ARTICLES
https://doi.org/10.1038/s41562-018-0527-9
nature human behaviour

# Behavioural evidence for parallel outcome-sensitive and outcome-insensitive Pavlovian learning systems in humans

Eva R. Pool[1]*, Wolfgang M. Pauli[1,2], Carolina S. Kress[1] and John P. O'Doherty[1,2]

**There is a dichotomy in instrumental conditioning between goal-directed actions and habits that are distinguishable on the basis of their relative sensitivity to changes in outcome value. It is less clear whether a similar distinction applies in Pavlovian conditioning, where responses have been found to be predominantly outcome-sensitive. To test for both devaluation-insensitive and devaluation-sensitive Pavlovian conditioning in humans, we conducted four experiments combining Pavlovian conditioning and outcome-devaluation procedures while measuring multiple conditioned responses. Our results suggest that Pavlovian conditioning involves two distinct types of learning: one that learns the current value of the outcome, which is sensitive to devaluation, and one that learns about the spatial localization of the outcome, which is insensitive to devaluation. Our findings have implications for the mechanistic understanding of Pavlovian conditioning and provide a more nuanced understanding of Pavlovian mechanisms that might contribute to a number of psychiatric disorders.**

A common symptom across many clinical disorders, such as drug addiction or binge eating, is the willingness to go to extraordinary lengths to obtain an object of desire, despite the object not being experienced as pleasurable once obtained[1,2]. Identifying the underlying mechanisms that lead to such paradoxical behaviour has been a major research focus. Of particular interest has been the role of stimulus–response habits, a form of instrumental responding that can persist even after the outcome of an action is no longer valued (for example, seeking snacks even when completely satiated)[2,3]. However, instrumental habits are only one of several systems known to exert influence on behaviour. Alongside instrumental conditioning, there exists an elaborate system for Pavlovian conditioning[4–11], whereby reflexive conditioned behaviours can come to be elicited by a conditioned stimulus (CS; for example, a metronome sound) that predicts the subsequent delivery of an affectively significant outcome (for example, food)[9,11–13].

The aim of the current study is to investigate whether there exist Pavlovian conditioned responses in humans that persist even after an associated outcome no longer has substantive affective significance to the organism. Such a form of Pavlovian conditioning could provide evidence for an important additional mechanism alongside instrumental habits, by which maladaptive inflexible behaviour can be generated. The bulk of behavioural evidence across animals and humans emphasizes the outcome-sensitive nature of Pavlovian conditioning, such that changes in the affective value of an associated outcome lead to an immediate and substantive change in the elicited conditioned response[7,14–19].

The apparent ubiquity of devaluation-sensitive behaviour in Pavlovian conditioning creates a paradox for popular theoretical models of Pavlovian conditioning. These models tend to describe Pavlovian conditioning as essentially a form of model-free reinforcement learning, analogous to that proposed to account for instrumental habits[20–22]. For instance, in model-free reinforcement-learning approaches to Pavlovian conditioning, such as the temporal difference algorithm or the Rescorla–Wagner rule, Pavlovian CSs become endowed with a 'cached' value by means of a reward prediction error that cannot be flexibly updated following changes in the value of the outcome responsible for stamping in the learned value[23].

However, Pavlovian conditioning does not appear to be a unitary process, but rather seems to involve several parallel associations between multiple aspects of the outcome[12,24]. Some associations are formed with the affective/motivational aspects of the outcome. These affective representations are independent of the specific perceptual properties of the outcome and are considered to be tracking the current value of the outcome[19]. At the same time, other associations are formed with the perceptual or sensory attributes of the outcome. These representations can be very specific to a particular sensorial property of the outcome[12,25]. Despite the long-standing conceptualization of multiple conditioned responses to a given Pavlovian stimulus, it is not clear whether these responses are always identical or can diverge by having, for instance, differential sensitivity to outcome devaluation. Evidence in favour of outcome-insensitive Pavlovian behaviour in animals is sparse, although it has been reported that some Pavlovian responses are more sensitive to outcome value changes[19] than others[26] and such differences have often been attributed to interindividual differences[26–28]. Here, we formulated the hypothesis that the class of Pavlovian response based on a representation of the current value of the outcome would by definition flexibly adapt to outcome devaluation; whereas the class of Pavlovian response based on some specific sensory aspect of an outcome would be resistant to outcome value changes.

A recent study by Zhang et al. successfully distinguished two different classes of Pavlovian response in humans[29] during pain conditioning: a class of responses reflecting the value of the outcome and a class of responses reflecting a specific sensory feature of the outcome (its spatial location). However, these authors did not address whether one or both of those Pavlovian responses are devaluation sensitive.

Here we employed an outcome-devaluation framework to test for the sensitivity of different classes of appetitive Pavlovian responses to outcome value. We used eye-tracking techniques combined with

[1]Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA, USA. [2]Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA, USA. *e-mail: Eva.pool@unige.ch

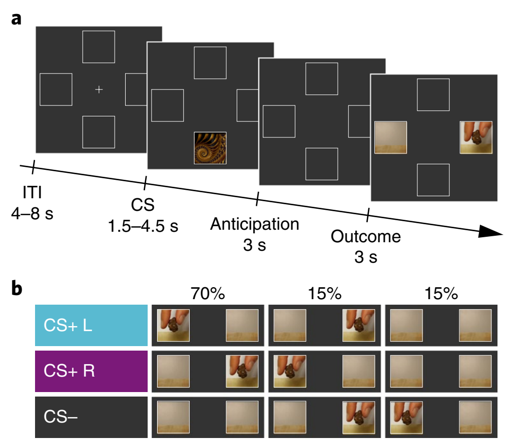


**Fig. 1 | Schematic representation of the experimental design. a**, Illustration of the sequence of events in a trial for Experiment 1, Experiment 2 and Experiment 4. At the beginning of each trial a CS was presented randomly in the upper or lower portion of the screen for 1.5–4.5 s (uniformly distributed). After an anticipation screen of 3 s, a video showing the snack delivery appeared either to the right or the left side of the screen for 3 s. The intertrial interval (ITI) lasted for 4–8 s (uniformly distributed). At the end of each session, participants received the actual snacks delivered during the task and allowed to eat them. **b**, Experiment 1 involved one of three CSs, each of which primarily predicted (70%) either snack delivery to the left (CS+ L), snack delivery to the right (CS+ R) or no snack delivery (CS−).

an appetitive Pavlovian conditioning task in which neutral images were associated with a video of food outcome delivery. Inspired by the laterality of stimulus presentation employed by Zhang et al.[29], we adapted this experimental feature to our appetitive conditioning framework: one image was more often associated with food outcome delivery on the left side of the screen (positive CS left (CS+ L)); one image was more often associated with outcome delivery on the right side of the screen (positive CS right (CS+ R)); and another image was more often associated with no outcome delivery (negative CS (CS−); see Fig. 1). We recorded eye gaze and pupil responses during the experiment. Pupil dilation during the CS onset was taken as a response reflecting the value representation of the outcome, as several studies have shown that pupil dilation is strongly influenced by value[8,30,31]. Anticipatory gaze direction (left versus right) was taken as a response reflecting a specific sensorial representation of the outcome (its spatial location). We therefore predicted that, in comparison to pupil dilation responses, gaze direction would be less sensitive to outcome devaluation. First, in Experiment 1 and Experiment 2, we tested the existence of two classes of Pavlovian response and their sensitivity to outcome devaluation. In Experiment 3 and Experiment 4, we addressed a potential confound in our interpretation of the first experiments. In Experiment 3 we tested whether the anticipatory gaze direction reflects the performance of an instrumental action instead of a Pavlovian conditioned response. In Experiment 4, we tested whether anticipatory gaze direction reflects a non-specific deployment of spatial attention towards a perceptually salient event, as opposed to a Pavlovian conditioned response established by learning about reward outcomes.

## Results

**Experiment 1.** We first tested whether the pupil dilation response and the anticipatory gaze direction reflected patterns of distinct classes of Pavlovian response as in Zhang et al.[29], the original study. We expected the pupil dilation to follow a value pattern (CS+ L and CS+ R different from CS−) and the gaze direction to follow a lateralized pattern (larger dwell time for CS+ L compared to CS+ R and CS− on the left side of the screen; larger dwell time for CS+ R compared to CS+ L and CS− on the right side of the screen).

Second, we tested the sensitivity of these two classes of Pavlovian response to outcome devaluation. After initial Pavlovian conditioning, the food outcome was devalued by feeding individuals on that outcome to satiety in half of the participants, while the remaining participants served as non-devalued controls. Subsequently, the CSs were presented under extinction (no outcome was delivered) and Pavlovian responses measured. We expected pupil dilation, but not gaze direction, to flexibly adapt to the decreased outcome value.

*Pavlovian learning.* Anticipatory gaze direction. As predicted, a first planned contrast analysis using *F*-tests conducted on the CS condition (CS+ L, CS+ R, CS–) with the following weights (+1, −0.5, −0.5) revealed an increased dwell time in the left region of interest (ROI) after the perception of CS+ L compared to CS+ R and CS– ($F(1,39) = 21.33$; $P < 0.001$; partial eta squared ($\eta^2_p$) = 0.354; 90% confidence interval (CI) (0.155, 0.503); see Fig. 2c). A second planned contrast analysis using *F*-tests conducted on the CS condition (CS+ L, CS+ R, CS–) with the following weights (−0.5, +1, −0.5) revealed an increased dwell time in the ROI after the perception of CS+ R compared to CS+ L and CS– ($F(1,39) = 27.10$; $P < 0.001$; $\eta^2_p = 0.41$; 90% CI (0.207, 0.550); see Fig. 2b).

Pupil dilation. As predicted, a planned contrast analysis using *F*-tests conducted on the CS condition (CS+ L, CS+ R, CS–) with the following weights (+0.5, +0.5, −1) revealed that the pupil was less constricted for CS+ L and CS+ R compared to CS– ($F(1,39) = 4.45$; $P = 0.041$; $\eta^2_p = 0.102$; 90% CI (0.002, 0.259); see Fig. 2a).

*Outcome devaluation.* Paired *t*-tests showed that hunger ($t(19) = 6.93$; $P < 0.001$; Cohen's *d* ($d$) = 1.367; 95% CI (0.779, 1.938)) and pleasantness of the favourite food outcome ($t(19) = 6.10$; $P < 0.001$; $d = 1.853$; 95% CI (1.005, 2.674)) significantly decreased after selective satiation compared to before (see Fig. 3a).

*Outcome-devaluation-induced changes.* Anticipatory gaze direction. We computed the average dwell time spent in the congruent ROI for both CSs+ (dwell time in the right ROI after CS+ R and dwell time on the left ROI after CS+ L) for the last session before satiation and during the first half of the extinction test for both the satiation and the control group. We used only the first half of the extinction test session to avoid confounding effects due to extinction processes. A 2 (session: pre- or post-satiation) × 2 (group: satiation or control) mixed repeated measures analysis of variance (ANOVA) applied to dwell time spent in the congruent ROIs revealed a significant main effect of session ($F(1,38) = 15.02$; $P < 0.001$; $\eta^2_p = 0.283$; 90% CI (0.095, 444)), but no significant interaction ($F(1,38) = 0.51$; $P = 0.478$; $\eta^2_p = 0.013$; 90% CI (0.000, 0.121)), suggesting that dwell time was rapidly modulated by extinction, but there was no statistically significant evidence that the dwell time was sensitive to outcome devaluation (see Fig. 4c).

Pupil dilation. We computed a similar index to that for dwell time by averaging pupil dilation during CS+ L and CS+ R for the last session before satiation and during the first half of the extinction session. A 2 (session: pre- or post-satiation) × 2 (group: satiation or control) mixed repeated measures ANOVA applied to pupil dilation revealed a significant session × outcome interaction ($F(1,38) = 4.93$; $P = 0.032$; $\eta^2_p = 0.115$; 90% CI (0.005, 0.276)), showing that the decrease in pupil dilation induced by satiation was significantly larger in the satiation group compared to the control group (see Fig. 4a). This suggests that pupil dilation flexibly adapted to outcome devaluation.

