Decisions bias future choices by modifying hippocampal associative memories

Lennart Luettgau[1,2*], Claus Tempelmann[3], Luca Franziska Kaiser[2] & Gerhard Jocham[1,2]

[1]Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[2]Center for Behavioral Brain Sciences, Otto-von-Guericke University, Magdeburg, Germany
[3]Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany

*Correspondence to: luettgau@hhu.de

Abstract

Decision making is guided by memories of option values. However, retrieving items from memory renders them malleable. Here, we show that merely retrieving values from memory and making a choice between options is sufficient both to induce changes to stimulus-reward associations in the hippocampus and to bias future decision making. After allowing participants to make repeated choices between reward-conditioned stimuli, in the absence of any outcome, we observed that participants preferred stimuli they had previously chosen, and neglected previously non-chosen stimuli, over otherwise identical-valued options. Using functional brain imaging, we show that decisions induced changes to hippocampal representations of stimulus-outcome associations. These changes predicted future decision biases. Our results indicate that choice-induced preference changes are driven by choice-induced modification of memory representations and suggest that merely making a choice - even without experiencing any outcomes - induces associative plasticity.

Introduction

According to neo-classical economic models of decision making, choices are guided by memories of option values (Ariely and Norton, 2008). However, this unidirectional view has been challenged by cognitive accounts of decision making, suggesting that memory representations of option values might themselves be subject to changes induced by an agent's choices. This suggests a bidirectional relationship between value representations in memory and decision making (Ariely and Norton, 2008; Riefer et al., 2017).

However, even though real-life decisions often involve memory retrieval of learned associations between reward-predictive cues and outcomes, memory mechanisms underlying choice-induced preference changes have, to the best of our knowledge, never been systematically studied (Ariely and Norton, 2008; Brehm, 1956; Sharot et al., 2010). This might be partially related to the fact that most studies on choice-induced preference changes employed direct presentation of the outcomes to be chosen. However, this approach by design obliterates and confounds underlying associative contributions to the revaluation process (Izuma et al., 2010), such as retrieval dynamics (Wimber et al., 2015).

We hypothesized that, when choosing between two reward-conditioned stimuli (CS), the stimulus-outcome association of the chosen option would be strengthened, whereas the association of the non-chosen option would be weakened, resulting in subsequent preference changes. Our hypothesis is based on three grounds. First, presentation of a CS reactivates the neural patterns coding outcomes it has been paired with (Barron et al., 2013; Boorman et al., 2016; Howard et al., 2016; Klein-Flugge et al., 2013; Tonegawa et al., 2018). Second, retrieving memories makes them malleable and requires subsequent reconsolidation of the original trace (Nader et al., 2000; Nader and Hardt, 2009). Third, while retrieval of an item from memory leads to improved remembering, memory for competing items, e.g. CS associated with the same outcome, is impaired simultaneously (Anderson et al., 1994; Wimber et al., 2015). We expected choice-induced preference changes to be driven by modifications of stimulus-outcome associations in the hippocampus and lateral

orbitofrontal cortex, two key regions for storing and updating associative representations (Boorman et al., 2016; Klein-Flugge et al., 2013).

Thus, the goal of the present study was twofold. First, we aimed at investigating how choice-related alterations of associative memories bias future decision making. Second, we sought to investigate a neurobiologically plausible mechanism underlying choice-induced preference changes. To test our key predictions, we designed a novel learning and decision making paradigm which we used in three independent behavioral experiments and one functional magnetic resonance imaging (fMRI) experiment. For the fMRI study, we exploited repetition suppression (RS) effects (Barron et al., 2016, 2013; Garvert et al., 2017; Grill-Spector and Malach, 2001; Klein-Flugge et al., 2013) to measure associative strength between conditioned (CS) and unconditioned stimuli (US) (Boorman et al., 2016; Klein-Flugge et al., 2013). Participants first established Pavlovian associations between CS and differently valued US. Next, in a choice-induced revaluation, participants made binary choices between differently valued CS, without observing the associated US. Finally, in a probe phase, where they made choices between all possible CS combinations, participants showed preference increases for previously chosen, and preferences decreases for previously non-chosen CS, compared to otherwise equivalent CS. These choice-induced alterations in decision behavior were accompanied by corresponding changes in CS-US associative strength in the hippocampus and lateral orbitofrontal cortex. Furthermore, the magnitude of the hippocampal effect predicted individual probe phase decision biases.

Results

Behavioral Experiments

First, we detailed the behavioral choice-induced revaluation effect in three independent experiments. In each experiment, participants learnt associations between neutrally rated CS and three food items (Blechert et al., 2014) serving as unconditioned stimuli (US, Fig. 1A). For each participant, the US were individually chosen based on a prior rating of subjective preference on a visual analog scale ranging from low to high preference. This rating procedure resulted in selection of a low-value (US^{-}), an intermediate-value (US^{0}) and a high-

value (US^+) food item. Next, participants rated kanji stimuli according to their subjective preference. Six of these kanjis rated in close proximity to “neutral” (center point of the visual analog scale) were selected as CS for Pavlovian learning. Two CS each were paired with one US, resulting in three categories of differently valued CS: $CS^+_{A/B}$, $CS^0_{A/B}$ and $CS^-_{A/B}$, for high-, intermediate- and low-value CS.

The Pavlovian learning phase was followed by a choice-induced revaluation. From two value categories, one CS each was selected, and participants made binary choices between them, interspersed with lure decisions between non-reward-predictive kanjis (Fig. 1B). Crucially, no associated US were presented following choices, excluding the possibility of alterations in strength of stimulus-outcome associations due to directly experienced outcomes. The choice-induced revaluation phase was followed by a decision probe phase in which participants chose repeatedly between all binary CS combinations to assess preferences (Fig. 1C). The key comparison was between CS presented during revaluation versus CS from the same value category that had not been presented. Again, no outcomes were presented.

Decisions are biased by past choices

There was evidence for value transfer from US to CS across all studies (Fig. 1E-H), as indicated by significant main effects of CS value on probe phase choice probabilities across all four experiments (all Fs > 94.99, Ps < 0.001, repeated-measure analysis of variance, rmANOVA). Decision making during the revaluation phase had clearly dissociable effects on choices during the probe phase. CS that were chosen during the revaluation phase were more likely to be selected during the later probe phase compared to the CS of equal value that were not presented during choice-induced revaluation.

In experiment 1, participants (N = 40) made choices between the intermediate-value CS^0_A and the high-value CS^+_A during the choice-induced revaluation phase. In the probe phase, participants preferred CS^+_A, the previously selected stimulus, compared to CS^+_B (Z = 3.98, P < 0.001, Wilcoxon signed-rank test). Conversely, participants were less likely to select CS^0_A, the previously non-selected stimulus, compared to CS^0_B (Z = 1.97, P = 0.048,

Wilcoxon signed-rank test). This dissociation in choice behavior was also evident in a significant interaction effect of CS value × CS type (A or B): $F_{2, 78} = 10.01$, $P < 0.001$ (rmANOVA, Fig. 1E). Thus, compared to otherwise equivalent CS, participants exhibited a systematic preference for CS they had previously chosen, whereas they displayed a diminished preference of CS they had previously not chosen.

After having established that an intermediate value CS (CS^0_A) could be devalued by non-choices, we next asked in experiment 2, whether we can induce the exact opposite, an increase in preference for CS^0_A. Therefore, participants (N = 40) were presented with decisions between intermediate-value CS^0_A and low-value CS^-_A during the choice-induced revaluation phase. Conceptually replicating the results of experiment 1, participants favored CS^0_A over CS^0_B (Z = 2.21, P = 0.027, Wilcoxon signed-rank test) and neglected CS^-_A compared to CS^-_B (Z = 1.92, P = 0.028, one-tailed, Wilcoxon signed-rank test) during the decision probe phase. Again, there was a significant interaction effect of CS value × CS type ($F_{2, 78} = 4.84$, P = 0.01, rmANOVA, Fig. 1F) indicating clearly dissociable choice behavior during the decision probe. This pattern of results suggests a ubiquitous, value-independent mechanism of choice-induced revaluation.

These results so far show that choices and non-choices act in opposite directions. Consequently, we predicted that it would be possible to cancel out choice-induced preference increases and devaluation of CS^0_A. To test this prediction, in experiment 3 (N = 44), we presented an equal amount of binary decisions between CS^0_A and CS^-_A as between CS^0_A and CS^+_A during choice-induced revaluation. As expected, there was no evidence for change in preference for CS^0_A compared to CS^0_B (Z = 0.41, P = 0.68, Wilcoxon signed-rank test, Fig. 1G), suggesting that the effects of choices and non-choices had indeed cancelled each other out (interaction effect of CS value × CS type: $F_{2, 86} = 1.31$, P = 0.28, rmANOVA).

Importantly, the dissociations observed in choice behavior in experiment 1 & 2 rule out alternative accounts for explaining choice behavior, such as extinction or mere exposure effects. Both accounts would predict unidirectional preference changes for the CS presented

during choice-induced revaluation, independent of the choices made (decreases or increases in preference, respectively), which is incompatible with the present results.

Additionally, we fit six different variants of a reinforcement learning model using Rescorla-Wagner-like delta update rules (Rescorla and Wagner, 1972). For each experiment, we compared the six models that implemented different ways by which participants could have learned CS-US associative strength - and updated associative strength during choice-induced revaluation based on "fictive" reward prediction errors. The “fictive” reward prediction errors were based on our reasoning that presentation of CS during revaluation would lead to retrieval of the associated US and that, consistent with our hypothesis, stimulus-outcome association of the chosen CS would be strengthened, whereas the association of the non-chosen CS would be weakened. As no objective feedback (US) was presented during choice-induced revaluation, we assumed that the “fictive” reward prediction error would be computed as the difference between the CS value during previous presentation of the respective CS and subjective value of the current trial’s associated US, retrieved from associative memory. Our behavioral results were best captured by a computational model that differentially updated the learned CS-US associative strengths using "fictive" reward prediction errors elicited by revaluation phase decisions (Methods and Supplementary Table S1).

Choices modify stimulus-outcome associations

Having established and replicated the behavioral effect of choice-induced revaluation in three independent behavioral samples, we next tested whether decisions induce changes of neural representations of CS-US associations.

