Interactions between mechanisms for memory and decision-making

Inaugural-Dissertation

zur Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Lennart Lüttgau

aus Oldenburg (Niedersachsen)

Düsseldorf, September 2020

Aus dem Institut für Experimentelle Psychologie,
Biologische Psychologie des Entscheidungsverhaltens
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Gerhard Jocham
2. Prof. Dr. Tobias Kalenscher

Tag der mündlichen Prüfung:

18.12.2020

Danksagung – Acknowledgements

In den vergangenen Tagen und Wochen habe ich viel über meine Promotionszeit nachgedacht. Es fühlt sich gut an, solch umfangreiche Forschungsprojekte zu Ende führen zu können und die Promotion abzuschließen. Glücklich bin ich über all das, was ich lernen durfte und über die Perspektive, meine wissenschaftliche Karriere fortsetzen zu können. Allerdings schaue ich auch ein wenig traurig auf das Ende einer guten, aber auch herausfordernden Phase meines Lebens zurück.

Ich bin unendlich dankbar für die Chance und die Möglichkeiten, die mir mit dieser Stelle geboten wurden. Ich habe so viel gelernt, über mich selbst, faszinierende neue methodische Ansätze, über die Psychologie und die kognitiven Neurowissenschaften – und auch über die Wissenschaft selbst. Mein größtmöglicher Dank geht daher an meinen Betreuer Gerhard Jocham, der in allen Aspekten der Promotion auch zu später Stunde für mich da war (ich glaube die späteste Antwort auf eine meiner Fragen habe ich mal nachts um halb 4 auf Skype von Gerhard bekommen) und dessen Leidenschaft für die Wissenschaft mich tief beeindruckt und geprägt hat. Ich habe selten jemanden kennengelernt, der so brennt für das was er tut und hätte mir keinen besseren Betreuer vorstellen können. Danke Gerhard, für alles was ich von dir lernen durfte, für die vielen guten Gespräche über Wissenschaft, Politik und Zeitgeschehen und für die Kritik und Diskussion all meiner mehr oder weniger guten Ideen.

Mir ist bewusst, dass ich heute nicht an diesem Punkt stünde, wenn nicht sehr viele mir wichtige Menschen, begünstigende gesellschaftliche Faktoren und Privilegien dazu beigetragen hätten. Ich hoffe, ich habe die mir gegebene Chance so gut wie möglich genutzt.

Meiner Familie, Sabine Hegeler-Lüttgau, Christian Lüttgau, Gunnar Lüttgau, Hanna Lüttgau und meiner Oma Renate Lüttgau möchte ich dafür danken, dass sie mich immer auf meinem Weg begleitet und unterstützt haben. Euer Interesse an meiner Forschung und eure Fragen beziehungsweise Perspektive darauf waren für mich immer eine Möglichkeit zur kritischen Reflektion und haben mir geholfen, meine Arbeit in einem anderen, größeren Kontext zu betrachten. Ich weiß, dass es nicht selbstverständlich ist mit beiden Elternteilen und Geschwistern aufzuwachsen und bin mir sicher, dass viele Aspekte meines Lebens heute anders wären, wenn es euch, meine Eltern, meinen Bruder und meine Schwester, nicht gäbe. Dafür bin ich euch, Mama, Papa, Gunnar und Hanna, für immer dankbar. Papa, ich weiß, dass du in diesem Moment gerne dabei wärst. Ich bin unendlich traurig, den Abschluss dieser wichtigen Phase meines Lebens nicht mit dir zusammen erleben zu können. Wo auch immer du das hier liest, ich will dass du weißt, wie wichtig du für mich und meine Entwicklung warst und dass ich unsere Diskussionen sehr vermisse. Ach ja, und deiner (und Gunnars) Elektroingenieurskunst verdanke ich den reibungslosen Ablauf von Experiment I.

Meine liebe Anne, ich weiß, dass ich dich mit meiner Promotion wahrscheinlich mehr als jeden anderen Menschen beansprucht habe. 4 Jahre Fernbeziehung, 3 Jahre mittlere Distanz (Magdeburg-Leipzig) und danach ein Jahr eine recht weite Fernbeziehung (Düsseldorf-Leipzig) sind nicht selbstverständlich und ich kann dir nicht genug dafür danken, dass du diesen weiten Weg mit mir zusammen gegangen bist. Ich bin so glücklich darüber, dich gefunden zu haben und bin dir unendlich dankbar dafür, dass du für mich immer ein Ort auf dieser Welt bist, an den ich gerne zurückkehre. Der Platz hier ist zu knapp, um dir zu sagen, was du für mich bedeutest, deshalb einfach: Danke, liebe Anne, ohne dich wäre das hier nicht möglich gewesen!

Luca, du warst für mich eine echte Partnerin „in crime“ während der Promotion, nicht nur im Lab sondern auch außerhalb habe ich mit dir am meisten geteilt und so viele coole, aber auch blöde Sachen erlebt. Deine unglaubliche Cleverness und deine kritische Sicht in allen wissenschaftlichen Belangen waren für mich immer ein Ansporn, Dinge besser machen zu wollen, vielleicht auch besser als ich es jemals ohne dich und deine Hilfe gekonnt hätte. Unvergessen bleiben für mich unsere

privaten Journal Clubs mit Hummus-Nudeln und Papern (eigentlich wollten wir dazu auch immer Bier trinken, haben wir aber irgendwie nie gemacht) und die vielen Abende mit Tatort, in Kneipen oder auf Partys. Ich hoffe, du bleibst der Wissenschaft erhalten und bin dankbar ein paar Schritte mit auf dem Weg eines großartigen Menschen gegangen zu sein.

Julia und Franziska, danke für die schöne Zeit mit euch. Unsere Mittagessen + Kaffeetrinken, unsere Gespräche und erkenntnistheoretischen Diskussionen vor der OVGU Mensa vermisse ich sehr. Schade, dass unsere Zeit schon so viel schneller als ich es mir gewünscht hätte vorbei gegangen ist! Ich bin froh, euch beiden als Freundinnen gewonnen zu haben und in Magdeburg und Leipzig besuchen zu können. Danke, dass ihr immer für mich da wart und meine Zeit in MD so bereichert habt.

Theo, danke für deine jederzeit verfügbare und unterstützende Art. Ich habe viel von dir bezüglich Coding, Cluster-Nutzung und vor allem über das kritische Hinterfragen von Annahmen und Ergebnissen gelernt. Danke, dass du in der schwierigen Zeit des Wechsels von Magdeburg nach Düsseldorf da warst und immer ein offenes Ohr für mich hattest, um mich mit guten Ratschlägen zu unterstützen. Es ist schade, dass sich unsere Wege schon getrennt haben, aber ich bin mir irgendwie fast schon sicher, dass wir uns nicht zum letzten Mal gesehen haben.

Karsta, deiner umfangreichen Kenntnisse über die Konstruktion und Anwendung von Gustometern, sowie die großzügige Spende gustatorischer Stimuli verdanke ich das Überwinden der anfänglichen Startschwierigkeiten von Experiment I. Danke für die coolen Zeiten mit Chinin-Hydrochlorid, Kaffee, Bier und guten Gesprächen.

Halla Mulla-Osman, Stefanie Linnhoff, Leonie Oevel, Carolin Kirchhoff, Nicola Harzen, Joshua Saal, Judith Geusen, Georg Schäfer, Christina Kalinicenko: Euch allen möchte ich ganz besonders für eure Unterstützung und für euer großes Engagement bei der Probanden-Akquise und Daten-Erhebung für meine Promotionsprojekte danken. Halla, Caro, Nicola und Joshua, danke für die super Zusammenarbeit und für euer Vertrauen in mich im Rahmen eurer Bachelor-Arbeiten. Es war eine große Bereicherung für mich, euch betreuen zu dürfen und eure großartige Entwicklung zu sehen. Ich habe immer versucht, euch so zu betreuen, wie ich es mir für mich selbst wünschen würde und auch selbst erfahren habe, sicherlich nicht immer perfekt und ohne Fehler. Ich bin mir absolut sicher, dass ihr bei allem was ihr euch vornehmt, erfolgreich sein werdet.

Denise Scheermann, vielen Dank für die unterhaltsamen Stunden am Scanner und die unkomplizierte (beinahe störungsfreie), aber unendlich erscheinende Erhebung der zwei fMRT-Studien. Das war ein großes Stück Arbeit und ich bin froh, den Weg mit dir zusammen gegangen zu sein.

Coconuts, I would like to thank all of you from the deepest bottom of my heart for your continued support and for the great team that we have been for the last one and a half years in Düsseldorf. Anna, your enthusiasm for Bayesian statistics, modelling and your willingness to go an extra mile (from Ghent) strongly impresses me. I am glad to have met you and hope our “rethinking” will remain a continued source of inspiration. Armin, it’s been amazing to see your development from M.Sc. to Ph.D. student. Your enthusiasm is an example of how to spark interest for cool topics in other people. Christiane, thank you very much for the generous supply with any kind of sweets one can imagine and for keeping the organizational work to be done at a tolerable level for all of us. I know it must be difficult with a bunch of weird people like us at times, that’s why I particularly enjoyed having you around for a chat and a cup of tea. Eduard, your skepticism and critical mindset combined with your unique sense of humor and blue Adidas “bicycle shorts” have made life in the lab both more enjoyable and challenging (in the good way) for me. Keep up that good spirit, even in times when science and academia might not seem to be on your side. In fact, they are. Always have been. Hannah, my roommate, thank you so much for a healthy mix of funny conversations and clear,

focused periods of work and your great supply with chocolate, cakes and other self-made sweets. I couldn't think of a better person to share an office with and I am super convinced that you will continue being a wonderful colleague, be it in academia or anywhere else. Mani, it was only a short period of us overlapping in the lab, but I greatly enjoyed having you around and hope you will have such a good time as I had, being a Coconut. Monja, thank you for keeping all of us together and for always having an open ear and open mind for discussions, both scientifically and on a personal level. Please keep on going, science and academia needs more people like you. Thank you, girls and boys, it was a fun ride!

Mario Borja Perez-Lopez, thanks for the super cool time collaborating during our project MAPLE. I only wished it had turned out more favorable for all of us. I am glad we have met and that I could learn a lot from your way of thinking and working. I have absolutely no doubt that you will continue being a great colleague and researcher.

Emanuele Porcu, thank you so much for all your support in MVPA- and other fMRI-related questions. It's funny how life intertwines people's roads several times. I hope that ours will cross again at one point.

Alex Waite, thank you for being the single best systems administrator one can imagine, for always helping me out with troubling Medusa-related issues and for providing such a generous extended storing time for my experimental and neuroimaging data. I greatly appreciated having you around during my American Football games in Magdeburg.

Mona Garvert möchte ich für die sehr hilfreichen Diskussionen und ihre Ideen zu meinen Daten aus Experiment II A-E während der Psychologie & Gehirn Konferenz im Juni 2019 in Dresden danken. Danke auch für eure Mind Meetings und die von dir ermöglichten Treffen mit den Vortragenden (in Prä-Corona-Zeiten) – eine stete Quelle der Inspiration und ein Grund für mich, in den letzten 1,5 Jahren noch lieber als ohnehin nach Leipzig zu kommen. Ich habe mich am MPI (wieder) ein bisschen zu Hause gefühlt.

Tom, mein bester Freund, ich danke dir für deine Freundschaft und deine Unterstützung während der letzten 8 Jahre. Ich bin so dankbar dafür, dass unsere Verbindung über jede zeitliche und räumliche Trennung erhaben ist.

Meinen beiden WGs möchte ich für die unglaublich schöne Zeit in Magdeburg danken. Johanna, Kathi, Ludwig und Martin, ich werde nie die super gute Zeit mit euch zusammen vergessen. Danke für all die Gespräche und Diskussionen zu später Stunde, für die aufmunternden Worte in schwierigen Phasen, die freitagmorgendlichen Deadlift-Sessions, die WG-Abende und -Partys, die Abende mit South Park, Family Guy, Game of Thrones und Nougat Bits und für das umfassende „casual" Medizin-Wissen, welches ich durch euch erlangen konnte. Ich bin froh, euch kennengelernt zu haben und unglaublich dankbar, dass ihr mich immer wieder als Probandinnen in meinen Projekten unterstützt habt. Inessa, Fred (und Jule), Lenard, Nadine, danke für ein cooles letztes Jahr in Magdeburg, ich habe die Zeit mit euch sehr genossen. Unsere CRISPR-Diskussionen, Gespräche über Epidemiologie und Statistik, die brutal geführten Partien (Street-)UNO, unsere WG-Abende und vor allem eure Überraschungs-Abschiedsparty für mich behalte ich für immer in sehr guter Erinnerung.

Marcus, Fabian, Robert, Sven (Motze) und Friedemann (Peaceman), a.k.a #hatesquad, danke ich für die zahlreichen unterhaltsamen, aufgebrachten Hin- und Rückfahrten zum Football-Training und all unsere Insider. Es war eine super Zeit mit euch. Ich möchte hier an dieser Stelle generell dem Football in Magdeburg und meinem Team, den Magdeburg Virgin Guards für 3 wunderbare Saisons danken. Vor allem Marko, der gesamten O-Line, Harry, Heiko, PW. Ohne euch, das Team und

diesen Sport wäre meine Zeit in Magdeburg nicht einmal halb so schön gewesen und ich hätte meine Promotion niemals mit derselben Energie angehen können. Ich danke euch für alles.

Der Barista in der Primo Cafebar im Gebäude 22 der OVGU Magdeburg möchte ich für den jederzeit guten Kaffee danken. Vor allem die diversen gratis Snacks haben mir schwierige Phasen ein bisschen besser erträglich gemacht.

Ein letzter Dank gilt den mannigfaltigen Tiramisu-Variationen aus der Mensa in der OVGU Magdeburg. Ich wünschte, wir hätten noch länger etwas voneinander gehabt. Möget ihr noch vielen Generationen von Doktorandinnen und Doktoranden den Alltag versüßen.

Zusammenfassung

Das Entscheidungsverhalten und das assoziative Gedächtnis sind allgegenwärtige kognitive Prozesse, die eine hohe Alltagsrelevanz für Menschen und andere Tiere aufweisen. Daher verwundert es nicht, dass beide Prozesse über lange Forschungstraditionen als separate Disziplinen in der Psychologie und in den (kognitiven) Neurowissenschaften verfügen. Allerdings entstand durch diese getrennte Betrachtung der Eindruck, beide Prozesse liefen vollkommen unabhängig beziehungsweise getrennt voneinander ab und die Forschung fokussierte sich auf eine isolierte Betrachtung beider Phänomene. Erst kürzlich entwickelte sich die Idee, dass sich Gedächtnis- und Entscheidungsprozesse gegenseitig beeinflussen und sogar systematisch verzerren könnten. Neoklassische ökonomische Modelle des Entscheidungsverhaltens nehmen an, dass Gedächtnisrepräsentationen von Präferenzen und der Wertigkeit von Entscheidungsoptionen einseitig das Entscheidungsverhalten bestimmen. Allerdings beinhalten viele Entscheidungssituationen komplett neue Stimuli, Ereignisse oder situationale Gegebenheiten, für die die Entscheiderin keinerlei Vorerfahrungen besitzt. Darüber hinaus stellen sich nicht nach allen Entscheidungen direkt belohnende oder bestrafende Konsequenzen ein und häufig entwickeln sich die Folgen von Entscheidungen erst sehr viel später als die konkrete Entscheidungssituation. Unter diesen Bedingungen ist es äußert schwierig, ausschließlich fehlerbasiert oder durch Belohnungslernmechanismen zu lernen. Es erscheint daher unplausibel, dass alle Entscheidungen auf die durchschnittliche, lerngeschichtlich erworbene Wertigkeit von Entscheidungsoptionen zurückzuführen sind. Wahrscheinlicher ist, dass andere Mechanismen, wie etwa assoziative Gedächtnisprozesse, genutzt werden, um Wertrepräsentationen in neue Entscheidungskontexte zu generalisieren beziehungsweise zu übertragen. Weiterhin ignorieren neoklassische ökonomische Modelle des Entscheidungsverhaltens die Möglichkeit, dass das Entscheidungsverhalten selbst einen verändernden Einfluss auf Wertigkeits- und Präferenz-Repräsentationen im Gedächtnis haben könnte. In der vorliegenden Arbeit wurde daher untersucht, ob es zu zweiseitigen Interaktionen zwischen dem Entscheidungsverhalten und assoziativen Gedächtnisprozessen kommt.

Um diese Interaktionen zu untersuchen wurden gesunde Normalprobandinnen mit Konditionierungsparadigmen höherer Ordnung und klassischen Konditionierungsexperimenten getestet. Zusätzlich kamen gedächtnisbasierte Entscheidungsaufgaben zum Einsatz, mit denen Präferenz und Präferenzänderungen auf der Verhaltensebene gemessen werden sollten. Wir nutzten computationale Modellierung um mögliche Lernmechanismen die das beobachtete Entscheidungsverhalten beschreiben können zu quantifizieren und um verschiedene Modelle zu vergleichen. Zur Messung von Gedächtnisrepräsentationen und der Veränderung der assoziativen Stärke zwischen Stimuli, nutzten wir drei repräsentationale Neurobildgebungstechniken auf funktionellen Magnetresonanztomografie (fMRT)-Daten: fMRT-Adaptation, Ähnlichkeitsanalysen multivariater neuronaler Muster und klassifikationsbasierte multivariate Musteranalyse.

Im Entscheidungsverhalten der Probandinnen nach einem Konditionierungsprozess höherer Ordnung fanden sich Belege für eine Übertragung von Wertrepräsentationen. Zu diesem

Entscheidungsmuster kam es obwohl Probandinnen über kein explizites Wissen über die (höhergeordnete) assoziative Struktur des Lernexperiment verfügten, was darauf hindeutet, dass Menschen Wertigkeiten von Entscheidungsoptionen implizit über Lernmechanismen höherer Ordnung erlernen können. Auf neuronaler Ebene gab es Belege dafür, dass die verwendeten gustatorischen unkonditionierten Stimuli (und damit einhergehend vermutlich der durch diese Stimuli transportierte motivationale Zustand oder die Wertigkeit) im linken lateralen orbito-frontalen Kortex durch die vorher damit gepaarten konditionierten Stimuli erster Ordnung reaktiviert wurde. Zusätzlich fanden sich Anzeichen dafür, dass eine direkte assoziative Verbindung zwischen dem durch den unkonditionierten Stimuli transportierten motivationalen Zustand oder dessen Wertigkeit und dem konditionierten Stimulus zweiter Ordnung in der Amygdala erzeugt wurde.

In einer weiteren Reihe von Experimenten mit einem neu entwickelten Lern- und Entscheidungsparadigma fanden sich konsistente Belege für entscheidungsinduzierte Präferenzänderungen: Zuvor ausgewählte Entscheidungsoptionen wurden häufiger gewählt, für nicht ausgewählte Optionen zeigten sich verringerte Präferenzen – im Vergleich zu ansonsten gleichwertigen Optionen. Diese Entscheidungseffekte wirken scheinbar in entgegengesetzte Richtungen und konnten entscheidungsinduzierte Präferenzänderungen bei Optionen, die gleich häufig ausgewählt und nicht ausgewählt wurden, aufheben. In einem zusätzlichen Experiment konnte zudem ausgeschlossen werden, dass die beobachteten entscheidungsinduzierten Präferenzänderungen ausschließlich auf akquirierte Stimulus-Reaktions-Tendenzen zurückführbar sind. Diese entscheidungsinduzierten Präferenzänderungen traten auf, ohne dass Probandinnen jemals die Konsequenzen ihrer Entscheidungen präsentiert wurden. Auf neuronaler Ebene gab es Belege dafür, dass die assoziative Stärke der Stimulus-Belohnungs-Verbindung zuvor gewählter Optionen gestärkt, wohingegen die assoziative Stärke der Stimulus-Belohnungs-Verbindung zuvor nicht gewählter Optionen abgeschwächt wurde. Diese Veränderungseffekte zeigten sich im linken Hippocampus und im rechten lateralen orbito-frontalen Kortex. Darüber hinaus waren stärkere entscheidungsbezogene Erhöhungen der assoziativen Stärke der Stimulus-Belohnungs-Verbindung assoziiert mit erhöhter Präferenz für zuvor ausgewählte Optionen.

Insgesamt deuten die Experimente der vorliegenden Arbeit daraufhin, dass es zu zweiseitigen Interaktionen zwischen dem Entscheidungsverhalten und assoziativen Gedächtnisprozessen kommt. Es konnte gezeigt werden, dass erlernte assoziative Strukturen höherer Ordnung für die Übertragung von Wertrepräsentationen auf Stimuli, die selbst nie direkt mit einer Belohnung oder Bestrafung gepaart wurden, genutzt werden können. Außerdem zeigte sich, dass entscheidungsinduzierte Präferenzänderungen durch die Veränderung von assoziativen Gedächtnisinhalten entstehen könnten. Entgegen der einseitigen Perspektive neoklassischer ökonomischer Modelle des Entscheidungsverhaltens, deuten beide Hauptergebnisse dieser Arbeit darauf hin, dass Werterepräsentationen im Gedächtnis und das Entscheidungsverhalten sich gegenseitig beeinflussen.

Abstract

Decision-making and associative memory represent ubiquitous cognitive processes that are highly relevant in humans' and other animals' everyday lives. Consequently, both processes have long standing traditions of investigation – yet within two separate fields of research in psychology and (cognitive) neuroscience. For decades of research, both processes have been viewed as distinct entities and were therefore studied mostly in isolation. Only recently it has been suggested that both memory and decision-making processes might influence and bias each other. Neo-classical economic models of decision-making have proposed that memory representations of preferences and choice option values unidirectionally guide or influence choice behavior. However, many decision scenarios involve entirely new stimuli, events or situational setups for which previous experience is lacking. Additionally, decisions are often not immediately followed by reinforcing stimuli or need to be taken in situations where the consequences arise at sometimes drastically long timescales. Both these circumstances render learning based on error-driven, reinforcement learning-based strategies challenging. Thus, it has been proposed that not all decisions are exclusively based on the averaged value resulting from a past learning history. Rather it seems that other mechanisms, like associative memory processes, are necessary for generalization and transfer of value to novel choice contexts. Moreover, neo-classical economic models neglect the possibility that choice behavior itself might have a modificatory influence on value and preference representations stored in the memory systems. In the present work, we therefore aimed at investigating bidirectional interactions between decision-making and associative memory processes.

To study these interactions, we investigated healthy human subjects and combined second-order conditioning and Pavlovian conditioning experiments with memory-based decision-making paradigms to assess preference (changes) on the behavioral level. We used computational modelling to quantify and compare model fits of candidate learning mechanisms that describe the observed choice behavior. To measure memory representations and changes of associative strength between stimuli, we employed three representational neuroimaging analysis techniques on functional magnetic resonance imaging (fMRI): fMRI repetition suppression, multivariate neural pattern similarity analyses and classification-based multivariate pattern analyses.

We found hallmarks of value transfer in the choice patterns following second-order conditioning. Importantly, this choice bias was present even though participants were unaware of the underlying (higher-order) associative learning structure of the experiment, indicating that humans implicitly acquire subjective value through higher-order learning mechanisms. The observed behavioral effect was paralleled by neuroimaging findings suggesting that neural patterns representing the administered gustatory unconditioned stimuli (and presumably the motivational state or value conveyed by these outcomes) are reinstated in the left lateral orbitofrontal cortex by previously paired first-order conditioned stimuli. Additionally, there was evidence for the formation of a direct associative link between the reinstated motivational state or value conveyed by outcomes to second-order conditioned stimuli in the amygdala.

Moreover, in a series of experiments employing a newly developed learning and decision-making paradigm, we found converging evidence for choice-induced preference changes: While previously chosen options were selected more frequently, unchosen options showed diminished preferences – compared to otherwise equivalent options. The choice effects were cancelled out when an option was chosen and unchosen equally often during choice-induced revaluation. Additional experimental results indicated that the observed choice-induced preference changes were unlikely to be exclusively driven by acquired stimulus-response tendencies. Importantly, preference changes occurred even without participants ever experiencing the consequences of their choice. There was neuroimaging evidence for choice-induced strengthening of stimulus-outcome associations of previously chosen options, whereas we observed weakening of stimulus-outcome associations of previously unchosen options in the left hippocampus and right lateral orbitofrontal cortex. Additionally, stronger hippocampal representations of stimulus-outcome associations were associated with higher preference for previously chosen options.

Taken together, the results of all experiments conducted in this work strongly suggest bidirectional interactions between decision-making and associative memory. Here we show that exploiting acquired higher-order associative structures stored in memory might support transfer of value to stimuli that had themselves never been directly coupled with reinforcement. Additionally, we found that choice-induced preference changes might arise from choice-related transformations of associative memories. Contrary to the unidirectional view of neo-classical economic models of decision-making, both our key findings suggest that value representations in memory and decision-making influence each other bidirectionally.

Contents

List of abbreviations

BOLD – Blood-oxygenation level-dependent

CP – Choice probability

CR – Conditioned responses

CS – Conditioned stimulus/stimuli

CS_1 – First-order conditioned stimulus/stimuli

CS_2 – Second-order conditioned stimulus/stimuli

fMRI – Functional magnetic resonance imaging

fMRT – Funktionelle Magnetresonanztomographie

fMRI-RS – fMRI repetition suppression

lOFC – Lateral orbitofrontal cortex

MVPA – Multivariate pattern analysis

ROI – Region(s)-of-interest

RPE – Reward prediction error

SOC – Second-order conditioning

US – Unconditioned stimulus/stimuli

UR – Unconditioned responses

1. Introduction

This work is dedicated to decision-making and associative memory processes, two cognitive abilities with tremendous implications for humans' and other animals' lives. Decision-making and associative memory represent two fields of research in psychology and (cognitive) neuroscience that each have long standing traditions. For the largest part of the past decades of research, however, both processes had been viewed as distinct entities and were therefore studied mostly in isolation. Only recently it has been proposed that both memory and decision-making processes might influence and bias each other (Shohamy & Daw, 2015). The present work therefore aimed at investigating interactions between decision-making and associative memory processes. In the following introductory section, I will first provide an overview of contemporary theories, models and (functional) neuroanatomical substrates of associative memory and highlight key empirical findings. Next, I turn to a selective description of (computational) theories of learning and decision-making and their neural correlates. Finally, I will highlight the literature investigating interactions between associative memory and decision-making and derive the main hypotheses of the present work.

1.1 Associative Memory

Associative learning and memory describe the acquisition and recall of relationships between stimuli, objects, events, or internal states (Mondragón, Alonso, & Kokkola, 2017), providing relational knowledge of the world. Since associative memory is ubiquitous and highly relevant for our everyday lives, most people have an intuitive understanding of what associative memory is and how associations might be acquired. Take the last time you met family or friends for instance: Upon recalling who exactly was present or remembering the place you met, one might also associatively retrieve the topics of conversation or what particularly funny joke your best friend made (e.g. "What's the best thing about Switzerland? I don't know, but the flag is a big plus"), easily enabling to tell this joke in front of colleagues.

The hippocampus is involved in the formation of associative links between items, stimuli or events, that occur in spatio-temporal proximity (Dusek & Eichenbaum, 1997; Eichenbaum & Cohen, 2001). Models and empirical findings have additionally suggested a pivotal role of the hippocampus in orchestrating the retrieval of episodic memory contents and associations between stimuli (Lisman et al., 2017; Staresina & Wimber, 2019; Tonegawa, Morrissey, & Kitamura, 2018; Yassa & Reagh, 2013). It has been proposed that the theta-dependent information flow from primary sensory areas across the sensory processing hierarchy towards the hippocampal formation observed during memory encoding is reversed during retrieval or reinstatement (Staresina & Wimber, 2019). Further, it is assumed that the hippocampus might indeed not represent episodic and associative memory content per se, but rather functions as a general purpose sequence generator storing memory-specific indices to preserve the spatio-temporal pattern of (sensory) neocortical neural populations that were activated during experience and encoding of events (Buzsáki, 2019; Lisman et al., 2017;

Staresina & Wimber, 2019). In this framework, the hippocampus stores and provides sequences organizing the memory-based reinstatement of cortical patterns representing sensory events in the neocortex, presumably tied to theta frequency oscillations (Buzsáki, 2019; Lisman et al., 2017; Liu, Dolan, Kurth-Nelson, & Behrens, 2019).

While the ability to encode and retrieve information is highly functional in everyday life, a counterintuitive finding from associative memory research suggests that remembering can cause forgetting (Wimber, Alink, Charest, Kriegeskorte, & Anderson, 2015). It has been shown that retrieval of a memory can selectively reduce the retrievability of competing memories (Anderson, Bjork, & Bjork, 1994; Hulbert & Norman, 2015; Storm, Bjork, & Bjork, 2008), for instance if two stimuli are associated with the exact same third stimulus and participants are asked to selectively recall one of these two associations (Anderson et al., 1994; Wimber et al., 2015). However, more recent accounts for memory retrieval dynamics, like the nonmonotonic plasticity hypothesis (Ritvo, Turk-Browne, & Norman, 2019), argue for unsupervised learning and U-shaped spreading activation underlying strengthening and weakening of retrieved memory contents: Inactive memories remain unaltered, whereas moderately activated memories are weakened and strongly retrieved memories are strengthened. The proposed U-shaped activation is motivated by the finding that moderate post-synaptic depolarization induces long-term depression, or synaptic weakening, whereas strong depolarization has been found to elicit long-term potentiation (Ritvo et al., 2019). While moderate activation might lead to differentiation of highly related memory contents, like competing memories, strong activation could facilitate integration of memories (Ritvo et al., 2019). Accordingly, retrieval-induced forgetting might result from strong activation of target associations and moderately activated competitor memories, leading to differentiation of the memory content and overall weakening of the competitor, impairing its future retrievability (Ritvo et al., 2019). Retrieval-induced forgetting might arise from prefrontal cortex-mediated suppression of neural patterns representing the competitor memory trace by the remembered target association and is assumed to have adaptive functionality (Wimber et al., 2015).

These theoretical perspectives and empirical findings suggest that the hippocampal formation does not act in isolation during memory encoding and retrieval. There is substantial evidence for strong bidirectional anatomical connections (Barbas & Blatt, 1995; Kondo & Witter, 2014) between the hippocampal formation and prefrontal cortical areas and for their functional interaction during memory encoding and retrieval (Wikenheiser & Schoenbaum, 2016). For instance, the lateral prefrontal cortex is involved in selectively suppressing neural patterns of competing memory during retrieval-induced forgetting (Wimber et al., 2015). Additionally, functional covariation between the hippocampus and individual change of reinstated future state representations in the orbitofrontal cortex has been observed in a maze-like task promoting learning about distal future rewards (Wimmer & Büchel, 2019). Moreover, both hippocampus and the medial prefrontal cortex seem to be involved in constructing prospective and compositional representations of novel rewarding stimuli (e.g. “tea jelly” – a linear combination of known stimuli: tea and jelly), dependent

on simultaneous activation of previously experienced rewards (Barron, Dolan, & Behrens, 2013). However, valuation of those novel goods is related to the medial prefrontal cortex only.

1.2 Decision-Making

Next to associative memory, decision-making is a vital part of our everyday lives. From deciding what kind of clothes to put on in the morning, which food to buy in the supermarket, to deciding for a career path and what kind of life one wants to live. A fundamental assumption in decision-making theories is that behavior aims at maximizing subjective utility (Ariely & Norton, 2008) and that choices can be predicted based on the difference between the (objective or subjective) values of available choice options (Padoa-Schioppa & Conen, 2017). Importantly, it is assumed that choices reveal preferences, i.e. that choice behavior reflects learned and remembered value structures in the world (Ariely & Norton, 2008). A plethora of seminal studies and theoretical models has implicated involvement of the orbitofrontal cortex/ventromedial prefrontal cortex and amygdala in the representation of value (Barron et al., 2013; Gottfried, O'Doherty, & Dolan, 2003; Jocham et al., 2016; Klein-Flugge, Barron, Brodersen, Dolan, & Behrens, 2013; Kringelbach, O'Doherty, Rolls, & Andrews, 2003; O'Doherty, 2004; Padoa-Schioppa & Assad, 2006; Padoa-Schioppa & Conen, 2017; Schultz, 2004; Seo & Sinha, 2011; Suzuki, Cross, & O'Doherty, 2017; Wang et al., 2018). Additionally, the orbitofrontal cortex/ventromedial prefrontal cortex plays a critical role in reward-guided decision-making, presumably by representing the difference between choice option values (Boorman, Behrens, Woolrich, & Rushworth, 2009; Jocham, Hunt, Near, & Behrens, 2012; Wunderlich, Range, & O'Doherty, 2009). This function might arise from neural dynamics underlying a biophysically plausible mechanism for choice behavior based on competition through mutual inhibition, that has been related to the balance of excitatory and inhibitory neurotransmission (Jocham et al., 2012; Kaiser, Gruendler, Speck, Luettgau, & Jocham, 2019; Soltani & Wang, 2010).

It has been proposed that value might be assigned to stimuli or choice options by error-driven learning mechanisms, like the reward prediction error (RPE). Computationally, RPEs are derived from the difference between the actually received reward and the expected value of a stimulus, an action, or a stimulus-action combination at a given time point (Sutton & Barto, 1998). Since its initial formulation, the reward prediction error theory of learning has been a powerful and influential framework for investigating how value is computed by dopamine neurons and used for decision-making in brains (Rescorla & Wagner, 1972; Schultz, Dayan, & Montague, 1997; Sutton & Barto, 1998). Dopaminergic ventral tegmental area and more downstream areas receiving dopaminergic innervation, like the ventral striatum, have consistently been associated with the neural computation of RPEs (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Deserno et al., 2015; Maes et al., 2020; Schultz et al., 1997; Sharpe et al., 2020; Takahashi et al., 2017).

Initially formulated as a quantitative account of Pavlovian conditioning (Rescorla & Wagner, 1972), the RPE computation enables individual tracking of environmental value functions in the form of running averages of values. Through experiencing both positive and negative mismatches in

varying magnitudes, RPEs allow approximation of the mean, or expected value, of the distribution of obtained outcomes, which is proportional to the respective stimulus/action/stimulus-action value. Despite the affordance of a quantitative account for behavior, the derivation of model-implied predictions for neural signals and its tremendous influence on (cognitive) neuroscience research in the past four decades, the canonical framework for value estimation has faced increasing challenges in recent years. More recent accounts have suggested that value assignment to stimuli or actions might not be driven by the mismatch between expected and received environmental outcomes, but could instead emerge from (active) reductions of the distance between current internal state variables (e.g. the level of hunger or thirst) and intrinsic goal states of the agent (Juechems & Summerfield, 2019). In this framework, the value or reward of stimuli, actions or stimulus-action combinations would be proportional to the expected minimization of the distance to the intrinsic goal. Furthermore, agents would aim at balancing the distances between a multitude of different intrinsic goal states and their respectively assigned internal state variables (Juechems & Summerfield, 2019).

Additionally, the implicit, canonical assumption that all dopamine neurons uniformly represent value to compute RPEs as a singular scalar quantity might be incompatible with recent empirical findings and computational models (Dabney et al., 2020). Firing rates of the many different dopamine neurons in the rat's ventral tegmental area have instead been found to vary drastically in response to reward receipt. This suggests that individual dopamine neurons might represent different information channels, conjunctively encoding a probability distribution over possible future rewards that represents a range of more "optimistic" to more "pessimistic" expected future outcome scenarios in parallel (Dabney et al., 2020). This distributional coding might facilitate weighting of different environmental scenarios with respect to their likelihood and uncertainty to motivate behaviors. Additionally, it has been suggested that neural RPE correlates might indeed encode more information than a scalar representation of the difference between the expected value and received rewards. For instance, it has been shown that value representations are most likely not learned in absolute terms (Klein, Ullsperger, & Jocham, 2017), but rather in relation to a reference point. This learning process shows hallmarks of adaptation to the range of presented outcome magnitudes in humans (Bavard, Lebreton, Khamassi, Coricelli, & Palminteri, 2018). Additionally, reward-responsive dopamine neurons in rats exhibit similar changes in firing rates as observed during violations of expected value, when sensory features of the reward (i.e. the reward identity), but not the value (i.e. the reward magnitude) itself is changed (Takahashi et al., 2017). These findings are difficult to reconcile with the canonical RPE framework.