*Experiment 1 discussion.* Results suggest two distinct classes of Pavlovian response: one reflecting outcome value, as measured by pupil dilation, and another reflecting the spatial localization

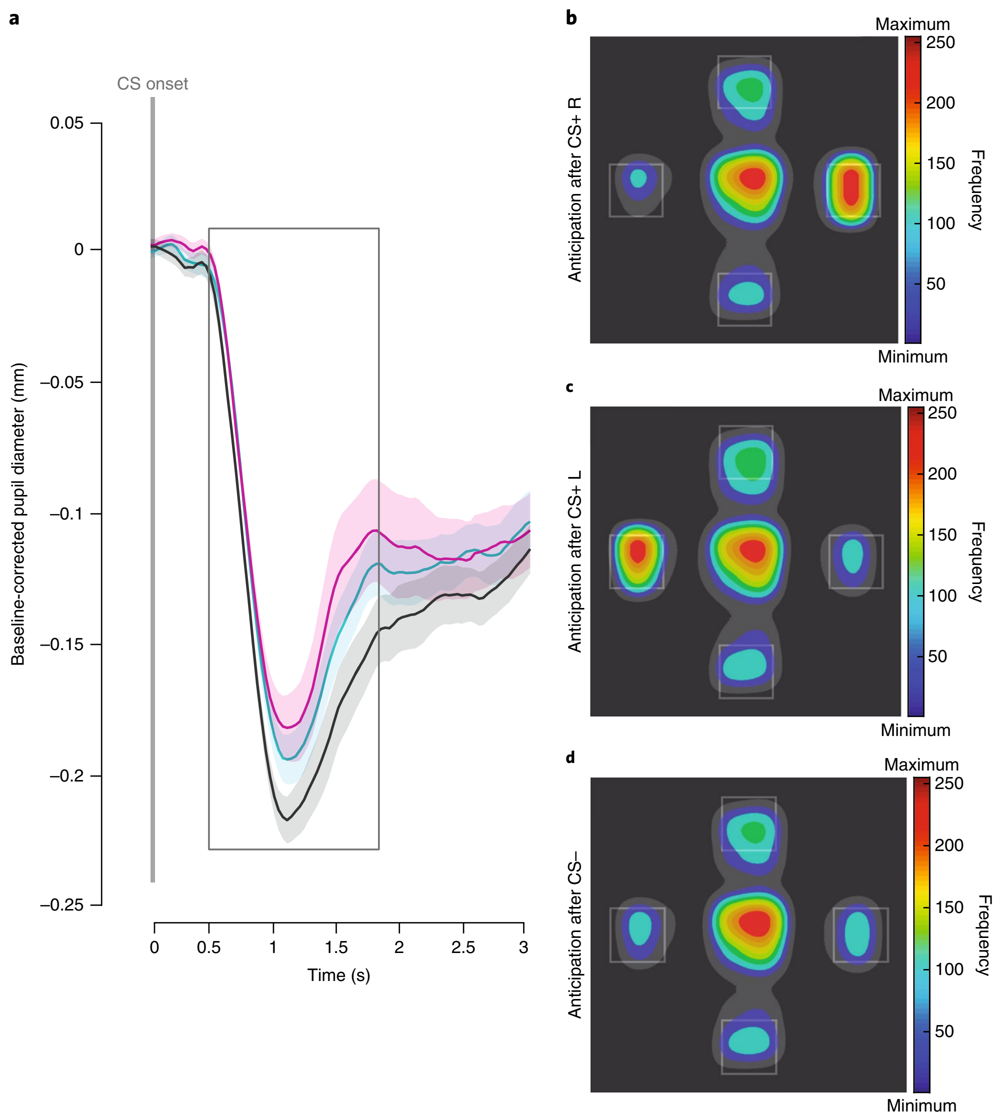


**Fig. 2 | Effect of conditioning during the learning phase of Experiment 1. a**, Plot of the averaged pupil response over time aligned to the onset of the CS and plotted separately for the CS predicting either the delivery of a snack to the left (CS+ L), the delivery of a snack to the right (CS+ R) or no snack delivery (CS–). **b**–**d**, Heatmaps of the fixation patterns during the anticipation screen (normalized frequency): after the offset of CS+ R (**b**), CS+ L (**c**) and CS– (**d**). Shaded areas indicate the within-subject standard error of the mean. All plots are based on data from 40 participants.

of the outcome, as measured by gaze direction. The pupil dilation responses flexibly adapted to changes in the outcome value, whereas responses based on spatial localization were seemingly not affected by outcome devaluation. Our findings suggest that perception of the same Pavlovian stimulus can trigger parallel responses in the same individual; some are more adapted than others to the current value of the associated outcome. However, in this experiment, because the CSs were associated with only one food outcome that was subsequently devalued, our results could reflect general motivational changes (a general decrease in the hunger level) rather than a specific change in outcome value (the specific pleasantness of the food outcome). A way to tackle this issue is to use multiple CSs associated with two different food outcomes (for example, a sweet food and a savoury food) and to devalue only one of the food outcomes[17,32].

**Experiment 2.** Experiment 2 aimed to extend and replicate the findings from Experiment 1 by using a more selective procedure for outcome devaluation. We introduced two different CSs+ L and two different CSs+ R: each one of the two CSs+ was associated with a specific food outcome that was either sweet or savoury. There were four different CSs+: a CS+ L and CS+ R associated with the sweet outcome and a CS+ L and CS+ R associated with the savoury outcome. After learning, only one of the two food outcomes was experimentally devalued by feeding that particular outcome to satiety. Thus, we were able to test the effect of a specific value change on the two classes of Pavlovian response that we identified in Experiment 1. We expected that conditioned pupil dilation would show sensitivity to changes in outcome value, but that conditioned responses based on spatial location (gaze direction) would be devaluation insensitive.

*Pavlovian learning.* Anticipatory gaze direction. We replicated the same findings as in Experiment 1, as shown by a planned contrast analysis using $F$-tests conducted on the CS condition (CS+ L, CS+ R, CS–) with the following weights (+1, −0.5, −0.5), which revealed an increased dwell time in the left ROI after the perception of the CS+ L compared to the CS+ R and the CS– ($F(1,19) = 13.15$; $P = 0.002$;

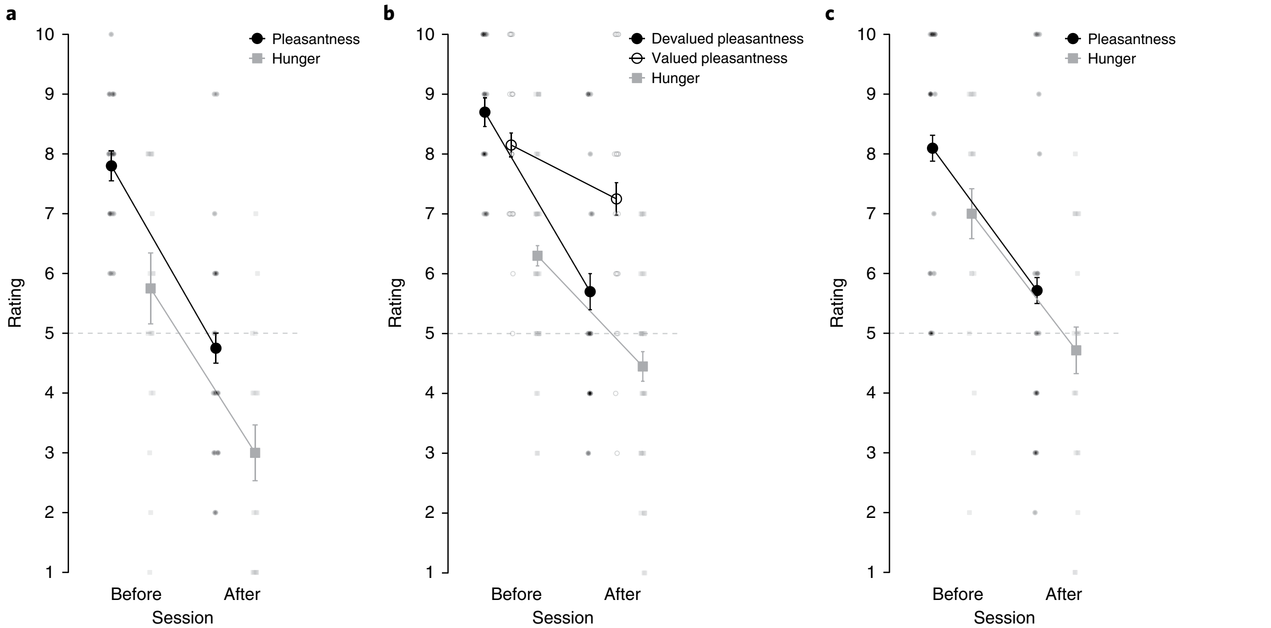


**Fig. 3 | Manipulation check of the outcome-devaluation procedure. a**,**c**, Mean ratings of hunger and the pleasantness of the snack before and after selective satiation for the group that underwent the outcome-devaluation procedure in Experiment 1 (**a**) and Experiment 3 (**c**). **b**, Mean ratings of hunger and of the pleasantness of the snack that was devalued through the selective satiation procedure (devalued pleasantness) and the snack that was not (valued pleasantness) in Experiment 2. Error bars indicate the within-subject s.e.m. Plots of Experiments 1 and 2 are based on two different sets of 20 participants. Plot from Experiment 3 is based on data from 21 participants.

$\eta^2_p = 0.409$; 90% CI (0.119, 0.590)). Likewise, a second planned contrast analysis using $F$-tests conducted on the CS condition (CS+ L, CS+ R, CS−) with the following weights (−0.5, +1, −0.5) revealed an increased dwell time in the ROI after the perception of the CS+ R compared to the CS+ L and to the CS− ($F(1,19) = 15.81$; $P = 0.001$; $\eta^2_p = 0.454$; 90% CI (0.157, 0.623).

Pupil dilation. We obtained a trend similar to the effect found in Experiment 1, as indicated by a planned contrast analysis using $F$-tests conducted on the CS condition (CS+ L, CS+ R, CS−) with the following weights (+0.5, +0.5, −1), which revealed that the pupil was less constricted for CS+ L and CS+ R compared to CS− ($F(1,19) = 3.27$; $P = 0.086$; $\eta^2_p = 0.147$; 90% CI (0.000, 0.368)), although this effect did not reach statistical significance.

*Outcome devaluation*. Paired $t$-tests showed that hunger ($t(19) = 5.52$; $P < 0.001$; $d = 1.07$; 95% CI (0.555, 1.573)), the pleasantness of the food outcome that had been eaten until satiety (devalued outcome ($t(19) = 7.19$; $P < 0.001$; $d = 1.760$; 95% CI (1.016, 2.483)) and the pleasantness of the food outcome that had not been eaten until satiety (valued outcome ($t(19) = 2.780$; $P = 0.012$; $d = 0.489$; 95% CI (0.106, 0.862)) significantly decreased after the selective satiation compared to before. Critically, a 2 (session: pre- or post-satiation) × 2 (outcome: valued or devalued) repeated measures ANOVA applied to the food pleasantness ratings revealed a significant interaction ($F(1,19) = 28.02$; $P < 0.001$; $\eta^2_p = 0.596$; 90% CI (0.331, 0.723)), showing that the decrease in pleasantness was significantly larger for the devalued food outcome compared to the valued food outcome (see Fig. 3b).

