Devaluation insensitivity of event related potentials associated with food cues

Thomas D. Sambrook [a,*] , Andy J. Wills [b], Ben Hardwick [b], Jeremy Goslin [b]

[a] School of Psychology, University of East Anglia, United Kingdom
[b] School of Psychology, University of Plymouth, United Kingdom

ABSTRACT

Eating in the absence of hunger represents a failure of homeostatic mechanisms responsible for energy balance and is a cause of obesity. The pervasive presence of food cues in the modern environment may play a role in this phenomenon. The present study used the technique of satiety-specific selective devaluation to investigate eating in the absence of hunger in the context of a reinforcement learning task. While participants' performance on the task suggested that food on which they had sated no longer held value for them, event related potentials following images of the food were unaffected by the devaluation. Food cues may thus serve as an entry point for over-eating in otherwise healthy individuals.

1. Introduction

The present experiment investigates the role that food cues might play in overeating. It seeks to establish whether a neural signature of reinforcement learning that is elicited by food cues continues to be elicited even after participants have sated on the food depicted. Such an effect of devaluation insensitivity might implicate learned responses to cues as a cause of eating in the absence of hunger.

Eating is determined by a number of factors including an individual's energy balance, the sensation of hunger, knowledge about the palatability of potential foods, estimates of effort-to-reward ratio for obtaining food, underlying genes, an individual's eating history and numerous cultural factors. The underlying homeostatic mechanism that governs eating and eating cessation is well understood. Secretion of ghrelin occurs when blood sugar levels are low producing a sensation of hunger that motivates an animal to eat, while rising blood sugar levels cause the secretion of leptin which acts on reward systems via the hypothalamus to reduce neural responses to food stimuli. Evidence that this mechanism does not reliably produce eating cessation comes from the rise in obesity in recent years. While a decline in physical activity has certainly contributed to this, epidemiological studies suggest that energy intake is the strongest determinant of BMI variance among European adults and furthermore that overeating is a learned rather than inherited behaviour (Herle et al., 2017).

Food cues may play a role in over-eating. Obese children recognise more food related advertisements than their non-obese counterparts and overweight children overeat after exposure to food cues (Halford et al.,

2004). Images of high calorie food sensitise eating: the neutral taste produced by electrical stimulation of the tongue is deemed pleasanter when paired with food images and this affects reward-related neural activity (Ohla et al., 2012). Images of high calorie food affect event related potentials (ERPs) even when they are incidental aspects of the task (Toepel et al., 2009) and ERPs that follow food cues are affected in binge eating disorder (Forester et al., 2024) as are those accompanying the inhibitory responses associated with forgoing food (Harris et al., 2013).

From the point of view of learning theory, food cues serve as conditioned stimuli which, by association with the food represented, come to elicit feeding behaviour. This occurs both in humans (Cornell et al., 1989) and animals (Boggiano et al., 2009; Petrovich et al., 2007). Food images provoke a response in the ventral striatum, an area associated with instrumental learning (Tang et al., 2012) and are as effective as the food itself in this regard (Boswell & Kober, 2016). Importantly, cues can elicit feeding in individuals who are sated (Volkow et al., 2011) suggesting they play a role in eating in the absence of hunger. However, the routes through which food cues affect feeding behaviour are complex. Broadly, a tendency to eat in response to cues in the environment has been described as "external eating" (Schachte et al., 1968). Such external eating begins with the capture of attention by food cues, an effect shown in experiments using event related potentials, eye-tracking and behavioural measures. (Heldman et al., 2025). Thereafter, food cues may affect multiple decision-making processes. They can act as either Pavlovian or instrumental cues, can interact via the phenomenon of Pavlovian instrumental transfer (Yin et al., 2008) and induce cognitively

---

elaborated craving (Kavanagh et al., 2005).

Establishing the cognitive processes between cue receipt and subsequent behaviour therefore requires carefully designed experiments. One method is the outcome devaluation technique (Adams & Dickinson, 1981) in which a previously valued good such as food is devalued, for example by allowing an individual to feed on it to satiety. The technique was developed to attempt to distinguish between goal directed and habitual control (Balleine et al., 2008). Individuals engaged in goal-directed behaviour pursue actions that are likely to achieve motivationally relevant goals such as food when they are hungry. Food ceases to be motivationally relevant for a sated individual and so when action is goal-directed food-related behaviour that is targeted at reducing hunger should terminate. However, in both humans and other animals it is suspected that there may exist a second basis for action selection that is blind to the likely outcome's current desirability: a habit-led system. This acts simply on the basis of prior reinforcement. Hungry individuals who eat in response to food cues derive reward and this strengthens the association between food cue and eating. Given a sufficiently strong association this could in principle lead to automatic eating in the presence of the food cue even in sated individuals. Because the homeostatic regulation of eating is built on the motivational value that food holds at the point of action selection it would be circumvented by such a habit-led system. Such insensitivity to devaluation is commonly used as a hallmark of habitual control in rodent studies though in humans the validity of the method remains a subject of debate.

The present study used the ERP technique to investigate the devaluation sensitivity of neural responses to food cues in the first 700 ms after presentation. Food cues were provided in the context of a probabilistic reinforcement learning task and served as rewards insofar as they increased the chance of being able to eat the food depicted later. Here we investigated whether the reward-related ERPs associated with food cues reduced after the participant had sated on the food depicted or remained unaffected by the devaluation. To the extent that the neural response showed devaluation insensitivity a cue-led basis for eating in the absence of hunger would be supported.

2. Methods

2.1. Participants

Ninety students at the University of Plymouth took part in the experiment. Participants were excluded if they were dieting, had an eating disorder, had experienced a neurological disorder, were on medication, were over twenty-nine or had a BMI outside the range of 18.5–45. The age restriction was included since components associated with cognitive control are known to be affected by aging from early adulthood onwards (Kropotov et al., 2016). Dieters were excluded to rule out the possibility that food cues would constitute ambivalent stimuli for those striving to restrict eating. During data pre-processing fourteen participants were excluded due to excessive artefacts in the EEG. Participants received course credit and the opportunity to eat desirable foods. The study was approved by the local ethics committee.

2.2. Materials

2.2.1. Foods

A pool of eleven foods were available for the experiment (see Supplementary Fig. 1), two of which were used for each participant's session. Participants rated all foods for desirability on a scale from 1 to 10 and two foods with very similar ratings were selected by the experimenter, one savoury and one sweet in order to maximise the possibilities for sensory-specific devaluation. Foods were chosen whose ratings lay in the range 5–8 since these were expected to produce satiety for the food consumed without actually filling up the participant. Each of the foods was privately allocated to one of the two levels of the *food value category*

independent variable, either valued or devalued. The food allocated to the devalued category was the food offered as a meal half-way through the experiment. For this reason, the effects of the food value category independent variable were expected to occur only after the meal. Across participants food allocation was made such as to ensure counterbalancing of food value category (valued/devalued) over the various actual foods on offer.

