RESEARCH ARTICLE

Reward signalling in brainstem nuclei under fluctuating blood glucose

Tobias Morville[1], Kristoffer H. Madsen[2], Hartwig R. Siebner[1,3], Oliver J. Hulme[1] *

1 Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark, **2** DTU Compute, Department of Informatics and Mathematical Modelling, Technical University of Denmark, Copenhagen, Denmark, **3** Department of Neurology, Copenhagen University Hospital Bispebjerg, Copenhagen, Denmark

* oliverh@drcmr.dk

Abstract

Phasic dopamine release from mid-brain dopaminergic neurons is thought to signal errors of reward prediction (RPE). If reward maximisation is to maintain homeostasis, then the value of primary rewards should be coupled to the homeostatic errors they remediate. This leads to the prediction that RPE signals should be configured as a function of homeostatic state and thus diminish with the attenuation of homeostatic error. To test this hypothesis, we collected a large volume of functional MRI data from five human volunteers on four separate days. After fasting for 12 hours, subjects consumed preloads that differed in glucose concentration. Participants then underwent a Pavlovian cue-conditioning paradigm in which the colour of a fixation-cross was stochastically associated with the delivery of water or glucose via a gustometer. This design afforded computation of RPE separately for better- and worse-than expected outcomes during ascending and descending trajectories of serum glucose fluctuations. In the parabrachial nuclei, regional activity coding positive RPEs scaled positively with serum glucose for both ascending and descending glucose levels. The ventral tegmental area and substantia nigra became more sensitive to negative RPEs when glucose levels were ascending. Together, the results suggest that RPE signals in key brainstem structures are modulated by homeostatic trajectories of naturally occurring glycaemic flux, revealing a tight interplay between homeostatic state and the neural encoding of primary reward in the human brain.

Introduction

A basic assumption of many models of adaptive behavior, is that the value of primary rewards are modulated by their capacity to rectify future homeostatic deficits [1, 2]. Compatible with this notion, deprivation-induced hypoglycaemia increases willingness to work for food in rats and humans [3], as well as the subjectively reported pleasure [2]. Dopamine is a neurotransmitter that plays a key role in signalling reward [4] and is involved in behavioural reinforcement, learning, and motivation [5, 6]. Via meso-cortical and mesolimbic dopaminergic projections, synaptic dopamine release modulates the plasticity of cortico-striatal networks and thereby sculpts behavioural policies according to their reward contingencies [4, 7].

Patterns of phasic dopaminergic firing have been demonstrated to follow closely the principles of reinforcement learning, encoding the errors in the prediction of reward [6, 8–10]. Reward prediction error (RPE) signals appear commensurate with the economic construct of marginal utility, defined as the additional utility obtained through additional units of consumption, where utility is a subjective value inferred from choice [7, 11, 12].

Although animals are motivated by a homeostatic deficit of thirst or hunger, homeostatic states are rarely considered as relevant modulators of dopaminergic signalling of reward prediction errors. In typical paradigms involving cumulative consumption, the homeostatic deficit gradually diminishes as the animal plays for consumption of water or sugar-containing juice. Eventually, the animal rejects further play, presumably because the marginal utility of consumption diminished to a point of indifference or even aversion. Interestingly, a recent electrophysiology study in rats, demonstrated that oral consumption of sodium solution causes phasic dopaminergic signals in the nucleus accumbens, that are modulated by sodium depletion [13].

There is now growing evidence for a multifaceted interface between dopamine mediated reward-signalling and the systems underpinning energy homeostasis. Firstly, dopamine neurons in the ventral tegmental area (VTA) express a suite of receptors targeted by energy-reporting hormones ghrelin, insulin, amylin, leptin and Glucagon Like Peptide 1 (GLP-1) [14, 15]. This provides numerous degrees of freedom for flexibly interfacing between homeostatic state and reward signalling. Although hormonal modulations of phasic dopamine are yet to be fully scrutinised, there is emerging evidence that circulating factors do indeed modulate its magnitude. For instance, amylin, a hormone co-released with insulin, acts on the VTA to reduce phasic dopamine release in its mesolimbic projection sites [16]. In terms of neuronal input, there are many such opportunities for the appetitive control of dopamine mediated signalling.

Appetitive control can be delineated into three interacting systems [17]. The first system generates a negative valence signal which involves activity of the Agouti-related peptide (AgRP) neurons of the arcuate nucleus of the hypothalamus (ARC). Activity of ARC_{AgRP} neurons reports on energy deficits, inhibits energy expenditure, and regulates glucose metabolism [18–21]. ARC neurons that contain peptide products of pro-opiomelanocortin (POMC) form an opponent code compared with ARC_{AgRP} neurons. The balance between the two neuronal ARC sub-populations putatively encodes the value of near-term energetic states, becoming rapidly modulated just prior to food consumption [22]. The second system codes positive valence signals and consists of circuits involving the lateral hypothalamus (LH). It is linked to positively reinforcing consummatory behaviours via its GABAergic projections to VTA dopamine neurons [23] assumed to trigger positive feedback to keep consumption going during feeding bouts. The third valuation system involves calcitonin gene-related protein (CGRP)-expressing neurons in the parabrachial nuclei (PBN) that potently suppress eating when activated, but do not increase food intake when inhibited. PBN_{CGRP} neurons are activated by signals associated with food intake, and they provide a signal of satiety that has negative valence when strongly activated [24]. The PBN has been characterised as a hedonic hotspot, the modulation of which by either GABA or Benzodiazepines potently modulates experienced reward [25]; ARC_{AgRP} neurons GABA-ergically inhibit PBN neurons, thus stimuli predicting glucose consumption should inhibit ARC_{AgRP}, releasing the PBN from inhibition [26]. Further, hormones related to hunger and feeding (GLP-1 & leptin) modulate PBN activity and subsequent behaviour [27, 28]. Of note, these three valuation systems all project to and modulate the dopaminergic neurons in the ventral tegmental area (VTA_{DA}). The interface between these hypothalamic-brainstem networks and the VTA_{DA}, is arguably the most important interface for mediating the dialogue between energy homeostasis and value computation.

While most evidence for encoding of RPEs is obtained under homeostatic deprivation, the modulation of RPE signalling triggered by physiological fluctuations in glucose availability (glycaemic flux) remains yet to be characterised in the human brain. This begs the questions, how are RPE signals modulated by these subcortical circuits that integrate, evaluate, and predict energy-homeostatic states? We hypothesize that glucose fluctuations above and below average levels of serum glucose, will down and up modulate RPE responses in these hypothalamic-brainstem networks. To test these hypotheses, we acquired a large volume of fMRI data in five participants during a simple Pavlovian cue-conditioning task, while their serum glucose was systematically manipulated.

Methods

Subjects

Five healthy (3 male), normal-weight subjects in the age range 23 to 29, participated in the study. Exclusion criteria were: $20 > BMI > 25$; $18 > Age > 32$ yrs; any metabolic or endocrine diseases or gastrointestinal disorder; any known medication that might interfere with the study; claustrophobia; and any metal implants or devices that could not be removed. Informed consent was obtained in writing from all subjects as approved by the Regional Ethics Committee of Region Hovedstaden (protocol H-4-2013-100) and in accordance with the declaration of Helsinki.

