Reinstatement of cortical outcome representations during higher-order learning

Lennart Luettgau[1,2*], Emanuele Porcu[3], Claus Tempelmann[4], Gerhard Jocham[1,2]

[1]Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University, Düsseldorf, Germany
[2]Center for Behavioral Brain Sciences, Otto-von-Guericke University, Magdeburg, Germany
[3]Department of Biological Psychology, Otto-von-Guericke University, Magdeburg, Germany
[4]Department of Neurology, Otto-von-Guericke University, Magdeburg, Germany
*Correspondence to: luettgau@hhu.de

Abstract

Naturalistic learning scenarios are characterized by infrequent experience of external feedback to guide behavior. Higher-order learning mechanisms like second-order conditioning (SOC) may allow stimuli that were never experienced together with reinforcement to acquire motivational value in such environments. Despite its explanatory potential for real-world learning phenomena, surprisingly little is known about the neural mechanism underlying such associative transfer of motivational value in SOC. Here, we used multivariate cross-session, cross-modality searchlight classification on functional magnetic resonance imaging data obtained from healthy humans during SOC. We show that visual first-order conditioned stimuli (CS) reinstate cortical patterns representing previously paired gustatory outcomes in the lateral orbitofrontal cortex (OFC). The same OFC region was found to increase functional coupling with amygdala, anterior hippocampus, and medial OFC during SOC. In the amygdala, neural pattern similarity specifically increased between second-order CS and outcomes from early to late stages of SOC. Our data suggest a mechanism by which value is conferred to stimuli that were never paired with reinforcement.

Introduction

Learning in naturalistic settings is characterized by infrequent direct encounters with rewarding or punishing stimuli (Gewirtz and Davis, 2000). Hence, several stimuli or actions often need to be chained together, such that one stimulus serves as a proxy for another stimulus that might eventually predict reward (Gewirtz and Davis, 2000). Exploiting such statistical regularities would allow agents to learn about the value of stimuli or actions that were never directly followed by reinforcement and support instrumental behavior in the absence of reinforcement or in new contexts. Second-order conditioning (SOC) is an example of higher-order learning that allows exactly such associative transfer to stimuli that have never been directly paired with reinforcement (Sharpe et al., 2017). In SOC, a first-order conditioned stimulus (CS_1) is first paired with a motivationally salient event or stimulus (unconditioned stimulus, US). By virtue of this pairing, the CS_1 is capable of evoking conditioned responses (CR, e.g. salivating, as in the presence of food), enabling it to function as conditioned reinforcer (Sharpe et al., 2017). Subsequently, another previously neutral conditioned stimulus (CS_2) is paired with the CS_1. Thereby, CS_2 acquires incentive properties and is, like CS_1, now able to elicit a CR. Despite its relevance for real-life learning phenomena (Gewirtz and Davis, 2000; Parkes and Westbrook, 2011) and decades of behavioral investigations (Barnet et al., 1991; Gewirtz and Davis, 2000; Parkes and Westbrook, 2011; Rizley and Rescorla, 1972), surprisingly little is known about the neural mechanism underlying associative transfer of motivational value in SOC (Parkes and Westbrook, 2011).

At least four potential mechanisms accounting for SOC have been proposed (Barnet et al., 1991; Gewirtz and Davis, 2000; Parkes and Westbrook, 2011; Rizley and Rescorla, 1972). Three of these suggest that CS_2 could become associated with a CR using (i) the associative link between CS_2 and CS_1, (ii) direct pairing of the CS_1-evoked CR with CS_2, or (iii) because CS_1 reactivates a representation of the US which then evokes a CR that becomes paired with CS_2. Behavioral evidence from previous studies suggests that neither of these three hypotheses can account for second-order learning effects (Barnet et al., 1991;

Rizley and Rescorla, 1972). This leaves open a fourth possibility: that CS_2 is directly paired with a neural representation of the US, or more precisely, with the motivational state conveyed by the US during second-order learning ("direct link" hypothesis) (Gewirtz and Davis, 2000; Parkes and Westbrook, 2011). Since the "direct link" hypothesis of second-order conditioning is the only remaining explanatory attempt and has never been tested to date, the present study sought to provide evidence that CS_2 is directly paired with a neural representation of the US. Recent evidence suggests that acquisition of incentive properties by second-order stimuli (CS_2) is related to activity of mesencephalic dopamine neurons that encode a reward prediction error elicited by CS_1 (Maes et al., 2020). Presumably, SOC critically depends on convergence of neural patterns representing CS_2 and US in the amygdala (Gewirtz and Davis, 2000, 1997; Parkes and Westbrook, 2011; Setlow et al., 2002) and hippocampus (Gilboa et al., 2014). However, it remains unclear, how exactly a "direct link" between CS_2 and US representations is formed at the neural level. Additionally, to the best of our knowledge, it is unknown whether human value-based decision-making is affected by motivational value acquired through SOC. This complicates cross-species translation of higher-order learning effects described in animal studies and highlights the need for using similar paradigms in humans.

Building on models of memory reinstatement (Tonegawa et al., 2018), we hypothesized that during SOC, presentation of the CS_1 would trigger reinstatement of the neural pattern representing the US with which it had previously been paired during first-order conditioning (FOC). This would allow linkage of "online" US and CS_2 representations by associative plasticity in amygdala (Gewirtz and Davis, 2000, 1997; Hatfield et al., 1996; Parkes and Westbrook, 2011; Setlow et al., 2002) and hippocampus (Gilboa et al., 2014). Both amygdala and hippocampus, as well as orbitofrontal cortex (OFC) and ventral striatum (Mcdannald et al., 2013) have consistently been shown as the key structures involved in second-order learning and conditioned reinforcement (Gewirtz and Davis, 2000, 1997; Gilboa et al., 2014; Hatfield et al., 1996; Parkes and Westbrook, 2011; Setlow et al., 2002). Our analyses therefore focused on these regions.

To test our predictions, we combine a SOC paradigm and choice preference tests with cross-session (Stokes et al., 2009), cross-modality searchlight (Kriegeskorte et al., 2006) classification of functional magnetic resonance imaging (fMRI) data in healthy human participants. Participants first established Pavlovian associations between visual CS_1 and appetitively or aversively valued gustatory US. During SOC, participants were exposed to associations between visual CS_2 and previously learned first-order CS_1. In a subsequent preference test phase, participants were more likely to select directly (CS_1) and indirectly (CS_2) appetitively paired stimuli over aversively paired stimuli. These behavioral associative transfer learning effects were accompanied by first-order CS-related reinstatement of the neural patterns representing US in the lateral OFC and increased functional coupling between lateral OFC, amygdala/anterior hippocampus, and medial OFC during SOC. Furthermore, representations of second-order CS in the amygdala became more similar to US representations from early to late phases of second-order conditioning, indicating the acquisition of an association between CS_2 and US patterns.

Results

Behavioral evidence for transfer of motivational value during second-order conditioning

Participants (*N* = 49, combined across two experiments) first established Pavlovian visuo-gustatory associations between a visual CS^+_1 and an appetitive gustatory US^+ (chocolate milk or orange juice), and between a visual CS^-_1 and an aversive gustatory US^- (quinine-HCl) during first-order conditioning (FOC, Fig. 1C). During SOC, participants were visually presented with CS_2^+ followed by CS_1^+, CS_2^- followed by CS_1^-, and a novel pairing of CS_2^n followed by CS_1^n (Fig. 1D). Since CS_1^n had not been presented during first-order conditioning, both CS_2^n and CS_1^n were considered neutral CS, serving as control stimuli for mere exposure effects and for correction of general visual activation effects in fMRI analyses (see Methods for details). Importantly, participants were not informed about the underlying associative structure of the tasks but were instructed to perform simple attentional control

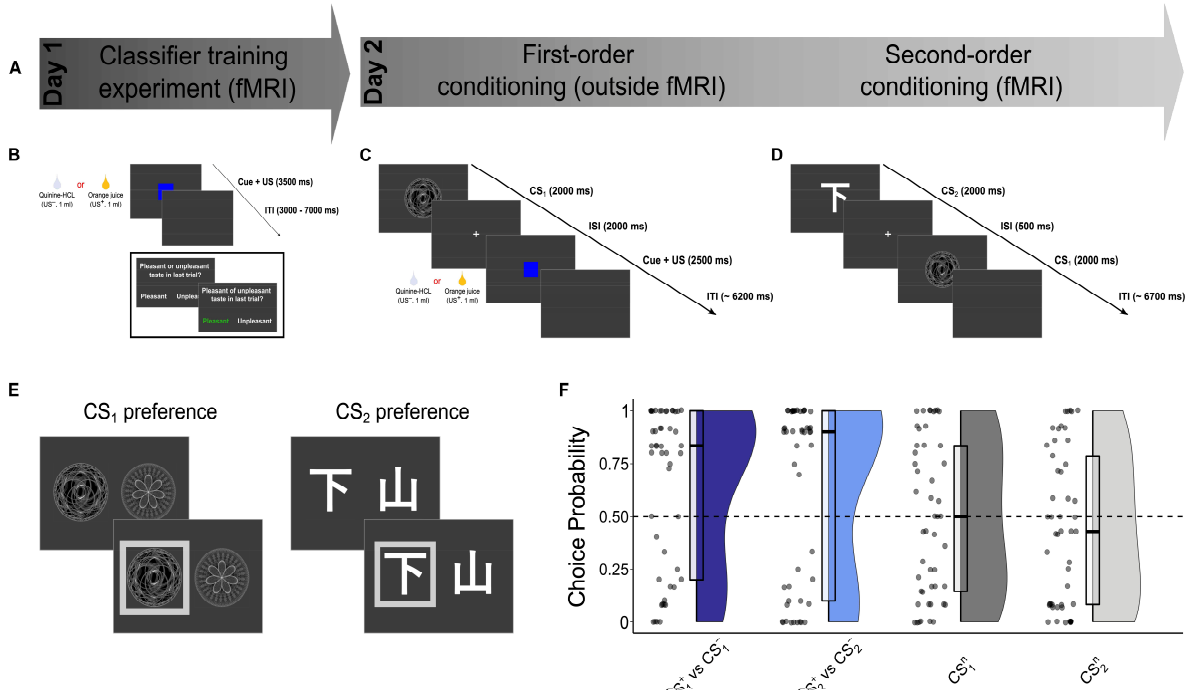


Fig. 1. Experimental procedure, task schematic and behavioral results.
A) Experimental procedure: Participants in the fMRI study (*N* = 29) performed both day 1 and day 2 (B-E), participants in the behavioral study (*N* = 20) only performed the procedures of day 2 (C-E, outside fMRI). B) Classifier training experiment: After rating of subjective value and intensity of both gustatory US (US^+: orange juice and US^-: quinine-HCL 0.1 mmol/l solved in purified water), in a total of five runs each US was administered twenty times (40 trials per run, 200 trials in total). Per trial, one US (1 ml bolus per trial) was delivered by a gustometer and infusion lines. The US bolus onset was preceded by a blue square. Trials were separated by an inter-trial-interval (ITI) marked by a grey screen. Participants performed a 0-back-style attentional control task. In 20% pseudo-randomly selected trials, participants were presented with probe trials in which they were asked to indicate which US they had received last ("pleasant" (US^+) or "unpleasant" (US^-)). C) First-order conditioning: In each trial, a CS_1 was followed by an inter-stimulus interval marked by a fixation cross, and oral infusion of one US (1 ml bolus per trial). Trials were separated by an ITI marked by a grey screen. Each CS_1 was presented 50 times, amounting to 100 trials total. CS_1 were followed by a US with 80% probability. D) Second-order conditioning: In each trial, a CS_2 was followed by an inter-stimulus interval marked by a fixation cross, and a CS_1. CS_2 were followed by a CS_1 fully deterministically. Each trial was separated by an ITI marked by a grey screen. Each CS_2-CS_1 pair was presented 50 times (150 trials total). E) Choice preference test: Following SOC, participants were presented with two separate test phases consisting of repeated binary choices between pairs of CS_1 (right) and pairs of CS_2 (left) to assess behavioral signatures of first- and second-order conditioning. F) Behavioral results (pooled across fMRI and behavioral study, *N* = 49): Raincloud plots (Allen et al., 2019) showing density of choice probability. CS_1^+ (dark blue, 89%-highest posterior density interval (89%-HPDI): [.58; .65]) and CS_2^+ (light blue, 89%-HPDI: [.60; .67]) were preferred over CS_1^- and CS_2^-, respectively. There was no evidence for choice probabilities of CS_1^n and CS_2^n being different from chance level in novel-to-lure comparison choice trials (dark and light gray, 89%-HPDIs: [.46; .52] and [.42; .48]). Box plot center lines represent sample medians and box bottom/top edges show 25th/75th percentile of the data, respectively.

tasks. This was aimed at leaving participants unaware of the associative learning process. Post-experimental tests for explicit knowledge indeed revealed that participants were

unaware of the (indirect) associations between CS_2 and US (Supplementary Results and Supplementary Table 1). In a subsequent choice preference test (Fig. 1E), participants performed binary decisions between pairs of CS_1 and CS_2. Choice trials were interspersed with lure decisions between CS_1^n or CS_2^n and novel fractals or kanjis, respectively, that had only been seen during pre-task rating. We reasoned that choice probabilities (CP) should exceed the indifference criterion (CP > 0.5) if associative direct (CS_1) and associative transfer (CS_2) learning effects occurred. Additionally, we assumed that choice probability should not exceed the indifference criterion in trials involving novel-to-lure stimuli, if choice behavior was not exclusively driven by mere exposure or novelty-dependent response biases. Combined across the two experiments, participants showed a preference both for the appetitive first- and second order stimuli, CS_1^+ and CS_2^+, respectively, over the aversive CS_1^- and CS_2^- (Fig. 1F). Highest posterior density intervals (HPDI) of choice probabilities (binomial model) for CS_1^+ versus CS_1^- [.58; .65] and for CS_2^+ versus CS_2^- [.60; .67] did not overlap (0% overlap) with the defined region of practical equivalence (ROPE, CP = [.45; .55]), an interval of parameter values representing the null hypothesis of no difference from chance level. These findings show that our conditioning procedure reliably induced both direct associative (CS_1) and associative transfer learning effects (CS_2). Importantly, there was no evidence for both choice probabilities of CS_1^n and CS_2^n being different from chance level in novel-to-lure comparison choice trials (Fig. 1F, HPDIs: [.46; .52] and [.42; .48], 100% and 48% ROPE overlap, respectively). This pattern of results rules out an explanation of CS_1^+ and CS_2^+ preference arising from mere-exposure or novelty. Importantly, very similar results were obtained when repeating the analyses separately for the behavioral and fMRI sample (Supplementary Figure 1).