2.2.2. Two-step task

The opportunity to eat the foods was earned on the basis of performance in the two-step task (Daw et al., 2011; Gillan et al., 2015). This task was used since it attempts to provide a behavioural assay of model-based vs. model-free learning while also presenting at the conclusion of each trial a reward or reward omission stimulus that can be expected to generate a model-free reward prediction error. The structure of the task is shown in Fig. 1a. On each trial of the task participants chose between two first step fractal stimuli. This led them to an intermediate step in which they saw one of two different fractals. The transition probability between the fractal chosen at the first step and that presented at the intermediate step was fixed throughout the experiment with each first step fractal tending to lead to one of the intermediate step fractals. Participants were required to observe the intermediate fractal for at least a second before making a confirmatory keypress and were then presented with the food outcome: a stimulus depicting either success (a picture of the food on offer in that trial) or failure (a picture of an empty cupboard or plate). Each success contributed a point towards the food on offer on that trial. Importantly, prior to the beginning of each block participants were briefed on which first step fractal tended to lead to which intermediate fractal and that this relationship was unchanging over the experiment. This information did not comprise part of the model-free learning aspect of the task but instead gave participants an explanatory model (albeit a probabilistic one) of events in the first part of each trial. In contrast, the probabilities relating the intermediate step fractal and the food outcome were unknown to the participants and were allowed to drift over the experiment. The participant therefore had to employ simple model-free reinforcement learning in order to track the relative profitability of the two intermediate fractals. Since they could not directly select the intermediate fractal they considered to be more profitable their best option was to employ model-based knowledge of the transition probabilities from first step to intermediate step in order to increase their likelihood of accessing the desired intermediate fractal.

The two-step task claims to be able to distinguish between model-free and model-based learners by using participants' choice behaviour at the first step to infer the values they assign to the fractals shown there. Importantly these values will come to differ depending on whether participants apply model-based learning at the first step or simply treat the whole task as a problem in model-free learning. This is because on a significant minority of occasions participants will experience an unexpected transition between first and intermediate step fractals. When this unexpected transition happens, model-based learners do not revise their estimate of the transition probabilities because these have been designated as immutable and indeed the participants have, over the course of the experiment, opportunity to witness the truth of this. Model-based learners do engage in model-free learning, but only where they have no choice, in the contingency between intermediate fractal and food cue that concludes the trial. In contrast, model-free learners are model-free for the entire task and simply update the value of the fractal chosen at the first step based on the final outcome without compensating for unexpected transitions. Importantly, outcomes that are reached via unexpected transitions should be used to update the value of the *unchosen* first step fractal, but model-free learners will update the chosen fractal. Whether participants are learning in a model-based or model-free way can be inferred by comparing the fit of participants' choices to a computational model simulating both model-based and model-free learning processes.
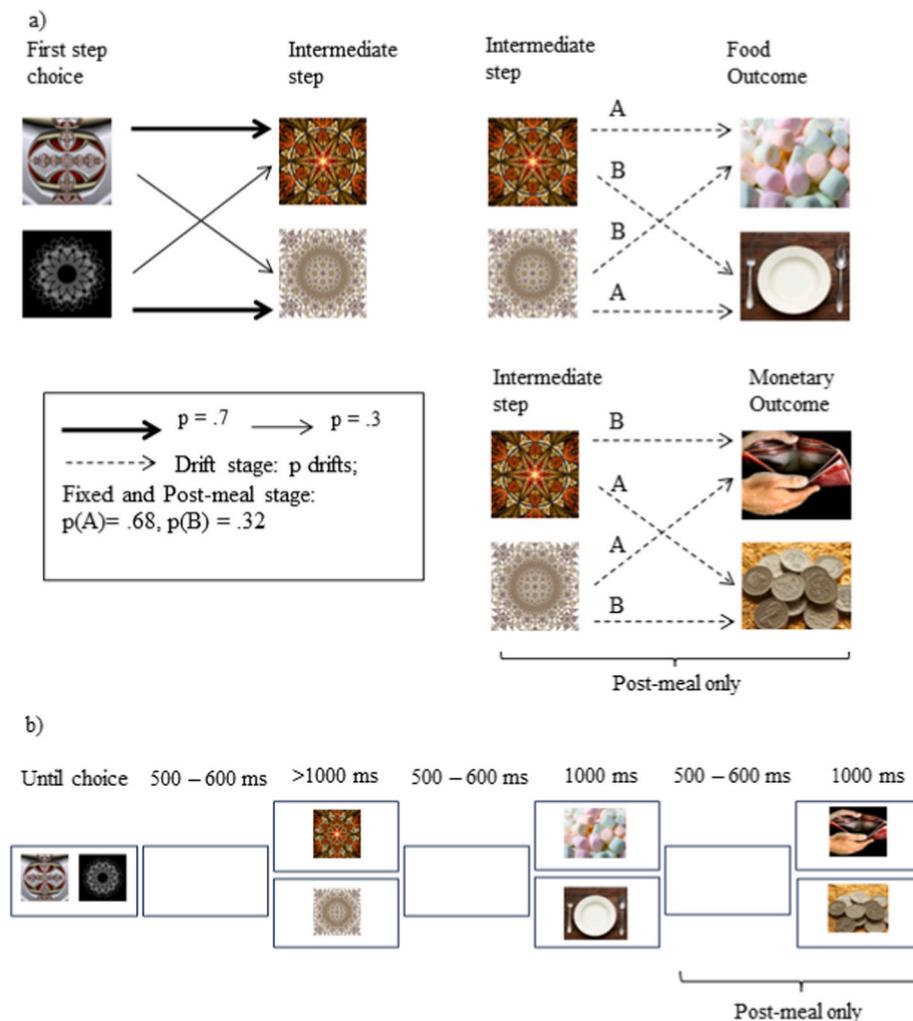
Fig. 1. a) Underlying structure of the two-step task. b) Sequence of events seen by participant on a single trial; paired boxes are alternatives.

2.3. Procedures

2.3.1. Devaluation procedure

On arrival at the laboratory hungry participants were weighed and measured for height. They rated their current hunger and the desirability of the eleven foods on a visual analogue scale and on this basis two foods of similar rating were selected. Participants were fitted with an elastic cap and EEG was recorded using 61 Ag/AgCl active electrodes (actiCAP, Brain Products, Gilching, Germany), arranged in a standard International 10–20 montage and with a left mastoid reference. Participants then engaged in the two-step task for approximately 180 trials in an attempt win the opportunity to access each of the foods for a later meal. At the end of this, the pre-meal part of the task, participants rated their hunger and the desirability of the two foods once again. They were then given a meal of the food assigned to the devalued food category. This was served in a series of small helpings with participants asked to continue requesting more helpings until they no longer wished to eat the food. Participants rated their hunger and the desirability of the two foods a third time and resumed the task for approximately a further sixty trials, once again attempting to win access to both foods but here, additionally, a sum of cash. They were then given a second meal of the devalued food, serving as a consumption test, which concluded the experiment.

2.3.2. Presentation of the two-step task

The two-step task was implemented largely as described by Gillan

et al. (2015) with some modifications to accommodate the devaluation procedure described above. Trials were arranged in blocks of twenty trials in which just one of the two foods was on offer. Each trial culminated with a food or no food image that represented either success or failure in winning a point towards the opportunity to later eat that food. At the end of each block participants were awarded a "lottery ticket" for the food if they had earned a point on more than ten of the twenty trials. Just before the mid-experiment meal an automated lottery was run to determine which of the two foods participants would have access to. Participants were told that the probability of winning access was proportional to the lottery tickets acquired and that draws for the foods were independent so they might have the opportunity to eat two, one or neither of the foods. In fact, the outcome was predetermined so that the food already privately allocated to the devalued food category was the only food accessed.

The two-step task was implemented in slightly different ways over different stages of its presentation. Before beginning the experiment proper, participants performed a truncated version of the task ending at the intermediate step. The purpose of this was to allow them to learn which first step fractal was associated with which intermediate step fractal. Participants repeated blocks until they achieved 100% accuracy in a test examining their understanding. Prior to instruction participants were explicitly shown the transition probabilities and could review these after each failed test. The instruction period served both to check participants' knowledge and to consolidate it in the context of actually performing the two-step task.