In Experiment 4, we used functional magnetic resonance imaging and leveraged repetition suppression (RS) effects (Barron et al., 2016, 2013; Garvert et al., 2017; Grill-Spector and Malach, 2001; Klein-Flugge et al., 2013) to measure CS-US associative strength (Boorman et al., 2016; Klein-Flugge et al., 2013). When a neural ensemble is activated twice in brief succession (e.g. by rapid sequential presentation of the same visual stimulus), the second stimulus causes a diminished response. Accordingly, after learning the association

between CS and US, the CS should elicit a representation of its associated US. Thus, presentation of the US itself, following the CS, should induce a diminished neural response. If the association between CS and US has been weakened by non-choices during revaluation, the CS is no longer capable of evoking the US representation to the same degree and should therefore elicit a stronger response (less repetition suppression). The same logic in reverse applies when the association has been strengthened by choices during revaluation.

As in the three behavioral experiments, participants first learned the six CS-US associations during Pavlovian learning. Following Pavlovian learning, we administered two RS blocks, one immediately before ("PRE") and one immediately after ("POST") the choice-induced revaluation phase, where participants (N = 42) made binary choices between CS^0_A and CS^+_A. Consistent with our hypothesis, we observed both a decrease in associative strength for CS^0_A-US^0 relative to CS^0_B-US^0, and an increase in associative strength between CS^+_A-US^+ compared to CS^+_B-US^+ (Fig. 2A, D) during POST but not during PRE (Fig. 2D inset, Z = 2.53, P = 0.01, Wilcoxon signed-rank test) in the left hippocampus (Fig. 2A). Detailed analysis showed that this conjunction effect was driven by a double dissociation (interaction effect CS value and time (PRE or POST), $F_{1, 41}$ = 4.51, p = 0.04, rmANOVA): we found decreased associative strength for CS^0_A (Z = 2.27, P = 0.023, Wilcoxon signed-rank test) and increased associative strength for CS^+_A (Z = 2.27, P = 0.023, Wilcoxon signed-rank test) during POST, without any evidence for non-zero differences in PRE (all Zs < 1.05, Ps > 0.3 , Wilcoxon signed-rank tests) or for the control contrast of either CS^-_A or CS^-_B (all Zs < 0.31, Ps > 0.76, Wilcoxon signed-rank tests, Fig 2D). Thus, CS choices during choice-induced revaluation increased, whereas non-choices decreased hippocampal CS-US associative strength. However, while there was evidence for significant PRE-POST reduction in associative strength for CS^0_A (Z = 1.84, P = 0.033, one-tailed, Wilcoxon signed-rank test), the PRE-POST increase in associative strength for CS^+_A was not significant (Z = 1.48, P = 0.07, one-tailed, Wilcoxon signed-rank test).

Furthermore, we found a conjunction effect of a decrease in associative strength for CS^0_A-US^0 relative to CS^0_B-US^0 and an increase in associative strength for CS^+_A-US^+ relative to CS^+_B-US^+ in the right lateral orbitofrontal cortex that survived small-volume correction (Fig. 2B). Extraction of parameter estimates from this cluster using an independent region of interest(Jocham et al., 2016) revealed that the interaction effect (CS value and time (PRE or POST), $F_{1, 41} = 5.57$, $p = 0.02$, rmANOVA) was driven by a significant PRE-POST reduction in associative strength for CS^0_A (Fig. 2E, $Z = 1.77$, $P = 0.038$, one-tailed, Wilcoxon signed-rank test) but not by a PRE-POST increase in associative strength for CS^+_A ($Z = 0.66$, $P = 0.26$, one-tailed, Wilcoxon signed-rank test). However, we found decreased associative strength for CS^0_A ($Z = 1.74$, $P = 0.04$, one-tailed, Wilcoxon signed-rank test) and increased associative strength for CS^+_A ($Z = 2.68$, $P = 0.004$, one-tailed, Wilcoxon signed-rank test) during POST, without evidence for non-zero differences in PRE (all $Zs < 1.77$, $Ps > 0.08$, Wilcoxon signed-rank tests). Exploratory analyses at a lenient, uncorrected threshold ($Z > 2.8$, uncorrected) yielded clusters in the left ventral tegmental area (VTA) and the right nucleus accumbens (NAcc, Fig 2C). Both effects were driven by significant reductions in associative strength for CS^0_A (VTA: $Z = 1.82$, $P = 0.034$, one-tailed; NAcc: $Z = 2.38$, $P = 0.009$, one-tailed, Wilcoxon signed-rank tests), but showed no evidence of increases in associative strength for CS^+_A (VTA: $Z = 0.42$, $P = 0.34$, one-tailed; NAcc: $Z = 1.53$, $P = 0.06$, one-tailed, Wilcoxon signed-rank tests). Overall, these results suggest that decisions during choice-induced revaluation had clearly dissociable effects on previously learned CS-US associations: While the previously chosen CS exhibited increased associative strength to its related US, the exact opposite effect was true for the previously non-chosen CS, exhibiting decreased associative strength to its related US.

Hippocampal CS-US associative strength predicts future choices

We next investigated whether the observed choice-induced modifications of hippocampal CS-US associative strength were predictive of choice biases during the probe phase. Unlike in Experiment 1, CS^0_A (non-chosen stimulus during revaluation) was not chosen at a lower rate than CS^0_B in the probe phase ($Z = 1.01$, $P = 0.32$, Wilcoxon signed-rank test, Fig. 1H) in

Experiment 4. We speculate that this might be due to reconsolidation of the originally learned CS-US associations in the POST choice-induced revaluation RS run (participants were explicitly instructed to judge whether the presented CS-US associations were correct). However, replicating Experiment 1, we observed a choice-induced increase in preference for CS^{+}_{A} over CS^{+}_{B} (Z = 3.03, P = 0.002, Wilcoxon signed-rank test, Fig. 1H). We therefore focused on this effect in brain-behavior correlations. The difference of hippocampal associative strength between CS^{+}_{A}-US^{+} and CS^{+}_{B}-US^{+} was positively correlated with the difference between choice probabilities of CS^{+}_{A} and CS^{+}_{B} (ρ_{41} = .31, P = 0.024, one-tailed, Spearman correlation; Fig. 3B). The more the hippocampal representation of the CS^{+}_{A}-US^{+} association had been strengthened by choices during revaluation, the more likely participants were to select CS^{+}_{A} compared to its non-revalued partner stimulus CS^{+}_{B}.

In line with this, in a whole-brain analysis we observed a positive relationship between the difference in choice preference for CS^{+}_{A} versus CS^{+}_{B} and changes in CS-US associative strength from PRE to POST in the left posterior hippocampus, extending to left occipito-temporal complex (Fig. 3C). A similar whole-brain analysis using only the choices between CS^{+}_{A}. and CS^{+}_{B} as behavioral covariate revealed areas in the bilateral anterior insula and orbitofrontal cortex (Supplementary Fig. S2B). Neither VTA, NAcc, nor the cluster in the lateral orbitofrontal cortex showed relationships with probe phase behavior.

As an alternative to the proposed associative mechanism, our results could also be explained based on cached values. Participants might acquire incentive ("cached") CS values and use these to guide their decisions, independent of CS-US associations. It is possible that revaluation choices changed these cached CS values, instead of the CS-US associations. Under this reasoning, one would assume that the value information conveyed by CS^{0}_{A}, the CS not chosen during revaluation, would become more similar to the value information of the low-valued US^{-} after choice-induced revaluation(Summerfield et al., 2019). The cached value account would thus predict larger repetition suppression effects for CS^{0}_{A} followed by US^{-} as for CS^{0}_{B} followed by US^{-} during POST. Our design allowed us to set up a contrast to test this possibility (Methods – fMRI contrasts). This indeed yielded effects in the

right posterior hippocampus and a midbrain region in the vicinity of the dorsolateral substantia nigra pars compacta (Supplementary Fig. S1) in POST. Importantly, the hippocampal effect did not differ PRE-POST (test on parameter estimates extracted from right hippocampal anatomical mask, $Z = 0.91$, $P = 0.37$, Wilcoxon signed-rank test) and was independent of the "associative" effect in the left hippocampus (all $\rho s < .18$, $Ps > 0.26$, Spearman correlations). According to the cached value account, hippocampal representations of CS^0_A-US^- should be inversely related to preferences of CS^0_A. However, directly opposing this prediction, relationships of right hippocampus parameter estimates, and choice behavior were positive (all $\rho s < .31$, $P > 0.05$, Spearman correlations), rendering an explanation of the observed behavioral results based on cached values unlikely.

Discussion

Using a novel paradigm, we showed that decisions bias future choices, even without participants directly experiencing the outcomes of their decisions. Participants were more likely to select CS they had previously chosen, and less likely to select CS they had not chosen previously, compared to otherwise equivalent CS. At the neural level, we found that choices introduced alterations to hippocampal and orbitofrontal representations of stimulus-outcome associations that predicted future decisions.

The idea that past decisions bias preferences was put forth decades ago (Ariely and Norton, 2008; Brehm, 1956) and compelling evidence for post-decision revaluation has accumulated since (Ariely and Norton, 2008; Izuma et al., 2010; Riefer et al., 2017; Sharot et al., 2010) (but see Chen and Risen, 2010 for critical discussion). Here, we present behavioral evidence from four independent data sets that reward-predictive CS that were chosen in the past are more likely to be selected during future choice contexts, compared to CS of equal value that were not presented during choice-induced revaluation. Conversely, we found a decreased preference for CS that were not chosen in the past, compared to otherwise equivalent CS, indicating dissociable effects of choice-induced revaluation. Most importantly, our behavioral findings are independent of experienced rewards, as participants were never presented with the outcomes of their choices. This suggests that the observed effect could

arise from associative retrieval mechanisms (Anderson et al., 1994; Wimber et al., 2015), as also indicated by computational model selection. Computational model fits favored a model that differentially updated CS-US associative strength of chosen and unchosen options based on “fictive” reward prediction errors elicited by associative retrieval during decisions in the choice-induced revaluation phase. Our results are in line with previous reports of choice-induced preference changes (Ariely and Norton, 2008; Brehm, 1956; Izuma et al., 2010; Sharot et al., 2010). However, as most previous studies directly presented participants with the outcomes to be chosen and thereby confounded individual contributions of memory and choice mechanisms, our study might be the first to provide evidence for a value-independent associative memory mechanism driving choice-induced preference changes.