Experimental procedure

The experimental design constituted a single-blinded, randomised control trial, with repeated measures crossover-design. On four separate days, subjects fasted for a minimum of twelve hours before testing. Compliance with the fasting instruction was based both on trust, and on the understanding that we would be able to detect if participants had not fasted via blood tests. Any participant that was not in the hypoglycemic range (defined here as <6mmol/L) at the start of the experiment would be assumed to have not fasted, and the session would be aborted. This was not necessary for any participant. At the beginning of an experimental session, subjects ingested either a hi-glucose (75 g, 300 kcal) or lo-glucose preload (10 g, 40 kcal) diluted to 100 ml with a non-caloric lemon juice, used in order to mask the taste of the glucose. The lo-glucose preload resulted in glucose ascending during the fMRI acquisition period, due to the consumption of glucose, whereas during the same period after the hi-glucose preload, glucose levels descended (Fig 1B). The preload conditions are thus referred to as ascending and descending conditions. The temporal order of the conditions was randomised within subjects, with each condition being performed twice. Both preloads were anecdotally reported by independent samplers to be highly palatable. Each delivery of glucose reward was 0.4ml, corresponding to 0.3 g of glucose (1.2 kcal) per delivery.

Experimental task

After consuming the preload, participants engaged in a simple Pavlovian cue-conditioning task. The colour of the fixation cross cued both the onset of each trial (Cue$_{onset}$), as well as stochastically predicting glucose delivery (Fig 1A), with one colour signalling a high probability of glucose delivery (Cue$_{high}$), and another signalling a low probability (Cue$_{low}$). 10–15 seconds after delivery of the liquid, a purple cross signalled that subjects were to swallow. The large temporal distance between the swallowing and the reward onset, as well as the levels of temporal jitter, was designed to mitigate the contamination of the reward signals by swallowing related artefacts. All probabilities and contingencies were implicitly revealed only through
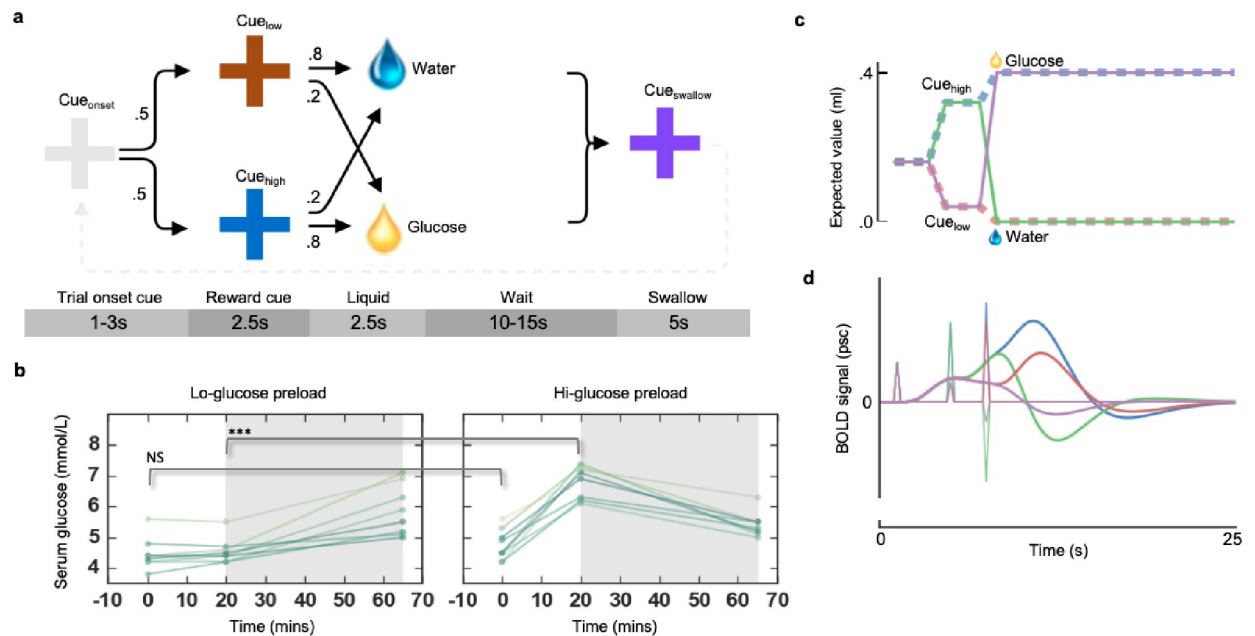
Fig 1. Experimental design, glucose trajectories, and expected reward signals. a, participants were presented with the Cue$_{onset}$ (grey fixation cross) for 1-3s after which either Cue$_{high}$ (blue cross) or Cue$_{low}$ (brown cross) is presented with a probability of 0.5 each. Cue$_{high}$ signalled a high probability (0.8) of glucose delivery and a low probability (0.2) of water delivery. Cue$_{low}$ signalled a low probability of glucose (0.2) and a high probability of water (0.8). The 0.4ml of the liquid were delivered over 2.5 seconds, followed by 10-15s wait period and a Cue$_{swallow}$ that cued the subject to swallow (here, purple) which lasted for 5s. All jitters are uniformly distributed within the ranges specified. b, serum glucose trajectories for the high and low glucose preload conditions. Grey shading indicates the period of fMRI acquisition for a single session. The different line plots indicate different sessions for all subjects. Glucose levels ascend during the fMRI acquisition period in the lo-glucose condition, and descend in the hi-glucose condition. c, graph depicts the objective reward expectations, expressed as the expected value in ml glucose, and the perturbation of these expectations under the onset of the experimental cues and outcomes. Note that reward expectations are updated three times per trial: at the onset of the Cue$_{onset}$; at the onset of Cue$_{high}$ or Cue$_{low}$; at the onset of Outcome$_{glucose}$ or Outcome$_{water}$. d, illustrates simulated BOLD responses to RPE signals resulting from the updated reward expectations shown in c, generated by convolving the canonical hemodynamic response function with the RPE stick functions evoked by changes to the reward expectations.

https://doi.org/10.1371/journal.pone.0243899.g001

experience in the scanner, and all were stationary over all test days. The mapping between colour and outcome probabilities was counterbalanced across subjects, while the mapping was stationary within and between sessions. Participants went through ~82 trials [82 ± 1.5 SEM] each day giving ~328 trials per subject. Serum glucose measurements were attained immediately before and 20 minutes after ingestion, using a Contour® Next glucose meter (Fig 1B). It should be noted that whilst there is no jitter between the cue and the outcome, independent estimation of each of the effects is achievable by virtue of their probabilistic transitions (Fig 1A).