Following Gillan et al. (2015) two stages were used prior to the meal. In the *drift* stage probabilities relating intermediate fractals to food outcome were initialised at hidden values between 0.25 and 0.75. These probabilities were then allowed to drift independently of each other within these starting boundaries following a Gaussian random walk with standard deviation of 0.025. On commencing the following *fixed* stage, whichever intermediate fractal of a pair currently had the higher probability of leading to the food outcome was given a new probability fixed to 0.68 and the other fixed to 0.32. This was not announced to the participants. The purpose of the fixed stage was to establish in participants a clear preference for one intermediate fractal (and therefore one first step fractal) in order to assess the consequent effect of devaluation on behaviour.

In the *post-meal* stage, the probabilities relating intermediate fractals to food outcomes remained at the same fixed values. Participants resumed the task, attempting to win lottery tickets towards a second meal of both foods. However, participants also had an opportunity to win money. The task was modified so that after the food outcome was displayed, a monetary outcome was shown indicating whether the participant had won points towards a cash prize. These points would be converted to lottery tickets for a cash prize on the same basis as the food lottery tickets. Participants were told that the intermediate fractals were now associated with two different probabilities, one reflecting the chance of winning a food point and one the chance of winning a cash point. Participants were told that they should attempt to track the profitability of the intermediate fractals on both accounts to make the best gains overall. In fact, these probabilities were not independent. The intermediate fractal that had a 0.68 chance to win a food point had a 0.32 chance to win a cash point and vice versa for the other fractal. The purpose of this was to sharpen the behavioural measure of food devaluation. As it stood, while participants had no particular reason to persist with a key choice that led to a food they had sated on, nor did they have a reason to switch. Money was introduced to provide that incentive. The monetary prize on offer was calibrated on an individual basis by asking participants how much they would hypothetically pay for unlimited access to the still-valued food. By contriving parity of valued food and monetary rewards in this way, participants had at least a partial incentive to persist with key presses appropriate for the valued food but an unmitigated incentive to switch keys for the devalued food. Persistence with the keypress would thus represent strong evidence for devaluation insensitivity.

The number of trials participants underwent varied for several reasons: trials were self-paced, participants varied greatly in the length of time they took to learn transition probabilities before starting the task and participants varied in the time they spent eating the half time meal. The trial means, standard deviations and ranges were as follows: drift stage M = 119.19, SD = 13.01, range = 60–140, fixed stage M = 58.95, SD = 7.93, range = 20–80, post-devaluation stage M = 74.47, SD = 12.90, range = 20–100). Blocks contained twenty trials and the number of blocks was always the same for valued and devalued foods.

The experimental task was presented using E-Prime 2.0. Eight fractals were used from http://www.fractalsciencekit.com. Two pairs of fractals were used for the first step choice for the two foods and another two pairs of fractals for the intermediate step for each food. Fractals were randomly assigned for each participant at the start of the experiment. Success on a trial was represented by a realistic image of the food being played for, failure by an image of an empty plate or empty cupboard. One of each failure stimulus was assigned to valued and devalued foods and this was counterbalanced across food value category over participants. Stimuli are shown in Supplementary Fig. 1. Fig. 1b shows the sequence of events on a trial.

2.4. Parameter fitting

As described above, the two-step task endeavours to generate different choice behaviour to the degree to which participants engage in

model-based vs. model-free learning. A computational model based on that employed by Gillan et al. (2015) and described in detail in Sambrook et al. (2018) was fitted to participants' choice behaviour in order to generate a single value, omega, representing the degree to which learning was model-based. The heart of the model is a SARSA(λ) model-free temporal difference learning algorithm (Sutton & Barto, 1998) which assumes that each of the two first step fractals is associated with an expected value (ranging from 0 to 1) that is continually updated on the basis of the reward (either 0 or 1) incurred at the end of the trial on which that fractal is selected. The reward prediction error (reward – expected value) is used to make this update, adjusted by a learning rate parameter α. The translation of these expected values into choice behaviour is modelled by a softmax choice rule in which the higher valued fractal is more likely to be chosen than the lower but with some stochasticity in this process represented by another parameter, the inverse temperature β.

This part of the model assumes that participants are engaging in a purely model-free learning process, associating first step fractals with rewards and ignoring the fixed transition probabilities to intermediate fractals. In parallel, a second model-based learning process is simulated. This uses the model-free learning process described above to assign expected values to the intermediate step fractals rather than first step fractals. Since the likelihood with which each first step fractal reaches each intermediate step fractal is known to the participant, the proper, model-based estimate of first step fractal value is determined by the fixed likelihood of it reaching each of the intermediate step fractals multiplied by their likelihood of reaching reward. This will return different expected values for the first step fractals than the model-free learning process.

Success of either part of the computational model can be quantified by the degree to which the fractal chosen by a participant on each trial corresponds with the fractal holding the higher value under that part of the model, with this choice expressed via the softmax choice rule. In practice it can be assumed that participants will show contributions of both model-free and model-based learning to their behaviour and so the fitting procedure takes the independently derived model-free and model-based estimates of value for each first step fractal and averages them. Because the contribution of the two processes will differ over participants this averaging was subject to a weighting term, omega, also fitted from choice data. This could vary from 1 in which only the model-based process contributed, to 0 where only the model-free process contributed. Fitting was performed on a participant-wise basis and was achieved by iterating the model through different candidate values of omega, learning rate and inverse temperature and selecting those values that returned the smallest deviation between observed and predicted behaviour. Modelling was implemented in R using the slpMBMF function of the *catlearn* package (Wills et al., 2023, see also Wills et al., 2017) and using the L-BFGS-B method (Byrd et al., 1995) of the optim function (R Core Team, 2017).

Because the fitting process proceeds trial by trial the basis by which the final omega value is generated is necessarily rather opaque for any individual participant. The success of modelling was checked by comparing omega estimates to a simpler index of model-based learning. Following Gillan et al. (2015), logistic regression was used to predict each participant's stay vs. switch behaviour on a given trial using the fixed effects of the previous trial's reward and its reward × transition interaction. Model-free learning should be affected only by the reward. Model-based learning should be affected only by the reward × transition interaction insofar as the reward reinforces different first step fractals depending on the transition. The beta value for the reward x transition effect was divided by that for the reward x transition effect plus reward effect to establish an index of model-based learning, running from 0 to 1, that could be compared to omega. Additionally, the presence of both model-free and model-based learning at the level of entire group was investigated by means of a mixed effects model with fixed effects of reward and reward x transition and participant as a random effect.

2.5. Preprocessing of EEG data

Data pre-processing was conducted in Brain Vision Analyzer software. Data was re-referenced to the average of the mastoids. Notch filters at 60 and 50 Hz and a bandpass filter of 0.1–30 Hz were applied. The EEG was epoched from 200 ms before to 700 ms after the food outcome, baseline corrected from −200 to 0 ms and downsampled to 125 Hz. Segments containing ocular artefacts (voltage change exceeding 75 μv per 200 ms at sites Rvb, Rc, Fp1 and Fp2) were discarded as were those containing non-specific artefacts at any point on the scalp (voltage greater than ±200μv compared to baseline or voltage change greater than 50 μv per ms). Topographic interpolation was used to replace channels showing sustained artefacts. Since the key measure in the experiment was the difference between food and no food outcomes, participants were discarded if artefact rejection resulted in the difference wave for any of the cells in the design being built from fewer than ten food or ten no food trials. The mean number of trials from which this difference wave was built were: pre-meal devalued food M = 137.45 (29.67); pre-meal valued food M = 134.97 (29.79); post-meal devalued food M = 53.91(13.35); post-meal valued food M = 54.16 (13.69).