Furthermore, even though memory retrieval dynamics appear as a likely candidate mechanism for choice-induced preference changes, the exact neural mechanisms underlying this decision making phenomenon have remained unknown. Here, we provide evidence that past choices bias future decision making by modifying the strength of neural stimulus-outcome associations. The present results suggests that, during decision making, reactivation of stimulus-outcome associations (Barron et al., 2013; Boorman et al., 2016; Klein-Flugge et al., 2013; Wimmer and Shohamy, 2012) renders them subject to plasticity, with the association of the chosen stimulus being strengthened, and the association of the non-chosen stimulus being diminished. Such an effect is in line both with the literature on reconsolidation of reactivated memory traces (Nader et al., 2000; Nader and Hardt, 2009; Tonegawa et al., 2018) and on retrieval competition, where recall of items induces forgetting of competing items (Anderson et al., 1994; Wimber et al., 2015). Our results suggest that merely making a choice induces plasticity of associative representations in the hippocampus and lateral orbitofrontal cortex. In line with our present results, a fronto-hippocampal network comprising hippocampus and lateral OFC seems to be critically involved in reward-related updating of stimulus-outcome associations (Boorman et al., 2016). Whereas the hippocampus has been shown to be a key brain structure supporting relational learning and memory processes, including value spread during reinforcement learning (Wimmer and

Shohamy, 2012), factorized replay of event trajectories (Liu et al., 2019), the lateral OFC has been additionally implicated in the resolution of credit assignment problems in reinforcement learning (Jocham et al., 2016; Walton et al., 2010).

Our study has at least three implications for current theories of decision making. First, it proposes a memory-based account for choice-induced revaluation and, more broadly, choice history bias, two well-known, but still poorly understood decision making phenomena. By teasing apart decisions and memory processes related to choice-induced revaluation, we show that choices are not only guided by associative representations of value stored in memory, but that decisions themselves dynamically transform associative memories. This suggests that relational structures stored in decision makers' cognitive maps (Behrens et al., 2018; Tolman, 1948) can be distorted through their very own choice behavior. Second, we provide evidence for involvement of the hippocampus and lateral OFC in maintaining and updating of stimulus-outcome associations. This extends previous findings (Boorman et al., 2016; Jocham et al., 2016; Klein-Flugge et al., 2013; Walton et al., 2010) by describing a functional role of the hippocampus in value-based decision making that is independent of experienced reward. Such a role may more closely resemble naturalistic decision situations where consequences of choices often unravel at distant future time points, rendering credit assignment challenging (Jocham et al., 2016; Walton et al., 2010). Third, we provide a mechanism underlying seemingly irrational choice behavior: Even though participants chose between equivalent options, choice-induced revaluation biased them towards preferring chosen, and towards neglecting non-selected options. The latter might have important implications for explaining subjective preferences, especially in consumer choice behavior (Riefer et al., 2017) and in understanding why humans tend to make coherent decisions, even in conditions characterized by maladaptive choice behavior such as substance dependence or obsessive compulsive disorder.

Taken together, both our behavioral and neural results support the key prediction that past choices bias future decision making by altering hippocampal and orbitofrontal representations of stimulus-outcome associations. Our study provides a memory-based

mechanism to account for choice-induced preferences change effects (Ariely and Norton, 2008; Brehm, 1956; Izuma et al., 2010; Sharot et al., 2010). The present study shows that merely making a choice is sufficient to induce plasticity in reward-predictive associations stored in memory and further suggests that relational structures constituting a decision makers' cognitive maps can be adaptively altered through their very own choice behavior.

Methods

Participants

Participants were recruited from the local student community of the Otto von Guericke University Magdeburg, Germany by public advertisements and via online announcements. Only participants indicating no history of psychiatric or neurological disorder and no regular intake of medication known to interact with the central nervous system were included. Participants in all experiments had normal or corrected-to-normal vision and did not report experience with Japanese kanjis or Chinese characters. All participants provided informed written consent before participation and received monetary compensation for taking part in the study. The study was approved by the local ethics committee at the medical faculty of the Otto von Guericke University Magdeburg, Germany (February 2nd 2018, reference number: 19/18) and conducted in accordance with the Declaration of Helsinki.

49 young, healthy volunteers (age: *M* = 23.93, *SD* = 2.9 years, 18 males) participated in experiment 1. Seven participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase that served as an attentional control (<10 % responses in trials that required to indicate the color of the square surrounding the conditioned stimuli. Two additional participants had to be excluded due to not passing the manipulation check (high-value option selected < 50 % during choice-induced revaluation), thus leaving a total of *n* = 40 participants for final analyses.

64 young, healthy volunteers (age: *M* = 23.47, *SD* = 3.79 years, 26 males), participated in experiment 2. Ten participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase, and thirteen subjects were excluded due to not passing the manipulation check (intermediate valued option selected < 50 % during choice-induced revaluation), one additional participant had to be excluded due to a technical error, leaving *n* = 40 participants for statistical analyses.

61 young, healthy volunteers (age: *M* = 23.26, *SD* = 3.27 years, 23 males), participated in experiment 3. Ten participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase, and six subjects were

excluded due to not passing the manipulation check (high-value option selected < 50 % during choice-induced revaluation), one additional participant had to be excluded due to a technical error (no data was recorded), leaving *n* = 44 participants for statistical analyses.

58 young, healthy and magnetic resonance imaging (MRI)-compatible volunteers (age: *M* = 24.61, *SD* = 4.01 years, 30 males) participated in experiment 4 (functional MRI (fMRI) experiment). One participant fell asleep during the post-revaluation fMRI-repetition suppression (fMRI-RS) run, three participants discontinued the MRI acquisition (one due to claustrophobia, two reported a headache during task performance). Twelve additional subjects were excluded due to not passing the manipulation check (high-value option selected < 50 % during choice-induced revaluation), leaving *n* = 42 participants for statistical analyses.

Behavioral task - ratings

Participants received written instructions for the experiment and were instructed once again on the computer screen. The experiment was programmed in MATLAB 2012b (MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, MA, USA), using Psychophysics Toolbox (Brainard, 1997). Before the task, participants rated twenty-five different sweet and high-caloric food items selected from an online database (Blechert et al., 2014). Subjects were instructed to indicate the subjective desirability of the food items by using the “y” and “m” button on a standard German (QWERTZ) computer keyboard to position a red slider bar on a white visual analog scale (VAS) between 0 (not liked) and 100 (very much liked). The lowest- (Experiment 1: *M* = 17.5, *SD* = 20.51; Experiment 2: *M* = 16.65, *SD* = 16.29; Experiment 3: *M* = 18.59, *SD* = 22.15; Experiment 4: *M* = 20.41, *SD* = 20.44), and highest-rated (Experiment 1: *M* = 96.55, *SD* = 7.10; Experiment 2: *M* = 94.85, *SD* = 9.08; Experiment 3: *M* = 95.21, *SD* = 14.52; Experiment 4: *M* = 97.79, *SD* = 4.96) food item as well as a food item rated with the median value of all ratings (Experiment 1: *M* = 57.43, *SD* = 11.55; Experiment 2: *M* = 56.23, *SD* = 9.80; Experiment 3: *M* = 57.14, *SD* = 13.95; Experiment 4: *M* = 60.07, *SD* = 11.71) were selected for Pavlovian conditioning. We explicitly decided for clearly differentiable pictures of food items in order to elicit activation of

differential neural ensembles coding for those stimuli and to facilitate learning of vivid memories of stimulus-outcome associations. Next, subjects rated twenty (eleven in Experiment 1) Japanese kanjis according to subjective value/liking by using the “y” and “m” button on a standard computer keyboard to position a red slider bar on a white visual analog scale (VAS) between 0 (not liked) and 100 (very much liked). The six kanjis rated closest to 50 (equivalent to “neutral”) were selected and their order was randomized before being associated with the food items in Pavlovian conditioning. The subjective values/liking of the six selected kanjis did not differ significantly from each other (main effect of stimulus: all Fs < 1.39, ps > 0.22, rmANOVA).

Behavioral task - Pavlovian conditioning

Participants learned to associate the selected kanjis (conditioned stimuli, CS) with differently valued outcomes (unconditioned stimuli, US) by repeatedly observing one CS (2000 ms), followed by an inter-stimulus interval (1000 ms) marked by a fixation cross, and presentation of one US (2000 ms). Each trial was separated by an inter-trial-interval (ITI) marked by a grey screen. The ITI per trial was drawn from a discretized γ-distribution (shape = 5, scale = 0.9) truncated for an effective range of values between 3000 ms and 8000 ms. CS-US couplings were interleaved with 20 % CS-no-US couplings. This was intended to create an association that would not rapidly extinguish due to extinction effects related to not being presented with the associated US during the subsequent choice phases. Each US was associated with two different CS, resulting in three pairs of differently valued CS: $CS^{+}_{A/B}$, $CS^{0}_{A/B}$ and $CS^{-}_{A/B}$. Participants completed 180 trials (19 of the participants included in Experiment 1 had performed 240 trials) of Pavlovian conditioning, 30 (40, respectively) trials per CS-US association. The participants were presented with a cover story for the experiment: They were told to imagine themselves in preparation for a journey to Japan during which they would need to learn kanjis associated with certain food items to be ready to select their favorite sweets in Japanese shops. As an attentional control, we introduced a simple binary classification task, presenting participants with a red or blue square surrounding the CS. Each CS was equally often presented with a red or blue square and

color of the square did not predict US contingency. Subjects were instructed to react as quickly and as correctly as possible by pressing “y” upon seeing a blue square and pressing “m” upon seeing a red square surrounding the CS. Pavlovian conditioning was split up in three blocks of 60 (80, respectively) trials, interleaved with self-paced breaks. In Experiment 4, both ratings and Pavlovian conditioning were performed outside the MRI scanner.

