



**HAL**  
open science

## Cross inhibition improves activity selection when switching incurs time costs

James A.R. Marshall, Angélique Favreau-Peigné, Lutz Fromhage, John M. Mcnamara, Lianne F.S. Meah, Alasdair I. Houston

► **To cite this version:**

James A.R. Marshall, Angélique Favreau-Peigné, Lutz Fromhage, John M. Mcnamara, Lianne F.S. Meah, et al.. Cross inhibition improves activity selection when switching incurs time costs. *Current zoology*, 2015, 61 (2), pp.242-250. hal-01144086

**HAL Id: hal-01144086**

**<https://hal.science/hal-01144086>**

Submitted on 20 Apr 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Cross inhibition improves activity selection when switching incurs time costs

James A. R. MARSHALL<sup>1\*</sup>, Angélique FAVREAU-PEIGNÉ<sup>2,6</sup>, Lutz FROMHAGE<sup>3</sup>, John M. MCNAMARA<sup>4</sup>, Lianne F. S. MEAH<sup>1</sup>, Alasdair I. HOUSTON<sup>5</sup>

<sup>1</sup> Department of Computer Science and University of Sheffield, Sheffield S1 4DP, UK

<sup>2</sup> INRA, UMR791 Modélisation Systémique Appliquée aux Ruminants, F-75005 Paris, France

<sup>3</sup> Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

<sup>4</sup> School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK

<sup>5</sup> School of Biological Sciences, University of Bristol, Bristol Life Sciences Building, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

<sup>6</sup> AgroParisTech, UMR Modélisation Systémique Appliquée aux Ruminants, F-75005 Paris, France

**Abstract** We consider a behavioural model of an animal choosing between two activities, based on positive feedback, and examine the effect of introducing cross inhibition between the motivations for the two activities. While cross-inhibition has previously been included in models of decision making, the question of what benefit it may provide to an animal's activity selection behaviour has not previously been studied. In neuroscience and in collective behaviour cross-inhibition, and other equivalent means of coupling evidence-accumulating pathways, have been shown to approximate statistically-optimal decision-making and to adaptively break deadlock, thereby improving decision performance. Switching between activities is an ongoing decision process yet here we also find that cross-inhibition robustly improves its efficiency, by reducing the frequency of costly switches between behaviours [*Current Zoology* 61 (2): 242–250, 2015].

**Keywords** Crossinhibition, Geometric framework, Foraging, Activity selection, Neuroscience, Behaviour

For decisions that are important during an animal's life history, either due to their consequences or the frequency with which they must be made, we expect natural selection to have shaped effective decision-making and behavioural mechanisms. Here we examine the behaviour of a hypothetical animal choosing between two activities, such as feeding and drinking, mediated by motivation levels or drives for those activities. The problem facing such an animal is to undertake efficiently the activities that reduce the corresponding motivations, and a key part of this is to avoid costly switching between behaviours. One simple behavioural strategy is to always perform the activity associated with the highest motivation; in this case the animal's behavioural patterns act to bring the two motivations to the same level, at which point if the animal is without cost able to service both motivations simultaneously, or to rapidly switch between the two, then this strategy is in fact probably optimal (Houston, 2011a). In many cases however simultaneous performance of the two behaviours will be impossible, and switching between them costly, because of physical distance between food and water sources for

example. In these situations, rapid alternation between two behaviours is sub-optimal, and a trade-off between servicing both motivations and reducing the costs of switching between behaviours must be managed. A previous model that addresses this problem made use of positive and negative feedback loops to reduce the frequency of behavioural switches (Houston, 1985). An alternative approach (Ludlow, 1976; Ludlow, 1982a) incorporates inhibition between motivations to ensure that only one activity is selected at any point in time. The comparison between these models was based on their ability to reproduce observed animal behaviour, and it was shown that both models produced similar patterns of behaviour making it hard to distinguish them on solely behavioural evidence (Houston, 1985). Here, drawing inspiration from computational neuroscience, we introduce coupling between the competing drives to the positive feedback model of (Houston, 1985) and investigate the effect of this on the overall efficiency of the behavioural process, thereby providing an evolutionary perspective on which mechanisms we might expect real animals to use.

---

Received Oct.29, 2014; accepted Jan.26, 2015.

