
White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/145649/

Version: Accepted Version

Article:
Bose, T., Reina, A. and Marshall, J.A.R. (2019) Inhibition and excitation shape activity selection: Effect of oscillations in a decision-making circuit. Neural Computation, 31 (5). pp. 870-896. ISSN: 0899-7667

https://doi.org/10.1162/neco_a_01185

© 2019 Massachusetts Institute of Technology. This is an author-produced version of a paper subsequently published in Neural Computation. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Inhibition and excitation shape activity selection: effect of oscillations in a decision-making circuit

Thomas Bose[1], **Andreagiovanni Reina**[1], **James A.R. Marshall**[1]
[1]Department of Computer Science, University of Sheffield, Sheffield, UK.

Keywords: neuroscience, decision-making, dietary choice, excitation-over-inhibition ratio, behaviour

Abstract

Decision-making is a complex task and its underlying mechanisms that regulate behaviour, such as the implementation of the coupling between physiological states and neural networks, are hard to decipher. To gain more insight into neural computations underlying ongoing binary decision-making tasks, here we consider a neural circuit that guides the feeding behaviour of a hypothetical animal making dietary choices. We adopt an inhibition motif from neural network theory and propose a dynamical system characterized by nonlinear feedback, which links mechanism (the implementation of the neural circuit and its coupling to the animal's nutritional state) and function (improving behavioural performance). A central inhibitory unit influences evidence-integrating excitatory units, which in our terms correspond to motivations competing for selection. We determine the parameter regime where the animal exhibits improved decision-making behaviour, and explain different behavioural outcomes by making the link between accessible states of the nonlinear neural circuit model and decision-making performance. We find that for given deficits in nutritional items the variation of inhibition strength and ratio of excitation and inhibition strengths in the decision circuit allows the animal to enter an oscillatory phase which describes its internal motivational state. Our findings indicate that this oscillatory phase may improve the overall performance of the animal in an ongoing foraging task, and underpin the importance of an integrated functional and mechanistic study of animal activity selection.

1 Introduction

Understanding systems that exhibit complex decision-making behaviour requires the integration of mechanism and function in a combined modelling framework (Gold and Shadlen, 2007; McNamara and Houston, 2009; Houston and McNamara, 1999). Animals making food choices may be fruitfully modelled using this approach (Simpson and

Raubenheimer, 2012; Fawcett et al., 2014). In natural scenarios animals are embedded in uncertain environments (Fawcett et al., 2014) and, for example, may be subject to predation risk (Lima and Dill, 1990). Such external influences affect the decision-making of animals in combination with the momentary nutritional requirements that need to be integrated in a multi-faceted physiologically- and neurobiologically-wired network (Morton et al., 2006; Vong et al., 2011; Atasoy et al., 2012; Wu et al., 2012; Williams and Elmquist, 2012; Rangel, 2013; Essner et al., 2017). The motivation to eat, for example, is related to peripheral signals provided by hormones (Vong et al., 2011; Morton et al., 2006; Williams and Elmquist, 2012) and populations of neurons that are distributed over different brain areas (Morton et al., 2006; Tong et al., 2008; Aponte et al., 2011; Essner et al., 2017). Furthermore, the modulation at corresponding neurobiological synapses involves both excitatory glutamatergic neurotransmitters (Liu et al., 2012; Wu et al., 2012) and inhibitory GABAergic neurotransmitters (Tong et al., 2008; Wu et al., 2009; Vong et al., 2011). However, given the complexity of interactions between homeostatic regulators and the decision-making circuitry in the modulation of dietary choices, unveiling a detailed picture of the underlying computational mechanisms is still at its beginning (Rangel, 2013).

In the present paper, we address the question of how nutritional deficits may induce feeding behaviour, by coupling nutritional state and animal behaviour through a decision-making circuit that implements the underlying neural computation. This circuit, which contains excitatory and inhibitory connections, guides a hypothetical animal making food choices in an ongoing binary decision-making task. Although we follow a coarse-grained approach that neglects biological detail, the decision-making circuit studied here may be considered as an abstraction of the overall neural hardware. To model the decision-making circuit we draw inspiration from mechanistic models which have been used to explain neural activity and behaviour in perceptual (Wang, 2002; Bogacz et al., 2006; Wong and Wang, 2006; Wong et al., 2007; Niyogi and Wong-Lin, 2013) and value-based decision-making tasks (Hunt et al., 2012). Adapting neural mechanisms from one decision paradigm to a different one is consistent with proposals for common decision-making mechanisms utilised in different scenarios (Krajbich et al., 2015).

To examine the utility of our proposed mechanism we embed it in the *geometric framework* – a well-studied nutritional theory able to capture real feeding behaviour of diverse species (Simpson and Raubenheimer, 2012). However, our results may be also applied to other possible scenarios involving conflicting needs, given the simple and minimal assumptions made. In the geometric framework animals perform actions (consume resources with different nutritional contents) to reach a preferred nutrient target, and derive utility according to how close to the target state they get. By reaching their target intake of nutrients, animals may maximise their reproductive success (Mayntz et al., 2005; Altaye et al., 2010; Dussutour et al., 2010; Houston et al., 2011; Jensen et al., 2012; Rho and Lee, 2016). An animal's inner drive (or motivation) to make food choices is influenced by the level of nutrients inside the body (Hinde, 1956; Sibly, 1975; Ludlow, 1976; Houston and Sumida, 1985; McFarland, 1999; Marshall et al., 2015; Bose et al., 2017). Thus, the internal nutritional state acts as excitatory input for the underlying circuit involved in the decision-making process. However, in neurobiological networks that regulate feeding behaviour, both excitatory and inhibitory inputs

are operative (Morton et al., 2006; Vong et al., 2011; Atasoy et al., 2012; Aponte et al., 2011; Wu et al., 2012; Williams and Elmquist, 2012; Liu et al., 2012; Rangel, 2013; Essner et al., 2017), and it has been shown in previous behavioural studies of foraging animals that inhibitory mechanisms between drives that stimulate different activities facilitate improved feeding behaviour (Ludlow, 1976; Marshall et al., 2015).

In our model-based study, we find that excitatory and inhibitory connections in the decision-making circuit regulate the food intake of an animal for given deficit levels. In particular, we show that the modulation of excitation and inhibition strengths can drive the animal through different internal motivational states, which may increase or decrease decision-making performance. Our results demonstrate that oscillatory regimes of the decision-making circuit may lead to improved feeding behaviour. We come to the conclusion that low-performance decision-making of an animal may emerge from suboptimal ratios of excitation and inhibition in the decision-making circuit.

2 Model and methods

2.1 Nutritional deficits and the geometric framework

We consider a model animal with deficits in two different nutrients each of which is exclusively contained in either of two different food types, in the following called food type 1 and food type 2. That is, by consuming food type 1 the animal cannot reduce the nutrient only contained in food type 2, and vice versa. Hence, the animal must take in both food types to satisfy its nutritional needs. We further assume that the nutritional state of the animal is described in the geometric framework (Simpson and Raubenheimer, 2012), which is empirically well-motivated (Chambers et al., 1995; Dussutour and Simpson, 2009; Behmer, 2009; Dussutour et al., 2010; Altaye et al., 2010; Jensen et al., 2012; Arganda et al., 2014; Rho and Lee, 2016; Simpson and Raubenheimer, 2012). In the geometric framework, animals (or social insect colonies) are considered in nutrient space, which, in general, is a n-dimensional space where each dimension corresponds to one of n mutually exclusive nutritional components, such as carbohydrates or proteins. The momentary nutritional deficits of an animal define its position in nutrient space relative to a desired target nutrient state. By consuming food items animals move along 'rails' in nutrient space, according to the nutritional composition of those items. This means that, in general, food items may contain several nutrients and the nutrient ratios determine the slopes of the rails in nutrient space. However, in the present study we focus on the special case of one food-one nutrient to simplify the analysis, although the framework could be extended to allow for food types containing mixtures of nutrients (e.g. see Houston et al., 2011).

Using the geometric framework we can also describe an animal moving in deficit space instead of nutrient space, which we do in the present paper. In Fig. 1A we show an illustration of the geometric framework in deficit space, where we assume that an animal has to decide about the sequence in which to consume food type 1 and food type 2 in order to reduce its corresponding nutritional deficits. Hence, we may consider the distance between final deficit state and desired target state as a measure characterising the animal's performance at the end of the foraging task. In deficit space the target

nutrient state is located at the origin of the diagram, as the animal aims at reducing all nutritional deficits to zero. The computation of the animal's performance in our model is explained in more detail in Sec. 2.4.

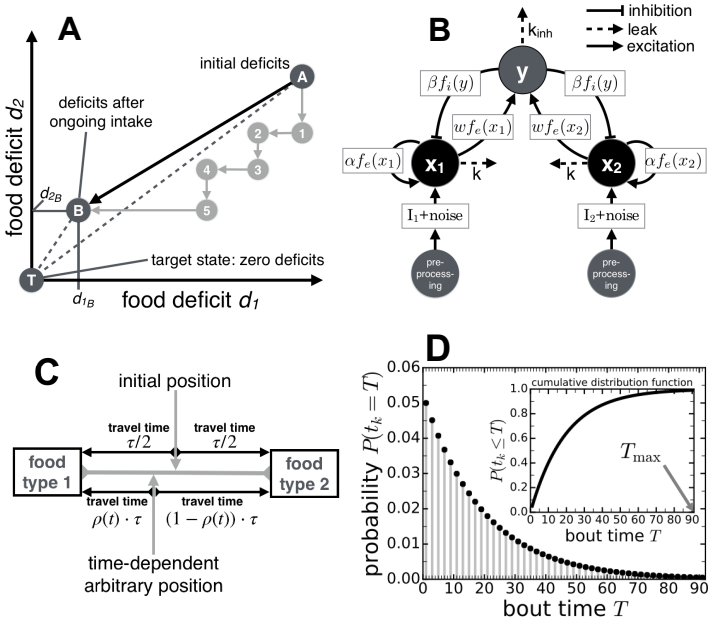

Figure 1: Overview of modelling assumptions. A: Illustration of the geometric framework in deficit space. The target state (T) is the origin of the diagram ($d_1 = 0, d_2 = 0$). Starting from state A, in a sequence of feeding bouts, the animal tries to reach state T but may end up in state B. Nutritional states are characterised by the Euclidean distance between points in the diagram (e.g. A or B) and the target state T. B: Schematic of the interneuronal inhibition motif. Inhibition is provided by neuronal unit y, which acts on evidence-integrating units x_1 and x_2. C: Food types 1 and 2 can be found in different locations. Travel times between arbitrary locations of the animal and the two food sources are modelled using the time-dependent ratio $\rho(t) \in [0, 1]$ and the switching cost τ (which equals the travel time needed to move from one food source to the other). The animal's initial position is given by $\rho(t = 0) = 1/2$. D: Plot of geometric distribution (Eq. (5)) with interruption probability $\lambda = 0.05$, including the corresponding cumulative distribution function (inset).