Behavioral task - choice-induced revaluation

After completion of Pavlovian conditioning, participants were presented with repeated choices (28 trials) between a CS^{+}_{A} versus a CS^{0}_{A} (Experiments 1 & 4), CS^{0}_{A} versus a CS^{-}_{A} (Experiment 2), and a CS^{0}_{A} versus either a CS^{-}_{A} (14 trials) or a CS^{+}_{A} (14 trials) (Experiment 3), interleaved with lure decisions (28 trials) between four other neutrally rated kanjis that had never been presented during Pavlovian learning and thus were not associated with any of the US. The choice-induced revaluation phase served as the crucial manipulation in all experiments and was systematically varied across studies. Choice probability (CP) for the high-value CS served as a control for learning and as a manipulation check. Only participants selecting the higher valued CS more than 50 % (CP ≥ 0.50) were included in the final analysis, as we reasoned that choice-induced revaluation choices would (1) represent a marker of having learned the true associative values of the CS, (2) be a measure for learning, independent of the actual decision probe phase data (avoiding biased and arbitrary decisions for exclusion of participants), and (3) allow us to exclude decision makers showing random, or arbitrary choice behavior. Choice options were presented for 1500 ms and the chosen option was highlighted by a grey square surrounding the chosen CS. If participants did not respond within the time-window, a time-out message was displayed, and the respective trial was repeated at the end of the choice-induced revaluation phase. Order (left/right) of choice options was counter-balanced to avoid simple response patterns or decision rules (e.g. “always press left”). Participants were instructed to imagine themselves in a Japanese shop, where they would like to buy their favorite food items based on the previously learned kanjis (CS). Participants were told that one of the choice trials would randomly be drawn and their choice would determine which food item (US) they would

receive as a bonus upon completion of the experiment. Participants selected choice options by pressing the “y” (left option) or “m” (right option) button (left or right index finger on an MRI-compatible response box in Experiment 4). Importantly, participants were not presented with the US related to their chosen or unchosen CS to dissociate the observed effects from outcome-related relearning of CS-US associations. We assumed that presentation of a CS would pre-activate neural ensembles coding for the associated US. Consequently, we expected that choosing a CS would induce strengthening of the chosen option's CS-US association, whereas not choosing a CS would weaken the non-chosen option's CS-US association.

Behavioral task - decision probe

Following choice-induced revaluation, participants were presented with repeated binary choices (120 trials) between all possible CS combinations to assess CS preferences. Every CS combination was presented eight times in pseudo-random order. Choice options were presented for 1500 ms. If participants did not respond within this time-window, a time-out message was displayed, and the respective trial was repeated at the end of the decision probe phase. Participants selected choice options by pressing the “y” (left option) or “m” (right option) button (left or right index finger on MR-compatible response box in Experiment 4). Order (left/right) of choice options was counter-balanced. Participants were instructed that their shopping bag was torn, and they had to return to the shop for buying their favorite food items based on the previously learned kanjis (CS). Again, participants were told that one of the choice trials would randomly be drawn and their choice would determine which food item (US) they would receive as a bonus upon completion of the experiment. Importantly, participants were again not presented with the US related to their chosen or unchosen CS.

fMRI repetition suppression task (Experiment 4)

After Pavlovian conditioning outside the MRI-scanner, we administered two fMRI-RS blocks, one immediately before ("PRE") and one immediately after ("POST") the revaluation phase to assess choice-induced effects of fMRI repetition suppression. Every possible combination of

CS and US (eighteen combinations) was presented twenty times each (360 trials in total). In one third of the trials, the originally learned CS-US associations were presented, the remaining two third of trials contained incorrect CS-US associations. In every trial, a CS was presented for 700 ms, followed by an interstimulus interval (fixation cross) for 400 ms and a US for 700 ms. The intertrial interval was drawn from a discretized γ-distribution (shape = 2.01, scale = 1), truncated for an effective range of values between 2000 ms and 6000 ms. Order of trials was pseudo-random, between-trial repetition of CS or US did not occur. Additionally, every batch of 18 trials contained every possible combination of CS-US-association to avoid comparison of temporally distal trials and between-trial biases in repetition suppression introduced by e.g. fluctuations of attention, “novelty” or surprise. During both runs of fMRI repetition suppression, participants performed an attentional control task. After a pseudo-random 20 % of trials, participants were presented with probe trials in which they were asked to indicate whether or not the previously seen CS-US-association matched the true CS-US-association learned during Pavlovian conditioning via button presses with their right and left index fingers on an MRI-compatible response box. Correct responses were rewarded with 0.05 € and incorrect responses or time-out trials (without a response by the participant within 2500 ms after onset of the probe trial) resulted in a 0.05 € penalty which would be summed up as a bonus upon completion of the experiment. On average, participants earned a bonus of 5.93 € (SD = 1.05). Performance during the attentional control task was generally high (overall probability of correct answers, excluding time-out trials: M = .92, SD = .06 (t_{41} = 41.54, P < 0.001, one-sample t-test vs. chance level (0.5)), with no evidence for a difference in performance between PRE and POST choice-induced revaluation run (PRE: M = .92, SD = .074; POST: M = .92, SD = .072; t_{41} = 0.06, P = 0.95, paired-samples t-test).

fMRI acquisition

Two runs of fMRI were recorded with a 3 Tesla Siemens PRISMA MR-system (Siemens, Erlangen, Germany), using a 64-channel head coil. Blood oxygenation level dependent (BOLD) signals were acquired using a multi-band accelerated T2*-weighted echo-planar

imaging (EPI) sequence (multi-band acceleration factor 2, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle = 80°, field of view (FoV) = 220 mm, voxel size = 2.2 × 2.2 × 2.2 mm, no gap). Per volume, 66 slices covering the whole brain, tilted by approximately 15° in z-direction relative to the anterior–posterior commissure plane were acquired in interleaved order. The first 5 volumes of the functional imaging time series were automatically discarded to allow for T1 saturation. After each run, a B0 magnitude and phase map was acquired to estimate field maps and B0 field distortion during preprocessing (TR = 660 ms, TE 1 = 4.92 ms, TE 2 = 7.38 ms, flip angle = 60°, FoV = 220 mm). Additionally, before the pre-choice-induced revaluation fMRI-RS run, a high-resolution three-dimensional T1-weighted anatomical map (TR = 2500 ms, TE = 2.82 ms, FoV = 256 mm, flip angle = 7°, voxel size = 1 × 1 × 1 mm, 192 slices) covering the whole brain was obtained using a magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence. This scan was used as anatomical reference to the EPI data during the registration procedure.

Behavioral analyses

Data were analyzed using MATLAB 2012b and 2017a (The MathWorks, Inc., Natick, MA, USA) using custom analysis scripts. For the manipulation check, indicating learning of the CS-US-associations, average choice probabilities (CPs) for the higher valued CS were calculated (one average in Experiment 1, 2 and 4, two averages in Experiment 3) by summing up choices of the higher valued CS and dividing by the number of choice trials with a recorded response. This CP was compared against the inclusion criterion of CP ≥ 0.50. If participants had chosen the high-valued CS in more than 50 % of the trials of choice-induced revaluation, they were included in the final analyses. For the decision probe phase, we computed an average overall CP per conditioned stimulus per subject including all binary decisions in which the respective CS was present (count data, 1 representing selection of the CS, 0 representing selection of the alternative CS). Lacking a non-parametric alternative to the parametric repeated-measures analysis of variances (rmANOVA), distributions of mean overall CPs were analyzed at the group level with a rmANOVA with factors CS valence (3) and stimulus type (2) and post-hoc Wilcoxon sign-rank tests for paired samples, focusing on

the pairwise comparison of stimulus types within each level of valence. We hypothesized a main effect of CS valence and an interaction effect of CS valence × CS type (A vs. B), resulting from higher CPs for previously chosen CS relative to the equivalent control CS and lower CPs for previously unchosen CS relative to the equivalent control CS (Experiment 1, 2 and 4). However, we expected absence/no evidence for such an interaction effect in Experiment 3 in which a CS^0_A was chosen and unchosen equally often (resulting in no change of the preference relative to the control CS). Additionally, we performed one-sample Wilcoxon sign-rank tests of overall CP against chance level (CP = 0.50). As an alternative measure of choice preference in addition to overall CP, we also computed pairwise choice probabilities by comparing choice ratios of the two CS within each valence category with one-sample Wilcoxon sign-rank against chance level (CP = 0.50).

Unless specified otherwise, we considered parameter estimates of our analyses two-tailed significant if $P < 0.05$.

fMRI data analysis

We exploited fMRI repetition suppression effects (rapid, repeated presentation of the same stimulus or pre-activation of a stimulus by associated stimuli elicits reduced neural responses, as stimuli are represented by overlapping neural ensembles (Barron et al., 2016; Grill-Spector and Malach, 2001)) to investigate choice-related changes in neural representations of CS-US associations. As conditioning enhances CS's ability to pre-activate neural ensembles coding for US, we expected a change of CS-US-associative strength after choice-induced revaluation. Strong CS-US associations should elicit high fMRI-RS effects (i.e. low activation), whereas weak CS-US associations should elicit low fMRI-RS effects (i.e. high activation). We expected decreased neural representations of CS-US associations for CS^+_A and increased choice-related neural representations of CS-US associations for CS^0_A after choice-induced revaluation.