Scanning procedure

Task related changes in regional brain activity were mapped with blood oxygen dependent (BOLD) MRI immediately after the second glucose measurement (t20). Functional MRI measurements were performed with a 3T Philips Achieva and a 32 channel receive head coil using a gradient echo T2* weighted echo-planar image (EPI) sequence with a repetition time of 2526 ms, and a flip-angle of 80°. Each volume consisted of 40 axial slices of 3 mm thickness and 3 mm in-plane resolution (220 x 220 mm). The axial field-of-view was 120 mm covering the whole brain, cutting off the medulla oblongata partially. This sequence was extensively piloted and optimised specifically for reducing distortion and maximising resolution with the hypothalamic and brainstem regions of interest. During each session, 800 EPI volumes were

acquired, resulting in 3200 EPI volumes per subject. Further, an anatomical T1-weighted image was recorded for each subject. Respiration and heart rate were measured to assess and model possible artefacts. Liquid tastants were contained in two 50 ml syringes, one containing water-only (water hence) the other containing glucose and water (glucose hence) solutions, attached to two programmable syringe pumps (AL1000-220, World Precision Instruments Ltd, Stevenage, UK), controlled by the stimulus paradigm script. The liquid was delivered orally via two separate 5m long 3mm wide silicone tubes. Each tube was attached to a gustatory manifold specifically built for the Philips head-coil (John B. Pierce Laboratory, Yale University). Visual stimuli were presented on a screen positioned ~30 cm away from the scanner.

Pre-processing

Pre-processing and image analysis were done using SPM12 software (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). To correct for motion, EPI scans were realigned to their mean using a two-step procedure and co-registered to the T1 weighted anatomical image. The realigned images were spatially normalised to the standard ICBM space template of European brains, with a resampled voxel size of 3 mm.

fMRI analysis

After model specification, the ascending condition sessions were concatenated using the function spm_fmri_concatenation (SPM 12) for each subject. Equivalently, the same concatenation was performed for the descending conditions sessions. A first-level fixed effects model was run over all subjects. The concatenation was performed to avoid state-dependent effects being expressed via the session-specific regression coefficients. All variables of interest were convolved with the canonical hemodynamic response function, along with their associated temporal and dispersion derivatives and fitted to the data using the specified GLM. The temporal evolution of cues and outcomes were modelled as separate conditions, each with state as parametric modulators. Regressors of no interest included a discrete cosine transform based 1/128 Hz cut-off frequency high-pass filter, rigid body realignment parameters using a 24 parameter Volterra expansion [29] and physiological noise from heart rate and respiration using the RETROICOR method [30]. We specified the striatum (caudate, putamen and nucleus accumbens), brainstem (pons, ventral tegmental area and substantia nigra) and hypothalamus as regions of interest (ROI). These ROIs were determined on the basis of the literature describing dopamine projections from midbrain to the striatum and its role in regulating behaviour as a function of reward. The pons was selected to accommodate the literature described above, which sets certain nuclei within the pons as important homeostatic modulators. All ROI were defined with the WFU pick atlas [31, 32] and cross checked against the book *Atlas of the human brain* [33]. All initial first-level analysis was performed as whole-brain uncorrected at $p < 0.001$. Significant clusters in regions of interest (ROI) are all reported as small-volume corrected with a family-wise threshold of $p < 0.05$ at cluster level (abbreviated SVC FWE), unless otherwise stated.

Modelling RPEs

At the first level, a general linear model (GLM) was set up to model cue and outcome related brain activity. We specified separate regressors which modelled the onset of cue events (Cue_{onset}, Cue_{high} and Cue_{low}) and outcome events ($Outcome_{gluc}$ & $Outcome_{water}$). Since there is no overt behavior in this task to which a temporal difference learning algorithm can be fit, we used a different approach based on how RPEs converge to changes in the conditional

expectation values of reward outcomes. As subjects learn the contingencies between the cues and the outcomes, the RPEs evoked by these events converge toward the change in expected value of the reward (here the volume of glucose), conditional on the events they have cumulatively experienced during the trial. This has been shown in single cell recordings, where the RPE signals of dopaminergic cells in the VTA, signal errors whose magnitudes reflect the change in expected value of juice volume, conditional on the events that have been experienced at that time [11]. Fig 1C illustrates how the expectation value of glucose volume evolves over time as a function of the cues observed. The conditional expectation value upon seeing the Cue$_{onset}$ is the expectation value of glucose volume for each trial, conditional on the fact the trial has started. This is shown in Fig 1C, since all lines begin from this starting point of 0.2ml. Thus, Cue$_{onset}$ triggers a small positive RPE, seen as the first spike on the left in Fig 1D. From there the expectation value of glucose increases or decreases depending on whether the Cue$_{high}$ or Cue$_{low}$ is experienced, as labelled in Fig 1C. This causes the second set of spikes seen in Fig 1D, where the green corresponds to the Cue$_{high}$ and the purple corresponds to Cue$_{low}$. Finally, the outcomes arrive changing again the conditional expectational value of glucose to either 0.4ml or 0ml. This can be thought of as the expectation value of the glucose that the agent can expect to metabolise having received the liquid in its mouth. This corresponds to the third set of spikes in Fig 1D. To approximate the RPEs without behavior, we specified contrasts which were formulated by linear combinations of these cue and outcome regressors, weighted as a function of the RPE values that would be expected from the temporal-difference learning algorithm, once converged [34]. In other words, the cue and outcome regressors were weighted by the change in conditional expectation of glucose volume caused by the regressor's event (either the cue or the outcome). A contrast of positive RPE signals (RPE$_{pos}$) was computed by assigning the positive valence cue and outcome regressors (i.e. Cue$_{onset}$, Cue$_{high,}$ and Outcome$_{gluc}$) contrast weights that were proportional to the change in expectation value of glucose volume that these events caused. Equivalently negative RPEs (RPE$_{neg}$) were computed equivalently as contrasts to include only the negative valence events (i.e. Cue$_{low}$ and Cue$_{water}$). It should be noted here that in this approach to modelling RPEs, the model does not incorporate any effect of learning, in effect modelling what signals are expected once learning has converged on the expected reward values.

Modelling RPE modulation by glucose

The effect of serum glucose on RPE was modelled via first order parametric modulator of the cue and outcome regressors, taking the interpolated serum glucose at each time point as the covariate (demeaned). This resulted in five parametric modulator regressors, namely Cue$_{onset^*state}$, Cue$_{high^*state}$, Cue$_{low^*state}$, Outcome$_{gluc^*state}$ & Outcome$_{water^*state}$. From these regressors, contrasts can be specified to model the effect of serum glucose on RPEs. The effect of glucose state on positive RPEs (RPE$_{pos^*state}$) was computed as a linear combination of Cue$_{onset^*state}$, Cue$_{high^*state}$, and Outcome$_{gluc^*state}$ regressors. Equivalently, the effect of glucose state on negative RPEs (RPE$_{neg^*state}$) was computed as a linear combination of (Cue$_{low^*state}$, Outcome$_{water^*state}$). All 1st-order parametric modulators are orthogonal to their associated onset regressors by construction. No other orthogonalization of regressors was performed.

