

Published in final edited form as:

*Nat Hum Behav.* 2019 March ; 3(3): 284–296. doi:10.1038/s41562-018-0527-9.

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

Eva R. Pool<sup>\*1</sup>, Wolfgang M. Pauli<sup>1,2</sup>, Carolina S. Kress<sup>1</sup>, and John P. O'Doherty<sup>1,2</sup>

<sup>1</sup>Human Reward and Decision Making Laboratory, Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, California

<sup>2</sup>Computation and Neural Systems Program, California Institute of Technology, Pasadena, California

### Abstract

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 localisation 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, even though once obtained, the object is not experienced as pleasurable<sup>1, 2</sup>. Identifying the underlying mechanisms leading to such paradoxical behaviour has been a major research focus. Of

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence concerning this article should be addressed to Eva R. Pool, Campus Biotech, University of Geneva - CISA, Case Postale 60, 1211 Geneva 20, Switzerland. Eva.pool@unige.ch.

#### Data Availability

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

#### Code Availability

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

#### Author Contributions

E.R.P., W.M.P., C.S.K. and J.O.D. 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.O.D. wrote the paper. All authors discussed the results and implications and commented on the manuscript at all stages.

#### Competing Interests statement

The authors declare no competing interests

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 (e.g., seeking snacks even when completely satiated)<sup>2, 3</sup>. 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<sup>4–11</sup>, whereby reflexive conditioned behaviours can come to be elicited by a conditioned stimulus (CS; e.g., a metronome sound) that predicts the subsequent delivery of an affectively significant outcome (e.g., food)<sup>9, 11–13</sup>.

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<sup>14–16, 7, 15, 17–19</sup>.

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<sup>20–22</sup>. For instance, in model-free Reinforcement Learning (RL) approaches to Pavlovian conditioning, such as the temporal difference algorithm or the Rescorla-Wagner rule, Pavlovian conditioned stimuli 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<sup>23</sup>.

However, Pavlovian conditioning does not appear to be a unitary process, but rather appears to involve several parallel associations between multiple aspects of the outcome<sup>12, 24</sup>. 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<sup>19</sup>. 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<sup>12, 25</sup>. Despite the long-standing conceptualisation 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<sup>19</sup> than others<sup>26</sup> and such differences have often been attributed to inter-individual differences<sup>26–28</sup>. 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 responses in humans<sup>29</sup> 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 (i.e., 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 paradigm to test for the sensitivity of different classes of appetitive Pavlovian responses to outcome value. We used eye-tracking techniques combined with 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.,<sup>29</sup> we adapted this experimental feature to our appetitive conditioning paradigm: one image was more often associated with the food outcome delivery on the left side of the screen (positive conditioned stimulus left; CS+ L); one image was more often associated with the outcome delivery on the right side of the screen (positive conditioned stimulus right; CS+ R) and another image was the more often associated with no outcome delivery (negative conditioned stimulus; CS-; see Figure 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<sup>8, 30, 31</sup>. Anticipatory gaze direction (left vs. right) was taken as a response reflecting a specific sensorial representation of the outcome (i.e., its spatial location). We therefore predicted that, in comparison to pupil dilation responses, gaze direction would be less sensitive to outcome devaluation. In the first experiments (Experiment 1 and Experiment 2) we tested the existence of two classes of Pavlovian responses and their sensitivity to outcome devaluation. In the last experiments (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 toward a perceptually salient event, as opposed to a Pavlovian conditioned response established by learning about reward outcomes.

## Experiment 1

We first tested whether the pupil dilation response and the anticipatory gaze direction reflected patterns of distinct classes of Pavlovian responses as in the Zhang et al.,<sup>29</sup> original study. We expected (a) the pupil dilation to follow a value pattern (i.e., CS+ L and CS+ R different from CS-); (b) the gaze direction to follow a lateralized pattern (i.e., 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).

Secondly, we tested the sensitivity of these two classes of Pavlovian responses 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 (i.e., no outcome was delivered) and Pavlovian responses measured. We expected pupil dilation, but not gaze direction, to flexibly adapt to the decreased outcome value.

## Results

**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 on the left Region of Interest (ROI) after the perception of the CS+ L compared to the CS+ R and the CS-,  $F(1,39) = 21.33$ ,  $p < .001$ ,  $\eta^2_p = .354$ , 90% CI [.155, .503] (see Figure 2 C). A second planned contrasts 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 on the ROI after the perception of the CS+ R compared to the CS+ L and to the CS-,  $F(1,39) = 27.10$ ,  $p < .001$ ,  $\eta^2_p = .41$ , 90% CI [.207, .550] (see Figure 2 B).

*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 the CS+ L and the CS+ R compared to the CS-,  $F(1,39) = 4.45$ ,  $p = .041$ ,  $\eta^2_p = .102$ , 90% CI [.002, .259] (see Figure 2 A).

**Outcome devaluation—Paired t-tests** showed that hunger ( $t(19) = 6.93$ ,  $p < .001$ ,  $d = 1.367$ , 95% CI [.779, 1.938]) and pleasantness of the favourite food outcome ( $t(19) = 6.10$ ,  $p < .001$ ,  $d = 1.853$ , 95% CI [1.005, 2.674]) significantly decreased after the selective satiation compared to before (see Figure 3 A).

**Outcome devaluation induced changes—Anticipatory gaze direction:** We computed the average dwell time spent on the congruent ROI for both CSs+ (i.e., the dwell time on the right ROI after the CS+ R and the dwell time on the left ROI after the CS+ L) for the last session before satiation and during the first half of the extinction test for both the satiation and 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)  $\times$  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  $F(1,38) = 15.02$ ,  $p < .001$ ,  $\eta^2_p = .283$ , 90% CI [.095, .444], but no significant interaction  $F(1,38) = .51$ ,  $p = .478$ ,  $\eta^2_p = .013$ , 90% CI [.000, .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 Figure 4 C).

*Pupil dilation:* We computed a similar index to that for dwell time, by averaging pupil dilation during the CS+ L and the CS+ R for the last session before satiation and during the first half of the extinction session. A 2 (session: pre- or post-satiation)  $\times$  2 (group: satiation or control) mixed repeated measures ANOVA applied to pupil dilation revealed a significant session  $\times$  outcome interaction  $F(1,38) = 4.93$ ,  $p = .032$ ,  $\eta^2_p = .115$ , 90% CI [.005, .276], showing that the decrease in pupil dilation induced by satiation was significantly larger in the satiation group compared to the control group (see Figure 4 A). This suggests that pupil dilation flexibly adapted to outcome devaluation.

## Discussion

Results suggest two distinct classes of Pavlovian response: one reflecting outcome-value, as measured by pupil dilation, and another reflecting the spatial localisation 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 localisation 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 that 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 (i.e., a general decrease in the hunger level) rather than a specific change in outcome-value (i.e., the specific pleasantness of the food outcome). A way to tackle this issue is to use multiple CSs associated with two different food outcomes (e.g., a sweet food and a savoury food) and to devalue only one of the food outcomes<sup>17, 32</sup>.

## Experiment 2

Experiment 2 aimed at extending and replicating 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 responses 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 (i.e., gaze direction) would be devaluation insensitive.

## Results

**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) revealed an increased dwell time on the left Region of Interest (ROI) after the perception of the CS+ L compared to the CS+ R and the CS-,  $F(1,19) = 13.15, p = .002, \eta^2_p = .409, 90\% \text{ CI } [.119, .590]$ . Likewise, a second planned contrasts 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 on the ROI after the perception of the CS+ R compared to the CS+ L and to the CS-,  $F(1,19) = 15.81, p = .001, \eta^2_p = .454, 90\% \text{ CI } [.157, .623]$ .

***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) revealed that the pupil was less constricted for the CS+ L and the CS+ R compared to the CS-,  $F(1,19) = 3.27, p = .086, \eta^2_p = .147, 90\% \text{ CI } [.000, 0.368]$ , although this effect did not reach statistical significance.

**Outcome devaluation**—Paired t-tests showed that hunger ( $t(19) = 5.52, p < .001, d = 1.07, 95\% \text{ CI } [.555, 1.573]$ ), the pleasantness of the food outcome that had been eaten until satiety (e.g., devalued outcome,  $t(19) = 7.19, p < .001, d = 1.760, 95\% \text{ CI } [1.016, 2.483]$ ) and the pleasantness of the food outcome that had not been eaten until satiety (e.g., valued outcome,  $t(19) = 2.780, p = .012, d = .489, 95\% \text{ CI } [.106, .862]$ ) significantly decreased after the selective satiation compared to before. Critically, a 2 (session: pre- or post-satiation)  $\times$  2 (outcome: valued or devalued) repeated measures ANOVA applied to the food pleasantness ratings revealed a significant interaction  $F(1,19) = 28.02, p < .001, \eta^2_p = .596, 90\% \text{ CI } [.331, .723]$  showing that a decrease in pleasantness was significantly larger for the devalued food outcome compared to the valued food outcome (see Figure 3 B).