2.2 Architecture of the decision-making circuit and its coupling with nutritional state

To specify the behaviour of the animal making foraging decisions we implemented a decision-making circuit, which is based on an interneuronal inhibition motif, as schematically illustrated in Fig. 1B. This model architecture has similarities with the inhibition motif applied in a biologically plausible cortical network model used to describe motion discrimination experiments (Wang, 2002). In a mean-field approach this network

model containing synaptic detail could be reduced to a two-variable model with effective cross-inhibition, while implicitly embedding an inhibitory interneuronal population (Wong and Wang, 2006). Fig. 1B is more closely related to the reduced model derived by Wong and Wang (2006), which has been studied by Niyogi and Wong-Lin (2013) to investigate the co-modulation of both excitatory and inhibitory neurons in a perceptual decision-making task, for example. The reduced model (Wong and Wang, 2006) has also been applied to economic choices (Hunt et al., 2012). We note that in our model interneuronal units are the sole source of inhibition, whereas previous models contained cross-inhibitory connections between evidence integrating units in addition to interneuronal inhibition (Wong and Wang, 2006; Niyogi and Wong-Lin, 2013). We also note that the architecture in Fig. 1B has previously been studied as a linear model in the context of perceptual decision making (Bogacz et al., 2006), and it has been shown that coupled nonlinear rate equations where inhibition is provided by interneuronal units may be reduced to a nonlinear diffusion equation (Roxin and Ledberg, 2008). In the present paper we use a nonlinear mathematical model to implement the interneuronal inhibition motif (see Eq. (3) below). Although mainly studied in perceptual decision-making tasks, we believe that the model architecture in Fig. 1B is rather generic, whereas its mathematical implementation may differ with the decision type.

Applied to our foraging animal model, evidence in favour of food type 1 (food type 2) is integrated by neuronal units x_1 (x_2). We interpret the activity level of x_1 (x_2) as motivation to feed at food source 1 (food source 2). The momentary nutritional state of the decision-maker generates a representation as neural activation. This is the function of the pre-processing units in Fig. 1B, which transform the physiological state into inputs I_1 and I_2 that feed their respective integrators x_1 and x_2. Here, the relationship between physiological levels characterising the nutritional state (deficits) and their representations in the neural circuit is given as

$$I_j(d_j) = q\, d_j\,, \qquad (j = 1, 2), \tag{1}$$

where q denotes the sensitivity of the animal to deficits d_j ($j = 1, 2$) in nutritional items.

We assume that inputs I_j ($j = 1, 2$) are polluted by processing noise with standard deviation σ, which may arise from currents originated in other circuits in the brain. Noise is included via Wiener processes W_1 and W_2. Recurrent excitation is taken into account in the self-excitatory terms with strength α. If activity levels of x_1 and x_2 are sufficiently large then the interneuronal inhibitory unit y becomes activated with strength w and in turn inhibits the evidence-integrating units with strength β. The functions $f_{e,i}(\cdot)$ appearing in different places in Fig. 1B are normalised nonlinear input-output functions with a typical sigmoidal shape. Here, f_e is involved in excitatory processes, whereas f_i represents inhibition. The form of the sigmoidal functions is given as

$$f_{e,i}(\xi) = \frac{1}{1 + \exp[-g_{e,i}\,(\xi - b_{e,i})]}\,, \tag{2}$$

where g_e and g_i are the gains, and b_e and b_i give the inflection points where f_e and f_i reach half-level, respectively. In addition, we assume that information may be lost by including leak-terms in the evidence-integrating units x_1 and x_2 (rate k), and in the

interneuronal inhibitory unit y (rate k_{inh}). This model is thus described by the following system of nonlinear stochastic differential equations

$$\begin{aligned} dx_1 &= [-k\,x_1 + \alpha\,f_e(x_1) - \beta\,f_i(y) + I_1(d_1)]\;dt + \sigma\,dW_1\,, \\ dx_2 &= [-k\,x_2 + \alpha\,f_e(x_2) - \beta\,f_i(y) + I_2(d_2)]\;dt + \sigma\,dW_2\,, \\ dy &= [-k_{inh}\,y + w\;(f_e(x_1) + f_e(x_2))]\;dt\,. \end{aligned} \tag{3}$$

In addition to the nonlinear functions in Eq. (3) we introduce an artificial nonlinearity to prevent x_1, x_2 and y from taking negative values, i.e. when numerically integrating system (3) from time t_n to obtain the state variables at the next timestep t_{n+1}, we reset $X_{t_{n+1}} = \max(0,\, X_{t_{n+1}})$, $X = x_1, x_2, y$; otherwise the leak terms (rates k and k_{inh}) would become positive when the activity levels of the state variables drop below zero (which may happen occasionally without the max-function filtering out negative values).

2.3 Switching cost, decision criterion and reduction of nutrient deficits

We assume that the two food sources are physically separated. Thus, there is a cost for the animal to switch between both food sources, as it has to move between the locations offering food type 1 or food type 2. Whilst the animal is moving it cannot consume nutrients. When there is no switching cost the optimal strategy is to eat exclusively one food type until a symmetric deficit state is reached then maintain a balanced nutrient intake (Houston et al., 2011). However, with switching costs suboptimal 'dithering' (i.e. frequent switches with little food intake) will result from this strategy (Marshall et al., 2015). In our model (as in the study by Marshall et al. (2015)) the cost for switching is represented by a time constant, denoted τ, which quantifies the travel time it takes the animal to move from one food source to the other (Fig. 1C). Note that the spatial position of the animal moving between both food locations is not modeled explicitly. To study the behaviour of the animal in our model it is sufficient to know at what point in time the animal is located at food source 1 or food source 2, and when it is moving between both food source locations. For this purpose, we introduce the ratio $\rho(t)$ to express the temporal distance between the animal's current position and the locations of the two food sources, i.e. $\rho(t)\,\tau$ gives the travel time between current position and food source 1, and $(1-\rho(t))\,\tau$ represents the travel time between current position and food source 2 (Fig. 1C). Thus, $\rho(t)$ is a time-dependent ratio that ranges in the interval $[0, 1]$; if $\rho(t) = 0$ ($\rho(t) = 1$) the animal is located at food source 1 (food source 2), and otherwise moves between both food source locations ($0 < \rho(t) < 1$).

The decision criterion in our model is based on the assumption that the animal performs the activity for which it has the greatest motivation. This assumption has also been applied in previous studies of ongoing foraging tasks (e.g. Marshall et al., 2015). We thus consider the time-dependent motivation difference $\Delta x(t) = x_1(t) - x_2(t)$ and assume that the animal feeds at food source 1 ($\rho(t) = 0$) or moves towards it ($0 < \rho(t) < 1$, $\rho(t)$ decreases) if $\Delta x(t) > 0$, or it feeds at food source 2 ($\rho(t) = 1$) or moves towards the location of food source 2 ($0 < \rho(t) < 1$, $\rho(t)$ increases) if $\Delta x(t) < 0$. Throughout the entire decision-making task motivations $x_1(t)$ and $x_2(t)$ (and hence also $\Delta x(t)$) as well as the inhibitory activity $y(t)$ are constantly updated at each timestep.

In addition, in the numerical simulation we monitor sign $[\Delta x(t)\,\Delta x(t+dt)]$ at each timestep dt to detect motivation changes; if sign $[\cdot] = +1$ then the motivation difference did not change sign, whereas if sign $[\cdot] = -1$ then the motivation difference changed sign within t and $t + dt$. As a sign change corresponds to a reversal of the travel direction, this allowed us to update momentary travel times and travel directions when the motivation difference changed from $\Delta x(t) \lesseqgtr 0$ at time t to $\Delta x(t+dt) \gtreqless 0$ at time $t + dt$. During the time interval when the sign of the motivation difference remained unchanged travel times corresponding to the travel towards the current target food source decreased by one dt at each timestep. For example, if we assume that the model animal moves towards food source 1, i.e. $\Delta x(t) > 0$ and $\Delta x(t+dt) > 0$, then $\rho(t+dt)\,\tau = \rho(t)\,\tau - dt$. Because we know the initial travel time (i.e. $\rho(t=0)\,\tau$), and keep track of the momentary motivation difference $\Delta x(t)$ and travel direction, at each point in time we can detect if the animal has reached either of the two food sources.

We emphasise that this is a generic approach, where we do not need to make specific assumptions on the position-time law capturing the momentary location of the animal performing the decision-making task. Hence, it would be possible to implement position-time laws with arbitrary functional dependence between location and time in our model but in the present study we focus on the effect of excitation and inhibition in the decision-making circuit and model the motion of the animal implicitly.

Initially, the animal is placed exactly midway between the two food sources ($\rho(t = 0) = 1/2$), i.e. the animal needs the same amount of time to travel to food source 1 or food source 2 (Fig. 1C). Initial food deficits, $d_1(t = 0)$ and $d_2(t = 0)$, are set to either equal or unequal values. To determine the initial motivational state of the animal we use Eq. (3) without noise (i.e. we set $\sigma = 0$) and integrate the dynamical system until it reaches a stable fixed point or closed orbit. This means, we consider initial motivations where fluctuations have been averaged out, which allows the system to be prepared in a well-defined state at $t = 0$ (the Wiener processes in Eq. (3) represent white noise with zero mean). When numerically integrating the deterministic equations ($\sigma = 0$) we make use of a fourth-order Runge-Kutta method and when simulating the stochastic differential equations ($\sigma = 0.01$) we apply a predictor-corrector method, where the deterministic part is calculated with second order of accuracy in timestep dt (Kloeden et al., 2002). For both methods we used a timestep of $dt = 0.005$ in the numerical integration. We found that this choice of dt gives a good compromise between computation time and accuracy when integrating the system, in particular with regard to the stochastic equations.

