Reward signalling in brainstem nuclei under glycemie flux

Tobias Morville^{1*}, Kristoffer Madsen³, Hartwig R. Siebner^{1,2}, Oliver J. Hulme

¹Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Kettegaard Allé 30, Hvidovre, 2650, Denmark ²DTU Compute, Department of Informatics and Mathematical Modeling, Technical University of Denmark, Copenhagen, Denmark ³Department of Neurology, Copenhagen University Hospital Bispebjerg, Copenhagen, 2400, Denmark.*
Corresponding author: tobiasm@drcmr.dk

Phasic dopamine release from mid-brain dopaminergic neurons signals 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 physiological serum glucose fluctuations. In the parabrachial nuclei, variations in regional activity coding positive RPEs scaled positively with serum glucose for 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 show that RPE signals in key brainstem structures are modulated by homeostatic trajectories of naturally occurring glycemie flux, revealing a tight interplay between homeostatic state and the neural encoding of primary reward in the human brain.

Keywords. Homeostasis, Reward prediction errors, Dopamine, Midbrain, Parabrachial nucleus.

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

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 (Pompilio et al. 2006; Cabanac 1971). Compatible with this notion, deprivation-induced hypoglycaemia increases willingness to work for food in rats (Sclafani et al. 1970), in humans (Pelchat 2009), as well as the subjectively reported pleasure (Cabanac 1971). Catecholamine dopamine is a neurotransmitter that plays a key role in signalling reward (Haber & Knutson 2010) and is involved in behavioural reinforcement, learning and motivation (Berridge 2006; Schultz et al. 1997). Via meso-cortical and mesolimbic dopaminergic projections, synaptic dopamine release modulates the plasticity of cortico-striatal networks and hereby sculpts behavioural policies according to their reward contingencies (Haber & Knutson 2010; Schultz 2015). Patterns of phasic dopaminergic firing have been demonstrated to follow closely the principles of reinforcement learning, encoding the errors in the prediction of reward (O'Doherty et al. 2004; Schultz et al. 1997; Rangel et al. 2008; Tobler et al. 2005). Reward prediction error (RPE) signals are 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 (Stauffer et al. 2014; Schultz 2005; Schultz 2015).

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. 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 (J. J. Cone et al. 2016).

54 There is now growing evidence for a multifaceted interface between dopamine mediated reward-
55 signalling and the systems underpinning energy homeostasis. Firstly, dopamine neurons in the ventral
56 tegmental area (VTA) express a suite of receptors targeted by energy-reporting hormones ghrelin,
57 insulin, amylin, leptin and Glucagon Like Peptide 1 (GLP-1, Ferrario et al. 2016; Palmiter 2007). This
58 provides numerous degrees of freedom for flexibly interfacing between homeostatic state and reward
59 signalling. Although hormonal modulations of phasic dopamine are yet to be fully scrutinised, there is
60 emerging evidence that circulating factors do indeed modulate its magnitude. For instance, amylin, a
61 hormone co-released with insulin, acts on the VTA to reduce phasic dopamine release in its mesolimbic
62 projection sites (Mietlicki-Baase et al. 2015). In terms of neuronal input, there are many such
63 opportunities for the appetitive control of dopamine mediated signalling.