2.6. Analysis plan

2.6.1. Behavioural and subjective devaluation sensitivity

Behavioural devaluation sensitivity is typically measured in terms of the likelihood of a participant persisting in the instrumental response (keypress) when the stimulus (first step fractal) is presented despite the reinforcer (food image) responsible for that S-R association having been devalued. Given that a two-action forced choice task was used here rather than a single key-press, following Friedel et al. (2014) food valuation was measured not in terms of the likelihood of a keypress but in terms of task accuracy: the proportion of trials on which the key chosen was the better for obtaining a food. Behavioural devaluation sensitivity therefore consists in this accuracy dropping after sating on a food. Behavioural devaluation sensitivity was assessed by means of the interaction of food value category (valued, devalued) and meal stage (drift, fixed, post-meal) with respect to task accuracy. Devaluation sensitivity would be indicated by an effect of food value category in the post-meal stage that was absent in the pre-meal stages and this was tested by means of an ANOVA and accompanying t-tests. Drift stage behavioural data were accidentally deleted for two participants reducing the sample size for behavioural devaluation analysis to 74.

For subjective devaluation sensitivity the food valuation was operationalised as food desirability ratings. Pre-meal desirability ratings were averaged into a single measure and an ANOVA run on food value category (valued food, devalued food) x stage (pre-meal, post-meal).

2.6.2. Testing for neural devaluation sensitivity using ANOVA and t-tests

Food valuation was operationalised as the voltage difference between food and non-food outcomes. This difference was calculated across the full epoch. A pre-defined area of interest was the RewP, operationalised here as the mean voltage difference in the interval 240–340 ms at FCz. Neural devaluation sensitivity was assessed, both in the RewP and the wider scalp activity by similar means as those used for behavioural and subjective measures above, that is, by the interaction of food value category (valued, devalued) and meal stage (pre-meal, post-meal) with respect to food valuation. Devaluation sensitivity would be indicated by an effect of food value category in the post-meal stage that was absent in the pre-meal stage and this was tested by means of an ANOVA and accompanying t-tests.

2.6.3. Bayesian t-tests for neural devaluation sensitivity vs insensitivity

Further tests benefited from expressing devaluation sensitivity as a single term. Following Friedel et al. (2014) this was achieved with the expression (pre-meal valued food - post-meal valued food) - (pre-meal devalued food - post-meal devalued food). Positive scores indicate devaluation sensitivity and scores close to zero indicate insensitivity. Since these are both of theoretical interest but evidence for devaluation insensitivity consists in demonstrating a null effect, Bayes factors are an appropriate measure here since they provide a straightforward assessment of the relative evidence for the null and alternative hypotheses. In the case of t-tests, Bayes factors can be calculated directly from t-values if the sample size is known (Francis & Jakicic, 2023). This conversion assumes a JZS prior and a scaling parameter which was set to $\sqrt{2}/2$ in the current paper. Bayes factors are conventionally thresholded at 3 and 1/3 (moderate evidence against or for the null) and 10 and 1/10 (strong evidence against or for the null). When converting t-values to Bayes factors, evidence in favour of the null is strongly limited by the sample size however. Given the sample size of 76 in the present study, evidence in favour of the null cannot be shown at a strength beyond 1/7. For the present paper Bayes factors for the null, i.e., devaluation insensitivity, are therefore thresholded at 1/7 and 1/3, and for the alternative, of devaluation sensitivity, at 3 and 7. These Bayes factors correspond to t values of 0.51, 1.43, 2.62 (p = .01) and 2.96 (p = .004). Conversions were performed using the ttest.tstat() function of the BayesFactor package (Morey & Rouder, 2024).

2.6.4. Correlating neural devaluation sensitivity with model-based learning

An individual's devaluation sensitivity might be expected to be associated with the degree to which they engaged in model-based learning insofar as both are related to goal-directed control of behaviour. Searching for locations in which devaluation sensitivity correlated with omega over participants might provide a more sensitive analysis capable of revealing areas in which devaluation sensitivity occurs for those experiencing it. Therefore, participants' omega value was correlated with neural devaluation sensitivity at each temporospatial location in an attempt to find these areas.

2.6.5. Testing for neural devaluation sensitivity with a linear mixed model

To further improve power, linear mixed models were performed in which the actual foods on offer were included as a random effect. In the present experiment, individual foods were counterbalanced with respect to whether they were allocated to devalued or valued food categories so the properties of individual foods could not serve as a confound. However, because the images of the foods varied greatly in physical properties such as luminance and contrast, neural activity associated with this might hide effects of devaluation sensitivity. A linear mixed model was constructed to test both for the presence of devaluation sensitivity and its association with omega. Fixed effects in this model were food value category (valued, devalued) and omega (continuous, centred), while the eleven possible foods were treated as a random variable fitted with a random intercept and the dependent variable was pre-meal vs. post-meal difference in food valuation. Devaluation sensitivity would be shown by the food value category fixed effect, and the effect of omega on devaluation sensitivity would be shown by the food value category × omega interaction.

2.6.6. Correlations between devaluation sensitivity measures and model-based learning

To permit correlations, both behavioural and subjective devaluation sensitivity were expressed as a single term by the same means described for neural devaluation sensitivity: (pre-meal valued food - post-meal valued food) - (pre-meal devalued food - post-meal devalued food). The values entered into this expression were task accuracy for behavioural devaluation sensitivity and desirability rating for subjective devaluation sensitivity. Correlations between each of the three measure of devaluation sensitivity and omega could then be performed.

2.6.7. Power analysis of EEG data

Probabilistic learning tasks that use the ERP method have focussed on a component, the reward positivity (RewP) referred to sometimes as feedback related negativity (Holroyd & Coles, 2002; Proudfit, 2015).

This consists in the voltage difference between two outcomes of different value, for example reward and reward omission. The interval and sites at which the RewP is measured differ but a meta-analysis by Sambrook and Goslin (2015) established 240–340 ms at FCz as a typical interval and that is used here. The component is robustly elicited. Power calculations based on this meta-analysis have suggested eighteen participants is sufficient to achieve 99% power (Stewardson & Sambrook, 2021). This is for the elicitation of the component, i.e., the reward – omission voltage difference wave, rather than treatment effect sizes on that component (here the devaluation procedure). However, similar food cue related designs to this study (Huvermann et al., 2021; Peterburs et al., 2019) have cited 30 participants as sufficient power.

3. Results

3.1. Parameter fits

Mean and standard deviations for parameters were as follows: omega 0.52 (0.31), alpha 0.53 (0.30), beta 9.66 (13.62). Distributions for these parameters are provided in Supplementary Fig. 2. The relatively uniform distribution for omega suggests that participants showed a mixture of both model-free and model-based learning. This was supported by the results of the mixed effects logistic regression on stay/switch behaviour which revealed strong effects of reward ($\beta = 2.32$, SE $= 0.15$, $z = 15.34$, $p < .0001$) and reward x transition ($\beta = 1.77$, SE $= 0.20$, $z = 8.84$, $p < .0001$). When comparing individual indices of model-based learning to omega values, the correlation was good ($r_p = 0.65$, $r_s = 0.72$, N = 76) and so the more sensitive measure of omega was used thereafter.

3.2. Behavioural and subjective devaluation

The mean accuracy of participant's choices and 95% confidence intervals were as follows: drift valued 0.53, [0.51, 0.55]; drift devalued 0.55 [0.53, 0.57]; fixed valued; 0.65 [0.60, 0.70]; fixed devalued 0.65 [0.61, 0.70]; post-meal valued 0.56 [0.52, 0.59]; post-meal devalued 0.47 [0.43, 0.51]. Participants success in choosing the better fractal for obtaining food was significantly above chance for each food in both drift and fixed stages (one sample t-tests, test value 0.50, $p < .005$). Behavioural devaluation was measured by the interaction of food value category and stage with respect to accuracy. This interaction was significant. $F_{2,146} = 6.36$, $p < .001$ $\eta_p^2 = .08$. There was no difference between valued and devalued foods in the drift stage t(73) = 0.95 p = .35, or fixed stage t (75) = 0.08, p = .94, but significantly worse performance for the devalued food in the post-meal stage t(75) = 3.65, p < .001.