\* Corresponding author. E-mail: James.Marshall@sheffield.ac.uk

© 2015 *Current Zoology*

In computational neuroscience great progress has been made in understanding the neural basis of the simplest decision tasks, known as two-alternative forced-choice tasks, in which a subject is given noisy evidence and must use this to attempt to correctly choose one of two alternatives (reviewed in (Bogacz, 2006a)). The simplest possible decision-strategy, known as the accumulator model or race model (Vickers, 1970a), uses uncoupled decision pathways; a recent review of various neurally plausible decision-models shows, however, that when the evidence-accumulating pathways are coupled in some way, such models can approximate the statistically-optimal decision-making strategy and thus exceed the performance of the naïve race model (Bogacz, 2006a). One method of coupling decision pathways is cross inhibition, in which different neural populations mutually inhibit each other's activation, as in the influential Usher-McClelland model (Usher, 2001a). Cross-inhibition also features in asymptotically optimal models of action selection in basal ganglia (e.g. Bogacz, 2007a). As part of the optimal decision-making process cross inhibition and other similar mechanisms lead to dynamics such that if one decision pathway has high activation then the other has low activation, which facilitates decision-making. The importance of cross-inhibition has also been identified in collective decision-making, where it can also allow statistically optimal decision-making (Marshall, 2009a) or value-based deadlock breaking (Seeley, 2012a; Pais, 2013). In forced-choice scenarios a decision-maker is required to make a single choice, with reaction time and decision accuracy determining their realised payoff; here we introduce cross inhibition between competing motivations into an investigation of performance in the activity model of (Houston, 1985), in which subjects switch repeatedly between two foraging options, with switches typically being costly and with realised payoff depending on how efficiently the subject manages its intake of forage in the time available to it.

## 1 Materials and Methods

The model for two competing motivations is specified by the following pair of coupled ordinary differential equations, where  $v_i$  is the level of the  $i$ -th motivation,  $x_i$  is the level of the corresponding deficit, and  $\dot{v}_i$  and  $\dot{x}_i$  are their rates of change with respect to time:

$$\begin{aligned}\dot{v}_1 &:= c_1\dot{x}_1 + c_2x_1 + c_3v_1 + c_4v_2, \\ \dot{v}_2 &:= c_1\dot{x}_2 + c_2x_2 + c_3v_2 + c_4v_1.\end{aligned}\quad (1)$$

The animal performs the action with the greatest motivation. If action  $i$  is performed then deficit  $x_i$  is re-

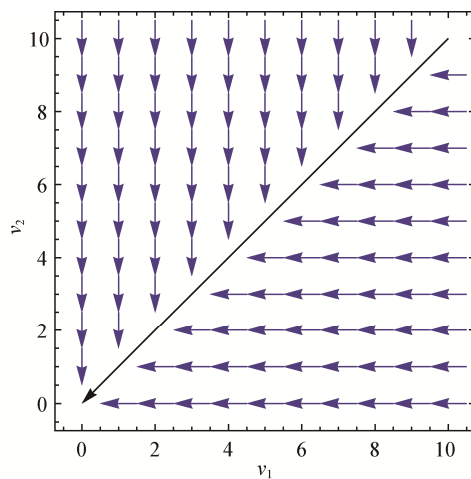
duced at rate  $\dot{x}_i = -(g - h)$ . If no action is taken to reduce it a deficit increases at rate  $\dot{x}_i = h$ , however in the current analysis we concentrate only on how an animal can efficiently reduce current levels of deficit, and thus assume that  $h = 0$ ; this corresponds to considering behaviour over a sufficiently short timeframe (Houston, 1985). Motivations and deficits are both prevented from going below zero when studied using numerical simulation. The model of equations (1) can be made equivalent to that of (Houston, 1985) by choosing  $c_2 > 0$ ,  $c_3 = -c_2$  and  $c_4 = 0$  (see Appendix A). Cross-inhibition of motivations can be added by choosing  $c_4 < 0$  in which case, with the other parameters as before, equations (1) correspond to a noise-free linear Usher-McClelland model (Usher, 2001a; Bogacz, 2006a), with the first two terms ( $c_1\dot{x}_i + c_2x_i$ ) corresponding to the sensory inputs in that model. Thus equations (1) demonstrate the close relationship between the model of activity selection of (Houston, 1985) and the connectionist model of decision-making of (Usher, 2001a). Conversely, by choosing  $c_4 > 0$  the decision-maker could exhibit cross-excitation of drives, so that an increase in one motivation also acted to increase the other motivation, and vice versa. We do not assume the particular form of cross-talk in our model, and make the minimal necessary assumptions about the remaining parameters in order to maintain equivalence with the basic positive feedback model of (Houston, 1985), including the possibility that such feedback is completely absent; indeed, the model is also able to implement the optimal strategy when switching is cost-free (Houston, 2011a), by setting  $c_1 = 1$  and  $c_2 = c_3 = c_4 = 0$  (see figure 1 and Appendix B). We then search for parameterisations that optimise the animal's behaviour under various interruption rates and costs for switching between behaviours.

To assess the performance of different parameterisations of equations (1) we model the hypothetical animal as moving in space between two locations corresponding to different activities; in this paper we will refer primarily to a source of water for drinking and a source of food for eating, but the approach is general and could equally be applied to choices between foods having different nutritional contents, or other activities for which physiological deficits and corresponding motivations should exist. The modelled animal moves towards the location corresponding to the greatest motivation. While at a location the animal reduces its corresponding deficit until the largest motivation changes, at which point the animal moves towards the other location. During the

time the animal is moving between locations it cannot take action to reduce either of its deficits, although the motivations are updated otherwise according to equations (1). The animal's initial position is set equidistant between the two locations, while the animal's initial deficits are set to not quite equal levels, since the model is deterministic and hence equal deficits would lead to equal motivations and to behavioural deadlock. The animal's behaviour is simulated for a geometrically-distributed period of time with interruption probability per-unit-time  $\lambda$ , and mean performance at interruption under this distribution is calculated. At interruption of the simulated bout the animal's performance in reducing its deficits is scored according to the penalty function

$$p := x_1^2 + x_2^2 \quad (2)$$

which has the biologically reasonable effect of rewarding the reduction of a large deficit more than an equal reduction of a small deficit. This penalty function has been used in models of animals balancing nutritional requirements (Sibly, 1976a; Houston, 2011a), and minimising it is equivalent to minimising the Euclidean distance (the square root of (2)) which has been invoked as the metric used by real organisms based on experimental data from diverse species, from unicellular organisms through to social insect colonies (Simpson,