First-order CS reinstate neural US patterns during second-order conditioning

After establishing behavioral evidence for associative transfer learning, we reasoned that a prerequisite for the observed learning effect would be reinstatement of the neural US pattern by the paired CS_1 to establish a direct associative link between CS_2 and US. Importantly, it

173 should be noted that US reinstatement during SOC could only be faithfully attributed to the
174 respective CS_1, but not to CS_2, since only CS_1 had been directly paired with the US, and CS_2
175 had not previously been experienced. To test for cortical reinstatement of neural patterns
176 representing US during second-order conditioning, we used functional magnetic resonance
177 imaging (fMRI) and a cross-session (Stokes et al., 2009), cross-modality searchlight
178 (Kriegeskorte et al., 2006) classification approach (multivariate pattern analysis, MVPA). To
179 obtain unbiased estimates of the neural patterns representing our gustatory US, without the
180 confounding influence of associations to a learned first-order CS, we first performed an fMRI
181 classifier training experiment on day 1 (Fig. 1B) during which the US^+ and US^- were
182 presented. A multivariate pattern classifier (linear support vector machine, C-SVM) was
183 trained on the fMRI data from this session. On day 2, during SOC, we used the weights of
184 the classifier trained on gustatory neural patterns to predict the class label of the visual CS
185 (Fig. 1D) that had been paired with the US during FOC (Fig. 1C, see Methods and Fig. 2A for
186 schematic of the classification approach). In other words, we used data from one sensory
187 modality, assessed during a first day to train a classification model and tested the model's
188 generalizability to unseen data from another sensory modality on a second day. We found
189 evidence for reinstatement of US patterns in a region in the left lateral orbitofrontal cortex
190 (lOFC). In this region, it was possible to predict the class labels of CS neural patterns based
191 on US pattern information obtained during classifier training on the previous day.
192 Classification with a 3-mm searchlight revealed a small-volume corrected (Neurosynth
193 (Yarkoni et al., 2011) meta-analytic functional mask for term "taste") cluster of above-chance
194 level (0.5) classification accuracy (extracted cluster mean = 0.56, *SD* = 0.07) in the left
195 lateral OFC (lOFC, Fig. 2B, peak voxel at MNI [x = –21, y = 30, z = –17], $z = 2.26$, $P = 0.012$,
196 random-effect cluster corrected (Stelzer et al., 2013), 50,000 iterations, one-tailed). The
197 location of this lOFC cluster is consistent with this region's well-documented role in gustatory
198 processing, particularly in representing motivational (Rolls, 2006, 2000; Small et al., 1999)
199 aspects of gustatory sensation and taste memory (Kobayashi et al., 2004). To ensure that
200 our results are not dependent on this particular choice of ROI, we repeated the same

analysis using two different, independent ROIs. The first was an anatomical mask of lateral orbitofrontal cortex (Harvard-Oxford atlas), the second was obtained from an independent gustatory mapping study by Benz and colleagues (K. Benz, personal communication, 12/2019). Again, we found significant CS decodability in both ROIs (Supplementary Results).

When we split participants into a high- and low-bias group, depending on their preference for the CS_2^+, we observed that significant classification of CS labels in lOFC was possible in the high bias group (Z = 3.43, P = .006, $U3_1$ = .88, one-sample Wilcoxon signed-rank test, two-tailed), but not in the low bias group (Z = 1.44, P = .151, $U3_1$ = .58, one-sample Wilcoxon signed-rank test, two-tailed, Fig. 2C).

Interaction between lOFC, amygdala, and medial OFC during second-order conditioning

To form an associative link between CS_2 and US, the reinstated US patterns need to be projected from their cortical storage site to regions like amygdala and hippocampus, allowing for convergence of US and CS_2 information. We investigated whole-brain BOLD signal covariation of the cluster in the left lOFC with the whole brain during SOC using two separate psychophysiological interaction (PPI) analyses. In the first PPI, we investigated general covariation differences in response to CS^- and CS^+ relative to CS^n, reasoning that in the former, but not in the latter, reinstated US representations need to be linked with the CS. To this end, we contrasted CS^- and CS^+ trials with CS^n trials., where the reinstated US is presumably paired with CS_2. We found that covariation of BOLD signal in the left lOFC with a cluster in the left hippocampus, extending to amygdala and medial temporal lobe (Fig. 3A, peak voxel at MNI [x = –38, y = --11, z = –24], Z = 3.75, P = 0.003, whole-brain corrected) and a cluster in the right inferior temporal gyrus, extending to the right hippocampus (peak

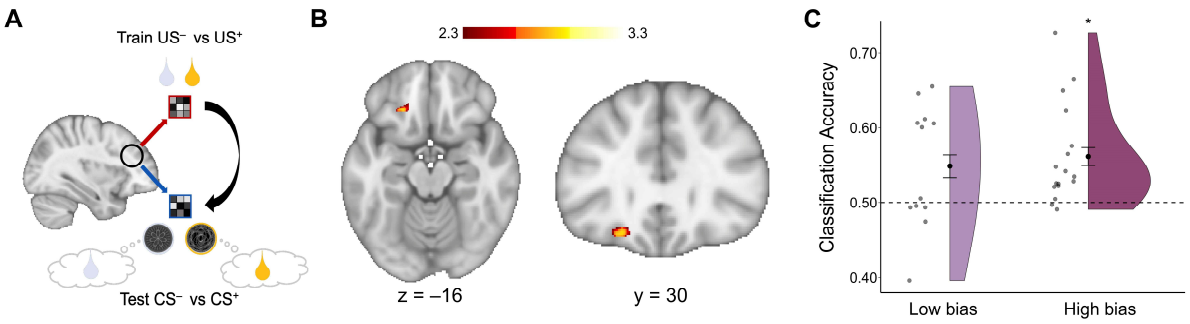


Figure 2. Cross-session, cross-modality classification analysis.
A) Analysis approach: For cross-session, cross-modality classification, a multivariate classifier was trained on the spatial activation patterns (*t* maps) of US^+ versus US^- and tested on CS_2^+/CS_1^+ versus CS_2^-/CS_1^- (CS_2^n/CS_1^n related activation were subtracted to control for general visual effects) during SOC within 3-mm searchlight spheres across the whole brain, resulting in one whole-brain map of classification accuracies per participant. Additionally, we repeated this procedure 100 times per participant with randomly permuted class labels in the training data set (US^+ versus US^-) to create 50,000 chance level group maps for random-effect cluster-statistics(Stelzer et al., 2013). B) Classification accuracy exceeded chance level (0.5) in a small-volume corrected cluster in the left lateral orbitofrontal cortex (lOFC, peak voxel at [x = –21, y = 30, z = –17], z = 2.26, P = 0.012, one-tailed, random-effect cluster-statistics, 50,000 iterations). C) Average classification accuracy in the lOFC cluster (extracted cluster mean = 0.56, SD = 0.07) was significantly above chance level in the high bias group of participants showing higher than chance preference for CS_2^+ (dark purple raincloud, Z = 3.43, P = 0.006, $U3_1$ = 0.88, one-sample Wilcoxon signed-rank test, two-tailed), but not in the low bias group (light purple raincloud, Z = 1.44, P = 0.151, $U3_1$ = 0.58, one-sample Wilcoxon signed-rank test, two-tailed). Color bar represents Z-values. Black dots indicate sample means and error bars represent standard errors of the means of the data, respectively.

voxel at MNI [x = 40, y = --32, z = –20], Z = 3.55, P = 0.007, whole-brain corrected) was higher in CS^- and CS^+ compared to CS^n trials. Despite the fact that we employed an approach with high statistical sensitivity for group-level cluster-based inference (FSL's FLAME1+2 (Beckmann et al., 2003)), a robust estimation technique that yields low family-wise errors rates even at liberal activation thresholds (Z > 2.3, corresponding to p < 0.05) (Eklund et al., 2016), we would like to note that the aforementioned result did not survive whole-brain correction at more conservative thresholds (e.g. Z > 2.6). Furthermore, in the same PPI analysis, we found a positive correlation between second-order choice preference and functional covariation of the left lOFC with a region in the right and left lateral prefrontal cortex (Supplementary Fig. 3, right peak voxel at MNI [x = 49, y = 41, z = 7], Z = 3.80, P < 0.001, whole-brain corrected; left peak voxel at MNI [x = –45, y = 46, z = –4], Z = 3.93, P =

257 0.010, whole-brain corrected). The more participants preferred CS_2^+ (versus CS_2^-), the
258 stronger these regions' BOLD signal covaried during SOC.

259 Some studies hint at complex interactions between OFC and amygdala that differ
260 between appetitive and aversive stimuli (Morrison et al., 2011). Therefore, we next asked
261 whether there were any specific differences between CS^- and CS^+ trials in lOFC BOLD signal
262 covariation with other regions. We found that covariation of the BOLD signal in the left lOFC
263 was higher in CS^+ compared to CS^- trials in a cluster located in the medial OFC, extending to
264 subgenual anterior cingulate cortex (Fig. 3B, peak voxel at MNI [x = 10, y = 50, z = –1], Z =
265 3.90, $P < 0.001$, whole-brain corrected) and a cluster in the left anterior insula, extending to
266 left caudo-lateral OFC and temporal pole (peak voxel at MNI [x = –51, y = 18, z = –11], Z =
267 3.33, $P = 0.004$, whole-brain corrected). Given both the difficulties in inferring directionality
268 from fMRI data and the reciprocal nature of connections between lOFC and both amygdala
269 and medial OFC, a directionality of our PPI results cannot be determined.

270

271 **Plasticity of association between second-order CS and US in the amygdala**

272 We reasoned that if the amygdala uses the reinstated cortical outcome pattern in lOFC to
273 acquire an association between CS_2 and the respective US, one would expect similarity
274 between the neural patterns evoked by CS_2 and US, respectively. This similarity should
275 increase over the course of second-order conditioning, as the association is being acquired.
276 Importantly, during the classifier training experiment, there was evidence for US
277 representations in the amygdala. We found differential neural patterns in response to US^-
278 and US^+ in the amygdala (among other regions) using an exploratory leave-one-run-out
279 cross-validated classification analysis on the neural pattern recorded during the classifier
280 training experiment (Supplementary Fig. 4). Using a least-squares separate (LS-S) approach
281 (Mumford et al., 2012) to deconvolve single-trial estimates, we first computed overall neural
282 pattern similarity (Kriegeskorte et al., 2008) between the pattern evoked by CS_2 (during SOC)
283 and their respective US (during classifier training) in the bilateral amygdala. There was
284 significant pattern similarity between CS_2^- and US^- ($t_{28} = 3.38$, $P = 0.002$, $U3_1 = 0.79$; but not