Devaluation sensitivity could nevertheless be described as partial. Since in the post-devaluation stage the correct key for winning money is the opposite to that for obtaining food, participants' accuracy for the keypress associated with the devalued food should be much lower than the observed 0.47. This implies a degree of devaluation insensitivity in some participants. To rule out the possibility that this persistence with the keypress associated with the devalued food was due to that food retaining value for these participants, data from the final consumption test was used. Those participants continuing to choose the key associated with the devalued food should show no greater tendency to eat it when presented as a subsequent meal if devaluation insensitivity underlies their behaviour. In order to quantify consumption of foods of very different energy, sweetness and bulkiness on a common scale the quantity consumed by a given participant was converted to a z-score based on the consumption of that food over all participants. These z-scores served as the regressor in a linear mixed model with accuracy as the dependent variable, and the actual food as a random effect. No significant effect of consumption on accuracy was found, t(76) = −0.50 p = .621, implying devaluation insensitivity.

With regard to subjective devaluation, means and standard deviations of ratings were as follows: pre-meal valued 6.69 (1.21), pre-meal devalued 6.82 (1.36), post-meal valued 6.00 (1.98), post-meal

devalued 4.21 (1.81). A similar interaction was seen between food value category and stage: $F_{1,75} = 50.20$, p < .001, $\eta_p^2 = .401$. Valued and devalued foods were subjectively rated similarly before the meal, t (75) = −1.28, p = .203 but the devalued food was rated significantly lower than the valued afterwards, t(75) = 6.70 p < .001. In order to correlate behavioural and subjective devaluation each was expressed as a single term as described earlier. No significant correlation was found between the two devaluation measures ($r_p = −0.048$, N = 74, p = .685). Nor was either measure correlated with omega (behavioural devaluation: $r_p = .09$, N = 74 p = .444; subjective devaluation $r_p = −0.137$, N = 76, p = .239).

3.3. Neural devaluation

In order to assess the strength and extent of pre-meal food valuation, the voltage difference between food and non-food outcomes was calculated at each temporospatial point with valued/devalued food categories collapsed together. This effect was strongly significant from 240 to 340 ms at FCz, the interval used to define the RewP: t(75) = 5.64 p < .001, see Fig. 2a. As expected, it did not differ for valued and devalued foods: t(75) = 1.39, p = .170. Food valuation was widespread beyond this interval and is shown in Fig. 2b.

Neural devaluation sensitivity in the RewP was assessed by inspecting the interaction of food value category and pre-meal vs. post-meal stage. No interaction was found ($F_{1,75} = 0.003$, p = .954, $\eta_p^2 < .001$) and therefore no evidence of neural devaluation sensitivity in the RewP. When expressed as a single term in the form (pre-meal valued food - post-meal valued food) - (pre-meal devalued food - post-meal devalued food), RewP devaluation showed no correlation with omega ($r_p = .016$, N = 76, p = .888), behavioural devaluation ($r_p = −0.092$, N = 74, p = .435) or subjective devaluation ($r_p = .109$, N = 76, p = .350).

The search for neural devaluation sensitivity was then extended to the full EEG. While the RewP is often the focus of attention of reward encoding, other sites are suspected to be valence sensitive, in particular parietal areas (San Martin, 2012; Stewardson & Sambrook, 2020). For this analysis the EEG was first masked to include only food valuation encoding areas, i.e., those showing a significant difference between food and non-food outcome in the pre-meal stage. One sample t-tests were run on the single term expression of neural devaluation sensitivity at each temporospatial point and t values were converted to Bayes factors to establish the relative evidence for devaluation sensitivity vs. insensitivity. No temporospatial points were found showing devaluation sensitivity at BF > 3 and in contrast wide areas showed moderate to strong evidence in favour of devaluation insensitivity at BF < 1/7: see Fig. 2c.

An attempt was made to isolate devaluation sensitivity by correlating neural devaluation sensitivity with omega. However, this correlation was found to be non-significant at all temporospatial locations. Finally, the linear mixed model, also run to establish the presence of devaluation sensitivity or its correlation with omega showed no evidence of either effect, with t values very similar to those shown in the fixed-effects analysis above and plotted in Fig. 2c.

4. Discussion

The devaluation procedure successfully produced behavioural and subjective devaluation of the sated food. Neurally however, robust insensitivity to devaluation was found throughout the EEG. Bayes factors for this null ranged from 1/3 to 1/7 and were particularly strong in temporospatial zones traditionally associated with reward processing.

How can this neural insensitivity to devaluation best be interpreted? Multiple forms of valuation are known to be held in the brain that are tied to distinct mechanisms of learning (Balleine et al., 2008). A food image, both in the present experiment and in the real-world context can be expected to be represented in a number of different forms simultaneously. One is as a conditioned stimulus in a Pavlovian association with
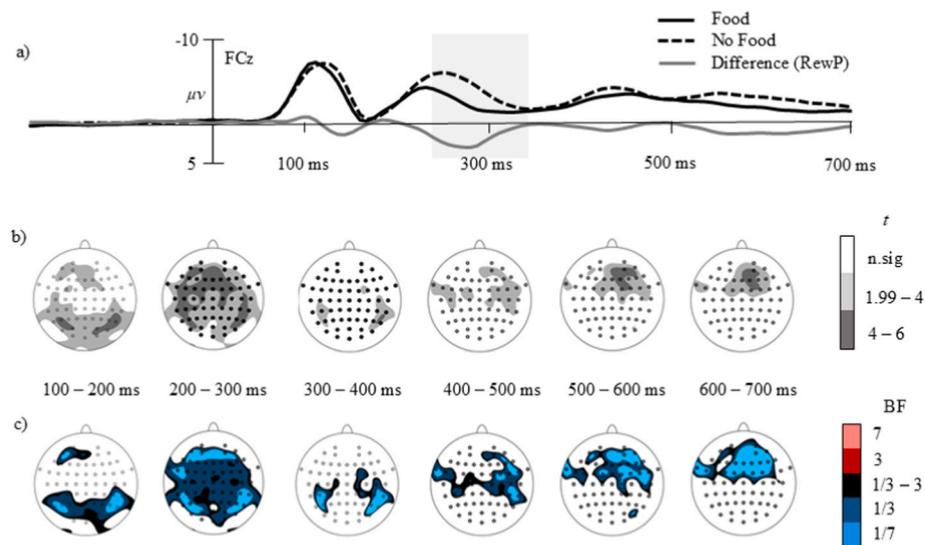
Fig. 2. a) Pre-meal RewP. b) Wider topography of food valuation (food – no food difference, one-sample t, μ = 0). c) Topography of devaluation insensitivity. White areas are outside the mask provided by 2b.

the unconditioned stimulus of the food represented (stimulus-stimulus association). In this capacity it is expected to generate anticipatory and consummatory effects in participants (e.g., sensation of hunger, increased vigour of action). A second form is as a conditioned stimulus in an instrumental model-free stimulus-response association that produces habitual behaviours which have previously resulted in food acquisition (e.g., food unwrapping where the cue is the wrapper). These two representations may interact in the form of Pavlovian-instrumental transfer which is known to be devaluation insensitive (Colagiuri & Lovibond, 2015; Hogarth, 2012; Hogarth & Chase, 2011; Watson et al., 2014). A third representation of the food image is as the outcome in an instrumental model-free learning process that reinforces the relationship between the first state fractal and a given keypress thus instilling a stimulus-response habit. It is this form of representation that has typically been invoked when interpreting the RewP. A fourth representation is as an outcome in an instrumental model-based learning process. This representation does not reinforce a stimulus-response link. Instead, it strengthens belief in the contingency between a keypress and the outcome which can be used in conjunction with knowledge about the current value of that outcome to direct future behaviour. This last representation is the basis for goal-directed behaviour and it is this representation alone that is expected to be responsive to the devaluation procedure (Yin et al., 2008).