*Outcome-devaluation-induced changes*. Anticipatory gaze direction. We computed the average dwell time allocated to the congruent ROI for all the CSs+ (dwell time in the right ROI after CS+ R and dwell time in the left ROI after CS+ L) for the CSs associated with the devalued outcome (CS devalued) and the CSs associated with the valued outcome (CS valued) at both times: the last session before satiation and the test session. Unlike Experiment 1, we could use the whole test session, because we used a manipulation to attenuate effects of extinction on responding (see Methods section). Using a 2 (session: pre- or post-satiation) × 2 (CS: valued or devalued) repeated measures ANOVA applied to dwell time spent in the congruent ROI, as in Experiment 1, we did not find a significant interaction ($F(1,19) = 0.04$; $P = 0.843$; $\eta^2_p = 0.002$; 90% CI (0.000, 0.100)), suggesting that there was no statistically significant evidence that dwell time was sensitive to outcome devaluation (see Fig. 4d).

Pupil dilation. We computed a similar index to the one for dwell time, by averaging pupil dilation during the CSs associated with the valued outcome and the CSs associated with the devalued outcome at two time points: the last session before satiation and the test session. A 2 (session: pre- or post-satiation) × 2 (CS: valued or devalued) repeated measures ANOVA applied to pupil dilation revealed a significant interaction ($F(1,19) = 8.08$; $P = 0.010$; $\eta^2_p = 0.298$; 90% CI (0.045, 0.504)), showing that the decrease in pupil dilation induced by satiation was significantly larger for the devalued CS compared to the valued CS (see Fig. 4b). This suggests that, as in Experiment 1, pupil dilation flexibly adapted to outcome devaluation.

*Experiment 2 discussion*. Experiment 2 replicated the main finding of Experiment 1: a CS can elicit multiple classes of Pavlovian response (as measured by pupil dilation and gaze direction) that are differentially sensitive to changes in the outcome value. Critically, Experiment 2 showed that changes in pupil dilation as a Pavlovian response reflect the value representation of the outcome, rather than being a consequence of an overall change in motivation or the physiological effects of generalized satiation.

However, there is a possible alternative explanation for the results of Experiment 1 and 2. While pupil dilation is evidently noninstrumental (for example, participants' pupil dilation could not influence the outcome delivery), it could be argued that gaze direction is an instrumental action as opposed to being a Pavlovian conditioned response. To counter this possibility, the task was

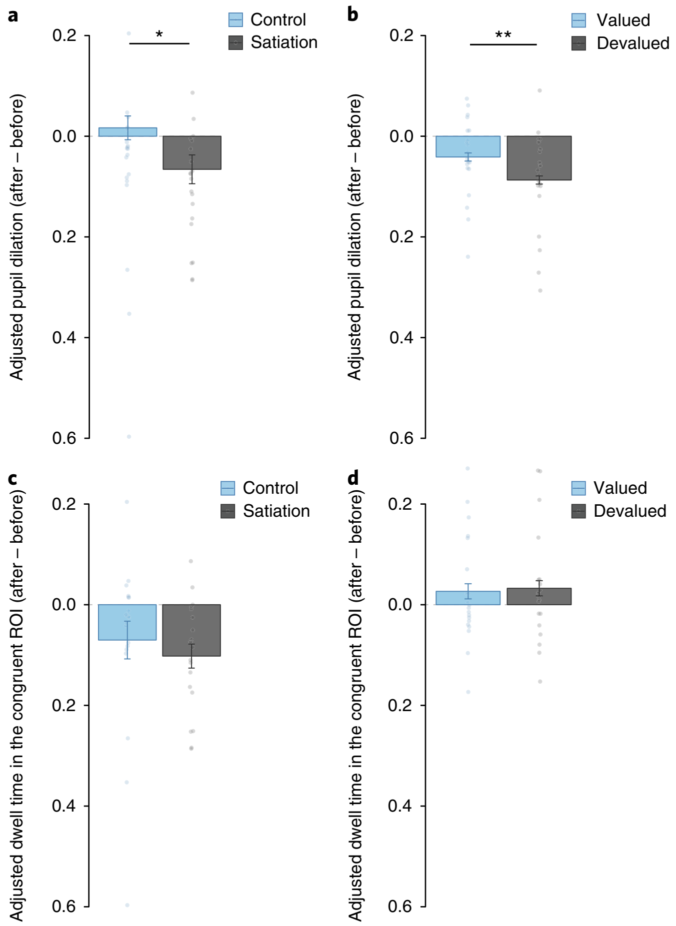


**Fig. 4 | Effects of the outcome-devaluation procedure on different conditioned responses during Experiment 1 and Experiment 2.** **a**,**b**, Adjusted pupil dilation before and after the outcome-devaluation procedure in Experiment 1 (**a**) and Experiment 2 (**b**). $F(1,38) = 4.93$; $^*P = 0.032$; $\eta^2_p = 0.115$; 90% CI (0.005, 0.276). $F(1,19) = 8.08$; $^{**}P = 0.010$; $\eta^2_p = 0.298$; 90% CI (0.045, 0.504). Bars represent a change score in the pupil dilation induced by the outcome-devaluation procedure (change score = adjusted pupil dilation after satiation – adjusted pupil dilation before satiation). **c**,**d**, Adjusted dwell time spent in the ROI congruent with the CS+ prediction (left ROI for the CS+ predicting outcome delivery to the left and right ROI for the CS+ predicting outcome delivery to the right) before and after the outcome-devaluation procedure in Experiment 1 (**c**) and Experiment 2 (**d**). Bars represent the change score in dwell time induced by outcome devaluation (change score = adjusted dwell time after satiation – adjusted dwell time before satiation). Results depicted from Experiment 1 are shown separately for the CS+ predictions in the satiation group that underwent outcome devaluation and the control group that did not, while results from Experiment 2 show effects from the CS+ associated with the devalued outcome and the CS+ associated with the outcome that was still valued. Plots of Experiment 1 are based on 40 participants and error bars indicate the between-subject s.e.m.; plots from Experiment 2 are based on 20 participants and error bars indicate the within-subject s.e.m.

programmed so that none of the participants' actions could influence the outcome delivery, which depended solely on the CS. However, participants might still have presumed that gazing towards the most likely location of the outcome would influence the delivery of the outcome. Moreover, the instrumental system might have been automatically invoked to learn a pseudo-contingency irrespective of a participant's subjective impressions. Under this interpretation, gaze direction effects would not reflect a Pavlovian conditioned response at all, but rather an instrumental response.

**Experiment 3.** In Experiment 3, we further investigated gaze direction behaviour to establish whether it genuinely reflects a Pavlovian conditioned response, or is instead an instrumentally controlled action. To address this question, we relied on a key behavioural distinction between instrumental and Pavlovian conditioning. An instrumental action is by definition fully flexible with regard to its directionality: one should be able to train the action to go in one direction for reward as easily as one can train the action to go in the opposite direction for the same reward[33,34]. On the other hand, if this behaviour is a Pavlovian response, the response itself is by definition inflexible as it is essentially a reflex. Thus, it will strongly resist being shaped to go in the opposite direction to that dictated by the reflex. A famous example is Hershberger's[35] 'room through a looking glass' experiment where food-deprived chicks were unable to learn that walking in the opposite direction to a food source would lead to actually gaining access to it, because approaching food (as opposed to moving away from it) is a strong Pavlovian conditioned response not amenable to reversible instrumental control.

In our specific gaze direction example, if gaze direction is solely under instrumental control it ought to be equally easy to train participants to gaze in the opposite direction to where the food pictures will be delivered as it is to train participants to gaze in the same direction. However, if gaze direction towards the outcome location is also under Pavlovian control, then gazing in the opposite direction should be more difficult than gazing in the same direction, reflecting a conflict between the Pavlovian and the instrumental system.

To address this we adapted the experimental task used previously. Stimuli were associated with a food outcome delivery on either the left or right side of the screen, but this time the outcome delivery was directly contingent on gaze behaviour. To successfully collect the food, participants had to look at a particular location depending on the cue they had just perceived. For some of the stimuli (congruent cues; see Fig. 5), participants had to look in the same direction as where the outcome delivery video was going to appear so that, for example, if the cue predicted the food picture would appear on the left, participants had to gaze to the left to obtain the food. However, for other stimuli (incongruent cues; see Fig. 5), participants had to look in the opposite direction to where the outcome delivery video was going to appear so that, for example, if the cue predicted outcome delivery on the left, participants had to gaze to the right to obtain that food outcome. Therefore, we fully orthogonalized the instrumental and the hypothesized Pavlovian influences on gaze behaviour in a 2 (action: look left, look right) by 2 (outcome: delivery left, delivery right) design. This design was similar to that used in a previous study[36] showing that conflicting Pavlovian expectations have a detrimental effect on human instrumental performance. We expected to observe a conflict effect that is reflected in a decreased dwell time on the opposing location that participants needed to look at to collect the food outcome during incongruent trials compared to congruent trials.

We also tested the sensitivity of this Pavlovian conflict effect to outcome devaluation. After two learning sessions, the food outcome was devalued by feeding participants on that outcome to satiety in half of the participants while the other half served as controls. Subsequently, cue stimuli were presented under extinction and gaze behaviour was measured. We expected the outcome devaluation to influence the instrumentally learned action more than the presumed Pavlovian conditioned response, because the instrumental action had undergone only moderate amounts of training (participants were not overtrained) and after modest training instrumental actions are generally found to be outcome-value sensitive[32,37–39].

*Pavlovian instrumental conflict.* Anticipatory gaze direction in the Pavlovian ROI. We defined the Pavlovian ROI as the location where the food outcome delivery video was most likely to be displayed given the specific contingencies for a particular CS. We expected

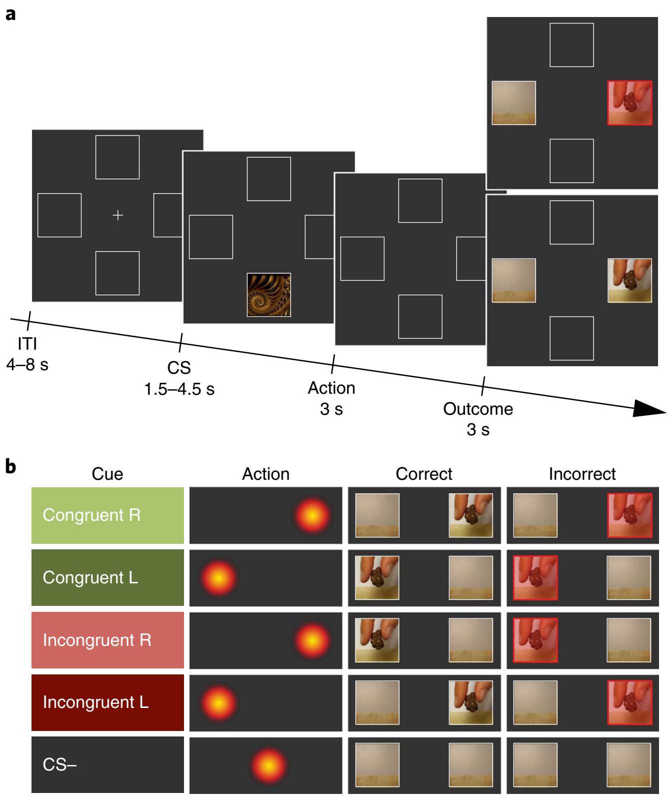


**Fig. 5 | Illustration of the sequence of events in a trial for Experiment 3.** **a**, At the beginning of the trial a cue was presented randomly in the upper or lower portion of the screen for 1.5–4.5 s (uniformly distributed). On the basis of this cue, participants were asked to look either to the left or right side of the screen to win a piece of their favourite snack. Then a video of the snack delivery was displayed on either the right or left side of the screen. If the participants looked at the correct side of the screen during the action screen, the video was displayed normally, whereas if the participants looked at the incorrect side of the action screen the video was displayed behind a transparent red square, indicating that no snack was successfully collected during that trial. The ITI lasted for 4–8 s (uniformly distributed). Participants were told that for each cue there was a correct gaze action to be performed to obtain the snack. At the end of each session they received the snacks that were successfully delivered during the task and allowed to eat them. **b**, Each trial involved three kinds of cues: congruent cues (either left or right), for which participants had to look at the same side as where the video was going to be displayed to obtain the snack; incongruent cues (either left or right), for which participants had to look at the opposite side to where the video was going to be displayed to obtain the snack; and CS− cues, for which participants were not required to do any specific action.

dwell time in the Pavlovian ROI during the anticipation to be larger after presentation of a congruent cue than an incongruent cue. A planned contrast analysis using $F$-tests conducted on the cue condition (congruent, incongruent) with the following weights (+1, −1) confirmed that the dwell time spent in the Pavlovian ROI was significantly larger after the presentation of a congruent cue compared to an incongruent cue ($F(1,41)=133.61$; $P<0.001$; $\eta^2_p=0.765$; 90% CI (0.646, 0.824)), suggesting that participants successfully learned where to look to obtain the food (see Fig. 6b).