Moreover, it is questionable whether all value learning and decision-making behavior observed in humans and other animals may exclusively be attributed to error-driven, or mismatch-based learning processes. Specifically, despite absence of external feedback there is evidence for improvement in perceptual decision-making performance driven by confidence-related signals, suggesting that mesolimbic confidence representations might be used as internally generated reinforcement signals for guidance and optimization of behavior (Guggenmos, Wilbertz, Hebart, &

Sterzer, 2016). Consistently, the internal representation of individuals' very own choice histories strongly biases subsequent decision-making processes (Urai, De Gee, Tsetsos, & Donner, 2019), and elicits adaptation to statistical properties of the environment even without experiencing the outcomes of choices (Braun, Urai, & Donner, 2018). Additionally, humans display preference increases for options they have chosen (Cockburn, Collins, & Frank, 2014; Leotti & Delgado, 2011) over and above what could be explained by reward history. It has been suggested that this valuation effect might result from dopamine-dependent amplification of RPEs elicited by free choices in the basal ganglia circuitry (Cockburn et al., 2014).

1.3 Interactions between Associative Memory and Decision-Making

Many decision scenarios that humans and other animals face every day involve entirely new stimuli, events or situational setups for which any previous experience is lacking. In other words, for many situations and stimuli encountered every day no value estimates based on error-driven learning are available (Biderman, Bakkour, & Shohamy, 2020). Moreover, decisions are often not immediately followed by reinforcing stimuli (Gewirtz & Davis, 2000) or need to be taken in situations where the consequences of the path taken unroll at sometimes drastically long timescales, rendering learning based on traditional error-driven strategies challenging, if not impossible. Despite these situational constraints affording tremendous cognitive flexibility, humans and other animals are – for the most part – capable of efficiently navigating through their lives. Thus, it has been proposed that not all decisions are exclusively based on the averaged value resulting from a past learning history (Biderman et al., 2020). Rather it seems that other mechanisms, like associative memory processes, are necessary for generalization and transfer of past experiences to novel choice contexts (Biderman et al., 2020) or for prospective planning in complex decision scenarios (Shohamy & Daw, 2015). Highlighting a potential way by which goal-directed behavior might be possible in novel contexts, empirical data show that decisions in humans are biased by reminder cues related to previously experienced episodes. These episodic reminder cues are assumed to associatively reinstate past choice contexts, performed actions and formerly received outcomes, over and above the influence of decision variables encoding the current choice context. Modelling results suggest that choice behavior might indeed be better explained by sampling of past actions from episodic memory than by error-driven computation and updating of running value averages for choice options (Bornstein, Khaw, Shohamy, & Daw, 2017; Bornstein & Norman, 2017). Additionally, it has been found that unexpected and strongly rewarded stimuli create event boundaries between environmental states in memory, presumably driven by the magnitude of expectancy violations, as encoded by RPEs (Rouhani, Norman, Niv, & Bornstein, 2020).

The fascinating (human) ability to make highly accurate predictions and perform choices based on one- or even zero-shot learning (in situations with very limited or no previous experience) has been proposed to emerge from acquiring, storing, application and manipulation of cognitive maps (Tolman, 1948) of the environment, containing relational knowledge of stimuli, actions and

potential outcomes (Behrens et al., 2018; Tolman, 1948; Wikenheiser & Schoenbaum, 2016). Cognitive maps are related to the place and grid cell architecture in the hippocampal formation and entorhinal cortex and have traditionally been thought of and investigated in the context of spatial navigation (O'Keefe & Nadel, 1978). Only recently this view has been extended by experimental work in non-spatial, abstract and conceptual frameworks (Behrens et al., 2018; Bellmund, Gärdenfors, Moser, & Doeller, 2018; Constantinescu, O'Reilly, & Behrens, 2016; Garvert, Dolan, & Behrens, 2017; Kurth-Nelson, Economides, Dolan, & Dayan, 2016; Schuck, Cai, Wilson, & Niv, 2016; Schuck & Niv, 2019), leading to a more general perspective on cognitive maps as representing predictive relationships between environmental states (Behrens et al., 2018; Momennejad et al., 2017; Niv, 2019; Stachenfeld, Botvinick, & Gershman, 2017). Importantly, theoretical models and empirical findings suggest that not only the hippocampal formation, but also the orbitofrontal cortex might encode a map, or state space of the environment (e.g. Niv, 2019; Schuck et al., 2016). These orbitofrontal cortex state space representations are thought to emerge from interactions with the hippocampus (Schuck & Niv, 2019) and might encode other aspects of abstract relational knowledge than the hippocampus, e.g. by focusing on environmental aspects with immediate biological relevance (Wikenheiser & Schoenbaum, 2016). Indeed, it has been shown that agents use their knowledge of relational structures in the environment to spread value from rewarding events to novel stimuli and guide their behavior (Kurth-Nelson, Barnes, Sejdinovic, Dolan, & Dayan, 2015; Wang, Schoenbaum, & Kahnt, 2020; Wimmer & Shohamy, 2012). Crucially, the observed spread of value in the employed sensory preconditioning paradigms is related to interactions between the hippocampus, ventral striatum (Wimmer & Shohamy, 2012) and orbitofrontal cortex (Wang et al., 2020).

Drawing on a rich literature and research tradition, the dominant – and perhaps also more intuitive – perspective on interactions between associative memory and decision-making is that associative memory processes, specifically storing of value and preference representations, are used to guide choice behavior, consistent with the views of neo-classical economic models of decision-making (Ariely & Norton, 2008; Padoa-Schioppa & Conen, 2017). However, this view neglects the possibility that choice behavior itself might have a modificatory influence on value and preference representations stored in the memory systems. Early studies on choice-induced preference changes found increased subjective value ratings for chosen options after subjects had made a free decision between two choice options with roughly the same a priori subjective value (Brehm, 1956). The general idea of this findings – that past choices bias future preferences and decision-making – has been replicated (conceptually) several times (Hornsby & Love, 2020; Izuma et al., 2010; Luettgau, Tempelmann, Kaiser, & Jocham, 2020; Riefer, Prior, Blair, Pavey, & Love, 2017; Sharot, Martino, & Dolan, 2009). Subsequently, the original finding has been criticized for mathematical inaccuracies, since the usually employed procedure – rating, choice, rating – might indeed not reveal preference changes, but could instead emerge naturally as a result of random error accumulation, regressions to the mean or simply inaccurate reflection of the "true" preference

during the initial rating (Chen & Risen, 2010). However, other experimental procedures overcoming this confound have found support for the reliability of choice-induced preference changes (Sharot, Velasquez, & Dolan, 2010).

1.4 Hypotheses

Given the assumed bidirectional relationship between associative memory and choice behavior, one of the two aims of the present work was to investigate how associative memory processes might serve goal-directed decision-making in novel situation. Specifically, we asked how value might be transferred from conditioned stimuli (CS) to other stimuli that had themselves never been directly followed or paired with reinforcement. We reasoned that one possibility to assign value to novel stimuli would be to chain together sequences of stimuli or actions that have themselves never been observed in direct contiguity with reinforcement, but eventually might lead to reward (or avoid punishment) in the future. Higher-order learning mechanisms like second-order conditioning (SOC) might support instrumental behavior despite absent reinforcement, by enabling spread of value to stimuli or actions that have never been directly paired with reinforcing stimuli (Gewirtz & Davis, 2000; Sharpe, Batchelor, & Schoenbaum, 2017). To the best of our knowledge, it is not known whether human decision-making might be influenced by value that is assigned through transfer learning mechanisms, like second-order conditioning. In Experiment I, we therefore employed a SOC paradigm and a memory-based choice preference test in healthy human participants. We hypothesized that choice behavior would exhibit hallmarks of value transfer to second-order CS and that there would be evidence for reinstatement of neural patterns representing unconditioned stimuli (US) by previously paired first-order CS during SOC. We additionally predicted that this US reinstatement would be paralleled by establishment of a direct associative link between the second-order CS and the US, as evidenced by learning-dependent increases in neural pattern similarity.

Even though most of the previously presented studies on choice-induced preference changes have implicitly assumed that choices might influence the underlying value representations stored in memory, this assumption has never been explicitly tested. This most likely results from the fact that former studies directly presented participants with the choice outcomes to be selected. Hereby, the decision-making process and the underlying changes in memory representations of value are by design obliterated. We reasoned that it would be necessary to disentangle both processes to investigate whether and how decision might bias future decision-making. Therefore, we developed a novel learning and memory-based decision-making paradigm presenting participants with binary decisions between CS that had previously been paired with differently valued US. Decisions were performed during a choice-induced revaluation phase and in a decision probe phase to assess final preferences. Crucially, participants were never presented with the consequences of their choices. This paradigm was then applied in several behavioral and a functional magnetic resonance imaging (fMRI) experimental setups (Experiment II A-E), allowing us to study and (conceptually) replicate the effects of choices on subsequent preferences and on the neural representations of associative

strength between CS and US. We hypothesized that CS chosen during choice-induced revaluation would be preferred over otherwise identical CS (paired with the same US) during the decision probe phase, whereas unchosen CS should be preferred less than equivalent CS. Additionally, we assumed that choices would induce alter the neural representation of associative strength between the chosen and unchosen CS and the respectively paired US.

2. Methods

The following section provides a short overview and description of the key experimental and analysis methods that were used in the present work.

2.1 Second-order conditioning and Pavlovian conditioning paradigms

In the present work, second-order conditioning and Pavlovian conditioning procedures (Pavlov, 1927) were used to establish associative relationships between CS and differently valued US. During the employed Pavlovian conditioning paradigms, previously neutral stimuli became CS by being repeatedly presented in close temporal succession of US, thus establishing expectations or predictions of the respectively paired US. By means of this contiguity, value is conferred from US to CS, so that conditioned responses (CR) towards the CS that closely mirror unconditioned responses (UR) can be observed (Pavlov, 1927). Quantitative accounts of Pavlovian conditioning suggest that CS develop the ability to pre-activate neural populations representing US during formation of an associative relationship between CS and US (Rescorla & Wagner, 1972; Sutton & Barto, 1998; Wagner, 1981).

Second-order conditioning is a higher-order learning mechanism, by which value may be spread to stimuli that have never been directly paired with US, or other stimuli with reinforcing properties. Like in Pavlovian conditioning, during the employed SOC paradigm, a first-order conditioned stimulus (CS_1) is initially linked associatively to a US. In a subsequent conditioning phase, another previously neutral second-order conditioned stimulus (CS_2) is presented in close temporal succession of the CS_1. Theoretically, CS_2 should acquire incentive properties similar to CS_1 and thus develops the ability to elicit a CR.

We administered gustatory reinforcers (quinine solution and orange juice or chocolate milk) in Experiment I or presented images of food items, drawn from an online database (Blechert, Meule, Busch, & Ohla, 2014) in Experiment II A-E as US. As CS_1, we used fractal images (Experiment I). Additionally, as CS_2 (Experiment I) or CS (Experiment II A-E) we used Japanese kanji character images drawn from an online database (Tamaoka, Makioka, Sanders, & Verdonschot, 2017).

The aims of this procedure were threefold: First, we sought to enhance motivational salience and behavioral relevance of the learned associations by employing (depictions of) primary reinforcers. Second, we planned to use representational fMRI methods (see detailed description below) to identify reinstated memory representations of US (Experiment I) and to test choice-related alterations of the associative strength between CS and US (Experiment II E). We reasoned that inferential power to detect neural pattern information and fMRI repetition suppression effects would be maximized if not only value, but also identity features (i.e. color, shape, size, texture) would be distinct for the respectively used US. Third, in Experiment I we trained a classification algorithm on gustatory neural patterns from one session to classify neural patterns related to visual CS recorded during another session. This procedure offered a way to avoid biased classification accuracies due

to same-session and same-modality variance, or due to confounding influence of associations to a learned first-order CS.

2.2 Memory-based decisions

In all our experiments, we used memory-based decision-making scenarios presenting participants with binary choices between two CS. CS had previously been paired with differently valued US, indirectly (Experiment I) or directly (Experiment I and II). Crucially, these decisions were always performed *in extinction*, i.e. participants never experienced the consequences of their choices, or the associated US. We reasoned that due to these constraints, participants would have to make their choices based on remembered CS-US associations and by exploiting the learned underlying value structure. Importantly, this memory-based procedure prevents additional outcome- or US-related learning processes during decision-making and might thus allow for reliable estimation of acquired subjective preferences.

Memory-based decisions were presented on computer screens or projected onto MRI-compatible projection screens (Experiment I and II E). Choice options were presented on the left and right side of a computer screen and participants selected the desired option via button presses on German standard keyboards (QWERTZ) or on fMRI-compatible button boxes (Experiment I and II E) using the index fingers of both right and left hand while lying in a MRI scanner. All decisions had to be performed under time-pressure, allowing participants no more than 1500 ms to respond. We quantified choice behavior by means of choice probability (CP) for a given stimulus, i.e. the proportion of trials during which the respective stimulus was selected divided by the number of times the respective stimulus was presented. Since we employed binary choice situations, it is possible to compare each stimulus' CP against a binary chance level criterion (CP = 0.50) and infer statistically significant deviations of each CP from chance level performance.

2.3 Computational modelling

In order to formally describe the participants' choice behavior and to formalize candidate learning mechanisms that participants could have employed to acquire CS-US associations – and update associative strength during choice-induced revaluation – we used Rescorla-Wagner-like models of reinforcement learning (Rescorla & Wagner, 1972; Sutton & Barto, 1998). Computational models allow to quantitatively fit different parameter value constellations to the data and optimize model parameters so that the likelihood of the observed data given the parameters is maximized. Additionally, since computational models are generative, optimized parameter values can be used to simulate behavior employing artificial agents.

In Experiment II A-C and E, we modelled six different learning mechanism that participants could have used to learn and update CS value. Four of these models incorporated the effect of choosing a CS (and not choosing the alternative CS) on the related CS-US associative strength by introducing "fictive reward prediction errors", the difference between the associatively retrieved

(subjective) US value and the expected stimulus value/associative strength during each choice trial, scaled by a learning parameter α. This discrepancy was then used to update the respective CS' value/associative strength after each choice. Estimated CS values were then be passed through a softmax decision function to generate trial-by-trial choice probabilities for the respectively presented CS.

We used the learned and altered associative strength between each CS and the respective US for fitting models to subsequent preference test behavior and for parameter optimization. We employed a two-stage parameter fitting procedure to minimize the negative log-likelihood estimate and find the model parameter constellation most compatible with the observed choice behavior. This procedure involved 1) a grid search on *n*-dimensional grid in log space (where *n* is the number of free, optimized parameters in the model) with 30 steps in each dimension to find the grid optimum as initial values for function optimization and 2) a constrained non-linear function optimization. Model fits, as represented by the optimized negative log likelihoods, were then compared with the sample-size corrected Akaike Information Criterion, which penalizes the likelihood for model complexity/number of free parameters. The model with the lowest sample-size corrected Akaike Information Criterion at the group level was considered most compatible with the empirical choice behavior observed in our participants. Additionally, we used the individual participants' model parameters of the best-fitting model at the group level to simulate choice behavior for the exact same sequence of choices our participants experienced (10,000 simulations for each participant in each experiment). The resulting simulated choice behavior was compared to the participants' choice behavior to identify whether the qualitative features of the empirical choice data (i.e. increased choice probability for previously chosen CS) could also be captured by the candidate computational models.

2.4 fMRI repetition suppression

fMRI repetition suppression (fMRI-RS) is a fMRI technique that can be used to measure neural representations of stimuli or cognitive variables (Barron, Garvert, & Behrens, 2016; Grill-Spector & Malach, 2001; Larsson, Solomon, & Kohn, 2016). fMRI-RS makes use of neurons' tendency to respond with reduced firing if stimuli or environmental states are repeatedly presented in rapid succession (Barron et al., 2016; Grill-Spector & Malach, 2001; Larsson et al., 2016; Summerfeld, Wyart, Johnen, & de Gardelle, 2011), which can be indirectly measured at the neural population-level by a reduction of the blood-oxygenation level-dependent (BOLD) signal. Crucially, fMRI-RS might be able to overcome a limiting factor in contemporary mass-univariate fMRI, where inference about neural activation is constrained by averaged BOLD signal within the acquired voxels. Typical isotropic voxel dimensions range between 2 and 4 mm, which averages activation states of approximately 10^5 neurons per voxels and hence impedes dissociation of activation levels in different neuronal populations within a voxel. The resolution of the acquired voxels thus constitutes the spatial resolution of activation patterns in the functional image. Contrarily, repeated presentation of the same stimulus should in principle repeatedly activate the same – or at least a large proportion of –

the same neural population that is responsive to the respective stimulus of interest. Thereby, changes in activation patterns driven by specific neuronal populations could be made observable even at below voxel size spatial resolutions (Barron et al., 2016).

Despite its wide range of implications and the explanatory potential for investigating neural representations, the exact neural mechanisms underlying repetition suppression and whether it reflects expectation violation or adaptation (Larsson et al., 2016) remain elusive. Biologically plausible fMRI forward modelling suggests that repetition suppression effects might arise from scaling of the response amplitudes for presented stimuli, locally affecting tuning-curves close to the adapting stimulus (Alink, Abdulrahman, & Henson, 2018). However, empirical work using optogenetic manipulations in monkey inferotemporal cortex provides evidence against the hypothesis of intrinsic firing rate changes and hints on a transsynaptic origin of repetition suppression effects (Fabbrini et al., 2019).

Recent studies have extended fMRI-RS's range of potential applications to contexts going beyond the initially postulated signal reductions reflecting repetition of stimuli (Barron et al., 2013; Garvert et al., 2017; Klein-Flügge et al., 2013). It has been suggested that repetition of shared features or the strength of associative links between stimuli might result in similar reductions of BOLD responses as repetition of stimuli, a phenomenon known as cross-stimulus suppression (Barron et al., 2016). Reinforcement learning accounts have proposed that Pavlovian conditioning procedures establish a predictive, associative relationship between the CS and the US, and that CS should be capable to pre-activate neural populations representing US after an association has been formed (Wagner, 1981). Building on these assumptions, we predicted that presentation of a US preceded by a previously associated CS should elicit reduced fMRI signal, i.e. increased repetition suppression, compared to presentation of a US preceded by a CS that was not associatively linked. In Experiment II E, we hence used fMRI-RS as a measure of the associative strength of previously learned CS-US associations, and how the associative strength between CS and US are altered by previously choosing or not choosing a CS (i.e. choice-induced revaluation).

In learning contexts, it is often challenging to design experiments and fMRI-RS sessions so that confounding effects, such as extinction, new learning or unlearning related to re-exposure to stimuli can be avoided. Since applying fMRI-RS requires many repeated presentations of stimuli, it is by design not always possible to harmonize the experimental (learning) paradigm with the affordances of the fMRI-RS session that will be used to measure neural representations. Even though studies have demonstrated that fMRI-RS may be applied to investigate trial-to-trial changes in neural representations of stimulus-outcome associations (Boorman, Rajendran, O'Reilly, & Behrens, 2016), experimental designs using fMRI-RS are often constrained to relatively stable representations. Additionally, canonical fMRI-RS makes use of mass-univariate statistical inference, which has been criticized for giving rise to inflated false-positive rates (Eklund, Nichols, & Knutsson, 2016) and might indeed not be able to adequately capture multivariate dependencies and spatial

correlations in the BOLD signal. We therefore also employed multivariate, information-based neuroimaging methods to identify and characterize neural representations distributed across voxels.

2.5 Cross-session, cross-modality searchlight multivariate pattern analysis

Another representational fMRI technique employed in the current work is multivariate pattern analysis (MVPA, Haxby, 2012; Haxby et al., 2001; Haynes & Rees, 2006), a pattern classification and machine learning based approach to the analysis of neuroimaging data. An important difference between MVPA and mass-univariate neuroimaging analysis approaches is that the direction of statistical inference is essentially reversed. Whereas mass-univariate approaches ask the question, where in the brain stimuli, conditions or cognitive variables (independent variable) are *encoded* in clusters of voxels (dependent variable), MVPA uses a *decoding* approach to make classification decisions about different conditions or trials (dependent variable) related to different multivariate activation patterns (independent variables, i.e. voxels). The technique has been proposed to be more sensitive to small deviations in activation patterns across different conditions or trials, since it uses the complete distribution of voxel activation levels to make predictions about the condition or class. This is an important feature to increase sensitivity, since even voxels with low activation levels might contribute meaningful information and could thus be diagnostic of condition differences. Such information conveyed by low activation levels of voxels might be lost in mass-univariate neuroimaging techniques, that traditionally "discard" activation below certain thresholds (Haxby, 2012; Haynes & Rees, 2006).

Most fMRI MVPA approaches employ cross-validation schemes, in which the complete data set is partitioned into independent training and test data sets (chunks). The training data set – often in combination with a cost function – is used to teach the regular features of the training data to a multivariate classification algorithm (Bishop, 2006; Hanke et al., 2009; Haxby, 2012). Then, the trained classification algorithm uses the weights that optimally distinguish between activation patterns from different conditions to make a prediction about the class or condition labels of the remaining test data set, thus generalizing to unseen data (Bishop, 2006). This procedure is repeated for $n - 1$ times or folds (where n is the number of data chunks) until each fold has been assigned to the test data set exactly one time. The resulting accuracy values for each test data set are used to score the overall out-of-sample predictive accuracy of the classification algorithm (Bishop, 2006). This out-of-sample predictive accuracy may then be tested against a chance-level classification criterion ($1/k$, where k is the number of classes/conditions in each classification decision, e.g. ½ = 0.50, for 2 conditions).

In Experiment I, we employed MVPA techniques to obtain estimates for neural pattern reinstatement of US by CS. Our main assumption was again that after Pavlovian conditioning, CS should be able to pre-activate neural ensembles representing US. We reasoned that it would be possible to use a multivariate classification algorithm trained on neural patterns representing the US to correctly predict class labels of previously paired CS neural patterns during second-order

conditioning. To this end, we employed a cross-session (Stokes, Thompson, Cusack, & Duncan, 2009) and cross-modality (gustatory to visual) multivariate pattern classification. During a first test session, we administered gustatory US (quinine and orange juice) to participants in the fMRI scanner and trained a support vector machine (default C = 1) algorithm on the neural pattern related to the US. We then asked whether the trained classifier could correctly predict the class of new, unseen data of the visually presented CS_1 in a second test session, after these CS_1 had been associatively paired with gustatory US in a Pavlovian conditioning phase. This procedure was based on the reasoning that for unbiased classification, the gustatory US neural patterns used as a training data set should not yet be associatively linked to a CS and that the possibility to correctly predict visual stimulus classes from gustatory stimulus information could not readily be explained by adaptation effects that might occur in same-modality classification.

To obtain predictions in an a priori spatially unconstrained fashion, we employed a searchlight classification approach (Kriegeskorte, Goebel, & Bandettini, 2006) that allowed to iterate through whole-brain patterns of activation using 3-mm searchlight spheres. Searchlight sphere classification accuracies were then mapped to the center voxel of each sphere, leading to a whole-brain map of classification accuracies. This procedure was then repeated 100 times per participant using shuffled class labels in the training data set to obtain chance level whole-brain accuracy maps. To correct for multiple comparisons we employed group-level random-effect cluster-statistics (Stelzer, Chen, & Turner, 2013). Here, for each of 50,000 iterations, one chance level per participant was randomly drawn and a group level z-statistic map was calculated. The resulting empirical null distribution (50,000 samples) was then compared to the clusters in the “real” whole-brain classification accuracy map.

Importantly, it would not have been possible to obtain comparable estimates of neural pattern reinstatement using other representational neuroimaging approaches, like fMRI-RS. In order to use fMRI-RS as an index of US neural pattern reinstatement during SOC, it would have been necessary to introduce trials presenting CS_1 followed by the US. However, this procedure would have invalidated the fundamental assumption of SOC: Second-order learning of CS_2 value should occur despite the fact that US are never presented during the second-order learning phase.

2.6 Neural pattern similarity analyses

Similar to fMRI-RS, neural pattern similarity analyses, a variant of representational similarity analysis (RSA, Kriegeskorte, Mur, & Bandettini, 2008) can be used to quantify the similarity (or dissimilarity) between BOLD activation patterns in fMRI data. Neural pattern similarity is mostly calculated in a distributional, multivariate fashion. This presents potential gains in sensitivity over the estimated representations resulting from mass-univariate fMRI-RS. Since the degree of similarity between brain activation patterns for different conditions, different trials etc. is calculated in unit-free representational space, the method is not limited to fMRI, but allows multi-modal and even cross-species comparisons of activation/activity similarity.

In Experiment I, we performed a template-based variant of neural pattern similarity analyses (Wimber et al., 2015), using a least-squares separate approach (Mumford, Turner, Ashby, & Poldrack, 2012) which aimed at deconvolving single trial estimates of CS_1 and CS_2 neural activation. This was done to compare early and late phase similarity between gustatory US neural patterns and visual CS_2 neural patterns in the second-order conditioning phase of Experiment I, as a measure of a developing associative link between CS_2 and the respective US. This approach also allowed us to estimate the associative relationship between CS_1 and the respective US, which was then used as a control condition to rule out non-specific changes in pattern similarity between CS_2/CS_1 (that were always presented in succession) and the respective US. We explicitly decided for this neural pattern similarity approach, since calculation of neural pattern similarity affords lower computational and storing demands than computationally intensive random-effect cluster-statistics that are necessary for multiple comparison corrections in MVPA. An alternative approach, classification of single trial CS_1/CS_2 neural patterns with a classifier trained on gustatory US neural patterns – similar to the cross-session, cross-modality classification approach described earlier – using adequate correction for multiple comparisons would have easily exceeded our available storing capacities (150 (trials) real classifications, and 150 x 100 = 15,000 classifications using shuffled class labels, total: 15,150 whole-brain maps per subject)

In Experiment II E, we used an ordinary least-squares approach to estimate the changes of neural pattern similarity between equivalent CS during trials in which they were followed by US that had not been paired with the respective CS during Pavlovian conditioning. This was based on our assumption that Pavlovian conditioning should establish predictive, associative CS-US and that CS should pre-activate neural populations representing US. Here, we calculated the PRE to POST choice-induced revaluation changes of neural pattern similarity. In both experiments, we conducted neural pattern similarity analyses in theoretically derived and anatomically defined regions-of-interest (ROI), spatially constraining the multivariate neural patterns under investigation, e.g. to the amygdala (Experiment I) or hippocampus and lateral orbitofrontal cortex (lOFC) (Experiment II E).

3. Results

3.1 Experiment I: Reinstatement of cortical pattern representing US during second-order conditioning

The following section is based on our submitted manuscript (see attachments), which is currently under review at *Nature Communications*:

Luettgau L., Porcu E., Tempelmann C., & Jocham G. Reinstatement of cortical outcome representations during higher-order learning. Preprint at *bioRxiv* (2020), doi: 10.1101/2020.05.28.121558

Building on the “direct link” hypothesis of second-order conditioning, we sought to investigate whether CS_2 are directly paired with a neural representation of the US, or more specifically, with the motivational state associated with the US (Gewirtz & Davis, 2000; Parkes & Westbrook, 2011). This direct link would constitute a neural mechanism underlying value transfer in SOC. In Experiment I, we hypothesized that presentation of a CS_1 should reinstate the cortical pattern representing the US that it had been paired with during first-order conditioning in the lOFC. The lOFC has been implicated in gustatory processing, specifically in higher-order motivational (Rolls, 2000, 2006; Small et al., 1999) aspects of gustatory perception and memory for taste (Kobayashi et al., 2004). This reinstatement could allow to associatively link CS_2 and US neural representations, using plasticity mechanisms in the amygdala and hippocampus.

In two separate samples (behavioral: $N = 20$, fMRI: $N = 29$), participants first rated fractal and kanji images serving as visual CS_1 and CS_2, respectively. Subsequently, subjects underwent Pavlovian first-order conditioning to establish associations between visual CS_1 (CS_1^+ and CS_1^-) and appetitive or aversive gustatory US (US^+ and US^-). In a second-order conditioning phase, participants learned to associate visual CS_2 (CS_2^+ and CS_2^-) with previously learned CS_1. Additionally, participants were presented with an association between two novel, neutral CS (CS_2^n and CS_1^n). During the SOC phase, fMRI data was recorded (in the fMRI sample). Finally, participants made choices between pairs of differently valued CS_1 or CS_2, respectively, to assess subjective preferences. These choices were interleaved with lure decisions between CS_1^n or CS_2^n and novel fractal or kanji images, which had only been presented during a pre-task rating. Importantly, participants were instructed to perform simple attentional control tasks, but were not informed about the underlying associative structure of the tasks, aiming at leaving them unaware of intended associative learning. We verified this assumption using post-experimental questionnaires. Indeed, participants did not report explicit knowledge of the (higher-order) associations between CS_2 and US. Participants in the fMRI sample additionally received gustatory US (quinine and orange juice) in a separate session performed one day before second-order conditioning. The recorded fMRI data from this session were used to train a multivariate classification algorithm.

Despite absent explicit knowledge of the associations, participants across both samples were more likely to choose both appetitive first- and second order stimuli, CS_1^+ and CS_2^+, over the aversive CS_1^- and CS_2^-. Importantly, both choice probabilities of CS_1^n and CS_2^n against novel fractal or kanji images did not differ from chance level. The observed high preference of both appetitively paired CS (or low preference for both aversively paired CS) suggests that both CS_1 and CS_2 acquired value during second-order conditioning. The absence of such clear preference patterns for both novel, neutral CS (CS_2^n and CS_1^n) suggests that the observed preferences cannot readily be explained by simple mere exposure or novelty-related effects.

A multivariate cross-session, cross-modality searchlight classification analysis provided evidence for reinstatement of US patterns by previously paired visual CS_1 in the left lOFC. In this region, the classification algorithm trained on US neural pattern information was able to make correct

predictions about CS identity with above chance level accuracy. We next asked whether the reinstated US neural pattern would be projected to other multimodal brain areas like amygdala and hippocampus convergence and linking of "online" CS_2 and US representations. In a psychophysiological interaction analysis, the covariation between BOLD signals in the left lOFC and a cluster in the hippocampus, amygdala and medial temporal lobe was found to be higher in trials presenting CS^- and CS^+ than during CS^n trials.

We subsequently investigated the putative direct associative link between CS_2 and the respective US. If the amygdala indeed uses the reinstated US pattern to form an association with CS_2, one would predict similarity between CS_2 and US neural pattern. Since the association between CS_2 and US is acquired with repeated presentations/trials, this similarity should show an increase over the course of second-order conditioning (e.g. from early to late stages of SOC). Indeed, averaged over the whole course of SOC, there was significant pattern similarity between CS_2^- and US^- in a bilateral amygdala ROI, however, we found no evidence for pattern similarity between CS_2^+ and US^+. This finding might be due to differential motivational salience of the employed gustatory US. Presumably, the administered bitter and aversively tasting quinine solution as US^- might have elicited innate avoidance responses (Yiannakas & Rosenblum, 2017), producing a salient gustatory sensation. However, we used orange juice as a US^+, a compound stimulus consisting of many different taste facets, which likely was familiar to (and differently valued by) most of our participants. This might have introduced high variability in taste responses and could have reduced the motivational relevance and salience of US^+. Since highly palatable and high caloric food and drinks are available abundantly in German and other Western societies, orange juice might additionally not have been considered particularly salient by our participants.

Consistent with the predicted formation of an associative link between CS_2 and US over the course of second-order conditioning, we observed an increase of neural pattern similarity between CS_2^- and US^- in the bilateral amygdala from early to late trials of SOC. This finding indicates the development of an associative link between CS_2 and US across trials. Importantly, there was no evidence for such change in similarity between first-order CS_1^+ and US^+ or between first-order CS_1^- and US^-. The latter two findings are consistent with the assumption that during SOC, the association between CS_1 and the US should remain unaltered (or become weaker), since CS_1 is presented in extinction, i.e. not followed by the previously paired US. Again, we also did not find evidence for neural pattern similarity increases between CS_2^+ and US^+ from early to late trials of SOC.

In sum, our findings support the "direct link" hypothesis of second-order conditioning (Barnet, Arnold, & Miller, 1991; Gewirtz & Davis, 2000; Rizley & Rescorla, 1972), proposing that the motivational state or value of outcomes might directly be linked to second-order CS by exploiting higher-order associative structures. The empirical results suggest a neural mechanism for propagation of outcome value to stimuli that had never been directly coupled with reinforcement. This mechanism might be used to enable proper credit assignment to reward-predictive stimuli in

real-world learning scenarios which are typically characterized by infrequent direct encounters with reinforcers.

3.2 Experiment II

The following section is based on our manuscript published in *Nature Communications* (see attachments):

Luettgau L., Tempelmann C., Kaiser L. F., & Jocham G. (2020). Decisions bias future choices by modifying hippocampal associative memories. *Nature Communications*. doi: 10.1038/s41467-020-17192-7. Preprint at *bioRxiv* (2019), doi: 10.1101/802462

3.2.1 Experiment II A: Choice-induced preference changes in high value and intermediate value conditioned stimuli

In Experiment I we had demonstrated how associative memory processes could support value transfer in second-order conditioning to influence choice behavior in humans. In the following five experiments (Experiment II A-E) we focused on the opposite direction of influence: We asked whether choice behavior itself could give rise to preference changes (choice-induced preference changes) and how this choice-induced preference changes might be related to altered representations of CS-US associations held in associative memory.

A recent theoretical framework has proposed nonmonotonic plasticity in associative memory: inactive memories remain unchanged, whereas moderately activated memories are weakened, and strongly activated memories get strengthened (Ritvo et al., 2019). Building on this framework, we hypothesized that both chosen and unchosen CS should moderately activate neural ensembles representing the respectively associated US. However, we further assumed that choosing a CS would elicit additional activation of the associated outcome, consistent with studies proposing heightened attentional weighting of chosen options, as reflected in higher learning rates (Klein et al., 2017; Palminteri, Khamassi, Joffily, & Coricelli, 2015). Contrarily, we assumed that the unchosen CS would retain a moderate activation level of the associated US, resulting in a weakened CS-US association.

In a newly developed paradigm, healthy human volunteers (N = 40 in final analyses) first established associations between six neutrally rated CS and three differently valued US (images of food items) in a Pavlovian conditioning phase. The three US were selected based on ratings of the subjective value prior to conditioning, so that a low value US^{-}, an intermediate value US^{0}, and a high value US^{+} were each associated with two CS, resulting in pairs of differently valued CS ($CS^{-}_{A/B}$, $CS^{0}_{A/B}$, $CS^{+}_{A/B}$). In Experiment II A, we presented participants with binary decisions between an intermediate value CS^{0}_{A} and a high value CS^{+}_{A} during a subsequent choice-induced revaluation phase. These choices were interleaved with lure decisions between novel kanji images, which had

only been presented during a pre-task rating. In a final decision probe phase, participants made binary choices between all possible combinations of CS to assess final preferences for all CS. According to our hypothesis, choice of CS^{+}_{A} should strengthen the association between CS^{+}_{A} and US^{+}, leading to increased preference for CS^{+}_{A} (compared to CS^{+}_{B}) during the decision probe phase. Contrarily, not choosing CS^{0}_{A} should weaken the association between CS^{0}_{A} and US^{0}, resulting in decreased preference for CS^{0}_{A} (compared to CS^{0}_{B}) in the decision probe phase.

We indeed observed an increased overall choice probability for CS^{+}_{A} (i.e. probability of choosing this stimulus, regardless of which other CS was the alternative option) compared to the overall choice probability for CS^{+}_{B}. More specifically, in trials directly contrasting CS^{+}_{A} and CS^{+}_{B}, participants favored CS^{+}_{A}. Conversely, there was a decreased overall choice probability for CS^{0}_{A} compared to the overall choice probability for CS^{0}_{B}. In trials directly comparing CS^{0}_{A} and CS^{0}_{B}, participants chose CS^{0}_{A} less likely.

Computational modelling indicated that among a set of candidate reinforcement learning models, an algorithm that differentially updated both chosen and unchosen CS associative strength to its respective US (or value updating) using “fictive reward prediction errors”, explained the observed decision probe phase behavior of the participants best. “Fictive reward prediction errors” were calculated as the difference between the associatively retrieved (subjective) US value and the expected stimulus value/associative strength during each choice trial. This discrepancy was then used to update the respective CS’ value/associative strength after each choice. Empirical choice patterns, i.e. increased choice probability for the previously chosen CS^{+}_{A} relative to CS^{+}_{B} and decreased choice for the previously unchosen CS^{0}_{A} relative to CS^{0}_{B}, could successfully be recovered in model simulations using the best-fitting model parameters.

These results suggest that choices can increase preferences for previously chosen options, whereas not choosing an option might diminish its subjective desirability.