285

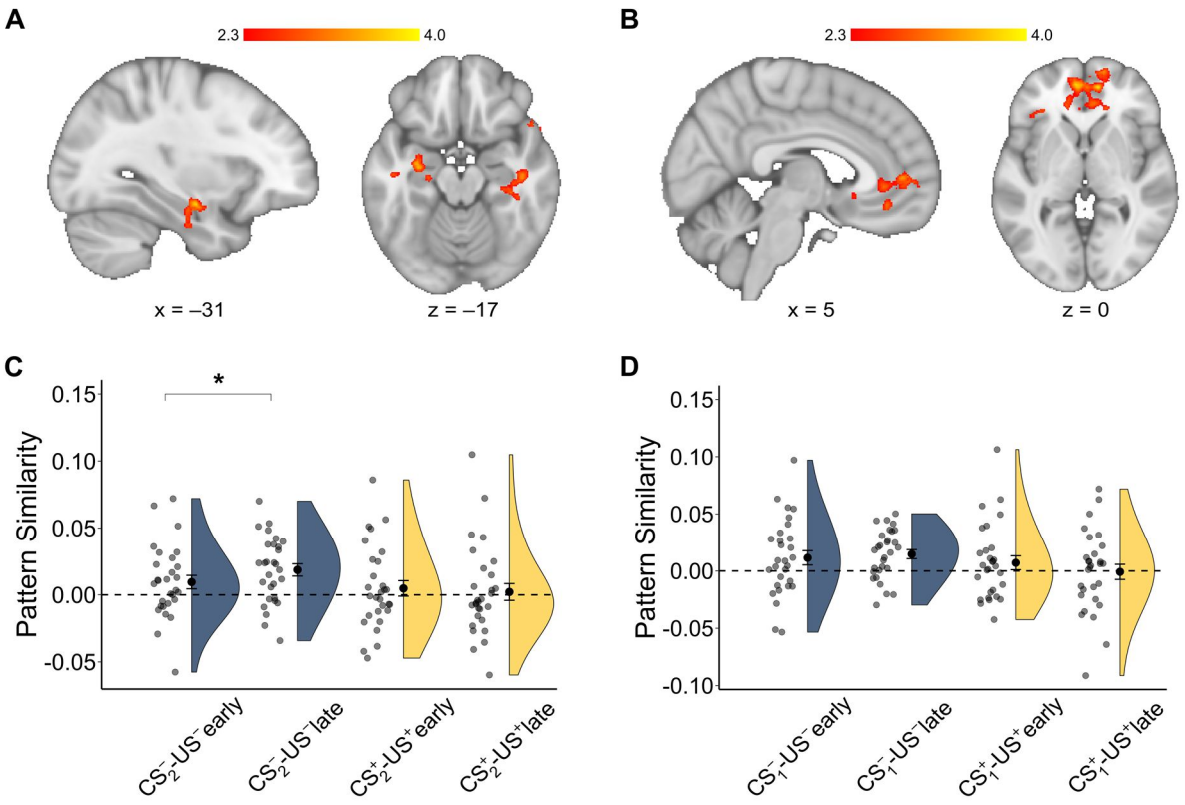


Figure 3. Psychophysiological interaction analyses and neural pattern similarity results.
A) BOLD covariation was higher in CS$^-$ and CS$^+$ trials than in CS^n trials between the left lOFC and a cluster in the anterior hippocampus, extending to (basolateral) amygdala and medial temporal lobe (peak voxel at MNI [x = –38, y = –11, z = –24], Z = 3.75, P = 0.003, whole-brain corrected) and a cluster in the right inferior temporal gyrus, extending to the right hippocampus (peak voxel at MNI [x = 40, y = --32, z = –20], Z = 3.55, P = 0.007, whole-brain corrected). B) Covariation of BOLD signal in the left lOFC was higher in CS$^+$ trials compared to CS$^-$ trials in a cluster located in the medial OFC, extending to subgenual anterior cingulate cortex (peak voxel at MNI [x = 10, y = 50, z = –1], Z = 3.90, P < 0.001, whole-brain corrected) and a cluster in the left anterior insula, extending to left caudo-lateral OFC and temporal pole (peak voxel at MNI [x = –51, y = 18, z = –11], Z = 3.33, P = 0.004, whole-brain corrected). C & D) Template-based neural pattern similarity between classifier training experiment patterns of both US and least-squares separate estimated, pooled (early 25 and late 25 SOC trials) single-trial estimates of neural patterns representing CS$_2$ (C) and CS$_1$ (D), corrected for the respective pooled CSn neural patterns in a bilateral amygdala ROI. C) CS$_2^-$ and US$^-$ patterns became more similar from early to late trials (t_{28} = 1.88, P = 0.035, Cohen's d = 0.35, paired-samples t-test, one-tailed). However, there was also no evidence for a difference between early and late trial similarity for CS$_2^+$ and US$^+$ (t_{28} = 0.50, P = 0.310, Cohen's d = 0.09, paired-samples t-test, one-tailed). D) There was no evidence for change in similarity for CS$_1^-$-US$^-$ (t_{28} = 0.50, P = 0.310, Cohen's d = 0.09, paired-samples t-test, one-tailed) or CS$_1^+$ and US$^+$ neural pattern similarity (t_{28} = 0.50, P = 0.310, Cohen's d = 0.09, paired-samples t-test, one-tailed). Black dots indicate sample means and error bars represent standard errors of the means of the data, respectively. Color bars represent Z-values.

between CS$_2^+$ and US$^+$, t_{28} = 0.74, P = 0.464, $U3_1$= 0.55). We reason that this could be due to differential motivational aspects of the gustatory outcomes we used. The US$^-$ consisted of

315 an aversively tasting quinine solution, which has been shown to produce innate avoidance
316 responses that are not thought to be generated by learning (Yiannakas and Rosenblum,
317 2017). In contrast, US^+ was orange juice, a compound stimulus that consists of many taste
318 features (e.g. sour and fruity components), next to sweet sensation, which also produces
319 innate responses (Yiannakas and Rosenblum, 2017). Presumably, this heterogeneity of taste
320 facets in US^+ had induced more variability in taste responses in our participants. Additionally,
321 it appears likely that in a society offering sugary and highly palatable foods and drinks in
322 abundance, US^+ (an orange juice, most likely a very familiar taste to most of our participants)
323 might not have been perceived as rewarding, or novel enough to represent a motivationally
324 relevant and salient event for our participants. It might well be that participants did not have
325 experience with the aversive taste of quinine, rendering it a novel and salient gustatory
326 sensation, presumably resulting in only the US^- being considered behaviorally and
327 motivationally relevant. Thus, it is possible that the observed preference for CS_1^+ and CS_2^+
328 was indeed due to an aversion against CS_1^- and CS_2^-.

329 If the observed neural pattern similarity between CS_2 and US indeed reflects the
330 formation of an associative link during second-order learning, one would expect the observed
331 similarity to increase from early to late stages of learning, as the association is being
332 acquired and the CS comes to predict the US. We therefore compared our measure of neural
333 pattern similarity between early and late trials of SOC. As expected, neural pattern similarity
334 between CS_2^- and US^- in the bilateral amygdala ROI increased from early to late trials of
335 SOC (t_{28} = 1.88, P = 0.035, Cohen's d = 0.35, paired-samples t-test, one-tailed, Fig. 3C).
336 This change in similarity was not observed for the first-order CS_1^+-US^+ or CS_1^--US^- pattern
337 similarity (Fig. 3D), nor for second-order CS_2^+-US^+ similarity (all P > .310, paired-samples t-
338 tests, one-tailed, Fig. 3C).

339

340 **Discussion**

341 Using a second-order conditioning paradigm, we show associative transfer effects during
342 higher-order conditioning, even when participants are unaware of the underlying associative

structure of the experiment. Participants were more likely to select directly and indirectly appetitively paired stimuli over aversively paired stimuli, closely resembling rodent studies describing choice biases consistent with second-order conditioning (Maes et al., 2020; Sharpe et al., 2017). This suggests that humans, similar to rodents, implicitly acquire preferences through higher-order transfer learning mechanisms – a finding that goes beyond previous studies promoting acquisition of explicit associative relationships between stimuli (Jara et al., 2006; Pauli et al., 2019; Wang et al., 2020; Wimmer and Shohamy, 2012). Notably, choice biases for second-order conditioned stimuli emerged in the absence of explicit knowledge of the underlying higher-order associative structure. To the best of our knowledge, this is the first report demonstrating that human value-based decision making is affected motivational value implicitly conferred via second-order conditioning. The associative transfer learning observed in the present study may be beneficial, as it allows an agent to exploit the underlying relational structure and assign value to states for which direct experience with reinforcement is not available. It may however also be maladaptive when it leads to pursuit of suboptimal options, for example if states only share contiguous occurrence but lack correlated reward-predictive properties.

Using a multivariate cross-session, cross-modality searchlight classification approach, we found reinstatement of neural outcome representations by visual first-order CS that had previously been directly paired with gustatory outcomes during first-order conditioning. This reinstatement of outcome patterns occurred in a region in the lateral OFC. Notably, the exact same rostrolateral portion of OFC has previously been implicated in representing stimulus-outcome associations (Jocham et al., 2016; Klein-Flügge et al., 2013; Luettgau et al., 2020) and in correctly assigning credit for a reward to the causal stimulus choice (Jocham et al., 2016). It is also part of the network most consistently involved in taste processing as obtained from the Neurosynth data base (Yarkoni et al., 2011) and plays a well-documented role in representing motivational (Rolls, 2006, 2000; Small et al., 1999) aspects of gustatory sensation and taste memory (Kobayashi et al., 2004). Importantly, we had trained the classifier on the gustatory US prior to pairing them with visual first-order CS (on a separate

371 day). Thus, decoding of the visual CS during second-order learning cannot be spuriously
372 driven by visual responses elicited by the US or the visually presented swallowing cue during
373 classifier training. Importantly, even though the classifier was tested on the entire activation
374 pattern elicited by the first-/second-order CS pairs, US reinstatement during SOC can only be
375 faithfully attributed to the respective CS_1, since only CS_1, but not CS_2, had been directly
376 paired with the outcome and should thus be capable of eliciting a US representation
377 (Wagner, 1981). In "direct link" models of second-order conditioning, reinstatement of the
378 neural pattern representing the outcome constitutes a necessary prerequisite to establish an
379 associative link between second-order CS and the motivational state conveyed by gustatory
380 outcomes. Our findings are consistent with a body of lesion studies in animals that indicate a
381 key role of OFC in second-order conditioning and in responding for conditioned
382 reinforcement (Cousens and Otto, 2003; Gallagher et al., 1999). They also align with the
383 account of OFC representing abstract relational structure, or a state space (Niv, 2019;
384 Wikenheiser and Schoenbaum, 2016) and recent evidence suggesting that OFC neurons
385 encode relevant second-order combinations of past state variables to guide decision-making
386 (Nogueira et al., 2017).

387 OFC clearly does not act in isolation. A plethora of studies suggests that interactions
388 between OFC, amygdala and ventral striatum (Baxter et al., 2000; Cousens and Otto, 2003;
389 Fiuzat et al., 2017; Morrison et al., 2011) might also support second-order learning
390 processes. We found evidence for the region of OFC that showed reinstatement of outcome
391 representations to increase BOLD signal covariation with amygdala, anterior hippocampus,
392 and medial OFC during second-order learning. Neurons in OFC and amygdala show
393 complex, bi-directional interactions during acquisition and reversal of outcome-predictive
394 associations (Morrison et al., 2011) and functional disconnection of OFC and amygdala using
395 asymmetric lesions produce deficits in flexibly adjusting behavior to changes in stimulus
396 value (Baxter et al., 2000; Fiuzat et al., 2017). We also observed that representations of
397 visual second-order CS in amygdala became more similar to gustatory US representations
398 from early to late phases of second-order conditioning. This indicates the gradual

development of an associative link between CS_2 and US in this region. Indeed, both amygdala and hippocampal lesions impair second-order learning (Gewirtz and Davis, 1997; Gilboa et al., 2014). The hippocampus may play a specific role, since animals with lesions show impaired acquisition of second-order contingencies, while acquisition of first-order stimulus-outcome associations is spared (Gilboa et al., 2014).

Our study has at least two implications for current theories of learning and decision-making. First, our study is – to the best of our knowledge – the only report so far of behavioral evidence for associative transfer of motivational value during human second-order conditioning. Although there is a rich literature on second-order conditioning, direct evidence for this phenomenon in humans using similar procedures as in previous animal work was lacking. The present study thus enables a closer comparison between humans and other species during higher-order learning. Similar to the present results, a recent fMRI study in humans employed a sequential Pavlovian conditioning paradigm to provide initial evidence that Pavlovian stimulus-stimulus associations might be learned and represented in a model-based fashion by the lateral OFC (Pauli et al., 2019). This result is consistent with evidence that rat midbrain dopamine neurons encode both reward and identity prediction errors (Takahashi et al., 2017). However, in the paradigm used by Pauli and colleagues (Pauli et al., 2019) two CS and a US are presented in close temporal proximity (or even overlapping) within the same trial. This differs from our study in two important ways. Firstly, unlike SOC, this procedure does not require any reinstatement of US representations, nor any transfer of value, but rather imposes different temporal offsets between CS and US. This contrasts with SOC, where the CS_2 is presented in a separate block that is entirely devoid of US presentations. Secondly, for the same reasons, it is difficult to dissociate the activation patterns related to the proximal CS (CS_1 in second-order conditioning) and the US. In contrast, in the present study, the classifier was trained on the (gustatory) US on a separate day and then tested on the (visual) CS presented during the second-order learning phase of the SOC paradigm.

427 A second implication of our study pertains to the neural mechanisms of associative
428 transfer learning in SOC, which thus far have remained elusive. Our results suggest that,
429 during second-order conditioning, outcome representations, and presumably features
430 pertaining to the motivational value of outcomes, are reinstated in the lateral OFC. Our data
431 suggests that reinstated outcome representations could be communicated between lOFC
432 and amygdala/anterior hippocampus for associative linking of previously neutral stimuli with
433 reinstated representations of the outcomes. It should be noted that the present design does
434 not allow us to dissociate whether the reinstated US representations detected by the
435 multivariate classification algorithm pertain to the reinforcement value, the intensity or
436 sensory features of the US – or a combination of these features. However, the question
437 which of these features was reinstated by CS_1 and contributed to the associative transfer of
438 motivational value to CS_2 is beyond the scope of the present study.

439 It has been proposed that relational knowledge of events and behavior characterized
440 by considering the state-transition probabilities of the environment might emerge from
441 associative learning mechanisms (Wang et al., 2018). Second-order conditioning can be
442 thought of as an instantiation of such a learning process in which a currently absent, but
443 previously experienced and motivationally relevant task state (receipt of an outcome) is
444 reinstated in lOFC to enable propagation of motivational relevance to task states that are
445 never directly experienced in contiguity with reward or punishment. Higher-order learning
446 paradigms impose a relational structure of events on agents and thus could be cast as an
447 inference problem, very much like in transitive inference (A→B, B→C, hence A→C) (Dusek
448 and Eichenbaum, 1997). Credit assignment processes of this kind have been shown to
449 involve the lOFC (Jocham et al., 2016; Walton et al., 2010). The lOFC reinstatement of
450 outcome representations observed in the present study could thus pose a necessary
451 prerequisite for value attribution to task states or stimuli that have never been directly
452 followed by an outcome, but nevertheless possess outcome-predictive properties.

453 How the reinstatement of cortical US patterns found in our study relates to the
454 observation in rats that midbrain dopamine neurons acquire temporal difference error signals

455 in response to CS_2 (Maes et al., 2020) presents an important question for future studies.
456 Furthermore, it would be of great interest to elucidate the directionality and exact content of
457 information flow between lOFC and amygdala/anterior hippocampus, or between lOFC and
458 medial OFC, and whether this transfer of information is supported by phase coherence in
459 theta oscillations (Benchenane et al., 2010; Knudsen and Wallis, 2020; Young and Shapiro,
460 2011).