**Outcome devaluation induced changes**—*Anticipatory gaze direction*: We computed the average dwell time allocated to the congruent ROI for all the CSs+ (i.e., the dwell time in the right ROI after the CS+ R and the dwell time in the left ROI after the 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)  $\times$  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  $F(1,19) = .04, p = .843, \eta^2_p = .002, 90\% \text{ CI } [.000, .100]$ , suggesting that there was no statistically significant evidence that dwell time was sensitive to outcome devaluation (see Figure 4 D).

*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)  $\times$  2 (CS: valued or devalued) repeated measures ANOVA applied to pupil dilation revealed a significant interaction  $F(1,19) = 8.08, p = .010, \eta^2_p = .298, 90\% \text{ CI } [.045, .504]$ , showing that the decrease in pupil dilation induced by satiation was significantly larger for the devalued CS compared to the valued CS (see Figure 4 B). This suggests that, as in Experiment 1, pupil dilation flexibly adapted to outcome devaluation.

## Discussion

Experiment 2 replicated the main finding of Experiment 1: a CS can elicit multiple classes of Pavlovian responses (as measured by in 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 generalised satiation.

However, there is a possible alternative explanation for the results of Experiment 1 and 2. While pupil dilation is evidently non-instrumental (e.g., 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 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 toward 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 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: in that one should be equally able to train the action to go in one direction for reward with equal ease as it is possible to train the action to go in the opposite direction for the same reward<sup>33, 34</sup>. 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<sup>35</sup> 'room through a looking glass' experiment where food-deprived chicks were unable to learn that walking in the opposite direction of 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 toward 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 just perceived. For some of the stimuli (i.e., congruent cues; see Figure 5), participants had to look in the same direction to where the outcome delivery video was going to appear, so that for example, if the cue predicted the food picture to appear on the left, participants had to gaze to the left to obtain the food. However, for other stimuli (i.e., incongruent cues; see Figure 5), participants had to look in the opposite direction to where the outcome delivery video was going to appear, so that if for example, the cue predicted outcome delivery on the left, the participants had to gaze to the right location to obtain that food outcome. Therefore, we fully orthogonalised 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 previous studies<sup>36</sup> 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 attend to collect the food outcome during incongruent trials compared to congruent trials.

Moreover, we 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 whereas 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 (i.e. participants were not over-trained), and after modest training instrumental actions are generally found to be outcome-value sensitive<sup>32, 37–39</sup>.

## Results

### **Pavlovian Instrumental Conflict**—*Anticipatory gaze direction on the Pavlovian ROI:*

We defined the Pavlovian ROI as the location where the food outcome delivery video was the most likely to be displayed given the specific contingencies for a given CS. We expected dwell time on the Pavlovian ROI during the anticipation to be larger after presentation of a congruent (i.e., the condition in which participants had to look at the same location to the one where the food delivery video was going to be displayed to obtain the food) than incongruent cue (i.e., in which participants had to look at the opposite location to the one where the food delivery video was going to be displayed to obtain the food). 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 on the Pavlovian ROI was significantly larger after the presentation of a congruent cue compared to an incongruent cue,  $F(1,41) = 133.61, p < .001, \eta_p^2 = .765, 90\% \text{ CI } [.646, .824]$ , suggesting that participants successfully learned where to look to obtain the food (see Figure 6B).

*Anticipatory gaze direction on the instrumental ROI:* We defined the instrumental ROI as the location that had to be attended to obtain the food outcome. We expected that dwell time on the instrumental ROI during the anticipation phase would be larger after the perception of a congruent 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,  $F(1,41) = 5.41, p = .025, \eta_p^2 = .117, 90\% \text{ CI } [.008, .272]$ , suggesting the presence of Pavlovian interference on the instrumental action (see Figure 6A).

**Outcome devaluation**—Paired t-tests showed that hunger ( $t(20) = 5.58, p < .001, d = 1.107, 90\% \text{ CI } [.582, 1.616]$ ) and the pleasantness of the favourite food outcome ( $t(20) = 5.49, p < .001, d = .988, 90\% \text{ CI } [.515, 1.447]$ ) significantly decreased after selective satiation compared to before satiation (see Figure 3 C).



**Satiation 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 on the instrumental ROI:* We expected devaluation to decrease the influence of the instrumental control over the 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)  $\times$  2 (group: satiation or control)  $\times$  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  $F(1,38) = 15.62, p < .001, \eta_p^2 = .291, 90\% \text{ CI } [.101, .451]$ , 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,  $F(1,40) = .81, p = .373, \eta_p^2 = .020, 90\% \text{ CI } [.000, .134]$ . Outcome devaluation did not seem to differentially affect dwell time on the instrumental ROI for the congruent and the incongruent cue (see Figure 6 C). Moreover, this analysis revealed a main effect of congruency ( $F(1,36) = 5.58, p < .024, \eta_p^2 = .134, 90\% \text{ CI } [.010, .302]$ ) that was not modulated by any kind of interaction. A follow-up 2 (group: satiation or control)  $\times$  2 (cue: congruent or incongruent) mixed repeated measures ANOVA on dwell time on the instrumental ROI during the test session revealed a main effect of congruency,  $F(1,40) = 7.36, p = .010, \eta_p^2 = .155, 90\% \text{ CI } [.022, .316]$ , but no interaction between congruency and group,  $F(1,40) = .16, p = .693, \eta_p^2 = .004, 90\% \text{ CI } [.000, .084]$ , suggesting that there was no statistically significant evidence that conflict effect was modulated by outcome devaluation.

*Anticipatory gaze direction on the Pavlovian ROI:* We expected outcome devaluation to decrease the influence of instrumental control more than Pavlovian control over the gaze behaviour. Therefore, we expected that outcome devaluation would decrease the dwell time on 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 (that solely reflects the Pavlovian influence). To formally test our hypothesis, we ran a 2 (session: pre- or post-satiation)  $\times$  2 (group: satiation or control)  $\times$  2 (cue: congruent or incongruent) mixed repeated measures ANOVA on dwell time on the Pavlovian ROI. As predicted, this analysis revealed a significant session by group by cue interaction ( $F(1,40) = 9.20, p = .004, \eta_p^2 = .187, 90\% \text{ CI } [.038, .350]$ ). This suggests that outcome devaluation differentially affects dwell time on the Pavlovian ROI after the perception of the congruent cue (that reflects the combined influence of the instrumental and Pavlovian systems) and after the perception of the incongruent cue (that solely reflected the Pavlovian influence; see Figure 6 D). Follow-up tests revealed that a 2 (group: satiation or control)  $\times$  2 (cue: congruent or incongruent) interaction was significant after devaluation,  $F(1,40) = 9.56, p = .004, \eta_p^2 = .193, 90\% \text{ CI } [.041, .356]$ , but not before,  $F(1,40) = .057, p = .813, \eta_p^2 = .001, 90\% \text{ CI } [.000, .060]$ . After devaluation, the satiation group's dwell time on the Pavlovian ROI significantly decreased compared with the control group after the perception of the congruent cue ( $F(1,40) = 17.89, p < .001, \eta_p^2 = .309, 90\% \text{ CI } [.120, .464]$ ) but not after the perception of the incongruent cue, which descriptively increased ( $F(1,40) = 1.70, p = .199, \eta_p^2 = .041, 90\% \text{ CI } [.000, .172]$ ).

## Discussion

We found that when the instrumental system was trained to go in the opposite direction than the Pavlovian system (e.g., gaze toward the left while expecting the outcome on the right), the execution of the instrumental action was impaired compared to when the instrumental system was trained to go in the same direction as the Pavlovian system (e.g., gaze toward the right while expecting the outcome on the right). This conflict effect supports the idea that the tendency to gaze toward 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<sup>11, 40, 41</sup>, unless they have been extensively trained<sup>42</sup>. It is also important to note that in our experiment, the instrumental system as a whole was trained to go in the opposite direction than 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 toward a perceptually salient event. In the paradigm 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 (i.e., CS+L and CS+R conditions) with a condition in which a non-event typically happens with no spatial predictability (i.e., CS-condition). Thus it remains possible that an affectively neutral event with similar perceptual features (e.g., 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 (i.e., the pupil dilation) and the Pavlovian responses based on

the spatial location of the outcome, (i.e., gaze direction toward the expected reward direction) to be enhanced for the CSs+ compared to the CSs control.

## Results

**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 (i.e., the dwell time in the right ROI after the CSs R and the 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,  $F(1,32) = 13.3, p = .001, \eta^2_p = .294, 90\% \text{ CI } [.088, .464]$  (see Figure 7 A-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,  $F(1,32) = 4.93, p = .034, \eta^2_p = .133, 90\% \text{ CI } [.006, .310]$ .

