Measuring human Pavlovian reward conditioning and memory retention after consolidation

Yanfang Xia[1-4*]

Huaiyu Liu[5*]

Oliver K. Kälin[1]

Samuel Gerster[1]

Dominik R. Bach[1,2,5]

[1]Computational Psychiatry Research, Department of Psychiatry, Psychotherapy, and Psychosomatics, Psychiatric University Hospital Zurich, University of Zurich, Zurich, Switzerland

[2]University of Bonn, Transdisciplinary Research Area Life and Health, Hertz Chair for Artificial Intelligence and Neuroscience, University of Bonn, Germany

[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, the Netherlands;

[4]Department of Psychiatry, Radboud University Medical Centre, Netherlands

[5]Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, University College London, United Kingdom;

[*]These authors contributed equally.

Corresponding author: Dominik R. Bach; d.bach@uni-bonn.de

Abstract

While a body of literature has addressed quantification of aversive conditioning, fewer studies have compared different conditioned response types during human reward learning and its retention after overnight consolidation. In consequence, it is unclear how to best quantify reward conditioning in humans. Here, we sought to fill this gap by investigating heart period response (HPR), skin conductance response (SCR), pupil size (PSR), and respiration amplitude response (RAR) during Pavlovian reward conditioning with a primary reinforcer. We conducted two independent experiments (N_1 = 37, N_2 = 34) with a learning phase and a recall phase 7 days later. A visual conditioned stimulus (CS+) predicted fruit juice reward (unconditioned stimulus, US), while a second CS- predicted US absence. In experiment 1, model-based analysis of HPR distinguished CS+/CS-, both during learning (Hedge's g = 0.56) and recall (g = 0.40). Furthermore, model-based analysis of PSR distinguished CS+/CS- in early trials during recall (g = 0.69). As an out-of-sample generalisation test, experiment 2 confirmed the result for HPR during learning (g = 0.78) and recall (g = 0.55), as well as for PSR during recall (g = .41). We conclude that in our Pavlovian reward conditioning paradigm, HPR is a robust measure of reward learning, while both HPR and PSR robustly index the retention of reward memory.

Keywords*:* Pavlovian reward conditioning, reward, associative learning, psychophysiological measure

1. Introduction

Learning to predict reward events plays an important role for many biological organisms. A quintessential paradigm to study this in the laboratory is Pavlovian reward conditioning (Martin-Soelch et al., 2007; Pavlov, 1927). Here, an initially neutral stimulus (conditioned stimulus, CS) is repeatedly paired with a naturally appetitive stimulus (unconditioned stimulus, US, e.g., primary vs. secondary reinforcers), and comes to elicit some behavioral or physiological responses (conditioned response, CR). In turn, this CR is taken to indicate learning the association, and when it is elicited after a delay, it is taken as indicative of reward memory. In human conditioning research, including reward and aversive conditioning (Lonsdorf et al., 2017), learning is usually indexed by the difference between responses elicited by CS+ with those elicited by a CS-, which predicts the absence of the US, in a paradigm known as differential learning.

Several psychophysiological CRs have been suggested to reflect human reward learning. Specifically, quantification of human reward conditioning has been based on skin conductance responses (SCR) (Exner et al., 2021; Klucken et al., 2019), cardiac responses (Ebrahimi et al., 2019, 2019; Hermann et al., 2000; Pietrock et al., 2019; Sayão et al., 2021; Wardle et al., 2018), startle-eyeblink (Andreatta & Pauli, 2015; Hermann et al., 2000; Stussi et al., 2018), postauricular reflex (Ebrahimi et al., 2019; Pietrock et al., 2019; Stussi et al., 2018), and pupil dilation (Bray et al., 2008; Pietrock et al., 2019; Pool et al., 2019; Prévost et al., 2013; Reinhard & Lachnit, 2002; Schad et al., 2020; Seymour et al., 2007). Notably, however, evidence for CS+/CS- differences in these observables is inconsistent between studies; but experimental protocols are also heterogenous and differ by the type of paradigm (delay or trace conditioning), the type of CS and US, including primary and secondary reinforcers, CS-US intervals, and reinforcement schedule (Exner et al., 2021; O'Doherty et al., 2003; Wardle et al., 2018). Consequently, interpretation of diverging results is difficult, which raises the question of what constitutes the most sensitive measure(s) to capture reward learning. Some previous work has directly compared different CRs, again with heterogeneous results. One study reported similarly small effect sizes in corrugator EMG, zygomaticus EMG, and SCR (Wardle et al., 2018); another one showed a large effect in HPR and smaller effects in pupil dilation, gaze patterns,

with no effect in SCR or startle eyeblink (Pietrock et al., 2019); and a third one a larger effect in SCR than in startle eyeblink (Andreatta & Pauli, 2019). Furthermore, it is unclear whether any of these results would generalise to the assessment of memory retention after overnight consolidation. This question is relevant in the context of pharmacological and non-invasive intervention studies (Ojala et al., 2022; Wehrli et al., 2023, 2024; Xia et al., 2024). It also holds significant importance for pre-clinical studies in the context of experimental psychopathology. Here, reward learning is often taken as a model of addiction symptoms, such as cue-induced drug craving (Keiflin & Janak, 2015).

Thus, the present work aimed to identify the most sensitive psychophysiological CR to quantify reward learning and retention after overnight consolidation. We employed a Pavlovian reward conditioning paradigm with primary reinforcer, in which participants gradually learned CS-US contingencies, and retention was tested after seven days. Our outcome variables were based on four observables: skin conductance response (SCR), pupil size responses (PSR), heart period response (HPR), and respiratory amplitude response (RAR).

To determine the most sensitive CR, we used a calibration approach. Put simply, this assumes a priori that participants do acquire a CS-US association, and evaluates different putative CR by their ability to reproduce this learning, as indexed by retrodictive validity (Bach, 2023; Bach et al., 2020). In our case, retrodictive validity can be expressed as the effect size to distinguish CS+/CS-. To protect ourselves against overfitting to peculiarities of small samples, we employed an exploration-confirmation approach as in previous work on these observables (Castegnetti et al., 2016, 2017; Korn et al., 2017; Xia et al., 2022). Specifically, in experiment 1 we explored different CR indices derived from all four psychophysiological observables (i.e., HPR, SCR, PSR, and RAR) in their sensitivity to distinguish CS+/CS-. We retained all indices that yielded an effect size of Hedge's $g > 0.5$, and confirmed them in experiment 2. This effect size was chosen a priori to be large enough to be usable in intervention studies, where the maximum achievable effect size is constrained by the effect size in the control group. In addition, we heuristically report smaller effect sizes in conjunction with significant p-values (without correction for multiple comparison).

2. Method

2.1. Sample size and participants

Power analysis for experiments 1 and 2 were performed using G*Power 3.1.9.7 (Faul et al., 2007). We analysed a one-sided paired *t*-test, which is appropriate since our study design involves repeated measures on the same subjects under different conditions. In result, 34 participants were needed to achieve our a priori chosen effect size of Cohen's $d > 0.5$ (which approximates Hedge's g) with 80% power and an alpha level at 0.05. We recruited 37 participants in experiment 1 to allow for drop-outs due to early termination and data exclusion. For experiment 2, power analysis showed that 32 participants were needed based on the HPR effect size reported in experiment 1 with 80% power. We recruited 34 participants in experiment 2 to allow for drop-out.

Healthy participants were recruited from the student and general population in Zurich. The governmental ethics committee approved the study (KEK-ZH-2013-0118). All participants provided informed consensus using a form approved by the ethics committee, and received monetary compensation based on experiment duration. See Table 1 for details of demographics and general information.

Table 1. Demographics and general information for experiments 1 and 2

	Phase	Experiment 1	Experiment 2
Participants completed per protocol	Learning	37 (26 women)	34 (20 women)
	Recall	37 (26 women)	33 (19 women)
Participants excluded per protocol	Learning	0	0
	Recall	0	1[a]
Age (full sample)	Learning	24.19 (4.20)	24.91 (4.11)
	Recall	24.19 (4.20)	24.94 (4.17)
Drink fasting time (hours)	Learning	7.70 (3.69)	7.46 (3.73)
	Recall	5.18 (3.05)	6.52 (3.71)
Food fasting time (hours)	Learning	11.05 (3.89)	10.84 (3.12)

	Recall	11.23 (5.43)	11.79 (3.25)
Hunger (%)	Learning	67.64 (26.06)	67.86 (20.00)
	Recall	70.38 (28.25)	78.99 (20.56)
Thirst (%)	Learning	69.85 (20.23)	67.94 (19.75)
	Recall	70.92 (23.75)	75.92 (15.92)
Favorite juice rating (%)	Learning	89.42 (13.38)	93.08 (8.89)
Selection of apple juice	Learning	8	4
Selection of mango juice	Learning	10	13
Selection of multivitamin juice	Learning	14	10
Selection of orange juice	Learning	5	7
	Recall		
Trait anxiety (STAI-T)		36.59 (8.03)	38.82 (10.43)
State anxiety (STAI-S) before learning		33.05 (6.98)	34.94 (8.89)
State anxiety (STAI-S) before recall test		31.84 (7.05)	34.00 (7.53)

Note: In columns Experiment 1 and Experiment 2, numbers are mean values with standard deviations inside parentheses, except the rows "Participants completed per protocol" and "Participants excluded per protocol". [a]One participant did not return to the recall test due to illness.