During feeding at one of the two food sources, the time-dependent deficits are reduced according to

$$d_j(t) = d_j(0) - \gamma\, t\,,\ j = 1,\, 2\ , \tag{4}$$

where $d_j(0)$ are the initial deficits at $t = 0$, and γ is the deficit decay parameter (Houston and Sumida, 1985; Marshall et al., 2015). During the time the animal does not feed, i.e. when the animal is moving between the two food sources, the nutritional state is assumed to remain constant. This is a valid assumption if feeding takes place within sufficiently short periods of time (Houston and Sumida, 1985; Marshall et al., 2015).

2.4 Interruption probability and evaluation of the animal's performance

We take into account that the animal may be interrupted whilst executing the sequence of feeding bouts. This interruption could be due to the presence of a predator, for instance. To model the probability of interruption, here we follow the approach presented in Marshall et al. (2015) by assuming that feeding activities are geometrically distributed over time, which is consistent with the concept of discounted utility of future rewards observed in foraging animals (e.g. see Stephens and Krebs, 1986). The geometric distribution is given as

$$P(t_k = T) = (1 - \lambda)^{(T-1)} \, \lambda \,, \tag{5}$$

where t_k and bout time T take integer values, i.e. $T = 1, 2, 3, \ldots$. With interruption probability per unit time λ the distribution $P(t_k = T)$ gives the probability that the ongoing decision-making task comes to an end at time $t_k = T$. In Fig. 1D the geometric distribution and the corresponding cumulative distribution function (inset) are displayed. The maximum bout time T_{max} is computed such that the cumulative probability of the foraging task being interrupted is at least 99%. In the following analysis, we assume an interruption probability of $\lambda = 0.05$, yielding a maximum bout time of $T_{max} = 91$. Hence, using the geometric distribution we can define an upper bound on the duration of the ongoing decision-making task solely defined by the interruption probability λ. This seems reasonable from a behavioural ecology point of view, as in a natural environment the presence of predators or other interruptions are likely to determine the end of feeding bouts rather than uninterrupted food consumption until all deficits are satisfied.

In the geometric framework a simple measure of performance in a nutritional decision-making task is the square of the Euclidean distance between current state and target state (Simpson and Raubenheimer, 2012). The larger this distance the smaller the reward, or phrased differently, the larger the penalty incurred by the animal. As we take into account possible interruptions at different points in time during the ongoing decision-making task, a quantity that characterises the overall performance of the animal is the expected penalty given as (Marshall et al., 2015)

$$E(p) = \sum_{T=1}^{T_{max}} p(T) \, P(t_k = T) \,, \qquad p(T) = d_1^2(T) + d_2^2(T) \,, \tag{6}$$

where $p(T)$ denotes the penalty if nutritional intake stops at time T, i.e. the square of the Euclidean distance between the nutritional state at time T and the target state. $P(t_k = T)$ is the probability representing the geometric distribution as introduced in Eq. (5).

3 Results

3.1 Temporal evolution of deficits and motivations

We begin our analysis by showing the feeding behaviour of the animal regulated by the decision-making circuit. For this purpose, we introduce the excitation-over-inhibition-

ratio (E/I-ratio) as

$$r = \frac{\alpha}{\beta} \ , \tag{7}$$

as well as the mean deficit, d_m, and the deficit difference, Δd, according to

$$d_m = \frac{1}{2}(d_1 + d_2) \ , \qquad \Delta d = d_1 - d_2 \ . \tag{8}$$

Using these definitions, we plot the change of the animal's motivational state and the reduction of the animal's deficits for different E/I-ratios, $r = 1$ and $r = 2$, in Fig. 2. Here we assumed that initial deficits in food type 1 and food type 2 are equal at $t = 0$, i.e. $d_m(t = 0) = 7.5$ and $\Delta d(t = 0) = 0$; below we also consider unequal initial deficits, $\Delta d(t = 0) > 0$, in Section 3.4.

In Fig. 2A, we show the temporal evolution of the motivational difference $\Delta x(t) = x_1(t) - x_2(t)$ for $r = 1$. If $\Delta x(t) > 0$ ($\Delta x(t) < 0$) the animal moves towards, or feeds at, food source 1 (food source 2). At $t = 0$ we have $\Delta x(t = 0) = 0$ due to the symmetry of the initial conditions but we point out that the absolute values of $x_1(t = 0)$ and $x_2(t = 0)$ are nonzero in this example (and in general). However, due to the presence of noise the motivation difference quickly becomes nonzero for $t > 0$, and if $\Delta x(t)$ gets sufficiently large it moves towards one of the accessible attracting states. In the following, accessible stable states expressed by the motivation difference are denoted $\Delta x^s = x_1^s - x_2^s$. We observe two types of attractors in our model – stable fixed points and stable periodic orbits. If an accessible state is a stable fixed point then x_1^s and x_2^s are simply the equilibrium values of the motivations, i.e. $x_j(t \to \infty) \to x_j^s$ ($j = 1, 2$). However, if an accessible state describes a stable periodic orbit then Δx^s oscillates between $\max(\Delta x^s)$ and $\min(\Delta x^s)$. Regarding the simulation in Fig. 2A ($r = 1$), the only accessible attracting states are stable limit cycles; one periodic orbit with amplitudes $\Delta x^s > 0$ and another periodic orbit with amplitudes $\Delta x^s < 0$. Hence, the symmetric initial condition $\Delta x(t = 0) = 0$ is a metastable state arising from equally strong attraction by both limit cycles. The sustained oscillations are inherent to the nonlinear decision-making circuit. This is further discussed below and illustrated in Fig. 3. If the animal reached food source 1 (food source 2) it could feed and reduce deficit d_1 (d_2) according to Eq. (4). Thus, during food intake the mean deficit decreased and the deficit difference increased or decreased as shown in Fig. 2C. A deficit reduction in turn means that the input to the decision-making circuit, which corresponds to the consumed food item, decreases. The animal can only consume one food item at a time which introduces an asymmetry in the system and causes another type of oscillation in the animal's response described by its motivational state – the oscillations around $\Delta x = 0$ which are due to food intake and correspond to the temporal evolution of $\Delta d(t)$ (compare Figs. 2A and C).

The behaviour for $r = 1$ in Figs. 2A and C is contrasted with the behaviour observed for the E/I-ratio $r = 2$ illustrated in Figs. 2B and D. As in the case $r = 1$ (Fig. 2A), we observe motivation differences oscillating around $\Delta x = 0$ for $r = 2$ but the amplitudes are much smaller (Fig. 2B). The reason for this behaviour is that $\Delta x^s = 0$ is now an accessible stable fixed point and motivation differences $\Delta x(t) \neq 0$ only arise from fluctuations due to noise in the system and from decreasing deficits as a result of

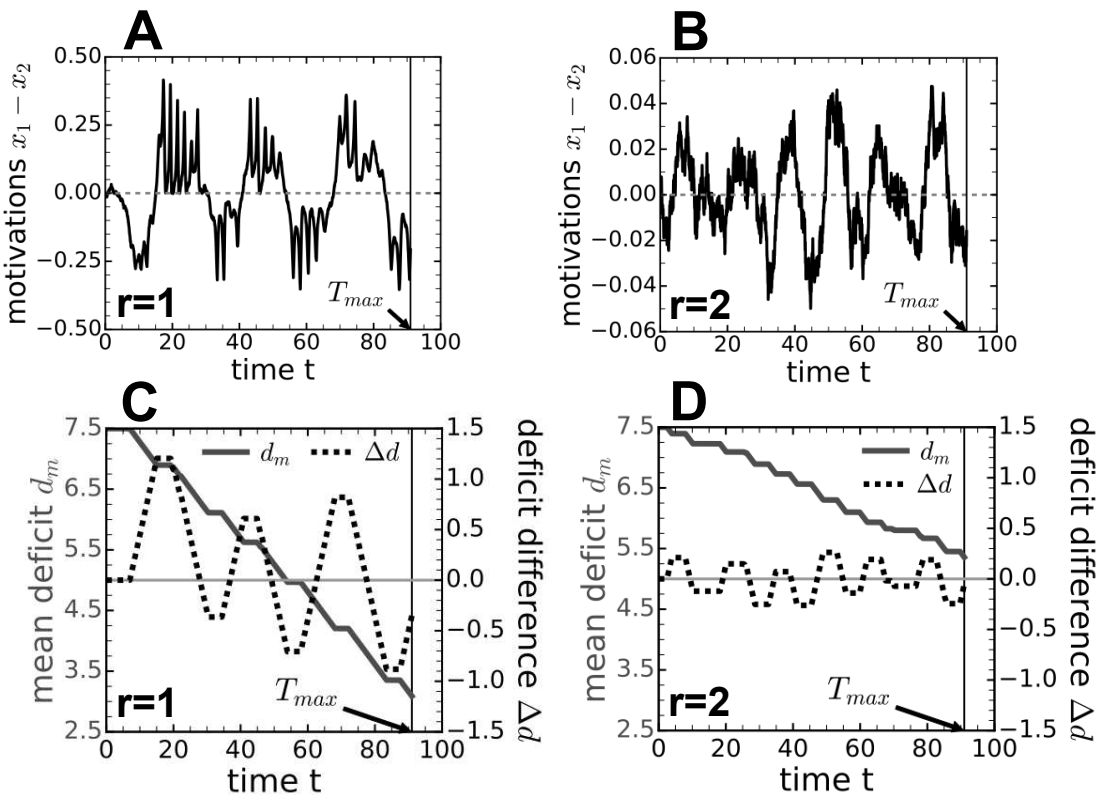

Figure 2: Simulation of ongoing decision-making process for different E/I-ratios. Shown are the change of motivation differences (A and B) and corresponding deficits, i.e. mean deficit d_m and deficit difference Δd, (C and D) over time. Amplitudes of motivation difference and number of motivation switches depend on the E/I-ratio r. For $r = 1$ (A) we observe fewer motivation switches compared with the case $r = 2$ (B), yielding a larger deficit reduction for $r = 1$ (C) compared with the $r = 2$ case (D). The task ends at $T_{\max} = 91$, as explained in the text. Parameter values: $d_1(t = 0) = 7.5 = d_2(t = 0)$, $\tau = 4$, $\gamma = 0.15$, $q = 0.1$, $\beta = 3$, $k = 0.8$, $k_{inh} = 0.8$, $w = 3$, $g_e = 10 = g_i$, $b_e = 0.5 = b_i$, and $\sigma = 0.01$.

feeding at a food source. However, as the magnitudes of $\Delta x(t)$ are small, we observe more frequent switches between activities (see also Fig. 4 below). This leads to a less effective deficit reduction for $r = 2$ (Fig. 2D) compared with the case $r = 1$ (Fig. 2C).