## Discussion

Experiment 4 showed that gaze direction toward the expected location of the rewarding outcome is greater than gaze direction toward 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 lateralisation of the reward outcome rather than a general tendency to allocate attention toward a perceptually salient event. These findings are consistent with findings in the animal literature describing the tendency to approach or orient toward an expected reward as a Pavlovian response<sup>13, 36, 43–45</sup>.

## General discussion

We combined Pavlovian conditioning with eye-tracking techniques to investigate the sensitivity of different classes of Pavlovian response to outcome devaluation. We 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 within the instrumental system (i.e., habitual controller<sup>41</sup>), evidence for devaluation insensitive Pavlovian responses is sparse even in animal studies, and typically observed in very specific paradigms such as sensory-specific Pavlovian instrumental transfer<sup>5, 18, 46</sup> and second-order conditioning<sup>47, 48</sup>.

We ran additional experiments to exclude alternative interpretations for our present findings. Experiment 3 showed that the tendency to gaze toward 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 generalised deployment of spatial attention toward 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 responses. The existence of multiple classes of Pavlovian responses triggered in parallel by the same stimulus is also consistent with recent evidence in humans<sup>29</sup> and with classical findings in animals<sup>13</sup>. This literature distinguishes between two classes of Pavlovian responses: “preparatory responses” that reflect the motivational properties of the outcome (e.g., heart rate) and “consummatory responses” that reflect the sensory properties of the outcome (e.g., chewing for a solid food vs. liking for a liquid food outcome<sup>45</sup>). Several studies showed how these different classes of Pavlovian responses are executed in parallel and are underlined by distinct neuronal networks<sup>29, 40</sup>. 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<sup>12, 25</sup>. We designed our experimental paradigms to obtain responses reflecting 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, involving other sensorial aspects of the outcome (e.g., sweet or savoury) and the temporal aspects of the outcome (e.g., temporal occurrence). It remains to be explored whether Pavlovian responses based on other sensorial representations of the outcome beyond spatial localisation, such as the savoury or sweet taste of the outcome, 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 not aversive. 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 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, since in Experiment 3, the tendency to gaze toward the outcome delivery location was present even when participants were instructed to look in the opposite location.

An important theoretical question raised by our findings is whether the co-existence of responses that are sensitive and responses that are insensitive to outcome devaluation within the Pavlovian system mirrors the co-existence of multiple controllers (i.e., habits and goal-directed) within 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 within the Pavlovian system<sup>8, 49</sup>. 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 – therefore they 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 – therefore they predict outcome devaluation insensitive behaviours<sup>49, 50</sup>. This proposal could account for the co-existence of parallel Pavlovian behaviours that differentially respond to changes in outcome value. Nonetheless, the typical conceptualisation of model-free RL as utilised within 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 (i.e., its spatial location). Such sensory information about an outcome cannot be learned in a model-free RL 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 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 vs model-free distinction utilised 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 which also 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, that 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 (i.e., habits)<sup>34, 37, 40</sup>. In instrumental conditioning the goal-directed and habitual influences target the same instrumental action (i.e., pressing on a button), whereas in our paradigms, multiple Pavlovian responses (i.e., 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 conceptualisation of parallel Pavlovian responses with different sensitivities to outcome devaluation could guide future 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.

## Method

### Participants

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. 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\*power51. The effect sizes of interest we focused on regarded the Pavlovian influence on pupil dilation. For Experiment 1 to 3, these effects were extracted from a previous study<sup>8</sup> and from an independent pilot study ( $n = 11$ ) using a paradigm similar to the one we used in Experiment 1 ( $d_z = .62$ ,  $d_z = .57$ ). 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 we obtained in Experiment 1 and Experiment 2 ( $d_z = .33$ ,  $d_z = .39$ ). 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, Experiment 4 was conducted at the University of Geneva, Switzerland.

For the four experiments: (a) all participants were pre-screened to ensure they were not dieting; (b) they were asked not to eat for at least 6 hours before the experimental session (but were allowed to drink water); (c) 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 (Pasadena, CA) 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; (d) 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. The reward outcome consisted of a 3s long video of 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 collected during the task to be consumed. 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 3s 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., Natick, MA, USA).

**Pupil dilation and gaze direction**—Pupil dilation and gaze direction were used to reflect two classes of Pavlovian responses. Pupil dilation upon cue presentation was used as an index reflecting a Pavlovian response based on the value representation of its associated outcome<sup>8, 30, 31</sup>. Anticipatory gaze direction was used as an index reflecting a Pavlovian response based on spatial localisation representation of its associated outcome. To obtain these measurements, an infrared camera continuously recorded a video of the participants' pupil at 30 frames per second. The eye-tracker was calibrated using a nine points 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/jmtyaska/mrgaze/>). Before statistical analysis, the pupil data were preprocessed to remove eye blinks and extreme variations. A pre-stimulus 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<sup>8, 30</sup>. 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<sup>8</sup>. The dwell time on the regions of interest (ROIs) was extracted through the EyeMMV toolbox<sup>52</sup>. The ROIs were defined as squares centred on the food outcome delivery video, but 25% bigger of than the actual video. Moreover, the index reflecting the pupil dilation was adjusted by regressing out the gaze position on the screen and the index reflecting the gaze direction was adjusted by regressing out the pupil size<sup>53, 54</sup>.

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 on the ROI.

In Experiment 3, the dwell time on 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 attended to obtain the food outcome. In contrast to Experiment 1 and 2, this experiment required the provision of an online feedback based on the participants' gaze direction, as we implemented an instrumental response contingency. In order to make this instrumental response 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 points calibration screen at the beginning of each session. Experiment 4 was conducted with the same eye-tracking methods as Experiment 3. To maintain the measures as comparable as possible with Experiment 1 and Experiment 2, we extracted the dwell time on the ROIs through the EyeMMV toolbox<sup>52</sup>.

## Statistical Analyses

All statistical analysis was conducted using the RStudio software 1.0.36 with R 3.4.3 (2009-2016 RStudio, Inc). 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 Experiment 1 and 2 we ran three planned contrasts analyses according to our a priori hypothesis. The first compared the dwell time on 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 on 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 contrasts analyses. The first compared the dwell time spent on the Pavlovian ROI after the perception of the congruent cue (weight contrast +1) to the incongruent cue (weight contrast -1), the second compared the dwell time spent in the instrumental ROI after the perception of 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 partial eta squared ( $\eta_p^2$ ) for the repeated measures ANOVA and planned contrasts and as Cohen's *d* (*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 participants underwent an outcome devaluation procedure (i.e., satiation group) while the other half of participants was asked to wait without performing any particular task (i.e., control group). Finally, all participants performed a test session under extinction.

*Snack selection:* Participants were presented with a selection of individual pieces of 6 snacks divided into two categories: sweet (M&M's<sup>®</sup>, Buncha Crunch Candy<sup>®</sup>, Almonds covered in cacao) and savoury (roasted cashews, roasted peanuts, Goldfish<sup>®</sup>). 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 minutes each. Each session was composed of 54 trials leading to a total of 162 trials. At the beginning of each trial four squares (6° visual angles 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.

On 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 only 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 Figure 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 Figure 1B). Specifically, one cue predicted the delivery of a specific outcome 70% of the time (e.g., outcome to the left); the remaining 30% of the time it was followed by one of two other possible outcomes (e.g., 15% outcome to the right and 15% no outcome; see Table 1).

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

Participants were instructed to focus on the cue and to try to predict what is going to happen next. They were also instructed to move their eyes freely around the computer screen, unless a fixation cross was present (i.e., during the inter-trial interval; ITI), in that 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 earned during the task to be consumed.

*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 it no longer palatable. The 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 minutes break. The allocation of participants to the groups was sequential: the first half was attributed 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 (e.g., 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 sweet favourite 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<sup>®</sup>, Cereal covered in chocolate, Raisins, Yogurt covered raisins, milk chocolate morsels<sup>®</sup>) and savoury (roasted cashews, roasted peanuts, Goldfish<sup>®</sup>, simply balanced popcorn<sup>®</sup>, cheese-flavored crackers, Ritz Bits cheese crackers<sup>®</sup>, Potato stick, Pretzel sticks). They were asked to taste each sample and to choose their favourite savoury snack and their favourite sweet snack. The participant's 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 minutes 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).

On 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 only 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 the sweet food outcome delivery on the left side of the screen (CS+ sweet L); one cue was more often associated with the sweet food outcome delivery on the right side of the screen (CS+ sweet R); one cue was more often associated with the savoury food outcome delivery on the left side of the screen (CS+ savoury L); one cue was more often associated with the 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 (e.g., sweet food outcome to the left), the remaining 30% of the time the cue was followed by one of the other three possible outcomes (e.g., 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 (i.e., during the ITI), in that case they were asked to look at the fixation cross.