**Fig. 1** The optimal strategy for an animal minimising its deficit in two nutrients, when there is no cost for switching between food types, is to calculate a switching line leading towards the animal's preferred nutrient levels

The appropriate food should be consumed exclusively to reach the switching line in the minimum time, then the two foods should be consumed in the optimal ratio in order to remain on the switching line (Houston, 2011a). As shown in Appendix B, for the assumptions of the present model, this switching line is the line  $x_1 = x_2$  and along the line both food types are consumed at the same rate. As also shown in Appendix B, the present model can be parameterised to implement this optimal strategy by setting  $v_1(0) = x_1(0)$ ,  $v_2(0) = x_2(0)$ ,  $c_1 = 1$ , and  $c_2 = c_3 = c_4 = 0$ .

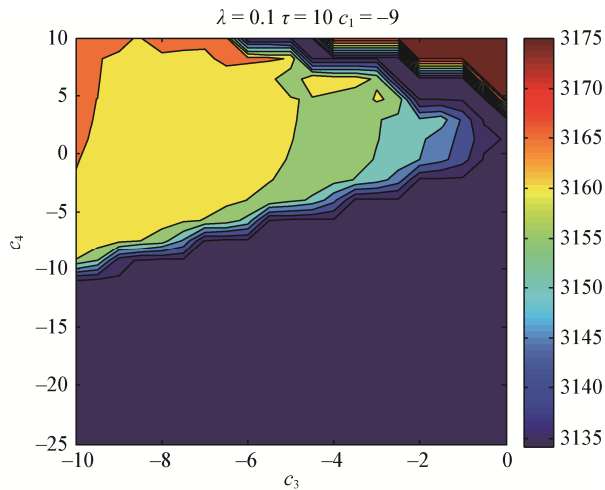
1993a; Simpson, 2004a; Dussutour, 2009a; Dussutour, 2010a; Simpson, 2012a).

To find model configurations that maximise performance under the penalty function (2) we systematically varied the parameters  $c_i$ , movement time between food and water sources  $\tau$ , and interruption probability per-unit-time  $\lambda$ . We refer to movement time  $\tau$  as the cost of switching between activities since, as described above, while moving the modelled animal is unable to reduce either of its deficits; time spent travelling thus represents an opportunity cost as the same time could have been spent performing the present activity if no decision to switch had been taken. Since equivalence with the positive feedback model of (Houston, 1985) requires that  $c_3 < 0$  and  $c_2 = -c_3$  (see Appendix A), we focussed on varying  $c_1$  (influence of rate of change of deficit on rate of change of motivation),  $c_3$  (strength of motivational decay leading to an equilibrium level of motivation for a given deficit) and  $c_4$  (strength of cross-inhibition between motivations), while treating  $\tau$  (distance between food and water sources) and  $\lambda$  (interruption probability) as parameters to be varied as part of a sensitivity analysis. The model and sensitivity parameterisations used are described in Appendix C. As described in the appendix, we calculated expected penalty under interruption probability  $\lambda$  by numerical integration over 99% of the geometric distributions' probability mass functions.

## 2 Results

We performed an extensive sensitivity analysis using numerical simulations of the model, by systematically varying the cost to switch behaviours and the interruption rate, each over an order of magnitude. Selected results from the sensitivity analysis using numerical simulations of the model are presented in figure 2. Inspection of these results (see Online Supplementary Information) reveals a pattern that is preserved throughout the costs and interruption rates studied; illustrated by a representative plot (figure 2), the lowest values of the penalty function are consistently associated with cross-inhibition ( $c_4 < 0$ ) and with parameterisations where the speed of equilibration of positive feedback is less than the strength of cross-inhibition ( $|c_3| < |c_4|$ ). An exhaustive search over all combinations of switching cost and interruption rate revealed no cases in which the global optimum within the studied parameter space violated the aforementioned inequalities (see Online Supplementary Information).