64 Appetitive control can be delineated into three interacting valuation systems (Sternson & Eisele 2017).
65 The first system generates a negative valence signal which involves activity of the Agouti-related peptide
66 (AgRP) neurons of the arcuate nucleus of the hypothalamus (ARC). Activity of ARC_{AgRP} neurons reports
67 on energy deficits, inhibits energy expenditure, and regulates glucose metabolism (e.g. Aponte et al.
68 2011; Dietrich et al. 2015; Luquet et al. 2005; Cansell et al. 2012). ARC neurons that contain peptide
69 products of pro-opiomelanocortin (POMC) form an opponent code compared with ARC_{AgRP} neurons. The
70 balance between the two neuronal ARC sub-populations putatively encodes the value of near-term
71 energetic states, becoming rapidly modulated just prior to food consumption (Mandelblat-Cerf et al.
72 2015). The second system codes positive valence signals and consists of circuits involving the lateral
73 hypothalamus (LH). It is linked to positively reinforcing consummatory behaviours via its dopaminergic
74 projections, assumed to trigger positive feedback to keep consumption going during feeding bouts. The
75 third valuation system involves calcitonin gene-related protein (CGRP)-expressing neurons in the (PBN)
76 that potently suppress eating when activated, but do not increase food intake when inhibited. PBN_{CGRP}
77 neurons are activated by signals associated with food intake, and they provide a signal of satiety that
78 has negative valence when strongly activated (Campos et al. 2016). The PBN has been characterised as a
79 hedonic hotspot, the modulation of which by either GABA or Benzodiazepines potently modulates

80 experienced reward (Söderpalm & Berridge 2000); ARC_{AgRP} neurons GABA-ergically inhibit PBN neurons,
 81 thus stimuli predicting glucose consumption should inhibit ARC_{AgRP}, releasing the PBN from inhibition
 82 (Qunli Wu et al. 2014). Further, hormones related to hunger and feeding (GLP-1 & leptin) modulate PBN
 83 activity and subsequent behaviour (e.g. Alhadeff, Baird, et al. 2014; Alhadeff, Hayes, et al. 2015). Of
 84 note, these three valuation systems all project to and modulate the dopaminergic neurons in the ventral
 85 tegmental area(VTA_{DA}). The interface between these hypothalamic-brainstem networks and the VTA_{DA},
 86 is arguably the most important interface for mediating the dialogue between energy homeostasis and
 87 value computation.

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

96

97

98 **Methods**

99 **Subjects.** Five healthy, normal-weight subjects in the age range 23 to 29, participated in the study.

100 Exclusion criteria were: $20 > \text{BMI} > 25$; $18 > \text{Age} > 32$ yrs; any metabolic or endocrine diseases or
 101 gastrointestinal disorder; any known medication that might interfere with the study; claustrophobia;
 102 and any metal implants or devices that could not be removed. Informed consent was obtained from all
 103 subjects as approved by the Regional Ethics Committee of Region Hovedstaden (protocol H-4-2013-100)
 104 and in accordance with the declaration of Helsinki.

105 **Experimental procedure.** The experimental design constituted a single-blinded, randomised control trial,
 106 with repeated measures crossover-design. On four separate days, subjects fasted for a minimum of ten
 107 hours before testing. At the beginning of an experimental session, subjects ingested either a hi-glucose
 108 (75 g, 300 kcal) or lo-glucose preload (10 g, 40 kcal) diluted to 100 ml with a 0-kcal lemon juice masking
 109 the taste of the glucose. Both preloads were anecdotally reported by independent samplers to be highly
 110 palatable.

111 **Experimental task.** After consuming the preload, participants engaged in a simple pavlovian cue-
 112 conditioning task. The colour of the fixation cross cued both the onset of each trial ($\text{Cue}_{\text{onset}}$), as well as
 113 stochastically predicting glucose delivery (Fig. 1a), with one colour signalling a high probability of
 114 glucose delivery (Cue_{high}), and another signalling a low probability (Cue_{low}). 10-15 seconds after delivery
 115 of oral stimulus, a purple cross signalled that subjects were allowed to swallow. All probabilities and
 116 contingencies were implicitly revealed only through experience in the scanner, and all were stationary
 117 over all test days. The mapping between colour and outcome probabilities was counterbalanced across
 118 subjects, while mapping was stationary within and between sessions. Participants went through ~82
 119 trials [82 ± 1.5 SEM] each day giving ~328 trials per subject. Serum glucose measurements were attained
 120 immediately before and 20 minutes after ingestion, using a Contour® Next glucose meter (Fig. 1b). As
 121 expected, prior to ingestion (t_0) there was no significant difference between hi- or lo-glucose days [$4.6 \pm$

0.4 SEM], whereas twenty minutes after ingestion (t20) there was [lo-glucose mean = 4.8, hi-glucose mean = 6.9].

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. 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 through a unified segmentation procedure (Ashburner & Friston 2005). The realigned images were spatially normalised to the standard ICBM space template of European brains (Mazziotta et al. 1995), with a resampled voxel size of 3 mm.