The order of the trial presentation was pseudo-randomised 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 outcome they predicted more frequently (e.g., savoury food to the left for the CS+ savoury left). The assignment of the neutral images to particular Pavlovian cue conditions (e.g., 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 be consumed.

*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<sup>55</sup>. 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 use a strategy to prevent extinction from occurring<sup>56</sup>. 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 delivery would be as they had been during the previous sessions. They were also asked to press on the keys to guess which one of the two black patches was obscuring the outcome delivery video. The reason for using this strategy is to allow measuring 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 (e.g., disappearance of the conditioned responses due to the lack of reinforcement) from happening too quickly<sup>56</sup>.

**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 participants underwent an outcome devaluation procedure (i.e., satiation group) while the other half of participants was asked to wait without performing any particular task (i.e., control group). Finally, all the participants performed a test session under extinction.

*Snack selection:* The snack selection was identical as in Experiment 2.

*Pavlovian-Instrumental Conflict task:* Participants learned associations between different cue stimuli, two gaze actions (i.e., looking on the right side or on the left side of the screen) and the delivery of their favourite food outcome. Unlike in Experiment 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 duration of the whole trial.

On 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 the left side of the screen based on the instrumental contingency associated with the cue they just saw. If they looked on 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 Figure 5A). If they looked on 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 Figure 5A). Critically, for some cues, participants had to look in the same location as the one where the outcome delivery video was going to appear (i.e., congruent trials) to obtain the food outcome. For other cues participants had to look in the opposite direction as the one where the outcome delivery video was going to

appear (i.e., incongruent trial). As illustrated in Figure 5B, one cue was more often associated with the food outcome delivery on the left side of the screen and required participants to look on the left side to obtain the food outcome (congruent cue L); one cue was more often associated with the outcome delivery on the left side of the screen and required participants to look on 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 the food outcome delivery on the right side of the screen and required participants to look on the right side to obtain the food outcome (congruent cue R); one cue was more often associated with the outcome delivery on the right side of the screen and required participants to look on 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, in each cue carried both instrumental (i.e., gaze action to the left or right) and Pavlovian (i.e., food outcome delivery displayed on the left or right) information. The instrumental contingencies (i.e., cue-action) were probabilistic: 70% of the time a particular action (e.g., look left) after the perception of a particular cue (e.g., 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 (i.e., 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-randomised 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 outcome they predicted more frequently (e.g., food to the left for the congruent L or incongruent L). The assignment of the neutral images to particular cue conditions (e.g., 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 Experiment 1.

*Extinction session:* The test session was composed of 60 trials, identical to the previous sessions except we used the same strategy as 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 in two categories: sweet (M&M's<sup>®</sup>, Maltesers<sup>®</sup>, Almonds covered in dark

chocolate, Skittles<sup>®</sup>, coconut covered in dark chocolate, Raisins) and savoury (roasted cashews, roasted peanuts, Goldfish<sup>®</sup>, organic salted popcorn, Ritz cracker<sup>®</sup>, Pretzel stick). They were asked to taste each sample and to choose their absolute favourite snack. The favourite snack participant's was used as 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.

On 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 only 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 the food outcome delivery on the left side of the screen (CS+ L); one cue was more often associated with the 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 (e.g., food outcome to the left), the remaining 30% of the time the cue was followed by one of the other three possible outcomes (e.g., 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 (i.e., during the ITI), in that case they were asked to look at the fixation cross.

The order of the trial presentation was pseudo-randomised 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 outcome they predicted more frequently (e.g., food to the left for the CS+ L). The assignment of the neutral images to particular Pavlovian cue conditions (e.g., CS+ L, CS control R) was counterbalanced across participants.

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

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by NIDA-NIH R01 grant (1R01DA040011-01A1) to J.P.O.D. 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 Omar D. Perez and Vanessa Sennwald for insightful comments on this manuscript.

