Neural Substrates Mediating the Utility of Instrumental Divergence

Kaitlyn G. Norton (kgnorton@uci.edu)
Department of Psychological Sciences, UC Irvine
Irvine, CA 92697

Mimi Liljeholm (m.liljeholm@uci.edu)
Department of Cognitive Sciences, UC Irvine
Irvine, CA 92697

Abstract

We assessed the neural substrates mediating a recently demonstrated preference for environments with high levels of instrumental divergence – a formal index of flexible operant control. Across choice scenarios, participants chose between gambling environments that differed in terms of both instrumental divergence and expected monetary pay-offs. Using model-based fMRI, we found that activity in the ventromedial prefrontal cortex scaled with a divergence-based measure of expected utility that reflected the value of both divergence and monetary reward. Implications for a neural common currency for information theoretic and economic variables are discussed.

Keywords: instrumental divergence; flexible control; utility; model-based fMRI

Introduction

A series of recent studies (Mistry & Liljeholm, 2016; Liljeholm et al., 2018) have demonstrated that individuals strongly prefer environments in which instrumental divergence – the degree to which alternative actions differ with respect to their outcome probability distributions – is relatively high. A high level of instrumental divergence is a necessary feature of flexible control: If all available action alternatives have identical, or very similar, outcome distributions, such that selecting one action over another does not significantly alter the probability of any given outcome state, an agent's ability to exert flexible control over its environment is considerably impaired. Conversely, when available action alternatives produce distinct outcomes, discrimination and selection between actions allow an agent to flexibly obtain the currently most desired outcome. Since subjective outcome utilities often change from one moment to the next, flexible instrumental control is essential for reward maximization and, as such, may have intrinsic value, serving to reinforce and motivate decisions that guide the organism towards high-agency environments (Liljeholm, 2018). In the current study, we investigate the neural substrates mediating the apparent preference for high instrumental divergence.

Previous work suggests that the ventromedial prefrontal cortex (vmPFC) retrieves and ranks the values of decision outcomes, and that these value signals are subsequently used to compute decision values (see O'doherty, 2011 for review). Intriguingly, activity in the vmPFC scales with the values of a wide variety of goods, including food, money, books DVDs, and clothes, suggesting a common neural value-scale for distinct stimulus categories (Chib et al., 2009; McNamee et al., 2013). It is unknown, however, whether this common value-scale might also extend to more abstract, cognitive, commodities, such as instrumental divergence. Here, using a task in which participants choose between gambling environments based on differences in both instrumental divergence and monetary pay-offs, we combine computational cognitive modeling with functional MRI to investigate neural representations of the utility of instrumental divergence.

Method

Participants

Twenty undergraduates at the University of California, Irvine (11 females; mean age = 21.2 ± 4.65) participated in the study for monetary compensation. The sample size was determined based on an a priori power analysis of data from a previously published study (Mistry & Liljeholm, 2016), indicating that 18 subjects were required to demonstrate a clear behavioral preference for high instrumental divergence at a power of 90% given a 0.05 threshold for statistical significance. All participants gave informed consent and the Institutional Review Board of the University of California, Irvine, approved the study.

Task & Procedure

The task is illustrated in Figure 1. At the start of the experiment, participants were instructed that they would assume the role of a gambler in a casino, playing a set of four slot machines (i.e., actions, respectively labeled A1, A2, A3, and A4) that yielded three different colored tokens (blue, green and red), each worth a particular amount of money, with different probabilities. They were further told that, in each of several blocks, they would be required to first select a room in which only two slot-machines were available, and that they could only choose between the two machines in the selected room on subsequent trials in that block. Finally, participants were instructed that, while the outcome probabilities would remain constant throughout the study, the values of the tokens would change at various times, and these changes might occur after the participant had already committed to a particular pair of machines in a given block. Consequently, although changes in value were explicitly announced, and the current values of tokens were always printed on their surface, a participant might find themselves in a room in which the values of the two available actions had suddenly been altered.