Results

Cue$_{onset}$ induced brain activity

An RPE signal should respond to the Cue$_{onset}$, with an error signal that signals the expected value of glucose reward for the whole trial [11]. Computing the main effect of this regressor, this was found to evoke an increase in activity in VTA bilaterally (Fig 2A). Thus cue-induced
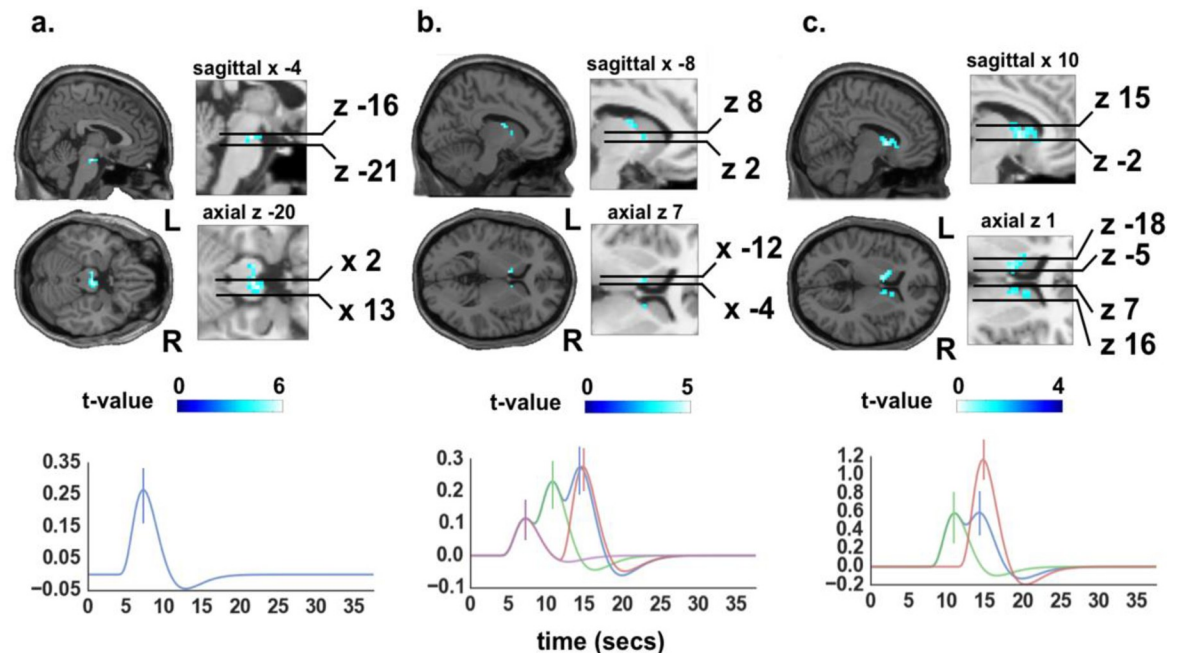
Fig 2. Statistical parametric maps of main effects of trial onset, positive and negative RPEs. a, main effect of Cue$_{onset}$, which reflects an RPE following the mean reward expectation for the whole trial, revealed activity in VTA bilaterally (β = 2.77) (R: [4–17–20] and L: [–8–17–20], FWE SVC). Further this revealed deactivation of precentral gyrus (primary somatosensory cortex), mediodorsal thalamus, and striatum (FWE whole brain, not shown). Lower panel shows fitted response for the Cue$_{onset}$ event within the same region. b, main effect of RPE$_{pos}$ revealed activity in left lateral caudate [β = 1.21; coordinates -8 4 7; FWE SVC]. Lower panel shows fitted response for the RPE$_{pos}$ contrast within the region. The colours of time courses follow the same meaning as shown in Fig 1C and 1D. c, main effect of RPE$_{neg}$ revealed bilateral activity in caudate (L: -11–2 13; R: 10 7 1; β = 12.2) medial dorsal thalamic nucleus [7, –2, 22], and lateral insula [43, –2, –17] (all FWE). Lower panel shows fitted response for the RPE$_{neg}$ contrast within the region. All fitted responses were generated by convolving the canonical hemodynamic response function with the RPE stick function multiplied by their respective beta-values extracted from the local maxima of the ROI in units of percent signal change. Again, the time courses follow the same meaning as shown in Fig 1C and 1D. Error bars show standard errors of the mean.

https://doi.org/10.1371/journal.pone.0243899.g002

VTA activation is consistent with existing evidence of VTA signalling RPEs [35–37]. Cue$_{onset}$ also led to the deactivation of the postcentral gyrus (primary somatosensory cortex), mediodorsal thalamus, and the striatum [whole brain, uncorrected p < 0.001] (not shown).

Positive and negative reward prediction error signals

In several brain regions, regional task-related activity changed in proportion with the magnitude of positive-going (i.e. better-than-expected) reward prediction errors (RPE$_{pos}$) or negative-going (i.e. worse-than-expected) reward prediction errors (RPE$_{neg}$). Task related activity scaling with the RPE$_{pos}$, formalized as an RPE-weighted linear combination of Cue$_{trial}$, Cue$_{high}$, and Outcome$_{gluc}$, was found in left lateral caudate nucleus (Fig 2B). Conversely, task related activity reflecting RPE$_{neg}$, formalized as an RPE-weighted linear combination of Cue$_{low}$ and Outcome$_{water}$, was located in the caudate nucleus bilaterally Fig 2C), the medial dorsal thalamic nucleus, and insula (not shown).