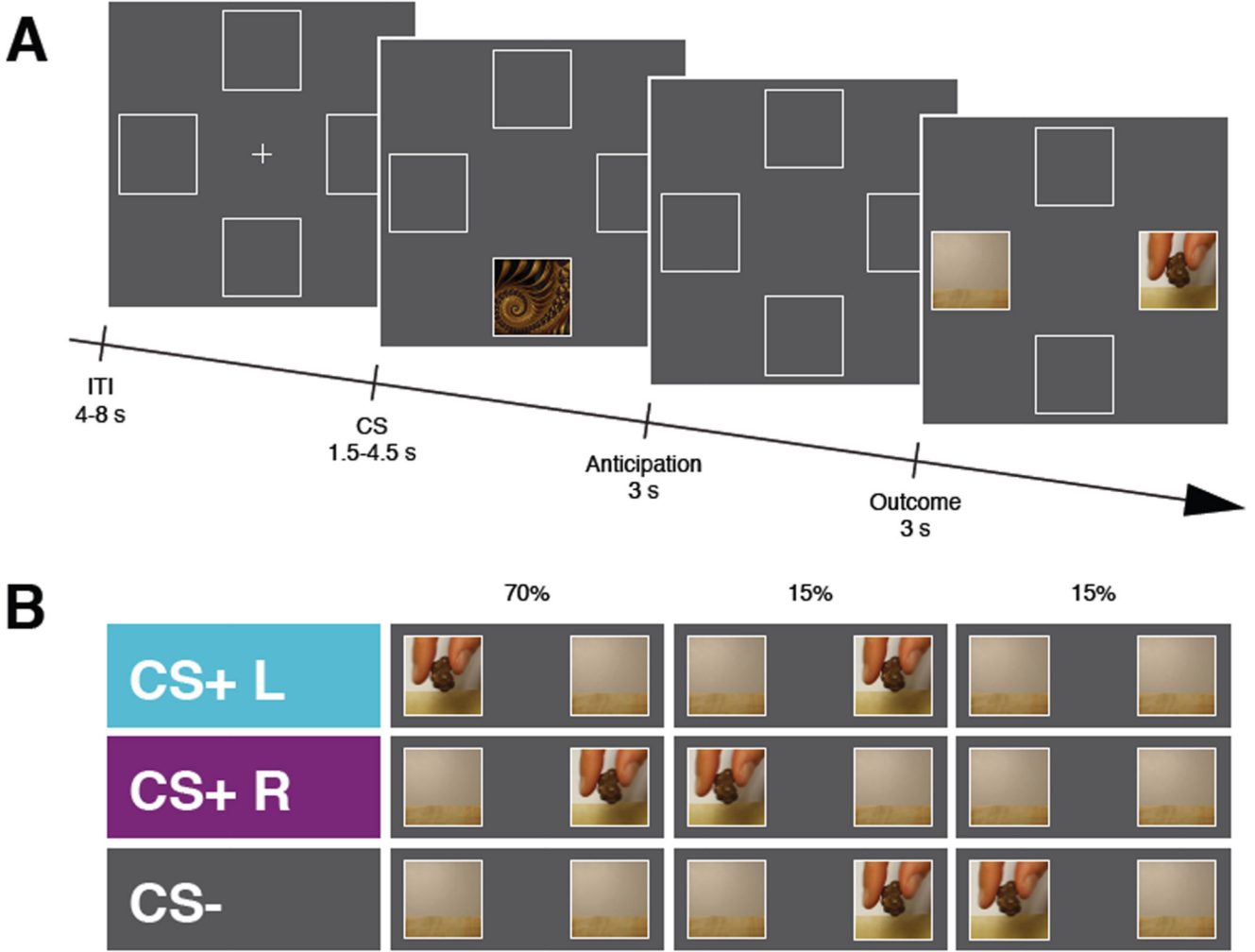
## References

1. Berridge KC, Robinson TE. Liking, wanting, and the incentive-sensitization theory of addiction. *American Psychologist*. 2016; 71:670. [PubMed: 27977239]
2. Everitt BJ, Robbins TW. Drug addiction: updating actions to habits to compulsions ten years on. *Annual Review of Psychology*. 2016; 67:23–50.
3. Voon V, Derbyshire K, Rück C, et al. Disorders of compulsivity: a common bias towards learning habits. *Molecular psychiatry*. 2015; 20:345. [PubMed: 24840709]
4. Anderson BA. The attention habit: how reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*. 2016; 1369:24–39. [PubMed: 26595376]
5. Eder AB, Dignath D. Cue-elicited food seeking is eliminated with aversive outcomes following outcome devaluation. *The Quarterly Journal of Experimental Psychology*. 2016; 69:574–88. [PubMed: 26089111]
6. Nadler N, Delgado MR, Delamater AR. Pavlovian to instrumental transfer of control in a human learning task. *Emotion*. 2011; 11:1112–23. [PubMed: 21534664]
7. Pool ER, Brosch T, Delplanque S, Sander D. Where is the chocolate? Rapid spatial orienting toward stimuli associated with primary rewards. *Cognition*. 2014; 130:348–59. [PubMed: 24389311]
8. Prévost C, McNamee D, Jessup RK, Bossaerts P, O'Doherty JP. Evidence for model-based computations in the human amygdala during Pavlovian conditioning. *PLOS Computational Biology*. 2013; 9
9. Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*. 2008; 9:545. [PubMed: 18545266]
10. Sali AW, Anderson BA, Yantis S. The role of reward prediction in the control of attention. *Journal of Experimental Psychology: Human Perception and Performance*. 2014; 40:1654–64. [PubMed: 24955700]
11. O'Doherty JP, Cockburn J, Pauli WM. Learning, reward, and decision making. *Annual Review of Psychology*. 2017; 68:73–100.
12. Delamater AR, Oakeshott S. Learning about multiple attributes of reward in Pavlovian conditioning. *Annals of the New York Academy of Sciences*. 2007; 1104:1–20. [PubMed: 17344542]
13. Balleine BW, Killcross S. Parallel incentive processing: an integrated view of amygdala function. *Trends Neuroscience*. 2006; 29:272–9.
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. *The Journal of Neuroscience*. 1996; 16:5256–65. [PubMed: 8756453]
15. Holland PC, Straub JJ. Differential effects of two ways of devaluing the unconditioned stimulus after Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1979; 5:65–78. [PubMed: 528879]
16. Ramachandran R, Pearce JM. Pavlovian analysis of interactions between hunger and thirst. *Journal of Experimental Psychology: Animal Behavior Processes*. 1987; 13:182–92. [PubMed: 3572308]
17. Gottfried JA, O'Doherty JP, Dolan RJ. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*. 2003; 301:1104–7. [PubMed: 12934011]
18. Holland PC, Lasseter H, Agarwal I. Amount of training and cue-evoked taste-reactivity responding in reinforcer devaluation. *Journal of Experimental Psychology Animal Behavior Processes*. 2008; 34:119–32. [PubMed: 18248119]
19. Robinson MJF, Berridge KC. Instant transformation of learned repulsion into motivational “wanting”. *Current Biology*. 2013; 23:282–9. [PubMed: 23375893]
20. Pearce JM, Hall G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*. 1980; 87:532. [PubMed: 7443916]
21. Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*. 1972; 2:64–99.

22. Sutton RS. Learning to predict by the methods of temporal differences. *Machine Learning*. 1988; 3:9–44.
23. Sharpe MJ, Schoenbaum G. Evaluation of the hypothesis that phasic dopamine constitutes a cached-value signal. *Neurobiology of Learning and Memory*. 2017
24. Cardinal RN, Parkinson JA, Hall J, Everitt BJ. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*. 2002; 26:321–52. [PubMed: 12034134]
25. Delamater AR. On the nature of CS and US representations in Pavlovian learning. *Learning & Behavior*. 2012; 40:1–23. [PubMed: 21786019]
26. Nasser HM, Chen Y-W, Fiscella K, Calu DJ. Individual variability in behavioral flexibility predicts sign-tracking tendency. *Frontiers in Behavioral Neuroscience*. 2015; 9:289. [PubMed: 26578917]
27. Ahrens AM, Singer BF, Fitzpatrick CJ, Morrow JD, Robinson TE. Rats that sign-track are resistant to Pavlovian but not instrumental extinction. *Behavioural Brain Research*. 2016; 296:418–30. [PubMed: 26235331]
28. Morrison SE, Bamkole MA, Nicola SM. Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Frontiers in Neuroscience*. 2015; 9:468. [PubMed: 26733783]
29. Zhang S, Mano H, Ganesh G, Robbins T, Seymour B. Dissociable learning processes underlie human pain conditioning. *Current Biology*. 2016; 26:52–8. [PubMed: 26711494]
30. Pauli WM, Larsen T, Collette S, Tyszka JM, Seymour B, O'Doherty JP. Distinct contributions of ventromedial and dorsolateral subregions of the human substantia nigra to appetitive and aversive learning. *The Journal of Neuroscience*. 2015; 35:14220–33. [PubMed: 26490862]
31. Seymour B, Daw ND, Dayan P, Singer T, Dolan RJ. Differential encoding of losses and gains in the human striatum. *The Journal of Neuroscience*. 2007; 27:4826. [PubMed: 17475790]
32. Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*. 2009; 29:2225–32. [PubMed: 19490086]
33. Dickinson A, Campos J, Varga ZI, Balleine B. Bidirectional Instrumental Conditioning. *The Quarterly Journal of Experimental Psychology Section B*. 1996; 49:289–306.
34. Grindley GC. The formation of a simple habit in guinea-pigs. *British Journal of Psychology General Section*. 2011; 23:127–47.
35. Hershberger WA. An approach through the looking-glass. *Animal Learning & Behavior*. 1986; 14:443–51.
36. Guitart-Masip M, Fuentemilla L, Bach DR, et al. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *The Journal of Neuroscience*. 2011; 31:7867–75. [PubMed: 21613500]
37. Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B*. 1982; 34:77–98.
38. Balleine BW, Dickinson A. Instrumental performance following reinforcer devaluation depends upon incentive learning. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*. 1991; 43:279–96.
39. Valentin VV, Dickinson A, O'Doherty JP. Determining the Neural Substrates of Goal-Directed Learning in the Human Brain. *The Journal of Neuroscience*. 2007; 27:4019–26. [PubMed: 17428979]
40. Balleine BW, O'Doherty JP. Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010; 35:48–69. [PubMed: 19776734]
41. Pauli WM, Cockburn J, Pool ER, Pérez OD, O'Doherty JP. Computational approaches to habits in a model-free world. *Current Opinion in Behavioral Sciences*. 2018; 20:104–9.
42. Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. *Animal Learning & Behavior*. 1995; 23:197–206.
43. Guitart-Masip M, Duzel E, Dolan RJ, Dayan P. Action versus valence in decision making. *Trends in Cognitive Sciences*. 2014; 18:194–202. [PubMed: 24581556]