All fMRI analyses steps were performed using tools from the Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL, v6.0) (Jenkinson et al., 2012). Preprocessing included motion correction using rigid-body registration to the central

volume of the functional time series(Jenkinson et al., 2002), correction for geometric distortions using the field maps and an n-dimensional phase-unwrapping algorithm (B_0 unwarping) (Jenkinson, 2003), slice timing correction using Hanning windowed sinc interpolation, high-pass filtering using a Gaussian-weighted lines filter of 1/100 Hz. EPI images were registered with the high-resolution brain images and normalized into standard (MNI) space using affine registration linear (boundary-based registration) as well as nonlinear registration (Andersson et al., 2007a, 2007b). Functional data were spatially smoothed using a Gaussian filter with 6 mm full-width at half maximum. We applied a conservative independent components analysis (ICA) to identify and remove obvious artefacts. Independent components were manually classified as signal or noise based on published guidelines (Griffanti et al., 2017), and noise components were removed from the functional time series. General linear models (GLMs) were fitted into prewhitened data space to account for local autocorrelations (Woolrich et al., 2001). The individual level (first level) GLM design matrix per run and participant was included fifty box-car regressors in total. Thirty-six regressors coded for onsets and durations of all eighteen presented CS-US-association trials (each modelled as single events of 1800 ms duration), two regressors coded for onsets and durations of the three within-run pauses (each 45 s), two regressors coded for onsets and durations of the attentional control task probe, four regressors coded onsets and durations of left and right button presses (delta stick functions on the recorded time of response button presses) and the six volume-to-volume motion parameters from motion correction during preprocessing were entered. Regressors were convolved with a hemodynamic response function (γ-function, mean lag = 6 s, SD = 3 s). Each first level GLM included five contrasts to estimate individual per run contrasts of parameter estimates for 1) lower CS^0_A-US^0 associative strength relative to the other equivalent CS^0_B, controlling for all other combinations of CS^0 presentations (Equation 1), 2) higher CS^+_A-US^+ associative strength relative to the other equivalent CS^+_B, controlling for all other combinations of CS^+ presentations (Equation 2), 3) higher CS^-_A associative strength relative to the other equivalent CS^-_B, controlling for all other combinations of CS^- presentations, 4) Conjunction of

1) & 2), i.e. voxels coding for both decrease of CS^0_A-US^0 associative strength AND increase of CS^+_A-US^+ associative strength (Equation 3), 5) right vs. left button press. Two separate pre and post choice-induced revaluation second level (group level) GLMs were carried out by submitting individual level parameter estimates to mixed-effects statistics and ordinary least squares (OLS) regression for higher-level COPE estimation. To control for multiple comparisons, cluster-based correction with an activation threshold of $Z > 2.3$ using a cluster-extent threshold of $P < 0.05$ was applied at the whole-brain level.

The key tests for our hypothesis were focused on the effect of revaluation choices on CS-US-associations during the POST run. The PRE run served as a control to rule out potential baseline differences in associative strength for CS^0_A or CS^+_A.

A priori regions-of-interest (ROIs) comprised the lateral orbitofrontal cortex (lOFC) and the hippocampus, as those regions have been implicated in representation and adaptive changes of stimulus-outcome associations, respectively (Boorman et al., 2016; Garvert et al., 2017; Jocham et al., 2016; Klein-Flugge et al., 2013). We investigated post choice-induced revaluation conjunction effects (Equation 9) to identify regions involved in processing of choice-induced changes to CS-US-associations. An independent functional mask of a contrast investigating stimulus-outcome-associations from a previous study (Jocham et al., 2016) (restricted along the z-direction from –6 to –14 to constrain spatial extent), was used for small-volume correction (P_{SVC}) of the bilateral lOFC. The small-volume corrected functional activation mask from the conjunction contrast was used to extract contrast parameter estimates of the CS^0_A contrast and the CS^+_A contrast. Additionally, an independent anatomical mask of the hippocampus (Harvard-Oxford Atlas) was used to extract pre and post choice-induced revaluation contrast parameter estimates of the CS^0_A contrast and the CS^+_A contrast. PRE versus POST comparisons of activation were carried out using a repeated-measures ANOVA and post-hoc Wilcoxon-sign rank tests. Additionally, we explored functional activation not surviving whole-brain or small-volume family-wise error corrections by thresholding activation maps at $Z = 2.8$ and extracting contrast estimates of the CS^0_A contrast and the CS^+_A contrast for brain-behavioral correlations.

Extracted parameter estimates were additionally used for brain-behavioral correlations using non-parametric Spearman correlations. For brain-behavior correlations, we had the directed hypotheses that associative strength should be positively related to the difference between overall CP of CS^{+}_{A} and overall CP CS^{+}_{B} and negatively related to difference between overall CP CS^{0}_{A} and overall CP CS^{0}_{B}. Additionally, to refine our insights in the relationship of neural and behavioral results, we also correlated associative strength with pairwise CP for CS^{+}_{A} versus CS^{+}_{B}, again assuming a positive relationship and pairwise CP for CS^{0}_{A} versus CS^{0}_{B}, predicting a negative relationship. Due to our directed hypotheses, and because parameter estimates were extracted from family-wise error corrected ROIs, we considered parameter estimates of our analyses one-tailed significant if $P < 0.05$, that the observed parameter value or a more extreme value could have been obtained under assumption of the null-hypothesis (no relationship between associative strength and CPs).

Furthermore, whole-brain voxel-wise regressions were applied at the group level for both PRE and POST run and also a group level analysis on the activation change from PRE to POST run. We used demeaned individual behavioral difference regressors of overall CP of CS^{+}_{A} – overall CP CS^{+}_{B} and overall CP CS^{0}_{A} – overall CP CS^{0}_{B} and demeaned individual pairwise CP of trials directly comparing $CS^{+}_{A} – CS^{+}_{B}$ and $CS^{0}_{A} – CS^{0}_{B}$, respectively.

For our proposed associative mechanism, we assumed that after Pavlovian conditioning, a CS would pre-activate the respective, associatively learned US and participants would make their decisions between CS based on the associated outcome. Our hypothesized mechanism relies on the assumption that participants form a (simple) model of the task, which is well in line with the literature on associative learning (Barron et al., 2013; Klein-Flugge et al., 2013). However, alternatively, the observed pattern of results could also be explained by a “cached value” account. According to this account, participants acquire “cached” values during Pavlovian conditioning and further use those “model-free” values to guide their decision, independent of associated outcomes or the learned relationships between CS and US. Our study was not designed to dissociate both mechanisms on a behavioral level. The prediction of the “cached value” account would be that decisions during

choice-induced revaluation lead to changes in “cached values” and leave CS-US association unaffected. In other words, CS^0_A should have a reduced “cached value”, whereas the “cached value” of CS^+_A should be increased following choice-induced revaluation. In another fMRI contrast, we tested whether the relationship/similarity between CS^0_A and US^- (Equation 4) would change due to the choice-induced revaluation phase. If our results can be explained by the “cached value” account, we would assume that CS^0_A followed by US^- would show reduced activation/higher similarity in the post choice-induced revaluation run, compared to its equivalent partner stimulus CS^0_B followed by US^-. As any other stimulus- or outcome-related effects are controlled for in the contrast, and CS^0_A was not learned to be associated with US^- during Pavlovian conditioning (i.e. not being able to pre-activate a US^- representation), we assumed that reductions in activation could only be interpreted as a repetition of a shared feature, namely the stimulus/outcome value. After testing the contrast, we extracted parameter estimates from an independent hippocampus mask and correlated the parameter estimates with CP. We hypothesized that CS^0_A and US^- similarity would be negatively related to overall CP CS^0_A – overall CP CS^0_B. The lower the CS^0_A and US^- similarity, the less likely participants would be to select CS^0_A.

fMRI contrasts

CS^0_A fMRI-RS contrast:

$$[\,2 \times (CS^0_A\text{-}US^0 - CS^0_B\text{-}US^0)\,] - [\,(CS^0_A\text{-}US^- - CS^0_B\text{-}US^-) + (CS^0_A\text{-}US^+ - CS^0_B\text{-}US^+)\,] \qquad (1)$$

CS^+_A fMRI-RS contrast:

$$[\,2 \times (CS^+_A\text{-}US^+ - CS^+_B\text{-}US^+)\,] - [\,(CS^+_A\text{-}US^- - CS^+_B\text{-}US^-) + (CS^+_A\text{-}US^0 - CS^+_B\text{-}US^0)\,] \qquad (2)$$

Conjunction fMRI-RS contrast:

$$[\,2 \times (CS^0_A\text{-}US^0 - CS^0_B\text{-}US^0)\,] - [\,(CS^0_A\text{-}US^- - CS^0_B\text{-}US^-) + (CS^0_A\text{-}US^+ - CS^0_B\text{-}US^+)\,]$$

AND (3)

$$[\,2 \times (CS^+_A\text{-}US^+ - CS^+_B\text{-}US^+)\,] - [\,(CS^+_A\text{-}US^- - CS^+_B\text{-}US^-) + (CS^+_A\text{-}US^0 - CS^+_B\text{-}US^0)\,]$$

“Cached” value control analysis fMRI-RS contrast:

$$[\,2 \times (CS^0_B\text{-}US^- - CS^0_A\text{-}US^-)\,] - [\,(CS^0_A\text{-}US^0 - CS^0_B\text{-}US^0) - (CS^0_B\text{-}US^+ - CS^0_A\text{-}US^+)\,] \qquad (4)$$

Computational models

To formally characterize behavior in our experiments, we fit six different variants of a reinforcement learning model using Rescorla-Wagner-like delta update rules (Rescorla and Wagner, 1972). For each experiment, we compared the six models that implemented different ways by which participants could have learned CS-US associative strength - and updated associative strength during choice-induced revaluation.

Model 1: Pavlovian learning only

In this model, CS values are exclusively acquired during Pavlovian learning, without any further update during the choice-induced revaluation. On each trial, the value of the stimulus currently presented was updated according to the following rule:

$$V_{t+1,i} = V_{t,i} + \alpha_{Learning}\left(R_t - V_{t,i}\right) \quad (5)$$

where $\alpha_{Learning}$ is the learning rate, $V_{t,i}$ is the value of the ith stimulus (1 to 6 for the six CS), and R_t is the reward value of the US (0, 0.5 and 1 for low-, intermediate, and high-value outcome, respectively, and 0 if no outcome was presented) in the Pavlovian conditioning phase. CS values were initialized at 0.5. The estimated associative strength for each CS after the learning phase was directly passed to a softmax choice rule (Equation 7), without further modulation of CS associative strength by choice-induced revaluation.

Model 2: Pavlovian learning and choice-induced revaluation (chosen CS)

Model 2 acquired stimulus values during Pavlovian learning exactly like model 1, but additionally updated CS-US associative strength by “fictive reward prediction errors” elicited by decisions in the choice-induced revaluation phase. The “fictive reward prediction errors” was based on our reasoning that presentation of CS during revaluation would lead to retrieval of the associated US and that, consistent with our hypothesis, stimulus-outcome association of the chosen CS would be strengthened, whereas the association of the non-chosen CS would be weakened. As no objective feedback (US) was presented during choice-induced revaluation, the reward prediction could only be derived by assuming associative retrieval. This model only updated associative strength of the chosen, but not the unchosen CS.

$$V_{t+1,ch} = V_{t,ch} + \alpha_{ch}\left(R_{t,ch} - V_{t,ch}\right) \quad (6)$$

where α_{ch} is the choice revaluation learning rate scaling the impact of “fictive" reward prediction errors elicited by fictive outcomes $R_{t,\,ch}$ (1 for chosen CS and −1 for unchosen CS) elicited by the US associated with each CS.