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_{onset}, Cue_{high} and Cue_{low} as well as outcome onsets for Outcome_{gluc} & Outcome_{water}. Fig. 2a illustrates how the expectation

148 value of glucose volume delivered evolves over time as a function of the cues observed. We specified
 149 RPE contrasts which were formulated by linear combinations of regressors, weighted as a function of
 150 the RPE values from the temporal-difference learning algorithm (Sutton & Barto 1998). As subjects learn
 151 the contingencies between visual stimuli (colour crosses) and outcome (juice or water) the RPE converge
 152 to the expected (average) value of the glucose content. This is conditioned on the cues that have been
 153 experienced and is illustrated in Fig. 2b. In this paradigm, there was no behaviour to fit a learning rate
 154 parameter to, so the steady-state values of the RPE was used instead. In effect, this assumes that the
 155 subjects learned the contingencies from the beginning. The effect of serum glucose on RPE was
 156 modelled by multiplying the resulting RPE by subject specific demeaned serum glucose (state hence),
 157 linearly interpolated between out-of-scan measurements. We specified the following contrasts of
 158 interest: RPE_{pos} , RPE_{neg} with their state-weighted counterparts $RPE_{pos*state}$, $RPE_{neg*state}$ computed as first
 159 order parametric modulators.

160 **fMRI analysis.** After model specification, the sessions were concatenated using the function
 161 `spm_fmri_concatenation` (SPM 12) for each subject and a standard first-level fixed effects models was
 162 run over all subjects. All variables of interest were convolved with the canonical hemodynamic response
 163 function and fitted to the data using the specified GLM. The temporal evolution of cues and outcomes
 164 were modelled as separate conditions, each with state as parametric modulators. Regressors of no
 165 interest included a discrete cosine transform based 1/128 Hz cut-off frequency high-pass filter, rigid
 166 body realignment parameters using a 24 parameter Volterra expansion (Friston et al. 1996) and
 167 physiological noise from heart rate and respiration using the RETROICOR method {Glover:2000wy}. We
 168 specified the striatum (caudate, putamen and nucleus accumbens), brainstem (pons, ventral tegmental
 169 area and substantia nigra) and hypothalamus as Regions of interest (ROI). These ROIs were determined
 170 on basis of the literature describing dopamine projections from midbrain to the striatum and its role in
 171 regulating behaviour as a function of reward. The pons was selected to accommodate the literature
 172 described above, which sets certain nuclei within the pons as important homeostatic modulators. All
 173 ROI were defined with the WFU pick atlas (Lancaster et al. 2000; Lancaster et al. 1997). All initial first-

174 level analysis was performed as whole-brain uncorrected at $p < 0.001$. Significant clusters in regions of
175 interest (ROI) are all reported as small-volume corrected with a family-wise threshold of $p < 0.05$ at
176 cluster level (abbreviated SVC FWE), unless otherwise stated.
177

Results

Cue induced brain activity. The “trial onset” cue signalled the expected value of glucose reward for the whole trial (Stauffer et al. 2014) and triggered an increase in activity in VTA bilaterally (Fig. 3a). Cue-induced VTA activation is consistent with existing evidence of VTA signalling RPE (e.g. D'Ardenne et al. 2008; Page et al. 2011; Eshel et al. 2016). The onset cue also led to deactivation of postcentral gyrus (primary somatosensory cortex), mediodorsal thalamus, and likewise in the striatum [whole brain, uncorrected $p < 0.001$] (not shown). In several brain regions, regional task-related activity changed in proportion with the magnitude of positive-going (i.e. better-than-expected) RPEs or negative-going (i.e. worse-than-expected) RPEs. 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. 3b). Conversely, task related activity reflecting RPE_{neg} , formalized an RPE-weighted linear combination of Cue_{low} and $Outcome_{water}$, was located in the caudate nucleus bilaterally (Fig. 3c), the medial dorsal thalamic nucleus, and insula.