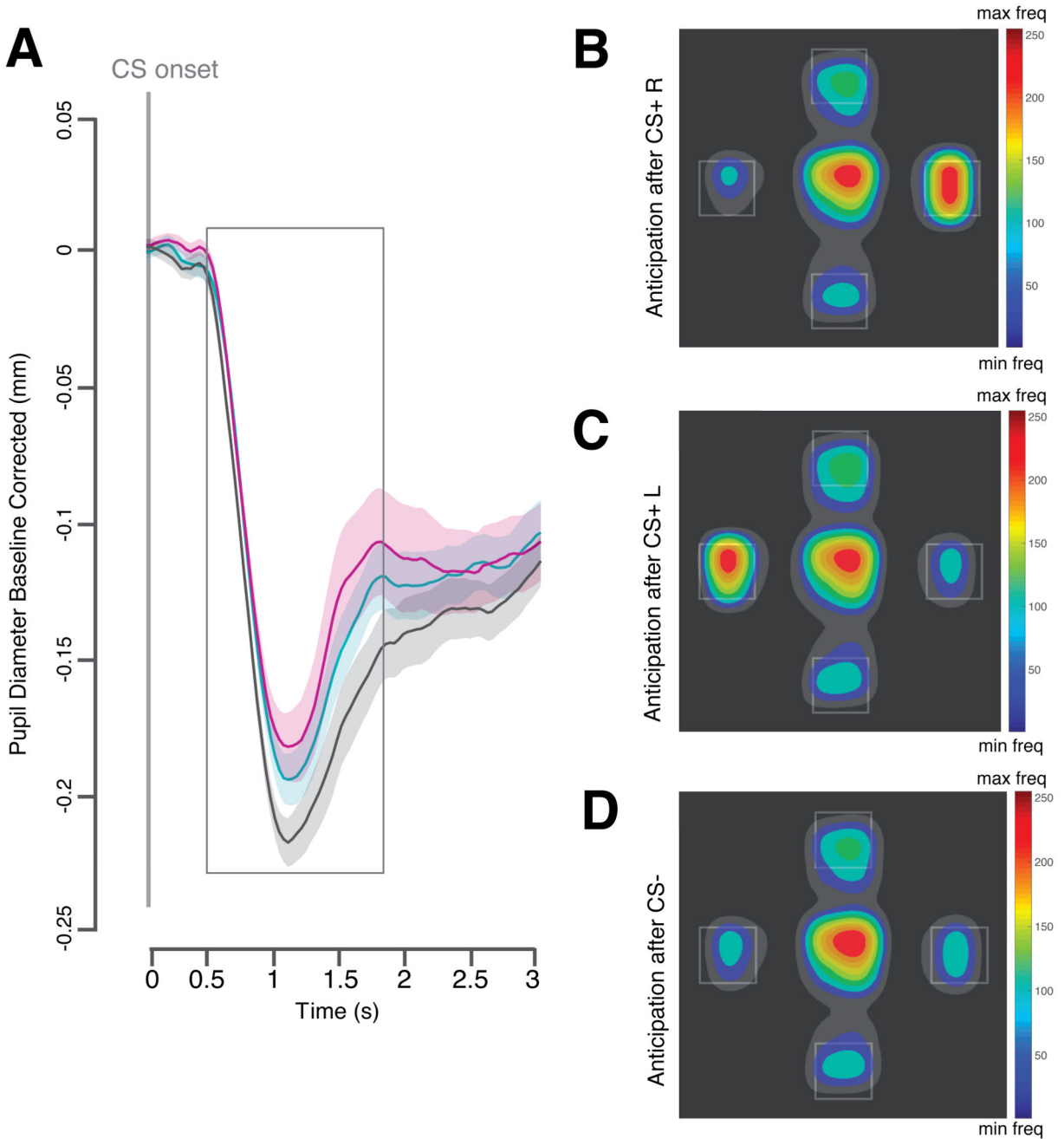
44. Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*. 2012; 62:154–66. [PubMed: 22548809]
45. Konorski, J. *Integrative activity of the brain: An interdisciplinary approach*. Chicago: University of Chicago; 1967.
46. De Tommaso M, Mastropasqua T, Turatto M. Working for beverages without being thirsty: Human Pavlovian-instrumental transfer despite outcome devaluation. *Learning and Motivation*. 2018; 63:37–48.
47. Holland PC. The effects of satiation after first—and second-order appetitive conditioning in rats. *The Pavlovian Journal of Biological Science: Official Journal of the Pavlovian*. 1981; 16:18–24.
48. Holland PC, Rescorla RA. The effect of two ways of devaluing the unconditioned stimulus after first-and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1:355. [PubMed: 1202141]
49. Dayan P, Berridge KC. Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*. 2014; 14:473–92.
50. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*. 2005; 8:1704. [PubMed: 16286932]
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. *Behavior Research Methods*. 2007; 39:175–91. [PubMed: 17695343]
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. *Journal of Eye Movement Research*. 2014; 7(1)
53. Choe KW, Blake R, Lee S-H. Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation. *Vision research*. 2016; 118:48–59. [PubMed: 25578924]
54. Nyström M, Hooge I, Andersson R. Pupil size influences the eye-tracker signal during saccades. *Vision Research*. 2016; 121:95–103. [PubMed: 26940030]
55. Reber J, Feinstein JS, O'Doherty JP, Liljeholm M, Adolphs R, Tranel D. Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain*. 2017; 140:1743–56. [PubMed: 28549132]
56. Prévost C, Liljeholm M, Tyszka JM, O'Doherty JP. Neural correlates of specific and general Pavlovian-to-instrumental transfer within human amygdalar subregions: A high-resolution fMRI study. *The Journal of Neuroscience*. 2012; 32:8383–90. [PubMed: 22699918]





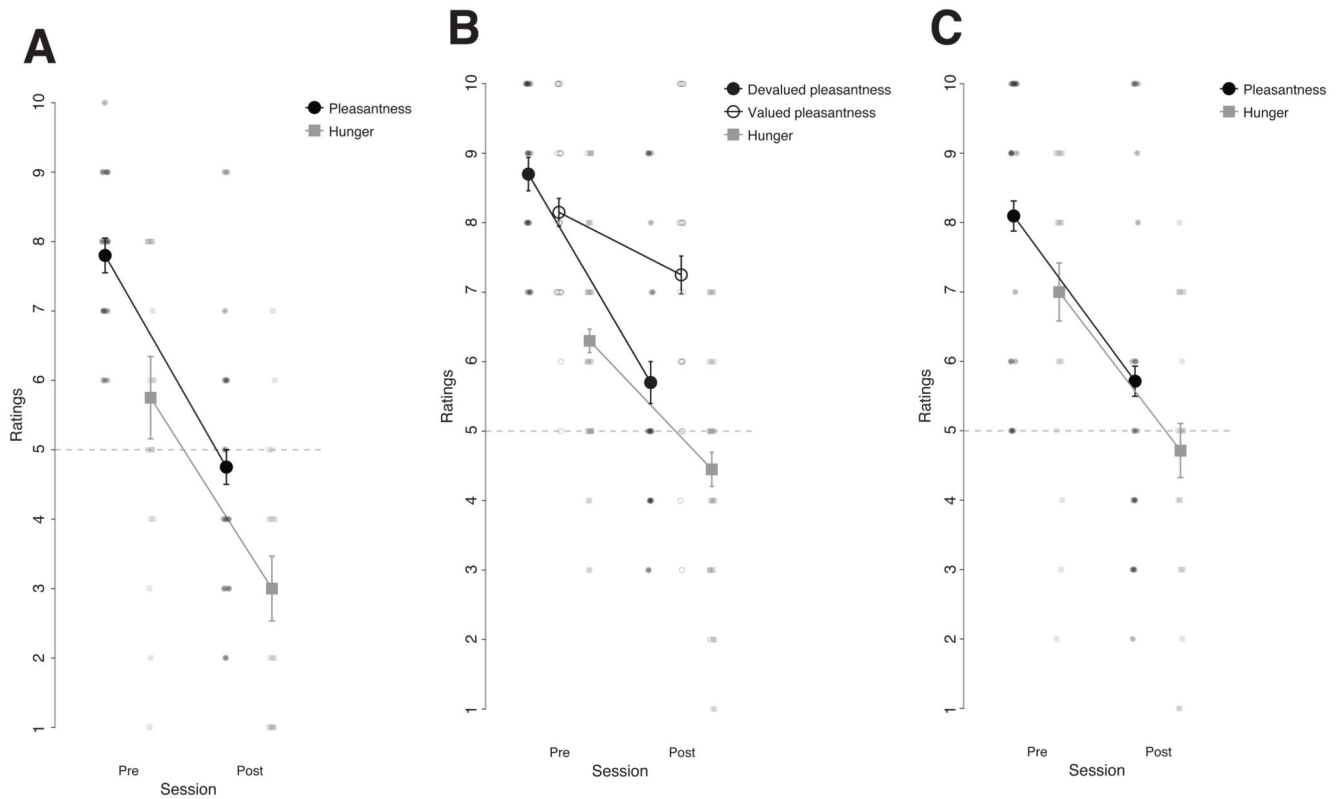
**Figure 1. Schematic representation of the experimental design.**

**A**, Illustration of the sequence of events within a trial for Experiment 1, Experiment 2 and Experiment 4. At the beginning of each trial a Conditioned Stimulus (CS) was presented randomly in the upper or the 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 3s. The inter-trial interval (ITI) lasted for 4 – 8 s (uniformly distributed). At the end of each session participants received the actual snacks delivered during the task to be consumed. **B**, In Experiment 1 trial involved one of three CS, each of which primarily predicted (70%) either the snack delivery to the left (CS+ L), the snack delivery to the right (CS+ R), or no snack delivery (CS-).



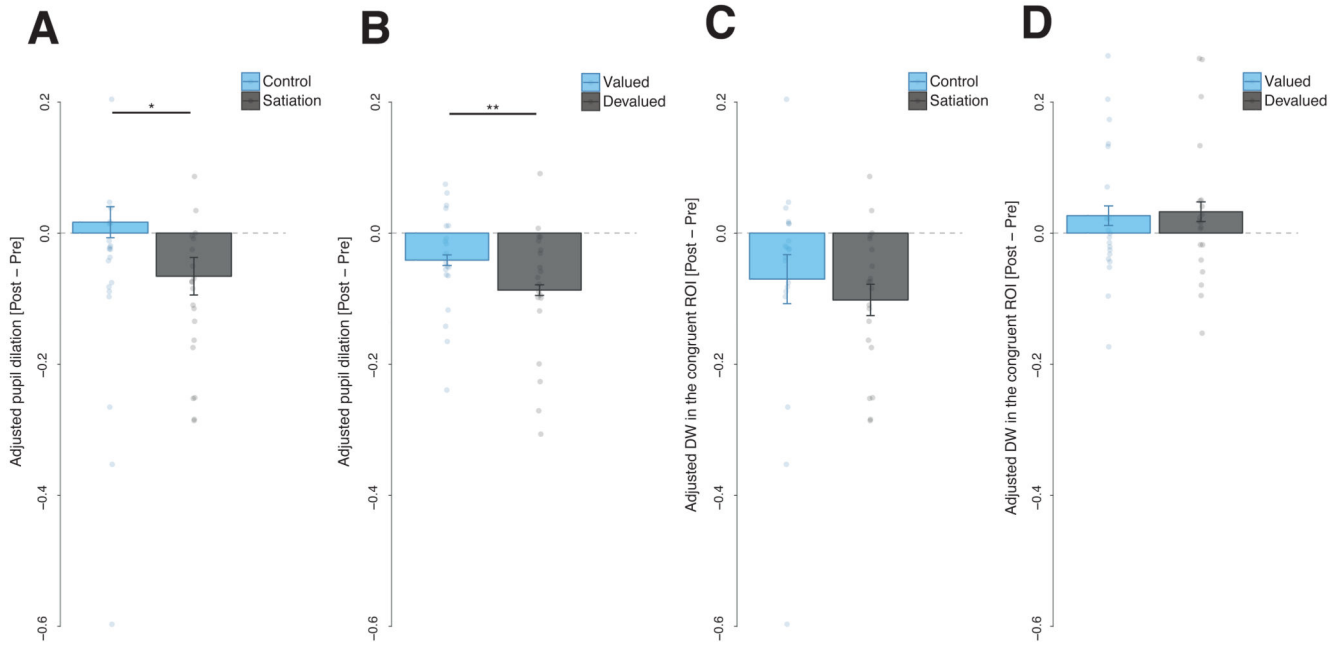
**Figure 2. Effect of conditioning during the learning phase of Experiment 1.**

