.2	0.000580
Medial Frontal Gyrus	32	left	-16	13	42	37.2	0.000289
Putamen		left	-13	7	-6	25.7	0.000967

Figure Legends

Figure 1. Set of playing cards. 1-3 cards were shown to form the card combinations described in Table 1.

Figure 2. Influence on learning. **(a)** Learning curves constructed from correct-versus-incorrect responses (see Methods) and averaged for all subjects are shown for the two card combinations with highest predictive value for sun (combination #10, red) and rain (combination #5, blue). Plateau performance is reached after approximately half the trials and is lower for rain than sun card combinations. Values on the y-axis represent the steepness, i.e., the first derivative of the cumulative performance curve, see Methods for details. **(b)** Plateau values estimated from the learning curves (examples in a) are shown for each card combination (bars indicate estimated plateau, error bars reflect SE, probabilities according to **Table 1**). The plateaus are lower for card combinations predicting rain (blue) than for those predicting sun (red). This indicates that the emotional value of the learned content has an influence on learning efficacy. **(c)** Plateau values also depend on the frequency by which a card combination occurs and **(D)** on the predictive value of the card combination (difference from chance).

Figure 3. **(a)** Smileys are better remembered than frowneys. If a smiley was presented, the subjects were more likely to give the same answer in the subsequent trial with the same card combination (yellow bars). In contrast, frowneys did not motivate the subjects to give the opposite answer (green bars, * $p < 0.001$). **(b)** The index of memory, computed as shown in the y-axis label, was related to how many smileys versus frowneys were received during the initial phase of WPT training and to whether sun or rain was predicted.

Figure 4. Brain activation during the feedback phase of the trial reflecting the contrast „card combinations with high (according to Table 1) > low predictive value“. High value combinations recruited the lateral cerebellar hemisphere (threshold $p < 0.05$, FDR corrected).

Figure 5. Brain activation during the feedback phase of the trial reflecting the contrast „smiley > frowney feedback“. Rewards (smileys) activated nucleus accumbens, sensorimotor and premotor cortices and cingulum more than negative (frowney) feedback (threshold $p < 0.05$, FDR corrected).