3.2.2 Experiment II B: Choice-induced preference changes in intermediate value and low value conditioned stimuli

Since Experiment II A provided evidence for decreased preference of a previously non-chosen CS^{0}_{A}, in Experiment II B we asked whether choices of CS^{0}_{A} during choice-induced revaluation could induce the exact opposite – preference increases for CS^{0}_{A} relative to CS^{0}_{B}. Importantly, this experiment was designed to (conceptually) replicate the choice-induced preference changes found in Experiment II A and to investigate whether the observed effects are independent of CS value. We again hypothesized that choice of CS^{0}_{A} should strengthen the association between CS^{0}_{A} and US^{0}, leading to increased preference for CS^{0}_{A} (compared to CS^{0}_{B}) during the decision probe phase. Contrarily, not choosing CS^{-}_{A} should weaken the association between CS^{-}_{A} and US^{-}, resulting in decreased preference for CS^{-}_{A} (compared to CS^{-}_{B}) in the decision probe phase.

Like in Experiment II A, an independent sample of healthy human volunteers (N = 40 in final analyses) first established associations between six neutrally rated CS and three differently valued

US during a Pavlovian conditioning phase. In the subsequent choice-induced revaluation phase, we presented participants with binary decisions between a low value CS^{-}_{A} and an intermediate value CS^{0}_{A}. In the final decision probe phase, participants again made binary choices between all possible combinations of CS.

Replicating the choice-induced preference change effects observed in Experiment II A, overall choice probability for CS^{0}_{A} was increased compared to the overall choice probability for CS^{0}_{B}. In trials directly contrasting CS^{0}_{A} and CS^{0}_{B}, participants preferred CS^{0}_{A}. Additionally, we found decreased overall choice probability for CS^{-}_{A} compared to the overall choice probability for CS^{-}_{B}. More specifically, in trials directly comparing CS^{-}_{A} and CS^{-}_{B}, there was descriptively reduced preference for CS^{-}_{A}. Like in Experiment II A, a computational model differentially updating both chosen and unchosen CS associative strength to its respective US using "fictive reward prediction errors" elicited by revaluation phase choices best captured the observed decision probe phase behavior. Moreover, simulations using the optimized parameter values could recover the empirical choice pattern.

These results, in conjunction with the results of Experiment II A, provide evidence that choices enhance preferences for previously chosen options, whereas not choosing an option might reduce its subjective desirability, independent of choice option value.

3.2.3 Experiment II C: Choice-induced preference changes can be cancelled out in intermediate value conditioned stimuli

The findings of Experiment II A and B suggest that choices and non-choices have opposite effects: While choosing increases preferences, not choosing decreases subjective desirability of choice options. In Experiment II C, we therefore investigated whether there was empirical support for the claim that choice-induced preference increases and decreases for CS^{0}_{A} should cancel each other out. Like in Experiment II A and B, an independent sample of healthy human volunteers (N = 44 in final analyses) first established associations between six neutrally rated CS and three differently valued US during a Pavlovian conditioning phase. In Experiment II C, we presented participants with an equal amount of binary decisions between an intermediate value CS^{0}_{A} and a low value CS^{-}_{A} as between the same CS^{0}_{A} and a high value CS^{+}_{A} during the subsequent choice-induced revaluation phase. In the final decision probe phase, participants again made binary choices between all possible combinations of CS. Since the equal amount of choosing and not choosing CS^{0}_{A} should even out putative preference changes, we predicted that choice probabilities for CS^{0}_{A} and CS^{0}_{B} would not differ during the decision probe phase.

Indeed, there was no evidence for differences between the overall choice probability for CS^{0}_{A} and the overall choice probability for CS^{0}_{B}. In trials directly contrasting CS^{0}_{A} and CS^{0}_{B}, participants did not show preference for either CS^{0}_{A} or CS^{0}_{B}. As before, the observed decision probe phase behavior was best explained by a computational model differentially updating both chosen and unchosen CS associative strength to its respective US using "fictive reward prediction errors" elicited

by revaluation phase choices. In simulations using the optimized parameter values, we found that it was possible to recover the empirical choice pattern.

Experiment II C provides evidence that previously found choice-induced preference increases for chosen and preference reductions for unchosen options (Experiment II A and B) work in opposite directions and that these opposing effects cancel each other out.

3.2.4 Experiment II D: Choice-induced preference changes are unlikely to be explained by choice heuristics

According to our hypothesis, choice-induced preference increases for previously chosen CS result from choice-induced strengthening of the association between the CS and its respectively associated US. Contrarily, not choosing a CS should weaken the association between the CS and its respectively associated US, resulting in decreased preference for the unchosen CS (associative hypothesis). However, the observed empirical choice behavior could alternatively be explained by acquired choice heuristics, or simple stimulus-response tendencies. Participants might have formed associations between each of the two CS presented during choice-induced revaluation and stimulus-response tendencies, i.e. choice heuristics. Specifically, chosen CS might have acquired *go tags* ("choose this stimulus"), whereas *no-go tags* could have been assigned to unchosen CS ("do not choose this stimulus"). According to the choice heuristic account, the observed increases and decreases of choice probabilities for chosen and unchosen CS, respectively, could thus be attributed to the expression of such acquired heuristics.

In Experiment II D, we therefore addressed this possibility by orthogonalizing the contributions of CS-US associative strength and choice heuristics. Participants (N = 40 in final analyses) first established associations between four neutrally rated CS and two differently valued US during a Pavlovian conditioning phase. Crucially, within each of the differently valued CS pairs, one CS had been associated strongly (80% association, CS^0_{80} or CS^+_{80}) to the respective US, whereas the other CS was weakly associated (20% association, CS^0_{20} or CS^+_{20}). In the subsequent choice-induced revaluation phase, participants were presented with binary choices between differently valued, but equally strongly associated CS: CS^0_{80} vs. CS^+_{80} and CS^0_{20} vs. CS^+_{20}. We reasoned that participants would be more likely to choose both CS^+_{80} and CS^+_{20} and would thus assign *go tags* to both chosen CS. Contrarily, non-choices of both CS^0_{80} and CS^0_{20} would assign *no-go tags* to both unchosen CS. During the decision probe phase, participants made binary choices between CS from the same value category, that exhibited a comparable choice history (i.e. frequency of being chosen or unchosen) but differed in associative strength with the US (CS^+_{80} vs. CS^+_{20} and CS^0_{80} vs. CS^0_{20}). We hypothesized that if participants made choices based on the learned associative strength and choice-induced strengthening/weakening of associations, more strongly associated CS (CS^+_{80} and CS^0_{80}) should be preferred. However, if choice behavior during the decision probe phase was exclusively driven by learned choice heuristics (*go* or *no-go tags*), there should be no clear

preference for either CS at choice, i.e. there should be no evidence for choice probabilities different from chance level (CP = 0.50).

Importantly, there was no evidence for differences between choice probabilities for both high value options, CS^{+}_{80} (vs. CS^{0}_{80}) and CS^{+}_{20} (vs. CS^{0}_{20}), during choice-induced revaluation, indicating that *go* or *no-go tags* had been assigned equivalently across both CS pairs. In the decision probe phase, participants were more likely to select the previously chosen and strongly associated CS^{+}_{80} against CS^{+}_{20}. However, only a descriptive increase of choice probability for CS^{0}_{80} over CS^{0}_{20} could be observed (*Median* = 0.60). These results strongly suggest that decision probe behavior was not exclusively driven by learned choice heuristics, although a potential role of *no-go tags* assigned to unchosen stimuli as an alternative explanation for the observed choice behavior cannot be ruled out based on Experiment II D.

3.2.5 Experiment II E: Choice-related modifications of neural CS-US representations as a mechanism for choice-induced preference changes

After having established and (conceptually) replicated behavioral choice-induced revaluation effects, and partially ruled out alternative explanations, we next asked whether decisions would change neural representations of CS-US associations and their associative strength.

In Experiment II E, healthy human volunteers (*N* = 42 in final analyses) first established associations between six neutrally rated CS and three differently valued US in a Pavlovian conditioning phase, resulting in pairs of differently valued CS ($CS^{-}_{A/B}$, $CS^{0}_{A/B}$, $CS^{+}_{A/B}$), as in Experiments II A-C. Next, participants underwent one run of fMRI repetition suppression (PRE), repeatedly presenting them with each CS followed by all possible US. In a subsequent choice-induced revaluation phase, we presented participants with binary decisions between an intermediate value CS^{0}_{A} and a high value CS^{+}_{A}. After choice-induced revaluation, participants underwent another run of fMRI repetition suppression (POST), again presenting them with all possible combinations of CS-US transitions. In a final decision probe phase, participants made binary choices between all possible combinations of CS to assess final preferences for all CS.

We predicted that presentation of a US preceded by a previously associated CS should elicit reduced fMRI signal, i.e. increased fMRI-RS, compared to presentation of a US preceded by a CS that was not associatively linked. According to our hypothesis inspired by nonmonotonic plasticity in associative memory (Ritvo et al., 2019), the fMRI-RS signal during the POST run should be increased for the previously chosen CS^{+}_{A} followed by US^{+} (relative to CS^{+}_{B} followed by US^{+}) and the fMRI-RS signal would be decreased for previously unchosen CS^{0}_{A} followed by US^{0} (relative to CS^{0}_{B} followed by US^{0}). However, such differences should be absent during the PRE fMRI-RS run. We expected to find choice-induced changes in representations of CS-US associations in the hippocampus and lOFC, brain regions that have consistently be involved in memory processing, storing and updating of stimulus-outcome associations (Boorman et al., 2016; Klein-Flügge et al., 2013).

Indeed there was evidence for both an increase of fMRI-RS for the previously chosen CS^{+}_{A} followed by US^{+} (compared with CS^{+}_{B} followed by US^{+}) and reduced fMRI-RS for the previously unchosen CS^{0}_{A} followed by US^{0} (compared with CS^{0}_{B} followed by US^{0}) during the POST fMRI-RS run in the left hippocampus (whole-brain corrected) and in the right lOFC (small-volume corrected for an independent functional mask from Jocham et al., 2016). The alterations in the fMRI-RS signal were not present during the PRE fMRI-RS run. Also, there was evidence for a PRE-POST decrease of RS signal for CS^{0}_{A} followed by US^{0} (compared with CS^{0}_{B} followed by US^{0}) and a qualitative increase of RS signal for CS^{+}_{A} followed by US^{+} (compared with CS^{+}_{B} followed by US^{+}). Importantly, we found no evidence for PRE-POST differences and also no significant effects during the PRE or the POST fMRI-RS run for the two unpresented control stimuli (CS^{-}_{A} and CS^{-}_{A}, followed by US^{-}).

These mass-univariate fMRI-RS based findings were further supported by converging evidence from multivariate neural pattern similarity analyses in the left hippocampus and right lOFC. For the PRE and the POST fMRI-RS runs, we separately calculated the pattern similarity between equivalent CS followed by US they had not been paired with during Pavlovian conditioning. This was based on our assumption that Pavlovian conditioning should establish predictive CS-US associations and that each CS should pre-activate neural populations representing US after learning. This pre-activation should also be observable in trials where the CS was not followed by the previously paired US. Our multivariate neural pattern similarity analyses thus focused only on trials during which the CS was followed by the two US it had not been associatively linked with during Pavlovian conditioning. We found that for both pairs of high value CS^{+} and intermediate value CS^{0} neural pattern similarity decreased from PRE to POST in the hippocampus, indicating that the pre-activation of neural populations representing the US might have changed due to decisions made in the choice-induced revaluation phase. However, in the pair of the low value CS^{-} similarity did not change from PRE to POST. Qualitatively similar results were obtained in the right lOFC. Nevertheless, these results should be interpreted with caution, since changes in neural pattern similarity might equally likely arise from both strengthening or weakening of CS-US associations of the chosen and unchosen CS. Unlike the mass-univariate fMRI-RS approach used before, the multivariate neural pattern similarity method employed here does not allow to tease apart these possibilities and thus thwarts the inference about the direction of changes.

We next asked whether the observed choice-induced changes in representations of CS-US associations were related to behavioral preference changes. Unlike in Experiment II A, participants did not select the previously unchosen intermediate CS^{0}_{A} less likely than the equivalent control stimulus CS^{0}_{B}. We reason that this might be related to the fact that participants in Experiment II E had to be re-exposed to the CS-US associations during the POST fMRI-RS run to measure CS-US associative strength. This restudying might have allowed them to relearn the putatively weakened association between CS^{0}_{A} and US^{0}, as consistently observed in studies reporting reversal of retrieval-induced forgetting effects following restudy of memorized materials (Hulbert & Norman, 2015; Storm et al., 2008). However, participants once again presented robust choice-induced

preference increases for CS^{+}_{A} compared with CS^{+}_{B}. We thus focused on the difference between overall choice probabilities of CS^{+}_{A} and CS^{+}_{B} for brain-behavioral correlations. The fMRI-RS signal difference between the previously chosen CS^{+}_{A} followed by US^{+} compared with CS^{+}_{B} followed by US^{+} was positively correlated with the difference between the overall choice probability of CS^{+}_{A} and the overall choice probability of CS^{+}_{B}. In other words, the more the hippocampal representation of the CS-US association between CS^{+}_{A} and US^{+} (relative to CS^{+}_{B} and US^{+}) had been strengthened during the choice-induced revaluation phase, the more likely participants were to select CS^{+}_{A} compared to CS^{+}_{B}.

These findings, also in conjunction with Experiment II A-D, suggest that choices change the neural representations of CS-US associations and the associative strength of stimulus-outcome associations. These choice-induced changes in neural representations of CS-US associations correlated with future decisions, suggesting mechanistic involvement in choice-induced preference changes.

4. General Discussion

The present work aimed at investigating interactions between decision-making and associative memory processes. These two cognitive domains have been studied in isolation quite extensively for many decades, however, only recently it has been proposed that both memory and decision-making processes might influence and bias each other (Shohamy & Daw, 2015). The traditional, neo-classical economic view on decision-making would solely predict a guiding influence of associative memory processes, specifically value representations in memory, on choice behavior (Ariely & Norton, 2008; Padoa-Schioppa & Conen, 2017). Contrary to this unidirectional view, we here provide evidence for both influences of value representations stored in associative memory on choice behavior and choice-induced modifications of associative memories, supporting the idea of bidirectional interactions.

Specifically, there was behavioral evidence for value transfer in choice patterns following second-order conditioning, a memory-dependent higher-order learning paradigm. Importantly, this choice bias was present even though participants were unaware of the underlying (higher-order) associative learning structure of the experiment, suggesting that humans implicitly acquire subjective value through higher-order learning mechanisms. This finding extends previous studies promoting the acquisition of explicit reward-predictive associations between stimuli (Jara, Vila, & Maldonado, 2006; Wang et al., 2020; Wimmer & Shohamy, 2012). To the best of our knowledge, this is the first demonstration that humans exhibit choice biases arising from second-order conditioning, using similar paradigms and procedures as previously employed in animal studies. The observed behavioral effect was paralleled by neuroimaging findings suggesting that neural patterns representing the administered gustatory outcomes (and presumably the motivational state or value conveyed by these outcomes) are reinstated in the left lOFC by previously paired first-order CS. The left lOFC has been implicated in higher-order processing of gustatory information, especially in motivational (Rolls, 2000, 2006; Small et al., 1999) aspects of gustatory perception and memory for taste (Kobayashi et al., 2004). Additionally, there was evidence for the formation of a direct association between the reinstated motivational state or value conveyed by outcomes and second-order CS in the amygdala, a multimodal brain area commonly implicated in associative plasticity. These results suggest that value might be propagated to stimuli that had never been directly coupled with reinforcement by exploiting acquired higher-order associative structures stored in memory.

Moreover, in a series of experiments employing a newly developed learning and decision-making paradigm, there was converging evidence for choice-induced preference changes: While previously chosen options were selected more frequently, unchosen options showed diminished preferences – compared to otherwise equivalent options. These choice effects seem to counteract and were found to cancel each other out when the same choice option was chosen and unchosen equally often during choice-induced revaluation. Additional experimental results indicate that the observed choice-induced preference changes are unlikely exclusively driven by acquired choice heuristics (*go* and *no-go tags*) for chosen and unchosen options, respectively. Importantly,

preference changes occurred even without participants experiencing the consequences of their choice (i.e. the reward/outcomes associated with the chosen or unchosen options was never presented). These behavioral findings suggest choice-related alterations of associative memory processes in choice-induced revaluation. There was neuroimaging evidence for choice-induced strengthening of stimulus-outcome associations of previously chosen options, whereas we observed weakening of stimulus-outcome associations of previously unchosen options in the left hippocampus and right lOFC, two key regions involved in the acquisition and updating of stimulus-outcome associations (Boorman et al., 2016; Klein-Flügge et al., 2013). Additionally, the altered hippocampal representations of stimulus-outcome associations were correlated with future decision-making: Higher preference for previously chosen options (compared to otherwise equivalent options) was associated with stronger hippocampal stimulus-outcome representations. These results suggest that merely retrieving outcome representations and making a choice – even without experiencing the outcome related to that decision – induces plasticity in stimulus-outcome associations stored in associative memory systems and biases future decision-making.

In conjunction, the results of all experiments conducted in this work strongly suggest bidirectional interactions between decision-making and associative memory. Importantly, in each experiment, participants made decisions without ever experiencing the outcomes related to those choices. This might closely resemble real-world decision-making scenarios, where direct and immediate exposure to the consequences of choices or rewarding and punishing stimuli is rare. Instead, humans and other animals might regularly make use of memory mechanisms to simulate prospective future outcomes to make decisions (Wikenheiser & Schoenbaum, 2016) and employ their relational knowledge of actions and consequences. Here we show that exploiting acquired higher-order associative structures stored in memory might support transfer of value to stimuli that had themselves never been directly coupled with reinforcement (Gewirtz & Davis, 2000). Additionally, we found that choice-induced preference changes, or more broadly choice history biases, might arise from nonmonotonic plasticity processes in memory retrieval (Ritvo et al., 2019) and choice-related transformations of associative memories. In summary, both our key findings suggest that value representations in memory and decision-making influence each other bidirectionally.

5. Future Research

The present work provides evidence for bidirectional influences of decision-making and associative memory processes. There is evidence that value is transferred to stimuli that were never directly paired with reinforcement and this transfer seems to be related to the outcome reinstatement in the lOFC and interactions between lOFC and the amygdala/anterior hippocampus. However, the direction of information flow remains elusive due to the poor temporal resolution of fMRI. It is unclear, whether pattern information is first reinstated by the lOFC and then represented in the amygdala to form an association, suggesting a projection of reinstated patterns from lOFC to amygdala, or vice

versa. To resolve this issue, asymmetric lesions or separate optogenetic inactivation of lOFC and amygdala neurons in rodents, or transcranial magnetic stimulation of the lOFC in humans (Howard et al., 2020) during second-order conditioning could be used to delineate between those possibilities and to overcome the so far only correlational evidence provided here. Future studies will also be necessary to elucidate whether this putative information transfer is supported by phase coherence in theta oscillations between lOFC and amygdala/anterior hippocampus, a plausible neural mechanism for inter-area communication (Benchenane et al., 2010; Knudsen & Wallis, 2020; Young & Shapiro, 2011).

Another promising avenue for future studies is to provide a more conclusive behavioral dissociation between the choice heuristic and associative strengthening account for choice-induced preference reductions of unchosen options, supporting the conclusions that can be drawn from neuroimaging results. Participants during Experiment II A-C and E pursued the same behavioral goal during the choice-induced revaluation phase and decision probe: choose the *more preferred* option. The choice heuristic and associative strengthening accounts could be delineated by specifically changing participants' aim during the decision probe phase, i.e. instruct them to choose the *less preferred* option (Frömer, Dean Wolf, & Shenhav, 2019). If indeed choices/non-choices strengthened/weakened the respective stimulus-outcome association and hence the value of choice options, the associative account would predict increased choice probabilities for the previously unchosen option and decreased choice probabilities for the previously chosen option. However, if the stimulus-outcome associations remain unaltered by choice and participants instead acquire *go* and *no-go tags*, the choice heuristic account would predict the exact opposite choice pattern: increased choice probabilities for previously chosen options and decreased choice probabilities for previously unchosen options.

Since we used unique stimuli as options to be chosen/unchosen in the present experiments, an important future perspective could be to investigate (stimulus similarity-dependent) generalization of the found choice-induced preference change effects to other neutral stimuli in novel choice situations. It appears plausible, that preference for stimuli exhibiting perceptual similarity with the chosen/unchosen option (but are unpaired with any outcome) should be increased/decreased over other neutral, but more dissimilar stimuli in novel choice situations, suggesting similarity-based spread of choice-induced revaluation effects and value (Onat & Büchel, 2015).

Pharmacological interventions in humans could be used to investigate a putative role of dopamine, e.g. using dopamine D2 receptor antagonist amisulpride, in mediating the choice-induced preference increase and strengthening of stimulus-outcome associations for chosen options, akin to model-implied dopamine-dependent amplification of reward prediction errors related to free choices in the basal ganglia circuitry (Cockburn et al., 2014). This hypothetical future investigation could provide a neurochemical mechanism for the choice-induced preference change effects (for chosen options) presented in the current work.

6. Conclusion

The present work focused on influences between two cognitive faculties, decision-making and (associative) memory processes, providing evidence for bidirectional interactions. Participants made binary memory-based decisions during which they never experienced the outcomes related to their choices. This set-up might closely mirror real-world decision-making scenarios where consequences of choices typically unroll at later time points (if ever), which renders learning from outcomes and credit assignment to stimuli or actions challenging (Jocham et al., 2016; Walton, Behrens, Buckley, Rudebeck, & Rushworth, 2010). Here, we provide evidence for value transfer and biases of decision-making by higher-order associative memory processes and for choice-induced changes of associative memory representations. The former finding might explain how humans (and other animals) use associative mechanisms to perform goal-directed decision-making in novel contexts without any prior experience. The latter result contrasts with the predominant neo-classical economic model of decision-making, proposing that memory representations of preferences and choice option values unidirectionally guide choice behavior (Ariely & Norton, 2008). Instead, we found that retrieval of outcome representations and making a choice might be sufficient to induce plasticity in value representations stored in memory, suggesting that value representations and decision-making influence each other bidirectionally.

The present work thus strongly supports adopting a more cognitively flavored perspective on decision-making. This perspective – considering the interactions between value representations and choice behavior – has important implications for theories of learning and decision-making. Our results might for instance explain how flexible decision-making is possible in new choice contexts despite lacking prior experience and could contribute to our understanding of why humans (and other animals) tend to make coherent decisions and assign high value to previously chosen options (Riefer et al., 2017), despite experiencing negative consequences of those choices. This bias might have drastic consequences, especially for consumers sticking to unhealthy food choices, in substance dependence and obsessive-compulsive disorder, psychiatric conditions characterized by maladaptive decision-making or in voting behavior, where a candidate could be (re-)elected due to long-lasting choice-induced preference changes (Hornsby & Love, 2020). In a broader perspective, our results suggest that choice behavior might provide a means for individuals to manipulate the environment (Cockburn et al., 2014) and internal state variables, i.e. the way they remember how stimuli, action and consequences in the environment are related, if no external feedback to guide behavior is available. Additionally, the current results indicate that choice behavior could serve as a brain mechanism providing a spatio-temporal grounding to neural firing patterns and representations of the environment (Buzsáki, 2019). The present work strongly suggests that relational structures constituting decision-makers' cognitive maps (Behrens et al., 2018; Tolman, 1948) might be biased and dynamically transformed by their very own choice behavior. This provides a novel, intriguing and cognitive perspective on decision-making, essentially reversing the long-held assumption of neo-classical economic theories that associative memory processes influence choice unidirectionally.

References

Alink, A., Abdulrahman, H., & Henson, R. N. (2018). Forward models demonstrate that repetition suppression is best modelled by local neural scaling. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05957-0

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering Can Cause Forgetting: Retrieval Dynamics in Long-Term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1063–1087. https://doi.org/10.1037/0278-7393.20.5.1063

Ariely, D., & Norton, M. I. (2008). How actions create - not just reveal - preferences. *Trends in Cognitive Sciences*, *12*(1), 13–16. https://doi.org/10.1016/j.tics.2007.10.008

Barbas, H., & Blatt, G. J. (1995). Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus*, *5*(6), 511–533. https://doi.org/10.1002/hipo.450050604

Barnet, R. C., Arnold, H. M., & Miller, R. R. (1991). Simultaneous conditioning demonstrated in second-order conditioning: Evidence for similar associative structure in forward and simultaneous conditioning. *Learning and Motivation*, *22*(3), 253–268. https://doi.org/10.1016/0023-9690(91)90008-V

Barron, H. C., Dolan, R. J., & Behrens, T. E. J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience*, *16*(10), 1492–1498. https://doi.org/10.1038/nn.3515

Barron, H. C., Garvert, M. M., & Behrens, T. E. J. (2016). Repetition suppression: a means to index neural representations using BOLD? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705), 20150355. https://doi.org/10.1098/rstb.2015.0355

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06781-2

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Bellmund, J. L. S., Gärdenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. *Science*, *362*(6415). https://doi.org/10.1126/science.aat6766

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal- Prefrontal Network upon Learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/j.neuron.2010.05.013

Biderman, N., Bakkour, A., & Shohamy, D. (2020). What Are Memories For? The Hippocampus Bridges Past Experience with Future Decisions. *Trends in Cognitive Sciences*, *24*(7), 542–556. https://doi.org/10.1016/j.tics.2020.04.004

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning* (Vol. 53). New York: Springer Science+Business Media. https://doi.org/10.1017/CBO9781107415324.004

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: An image database for experimental research on eating and appetite. *Frontiers in Psychology*, *5*(JUN). https://doi.org/10.3389/fpsyg.2014.00617

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, *62*(5), 733–743. https://doi.org/10.1016/j.neuron.2009.05.014

Boorman, E. D., Rajendran, V. G., O'Reilly, J. X., & Behrens, T. E. (2016). Two Anatomically and Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome Associations in Hippocampus. *Neuron*, *89*(6), 1343–1354. https://doi.org/10.1016/j.neuron.2016.02.014

Bornstein, A. M., Khaw, M. W., Shohamy, D., & Daw, N. D. (2017). Reminders of past choices bias decisions for reward in humans. *Nature Communications*, *8*(May 2015), 1–9. https://doi.org/10.1038/ncomms15958

Bornstein, A. M., & Norman, K. A. (2017). Reinstated episodic context guides sampling-based decisions for reward. *Nature Neuroscience*, *20*(7), 997–1003. https://doi.org/10.1038/nn.4573

Braun, A., Urai, A. E., & Donner, T. H. (2018). Adaptive history biases result from confidence-weighted accumulation of past choices. *Journal of Neuroscience*, *38*(10), 2418–2429. https://doi.org/10.1523/JNEUROSCI.2189-17.2017

Brehm, J. W. (1956). Postdecision changes in the desireability of alternatives. *Journal of Abnormal and Social Psychology*, *52*(3), 384–389. https://doi.org/10.1037/h0041006

Buzsáki, G. (2019). *The Brain from Inside Out*. Oxford University Press.

Chen, M. K., & Risen, J. L. (2010). How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm. *Journal of Personality and Social Psychology*, *99*(4), 573–594. https://doi.org/10.1037/a0020217

Cockburn, J., Collins, A. G. E., & Frank, M. J. (2014). A Reinforcement Learning Mechanism Responsible for the Valuation of Free Choice. *Neuron*, *83*(3), 551–557. https://doi.org/10.1016/j.neuron.2014.06.035

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science (New York, N.Y.)*, *319*(5867), 1264–1267. https://doi.org/10.1126/science.1150605

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C. K., Hassabis, D., Munos, R., & Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. *Nature*, *577*(7792), 671–675. https://doi.org/10.1038/s41586-019-1924-6

Deserno, L., Huys, Q. J. M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A. A., … Schlagenhauf, F. (2015). Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5). https://doi.org/10.1073/pnas.1417219112

Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 7109–7114. https://doi.org/10.1073/pnas.94.13.7109

Eichenbaum, H., & Cohen, N. (2001). *From Conditioning to Conscious Recollection: Memory Systems of the Brain.* Oxford University Press, USA.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Fabbrini, F., Van den Haute, C., De Vitis, M., Baekelandt, V., Vanduffel, W., & Vogels, R. (2019). Probing the Mechanisms of Repetition Suppression in Inferior Temporal Cortex with Optogenetics. *Current Biology*, *29*(12), 1988-1998.e4. https://doi.org/10.1016/j.cub.2019.05.014

Frömer, R., Dean Wolf, C. K., & Shenhav, A. (2019). Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-12931-x

Garvert, M. M., Dolan, R. J., & Behrens, T. E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *ELife*, *6*, 1–20. https://doi.org/10.7554/elife.17086

Gewirtz, J. C., & Davis, M. (2000). Using Pavlovian higher-order conditioning paradigms to investigate the neural substrates of emotional learning and memory. *Learning and Memory*, *7*(5), 257–266. https://doi.org/10.1101/lm.35200

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, *301*(5636), 1104–1107. https://doi.org/10.1126/science.1087919

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*(1–3), 293–321. https://doi.org/10.1016/S0001-6918(01)00019-1

Guggenmos, M., Wilbertz, G., Hebart, M. N., & Sterzer, P. (2016). Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *ELife*, *5*(MARCH2016), 1–19. https://doi.org/10.7554/eLife.13388

Hanke, M., Halchenko, Y. ., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollman, S. (2009). PyMVPA: A Python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics.*, *7*(1), 37–53. https://doi.org/doi:10.1007/s12021-008-9041-y

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage*, *62*(2), 852–855. https://doi.org/10.1016/j.neuroimage.2012.03.016.Multivariate

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal corten. *Science*, *293*(September), 2425–2430. https://doi.org/10.4324/9780203496190

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523–534. https://doi.org/10.1038/nrn1931

Hornsby, A. N., & Love, B. C. (2020). How decisions and the desire for coherency shape subjective preferences over time. *Cognition*, *200*(August 2019), 104244. https://doi.org/10.1016/j.cognition.2020.104244

Howard, J. D., Reynolds, R., Smith, D. E., Voss, J. L., Schoenbaum, G., & Kahnt, T. (2020). Targeted Stimulation of Human Orbitofrontal Networks Disrupts Outcome-Guided Behavior. *Current Biology*, *30*(3), 490-498.e4. https://doi.org/10.1016/j.cub.2019.12.007

Hulbert, J. C., & Norman, K. A. (2015). Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Izuma, K., Matsumoto, M., Murayama, K., Samejima, K., Sadato, N., & Matsumoto, K. (2010). Neural correlates of cognitive dissonance and choice-induced preference change. *Proceedings of the National Academy of Sciences*, *107*(51), 22014–22019. https://doi.org/10.1073/pnas.1011879108

Jara, E., Vila, J., & Maldonado, A. (2006). Second-order conditioning of human causal learning. *Learning and Motivation*, *37*(3), 230–246. https://doi.org/10.1016/j.lmot.2005.12.001

Jocham, G., Brodersen, K. H. H., Constantinescu, A. O. O., Kahn, M. C. C., Ianni, A. M., Walton, M. E. E., … Behrens, T. E. E. J. (2016). Reward-Guided Learning with and without Causal Attribution. *Neuron*, *90*(1), 177–190. https://doi.org/10.1016/j.neuron.2016.02.018

Jocham, G., Hunt, L. T., Near, J., & Behrens, T. E. J. (2012). A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nature Neuroscience*, *15*(7), 960–961. https://doi.org/10.1038/nn.3140

Juechems, K., & Summerfield, C. (2019). Where Does Value Come From? *Trends in Cognitive Sciences*, *23*(10), 836–850. https://doi.org/10.1016/j.tics.2019.07.012

Kaiser, L., Gruendler, T., Speck, O., Luettgau, L., & Jocham, G. (2019). *Dissociable roles of cortical excitation-inhibition balance during patch-leaving versus value-guided decisions*. *bioRxiv*. https://doi.org/10.1101/2019.12.15.877019

Klein-Flügge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., & Behrens, T. E. J. (2013). Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex. *Journal of Neuroscience*, *33*(7), 3202–3211. https://doi.org/10.1523/JNEUROSCI.2532-12.2013

Klein, T. A., Ullsperger, M., & Jocham, G. (2017). Learning relative values in the striatum induces violations of normative decision making. *Nature Communications*, *8*, 1–12. https://doi.org/10.1038/ncomms16033

Knudsen, E. B., & Wallis, J. D. (2020). Closed-Loop Theta Stimulation in the Orbitofrontal Cortex Prevents Reward-Based Learning. *Neuron*, *106*(3), 537-547.e4. https://doi.org/10.1016/j.neuron.2020.02.003

Kobayashi, M., Takeda, M., Hattori, N., Fukunaga, M., Sasabe, T., Inoue, N., … Watanabe, Y. (2004). Functional imaging of gustatory perception and imagery: "Top-down" processing of gustatory signals. *NeuroImage*, *23*(4), 1271–1282. https://doi.org/10.1016/j.neuroimage.2004.08.002

Kondo, H., & Witter, M. P. (2014). Topographic organization of orbitofrontal projections to the parahippocampal region in rats. *Journal of Comparative Neurology*, *522*(4), 772–793. https://doi.org/10.1002/cne.23442

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(November), 1–28. https://doi.org/10.3389/neuro.06.004.2008

Kringelbach, M. L., O'Doherty, J., Rolls, E. T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex*, *13*(10), 1064–1071. https://doi.org/10.1093/cercor/13.10.1064

Kurth-Nelson, Z., Barnes, G., Sejdinovic, D., Dolan, R., & Dayan, P. (2015). Temporal structure in associative retrieval. *ELife*, *4*, 1–18. https://doi.org/10.7554/eLife.04919

Kurth-Nelson, Z., Economides, M., Dolan, R. J., & Dayan, P. (2016). Fast Sequences of Non-spatial State Representations in Humans. *Neuron*, *91*(1), 194–204. https://doi.org/10.1016/j.neuron.2016.05.028

Larsson, J., Solomon, S. G., & Kohn, A. (2016). fMRI adaptation revisited. *Cortex*, *80*, 154–160. https://doi.org/10.1016/j.cortex.2015.10.026

Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological Science*, *22*(10), 1310–1318. https://doi.org/10.1177/0956797611417005

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Rangananth, C., & Redish, A. D. (2017). Viewpoints: How the hippocampus contributes to memory, navigation and cognition. *Nature Neuroscience*, *20*(11), 1434–1447. https://doi.org/10.1038/nn.4661

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

Luettgau, L., Tempelmann, C., Kaiser, L. F., & Jocham, G. (2020). Decisions bias future choices by modifying hippocampal associative memories. *Nature Communications*, *11*, 3318. https://doi.org/10.1101/802462

Maes, E. J. P., Sharpe, M. J., Usypchuk, A. A., Lozzi, M., Chang, C. Y., Gardner, M. P. H., … Iordanova, M. D. (2020). Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. *Nature Neuroscience*, *23*(2), 176–178. https://doi.org/10.1038/s41593-019-0574-1

Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., & Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nature Human Behaviour*, *1*(9), 680–692. https://doi.org/10.1038/s41562-017-0180-8

Mondragón, E., Alonso, E., & Kokkola, N. (2017). Associative Learning Should Go Deep. *Trends in Cognitive Sciences*, *21*(11), 822–825. https://doi.org/10.1016/j.tics.2017.06.001

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, *59*(3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076

Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, *22*(10), 1544–1553. https://doi.org/10.1038/s41593-019-0470-8

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, *14*(6), 769–776. https://doi.org/10.1016/j.conb.2004.10.016

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. *Hippocampus*. https://doi.org/10.1177/1350507609340812

Onat, S., & Büchel, C. (2015). The neuronal basis of fear generalization in humans. *Nature Neuroscience*, *18*(12), 1811–1818. https://doi.org/10.1038/nn.4166

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676

Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron*, *96*(4), 736–754. https://doi.org/10.1016/j.neuron.2017.09.031

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms9096

Parkes, S. L., & Westbrook, R. F. (2011). Role of the basolateral amygdala and NMDA receptors in higher-order conditioned fear. *Reviews in the Neurosciences*, *22*(3), 317–333. https://doi.org/10.1515/RNS.2011.025

Pavlov, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press. Retrieved from https://books.google.de/books?id=XbOjAQAAQBAJ

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

Riefer, P. S., Prior, R., Blair, N., Pavey, G., & Love, B. C. (2017). Coherency-maximizing exploration in the supermarket. *Nature Human Behaviour*, *1*(1), 1–4. https://doi.org/10.1038/s41562-016-0017

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic Plasticity: How Memory Retrieval Drives Learning. *Trends in Cognitive Sciences*, *23*(9), 726–742. https://doi.org/10.1016/j.tics.2019.06.007

Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology*, *81*(1), 1–11. https://doi.org/10.1037/h0033333

Rolls, E. T. (2000). The Orbitofrontal Cortex and Reward. *Cerebral Cortex*, *10*, 284–294. https://doi.org/10.1093/cercor/10.3.284

Rolls, E. T. (2006). Brain mechanisms underlying flavour and appetite. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1471), 1123–1136. https://doi.org/10.1098/rstb.2006.1852

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*(March), 104269. https://doi.org/10.1016/j.cognition.2020.104269

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447). https://doi.org/10.1126/science.aaw5181

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, *14*(2), 139–147. https://doi.org/10.1016/j.conb.2004.03.017

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Seo, D., & Sinha, R. (2011). Neural Mechanisms of Stress and Addiction. In B. Adinoff & E. A. Stein (Eds.), *Neuroimaging in Addiction* (pp. 209–233). Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119998938.ch9

Sharot, T., Martino, B. De, & Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *Journal of Neuroscience*, *29*(12), 3760–3765. https://doi.org/10.1523/JNEUROSCI.4972-08.2009

Sharot, T., Velasquez, C. M., & Dolan, R. J. (2010). Do decisions shape preference? Evidence from blind choice. *Psychological Science*, *21*(9), 1231–1235. https://doi.org/10.1177/0956797610379235

Sharpe, M. J., Batchelor, H. M., Mueller, L. E., Yun Chang, C., Maes, E. J. P., Niv, Y., & Schoenbaum, G. (2020). Dopamine transients do not act as model-free prediction errors during associative learning. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-019-13953-1

Sharpe, M. J., Batchelor, H. M., & Schoenbaum, G. (2017). Preconditioned cues have no value. *ELife*, *6*. https://doi.org/10.7554/eLife.28362

Shohamy, D., & Daw, N. D. (2015). Integrating memories to guide decisions. *Current Opinion in Behavioral Sciences*, *5*, 85–90. https://doi.org/10.1016/j.cobeha.2015.08.010

Small, D. M., Zald, D. H., Jones-Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., & Petrides, M. (1999). Human cortical gustatory areas: A review of functional neuroimaging data. *NeuroReport*, *10*(1), 7–14. https://doi.org/10.1097/00001756-199901180-00002

Soltani, A., & Wang, X. J. (2010). Synaptic computation underlying probabilistic inference. *Nature Neuroscience*, *13*(1), 112–119. https://doi.org/10.1038/nn.2450

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

Staresina, B. P., & Wimber, M. (2019). A Neural Chronometry of Memory Recall. *Trends in Cognitive Sciences*, *23*(12), 1071–1085. https://doi.org/10.1016/j.tics.2019.09.011

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

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, *29*(5), 1565–1572. https://doi.org/10.1523/JNEUROSCI.4657-08.2009

Storm, B. C., Bjork, E. L., & Bjork, R. A. (2008). Accelerated Relearning After Retrieval-Induced Forgetting: The Benefit of Being Forgotten. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(1), 230–236. https://doi.org/10.1037/0278-7393.34.1.230

Summerfeld, C., Wyart, V., Johnen, V. M., & de Gardelle, V. (2011). Human scalp electroencephalography reveals that repetiti suppression varies with expectation. *Frontiers in Human Neuroscience*, *5*(JULY), 1–13. https://doi.org/10.3389/fnhum.2011.00067

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press.