2.2. Experimental procedure

2.2.1. Overview

Both experiments followed the same procedure. Before arrival, to enhance US craving, participants were asked to fast from food for at least six hours and from drinks for at least four hours before arrival (see Table 1 for a summary of fasting statistics). Upon arrival, participants provided informed consent and were given instructions about the entire experiment. Next, participants completed the State-

Trait-Anxiety-Inventory (STAI) and watched a 4-min priming video with presentation of delicious food and drink images. After the video, they reported their food- and drink-fasting duration (Table 1) and rated their hunger and thirst levels on a visual analogue scale (VAS) with endpoints 0-100. They chose their favorite flavor among apple, orange, mango, and multivitamin juices, and then rated all four juices on a VAS with endpoints 0-100, which was consistent with the categorical selection for all participants. The chosen favorite juice was then used as US in the subsequent Pavlovian reward conditioning task (Figure 1). This conditioning task was conducted in the morning between 8:00 a.m. to 12:00 p.m. After conditioning, participants completed a computer-based questionnaire about awareness of CS-US contingency ("How likely were you to receive a sip of juice when looking at this triangle today", "How likely were you to receive a sip of juice when looking at this triangle last week?") as well as experienced arousal and valence to all CS. Seven days later, participants came back for a recall test after the same fasting procedure, and this recall test was conducted in the afternoon between 1:00 pm to 5:00 pm. They completed the state anxiety part of the STAI and watched the same priming video as describe earlier, followed by fasting time reporting and ratings for hunger and thirst. Subsequently, they completed the recall task with a follow-up computer-based questionnaire about CS-US contingency ('How likely were you to receive a sip of juice when looking at this triangle last week?') and ratings on arousal and valence to all CS conditions.

2.2.2. Pavlovian reward conditioning task

All experiments were presented via MATLAB R2021a (The Math Works; https://www.mathworks.com/products/matlab.html) using the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/Cogent). The reward conditioning task included two blocks, with 24 CS+ and 24 CS- trials per block, resulting in 96 trials in total. Trial order was randomized with the constraint that were no more than three trials with the same CS, or more than three US, in a row. In the learning phase, the CS+ was reinforced 50% of the time, whereas the CS- was never reinforced. The first CS+ trial in each block was always reinforced. In the recall task, both CS conditions were never reinforced. The instruction for the learning phase was "In this experiment, you will see differently colored triangles and receive a sip of juice now and then. You will notice that depending on the triangle you will receive a sip of

juice more or less frequently"; the instruction for the recall phase was "Today the same triangles will be presented again". In both phases, no explicit CS-US contingency instructions were given to participants.

Each trial started with a 6-s CS presentation, followed by an inter-trial interval (ITI) during which a fixation cross was presented. ITI duration was a random integer between 9-16 s. In reinforced CS+ trials, a sip of fruit juice was automatically delivered into participants' mouth 5 s after CS+ onset. To avoid artifacts in the psychophysiological recordings, participants were tasked to keep the juice in their mouth and swallow during the ITI when signaled. This signal appeared at random between 2-5 s after CS offset for 2 s. In order to keep the participants attentive during the task, they were asked to press a specific key associated with each CS. If participants did not respond, or pressed the incorrect key, error feedback was given during the first 2 s of the ITI.

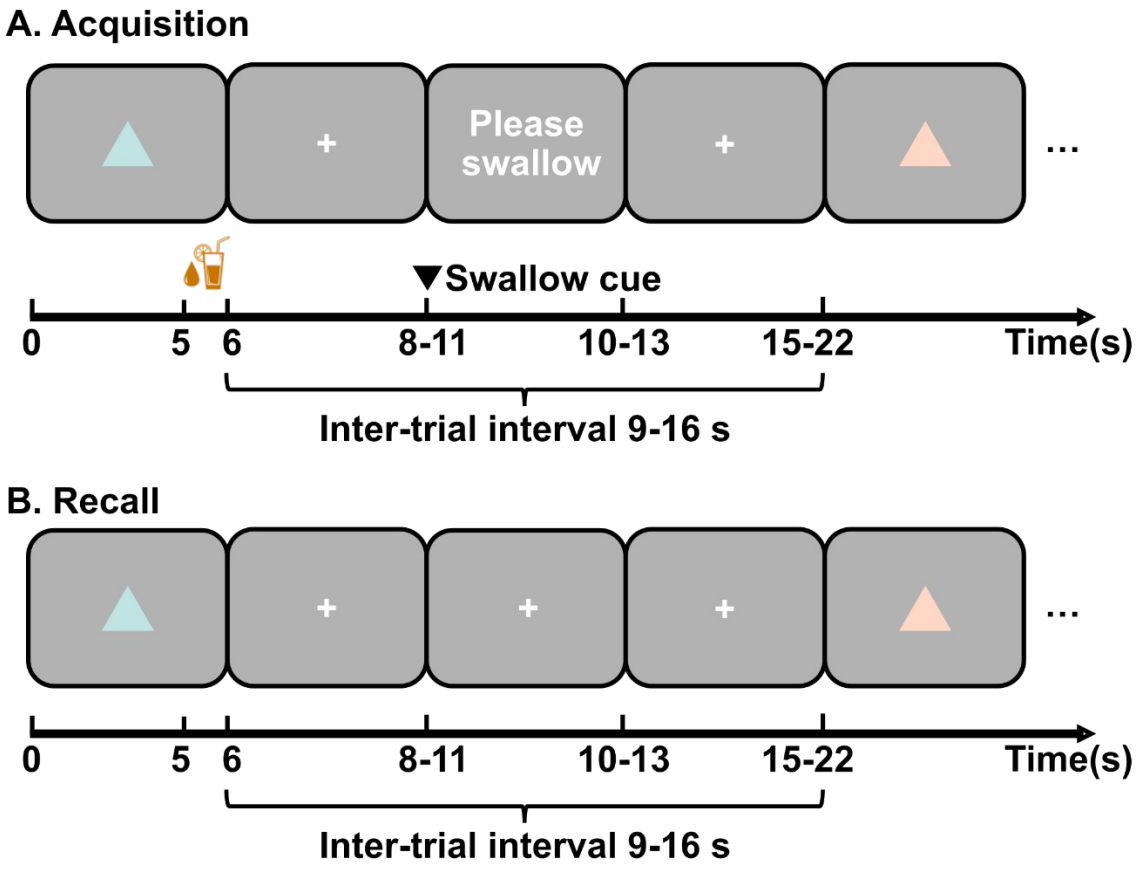

Figure.1. (A) A sample trial of the reward conditioning paradigm, where participants learned to associate a colored triangle (CS+) with a fruit juice reward

(US). The reinforcement rate was 50% and another colored triangle was never reinforced (CS-). In each trial, participants had to press the right or left arrow key to indicate the presented CS. Participants were informed about wrong or no key presses during the first 2 s of the inter-trial interval (ITI). The signal to swallow was presented at a random time between 2 – 5 s after CS offset for 2 s. The duration of the ITI was randomly sampled between 9 – 16 s. (B) A sample trial of the recall phase, which is identical to the learning phase except that there was no US. In other words, the recall test was conducted under extinction learning.

2.2.4. Stimuli and apparatus

The 4-min priming video consisted of 63 appetitive food and drinks images; each image was displayed for 4 s. All images were selected from an online image repository (https://unsplash.com/t/food-drink) with a CC0 license.

Red and blue isoluminant triangles (RGB: [0.753, 0.894, 0.894], [1, 0.843, 0.776]) were randomly allocated to CS+ and CS- across participants. Both CS were presented for 6 s at the center of an isoluminant grey computer screen (RGB: [175, 175, 175], screen size 318 mm x 256 mm) with a size of 3° visual angle at 67.2-cm distance from the participant (figure 1), During the ITI, a white (RGB: [255, 255, 255]) fixation cross (0.8° visual angle) was presented at the center of the same grey background screen.

An automatic pump (AL-8000 Syringe Pump, World Precision Instruments, Sarasota FL, USA) dispensed US via a 5-m PVC tube (Faust Laborbedarf AG, Schaffhausen, CH) with an inner diameter of 2 mm and an outer diameter of 4 mm. This tube was positioned in such a way that it sat easily within the participants' mouths and was affixed to a chin rest, on which participants were asked to place their chins throughout the tasks.

2.2.5 Recording of psychophysiological indices

We collected ECG data with three pre-gelled Ag/AgCl adhesive snap electrodes (01-7500, TIGA-MED; FS-TC1, Skintact; and EL503, Biopac Systems Inc.) attaching to the outsides of wrists and right ankle, respectively. For each participant, we used and recorded the lead configuration that yielded the clearest R spikes (ECG100C, Biopac Systems Inc.). To record skin conductance, two 8-mm disk Ag/AgCl cup electrodes (EL258, Biopac, Goleta, CA) filled with 0.5% NaCl gel

(GEL101, Biopac; Hygge & Hugdahl, 1985) were applied to the thenar and hypothenar eminence of the non-dominant hand and connected to a constant voltage coupler/amplifier (EDA100C, Biopac). We measured respiration using a single-belt cushion system (RSP100C, Biopac). All data were amplified and digitised (MP160, Biopac), and recorded with Acknowledge (version 5.0, Biopac). Pupil size and gaze direction were recorded for both eyes with an Eyelink 1000 system (SR Research, Ottawa, ON, Canada) at a sampling rate of 500 Hz in a distance of 52.2 cm, after gaze calibration using the manufacturer's 9-point procedure.