One interpretation is that the EEG in this short interval reflects habit values. The RewP as elicited in probabilistic learning tasks such as that used here has tended to be interpreted as a valuation in the context of model-free learning (Chase et al., 2011; Cohen & Ranganath, 2007; Walsh & Anderson, 2011, 2012), with its amplitude predicting behaviour consistent with model-free learning. Neural encoding of habit values can persist even when this does not affect an individual's behavioural policy. Bayer and Glimcher (2005) showed that rats in a reversal learning task were sensitive to contingency degradation and their behaviour was best explained by model-based rather than habit learning. Nevertheless, food cues produced phasic dopaminergic responses in the midbrain that were better modelled as simple habit values suggesting they continued to be computed even if they did not have access to behaviour. A comparable pattern was shown in in the present study in which the failure of any part of the EEG to devalue in response to the sated food was at odds with participants loss of interest in the food behaviourally and subjectively.

While this dissociation is certainly of interest there are limits to how far it can be used to infer habit learning. Although participants in the post-devaluation stage were evidently not showing habit-led behaviour, in the predevaluation stage it is not possible to gauge the degree to which responses were goal-directed or habit led since they predicted the same behaviour. Thus, it is not possible to say for certain that habits or neural habit values were ever formed.

The present findings contrast with those from a study by Huvermann et al. (2021) in which neural devaluation sensitivity was shown in frontal and parietal areas in the interval 340–400 ms. This is a temporospatial interval that shows strong devaluation insensitivity in the present study. Huvermann et al. also presented food images as outcomes in a probabilistic task though this was a guessing rather than learning task ("doors task" Hajcak, Moser et al., 2006) and the outcomes were graded rather than dichotomous. Both these properties are thought to reduce the strength of reward signals (Cockburn & Holroyd, 2018; Sambrook & Goslin, 2015). There is also some evidence that valuation of a graded rather than dichotomous reward occurs later (Pedroni et al., 2011; Stewardson & Sambrook, 2023) suggesting rather different neural processes may be elicited. One methodological difference between the experiments was that rather than participants eating a meal in the course of the session they attended two different sessions, eating the meal at the start of one of them.

The validity of the devaluation method is subject to some debate, much of which centres on the dangers of inference from a null result, i.e., devaluation insensitivity. Because the method reciprocally operationalises habitual and goal directed behaviour, placing them at either end of a single measure of behaviour this means that failures of goal-directed behaviour may be wrongly attributed to the influence of habit (Buabang et al., 2021). Habit may then be wrongly inferred from shortcomings in the experiment: from weak devaluation, poor learning or a failure to devalue the outcome to which goal directed action is actually oriented (De Houwer et al., 2023; de Wit et al., 2018; Nebe et al., 2024; De Houwer et al., 2018 note that the technique assumes two things have been successfully achieved by the devaluation procedure. First, since any given event can constitute a number of different outcomes for an individual that is subject to the pursuit of multiple goals, experimenters must ensure that devaluation targets the outcome that controls behaviour. Thus, in the present case after devaluing the palatability outcome associated with the presentation of food or food cues, there must remain no other outcomes whose desirability is unaffected and which might produce goal-directed behaviour mistaken for habitual behaviour. Second, there should be a basis for showing that the devaluation is genuinely impactful if there is to be any hope of treating the null behavioural

result as evidence of habit rather than simply a weak experimental manipulation. In the present case this means ensuring that participants are genuinely sated on the devalued food. Drawing from the literature on unconscious learning these are labelled the information and sensitivity criteria respectively (Shanks & St. John, 1994). The sensitivity criterion in the present study appears to have been satisfactory insofar as, both behaviourally and subjectively, there was clear devaluation of the sated food. It likewise suggests that there was no failure to learn the task. With the information criterion it is less easy to rule out alternative explanations however. In particular, by the time participants reached the post-devaluation stage they had undergone a number of trials in which food images had constituted "correct" outcomes relative to a designated "wrong" outcome, an empty cupboard or table. Even in hungry participants pursuing the goal of winning desirable foods obtaining the "correct" outcome might have been incorporated as an additional goal in the pre-devaluation stage, carried over to the post-devaluation stage in a neural signal. As such we cannot be certain that habit developed.

The two-step task was used because it purports to measure model-based learning at the first step while still providing an entirely model-free second step that could serve as a basis for habit formation. In fact, the capacity of the two-step task to distinguish model-based and model-free learning has been criticised on a number of accounts: that apparent model-free learning reflects confusion about the task (da Silva & Hare, 2020), that the small advantage the task gives to model-based learners over model-free learners disincentivises them from using model-based learning (Akam et al., 2015; Kool et al., 2016, 2017) and that apparent model-based learning might actually reflect statistical artefacts or latent state variants of model-free learning (Akam et al., 2015). In defence of the task as implemented here, the reduced Gillan et al. (2015) version which presents no choice at the intermediate stage improves the return on model-based learning (Kool et al., 2016), the model-free latent state learning described by Akam et al. arises in artificial agents exposed to many more trials than used here (~200) and participants in the present study were trained to the point that they were able to pass a test on the transition structure, increasing our confidence that they were not confused about this aspect as described by da Silva & Hare. Nevertheless, it remains the case that no correlations were found between the two step task's measure of model-based vs. model-free learning, and behavioural, subjective or neural devaluation sensitivity. Given the uncertainty surrounding the two-step task the safest conclusion is to assume that the attempt to discriminate model-free and model-based learning was unsuccessful.

Participants were exposed to food images over a number of trials without being able to gain access to the foods represented. If participants in the post-meal stage were required to eat a piece of the devalued food each time they received a food image they would be expected to devalue the food image through a process of model-free learning. There are practical and ethical objections to this but more importantly the aim of the experiment was to see whether the food image could be neurally devalued without the requirement for negative reward prediction errors, as a goal-directed learning process should be able to do. It is also not clear that separating out extended exposure to food images and consumption is unusual in real world eating behaviour, when cues may be persistent, occur some way from the actual food and may be internally maintained in the form of craving.

A potential limitation of the experiment is the absence of any aversive outcome. While a voltage difference between reward and omission is often interpreted as reward encoding it may in fact represent an event's motivational salience (Bromberg-Martin et al., 2010; Talmi et al., 2013) and indeed this effect has been shown in the context of food (Hird et al., 2018; Springer et al., 2021). Aversive outcomes, the reward value of which decreases as their motivational salience increases, are needed to control for this possibility. They were not employed here in order to achieve sufficient trials to investigate devaluation of appetitive foods. There appears to be an absence of experimental work investigating the relative effect of devaluation on motivational salience and motivational value and it might be expected that both processes always occur in tandem.

While the interpretation of the non-devaluing RewP remains open to debate, there was an absence of devaluation throughout the EEG. Other EEG studies analysing effects outside the traditional interval of the RewP have suggested the sensitivity of the EEG to the contingencies represented in model-based learning (e.g., Chase, Swainson et al., 2011; Luque, Morís et al. 2015; Collins & Frank, 2016; Reiter, Koch et al. 2016; Sambrook, Hardwick et al., 2018) and thus provide evidence of goal directed processes in the EEG when reward feedback is delivered. The present study did not attempt to demonstrate neural corelates of model-based contingencies, focussing instead on the other facet of goal-directed behaviour, devaluation sensitivity, for which it found no neural evidence. In an experiment using symbols to stand for points towards an end-of-session payment Luque et al. (2017) were able to show devaluation sensitivity in the interval 550–700 ms. In a similar vein, Yousuf et al. (2019) found that in a slips of action task, designed to measure devaluation sensitivity to recently learned associations, the N2 was modulated by goal-directed vs. habitual control. An important respect in which the present experiment differed from these was that the outcome stimuli were veridical representations of foods rather than symbols. This was deemed important in order to draw on the kinds of long-term associations that would underlie cue-reactivity in a real-world context and increase the ecological validity of the task. It is possible that rather different processes are in play for stimuli that have undergone very transient conditioning within an experiment and there would be more scope for participants to successfully take a model-based approach with these.