Modulation of task-related brain activity by glycaemic state

We were interested to identify changes in RPE signalling over time as serum glucose either ascended or descended. A bilateral cluster, including the parabrachial nuclei (PBN), showed a modulation of the regional neural responses to RPEs by the glycaemic state dynamics (Fig 4A).
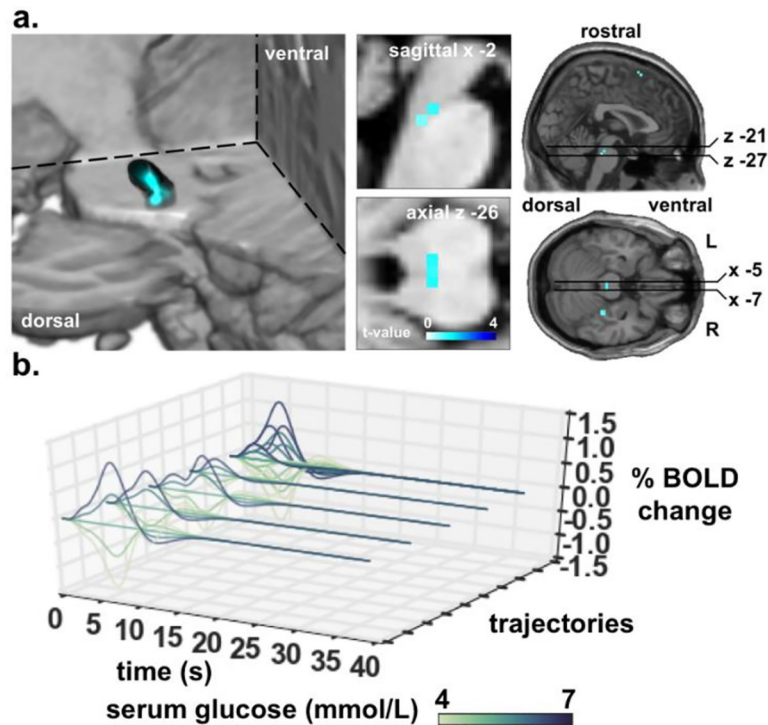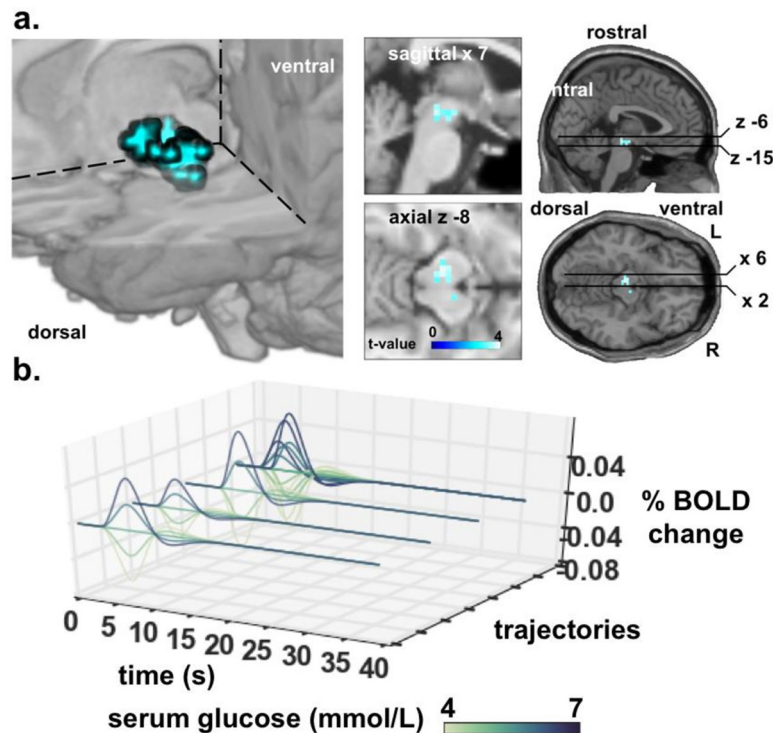
Fig 3. Statistical parametric maps of RPE$_{pos*state}$ and fitted responses over varying glycaemic state. a, main effect of RPE$_{pos*state}$ revealed bilateral activity in the PBN [-2–29–26; FWE SVC]. The colour bar indicates the t-value on a scale of white to blue. b, fitted response (β = 1.66) of the local maxima of PBN cluster (7 voxels) to the possible trajectories that RPE$_{pos*state}$, yield (Fig 1D) modulated by serum glucose state. There are four possible trajectories of the positive RPE, according to the four different possible trial types depicted in Fig 1C. These trajectories are modulated by serum glucose and shown as different colours with 5 different equally spaced glycemic states. There are thus 4 sets of these trajectories (the four closest to the viewer), each showing their modulation by glucose in the different colours. In order of closeness to the viewer, the first set of trajectories is for the Cue$_{high}$- Outcome$_{gluc}$ trial; the second set is for the Cue$_{high}$−Outcome$_{water}$ trial; the third set is for Cuelow−Outcome$_{gluc}$ trial; the fourth set is for the Cuelow −Outcome$_{water}$ trial. The fifth set of trajectories (furthest from the viewer) superimpose together all possible trajectories depicting the complexity of how the trajectories are patterned according to their modulation by glucose state. For comparison, this presentation is analogous to the superimposed trajectories shown in Fig 1D. Note that the colour bar for the graphs in (b) is serum glucose and is distinct from the colour bar in (a) used to indicate t-values.

https://doi.org/10.1371/journal.pone.0243899.g003

Higher levels of serum glucose amplified the response to RPE$_{pos}$ in the PBN region (Fig 3B). The main effect of RPE$_{neg*state}$, which models the interaction between RPE$_{neg}$ and state, did not yield any significant results in any ROI, or in exploratory analyses using uncorrected thresholds, in positive or negative contrasts. When considering both ascending and descending serum glucose fluctuations together, there was no detectable region where the RPE$_{neg}$ signal was either positively or negatively modulated by serum glucose. Brain responses to Cue$_{onset}$ were also not altered by glycaemic state.

We also tested for state-dependent modulatory effects on RPE processing which depends on whether serum glucose was ascending (Fig 1B, left) or descending (Fig 1B, right) over time. This yields four different contrasts (ascending vs. descending and the converse, for RPE$_{pos*state}$ and RPE$_{neg*state}$) that are directly relevant to glucose state. Subtracting descending trajectories from ascending and vice versa, revealed no significant activity changes for RPE$_{pos*state}$ [whole brain, uncorrected]. The same comparisons for RPE$_{neg*state}$ did reveal significant effects in VTA and substantia nigra for ascending trajectories relative to descending trajectories (Fig 4A). This result shows a relative amplification of the RPE$_{neg*state}$ signal as glucose state increases. In instances where reward was lower-than-expected (thus yielding negative RPE),

Fig 4. Statistical parametric maps of RPE$_{neg*state}$ subtracted for increasing minus decreasing. a, negative reward prediction error RPE$_{neg*state}$ revealed glucose modulated activity in SN [±12, -22, -10] and VTA [0, –15, –9] when subtracting the effect of descending from the ascending glucose state [FWE SVC]. b, fitted response (β = 0.34) of the local maxima of cluster [7, -11, 8; 52 voxels] to the three possible trajectories that RPE$_{neg*state}$ yield modulated by serum glucose state. Onsets are not at zero because the negative trajectories do not envelop the trial mean which has a positive expectation. In order of closeness to the viewer, the first set of trajectories is for the Cue$_{high}$- Outcome$_{gluc}$ trial; the second set is for the Cue$_{high}$–Outcome$_{water}$ trial; the third set is for Cue$_{low}$–Outcome$_{gluc}$ trial; the fourth set is for the Cue$_{low}$–Outcome$_{water}$ trial. The fifth set of trajectories (furthest from the viewer) superimpose together all possible trajectories depicting the complexity of how the trajectories are patterned according to their modulation by glucose state. The set of trajectories furthest away from the viewer superimposes the trajectories from all trial types into one plot.

the glucose state modulated the RPE$_{neg}$ signal in VTA and SN more so when glucose levels were ascending than descending.

Discussion

Participants performed a simple cue-conditioning task involving the probabilistic delivery of glucose or water, whilst their blood glucose fluctuated over the course of an hour. We had hypothesized that low levels of serum glucose will positively modulate the scale of positive RPE responses in hypothalamic-brainstem networks, reflecting the marginal utility of glucose as a function of homeostatic needs. Contrary to this hypothesis, we did not observe any positive RPE that increased its scaling with decreasing levels of serum glucose levels. In exploratory analyses there were however several observations of a dependency between serum glucose and RPE signals. Reward prediction error signalling in the parabrachial nuclei scaled positively with serum glucose levels, and this was true whether glucose was ascending or descending over time. We found that both the VTA and SN became more sensitive to negative RPEs for ascending compared to descending glycaemic trajectories. We begin by discussing the interpretation of these state modulated RPE effects, before considering other effects, and the limitations inherent under this paradigm.

In rodent models, the PBN acts as a 2nd order relay of inputs from the nucleus tractus solitarius, and is critical in the control of energy homeostasis via its projections to amygdala [38, 39], VTA [40], hypothalamus [39, 41] and the nucleus accumbens [42]. Subnuclei of the PBN are targeted by descending projections from several nuclei implicated in energy homeostasis, including hypothalamus, amygdala, and the bed nucleus of the stria terminalis [39, 43]. The PBN is known to be a potent site of reward modulation and subsequent behaviour in rodents. Microinjection of benzodiazepines [44–46], endocannabinoids [47], opioids [48, 49] and melanocortin agonists [50] into the PBN, all evoke hyperphagias. To our knowledge, the involvement of PBN in context of hedonics and reward signalling in the human brain remains to be charted. Here we provide tentative evidence that PBN activity generates a positive RPE-like signal that is sensitive to blood glucose and is time-locked to both the sensory cues predicting glucose, as well as glucose consumption.