461 Taken together, our data lend support to the "direct link" hypothesis of second-order
462 conditioning (Barnet et al., 1991; Gewirtz and Davis, 2000; Rizley and Rescorla, 1972),
463 stating that second-order CS and representations of outcomes – events that had never been
464 explicitly paired during the individual's learning history – can be linked by exploiting the
465 relational structure of events. Importantly, the present study enables a closer comparison
466 between humans and other species during higher-order learning. In conclusion, our results
467 suggest a neural mechanism by which outcome representations can be propagated to stimuli
468 that are never experienced in contiguity with reinforcement, allowing for credit assignment in
469 real-world learning scenarios with infrequent direct encounters with rewards or punishments.
470

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, no regular intake of medication known to interact with the central nervous system and participants not reporting quinine intolerance were included. Participants in both samples 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 (reference number: 101/15) and conducted in accordance with the Declaration of Helsinki.

32 healthy adult volunteers (age: M = 24.16, SD = 3.61 years, 15 males) participated in the fMRI study, 20 healthy adult volunteers (age: M = 23.40, SD = 3.07 years, 9 males), participated in the behavioral study. In the fMRI study, two participants were excluded from statistical analyses due to self-reports of having fallen asleep during the second-order conditioning scanning run. One additional participant had to be excluded due to a scanner malfunction and corruption of three of the five classifier training experiment scanning runs, thus leaving a total of N = 29 participants for final analyses.

Learning experiment – ratings

The learning experiment was performed on day 2 in the fMRI sample, in the behavioral sample, only the learning experiment was performed in a single session. Participants received written instructions for the experiment and were instructed once again on the computer screen. All experiments were programmed in MATLAB 2012b (v8.0.0.783, MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, MA, USA), using Psychophysics Toolbox (Brainard, 1997) (version 3). Before and after the learning

experiment, participants rated ten different round, greyscale fractal images (300 x 300 pixels) serving as first-order conditioned stimuli (CS_1), ten white Japanese kanjis (Tamaoka et al., 2017) (250 x 250 pixels) serving as second-order conditioned stimuli (CS_2) and gustatory stimuli (aversive: quinine-HCl 0.2 mmol/l solved in purified water, appetitive: either chocolate milk (Nesquik, Nestlé, Switzerland) or orange juice (Milde Orange, EDEKA, Germany) in the behavioral study, only orange juice in the fMRI study, serving as unconditioned stimuli (US). Each participant received the same kind and amount of US per trial during the experiment. Ratings of subjective value/liking were assessed with the number buttons on a German (QWERTZ) computer keyboard from 1 (not liked) to 9 (very much liked). In the fMRI sample, participants additionally rated gustatory stimuli regarding their subjective intensity levels from 1 (low intense) to 9 (very intense). Three fractals and kanjis rated closest to 5 (equivalent to "neutral") were selected for first- and second order conditioning and their order was randomized before being associated with the US in first-order conditioning or with the first-order CS_1 in second-order conditioning. To ensure motivational salience of the gustatory US, participants were instructed to abstain from food for 12 hours and on average reported having fasted for 13.28 (*SD* = 2.71) hours. Participants reported intermediate levels of hunger before the task on a paper-pencil visual analog scale (VAS), ranging from 0 ("not hungry") to 100 ("very hungry"), *M* = 58.38 (*SD* = 29.61).

Learning experiment – first-order conditioning

In first-order conditioning (FOC), participants were presented with CS^+_1 followed by the appetitive US^+, and CS^-_1 followed by the aversive US^-. In a typical trial, a CS_1 (2000 ms) was followed by an inter-stimulus interval (1000 ms) marked by a fixation cross, and oral infusion of one US (1 ml bolus per trial). Each CS_1 was presented 50 times, amounting to 100 trials total. CS_1 were followed by a US with 80% probability (40 trials of each CS_1-US pair, 10 trials of CS_1-no US per CS). US were delivered by a MATLAB code-controlled custom-made gustometer consisting of two high pressure single syringe pumps (AL-1000HP, World Precision Instruments, Saratoga, FL) operating 50 ml Luer lock syringes. Syringes were

527 attached to Luer lock infusion lines (1,40 m length, 2 mm inner diameter) that participants
528 held centrally in their mouths like drinking straws. In the fMRI study, infusion line position
529 order (Q-O (*N* = 17) and O-Q (*N* = 15) for quinine (Q) and orange juice (O)) was
530 counterbalanced across participants. In the fMRI study, the US bolus onset was preceded
531 (500 ms) by a blue square that was presented for 2500 ms on the screen (Figure 1B).
532 Participants were instructed to only swallow the US bolus upon offset of the blue square.
533 Each trial was separated by an inter-trial-interval (ITI) marked by a grey screen. The ITI per
534 trial was drawn from a discretized γ-distribution (shape = 6, scale = 1) truncated for an
535 effective range of values between 3500 ms and 10,000 ms. Participants took self-paced
536 breaks after each 10th trial during which they could drink water. Importantly, the instructions
537 did not contain information about the underlying associative structure of the experiment,
538 aiming at leaving participants unaware of the associative learning process. Instead,
539 participants were instructed to perform a simple attentional control task, during which they
540 should react as quickly and as correctly as possible by pressing the "y" button upon seeing
541 the CS_1 colored in red. Each CS_1 was colored red in 10 % of the trials (90 % of trials
542 grayscale image) and color did not predict US contingency. Performance during the task was
543 rewarded with a bonus of 1 € (if > 70 % correct answers). Participants performed very well in
544 the attentional control task (overall probability of correct answers: *M* = .99, *SD* = .05). In the
545 fMRI sample, both ratings and first-order conditioning were performed outside the MRI
546 scanner.
547
548 **Learning experiment – second-order conditioning**
549 For second-order conditioning (SOC), participants were presented with CS_2^+ followed by
550 CS_1^+, CS_2^- followed by CS_1^- and CS_2^n followed by CS_1^n. CS_1^n had not been presented during
551 first-order conditioning and thus was not paired with any US. The novel CS_2^n-CS_1^n pair
552 served to control for attentional and novelty effects. In each trial (Figure 1D), a CS_2 (2000
553 ms) was followed by an inter-stimulus interval (500 ms) marked by a fixation cross, and a
554 CS_1 (2000 ms). CS_2 were followed by a CS_1 deterministically. Each trial was separated by an

555 ITI marked by a grey screen. The ITI per trial was drawn from a discretized γ-distribution
556 (shape = 7, scale = 1) truncated for an effective range of values between 4000 ms and
557 10,000 ms. Each CS_2-CS_1 pair was presented 50 times, amounting to 150 trials total. Again,
558 instructions did not explicitly mention relational structures to be learned, but participants were
559 instructed to perform a simple attentional control task. Participants were instructed to
560 respond as quickly and as correctly as possible by pressing the "y" button (behavioral study)
561 or with the right index finger on an MRI-compatible 4-button response box (fMRI study) upon
562 seeing the CS_2 tilted by a 45° angle. Each CS_2 was tilted in 10 % of its presentations.
563 Performance during the task was rewarded with a bonus of 1 € (if > 70 % correct answers).
564 Participants performed very well in the attentional control task (overall probability of correct
565 answers: *M* = .97, *SD* = .07). In the fMRI study, SOC was performed inside the MRI scanner.
566

567 **Learning experiment – choice preference test**

568 Following SOC, participants were presented with two separate test phases consisting of
569 repeated binary choices (30 trials) between pairs of CS_1 and pairs of CS_2 to assess
570 behavioral signatures of first- and second-order conditioning. Each choice between CS_1^+ and
571 CS_1^-, and between CS_2^+ and CS_2^- was presented ten times (twelve times in the fMRI
572 sample). We hypothesized preference for both CS_1^+ and CS_2^+ over CS_1^- and CS_2^-,
573 respectively, as expressed by choice probabilities above indifference criterion (CP > 0.5).
574 Decision trials were interleaved with choices between the novel CS_1^n or CS_2^n and lure stimuli
575 (fractal or kanjis, respectively) that had only been seen during pre-task rating in pseudo-
576 random order. These novel-to-lure comparison trials were intended to rule out response
577 biases related to mere exposure to the stimuli experienced during conditioning. We reasoned
578 that if choices were reflecting acquired and transferred value and could not only be attributed
579 to mere exposure-dependent response biases, choice probability in trials involving novel-to-
580 lure stimuli should not exceed the indifference criterion. Choice options were presented for
581 1500 ms on the right- and left-hand side of the screen. Order (left/right) of choice options was
582 counterbalanced. If participants did not respond within this time-window, a time-out message

583 was displayed, and the respective trial was repeated at the end of the choice preference test.
584 Participants selected choice options by pressing the “y” (left option) or “m” (right option)
585 button (behavioral sample, left or right index finger on MR-compatible response box in the
586 fMRI sample). Participants were instructed to select the CS they preferred/liked more.
587 Importantly, participants were never presented with the US related to their chosen or
588 unchosen CS.
589

590 **Behavioral analyses**

591 Data were analyzed using MATLAB 2019a (v9.6.0.1072779, The MathWorks, Inc., Natick,
592 MA, USA) and RStudio (RStudioTeam, 2019) (version 3.6.3, RStudio Team, Boston, MA)
593 using custom analysis scripts. We calculated an average choice probability (CP) per first-
594 and second-order CS per subject including all binary decisions in which the respective CS
595 was present (1 = selection of the respective target CS, 0 = selection of the alternative CS) for
596 CS_1^+ and CS_2^+ (versus CS_1^- and CS_2^-, respectively). We hypothesized above-chance level
597 CPs for both CS_1^+ and CS_2^+. For both first-order and second-order novel-to-lure comparison
598 trials, we calculated an average CP per stimulus and expected no difference from chance
599 level for these choices. CPs for each CS in the fMRI and behavioral study were pooled and
600 jointly analyzed. For hypothesis testing, we calculated 89%-highest posterior density
601 intervals (89%-HPDI) for each choice probability using the R package rethinking (McElreath,
602 2020). The HPDI represents the parameter values most consistent with the data. We
603 specified a weakly informative prior probability distribution (10000 steps), an aggregated
604 binomial distribution including the number of choices for the respective stimulus and the total
605 number of choice trials involving the respective CS as likelihood function (Equation 1) and
606 randomly drew 100000 samples from a grid-approximated (10000 steps) normalized
607 posterior probability distribution. For hypothesis testing, we specified a region of practical
608 equivalence (ROPE), i.e. an interval of parameter values (CP = [.45; .55]) representing the
609 null hypothesis. Specifically, we expected that the 89%-HPDI would not overlap with the
610 ROPE for choices of CS_1^+ and CS_2^+ (versus CS_1^- and CS_2^-, respectively), allowing to reject

611 the null hypothesis. However, the ROPE would overlap with the 89%-HPDI for both first-
612 order and second-order novel-to-lure comparison trials. Overlap between HPDIs and ROPEs
613 was quantified using the R package bayestestR (Makowski et al., 2019).

$$CP_i \sim \text{Binomial}(N_i, p_i)$$
$$p_i \sim \text{Normal}(0.5, 0.1) \quad (1)$$

614 where CP_i is the binomially distributed choice probability for CS_i (CS_1^+, CS_2^+, CS_1^n or CS_2^n).
615 N_i is the total number of trials in which CS_i was present and p_i is the proportion of choices of
616 CS_i.
617

618 **fMRI – classifier training experiment**

619 The classifier training experiment was always performed on day 1 of two consecutive testing
620 days. This was done to acquire unbiased estimates of the neural patterns representing the
621 two US for training of a multivariate classification algorithm (multivariate pattern analysis,
622 MVPA), before any association of the US with first- or second order CS had been acquired.
623 Participants received oral instructions for the classifier training experiment and were
624 instructed once again on the projection screen in the MRI. Before and after the experiment,
625 participants rated gustatory US (aversive: quinine-HCL 0.1 mmol/l solved in purified water
626 and appetitive: orange juice) according to subjective value/liking and intensity via button
627 presses with their right and left index and middle fingers on a MRI-compatible 4-button
628 response box. Ratings ranged from 1 (not liked/not intense) to 4 (very much liked/very
629 intense) and were indicated by e.g. pressing the left middle finger corresponding to rating 1
630 or pressing the right index finger corresponding to rating 3. As on day 2, participants were
631 instructed to abstain from food for 12 hours and on average reported to have fasted for 13.59
632 (*SD* = 2.18) hours. Participants reported intermediate levels of hunger before the experiment
633 on a paper-pencil VAS, *M* = 48.50 (*SD* = 31.28). In each of the five total runs of the classifier
634 training experiment, each US was administered twenty times (40 trials per run, 200 trials in
635 total). Per trial, one US (1 ml bolus per trial) was delivered by a MATLAB code-controlled
636 custom-made gustometer and Luer lock infusion lines (8.30 m length, 2 mm inner diameter).