Model 3: Pavlovian learning and choice-induced revaluation (chosen CS)

This model was set up to account for the possibility that updating of both the chosen and unchosen CS association could occur during the choice-induced revaluation. It was identical to model 2, with the exception that it performed an update to the associative strength of the unchosen stimulus, using a separate learning rate α_{unch}, in addition to updating the chosen CS associative strength.

These same three models were set up as variants ("associative value models") that were identical in all respects except for (1) the outcomes during Pavlovian learning, which were modelled as 0 (no outcome presented) or 1 (outcome presented) and CS associative strength values were scaled with the normalized (0-1), individually rated subjective value of the outcome (pre-rating) at the end of the Pavlovian learning phase.

The estimated associative strengths for all CS after the choice-induced revaluation phase were passed to a softmax decision function to generate choice probabilities for each option on each trial:

$$P_{C,t} = \frac{1}{1 + \exp(-VD_t \,/\, \tau)} \qquad (7)$$

where $P_{C,t}$ is the model's probability to select option C on trial t, the choice the participant actually made on trial t. $VD_{t,}$ is the value difference (or difference in associative strength) between the chosen and unchosen CS on trial t, and τ is a temperature parameter that determines the degree of stochasticity in participants' choice behavior.

To find the free parameters that best described participants' behavior, we used a two-step fitting procedure to minimize the negative log likelihood estimate (−LLE):

$$-LLE = -\Sigma \log\left(P_{C,t}\right) \tag{8}$$

First, we ran a grid search on an n-dimensional grid in log space (where n = number of free parameters for each model), with 30 steps along each dimension. The grid optimum was then used as initial value and passed to constrained non-linear optimization using the MATLAB function fmincon. Optimized negative log likelihoods were compared by means of the sample-size corrected Akaike Information Criterion (AICc, Equation 9 and 10). The model with the lowest AICc value was considered to account best for the observed participants' choice data, penalized for model complexity and sample size (number of participants).

$$\mathrm{AIC} = 2k - 2(-LLE) \tag{9}$$

Where k is the number of free parameters of the model and −LLE is the negative log likelihood of the model given the data.

$$\mathrm{AIC_C} = \mathrm{AIC} + \frac{2k^2 + 2k}{n - k - 1} \tag{10}$$

Where k is the number of free parameters in the model, and n is the sample size.

Acknowledgments

The authors thank all volunteers who participated in this study. The authors would like to thank Halla Mulla-Osman, Stefanie Linnhoff and Denise Scheermann for their invaluable support during data acquisition. We thank the Center for Magnetic Resonance Research of the University of Minnesota for providing the Multiband accelerated fMRI sequence. The authors additionally thank Theo O.J. Gruendler, Lukas M. Neugebauer and Daniel Priegnitz for helpful discussions during set-up of the study and for sharing custom MATLAB code. **Funding:** This work was funded by the federal state of Saxony-Anhalt and the „European Regional Development Fund“ (ERDF 2014-2020), Vorhaben: Center for Behavioral Brain Sciences (CBBS), FKZ: ZS/2016/04/78113; **Author contributions:** LL designed the study and conceptualized research, acquired the data, analyzed the data, drafted the manuscript, read and edited versions of the manuscript and approved the final version of the manuscript. CT set up the MRI acquisition protocol, read and edited versions of the manuscript and approved the final version of the manuscript. LFK read and edited versions of the manuscript and approved the final version of the manuscript. GJ designed the study and conceptualized research, analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript. **Competing interests:** Authors declare no competing interests.

References

Anderson, M.C., Bjork, R.A., Bjork, E.L., 1994. Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory. J. Exp. Psychol. Learn. Mem. Cogn. 20, 1063–1087. doi:10.1037/0278-7393.20.5.1063

Andersson, J.L.., Jenkinson, M., Smith, S., 2007a. Non-linear registration aka spatial normalisation. FMRIB Tech. Rep. TRO7JA2 22.

Andersson, J.L.., Jenkinson, M., Smith, S.M., 2007b. Non-linear optimisation. FMRIB Tech. Rep. TRO7JA1.

Ariely, D., Norton, M.I., 2008. How actions create - not just reveal - preferences. Trends Cogn. Sci. 12, 13–16. doi:10.1016/j.tics.2007.10.008

Barron, H.C., Dolan, R.J., Behrens, T.E.J., 2013. Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. 16, 1492–1498. doi:10.1038/nn.3515

Barron, H.C., Garvert, M.M., Behrens, T.E.J., 2016. Repetition suppression: a means to index neural representations using BOLD? Philos. Trans. R. Soc. B Biol. Sci. 371, 20150355. doi:10.1098/rstb.2015.0355

Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., Mark, S., Baram, A.B., Stachenfeld, K.L., Kurth-Nelson, Z., 2018. What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. Neuron 100, 490–509. doi:10.1016/j.neuron.2018.10.002

Blechert, J., Meule, A., Busch, N.A., Ohla, K., 2014. Food-pics: An image database for experimental research on eating and appetite. Front. Psychol. 5. doi:10.3389/fpsyg.2014.00617

Boorman, E.D., Rajendran, V.G., O'Reilly, J.X., Behrens, T.E., 2016. Two Anatomically and Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome Associations in Hippocampus. Neuron 89, 1343–1354. doi:10.1016/j.neuron.2016.02.014

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Brehm, J.W., 1956. Postdecision changes in the desireability of alternatives. J. Abnorm. Soc. Psychol. 52, 384–389. doi:10.1037/h0041006

Chen, M.K., Risen, J.L., 2010. How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. J. Pers. Soc. Psychol. 99, 573–594. doi:10.1037/a0020217

Garvert, M.M., Dolan, R.J., Behrens, T.E., 2017. A map of abstract relational knowledge in

the human hippocampal–entorhinal cortex. Elife 6, 1–20. doi:10.7554/elife.17086

Griffanti, L., Douaud, G., Bijsterbosch, J., Evangelisti, S., Alfaro-Almagro, F., Glasser, M.F., Duff, E.P., Fitzgibbon, S., Westphal, R., Carone, D., Beckmann, C.F., Smith, S.M., 2017. Hand classification of fMRI ICA noise components. Neuroimage 154, 188–205. doi:10.1016/J.NEUROIMAGE.2016.12.036

Grill-Spector, K., Malach, R., 2001. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst). 107, 293–321. doi:10.1016/S0001-6918(01)00019-1

Howard, J.D., Kahnt, T., Gottfried, J.A., 2016. Converging prefrontal pathways support associative and perceptual features of conditioned stimuli. Nat. Commun. 7, 1–11. doi:10.1038/ncomms11546

Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., Matsumoto, K., 2010. Neural correlates of cognitive dissonance and choice-induced preference change. Proc. Natl. Acad. Sci. 107, 22014–22019. doi:10.1073/pnas.1011879108

Jenkinson, M., 2003. Fast, automated,N-dimensional phase-unwrapping algorithm. Magn. Reson. Med. 49, 193–197. doi:10.1002/mrm.10354

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. Neuroimage 17, 825–841. doi:10.1006/NIMG.2002.1132

Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. Neuroimage 62, 782–790. doi:10.1016/J.NEUROIMAGE.2011.09.015

Jocham, G., Brodersen, K.H.H., Constantinescu, A.O.O., Kahn, M.C.C., Ianni, A.M., Walton, M.E.E., Rushworth, M.F.F.S., Behrens, T.E.E.J., 2016. Reward-Guided Learning with and without Causal Attribution. Neuron 90, 177–190. doi:10.1016/j.neuron.2016.02.018

Klein-Flugge, M.C., Barron, H.C., Brodersen, K.H., Dolan, R.J., Behrens, T.E.J., 2013. Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex. J. Neurosci. 33, 3202–3211. doi:10.1523/JNEUROSCI.2532-12.2013

Liu, Y., Dolan, R.J., Kurth-Nelson, Z., Behrens, T.E.J., 2019. Human Replay Spontaneously Reorganizes Experience. Cell 178, 1–13. doi:10.1016/j.cell.2019.06.012

Nader, K., Hardt, O., 2009. A single standard for memory□: the case for reconsolidation. Nat. Rev. Neurosci. 10, 224–234. doi:10.1038/nrn2590

Nader, K., Schafe, G.E., Doux, J.E. Le, 2000. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature 406, 722–726.

Rescorla, R.A., Wagner, A.R., 1972. A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement, in: Black, A.H., Prokasy, W.F. (Eds.), Classical Conditioning II: Current Research and Theory. Appleton-Century-Crofts, New York, pp. 64–99.

Riefer, P.S., Prior, R., Blair, N., Pavey, G., Love, B.C., 2017. Coherency-maximizing exploration in the supermarket. Nat. Hum. Behav. 1, 1–4. doi:10.1038/s41562-016-0017

Sharot, T., Velasquez, C.M., Dolan, R.J., 2010. Do decisions shape preference? Evidence from blind choice. Psychol. Sci. 21, 1231–1235. doi:10.1177/0956797610379235

Summerfield, C., Luyckx, F., Sheahan, H., 2019. Structure Learning and the Parietal Cortex. PsyArXiv. doi:10.31234/osf.io/zfxj2

Tolman, E.C., 1948. Cognitive Maps in Rats and Man. Psychol. Rev. 55, 189–208.