We found no state modulation of RPE_{neg} signalling ($RPE_{neg*state}$), expressed during both ascending and descending glycaemic trajectories. For the RPE_{neg} signal, the modulatory effect of the glycaemic trajectory depended on whether glucose trajectories were ascending or descending. Regional activity scaling with RPE_{neg}, the VTA and SN showed significantly higher state modulation effects during ascending compared to the descending glycaemic paths. In our experiment, the ascending glucose trajectory resulted from a low-glucose preload with the subsequent increase over time likely occurring by virtue of the continual ingestion of glucose throughout the paradigm (Fig 1B). In the ascending condition, the neural response to RPE_{neg} is attenuated at lower levels of serum glucose, while it becomes amplified by the transition to higher serum glucose. Given that there is some evidence that dopaminergic neurons of the VTA and SN are directly inhibited by insulin [15], it is possible that the insulin release following hi-glucose preload was highest at the start of the paradigm, decreasing over time, and thus resulting in a gradual decrease in inhibition. However, it should be noted insulin can have a stimulatory effect on dopaminergic firing rates [51]. The difference in RPE_{neg} in its state modulation ($RPE_{neg*state}$) between ascending and descending may therefore be attributed to differential dynamics of insulin secretion [52], though other hormones such as ghrelin [52–54] or leptin [55–58] may play a role. It is not presently known, why it would make sense that a behaviorally reinforcing signal, such as phasic RPE, is expressed less as glucose levels are decreasing in a situation where the body is moving towards a state of potential dyshomeostasis.

Our finding that the VTA and SN responses are linked to RPE_{neg} may appear counterintuitive, given that these midbrain regions are typically associated with BOLD responses signalling positive-going RPEs. This is assumed to be by virtue of the fact that a greater range of firing rates can be devoted to the better-than-expected range, signalled by above baseline firing. This is contrasted to the worse-than-expected range, which can only be signalled by a decrease from an already low baseline frequency. It is conceivable that what we are asserting as being RPE_{neg} is in fact a positive RPE resulting from the gradual avoidance of glucose, which increases in magnitude with increasing levels of serum glucose as reported in humans [2] and rats [59]. Thus, as the experimental paradigm continues, especially under the conditions of glucose preload, serum glucose increases, and this may change the valence of the outcome, switching the affective connotation of glucose from palatable to aversive.

As detailed in the introduction, little is known about how the interface between dopaminergic RPE signalling and energy homeostasis is implemented in the human brain. While there are many means by which circulating factors can modulate activity in the VTA and SN, the mechanisms by which this is mediated cannot be revealed without wider hormonal assays. Contemporaneous hormonal sampling, as well as continuous glucose monitoring in the scanner will prove an important step in revealing these latent factors.

There are several technical limitations that should be noted in discussing this experiment. Though relatively high volumes of functional data (150 minutes per subject) were acquired in each subject, the total number of subjects was small. The reason for this was a focusing on maximising experimental power within subjects, for finer scale inference on the longer time-scale glucose dynamics. It was not known ahead of time how large the modulatory effects would be, and thus we deployed a conservative strategy of testing fewer subjects for longer. The long regressors that result from the concatenation of sessions may have meant that the high pass filtering would have reduced the effects of ascending and descending glucose levels. Inferring slow timescale dynamics is generally a problem for fMRI, however it is circumvented to some degree here, insofar as we are inferring the modulatory effect of glucose on faster RPE signalling dynamics which occur on a faster timescale than that which is filtered by the high-pass filter. Due to the small number of subjects, we deployed a fixed effects analysis over all subjects. It should be noted that this makes assumptions about the nature of the noise that might not be compatible with the repeated measures design, since it is difficult to correct for non-sphericity in this setting. Future work will expand this paradigm with a larger group of subjects to afford random effects modelling, and thus generalisation to the population sampled from. Contrary to our hypotheses, we found no modulatory effect of hypothalamic nuclei on RPE signalling. We stress that the current imaging protocols and field-strength (3T) were not optimal to dissociate neural activity in the hypothalamic nuclei. Due to the proximity of air sinuses adjacent to the hypothalamus and the effective resolution available, the present study most likely had insufficient sensitivity to capture activity in hypothalamic regions of interest. Another issue is whether the hemodynamic modelling was appropriate for detected evoked responses in subcortical regions which may deviate from the canonical hemodynamic response functions typically used. The regression model we used deployed temporal and dispersion derivatives for all regressors of interest in order to account for idiosyncratic variance in the timing and temporal spread of the hemodynamic response function. Finally, it should be noted that the cue-conditioning employed in this study was passive. Hence, subjects produced no overt choice behaviour against which to fit learning rate parameters for the RPE model, instead we relied on the asymptote values for the RPE signals. The problem of modelling RPEs in the absence of choice behaviour, motivates fitting learning rate parameters directly to brain data, a computational imaging approach that future work will exploit [60].

Recent literature on the computational neuroscience of reinforcement Learning (RL) has evidenced how decision-making in the mammalian brain is driven by optimizing the net value of both primary and non-primary rewards. Such reward computations have been shown to rest on a comparison between the expectation and outcome of external environmental cues, integrating both the physical and cognitive effort costs of the agent [61–63]. The work presented here tentatively expands this perspective this by showing that reward and RPE signals are dependent on internal homeostatic states, which may serve to modify the motivational values according to the personal and time varying homeostatic needs of the organism.

In conclusion, we exploited a simple paradigm, capable of eliciting RPEs under differential glycaemic trajectories, to identify brain stem structures that show a modulation of RPE signalling depending on the glycaemic state. We found that the PBN signals a positive-going reward prediction that is subject to systematic modulation by serum glucose. In the VTA and SN, negative-going RPEs were modulated by serum glucose trajectories, but in a way that was specific to an ascending glycaemic slope. Together the results show that RPE signals in key brainstem structures can be modulated by homeostatic trajectories inherent in naturally occurring glycaemic flux, revealing a potentially tight interplay between homeostatic state and the signalling of primary reward in the human brain.

Acknowledgments

Author Contributions

References

1. Pompilio L., Kacelnik A. & Behmer S.T., 2006. State-dependent learned valuation drives choice in an invertebrate. *Science (New York, N.Y.)*, 311(5767), pp.1613–1615.

2. Cabanac M., 1971. Physiological Role of Pleasure. *Science*, 173(4002), pp.1103–1107. https://doi.org/10.1126/science.173.4002.1103 PMID: 5098954

3. Sclafani A. Learned controls of ingestive behaviour. Appetite. 1997 Oct; 29(2):153–8. https://doi.org/10.1006/appe.1997.0120 PMID: 9344424.

4. Haber S.N. & Knutson B., 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*, 35(1), pp.4–26. https://doi.org/10.1038/npp.2009.129 PMID: 19812543

5. Berridge K.C., 2006. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, 191(3), pp.391–431. https://doi.org/10.1007/s00213-006-0578-x PMID: 17072591

6. Schultz W., Dayan P. & Montague P.R., 1997. A neural substrate of prediction and reward. *Science (New York, N.Y.)*, 275(5306), pp.1593–1599.