Our finding that oscillatory regimes, which arise from nonlinearity in the underlying decision-making circuit, may facilitate the continuous decision-making process is also evident in the bifurcation diagrams in Fig. 3. Here, we plotted all accessible stationary states of the dynamical system (3) with $d_m \in [2.5, 7.5]$ as the critical parameter, including bifurcation points[1]. We chose different values of Δd that are representative for the entire decision-making task, i.e. deficits change over time and are frequently equal or characterised by small differences (the animal feeds at food source 1, then at

[1] Bifurcation points were computed using the numerical continuation tool *MatCont* (Dhooge et al., 2003, 2008).

food source 2, and so on; cf. Figs. 2C and D). This means that the mean deficit, $d_m(t)$, decreases over time whilst the deficit difference, $\Delta d(t)$, alternates between positive and negative values and zero.

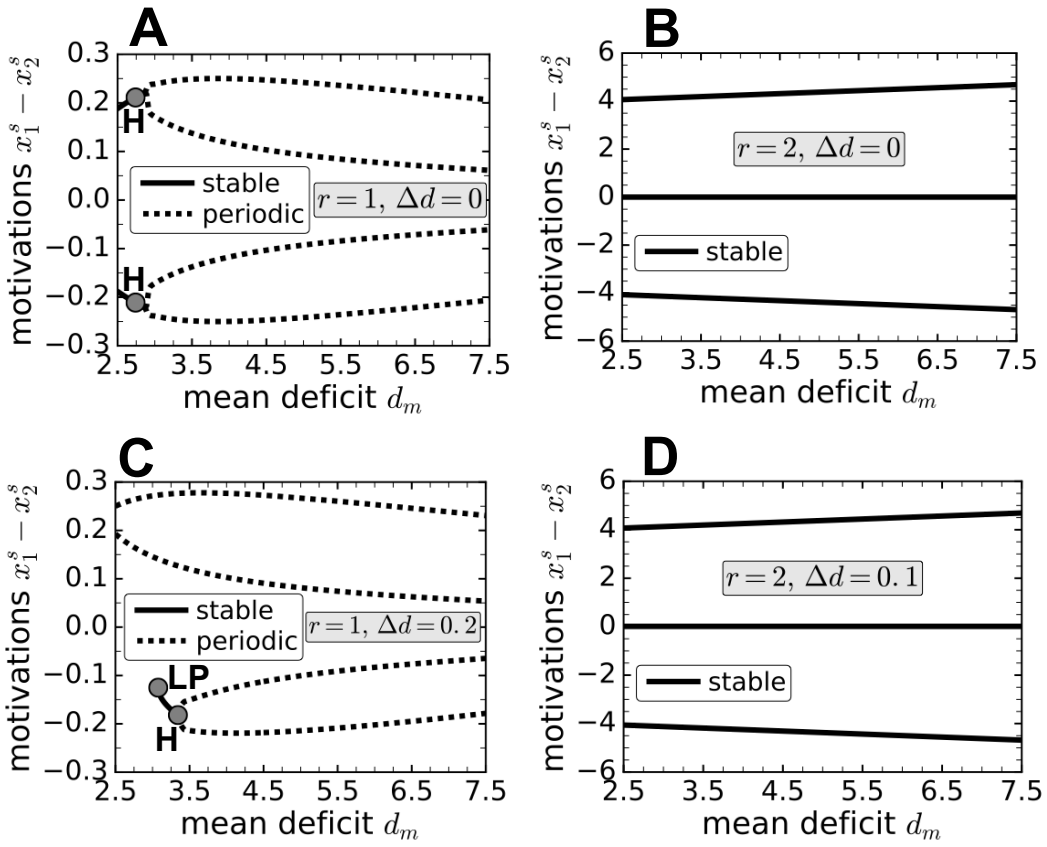

Figure 3: Plot of stable motivational states, $\Delta x^s = x_1^s - x_2^s$, depending on mean deficits, $d_m \in [2.5,\ 7.5]$, that correspond to Fig. 2. Different initial deficits, $\Delta d(t = 0)$, and E/I-ratios, r, are considered. For $r = 1$ (A and C) stable periodic orbits are the only accessible states over a wide range of relevant d_m-values, whereas for $r = 2$ (B and D) periodic orbits do not exist in the same range of d_m-values. Bifurcation points are indicated (H: Hopf bifurcation, LP: limit point). Only accessible states are shown – either stable fixed points (solid lines) or stable limit cycles. Maximum and minimum amplitudes in a limit cycle are plotted for the periodic solutions (dashed lines). Parameter values: $\tau = 4$, $\gamma = 0.15$, $q = 0.1$, $\beta = 3$, $k = 0.8$, $k_{inh} = 0.8$, $w = 3$, $g_e = 10 = g_i$, $b_e = 0.5 = b_i$, $\sigma = 0$.

Figs. 3A and C show that stable limit cycles are the only attracting states for $r = 1$ over a wide range of d_m-values, except for small d_m below the Hopf-bifurcation point at $d_m \approx 2.74$ in case $\Delta d = 0$ (Fig. 3A) and at $d_m \approx 3.08$ in case $\Delta d = 0.2$ (Fig. 3C). Below this bifurcation point we observe two stable fixed points with nonzero Δx^s. Both Hopf-bifurcations are supercritical (the first Lyapunov-exponent is negative). If we introduce a nonzero deficit difference (Fig. 3C), we see that the Hopf-bifurcation points are shifted; the Hopf-bifurcation point for which $\Delta x^s > 0$ moves to smaller d_m-

values (out of the range plotted), and the Hopf-bifurcation point for which $\Delta x^s < 0$ moves to larger d_m-values. In addition, the amplitudes corresponding to a motivation difference below zero, i.e. the motivation to eat the food type in which the animal has the lower deficit is greater, are slightly smaller than those corresponding to a motivation difference larger zero. However, in our simulations we observed that although both limit cycles are orbitally stable, oscillating motivation differences move quickly onto the limit cycle for which $\Delta x^s > 0$.

In contrast, if we consider the case $r = 2$ (Figs. 3B and D), we see that there is a stable fixed point characterised by $\Delta x^s = 0$. The other two stable fixed points shown in Fig. 3B cannot be reached due to the symmetric initial deficits $d_1(t = 0) = d_2(t = 0)$ (cf. the small amplitudes of $\Delta x(t)$ Fig. 2B). Even if the animal feeds at one of the food sources, which yields asymmetric deficit inputs, we do not observe a noticeable change in the plot of the accessible states in Fig. 3D ($\Delta d = 0.1$) compared with Fig. 3B ($\Delta d = 0$). Therefore, it is the stable equilibrium with $\Delta x^s = 0$ that pulls back the system to a symmetric state, which makes the foraging task less effective.

3.2 Time intervals between motivation changes

The higher efficiency in case $r = 1$ compared with the case $r = 2$ is also highlighted in Fig. 4. Here we plotted the time interval between two consecutive motivation changes, denoted ΔT_{change}, which is defined according to

$$\Delta T_{\text{change}}(n) = t^c_n - t^c_{n-1} \quad \text{if} \;\; t^c_n - t^c_{n-1} > \varepsilon \,, \;\; t^c_n > t^c_{n-1} \,, \tag{9}$$

where t^c_n denotes the point in time of the n-th motivational change observed in the simulation, i.e. $t^c_n = t$ if $\text{sign}[\Delta x(t - dt)\, \Delta x(t)] = -1$ (dt is the timestep in the numerical integration). In the definition of ΔT_{change} in Eq. (9) we only count time intervals that are larger than the threshold value ε. Otherwise noisy fluctuations would lead to a large number of motivation changes with $\Delta T_{\text{change}} \approx 0$. The shaded area where $\Delta T_{\text{change}} < \tau$ indicates inefficient decision-making. As τ represents the time it takes to travel from one food source to the other, values of $\Delta T_{\text{change}} \leq \tau$ indicate dithering between the two available options, that is frequently changing motivations lead the animal to travel back and forth between both food sources with little or no food intake. In addition, we can also see that, on average, ΔT_{change} increases with increasing number of motivation changes for $r = 1$, whereas it only increases slightly for $r = 2$ for small numbers of motivation changes and remains almost constant afterwards. Notably, the averaged curve for $r = 2$ remains below τ for all motivation changes, whereas the averaged curve for $r = 1$ is always larger than τ (averaging included 10^3 simulations). However, we also point out that ΔT_{change} should not increase too much, as otherwise efficiency drops again. This happens if ΔT_{change} approaches the maximum bout time $T_{\max}$. Therefore, we conjecture that improved decision-making of the animal is possible only if the decision-making circuit regulates the animal's behaviour such that $\tau < \Delta T_{\text{change}} < T_{\max}$, suggesting that ΔT_{change} may also be considered as an indicator of efficient food intake.

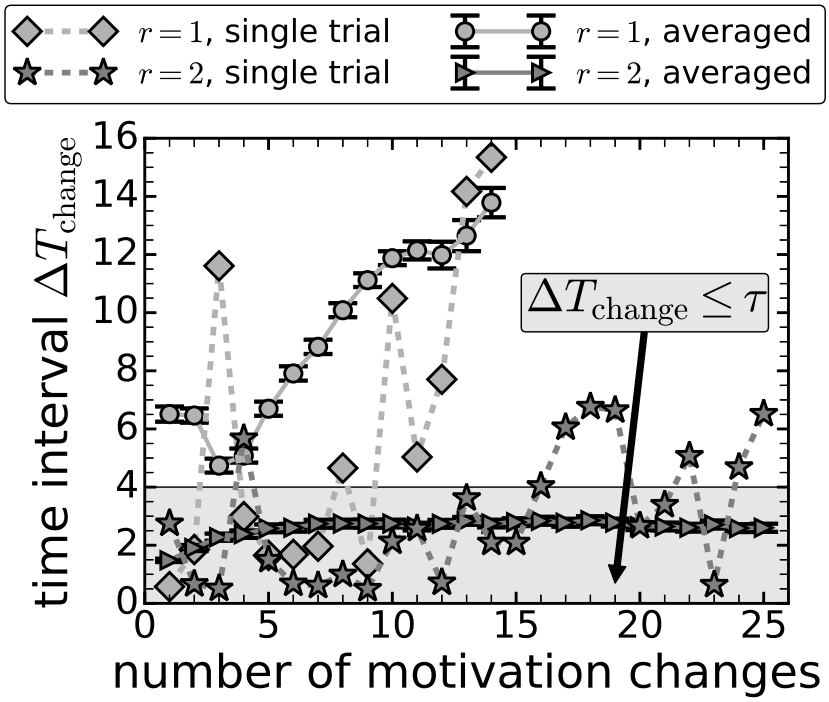

Figure 4: Plot of time intervals between two consecutive motivation changes depending on the number of motivation changes that occur during the ongoing decision-making task. The shaded area ($\Delta T_{\text{change}} \leq \tau$) is the area where dithering might occur, i.e. switching motivations in a time interval smaller than is needed to travel from food source 1 to food source 2. Averaged curves show mean values obtained from simulating 10^3 independent trials; on average, foraging is more effective for $r = 1$ (averaged curve above the shaded area) than it is for $r = 2$ (averaged curve within the shaded area). Single trial curves correspond to motivation changes shown in Figs. 2A ($r = 1$) and C ($r = 2$) and fluctuate around the averaged curves. Bars represent 95% confidence intervals. Model parameters as in Fig. 2, and $\varepsilon = 0.5$.