Anticipatory gaze direction in the instrumental ROI. We defined the instrumental ROI as the location that had to be looked at to obtain the food outcome. We expected that dwell time in the instrumental ROI during the anticipation phase would be larger after the perception of a congruent than after an incongruent cue. A planned contrast analysis using $F$-tests conducted on the cue condition (congruent, incongruent) with the following weights (+1, −1) confirmed that dwell time spent on the instrumental location was significantly larger after the perception of a congruent cue compared to an incongruent cue ($F(1,41)=5.41$; $P=0.025$; $\eta^2_p=0.117$; 90% CI (0.008, 0.272)), suggesting the presence of Pavlovian interference on the instrumental action (see Fig. 6a).

*Outcome devaluation.* Paired $t$-tests showed that hunger ($t(20)=5.58$; $P<0.001$; $d=1.107$; 90% CI (0.582, 1.616)) and the pleasantness of the favourite food outcome ($t(20)=5.49$; $P<0.001$; $d=0.988$; 90% CI (0.515, 1.447)) significantly decreased after selective satiation compared to before satiation (see Fig. 3c).

*Outcome-devaluation-induced changes.* Outcome-devaluation-induced changes were measured by comparing the dwell time during the last session before satiation with those during the test session administered after selective satiation.

Anticipatory gaze direction in the instrumental ROI. We expected devaluation to decrease the influence of instrumental control over gaze direction behaviour and thereby to globally decrease the dwell time spent in the instrumental ROI after the perception of both the congruent and the incongruent cues. To formally test our hypothesis, we ran a 2 (session: pre- or post-satiation) × 2 (group: satiation or control) × 2 (cue: congruent or incongruent) mixed repeated measures ANOVA on dwell time in the instrumental ROI. As predicted, we found a significant session by group interaction ($F(1,38)=15.62$; $P<0.001$; $\eta^2_p=0.291$; 90% CI (0.101, 0.451)), indicating that devaluation decreased the dwell time in the instrumental ROI significantly more for the satiation group compared to the controls. However, there was no significant session by group by cue interaction ($F(1,40)=0.81$; $P=0.373$; $\eta^2_p=0.020$; 90% CI (0.000, 0.134)). Outcome devaluation did not seem to differentially affect dwell time on the instrumental ROI for the congruent and incongruent cues (see Fig. 6c). Moreover, this analysis revealed the main effect of congruency ($F(1,36)=5.58$; $P<0.024$; $\eta^2_p=0.134$; 90% CI (0.010,0.302)) that was not modulated by any kind of interaction. A follow-up 2 (group: satiation or control) × 2 (cue: congruent or incongruent) mixed repeated measures ANOVA on dwell time in the instrumental ROI during the test session revealed the main effect of congruency ($F(1,40)=7.36$; $P=0.010$; $\eta^2_p=0.155$; 90% CI (0.022, 0.316)), but no interaction between congruency and group ($F(1,40)=0.16$; $P=0.693$; $\eta^2_p=0.004$; 90% CI (0.000, 0.084)), suggesting that there was no statistically significant evidence that conflict effect was modulated by outcome devaluation.

Anticipatory gaze direction in the Pavlovian ROI. We expected outcome devaluation to decrease the influence of instrumental control more than that of Pavlovian control over gaze behaviour. Therefore, we expected that outcome devaluation would decrease the dwell time in the Pavlovian ROI after the perception of the congruent cue (because of the reduction of the instrumental influence) but not after the perception of the incongruent cue (which solely reflects the Pavlovian influence). To formally test our hypothesis, we ran a 2 (session: pre- or post-satiation) × 2 (group: satiation or control) × 2 (cue: congruent or incongruent) mixed repeated measures ANOVA on dwell time in the Pavlovian ROI. As predicted, this analysis revealed a significant session by group by cue interaction ($F(1,40)=9.20$; $P=0.004$; $\eta^2_p=0.187$; 90% CI (0.038, 0.350)). This suggests that outcome devaluation differentially affects dwell time in the Pavlovian ROI after the perception of the congruent cue (which reflects the combined influence of the instrumental and Pavlovian systems) and after the perception of the incongruent cue (which solely reflected the Pavlovian influence; see Fig. 6d). Follow-up tests revealed that a 2 (group: satiation or control) × 2

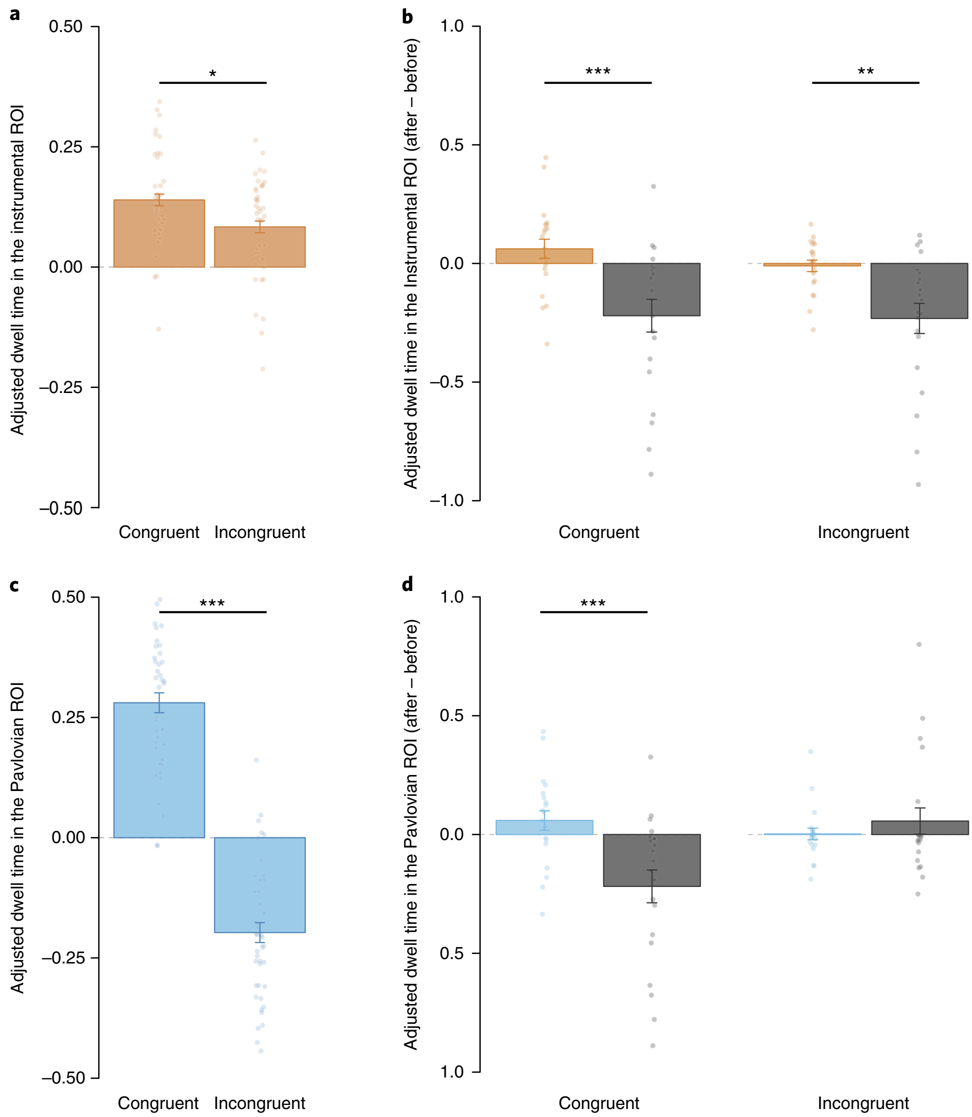


**Fig. 6 | Illustration of the main effects during Experiment 3. a**,**c**, Adjusted dwell time spent in the instrumental ROI (**a**) and the Pavlovian ROI (**c**) during the action screen after the perception of a congruent or incongruent cue. $F(1,41)=5.41$; $^{*}P=0.025$; $\eta^2_p=0.117$; 90% CI (0.008, 0.272). $F(1,41)=133.61$; $^{***}P<0.001$; $\eta^2_p=0.765$; 90% CI (0.646, 0.824). **b**,**d**, Influence of the outcome-devaluation procedure on the dwell time in the instrumental ROI (**b**) and the Pavlovian ROI (**d**) for the satiation group that underwent the devaluation procedure and for the control group that did not: bars represent a change score in the dwell time induced by the outcome-devaluation procedure (change score = dwell time in a specific ROI after satiation − dwell time in a specific ROI before satiation). $F(1,40)=12.02$; $^{***}P=0.001$; $\eta^2_p=0.231$; 90% CI (0.063, 0.393). $F(1,40)=10.75$; $^{**}P=0.002$; $\eta^2_p=0.212$; 90% CI (0.051, 0.374). Error bars represent the between-subject s.e.m.; all plots are based on data from 42 participants.

(cue: congruent or incongruent) interaction was significant after devaluation ($F(1,40)=9.56$; $P=0.004$; $\eta^2_p=0.193$; 90% CI (0.041, 0.356)) but not before ($F(1,40)=0.057$; $P=0.813$; $\eta^2_p=0.001$; 90% CI (0.000, 0.060)). After devaluation, the satiation group's dwell time in the Pavlovian ROI significantly decreased compared with the control group after the perception of the congruent cue ($F(1,40)=17.89$; $P<0.001$; $\eta^2_p=0.309$; 90% CI (0.120, 0.464)) but not after the perception of the incongruent cue, which descriptively increased ($F(1,40)=1.70$; $P=0.199$; $\eta^2_p=0.041$; 90% CI (0.000, 0.172)).

*Experiment 3 discussion.* We found that when the instrumental system was trained to go in the opposite direction to the Pavlovian system (for example, gaze towards the left while expecting the outcome on the right), the execution of the instrumental action was impaired compared with when the instrumental system was trained to go in the same direction as the Pavlovian system (for example, gaze towards the right while expecting the outcome on the right). This conflict effect supports the idea that the tendency to gaze towards the outcome delivery's expected location is a Pavlovian response that works in parallel to the instrumental system.

Our findings suggest that when gaze direction is overtly controlled by an instrumental action alongside the contribution of the Pavlovian system, the outcome-devaluation procedure impacts the instrumental gaze response much more than the Pavlovian gaze response. The ability of the instrumental influence to flexibly adapt to outcome devaluation without any additional learning is consistent with the interpretation that instrumentally trained actions remain under goal-directed control[11,40,41], unless they have been extensively trained[42]. It is also important to note that in our experiment, the instrumental system as a whole was trained to go in the opposite direction to the Pavlovian system from the outset. Thus, any putative instrumental habits would have also been in conflict with the Pavlovian influence, thereby allowing us to disentangle the Pavlovian response and its sensitivity to outcome devaluation even if instrumental behaviour was under habitual and not goal-directed control.