7. Schultz W., 2015. Neuronal Reward and Decision Signals: From Theories to Data. *Physiological reviews*, 95(3), pp.853–951. https://doi.org/10.1152/physrev.00023.2014 PMID: 26109341

8. O'Doherty J. et al., 2004. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science (New York, N.Y.)*, 304(5669), pp.452–454. https://doi.org/10.1126/science.1094285 PMID: 15087550

9. Rangel A., Camerer C. & Montague P.R., 2008. A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), pp.545–556. Available at: http://www.nature.com/doifinder/10.1038/nrn2357. PMID: 18545266

10. Tobler P.N., Fiorillo C.D. & Schultz W., 2005. Adaptive coding of reward value by dopamine neurons. *Science (New York, N.Y.)*, 307(5715), pp.1642–1645. https://doi.org/10.1126/science.1105370 PMID: 15761155

11. Stauffer W.R., Lak A. & Schultz W., 2014. Dopamine Reward Prediction Error Responses Reflect Marginal Utility. *Current Biology*, 24(21), pp.2491–2500. https://doi.org/10.1016/j.cub.2014.08.064 PMID: 25283778

12. Schultz W., 2005. Behavioral Theories and the Neurophysiology of Reward. dx.doi.org.ep.fjernadgang. *kb.dk*, 57(1), pp.87–115.

13. Cone J.J. et al., 2016. Physiological state gates acquisition and expression of mesolimbic reward prediction signals. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (7), pp.1943–1948. https://doi.org/10.1073/pnas.1519643113 PMID: 26831116

14. Ferrario C.R. et al., 2016. Homeostasis Meets Motivation in the Battle to Control Food Intake. *Journal of Neuroscience*, 36(45), pp.11469–11481. https://doi.org/10.1523/JNEUROSCI.2338-16.2016 PMID: 27911750

15. Palmiter R.D., 2007. Is dopamine a physiologically relevant mediator of feeding behavior? *Trends in Neurosciences*, 30(8), pp.375–381. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0166223607001336. https://doi.org/10.1016/j.tins.2007.06.004 PMID: 17604133

16. Mietlicki-Baase E.G. et al., 2015. Amylin modulates the mesolimbic dopamine system to control energy balance. *Neuropsychopharmacology*, 40(2), pp.372–385. https://doi.org/10.1038/npp.2014.180 PMID: 25035079

17. Sternson S.M. & Eiselt A.-K., 2017. Three Pillars for the Neural Control of Appetite. 79(1), pp.401–423. Available at: http://www.annualreviews.org/doi/10.1146/annurev-physiol-021115-104948.

18. Aponte Y., Atasoy D. & Sternson S.M., 2011. AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nature Neuroscience*, 14(3), pp.351–355. https://doi.org/10.1038/nn.2739 PMID: 21209617

19. Dietrich M.O. et al., 2015. Hypothalamic Agrp neurons drive stereotypic behaviors beyond feeding. *Cell*, 160(6), pp.1222–1232. https://doi.org/10.1016/j.cell.2015.02.024 PMID: 25748653

20. Luquet S. et al., 2005. NPY/AgRP neurons are essential for feeding in adult mice but can be ablated in neonates. *Science (New York, N.Y.)*, 310(5748), pp.683–685. https://doi.org/10.1126/science.1115524 PMID: 16254186

21. Cansell C. et al., 2012. Arcuate AgRP neurons and the regulation of energy balance. *Frontiers in Endocrinology*, 3, p.169. https://doi.org/10.3389/fendo.2012.00169 PMID: 23293630

22. Mandelblat-Cerf Y. et al., 2015. Arcuate hypothalamic AgRP and putative POMC neurons show opposite changes in spiking across multiple timescales. *eLife*, 4, p.351. Available at: http://elifesciences.org/lookup/doi/10.7554/eLife.07122 PMID: 26159614

23. Nieh E.H. et al. 2015. Decoding Neural Circuits that Control Compulsive Sucrose Seeking. *Cell*, 160, 528–541 https://doi.org/10.1016/j.cell.2015.01.003 PMID: 25635460

24. Campos C.A. et al., 2016. Parabrachial CGRP Neurons Control Meal Termination. *Cell metabolism*, 23 (5), pp.811–820. https://doi.org/10.1016/j.cmet.2016.04.006 PMID: 27166945

25. Söderpalm A.H.V. & Berridge K.C., 2000. The hedonic impact and intake of food are increased by midazolam microinjection in the parabrachial nucleus. *Brain Research*, 877(2), pp.288–297. https://doi.org/10.1016/s0006-8993(00)02691-3 PMID: 10986343

26. Qunli et al., 2014. The Temporal Pattern of cfos Activation in Hypothalamic, Cortical, and Brainstem Nuclei in Response to Fasting and Refeeding in Male Mice. *Endocrinology*, 155(3), pp.840–853. https://doi.org/10.1210/en.2013-1831 PMID: 24424063

27. Alhadeff A.L., Baird J.-P., et al., 2014. Glucagon-like Peptide-1 receptor signaling in the lateral parabrachial nucleus contributes to the control of food intake and motivation to feed. *Neuropsychopharmacology*, 39(9), pp.2233–2243. https://doi.org/10.1038/npp.2014.74 PMID: 24681814

28. Alhadeff A.L., Hayes M.R. & Grill H.J., 2014. Leptin receptor signaling in the lateral parabrachial nucleus contributes to the control of food intake. *American journal of physiology. Regulatory, integrative and comparative physiology*, 307(11), pp.R1338–R1344. https://doi.org/10.1152/ajpregu.00329.2014 PMID: 25298514

29. Friston K.J. et al., 1996. Movement-related effects in fMRI time-series. *Magnetic resonance in medicine*, 35(3), pp.346–355. https://doi.org/10.1002/mrm.1910350312 PMID: 8699946

30. Glover GH, Li TQ, Ress D. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magnetic Resonance in Medicine. 2000; 44:162–167. https://doi.org/10.1002/1522-2594(200007)44:1<162::aid-mrm23>3.0.co;2-e PMID: 10893535

31. Lancaster J.L. et al., 2000. Automated Talairach atlas labels for functional brain mapping. *Human brain mapping*, 10(3), pp.120–131. Available at: http://onlinelibrary.wiley.com/doi/10.1002/1097-0193. PMID: 10912591

32. Lancaster J.L. et al., 1997. *The Talairach Daemon, a database server for Talairach atlas labels*, Neuroimage.

33. Mai JK, Majtanik M, Paxinos G. Atlas of the human brain. Academic Press; 2015 Dec 2.

34. Sutton R.S. & Barto A.G., 1998. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks*, 9(5), pp.1054–1054.

35. D'Ardenne K. et al., 2008. BOLD Responses Reflecting Dopaminergic Signals in the Human Ventral Tegmental Area. *Science (New York, N.Y.)*, 319(5867), pp.1264–1267. https://doi.org/10.1126/science.1150605 PMID: 18309087

36. Page K.A. et al., 2011. Circulating glucose levels modulate neural control of desire for high-calorie foods in humans. *Journal of Clinical Investigation*, 121(10), pp.4161–4169. https://doi.org/10.1172/JCI57873 PMID: 21926468