In conclusion, the absence of any effect of the devaluation procedure in the EEG immediately following image presentation, despite clear behavioural and subjective effects of devaluation, raises the possibility that the food's motivational value is not well represented in this time window. If so, inhibition of eating in the presence of food cues may be dependent on goal-directed processes that follow. Food cues are known to increase the likelihood of eating in sated individuals. While the effects of food cues on behaviour are likely complex the present study proposes a role for learned responses to cues.

CRediT authorship contribution statement

Thomas D. Sambrook: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Andy J. Wills:** Writing – review & editing, Software, Conceptualization. **Ben Hardwick:** Investigation. **Jeremy Goslin:** Writing – review & editing, Software.

Ethics

Institutional review board approval for the study was obtained and all participants provided informed consent.

Ethical statement

This study was approved by the School of Psychology research ethics committee of the University of Plymouth.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.appet.2025.108390.

Data availability

Available at https://osf.io/yn9z6/

References

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology Section B Comparative and Physiological Psychology, 33*(May), 109–121. https://doi.org/10.1080/14640748108400816

Akam, T., Costa, R., & Dayan, P. (2015). Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS Computational Biology, 11*(12). https://doi.org/ARTNe1004648.

Balleine, B. W., Daw, N. D., & O'Doherty, J. P. (2008). *Multiple forms of value learning and the function of dopamine* (pp. 367–385). Neuroeconomics: Decision making and the brain.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Boggiano, M. M., Dorsey, J. R., Thomas, J. M., & Murdaugh, D. L. (2009). The Pavlovian power of palatable food: lessons for weight-loss adherence from a new rodent model of cue-induced overeating. *International Journal of Obesity, 33*(6), 693–701. https://doi.org/10.1038/ijo.2009.57

Boswell, R. G., & Kober, H. (2016). Food cue reactivity and craving predict eating and weight gain: a meta-analytic review. *Obesity Reviews, 17*(2), 159–177. https://doi.org/10.1111/obr.12354

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron, 68*(5), 815–834. https://doi.org/10.1016/j.neuron.2010.11.022

Buabang, E. K., Boddez, Y., De Houwer, J., & Moors, A. (2021). Don't make a habit out of it: Impaired learning conditions can make goal-directed behavior seem habitual. *Motivation Science, 7*(3), 252–263. https://doi.org/10.1037/mot0000218

Byrd, R. H., Lu, P. H., Nocedal, J., & Zhu, C. Y. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing, 16*(5), 1190–1208. https://doi.org/10.1137/0916069

Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience, 23*(4), 936–946. https://doi.org/10.1162/jocn.2010.21456

Cockburn, J., & Holroyd, C. B. (2018). Feedback information and the reward positivity. *International Journal of Psychophysiology, 132*, 243–251. https://doi.org/10.1016/j.ijpsycho.2017.11.017

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience, 27*(2), 371–378. https://doi.org/10.1523/Jneurosci.4421-06.2007

Colagiuri, B., & Lovibond, P. F. (2015). How food cues can enhance and inhibit motivation to obtain and consume food. *Appetite, 84*, 79–87. https://doi.org/10.1016/j.appet.2014.09.023

Collins, A. G. E., & Frank, M. J. (2016). Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition, 152*, 160–169. https://doi.org/10.1016/j.cognition.2016.04.002

Cornell, C. E., Rodin, J., & Weingarten, H. (1989). Stimulus-induced eating when satiated. *Physiol Behav, 45*(4), 695–704. https://doi.org/10.1016/0031-9384(89)90281-3

da Silva, C. F., & Hare, T. A. (2020). Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour, 4*(10), 1053. https://doi.org/10.1038/s41562-020-0905-y

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

De Houwer, J., Buabang, E. K., Boddez, Y., Köster, M., & Moors, A. (2023). Reasons to Remain Critical About the Literature on Habits: A Commentary on Wood et al. (2022). *Perspectives on Psychological Science, 18*(4), 871–875. https://doi.org/10.1177/17456916221131508

de Wit, S., Kindt, M., Knot, S. L., Verhoeven, A. A. C., Robbins, T. W., Gasull-Camos, J., Evans, M., Mirza, H., & Gillan, C. M. (2018). Shifting the balance between goals and habits: Five failures in experimental habit induction. *Journal of Experimental Psychology: General, 147*(7), 1043–1065. https://doi.org/10.1037/xge0000402

Forester, G., Schaefer, L. M., Johnson, J. S., Amponsah, T., Dvorak, R. D., & Wonderlich, S. A. (2024). Neurocognitive reward processes measured via event-related potentials are associated with binge-eating disorder diagnosis and ecologically-assessed behavior. *Appetite, 193*. https://doi.org/ARTN107151.

Francis, G., & Jakicic, V. (2023). Equivalent statistics for a one-sample -test. *Behavior Research Methods, 55*(1), 77–84. https://doi.org/10.3758/s13428-021-01775-3

Friedel, E., Koch, S. P., Wendt, J., Heinz, A., Deserno, L., & Schlagenhauf, F. (2014). Devaluation and sequential decisions: Linking goal-directed and model-based behavior. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00587

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience, 15*(3), 523–536. https://doi.org/10.3758/s13415-015-0347-6

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology, 71*(2), 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

Halford, J. C. G., Gillespie, J., Brown, V., Pontin, E. E., & Dovey, T. M. (2004). Effect of television advertisements for foods on food consumption in children. *Appetite, 42*(2), 221–225. https://doi.org/10.1016/j.appet.2003.11.006

Harris, A., Hare, T., & Rangel, A. (2013). Temporally dissociable mechanisms of self-control: Early attentional filtering versus late value modulation. *Journal of Neuroscience, 33*(48), 18917–18931. https://doi.org/10.1523/Jneurosci.5816-12.2013

Heldmann, M., Müller-Miny, L., Wagner-Altendorf, T., & Münte, T. F. (2025). Automatic attentional capture by food items in a visuospatial attention task–A study with event-related brain potentials. *Behavioural Brain Research, 484*. https://doi.org/10.1016/j.bbr.2025.115514

Herle, M., Fildes, A., Steinsbekk, S., Rijsdijk, F., & Llewellyn, C. H. (2017). Emotional over- and under-eating in early childhood are learned not inherited. *Scientific Reports, 7*. https://doi.org/10.1038/s41598-017-09519-0

Hird, E. J., El-Deredy, W., Jones, A., & Talmi, D. (2018). Temporal dissociation of salience and prediction error responses to appetitive and aversive taste. *Psychophysiology, 55*(2). https://doi.org/10.1111/psyp.12976

Hogarth, L. (2012). Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: Evidence for independent additive controllers. *Journal of Experimental Psychology-Animal Behavioral Processes, 38*(3), 266–278. https://doi.org/10.1037/a0028914

Hogarth, L., & Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: Implications for dependence vulnerability. *Journal of Experimental Psychology-Animal Behavioral Processes, 37*(3), 261–276. https://doi.org/10.1037/a0022913

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679–709. https://doi.org/10.1037/0033-295x.109.4.679

Huvermann, D. M., Bellebaum, C., & Peterburs, J. (2021). Selective devaluation affects the processing of preferred rewards. *Cognitive, Affective, & Behavioral Neuroscience, 21*(5), 1010–1025. https://doi.org/10.3758/s13415-021-00904-x