637 Participants held infusion lines centrally in their mouths like drinking straws. Additionally,
638 infusion lines were fixed at absorbent materials attached to throat, chin and cheeks. The US
639 bolus onset was preceded (500 ms) by a blue square that was presented in total for 4000 ms
640 on the screen. Participants were instructed to only swallow the US bolus upon offset of the
641 blue square. Each trial was separated by an ITI marked by a grey screen. The ITI was drawn
642 from a uniform distribution with 5 discrete steps (range: 3000 – 7000 ms). Both US were
643 presented in a pseudo-random order, thus reducing the influence of potentially confounding
644 low-level features of the schedule (e.g. number of same/different US repetitions, different ITI
645 lengths following each US). During each run of the classifier training experiment, participants
646 performed a 0-back style attentional control task. After a pseudo-random 20% of trials,
647 participants were presented with probe trials in which they were asked to indicate which US
648 they had received last ("pleasant" (US^+) or "unpleasant" (US^-)) via button presses with their
649 right and left index fingers on an MRI-compatible response box. Correct responses were
650 rewarded with 0.05 € and incorrect responses or time-out trials (without a response by the
651 participant within 2500 ms after onset of the probe trial) resulted in a 0.05 € penalty which
652 would be summed up as a bonus upon completion of the experiment. On average,
653 participants earned a bonus of 1.86 € (*SD* = .11) during the classifier training experiment.
654 Performance during the 0-back attentional control task was generally high (overall probability
655 of correct answers, excluding time-out trials: *M* = .93, *SD* = .05).
656
657 **fMRI – acquisition**
658 During SOC, one run and during the classifier training experiment, five runs of fMRI were
659 acquired on a 3 Tesla Siemens PRISMA MR-system (Siemens, Erlangen, Germany), using a
660 64-channel head coil. Blood oxygenation level dependent (BOLD) signals were acquired
661 using a multi-band accelerated T2*-weighted echo-planar imaging (EPI) sequence (multi-
662 band acceleration factor 2, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle
663 = 80°, field of view (FoV) = 220 mm, voxel size = 2.2 × 2.2 × 2.2 mm, no gap). Per volume,
664 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. To ensure close spatial alignment of the acquired slices during both sessions, the AutoAlign technique provided by the vendor was applied. At the end of both testing days, a B_0 magnitude and phase map was acquired to estimate field maps and B_0 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 task-based fMRI on both days, 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. For all cross-session classification analyses, SOC EPI data was referenced to orientation of the classifier training experiment.

fMRI – data preprocessing

All fMRI preprocessing steps were performed using tools from the Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL, v5.0 and v6.0) (Jenkinson et al., 2012). Preprocessing for each run of both classifier training task and SOC task 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 and high-pass filtering using a Gaussian-weighted lines filter of 1/100 Hz. EPI images were registered to the high-resolution structural image using affine linear registration (boundary-based registration) and then to standard (MNI) space using linear (12 degrees of freedom) and nonlinear registration (Andersson et al., 2007a, 2007b). Functional data was not spatially smoothed. 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

693 published guidelines (Griffanti et al., 2017), and noise components were removed from the
694 functional time series.
695
696 **fMRI – searchlight classification analyses**
697 As conditioning enhances CS's ability to pre-activate neural ensembles representing an
698 associated US, we expected that a multivariate classification algorithm trained on neural
699 patterns representing the US (during the classifier training task) would be able to correctly
700 predict the class of a paired CS_1 during SOC. Hence, we used a cross-session (Stokes et al.,
701 2009), cross-modality searchlight (Kriegeskorte et al., 2006) classification approach to
702 identify brain regions in which the CS class during SOC could be predicted based on training
703 the classifier during the classifier training experiment on day 1. General linear models
704 (GLMs) were fitted into pre-whitened data space to account for local autocorrelations
705 (Woolrich et al., 2001). For the searchlight classification analyses, the individual level (first
706 level) GLM design matrix per run and participant of the classifier training experiment included
707 four box-car regressors in total. Two regressors coded for onsets and durations of both US
708 (each modelled as single events of 4000 ms duration) and two regressors coded onsets and
709 durations of left and right button presses (delta stick functions on the recorded time of
710 response button presses) and the six volume-to-volume motion parameters from motion
711 correction during preprocessing were entered to account for residual head motion. The SOC
712 run per participant included five box-car regressors in total. Three regressors coding for
713 onsets and durations of all combinations of CS_2-CS_1 (CS_2^+/CS_1^+, CS_2^-/CS_1^-, CS_2^n/CS_1^n) trials
714 (each modelled as single events of 4500 ms duration, due to the 100 % contingency of CS-
715 CS pairs), one regressor coding onsets and durations of right button presses (delta stick
716 functions on the recorded time of response button presses), one regressor coding onset and
717 duration of the within-run pause (45 sec), and the six volume-to-volume motion parameters
718 from motion correction during preprocessing were entered.
719 Additionally, we performed an exploratory multivariate searchlight classification to
720 distinguish between the spatial activation patterns (*t* maps) of US^+ versus US^- across the

721 whole brain during the classifier training experiment. This analysis resulted in one whole-
722 brain map of classification accuracies per participant. The classification analysis was
723 performed using a 5-fold cross-validation scheme, iteratively training the classification
724 algorithm on the difference between US^+ versus US^- on 4 (out of 5 runs) labeled training data
725 sets and then assessing the out-of-sample predictive accuracy of the classifier on the
726 remaining, left-out data set (1 out of 5 runs). For this analysis, the individual level (first level)
727 GLM design matrix per run and participant of the classifier training experiment included four
728 box-car regressors in total. Two regressors coded for onsets and durations of both US (each
729 modelled as single events of 4000 ms duration) and two regressors coded onsets and
730 durations of left and right button presses (delta stick functions on the recorded time of
731 response button presses) and the six volume-to-volume motion parameters from motion
732 correction during preprocessing were entered to account for residual head motion.

733 Regressors were convolved with a hemodynamic response function (γ-function, mean
734 lag = 6 s, SD = 3 s). Each first level GLM included two or three (for classifier training
735 experiment or SOC, respectively) contrasts to estimate individual per run *t*-statistic maps for
736 each US or each CS_2-CS_1 pair (for classifier training experiment or SOC, respectively). For
737 SOC, the activation pattern of CS_2^n/CS_1^n was subtracted from CS_2^+/CS_1^+ and CS_2^-/CS_1^-
738 activation patterns. This was intended to account for general activation expected to result
739 from visually presented CS and the visually presented swallowing cue coinciding with US
740 presentation, which could have confounded classification accuracies.

741 All classification-based analyses were conducted in subject native space. Per
742 participant *t*-statistic maps were subjected to linear support vector machine (C-SVM, cost
743 parameter C = 1) classification using the MATLAB-based multivariate pattern analysis
744 toolbox CoSMoMVPA (Oosterhof et al., 2016). For cross-session, cross-modality
745 classification, a classifier was trained on the spatial activation patterns of US^+ versus US^- and
746 tested on CS_2^+/CS_1^+ versus CS_2^-/CS_1^- during SOC within 3-mm searchlight spheres across
747 the whole brain. For the exploratory multivariate searchlight classification between US^+
748 versus US^- patterns, we similarly employed a 3-mm searchlight sphere across the whole

749 brain. Each searchlight sphere classification accuracy was mapped to the center voxel of the
750 sphere, resulting in one whole-brain map of classification accuracies per participant.
751 Additionally, in cross-session, cross-modality classification we repeated this procedure 100
752 times per participant with randomly permuted class labels in the training data set (US^+ versus
753 US^-) to create chance level maps. Before group level statistics, the normalization parameters
754 obtained from preprocessing were applied to both whole-brain classification accuracy maps
755 and chance level maps for normalization to MNI space. The resulting normalized whole-brain
756 maps were spatially smoothed with a Gaussian kernel (5 mm full-width half-maximum).

757 In cross-session, cross-modality classification, we employed small-volume correction
758 (P_{SVC}) to assess significance of clusters with above-chance level (0.5) classification accuracy
759 in the lateral orbitofrontal cortex (lOFC), our a priori region-of-interest (ROI). This region has
760 been implicated in higher-order gustatory processing, such as motivational aspects and
761 discrimination of taste (Kobayashi et al., 2004; Miranda, 2012; Small et al., 1999), but also
762 representation and adaptive changes of stimulus-outcome associations (Boorman et al.,
763 2016; Jocham et al., 2016; Klein-Flügge et al., 2013; Luettgau et al., 2020). We thus
764 reasoned that this brain region would be a well-suited candidate for associative coupling
765 between visual conditioned stimuli and gustatory outcomes and the proposed associative
766 transfer learning. We employed two different approaches for ROI definition: 1) an
767 independent anatomical mask of the lateral OFC (Harvard-Oxford Atlas) and 2) a meta-
768 analysis-based explicit mask for functional activations related to the term "taste" in the
769 Neurosynth data base (Yarkoni et al., 2011) (www.neurosynth.org), thresholded at $z > 7$. This
770 value was chosen so that only meta-analytic clusters of activation, but not single voxels,
771 survived thresholding. The Neurosynth-based mask encompassed, aside from the lateral
772 OFC, functional activation clusters in bilateral anterior insula. This approach aimed at
773 reducing inferential limitations related to arbitrary ROI definition.

774 Within these ROIs, we computed group-level random-effect cluster-statistics
775 corrected for multiple comparisons in small volumes as implemented in CoSMoMVPA
776 (Oosterhof et al., 2016). In brief, for 50,000 iterations, we randomly selected one chance

level map per participant, calculated a group level z-statistic map for the respective permutation and finally compared the resulting cluster sizes drawn from this empirical null distribution (50,000 samples) to the clusters in the "real" classification accuracy map (Stelzer et al., 2013). Since classifications accuracies below chance level are generally limited in interpretability, we considered clusters significant if $P_{SVC} < .05$ (one-tailed). Please note that due to the exploratory nature and illustrative purpose of the exploratory multivariate searchlight classification between US^+ versus US^- patterns, no multiple comparisons correction using random-effects cluster-statistics was performed in this analysis.

Additionally, we split participants into a high- or low-bias group, depending on their preference for the CS_2^+ (vs. CS_2^-) and compared classification accuracies within the small-volume corrected cluster. The cluster ROI was built in MNI space and the ROI was then back-projected into subject native space using inverse normalization parameters obtained during preprocessing to extract individual averaged classification accuracies from the ROI. Both groups' average extracted classification accuracies were separately tested against chance level (0.5) using one-sample Wilcoxon signed-rank tests. We report measures of effect size Cohen's $U3_1$ for one-sample Wilcoxon signed-rank tests (range: 0 – 1, .5 indicating no effect), calculated in the MATLAB-based Measures-of-Effect-Size-toolbox (Hentschke, 2020).

fMRI – functional covariation analyses

We investigated functional covariation between the MVPA-identified lOFC cluster surviving multiple comparisons (thresholded at $z > 1.65$, or $P < .05$, one-tailed) during SOC as a seed region and the whole-brain using a psychophysiological interaction (PPI) analysis in FSL. For each participant, we set up two separate first-level PPI GLMs. The first PPI GLM was aimed at investigating general differences in BOLD signal covariation for both CS_2^+/CS_1^+ and CS_2^-/CS_1^- compared with CS_2^n/CS_1^n. The first PPI GLM contained the following regressors: 1) the BOLD timeseries (preprocessed functional time series) of the lOFC seed as physiological regressor, 2) onsets and durations of CS_2^+/CS_1^+ and CS_2^-/CS_1^- pairs (magnitude coded as 1)

805 and CS_2^n/CS_1^n (magnitude coded as –1) as psychological regressor and 3) the interaction
806 between 1) and 2) as psychophysiological interaction regressor. First-level regressor 2) was
807 convolved with a hemodynamic response function (γ-function, mean lag = 6 s, *SD* = 3 s). The
808 second PPI GLM aimed at investigating specific differences in covariation between
809 CS_2^+/CS_1^+ and CS_2^-/CS_1^-. The second PPI GLM contained the following regressors: 1) the
810 BOLD timeseries (preprocessed functional time series) of the lOFC seed as physiological
811 regressor, 2) onsets and durations of CS_2^+/CS_1^+ (magnitude coded as 1) and CS_2^-/CS_1^- pairs
812 (magnitude coded as –1) as psychological regressor and 3) the interaction between 1) and 2)
813 as psychophysiological interaction regressor. First-level regressor 2) was convolved with a
814 hemodynamic response function (γ-function, mean lag = 6 s, *SD* = 3 s).

815 In addition to the main effects of the PPI regressors, we also investigated whether
816 task-related functional covariation of lOFC with the rest of the brain was related to our
817 behavioral index of second-order conditioning (choice probability of CS_2^+ vs. CS_2^-). For this,
818 we entered the averaged second-order choice preference test data per subject as an
819 additional regressor to the group level GLMs. In all analyses, contrast images from the first
820 level were then taken to the group level and analyzed using mixed-effects analyses in
821 FLAME1+2 (Beckmann et al., 2003). We used cluster-based correction with an activation
822 threshold of Z > 2.3 and a cluster-extent threshold of $P < .05$ at whole-brain level to control
823 for multiple comparisons.

824

825 **fMRI – neural pattern similarity analyses**

826 In addition to our classification-based approach, we investigated whether there was evidence
827 for changes of neural pattern similarity across second-order conditioning. Specifically, we
828 used a least-squares separate (LS-S) approach (Mumford et al., 2012) to deconvolve single-
829 trial estimates of neural patterns representing CS_2 and CS_1 and performed a template-based
830 neural pattern similarity analysis (Wimber et al., 2015) (a variant of representational similarity
831 analysis, RSA (Kriegeskorte et al., 2008)) between patterns of the two US during the
832 classifier training experiment and all trial-specific CS_2/CS_1 patterns during the SOC run. We

833 estimated two subject-specific GLMs per trial, one for both CS_1 and CS_2, containing the
834 following regressors: 1) onset and duration of the respective trial to be estimated (each
835 modelled as a single event of 2000 ms duration), 2) onsets and durations of the remaining
836 CS_1 or CS_2 trials, respectively, 3) onsets and durations of all other CS_1 or CS_2 trials,
837 respectively. Additionally, one regressor coding onsets and durations of right button presses
838 (delta stick functions on the recorded time of response button presses), one regressor coding
839 onset and duration of the within-run pause (45 sec), and the six volume-to-volume motion
840 parameters from motion correction during preprocessing were entered. For example, the
841 GLM of the 25th trial of CS_1^+, contained one regressor (onset and duration) modelling this
842 particular trial and one separate regressor (onsets and durations) for all remaining CS_1 trials
843 but the 25th trial of CS_1^+ (149 other trials in total) and one separate regressor (onsets and
844 durations) modelling all CS_2 trials. This procedure is assumed to allow for better
845 deconvolution of events happening within close temporal proximity than standard least-
846 squares-based approaches, which generally do not distinguish between overall and trial-
847 specific error terms (Abdulrahman and Henson, 2016; Mumford et al., 2012).