Modulation of task-related brain activity by homeostatic glycemic state. We were interested to identify changes in RPE processing 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 positive RPEs by the glycemic state dynamics (Fig 4a). Higher levels of serum glucose amplified the response to RPE_{pos} in the PBN region (Fig. 4b). The main effect of $RPE_{neg} \times state$, 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 the “onset cue” were also not altered by glycemic state dynamics.

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 over $RPE_{pos} \times state$ and $RPE_{pos} \times state$) 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. 5a). 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), the glucose state modulated the RPE_{neg} signal in VTA and SN more so when glucose levels were ascending than descending.

Discussion

We studied five individuals repeatedly with fMRI under increasing or decreasing levels of glucose, while participants performed a simple cue-conditioning task involving the probabilistic delivery of glucose or water in a single trial. Reward prediction error signalling in the parabrachial nuclei scaled positively with serum glucose levels during ascending and descending glycemic trajectories. The VTA and SN became more sensitive to negative RPEs for ascending compared to descending glycemic 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 (Norgren 1978; Loewy 1998), VTA (Miller et al. 2011), hypothalamus (Norgren 1976; Loewy 1998) and the nucleus accumbens (Li et al. 2012). 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 (Zhang et al. 2011; Loewy 1998). The PBN is known to be a potent site of reward modulation and subsequent behaviour in rodents. Microinjection of benzodiazepines (Söderpalm & Berridge 2000; Qi Wu et al. 2009; De Oliveira et al. 2011), endocannabinoids (DiPatrizio & Simansky 2008), opioids (Wilson et al. 2003; Chaijale et al. 2013) and melanocortin agonists (Skibicka & Grill 2009) into the PBN, all evoke hyperphagia. To the best of our knowledge, the involvement of PBN in context of hedonics and reward signalling in the human brain remains yet to be charted. Here we provide novel evidence that PBN activity generates a gluco-sensory scaled positive RPE signal which is time-locked to both the sensory cues predicting glucose, as well as glucose outcomes.

Unlike the state modulation of serum glucose trajectories on the RPE_{pos} signal, we found no general state modulation of RPE_{neg} signalling, expressed during ascending and descending glycemic trajectories. Here the modulatory effect of the glycemic 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 vs. descending glycemic paths. In our experiment, the

238 ascending glucose trajectory resulted from a low-glucose preload with the subsequent increase over
239 time likely occurring by virtue of the continual ingestion of glucose throughout the paradigm (Fig. 1b). In
240 the ascending condition, the neural response to RPE_{neg} is attenuated at lower levels of serum glucose,
241 while it becomes amplified by the transition to higher serum glucose. Given that dopaminergic neurons
242 of the VTA and SN are directly inhibited by insulin (Palmiter 2007), it is likely that the insulin release
243 following hi-glucose preload was highest at the start of the paradigm, decreasing over time, and thus
244 resulting in a gradual decrease in inhibition. The difference in RPE_{neg} in its state modulation between
245 ascending and descending may therefore be attributed to differential dynamics of insulin secretion (see
246 (Sun et al. 2014), though other hormones such as ghrelin (Malik et al. 2008; Kroemer et al. 2013; Sun et
247 al. 2014) or leptin (Domingos et al. 2011; Figlewicz et al. 2003; Fulton 2000; Alhadeff, Hayes, et al. 2014;
248 Takahashi & R. D. Cone 2005) may play a role.