3.3 Performance under the modulation of inhibition strength and excitation/inhibition-ratio

In this section, we highlight the significance of excitatory and inhibitory connections in the decision-making circuit. In Fig. 5, we have simulated our model for inhibition strengths in the range $0 < \beta \leq 5$ and E/I-ratios varied between $0 < r \leq 2.5$. Fig. 5C depicts the performance of the hypothetical animal measured by the expected penalty (cf. Eq. (6)). Additionally, we show the bifurcations and the relevant stable equilibria and closed orbits that occur when the values of β and r are varied. An area of improved performance is clearly recognisable (smallest values of expected penalty) in Fig. 5C. The shape of this area may be related to the corresponding bifurcation diagrams. We show the bifurcation diagram when $\beta = 3$ is kept constant and r is varied (Fig. 5D), and the bifurcation diagram when $r = 1$ is kept constant and β is varied (Fig. 5A).

Both bifurcation diagrams correspond to the initial deficit condition at $t = 0$. As time progresses, bifurcation diagrams will be updated, so that at every instant in time the bifurcation diagrams change. However, as indicated in Fig. 3, relevant accessible states do not seem to change significantly with decreasing deficits over a wide range of possible deficit values. Therefore, we assume that the bifurcation diagram at $t = 0$ may be considered as a suitable indicator of the expected decision-dynamics.

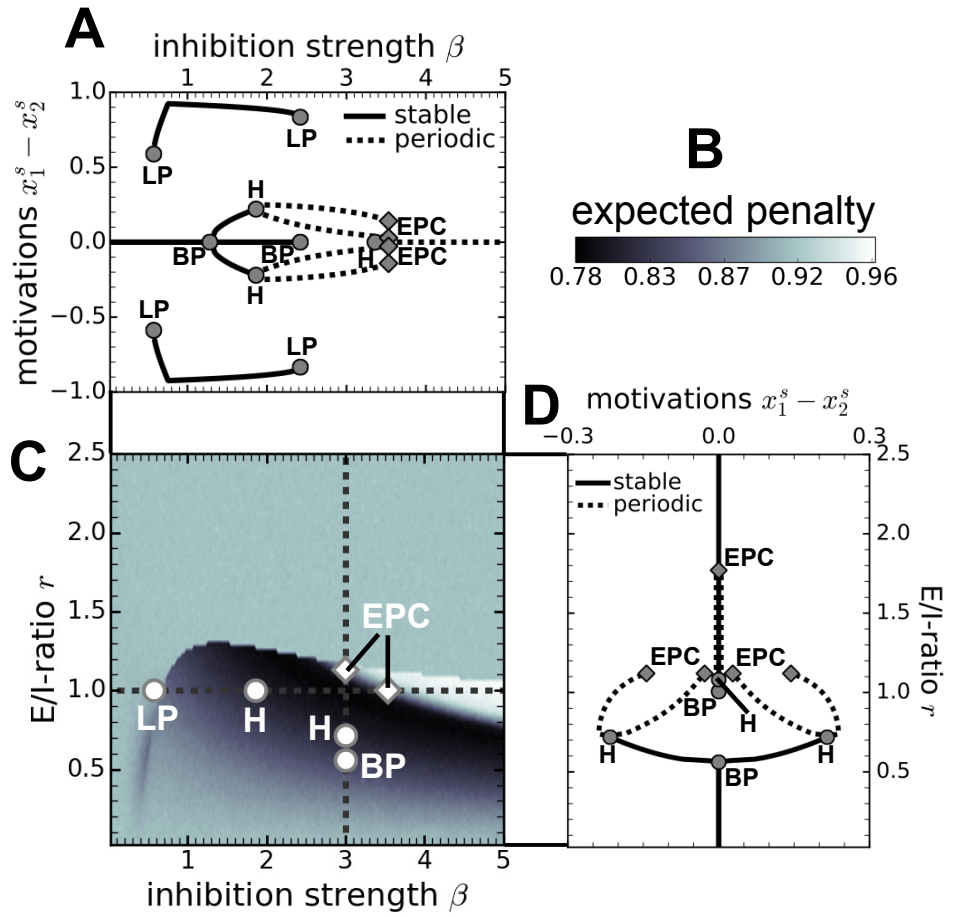

Figure 5: Depiction of the expected penalty with plots of accessible stationary states for $d_m(t = 0) = 7.5$ and $\Delta d(t = 0) = 0$. Colourbar in B corresponds to plot of expected penalty in C. Bifurcation diagrams in A and D correspond to dashed lines in C. Only stable stationary states are shown: either stable fixed points or stable limit cycles. Maximum and minimum amplitudes are plotted for periodic solutions (A and D). The area of best performance seems to coincide with the occurrence of stable limit cycles where $x_1^s - x_2^s \neq 0$. Abbreviations: LP: limit point, BP: branch point, H: Hopf bifurcation point, EPC: end point of cycle (diamonds). Selected bifurcation points are re-plotted in C. Other model parameters as in Fig. 2.

Inspecting the bifurcation diagram when β is the critical parameter (Fig. 5A), we can see that for low values of the inhibition strength ($\beta < 0.57$) the only stable fixed point is given by a decision deadlock state ($\Delta x^s = 0$). Increasing β to larger values,

we observe possible decision deadlock-breaking indicated by the existence of stable equilibria with $\Delta x^s \neq 0$. With the occurrence of decision deadlock-breaking the performance of the animal improves (compare bifurcation diagram in Fig. 5A and performance plot in Fig. 5C). The performance improves even more with the emergence of two stable periodic orbits with amplitudes $\Delta x^s > 0$ and $\Delta x^s < 0$, respectively (note the two supercritical Hopf-bifurcations points at $\beta \approx 1.9$). However, at $\beta \approx 3.4$ we observe another supercritical Hopf-bifurcation with $\Delta x^s = 0$. In addition, the stable orbits for which $\Delta x^s \neq 0$ cease to exist at $\beta \approx 3.54$, at which point the performance of the animal drops significantly (compare bifurcation diagram in Fig. 5A and performance plot in Fig. 5C). In Fig. 5 we use the label EPC (end point of cycle) to indicate that stable closed orbits vanish. This can either be due to the existence of a limit point of cycles where stable and unstable periodic orbits meet or the collision of the limit cycle with a saddle point (homoclinic bifurcation). We observe both events in our analysis.

Similar qualitative behaviour can be observed in the bifurcation digram with r as the critical parameter (see Fig. 5D). The performance improves as soon as the decision-deadlock state is broken (see branch point at $r \approx 0.56$), and is even further enhanced with the emergence of stable periodic orbits with $\Delta x^s \neq 0$ (see Hopf-bifurcations (supercritical) at $r \approx 0.71$). However, again we observe a clear drop in performance when these periodic orbits vanish at $r \approx 1.13$. For larger r-values the relevant accessible solutions for the decision-making circuit are a stable fixed point and another stable periodic orbit (which exists until $r \approx 1.77$), both characterised by $\Delta x^s = 0$.

Our results in Fig. 5 underpin that the occurrence of stable periodic orbits characterised by motivation differences $\Delta x^s \neq 0$ may enhance decision-making performance. The size of the area of improved performance is more extended along the β-axis and narrower along the r-axis, which seems to be strongly correlated with the range for which these periodic solutions exist. In contrast, stable fixed points and periodic oscillations for which $\Delta x^s = 0$ lead to a drop in performance. If the motivational state is attracted by these solutions that relate to a decision-deadlock, then frequent changes in motivation difference with small amplitudes may occur. The temporal evolution of $\Delta x(t)$ is thus prevented from gaining large motivation differences because it is driven back to the symmetric state $x_1^s = x_2^s$. In contrast, when the motivations move along the asymmetric orbits with $\Delta x^s \neq 0$ (compare also Fig. 2A) the periodic orbit allows the motivations to achieve sufficiently large differences, so that the animal can feed effectively. However, within one oscillation period motivation differences always come close to the switching line $\Delta x = 0$. Due to the reduction of deficits (whilst feeding) and the presence of noise, this facilitates activity switching in an efficient way. In Fig. A1 in Appendix A we also show that with increasing travel time between food sources (i.e. increasing switching cost τ), the expected penalty increases as well. However, the shape of the performance plots remains similar compared with Fig. 5C.

3.4 Dependence of expected penalty on initial deficits

To investigate the dependence of the expected penalty on the initial deficit difference at $t = 0$ we refer to Fig. 6, where expected penalties are plotted for different E/I-ratios, r, alongside examples of the temporal evolution of motivations for selected $\Delta d(t = 0)$. To simplify the comparison among different $\Delta d(t = 0)$, we chose the initial deficits,

$d_1(t=0)$ and $d_2(t=0)$, such that the value of the initial penalty, $p(t=0) = p_0 = d_1^2(t=0) + d_2^2(t=0)$, remains constant for all $\Delta d(t=0)$ considered in Fig. 6. Hence, in all cases the animal's deficit state is characterised by identical initial penalties but different initial deficits.