Tonegawa, S., Morrissey, M.D., Kitamura, T., 2018. The role of engram cells in the systems consolidation of memory. Nat. Rev. Neurosci. 19, 485–498. doi:10.1038/s41583-018-0031-2

Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rudebeck, P.H., Rushworth, M.F.S., 2010. Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. Neuron 65, 927–939. doi:10.1016/j.neuron.2010.02.027

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., Anderson, M.C., 2015. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. Nat. Neurosci. 18, 582–589. doi:10.1038/nn.3973

Wimmer, G.E., Shohamy, D., 2012. Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. Science 338, 270–3. doi:10.1126/science.1223252

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. Neuroimage 14, 1370–1386. doi:10.1006/NIMG.2001.0931

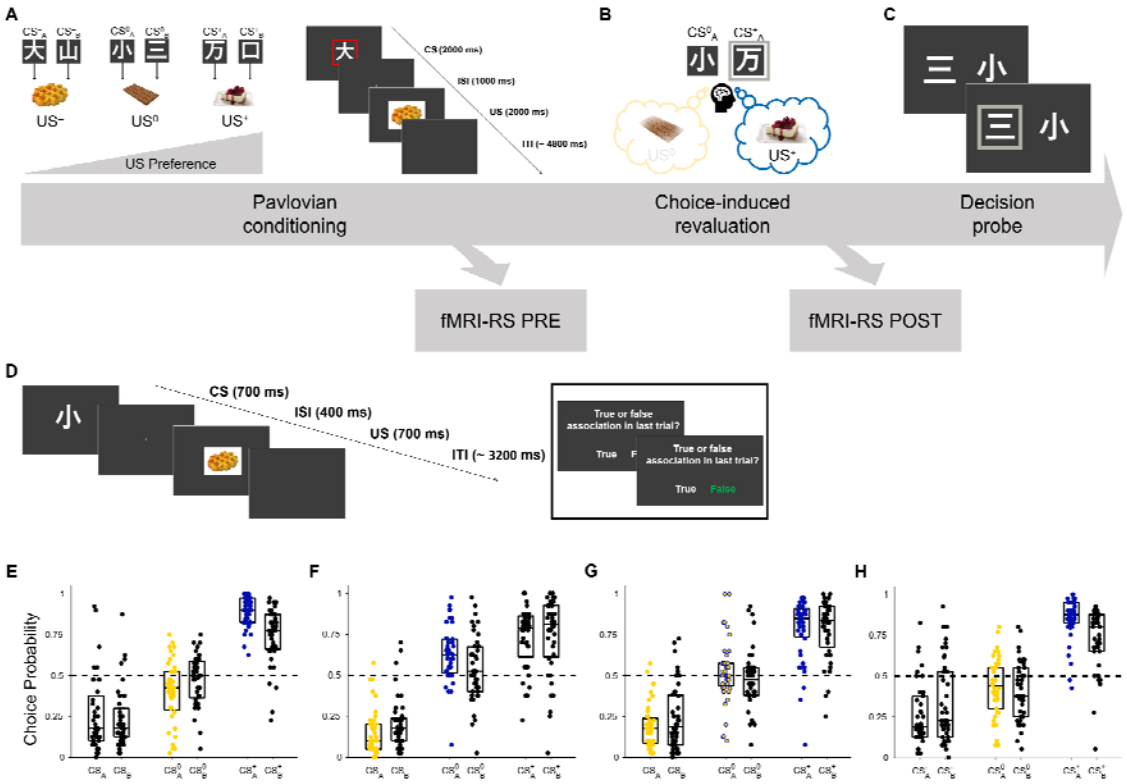


Fig. 1. Task schematic and behavioral results.

A) Participants rated subjective desirability of conditioned stimuli (CS, kanjis) and unconditioned stimuli (US, food items). During Pavlovian conditioning, participants learned to associate six CS with three US. Each US was associated with two CS. B) Choice-induced revaluation: After Pavlovian conditioning, participants made choices between CS^{+}_{A} versus CS^{0}_{A} (Experiments 1 & 4), CS^{0}_{A} versus CS^{-}_{A} (Experiment 2), and CS^{0}_{A} versus either CS^{-}_{A} or CS^{+}_{A} (Experiment 3), interleaved with lure decisions. C) Decision probe: Following choice-induced revaluation, participants made binary choices between all possible combinations of CS to assess preferences. D) Attentional control task performed during fMRI repetition suppression (Experiment 4). E-H) Previously chosen CS (blue scatter) are selected more often compared to equivalent CS (black scatter) in Experiments 1, 2, 4 (E, F, H) and previously non-chosen CS (yellow scatter) are selected less often compared to equivalent CS (black scatter) in Experiment 1, 2 (E, F) during decision probe. The effect is not present in Experiment 3 (G), indicating that the roughly equal proportion of choices and non-choices of CS^{0}_{A} during revaluation had cancelled each other out.

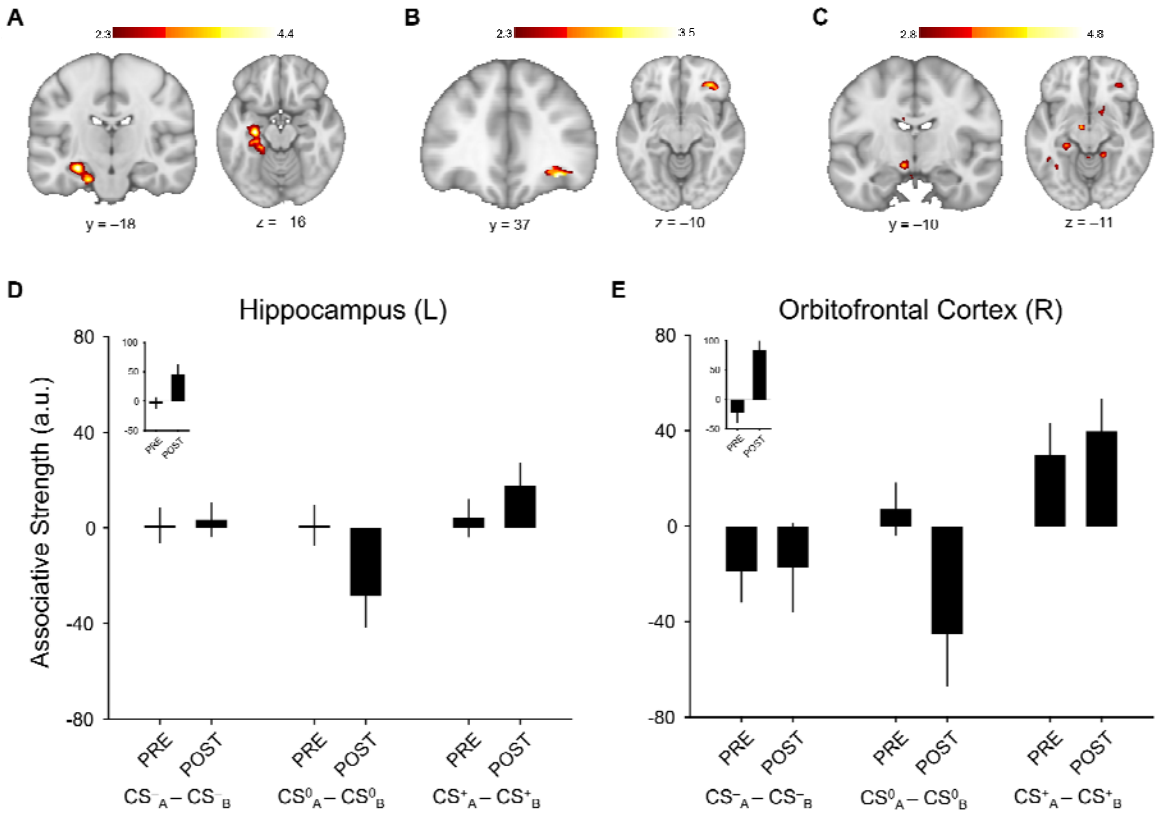


Fig. 2. Whole-brain effects of choice-induced revaluation.

A) Left hippocampus encodes conjunction of decreased associative strength for CS^0_A-US^0 relative to CS^0_B-US^0 (Equation 1) AND increased CS^+_A-US^+ associative strength relative to CS^+_B-US^+ (Equation 2). B) Using small-volume correction (independent mask from Jocham et al., 2016), the same effect was found in the right lateral orbitofrontal cortex (lOFC). C) At a lenient threshold of $Z > 2.8$ (uncorrected) we observed the same effect in left ventral tegmental area and right nucleus accumbens. D) Extracted parameter estimates of the effect in left hippocampus. E) Extracted parameter estimates of the effect in right lOFC.

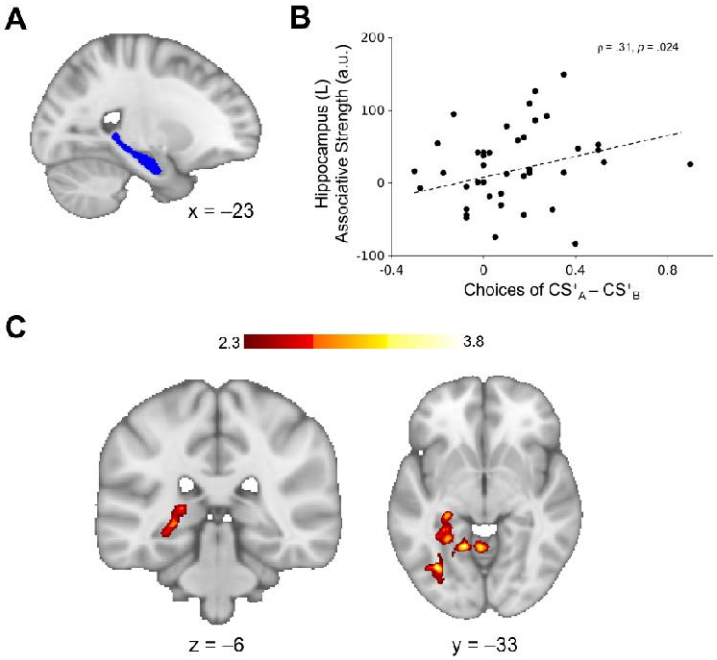


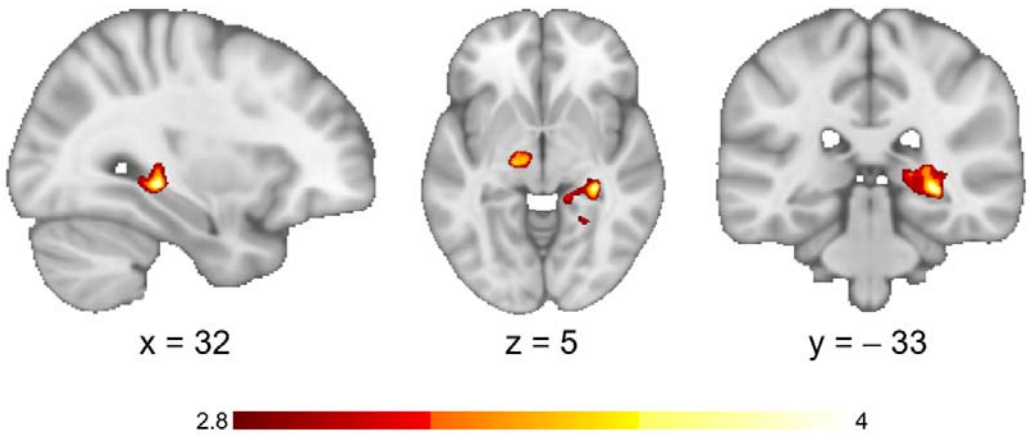
Fig. 3. Brain-behavior correlation and whole-brain regressions.