249 Our finding that the VTA and SN responses are linked to RPE_{neg} may appear counterintuitive, given that
250 these midbrain regions are typically associated with BOLD responses signalling positive-going RPEs. This
251 is assumed to be by virtue of the fact that a greater range of firing rates can be devoted to the better-
252 than-expected range, signalled by above baseline firing. This is contrasted to the worse-than-expected
253 range, which can only be signalled by a decrease from an already low baseline frequency. It is
254 conceivable that what we are asserting as being RPE_{neg} is in fact a positive RPE resulting from the
255 gradual avoidance of glucose, which increases in magnitude with increasing levels of serum glucose as
256 reported in humans (Cabanac 1971) and rats (Berridge 1991). Thus, as the experimental paradigm
257 continues, especially under the conditions of glucose preload, serum glucose increases, and this may
258 change the valence of the outcome, switching the affective connotation of glucose from palatable to
259 aversive.

260 As detailed in the introduction, little is known about the principles how the interface between
261 dopaminergic RPE signalling and energy homeostasis is implemented in the human brain. While there
262 are many means by which circulating factors can modulate activity in the VTA and SN, the mechanisms
263 by which this is mediated cannot be revealed without wider hormonal assays. Contemporaneous

264 hormonal sampling, as well as continuous glucose monitoring in the scanner will prove an important
265 step in revealing these hidden mediating factors.

266 From a theoretical perspective, results as presented here could be predicted by any model that invokes
267 the notion of RPEs in service of homeostatic regulation. For example, models inspired by optimal control
268 theory such as Homeostatic Reinforcement Learning ((Keramati & Gutkin 2014) or MOTIVATOR theory
269 (Draniias et al. 2008). Alternatively, under the theory of Active Inference, phasic dopamine is recast as
270 encoding updates to the precision assigned to the behavioral policies that lead to desired outcomes,
271 that (in this context) remediate long-run homeostatic error (Schwartenbeck et al. 2015).

272 There are several technical limitations that should be noted in discussing this experiment. Though
273 relatively high volumes of functional data (150 minutes per subject) were acquired in each subject, the
274 total number of subjects was low. Future work will expand this paradigm with a larger group of to afford
275 random effects modelling, and thus generalisation to the population sampled from. In contrast to our
276 hypotheses, we found no modulatory effect of hypothalamic nuclei on RPE signalling. We would like to
277 stress that the current imaging protocols and field-strength (3T) were not optimal to dissociate neural
278 activity in the hypothalamic nuclei. Due to the proximity of air sinuses adjacent to the hypothalamus and
279 the effective resolution available, the present study most likely had insufficient sensitivity to capture
280 activity in hypothalamic regions of interest. Future work at higher field strengths (7T) may overcome
281 these limitations. Finally, the cue-conditioning employed in this study was passive. Hence, subjects
282 produced no overt choice behaviour against which to fit learning rate parameters for the RPE model,
283 instead we relied on the asymptote values for the RPE signals. The problem of modelling RPEs in the
284 absence of choice behaviour, motivates fitting learning rate parameters directly to brain data, a
285 computational imaging approach that future work will exploit (Meder et al. 2017)

286 In conclusion, we exploited a simple paradigm, capable of eliciting RPEs under differential glycemic
287 trajectories, to identify brain stem structures that show a modulation of RPE signalling dependning on
288 the glycemic homeostatic state. We found that the PBN signals a positive-going reward prediction that is
289 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 glycemic slope. Together the results show that RPE signals in key brainstem structures are modulated by homeostatic trajectories inherent in naturally occurring glycemic flux, revealing a tight interplay between homeostatic state and the neural processing of primary reward in the human brain.

Acknowledgements. We thank Mehdi Keramati and Boris Gutkin for several helpful discussions. This work was supported by the following funders: H.R.S (Lundbeck Foundation Grant of Excellence “ContAct” ref: R59 A5399 ; Novo Nordisk Foundation Interdisciplinary Synergy Programme Grant “BASICS” ref: NNF14OC0011413) O.J.H (Lundbeck Foundation, ref: R140-2013-13057; Danish Research Council ref: 12-126925) T.M (Lundbeck Foundation ref: R140-2013-13057).