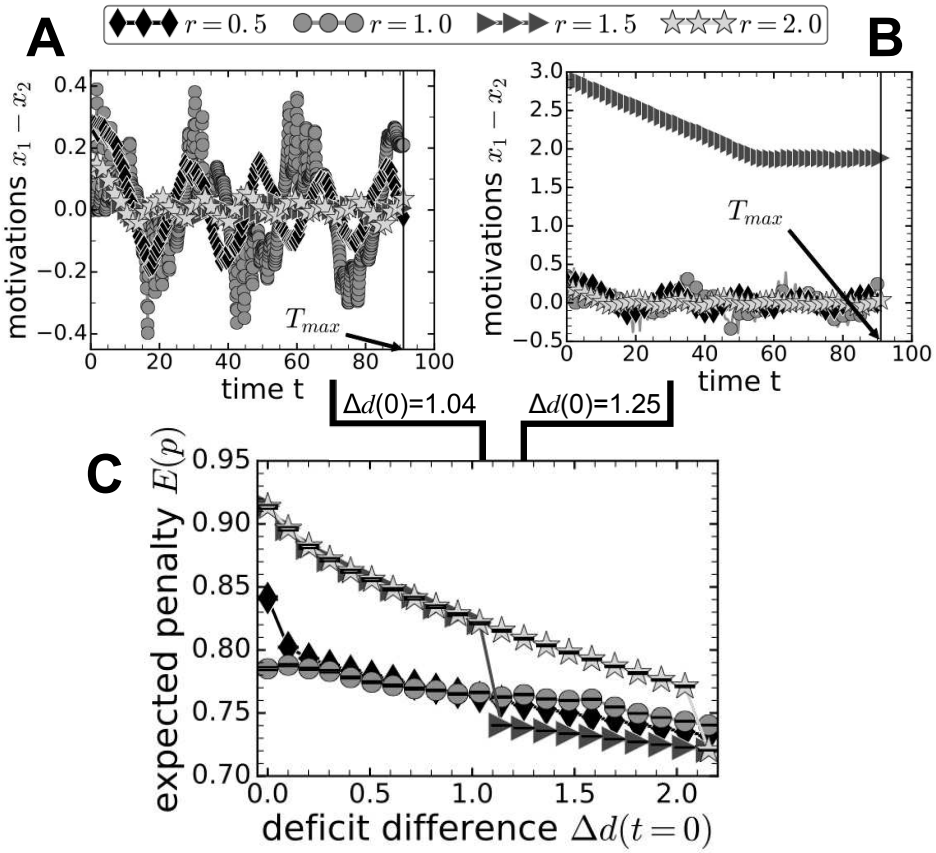

Figure 6: Plot of expected penalties averaged over 1000 trials (C) and motivation differences for single trial examples (A and B) depending on the initial deficit difference, $\Delta d(t=0)$. The expected penalty is normalised with respect to the initial penalty at $t=0$: $p_0 = d_1^2(0) + d_2^2(0)$. For varying $\Delta d(t=0)$ the initial penalty was kept constant. Performance improves (expected penalty decays) with increasing $\Delta d(t=0)$. Sudden jumps of $E(p)$, which also depend on the value of the E/I-ratio r, are observed for sufficiently large $\Delta d(t=0)$; more details can be found in the text. Error bars in C represent 95% confidence intervals (errors are small). Other parameter values: $\tau = 4$, $\beta = 3$, $\gamma = 0.15$, $q = 0.1$, $k = 0.8$, $k_{inh} = 0.8$, $w = 3$, $g_e = 10 = g_i$, $b_e = 0.5 = b_i$, and $\sigma = 0.01$.

In line with our results reported in Sec. 3.3, a variation of the E/I-ratio has a significant effect on the performance of the animal. For instance, the $r = 1$ curve in Fig. 6C shows a lower penalty value compared with both smaller ($r = 0.5$) and larger ($r = 1.5$ and $r = 2$) values of the E/I-ratio for sufficiently small differences in the initial deficits. In contrast, when increasing the initial deficit differences we can see that first the $r = 0.5$ and $r = 1.5$ curves (at $\Delta d(t=0) \approx 1.1$) and later the $r = 2$

curve (at $\Delta d(t=0) \approx 2.1$) drop below the $r = 1$ curve. However, the $r = 0.5$ and $r = 1.0$ curves show only small differences in performance in the whole Δd-interval, except for very small $\Delta d(t=0)$ (Fig. 6C). Thus, we find that adjusting the E/I-ratio according to the initial deficit state may help the animal to improve its food intake. The drop of the expected penalty we observe on the $r = 1.5$ and $r = 2$ curves in Fig. 6C is a direct consequence of the interplay between switching cost τ and the coexistence of different stable stationary motivational states, briefly described in the following. For $\Delta d(t=0) > 0$ there are two different stable fixed points available with $\Delta x^s > 0$; one characterised by a large difference in motivations and another fixed point characterised by a small motivational difference. In what follows, the value of the initial deficit difference quantifying the switch from small-Δx^s to large-Δx^s stable fixed points is denoted Δd_{switch}. Consider, for example, the $r = 1.5$ curves in Figs. 6A and C. If the initial deficit difference is small ($0 \leq \Delta d(t=0) < \Delta d_{switch} \approx 1.1$, Fig. 6C), then motivational differences are small, too (see initial motivations for $r = 1.5$ at $t = 0$ in Fig. 6A). However, if initial deficit differences are larger than Δd_{switch}, the initial motivational states make a transition from the small-Δx^s fixed point to the large-Δx^s equilibrium (cf. initial motivations for $r = 1.5$ at $t = 0$ in Figs. 6A and 6B). If this occurs, then the motivational differences are so far away from the switching condition for motivation changes ($\Delta x = 0$) that the animal only consumes one food type over the entire course of the ongoing decision-making task. Even when the animal has reduced all deficits of that one type to zero, its motivations reach a new steady state which is still too far from the switching condition, as shown in Fig. 6B (see curve labelled $r = 1.5$). The explanation for the drop of the $r = 2$ curve at $\Delta d(t=0) \approx 2.1$ in Fig. 6C is equivalent to that for the behaviour of the $r = 1.5$ curve. Hence, for sufficiently large differences of the initial deficits the animal may consume only one nutritional item, and by doing so, may achieve the lowest penalty value. However, this is only beneficial if the corresponding time frame is sufficiently small. Otherwise it would be detrimental for the animal to only focus on balancing one of its deficits, and neglecting the other one. We also note that for sufficiently small switching costs, the penalty for consuming exclusively one food type would be higher compared with switching between the two activities (Houston et al., 2011; Marshall et al., 2015). We confirm this and present more details about the reduction of deficits for $\tau = 0.05$ and $\tau = 4$ in Fig. B1 in Appendix B, including the deficit plots corresponding to Fig. 6B.

4 Discussion

Using an interneuronal inhibition motif implemented in a decision-making circuit on the behavioural level, we demonstrated that modulating inhibition strength and E/I-ratio may enhance decision-making performance in an ongoing binary choice task. Applied to a model animal performing a foraging task we found that the feeding behaviour of the animal improved if its internal motivations were characterised by periodic oscillations inherent to the nonlinear decision-making circuit (Figs. 2, 3, 5 and 6). Entering oscillatory internal states may be achieved by tuning inhibition strength and E/I-ratio in accordance with given nutrient deficits.

Our result that a modulation of the E/I-ratio, r, may improve behavioural perfor-

mance was further underpinned by the observation that time intervals between two motivation changes, $\Delta T_{\mathrm{change}}$, may increase when the number of motivation changes increases (Fig. 4). For $r = 1$ we found that, except for the first few (in the beginning of the task) and the last few motivation changes (towards the end of the task), $\Delta T_{\mathrm{change}}$ increases monotonically with increasing number of motivation changes, whereas for $r = 2$ we did not observe this effect (Fig. 4). The increase of $\Delta T_{\mathrm{change}}$ for $r = 1$ was caused by a sufficient decrease of food deficits, which are the inputs to the decision-making circuit. Deficit reduction in case $r = 2$ was less effective (cf. Figs. 2C and D). As a change of motivation corresponds to the decision to stop the current and perform the alternative activity, $\Delta T_{\mathrm{change}}$ may be compared with reaction times in other choice paradigms, such as the free-response paradigm in perceptual decision-making, where evidence is integrated until a threshold criterion is met (e.g. Bogacz et al., 2006; Ratcliff and McKoon, 2008). This comparison is non-trivial but should be sensible if $\tau < \Delta T_{\mathrm{change}} < T_{\mathrm{max}}$, as discussed at the end of Sec. 3.2. For example, in a reduced cortical network model applied to investigate a perceptual decision-making task, reaction times decreased when the stimulus strength increased (Wong and Wang, 2006). Although decision type, choice paradigm and mathematical equations in the present paper and in the study by Wong and Wang (2006) are different, the finding of slower responses with decreasing absolute stimulus strengths reported by Wong and Wang (2006) seems to show similarities with our observation of increasing $\Delta T_{\mathrm{change}}$ with decreasing food deficits, at least on the behavioural level. We note, however, that the model by Wong and Wang (2006) represents a biophysically plausible network with synaptic currents, whereas in the present paper we investigated a coarse-grained macroscopic model that focuses on the inhibition mechanism and not on synaptic detail. Furthermore, reaction times in the work by Wong and Wang (2006) could be explained by local dynamics around a saddle point and did not involve oscillating activity levels of excitatory populations. Interestingly, decreasing reaction times with stronger input values have also been observed in other studies of perceptual decision-making (Pins and Bonnet, 1996; Polanía et al., 2014; Teodorescu et al., 2016; Pirrone et al., 2018), and value-based decision-making (Hunt et al., 2012; Polanía et al., 2014; Pirrone et al., 2018; Reina et al., 2018).

Our nonlinear implementation of the interneuronal inhibition motif could also have potential applications in behavioural resonance (Wiesenfeld and Moss, 1995; Russell et al., 1999). Inside the brain, noise is present at all stages of the sensorimotor loop and has immediate behavioural consequences (Faisal et al., 2008). It is known that a variation of noise strengths may induce transitions between different dynamical regimes (Juel et al., 1997; Yang et al., 1999; Gao et al., 2002). For example, it has been shown that the presence of noise in nonlinear dynamical systems may shift Hopf-bifurcation points (Juel et al., 1997), and can lead to stochastic resonance-like behaviour even in the absence of external periodic signals, when the system is close to a Hopf-bifurcation point (Yang et al., 1999). This seems to be particularly relevant for our study, as we have demonstrated that stable limit cycles born at Hopf-bifurcation points may improve decision-making and feeding behaviour. However, performing a bifurcation analysis in the presence of noise is a subtle issue and deserves to be investigated in a separate study, as noise-induced Hopf-bifurcation-type sequences may also arise in parameter regimes, where noise-free equations do not exhibit periodic solutions (Gao et al., 2002).