B) Left hippocampus associative strength of CS^{+}_{A}-US^{+} relative to CS^{+}_{B}-US^{+} (Equation 3), extracted from an independent anatomical mask (A) during POST was positively related to decision probe difference between overall choice probability of CS^{+}_{A} and CS^{+}_{B}. With increasing associative strength, participants were more likely to prefer CS^{+}_{A} over CS^{+}_{B}. C) Whole-brain regression showing a positive relationship between the difference of CS^{+}_{A} and CS^{+}_{B} preference and PRE-POST change of associative strength in the left posterior hippocampus. With more positive PRE-POST associative strength change, participants are more likely to prefer CS^{+}_{A} over CS^{+}_{B}.

Supplementary Materials

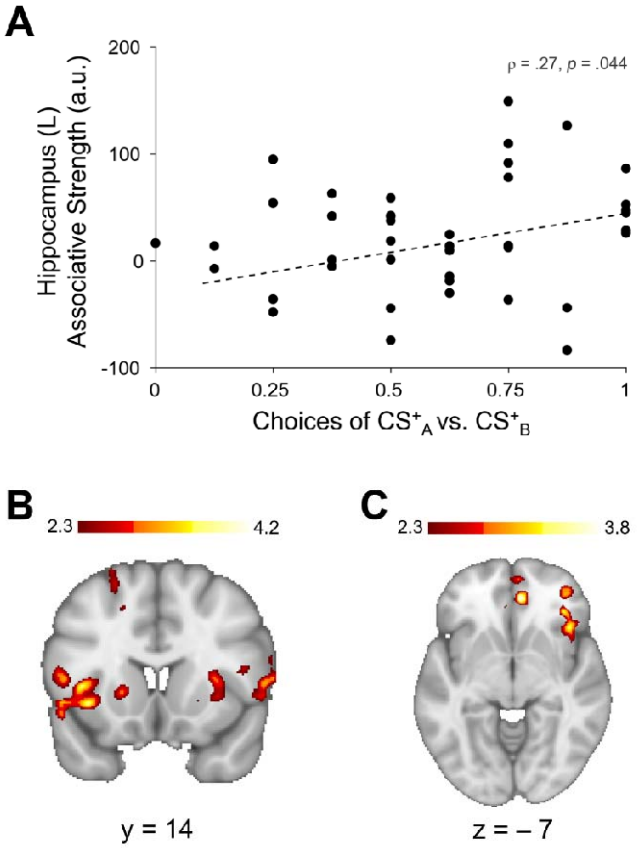
Decisions bias future choices by modifying hippocampal associative memories

Lennart Luettgau, Claus Tempelmann, Luca Franziska Kaiser & Gerhard Jocham



Supplementary Fig. S1. "Cached value" control analysis results.

Whole-brain corrected pattern of activation (activation threshold $Z > 2.3$, cluster-forming threshold $P < 0.05$, thresholded at $Z > 2.8$ for display purposes) for the cached value control analysis during the POST choice-induced revaluation fMRI run. Two clusters in the right posterior hippocampus and in a left midbrain region in the vicinity of the dorsolateral substantia nigra pars compacta survived whole-brain correction. Extracting hippocampal activation with an independent anatomical right hippocampus mask and comparison of parameter estimates for PRE and POST yielded no significant difference ($Z = 0.91$, $P = 0.37$), suggesting no evidence for an influence of choice-induced revaluation decision on activation. Additionally, we did not observe significant correlations between the cluster in the right posterior hippocampus from the "cached value" control analysis and the left hippocampus result from the associative analyses (all $\rho s < .18$, $Ps > 0.26$), suggesting different processes. Parameter estimates from the right posterior hippocampus were only marginally related with overall CP CS^0_A – overall CP CS^0_B: $\rho_{40} = .30$, $P = 0.050$ (two-tailed) and the within-category CP CS^0_A vs. CS^0_B: $\rho_{40} = .29$, $P = 0.062$ (two-tailed). However, the direction of the correlation was exactly opposite to the predictions of the cached value account. As there was no apriori anatomical hypothesis for the SNpc, we were unable to extract parameter estimates in an unbiased fashion. Thus, PRE and POST parameter estimates could not be compared. Left SNpc parameter estimates were not significantly related to left VTA parameter estimates from the associative analysis (all $\rho s < .08$, $Ps > 0.63$), nor with decision probe behavior (all $\rho s < .20$, $Ps > 0.20$).



Supplementary Fig. S2. Brain-behavior correlation and whole-brain regressions.

A) Left hippocampus parameter estimates (independent anatomical mask, Fig. 3A) for associative strength of CS^{+}_{A} relative to CS^{+}_{B} (see Equation 3) during the POST choice-induced revaluation fMRI run positively predict decision probe probability of CS^{+}_{A} vs. CS^{+}_{B}. The higher associative strength was after choice-induced revaluation, the more likely participants were to prefer CS^{+}_{A} over CS^{+}_{B}. B) CS^{+}_{A} preference vs. CS^{+}_{B} was predicted by post choice-induced revaluation run associative strength in the bilateral anterior insula, suggesting choice-induced strengthening of CS-related pre-activation of neural ensembles coding for the used food items. C) We observed a positive relationship between PRE-POST changes in associative strength of CS^{0}_{A} and pairwise CP for CS^{0}_{A} vs. CS^{0}_{B}, in two clusters in the medial orbitofrontal cortex and lOFC, extending to the anterior insula.

Supplementary Table S1.

Exp. 1	Update	Model	$\alpha_{Learning}$	α_{ch}	α_{unch}	τ	–LLE	AIC_c
	RW	Model 1	< .001 (0 – 1)		-	.0002 (0 – 1.55)	63.43 (37.50 – 81.58)	133.59
		Model 2	.04 (0 – 1)	< .001 (0 – .78)	-	.20 (0 – 1.48)	40.52 (26.19 – 80.67)	98.10
		Model 3	.02 (0 – .58)	.50 (0 – .85)	.006 (0 – .86)	.15 (0 – .77)	37.27 (25.62 – 79.88)	**94.62**
	Asso. Value	Model 1	.002 (0 – 1)		-	.11 (.05 – 2.65)	41.35 (27.38 – 80.9)	97.10
		Model 2	.02 (0 – 1)	.002 (0 – .84)	-	.14 (.06 – 1.43)	40.60 (20.64 – 77.37)	95.87
		Model 3	.03 (0 – 1)	.42 (0 – .84)	.005 (0 – 1)	.17 (.06 – 3.37)	37.08 (20.60 – 81.88)	96.30
Exp. 2	RW	Model 1	.55 (0 – .95)		-	.45 (0 – 66.59)	53.41 (31.86 – 83.18)	115.99
		Model 2	.18 (0 – .72)	.005 (0 – .95)	-	.24 (0 – 3.44)	45.42 (15.43 – 82.02)	103.18
		Model 3	.04 (0 – .85)	.02 (0 – .88)	.005 (0 – .99)	.24 (0 – .77)	44.67 (5.55 – 77.11)	98.98
	Asso. Value	Model 1	.003 (0 – 1)		-	.15 (.05 – 2.65)	48.77 (15.81 – 83.23)	101.68
		Model 2	.04 (0 – 1)	.005 (0 – .92)	-	.20 (.002 – 6.11)	44.53 (15.71 – 82.81)	100.51
		Model 3	.03 (0 – 1)	.02 (0 – .99)	.005 (0 – .74)	.20 (.004 – 2.08)	44.38 (5.55 – 82.4)	**97.43**
Exp. 3	RW	Model 1	.03 (0 – .88)	-		.18 (0 – 1.55)	50.60 (22.15 – 82.66)	107.92
		Model 2	.02 (0 – .85)	< .001 (0 – .70)	-	.07 (0 – 16.92)	41.06 (14.50 – 82.99)	95.31
		Model 3	.01 (0 – .83)	.009 (0 – .55)	< .001 (0 – .40)	.08 (0 – 24.81)	40.47 (13.64 – 83.07)	93.15
	Asso. Value	Model 1	.005 (0 – 1)	-		.11 (0 – 1.44)	41.12 (13.65 – 82.64)	93.78
		Model 2	.02 (0 – 1)	.008 (0 – .76)	-	.16 (0 – 1.28)	40.29 (13.65 – 82.62)	93.84
		Model 3	.02 (0 – 1)	.012 (0 – .86)	.009 (0 – .80)	.13 (.004 – 2.01)	44.38 (5.55 – 82.4)	**91.96**
Exp. 4	RW	Model 1	< .001 (0 – 1)	-		< .001 (0 – 3.89)	71.61 (49.60 – 82.72)	143.94
		Model 2	< .001 (0 – .85)	< .001 (0 – 1)	-	.16 (0 – 5.05)	54.25 (22.54 – 82.64)	114.94
		Model 3	.04 (0 – .61)	.52 (0 – .85)	.008 (0 – .72)	.24 (0.03 – 1.02)	51.11 (22.54 – 81.86)	110.71
	Asso. Value	Model 1	.002 (0 – 1)	-		.16 (0.04 – 4.78)	55.11 (23.54 – 82.86)	112.97
		Model 2	.13 (0 – 1)	.007 (0 – 1)	-	.24 (0.05 – 4.81)	50.79 (23.37 – 82.61)	110.69
		Model 3	.18 (0 – 1)	.50 (0 – 1)	.009 (0 – .97)	.22 (.04 – 2.98)	49.25 (22.90 – 82.51)	**107.85**

Note: Median and range of computational parameters, negative log likelihood estimates (–LLE) and corrected Akaike Information Criteria (AICc) of the Rescorla-Wagner like (RW)(Rescorla and Wagner, 1972) and Associative Value (Asso. Value) reinforcement learning models for the four experiments. Lowest AIC_c per experiment, which was considered to account best for the observed participants' choice data, penalized for model complexity and sample size (number of participants) in bold letters.