Suzuki, S., Cross, L., & O'Doherty, J. P. (2017). Elucidating the underlying components of food valuation in the human orbitofrontal cortex. *Nature Neuroscience*, *20*(12), 1780–1786. https://doi.org/10.1038/s41593-017-0008-x

Takahashi, Y. K., Batchelor, H. M., Liu, B., Khanna, A., Morales, M., & Schoenbaum, G. (2017). Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards. *Neuron*, *95*(6), 1395-1405.e3. https://doi.org/10.1016/j.neuron.2017.08.025

Tamaoka, K., Makioka, S., Sanders, S., & Verdonschot, R. G. (2017). www.kanjidatabase.com: a new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychological Research*, *81*(3), 696–708. https://doi.org/10.1007/s00426-016-0764-3

Tolman, E. C. (1948). Cognitive Maps in Rats and Man. *Psychological Review*, *55*(4), 189–208.

Tonegawa, S., Morrissey, M. D., & Kitamura, T. (2018). The role of engram cells in the systems consolidation of memory. *Nature Reviews Neuroscience*, *19*(8), 485–498. https://doi.org/10.1038/s41583-018-0031-2

Urai, A. E., De Gee, J. W., Tsetsos, K., & Donner, T. H. (2019). Choice history biases subsequent evidence accumulation. *ELife*, *8*, 1–34. https://doi.org/10.7554/eLife.46331

Wagner, A. R. (1981). SOP : A model of automatic processing in animal behavior. In *Information Processing in Animals: Conditioned Inhibition* (pp. 5–24).

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. *Neuron*, *65*(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027

Wang, F., Schoenbaum, G., & Kahnt, T. (2020). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biology*, *18*(1), 1–24. https://doi.org/10.1371/journal.pbio.3000578

Wang, L., Gillis-Smith, S., Peng, Y., Zhang, J., Chen, X., Salzman, C. D., … Zuker, C. S. (2018). The coding of valence and identity in the mammalian taste system. *Nature*, *558*(7708), 127–131. https://doi.org/10.1038/s41586-018-0165-4

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, *17*(8), 513–523. https://doi.org/10.1038/nrn.2016.56

Wimber, M., Alink, A., Charest, I., Kriegeskorte, N., & Anderson, M. C. (2015). Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nature Neuroscience*, *18*(4), 582–589. https://doi.org/10.1038/nn.3973

Wimmer, G. E., & Büchel, C. (2019). Learning of distant state predictions by the orbitofrontal cortex in humans. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-10597-z

Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science (New York, N.Y.)*, *338*(October), 270–273. https://doi.org/10.1126/science.1223252

Wunderlich, K., Range, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(40), 17199–17204. https://doi.org/10.1073/pnas.0901077106

Yassa, M. A., & Reagh, Z. M. (2013). Competitive trace theory: A role for the hippocampus in contextual interference during retrieval. *Frontiers in Behavioral Neuroscience*, *7*(JUL), 1–13. https://doi.org/10.3389/fnbeh.2013.00107

Yiannakas, A., & Rosenblum, K. (2017). The insula and taste learning. *Frontiers in Molecular Neuroscience*, *10*(November), 1–24. https://doi.org/10.3389/fnmol.2017.00335

Young, J. J., & Shapiro, M. L. (2011). Dynamic coding of goal-directed paths by orbital prefrontal cortex. *Journal of Neuroscience*, *31*(16), 5989–6000. https://doi.org/10.1523/JNEUROSCI.5436-10.2011

List of Publications

The following section lists the publications and manuscripts submitted to journals with peer-review process that this thesis was based upon. The respectively cited literature can be found at the end of each manuscript in the reference list.

1.) Luettgau L., Tempelmann C., Kaiser L. F., & Jocham G. (2020). Decisions bias future choices by modifying hippocampal associative memories. *Nature Communications*. doi: 10.1038/s41467-020-17192-7. Preprint at *bioRxiv* (2019), doi: 10.1101/802462

Author contributions: Lennart Lüttgau (please note the transcription of German “ü” to English “ue”) 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. Claus Tempelmann set up the MRI acquisition protocol, read and edited versions of the manuscript and approved the final version of the manuscript. Luca Franziska Kaiser analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript. Gerhard Jocham designed the study and conceptualized research, analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript.

2.) Luettgau L., Porcu E., Tempelmann C., & Jocham G. (under review at *Nature Communications*). Reinstatement of cortical outcome representations during higher-order learning. Preprint at *bioRxiv* (2020), doi: 10.1101/2020.05.28.121558

Author contributions: Lennart Lüttgau 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. Emanuele Porcu analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript. Claus Tempelmann set up the MRI acquisition protocol, read and edited versions of the manuscript and approved the final version of the manuscript. Gerhard Jocham designed the study and conceptualized research, analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript.

Eidesstattliche Erklärung

Ich versichere an Eides statt, dass die Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Die Dissertation wurde in der vorliegenden oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 22.09.2020

Lennart Lüttgau

Attachments

Publications and submitted manuscripts under review

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-17192-7 OPEN

Decisions bias future choices by modifying hippocampal associative memories

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

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

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

According to neo-classical economic models of decision-making, choices are guided by memories of option values[1]. This unidirectional view has been challenged by cognitive accounts of decision-making, suggesting that memory representations of option values might themselves be subject to changes induced by an agent's choices. This suggests a bidirectional relationship between value representations in memory and decision-making[1,2].

Even though real-life decisions often involve memory retrieval of learned associations between reward-predictive cues and outcomes, memory mechanisms underlying choice-induced preference changes have, to the best of our knowledge, never been systematically studied[1,3–5]. This might be partially related to the fact that most studies on choice-induced preference changes employed direct presentation of the outcomes to be chosen. However, this approach by design obliterates and confounds underlying associative learning contributions to the revaluation process[6], and is blind to related memory processes, such as retrieval competition[7].

In naturalistic decision-making scenarios, choices often have to be made without direct experience of feedback. Instead, decision makers have to rely on relational knowledge of actions and outcomes. Likely candidate mechanisms for behavioral adaptation without direct external feedback are memory retrieval dynamics. It is well established that retrieval of an item from memory, e.g. a conditioned stimulus (CS) triggering retrieval of an associated outcome, leads to improved remembering. However, memory for competing items, e.g. another CS associated with the same outcome, is impaired simultaneously[7–9]. Such retrieval-induced forgetting[8] would predict choice biases towards previously chosen CS based on retrieval-related strengthening of CS–US association. However, the same effect would be predicted for a previously presented, but unchosen CS, since both chosen and unchosen CS activate neural populations representing the associated outcome[10–15]. A recent theoretical framework[16] suggests nonmonotonic plasticity during associative memory retrieval: Inactive memories remain unaltered, moderately activated associative memories are weakened, and high activation strengthens memories[16]. Translating this idea to memory-based decisions between two CS, we assume that both CS will moderately activate neural populations representing the associated outcome (as the outcome is never presented). However, consistent with the finding that chosen options receive higher attentional weighting than unchosen options (as reflected in learning rates[17,18]), we further assume that choices of a CS will induce additional activation of the associated outcome, whereas this will not be the case for unchosen CS, retaining an intermediate activation state of the associated outcome. Thus, we hypothesize that choosing a CS will strengthen the related stimulus-outcome association. Conversely, not choosing a CS will weaken the respective stimulus-outcome association. We expect that these choice-induced alterations of the associative memory structure will result in subsequent preference changes.

We expect choice-induced preference changes to be driven by modifications of stimulus-outcome associations in the hippocampus and lateral orbitofrontal cortex, two key regions for storing and updating associative representations[10,14].

Thus, the goal of the present study is twofold. First, we aim at investigating how choice-related alterations of associative memories bias future decision-making. Second, we seek to investigate a neurobiologically plausible mechanism underlying choice-induced preference changes. To test our key predictions, we designed a learning and decision-making paradigm which we use in three independent behavioral experiments and one functional magnetic resonance imaging (fMRI) experiment. For the fMRI study, we exploit repetition suppression (RS) effects[10,11,19–21] to measure associative strength between conditioned (CS) and unconditioned stimuli (US)[10,14]. Participants first establish Pavlovian associations between CS and differently valued US. Next, in a choice-induced revaluation, participants make binary choices between differently valued CS, without observing the associated US. Finally, in a probe phase, where they make choices between all possible CS combinations, participants show preference increases for previously chosen, and preferences decreases for previously unchosen CS, compared to otherwise equivalent CS. These choice-induced alterations in decision behavior are accompanied by corresponding changes in CS–US RS effects in the hippocampus and lateral orbitofrontal cortex. Our findings are corroborated by multivariate pattern similarity analyses (a variant of representational similarity analysis, RSA[22]). Furthermore, the magnitude of the hippocampal RS effect correlates with individual probe phase decision biases.

Results

Behavioral experiments. First, we detailed the behavioral choice-induced revaluation effect in three independent experiments. In each experiment, participants learnt associations between neutrally rated CS[23] and three food items[24] serving as unconditioned stimuli (US, Fig. 1a). For each participant, the US were individually chosen based on a prior rating of subjective preference, and a low-value (US^-), an intermediate-value (US^0), and a high-value (US^+) food item was selected. Next, participants rated kanji stimuli[23] according to their subjective preference. Six of these kanjis rated in close proximity to neutral were selected as CS for Pavlovian learning. Two CS each were paired with one US, resulting in three categories of differently valued CS: $CS^+_{A/B}$, $CS^0_{A/B}$, and $CS^-_{A/B}$, for high-value, intermediate-value, and low-value CS.

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

Decisions are biased by past choices. There was evidence for value transfer from US to CS across all studies (Fig. 1e–h), as indicated by significant main effects of CS value on probe phase choice probabilities (CP, all Fs > 94.99, Ps < 0.001, η_p^2s > 0.69, $1-\beta$s > 0.99, repeated-measure analyses of variance, rmANOVA). Decision-making during the revaluation phase had clearly dissociable effects on choices during the probe phase. CS that were chosen during the revaluation phase were more likely to be selected during the later probe phase compared to the CS of equal value that were not presented during choice-induced revaluation.

In Experiment 1, participants ($N = 40$) made choices between the intermediate-value CS^0_A and the high-value CS^+_A during the choice-induced revaluation phase. As we had directed hypotheses for the choice effects (increased CP for the previously chosen and decreased CP for the previously unchosen CS), we used one-tailed post hoc tests. In the probe phase, participants preferred CS^+_A, the previously selected stimulus, compared to CS^+_B ($Z = 3.98$,

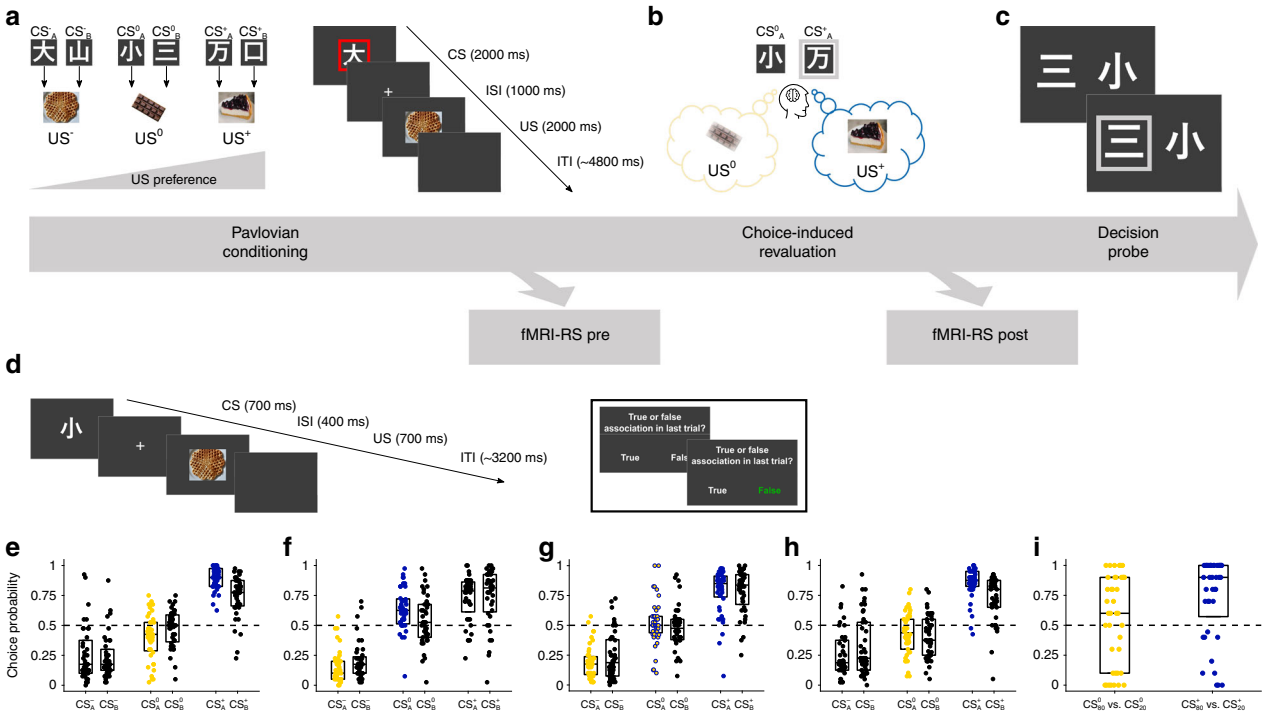


Fig. 1 Task schematic and behavioral results. a Participants rated subjective desirability of conditioned stimuli (CS, kanjis) and unconditioned stimuli (US, food items). Kanji images were selected from an online database (https://www.kanjidatabase.com/)[23]. During Pavlovian conditioning, participants learned to associate six CS with three US. Each US was associated with two CS. **b** Choice-induced revaluation: after Pavlovian conditioning, participants made choices between CS^{+}_{A} versus CS^{0}_{A} (Experiments 1 and 5), CS^{0}_{A} versus CS^{-}_{A} (Experiment 2), CS^{0}_{A} versus either CS^{-}_{A} or CS^{+}_{A} (Experiment 3), and CS^{+}_{80} versus CS^{0}_{80} or CS^{+}_{20} versus CS^{0}_{20} (Experiment 4), interleaved with lure decisions. **c** Decision probe: following choice-induced revaluation, participants made binary choices between all possible combinations of CS, or CS^{+}_{80} versus CS^{+}_{20} and CS^{0}_{80} versus CS^{0}_{20} (Experiment 4), to assess preferences. **d** Attentional control task performed during fMRI repetition suppression (Experiment 5). **e**–**h** Previously chosen CS (blue dots) are selected more often compared to equivalent CS (black dots) in Experiment 1 (**e**, $N = 40$, $Z = 3.98$, $P < 0.001$, Cohen's $U_3 = 0.85$, Wilcoxon signed-rank test, one-tailed), Experiment 2 (**f**, $N = 40$, $Z = 2.20$, $P = 0.014$, $U_3 = 0.68$, Wilcoxon signed-rank test, one-tailed), Experiment 5 (**h**, $N = 42$, $Z = 3.03$, $P = 0.001$, $U_3 = 0.76$, Wilcoxon signed-rank test, one-tailed) and previously unchosen CS (yellow dots) are selected less often compared to equivalent CS (black dots) in Experiment 1 (**e**, $N = 40$, $Z = 1.97$, $P = 0.025$, $U_3 = 0.70$, Wilcoxon signed-rank test, one-tailed) and Experiment 2 (**f**, $N = 40$, $Z = 1.91$, $P = 0.028$, $U_3 = 0.66$, Wilcoxon signed-rank test, one-tailed) during decision probe. The effect is not present in Experiment 3 (**g**, $N = 44$, $Z = 0.41$, $P = 0.680$, $U_3 = 0.55$, Wilcoxon signed-rank test, two-tailed), indicating that the roughly equal proportion of choices and non-choices of CS^{0}_{A} during revaluation had canceled each other out. **i** Behavioral control experiment (Experiment 4), orthogonalizing contributions of go and no-go tagging and associative strength between CS and US to choice probabilities. Previously chosen (go tag) and strongly associated CS^{+}_{80} is preferred over previously chosen and weakly associated CS^{+}_{20} (blue dots, $N = 40$, $Z = 3.55$, $P < 0.001$, $U3_1 = 0.75$, $1\text{-}\beta > 0.99$, one-sample Wilcoxon signed-rank test, one-tailed), while there is only descriptive evidence for preference of previously unchosen (no-go tag) and strongly associated CS^{0}_{80} over previously unchosen and weakly associated CS^{0}_{20} (yellow dots, $N = 40$, $Z = 0.61$, $P = 0.271$, $U3_1 = 0.61$, $1\text{-}\beta = 0.23$, one-sample Wilcoxon signed-rank test, one-tailed). Box plot center lines represent sample medians and box bottom/top edges show 25th/75th percentile of the data, respectively. Source data are provided as a Source Data file.

$P < 0.001$, Cohen's $U_3 = 0.85$, Wilcoxon signed-rank test, one-tailed). This effect was mainly driven by preference for CS^{+}_{A} in pairwise within-category choice trials between CS^{+}_{A} and CS^{+}_{B} ($Z = 3.43$, $P < 0.001$, Cohen's $U3_1 = 0.69$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2e). Conversely, participants selected CS^{0}_{A}, the previously non-selected stimulus, compared to CS^{0}_{B} less likely ($Z = 1.97$, $P = 0.025$, $U_3 = 0.70$, Wilcoxon signed-rank test, one-tailed). Again, this effect was mainly driven by reduced choice of CS^{0}_{A} in pairwise within-category choice trials contrasting CS^{0}_{A} and CS^{0}_{B} ($Z = 2.05$, $P = 0.020$, $U3_1 = 0.68$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2e). The observed dissociation in choice behavior was also evident in a significant interaction effect of CS value × CS type (A or B): $F_{2,\ 78} = 10.01$, $P < 0.001$, $\eta^2_p = 0.20$, $1\text{-}\beta > 0.99$ (rmANOVA, Fig. 1e). Thus, compared to equivalent CS, participants exhibited a systematic preference for CS they had previously chosen, whereas they displayed a diminished preference of CS they had previously not chosen.

After having established that an intermediate value CS (CS^{0}_{A}) could be devalued by non-choices, we next asked in Experiment 2, whether we could induce the exact opposite — increased preference for CS^{0}_{A}. Therefore, participants ($N = 40$) were presented with decisions between intermediate-value CS^{0}_{A} and low-value CS^{-}_{A} during the choice-induced revaluation phase. Conceptually replicating the results of Experiment 1, participants in Experiment 2 favored the previously chosen CS^{0}_{A} over CS^{0}_{B} ($Z = 2.20$, $P = 0.014$, $U_3 = 0.68$, Wilcoxon signed-rank test, one-tailed) during the decision probe, resulting from preference for CS^{0}_{A} in pairwise within-category choice trials between CS^{0}_{A} and CS^{0}_{B} ($Z = 1.93$, $P = 0.027$, $U3_1 = 0.68$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2f). Conversely, participants neglected CS^{-}_{A} compared to CS^{-}_{B} ($Z = 1.91$, $P = 0.028$, $U_3 = 0.66$, Wilcoxon signed-rank test, one-tailed),

resulting from descriptively reduced preference for CS^-_A in pairwise within-category choice trials between CS^-_A and CS^-_B ($Z = 1.41$, $P = 0.079$, $U3_1 = 0.63$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2f). Again, there was a significant interaction effect of CS value × CS type ($F_{2,\ 78} = 4.84$, $P = 0.010$, $\eta_p^2 = 0.11$, $1\text{–}\beta > 0.99$, rmANOVA, Fig. 1f) indicating clearly dissociable choice behavior during the decision probe. This pattern of results suggests a value-independent mechanism of choice-induced revaluation.

These results so far show that choices and non-choices act in opposite directions. Consequently, we predicted that choice-induced preference increases and devaluation of CS^0_A would cancel out. To test this prediction, in Experiment 3 ($N = 44$), we presented an equal number of binary decisions between CS^0_A and CS^-_A as between CS^0_A and CS^+_A during choice-induced revaluation. Since choice-induced revaluation effects for CS^0_A should cancel each other out, we did not have directed hypotheses for the choice effects. We thus used two-tailed post hoc tests. As expected, there was no evidence for change in preference for CS^0_A compared to CS^0_B ($Z = 0.41$, $P = 0.680$, $U_3 = 0.55$, Wilcoxon signed-rank test, two-tailed, Fig. 1g). Consistently, there was no evidence for CS^0_A preference changes in pairwise within-category choices between CS^0_A and CS^0_B ($Z = 0.12$, $P = 0.905$, $U3_1 = 0.57$, one-sample Wilcoxon signed-rank test vs. 0.5, two-tailed; Supplementary Fig. 2g), suggesting that the effects of choices and non-choices had indeed canceled each other out (interaction effect of CS value × CS type: $F_{2,\ 86} = 1.31$, $P = 0.280$, $\eta_p^2 = 0.03$, $1\text{–}\beta = 0.71$, rmANOVA).

Importantly, the dissociations observed in choice behavior in Experiments 1 and 2 rule out alternative accounts for explaining choice behavior, such as extinction or mere exposure effects. Both accounts would predict unidirectional preference changes for the CS presented during choice-induced revaluation, independent of the choices made (decreases or increases in preference, respectively), which is incompatible with the present results.

However, an alternative explanation for the observed choice pattern is that participants learned simple choice rules for the two CS presented during the revaluation phase, akin to go tags for chosen CS ("choose this stimulus") and no-go tags for unchosen CS ("do not choose this stimulus"). Accordingly, the observed changes in preferences could be attributed to repeating such choice heuristics acquired during revaluation. An additional behavioral experiment (Experiment 4, $N = 40$) was specifically designed to address this possibility. We orthogonalized contributions of associative strength and choice rule by letting participants assign choice-induced go tags to two chosen CS^+ that differed in their associative strength to US^+ (80% vs. 20% association) and no-go tags to two unchosen CS^0 that likewise differed in their associative strength to US^0 (80% vs. 20%). According to our hypothesis (associative account), probe phase decisions are guided by the learned associations and the strengthening/weakening of this association during revaluation. Therefore, we expected significantly increased CP for both highly associated stimuli: CS^0_{80} should be preferred over CS^0_{20}, and CS^+_{80} should be preferred over CS^+_{20}. Contrarily, if choice behavior was instead exclusively driven by learned go and no-go tags (heuristic account), both same-value pairwise CP should be at chance level (CP = 0.50). Due to the directionality of our hypothesis, we used one-tailed tests.

Importantly, there was no significant difference between revaluation CP of CS^+_{80} versus CS^0_{80} and CS^+_{20} versus CS^0_{20} ($Z = 1.19$, $P = 0.234$, $U_3 = 0.53$, Wilcoxon signed-rank test, two-tailed), ruling out unequal assignment of go and no-go tags across CS pairs. We observed that participants favored the previously chosen and strongly associated CS^+_{80} over the previously chosen and weakly associated CS^+_{20} ($Z = 3.55$, $P < 0.001$, $U3_1 = 0.75$, $1\text{–}\beta > 0.99$, one-sample Wilcoxon signed-rank test, one-tailed). Descriptively, participants also tended to favor the previously unchosen and strongly associated CS^0_{80} over the previously unchosen and weakly associated CS^0_{20} ($Z = 0.61$, $P = 0.271$, $U3_1 = 0.61$, $1\text{–}\beta = 0.23$, one-sample Wilcoxon signed-rank test, one-tailed, Fig. 1i) during the decision probe phase. This pattern of results favors an explanation based on associative strengthening of the memory trace between CS^+ and US^+, rather than on merely expressing a go tag. However, there is no definite evidence against the alternative explanation that participants learned a no-go tag for the unchosen stimuli. This asymmetric expression of response tendencies might result from differential acquisition of go and no-go choice rules, akin to well-described Pavlovian biases[25]. Presumably, during high-value (CS^+_{80} vs. CS^+_{20}) choices, most participants used the learned CS–US associative strength instead of go response tendencies to guide their decisions, while this only tended to be the case for intermediate-value (CS^0_{80} vs. CS^0_{20}) choices (Median CP = 0.60). Consistent with modeling and empirical evidence for asymmetric action and inaction learning[26], reverting the initially learned action tendency for CS^+_{20} could have less of an impeding effect on re-acquisition of a no-go response during decision probe, than re-acquisition of a go response for CS^0_{80}, which was initially learned with an inaction choice rule.

For each experiment (Experiments 1, 2, 3, and 5), we compared six reinforcement learning models[27] that implemented different ways by which participants could have learned CS–US associations—and updated associative strength during choice-induced revaluation based on fictive reward prediction errors (RPE). The fictive RPE were based on our reasoning that presentation of CS during revaluation would lead to retrieval of the associated US and strengthening/weakening of the chosen/unchosen CS–US association, respectively. Our behavioral results were best captured by a model that differentially updated the learned CS–US associative strengths using fictive RPE elicited by revaluation phase decisions (see Methods section and Supplementary Table 1). Simulations using the best-fitting parameters successfully reproduced the observed empirical choice pattern (with the exception of the observed reduced CP of CS^0_B in Experiment 5, Supplementary Fig. 1e–h). Thus, the best fitting models likely incorporate candidate computational mechanisms underlying the observed choice biases.

Choices modify univariate neural measures of stimulus-outcome associations. Having established and replicated the behavioral effect of choice-induced revaluation in three independent behavioral samples, we next tested whether decisions induce changes of neural representations of CS–US associations. In Experiment 5, we used fMRI and leveraged RS effects[10,11,19–21] to measure CS–US associative strength[10,14]. When a neural ensemble is activated twice in brief succession (e.g. by rapid sequential presentation of the same visual stimulus), the second stimulus causes a diminished response. Accordingly, after learning the association between CS and US, the CS should elicit a representation of its associated US. Thus, presentation of the US itself, following the CS, should induce a diminished neural response. If the association between CS and US has been weakened by non-choices during revaluation, the CS is no longer capable of evoking the US representation to the same degree and should therefore elicit a stronger response (less RS). The same logic in reverse applies when the association has been strengthened by choices during revaluation.

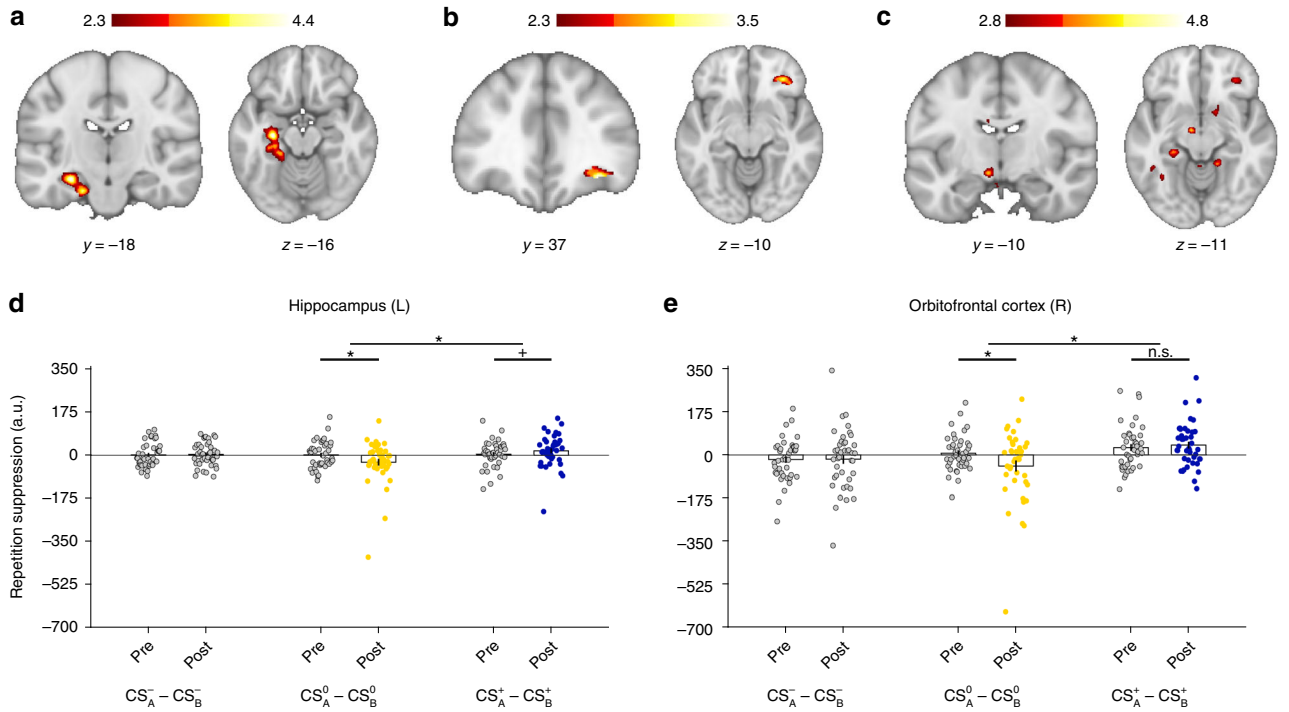


Fig. 2 Whole-brain effects of choice-induced revaluation. a Left hippocampus encodes conjunction of decreased repetition suppression for $CS^0_A - US^0$ relative to $CS^0_B - US^0$, controlling for activation elicited by CS^0_A and CS^0_B followed by both incorrect outcomes (US^- and US^+) (Eq. (1)) AND increased $CS^+_A - US^+$ repetition suppression relative to $CS^+_B - US^+$, controlling for activation elicited by CS^+_A and CS^+_B followed by both incorrect outcomes (US^- and US^0) (Eq. (2)). **b** Using small-volume correction (independent mask from ref. [28]), the same effect was found in the right lateral orbitofrontal cortex (lOFC). **c** At a lenient threshold of $Z > 2.8$ (uncorrected) we observed the same effect in left ventral tegmental area and right nucleus accumbens. **d** Extracted parameter estimates of the effect in left hippocampus ($N = 42$), showing an interaction effect of CS value and time (PRE or POST), $F_{1,\ 41} = 4.51$, $P = 0.040$, $\eta_p^2 = 0.10$, $1\text{-}\beta = 0.99$, repeated-measures ANOVA (rmANOVA). There was evidence for significant PRE–POST reduction in repetition suppression for CS^0_A ($Z = 1.84$, $P = 0.033$, $U_3 = 0.69$, Wilcoxon signed-rank test, one-tailed). However, PRE–POST increase in repetition suppression for CS^+_A was not significant ($Z = 1.48$, $P = 0.070$, $U_3 = 0.62$, Wilcoxon signed-rank test, one-tailed). We observed decreased repetition suppression for CS^0_A ($Z = 2.26$, $P = 0.012$, $U3_1 = 0.67$, one-sample Wilcoxon signed-rank test, one-tailed) and increased repetition suppression for CS^+_A ($Z = 2.26$, $P = 0.012$, $U3_1 = 0.69$, one-sample Wilcoxon signed-rank test, one-tailed) during POST, without any evidence for non-zero differences in PRE (all Zs < 1.05, Ps > 0.296, $U3_1 < 0.62$, one-sample Wilcoxon signed-rank tests, two-tailed) or for the control contrast of either CS^-_A or CS^-_B (all Zs < 0.31, Ps > 0.760, $U3_1 < 0.55$, one-sample Wilcoxon signed-rank tests, two-tailed). **e** Extracted parameter estimates of the effect in right lOFC ($N = 42$). The interaction effect of CS value and time (PRE or POST) was significant, $F_{1,\ 41} = 5.57$, $P = 0.023$, $\eta_p^2 = 0.12$, $1\text{-}\beta = 0.99$, rmANOVA. We found a significant PRE–POST reduction in repetition suppression for CS^0_A ($Z = 1.77$, $P = 0.039$, $U_3 = 0.55$, Wilcoxon signed-rank test, one-tailed), but no evidence of a PRE–POST increase in repetition suppression for CS^+_A ($Z = 0.65$, $P = 0.260$, $U_3 = 0.50$, Wilcoxon signed-rank test, one-tailed). However, we found decreased repetition suppression for CS^0_A ($Z = 1.74$, $P = 0.040$, $U3_1 = 0.55$, Wilcoxon signed-rank test, one-tailed) and increased repetition suppression for CS^+_A ($Z = 2.68$, $P = 0.004$, $U3_1 = 0.67$, Wilcoxon signed-rank test, one-tailed) during POST, without evidence for non-zero differences in PRE (all Zs < 1.77, Ps > 0.077, $U3_1 < 0.62$, Wilcoxon signed-rank tests, two-tailed). Bar plots represent sample means. Error bars indicate standard errors of the means. Asterisks indicate P-values < 0.05, plus signs represent P-values > 0.05 and < 0.10. Color bars indicate Z-values. Source data are provided as a Source Data file.