848 Each first level GLM included one contrast to model activation related to the
849 respective trial of interest versus baseline. The a priori ROIs of the bilateral amygdala were
850 built in MNI space and back-projected into subject native space using inverse normalization
851 parameters obtained during preprocessing. We used these individual ROIs for spatially
852 constrained multivoxel pattern extraction from the respective contrast *t*-value maps.
853 Similarity-based analyses were carried out using CoSMoMVPA (Oosterhof et al., 2016). We
854 employed 1−Pearson's product-moment correlation coefficient (1−*r*) as a measure of
855 pairwise dissimilarity between trial-specific neural patterns and the US template. Within-
856 subject pairwise neural dissimilarity was subtracted from 1 (to create a measure of neural
857 pattern similarity) and Fisher-*Z* transformed to closer approximate normally distributed data.
858 CS_1^n-US^-/CS_2^n-US^- and CS_1^n-US^+/CS_2^n-US^+ early and late neural pattern similarity were
859 subtracted from CS_1^--US^-/CS_2^--US^- and CS_1^+-US^+/CS_2^+-US^+ early and late neural pattern
860 similarity, respectively. We hypothesized that neural pattern similarity would show

861 dissociable changes from early (first 25) to late (last 25) trials, as both CS_2^+ and CS_2^- should
862 become more similar to the respective indirectly associated US, indicating associative
863 learning transfer from CS_1 to CS_2. Contrarily, CS_1^+ and CS_1^- should, if anything, become less
864 similar to the respectively paired US (e.g. due to extinction). Average change of early to late
865 CS-US neural pattern similarity was analyzed at the group level with paired-samples t-tests.
866 Due to the expected negative change for CS_1^+ and CS_1^- and the expected positive change
867 for CS_2^+ and CS_2^- across trials, we used one-tailed tests accordingly.
868

Acknowledgments

The authors thank all volunteers who participated in this study. The authors would like to thank Halla Mulla-Osman, Stefanie Linnhoff, Leonie Oevel and Denise Scheermann for their invaluable support during data acquisition. Additionally, we thank Gunnar and Christian Lüttgau for electrical engineering support during construction of the gustometer, Karsta Benz for helpful discussions during the set-up of the study and Denise Kramer for production of the gustatory stimuli. We thank the Center for Magnetic Resonance Research of the University of Minnesota for providing the Multiband accelerated fMRI sequence. Computational infrastructure and support were provided by the Centre for Information and Media Technology at Heinrich Heine University Düsseldorf. **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. EP analyzed the data, 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. 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.

Data availability

The behavioral data, extracted classification accuracies and neural pattern similarity matrices that support the findings of this study will be made publicly available at GitHub upon publication. The neuroimaging raw data that support the findings of this study are available upon reasonable request from the corresponding author (LL). The neuroimaging raw data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Code availability

Custom analysis code for the reported behavioral data analyses and multivariate fMRI analyses will be made publicly available at GitHub upon publication.

904 **References**

905 Abdulrahman H, Henson RN. 2016. Effect of trial-to-trial variability on optimal event-related
906 fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis.
907 *Neuroimage* **125**:756–766. doi:10.1016/j.neuroimage.2015.11.009

908 Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA. 2019. Raincloud plots: A multi-
909 platform tool for robust data visualization [version 1; peer review: 2 approved].
910 *Wellcome Open Res* **4**:1–41. doi:10.12688/wellcomeopenres.15191.1

911 Andersson JL., Jenkinson M, Smith S. 2007a. Non-linear registration aka spatial
912 normalisation. *FMRIB Tech Rep TRO7JA2* 22.

913 Andersson JL., Jenkinson M, Smith SM. 2007b. Non-linear optimisation. *FMRIB Tech Rep*
914 *TRO7JA1*.

915 Barnet RC, Arnold HM, Miller RR. 1991. Simultaneous conditioning demonstrated in second-
916 order conditioning: Evidence for similar associative structure in forward and
917 simultaneous conditioning. *Learn Motiv* **22**:253–268. doi:10.1016/0023-9690(91)90008-
918 V

919 Baxter MG, Parker A, Lindner CCC, Izquierdo AD, Murray EA. 2000. Control of response
920 selection by reinforcer value requires interaction of amygdala and orbital prefrontal
921 cortex. *J Neurosci* **20**:4311–4319. doi:10.1523/jneurosci.20-11-04311.2000

922 Beckmann CF, Jenkinson M, Smith SM. 2003. General multilevel linear modeling for group
923 analysis in FMRI. *Neuroimage* **20**:1052–1063. doi:10.1016/S1053-8119(03)00435-X

924 Benchenane K, Peyrache A, Khamassi M, Tierney PL, Gioanni Y, Battaglia FP, Wiener SI.
925 2010. Coherent Theta Oscillations and Reorganization of Spike Timing in the
926 Hippocampal- Prefrontal Network upon Learning. *Neuron* **66**:921–936.
927 doi:10.1016/j.neuron.2010.05.013

928 Boorman ED, Rajendran VG, O'Reilly JX, Behrens TE. 2016. Two Anatomically and
929 Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome
930 Associations in Hippocampus. *Neuron* **89**:1343–1354.
931 doi:10.1016/j.neuron.2016.02.014

932 Brainard DH. 1997. The Psychophysics Toolbox. *Spat Vis* **10**:433–436.

933 Cousens GA, Otto T. 2003. Neural Substrates of Olfactory Discrimination Learning with
934 Auditory Secondary Reinforcement. I. Contributions of the Basolateral Amygdaloid
935 Complex and Orbitofrontal Cortex. *Integr Physiol Behav Sci* **38**:272–294.
936 doi:10.1007/BF02688858

937 Dusek JA, Eichenbaum H. 1997. The hippocampus and memory for orderly stimulus
938 relations. *Proc Natl Acad Sci U S A* **94**:7109–7114. doi:10.1073/pnas.94.13.7109

939 Eklund A, Nichols TE, Knutsson H. 2016. Cluster failure: Why fMRI inferences for spatial
940 extent have inflated false-positive rates. *Proc Natl Acad Sci* **113**:7900–7905.
941 doi:10.1073/pnas.1602413113

942 Fiuzat EC, Rhodes SEV, Murray EA. 2017. The role of orbitofrontal-amygdala interactions in
943 updating action-outcome valuations in macaques. *J Neurosci* **37**:2463–2470.
944 doi:10.1523/JNEUROSCI.1839-16.2017

945 Gallagher M, McMahan RW, Schoenbaum G. 1999. Orbitofrontal cortex and representation
946 of incentive value in associative learning. *J Neurosci* **19**:6610–6614.
947 doi:10.1523/jneurosci.19-15-06610.1999

948 Gewirtz JC, Davis M. 2000. Using Pavlovian higher-order conditioning paradigms to
949 investigate the neural substrates of emotional learning and memory. *Learn Mem* **7**:257–
950 266. doi:10.1101/lm.35200

951 Gewirtz JC, Davis M. 1997. Second-order fear conditioning prevented by blocking NMDA
952 receptors in amygdala. *Nature* **388**:471–474. doi:10.1038/41325

953 Gilboa A, Sekeres M, Moscovitch M, Winocur G. 2014. Higher-order conditioning is impaired
954 by hippocampal lesions. *Curr Biol* **24**:2202–2207. doi:10.1016/j.cub.2014.07.078

955 Griffanti L, Douaud G, Bijsterbosch J, Evangelisti S, Alfaro-Almagro F, Glasser MF, Duff EP,
956 Fitzgibbon S, Westphal R, Carone D, Beckmann CF, Smith SM. 2017. Hand
957 classification of fMRI ICA noise components. *Neuroimage* **154**:188–205.
958 doi:10.1016/J.NEUROIMAGE.2016.12.036

959 Hatfield T, Han JS, Conley M, Gallagher M, Holland P. 1996. Neurotoxic lesions of
960 basolateral, but not central, amygdala interfere with pavlovian second-order conditioning
961 and reinforcer devaluation effects. *J Neurosci* **16**:5256–5265. doi:10.1523/jneurosci.16-
962 16-05256.1996

963 Hentschke H. 2020. hhentschke/measures-of-effect-size-toolbox
964 (https://www.github.com/hhentschke/measures-of-effect-size-toolbox), GitHub.
965 Retrieved January 10, 2020.

966 Jara E, Vila J, Maldonado A. 2006. Second-order conditioning of human causal learning.
967 *Learn Motiv* **37**:230–246. doi:10.1016/j.lmot.2005.12.001

968 Jenkinson M. 2003. Fast, automated,N-dimensional phase-unwrapping algorithm. *Magn*
969 *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 CF, Behrens TEJ, Woolrich MW, Smith SM. 2012. FSL. *Neuroimage* **62**:782–790. doi:10.1016/J.NEUROIMAGE.2011.09.015

Jocham G, Brodersen KHH, Constantinescu AOO, Kahn MCC, Ianni AM, Walton MEE, Rushworth MFFS, Behrens TEEJ. 2016. Reward-Guided Learning with and without Causal Attribution. *Neuron* **90**:177–190. doi:10.1016/j.neuron.2016.02.018

Klein-Flügge MC, Barron HC, Brodersen KH, Dolan RJ, Behrens TEJ. 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

Knudsen EB, Wallis JD. 2020. Closed-Loop Theta Stimulation in the Orbitofrontal Cortex Prevents Reward-Based Learning. *Neuron* **106**:537-547.e4. doi:10.1016/j.neuron.2020.02.003

Kobayashi M, Takeda M, Hattori N, Fukunaga M, Sasabe T, Inoue N, Nagai Y, Sawada T, Sadato N, Watanabe Y. 2004. Functional imaging of gustatory perception and imagery: “Top-down” processing of gustatory signals. *Neuroimage* **23**:1271–1282. doi:10.1016/j.neuroimage.2004.08.002

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. *Proc Natl Acad Sci U S A* **103**:3863–3868. doi:10.1073/pnas.0600244103

Kriegeskorte N, Mur M, Bandettini P. 2008. Representational similarity analysis – connecting the branches of systems neuroscience. *Front Syst Neurosci* **2**:1–28. doi:10.3389/neuro.06.004.2008

Luettgau L, Tempelmann C, Kaiser LF, Jocham G. 2020. Decisions bias future choices by modifying hippocampal associative memories. *Nat Commun* **11**:3318. doi:10.1101/802462

Maes EJP, Sharpe MJ, Usypchuk AA, Lozzi M, Chang CY, Gardner MPH, Schoenbaum G, Iordanova MD. 2020. Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. *Nat Neurosci* **23**:176–178. doi:10.1038/s41593-019-0574-1

Makowski D, Ben-Schachar M, Lüdecke D. 2019. bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *J Open Source Softw* **4**:1541.

1003 Mcdannald MA, Setlow B, Holland PC. 2013. Effects of ventral striatal lesions on first- and
1004 second-order appetitive conditioning. *Eur J Neurosci* **38**:2589–2599.
1005 doi:10.1111/ejn.12255

1006 McElreath R. 2020. rethinking: Statstical rethinking book package.

1007 Miranda MI. 2012. Taste and odor recognition memory: The emotional flavor of life. *Rev*
1008 *Neurosci* **23**:481–499. doi:10.1515/revneuro-2012-0064

1009 Morrison SE, Saez A, Lau B, Salzman CD. 2011. Different Time Courses for Learning-
1010 Related Changes in Amygdala and Orbitofrontal Cortex. *Neuron* **71**:1127–1140.
1011 doi:10.1016/j.neuron.2011.07.016

1012 Mumford JA, Turner BO, Ashby FG, Poldrack RA. 2012. Deconvolving BOLD activation in
1013 event-related designs for multivoxel pattern classification analyses. *Neuroimage*
1014 **59**:2636–2643. doi:10.1016/j.neuroimage.2011.08.076

1015 Niv Y. 2019. Learning task-state representations. *Nat Neurosci* **22**:1544–1553.
1016 doi:10.1038/s41593-019-0470-8

1017 Nogueira R, Abolafia JM, Drugowitsch J, Balaguer-Ballester E, Sanchez-Vives M V.,
1018 Moreno-Bote R. 2017. Lateral orbitofrontal cortex anticipates choices and integrates
1019 prior with current information. *Nat Commun* **8**. doi:10.1038/ncomms14823

1020 Oosterhof NN, Connolly AC, Haxby J V, Rosa MJ. 2016. CoSMoMVPA : Multi-Modal
1021 Multivariate Pattern Analysis of Neuroimaging Data in Matlab / GNU Octave **10**:1–27.
1022 doi:10.3389/fninf.2016.00027

1023 Parkes SL, Westbrook RF. 2011. Role of the basolateral amygdala and NMDA receptors in
1024 higher-order conditioned fear. *Rev Neurosci* **22**:317–333. doi:10.1515/RNS.2011.025

1025 Pauli WM, Gentile G, Collette S, Tyszka JM, O'Doherty JP. 2019. Evidence for model-based
1026 encoding of Pavlovian contingencies in the human brain. *Nat Commun* **10**:1099.
1027 doi:10.1038/s41467-019-08922-7

1028 Rizley RC, Rescorla RA. 1972. Associations in second-order conditioning and sensory
1029 preconditioning. *J Comp Physiol Psychol* **81**:1–11. doi:10.1037/h0033333

1030 Rolls ET. 2006. Brain mechanisms underlying flavour and appetite. *Philos Trans R Soc B*
1031 *Biol Sci* **361**:1123–1136. doi:10.1098/rstb.2006.1852

1032 Rolls ET. 2000. The Orbitofrontal Cortex and Reward. *Cereb Cortex* **10**:284–294.
1033 doi:10.1093/cercor/10.3.284