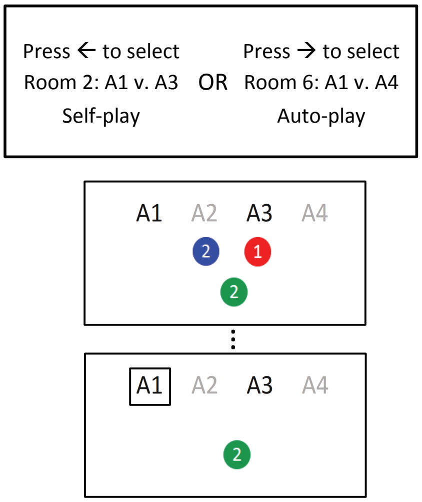

Figure 1: Task illustration showing the room-choice screen at the beginning of a block (top), and the choice (middle) and feedback (bottom) screens on a trial in the selected room.

Two distinct probability distributions over the three possible token outcomes were used and the assignment of outcome distributions to slot machines was such that two of the machines (either A1 and A2 or A1 and A3, counterbalanced across subjects) always shared one distribution, while the other two machines shared the other distribution. This yielded a low (zero) divergence for rooms in which the two available slot machines shared the same probability distribution, and a relatively high divergence for rooms in which slot machines had different outcome probability distribution. The unpredictability (i.e., Shannon entropy) of outcomes given a particular machine was held constant for all machines. Three token-reward distributions were used, changing intermittently across blocks, such that expected monetary pay-offs were either the same across rooms, or differed across rooms in either the same or opposite direction of instrumental divergence. In addition to mimicking dynamic changes in the utilities of natural rewards, the sporadic changes in token reward values across blocks allowed us to pit the value of instrumental divergence against that of monetary reward.

Given a constant outcome entropy level, increases in instrumental divergence are accompanied by increases in the perceptual diversity of obtainable outcomes – a variable previously shown to elicit preferences in economic tasks (Ayal & Zakay, 2009). To rule out perceptual diversity as an explanation for any effects of instrumental divergence, gambling rooms differed in terms of whether the participant was allowed to chose freely between slot machines in the room (self-play) or a computer algorithm alternated between machines across trials in that room (auto-play). In auto-play rooms, participants were still required to press a key corresponding to the slot machine indicated by the computer, to control for movement execution. Critically, in the absence of voluntary choice, high-divergence no longer yields flexible instrumental control. However, the alternating computer algorithm still yields greater perceptual diversity in high- than in low-divergence rooms. Consequently, if choices were driven by a desire to maximize perceptual diversity, rather than instrumental divergence, they should not differ depending on whether the participant or an alternating computer algorithm choose between the slot machines in a room. In addition to controlling for perceptual diversity, this self- vs. auto-play manipulation relates the preference for instrumental divergence to a well-established preference for free over forced choice (e.g., Leotti & Delgado, 2011).

There were a total of 44 blocks, with participants choosing between two gambling rooms at the start of each block (the decision of interest), followed by 3-5 gambling trials within the selected room. The order different reward distributions, and of room choice scenarios, was counterbalanced across subjects. Before starting the gambling task participants were given a practice session in order to learn the probabilities with which each slot machine produced the different colored tokens. If a participants' estimate of any given probability deviated by more than 0.2 from the programmed probability, they were returned to the beginning of the practice phase, and this continued until all rated probabilities were within 0.2 points of programmed probabilities. At the end of the study, participants again provided estimates of the action-token probabilities.

Computational Models

Instrumental divergence is formalized as the Jensen-Shannon divergence of instrumental sensory-specific outcome probability distributions (Liljeholm et al., 2013). Let P_1 and P_2 be the respective outcome probability distributions for two available actions, let O be the set of possible outcomes, and $P(o)$ the probability of a particular outcome, o. The instrumental divergence (ID) is:

$$ID = \frac{1}{2}\sum_{o \in O} \log\left(\frac{P_1(o)}{P_*(o)}\right) P_1(o) + \frac{1}{2}\sum_{o \in O} \log\left(\frac{P_2(o)}{P_*(o)}\right) P_2(o),$$