As in the first three behavioral experiments, participants first learned the six CS–US associations during Pavlovian learning. Following Pavlovian learning, we administered two RS blocks, one immediately before (PRE) and one immediately after (POST) the choice-induced revaluation phase, where participants ($N = 42$) made binary choices between CS^0_A and CS^+_A. For the RS effects during POST we had directed hypotheses (increased RS for the previously chosen, and decreased RS for the previously unchosen CS). Therefore, we used one-tailed post hoc tests. For the PRE phase, as well as for the control RS effects of CS^-_A relative to CS^-_B, there was no such directed hypothesis and we used two-tailed tests accordingly. Consistent with our hypothesis, we observed both a decrease in RS for $CS^0_A - US^0$ relative to $CS^0_B - US^0$, and an increase in RS between $CS^+_A - US^+$ compared to $CS^+_B - US^+$ during POST but not during PRE ($Z = 2.53$, $P = 0.006$, $U_3 = 0.67$, Wilcoxon signed-rank test, one-tailed) in the left hippocampus (Fig. 2a). Detailed analyses showed that this effect was driven by dissociable effects (interaction effect CS value and time (PRE or POST), $F_{1,\ 41} = 4.51$, $P = 0.040$, $\eta_p^2 = 0.10$, $1\text{-}\beta = 0.99$, rmANOVA): we found decreased RS for CS^0_A ($Z = 2.26$, $P = 0.012$, $U3_1 = 0.67$, one-sample Wilcoxon signed-rank test, one-tailed) and increased RS for CS^+_A ($Z = 2.26$, $P = 0.012$, $U3_1 = 0.69$, one-sample Wilcoxon signed-rank test, one-tailed) during POST, without any evidence for non-zero differences in PRE (all Zs < 1.05, Ps > 0.296, $U3_1 < 0.62$, one-sample Wilcoxon signed-rank tests, two-tailed) or for the control contrast of either CS^-_A or CS^-_B (all Zs < 0.31, Ps > 0.760, $U3_1 < 0.55$, one-sample Wilcoxon signed-rank tests, two-tailed, Fig. 2d). These effects arose from (numerically) reduced RS elicited by CS^0_A and increased RS for CS^+_A in separate analyses for $CS^0_A - US^0$ and $CS^+_A - US^+$ (Supplementary Fig. 3, Supplementary Notes 2 and 3). Thus, CS choices during choice-induced revaluation increased, whereas non-choices decreased hippocampal CS–US associative strength. However, while there was evidence for significant PRE–POST reduction in RS for CS^0_A ($Z = 1.84$, $P = 0.033$, $U_3 = 0.69$, Wilcoxon signed-rank test, one-tailed), the PRE–POST increase

in RS for CS^{+}_{A} was not significant ($Z = 1.48$, $P = 0.070$, $U_3 = 0.62$, Wilcoxon signed-rank test, one-tailed).

Furthermore, we found decreased RS for $CS^{0}_{A} - US^{0}$ relative to $CS^{0}_{B} - US^{0}$ and increased RS for $CS^{+}_{A} - US^{+}$ relative to $CS^{+}_{B} - US^{+}$ in the right lateral orbitofrontal cortex that survived small-volume correction (Fig. 2b). Extraction of parameter estimates from this cluster using an independent region of interest[28] revealed that the interaction effect (CS value and time (PRE or POST), $F_{1,\ 41} = 5.57$, $P = 0.023$, $\eta^2_p = 0.12$, $1–\beta = 0.99$, rmA-NOVA) was driven by a significant PRE–POST reduction in RS for CS^{0}_{A} (Fig. 2e, $Z = 1.77$, $P = 0.039$, $U_3 = 0.55$, Wilcoxon signed-rank test, one-tailed) but not by a PRE–POST increase in RS CS^{+}_{A} ($Z = 0.65$, $P = 0.260$, $U_3 = 0.50$, Wilcoxon signed-rank test, one-tailed). However, we found decreased RS for CS^{0}_{A} ($Z = 1.74$, $P = 0.040$, $U3_1 = 0.55$, Wilcoxon signed-rank test, one-tailed) and increased RS for CS^{+}_{A} ($Z = 2.68$, $P = 0.004$, $U3_1 = 0.67$, Wilcoxon signed-rank test, one-tailed) during POST, without evidence for non-zero RS in PRE (all $Zs < 1.77$, $Ps > 0.077$, $U3_1 < 0.62$, Wilcoxon signed-rank tests, two-tailed).

Exploratory analyses at a lenient, uncorrected threshold ($Z > 2.8$, uncorrected) yielded clusters in the left ventral tegmental area (VTA) and the right nucleus accumbens (NAcc, Fig. 2c). Both effects were driven by significantly reduced RS for CS^{0}_{A} (VTA: $Z = 1.81$, $P = 0.035$, $U_3 = 0.79$; NAcc: $Z = 2.38$, $P = 0.009$, $U_3 = 0.64$, Wilcoxon signed-rank tests, one-tailed), but only NAcc showed evidence of increased RS for CS^{+}_{A} (VTA: $Z = 0.41$, $P = 0.340$, $U_3 = 0.55$; NAcc: $Z = 1.53$, $P = 0.064$, $U_3 = 0.64$, Wilcoxon signed-rank tests, one-tailed). Overall, these RS results suggest that decisions during choice-induced revaluation had clearly dissociable effects on the neural representation of previously learned CS–US associations: While the previously chosen CS exhibited increased associative strength to its related US, the exact opposite effect was true for the previously unchosen CS. Importantly, the observed dissociation of choice-induced increase of RS effects for CS^{+}_{A} and decrease of RS effects for CS^{0}_{A} and the absence of PRE–POST differences of RS effects for the CS^{-}_{A} relative to CS^{-}_{B} cannot be explained by general extinction effects resulting from exposition to CS–US associations other than the initially learned associations. Extinction would imply equidirectional PRE–POST changes of all CS–US associations, which is incompatible with the observed results.

Choice-induced decrease of multivariate neural pattern similarity. Complementary to the mass-univariate RS-based approach, we performed multivariate fMRI analyses, employing a neural pattern similarity analysis[22] in the left hippocampus and right lateral OFC. Using the same logic as for the RS-based analyses, we reasoned that presentation of a CS would activate neural ensembles representing the associated US. This mnemonic pre-activation should not only be present in trials where the CS was followed by the originally learned US, but also in trials where the CS was followed by any of the other two possible, but not associatively linked US. Similarity of neural patterns related to two CS from the same value category could thus be indicative of associative memory retrieval of a US representation. According to the idea of choice-induced weakening of CS^{0}_{A} association with US^{0} and strengthening of CS^{+}_{A} association with US^{+}, our hypothesis therefore was that neural pattern similarity between same-value stimulus–outcome pairs ($CS^{0}_{A} - US^{-}/CS^{0}_{B} - US^{-}$ and $CS^{0}_{A} - US^{+}/CS^{0}_{B} - US^{+}$; $CS^{+}_{A} - US^{-}/CS^{+}_{B} - US^{-}$ and $CS^{+}_{A} - US^{0}/CS^{+}_{B} - US^{0}$) should decrease from PRE to POST, indicating less similarity between patterns of interest (i.e. the weakened/strengthened CS and the respective same-value CS, see Methods section for a detailed description). Therefore, we used one-tailed tests accordingly. For the pairs of control stimuli ($CS^{-}_{A} - US^{0}/CS^{-}_{B} - US^{0}$ and $CS^{-}_{A} - US^{+}/CS^{-}_{B} - US^{+}$), we did not expect changes in neural pattern similarity and thus employed two-tailed tests.

In the left hippocampus ROI, we observed negative PRE–POST change in neural pattern similarity when averaging across all patterns of interest ($t_{41} = 2.09$, $P = 0.021$, $U3_1 = 0.64$, $1–\beta = 0.63$, one-sample t-test, one-tailed) and for CS^{+}_{A}/CS^{+}_{B} pairs ($t_{41} = 1.81$, $P = 0.039$, $U3_1 = 0.57$, $1–\beta = 0.53$, one-sample t-test, one-tailed), but only a numerically decreased neural pattern similarity from PRE to POST for CS^{0}_{A}/CS^{0}_{B} pairs ($t_{41} = 1.01$, $P = 0.144$, $U3_1 = 0.50$, $1–\beta = 0.28$, one-sample t-test, one-tailed). Importantly, change of neural pattern similarity for the control stimulus pairs CS^{-}_{A}/CS^{-}_{B} was positive and not significant ($t_{41} = 0.76$, $P = 0.451$, $U3_1 = 0.57$, $1–\beta = 0.12$, one-sample t-test, two-tailed, Fig. 3a).

In the right lOFC ROI, we observed qualitatively similar results as in the left hippocampus: There was significant negative PRE–POST change in neural pattern similarity when averaging across all patterns of interest ($t_{41} = 1.70$, $P = 0.049$, $U3_1 = 0.62$, $1–\beta = 0.40$, one-sample t-test, one-tailed). However, there was no evidence of significant change in pattern similarity for CS^{+}_{A}/CS^{+}_{B} pairs ($t_{41} = 1.23$, $P = 0.113$, $U3_1 = 0.64$, $1–\beta = 0.23$, one-sample t-test, one-tailed). There was also no evidence for decreased neural pattern similarity from PRE to POST for CS^{0}_{A}/CS^{0}_{B} pairs ($t_{41} = 1.55$, $P = 0.061$, $U3_1 = 0.62$, $1–\beta = 0.13$, one-sample t-test, one-tailed). Only descriptively, both CS^{+}_{A}/CS^{+}_{B} pairs and CS^{0}_{A}/CS^{0}_{B} pairs became less similar from PRE to POST. The change of neural pattern similarity for the control stimulus pairs CS^{-}_{A}/CS^{-}_{B} was positive, but not significantly different from 0 ($t_{41} = 0.04$, $P = 0.969$, $U3_1 = 0.43$, $1–\beta = 0.05$, one-sample t-test, two-tailed, Fig. 3b).

Taken together, these multivariate results conceptually confirm the findings from the mass-univariate RS-based analyses and further support the interpretation that the observed choice effects could be explained by choice-induced changes of associative strength. However, these results should be interpreted with caution, as power was generally low ($1–\beta < 0.80$), most likely resulting from the reduced number of trials included in the analyses (40 trials per CS pair in PRE and POST). Additionally, unlike our RS-based results, changes in neural pattern similarity do not allow to infer the directionality of the effects (i.e. patterns may become more dissimilar both due to strengthening or weakening of the associative trace). Nevertheless, the results from both sets of analyses provide convergent evidence for our hypothesis that revaluation choices changed the degree to which neural US representations were pre-activated by their associated CS.

Hippocampal CS–US RS correlates with future choices. We next investigated whether the observed choice-induced modifications of hippocampal CS–US RS were correlated with choice biases during the probe phase. As in Experiments 1 and 2, we had directed hypotheses for the choice effects and thus used one-tailed post-hoc tests.

Unlike in Experiment 1, CS^{0}_{A} (unchosen stimulus during revaluation) was not chosen less likely than CS^{0}_{B} in the decision probe ($Z = 1.01$, $P = 0.844$, $U_3 = 0.55$, Wilcoxon signed-rank test, one-tailed, Fig. 1h) in Experiment 5. Relatedly, there was no evidence for preference differences in within-category choice trials directly comparing CS^{0}_{A} and CS^{0}_{B} ($Z = 1.07$, $P = 0.857$, $U3_1 = 0.62$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2h). To assess PRE–POST changes of associative strength, participants had to be re-exposed to the initially learned CS–US associations and were explicitly instructed to judge whether the presented CS–US associations were correct.

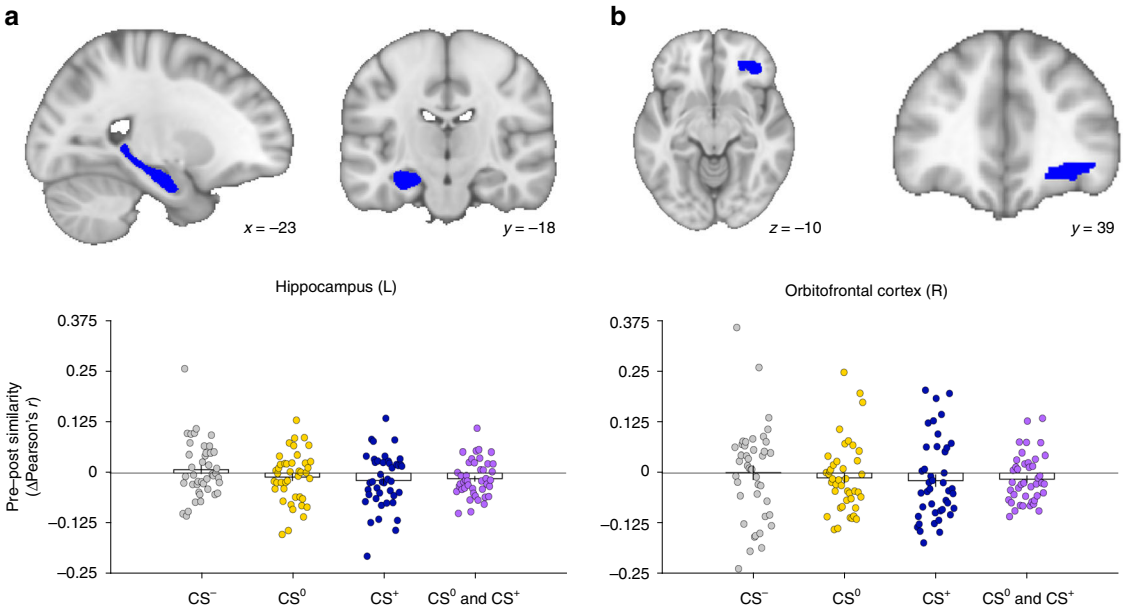


Fig. 3 Multivariate neural pattern similarity analysis results. a Multivariate neural pattern similarity analyses in the left hippocampus and **b** right lateral OFC, showing that neural similarity (POST *r* - PRE *r*, ΔPearson's *r*) between same value stimulus-outcome pairs decreased from PRE to POST ($N = 42$). CS^0 (yellow): $CS^0_A - US^-/CS^0_B - US^-$ and $CS^0_A - US^+/CS^0_B - US^+$ (Hippocampus: $t_{41} = 1.01$, $P = 0.144$, $U3_1 = 0.50$, 1-$\beta = 0.28$, one-sample *t*-test, one-tailed; lateral OFC: $t_{41} = 1.55$, $P = 0.061$, $U3_1 = 0.62$, 1-$\beta = 0.13$, one-sample *t*-test, one-tailed); CS^+ (blue): $CS^+_A - US^-/CS^+_B - US^-$ and $CS^+_A - US^0/CS^+_B - US^0$ (Hippocampus: $t_{41} = 1.81$, $P = 0.039$, $U3_1 = 0.57$, 1-$\beta = 0.53$, one-sample *t*-test, one-tailed; lateral OFC: $t_{41} = 1.23$, $P = 0.113$, $U3_1 = 0.64$, 1-$\beta = 0.23$, one-sample *t*-test, one-tailed); CS^0 and CS^+ (purple): average of CS^0 and CS^+ (Hippocampus: $t_{41} = 2.09$, $P = 0.021$, $U3_1 = 0.64$, 1-$\beta = 0.63$, one-sample *t*-test, one-tailed; lateral OFC: $t_{41} = 1.70$, $P = 0.049$, $U3_1 = 0.62$, 1-$\beta = 0.40$, one-sample *t*-test, one-tailed). Importantly, change of neural pattern similarity for the control stimulus pairs CS^-_A/CS^-_B (CS^-, gray) was numerically positive (Hippocampus: $t_{41} = 0.76$, $P = 0.451$, $U3_1 = 0.57$, 1-$\beta = 0.12$, one-sample *t*-test, two-tailed; lateral OFC: $t_{41} = 0.04$, $P = 0.969$, $U3_1 = 0.43$, 1-$\beta = 0.05$, one-sample *t*-test, two-tailed), indicating higher similarity from PRE to POST. Bar plots represent sample means. Error bars indicate standard errors of the mean. Source data are provided as a Source Data file.

It is well established that restudying of memorized material reverses retrieval-induced forgetting effects[9,29]. Thus, the observed behavioral null effect for CS^0_A might be due to re-exposure and restudy of the original CS–US association, allowing the weakened association between $CS^0_A - US^0$ to regain its original associative strength.

However, replicating Experiment 1, we observed a choice-induced increase in preference for CS^+_A over CS^+_B ($Z = 3.03$, $P = 0.001$, $U_3 = 0.76$, Wilcoxon signed-rank test, one-tailed, Fig. 1h). This effect was mainly driven by preference for CS^+_A in within-category choices contrasting CS^+_A and CS^+_B ($Z = 1.93$, $P = 0.027$, $U3_1 = 0.62$, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; Supplementary Fig. 2h). We therefore focused on this effect in brain–behavior correlations. We hypothesized a positive linear relationship between the difference between choice probabilities of CS^+_A and CS^+_B and the magnitude of hippocampal RS between $CS^+_A - US^+$ and $CS^+_B - US^+$), and thus tested the Spearman correlation coefficient one-tailed. The difference of hippocampal RS between $CS^+_A - US^+$ and $CS^+_B - US^+$ was positively correlated with the difference between choice probabilities of CS^+_A and CS^+_B ($\rho_{40} = 0.31$, $P = 0.024$, one-tailed; Fig. 4b). The more hippocampal representations of the $CS^+_A - US^+$ association had been strengthened by choices during revaluation, the more likely participants were to select CS^+_A compared to its non-revalued partner stimulus CS^+_B.

Consistently, in a whole-brain analysis we observed a positive relationship between the difference in choice preference for CS^+_A versus CS^+_B and changes in CS–US RS from PRE to POST in the left posterior hippocampus, extending to occipito-temporal complex (Fig. 4c). A similar whole-brain analysis using only the choices between CS^+_A and CS^+_B as behavioral covariate revealed areas in the bilateral anterior insula and orbitofrontal cortex (Supplementary Fig. 5b). Neither VTA, NAcc, nor the cluster in the lateral orbitofrontal cortex showed relationships with probe phase behavior.

An alternative explanation of our results is based on cached values, a possibility that we address in the Supplementary Methods, Supplementary Notes 1 and 4, and Supplementary Fig. 4.

Discussion

Using a carefully designed paradigm, we show that decisions bias future choices, even without participants directly experiencing the outcomes of their decisions. Participants were more likely to select CS they had previously chosen, and less likely to select CS they had not chosen, compared to equivalent CS. At the neural level, we found that choices induced alterations to hippocampal and orbitofrontal representations of stimulus–outcome associations that were correlated with future decisions.

The idea that past decisions bias preferences was put forth decades ago[1,4] and evidence for post-decision revaluation has accumulated since[1–3,6] (see ref. [30] for critical discussion). Here, we present behavioral evidence that reward-predictive CS that were chosen in the past are more likely to be selected during future choices, compared to CS of equal value that were not presented. Conversely, we found decreased preferences for CS that were not chosen in the past, compared to equivalent CS, indicating bidirectional effects of choice-induced revaluation. Most importantly, our behavioral findings are independent of rewards, as participants never experienced the outcomes of their

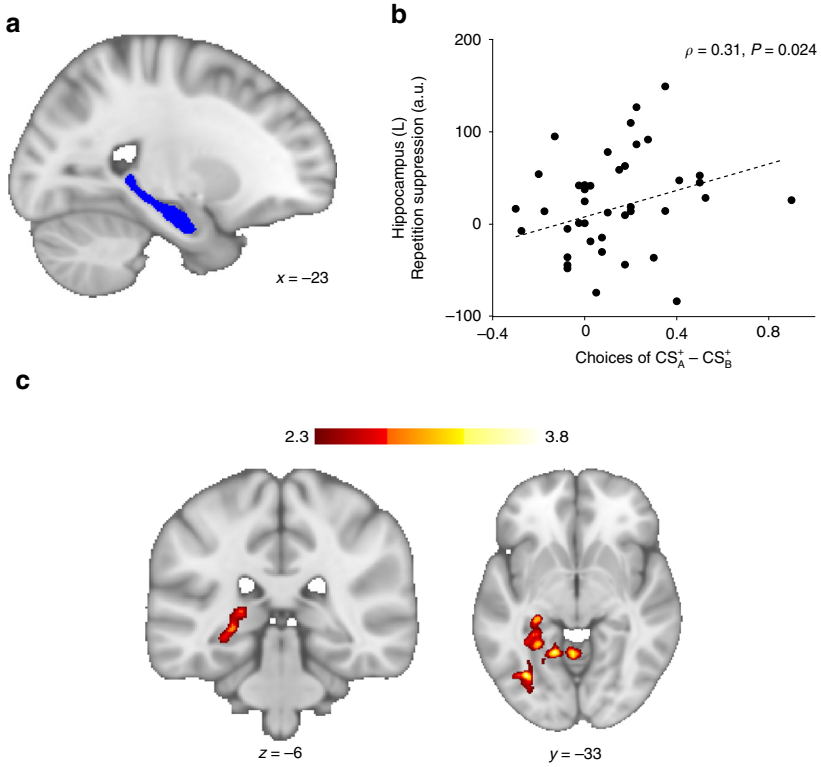


Fig. 4 Brain–behavior correlation and whole-brain regressions. b Left hippocampus repetition suppression of $CS^{+}_{A} - US^{+}$ relative to $CS^{+}_{B} - US^{+}$, controlling for activation elicited by CS^{+}_{A} and CS^{+}_{B} followed by both incorrect outcomes (US^{-} and US^{0}) (Eq. (2)), extracted from an independent anatomical mask (**a**) during POST was positively related to the decision probe difference between overall choice probability of CS^{+}_{A} and overall choice probability of CS^{+}_{B} ($N = 42$, $\rho_{40} = 0.31$, $P = 0.024$, Spearman correlation, one-tailed). With increasing repetition suppression, participants were more likely to prefer CS^{+}_{A} over CS^{+}_{B}. **c** Whole-brain regression ($N = 42$) showing a positive relationship between the difference of CS^{+}_{A} and CS^{+}_{B} preference and PRE–POST change of repetition suppression in the left posterior hippocampus. With more positive PRE–POST repetition suppression change, participants are more likely to prefer CS^{+}_{A} over CS^{+}_{B}. Color bar indicates *Z*-values. Source data are provided as a Source Data file.

choices. This suggests that the observed effect could arise from associative memory mechanisms, as also indicated by an additional control experiment designed to rule out the alternative explanation that the observed choice biases could result from learned choice heuristics. Additionally, computational models implementing differentially updated CS–US associative strength of chosen and unchosen options based on fictive prediction errors explained our data best.

Our results are in line with previous reports of choice-induced preference changes[1,3,4,6] and conceptually replicate studies showing changes in stimulus valuation by cued approach training (CAT)[5,31–33]. Similar to the present approach, performance of a button press (go response) upon presentation of go stimuli during CAT induces long-term[31] non-reinforced changes of desirability of go stimuli over no-go stimuli. CAT effects are independent of initial value of the stimuli and rely on integrity[33] and activation[5,31] of ventromedial prefrontal cortex, and interactions between orbitofrontal cortex and ventral striatum[31]. Importantly, the results of Experiment 4 suggest that choice-induced revaluation effects, at least for previously chosen options, seem to go beyond a trained action tendency or choice rule (go response), as observed in CAT.

As most previous studies have presented participants directly with the choice outcomes and thereby confounded contributions of memory and choice mechanisms, our study might be the first to provide evidence for a value-independent associative memory mechanism driving choice-induced preference changes. Furthermore, although associative memory dynamics are a likely candidate mechanism of choice-induced preference changes, the exact neural mechanisms underlying this phenomenon have remained unknown. Here, we provide evidence that past choices bias future decision-making, in part by modifying the strength of neural stimulus–outcome associations. The present results suggest that, during decision-making, reactivation of stimulus–outcome associations[10,11,14,34], and making a choice renders them subject to nonmonotonic plasticity[16], with the association of the chosen stimulus being strengthened, and the association of the unchosen stimulus being diminished. Since both chosen and unchosen CS activate neural populations representing the respective-associated outcome[10–15], we reason that the observed opposing decision biases are presumably related to additional choice-induced activation of the chosen CS–US association, and absence of such choice-induced activation of the unchosen CS–US associations. Our results suggest that choices can act as self-generated teaching signals[18,35], dynamically altering stimulus–outcome associations stored in memory. However, it should be noted that we did not observe behavioral evidence for choice-induced weakening of the unchosen CS–US association in Experiment 5. This might be due to re-exposure to the initially learned CS–US associations during the POST fMRI run, which presumably allowed the weakened association between $CS^{0}_{A} - US^{0}$ to regain its original associative strength, in line with studies showing that restudying of memorized material reverses retrieval-induced forgetting effects[9,29].

Even though our data provide evidence for choice-induced changes to associative strength, the current approach does not allow to dissociate which exact features of the US contribute to RS effects. As US value and identity are inextricably linked in our experiment, the observed effects could be related to changes in

CS-dependent pre-activation of identity-related or value-related features of the US. However, both identity-related and value-related features of the US are learned associatively in the present study and are therefore likely retrieved in an associative fashion to guide choices. Alternatively, the observed change in the hippocampal RS could be due to choice-induced alterations of CS representations per se. However, changes of CS representation cannot explain the specificity of RS signals to presentations of the learnt CS–US associations, and, perhaps more importantly, cannot account for the reversed directionality of RS effects depending on choice during revaluation.

Our results suggest that merely making a choice induces plasticity of associative representations in the hippocampus and lateral OFC. In line with the present results, a fronto-hippocampal network comprising hippocampus and lateral OFC seems to be critically involved in reward-related updating of stimulus–outcome associations[14,28]. Whereas the hippocampus has been found to support relational learning and memory processes, including value spread[34] and factorized replay of event trajectories[36], the lateral OFC has been additionally implicated in the resolution of credit-assignment problems in reinforcement learning[28,37].

Our study has at least three implications for current theories of decision-making. First, it proposes a memory-based account for choice-induced revaluation and, more broadly, choice history bias, two well-known, but still poorly understood decision-making phenomena. We show that choices are not only guided by associative representations of value stored in memory, but that decisions themselves dynamically transform associative memories. This suggests that relational structures constituting decision-makers' cognitive maps[38,39] can be distorted through their very own choice behavior. Second, we provide evidence for involvement of the hippocampus and lateral OFC in maintaining and updating of stimulus–outcome associations. This extends previous findings[10,14,28,37] by describing a functional role of the hippocampus in value-based decision-making that is independent of experienced reward. Such a role may more closely resemble naturalistic decision situations where consequences of choices often unravel at distant future time points, rendering credit assignment challenging[28,37]. Third, we provide a mechanism underlying seemingly irrational choice behavior: Even though participants chose between equivalent options, they were biased to prefer chosen, and to neglect non-selected options. The latter might have important implications for explaining subjective preferences, especially in consumer choice behavior[2] and in understanding why humans tend to make coherent decisions, even in conditions characterized by maladaptive choice behavior, such as substance dependence or obsessive compulsive disorder.

Taken together, both our behavioral and neural results support the key prediction that past choices bias future decision-making, partially by altering hippocampal and orbitofrontal representations of stimulus–outcome associations. Our study provides a memory-based mechanism for choice-induced preference change effects[1,3,4,6]. The present study shows that merely retrieving stimulus–outcome associations and making a choice is sufficient to induce plasticity in reward-predictive associations stored in memory.

Methods

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

Forty-nine young, healthy volunteers (age: $M = 23.93$, $SD = 2.90$ years, 18 males) participated in Experiment 1. Seven participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase that served as an attentional control (<10% responses in trials that required to indicate the color of the square surrounding the conditioned stimuli. Two additional participants had to be excluded due to not passing the manipulation check (high-value option selected <50% during choice-induced revaluation), thus leaving a total of $N = 40$ participants for final analyses.

Sixty-four young, healthy volunteers (age: $M = 23.47$, $SD = 3.79$ years, 26 males), participated in Experiment 2. Ten participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase, and 13 subjects were excluded due to not passing the manipulation check (intermediate valued option selected <50% during choice-induced revaluation), one additional participant had to be excluded due to a technical error, leaving $N = 40$ participants for statistical analyses.

Sixty-one young, healthy volunteers (age: $M = 23.26$, $SD = 3.27$ years, 23 males), participated in Experiment 3. Ten participants were excluded from statistical analyses due to lacking engagement in the cover task during the Pavlovian conditioning phase, and six subjects were excluded due to not passing the manipulation check (high-value option selected <50% during choice-induced revaluation), one additional participant had to be excluded due to a technical error (no data was recorded), leaving $N = 44$ participants for statistical analyses.

Fifty-two young, healthy volunteers (age: $M = 22.06$, $SD = 3.69$ years, 20 males), participated in Experiment 4. Twelve subjects were excluded due to not passing the manipulation check (high-value option selected <50% during choice-induced revaluation), leaving $N = 40$ participants for statistical analyses.

Fifty-eight young, healthy and magnetic resonance imaging (MRI)-compatible volunteers (age: $M = 24.61$, $SD = 4.01$ years, 30 males) participated in Experiment 5 (functional MRI (fMRI) experiment). One participant fell asleep during the POST revaluation fMRI-RS run, three participants discontinued the MRI acquisition (one due to claustrophobia, two reported a headache during task performance). Twelve additional subjects were excluded due to not passing the manipulation check (high-value option selected <50% during choice-induced revaluation), leaving $N = 42$ participants for statistical analyses.

Behavioral task—ratings. Participants received written instructions for the experiment and were instructed once again on the computer screen. The experiments were programmed in MATLAB 2012b (MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, MA, USA, v8.0.0.783), using Psychophysics Toolbox[40] (version 3) and MATLAB 2019a (v9.6.0.1072779). Before the task, participants rated 25 different sweet and high-caloric food items selected from an online database[24]. Subjects were instructed to indicate the subjective desirability of the food items by using the "y" and "m" button on a standard German (QWERTZ) computer keyboard to position a red slider bar on a white visual analog scale (VAS) between 0 (not liked) and 100 (very much liked). The lowest- (Experiment 1: $M = 17.5$, $SD = 20.51$; Experiment 2: $M = 16.65$, $SD = 16.29$; Experiment 3: $M = 18.59$, $SD = 22.15$; Experiment 5: $M = 20.41$, $SD = 20.44$), and highest-rated (Experiment 1: $M = 96.55$, $SD = 7.10$; Experiment 2: $M = 94.85$, $SD = 9.08$; Experiment 3: $M = 95.21$, $SD = 14.52$; Experiment 4: $M = 96.63$, $SD = 6.62$; Experiment 5: $M = 97.79$, $SD = 4.96$) food item as well as a food item rated with the median value of all ratings (Experiment 1: $M = 57.43$, $SD = 11.55$; Experiment 2: $M = 56.23$, $SD = 9.80$; Experiment 3: $M = 57.14$, $SD = 13.95$; Experiment 4: $M = 56.48$, $SD = 9.95$; Experiment 5: $M = 60.07$, $SD = 11.71$) were selected for Pavlovian conditioning. We explicitly decided for clearly differentiable pictures of food items in order to elicit activation of differential neural ensembles coding for those stimuli and to facilitate learning of vivid memories of stimulus–outcome associations. Next, subjects rated 20 (11 in Experiment 1) Japanese kanjis[23] according to subjective value/liking by using the "y" and "m" button on a standard computer keyboard to position a red slider bar on a VAS between 0 (not liked) and 100 (very much liked). The six kanjis rated closest to 50 (equivalent to neutral) were selected and their order was randomized before being associated with the food items in Pavlovian conditioning. The subjective values/liking of the six selected kanjis did not differ significantly from each other in Experiments 1, 3, 4, and 5 (main effects of stimulus: all Fs < 1.14, Ps > 0.344, η_p^2s < 0.03, rmANOVA). However, there was a main effect of stimulus in Experiment 2 ($F_{5,\ 195} = 2.65$, $P = 0.024$, $\eta_p^2 = 0.064$, rmANOVA), resulting from a significantly higher pre-rating for control stimulus CS_A^+ compared to control stimulus CS_B^+ ($t_{39} = 3.13$, $P = 0.003$, paired-samples t-test). More importantly, both critical pairs (CS^- and CS^0) did not differ significantly (all ts < 0.79, Ps > 0.437, paired-samples t-tests).

Behavioral task—Pavlovian conditioning. Participants learned to associate the selected kanjis (conditioned stimuli, CS) with differently valued outcomes (unconditioned stimuli, US) by repeatedly observing one CS (2000 ms), followed by an inter-stimulus interval (1000 ms) marked by a fixation cross, and presentation of one US (2000 ms). Each trial was separated by an inter-trial-interval (ITI) marked by a gray screen. The ITI per trial was drawn from a discretized γ-distribution

(shape = 5, scale = 0.9) truncated for an effective range of values between 3000 and 8000 ms. Across all CS, CS–US couplings were interleaved with 20% CS–no-US couplings. This was intended to create an association that would not rapidly extinguish due to extinction effects related to not being presented with the associated US during the subsequent choice phases. Each US was associated with two different CS, resulting in three pairs of differently valued CS: $CS^{+}_{A/B}$, $CS^{0}_{A/B}$ and $CS^{-}_{A/B}$. Participants completed 180 trials (19 of the participants included in Experiment 1 had performed 240 trials) of Pavlovian conditioning, 30 (40, respectively) trials per CS–US association. In Experiment 4, participants performed 160 trials and learned associations between four neutrally rated CS and an intermediate value outcome and a high value outcome (40 trials per CS–US association), resulting in two pairs of differently valued CS: CS^{+}_{80}, CS^{+}_{20} and CS^{0}_{80}, CS^{0}_{20}. Importantly, one of the two CS in each pair was followed by the outcome with a probability of 80% (CS^{+}_{80} and CS^{0}_{80}) while the other CS was followed by the US in 20% (CS^{+}_{20} and CS^{0}_{20}) of the trials. The participants were presented with a cover story for the experiment: They were told to imagine themselves in preparation for a journey to Japan during which they would need to learn kanjis associated with certain food items to be ready to select their favorite sweets in Japanese shops. As an attentional control, we introduced a simple binary classification task, presenting participants with a red or blue square surrounding the CS. Each CS was equally often presented with a red or blue square and color of the square did not predict US contingency. Subjects were instructed to react as quickly and as correctly as possible by pressing "y" upon seeing a blue square and pressing "m" upon seeing a red square surrounding the CS. Pavlovian conditioning was split up in three blocks of 60 (80, respectively) trials, interleaved with self-paced breaks. In Experiment 5, both ratings and Pavlovian conditioning were performed outside the MRI scanner.

Due to a coding error in the script creating pseudo-randomized reward schedules in Experiment 1, 2, and 3 (which we spotted during setup of Experiment 4), CS were not followed by outcomes in exactly 80% of trials equivalently across all CS. For each experiment, we had set up five different reward schedules, assigning different outcome probabilities to each CS. In Experiment 1, participants received the following outcome probabilities on average: CS^{-}_{A} (Probability: $M = 0.80$, range = 0.70–0.87), CS^{-}_{B} ($M = 0.78$, range = 0.73–0.88), CS^{0}_{A} ($M = 0.78$, range = 0.70–0.87), CS^{0}_{B} ($M = 0.78$, range = 0.63–0.90), CS^{+}_{A} ($M = 0.80$, range = 0.70–0.90), CS^{+}_{B} ($M = 0.86$, range = 0.73–0.93). In Experiment 2, participants received the following outcome probabilities on average: CS^{-}_{A} ($M = 0.80$, range = 0.73–0.87), CS^{-}_{B} ($M = 0.76$, range = 0.73–0.83), CS^{0}_{A} ($M = 0.78$, range = 0.70–0.87), CS^{0}_{B} ($M = 0.78$, range = 0.63–0.90), CS^{+}_{A} ($M = 0.80$, range = 0.70–0.90), CS^{+}_{B} ($M = 0.88$, range = 0.83–0.93). In Experiment 3, participants received the following outcome probabilities on average: CS^{-}_{A} ($M = 0.80$, range = 0.73–0.87), CS^{-}_{B} ($M = 0.77$, range = 0.73–0.83), CS^{0}_{A} ($M = 0.79$, range = 0.70–0.87), CS^{0}_{B} ($M = 0.77$, range = 0.63–0.90), CS^{+}_{A} ($M = 0.79$, range = 0.70–0.90), CS^{+}_{B} ($M = 0.88$, range = 0.83–0.93). It should be noted that these minor differences in outcome probabilities between CS cannot account for the observed choice-induced revaluation effects, as the respective chosen CS on average received less outcomes (CS^{+}_{A} in Experiment 1) than or an equal number of outcomes (CS^{0}_{A} in Experiment 2) as their same-valued partner stimuli. Contrarily, the respective unchosen CS on average received more outcomes (CS^{-}_{A} in Experiment 2) than or an equal number of outcomes (CS^{0}_{A} in Experiment 1) as their same-valued partner stimuli. Thus, the outcome probabilities assigned to each CS would have worked against the hypothesized effects. Consistently, there were no significant correlations between the outcome probability during Pavlovian conditioning and decision probe overall or pairwise within-category choice probability ($\rho s < 0.29$, $Ps > 0.067$, Spearman correlations, two-tailed). Importantly, Experiment 5 was not affected from this error, as reward schedules were created with a different script in which we correctly coded that each CS would be followed by an outcome in 80% of the trials.