2.3. Preprocessing and modeling of psychophysiological indices

We analysed all data in the framework of psychophysiological modelling (PsPM). Additionally, we report peak-scoring analyses of pupil size responses and grand means of intra-trial time-courses.

2.3.1. Psychophysiological modelling

Psychophysiological (forward) models mathematically describe how a neural input generates a peripheral physiological response (Bach et al., 2018). The amplitudes of the input into this system are assumed to reflect the psychological latent variable, which in the context of associative learning is the CS-US association. Given a psychophysiological model, the latent variable can be estimated from physiological data by means of model inversion (Bach et al., 2018). Similar to previous work on these observables (Bach, Flandin, et al., 2010; Castegnetti et al., 2016, 2017; Korn et al., 2017), we constructed our psychophysiological models as linear time invariant systems (LTI), which are fully defined by their response functions (RF). Two approaches are often used to construct the RF for an LTI. On the one hand, one may formalize the RF from identified biophysical relations between inputs and outputs. This approach assumes that researchers already understand the underlying (Friston et al., 2000). On the other hand, one may construct a phenomenological RF from the empirical data, even if the underlying biophysical systems are unknown (Castegnetti et al., 2016). In the present work, we employed the second approach. For SCR, the existing phenomenological model is split into an invariant peripheral system that does not depend on experimental paradigm, and a flexible model of the neural input, which can be estimated from data. The peripheral model has been validated by direct intraneural recordings

(Gerster et al., 2018), such that there was no need to develop a new RF for SCR. On the other hand, existing RF for HPR, PSR, and RAR collapse paradigm-specific central processes and the peripheral system, such that they are not necessarily applicable to the current experimental paradigm. Thus, we used data from experiment 1 to develop RF for these modalities.

Once the shape of the RF is defined, the next step is to estimate the system's input to best explain data. To obtain input amplitude estimates, we inverted general linear convolution models (GLM) to fit the predicted timeseries (obtained through the convolution of the RF with a constant input shape) to the empirical data timeseries (Bach et al., 2018). The GLMs are either trial-wise or condition-wise, depending on the modality of the data. Trial-wise GLMs were used for PSR, which have a time course that does not overlap between trials (Korn et al., 2017), whereas condition-wise GLMs were used for HPR and RAR. For SCR, we conducted trial-wise estimation using the non-linear model in PsPM (Bach, Daunizeau, et al., 2010). See the section 2.3.3 below for details.

All preprocessing and modeling of psychophysiological data were conducted using the PsPM toolbox 4.1.1 (https://bachlab.github.io/PsPM/) (Bach et al., 2018) in MATLAB R2021a.

2.3.2. Heart period responses (HPR)

Heart beats were detected in ECG signal by a modified version of the Pan-Tompkins algorithm (Pan & Tompkins, 1985) as implemented in PsPM (Paulus et al., 2016). The presence of artefacts was further controlled by visually inspecting each participant's time-series, and removing artefacts due to clipping, movement, or electrode detachment. For each detected heart beat, we computed the preceding inter beat interval. Inter beat intervals corresponding to a heart rate outside 50-150 beats per minute were automatically excluded. In line with previous work (Paulus et al., 2016), our analyses were based on heart period rather than heart rate, because heart period and autonomic input are linearly related in stimulation studies (Cacioppo et al., 2007). The remaining data points were linearly interpolated in chronological time at 100 Hz and filtered with a 4^{th}-order bidirectional band-pass Butterworth filter (cut-off frequencies: 0.015–0.5 Hz) as in previous work (Castegnetti et al., 2016). To build the RF, we extracted trial-wise segments, and baseline-corrected single-trial responses by subtracting the heart

period average during 5 s before the CS onset (Pollatos et al., 2007). Afterwards, responses were averaged first within each condition, and then over participants. In line with previous work (Castegnetti et al., 2016), we fitted the difference between the mean over all CS+ and the mean over all CS- trials with a gamma probability density function (= 1.72, [s^{-1}] = 0.14, = 60.10, [s] = -17.61). As the duration of typical HPR is much longer than the CS-US interval, only CS+ trials not reinforced by a US were considered for modelling and analysis.

2.3.3. Skin conductance responses (SCR)

SCR data quality was assessed by the SCR preprocessing function implemented in PsPM. Raw data outside 0.05-100 µS or with an absolute slope over 10 µS/s were automatically marked as missing data. The presence of artefacts was further controlled by visually inspecting each participant's SCR time-series and removing artefacts due to clipping, movement, or electrode detachment. All such missing data were linearly interpolated for filtering and removed from analysis. Data were then filtered with a bidirectional 1st-order band-pass Butterworth filter with the cut-off frequencies 0.0159-Hz and 5-Hz and downsampled to 10 Hz.

We then estimated conditioned and unconditioned responses using a non-linear model implemented in PsPM (Bach, 2014; Bach & Melinscak, 2020). We modeled a response evoked by CS onsets with fixed latency, and one evoked by US or US omission. Amplitude estimates were normalized by dividing through the average of all CS- trials from the corresponding participant.

2.3.4. Pupil diameter

Pupil diameter data were converted to metric units and pre-processed with the algorithm implemented in PsPM (Kret & Sjak-Shie, 2019). This algorithm excludes data points outside the biological range of pupil size and its time derivative. Furthermore, it excludes isolated data points, outliers, and data points at the beginning and the end of temporal gaps, interpolates the data, and combines data from both pupils. Next, data points were excluded if gaze point deviated more than 5° visual angle from screen center (Korn et al., 2017). We then corrected for the pupil foreshortening error (Hayes & Petrov, 2016) and downsampled to 100 Hz after a low-pass filter with a cutoff of 50 Hz. We developed and tested several RF (Table 2). For all models, pupil diameter time series for each participant and each

block were z-scored. Then, the first 3 (models 3, 4, and 5), 3.5 (models 2 and 6) or 15 (model 1) s of each trial were extracted, and each timepoint averaged first within condition, then over all participants. In line with previous work (Korn et al., 2017), we fitted the difference between the mean over all CS+ and the mean over all CS- trials. To avoid contamination of the PSR by overlapping US responses, we only considered CS+ trials that were not reinforced by a US for RF1 that extended beyond the CS-US interval. For RF2-6, we used all trials. At the request of an anonymous reviewer, we compared this approach to standard peak-scoring analysis. Specifically, we first extracted the preprocessed pupil size within each trial from CS onset to US onset, which spans a duration of 5 seconds. Next, we subtracted the baseline value from the pupil data. The baseline is defined as the mean pupil size, excluding any missing values, during the 1-second period preceding each CS onset. From this baseline-corrected pupil data we took the maximum value as the peak-scored pupil dilation of a trial. Trials with unavailable baseline (i.e., with all missing values during the 1-second pre-CS period) were marked as missing data(Finke et al., 2021).

Table 2. Pupil size response functions

Model	Type	Specification
RF1	Gamma	= 3.534, [s^{-1}] = 1.946, = -1.183, [s] = 1.712
RF2	Gamma	= 30.781, [s^{-1}] = 0.042, = 0.033, [s] = 0.506
RF3	Gaussian	[s] = 1.784, = 0.246, = 0.035
RF4	Low-Pass Filtered	Bidirectional Butterworth 2 Hz Low-Pass Filter of the first 2.88 seconds of the mean difference of CS+ and CS-
RF5	Mixture of two Gammas	= 12.998, = 1.355, = 0.750, = 0.059, = 0.178, = 2.715, = 0.439

RF6	Difference of the mean prediction of 20 Gaussian Processes fit to all CS+ and CS- trials separately	CS+: Matern kernel: = 1.5, = 1.031 (*sd* = 0.000), White noise kernel: 0.922(*sd* = 0.005) CS-: Matern kernel: = 1.5, = 1.031 (*sd* = 0.000), White noise kernel: 0.903 (*sd* = 0.007)

Note: RF refers to response function.

2.3.5. Respiration amplitude (RAR)

To extract respiration amplitude from the raw time series, we used an established algorithm implemented in PsPM which automatically detects respiration cycles (Bach et al., 2016). The respiration amplitude values were interpolated at 10 Hz, and band-pass filtered with a bidirectional Butterworth filter, with 2-Hz low-pass and 0.01-Hz high-pass cutoffs. As for the other modalities, and in line with previous work (Castegnetti et al., 2017), we fitted the difference between the mean over all CS+ and the mean over all CS- trials with a gamma pdf (= 40.87, [s^{-1}] = 0.29, = 0.14, [s] = 2.09).

2.4. Statistical analysis

All data analysis was conducted in MATLAB R2021a (The Math Works, https://www.mathworks.com/products/matlab.html) and statistical software R 4.2.1 (Ihaka & Gentleman, 1996).