Together, the three studies provide evidence demonstrating that gaze direction elicited in a Pavlovian conditioning context is a response that is insensitive to outcome devaluation; however, it remains unclear whether this gaze response needs to be associated with a rewarding outcome to be acquired. An alternative explanation for our outcome-devaluation-insensitivity findings could be that what is reflected in the gaze direction is not a Pavlovian response, but rather the spatial allocation of attention towards a perceptually salient event. In the framework used in these studies, the acquisition of Pavlovian responses is tested by contrasting conditions in which a reward appears with greater regularity in a given spatial location (CS+ L and CS+ R conditions) with a condition in which a non-event typically happens with no spatial predictably (the CS– condition). Thus, it remains possible that an affectively neutral event with similar perceptual features (for example, luminance, dynamic, contrast) and predictability with regard to spatial location would have had the same effect.

**Experiment 4.** Experiment 4 contrasts an affectively neutral, perceptually salient event with a rewarding event to determine the extent to which the anticipatory gaze response is driven by learning about rewards as opposed to perceptually salient events more generally. We adapted the Pavlovian conditioning procedure from Experiment 2 to have two different CSs L and two different CSs R: each of the two CSs were either associated with the food outcome (CS+) or with the neutral outcome (CS control). There were two CSs+ (a CS+ L and a CS+ R associated with the food outcome) and two CSs control (a CS control L and a CS control R associated with the neutral outcome). We expected Pavlovian responses based on the outcome value representation (pupil dilation) and Pavlovian responses based on the spatial location of the outcome (gaze direction towards the expected reward direction) to be enhanced for the CSs+ compared to the CSs control.

*Pavlovian learning.* Anticipatory gaze direction. To directly compare the CS+ condition with the CS control, we computed the average dwell time allocated to the congruent ROI (dwell time in the right ROI after the CSs R and dwell time in the left ROI after the CSs L) for the CSs associated with the food outcome (CS+) and the CSs associated with the control outcome (CS control). A planned contrast analysis using *F*-tests conducted on the CS condition (CS+, CS control) with the following weights (+1, −1) showed an increased dwell time in the congruent ROI after the perception of the CS+ compared to the CS control ($F(1,32) = 13.3$; $P = 0.001$; $\eta^2_p = 0.294$; 90% CI (0.088, 0.464)); see Fig. 7a–f.

Pupil dilation. We applied the same contrast to the pupil dilation on the onset of the CS. The analysis revealed that the pupil was less constricted at the onset of the CS+ compared to the onset of the CS control ($F(1,32) = 4.93$; $P = 0.034$; $\eta^2_p = 0.133$; 90% CI (0.006, 0.310)).

*Experiment 4 discussion.* Experiment 4 showed that gaze direction towards the expected location of the rewarding outcome is greater than gaze direction towards the expected location of a neutral outcome that is perceptually matched except for the absence of the food reward. This suggests that gaze direction is a Pavlovian conditioned response that reflects the spatial lateralization of the reward outcome rather than a general tendency to allocate attention towards a perceptually salient event. These findings are consistent with findings in the literature describing the tendency to approach or orient towards an expected reward as a Pavlovian response[13,36,43–45].

## Discussion

We combined Pavlovian conditioning with eye-tracking techniques to investigate the sensitivity of different classes of Pavlovian response to outcome devaluation, and found evidence for the differential sensitivity of distinct Pavlovian responses to outcome devaluation. Whereas conditioned pupil dilation flexibly adapted to changes in outcome value without the need to resample environmental contingencies, anticipatory gaze direction was resistant to changes in outcome value. Although responses insensitive to outcome devaluation have been demonstrated many times in the instrumental system (habitual controller[41]), evidence for devaluation-insensitive Pavlovian responses is sparse, even in animal studies, and typically observed in very specific paradigms such as outcome-specific Pavlovian instrumental transfer[5,18,46] and second-order conditioning[47,48].

We ran additional experiments to exclude alternative interpretations for our present findings. Experiment 3 showed that the tendency to gaze towards the expected location of outcome delivery was present even when the instrumental system mandated gazes to go in the opposite direction. This response tendency persisted despite outcome devaluation, thereby supporting the idea that the gaze direction effects in our first two experiments indeed reflected a Pavlovian response resistant to changes in outcome value. Experiment 4 showed that the gaze response is strongly affected by the extent to which the anticipated outcome is a reward as opposed to merely a perceptually salient event, thereby excluding the possibility that gaze direction solely reflects a more generalized deployment of spatial attention towards perceptually salient events.

Our findings support the idea that Pavlovian conditioning is not a unitary process, but rather involves parallel forms of associative learning involving multiple types of Pavlovian response. The existence of multiple classes of Pavlovian response triggered in parallel by the same stimulus is also consistent with recent evidence in humans[29] and with classical findings in animals[13]. This literature distinguishes between two classes of Pavlovian response: 'preparatory responses' that reflect the motivational properties of the outcome (for example, heart rate) and 'consummatory responses' that reflect the sensory properties of the outcome (for example, chewing for a solid food versus licking for a liquid food outcome[45]). Several studies showed how these different classes of Pavlovian response are executed in parallel and are underlined by distinct neuronal networks[29,40]. Others have suggested that associations between a CS and different aspects of the outcome could be even more extensive, involving associations with sensory, motivational, hedonic and even temporal aspects of the outcome[12,25]. We designed our experimental tasks to obtain responses that reflect two aspects of outcome representation: its current value and its spatial location. However, it is likely that other associations were also being formed during our studies. For instance, these could involve other sensorial aspects of the outcome (such as sweet or savoury taste) or temporal aspects of the outcome (such as the timing of occurrence). It remains to be explored whether Pavlovian responses based on other sensorial representations of the outcome beyond spatial localization are sensitive or insensitive to outcome devaluation.

One possible objection to our conclusions is that the anticipatory gaze response might have remained intact after devaluation not because of the insensitivity of the conditioned response, but because the devaluation procedure had rendered the food outcomes aversive. This is unlikely, because the pleasantness ratings of the food outcomes decreased from pleasant to affectively neutral but did not reach aversive levels. Furthermore, if the CSs took on aversive properties, pupil dilation would have responded equally strongly to the CSs predicting the devalued outcomes and the CSs predicting

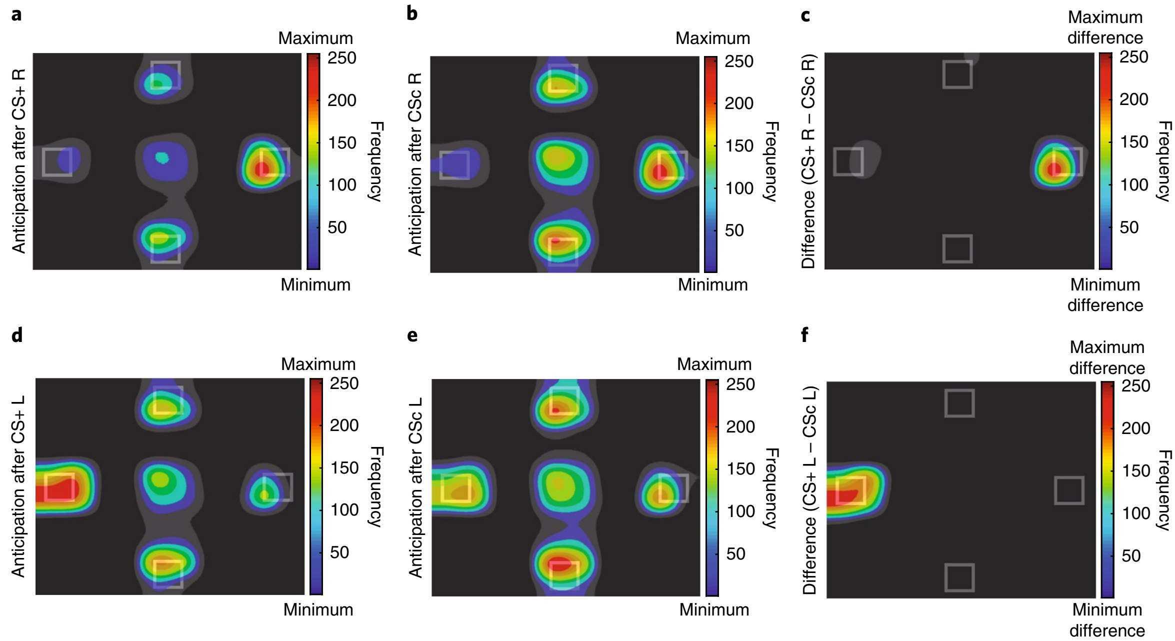


**Fig. 7 | Effect of conditioning during Experiment 4. a**,**b**,**d**,**e**, Heatmaps of the fixation patterns during the anticipation screen (normalized frequency) after the offset of the CS that predicted: the video of a hand delivering a snack to the right side of the screen (CS+ R; **a**); the video of an empty hand to the right side (CS-control (CSc) R; **b**); the video of a hand delivering a snack to the left side of the screen (CS+ L; **d**); and the video of an empty hand to the left side of the screen (CSc L; **e**). **c**,**f**, Heatmaps of the normalized difference between the fixation pattern during the anticipation screen after the offset of CS+ R and the offset of CSc R (**c**) and after the offset of CS+ L and the offset of CSc L (**f**). All plots are based on data from 33 participants.

the valued outcomes, as both CSs would have had strong affective significance for the organism. However, the dilatory CS responses to the devalued outcome were decreased following the devaluation procedure, suggesting that the devalued CSs elicited reduced arousal. A second possible objection is that the anticipatory gaze response might have been driven by the instructions asking participants to focus on the cue and watch what happens next. This is unlikely, as in Experiment 3 the tendency to gaze towards the outcome delivery location was present even when participants were instructed to look at the opposite location.

An important theoretical question raised by our findings is whether the coexistence of responses that are sensitive or insensitive to outcome devaluation in the Pavlovian system mirrors the coexistence of multiple controllers (habits and goal-directed) in the instrumental system. Recently, it has been proposed that model-based and model-free algorithms used to describe the goal-directed and habitual controllers in the instrumental system could also potentially be applied to describe multiple controllers in the Pavlovian system[8,49]. In model-based reinforcement learning algorithms, the value of an instrumental action is computed on the basis of a rich knowledge of the states of the world, including the value of outcomes in those states—they therefore predict outcome-devaluation-sensitive behaviours. On the other hand, in model-free reinforcement learning algorithms the value of an instrumental action is updated incrementally via prediction error, without an internal representation of the states of the world—they therefore predict outcome-devaluation-insensitive behaviours[49,50]. This proposal could account for the coexistence of parallel Pavlovian behaviours that respond differentially to changes in outcome value. Nonetheless, the typical conceptualization of model-free reinforcement learning as utilized in the instrumental domain does not seem to provide a satisfactory account of our findings. In our findings, the outcome devaluation insensitive Pavlovian responses seemed to encode information about a particular sensory property of the outcome (its spatial location). Such sensory information about an outcome cannot be learned in a model-free reinforcement learning algorithm, at least as it is typically conceived. The model-free algorithm learns a cached value for the cue based on the extent to which that cue predicted reward in the past, but such a cached value signal does not encode any information about the cue's sensory features. Instead, it appears that a form of stimulus–stimulus (features) association must be driving the devaluation-insensitive Pavlovian phenomenon. Stimulus–stimulus learning would typically be more associated with a model-based framework, as such learning would underpin the state–space transition model needed for model-based inference. As a result, the model-based versus model-free distinction utilized in instrumental conditioning to account for the distinction between goal-directed and habitual learning may not readily apply to the two classes of Pavlovian conditioned response described here. When taken alongside the fact that typical models of Pavlovian conditioning are model-free, and therefore cannot account for devaluation-sensitive Pavlovian behaviour, our findings highlight the need to develop new computational approaches that might better capture the distinction between different forms of Pavlovian conditioning, which vary in their devaluation sensitivity.