Behavioral task—choice-induced revaluation. After completion of Pavlovian conditioning, participants were presented with repeated choices (28 trials) between a CS^{+}_{A} versus a CS^{0}_{A} (Experiments 1 and 5), CS^{0}_{A} versus a CS^{-}_{A} (Experiment 2), a CS^{0}_{A} versus either a CS^{-}_{A} (14 trials) or a CS^{+}_{A} (14 trials) (Experiment 3), or a CS^{+}_{80} versus a CS^{0}_{80} and a CS^{+}_{20} versus a CS^{0}_{20} (Experiments 4), interleaved with lure decisions (28 trials) between four other neutrally rated kanjis that had never been presented during Pavlovian learning and thus were not associated with any of the US. The choice-induced revaluation phase served as the crucial manipulation in all experiments and was systematically varied across studies. Choice probability (CP) for the high-value CS served as a control for learning and as a manipulation check. Only participants selecting the higher valued CS more than 50% (CP ≥ 0.50) were included in the final analysis, as we reasoned that choice-induced revaluation choices would (1) represent a marker of having learned the true associative values of the CS, (2) be a measure for learning, independent of the actual decision probe phase data (avoiding biased and arbitrary decisions for exclusion of participants), and (3) allow us to exclude decision makers showing random, or arbitrary choice behavior. Choice options were presented for 1500 ms and the chosen option was highlighted by a gray square surrounding the chosen CS. If participants did not respond within the time-window, a time-out message was displayed, and the respective trial was repeated at the end of the choice-induced revaluation phase. Order (left/right) of choice options was counterbalanced to avoid simple response patterns or decision rules (e.g. "always press left"). Participants were instructed to imagine themselves in a Japanese shop, where they would like to buy their favorite food items based on the previously learned kanjis (CS). Participants were told that one of the choice trials would randomly be drawn and their choice would determine which food item (US) they would receive as a bonus upon completion of the experiment. Participants selected choice options by pressing the "y" (left option) or "m" (right option) button (left or right index finger on an MRI-compatible response box in Experiment 5). Importantly, participants were not presented with the US related to their chosen or unchosen CS to dissociate the observed effects from outcome-related relearning of CS–US associations. We assumed that presentation of a CS would pre-activate neural ensembles coding for the associated US. Consequently, we expected that choosing a CS would induce strengthening of the chosen option's CS–US association, whereas not choosing a CS would weaken the unchosen option's CS–US association.

Behavioral task—decision probe. Following choice-induced revaluation, participants were presented with repeated binary choices (120 trials) between all possible CS combinations to assess CS preferences. Every CS combination was presented eight times in pseudo-random order. In Experiment 4, participants made choices between the two same-value pairs of stimuli that were differently strong associated to their respective outcomes (CS^{+}_{80} versus CS^{+}_{20} and CS^{0}_{80} versus CS^{0}_{20}). Here, every CS pair was presented 10 times in pseudo-random order. Choice options were presented for 1500 ms. If participants did not respond within this time-window, a time-out message was displayed, and the respective trial was repeated at the end of the decision probe phase. Participants selected choice options by pressing the "y" (left option) or "m" (right option) button (left or right index finger on MR-compatible response box in Experiment 5). Order (left/right) of choice options was counterbalanced. Participants were instructed that their shopping bag was torn, and they had to return to the shop for buying their favorite food items based on the previously learned kanjis (CS). Again, participants were told that one of the choice trials would randomly be drawn and their choice would determine which food item (US) they would receive as a bonus upon completion of the experiment. Importantly, participants were again not presented with the US related to their chosen or unchosen CS.

fMRI-RS task (Experiment 5). After Pavlovian conditioning outside the MRI-scanner, we administered two fMRI-RS blocks, one immediately before (PRE) and one immediately after (POST) the revaluation phase to assess choice-induced effects of fMRI-RS. Every possible combination of CS and US (18 combinations) was presented 20 times each (360 trials in total). In one-third of the trials, the originally learned CS–US associations were presented, the remaining two-third of trials contained incorrect CS–US associations. In every trial, a CS was presented for 700 ms, followed by an interstimulus interval (fixation cross) for 400 ms and a US for 700 ms. The intertrial interval was drawn from a discretized γ-distribution (shape = 2.01, scale = 1), truncated for an effective range of values between 2000 and 6000 ms. Order of trials was pseudo-random, between-trial repetition of CS or US did not occur. Additionally, every batch of 18 trials contained every possible combination of CS–US association to avoid comparison of temporally distal trials and between-trial biases in RS introduced by, e.g. fluctuations of attention, "novelty" or surprise. During both runs of fMRI-RS, participants performed an attentional control task. After a pseudo-random 20% of trials, participants were presented with probe trials in which they were asked to indicate whether or not the previously seen CS–US-association matched the true CS–US-association learned during Pavlovian conditioning via button presses with their right and left index fingers on an MRI-compatible response box. Correct responses were rewarded with 0.05€ and incorrect responses or time-out trials (without a response by the participant within 2500 ms after onset of the probe trial) resulted in a 0.05€ penalty which would be summed up as a bonus upon completion of the experiment. On average, participants earned a bonus of 5.93€ ($SD = 1.05$). Performance during the attentional control task was generally high (overall probability of correct answers, excluding time-out trials: $M = 0.92$, $SD = 0.06$ ($t_{41} = 41.54$, $P < 0.001$, one-sample t-test vs. chance level (0.5)), with no evidence for a difference in performance between PRE and POST choice-induced revaluation run (PRE: $M = 0.92$, $SD = 0.07$; POST: $M = 0.92$, $SD = 0.072$; $t_{41} = 0.06$, $P = 0.95$, paired-samples t-test).

fMRI acquisition. Two runs of fMRI were recorded with a 3 Tesla Siemens PRISMA MR-system (Siemens, Erlangen, Germany), using a 64-channel head coil. Blood oxygenation level-dependent (BOLD) signals were acquired using a multi-band accelerated T2*-weighted echo-planar imaging (EPI) sequence (multi-band acceleration factor 2, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle = 80°, field of view (FoV) = 220 mm, voxel size = 2.2 × 2.2 × 2.2 mm, no gap). Per volume, 66 slices covering the whole brain, tilted by ~15° in z-direction relative to the anterior–posterior commissure plane were acquired in interleaved order. The first five volumes of the functional imaging time series were automatically discarded to allow for T1 saturation. After each run, a $\mathbf{B_0}$ magnitude and phase map was acquired to estimate field maps and $\mathbf{B_0}$ 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 the PRE choice-induced revaluation

fMRI-RS run, a high-resolution three-dimensional T1-weighted anatomical map (TR = 2500 ms, TE = 2.82 ms, FoV = 256 mm, flip angle = 7°, voxel size = 1 × 1 × 1 mm, 192 slices) covering the whole brain was obtained using a magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence. This scan was used as anatomical reference to the EPI data during the registration procedure.

Behavioral analyses. Data were analyzed in MATLAB 2012b (v8.0.0.783), 2017a (v9.2.0.556344), and 2019a (v9.6.0.1072779) (The MathWorks, Inc., Natick, MA, USA) using custom analysis scripts. For the manipulation check, indicating learning of the CS–US-associations, average choice probabilities (CPs) for the higher valued CS were calculated (one average in Experiments 1, 2, and 5, two averages in Experiments 3 and 4) by summing up choices of the higher-valued CS and dividing by the number of choice trials with a recorded response. This CP was compared against the inclusion criterion of CP ≥ 0.50. If participants had chosen the high-valued CS in more than 50% (or in exactly 50%) of the trials of choice-induced revaluation, they were included in the final analyses. For the decision probe phase, we computed an average overall CP per CS per subject including all binary decisions in which the respective CS was present (count data, 1 representing selection of the CS, 0 representing selection of the alternative CS). Lacking a non-parametric alternative to the parametric two-way repeated-measures analysis of variances (rmANOVA), distributions of mean overall CPs were analyzed at the group level with a rmANOVA with factors CS valence (3) and stimulus type (2) and post-hoc Wilcoxon sign-rank tests for paired samples, focusing on the pairwise comparison of stimulus types within each level of valence. We hypothesized a main effect of CS valence and an interaction effect of CS valence × CS type (A vs. B), resulting from higher CPs for previously chosen CS relative to the equivalent control CS and lower CPs for previously unchosen CS relative to the equivalent control CS (Experiments 1, 2, and 5). However, we expected absence/no evidence for such an interaction effect in Experiment 3 in which a CS^0_A was chosen and unchosen equally often (resulting in no change of the preference relative to the control CS). Additionally, we performed one-sample Wilcoxon sign-rank tests of overall CP against chance level (CP = 0.50). As an alternative measure of choice preference in addition to overall CP, we also computed pairwise choice probabilities by comparing choice ratios of the two CS within each valence category with one-sample Wilcoxon sign-rank against chance level (CP = 0.50).

Experiment 4 was specifically designed to rule out the possibility that participants did not guide their choices based on (altered) associative strength between CS and US but had simply learned choice rules for the two CS presented during the revaluation phase (go tag for chosen CS, no go tag for unchosen CS). We thus aimed to orthogonalize contributions of associative strength and choice rule, by assigning go tags to the two chosen CS^+ that differed in their associative strength to US^+ and no-go tags to the two unchosen CS^0 that differed in their associative strength to US^0. Our hypothesis was that if choice behavior was exclusively driven by these tags participants had learned, go tags for both chosen CS^+_{80} and CS^+_{20} and no-go tags for both unchosen CS^0_{80} and CS^0_{20}, there should be no evidence for both same-value pairwise choice probabilities different from chance level (CP = 0.50). However, if the choices were made based on the learned associations and the associative strengthening/weakening of the memory trace between CS and US, there should be a significantly increased choice probability for CS^0_{80}/CS^+_{80} that were more strongly associated with their respective outcomes.

As we had formulated directional hypotheses for the choice effects, we performed one-tailed (post-hoc) tests. In Experiment 3, there were no directional hypotheses for the choice effects, therefore, we used two-tailed post-hoc tests. We report effect sizes η_p^2 for rmANOVAs, Cohen's U_3 for Wilcoxon signed-rank tests and Cohen's $U3_1$ for one-sample Wilcoxon signed-rank tests (range: 0–1, 0.5 indicating no effect), calculated in the MATLAB-based Measures-of-Effect-Size-toolbox[41]. Based on the reported effect sizes η_p^2 we additionally indicate post-hoc achieved power ($1-\beta$) for the hypothesized interaction effects of CS valence × CS type in rmANOVAs across behavioral analyses in Experiments 1, 2, 3, and 5. Based on the means and standard deviations for both choice probabilities, we indicate post-hoc achieved power for Experiment 4. All power analyses were conducted in G*Power[42,43] (v3.1.9.2).

Univariate fMRI data analysis. We exploited fMRI-RS effects (rapid, repeated presentation of the same stimulus or pre-activation of a stimulus by associated stimuli elicits reduced neural responses, as stimuli are represented by overlapping neural ensembles[19,20]) to investigate choice-related changes in neural representations of CS–US associations. As conditioning enhances CS's ability to pre-activate neural ensembles coding for US, we expected a change of CS–US associative strength, as measured by RS after choice-induced revaluation. Strong CS–US associations should elicit high fMRI-RS effects (i.e. low activation), whereas weak CS-US associations should elicit low fMRI-RS effects (i.e. high activation). We expected decreased neural representations of CS–US associations for CS^+_A and increased choice-related neural representations of CS–US associations for CS^0_A after choice-induced revaluation.

All univariate fMRI analyses steps were performed using tools from the functional magnetic resonance imaging of the brain (FMRIB) Software Library (FSL, v6.0)[44]. Preprocessing included motion correction using rigid-body registration to the central volume of the functional time series[45], correction for geometric distortions using the field maps and an *n*-dimensional phase-unwrapping algorithm ($\mathbf{B_0}$ unwarping)[46], slice timing correction using Hanning windowed sinc interpolation, high-pass filtering using a Gaussian-weighted lines filter of 1/100 Hz. EPI images were registered with the high-resolution brain images and normalized into standard (MNI) space using affine linear registration (boundary-based registration) as well as nonlinear registration[47,48]. Functional data were spatially smoothed using a Gaussian filter with 6 mm full-width at half maximum. We applied a conservative independent components analysis (ICA) to identify and remove obvious artefacts. Independent components were manually classified as signal or noise based on published guidelines[49], and noise components were removed from the functional time series. General linear models (GLMs) were fitted into pre-whitened data space to account for local autocorrelations[50]. The individual level (first level) GLM design matrix per run and participant included 22 box-car regressors in total. Eighteen regressors coded for onsets and durations of all 18 presented CS–US association trials (each modeled as single events of 1800 ms duration), one regressor coded for onsets and durations of the three within-run pauses (each 45 s), one regressor coded for onsets and durations of the attentional control task probes, two regressors coded onsets and durations of left and right button presses (delta stick functions on the recorded time of response button presses) and the six volume-to-volume motion parameters from motion correction during preprocessing were entered. Regressors were convolved with a hemodynamic response function (γ-function, mean lag = 6 s, SD = 3 s). Each first level GLM included five contrasts to estimate individual per run contrasts of parameter estimates for (1) lower $CS^0_A - US^0$ RS relative to the other equivalent CS^0_B, controlling for all other combinations of CS^0 presentations (Eq. (1)), (2) higher $CS^+_A - US^+$ RS relative to the other equivalent CS^+_B, controlling for all other combinations of CS^+ presentations (Eq. (2)), (3) higher CS^-_A RS relative to the other equivalent CS^-_B, controlling for all other combinations of CS^- presentations, (4) Conjunction of (1) and (2), i.e. voxels coding for both decrease of $CS^0_A - US^0$ RS and increase of $CS^+_A - US^+$ RS (Eq. (3)), (5) right vs. left button press. Two separate PRE and POST choice-induced revaluation second level (group level) GLMs were carried out by submitting individual level parameter estimates to mixed-effects statistics and ordinary least-squares (OLS) regression for higher-level contrast of parameter estimates (COPE) estimation. To control for multiple comparisons, cluster-based correction with an activation threshold of $Z > 2.3$ using a cluster-extent threshold of $P < 0.05$ was applied at the whole-brain level.

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

A priori regions-of-interest (ROIs) comprised the lateral orbitofrontal cortex (lOFC) and the hippocampus, as those regions have been implicated in representation and adaptive changes of stimulus–outcome associations, respectively[10,14,21,28]. We investigated POST choice-induced revaluation conjunction effects (Eq. (3)) to identify regions involved in processing of choice-induced changes to CS–US associations. An independent functional mask of a contrast investigating stimulus–outcome associations from a previous study[28] (restricted along the *z*-direction from −6 to −14 to constrain spatial extent), was used for small-volume correction (P_{SVC}) of the bilateral lOFC. The small-volume corrected functional activation mask from the conjunction contrast was used to extract contrast parameter estimates of the CS^0_A contrast and the CS^+_A contrast. Additionally, an independent anatomical mask of the hippocampus (Harvard–Oxford Atlas) was used to extract PRE and POST choice-induced revaluation contrast parameter estimates of the CS^0_A contrast and the CS^+_A contrast. PRE versus POST comparisons of activation were carried out using a repeated-measures ANOVA and post-hoc Wilcoxon-sign rank tests. As we had formulated directed hypotheses, and because parameter estimates were extracted from family-wise error corrected ROIs, we used one-tailed post-hoc tests. We report effect sizes η_p^2 for rmANOVAs, Cohen's U_3 for Wilcoxon signed-rank tests and Cohen's $U3_1$ for one-sample Wilcoxon signed-rank tests (range: 0–1, 0.5 indicating no effect), calculated in the MATLAB-based Measures-of-Effect-Size-toolbox[41]. Based on the reported effect sizes η_p^2 we additionally indicate post-hoc achieved power ($1-\beta$) for the hypothesized interaction effects of CS valence × time in rmANOVAs, calculated with G*Power[42,43] (v3.1.9.2). Additionally, we explored functional activation not surviving whole-brain or small-volume family-wise error corrections by thresholding activation maps at $Z = 2.8$ and extracting contrast estimates of the CS^0_A contrast and the CS^+_A contrast for brain–behavioral correlations.

Extracted parameter estimates were additionally used for brain–behavioral correlations using non-parametric Spearman correlations. For brain–behavior correlations, we had the directed hypotheses that associative strength, as measured by RS should be positively related to the difference between overall CP of CS^+_A and overall CP CS^+_B and negatively related to difference between overall CP CS^0_A and overall CP CS^0_B. Additionally, to refine our insights in the relationship of neural and behavioral results, we also correlated RS with pairwise CP for CS^+_A versus CS^+_B, again assuming a positive relationship and pairwise CP for CS^0_A versus CS^0_B, predicting a negative relationship. Due to our directed hypotheses, and because parameter estimates were extracted from family-wise error corrected ROIs, we performed one-tailed tests on Spearman correlation coefficients.

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

Multivariate fMRI data analysis. Complementary to the univariate RS-based approach, we performed confirmatory multivariate fMRI analyses, employing a neural pattern similarity analysis (a variant of RSA[22]) in the left hippocampus and right lateral OFC to further support the hypothesized associative strengthening/weakening mechanism. As for the RS-based analyses we assumed that presentation of a CS should induce pre-activation of neural ensembles coding the respective associated US. This mnemonic pre-activation should not only be present in trials where the CS was followed by the originally learned US, but also in trials where the CS was followed by any of the other two possible, but not associatively linked US. Similarity of neural patterns related to two CS from the same value category could thus potentially be indicative of associative memory retrieval of a US representation. According to the idea of choice-induced weakening of the CS^{0}_{A} association with US^{0} and strengthening of the CS^{+}_{A} association with US^{+}, our hypothesis was that neural similarity between same value stimulus–outcome pairs ($CS^{0}_{A}-US^{-}/CS^{0}_{B}-US^{-}$ and $CS^{0}_{A}-US^{+}/CS^{0}_{B}-US^{+}$; $CS^{+}_{A}-US^{-}/CS^{+}_{B}-US^{-}$ and $CS^{+}_{A}-US^{0}/CS^{+}_{B}-US^{0}$) should decrease from PRE to POST, indicating less similarity between patterns of interest (i.e. the weakened/strengthened stimulus and the respective partner stimulus). The assumption of decreased neural pattern similarity for pairs of CS^{+}_{A} and CS^{+}_{B} as a result of choice-induced strengthening of CS^{+}_{A} is based on two grounds. Firstly, we assumed that both CS^{+}_{A} and CS^{+}_{B} would equivalently, and partially reinstate the pattern representing US^{+} during PRE, but repeated retrieval of US^{+} and choice of CS^{+}_{A} should selectively strengthen the association between CS^{+}_{A} and US^{+} during POST. Since CS^{+}_{B} is not presented during revaluation and thus not actively rehearsed, the memory trace between CS^{+}_{B} and US^{+} should be subject to passive decay, presumably resulting in a slightly weakened association. Secondly, consistent with the literature on retrieval-induced forgetting[7–9], it is plausible to assume that actively retrieving the memory trace between the target stimulus CS^{+}_{A} and US^{+} would additionally weaken the memory trace between the competitor stimulus CS^{+}_{B} and US^{+}. Both of the above mechanisms would lead to a differentiation of the memory engrams encoding CS^{+}_{A} and US^{+}, and CS^{+}_{B} and US^{+} and should be reflected in diminished PRE to POST similarity. For the pair of control stimuli ($CS^{-}_{A}-US^{0}/CS^{-}_{B}-US^{0}$ and $CS^{-}_{A}-US^{+}/CS^{-}_{B}-US^{+}$), we did not expect changes in neural pattern similarity.

Preprocessing steps for multivariate fMRI analyses were identical as for previously mentioned univariate fMRI analyses, with the only exception that the functional imaging timeseries were not spatially smoothed. As for univariate fMRI analyses, we applied a conservative ICA to identify and remove obvious artefacts. GLMs were fitted into pre-whitened data space to account for local autocorrelations[50]. The individual level (first level) GLM design matrix per run and participant included 22 box-car regressors in total. Eighteen regressors coded for onsets and durations of all 18 presented CS–US-association trials (each modeled as single events of 1800 ms duration), one regressor coded for onsets and durations of the three within-run pauses (each 45 s), one regressor coded for onsets and durations of the attentional control task probes, two regressors coded onsets and durations of left and right button presses (delta stick functions on the recorded time of response button presses) and the six volume-to-volume motion parameters from motion correction during preprocessing were entered. Regressors were convolved with a hemodynamic response function (γ-function, mean lag = 6 s, SD = 3 s). Each first level GLM included one contrast to model activation related to each of the 18 presented CS–US-associations versus baseline (18 contrasts in total). The a priori ROIs were built in MNI space and back-projected into subject native space using inverse normalization parameters obtained from FSL during preprocessing procedures. We used these individual ROIs for spatially constrained multivoxel pattern extraction from the respective contrast t-value maps. Similarity-based analyses were carried out using the MATLAB-based multivariate pattern analysis toolbox CoSMoMVPA[51]. We employed 1−Pearson's product–moment correlation coefficient ($1-r$) as a measure of pairwise dissimilarity between neural patterns of interest, separately for PRE and POST and the two ROIs. Within-subject pairwise neural dissimilarity was subtracted from 1 (to create a measure of neural pattern similarity) and Fisher-Z transformed to closer approximate normally distributed data. We then calculated the within-subject PRE–POST change between the resulting pairwise neural pattern similarity measures (POST r – PRE r, ΔPearson's r). As we had a directional hypothesis of negative PRE–POST change of both CS^{0}_{A}/CS^{0}_{B} and CS^{+}_{A}/CS^{+}_{B} and to increase the number of trials included in the inference, we pooled neural pattern similarity measures across all patterns of interest. Lastly, average neural similarity changes were analyzed at the group level with one-sample t-tests against 0. Due to the expected negative effects of neural pattern similarity changes in patterns of interest, we used one-tailed tests accordingly. As there was no such directional hypothesis for the pairs of CS^{-}_{A}/CS^{-}_{B}, we employed two-tailed tests. We report effect sizes Cohen's $U3_1$ for one-sample t-tests against 0 (range: 0–1, 0.5 indicating no effect), calculated in the MATLAB-based Measures-of-Effect-Size-toolbox[41]. Based on the mean and standard deviation of pattern similarity measures, we report post-hoc achieved power of all analyses.

Univariate fMRI contrasts.

CS^{0}_{A} fMRI-RS contrast:

$$[2\times(CS^{0}_{A}-US^{0}-CS^{0}_{B}-US^{0})]-[(CS^{0}_{A}-US^{-}-CS^{0}_{B}-US^{-})+(CS^{0}_{A}-US^{+}-CS^{0}_{B}-US^{+})] \quad (1)$$

CS^{+}_{A} fMRI-RS contrast:

$$[2\times(CS^{+}_{A}-US^{+}-CS^{+}_{B}-US^{+})]-[(CS^{+}_{A}-US^{-}-CS^{+}_{B}-US^{-})+(CS^{+}_{A}-US^{0}-CS^{+}_{B}-US^{0})] \quad (2)$$

Conjunction fMRI-RS contrast:

$$[2\times(CS^{0}_{A}-US^{0}-CS^{0}_{B}-US^{0})]-[(CS^{0}_{A}-US^{-}-CS^{0}_{B}-US^{-})+(CS^{0}_{A}-US^{+}-CS^{0}_{B}-US^{+})]$$

and

$$[2\times(CS^{+}_{A}-US^{+}-CS^{+}_{B}-US^{+})]-[(CS^{+}_{A}-US^{-}-CS^{+}_{B}-US^{-})+(CS^{+}_{A}-US^{0}-CS^{+}_{B}-US^{0})]. \quad (3)$$

Computational models. To formally characterize behavior in experiments 1, 2, 3, and 5, we fit six different variants of a reinforcement learning model using Rescorla–Wagner-like delta update rules[27]. For each experiment, we compared the six models that implemented different ways by which participants could have learned CS–US associative strength—and updated associative strength during choice-induced revaluation.

In model 1 (Pavlovian learning only), CS values are exclusively acquired during Pavlovian learning, without any further update during the choice-induced revaluation. On each trial, the value of the stimulus currently presented was updated according to the following rule:

$$V_{t+1,i}=V_{t,i}+\alpha_{\text{Learning}}(R_t-V_{t,i}) \quad (4)$$

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

Model 2 (Pavlovian learning and choice-induced revaluation, chosen CS) acquired stimulus values during Pavlovian learning exactly like model 1, but additionally updated CS–US associative strength by fictive reward prediction errors elicited by decisions in the choice-induced revaluation phase. The fictive reward prediction errors was based on our reasoning that presentation of CS during revaluation would lead to retrieval of the associated US and that, consistent with our hypothesis, stimulus–outcome association of the chosen CS would be strengthened, whereas the association of the unchosen CS would be weakened, resulting from nonmonotonic memory plasticity[16]. As no objective feedback (US) was presented during choice-induced revaluation, the reward prediction could only be derived by assuming associative retrieval. This model only updated associative strength of the chosen, but not the unchosen CS.

$$V_{t+1,\text{ch}}=V_{t,\text{ch}}+\alpha_{\text{ch}}(R_{t,\text{ch}}-V_{t,\text{ch}}) \quad (5)$$

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

Model 3 (Pavlovian learning and choice-induced revaluation, chosen, and unchosen CS) was set up to account for the possibility that updating of both the chosen and unchosen CS association could occur during the choice-induced revaluation. It was identical to model 2, with the exception that it performed an update to the associative strength of the unchosen stimulus, using a separate learning rate α_{unch}, in addition to updating the chosen CS associative strength.

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

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

$$P_{C,t}=\frac{1}{1+\exp(-\text{VD}_t/\tau)} \quad (6)$$

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

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

$$-\mathrm{LLE} = -\Sigma \log(P_{C,t}) \quad (7)$$

First, we ran a grid search on an n-dimensional grid in log space (where $n =$ number of free parameters for each model), with 30 steps along each dimension. The grid optimum was then used as initial value and passed to constrained non-linear optimization using the MATLAB function fmincon. Optimized negative log likelihoods were compared by means of the sample-size corrected Akaike Information Criterion (AICc, Eqs. (8) and (9)). The model with the lowest AICc value was considered to account best for the observed participants' choice data, penalized for model complexity and sample size (number of participants).

$$\mathrm{AIC} = 2k - 2(-\mathrm{LLE}) \quad (8)$$

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

$$\mathrm{AIC_C} = \mathrm{AIC} + \frac{2k^2 + 2k}{n - k - 1} \quad (9)$$

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

Additionally, we performed 10,000 simulations of choice behavior per participant. After value estimation as described above, using individual parameters of the best-fitting models at group level, the resulting value estimates for each stimulus were entered in the exact same sequence of 120 choices that the participants individually experienced during the decision probe phase. Simulated overall choice probabilities were averaged per participant across 10,000 simulations.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw behavioral data, univariate parameter estimates extracted from regions of interest, thresholded and unthresholded univariate *Z*-maps, neural pattern similarity correlation matrices and brain–behavioral correlation data that support the findings of this study are publicly available at GitHub (https://github.com/LLuettgau/revaluation). The neuroimaging raw data that support the findings of this study are available upon reasonable request from the corresponding author (L.L.). The neuroimaging raw data are not publicly available due to them containing information that could compromise research participant privacy/consent. The source data underlying Figs. 1e–i, 2d, e, 3a, b, 4b and Supplementary Figs. 1a–h, 2a–h, 3a, b, and 5a are provided as a Source Data file. A reporting summary for this article is available as a Supplementary Information file. Source data are provided with this paper.

Code availability

Custom analysis code for the reported behavioral data analyses, univariate fMRI analyses on extracted parameters, multivariate fMRI analyses on neural pattern similarity correlation matrices, brain–behavioral correlational data analyses and computational modeling/simulations are publicly available at GitHub (https://github.com/LLuettgau/revaluation). Source data are provided with this paper.

Received: 11 October 2019; Accepted: 15 June 2020;
Published online: 03 July 2020

References

1. Ariely, D. & Norton, M. I. How actions create—not just reveal—preferences. *Trends Cogn. Sci.* **12**, 13–16 (2008).
2. Riefer, P. S., Prior, R., Blair, N., Pavey, G. & Love, B. C. Coherency-maximizing exploration in the supermarket. *Nat. Hum. Behav.* **1**, 1–4 (2017).
3. Sharot, T., Velasquez, C. M. & Dolan, R. J. Do decisions shape preference? Evidence from blind choice. *Psychol. Sci.* **21**, 1231–1235 (2010).
4. Brehm, J. W. Postdecision changes in the desireability of alternatives. *J. Abnorm. Soc. Psychol.* **52**, 384–389 (1956).
5. Schonberg, T. et al. Changing value through cued approach: an automatic mechanism of behavior change. *Nat. Neurosci.* **17**, 625–630 (2014).
6. Izuma, K. et al. Neural correlates of cognitive dissonance and choice-induced preference change. *Proc. Natl Acad. Sci. USA* **107**, 22014–22019 (2010).
7. Wimber, M., Alink, A., Charest, I., Kriegeskorte, N. & Anderson, M. C. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nat. Neurosci.* **18**, 582–589 (2015).
8. Anderson, M. C., Bjork, R. A. & Bjork, E. L. Remembering can cause forgetting: retrieval dynamics in long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1063–1087 (1994).
9. Hulbert, J. C. & Norman, K. A. Neural differentiation tracks improved recall of competing memories following interleaved study and retrieval practice. *Cereb. Cortex* **25**, 3994–4008 (2015).
10. Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J. & Behrens, T. E. J. Segregated encoding of reward-identity and stimulus–reward associations in human orbitofrontal cortex. *J. Neurosci.* **33**, 3202–3211 (2013).
11. Barron, H. C., Dolan, R. J. & Behrens, T. E. J. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* **16**, 1492–1498 (2013).
12. Howard, J. D., Kahnt, T. & Gottfried, J. A. Converging prefrontal pathways support associative and perceptual features of conditioned stimuli. *Nat. Commun.* **7**, 1–11 (2016).
13. Tonegawa, S., Morrissey, M. D. & Kitamura, T. The role of engram cells in the systems consolidation of memory. *Nat. Rev. Neurosci.* **19**, 485–498 (2018).
14. Boorman, E. D., Rajendran, V. G., O'Reilly, J. X. & Behrens, T. E. Two anatomically and computationally distinct learning signals predict changes to stimulus–outcome associations in hippocampus. *Neuron* **89**, 1343–1354 (2016).
15. Onat, S. & Büchel, C. The neuronal basis of fear generalization in humans. *Nat. Neurosci.* **18**, 1811–1818 (2015).
16. Ritvo, V. J. H., Turk-Browne, N. B. & Norman, K. A. Nonmonotonic plasticity: how memory retrieval drives learning. *Trends Cogn. Sci.* **23**, 726–742 (2019).
17. Klein, T. A., Ullsperger, M. & Jocham, G. Learning relative values in the striatum induces violations of normative decision making. *Nat. Commun.* **8**, 1–12 (2017).
18. Palminteri, S., Khamassi, M., Joffily, M. & Coricelli, G. Contextual modulation of value signals in reward and punishment learning. *Nat. Commun.* **6**, 8096 (2015).
19. Grill-Spector, K. & Malach, R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol.* **107**, 293–321 (2001).
20. Barron, H. C., Garvert, M. M. & Behrens, T. E. J. Repetition suppression: a means to index neural representations using BOLD? *Philos. Trans. R. Soc. B* **371**, 20150355 (2016).
21. Garvert, M. M., Dolan, R. J. & Behrens, T. E. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife* **6**, 1–20 (2017).
22. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 1–28 (2008).
23. Tamaoka, K., Makioka, S., Sanders, S. & Verdonschot, R. G. A new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychol. Res.* **81**, 696–708, www.kanjidatabase.com (2017).
24. Blechert, J., Meule, A., Busch, N. A. & Ohla, K. Food-pics: an image database for experimental research on eating and appetite. *Front. Psychol.* **5**, 617 (2014).
25. Guitart-Masip, M. et al. Go and no-go learning in reward and punishment: Interactions between affect and effect. *Neuroimage* **62**, 154–166 (2012).
26. Swart, J. C. et al. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife* **6**, 1–36 (2017).
27. Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. & Prokasy, W. F.) 64–99 (Appleton-Century-Crofts, 1972).
28. Jocham, G. et al. Reward-Guided Learning With And Without Causal Attribution. *Neuron* **90**, 177–190 (2016).
29. Storm, B. C., Bjork, E. L. & Bjork, R. A. Accelerated relearning after retrieval-induced forgetting: the benefit of being forgotten. *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 230–236 (2008).
30. Chen, M. K. & Risen, J. L. How choice affects and reflects preferences: revisiting the free-choice paradigm. *J. Pers. Soc. Psychol.* **99**, 573–594 (2010).
31. Salomon, T., Botvinik-Nezer, R., Oren, S. & Schonberg, T. Enhanced striatal and prefrontal activity is associated with individual differences in nonreinforced preference change for faces. *Hum. Brain Mapp.* 1–18 https://doi.org/10.1002/hbm.24859 (2019).
32. Botvinik-Nezer, R., Bakkour, A., Salomon, T., Shohamy, D. & Schonberg, T. Memory for individual items is related to non-reinforced preference change. Preprint at https://doi.org/10.1101/621292v1 (2019).
33. Aridan, N., Pelletier, G., Fellows, L. K. & Schonberg, T. Is ventromedial prefrontal cortex critical for behavior change without external reinforcement? *Neuropsychologia* **124**, 208–215 (2019).
34. Wimmer, G. E. & Shohamy, D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* **338**, 270–273 (2012).
35. Guggenmos, M., Wilbertz, G., Hebart, M. N. & Sterzer, P. Mesolimbic confidence signals guide perceptual learning in the absence of external feedback. *Elife* **5**, 1–19 (2016).
36. Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. J. Human replay spontaneously reorganizes experience. *Cell* **178**, 1–13 (2019).

37. Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H. & Rushworth, M. F. S. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* **65**, 927–939 (2010).
38. Tolman, E. C. Cognitive maps in rats and man. *Psychol. Rev.* **55**, 189–208 (1948).
39. Behrens, T. E. J. et al. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
40. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
41. Hentschke, H. hhentschke/measures-of-effect-size-toolbox (https://www.github.com/hhentschke/measures-of-effect-size-toolbox), GitHub. Retrieved 10 January 2020 (2020).
42. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).
43. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
44. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).
45. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
46. Jenkinson, M. Fast, automated, N-dimensional phase-unwrapping algorithm. *Magn. Reson. Med.* **49**, 193–197 (2003).
47. Andersson, J. L., Jenkinson, M. & Smith, S. *Non-linear Registration Aka Spatial Normalisation*. FMRIB Technical Report TRO7JA2 22. https://www.fmrib.ox.ac.uk/datasets/techrep/ (2007).
48. Andersson, J. L., Jenkinson, M. & Smith, S. M. *Non-linear Optimisation*. FMRIB Technical Report TRO7JA1. https://www.fmrib.ox.ac.uk/datasets/techrep/ (2007).
49. Griffanti, L. et al. Hand classification of fMRI ICA noise components. *Neuroimage* **154**, 188–205 (2017).
50. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* **14**, 1370–1386 (2001).
51. Oosterhof, N. N., Connolly, A. C., Haxby, J. V & Rosa, M. J. CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Front. Neuroinform.* **10**, 1–27 (2016).

Acknowledgements

The authors thank all volunteers who participated in this study. The authors would like to thank Halla Mulla-Osman, Stefanie Linnhoff, Nicola Harzen and Denise Scheermann for their invaluable support during data acquisition. We thank the Center for Magnetic Resonance Research of the University of Minnesota for providing the Multiband accelerated fMRI sequence. The authors additionally thank Theo O.J. Gruendler, Lukas M. Neugebauer, and Daniel Priegnitz for helpful discussions during set-up of the study and for sharing custom MATLAB code. 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

L.L. 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. C.T. set up the MRI acquisition protocol, read and edited versions of the manuscript and approved the final version of the manuscript. L.F.K. analyzed the data, read and edited versions of the manuscript and approved the final version of the manuscript. G.J. 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

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-17192-7.

Correspondence and requests for materials should be addressed to L.L.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

Decisions bias future choices by modifying hippocampal associative memories

Luettgau et al.