2.4.1. Exclusion criteria

For SCR, HPR, and RAR, we excluded participants if their estimated data were outside three standard deviations around the corresponding condition group mean. For PSR, trials with unreasonable pupil dilation estimates (estimates exceeding ±6 mm) (Spector, 1990) or with more than 50% missing values were excluded, and participants were excluded if they had more than 50% trials removed. Please see supplemental materials (Table S1) for a summary of number of excluded trials and participants for each measure in each phase for each experiment.

2.4.2. Data analysis

For all models in which the RF overlapped with the US presentation (HPR, HPR, RAR, PSR RF1), we retained data from CS- trials and non-reinforced CS+ trials only (to avoid biasing the estimated CRs by the US response). Next, we obtained condition-wise estimates by averaging data across all trials within each CS condition. Finally, we performed pairwise *t*-tests to examine the CS+/CS- difference.

As the recall test was done without US reinforcement, CS+/CS- differences are likely to extinguish over the course of the recall test. Thus, including all trials into the analysis might reduce a CS+/CS- difference seen in early trials. On the other hand, including fewer trials might increase the impact of trial-by-trial variation due to experiment-unrelated factors, and the optimal balance is difficult to intuit. Hence, for data available on a trial-by-trial basis, we approached this in a data-driven way by analysing the condition average over 1…n trials, with n ranging from 1 to the number of trials per condition. Similarly, at least for some observables, it is speculated that CS-US association is learned relatively quickly, but that the CS+/CS- differences might decay over time (Tzovara et al., 2018). Hence, we did a similar analysis for the learning phase, excluding the first pair of trials. For SCR and PSR BF1, where we retained only non-reinforced CS+ trials, we would average over 1..n CS+ trials and 2..2n CS- trials, where n ranges from 1 to the number of non-reinforced CS+ trials. For PSR (with the exception of BF1), CS+ refers to both reinforced and non-reinforced CS+ trials, for which the estimated CR do not overlap with US presentation. Finally, for observables unavailable on a trial-by-trial basis, we analysed the first and the second half of the phases separately.

We computed effect sizes to compare and find the optimal psychophysiological measure(s) for reward learning. For all models the Hedge's *g* was computed using the following formula (Lakens, 2013).

2.5. Data and code accessibility

Anonymised data, experimental materials (stimuli and MATLAB scripts), and scripts of data-analysis will be made available (on Zenodo and Open Science Framework) for a revision, and will be made publicly available upon acceptance of

the manuscript. An updated heart period response function, fitted on data from both experiments, is available in the PsPM toolbox from version 7 onwards as pspm_bf_hprf_rew (https://bachlab.github.io/pspm).

3. Results

3.1. Experiment 1

Grand means for the intra-trial time course, as well as the two relevant RF, are shown in figure 2. In the learning phase, the condition difference between CS+ and CS- exceeded our a priori effect-size threshold (Hedge's $g > 0.5$) for the HPR index only (Table 3, figure 3). For PSR, indices from RF4 and RF5 were close to the threshold with g-values around 0.4, while there was no apparent CS+/CS- difference for SCR and RAR (Table 3). Notably, this analysis is biased for HPR, PSR and RAR because it is based on RF fitted to the same data sets. Finally, there was no CS+/CS- difference in peak-scoring analysis of PSR. Analysis of trial-by-trial responses revealed no subsets of trials that exceeded the threshold in any PSR metric or SCR, and there were no additional insights from analysing first and second half of the learning phase for HPR and SCR.

For the recall phase, we found an above-threshold condition difference between CS+ and CS- for HPR (Table 3, figure 3). Because the RF was fit on data from the learning phase only, this was an unbiased analysis and therefore suggests that participants indeed retained reward memory. When averaging over all trials, there were no CS+/CS- differences in any other metric. However, trial-by-trial analysis of PSR revealed large effects in the first part of the recall phase. The largest effect size was achieved by PSR RF5 for the first 7 CS+ and CS- trials ($g = 0.68$) (figure 3). In SCR, effect sizes up to around $g = 0.35$ were observed in the first 3-4 CS+ and CS- trials. See supplementary material for trial-by-trial results for all metrics from the recall phase.

Table 3. Results of pairwise *t*-tests for HPR, SCR, PSR, and RAR

Phase	Condition Comparison	HPR	SCR	PSR RF1	PSR RF2	PSR RF3	PSR RF
Learning (Exp1)	CS+	7.69 (26.17)	0.96 (0.15)	0.01 (0.08)	0.05 (0.16)	0.05 (0.16)	0.10 (0.17)
	CS-	-9.34 (16.73)	0.95 (0.11)	-0.02 (0.07)	0.02 (0.14)	0.03 (0.14)	0.05 (0.15)
	CS+ vs. CS-	$t(36) = 3.51$, $p = .001$**, $g = 0.56$	$t(36) = 0.42$, $p = .68$, $g = 0.07$	$t(33) = 1.62$, $p = .12$, $g = 0.27$	$t(33) = 1.53$, $p = .14$, $g = 0.26$	$t(33) = 1.19$, $p = .24$, $g = 0.20$	$t(33) = 2.24$, $p = .03$ $g = 0.3$
Recall (Exp1)	CS+	-6.42 (15.47)	0.93 (0.21)	-0.03 (0.07)	0.04 (0.14)	0.04 (0.15)	0.11 (0.16)
	CS-	-14.07 (14.51)	0.92 (0.18)	-0.03 (0.08)	0.03 (0.10)	0.03 (0.10)	0.09 (0.15)
	CS+ vs. CS-	$t(36) = 2.48$, $p = .02$*, $g = 0.40$	$t(36) = 0.50$, $p = .62$, $g = 0.08$	$t(34) = -0.38$, $p = .71$, $g = -0.06$	$t(34) = 0.19$, $p = .85$, $g = 0.03$	$t(34) = 0.24$, $p = .81$, $g = 0.04$	$t(34) = 0.91$, $p = .37$ $g = 0.1$
Learning (Exp2)	CS+	16.57 (22.90)	0.97 (0.11)	0.02 (0.11)	0.03 (0.14)	0.03 (0.14)	0.06 (0.19)
	CS-	-3.29 (11.26)	0.99 (0.03)	0.01 (0.07)	0.01 (0.10)	0.01 (0.10)	0.04 (0.13)
	CS+ vs. CS-	$t(32) = 4.56$, $p < .001$***, $g = 0.78$	$t(33) = -0.84$, $p = .41$, $g = -0.14$	$t(31) = 0.41$, $p = .68$, $g = 0.07$	$t(30) = 0.83$, $p = .41$, $g = 0.15$	$t(30) = 1.23$, $p = .23$, $g = 0.22$	$t(31) = 0.76$, $p = .45$ $g = 0.1$
Recall (Exp2)	CS+	1.66 (16.31)	0.93 (0.14)	-0.02 (0.06)	0.02 (0.12)	0.02 (0.12)	0.04 (0.14)

	CS-	-7.96 (13.55)	0.97 (0.08)	-0.03 (0.08)	0.02 (0.11)	0.02 (0.11)	0.04 (0.14)
	CS+ vs. CS-	$t(32)$ = 3.26, p = .002**, g = 0.55	$t(32)$ = -1.67, p = .10, g = -0.28	$t(31)$ = 0.58, p = .57, g = 0.10	$t(31)$ = 0.08, p = .93, g = 0.01	$t(31)$ = 0.25, p = .81, g = 0.04	$t(31)$ = 0.13, p = .90 g = 0.0

Notes: In the Measure column, RF refers to response function; The numbers in the columns CS+ and CS- are group means with standard deviations inside the parentheses; In the t-value column, the numbers inside the parentheses are degrees of freedom; CS+ refers to both reinforced and non-reinforced CS+ trials in columns PSR RF2-7 only. *$p < .05$; **$p < .01$; ***$p < .001$. P-values are uncorrected and presented for illustration only; our a priori decision criterion to retain an index was the effect size Hedge's g value as shown above.