Interestingly, the devaluation-insensitive responses that we found in Pavlovian conditioning do not seem to require overtraining to manifest, which is different from the devaluation-insensitive responses classically found in instrumental conditioning (habits)[34,37,40]. In instrumental conditioning, the goal-directed and habitual influences target the same instrumental action (pressing a button), whereas in our experiments, multiple Pavlovian responses (anticipatory gaze behaviour and pupil dilation) were executed in parallel without being in conflict with each other. The absence of shared/conflicting response pathways might potentially mitigate against the need to arbitrate between the two Pavlovian strategies, allowing both to independently operate in parallel irrespective of training duration.

The conceptualization of parallel Pavlovian responses with different sensitivities to outcome devaluation could guide future

**Table 1 | Summary of Pavlovian contingencies across the four experiments**

| | | Outcome 1 | | Outcome 2 | | No snack |
|---|---|---|---|---|---|---|
| | | Left (%) | Right (%) | Left (%) | Right (%) | |
| Experiments 1 and 3 | CS+ L | 70 | 15 | 0 | 0 | 15 |
| | CS+ R | 15 | 70 | 0 | 0 | 15 |
| | CS– | 15 | 15 | 0 | 0 | 70 |
| Experiments 2 and 4 | CS1+ L | 70 | 10 | 10 | 0 | 10 |
| | CS1+ R | 10 | 70 | 0 | 10 | 10 |
| | CS2+ L | 10 | 0 | 70 | 10 | 10 |
| | CS2+ R | 0 | 10 | 10 | 70 | 10 |
| | CS– | 10 | 10 | 10 | 0 | 70 |
| | CS– | 0 | 10 | 10 | 10 | 70 |

In Experiment 2, outcome 1 was a video of the delivery of a salty snack and outcome 2 was a video of the delivery of a sweet snack, whereas in Experiment 3, outcome 1 was a video of the delivery of a snack and outcome 2 was a video of the experimenter's empty hand.

attempts to find evidence for devaluation-insensitive Pavlovian responses. The existence of Pavlovian responses that persist in spite of the fact that the outcome is no longer valued could provide additional insight into pathological situations where undesirable outcomes are nevertheless assigned high behavioural priority.

## Methods

**Participants.** For Experiment 1, which was a between-subject design, 40 participants (24 females) with a mean age of 26 years (s.d. = 6.95 years) were recruited. For Experiment 2, which was a within-subject design, 20 participants (14 females, 1 agender) with a mean age of 25.1 years (s.d. = 9 years) were recruited. For Experiment 3, which was a between-subject design, 42 participants (23 females) with a mean age of 25.7 years (s.d. = 8.6 years) were recruited. For Experiment 4, 34 participants (23 females) with a mean age of 28 years (s.d. = 10.57 years) were recruited. One participant was excluded from the analysis for not liking any of the snack options proposed (the most liked option for that participant was rated 3 out of 10).

The planned sample size was motivated by a power analysis conducted with G*power[51]. The effect sizes of interest that we focused on regarded the Pavlovian influence on pupil dilation. For Experiments 1 to 3, these effects were extracted from a previous study[8] and from an independent pilot study ($n = 11$) using a framework similar to the one we used in Experiment 1 (Cohen's $d_z = 0.62$; $d_z = 0.57$). The analysis revealed that a sample size of 20 participants per group was required to obtain a power of 80%. For Experiment 4, we averaged the previous effect sizes with the effect size that we obtained in Experiment 1 and Experiment 2 ($d_z = 0.33$; $d_z = 0.39$). The analysis revealed that a sample size of 34 participants was required to obtain a power of 80%. Note that while Experiments 1 to 3 were conducted at the California Institute of Technology in Pasadena, CA, USA, Experiment 4 was conducted at the University of Geneva, Switzerland.

For the four experiments: (1) all participants were prescreened to ensure they were not dieting; (2) they were asked not to eat for at least 6 h before the experimental session (but were allowed to drink water); (3) written informed consent was obtained from all the participants, according to a protocol approved by the Human Subject Protection committee of the California Institute of Technology for Experiments 1–3; for Experiment 4, the protocol was approved by the Faculty of Psychology and Educational Sciences committee of the University of Geneva; (4) before the beginning of the experimental procedure the participants completed demographic and personality questionnaires.

**Materials.** *Stimuli.* For the three experiments the cues consisted of three neutral fractal images (see Supplementary material). The reward outcome consisted of a 3-s long video of the experimenter's hand delivering the participant's favourite snack into a small bag. At the end of each session, participants received the bag containing the snacks they had collected during the task, to consume. The correspondence between the amount of food consumed at the end of each session was not identical (1 video:1 piece of snack) but proportional. This proportion varied from 1:2 to 1:6 according to the amount of calories per individual piece of the snack selected by the participant. The neutral outcome used in Experiment 4 consisted of a 3-s long video of the experimenter's hand approaching the bag in a highly similar fashion to the reward outcome video but without any snack. All stimuli were displayed on a computer screen with a visual angle of 6° using Psychtoolbox 3.0, a visual interface implemented on Matlab (version 8.6; The Mathworks Inc.).

*Pupil dilation and gaze direction.* Pupil dilation and gaze direction were used to reflect two classes of Pavlovian responses. Pupil dilation on cue presentation was used as an index reflecting a Pavlovian response based on the value representation of its associated outcome[8,30,31]. Anticipatory gaze direction was used as an index reflecting a Pavlovian response based on the spatial localization representing its associated outcome. To obtain these measurements, an infrared camera continuously recorded a video of the participant's pupil at 30 frames per second. The eye-tracker was calibrated using a nine-point calibration screen at the beginning of each session. Pupil diameter and the *xy* coordinates of the pupil on the screen were extracted using the open source eye-tracking software MrGaze (https://github.com/jmtyszka/mrgaze/). Before statistical analysis, the pupil data were preprocessed to remove eye blinks and extreme variations. A prestimulus baseline pupil size average of 1 s was calculated for each trial and subtracted from each subsequent data point to establish baseline-corrected pupil response. The statistical analysis was conducted using the average pupil diameter between 0.5 and 1.8 s after stimulus onset. This is the time window after stimulus presentation that was previously found to be responsive during conditioning[8,30]. The averaged pupil diameter was adjusted to account for linear trends independently of the trial type and changes related to switching responses from one side of the screen to the other[8]. The dwell time in the ROIs was extracted through the EyeMMV toolbox[52]. The ROIs were defined as squares centred on the food outcome delivery video, but 25% bigger than the actual video. Moreover, the index reflecting pupil dilation was adjusted by regressing out the gaze position on the screen and the index reflecting gaze direction was adjusted by regressing out the pupil size[53,54].

In Experiment 2, eye data were down sampled to 15 frames per second because of a technical problem. This resolution was still sufficient for the analysis of the pupil dilation and the dwell time in the ROI.

In Experiment 3, the dwell time in the particular ROIs during anticipation was used as the measure of interest. We defined the Pavlovian ROI as being the most likely location of the food outcome delivery, and the instrumental ROI as being the location that had to be looked at to obtain the food outcome. In contrast to Experiments 1 and 2, this experiment required the provision of online feedback based on the participants' gaze direction, as we implemented an instrumental response contingency. To ensure that this instrumental response was not overly difficult for participants to implement, we defined bigger ROIs: 50% bigger than the actual squares displayed on the screen and we recorded eye movements at 500 Hz using an EyeLink 1000 Plus desktop-mounted eye tracker. The eye tracker was calibrated using a five-point calibration screen at the beginning of each session. Experiment 4 was conducted with the same eye-tracking methods as Experiment 3. To keep the measurements as comparable as possible with Experiment 1 and Experiment 2, we extracted the dwell time in the ROIs through the EyeMMV toolbox[52].

**Statistical analyses.** All statistical analysis was conducted using RStudio v. 1.0.36 with R 3.4.3 (RStudio, Inc., 2015). We used a repeated measures ANOVA and planned contrasts according to the a priori hypotheses. When necessary, we verified homogeneity of variance, and the normality of the residuals distribution was verified through visual inspection but not formally tested.

Specifically, in Experiments 1 and 2 we ran three planned contrast analyses according to our a priori hypothesis. The first compared the dwell time in the left ROI after the perception of the CS+ L (weight contrast + 1) to the CS+ R (weight contrast −0.5) and the CS– (weight contrast −0.5). The second compared the dwell time in the right ROI after the perception of the CS+ R (weight contrast + 1) to the CS+ L (weight contrast −0.5) and the CS– (weight contrast −0.5). The third compared the pupil dilation during the perception of the CS+ R (weight contrast +0.5) and the CS+ L (weight contrast +0.5) to the CS– (weight contrast −1). In Experiment 3, we ran two planned contrast analyses. The first compared the dwell time spent in the Pavlovian ROI after the perception of the congruent cue (weight contrast +1) to the incongruent cue (weight contrast −1) and the second compared the dwell time spent in the instrumental ROI after the perception of the congruent cue (weight contrast +1) to the incongruent cue (weight contrast −1). In Experiment 4, we ran two planned contrast analyses, comparing the CS+ (weight contrast +1) to the CS control (weight contrast −1) on the pupil dilation and the dwell time in the congruent ROI.

Effect sizes were measured as $\eta^2_p$ for the repeated measures ANOVA and planned contrasts and as $d$ for the $t$-tests. All $t$-tests were two-tailed.

Data collection and analysis were not performed blind to the conditions of the experiments.

**Procedure.** *Experiment 1.* The experimental procedure involved four main parts. First, participants selected their favourite snack. Second, they completed a Pavlovian conditioning task. Third, half of the participants underwent an outcome devaluation procedure (satiation group) while the other half of the participants were asked to wait without performing any particular task (control group). Finally, all participants performed a test session under extinction.

Snack selection. Participants were presented with a selection of individual pieces of six snacks divided into two categories: sweet (M&M's, Buncha Crunch candy, almonds covered in cacao) and savoury (roasted cashews, roasted peanuts, Goldfish). They were asked to taste each sample and to choose the snack they liked the most and felt like eating during the experiment. Each participant's favourite snack from the selection was used as a food outcome during the Pavlovian conditioning task.

Pavlovian conditioning task**.** Participants learned associations between the delivery of their favourite food outcome and three different cues while their eye movements and pupil responses were being recorded. The task consisted of three learning sessions lasting approximately 12 min each. Each session comprised 54 trials leading to a total of 162 trials. At the beginning of each trial, four squares (6° visual angle each) highlighted by a white frame were displayed at the top and bottom horizontal centre (15° visual angle on the *x* axis from the centre) and the left and right vertical centre (7° visual angle on the *y* axis from the centre). These squares stayed on the screen for the duration of the whole trial.

In each trial, participants first saw a cue either in the upper or lower white frames, then an empty screen with only the background white frames and, finally, a video of the experimenter's hand delivering their favourite snack into a small bag. The video appeared either in the left or the right white frame (see Fig. 1a). Critically, one cue was more often associated with the food outcome delivery on the left side of the screen (CS+ L); one cue was more often associated with the outcome delivery on the right side of the screen (CS+ R); and another cue was more often associated with no outcome delivery (CS–; see Fig. 1b). Specifically, one cue predicted the delivery of a specific outcome 70% of the time (for example, outcome to the left); the remaining 30% of the time it was followed by one of two other possible outcomes (for example, 15% outcome to the right and 15% no outcome; see Table 1).

The order of the trial presentation was fully randomized within participants, whereas the assignment of the neutral images to particular Pavlovian cue conditions (CS+ L, CS+ R, CS–) was counterbalanced across participants.

Participants were instructed to focus on the cue and to try to predict what would happen next. They were also instructed to move their eyes freely around the computer screen, unless a fixation cross was present (during the ITI), in which case they were asked to look at the fixation cross. At the end of each session, the participants received a bag containing the snacks they had earned during the task, to consume.