**A**, Plot of the averaged pupil response over time aligned to the onset of the conditioned stimulus (CS) and plotted separately for the CS that could predict 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; freq): after the offset of the CS+R (**B**) the CS+ L (**C**) and the CS- (**D**). Shaded areas indicate the within subject standard error of the mean. All plots are based on data from 40 participants.



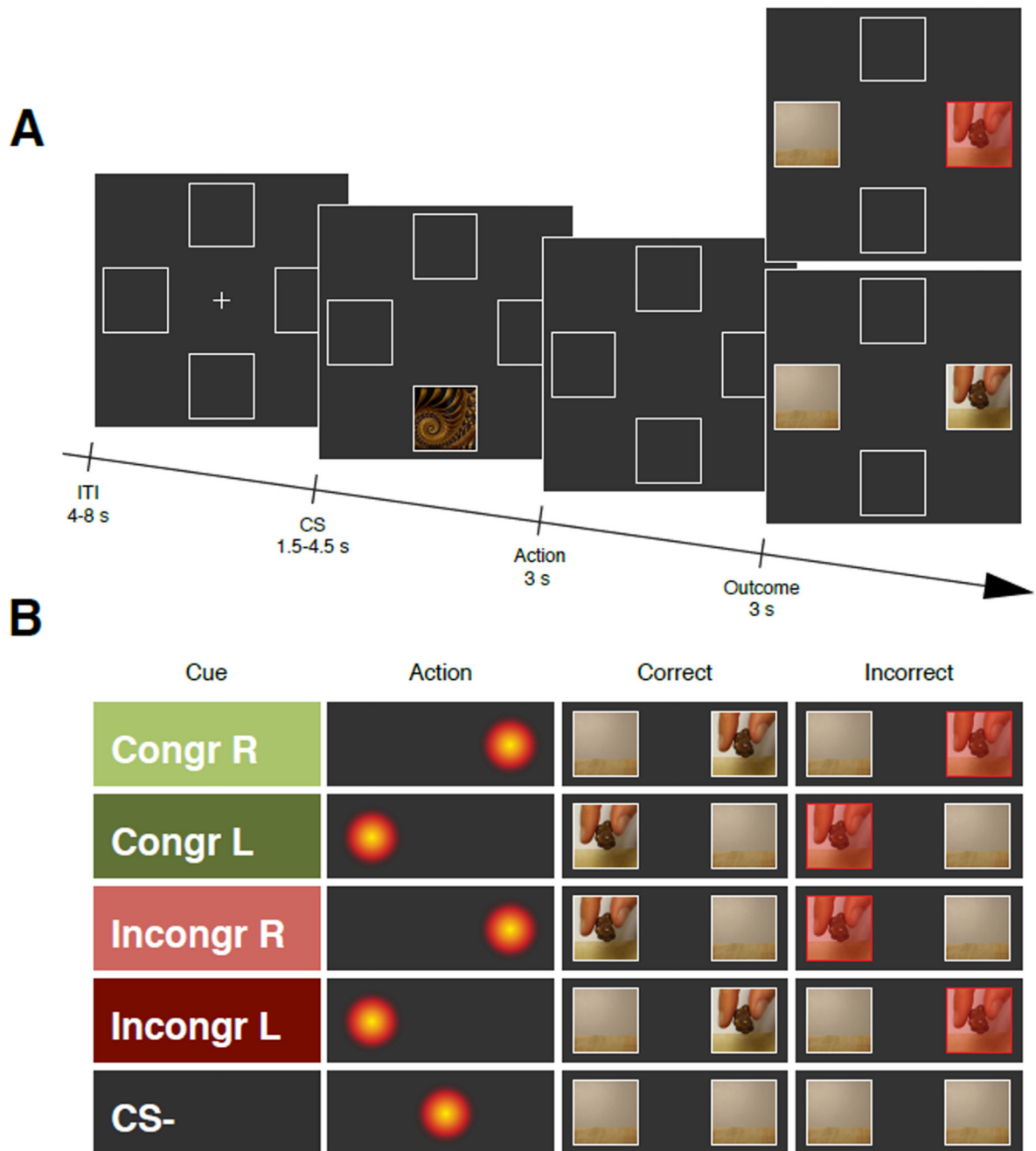
**Figure 3. Manipulation check of the outcome devaluation procedure.**

**A, C,** Mean ratings of hunger and the pleasantness of the snack before and after selective satiation for the group that underwent to 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 (i.e., Devalued) and the snack that was not (i.e., Valued) in Experiment 2. Error bars indicate the within subject standard error of the mean. Plots of Experiments 1 to 2 are based on two different sets of 20 participants. Plot from Experiment 3 is based on data from 21 participants.



**Figure 4. Effects of the outcome devaluation procedure on different conditioned responses during Experiment 1 and Experiment 2.**

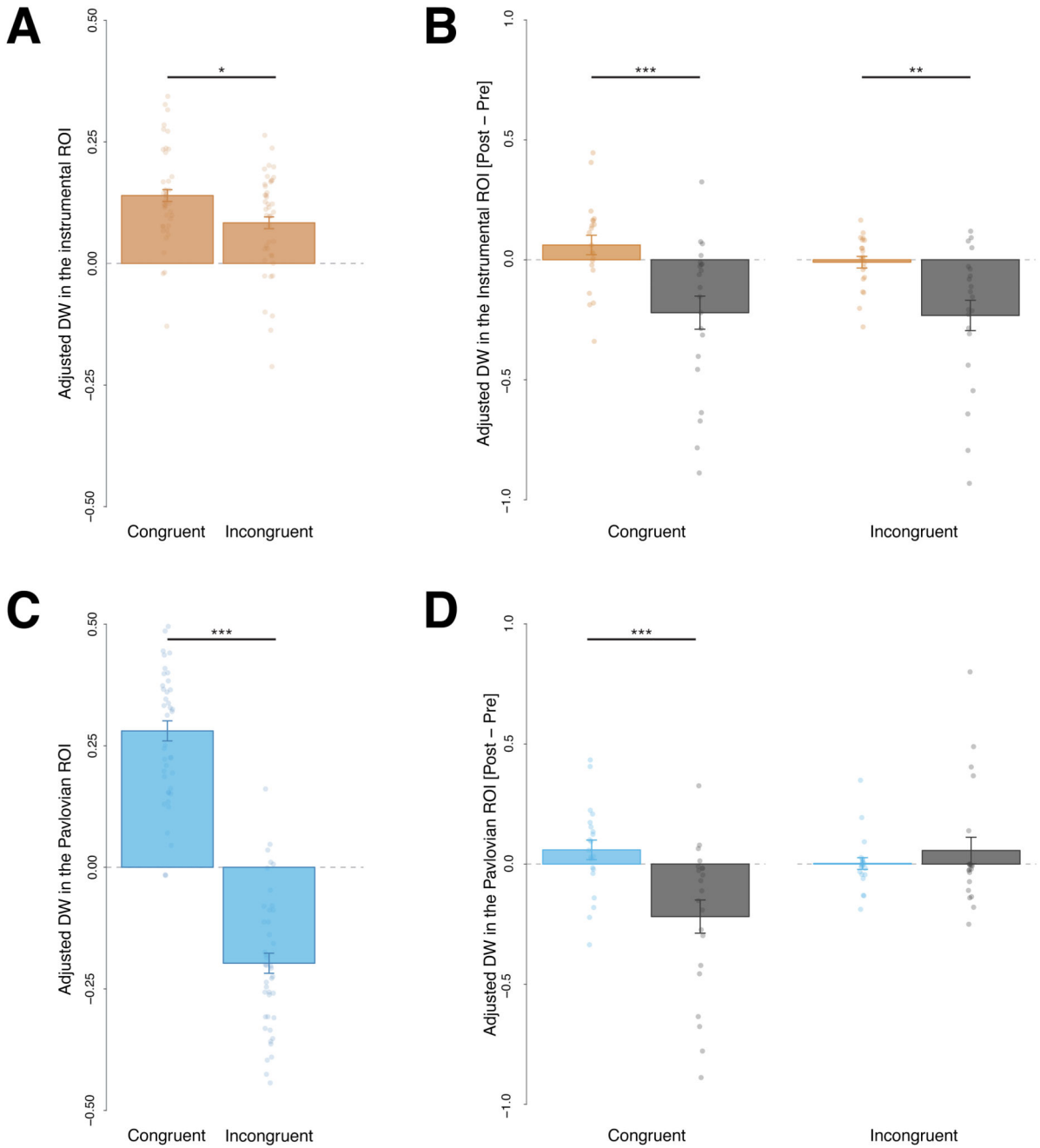
**A-B,** Adjusted pupil dilation before and after the outcome devaluation procedure: in (A) Experiment 1,  $*F(1,38) = 4.93, p = .032, \eta^2_p = .115, 90\% \text{ CI } [.005, .276]$  and (B) Experiment 2,  $**F(1,19) = 8.08, p = .010, \eta^2_p = .298, 90\% \text{ CI } [.045, .504]$ . Bars represent 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 (DW) spent in the Region of Interest (ROI) congruent with the CS+ prediction (i.e., left ROI for the CS+ predicting the outcome delivery to the left and right ROI for the CS+ predicting the 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 in the after satiation – adjusted dwell time before satiation). Results depicted from Experiment 1 are shown separately for the CS+ 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 subjects standard error of the mean; plots from Experiment 2 are based on 20 participants and error bars indicate the within subject standard errors of the mean.