Author contributions. T.M. & O.H conceived of the study, and all authors designed the study. T.M collected the data and visualised the results. All authors contributed to analysis and interpretation of data, and to the writing of the manuscript.

Declaration of interests. The authors declare no competing interests. H.R.S. has received honoraria as speaker from Genzyme, Denmark and as senior editor of NeuroImage from Elsevier Publishers, Amsterdam, The Netherlands. H.R.S. has received a research fund from Biogen-idec, Denmark.

Figure 1 | Experimental design and glucose trajectories. **a**, At Cue_{onset} participants are presented with a neutrally coloured fixation cross (grey) for 1-3s after which either Cue_{high} (here illustrated as blue cross) or Cue_{low} (brown cross) is presented with 0.5 probability each. Each cue signalled either high probability (0.8) of glucose delivery (0.4 ml) and low probability (0.2) of water delivery (0.4 ml), or the converse probabilities, respectively. A fixed duration after presentation of either cue (2.5 seconds), the liquid stimuli was delivered over 2.5 seconds. This was followed by 10-15s jitter and a cue for swallowing (purple) that lasted for 5s. after which a new trial initiated with the onset of the neutral cross. **b**, Plot of measured serum glucose (y-axis) over each session (x-axis) that lasted approximately 65 minutes. Left shows the lo-glucose preload sessions, while right shows hi-glucose preload sessions. The grey shading indicates the duration of the fMRI acquisition of a single session.

Figure 2 | Expectations and fitted responses for reward prediction. **a**, Line 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. The dashed transparent lines illustrate when cues signalled high (or low) outcomes truthfully (blue high, orange low). The solid line illustrates when cues were paired with low-probability outcomes (green for high to low & purple from low to high). Note that reward expectations are updated three times: 1) at the onset of the Cue_{trial}, 2) at the onset of Cue_{high} or Cue_{low}, and 3) at the onset of Outcome_{gluc} or Outcome_{water}. **b**, Illustrates simulated BOLD responses to RPE signals resulting from the updated reward expectations shown in Fig. 2a, generated by convolving the canonical hemodynamic response function with the RPE stick functions evoked by changes to the reward expectations with $\beta = 1$.

Figure 3 | Statistical parametric maps of main effects of trial onset and RPE and fitted response. **a**, Main effect of trial onset cue, which reflects an RPE following the mean reward expectation for the whole trial, revealed activity in VTA bilaterally ($\beta = 2.77$) (R: [4 -17 -20] and L: [-8 -17 -20], FWE SVC). Further this revealed deactivation of precentral gyrus (primary somatosensory cortex), mediadorsal thalamus, and striatum (FWE whole brain, not shown). **b**, The main effect of RPE_{pos} revealed activity in left lateral caudate [$\beta = 1.21$; coordinates -8 4 7; FWE SVC]. **c**, The main effect of RPE_{neg} revealed bilateral activity in caudate (L: -11 -2 13; R: 10 7 1; $\beta = 12.2$) medial dorsal thalamic nucleus [7, -2, 22], and lateral insula [43, -2, -17] (all FWE). 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.

Figure 4 | Statistical parametric maps of $RPE_{pos*state}$ and fitted responses over varying glucose state. **a**, The main effect of $RPE_{pos*state}$ revealed bilateral activity in the PBN [-2 -29 -26; FWE SVC]. **b**, Fitted response ($\beta = 1.66$) of the local maxima of PBN cluster (7 voxels) to the four possible trajectories that $RPE_{pos*state}$, yield (see Fig. 2b) modulated by serum glucose state. The furthest trajectory on the y-axis are all four trajectories superimposed on to each other and is the signal which statistics is shown above in Fig. 2a.