Having established the general location of the global minima, we next considered the sensitivity of the op-



**Fig. 2 Representative sensitivity analysis using numerical simulation of the model**

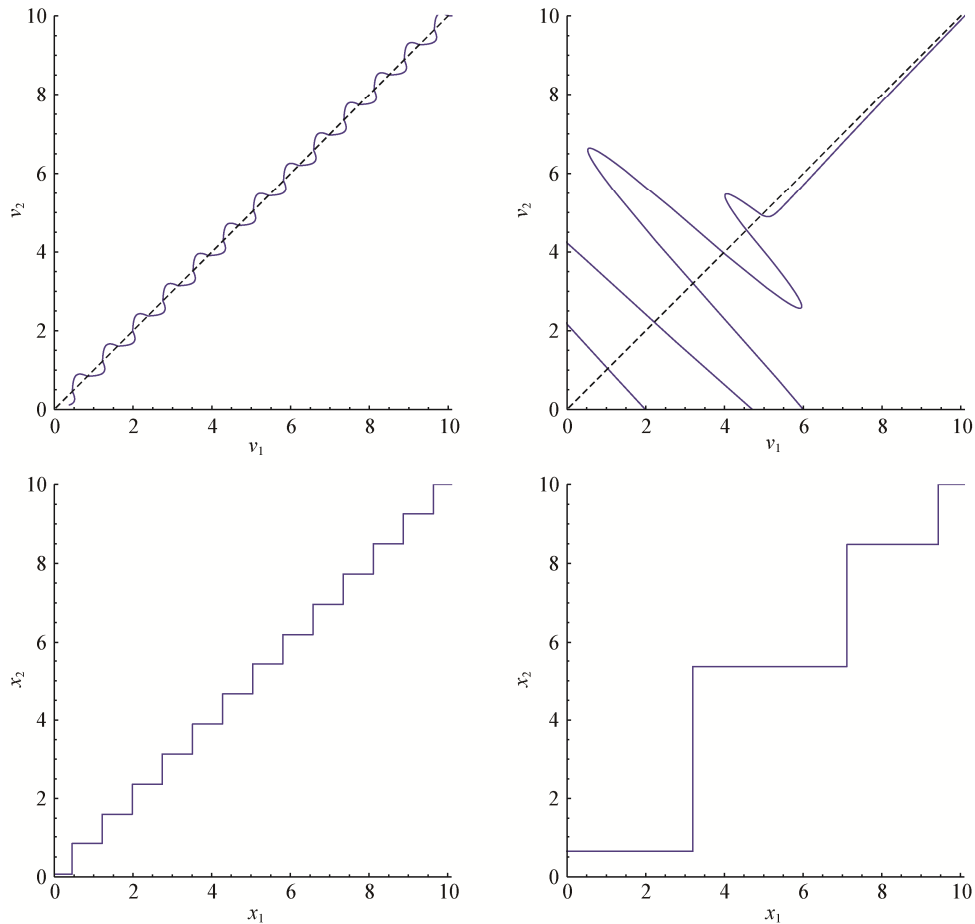
Parameter  $c_2 = -c_3$  as described in ‘Methods’. Other parameters were varied as shown in the figures. Darker colours correspond to lower values of the penalty function at the end of the simulated time period, and hence better performance of the modelled animal. In all parameter spaces a ‘wedge’ of low penalty parameterisations is observed, covering the space where  $c_4 < c_3 < 0$ .

timal parameterisation of an animal’s behavioural strategy under changes in switching cost and interruption rate. That is, we examined whether the optimal values  $c_1^*$ ,  $c_3^*$  and  $c_4^*$  could easily be found given different values of  $\tau$  and  $\lambda$ . We discovered multiple points within the parameter space produced the lowest global penalty so, assuming increased values of  $c_1$  through  $c_4$  to be physiologically costly for an animal to implement, chose the global optimum with the lowest Euclidean distance to the  $c_1 = c_2 = c_3 = c_4 = 0$  origin of the parameter space. Unfortunately, the exact position of this efficient global optimum varies unpredictably as switching costs and interruption rate change (see Online Supplementary Information). However, across all parameters studied in the sensitivity analysis, all the points within the ‘wedge’ of the parameter space described by  $c_1 < 1$ ,  $c_3 < 0$ ,  $c_4 < c_3$  were within 3% of the discovered optimum, whereas the distance between the best and worst performances in the studied parameter space, averaged across all scenarios in the sensitivity analysis, was over 12%.

The reason why cross inhibition improves the efficiency of the model can be understood by examining the dynamics of the motivations with and without cross inhibition. Figure 3 (left) shows how the standard positive feedback model of (Houston, 1985), which has no cross inhibition, functions to reduce deficits. The model’s behaviour is characterised by regular and frequent

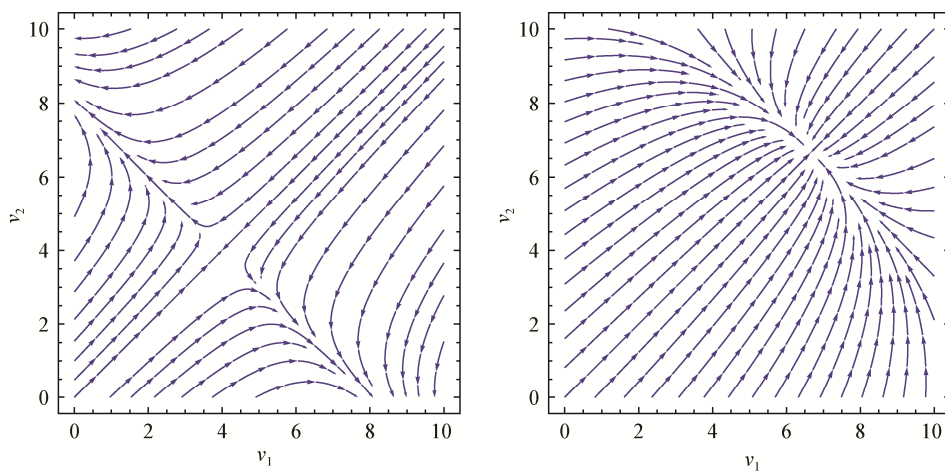
switches between the two activities which, since switching is costly in the present model, reduces efficiency. In contrast, figure 3 (right), which has the same behavioural parameters except for the inclusion of cross-inhibition, reduces the frequency of switching between activities, because motivations act to suppress each other. The reason for this suppression is explored in more detail below, in discussing the second pattern observed in the sensitivity analysis of figure 2.