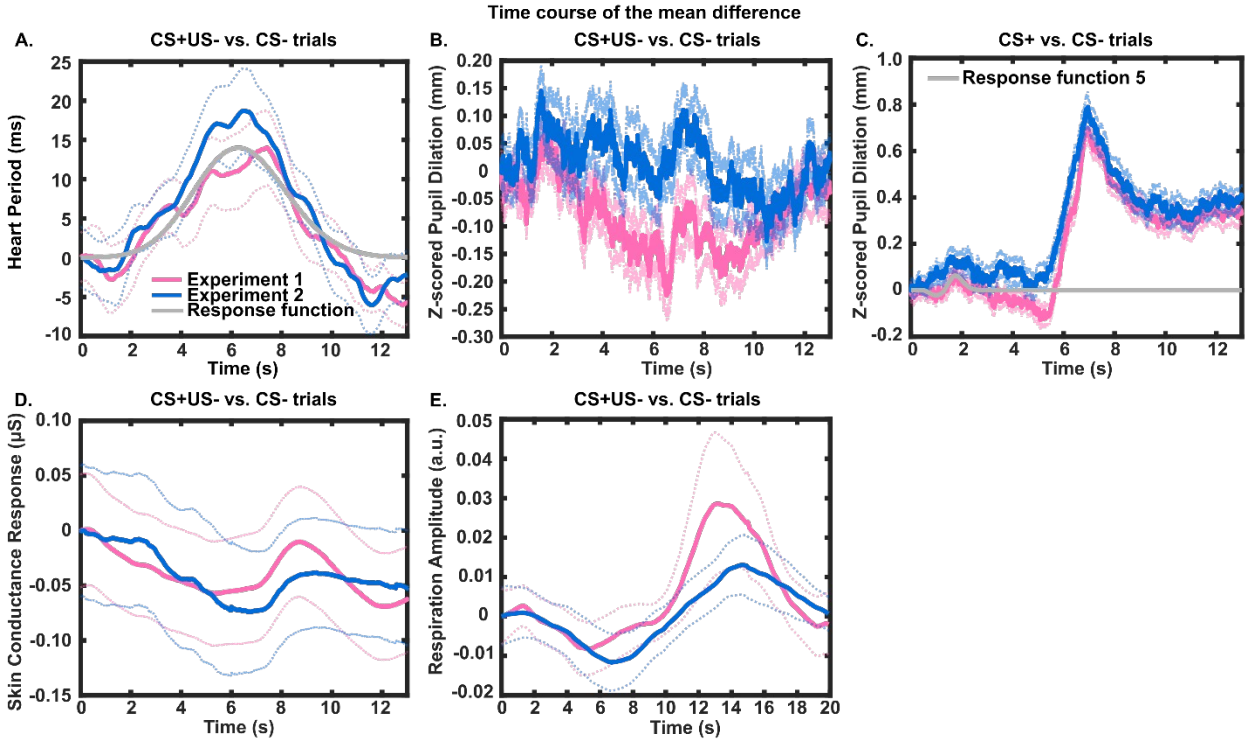

Figure 2. Responses to four modalities and relevant response functions. All panels show the condition differences between CS+ and CS- trials during learning for experiments 1 (pink) and 2 (blue). (A) HPR and corresponding RF (non-reinforced CS+ trials). (B) PSR (non-reinforced CS+ trials, used for RF1). (C) PSR (all CS+ trials, used for RF2-6 and peak scoring) and RF5. (D) SCR (non-reinforced CS+ trials), and (E) RAR (non-reinforced CS+ trials). Solid lines indicate mean and dotted lines represent ± SEM. Data from experiment 2 are shown for illustration only. The start of the X-axis represents trial onset (i.e., CS onset); US onset is after 5 s.

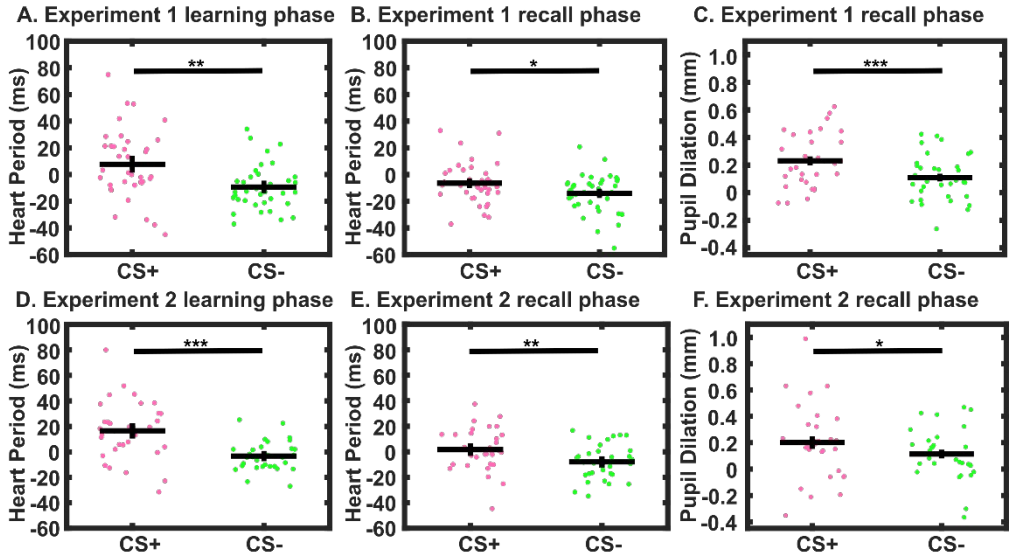

Figure 3. Condition-wise mean heart period responses for learning phase (A) and recall phase (B) in experiment 1, and learning phase (C) and recall phase (D) in experiment 2; condition-wise mean PSR (RF5) for recall phase (C) in experiment 1 with a subset of 7 trials, and recall phase (F) in experiment 2 with a subset of 7 trials. Note that CS+ refers to non-reinforced CS+ trials without US presence. The statistical test in A is biased (RF generated from the same data), but tests in B, D, and E are unbiased (RF generated from learning data in experiment 1). Dots represent individual condition-wise mean HPR (A, B, D, and D) and condition-wise mean PSR (C and F). Black crosses represent condition mean ± SEM. *$p < .05$; **$p < .01$; ***$p < .001$ (uncorrected).

3.2. Experiment 2

We analysed experiment 2 with RF developed in experiment 1, so this represents an unbiased out-of-sample generalisation analysis. For HPR, the condition differences between CS+ and CS- were significant for learning and recall after Holm-Bonferroni correction for multiple comparison across two tests (learning and retention), see Table 3 and Figure 3 for details. For PSR, the effects observed in the learning phase of experiment 1 in RF4/5 were not replicated; effect sizes were $g < 0.15$ when averaging over all trials. The trial-by-trial analysis that yielded the largest effect size in experiment 1 (PSR RF5 for the first 7 CS+ and CS- trials) also yielded the largest effect size in experiment 2 ($g = 0.41$, $p < .05$, figure 3, see also

supplementary material for details. Trial-by-trial results observed for SCR in experiment 1 were not replicated in experiment 2.

3.3. Subjective ratings

Tables 4 displays the summary statistics for post-experiment questionnaire data in experiments 1 and 2. Overall, the data patterns are consistent across these two experiments. The contingency awareness and valence were higher for CS+ (vs. CS-) in both learning and recall phases, the arousal was larger for CS+ (vs. CS-) in the learning phase only.

Table 4. Summary statistics and results of paired *t*-tests for questionnaire data in experiments 1 and 2

Experiment	Phase	Condition Comparison	Contingency Awareness (%)	Arousal Rating (%)	Valence Rating (%)
1	Learning	CS+	45.92 (32.66)	54.36 (29.41)	58.59 (22.38)
		CS-	10.85 (21.99)	26.12 (23.55)	38.80 (19.18)
		CS+ vs. CS-	$t(36) = 17.50$, $p < .001$***	$t(36) = 10.36$, $p < .001$***	$t(36) = 3.28$, $p < .01$**
1	Recall	CS+	48.18 (33.47)	2.99 (7.62)	35.28 (26.12)
		CS-	13.73 (19.44)	1.33 (4.53)	12.60 (14.75)
		CS+ vs. CS-	$t(36) = 9.86$, $p < .001$***	$t(36) = 2.01$, $p = .05$	$t(36) = 4.59$, $p < .001$**
2	Learning	CS+	42.73	50.88	59.69

			(33.70)	(29.83)	(21.26)
		CS-	6.37 (11.42)	21.39 (14.35)	32.85 (18.05)
		CS+ vs. CS-	$t(33) = 16.48$, $p < .001$***	$t(33) = 11.08$, $p < .001$***	$t(33) = 4.69$, $p < .001$***
2	Recall	CS+	44.71 (33.61	1.12 (3.13)	31.35 (27.89)
		CS-	9.41 (24.08)	4.24 (12.87)	25.17 (21.84)
		CS+ vs. CS-	$t(32) = 12.55$, $p < .001$***	$t(32) = -.82$, $p = .42$	$t(32) = 6.30$, $p < .001$***

Notes: The numbers in the columns Contingency Awareness, Arousal Rating, and Valence Rating are condition means and standard deviations inside the parentheses. *$p < .05$; **$p < .01$; ***$p < .001$ (uncorrected).

4. Discussion

Pavlovian reward conditioning is an important basic learning paradigm but its optimal quantification in humans remains unclear, in particular for recall after overnight consolidation. Here, we pitched different CR against each other in the same experimental paradigm. We used a rigorous exploration-confirmation approach to establish the best psychophysiological indices for measuring reward learning and memory retention over seven days. Our key finding is that among candidate psychophysiological indices (based on HPR, SCR, PSR, and RAR), model-based HPR analysis distinguished CS+/CS- in the learning phase across both experiments, with Hedge's $g = 0.56$ in experiment 1 and $g = 0.78$ in experiment 2. It also distinguished CS+/CS- in the recall test after 7 days, with $g = 0.40$ in experiment 1 and $g = 0.55$ in experiment 2. Furthermore, model-based PSR analysis distinguished CS+/CS- in the recall phase with largest effect size when

averaging over 7 trials in experiment 1 (g = 0.69) and experiment 2 (g = 0.41). Thus, HPR and PSR as analysed with our new response functions appear to be replicable measures of classical reward conditioning and memory retention over 7 days.