1034 RStudioTeam. 2019. RStudio: Integrated Development for R.

1035 Setlow B, Gallagher M, Holland PC. 2002. The basolateral complex of the amygdala is
1036 necessary for acquisition but not expression of CS motivational value in appetitive
1037 Pavlovian second-order conditioning. *Eur J Neurosci* **15**:1841–1853. doi:10.1046/j.1460-
1038 9568.2002.02010.x

1039 Sharpe MJ, Batchelor HM, Schoenbaum G. 2017. Preconditioned cues have no value. *Elife*
1040 **6**. doi:10.7554/eLife.28362

1041 Small DM, Zald DH, Jones-Gotman M, Zatorre RJ, Pardo J V., Frey S, Petrides M. 1999.
1042 Human cortical gustatory areas: A review of functional neuroimaging data. *Neuroreport*
1043 **10**:7–14. doi:10.1097/00001756-199901180-00002

1044 Stelzer J, Chen Y, Turner R. 2013. Statistical inference and multiple testing correction in
1045 classification-based multi-voxel pattern analysis (MVPA): Random permutations and
1046 cluster size control. *Neuroimage* **65**:69–82. doi:10.1016/J.NEUROIMAGE.2012.09.063

1047 Stokes M, Thompson R, Cusack R, Duncan J. 2009. Top-down activation of shape-specific
1048 population codes in visual cortex during mental imagery. *J Neurosci* **29**:1565–1572.
1049 doi:10.1523/JNEUROSCI.4657-08.2009

1050 Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, Schoenbaum G. 2017.
1051 Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of
1052 Expected Rewards. *Neuron* **95**:1395-1405.e3. doi:10.1016/j.neuron.2017.08.025

1053 Tamaoka K, Makioka S, Sanders S, Verdonschot RG. 2017. www.kanjidatabase.com: a new
1054 interactive online database for psychological and linguistic research on Japanese kanji
1055 and their compound words. *Psychol Res* **81**:696–708. doi:10.1007/s00426-016-0764-3

1056 Tonegawa S, Morrissey MD, Kitamura T. 2018. The role of engram cells in the systems
1057 consolidation of memory. *Nat Rev Neurosci* **19**:485–498. doi:10.1038/s41583-018-0031-
1058 2

1059 Wagner AR. 1981. SOP : A model of automatic processing in animal behaviorInformation
1060 Processing in Animals: Conditioned Inhibition. pp. 5–24.

1061 Walton ME, Behrens TEJ, Buckley MJ, Rudebeck PH, Rushworth MFS. 2010. Separable
1062 Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in
1063 Contingent Learning. *Neuron* **65**:927–939. doi:10.1016/j.neuron.2010.02.027

1064 Wang F, Schoenbaum G, Kahnt T. 2020. Interactions between human orbitofrontal cortex
1065 and hippocampus support model-based inference. *PLoS Biol* **18**:1–24.
1066 doi:10.1371/journal.pbio.3000578

1067 Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, Hassabis D, Botvinick

1068 M. 2018. Prefrontal cortex as a meta-reinforcement learning system. *Nat Neurosci*
1069 **21**:860–868. doi:10.1038/s41593-018-0147-8

1070 Wikenheiser AM, Schoenbaum G. 2016. Over the river, through the woods: Cognitive maps
1071 in the hippocampus and orbitofrontal cortex. *Nat Rev Neurosci* **17**:513–523.
1072 doi:10.1038/nrn.2016.56

1073 Wimber M, Alink A, Charest I, Kriegeskorte N, Anderson MC. 2015. Retrieval induces
1074 adaptive forgetting of competing memories via cortical pattern suppression. *Nat*
1075 *Neurosci* **18**:582–589. doi:10.1038/nn.3973

1076 Wimmer GE, Shohamy D. 2012. Preference by Association: How Memory Mechanisms in
1077 the Hippocampus Bias Decisions. *Science* **338**:270–3. doi:10.1126/science.1223252

1078 Woolrich MW, Ripley BD, Brady M, Smith SM. 2001. Temporal Autocorrelation in Univariate
1079 Linear Modeling of FMRI Data. *Neuroimage* **14**:1370–1386.
1080 doi:10.1006/NIMG.2001.0931

1081 Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale
1082 automated synthesis of human functional neuroimaging data. *Nat Methods* **8**:665–670.
1083 doi:10.1038/nmeth.1635

1084 Yiannakas A, Rosenblum K. 2017. The insula and taste learning. *Front Mol Neurosci* **10**:1–
1085 24. doi:10.3389/fnmol.2017.00335

1086 Young JJ, Shapiro ML. 2011. Dynamic coding of goal-directed paths by orbital prefrontal
1087 cortex. *J Neurosci* **31**:5989–6000. doi:10.1523/JNEUROSCI.5436-10.2011

1088

1089

1090

1091

Supplementary Information

Reinstatement of cortical outcome representations during higher-order learning

Luettgau et al.

Supplementary Results

Explicit knowledge of associative structure

Post-experimental paper-pencil tests for explicit knowledge of the associative structure revealed that the number of participants that had indeed realized the indirect associations between CS_2 and US was significantly below the value that would be expected under a random guessing assumption (29%, $P = 0.004$, binomial test vs. 0.5). Participants could also not reliably indicate which CS_2 and which US had been indirectly linked in the experiment (mean correct responses: $M = 0.31$, $SD = 0.68$, maximum of 2 correct answers possible, Supplementary Table 1).

CS and US ratings

During the ratings prior to the learning experiment, the subjective values/liking of the three kanjis selected as CS_2 did not differ significantly from each other (main effect of stimulus: $F_{2,96} = 0.07$, $P = .993$, $\eta^2_p < .001$, rmANOVA, Supplementary Fig. 2A). However, there was a significant main effect of stimulus for the three fractals selected as CS_1 (main effect of stimulus: $F_{2,96} = 3.25$, $P = .043$, $\eta^2_p = .06$, rmANOVA, Supplementary Fig. 2B) that was driven by higher rating of CS_1^n than CS_1^- ($t_{48} = 2.26$, $P = .028$, paired-samples t-test). This difference most likely resulted from the selection process and the inherent rank-ordering of the CS according to subjective value ratings. However, there were no significant differences between CS_1^n and CS_1^+ ($t_{48} = 1.68$, $P = .101$, paired-samples t-test). Most importantly, no rating differences were observed between CS_1^- and CS_1^+ ($t_{48} = 0.44$, $P = .659$, paired-samples t-test). Overall, US valence ratings did not differ from pre to post rating across both US (main effect of time (pre/post): $F_{1,48} = 0.34$, $P = .564$, $\eta^2_p = .01$, rmANOVA), but the appetitive US^+ was – as expected – rated as significantly more pleasant than the aversive US^- (main effect of stimulus: $F_{1,48} = 466.64$, $P < .001$, $\eta^2_p = .91$, rmANOVA). There was also a significant interaction effect between stimulus and time, indicating that US^+ was rated as more pleasant and US^- was rated as less pleasant from pre to post (interaction effect of stimulus x time: $F_{1,48} = 7.32$ $P = .009$, $\eta^2_p = .13$, rmANOVA, Supplementary Fig. 2C). Importantly, no subject rated US^- as more pleasant than US^+ (pre or post). In the fMRI study, intensity ratings did not differ between US overall (main effect of stimulus: $F_{1,28} = 1.05$, $P = .315$, $\eta^2_p = .04$, rmANOVA), but US were rated as significantly more intense at the post rating (main effect of time: $F_{1,28} = 4.89$, $P = .035$, $\eta^2_p = .15$, rmANOVA, Supplementary Fig. 2D). There was also a significant interaction effect between stimulus and time, indicating that US^- intensity rating increased more strongly than US^+ from pre to post (interaction effect of stimulus x time: $F_{1,28} = 5.62$, $P = .025$, $\eta^2_p = .17$, rmANOVA). The average temperature (°Celsius) of US^+ ($M = 22.49$, $SD = 1.46$) was significantly higher than temperature of US^- (M

= 22.37, *SD* = 1.52) before being loaded into the syringes (t_{27} = 2.22, *P* = .035, paired-samples t-test). However, we would argue that this minor difference in average temperatures (*M* = .12) is unlikely to be perceived by participants. Additionally, this difference does not meaningfully influence the observed results, since no neural activation was recorded during first-order conditioning and the multivariate classification analysis was trained on the data resulting from the fMRI classifier training experiment.

In the fMRI classifier training experiment, overall US valence ratings were higher in pre than in post rating (main effect of time (pre/post): $F_{1,28}$ = 11.67, *P* = .002, η^2_p = .29, rmANOVA), but the appetitive US^+ was, as intended, rated as significantly more pleasant than the aversive US^- across ratings (main effect of stimulus: $F_{1,28}$ = 185.28, *P* < .001, η^2_p = .87, rmANOVA). There was no interaction effect between stimulus and time ($F_{1,28}$ = 0.28, *P* = .602, η^2_p = .01, rmANOVA). Importantly, no subject rated US^- as more pleasant than US^+ (pre or post). Intensity ratings for US^+ were higher than for US^- across ratings (main effect of stimulus: $F_{1,28}$ = 15.14, *P* < .001, η^2_p = .35, rmANOVA). There was no difference between US ratings from pre to post (main effect of time: $F_{1,28}$ = 0.02, *P* = .899, η^2_p < .001, rmANOVA). However, there was also a marginal interaction effect between stimulus and time, indicating that US^- intensity rating increased while US^+ intensity rating decreased from pre to post (interaction effect of stimulus x time: $F_{1,28}$ = 3.93, *P* = .057, η^2_p = .12, rmANOVA). Average temperature of US^+ (*M* = 21.71, *SD* = 1.93 °C) and US^- (*M* = 21.68, *SD* = 1.95) did not differ before being loaded into the syringes (t_{27} = 0.71, *P* = .486, paired-samples t-test).

fMRI searchlight classification

In the main article, we present evidence for reinstatement of US patterns in a region in the left lateral orbitofrontal cortex (lOFC). In this region, significant decoding of CS identity was possible on the basis of the US pattern obtained during classifier training on the previous day using classification with a 3-mm searchlight and small-volume correction in a ROI defined by a meta-analysis based on the term "taste" in the Neurosynth data base. To ensure that our results are not dependent on this particular choice of ROI, we repeated the same analysis using two different, independent ROI. The first was an anatomical mask of the left lateral orbitofrontal cortex (Harvard-Oxford atlas), the second was obtained from an independent gustatory mapping study by Benz and colleagues (K. Benz, personal communication, 12/2019). Using random-effect cluster-statistics with 50,000 iterations, we found similar results in the left lOFC anatomical ROI (peak voxel at MNI [x = –21, y = 30, z = –17], *z* = 1.96, *P* = .024, corrected, one-tailed) and in the mask from Benz and colleagues (peak voxel at MNI [x = –21, y = 30, z = –17], *z* = 1.84, *P* = .033, corrected, one-tailed).

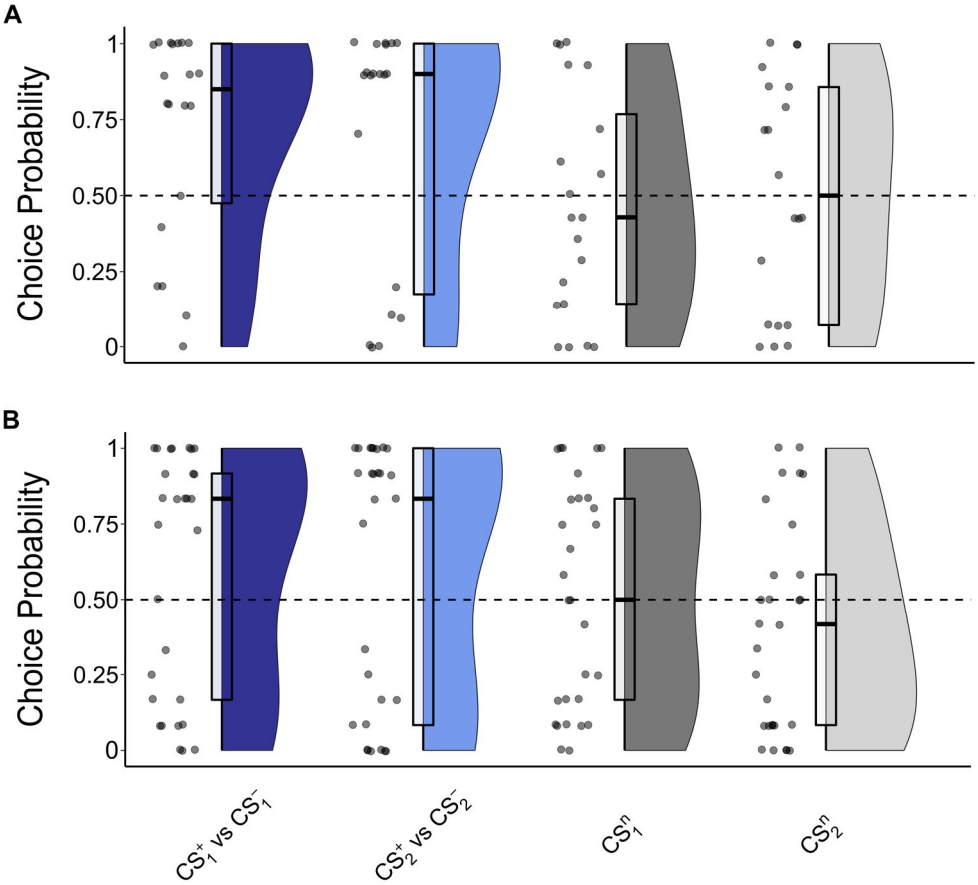
1171 **Association between second-order CS and US in the lOFC**

1172 When repeating the analyses performed in the amygdala within the lOFC cluster in which it
1173 was possible to predict CS labels based on US patterns, we did not observe significant
1174 pattern similarity between CS_2^- and US^- (t_{28} = 0.06, P = 0.955, $U3_1$ = 0.52), nor between CS_2^+
1175 and US^+ (t_{28} = 0.71, P = 0.489, $U3_1$= 0.52). There was no indication of increases of pattern
1176 similarity in this region (CS_2^- and US^-: t_{28} = 0.02, P = 0.491, Cohen's d = 0.005; CS_2^+ and
1177 US^+: t_{28} = 0.33, P = 0.629, Cohen's d = 0.06).