The second pattern observed is that, when  $c_4$  is negative and hence there is some level of cross-inhibition, having  $c_3 > c_4$  improves behavioural performance (Fig. 2). This is explained due to the parallels of the behavioural model with that of the linear Usher-McClelland model (Usher, 2001a; Bogacz, 2006a); multiplying the parameters  $c_3$  and  $c_4$  by minus one gives, respectively, the decay and the inhibition rates of the Usher-McClelland model, and it has been shown that when decay is less than inhibition the system becomes unstable, driving one of its integrators to maximal activation and the other to minimal activation (Bogacz, 2006a). Since integrators in the Usher-McClelland model correspond to motivations in the present model, we can see why this behaviour helps reduce rapid switching, since it moves the motivations apart so that one is low and the other is high (Fig. 4A); note that (Bogacz, 2006a) also showed that when the converse condition holds, i.e. decay exceeds inhibition in the Usher-McClelland model, or  $c_3 < c_4$  in the present model, activations (motivations) converge to stable non-zero levels (Fig. 4B) which will increase the frequency of behavioural switching since the level of motivations are held closer together by the dynamics of the model. The superiority of inhibition-dominance for activity selection in the Usher-McClelland model contrasts with its parameterisation for approximating statistically-optimal sequential tests, which is shown to occur when decay equals inhibition ( $c_3 = c_4$ ) and the two-dimensional model can be reduced to the one-dimensional drift diffusion process (Bogacz, 2006a). The analytic approach of Bogacz et al. makes use of a rotation of the co-ordinate system, to track the summed activation of the integrators (our motivations, i.e.  $v_1 + v_2$ ) and the difference in activation of the integrators (i.e.  $v_1 - v_2$ ). Given we make use of the same equations, and given the importance of the switching line  $v_1 - v_2 = 0$  in our model, this same rotation would provide an alternative way of conceptualising the dynamics presented above.



**Fig. 3** Top row: illustrative trajectories of the motivations over time with (right) and without (left) cross inhibition (dashed diagonal line represents point of switching between activities). Bottom row: corresponding trajectories of the deficits over time with (right) and without (left) cross inhibition

See Houston(1985) for introduction of the plotting methodology; time starts at the top-right of the deficit plots, and progresses downwards and leftwards; since intake rate is fixed and identical at the two patches, the length of vertical and horizontal lines in the deficit plots is exactly proportional to the duration of the corresponding bout). Without cross inhibition (left) behaviour is characterised by regular bouts of each activity, and frequent switches between activities. With cross inhibition (right) activity bouts become irregular, and frequency of costly activity-switching is reduced resulting in improved behavioural performance. Parameterisations: (all)  $g = 0.2$ ,  $c_1 = -1$ ,  $c_2 = 1$ ,  $c_3 = -1$  (left)  $x_1(0) = v_1(0) = 10$ ,  $x_2(0) = v_2(0) = 10.1$ ,  $c_4 = 0$  (right)  $x_1(0) = v_1(0) = 10$ ,  $x_2(0) = v_2(0) = 10.1$ ,  $c_4 = -1$ .



**Fig. 4** Dynamics of the motivations over time when motivational decay is less than cross inhibition (left) and greater than cross inhibition (right)

When decay is less than cross inhibition the motivational levels are unstable and, excluding the influence of feeding or drinking on drives, will converge to a situation where one is high and the other low, which reduces the tendency for behavioural switching. Parameterisations: (both)  $x_1 = x_2 = 10$ ,  $c_1 = 0$ ,  $c_2 = 2$ ,  $c_3 = -2$  (left)  $c_4 = -3$  (right)  $c_4 = -1$ .

### 3 Discussion

Our results suggest consistent selective pressure for the evolution of activity selection mechanisms utilising cross-inhibition of drives. To conclude we now relate these findings to previous models of activity selection. A previous model by Houston and Sumida (Houston, 1985) was proposed to reduce dithering in behavioural switches, which was also our aim, but did not consider the possibility for interaction between drives. We find cross-inhibition, combined with Houston and Sumida's 'positive feedback' is optimal. Houston and Sumida also wished to reproduce behavioural data on which behaviour follows an interruption of variable length. They concluded that both their model, and an earlier model proposed by Ludlow (Ludlow, 1976), explain such data.