4.1.1. HPR discriminated CS+/CS- difference in learning and recall

HPR has been examined in previous reward conditioning work, however with conflicting results. Our present observation of reward-conditioned bradycardia replicates a previous study (Pietrock et al. 2019). This work employed a highly similar Pavlovian reward conditioning paradigm, and analyzed HPR using PsPM's general linear convolution model (GLM). In contrast, in other previous work, cardiac responses did not discriminate CS+/CS- (Ebrahimi et al., 2019; Exner et al., 2021; Hermann et al., 2000; Sayão et al., 2021; Wardle et al., 2018). There appear two main differences between these studies and ours. First, these studies used heart rate as the CR and/or analyzed HPR using different approaches (e.g., mean change, heart index, and mean level) (Ebrahimi et al., 2019; Sayão et al., 2021; Wardle et al., 2018). Compared to heart rate, HPR has been shown to linearly relate to neural input into the heart and is therefore more likely to linearly relate to psychological variables (Bach & Melinscak, 2020; Berntson et al., 1995). Also, work on fear conditioning suggests that a model-based approach might be more sensitive to discriminate CS+/CS- differences based on HPR than peak-scoring analysis (Castegnetti et al., 2016; Paulus et al., 2016). Hence, the observed null results in these studies may be due to the selection of CR index. Second, in some studies using odor as the US, a conditioning effect was not only absent for the HPR, but also for any other psychophysiological measures (Exner et al., 2021; Hermann et al., 2000). A potential interpretation is that participants simply did not learn, possibly due to reduced associability of CS and odour (Kokkola et al., 2019).

Another interesting question is whether reward and fear conditioning affect HPR differently. Descriptively, the effect size for reward-conditioned HPR (Cohen's d = 0.79 in learning phase of experiment 2) is smaller than that of fear-conditioned HPR (Cohen's d = 0.97) (Bach & Melinscak, 2020). Also, the response functions of these two differ from one another (RF for reward conditioning: = 1.72, [s^{-1}] = 0.14, = 60.10, [s] = -17.61; RF for fear conditioning: = 48.5, [s^{-1}] = 0.182, = 1, [s] = -7.36) (Figure 4). Compared to reward conditioning, the fear-conditioned responses

appear narrower and returns to baseline more quickly. This could suggest a more vigilant preparatory reaction to successive behavioral response, which might be evolutionarily adaptive (Andreatta & Pauli, 2015). Finally, there is evidence that fear-conditioned bradycardia is time-locked to US onset, that is, the onset of fear-conditioned bradycardia moves as the time point of possible US delivery changes after CS onset (Castegnetti et al., 2016). In the present work, CS-US interval was not varied, such that we could not investigate this question for reward conditioning. Overall, it remains unclear whether reward- and fear-conditioned HPR operate through different mechanisms. Fear-conditioned bradycardia has been linked to behavioural freezing (immobility), which might have adaptive value in certain defensive situations (Roelofs & Dayan, 2022), although we note other work has demonstrated that freezing can co-occur with tachycardia as well as with bradycardia (Signoret-Genest et al., 2023), such that this link remains speculative.

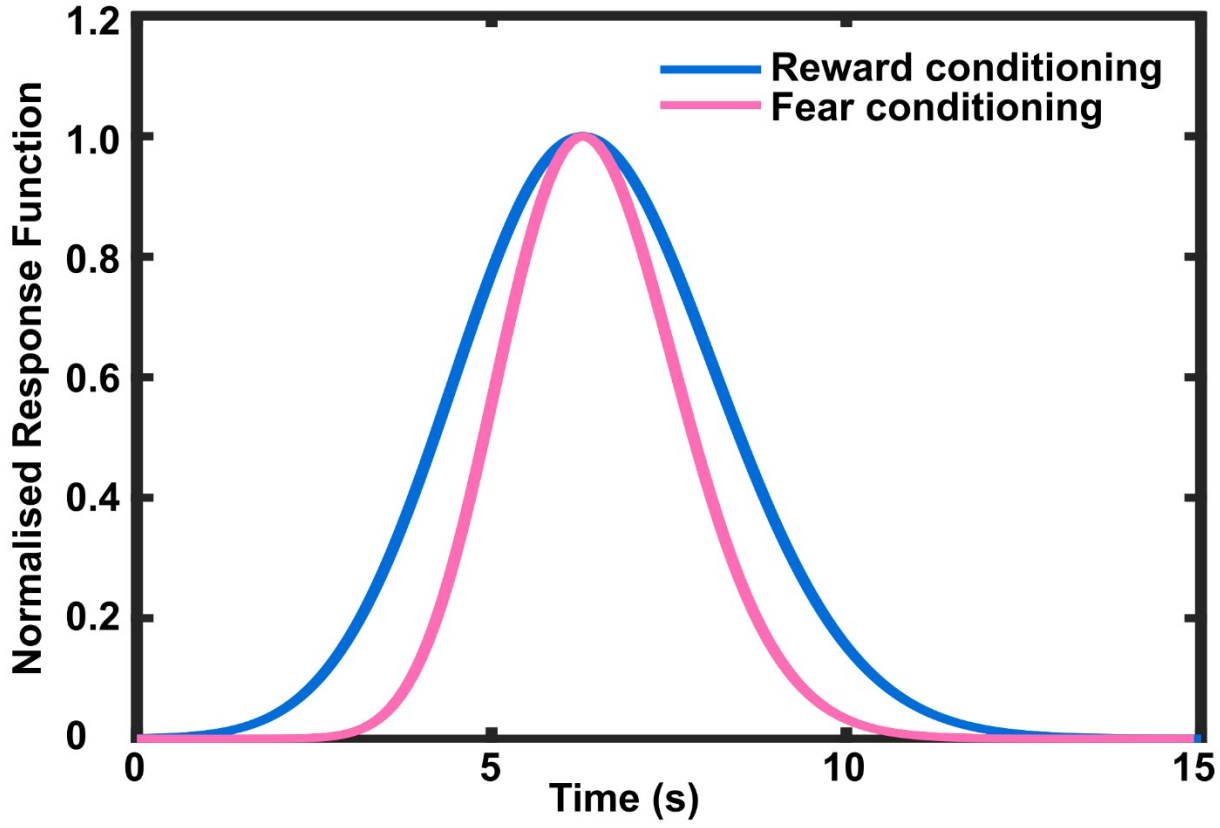

Figure 4. Heart period response function (normalised) with the same SOA (5 s) for reward and fear conditioning, respectively.

Our recall test after 7 days revealed retention of the conditioned HPR in both experiments. To our best knowledge, no prior studies have assessed the retention of human reward learning memory after a delay of several days. The recall phase can be viewed as an extinction training due to the lack of reinforcers. Hence, it is interesting that there appears to be only limited extinction for reward learning. Several reasons might be plausible. First, previous research on fear learning found that HPR is more resistant to extinction compared to SCR and saccadic scanpath length (Xia et al., 2021). It is possible that in the current paradigm, the number of trials (96) is not sufficient to extinguish reward conditioning. Another potential explanation is that HPR becomes habitual during learning and looses its dependence on US outcome predictions (Pool et al., 2019).

4.1.2. SCR did not discriminate CS+/CS- difference in learning

Interestingly, we did not find evidence that SCR discriminated CS+/CS- difference during the learning phase, although this measure was indeed sensitive to reward conditioning in most previous work (Andreatta & Pauli, 2015, 2019; Ebrahimi et al., 2019; Klucken et al., 2019; Kruse et al., 2017, 2020; Tapia León et al., 2018; van den Akker et al., 2017; Wardle et al., 2018). Upon a closer look, it seems that the US type and response quantification approach influenced whether SCR discerned CS+/CS- difference. There were two main US types (primary reinforcers such as snacks, fruit juice; secondary reinforcers such as monetary reward) and two response quantification approaches (model-based approach such as psychophysiological models, non-model-based approach such as through-to-peak approach) in previous work (Andreatta & Pauli, 2019; Ebrahimi et al., 2019; Kruse et al., 2020; van den Akker et al., 2017). All previous work employing non-model-based approaches, particularly the through-to-peak and baseline-to-peak methods, have found CS+/CS- difference in SCR, regardless of the US type (Andreatta & Pauli, 2015, 2019; Klucken et al., 2019; Kruse et al., 2017, 2020; Tapia León et al., 2018; Wardle et al., 2018). In contrast, research utilizing model-based approaches (PsPM GLM and Ledalab), with primary and secondary reinforcers, presents mixed evidence. Among these, only one study out of five (including the present work) has found a difference in SCR between CS+ and CS- (Ebrahimi et al., 2017, 2019; Pietrock et al., 2019; van den Akker et al., 2017). Collectively, these findings could suggest that model-based approaches may be less effective in detecting CS+/CS-

differences in SCR, for example because they did not assume the correct underlying model of SCR generation. However, the heterogeneity of peak-scoring schemes in previous work and the limited number of studies warrant caution. Another interpretation is that different primary reinforcers (fruit juice in the present work, snacks in some previous work) elicit different CR, potentially due to differences in arousal elicited by the US.