Figure 5 | 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 [$\pm 12, -22, -10$] and VTA [$0, -15, -6$] when subtracting the effect of descending from the ascending glucose state [FWE SVC]. **b)** Fitted response ($\beta = 0.34$) 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.

References

- 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.
- 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.
- 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.
- Ashburner, J. & Friston, K., 2005. Unified segmentation. *NeuroImage*, 26(3), pp.839–851.
- Berridge, K.C., 2006. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, 191(3), pp.391–431.
- Cabanac, M., 1971. Physiological Role of Pleasure. *Science*, 173(4002), pp.1103–1107.
- Campos, C.A. et al., 2016. Parabrachial CGRP Neurons Control Meal Termination. *Cell metabolism*, 23(5), pp.811–820.
- Cansell, C. et al., 2012. Arcuate AgRP neurons and the regulation of energy balance. *Frontiers in Endocrinology*, 3, p.169.
- 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.
- 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.
- 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.
- 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.
- Dietrich, M.O. et al., 2015. Hypothalamic Agrp neurons drive stereotypic behaviors beyond feeding. *Cell*, 160(6), pp.1222–1232.
- 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.
- 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/doi/10.1038/nn.2977>.
- Dranias, M.R., Grossberg, S. & Bullock, D., 2008. Dopaminergic and non-dopaminergic value systems in conditioning and outcome-specific revaluation. *Brain Research*, 1238, pp.239–287.
- Eshel, N. et al., 2016. Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*, 19(3), pp.479–486.

391 Ferrario, C.R. et al., 2016. Homeostasis Meets Motivation in the Battle to Control Food Intake. *Journal of*
392 *Neuroscience*, 36(45), pp.11469–11481.

393 Figlewicz, D.P. et al., 2003. Expression of receptors for insulin and leptin in the ventral tegmental
394 area/substantia nigra (VTA/SN) of the rat. *Brain Research*, 964(1), pp.107–115.

395 Friston, K.J. et al., 1996. Movement-related effects in fMRI time-series. *Magnetic resonance in medicine*,
396 35(3), pp.346–355.

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

399 Haber, S.N. & Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging.
400 *Neuropsychopharmacology*, 35(1), pp.4–26.

401 Keramati, M. & Gutkin, B., 2014. Homeostatic reinforcement learning for integrating reward collection
402 and physiological stability. *eLife*, 3, p.475.

403 Kroemer, N.B. et al., 2013. Fasting levels of ghrelin covary with the brain response to food pictures.
404 *Addiction Biology*, 18(5), pp.855–862.

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

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

409 Li, C.-S. et al., 2012. Descending projections from the nucleus accumbens shell suppress activity of taste-
410 responsive neurons in the hamster parabrachial nuclei. *Journal of Neurophysiology*, 108(5), pp.1288–
411 1298.

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

414 Luquet, S. et al., 2005. NPY/AgRP neurons are essential for feeding in adult mice but can be ablated in
415 neonates. *Science (New York, N.Y.)*, 310(5748), pp.683–685.

416 Malik, S. et al., 2008. Ghrelin modulates brain activity in areas that control appetitive behavior. *Cell*
417 *metabolism*, 7(5), pp.400–409.

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

421 Mazziotta, J.C. et al., 1995. A Probabilistic Atlas of the Human Brain: Theory and Rationale for Its
422 Development. *NeuroImage*, 2(2), pp.89–101.

423 Meder, D. et al., 2017. Simultaneous representation of a spectrum of dynamically changing value
424 estimates during decision making. *Nature Communications*, 8(1), p.1942.

425 Mietlicki-Baase, E.G. et al., 2015. Amylin modulates the mesolimbic dopamine system to control energy
426 balance. *Neuropsychopharmacology*, 40(2), pp.372–385.