Ludlow (Ludlow, 1976) described a model in which the current activity inhibits others, until such point as another activity exceeds this inhibition, at which point it becomes the inhibiting activity, and the other activities are inhibited. This differs from our model in which cross-inhibition between motivations always occurs, but the strength with which it occurs varies according to the strength of each motivation.

More recently, Houston and colleagues (Houston, 2011a) showed how, without a cost for switching between food types, an animal should move towards a switching line in nutrient space, then move along it. This differs from our model, because we include costs for switching between behaviours. If costs are negligible, moving towards and along the switching line should indeed be optimal in the current model. Since the model of (Houston, 2011a) was motivated by the geometric framework for nutritional decision-making (Simpson, 1993a; Simpson, 2012a), and since switching between food types may well incur a time or other cost for many animals, our results may have relevance for that literature; various experimental studies have shown that increasing physical distances between nutritional sources leads to longer visits to those sources and reduced switching frequency, indicating that these species' behavioural strategies do indeed take account of switching costs (Larkin, 1978a; Bernays, 1997a; Zee, 2002a; Behmer, 2003a; Seaman, 2008a). There are other contexts in which the cost of switching influences the frequency of switches. A well-studied case is the decision to leave a depleting patch in order to find a new patch. The marginal value theorem (Charnov, 1976a; Stephens, 1986a) specifies the time to leave that results in the highest long-term rate of energy gain. This optimal time in-

creases as the time to find a new patch increases; in this case, the animal is not necessarily switching between two alternatives although some lab tests might use two patches. Another example involves choice between two operant schedules. On variable interval schedules, the probability of reward on choosing a schedule increases with time since that schedule was last chosen. The pattern of choice between two variable interval schedules that gives the highest long-term rate of energy gain involves a decrease in switching rate as the time required to make a switch increases (Houston, 1981a), and empirical data show this trend (Baum, 1982a; Boelens, 1983a). While there are similarities in these approaches and results, there may also be important differences. In the case of patch-use, it is optimal to switch because the rate of gain in a patch is decreasing, whereas in the case of variable interval schedules it is optimal to switch because the probability of reward on the new schedule is increasing. In the environment that we have modelled, however, switching avoids large costs following interruption. Since there are different functional reasons for switching behaviour in these examples, there may well also be different mechanisms.

Our model assumes that one resource satisfies only one deficit. This is unrealistic, and is relaxed in (Houston, 2011a). Having different resources satisfy multiple deficits may help explain observations such as deprived animals feeding before drinking (since many foods contain water as well) (Mayer, 1972a; McFarland, 1973a), when taken together with different weightings in the utility of the different resources to the animal (since in our model both are assumed to be equally valuable; because of this weights for the two motivations are identical). Our model also assumes, as others before have, that resource levels do not interact in determining the eventual fitness of an animal; relaxing this assumption may lead to benefits of, for example, cross-excitation of drives. Our model is also unrealistic in that interactions between deficits and motivations, and between motivations, are all linear; linearity is the exception rather than the rule in physiology and neurophysiology, and it could be of interest to consider non-linear models in future work.

Finally, it is worth considering the behavioural predictions of our model. In particular, the model predicts irregular activity bouts. Data on these have previously been collected for barbery doves (Sibly, 1975a; Sibly, 1976a) and in some cases exhibit qualitatively similar behaviour to that exhibited by our model (e.g. figure 1 in Sibly (1976b) and figure 3 herein), albeit under a

different interruption schedule. We have not addressed the issue of what the optimal strategy is, in the absence of mechanistic constraints on behaviour, for an animal seeking to minimise expected penalty in our scenario. However the pattern of behaviour exhibited in figure 3 (lower-right), of progressively increasing bout durations, contrasts with models of activity selection proposed in the computational neuroscience literature; for example (Humphries, 2002a) shows a pattern of decreasing activity bout durations as two motivations compete. In fact, as we show in Online Supplementary Information, our model is also capable of exhibiting this pattern of behaviour. We leave the problem of determining the optimal mechanism-free strategy for our scenario to possible future work.

**Acknowledgements** The authors thank Alex Cope, Tim Fawcett and Andy Higginson for helpful discussions. This work was supported in part by INRA PHASE (mobility grant to AF-P) and the European Research Council (Advanced Grant 250209 to AIH).