Kavanagh, D. J., Andrade, J., & May, J. (2005). Imaginary relish and exquisite torture: The elaborated intrusion theory of desire. *Psychological Review, 112*(2), 446–467. https://doi.org/10.1037/0033-295x.112.2.446

Kool, W., Cushman, A., & Gershman, S. J. (2016). When does model-based control pay off? *PLoS Computational Biology, 12*(8). https://doi.org/10.1371/journal.pcbi.1005090

Kool, W., Gershman, S. J., & Cushman, F. A. (2017). Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychological Science, 28*(9), 1321–1333. https://doi.org/10.1177/0956797617708288

Kropotov, J., Ponomarev, V., Tereshchenko, E. P., Müller, A., & Jäncke, L. (2016). Effect of aging on ERP components of cognitive control. *Frontiers in Aging Neuroscience, 8*. https://doi.org/10.3389/fnagi.2016.00069

Luque, D., Beesley, T., Morris, R. W., Jack, B. N., Griffiths, O., Whitford, T. J., & Le Pelley, M. E. (2017). Goal-directed and habit-like modulations of stimulus processing during reinforcement learning. *Journal of Neuroscience, 37*(11), 3009–3017. https://doi.org/10.1523/Jneurosci.3205-16.2017

Luque, D., Morís, J., Rushby, J. A., & Le Pelley, M. E. (2015). Goal-directed EEG activity evoked by discriminative stimuli in reinforcement learning. *Psychophysiology, 52*(2), 238–248. https://doi.org/10.1111/psyp.12302

Morey, R. D., & Rouder, J. N. (2024). BayesFactor: Computation of bayes factors for common designs. https://CRAN.R-project.org/package=BayesFactor.

Nebe, S., Kretzschmar, A., Brandt, M. C., & Tobler, P. N. (2024). Characterizing human habits in the lab. *Collabra-Psychology, 10*(1). https://doi.org/ARTN9294910.1525/collabra.92949.

Ohla, K., Toepel, U., le Coutre, J., & Hudry, J. (2012). Visual-gustatory interaction: Orbitofrontal and Insular cortices mediate the effect of high-calorie visual food cues on taste pleasantness. *PLoS One, 7*(3). https://doi.org/10.1371/journal.pone.0032434

Pedroni, A., Langer, N., Koenig, T., Allemand, M., & Jancke, L. (2011). Electroencephalographic topography measures of experienced utility. *Journal of Neuroscience, 31*(29), 10474–10480. https://doi.org/10.1523/Jneurosci.5488-10.2011

Peterburs, J., Sannemann, L., & Bellebaum, C. (2019). Subjective preferences differentially modulate the processing of rewards gained by own vs. observed choices. *Neuropsychologia, 132*. https://doi.org/10.1016/j.neuropsychologia.2019.107139

Petrovich, G. D., Ross, C. A., Gallagher, M., & Holland, P. C. (2007). Learned contextual cue potentiates eating in rats. *Physiol Behav, 90*(2–3), 362–367. https://doi.org/10.1016/j.physbeh.2006.09.031

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449–459. https://doi.org/10.1111/psyp.12370

R Core Team. (2017). R: A language and environment for statistical computing. https://www.r-project.org/.

Reiter, A. M. F., Koch, S. P., Schröger, E., Hinrichs, H., Heinze, H. J., Deserno, L., & Schlagenhauf, F. (2016). The feedback-related negativity codes components of abstract inference during reward-based decision-making. *Journal of Cognitive Neuroscience, 28*(8), 1127–1138. https://doi.org/10.1162/jocn_a_00957

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin, 141*(1), 213–235. https://doi.org/10.1037/bul0000006

Sambrook, T. D., Hardwick, B., Wills, A. J., & Goslin, J. (2018). Model-free and model-based reward prediction errors in EEG. *NeuroImage, 178*, 162–171. https://doi.org/10.1016/j.neuroimage.2018.05.023

San Martin, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience, 6*. https://doi.org/10.3389/fnhum.2012.00304

Schachte, S., Goldman, R., & Gordon, A. (1968). Effects of fear food deprivation and obesity on eating. *Journal of personality and social psychology, 10*(2). https://doi.org/10.1037/h0026284, 91-&.

Shanks, D. R., & St John, M. F. (1994). Characteristics of dissociable human learning-systems. *Behavioral and Brain Sciences, 17*(3), 367–395. https://doi.org/10.1017/S0140525x00035032

Springer, A., Ohlendorf, F., Schober, J., Lange, L., & Osinsky, R. (2021). Do food images as action outcomes evoke a reward positivity? *Brain Cogn, 154*. https://doi.org/10.1016/j.bandc.2021.105804, 10.1016/j.bc.2021.105804.

Stewardson, H. J., & Sambrook, T. D. (2020). Evidence for parietal reward prediction errors using great grand average meta-analysis. *International Journal of Psychophysiology, 152*, 81–86. https://doi.org/10.1016/j.ijpsycho.2020.03.002

Stewardson, H. J., & Sambrook, T. D. (2021). Reward prediction error in the ERP following unconditioned aversive stimuli. *Scientific Reports, 11*(1). https://doi.org/10.1038/s41598-021-99408-4

Stewardson, H., & Sambrook, T. D. (2023). Valence precedes value in neural encoding of prediction error. *Psychophysiology, 60*(7). https://doi.org/10.1111/psyp.14266

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning.* MIT Press.

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience, 33* (19), 8264–8269. https://doi.org/10.1523/Jneurosci.5695-12.2013

Tang, D. W., Fellows, L. K., Small, D. M., & Dagher, A. (2012). Food and drug cues activate similar brain regions: A meta-analysis of functional MRI studies. *Physiol Behav, 106*(3), 317–324. https://doi.org/10.1016/j.physbeh.2012.03.009

Toepel, U., Knebel, J. F., Hudry, J., le Coutre, J., & Murray, M. M. (2009). The brain tracks the energetic value in food images. *NeuroImage, 44*(3), 967–974. https://doi.org/10.1016/j.neuroimage.2008.10.005

Volkow, N. D., Wang, G. J., & Baler, R. D. (2011). Reward, dopamine and the control of food intake: Implications for obesity. *Trends in Cognitive Sciences, 15*(1), 37–46. https://doi.org/10.1016/j.tics.2010.11.001

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America, 108*(47), 19048–19053. https://doi.org/10.1073/pnas.1117189108

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews, 36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Watson, P., Wiers, R. W., Hommel, B., & de Wit, S. (2014). Working for food you don't desire. Cues interfere with goal-directed food-seeking. *Appetite, 79*, 139–148. https://doi.org/10.1016/j.appet.2014.04.005

Wills, A., Dome, L., Edmunds, C., Honke, G., Inkster, A., Schlegelmilch, R., & Spicer, S. (2023). Catlearn: Formal psychological models of categorization and learning. https://cran.r-project.org/package=catlearn.

Wills, A. J., O'Connell, G., Edmunds, C. E. R., & Inkster, A. B. (2017). Progress in modeling through distributed collaboration: Concepts, tools and category-learning examples. *Psychology of Learning and Motivation, 66*(66), 79–115. https://doi.org/10.1016/bs.plm.2016.11.007

Yin, H. H., Ostlund, S. B., & Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: The integrative functions of cortico-basal ganglia networks. *European Journal of Neuroscience, 28*(8), 1437–1448. https://doi.org/10.1111/j.1460-9568.2008.06422.x

Yousuf, M., Heldmann, M., Münte, T. F., & Doñamayor, N. (2019). Electrophysiology of goal-directed versus habitual control during outcome devaluation. *Cortex, 119*, 401–416. https://doi.org/10.1016/j.cortex.2019.08.002