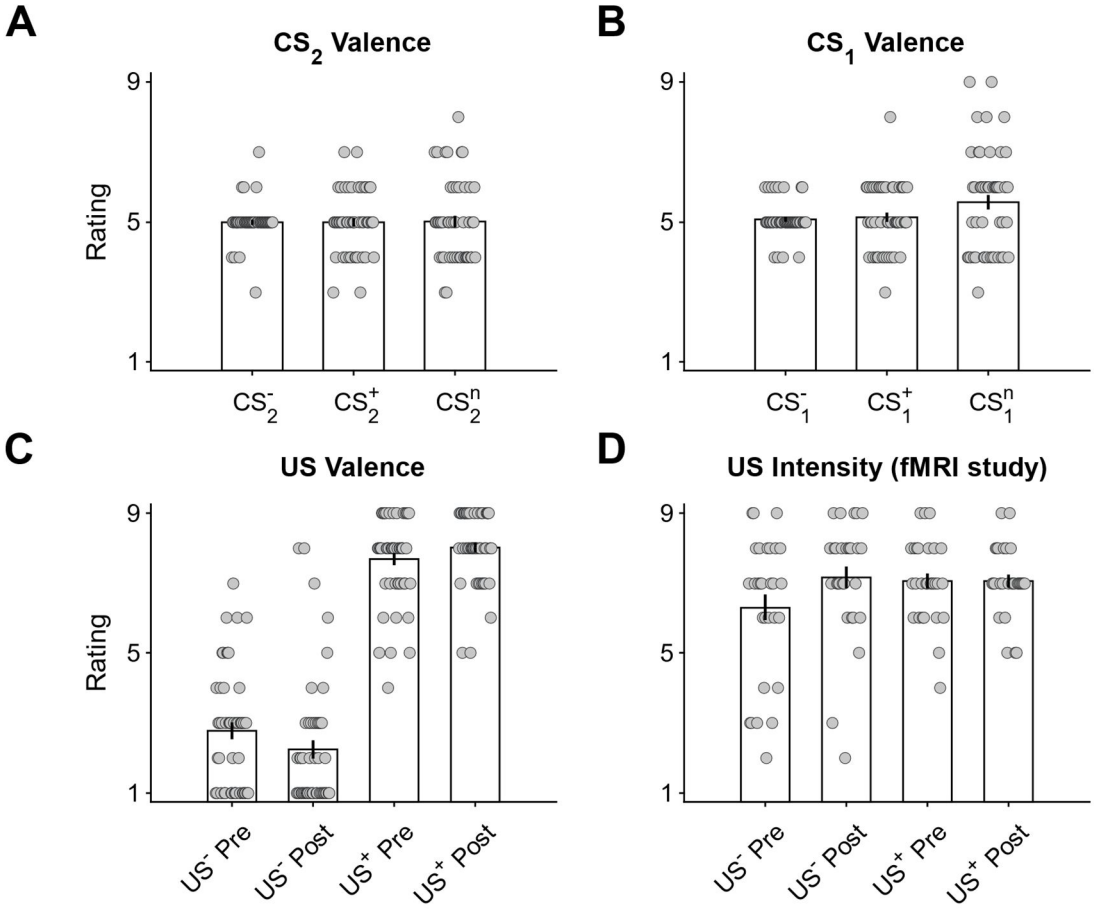
1178

1179

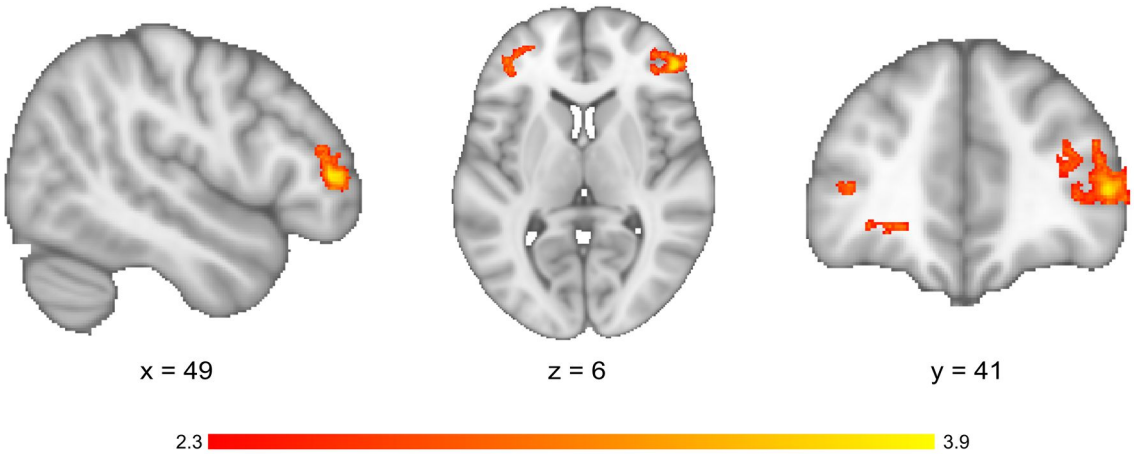
1180



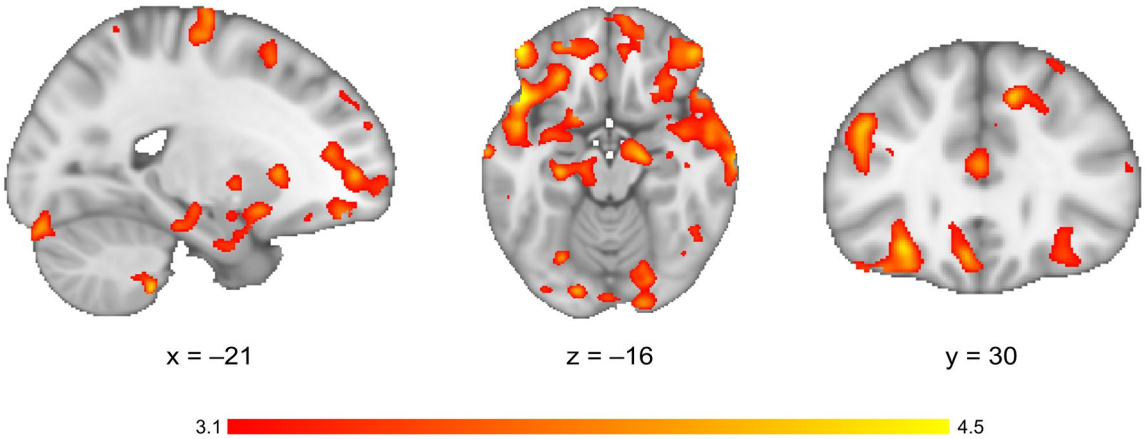
Supplementary Fig. 1. Extended behavioral results.
A) Choice probabilities in the behavioral study (*N* = 20): Raincloud plots showing density of choice probability. CS_1^+ (dark blue, 89%-highest posterior density interval (89%-HPDI): [.62; .72]) and CS_2^+ (light blue, 89%-HPDI: [.65; .77]) were preferred over CS_1^- and CS_2^-, respectively. HPDIs did not overlap (0% overlap) with the defined region of practical equivalence (ROPE, CP = [.45; .55]). There was no evidence for choice probabilities of CS_1^n and CS_2^n being different from chance level in novel-to-lure comparison choice trials (dark and light gray, 89%-HPDIs: [.42; .52] and [.46; .56], 68.3% and 94.8% ROPE overlap, respectively). B) Choice probabilities in the fMRI study (*N* = 29). CS_1^+ (dark blue, 89%-HPDI: [.54; .63]) and CS_2^+ (light blue, 89%-HPDI: [.65; .77]) were preferred over CS_1^- and CS_2^-, respectively. HPDIs did only minimally overlap (4.6% and 0.8% overlap, respectively) with the defined ROPE. There was no evidence for choice probabilities of CS_1^n being different from chance level in novel-to-lure comparison choice trials (dark gray, 89%-HPDIs: [.47; .55], 99.2% ROPE overlap). However, participants neglected CS_2^n compared to control stimuli (light gray, 89%-HPDI [.36; .44], 0% ROPE overlap). Box plot center lines represent sample medians and box bottom/top edges show $25^{th}/75^{th}$ percentile of the data, respectively.



Supplementary Fig. 2. CS and US ratings.
A) During the ratings prior to the learning experiment, pooled across both studies (N = 49) the subjective values/liking of the three kanjis selected as CS_2 did not differ significantly from each other ($F_{2,96}$ = 0.07, P = .993, η^2_p < .001, rmANOVA). B) There was a significant main effect of stimulus for the three fractals selected as CS_1 ($F_{2,96}$ = 3.25, P = .043, η^2_p = .06, rmANOVA) that was driven by higher rating of CS_1^n than CS_1^- (t_{48} = 2.26, P = .028, paired-samples t-test). However, there were no significant differences between CS_1^n and CS_1^+ (t_{48} = 1.68, P = .101, paired-samples t-test). Most importantly, no rating differences were observed between CS_1^- and CS_1^+ (t_{48} = 0.44, P = .659, paired-samples t-test). C) US valence ratings did not differ from pre to post rating across both US (main effect of time (pre/post): $F_{1,48}$ = 0.34, P = .564, η^2_p = .01, rmANOVA), but the appetitive US^+ was – as expected – rated as significantly more pleasant than the aversive US^- (main effect of stimulus: $F_{1,48}$ = 466.64, P < .001, η^2_p = .91, rmANOVA). There was also a significant interaction effect between stimulus and time, indicating that US^+ was rated as more pleasant and US^- was rated as less pleasant from pre to post (interaction effect of stimulus x time: $F_{1,48}$ = 7.32 P = .009, η^2_p = .13, rmANOVA). D) In the fMRI study (N = 29), intensity ratings did not differ between US overall (main effect of stimulus: $F_{1,28}$ = 1.05, P = .315, η^2_p = .04, rmANOVA), but US were rated as significantly more intense at the post rating (main effect of time: $F_{1,28}$ = 4.89, P = .035, η^2_p = .15, rmANOVA). There was also a significant interaction effect between stimulus and time, indicating that US^- intensity rating increased more strongly than US^+ from pre to post (interaction effect of stimulus x time: $F_{1,28}$ = 5.62, P = .025, η^2_p = .17, rmANOVA). Bar plots represent sample means and error bars indicate standard errors of the means, respectively.



Supplementary Fig. 3. Psychophysiological interaction analyses: Association with CS_2^+ versus CS_2^- choice probability.
A positive correlation was observed between second-order choice preferences (CS_2^+ versus CS_2^-) and functional covariation between the left lOFC and a region in the right and left lateral prefrontal cortex (right peak voxel at MNI [x = 49, y = 41, z = 7], Z = 3.80, $P < 0.001$, whole-brain corrected; left peak voxel at MNI [x = –45, y = 46, z = –4], Z = 3.93, $P = 0.010$, whole-brain corrected). The more participants preferred CS_2^+ (versus CS_2^-), the higher these regions' activation covaried during second-order conditioning. Color bar represents *Z*-values.



Supplementary Fig. 4. US$^+$ versus US$^-$ classification.
Group average (*N* = 29) cross-validated predictive accuracy for US$^+$ versus US$^-$ activation patterns during the classifier training experiment. As an exploratory analysis, we employed multivariate searchlight classification (3-mm searchlight spheres) to distinguish between the spatial activation patterns (*t* maps) of US$^+$ versus US$^-$ across the whole brain, resulting in one whole-brain map of classification accuracies per participant. This classification was performed using a 5-fold cross-validation scheme, iteratively training the classification algorithm on the difference between US$^+$ versus US$^-$ on labeled training data sets (4 out of 5 runs) and then assessing the out-of-sample predictive accuracy of the classifier on the remaining, left-out data set (1 out of 5 runs). Data was averaged across participants and smoothed with a Gaussian kernel with 5 mm FWHM. Please note that due to the exploratory nature and illustrative purpose of this analysis, no multiple comparisons correction using random-effects cluster-statistics was performed. Color bar represents uncorrected *Z*-values.

1250 **Supplementary Table 1. Explicit knowledge of CS-US associations**

CS_1–US (*N* = yes)	CS_1–US (*Mean* + *SD*)	CS_2–US (*N* = yes)	CS_2–US (*Mean* + *SD*)
38 (78 %)	.98 (.95)	14 (29 %)	.31 (.68)

1251

1252 *Note:* Number (*N)* and percentage (in parentheses) of participants indicating explicit
1253 knowledge of CS–US associations between directly paired CS_1 and US (providing an answer
1254 to the question which CS_1 was associated with which US, regardless of being correct or
1255 incorrect) and indirectly paired CS_2 and US (responding "yes" to the question whether there
1256 was a relationship between CS_2 and US). These numbers of participants were significantly
1257 above (CS_1–US) and below (CS_2–US) the numbers expected under the assumption of
1258 random guessing, $P < .001$ and $P = .004$, respectively (Binomial test vs. 0.5). Additionally, we
1259 provide the mean (and standard deviation, *SD*) number of correct answers for CS–US
1260 associations between CS_1 and US, and CS_2 and US. Since participants were presented with
1261 two CS–US associations during conditioning, the maximum number of correct answers is 2 in
1262 both cases.

1263
1264