## References

- Baum WM, 1982. Choice, changeover, and travel. *Journal of the Experimental Analysis of Behavior* 38: 35–49.
- Behmer ST, Cox E, Raubenheimer E, Simpson SJ, 2003. Food distance and its effect on nutrient balancing in a mobile insect herbivore. *Animal Behaviour* 66: 665–675.
- Bernays EA, Angel JE, Augner M, 1997. Foraging by a generalist grasshopper: The distance between food resources influences diet mixing and growth rate (Orthoptera: Acrididae). *Journal of Insect Behavior* 10: 829–840.
- Boelens H, Ko PFM, 1983. Concurrent schedules: Spatial separation of response alternatives. *Journal of the Experimental Analysis of Behavior* 40: 35–45.
- Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD, 2006. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review* 113: 700.
- Bogacz R, Gurney K, 2007. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation* 19: 442–477.
- Charnov EL, 1976. Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9: 129–136.
- Dussutour A, Latty T, Beekman M, Simpson SJ, 2010. Amoeboid organism solves complex nutritional challenges. *Proceedings of the National Academy of Sciences* 107: 4607–4611.
- Dussutour A, Simpson SJ, 2009. Communal nutrition in ants. *Current Biology* 19: 740–744.
- Houston AI, Higginson AD, McNamara JM, 2011. Optimal foraging for multiple nutrients in an unpredictable environment. *Ecology Letters* 14: 1101–1107.
- Houston AI, Sumida B, 1985. A positive feedback model for switching between two activities. *Animal Behaviour* 33: 315–325.
- Houston AI, McNamara JM, 1981. How to maximize reward rate on two variable-interval paradigms. *Journal of the Experimental Analysis of Behavior* 35: 367–396.
- Humphries MD, Gurney KN, 2001. The role of intra-thalamic and thalamocortical circuits in action selection. *Network: Computation in Neural Systems* 13: 131–156.
- Larkin S, McFarland D, 1978. The cost of changing from one activity to another. *Animal Behaviour* 26: 1237–1246.
- Ludlow AR, 1976. The behaviour of a model animal. *Behaviour* 58: 131–172.
- Ludlow AR, 1982. Towards a theory of thresholds. *Animal Behaviour* 30: 253–267.
- Marshall JAR, Bogacz R, Dornhaus A, Planqué R, Kovacs T et al., 2009. On optimal decision-making in brains and social insect colonies. *Journal of the Royal Society Interface* 6: 1065–1074.
- Mayer AJ, Cox VC, 1972. Serum osmolality and meal initiation in food and water-deprived rats. *Behavioral Biology* 7: 435–439.
- McFarland DJ, Lloyd IH, 1973. Time-shared feeding and drinking. *The Quarterly Journal of Experimental Psychology* 25: 48–61.
- Pais D, Hogan PM, Schlegel T, Franks NR, Leonard NE et al., 2013. Value-sensitive decision-making in a model of collective choice. *PLoS ONE* 8(9): e73216.
- Seaman SC, Waran NK, Mason G, D'Eath RB, 2008. Animal economics: Assessing the motivation of female laboratory rabbits to reach a platform, social contact and food. *Animal Behaviour* 75: 31–42.
- Seeley TD, Visscher PK, Schlegel T, Hogan PM, Franks NR et al., 2012. Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science* 335: 108–111.
- Sibly RM, McCleery RH, 1976. The dominance boundary method of determining motivational state. *Animal Behaviour* 24: 108–124.
- Sibly R, 1975. How incentive and deficit determine feeding tendency. *Animal Behaviour* 23: 437–446.
- Sibly R, McFarland D, 1976. On the fitness of behavior sequences. *American Naturalist* 110: 601–617.
- Simpson SJ, Raubenheimer D, 1993. A multi-level analysis of feeding behaviour: The geometry of nutritional decisions. *Philosophical Transactions of the Royal Society B* 342: 381–402.
- Simpson SJ, Sibly RM, Lee PK, Behmer ST, Raubenheimer D, 2004. Optimal foraging when regulating intake of multiple nutrients. *Animal Behaviour* 68: 1299–1311.
- Simpson SJ, Raubenheimer D, 2012. *The Nature of Nutrition: A Unifying Framework from Animal Adaptation to Human Obesity*. Princeton: Princeton University Press.
- Stephens DW, Krebs JR, 1986. *Foraging Theory*. Princeton: Princeton University Press.
- Usher M, McClelland JL, 2001. The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review* 108: 550.
- Vickers D, 1970. Evidence for an accumulator model of psychophysical discrimination. *Ergonomics* 13: 37–58.
- Zee B, Behmer ST, Simpson SJ, 2002. Food mixing strategies in the desert locust: Effects of phase, distance between foods, and food nutrient content. *Entomologia Experimentalis et Applicata* 103: 227–237.



## Appendix A: Equivalence with Houston and Sumida's Positive-Feedback Model

Here we show how the model of equations 1 can be parameterised for equivalence with (Houston1985)'s model of activity-selection mediated by positive feedback. Houston and Sumida consider a 'tendency' to eat or drink (equivalent to drives in the current work) which is equal to the sum of the current deficit (for food or water respectively) and a positive feedback term ((Houston, 1985), equation 1). In the notation of the current paper, for a motivation  $i$  we thus have

$$v_i(t) := x_i(t) + f_i(t) \quad (3)$$

where  $f_i(t)$  is the strength of positive feedback at time  $t$ . Differentiating with respect to time gives

$$\dot{v}_i := \dot{x}_i + \dot{f}_i. \quad (4)$$

Full equivalence with the model of (Houston, 1985) requires that motivations and deficits are set to the same initial levels; i.e.  $v_i(0) = x_i(0)$ . The change in positive feedback over time  $\dot{f}_i$  is defined as