Outcome devaluation. Participants in the satiation group ($n=20$) were presented with a large bowl containing a very large amount of the food outcome used in the Pavlovian conditioning task. They were asked to eat until they found the food to be no longer palatable. Levels of hunger and food pleasantness were measured through visual analogue scales before and after the outcome-devaluation procedure. Participants in the control group ($n=20$) were asked to take a 5-min break. The allocation of participants to the groups was sequential: the first half were assigned to the control group and the second half to the satiation group.

Test session**.** The test session was composed of 42 trials identical to the Pavlovian conditioning session, except that they were administered under extinction, meaning that no food outcome was delivered for any of the cues. The reason for administering this session under extinction (no outcome delivery) was to assess the influence of the outcome devaluation on the conditioned responses without the confounding effects of the outcome itself.

*Experiment 2.* The experimental procedure involved four main parts. First, participants selected their favourite sweet snack and their favourite savoury snack. Second, they completed a Pavlovian conditioning task. Third, they underwent an outcome-devaluation procedure. Finally, they performed the test session under extinction.

Snack selection. Participants were presented with a selection of individual pieces of 16 snacks divided in two categories: sweet (M&M's, Buncha Crunch candy, almonds covered in cacao, Skittles, cereal covered in chocolate, raisins, yogurt-covered raisins, Milk Chocolate Morsels) and savoury (roasted cashews, roasted peanuts, Goldfish, Simply Balanced Popcorn, cheese-flavoured crackers, Ritz Bits cheese crackers, potato sticks, pretzel sticks). They were asked to taste each sample and to choose their favourite savoury snack and their favourite sweet snack. The participants' favourite snacks were used as outcomes during the Pavlovian conditioning task.

Pavlovian conditioning task. The task was similar to Experiment 1 but consisted of two learning sessions lasting approximately 15 min each. Each session was composed of 60 trials leading to a total of 120 trials. The four squares highlighted by a white frame were slightly more distant: they were displayed at the top and bottom horizontal centre (18° visual angle on the *x* axis from the centre) and the left and right vertical centre (9° visual angle on the *y* axis from the centre).

In each trial, participants first saw a cue either in the upper or lower white frames, then an empty screen with only the background white frames and, finally, a video of the experimenter's hand delivering their favourite snack into a small bag. The video appeared either in the left or the right white frame. Critically, one cue was more often associated with sweet food outcome delivery on the left side of the screen (CS+ sweet L); one cue was more often associated with sweet food outcome delivery on the right side of the screen (CS+ sweet R); one cue was more often associated with savoury food outcome delivery on the left side of the screen (CS+ savoury L); one cue was more often associated with savoury food outcome delivery on the right side of the screen (CS+ savoury R); and another cue was more often associated with no outcome delivery (CS–; see Table 1). Specifically, one cue predicted the delivery of a specific outcome 70% of the time (for example, sweet food outcome to the left), the remaining 30% of the time the cue was followed by one of the other three possible outcomes (for example, 10% sweet food outcome on the right; 10% savoury food outcome on the left; 10% no outcome; see Table 1). Participants were instructed to focus on the image and to try to predict what was going to happen next. They were instructed to move their eyes freely around the computer screen, unless a fixation cross was present (during the ITI), in that case they were ask to look at the fixation cross.

The order of the trial presentation was pseudo-randomized within participants with a maximum of three consecutive repetitions of the same kind of trial and with the first ten trials of the first session to be reinforced with the outcome they predicted more frequently (for example, savoury food to the left for CS+ savoury left). The assignment of the neutral images to particular Pavlovian cue conditions (for example, CS+ savoury L, CS–) was counterbalanced across participants.

At the end of each session, the participants received a bag containing the snacks they collected during the task, to consume.

Outcome devaluation**.** Participants were presented with a large bowl containing a very large amount of one of the two food outcomes used in the Pavlovian conditioning task. They were asked to eat it until they found the target food no longer palatable. The level of hunger and food pleasantness was measured through a visual analogue scale before and after the selective satiation procedure[55]. The food chosen for the devaluation procedure was counterbalanced across participants.

Extinction session. The test session was composed of 60 trials identical to the Pavlovian conditioning session, except we used a strategy to prevent extinction from occurring[56]. Participants were explicitly told that they would not be able to see any food outcome delivery video during this phase, because the area where they were usually displayed would be hidden by two black patches for the whole duration of the session, but that they should assume that all the outcome deliveries would still occur as they had during the previous sessions. They were also asked to press a key to indicate which one of the two black patches was obscuring the outcome delivery video. The reason for using this strategy was to measure the influence of the outcome devaluation on the Pavlovian responses without confounding effects of the outcome itself, and at the same time to prevent the effects of behavioural extinction (for example, disappearance of the conditioned responses due to lack of reinforcement) from happening too quickly[56].

*Experiment 3.* The experimental procedure involved four main parts. First, participants selected their favourite snack. Second, they completed a Pavlovian instrumental conflict task. Third, half of the participants underwent an outcome-devaluation procedure (satiation group) while the other half of the participants were asked to wait without performing any particular task (control group). Finally, all the participants performed a test session under extinction.

Snack selection. The snack selection was identical to that in Experiment 2.

Pavlovian instrumental conflict task. Participants learned associations between different cue stimuli, two gaze actions (looking at the right side or the left side of the screen) and the delivery of their favourite food outcome. Unlike Experiments 1 and 2, the outcome delivery was contingent on the gaze behaviour so as to introduce an instrumental action. As in Experiment 2, the task consisted of two learning sessions composed of 60 trials each and four squares highlighted by a white frame were displayed on the screen for the duration of the whole trial.

In each trial, participants first saw a cue either in the upper or lower white frames, then an empty screen with only the background white frames. During the empty screen they had to look either to the right or left side of the screen based on the instrumental contingency associated with the cue they had just seen. If they looked at the correct side of the screen, a video depicting the experimenter's hand delivering the food outcome in a small bag was displayed on either the right or left side of the screen, indicating that they had just collected a piece of their favourite snack (see Fig. 5a). If they looked at the incorrect side of the screen, the video of the food outcome delivery was displayed behind a transparent red square either to the left or the right side of the screen, indicating that the participants did not successfully collect a piece of their favourite snack (see Fig. 5a). Critically, for some cues, participants had to look at the same location as the one where the outcome delivery video was going to appear (congruent trials) to obtain the food outcome. For other cues, participants had to look in the opposite direction to the one where the outcome delivery video was going to appear (incongruent trial). As illustrated in Fig. 5b, one cue was more often associated with food outcome delivery on the left side of the screen and required participants to look at the left side to obtain the food outcome (congruent cue L); one cue was more often associated with outcome delivery on the left side of the screen and required participants to look at the right side of the screen to obtain the food outcome (incongruent cue L); following the same logic, one cue was more often associated with food outcome delivery on the right side of the screen and required participants to look at the right side to obtain the food outcome (congruent cue R); one cue was more often associated with outcome delivery on the right side of the screen and required participants to look at the left side of the screen to obtain the food outcome (incongruent cue R); the last cue was simply associated with the absence of the food outcome delivery (CS–). In summary, each cue carried both instrumental (gaze action to the left or right) and Pavlovian (food outcome delivery displayed on the left or right) information. The instrumental contingencies (cue−action) were probabilistic: 70% of the time a particular action (for example, look left) after the perception of a particular cue (for example, congruent L) led to a particular food outcome (successful food outcome delivery on the left side of the screen) and 30% of the time it led to no outcome delivery; the Pavlovian contingencies (cue−outcome) were probabilistic and were exactly the same as Experiment 1 (see Table 1). The participants were instructed to focus on the cue image and to try to obtain as many food outcomes as possible. They were also instructed that, for each cue, there was a correct action to be performed to collect the food outcome, however, if a red square appeared on top of the outcome delivery video, it indicated that a piece of their favourite snack was not successfully collected. Participants were instructed to look at the fixation cross, when the fixation cross was presented on the screen.

The order of the trial presentation was pseudo-randomized within participants with a maximum of three consecutive repetitions of the same kind of trial and with the first ten trials of the first session to be reinforced with the outcome they predicted more frequently (for example, food to the left for the congruent L or incongruent L). The assignment of the neutral images to particular cue conditions (for example, congruent L, incongruent R) was counterbalanced across participants.

At the end of each session, the participants received the bag containing the snacks they collected during the task, and they were invited to consume those snacks.

Outcome devaluation. The outcome-devaluation procedure was identical to that in Experiment 1.

Extinction session. The test session was composed of 60 trials, identical to the previous sessions except that we used the same strategy as in Experiment 2 to mitigate the effects of extinction on responding.

*Experiment 4.* The experimental procedure involved two main parts. First, participants selected their favourite snack. Second, they completed a Pavlovian conditioning task.

Snack selection. Participants were presented with a selection of individual pieces of 12 snacks divided into two categories: sweet (M&M's, Maltesers, almonds covered in dark chocolate, Skittles, coconut covered in dark chocolate, raisins) and savoury (roasted cashews, roasted peanuts, Goldfish, organic salted popcorn, Ritz crackers, pretzel sticks). They were asked to taste each sample and to choose their absolute favourite snack. The participant's favourite snack was used as the outcome during the Pavlovian conditioning task.

Pavlovian conditioning task. The task was similar to Experiment 2 but consisted of three learning sessions instead of two.

In each trial, participants first saw a cue either in the upper or lower white frames, then an empty screen with only the background white frames and, finally, a video of the experimenter's hand delivering their favourite snack into a small bag. The video appeared either in the left or the right white frame. Critically, there was a neutral outcome consisting of a video of the experimenter's hand approaching the small bag without any snack. One cue was more often associated with food outcome delivery on the left side of the screen (CS+ L); one cue was more often associated with food outcome delivery on the right side of the screen (CS+ R); one cue was more often associated with the neutral outcome on the left side of the screen (CS control L); one cue was more often associated with the neutral outcome on the right side of the screen (CS control R); and another cue was more often associated with no outcome delivery (CS–; see Table 1). Specifically, one cue predicted the delivery of a specific outcome 70% of the time (for example, food outcome to the left), the remaining 30% of the time the cue was followed by one of the other three possible outcomes (for example, 10% food outcome on the right; 10% control outcome on the left; 10% no outcome; see Table 1). Participants were instructed to focus on the image and to try to predict what was going to happen next. They were instructed to move their eyes freely around the computer screen, unless a fixation cross was present (during the ITI), in which case they were ask to look at the fixation cross.

The order of the trial presentation was pseudo-randomized within participants with a maximum of three consecutive repetitions of the same kind of trial and with the first ten trials of the first session to be reinforced with the outcome they predicted more frequently (for example, food to the left for CS+ L). The assignment of the neutral images to particular Pavlovian cue conditions (for example, CS+ L, CS control R) was counterbalanced across participants.

At the end of each session, the participants received a bag containing the snacks they had collected during the task, to consume.