4.1.3. PSR discriminates CS+/CS- difference in recall but not during learning

Unexpectedly, we did not find replicable CS+/CS- difference based on PSR during learning (Bray et al., 2008; Pietrock et al., 2019; Pool et al., 2019; Prévost et al., 2013; Seymour et al., 2007), despite our reward-conditioning paradigm being closely modelled on Pietrock et al. (2019), who did report a difference. There are two main differences between our study and theirs. One is the trial-by-trial collection of US expectancy ratings in their study. This might strengthen contingency awareness, which in turn could affect PSR (Van Dessel et al., 2019). The second is the CS sensory modality. Pietrock et al. (2019) used compound CS (i.e., visual and auditory stimuli presented simultaneously). Finally, other previous reward-conditioning studies that revealed CS+/CS- difference for PSR employed a reinforcement rate larger than the 50% used here (O'Doherty et al., 2003; Pool et al., 2019; Prévost et al., 2013; Reinhard & Lachnit, 2002; Schad et al., 2020). On the other hand, PSR distinguished CS+/CS- early during recall, with the largest effect size in both experiments when averaging over 7 CS+ and 7 CS- trials. How this discrepancy between initial learning and recall can be reconciled is unclear at this point.

4.2. Future directions

The present work raises several important questions. First, it remains unclear how reward and fear conditioned HPR differ from each other, and what their adaptive value might be. Second, variability of paradigm characteristics may influence the reliability of measurement. Future work may systematically examine the roles of these characteristics (e.g., reinforcement rate, US expectancy, type of CS and US stimuli, SOA, etc.) in reward learning. Third, in both our experiments, the effect size of the HPR largely decreased from the learning to the memory retention

phase. On the other hand, PSR robustly differentiated CS+/CS- during recall but not early learning. Hence future research could investigate the temporal dynamics of PSR and HPR as markers of reward learning.

4.3. Conclusion

In conclusion, the current study explored and validated a human Pavlovian reward conditioning paradigm and HPR as a robust marker of reward learning and retention of reward memory, as well as PSR as a robust marker of reward memory only. These findings may be beneficial for studies that involve reward learning and memory processes. Our work may also facilitate the identification of individual's exhibiting atypical reward learning patterns, thereby enabling the development of targeted treatment strategies.

Acknowledgements

This work was supported by the Clinical Research Priority Program of the University of Zurich for the CRPP "Synapse & Trauma" and the Economic and Social Research Council (ES/W000776/1). Dominik R. Bach receives funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-2018 CoG-816564 ActionContraThreat). The Wellcome Centre for Human Neuroimaging is funded by core funding from the Wellcome (203147/Z/16/Z). The Hertz Chair for Artificial Intelligence and Neuroscience in the Transdisciplinary Research Area Life and Health, Bonn University, is funded as part of the Excellence Strategy of the German federal and state governments.

ORCID

Yanfang Xia, https://orcid.org/0000-0002-9885-1748

Huaiyu Liu, https://orcid.org/0009-0006-5314-7008

Dominik R. Bach, https://orcid.org/0000-0003-3717-2036

Reference

Andreatta, M., & Pauli, P. (2015). Appetitive vs. Aversive conditioning in humans. *Frontiers in Behavioral Neuroscience*, *9*. https://www.frontiersin.org/articles/10.3389/fnbeh.2015.00128

Andreatta, M., & Pauli, P. (2019). Generalization of appetitive conditioned responses. *Psychophysiology*, *56*(9), e13397. https://doi.org/10.1111/psyp.13397

Bach, D. R. (2014). A head-to-head comparison of SCRalyze and Ledalab, two model-based methods for skin conductance analysis. *Biological Psychology*, *103*, 63–68. https://doi.org/10.1016/j.biopsycho.2014.08.006

Bach, D. R. (2023). Psychometrics in experimental psychology: A case for calibration. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02421-z

Bach, D. R., Castegnetti, G., Korn, C. W., Gerster, S., Melinscak, F., & Moser, T. (2018). Psychophysiological modeling: Current state and future directions. *Psychophysiology*, *55*(11), e13214. https://doi.org/10.1111/psyp.13209

Bach, D. R., Daunizeau, J., Friston, K. J., & Dolan, R. J. (2010). Dynamic causal modelling of anticipatory skin conductance responses. *Biological Psychology*, *85*(1), 163–170. https://doi.org/10.1016/j.biopsycho.2010.06.007

Bach, D. R., Flandin, G., Friston, K. J., & Dolan, R. J. (2010). Modelling event-related skin conductance responses. *International Journal of Psychophysiology*, *75*(3), 349–356. https://doi.org/10.1016/j.ijpsycho.2010.01.005

Bach, D. R., Gerster, S., Tzovara, A., & Castegnetti, G. (2016). A linear model for event-related respiration responses. *Journal of Neuroscience Methods*, *270*, 147–155. https://doi.org/10.1016/j.jneumeth.2016.06.001

Bach, D. R., & Melinscak, F. (2020). Psychophysiological modelling and the measurement of fear conditioning. *Behaviour Research and Therapy*, *127*, 103576. https://doi.org/10.1016/j.brat.2020.103576

Bach, D. R., Melinščak, F., Fleming, S. M., & Voelkle, M. C. (2020). Calibrating the experimental measurement of psychological attributes. *Nature Human Behaviour*, *4*(12), Article 12. https://doi.org/10.1038/s41562-020-00976-8

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1995). The metrics of cardiac chronotropism: Biometric perspectives. *Psychophysiology*, *32*(2), 162–171. https://doi.org/10.1111/j.1469-8986.1995.tb03308.x

Bray, S., Rangel, A., Shimojo, S., Balleine, B., & O'Doherty, J. P. (2008). The Neural Mechanisms Underlying the Influence of Pavlovian Cues on Human Decision Making. *Journal of Neuroscience*, *28*(22), 5861–5866. https://doi.org/10.1523/JNEUROSCI.0897-08.2008

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). *Handbook of Psychophysiology*. Cambridge University Press.

Castegnetti, G., Tzovara, A., Staib, M., Gerster, S., & Bach, D. R. (2017). Assessing fear learning via conditioned respiratory amplitude responses. *Psychophysiology*, *54*(2), 215–223. https://doi.org/10.1111/psyp.12778

Castegnetti, G., Tzovara, A., Staib, M., Paulus, P. C., Hofer, N., & Bach, D. R. (2016). Modeling fear-conditioned bradycardia in humans. *Psychophysiology*, *53*(6), 930–939. https://doi.org/10.1111/psyp.12637

Ebrahimi, C., Koch, S. P., Friedel, E., Crespo, I., Fydrich, T., Ströhle, A., Heinz, A., & Schlagenhauf, F. (2017). Combining D-cycloserine with appetitive extinction learning modulates amygdala activity during recall. *Neurobiology of Learning and Memory*, *142*, 209–217. https://doi.org/10.1016/j.nlm.2017.05.008

Ebrahimi, C., Koch, S. P., Pietrock, C., Fydrich, T., Heinz, A., & Schlagenhauf, F. (2019). Opposing roles for amygdala and vmPFC in the return of appetitive conditioned responses in humans. *Translational Psychiatry*, *9*(1), Article 1. https://doi.org/10.1038/s41398-019-0482-x

Exner, A., Tapia León, I., Mueller, E. M., & Klucken, T. (2021). Cardiac response in aversive and appetitive olfactory conditioning: Evidence for a valence-independent CS-elicited bradycardia. *Psychophysiology*, *58*(11), e13912. https://doi.org/10.1111/psyp.13912

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Finke, J. B., Roesmann, K., Stalder, T., & Klucken, T. (2021). Pupil dilation as an index of Pavlovian conditioning. A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, *130*, 351–368. https://doi.org/10.1016/j.neubiorev.2021.09.005

Friston, K. J., Mechelli, A., Turner, R., & Price, C. J. (2000). Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and Other Hemodynamics. *NeuroImage*, *12*(4), 466–477. https://doi.org/10.1006/nimg.2000.0630

Gerster, S., Namer, B., Elam, M., & Bach, D. R. (2018). Testing a linear time invariant model for skin conductance responses by intraneural recording and stimulation. *Psychophysiology*, *55*(2), e12986. https://doi.org/10.1111/psyp.12986

Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, *48*(2), 510–527. https://doi.org/10.3758/s13428-015-0588-x

Hermann, C., Ziegler, S., Birbaumer, N., & Flor, H. (2000). Pavlovian aversive and appetitive odor conditioning in humans: Subjective, peripheral, and electrocortical changes. *Experimental Brain Research*, *132*(2), 203–215. https://doi.org/10.1007/s002210000343

Ihaka, R., & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, *5*(3), 299–314. https://doi.org/10.1080/10618600.1996.10474713

Keiflin, R., & Janak, P. H. (2015). Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*, *88*(2), 247–263. https://doi.org/10.1016/j.neuron.2015.08.037

Klucken, T., Kruse, O., Klein, S., Kampa, M., Tapia León, I., & Stark, R. (2019). The relationship between neuroticism and appetitive conditioning. *Neurobiology of Learning and Memory*, *164*, 107068. https://doi.org/10.1016/j.nlm.2019.107068