$$\dot{f}_i := -\beta_i \dot{x}_i - \alpha_i f_i(t), \quad (5)$$

where  $\beta_i$  and  $\alpha_i$  are parameters controlling the positive-feedback loop through, respectively, its strength and the speed with which it reaches equilibrium (Houston, 1985). Now we take the implicit definition of  $f_i(t)$  from equation 3, and substitute this and equation 5 into equation 4 to obtain

$$\dot{v}_i := \dot{x}_i(1 - \beta_i) + \alpha_i x_i - \alpha_i v_i. \quad (6)$$

Note that we have now omitted time indices for  $v_i$  and  $x_i$ , as in the main paper text. Equation 6 with  $i \in \{1, 2\}$  gives equations 1 provided that  $c_1 = 1 - \beta$ ,  $c_2 = \alpha$ ,  $c_3 = -c_2$  and  $c_4 = 0$ . Note that we have also here omitted subscripts from  $\alpha$  and  $\beta$  and thus assume that these are equal for the two activities.

Defining our cross-inhibition model as an extension of the positive-feedback model of (Houston, 1985) constrains our search for optimal parameterisations. From the equivalence shown above we see that, since  $\alpha > 0$  and  $\beta > 0$  in the original model (Houston, 1985), we must have that  $c_1 < 1$  and  $c_3 < 0$ , and furthermore  $c_3$  uniquely defines  $c_2 = -c_3$ . The parameter  $c_1$  can take positive or negative values, since there is no upper bound on  $\beta$ . If  $c_1 = 1$  this corresponds to a model with no positive feedback, whereas if  $c_1 < 1$  but  $c_2 = c_3 = 0$  the model exhibits runaway positive feedback, that does not reach equilibrium.

## Appendix B: Equivalence with Houston et al.'s Zero Switching Cost Model

Houston et al. consider an idealised model in which an animal can simultaneously consume two foods, each of which provides a different ratio of two required nutrients (Houston, 2011a), and there is no cost for doing so. The objective of the animal is to minimise their Euclidean distance in nutrient space from their target nutrient levels, when their feeding is interrupted. Houston et al.'s analysis shows that the optimal strategy is to consume the food that is richest in the nutrient the animal has the greatest deficit in, thereby moving towards a switching line in nutrient space. The switching line leads towards the target nutrient levels; once on this switching line it is optimal to remain on it by consuming both food types in a ratio determined by the slope of the switching line in nutrient space (Houston, 2011a). Our model is a simplified version of that of Houston et al. in which the two foods each contain only one nutrient, food and water respectively. Furthermore each nutrient is assumed to be equally valuable to the animal in the present model, since the nutrient target to reach is  $x_1^* = x_2^* = 0$ , and consuming each food type reduces its corresponding deficit at the same rate  $-g$ . If  $x_{1A}$  is the amount of nutrient 1 associated with consuming food type  $A$ ,  $x_{2A}$  the amount of nutrient 2 associated with consumption of food A, and so on, then in the present model we have  $x_{1A} = x_{2B} = -g$  and  $x_{1B} = x_{2A} = 0$ .

Houston et al. show that the slope of the switching line in nutrient space is given by

$$\frac{x_{1A} - x_{1B}}{x_{2B} - x_{2A}} \quad (7)$$

(their eq. 8 and immediately following). Substituting in the values defined above, the slope is seen to be 1. This translates into a ratio of resource consumption when on the switching line given by

$$\frac{x_{2B}(x_{2B} - x_{2A}) - x_{1B}(x_{1A} - x_{1B})}{(x_{1A} - x_{1B})^2 + (x_{2A} - x_{2B})^2} \quad (8)$$

(eq. 11 in (Houston, 2011a)). Substituting in the values defined above, equation 8 simplifies to  $1/2$ , indicating that the two food types are consumed at equal rates.

The optimal strategy for the present model, when there is no cost for switching between food types, is thus to converge as quickly as possible to the line  $x_1 = x_2$ , then to move along this equally exploiting the two available food types. Given that the initial conditions of the model are  $v_i(0) = x_i(0)$  for  $i \in \{1, 2\}$ , this can be achieved by choosing parameters  $c_1 = 1$ , and  $c_2 = c_3 = c_4 = 0$  as illustrated in figure 1.

## Appendix C: Numerical Simulations and Results

To generate the results presented in the main paper we simulated an animal behaving under an interruption protocol, with geometrically-distributed bouts defined by an interruption probability per-unit-time  $\lambda$ . The expected penalty for an animal was then calculated as

$$E(p) := \sum_{T=1}^{\tilde{T}} p(T)(1-\lambda)^{(T-1)} \lambda \quad (9)$$

where  $T$  is the terminal time of the bout, and  $p(T)$  the penalty (defined according to equation 2) experienced by the animal behaving according to its decision-rule up to time  $T$ . Since our numerical simulation is deterministic a single run is sufficient for each value of  $T$  in the calculation of  $E(p)$  according to equation 9. The maximum terminal time  $\tilde{T}$  was chosen such that 99 percent of the geometric probability distribution was included in the calculation of equation 9. Initial deficits were chosen to be sufficiently large that consistent satisfaction of one over an entire simulated epoch of 'typical' duration (up to 100s) would not result in that deficit being reduced to 0.