Although our macroscopic decision-making circuit allows the identification of all accessible motivational states of the behaving model animal, it does not include biological detail on the cellular or molecular level. In a physiologically more detailed picture, the motivation to eat involves signals from the periphery transmitted by hormones, such as leptin, insulin and ghrelin (Vong et al., 2011; Morton et al., 2006; Williams and Elmquist, 2012), neurotransmission in hypothalamic neurocircuits (Morton et al., 2006; Tong et al., 2008; Aponte et al., 2011) and the relative balance of activity in distinct brain areas (Essner et al., 2017). Agouti-related protein (AgRP) neurons and neurons that express pro-opiomelanocortin (POMC) located in the arcuate nucleus play pivotal roles in regulating food intake – AgRP neurons stimulate food intake whereas POMC neurons reduce the intake of food (Morton et al., 2006; Vong et al., 2011; Atasoy et al., 2012; Aponte et al., 2011; Wu et al., 2012; Williams and Elmquist, 2012; Liu et al., 2012; Rangel, 2013; Essner et al., 2017). Excitatory and inhibitory neurotransmitters are modulators of signals at corresponding neurobiological synapses. More precisely, there is evidence that excitatory glutamatergic input and its modulation by NMDA receptors play key roles in controlling AgRP neurons (Liu et al., 2012). Glutamatergic neurons in other brain regions have also been identified to affect food intake (Wu et al., 2012). Furthermore, it has been observed that leptin-responsive GABAergic presynaptic neurons mediate the response of postsynaptic POMC neurons (Vong et al., 2011) and it has been shown that GABAergic signalling by AgRP neurons is required to regulate feeding behaviour (Tong et al., 2008; Wu et al., 2009).

Although providing a simplified picture of reality, our modelling approach may give further insights on the behavioural level, as it combines a neurally-inspired circuit architecture combining mechanism and function; function in this context means that an optimal diet, i.e. achieving the target nutrient intake, is related to maximising reproductive value (Mayntz et al., 2005; Altaye et al., 2010; Dussutour et al., 2010; Houston et al., 2011; Jensen et al., 2012; Rho and Lee, 2016). As the decision-making circuit which underlies the neural computation regulates choice behaviour based on nutritional needs, its excitatory and inhibitory couplings are of paramount importance to advance our understanding of dietary choices. On the molecular and neuroanatomical level, progress has been made to reveal underlying neural circuits for hunger (Atasoy et al., 2012) and for mediating appetite (Wu et al., 2012), for example, which could build the basis for a biologically more refined network-based computational model of dietary choice. However, whether or not a biologically-based network model can attain sufficiently slow switching dynamics on the behavioural level, as observed in the macroscopic decision-making circuit in the present paper, and adapt to realistically large physical distances (i.e. large switching costs), requires further investigation. Potentially, this could also be of interest for applications in artificial decision-making systems, such as robots implementing brain-inspired mechanisms to perform activity selection tasks (Girard et al., 2003).

Acknowledgments

The authors thank Philip Holmes, Jonathan Cohen, Naomi Leonard and Sebastian Musslick (all at Princeton University, NJ, US), and Benoît Girard (CNRS, Sorbonne Uni-

versité, Paris, France) for fruitful discussions of the initial results of this study. The authors are also grateful for very helpful comments and suggestions of two anonymous reviewers.

Funding

The authors acknowledge funding by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 647704). The funders had no role in study design, data generation and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Computer code for data generation is open source and available under: https://github.com/DiODeProject/Inhi and-excitation-shape-activity-selection.

Appendix

A Dependence of expected penalty on switching cost τ

To show the effect of switching cost τ on the expected penalty defined in Eq. (6), we assumed initial deficits $d_1(t = 0) = 7.5 = d_2(t = 0)$ and compared the expected penalties for five different values of τ, i.e. $\tau = 2,\ 4,\ 8,\ 16$ and 32. The corresponding results are depicted in Fig. A1.

We can recognise that, although the shape of the penalty landscape in Figs. A1A-A1E remains very similar under variation of τ, the whole process becomes less effective. A numerical comparison of the minimum values of the normalised expected penalties $\min(E(p)/p_0)$ after terminal time T_{max}, is shown in Fig. A1F. The initial penalty at $t = 0$ is given as $p_0 = d_1^2(t = 0) + d_2^2(t = 0)$. The shape of the curve in this diagram confirms that performance decreases with increasing τ. Fig. A1F also illustrates that an increase of τ leads to a nonlinear relationship between expected penalty and switching cost.

B Comparison of deficit reduction for different switching costs

In Fig. B1 we show a comparison of the deficit reduction in the ongoing decision making task of the model animal in dependence on the travel time between both food sources (i.e. the switching cost) τ. In agreement with other work (Houston et al., 2011; Marshall et al., 2015) the animal performs better when the switching cost is lower; compare the plots in Figs. B1A-D with their counterparts in Figs. B1E-H. The animal also performs better when it frequently alternates between between both food types, if τ is sufficiently low. However, if the opposite applies and the switching cost is significantly higher, then animals performing exclusively one activity could improve their performance at the end of the ongoing foraging task. This behaviour can be achieved by modulating

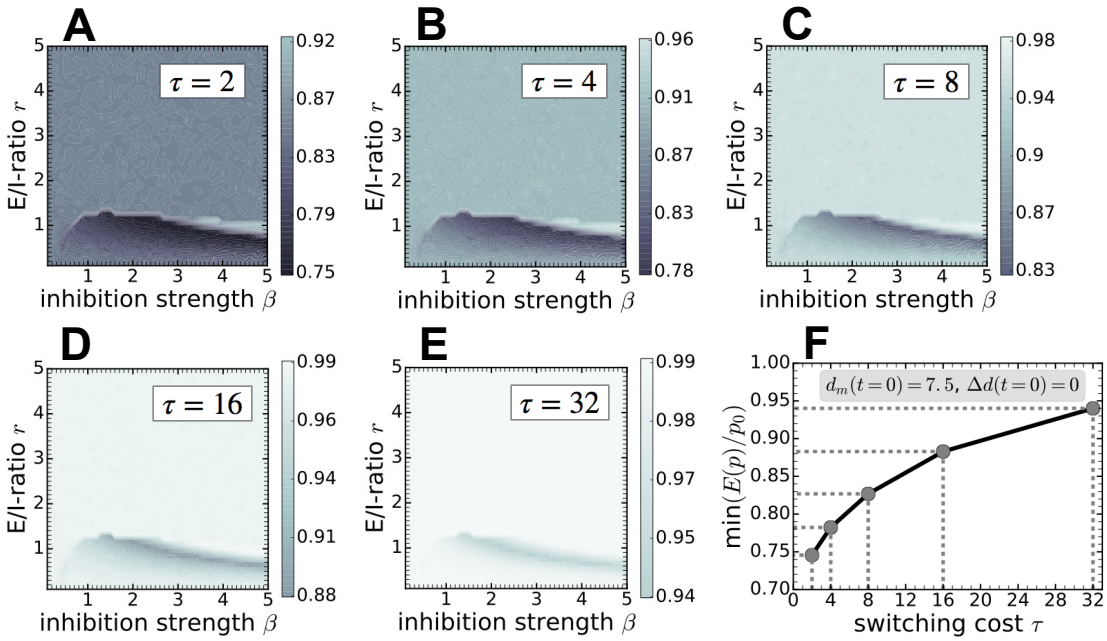

Figure A1: Dependence of expected penalty on switching cost τ. We chose a symmetric starting point at $\tau/2$ in all plots. Areas characterised by the lowest penalty values mirror the best performance of the model animal. Parameter values: $d_m(t=0) = 7.5$, $\Delta d(t=0) = 0$, $\gamma = 0.15$, $q = 0.1$, $k = 0.8$, $k_{inh} = 0.8$, $w = 3$, $g_e = 10 = g_i$, $b_e = 0.5 = b_i$, and $\sigma = 0.01$.

the E/I-ratio accordingly. Figs. B1E-H illustrate this result. An animal performing only one activity may achieve the best performance for $\tau = 4$ (see Fig. B1G). As discussed in the main text, this observation is a direct consequence of the nonlinearity of the underlying Eq. (3), and is, of course, only reasonable in the short-term, to which our study refers. In contrast, over longer periods of time the animal needs to perform both activities to survive.

References

Altaye, S. Z., Pirk, C. W. W., Crewe, R. M., and Nicolson, S. W. (2010). Convergence of carbohydrate-biased intake targets in caged worker honeybees fed different protein sources. *J. Experiment. Biol.*, 213(19):3311–3318.

Aponte, Y., Atasoy, D., and Sternson, S. M. (2011). AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat. Neurosci.*, 14(3):351–355.

Arganda, S., Nicolis, S., Perochain, A., Péchabadens, C., Latil, G., and Dussutour, A. (2014). Collective choice in ants: The role of protein and carbohydrates ratios. *J. Ins. Physiol.*, 69:19–26.

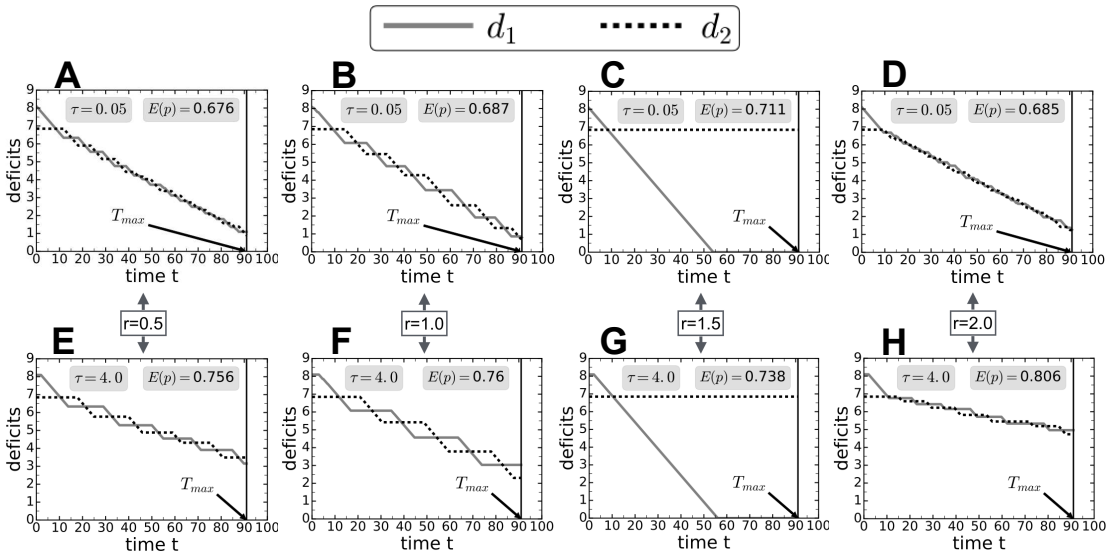

Figure B1: Effect of switching cost τ and E/I-ration r on deficit reduction. We chose a symmetric starting point at $\tau/2$. The expected penalties, $E(p)$, and switching costs (travel time between food sources), τ, are given in each plot. Lower penalty values mean better performance of the model animal. Parameter values: $d_m(t = 0) = 7.47$, $\Delta d(t = 0) = 1.25$, $\tau = 4$, $\beta = 3$, $\gamma = 0.15$, $q = 0.1$, $k = 0.8$, $k_{inh} = 0.8$, $w = 3$, $g_e = 10 = g_i$, $b_e = 0.5 = b_i$, and $\sigma = 0.01$.