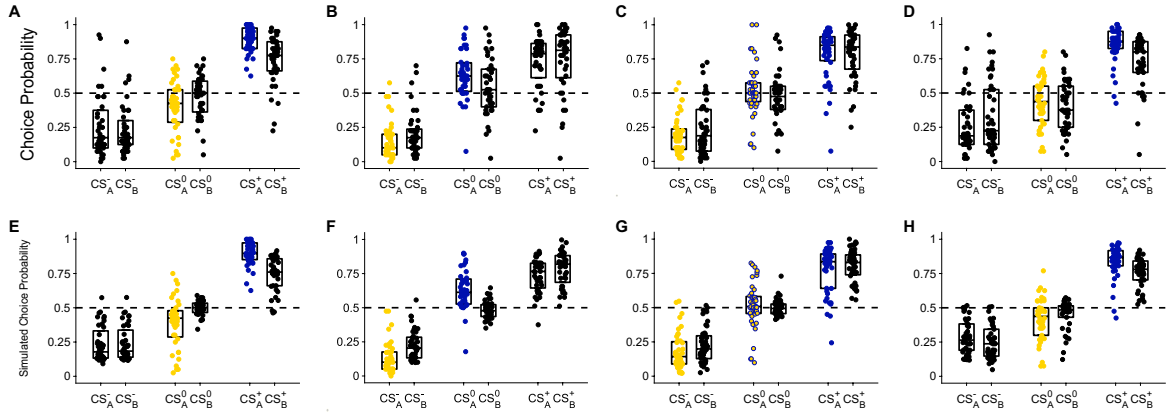
Supplementary Methods

For our proposed associative mechanism, we assumed that after Pavlovian conditioning, a CS would pre-activate the respective, associatively learned US and participants would make their decisions between CS based on the associated outcome. Our hypothesized mechanism relies on the assumption that participants form a (simple) model of the task, which is well in line with the literature on associative learning[1,2]. However, alternatively, the observed pattern of results could also be explained by a cached value account. According to this account, participants acquire cached values during Pavlovian conditioning and further use those model-free values to guide their decision, independent of associated outcomes or the learned relationships between CS and US. Our study was not designed to dissociate both mechanisms on a behavioral level. The prediction of the cached value account would be that decisions during choice-induced revaluation lead to changes in cached values and leave CS-US association unaffected. In other words, CS^0_A should have a reduced cached value, whereas the cached value of CS^+_A should be increased following choice-induced revaluation. In another fMRI contrast, we tested whether the relationship/similarity between CS^0_A and US^- (Supplementary Note 4) would change due to the choice-induced revaluation phase. If our results can be explained by the cached value account, we would assume that CS^0_A followed by US^- would show reduced activation/higher similarity in the post choice-induced revaluation run, compared to its equivalent partner stimulus CS^0_B followed by US^-. As any other stimulus- or outcome-related effects are controlled for in the contrast, and CS^0_A was not learned to be associated with US^- during Pavlovian conditioning (i.e. not being able to pre-activate a US^- representation), we assumed that reductions in activation could only be interpreted as a repetition of a shared feature, namely the stimulus/outcome value. After testing the contrast, we extracted parameter estimates from an independent hippocampus mask and correlated the parameter estimates with CP. We hypothesized that CS^0_A and US^- similarity would be negatively related to overall CP CS^0_A – overall CP CS^0_B. The lower the CS^0_A and US^- similarity, the less likely participants would be to select CS^0_A.

Supplementary Note 1

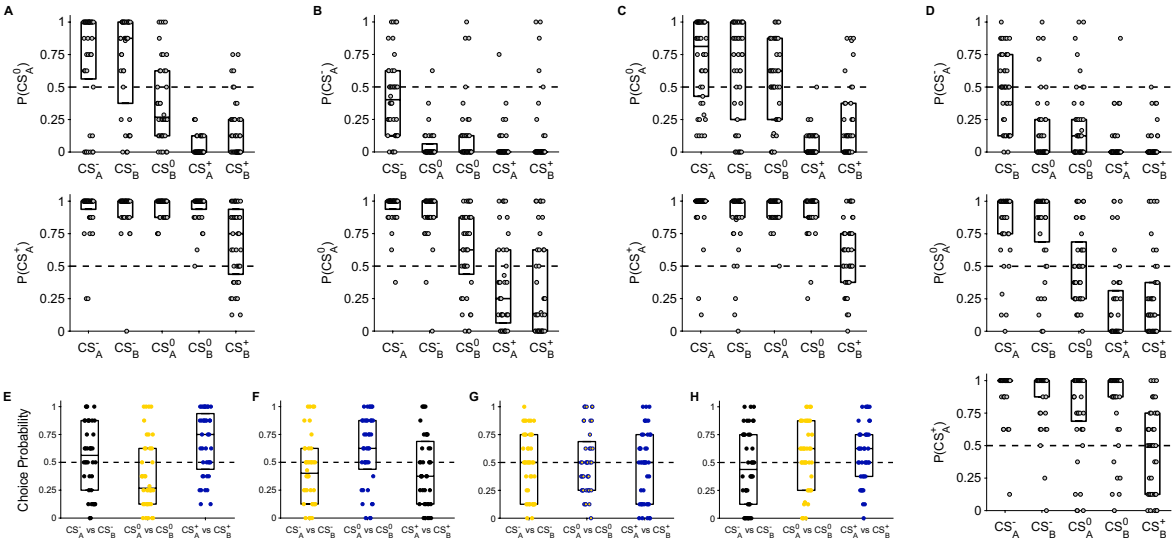
As an alternative to the proposed associative mechanism, parts of our functional neuroimaging results could also be explained based on cached values. During Pavlovian conditioning, participants might acquire incentive (cached) CS values and use these to guide their decisions, independent of CS-US associations. It is possible that revaluation choices changed these cached CS values, instead of the CS-US associations. Under this reasoning, one would assume that the value information conveyed by CS^0_A, the CS not chosen during revaluation and supposedly devalued, would become more similar to the value information of the low-valued US^- after choice-induced revaluation[3]. By its very nature, the fMRI-RS signal for learned CS-US associations represents BOLD signal reductions due to repetition of both value and identity features shared by CS and US, alongside the associative strength between CS and US[1]. However, CS^0_A should by design of the experiment not be capable to elicit any associative strength- or identity-related RS effects when followed by US^-, as it was never coupled with US^- during Pavlovian conditioning. Since value and identity of the US are inextricably linked in the present design, we reasoned that any RS signal changes for CS^0_A followed by US^- from PRE to POST would most likely be attributable to changes in valuation of CS^0_A. The cached value account would predict larger repetition suppression effects for CS^0_A followed by US^- as for CS^0_B followed by US^- during POST. Our design allowed us to set up a contrast to test this possibility (Supplementary Note 4). This indeed yielded effects in the right posterior hippocampus and a midbrain region in the vicinity of the dorsolateral substantia nigra pars compacta (Supplementary Figure 4) in POST. However, consistent with the absence of a behavioral revaluation effect for CS^0_A, the observed hippocampal effect did not differ PRE-POST (test on parameter estimates extracted from right hippocampal anatomical mask, Z = 0.90, P = .184, U_3 = .64, Wilcoxon signed-rank test, one-tailed). Importantly, the parameter estimates of the cached value effect in the right hippocampus and the associative effect in the left hippocampus were not correlated (all $\rho s < .18$, $Ps > .260$, Spearman correlations, two-tailed). Even when extracting parameter estimates for cached value (Supplementary Note 4) and associative effect (Equation 1) from the exact same anatomical mask of the left

hippocampus, we did not observe significant correlations between the two contrasts (PRE: ρ = .22, *P* = .170; POST: ρ = –.03, *P* = .852; Spearman correlations, two-tailed), which would have been expected if both contrasts measure the same with flipped signs. According to the cached value account, hippocampal representations of CS^{0}_{A}-US^{-} should be inversely related to preferences of CS^{0}_{A}. However, directly opposing this prediction, relationships of right hippocampus parameter estimates, and choice behavior were positive (all ρ*s* < .31, *P* > 0.05, Spearman correlations, two-tailed), rendering an explanation of the observed behavioral results based on cached values unlikely. It should be noted that the cached value effects would have critically depended on choice-induced devaluation of CS^{0}_{A}. However, as we did not find behavioral support for the hypothesized devaluation effect, the cached value effects should be interpreted with caution.



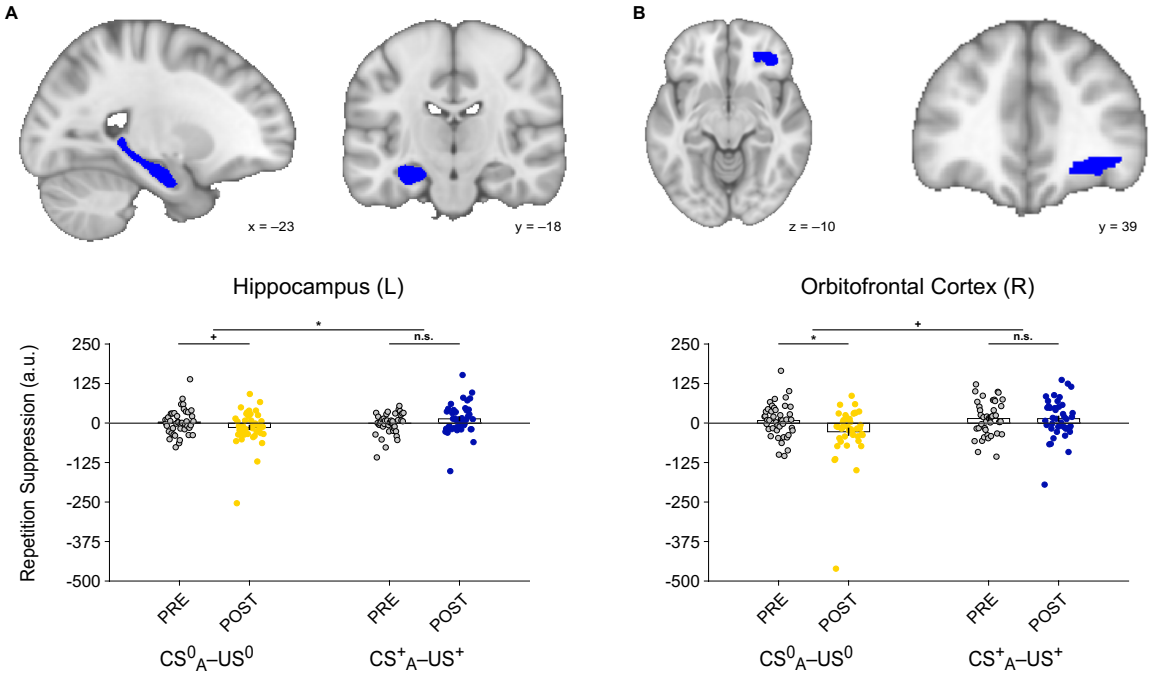
Supplementary Figure 1. Behavioral and simulation results.

A-D) Empirical choice results (as in Fig. 1E-H). Previously chosen CS (blue dots) are selected more often compared to equivalent CS (black dots) in Experiment 1 (A, $N = 40$, $Z = 3.98$, $P < .001$, Cohen's $U_3 = .85$, Wilcoxon signed-rank test, one-tailed), Experiment 2 (B, $N = 40$, $Z = 2.20$, $P = .014$, $U_3 = .68$, Wilcoxon signed-rank test, one-tailed), Experiment 5 (D, $N = 42$, $Z = 3.03$, $P = .001$, $U_3 = .76$, Wilcoxon signed-rank test, one-tailed) and previously unchosen CS (yellow dots) are selected less often compared to equivalent CS (black dots) in Experiment 1 (A, $N = 40$, $Z = 1.97$, $P = .025$, $U_3 = .70$, Wilcoxon signed-rank test, one-tailed) and Experiment 2 (B, $N = 40$, $Z = 1.91$, $P = .028$, $U_3 = .66$, Wilcoxon signed-rank test, one-tailed) during decision probe. The effect is not present in Experiment 3 (C, $N = 44$, $Z = 0.41$, $P = .68$, $U_3 = .55$, Wilcoxon signed-rank test, two-tailed), indicating that the roughly equal proportion of choices and non-choices of CS^0_A during revaluation had cancelled each other out. E-H) Averaged simulated choice probabilites (10,000 simulations per participant, Experiment 1: $N = 40$ (E), Experiment 2: $N = 40$ (F), Experiment 3: $N = 44$ (G), Experiment 5: $N = 42$ (H)), recapitulating observed empirical choice patterns. Please note that no statistical tests were performed on simulated data. Box plot center lines represent sample medians and box bottom/top edges show $25^{th}/75^{th}$ percentile of the (simulated) data, respectively. Source data are provided as a Source Data file.



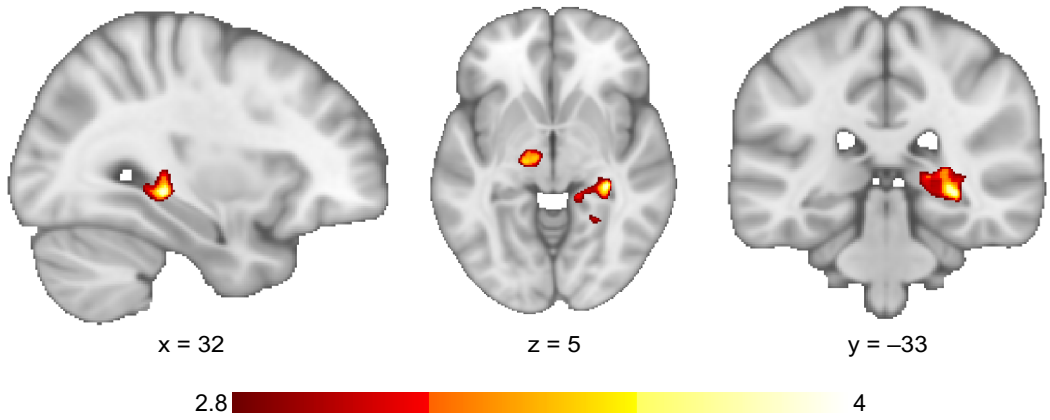
Supplementary Figure 2. Extended Behavioral Results and Pairwise Choice Probabilities.

Decision probe pairwise choice probabilities for CS that were presented during revaluation against every other CS. A) In Experiment 1 (N = 40), CS^0_A is only preferred over CS^-_A and CS^-_B (top), CS^+_A is the most preferred CS (bottom). B) In Experiment 2 (N = 40), CS^-_A is the least preferred CS (top), CS^0_A is chosen more frequently than CS^-_A, CS^-_B, and CS^0_B (bottom). C) In Experiment 5 (N = 42), CS^0_A is only preferred over CS^-_A and CS^-_B (top), CS^+_A is the most preferred CS (bottom). D) In Experiment 3 (N = 44), CS^0_A is chosen at the same frequency as CS^0_B (middle). Please note that since these plots serve descriptive purposes only, no statistical tests were performed. E-H) Pairwise within-category choice probabilities displaying the pairwise comparison that are most indicative of choice-induced preference changes. E) Experiment 1 (N = 40): CS^+_A and CS^+_B (Z = 3.43, P < .001, Cohen's $U3_1$ = .69, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed; CS^0_A and CS^0_B (Z = 2.05, P = .020, $U3_1$ = .68, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed), F) Experiment 2 (N = 40): CS^0_A and CS^0_B (Z = 1.93, P = .027, $U3_1$ = .68, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed); CS^-_A and CS^-_B (Z = 1.41, P = .079, $U3_1$ = .63, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed), G) Experiment 3 (N = 44): CS^0_A and CS^0_B (Z = 0.12, P = .905, $U3_1$ = .57, one-sample Wilcoxon signed-rank test vs. 0.5, two-tailed), H) Experiment 5 (N = 42): CS^+_A and CS^+_B (Z = 1.93, P = .027, $U3_1$ = .62, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed); CS^0_A and CS^0_B (Z = 1.07, P = .857, $U3_1$ = .62, one-sample Wilcoxon signed-rank test vs. 0.5, one-tailed). These results suggest that initially conditioned value was not overridden by revaluation choices and that the choice bias was mostly driven by the pairwise decisions of the respective revaluation CS against the same-value CS. However, repetitions of choices between the two revaluation CS also contributed to the observed overall choice probability effects. Box plot center lines represent sample medians and box bottom/top edges show 25th/75th percentile of the data, respectively. Source data are provided as a Source Data file.



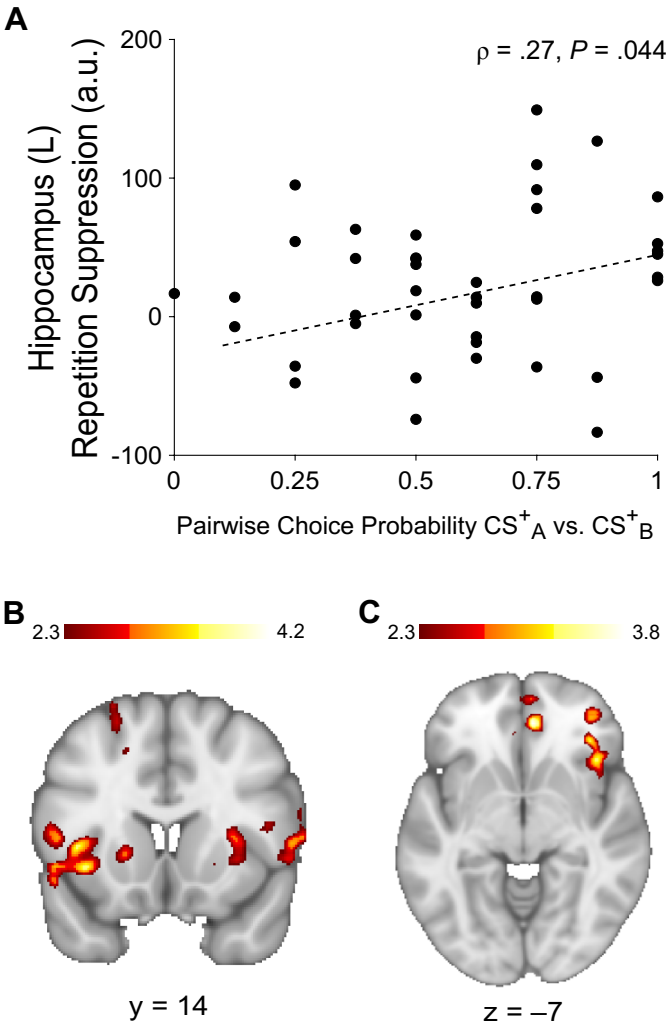
Supplementary Figure 3. Repetition suppression effects, separately for CS^0_A-US^0 and CS^+_A-US^+.

Extracted parameter estimates of the repetition suppression effects for CS^0_A-US^0, controlling for activation elicited by CS^0_A followed by both incorrect outcomes (US^- and US^+) (Supplementary Note 2) and CS^+_A-US^+ repetition suppression, controlling for activation elicited by CS^+_A followed by both incorrect outcomes (US^- and US^0) (Supplementary Note 3). A) Extracted parameter estimates from the left hippocampus (interaction effect: $F_{1, 41} = 4.31$, $P = .044$, $\eta^2_p = .10$, $1–\beta = .99$, rmANOVA). CS^0_A-US^0 showed a marginal effect from PRE to POST ($Z = 1.82$, $P = .069$, $U3_1 = .64$, Wilcoxon signed-rank test, two-tailed), but only a numerical difference was observed for PRE-POST change of CS^+_A-US^+ ($Z = 1.56$, $P = .120$, $U3_1 = .69$, Wilcoxon signed-rank test, two-tailed). B) Extracted parameter estimates from the right lateral orbitofrontal cortex (interaction effect: $F_{1, 41} = 3.51$, $P = .068$, $\eta^2_p = .08$, $1–\beta = .99$, rmANOVA). CS^0_A-US^0 showed a significant decrease from PRE to POST ($Z = 2.44$, $P = .015$, $U3_1 = .71$, Wilcoxon signed-rank test, two-tailed), but there was no evidence of a PRE-POST difference for CS^+_A-US^+ ($Z = .01$, $P = .995$, $U3_1 = .52$, Wilcoxon signed-rank test, two-tailed). Bar plots represent sample means. Error bars indicate standard errors of the mean. Asterisks indicate P-values < .05, plus signs represent P-values > .05 and < .10. Color bars indicate Z-values. Source data are provided as a Source Data file.



Supplementary Figure 4. Cached value control analysis results.

Whole-brain corrected pattern of activation (activation threshold $Z > 2.3$, cluster-forming threshold $P < 0.05$, thresholded at $Z > 2.8$ for display purposes) for the cached value control analysis (Supplementary Note 4) during the POST choice-induced revaluation fMRI run ($N = 42$). Two clusters in the right posterior hippocampus and in a left midbrain region in the vicinity of the dorsolateral substantia nigra pars compacta survived whole-brain correction. Extracting hippocampal activation with an independent anatomical right hippocampus mask and comparison of parameter estimates for PRE and POST yielded no significant difference ($Z = 0.90$, $P = .184$, $U_3 = .64$, Wilcoxon signed-rank test, one-tailed), suggesting no evidence for an influence of choice-induced revaluation decisions on activation. Additionally, we did not observe significant correlations between the cluster in the right posterior hippocampus from the cached value control analysis and the left hippocampus result from the associative analyses (all $\rho s < .18$, $Ps > .260$, Spearman correlations, two-tailed), suggesting different processes. Even when extracting parameter estimates for cached value (Supplementary Note 4) and associative effect (Equation 1) from the exact same anatomical mask of the left hippocampus, we did not observe significant correlations between the two contrasts (PRE: $\rho = .22$, $P = .170$; POST: $\rho = -.03$, $P = .852$; Spearman correlations, two-tailed), which would have been expected if both contrasts measure the same with flipped signs. Parameter estimates from the right posterior hippocampus were only marginally related with overall CP CS^0_A – overall CP CS^0_B: $\rho_{40} = .30$, $P = .050$ (two-tailed) and the within-category CP CS^0_A vs. CS^0_B: $\rho_{40} = .29$, $P = .062$ (two-tailed). However, the direction of the correlation was exactly opposite to the predictions of the cached value account. As there was no apriori anatomical hypothesis for the SNpc, we were unable to extract parameter estimates in an unbiased fashion. Thus, PRE and POST parameter estimates could not be compared. Left SNpc parameter estimates were not significantly related to left VTA parameter estimates from the associative analysis (all $\rho s < .08$, $Ps > .63$), nor with decision probe behavior (all $\rho s < .20$, $Ps > .20$). Color bar indicates Z-values.



Supplementary Figure 5. Brain-behavior correlation and whole-brain regressions.

A) Left hippocampus parameter estimates (independent anatomical mask, Fig. 4A) for repetition suppression of CS^{+}_{A} relative to CS^{+}_{B} relative to CS^{+}_{B}-US^{+}, controlling for activation elicited by CS^{+}_{A} and CS^{+}_{B} followed by both incorrect outcomes (US^{-} and US^{0}) (Equation 2) during the POST choice-induced revaluation fMRI run positively correlate with decision probe pairwise choice probability of CS^{+}_{A} vs. CS^{+}_{B} (N = 42, ρ_{41} = .27, P = .044, Spearman correlation, one-tailed). The higher repetition suppression was after choice-induced revaluation, the more likely participants were to prefer CS^{+}_{A} over CS^{+}_{B}. B) CS^{+}_{A} preference vs. CS^{+}_{B} was correlated with POST choice-induced revaluation run repetition suppression (N = 42) in the bilateral anterior insula, suggesting choice-induced strengthening of CS-related pre-activation of neural ensembles coding for the used food items. C) We observed a positive relationship between PRE-POST changes in associative strength of CS^{0}_{A} and pairwise CP for CS^{0}_{A} vs. CS^{0}_{B}, in two clusters in the medial orbitofrontal cortex and lOFC, extending to the anterior insula (N = 42). Color bars indicate Z-values. Source data are provided as a Source Data file.

Supplementary Table 1. Computational Parameters and Model Fits

Exp. 1	Update	Model	$\alpha_{Learning}$	α_{ch}	α_{unch}	τ	*–LLE*	AIC_c
	RW	Model 1	< .001 (0 – 1)		-	.0002 (0 – 1.55)	63.43 (37.50 – 81.58)	133.59
		Model 2	.04 (0 – 1)	< .001 (0 – .78)	-	.20 (0 – 1.48)	40.52 (26.19 – 80.67)	98.10
		Model 3	.02 (0 – .58)	.50 (0 – .85)	.006 (0 – .86)	.15 (0 – .77)	37.27 (25.62 – 79.88)	**94.62**
	Asso. Value	Model 1	.002 (0 – 1)		-	.11 (.05 – 2.65)	41.35 (27.38 – 80.9)	97.10
		Model 2	.02 (0 – 1)	.002 (0 – .84)	-	.14 (.06 – 1.43)	40.60 (20.64 – 77.37)	95.87
		Model 3	.03 (0 – 1)	.42 (0 – .84)	.005 (0 – 1)	.17 (.06 – 3.37)	37.08 (20.60 – 81.88)	96.30
Exp. 2	RW	Model 1	.55 (0 – .95)		-	.45 (0 – 66.59)	53.41 (31.86 – 83.18)	115.99
		Model 2	.18 (0 – .72)	.005 (0 – .95)	-	.24 (0 – 3.44)	45.42 (15.43 – 82.02)	103.18
		Model 3	.04 (0 – .85)	.02 (0 – .88)	.005 (0 – .99)	.24 (0 – .77)	44.67 (5.55 – 77.11)	98.98
	Asso. Value	Model 1	.003 (0 – 1)		-	.15 (.05 – 2.65)	48.77 (15.81 – 83.23)	101.68
		Model 2	.04 (0 – 1)	.005 (0 – .92)	-	.20 (.002 – 6.11)	44.53 (15.71 – 82.81)	100.51
		Model 3	.03 (0 – 1)	.02 (0 – .99)	.005 (0 – .74)	.20 (.004 – 2.08)	44.38 (5.55 – 82.4)	**97.43**
Exp. 3	RW	Model 1	.03 (0 – .88)	-		.18 (0 – 1.55)	50.60 (22.15 – 82.66)	107.92
		Model 2	.02 (0 – .85)	< .001 (0 – .70)	-	.07 (0 – 16.92)	41.06 (14.50 – 82.99)	95.31
		Model 3	.01 (0 – .83)	.009 (0 – .55)	< .001 (0 – .40)	.08 (0 – 24.81)	40.47 (13.64 – 83.07)	93.15
	Asso. Value	Model 1	.005 (0 – 1)	-		.11 (0 – 1.44)	41.12 (13.65 – 82.64)	93.78
		Model 2	.02 (0 – 1)	.008 (0 – .76)	-	.16 (0 – 1.28)	40.29 (13.65 – 82.62)	93.84
		Model 3	.02 (0 – 1)	.012 (0 – .86)	.009 (0 – .80)	.13 (.004 – 2.01)	44.38 (5.55 – 82.4)	**91.96**
Exp. 5	RW	Model 1	< .001 (0 – 1)	-		< .001 (0 – 3.89)	71.61 (49.60 – 82.72)	143.94
		Model 2	< .001 (0 – .85)	< .001 (0 – 1)	-	.16 (0 – 5.05)	54.25 (22.54 – 82.64)	114.94
		Model 3	.04 (0 – .61)	.52 (0 – .85)	.008 (0 – .72)	.24 (0.03 – 1.02)	51.11 (22.54 – 81.86)	110.71
	Asso. Value	Model 1	.002 (0 – 1)	-		.16 (0.04 – 4.78)	55.11 (23.54 – 82.86)	112.97
		Model 2	.13 (0 – 1)	.007 (0 – 1)	-	.24 (0.05 – 4.81)	50.79 (23.37 – 82.61)	110.69
		Model 3	.18 (0 – 1)	.50 (0 – 1)	.009 (0 – .97)	.22 (.04 – 2.98)	49.25 (22.90 – 82.51)	**107.85**

Note: Median and range of computational parameters, negative log likelihood estimates (*–LLE*) and corrected Akaike Information Criteria (AICc) of the Rescorla-Wagner like (RW)[4] and Associative Value (Asso. Value) reinforcement learning models for the four experiments. Lowest AIC_c per experiment, which was considered to account best for the observed participants' choice data, penalized for model complexity and sample size (number of participants) in bold letters.

Supplementary Note 2

CS^{0}_{A} simple effect contrast:

$$[\ 2 \times CS^{0}_{A}\text{-}US^{0}\ ] - [\ CS^{0}_{A}\text{-}US^{-} + CS^{0}_{A}\text{-}US^{+}\ ]$$

Supplementary Note 3

CS^{+}_{A} simple effect contrast:

$$[\ 2 \times CS^{+}_{A}\text{-}US^{+}\ ] - [\ CS^{+}_{A}\text{-}US^{-} + CS^{+}_{A}\text{-}US^{0}\ ]$$

Supplementary Note 4

Cached value control analysis contrast:

$$[\ 2 \times (CS^{0}_{B}\text{-}US^{-} - CS^{0}_{A}\text{-}US^{-})\ ] - [\ (CS^{0}_{A}\text{-}US^{0} - CS^{0}_{B}\text{-}US^{0}) - (CS^{0}_{B}\text{-}US^{+} - CS^{0}_{A}\text{-}US^{+})\ ]$$

Supplementary References

1. Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J. & Behrens, T. E. J. Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex. *J. Neurosci.* **33**, 3202–3211 (2013).
2. Barron, H. C., Dolan, R. J. & Behrens, T. E. J. Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat. Neurosci.* **16**, 1492–1498 (2013).
3. Summerfield, C., Luyckx, F. & Sheahan, H. Structure learning and the posterior parietal cortex. *Prog. Neurobiol.* **184**, (2020).
4. Rescorla, R. A. & Wagner, A. R. A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. in *Classical conditioning II: current research and theory* (eds. Black, A. H. & Prokasy, W. F.) 64–99 (Appleton-Century-Crofts, 1972).

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 agents to infer value in environments with sparse rewards. Despite its explanatory potential for real-world learning phenomena, surprisingly little is known about the neural mechanism underlying such value transfer 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 and anterior hippocampus 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.

Main text

Learning in naturalistic settings is characterized by infrequent direct encounters with rewarding or punishing stimuli[1]. 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[1]. Exploiting such statistical regularities would allow agents to infer 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 spread of value to stimuli that have never been directly paired with reinforcement[2]. 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[2]. 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[1,3] and decades of behavioral investigations[1,3–5], surprisingly little is known about the neural mechanism underlying the transfer of value in SOC[3].

At least four potential mechanisms accounting for SOC have been proposed[1,3–5]. 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 suggests that neither of these three hypotheses can account for second-order learning effects[4,5]. 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 or value conveyed by the US during second-order learning ("direct link" hypothesis)[1,3]. 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[6]. Presumably, SOC critically depends on convergence of

neural patterns representing CS_2 and US in the amygdala[1,3,7,8] and hippocampus[9]. However, it remains unclear, how exactly a "direct link" between CS_2 and US-related motivational value 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 value conferred through second-order learning. 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[10], 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[1,3,7,8,11] and hippocampus[9]. Both amygdala and hippocampus, as well as orbitofrontal cortex (OFC) and ventral striatum[12] have consistently been shown as the key structures involved in second-order learning and conditioned reinforcement[1,3,7–9,11]. Our analyses therefore focused on these regions.

To test our predictions, we combine a SOC paradigm and choice preference tests with cross-session[13], cross-modality searchlight[14] 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 value transfer effects were accompanied by first-order CS-related reinstatement of the neural patterns representing US in the lateral OFC and increased functional coupling between OFC and amygdala/anterior hippocampus 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.

----------Insert Figure 1 about here----------

Results

Evidence for value transfer 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. Importantly, participants were not informed about the underlying associative structure of the tasks but were instructed to perform simple attentional control 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 they reflect acquired (CS_1) and transferred (CS_2) value. 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. 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 conferred value to both CS_1 and 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.

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

After establishing behavioral evidence for associative value transfer, 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. To test for cortical reinstatement of neural patterns representing US during second-order conditioning, we used functional magnetic resonance imaging (fMRI) and a cross-session[13], cross-modality searchlight[14] classification approach (multivariate pattern analysis, MVPA). To obtain unbiased estimates of the neural patterns representing our gustatory US, without the confounding influence of associations to a learned first-order CS, we first performed an fMRI classifier training experiment on day 1 (Fig. 1B) during which the US^+ and US^- were presented. A multivariate pattern classifier (linear support vector machine, C-SVM) was trained on the fMRI data from this session. On day 2, during SOC, we used the weights of the classifier trained on gustatory neural patterns to predict the identity of the visual CS (Fig. 1D) that had been paired with the US during FOC (Fig. 1C, see Methods and Fig. 2A for schematic of the classification approach). In other words, we used data from one sensory modality, assessed during a first day to train a classification model and tested the model's generalizability to unseen data from another sensory modality on a second day. We found 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. Classification with a 3-mm

searchlight revealed a small-volume corrected (Neurosynth[15] meta-analytic functional mask for term “taste”) cluster of above-chance level (0.5) classification accuracy (extracted cluster mean = 0.56, *SD* = 0.07) in the left lateral OFC (lOFC, Fig. 2B, peak voxel at MNI [x = –21, y = 30, z = –17], *z* = 2.26, *P* = 0.012, random-effect cluster corrected[16], 50,000 iterations, one-tailed). The location of this lOFC cluster is consistent with this region's well-documented role in gustatory processing, particularly in representing motivational[17–19] aspects of gustatory sensation and taste memory[20]. To ensure that 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 et al. (personal correspondence). 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 identity 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).

----------Insert Figure 2 about here----------

Interaction between lOFC and amygdala 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 a psychophysiological interaction (PPI) analysis. We contrasted CS^n trials with CS^- and CS^+ 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 hippocampus, extending to amygdala and

medial temporal lobe, was higher in CS^- and CS^+ compared to CS^n trials (Fig. 3A, peak voxel at MNI [x = –60, y = 3, z = –24], Z = 3.97, P = 0.029, whole-brain corrected).

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 lateral prefrontal cortex (Fig. 3B, peak voxel at MNI [x = 49, y = 43, z = 6], Z = 4.21, P = 0.034, whole-brain corrected). The more participants preferred CS_2^+ (versus CS_2^-), the stronger these regions' BOLD signal covaried during SOC.

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

We reasoned that if the amygdala uses the reinstated cortical outcome pattern in lOFC to acquire an association between CS_2 and the respective US, one would expect similarity between the neural patterns evoked by CS_2 and US, respectively. This similarity should increase over the course of second-order conditioning, as the association is being acquired. Using a least-squares separate (LS-S) approach[21] to deconvolve single-trial estimates, we first computed overall neural pattern similarity[22] between the pattern evoked by CS_2 (during SOC) and their respective US (during classifier training) in the bilateral amygdala. There was significant pattern similarity between CS_2^- and US^- (t_{28} = 3.38, P = 0.002, $U3_1$ = 0.79; but not 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 an aversively tasting quinine solution, which has been shown to produce innate avoidance responses that are not thought to be generated by learning[23]. In contrast, US^+ was orange juice, a compound stimulus that consists of many taste features (e.g. sour and fruity components), next to sweet sensation, which also produces innate responses[23]. Presumably, this heterogeneity of taste facets in US^+ had induced more variability in taste responses in our participants. Additionally, it appears likely that in a society offering sugary and highly palatable foods and drinks in abundance, US^+ (an orange juice, most likely a very familiar taste to most of our participants) might not have been perceived rewarding, novel enough to represent a motivationally relevant and salient event for our participants. It might well be that

participants did not have experience with the aversive taste of quinine, rendering it a novel and salient gustatory sensation, presumably resulting in only the US^- being considered behaviorally and motivationally relevant. Thus, it is possible that the observed preference for CS_1^+ and CS_2^+ was indeed due to an aversion against CS_1^- and CS_2^-.

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

----------Insert Figure 3 about here----------

Discussion

Using a second-order conditioning paradigm, we show transfer of outcome value during higher-order learning, 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[2,6]. 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[24–26]. 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 by value conferred

through second-order conditioning. Such transfer of value may be beneficial, as it allows an agent to exploit the underlying relational structure and infer value in 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[27–29] and in correctly assigning credit for a reward to the causal stimulus choice[29]. It is also part of the network most consistently involved in taste processing as obtained from the Neurosynth data base[15] and plays a well-documented role in representing motivational[17–19] aspects of gustatory sensation and taste memory[20]. Importantly, we had trained the classifier on the gustatory US prior to pairing them with visual first-order CS (on a separate day). Thus, decoding of the visual CS during second-order learning cannot be spuriously driven by visual responses elicited by the US during classifier training. Reinstatement of the neural pattern representing the outcome constitutes a necessary prerequisite to establish an associative link between second-order CS and the motivational state or value conveyed by gustatory outcomes. Our findings are consistent with a body of lesion studies in animals that indicate a key role of OFC in second-order conditioning and in responding for conditioned reinforcement[30,31]. They also align with the account of OFC representing abstract relational structure, or a state space[32,33] and recent evidence suggesting that OFC neurons encode relevant second-order combinations of past state variables to guide decision-making[34].