37. Eshel N. et al., 2016. Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*, 19(3), pp.479–486. https://doi.org/10.1038/nn.4239 PMID: 26854803

38. Norgren R., 1978. Projections from the nucleus of the solitary tract in the rat. *Neuroscience*, 3 (2), pp.207–218. https://doi.org/10.1016/0306-4522(78)90102-1 PMID: 733004

39. Loewy A.D., 1998. The Lower Brainstem and Bodily Homeostasis. *Trends in Neurosciences*, 21 (6), pp.270–271.

40. Miller R.L., Stein M.K. & Loewy A.D., 2011. Serotonergic inputs to FoxP2 neurons of the pre-locus coeruleus and parabrachial nuclei that project to the ventral tegmental area. *Neuroscience*, 193, pp.229–240. https://doi.org/10.1016/j.neuroscience.2011.07.008 PMID: 21784133

41. Norgren R., 1976. Taste pathways to hypothalamus and amygdala. *The Journal of comparative neurology*, 166(1), pp.17–30. https://doi.org/10.1002/cne.901660103 PMID: 1262547

42. Li C.-S. et al., 2012. Descending projections from the nucleus accumbens shell suppress activity of taste-responsive neurons in the hamster parabrachial nuclei. *Journal of Neurophysiology*, 108 (5), pp.1288–1298. https://doi.org/10.1152/jn.00121.2012 PMID: 22696536

43. Zhang C., Kang Y. & Lundy R.F., 2011. Terminal field specificity of forebrain efferent axons to the pontine parabrachial nucleus and medullary reticular formation. *Brain Research*, 1368(2), pp.108–118. https://doi.org/10.1016/j.brainres.2010.10.086 PMID: 21040715

44. Söderpalm A.H.V. & Berridge K.C., 2000. The hedonic impact and intake of food are increased by midazolam microinjection in the parabrachial nucleus. *Brain Research*, 877(2), pp.288–297. https://doi.org/10.1016/s0006-8993(00)02691-3 PMID: 10986343

45. Wu Qi, Boyle M.P.& Palmiter R.D., 2009. Loss of GABAergic signaling by AgRP neurons to the parabrachial nucleus leads to starvation. *Cell*, 137(7), pp.1225–1234. https://doi.org/10.1016/j.cell.2009.04.022 PMID: 19563755

46. De Oliveira L.B. et al., 2011. Baclofen into the lateral parabrachial nucleus induces hypertonic sodium chloride and sucrose intake in rats. *Neuroscience*, 183, pp.160–170. https://doi.org/10.1016/j.neuroscience.2011.02.019 PMID: 21419196

47. DiPatrizio N.V. & Simansky K.J., 2008. Activating parabrachial cannabinoid CB1 receptors selectively stimulates feeding of palatable foods in rats. *The Journal of neuroscience*: *the official journal of the Society for Neuroscience*, 28(39), pp.9702–9709.

48. Wilson J.D. et al., 2003. An orexigenic role for μ-opioid receptors in the lateral parabrachial nucleus. *American journal of physiology. Regulatory, integrative and comparative physiology*, 285(5), pp. R1055–R1065. https://doi.org/10.1152/ajpregu.00108.2003 PMID: 14557237

49. Chaijale N.N., Aloyo V.J. & Simansky K.J., 2013. The stereoisomer (+)-naloxone potentiates G-protein coupling and feeding associated with stimulation of mu opioid receptors in the parabrachial nucleus. *Journal of Psychopharmacology*, 27(3), pp.302–311. https://doi.org/10.1177/0269881112472561 PMID: 23348755

50. Skibicka K.P. & Grill H.J., 2009. Hypothalamic and hindbrain melanocortin receptors contribute to the feeding, thermogenic, and cardiovascular action of melanocortins. *Endocrinology*, 150(12), pp.5351–5361. https://doi.org/10.1210/en.2009-0804 PMID: 19854868

51. Konner A.C et al. 2011. Role for insulin signaling in catecholaminergic neurons in control of energy homeostasis.*Cell Metabolism*, 3(6):720–8 https://doi.org/10.1016/j.cmet.2011.03.021 PMID: 21641553

52. Sun X. et al., 2014. The neural signature of satiation is associated with ghrelin response and triglyceride metabolism. *Physiology & Behavior*, 136, pp.63–73. https://doi.org/10.1016/j.physbeh.2014.04.017 PMID: 24732416

53. Malik S. et al., 2008. Ghrelin modulates brain activity in areas that control appetitive behavior. *Cell metabolism*, 7(5), pp.400–409. https://doi.org/10.1016/j.cmet.2008.03.007 PMID: 18460331

54. Kroemer N.B. et al., 2013. Fasting levels of ghrelin covary with the brain response to food pictures. *Addiction Biology*, 18(5), pp.855–862. https://doi.org/10.1111/j.1369-1600.2012.00489.x PMID: 22974271

55. Domingos A.I. et al., 2011. Leptin regulates the reward value of nutrient. *Nature Neuroscience*, 14 (12), pp.1562–1568. Available at: http://www.nature.com/doifinder/10.1038/nn.2977. PMID: 22081158

56. Figlewicz D.P. et al., 2003. Expression of receptors for insulin and leptin in the ventral tegmental area/ substantia nigra (VTA/SN) of the rat. *Brain Research*, 964(1), pp.107–115. https://doi.org/10.1016/ s0006-8993(02)04087-8 PMID: 12573518

57. Fulton S., 2000. Modulation of Brain Reward Circuitry by Leptin. 287(5450), pp.125–128. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.287.5450.125. PMID: 10615045

58. Takahashi K.A. & Cone R.D., 2005. Fasting induces a large, leptin-dependent increase in the intrinsic action potential frequency of orexigenic arcuate nucleus neuropeptide Y/Agouti-related protein neurons. *Endocrinology*, 146(3), pp.1043–1047. https://doi.org/10.1210/en.2004-1397 PMID: 15591135

59. Berridge KC. Modulation of taste affect by hunger, caloric satiety, and sensory-specific satiety in the rat. Appetite. 1991 Apr; 16(2):103–20. https://doi.org/10.1016/0195-6663(91)90036-r PMID: 2064389.

60. Meder D. et al., 2017. Simultaneous representation of a spectrum of dynamically changing value estimates during decision making. *Nature Communications*, 8(1), p.1942. https://doi.org/10.1038/s41467-017-02169-w PMID: 29208968

61. Alexander W. H., & Brown J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Publishing Group*, 14(10), 1338–1344. https://doi.org/10.1038/nn.2921 PMID: 21926982

62. Silvetti M., Alexander W., Verguts T., & Brown J. W. (2014). From conflict management to reward-based decision making: Actors and critics in primate medial frontal cortex. *Neuroscience and Biobehavioral Reviews*, 46 Pt 1, 44–57. https://doi.org/10.1016/j.neubiorev.2013.11.003 PMID: 24239852

63. Verguts T., Vassena E., & Silvetti M. (2015). Adaptive effort investment in cognitive and physical tasks: A neurocomputational model. *Frontiers in Behavioral Neuroscience*, 9. https://doi.org/10.3389/fnbeh.2015.00057 PMID: 25805978