Atasoy, D., Betley, J. N., Su, H. H., and Sternson, S. M. (2012). Deconstruction of a neural circuit for hunger. *Nature*, 488(7410):172–177.

Behmer, S. T. (2009). Insect Herbivore Nutrient Regulation. *Annu. Rev. Entomol.*, 54(1):165–187.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.*, 113(4):700–765.

Bose, T., Reina, A., and Marshall, J. A. R. (2017). Collective decision-making. *Curr. Opin. Behav. Sci.*, 16:30–34.

Chambers, P., Simpson, S., and Raubenheimer, D. (1995). Behavioural mechanisms of nutrient balancing in Locusta migratoria nymphs. *Anim. Behav.*, 50(6):1513–1523.

Dhooge, A., Govaerts, W., and Kuznetsov, Y. (2003). MatCont: A MATLAB package for numerical bifurcation analysis of ODEs. *ACM TOMS*, 29:141–164.

Dhooge, A., Govaerts, W., Kuznetsov, Y., Meijer, H., and Sautois, B. (2008). New features of the software MatCont for bifurcation analysis of dynamical systems. *MCMDS*, 14:147–175.

Dussutour, A., Latty, T., Beekman, M., and Simpson, S. J. (2010). Amoeboid organism solves complex nutritional challenges. *Proc. Nat. Acad. Sci.*, 107(10):4607–4611.

Dussutour, A. and Simpson, S. J. (2009). Communal Nutrition in Ants. *Curr. Biol.*, 19(9):740–744.

Essner, R. A., Smith, A. G., Jamnik, A. A., Ryba, A. R., Trutner, Z. D., and Carter, M. E. (2017). AgRP Neurons Can Increase Food Intake during Conditions of Appetite Suppression and Inhibit Anorexigenic Parabrachial Neurons. *J. Neurosci.*, 37(36):8678–87.

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.*, 9(4):292–303.

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E., Trimmer, P. C., and McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends Cogn. Sci.*, 18(3):153–161.

Gao, J. B., Tung, W.-W., and Rao, N. (2002). Noise-Induced Hopf-Bifurcation-Type Sequence and Transition to Chaos in the Lorenz Equations. *Phys. Rev. Lett.*, 89(25):254101.

Girard, B., Cuzin, V., Guillot, A., Gurney, K. N., and Prescott, T. J. (2003). A basal ganglia inspired model of action selection evaluated in a robotic survival task. *J. Integrat. Neurosci.*, 2:179–200.

Gold, J. I. and Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annu. Rev. Neurosci.*, 30(1):535–574.

Hinde, R. A. (1956). Ethological models and the concept of drive. *Br. J. Philos. Sci.*, 6:321–331.

Houston, A. and McNamara, J. (1999). *Models of Adaptive Behaviour: An Approach Based on State*. Cambridge University Press, Cambridge.

Houston, A. and Sumida, B. (1985). A positive feedback model for switching between two activities. *Anim. Behav.*, 33(1):315–325.

Houston, A. I., Higginson, A. D., and McNamara, J. M. (2011). Optimal foraging for multiple nutrients in an unpredictable environment. *Ecol. Lett.*, 14(11):1101–1107.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., and Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.*, 15(3):470–476.

Jensen, K., Mayntz, D., Toft, S., Clissold, F. J., Hunt, J., Raubenheimer, D., and Simpson, S. J. (2012). Optimal foraging for specific nutrients in predatory beetles. *Proc. R. Soc. B*.

Juel, A., Darbyshire, A. G., and Mullin, T. (1997). The effect of noise on pitchfork and Hopf bifurcations. *Proc. R. Soc. A*, 453(1967):2627–2647.

Kloeden, P., Platen, E., and Schurz, H. (2002). *Numerical Solution of SDE Through Computer Experiments*. Universitext. Springer, Berlin.

Krajbich, I., Hare, T., Bartling, B., Morishima, Y., and Fehr, E. (2015). A Common Mechanism Underlying Food Choice and Social Decisions. *PLoS Comput. Biol.*, 11(10):e1004371.

Lima, S. L. and Dill, L. M. (1990). Behavioral decisions made under the risk of predation: a review and prospectus. *Can. J. Zool.*, 68(4):619–640.

Liu, T., Kong, D., Shah, B. P., Ye, C., Koda, S., Saunders, A., Ding, J. B., Yang, Z., Sabatini, B. L., and Lowell, B. B. (2012). Fasting Activation of AgRP Neurons Requires NMDA Receptors and Involves Spinogenesis and Increased Excitatory Tone. *Neuron*, 73(3):511–522.

Ludlow, A. R. (1976). The Behaviour of a Model Animal. *Behaviour*, 58:131–172.

Marshall, J. A. R., Favreau-Peigne, A., Fromhage, L., McNamara, J. M., Meah, L. F. S., and Houston, A. I. (2015). Cross inhibition improves activity selection when switching incurs time costs. *Curr. Zool.*, 61(2):242–250.

Mayntz, D., Raubenheimer, D., Salomon, M., Toft, S., and Simpson, S. J. (2005). Nutrient-specific foraging in invertebrate predators. *Science*, 307(5706):111–113.

McFarland, D. (1999). *Animal Behaviour: Psychobiology, Ethology and Evolution*. Longman, Harlow, 3rd edition.

McNamara, J. M. and Houston, A. I. (2009). Integrating function and mechanism. *Trends Ecol. Evol.*, 24(12):670–675.

Morton, G. J., Cummings, D. E., Baskin, D. G., Barsh, G. S., and Schwartz, M. W. (2006). Central nervous system control of food intake and body weight. *Nature*, 443(7109):289–295.

Niyogi, R. K. and Wong-Lin, K. (2013). Dynamic Excitatory and Inhibitory Gain Modulation Can Produce Flexible, Robust and Optimal Decision-making. *PLoS Comput. Biol.*, 9(6):e1003099.

Pins, D. and Bonnet, C. (1996). On the relation between stimulus intensity and processing time: Piéron's law and choice reaction time. *Percept. Psychophys.*, 58(3):390–400.

Pirrone, A., Azab, H., Hayden, B., Stafford, T., and Marshall, J. (2018). Evidence for the speed-value trade-off: human and monkey decision making is magnitude sensitive. *Decision*, 5(2):129–142.

Polanía, R., Krajbich, I., Grueschow, M., and Ruff, C. C. (2014). Neural Oscillations and Synchronization Differentially Support Evidence Accumulation in Perceptual and Value-Based Decision Making. *Neuron*, 82(3):709–720.

Rangel, A. (2013). Regulation of dietary choice by the decision-making circuitry. *Nat. Neurosci.*, 16(12):1717–1724.

Ratcliff, R. and McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput.*, 20(4):873–922.

Reina, A., Bose, T., Trianni, V., and Marshall, J. A. R. (2018). Psychophysical laws and the superorganism. *Scientific Reports*, 8(1):4387.

Rho, M. S. and Lee, K. P. (2016). Balanced intake of protein and carbohydrate maximizes lifetime reproductive success in the mealworm beetle, Tenebrio molitor (Coleoptera: Tenebrionidae). *J. Ins. Physiol.*, 91-92:93–99.

Roxin, A. and Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.*, 4(3):e1000046.

Russell, D. F., Wilkens, L. A., and Moss, F. (1999). Use of behavioural stochastic resonance by paddle fish for feeding. *Nature*, 402(6759):291–4.

Sibly, R. (1975). How incentive and deficit determine feeding tendency. *Anim. Behav.*, 23:437–446.

Simpson, S. and Raubenheimer, D. (2012). *The Nature of Nutrition: A Unifying Framework from Animal Adaptation to Human Obesity*. Princeton University Press, Princeton.

Stephens, D. and Krebs, J. (1986). *Foraging Theory*. Monographs in behavior and ecology. Princeton University Press, Princeton.

Teodorescu, A. R., Moran, R., and Usher, M. (2016). Absolutely relative or relatively absolute: violations of value invariance in human decision making. *Psychon. Bull. Rev.*, 23(1):22–38.

Tong, Q., Ye, C.-P., Jones, J. E., Elmquist, J. K., and Lowell, B. B. (2008). Synaptic release of GABA by AgRP neurons is required for normal regulation of energy balance. *Nat. Neurosci.*, 11(9):998–1000.

Vong, L., Ye, C., Yang, Z., Choi, B., Chua, S., and Lowell, B. B. (2011). Leptin Action on GABAergic Neurons Prevents Obesity and Reduces Inhibitory Tone to POMC Neurons. *Neuron*, 71(1):142–154.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968.

Wiesenfeld, K. and Moss, F. (1995). Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. *Nature*, 373(6509):33–36.

Williams, K. W. and Elmquist, J. K. (2012). From neuroanatomy to behavior: central integration of peripheral signals regulating feeding behavior. *Nat. Neurosci.*, 15(10):1350–1355.

Wong, K.-F., Huk, A. C., Shadlen, M. N., and Wang, X.-J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Frontiers in Computational Neuroscience*, 1:6.

Wong, K.-F. and Wang, X.-J. (2006). A Recurrent Network Mechanism of Time Integration in Perceptual Decisions. *J. Neurosci.*, 26(4):1314–1328.

Wu, Q., Boyle, M. P., and Palmiter, R. D. (2009). Loss of GABAergic signaling by AgRP neurons to the parabrachial nucleus leads to starvation. *Cell*, 137(7):1225–34.

Wu, Q., Clark, M. S., and Palmiter, R. D. (2012). Deciphering a neuronal circuit that mediates appetite. *Nature*, 483(7391):594–597.

Yang, L., Hou, Z., and Xin, H. (1999). Stochastic resonance in the absence and presence of external signals for a chemical reaction. *J. Chem. Phys.*, 110(7):3591–3595.

Disclosure Statement

We have no disclosures or conflict of interest.