427 Miller, R.L., Stein, M.K. & Loewy, A.D., 2011. Serotonergic inputs to FoxP2 neurons of the pre-locus
428 coeruleus and parabrachial nuclei that project to the ventral tegmental area. *Neuroscience*, 193,
429 pp.229–240.

430 Norgren, R., 1978. Projections from the nucleus of the solitary tract in the rat. *Neuroscience*, 3(2),
431 pp.207–218.

432 Norgren, R., 1976. Taste pathways to hypothalamus and amygdala. *The Journal of comparative*
433 *neurology*, 166(1), pp.17–30.

434 O'Doherty, J. et al., 2004. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning.
435 *Science (New York, N.Y.)*, 304(5669), pp.452–454.

436 Page, K.A. et al., 2011. Circulating glucose levels modulate neural control of desire for high-calorie foods
437 in humans. *Journal of Clinical Investigation*, 121(10), pp.4161–4169.

438 Palmiter, R.D., 2007. Is dopamine a physiologically relevant mediator of feeding behavior? *Trends in*
439 *Neurosciences*, 30(8), pp.375–381. Available at:
440 <http://linkinghub.elsevier.com/retrieve/pii/S0166223607001336>.

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

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

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

448 Schultz, W., 2015. Neuronal Reward and Decision Signals: From Theories to Data. *Physiological*
449 *reviews*, 95(3), pp.853–951.

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

452 Schwartenbeck, P. et al., 2015. The Dopaminergic Midbrain Encodes the Expected Certainty about
453 Desired Outcomes. *Cerebral cortex (New York, N.Y. : 1991)*, 25(10), pp.3434–3445.

454 Skibicka, K.P. & Grill, H.J., 2009. Hypothalamic and hindbrain melanocortin receptors contribute to the
455 feeding, thermogenic, and cardiovascular action of melanocortins. *Endocrinology*, 150(12), pp.5351–
456 5361.

457 Söderpalm, A.H.V. & Berridge, K.C., 2000. The hedonic impact and intake of food are increased by
458 midazolam microinjection in the parabrachial nucleus. *Brain Research*, 877(2), pp.288–297.

459 Stauffer, W.R., Lak, A. & Schultz, W., 2014. Dopamine Reward Prediction Error Responses Reflect
460 Marginal Utility. *Current Biology*, 24(21), pp.2491–2500.

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

463 Sun, X. et al., 2014. The neural signature of satiation is associated with ghrelin response and triglyceride
464 metabolism. *Physiology & Behavior*, 136, pp.63–73.

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

467 Takahashi, K.A. & Cone, R.D., 2005. Fasting induces a large, leptin-dependent increase in the intrinsic
468 action potential frequency of orexigenic arcuate nucleus neuropeptide Y/Agouti-related protein
469 neurons. *Endocrinology*, 146(3), pp.1043–1047.

470 Tobler, P.N., Fiorillo, C.D. & Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons.
471 *Science (New York, N.Y.)*, 307(5715), pp.1642–1645.

472 Wilson, J.D. et al., 2003. An orexigenic role for μ -opioid receptors in the lateral parabrachial nucleus.
473 *American journal of physiology. Regulatory, integrative and comparative physiology*, 285(5),
474 pp.R1055–R1065.

475 Wu, Qi, Boyle, M.P. & Palmiter, R.D., 2009. Loss of GABAergic signaling by AgRP neurons to the
476 parabrachial nucleus leads to starvation. *Cell*, 137(7), pp.1225–1234.

477 Wu, Qunli et al., 2014. The Temporal Pattern of cfos Activation in Hypothalamic, Cortical, and
478 Brainstem Nuclei in Response to Fasting and Refeeding in Male Mice. *Endocrinology*, 155(3),
479 pp.840–853.

480 Zhang, C., Kang, Y. & Lundy, R.F., 2011. Terminal field specificity of forebrain efferent axons to the
481 pontine parabrachial nucleus and medullary reticular formation. *Brain Research*, 1368(2), pp.108–
482 118.

483