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

## Code availability

Code used to generate the figures and the results of the four studies reported in this manuscript is available through the Open Science Framework repository: https://osf.io/rve2p/

## Data availability

Data from the four studies reported in this manuscript are available through the Open Science Framework repository: https://osf.io/rve2p/

Received: 20 March 2018; Accepted: 21 December 2018;
Published online: 25 February 2019

## References

1. Berridge, K. C. & Robinson, T. E. Liking, wanting, and the incentive-sensitization theory of addiction. *Am. Psychol.* **71**, 670–679 (2016).
2. Everitt, B. J. & Robbins, T. W. Drug addiction: updating actions to habits to compulsions ten years on. *Annu. Rev. Psychol.* **67**, 23–50 (2016).
3. Voon, V. et al. Disorders of compulsivity: a common bias towards learning habits. *Mol. Psychiatr.* **20**, 342–352 (2015).
4. Anderson, B. A. The attention habit: how reward learning shapes attentional selection. *Ann. NY Acad. Sci.* **1369**, 24–39 (2016).
5. Eder, A. B. & Dignath, D. Cue-elicited food seeking is eliminated with aversive outcomes following outcome devaluation. *Q. J. Exp. Psychol.* **69**, 574–588 (2016).
6. Nadler, N., Delgado, M. R. & Delamater, A. R. Pavlovian to instrumental transfer of control in a human learning task. *Emotion* **11**, 1112–1123 (2011).
7. Pool, E. R., Brosch, T., Delplanque, S. & Sander, D. Where is the chocolate? Rapid spatial orienting toward stimuli associated with primary rewards. *Cognition* **130**, 348–359 (2014).
8. Prévost, C., McNamee, D., Jessup, R. K., Bossaerts, P. & O'Doherty, J. P. Evidence for model-based computations in the human amygdala during Pavlovian conditioning. *PLoS Comput. Biol.* **9**, e1002918 (2013).
9. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
10. Sali, A. W., Anderson, B. A. & Yantis, S. The role of reward prediction in the control of attention. *J. Exp. Psychol. Human* **40**, 1654–1664 (2014).
11. O'Doherty, J. P., Cockburn, J. & Pauli, W. M. Learning, reward, and decision making. *Annu. Rev. Psychol.* **68**, 73–100 (2017).
12. Delamater, A. R. & Oakeshott, S. Learning about multiple attributes of reward in Pavlovian conditioning. *Ann. NY Acad. Sci.* **1104**, 1–20 (2007).
13. Balleine, B. W. & Killcross, S. Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci.* **29**, 272–279 (2006).
14. 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).
15. Holland, P. C. & Straub, J. J. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *J. Exp. Psychol. Anim. B.* **5**, 65–78 (1979).
16. Ramachandran, R. & Pearce, J. M. Pavlovian analysis of interactions between hunger and thirst. *J. Exp. Psychol. Anim. B.* **13**, 182–192 (1987).
17. Gottfried, J. A., O'Doherty, J. P. & Dolan, R. J. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* **301**, 1104–1107 (2003).
18. Holland, P. C., Lasseter, H. & Agarwal, I. Amount of training and cue-evoked taste-reactivity responding in reinforcer devaluation. *J. Exp. Psychol. Anim. B.* **34**, 119–132 (2008).
19. Robinson, M. J. F. & Berridge, K. C. Instant transformation of learned repulsion into motivational "wanting". *Curr. Biol.* **23**, 282–289 (2013).
20. Pearce, J. M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532–552 (1980).
21. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds. Black, A.H. & Prokasy, W.F.) 64–99 (Appleton Century Crofts, New York, 1972).
22. Sutton, R. S. Learning to predict by the methods of temporal differences. *Mach. Learn.* **3**, 9–44 (1988).
23. Sharpe, M. J. & Schoenbaum, G. Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiol. Learn. Mem.* (2017).
24. Cardinal, R. N., Parkinson, J. A., Hall, J. & Everitt, B. J. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* **26**, 321–352 (2002).
25. Delamater, A. R. On the nature of CS and US representations in Pavlovian learning. *Learn. Behav.* **40**, 1–23 (2012).
26. Nasser, H. M., Chen, Y.-W., Fiscella, K. & Calu, D. J. Individual variability in behavioral flexibility predicts sign-tracking tendency. *Front. Behav. Neurosci.* **9**, 289 (2015).
27. Ahrens, A. M., Singer, B. F., Fitzpatrick, C. J., Morrow, J. D. & Robinson, T. E. Rats that sign-track are resistant to Pavlovian but not instrumental extinction. *Behav. Brain. Res.* **296**, 418–430 (2016).
28. Morrison, S. E., Bamkole, M. A. & Nicola, S. M. Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Front. Neurosci.* **9**, 468 (2015).

29. Zhang, S., Mano, H., Ganesh, G., Robbins, T. & Seymour, B. Dissociable learning processes underlie human pain conditioning. *Curr. Biol.* **26**, 52–58 (2016).
30. Pauli, W. M. et al. Distinct contributions of ventromedial and dorsolateral subregions of the human substantia nigra to appetitive and aversive learning. *J. Neurosci.* **35**, 14220–14233 (2015).
31. Seymour, B., Daw, N. D., Dayan, P., Singer, T. & Dolan, R. J. Differential encoding of losses and gains in the human striatum. *J. Neurosci.* **27**, 4826–4831 (2007).
32. Tricomi, E., Balleine, B. W. & O'Doherty, J. P. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* **29**, 2225–2232 (2009).
33. Dickinson, A., Campos, J., Varga, Z. I. & Balleine, B. Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. B* **49**, 289–306 (1996).
34. Grindley, G. C. The formation of a simple habit in guinea-pigs. *B. J. Psychol. Gen. Sect.* **23**, 127–147 (2011).
35. Hershberger, W. A. An approach through the looking-glass. *Anim. Learn. Behav.* **14**, 443–451 (1986).
36. Guitart-Masip, M. et al. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J. Neurosci.* **31**, 7867–7875 (2011).
37. Adams, C. D. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol.-B* **34**, 77–98 (1982).
38. Balleine, B. W. & Dickinson, A. Instrumental performance following reinforcer devaluation depends upon incentive learning. *Q. J. Exp. Psychol. B* **43**, 279–296 (1991).
39. Valentin, V. V., Dickinson, A. & O'Doherty, J. P. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* **27**, 4019–4026 (2007).
40. Balleine, B. W. & O'Doherty, J. P. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* **35**, 48–69 (2010).
41. Pauli, W. M., Cockburn, J., Pool, E. R., Pérez, O. D. & O'Doherty, J. P. Computational approaches to habits in a model-free world. *Curr. Opin. Behav. Sci.* **20**, 104–109 (2018).
42. Dickinson, A., Balleine, B. W., Watt, A., Gonzalez, F. & Boakes, R. A. Motivational control after extended instrumental training. *Anim. Learn. Behav.* **23**, 197–206 (1995).
43. Guitart-Masip, M., Duzel, E., Dolan, R. J. & Dayan, P. Action versus valence in decision making. *Trends. Cogn. Sci.* **18**, 194–202 (2014).
44. Guitart-Masip, M. et al. Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage* **62**, 154–166 (2012).
45. Konorski, J. *Integrative Activity of the Brain: An Interdisciplinary Approach* (Univ. Chicago, Chicago, 1967).
46. De Tommaso, M., Mastropasqua, T. & Turatto, M. Working for beverages without being thirsty: human Pavlovian-instrumental transfer despite outcome devaluation. *Learn. Motiv.* **63**, 37–48 (2018).
47. Holland, P. C. The effects of satiation after first- and second-order appetitive conditioning in rats. *Pavlovian J. Biol. Sci.* **16**, 18–24 (1981).
48. Holland, P. C. & Rescorla, R. A. The effect of two ways of devaluing the unconditioned stimulus after first-and second-order appetitive conditioning. *J. Exp. Psychol. Anim. B.* **1**, 355–363 (1975).
49. Dayan, P. & Berridge, K. C. Model-based and model-free Pavlovian reward learning: revaluation, revision, and revelation. *Cogn. Affect. Behav. Neurosci.* **14**, 473–492 (2014).
50. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1707–1711 (2005).
51. 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).
52. Krassanakis, V., Filippakopoulou, V. & Nakos, B. EyeMMV toolbox: an eye movement post-analysis tool based on a two-step spatial dispersion threshold for fixation identification. *J. Eye Movement Res.* https://doi.org/10.16910/jemr.7.1.1 (2014).
53. Choe, K. W., Blake, R. & Lee, S.-H. Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation. *Vision Res.* **118**, 48–59 (2016).
54. Nyström, M., Hooge, I. & Andersson, R. Pupil size influences the eye-tracker signal during saccades. *Vision Res.* **121**, 95–103 (2016).
55. Reber, J. et al. Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain* **140**, 1743–1756 (2017).
56. Prévost, C., Liljeholm, M., Tyszka, J. M. & O'Doherty, J. P. Neural correlates of specific and general Pavlovian-to-instrumental transfer within human amygdalar subregions: a high-resolution fMRI study. *J. Neurosci.* **32**, 8383–8390 (2012).

## Acknowledgements
This work was supported by a NIDA-NIH R01 grant (1R01DA040011-01A1) to J.P.O. and W.M.P. and by an Early Postdoctoral Mobility fellowship from the Swiss National Science Foundation (P2GEP1162079) to E.R.P. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors thank O. D. Perez and V. Sennwald for insightful comments on this manuscript.

## Author contributions
E.R.P., W.M.P., C.S.K. and J.P.O. designed the experiments. E.R.P. and C.S.K. collected and analysed the data. E.R.P., W.M.P., C.S.K. and J.P.O. wrote the paper. All authors discussed the results and implications and commented on the manuscript at all stages.

## Competing interests
The authors declare no competing interests.

## Additional information
**Supplementary information** is available for this paper at https://doi.org/10.1038/s41562-018-0527-9.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Correspondence and requests for materials** should be addressed to E.R.P.

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

© The Author(s), under exclusive licence to Springer Nature Limited 2019

naturereseach

Corresponding author(s): E. R. Pool

# Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

## Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |
| ☐ | ☒ | Clearly defined error bars<br>*State explicitly what error bars represent (e.g. SD, SE, CI)* |

*Our web collection on statistics for biologists may be useful.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Behavioural data were collected using Psychtoolbox 3.0 implemented on Matlab (version 8.6; The Mathworks Inc., Natick, MA, USA).<br>In Experiment 1 and Experiment 2, eye tracking data were extracted using the open source eye tracking software MrGaze (https://github.com/jmtyszka/mrgaze/) and the EyeMMV toolbox (Krassanakis et al., 2014)<br>In Experiment 3 and Experiment4 , eye tracking data were collected with an EyeLink 1000 Plus desktop-mounted eye tracker. |
| Data analysis | All statistical analyses were conducted using the RStudio software 1.0.36 with R 3.4.3 (2009-2016 RStudio, Inc) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data from the four studies reported in this manuscript are available through the Open Science Framework repository: https://osf.io/rve2p/

# Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | The planned sample size was motivated by a power analysis conducted with G*power. The effect sizes of interest we focused on was the Pavlovian influence on pupil dilation. For Experiment 1 to 3, these effects were extracted from a previous study and from an independent pilot study (n = 11) using a paradigm similar to the one we used in Experiment 1 (dz = .62, dz = .57). The analysis revealed that sample size of 20 participants per group was required to obtain a power of 80%. For Experiment 4, we averaged the previous effect sizes with the effect size we obtained in Experiment 1 and Experiment 2 (dz = .33, dz = .39). The analysis revealed that a sample size of 34 participants was required to obtain a power of 80%. |
| Data exclusions | In Experiment 4 data from one participant was excluded from the analysis for not liking any of the snack options proposed (the most liked option for that participant was rated 3 out of 10). |
| Replication | We ran 4 variations of the same task and replicated the main finding each time. These 4studies are reported in the main text. |
| Randomization | Allocation to experimental conditions was either randomized, counterbalanced or sequential. |
| Blinding | Blinding was not possible: The experimenter administered the devaluation procedure and therefore was aware of the outcome stimulus that was being devalued. |

# Reporting for specific materials, systems and methods

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique biological materials |
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology |
| ☒ | ☐ Animals and other organisms |
| ☐ | ☒ Human research participants |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |

## Human research participants

Policy information about studies involving human research participants

| | |
|---|---|
| Population characteristics | Forty participants (24 females) with a mean age of 26 years (SD = 6.95 years) were recruited for Experiment 1, which was a between subjects design. Twenty participants (14 females, 1 agender) with a mean age of 25.1 years (SD = 9 years) were recruited for Experiment 2, which was a within subjects design. Forty-two participants (23 females) with a mean age of 25.7 years (SD = 8.6 years) were recruited for Experiment 3, which was a between subjects design. Thirty-four participants (23 females) with a mean age of 28 years (SD = 10.57 years) were recruited for Experiment 4. |

Recruitment

Participants were recruited through flyers posted on campus and librarie. Note that while Experiments 1 to 3 were conducted at the California Institute of Technology in Pasadena, CA, Experiment 4 was conducted at the University of Geneva, Switzerland.