**Figure 5. Illustration of the sequence of events within a trial for Experiment 3.**

**A**, at the beginning of the trial a cue was presented randomly in the upper or the lower portion of the screen for 1.5 – 4.5 s (uniformly distributed). Based on this cue participants were asked to look either to the left or the right side of the screen during 3 s to win a piece of their favourite snack. Then a video of the snack delivery was displayed either on the right or on the left side of the screen. If the participants looked on the correct side of the screen during the action screen the video was normally displayed, whereas if the participants looked on 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 to be consumed. **B**, Each trial involved three kinds of cues: (1) the congruent cues (either left or right: Congr L and Congr R) where participants had to look on the same side where the video was going to be displayed to obtain the snack (2) the incongruent cues (either left or right: Incongr L and Incongr R) where participants had to look on the opposite side as where the video was going to be displayed to obtain the snack and (3) the CS- cue where participants were not required to do any specific action.

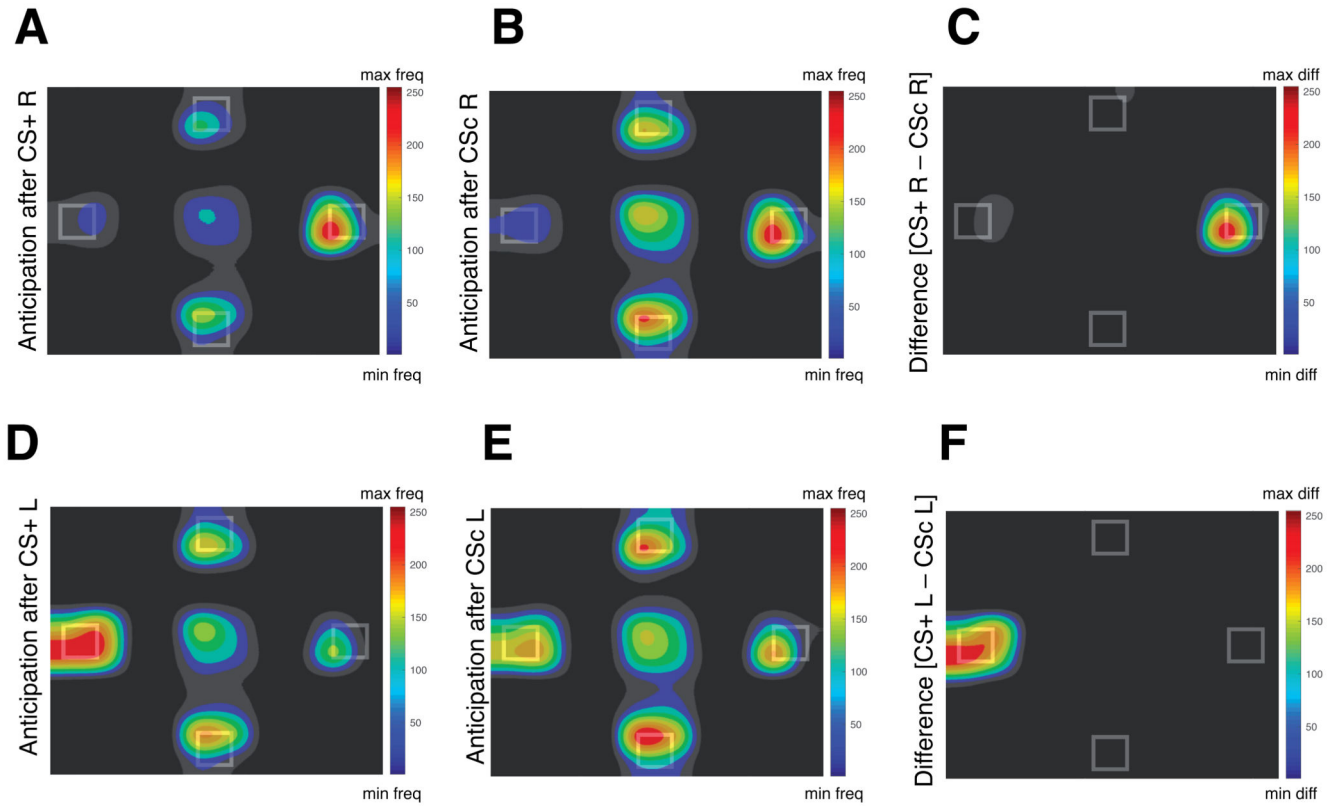


**Figure 6. Illustration of the main effects during Experiment 3.**

**A, C,** Adjusted Dwell Time (DW) spent in **(A)** the instrumental Region of interests (ROI)  $*F(1,41) = 5.41, p = .025, \eta^2_p = .117, 90\% \text{ CI } [.008, .272]$  and **(C)** the Pavlovian ROI  $***F(1,41) = 133.61, p < .001, \eta^2_p = .765, 90\% \text{ CI } [.646, .824]$  during the action screen after the perception of a congruent or incongruent cue. **B, D,** Influence of the outcome devaluation procedure on the dwell time in these ROI for the satiation group that underwent the devaluation procedure and 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); \*\*\* $F(1,40) = 12.02, p = .001, \eta^2_p = .231$ , 90% CI [.063, .393], \*\* $F(1,40) = 10.75, p = .002, \eta^2_p = .212$ , 90% CI [.051, .374]. Error bar represents the between subject standard error of the mean, all plots are based on data from 42 participants.





**Figure 7. Effect of conditioning during Experiment 4.**

*A, B, D, E*, Heatmaps of the fixation patterns during the anticipation screen (normalized frequency; freq) after the offset of the conditioned stimulus (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 of the screen (CSc R; *B*); the video of a hand delivering a snack to the left side of the screen (CS+ L; *D*); the video of an empty hand to the left side of the screen (CSc L; *E*). *C, F*, Heatmaps of the normalized difference (diff) between the fixation pattern during the anticipation screen after the offset of the CS+ R and the offset of the CSc R (*C*) and after the offset of the CS + L and the offset of CSc L (*F*).

**Table 1**  
**Summary of the Pavlovian Contingencies Across the Four Experiments**

		Outcome 1		Outcome 2		No Snack
		Left	Right	Left	Right	
Exp. 1 & 3	CS+ L	70 %	15 %	0 %	0 %	15 %
	CS+ R	15 %	70 %	0 %	0 %	15 %
	CS -	15 %	15 %	0 %	0 %	70 %
Exp. 2 & 4	CS 1+ L	70 %	10 %	10 %	0 %	10 %
	CS 1+ R	10 %	70 %	0 %	10 %	10 %
	CS 2+ L	10 %	0 %	70 %	10 %	10 %
	CS 2+ R	0 %	10 %	10 %	70 %	10 %
	CS -	10 %	10 %	10 %	0 %	70 %
	CS -	0 %	10 %	10 %	10 %	70 %

*Note.* CS+ L = Positively conditioned stimulus left; CS+ R = Positively conditioned stimulus right; CS- = Negatively conditioned stimulus; Exp. = Experiment. 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 empty experimenter's hand.