Kokkola, N. H., Mondragón, E., & Alonso, E. (2019). A double error dynamic asymptote model of associative learning. *Psychological Review*, *126*(4), 506–549. https://doi.org/10.1037/rev0000147

Korn, C. W., Staib, M., Tzovara, A., Castegnetti, G., & Bach, D. R. (2017). A pupil size response model to assess fear learning. *Psychophysiology*, *54*(3), 330–343. https://doi.org/10.1111/psyp.12801

Kret, M. E., & Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior Research Methods*, *51*(3), 1336–1342. https://doi.org/10.3758/s13428-018-1075-y

Kruse, O., Klein, S., Tapia León, I., Stark, R., & Klucken, T. (2020). Amygdala and nucleus accumbens involvement in appetitive extinction. *Human Brain Mapping*, *41*(7), 1833–1841. https://doi.org/10.1002/hbm.24915

Kruse, O., Tapia León, I., Stark, R., & Klucken, T. (2017). Neural correlates of appetitive extinction in humans. *Social Cognitive and Affective Neuroscience*, *12*(1), 106–115. https://doi.org/10.1093/scan/nsw157

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2013.00863

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., Heitland, I., Hermann, A., Kuhn, M., Kruse, O., Meir Drexler, S., Meulders, A., Nees, F., Pittig, A., Richter, J., Römer, S., Shiban, Y., Schmitz, A., Straube, B., … Merz, C. J. (2017). Don’t fear ‘fear conditioning’: Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Martin-Soelch, C., Linthicum, J., & Ernst, M. (2007). Appetitive conditioning: Neural bases and implications for psychopathology. *Neuroscience & Biobehavioral Reviews*, *31*(3), 426–440. https://doi.org/10.1016/j.neubiorev.2006.11.002

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal Difference Models and Reward-Related Learning in the Human Brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/S0896-6273(03)00169-7

Ojala, K. E., Tzovara, A., Poser, B. A., Lutti, A., & Bach, D. R. (2022). Asymmetric representation of aversive prediction errors in Pavlovian threat conditioning. *NeuroImage*, *263*, 119579. https://doi.org/10.1016/j.neuroimage.2022.119579

Pan, J., & Tompkins, W. J. (1985). A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, *BME-32*(3), 230–236. IEEE Transactions on Biomedical Engineering. https://doi.org/10.1109/TBME.1985.325532

Paulus, P. C., Castegnetti, G., & Bach, D. R. (2016). Modeling event-related heart period responses. *Psychophysiology*, *53*(6), 837–846. https://doi.org/10.1111/psyp.12622

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (pp. xv, 430). Oxford Univ. Press.

Pietrock, C., Ebrahimi, C., Katthagen, T. M., Koch, S. P., Heinz, A., Rothkirch, M., & Schlagenhauf, F. (2019). Pupil dilation as an implicit measure of appetitive Pavlovian learning. *Psychophysiology*, *56*(12). https://doi.org/10.1111/psyp.13463

Pollatos, O., Herbert, B. M., Matthias, E., & Schandry, R. (2007). Heart rate response after emotional picture presentation is modulated by interoceptive awareness. *International Journal of Psychophysiology*, *63*(1), 117–124. https://doi.org/10.1016/j.ijpsycho.2006.09.003

Pool, E. R., Pauli, W. M., Kress, C. S., & O'Doherty, J. P. (2019). Behavioural evidence for parallel outcome-sensitive and outcome-insensitive Pavlovian learning systems in humans. *Nature Human Behaviour*, *3*(3), Article 3. https://doi.org/10.1038/s41562-018-0527-9

Prévost, C., McNamee, D., Jessup, R. K., Bossaerts, P., & O'Doherty, J. P. (2013). Evidence for Model-based Computations in the Human Amygdala during Pavlovian Conditioning. *PLOS Computational Biology*, *9*(2), e1002918. https://doi.org/10.1371/journal.pcbi.1002918

Reinhard, G., & Lachnit, H. (2002). Differential conditioning of anticipatory pupillary dilation responses in humans. *Biological Psychology*, *60*(1), 51–68. https://doi.org/10.1016/S0301-0511(02)00011-X

Roelofs, K., & Dayan, P. (2022). Freezing revisited: Coordinated autonomic and central optimization of threat coping. *Nature Reviews Neuroscience*, *23*(9), Article 9. https://doi.org/10.1038/s41583-022-00608-2

Sayão, A., Alves, H., Furukawa, E., Schultz Wenk, T., Cagy, M., Gutierrez-Arango, S., Tripp, G., & Caparelli-Dáquer, E. (2021). Development of a Classical Conditioning Task for Humans Examining Phasic Heart Rate Responses to Signaled Appetitive Stimuli: A Pilot Study. *Frontiers in Behavioral Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnbeh.2021.639372

Schad, D. J., Rapp, M. A., Garbusow, M., Nebe, S., Sebold, M., Obst, E., Sommer, C., Deserno, L., Rabovsky, M., Friedel, E., Romanczuk-Seiferth, N., Wittchen, H.-U., Zimmermann, U. S., Walter, H., Sterzer, P., Smolka, M. N., Schlagenhauf, F., Heinz, A., Dayan, P., & Huys, Q. J. M. (2020). Dissociating neural learning signals in human sign- and goal-trackers. *Nature Human Behaviour*, *4*(2), Article 2. https://doi.org/10.1038/s41562-019-0765-5

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential Encoding of Losses and Gains in the Human Striatum. *Journal of Neuroscience*, *27*(18), 4826–4831. https://doi.org/10.1523/JNEUROSCI.0400-07.2007

Signoret-Genest, J., Schukraft, N., L. Reis, S., Segebarth, D., Deisseroth, K., & Tovote, P. (2023). Integrated cardio-behavioral responses to threat define defensive states. *Nature Neuroscience*, *26*(3), 447–457. https://doi.org/10.1038/s41593-022-01252-w

Spector, R. H. (1990). The Pupils. In H. K. Walker, W. D. Hall, & J. W. Hurst (Eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations* (3rd ed.). Butterworths. http://www.ncbi.nlm.nih.gov/books/NBK381/

Stussi, Y., Delplanque, S., Coraj, S., Pourtois, G., & Sander, D. (2018). Measuring Pavlovian appetitive conditioning in humans with the postauricular reflex. *Psychophysiology*, *55*(8), e13073. https://doi.org/10.1111/psyp.13073

Tapia León, I., Kruse, O., Stalder, T., Stark, R., & Klucken, T. (2018). Neural correlates of subjective CS/UCS association in appetitive conditioning. *Human Brain Mapping*, *39*(4), 1637–1646. https://doi.org/10.1002/hbm.23940

Tzovara, A., Korn, C. W., & Bach, D. R. (2018). Human Pavlovian fear conditioning conforms to probabilistic learning. *PLOS Computational Biology*, *14*(8), e1006243. https://doi.org/10.1371/journal.pcbi.1006243

van den Akker, K., Nederkoorn, C., & Jansen, A. (2017). Electrodermal responses during appetitive conditioning are sensitive to contingency instruction ambiguity. *International Journal of Psychophysiology*, *118*, 40–47. https://doi.org/10.1016/j.ijpsycho.2017.06.004

Van Dessel, P., Hughes, S., & De Houwer, J. (2019). How Do Actions Influence Attitudes? An Inferential Account of the Impact of Action Performance on Stimulus Evaluation. *Personality and Social Psychology Review*, *23*(3), 267–284. https://doi.org/10.1177/1088868318795730

Wardle, M. C., Lopez-Gamundi, P., & Flagel, S. B. (2018). Measuring appetitive conditioned responses in humans. *Physiology & Behavior*, *188*, 140–150. https://doi.org/10.1016/j.physbeh.2018.02.004

Wehrli, J. M., Xia, Y., Meister, L., Tursunova, S., Kleim, B., Bach, D. R., & Quednow, B. B. (2024). Forget me not: The effect of doxycycline on human declarative memory. *European Neuropsychopharmacology*, *89*, 1–9. https://doi.org/10.1016/j.euroneuro.2024.08.006

Wehrli, J. M., Xia, Y., Offenhammer, B., Kleim, B., Müller, D., & Bach, D. R. (2023). Effect of the Matrix Metalloproteinase Inhibitor Doxycycline on Human Trace Fear Memory. *eNeuro*, *10*(2). https://doi.org/10.1523/ENEURO.0243-22.2023

Xia, Y., Melinscak, F., & Bach, D. R. (2021). Saccadic scanpath length: An index for human threat conditioning. *Behavior Research Methods*, *53*(4), 1426–1439. https://doi.org/10.3758/s13428-020-01490-5

Xia, Y., Wehrli, J., Abivardi, A., Hostiuc, M., Kleim, B., & Bach, D. R. (2024). Attenuating human fear memory retention with minocycline: A randomized placebo-controlled trial. *Translational Psychiatry*, *14*(1), Article 1. https://doi.org/10.1038/s41398-024-02732-2

Xia, Y., Wehrli, J., Gerster, S., Kroes, M., Houtekamer, M., & Bach, D. (2022). *Measuring human context fear conditioning and retention after consolidation*. PsyArXiv. https://doi.org/10.31234/osf.io/ypu4j