OFC clearly does not act in isolation. A plethora of studies suggests that interactions between OFC, amygdala and ventral striatum[30,35–37] might also support second-order learning processes. We found evidence for the region of OFC that showed reinstatement of outcome

representations to increase BOLD signal covariation with amygdala and anterior hippocampus during second-order learning. Neurons in OFC and amygdala show complex, bi-directional interactions during acquisition and reversal of outcome-predictive associations[36] and functional disconnection of OFC and amygdala using asymmetric lesions produce deficits in flexibly adjusting behavior to changes in stimulus value[35,37]. We also observed that representations of visual second-order CS in amygdala became more similar to gustatory US representations 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[7,9]. 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[9].

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 value transfer during human second-order learning. 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. Second, the neural mechanisms of value transfer in SOC have remained elusive. The present study suggests that during second-order conditioning, outcome representations, and presumably features pertaining to the motivational value of outcomes, are reinstated in the lateral OFC. Our data suggests that reinstated outcome representations could be communicated between lOFC and amygdala/anterior hippocampus for associative linking of previously neutral stimuli with the motivational value conveyed by the reinstated outcomes.

It has been proposed that relational knowledge of events and behavior characterized by considering the state-transition probabilities of the environment might emerge from associative learning mechanisms[38]. Second-order conditioning can be thought of as an instantiation of such a learning process in which a currently absent, but previously experienced and motivationally relevant task state (receipt of an outcome) is reinstated in

lOFC to enable propagation of motivational value to task states that are never directly experienced in contiguity with reward or punishment. Higher-order learning paradigms impose a relational structure of events on agents and thus can be cast as an inference problem, very much like in transitive inference (A→B, B→C, hence A→C)[39]. Credit assignment processes of this kind have been shown to involve the lOFC[29,40]. The lOFC reinstatement of outcome representations observed in the present study could thus pose a necessary prerequisite for value attribution to task states or stimuli that have never been directly followed by an outcome, but nevertheless possess outcome-predictive properties.

How the reinstatement of cortical US patterns found in our study relates to the observation in rats that midbrain dopamine neurons acquire temporal difference error signals in response to CS_2[6] presents an important question for future studies. Furthermore, it would be of great interest to elucidate the directionality and exact content of information flow between lOFC and amygdala/anterior hippocampus and whether this information transfer is supported by phase coherence in theta oscillations[41–43].

Taken together, our data support the "direct link" hypothesis of second-order conditioning[1,4,5], stating that second-order CS and the motivational value of outcomes – events that had never been explicitly paired during the individual's learning history – can be linked by exploiting the relational structure of events. Importantly, the present study enables a closer comparison between humans and other species during higher-order learning. In conclusion, our results suggest a neural mechanism by which the motivational value of outcomes can be propagated to stimuli that are never experienced in contiguity with reinforcement, allowing for credit assignment in real-world learning scenarios with infrequent direct encounters with rewards or punishments.

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[44] (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[45] (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

attached to Luer lock infusion lines (1,40 m length, 2 mm inner diameter) that participants held centrally in their mouths like drinking straws. In the fMRI study, infusion line position order (Q-O (*N* = 17) and O-Q (*N* = 15) for quinine (Q) and orange juice (O)) was counterbalanced across participants. In the fMRI study, the US bolus onset was preceded (500 ms) by a blue square that was presented for 2500 ms on the screen (Figure 1B). Participants were instructed to only swallow the US bolus upon offset of the blue square. Each trial was separated by an inter-trial-interval (ITI) marked by a grey screen. The ITI per trial was drawn from a discretized γ-distribution (shape = 6, scale = 1) truncated for an effective range of values between 3500 ms and 10,000 ms. Participants took self-paced breaks after each 10^{th} trial during which they could drink water. Importantly, the instructions did not contain information about the underlying associative structure of the experiment, aiming at leaving participants unaware of the associative learning process. Instead, participants were instructed to perform a simple attentional control task, during which they should react as quickly and as correctly as possible by pressing the "y" button upon seeing the CS_1 colored in red. Each CS_1 was colored red in 10 % of the trials (90 % of trials grayscale image) and color did not predict US contingency. Performance during the task was rewarded with a bonus of 1 € (if > 70 % correct answers). Participants performed very well in the attentional control task (overall probability of correct answers: *M* = .99, *SD* = .05). In the fMRI sample, both ratings and first-order conditioning were performed outside the MRI scanner.

Learning experiment – second-order conditioning

For second-order conditioning (SOC), participants were presented with CS_2^+ followed by 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 first-order conditioning and thus was not paired with any US. The novel CS_2^n-CS_1^n pair served to control for attentional and novelty effects. In each trial (Figure 1D), a CS_2 (2000 ms) was followed by an inter-stimulus interval (500 ms) marked by a fixation cross, and a CS_1 (2000 ms). CS_2 were followed by a CS_1 deterministically. Each trial was separated by an

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

Learning experiment – choice preference test

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

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

Behavioral analyses

Data were analyzed using MATLAB 2019a (v9.6.0.1072779, The MathWorks, Inc., Natick, MA, USA) and RStudio[46] (version 3.6.3, RStudio Team, Boston, MA) using custom analysis scripts. We calculated an average choice probability (CP) per first- and second-order CS per subject including all binary decisions in which the respective CS was present (1 = selection of the respective target CS, 0 = selection of the alternative CS) for CS_1^+ and CS_2^+ (versus CS_1^- and CS_2^-, respectively). We hypothesized above-chance level CPs for both CS_1^+ and CS_2^+. For both first-order and second-order novel-to-lure comparison trials, we calculated an average CP per stimulus and expected no difference from chance level for these choices. CPs for each CS in the fMRI and behavioral study were pooled and jointly analyzed. For hypothesis testing, we calculated 89%-highest posterior density intervals (89%-HPDI) for each choice probability using the R package rethinking[47]. The HPDI represents the parameter values most consistent with the data. We specified a weakly informative prior probability distribution (10000 steps), an aggregated binomial distribution including the number of choices for the respective stimulus and the total number of choice trials involving the respective CS as likelihood function (Equation 1) and randomly drew 100000 samples from a grid-approximated (10000 steps) normalized posterior probability distribution. For hypothesis testing, we specified a region of practical equivalence (ROPE), i.e. an interval of parameter values (CP = [.45; .55]) representing the null hypothesis. Specifically, we expected that the 89%-HPDI would not overlap with the ROPE for choices of CS_1^+ and CS_2^+ (versus CS_1^- and CS_2^-, respectively), allowing to reject the null hypothesis. However, the ROPE

would overlap with the 89%-HPDI for both first-order and second-order novel-to-lure comparison trials. Overlap between HPDIs and ROPEs was quantified using the R package bayestestR[48].

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

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

fMRI – classifier training experiment

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

Participants held infusion lines centrally in their mouths like drinking straws. Additionally, infusion lines were fixed at absorbent materials attached to throat, chin and cheeks. The US bolus onset was preceded (500 ms) by a blue square that was presented in total for 4000 ms on the screen. Participants were instructed to only swallow the US bolus upon offset of the blue square. Each trial was separated by an ITI marked by a grey screen. The ITI was drawn from a uniform distribution with 5 discrete steps (range: 3000 – 7000 ms). Both US were presented in a pseudo-random order, thus reducing the influence of potentially confounding low-level features of the schedule (e.g. number of same/different US repetitions, different ITI lengths following each US). During each run of the classifier training experiment, participants performed a 0-back style attentional control task. After a pseudo-random 20% of 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^-)) via button presses with their right and left index fingers on an MRI-compatible response box. Correct responses were rewarded with 0.05 € and incorrect responses or time-out trials (without a response by the participant within 2500 ms after onset of the probe trial) resulted in a 0.05 € penalty which would be summed up as a bonus upon completion of the experiment. On average, participants earned a bonus of 1.86 € (*SD* = .11) during the classifier training experiment. Performance during the 0-back attentional control task was generally high (overall probability of correct answers, excluding time-out trials: *M* = .93, *SD* = .05).

fMRI – acquisition

During SOC, one run and during the classifier training experiment, five runs of fMRI were acquired on a 3 Tesla Siemens PRISMA MR-system (Siemens, Erlangen, Germany), using a 64-channel head coil. Blood oxygenation level dependent (BOLD) signals were acquired using a multi-band accelerated T2*-weighted echo-planar imaging (EPI) sequence (multi-band acceleration factor 2, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle = 80°, field of view (FoV) = 220 mm, voxel size = 2.2 × 2.2 × 2.2 mm, no gap). Per volume, 66 slices covering the whole brain, tilted by approximately 15° in *z*-direction relative to the

anterior–posterior commissure plane were acquired in interleaved order. The first 5 volumes of the functional imaging time series were automatically discarded to allow for T1 saturation. 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)[49]. 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[50], correction for geometric distortions using the field maps and an n-dimensional phase-unwrapping algorithm (B_0 unwarping)[51], 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[52,53]. 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 published guidelines[54], and noise components were removed from the functional time series.

fMRI – searchlight classification analyses

As conditioning enhances CS's ability to pre-activate neural ensembles representing an associated US, we expected that a multivariate classification algorithm trained on neural patterns representing the US (during the classifier training task) would be able to correctly predict the class of a paired CS during SOC. Hence, we used a cross-session[13], cross-modality searchlight[14] classification approach to identify brain regions in which the CS class during SOC could be predicted based on training the classifier during the classifier training experiment on day 1. General linear models (GLMs) were fitted into pre-whitened data space to account for local autocorrelations[55]. For the searchlight classification analyses, the individual level (first level) GLM design matrix per run and participant of the classifier training experiment included four box-car regressors in total. Two regressors coded for onsets and durations of both US (each modelled as single events of 4000 ms duration) and two regressors coded onsets and durations of left and right button presses (delta stick functions on the recorded time of response button presses) and the six volume-to-volume motion parameters from motion correction during preprocessing were entered to account for residual head motion. The SOC run per participant included five box-car regressors in total. Three regressors coding for 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 (each modelled as single events of 4500 ms duration, due to the 100 % contingency of CS-CS pairs), one regressor coding onsets and durations of right button presses (delta stick functions on the recorded time of response button presses), one regressor coding onset and duration of the within-run pause (45 sec), and the six volume-to-volume motion parameters from motion correction during preprocessing were entered.

Regressors were convolved with a hemodynamic response function (γ-function, mean lag = 6 s, SD = 3 s). Each first level GLM included two or three (for classifier training experiment or SOC, respectively) contrasts to estimate individual per run *t*-statistic maps for

each US or each CS_2-CS_1 pair (for classifier training experiment or SOC, respectively). For SOC, the activation pattern of CS_2^n/CS_1^n was subtracted from CS_2^+/CS_1^+ and CS_2^-/CS_1^- activation patterns to account for general visual activation that was common to both CS and US (i.e., swallowing cue) presentation and could thus confound classification accuracies.

All classification-based analyses were conducted in subject native space. Per participant *t*-statistic maps were subjected to linear support vector machine (C-SVM, default C = 1) classification using the MATLAB-based multivariate pattern analysis toolbox CoSMoMVPA[56]. For cross-session, cross-modality classification, a classifier was trained on the spatial activation patterns of US^+ versus US^- and tested on CS_2^+/CS_1^+ versus CS_2^-/CS_1^- during SOC within 3-mm searchlight spheres across the whole brain. Each searchlight sphere classification accuracy was mapped to the center voxel of the sphere, 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 chance level maps. Before group level statistics, the normalization parameters obtained from preprocessing were applied to both whole-brain classification accuracy maps and chance level maps for normalization to MNI space. The resulting normalized whole-brain maps were spatially smoothed with a Gaussian kernel (5 mm full-width half-maximum).

We employed small-volume correction (P_{SVC}) to assess significance of clusters with above-chance level (0.5) classification accuracy in the lateral orbitofrontal cortex (lOFC), our a priori region-of-interest (ROI). This region has been implicated in higher-order gustatory processing, such as motivational aspects and discrimination of taste[17,20,57], but also representation and adaptive changes of stimulus-outcome associations[27–29,58,59]. We thus reasoned that this brain region would be a well-suited candidate for associative coupling between visual conditioned stimuli and gustatory outcomes and the proposed associative transfer learning. We employed two different approaches for ROI definition: 1) an independent anatomical mask of the lateral OFC (Harvard-Oxford Atlas) and 2) a meta-analysis-based explicit mask for functional activations related to the term “taste” in the

Neurosynth data base[15] (www.neurosynth.org), thresholded at $z > 7$. This value was chosen so that only meta-analytic clusters of activation, but not single voxels, survived thresholding. The Neurosynth-based mask encompassed, aside from the lateral OFC, functional activation clusters in bilateral anterior insula. This approach aimed at reducing inferential limitations related to arbitrary ROI definition.

Within these ROIs, we computed group-level random-effect cluster-statistics corrected for multiple comparisons in small volumes as implemented in CoSMoMVPA[56]. 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[16]. Since classifications accuracies below chance level are generally limited in interpretability, we considered clusters significant if $P_{SVC} < .05$ (one-tailed).

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[60].

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 a first-level PPI GLM with 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) and CS_2^n/CS_1^n (magnitude coded as –1) as psychological regressor and 3) the interaction between 1) and 2) as psychophysiological interaction regressor. First-level regressor 2) was convolved with a hemodynamic response function (γ-function, mean lag = 6 s, SD = 3 s). Contrast images from the first level were then taken to the group level using a random effects analysis. In addition to the main effect of the PPI regressor, we also investigated whether task-related functional covariation of lOFC with the rest of the brain was related to second-order conditioning. For this, we entered the averaged second-order choice preference test data per subject as an additional regressor to the group level GLM. We used cluster-based correction with an activation threshold of $Z > 2.3$ and a cluster-extent threshold of $P < .05$ at whole-brain level to control for multiple comparisons.

fMRI – neural pattern similarity analyses

In addition to our classification-based approach, we investigated whether there was evidence for changes of neural pattern similarity across second-order conditioning. Specifically, we used a least-squares separate (LS-S) approach[21] to deconvolve single-trial estimates of neural patterns representing CS_2 and CS_1 and performed a template-based neural pattern similarity analysis[61] (a variant of representational similarity analysis, RSA[22]) between patterns of the two US during the classifier training experiment and all trial-specific CS_2/CS_1 patterns during the SOC run. We estimated two subject-specific GLMs per trial, one for both CS_1 and CS_2, containing the following regressors: 1) onset and duration of the respective trial to be estimated (each modelled as a single event of 2000 ms duration), 2) onsets and durations of the remaining CS_1 or CS_2 trials, respectively, 3) onsets and durations of all other CS_1 or CS_2 trials, respectively. Additionally, one regressor coding onsets and durations of right button presses (delta stick functions on the recorded time of response button presses), one regressor coding onset and duration of the within-run pause (45 sec), and the six volume-to-volume motion parameters from motion correction during preprocessing were entered. For

example, the GLM of the 25th trial of CS_1^+, contained one regressor (onset and duration) modelling this particular trial and one separate regressor (onsets and durations) for all remaining CS_1 trials but the 25th trial of CS_1^+ (149 other trials in total) and one separate regressor (onsets and durations) modelling all CS_2 trials. This procedure is assumed to allow for better deconvolution of events happening within close temporal proximity than standard least-squares-based approaches, which generally do not distinguish between overall and trial-specific error terms[21,62].

Each first level GLM included one contrast to model activation related to the respective trial of interest versus baseline. The a priori ROIs of the bilateral amygdala were built in MNI space and back-projected into subject native space using inverse normalization parameters obtained during preprocessing. We used these individual ROIs for spatially constrained multivoxel pattern extraction from the respective contrast *t*-value maps. Similarity-based analyses were carried out using CoSMoMVPA[56]. We employed 1−Pearson's product-moment correlation coefficient (1−*r*) as a measure of pairwise dissimilarity between trial-specific neural patterns and the US template. Within-subject pairwise neural dissimilarity was subtracted from 1 (to create a measure of neural pattern similarity) and Fisher-*Z* transformed to closer approximate normally distributed data. CS_1^n-US^-/CS_2^n-US^- and CS_1^n-US^+/CS_2^n-US^+ early and late neural pattern similarity were subtracted from CS_1^--US^-/CS_2^--US^- and CS_1^+-US^+/CS_2^+-US^+ early and late neural pattern similarity, respectively. We hypothesized that neural pattern similarity would show dissociable changes from early (first 25) to late (last 25) trials, as both CS_2^+ and CS_2^- should become more similar to the respective indirectly associated US, indicating associative learning transfer from CS_1 to CS_2. Contrarily, CS_1^+ and CS_1^- should, if anything, become less similar to the respectively paired US (e.g. due to extinction). Average change of early to late CS-US neural pattern similarity was analyzed at the group level with paired-samples t-tests. Due to the expected negative change for CS_1^+ and CS_1^- and the expected positive change for CS_2^+ and CS_2^- across trials, we used one-tailed tests accordingly.

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. We thank the Center for Magnetic Resonance Research of the University of Minnesota for providing the Multiband accelerated fMRI sequence. 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. **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.

References

1. Gewirtz, J. C. & Davis, M. Using Pavlovian higher-order conditioning paradigms to investigate the neural substrates of emotional learning and memory. *Learn. Mem.* **7**, 257–266 (2000).

2. Sharpe, M. J., Batchelor, H. M. & Schoenbaum, G. Preconditioned cues have no value. *Elife* **6**, (2017).

3. Parkes, S. L. & Westbrook, R. F. Role of the basolateral amygdala and NMDA receptors in higher-order conditioned fear. *Rev. Neurosci.* **22**, 317–333 (2011).

4. Rizley, R. C. & Rescorla, R. A. Associations in second-order conditioning and sensory preconditioning. *J. Comp. Physiol. Psychol.* **81**, 1–11 (1972).

5. Barnet, R. C., Arnold, H. M. & Miller, R. R. Simultaneous conditioning demonstrated in second-order conditioning: Evidence for similar associative structure in forward and simultaneous conditioning. *Learn. Motiv.* **22**, 253–268 (1991).

6. Maes, E. J. P. *et al.* Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors. *Nat. Neurosci.* **23**, 176–178 (2020).

7. Gewirtz, J. C. & Davis, M. Second-order fear conditioning prevented by blocking NMDA receptors in amygdala. *Nature* **388**, 471–474 (1997).

8. Setlow, B., Gallagher, M. & Holland, P. C. The basolateral complex of the amygdala is necessary for acquisition but not expression of CS motivational value in appetitive Pavlovian second-order conditioning. *Eur. J. Neurosci.* **15**, 1841–1853 (2002).

9. Gilboa, A., Sekeres, M., Moscovitch, M. & Winocur, G. Higher-order conditioning is impaired by hippocampal lesions. *Curr. Biol.* **24**, 2202–2207 (2014).

10. Tonegawa, S., Morrissey, M. D. & Kitamura, T. The role of engram cells in the systems consolidation of memory. *Nat. Rev. Neurosci.* **19**, 485–498 (2018).

11. Hatfield, T., Han, J. S., Conley, M., Gallagher, M. & Holland, P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with pavlovian second-order conditioning and reinforcer devaluation effects. *J. Neurosci.* **16**, 5256–5265 (1996).

12. Mcdannald, M. A., Setlow, B. & Holland, P. C. Effects of ventral striatal lesions on first- and second-order appetitive conditioning. *Eur. J. Neurosci.* **38**, 2589–2599 (2013).

13. Stokes, M., Thompson, R., Cusack, R. & Duncan, J. Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.* **29**, 1565–1572 (2009).

14. Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3863–3868 (2006).

15. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).

16. Stelzer, J., Chen, Y. & Turner, R. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *Neuroimage* **65**, 69–82 (2013).

17. Small, D. M. *et al.* Human cortical gustatory areas: A review of functional neuroimaging data. *Neuroreport* **10**, 7–14 (1999).

18. Rolls, E. T. The Orbitofrontal Cortex and Reward. *Cereb. Cortex* **10**, 284–294 (2000).

19. Rolls, E. T. Brain mechanisms underlying flavour and appetite. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 1123–1136 (2006).

20. Kobayashi, M. *et al.* Functional imaging of gustatory perception and imagery: 'Top-down' processing of gustatory signals. *Neuroimage* **23**, 1271–1282 (2004).

21. Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636–2643 (2012).

22. Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 1–28 (2008).

23. Yiannakas, A. & Rosenblum, K. The insula and taste learning. *Front. Mol. Neurosci.* **10**, 1–24 (2017).

24. Wimmer, G. E. & Shohamy, D. Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science* **338**, 270–3 (2012).

25. Wang, F., Schoenbaum, G. & Kahnt, T. Interactions between human orbitofrontal cortex and hippocampus support model-based inference. *PLoS Biol.* **18**, 1–24 (2020).

26. Jara, E., Vila, J. & Maldonado, A. Second-order conditioning of human causal learning. *Learn. Motiv.* **37**, 230–246 (2006).

27. Klein-Flugge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J. & Behrens, T. E. J. Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex. *J. Neurosci.* **33**, 3202–3211 (2013).

28. Luettgau, L., Tempelmann, C., Kaiser, L. F. & Jocham, G. Decisions bias future choices by modifying hippocampal associative memories. *Nat. Commun.* **11**, 3318 (2020).

29. Jocham, G. *et al.* Reward-Guided Learning with and without Causal Attribution. *Neuron* **90**, 177–190 (2016).

30. Cousens, G. A. & Otto, T. Neural Substrates of Olfactory Discrimination Learning with Auditory Secondary Reinforcement. I. Contributions of the Basolateral Amygdaloid Complex and Orbitofrontal Cortex. *Integr. Physiol. Behav. Sci.* **38**, 272–294 (2003).

31. Gallagher, M., McMahan, R. W. & Schoenbaum, G. Orbitofrontal cortex and representation of incentive value in associative learning. *J. Neurosci.* **19**, 6610–6614 (1999).

32. Niv, Y. Learning task-state representations. *Nat. Neurosci.* **22**, 1544–1553 (2019).

33. Wikenheiser, A. M. & Schoenbaum, G. Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nat. Rev. Neurosci.* **17**, 513–523 (2016).

34. Nogueira, R. *et al.* Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. *Nat. Commun.* **8**, (2017).

35. Fiuzat, E. C., Rhodes, S. E. V. & Murray, E. A. The role of orbitofrontal-amygdala interactions in updating action-outcome valuations in macaques. *J. Neurosci.* **37**, 2463–2470 (2017).

36. Morrison, S. E., Saez, A., Lau, B. & Salzman, C. D. Different Time Courses for Learning-Related Changes in Amygdala and Orbitofrontal Cortex. *Neuron* **71**, 1127–1140 (2011).

37. Baxter, M. G., Parker, A., Lindner, C. C. C., Izquierdo, A. D. & Murray, E. A. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J. Neurosci.* **20**, 4311–4319 (2000).

38. Wang, J. X. *et al.* Prefrontal cortex as a meta-reinforcement learning system. *Nat. Neurosci.* **21**, 860–868 (2018).

39. Dusek, J. A. & Eichenbaum, H. The hippocampus and memory for orderly stimulus relations. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7109–7114 (1997).

40. Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H. & Rushworth, M. F. S. Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal

Cortex in Contingent Learning. *Neuron* **65**, 927–939 (2010).

41. Knudsen, E. B. & Wallis, J. D. Closed-Loop Theta Stimulation in the Orbitofrontal Cortex Prevents Reward-Based Learning. *Neuron* **106**, 537-547.e4 (2020).

42. Benchenane, K. *et al.* Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal- Prefrontal Network upon Learning. *Neuron* **66**, 921–936 (2010).

43. Young, J. J. & Shapiro, M. L. Dynamic coding of goal-directed paths by orbital prefrontal cortex. *J. Neurosci.* **31**, 5989–6000 (2011).

44. Brainard, D. H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).

45. Tamaoka, K., Makioka, S., Sanders, S. & Verdonschot, R. G. www.kanjidatabase.com: a new interactive online database for psychological and linguistic research on Japanese kanji and their compound words. *Psychol. Res.* **81**, 696–708 (2017).

46. RStudioTeam. RStudio: Integrated Development for R. (2019).

47. McElreath, R. rethinking: Statstical rethinking book package. (2020).

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

49. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).

50. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage* **17**, 825–841 (2002).

51. Jenkinson, M. Fast, automated,N-dimensional phase-unwrapping algorithm. *Magn. Reson. Med.* **49**, 193–197 (2003).

52. Andersson, J. L. ., Jenkinson, M. & Smith, S. Non-linear registration aka spatial normalisation. *FMRIB Tech. Rep. TRO7JA2* 22 (2007).

53. Andersson, J. L. ., Jenkinson, M. & Smith, S. M. Non-linear optimisation. *FMRIB Tech. Rep. TRO7JA1* (2007).

54. Griffanti, L. *et al.* Hand classification of fMRI ICA noise components. *Neuroimage* **154**, 188–205 (2017).

55. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *Neuroimage* **14**, 1370–1386 (2001).

56. Oosterhof, N. N., Connolly, A. C., Haxby, J. V & Rosa, M. J. CoSMoMVPA : Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab / GNU Octave. **10**, 1–27 (2016).

57. Miranda, M. I. Taste and odor recognition memory: The emotional flavor of life. *Rev. Neurosci.* **23**, 481–499 (2012).

58. Garvert, M. M., Dolan, R. J. & Behrens, T. E. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife* **6**, 1–20 (2017).

59. Boorman, E. D., Rajendran, V. G., O'Reilly, J. X. & Behrens, T. E. Two Anatomically and Computationally Distinct Learning Signals Predict Changes to Stimulus-Outcome Associations in Hippocampus. *Neuron* **89**, 1343–1354 (2016).

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

61. Wimber, M., Alink, A., Charest, I., Kriegeskorte, N. & Anderson, M. C. Retrieval induces adaptive forgetting of competing memories via cortical pattern suppression. *Nat. Neurosci.* **18**, 582–589 (2015).

62. Abdulrahman, H. & Henson, R. N. Effect of trial-to-trial variability on optimal event-related fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis. *Neuroimage* **125**, 756–766 (2016).

63. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. A. Raincloud plots: A multi-platform tool for robust data visualization [version 1; peer review: 2 approved]. *Wellcome Open Res.* **4**, 1–41 (2019).

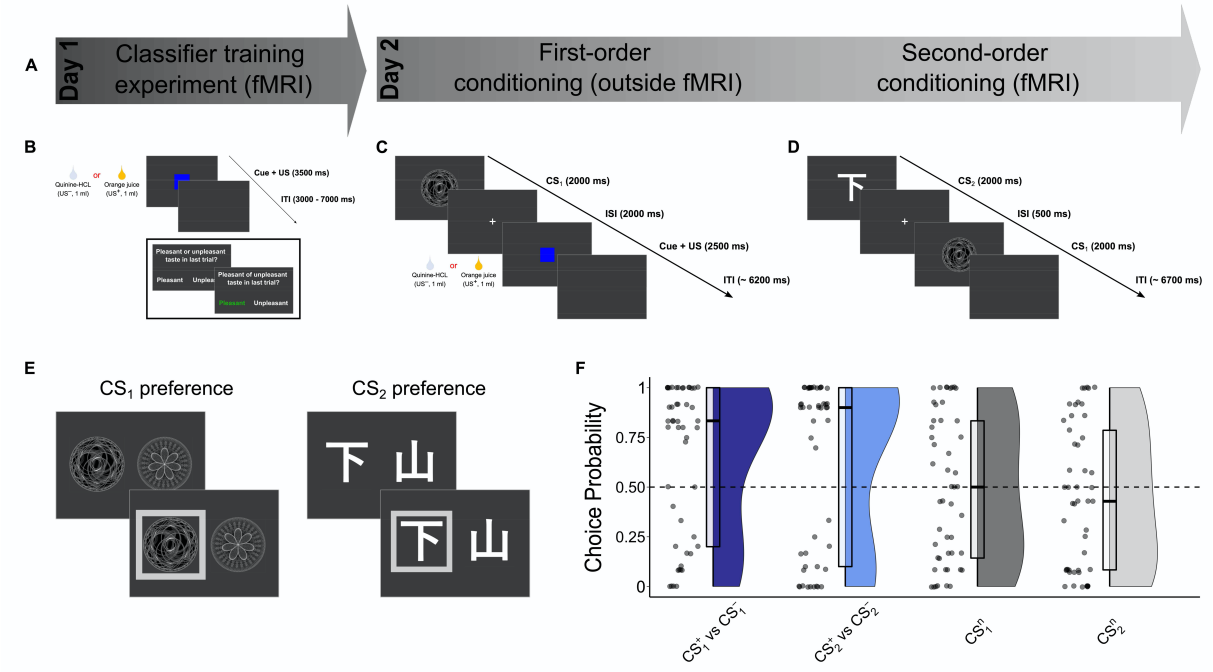


Fig. 1. Experimental procedure, task schematic and behavioral results.

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[63] showing posterior 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.

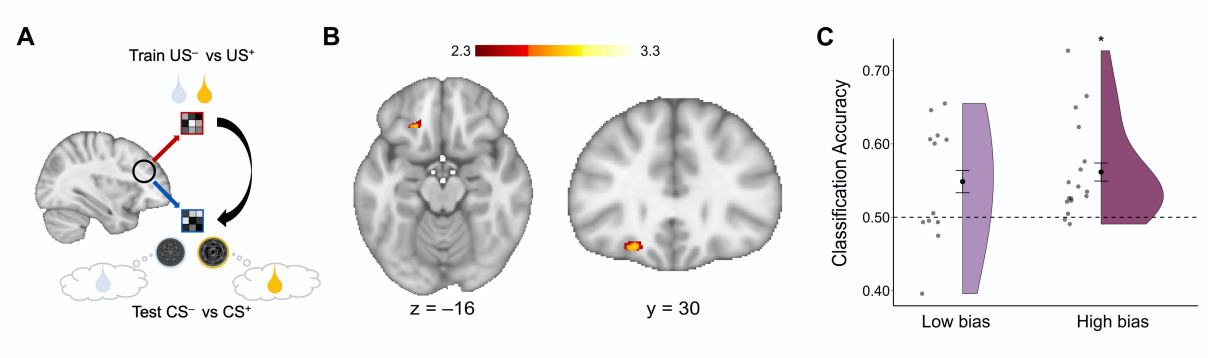


Figure 2. Cross-session, cross-modality classification analysis.

A) Analytic 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[16]. 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 = .012, one-tailed, random-effect cluster-statistics, 50,000 iterations). C) Average classification accuracy in the lOFC cluster (extracted cluster mean = .56, SD = .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 = .006, $U3_1$ = .88, one-sample Wilcoxon signed-rank test, two-tailed), but not in the low bias group (light purple raincloud, Z = 1.44, P = .151, $U3_1$ = .58, one-sample Wilcoxon signed-rank test, two-tailed). Black dots indicate sample means and error bars represent standard errors of the means of the data, respectively.

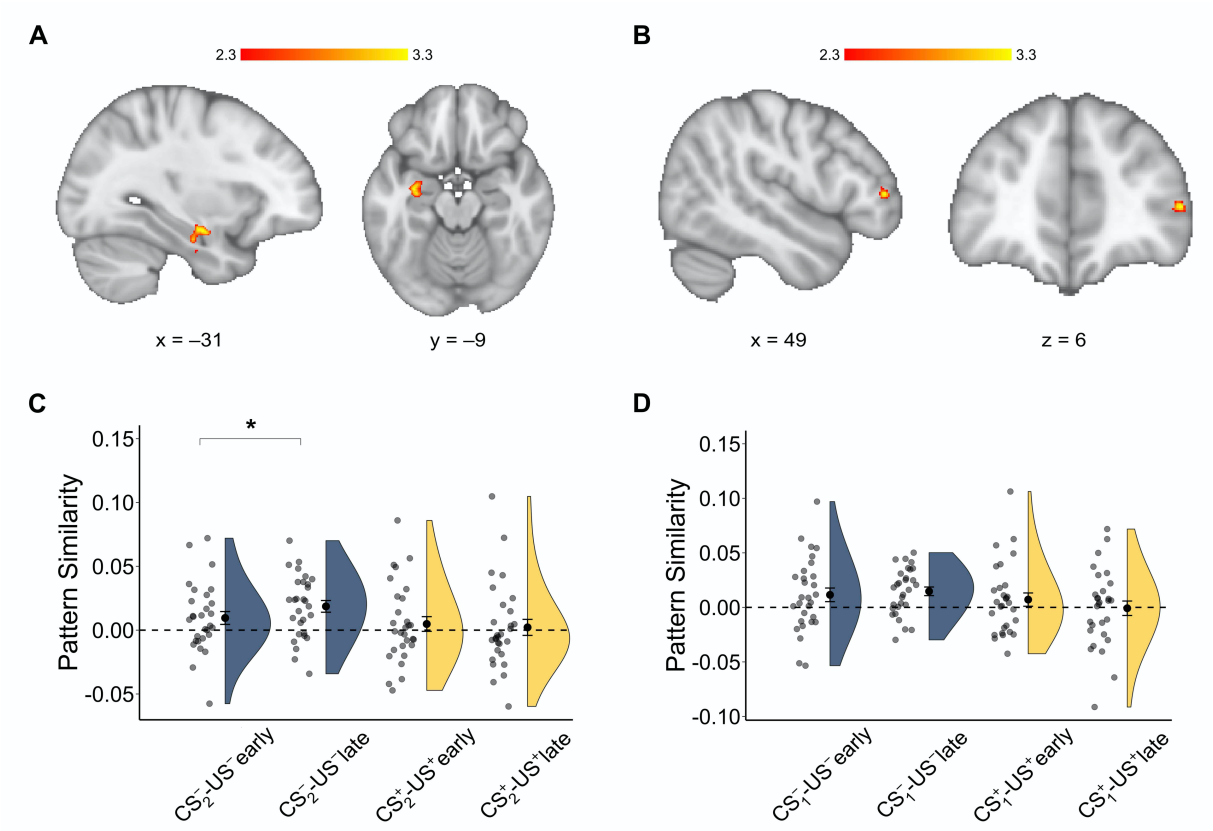


Figure 3. Psychophysiological interaction analysis and neural pattern similarity results.

A) BOLD covariation in the left lOFC and a cluster in the anterior hippocampus, extending to (basolateral) amygdala and medial temporal lobe was higher in CS^{-} and CS^{+} trials than in CS^{n} trials (peak voxel at [x = –60, y = 3, z = –24], Z = 3.97, *P* = .029, whole-brain corrected). B) A positive correlation was observed between second-order choice preferences and functional covariation between the left lOFC and a region in the right lateral prefrontal cortex (peak voxel at [x = 49, y = 43, z = 6], Z = 4.21, *P* = .034, whole-brain corrected). The more participants preferred CS_2^{+} (versus CS_2^{-}), the higher these regions' activation covaried during second-order conditioning. C & D) Template-based neural pattern similarity between classifier training experiment patterns of both US and least-squares separate (LS-S)[41] 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* = .035, Cohen's *d* = .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} = .50, *P* = .310, Cohen's *d* = .09, paired-samples t-test, one-tailed). D) There was no evidence for change in similarity for CS_1^{-}-US^{-} (t_{28} = .50, *P* = .310, Cohen's *d* = .09, paired-samples t-test, one-tailed) or CS_1^{+} and US^{+} neural pattern similarity (t_{28} = .50, *P* = .310, Cohen's *d* = .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.

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, η^2_p < .001, rmANOVA). 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, 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). 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, 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). 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, 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). Average temperature (°Celsius) of US^+ (M = 22.53, SD = 1.52) and US^- (M = 22.44, SD = 1.47) did not differ before being loaded into the syringes (t_{27} = 1.36, P = .184, paired-samples t-test).

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 = 1.94°C) and US^- (M = 21.68, SD = 1.95) did not differ before being loaded into the syringes (t_{27} = 0.61, P = .544, 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 lateral orbitofrontal cortex (Harvard-Oxford atlas), the second was obtained from an independent gustatory mapping study by Benz et al. (personal correspondence). Using random-effect cluster-statistics[1] 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 et al. (peak voxel at MNI [x = –21, y = 30, z = –17], z = 1.84, P = .033, corrected, one-tailed).

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)

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

References

1. Rizley, R.C. & Rescorla, R.A. *J. Comp. Physiol. Psychol.* **81**, 1–11 (1972).

2. Barnet, R.C., Arnold, H.M. & Miller, R.R. *Learn. Motiv.* **22**, 253–268 (1991).

3. Gewirtz, J.C. & Davis, M. *Learn. Mem.* **7**, 257–266 (2000).

4. Parkes, S.L. & Westbrook, R.F. *Rev. Neurosci.* **22**, 317–333 (2011).