where

$$P_* = \frac{1}{2}\left(P_1 + P_2\right)$$

We defined the *expected value* (EV) of each slot machine as the sum over the products of its transition probabilities and token utilities. In turn, the expected *monetary* value of a gambling room is simply the mean of the EVs of slot machines in that room. To model the utility of instrumental divergence, a second variant of EV was specified by adding the term $w*ID$ to the expected monetary value of a room, where the free parameter w represents the subjective utility of instrumental divergence and ID is the divergence of the particular room. Thus, in this variant, the EV of a room reflects *both* the monetary pay-off and the instrumental divergence associated with that room. For both models, a softmax distribution with a noise parameter, τ, was used to translate expected room values into choice probabilities, and free parameters were fit to behavioral data by minimizing the negative log likelihood of observed choices. Choice scenarios in which at least one room option was both high divergence and self-play (HDSP), yielding high *instrumental* divergence, and those in which the high-divergence room option was auto-play, or both rooms had zero divergence (HDAP), were modeled separately. The corrected Akaike Information Criterion (AICc) was used for behavioral model comparisons.

Neuroimaging Acquisition & Analyses

All MR images were obtained in a 3T Siemens Prisma Scanner, fitted with a 32-channel RF receiver head coil, padded to minimize head motion, at the facility for imaging and brain research (FIBRE) at the University of California, Irvine. Functional images covered the whole brain with 48 continuous 3-mm thick axial slices with T2*-weighted gradient echoplanar imaging (TR=2.65s, TE=28ms, 3-mm^2 in-plane voxel size, 64 x 64 matrix). All participants had a high-resolution structural image taken before functional scanning commenced (T1-weighted FSPGR sequence: 208 continuous 0.8-mm axial slices 0.4-mm^2 in-plane voxel size; 640 x 640 matrix). All stimulus materials were presented, and all responses recorded, using MATLAB. All imaging data was preprocessed with MATLAB and SPM12. Functional images were preprocessed with standard parameters, including slice timing correction, spatial realignment, coregistration of the high-resolution structural image to functional images, segmentation of the structural image into tissue types, spatial normalization of functional images into MNI space, and spatial smoothing with an 8mm FWHM kernel.

All imaging data was analyzed using MATLAB and SPM12. At the first level, two general linear models (GLMs) were specified for each participant. In both GLMs, two regressors respectively specified the onsets of room choice screens for HDSP and HDAP choice scenarios. In the first GLM, these onsets were parametrically modulated by the absolute difference between rooms in their expected monetary pay-offs; in the second GLM these onsets were parametrically modulated by the absolute difference between rooms in their divergence-based utility, which reflected both the monetary pay-off and the level of divergence associated with each room. In addition, in both GLMs, two regressors indicated the onsets of choice screens on each trial within a selected room, for self-play and auto-play rooms respectively, and each of these were parametrically modulated by the expected monetary value of the chosen slot machine. Finally, both GLMs included a single regressor indicating the onsets of trial feedback screens, modulated by the monetary reward obtained on each trial, as well as regressors indicating separate scanning runs and accounting for the residual effects of head motion.

Fixed effects models were estimated using restricted maximum likelihood and an AR(1) model for temporal autocorrelation. Group-level statistics were generated by entering contrasts of first level parameter estimates into between-subject analyses. All effects are reported at a whole brain corrected $p < 0.05$ level, using cluster size thresholding (CST) to adjust for multiple comparisons. AlphaSim, a Monte Carlo simulation, was used to determine cluster size and significance. For an individual voxel probability threshold of $p=0.005$, a minimum cluster size of 148 MNI transformed voxels resulted in an overall significance of $p < 0.05$.

Results

Behavioral Results

Participants required on average 2.1 (*SD*=0.3) cycles of practice on the action-token probabilities. Mean probability ratings, obtained right before and right after the gambling phase, are shown in Table 1.

Table 1: Mean probability ratings with standard deviations. Programmed probabilities are shown in the top row. Mean ratings, obtained before and after the gambling task, are averaged across identical objective probabilities, yielding three unique values.

	0.7	**0.0**	**0.3**
Before	0.69 ± 0.06	0.00 ± 0.00	0.31 ± 0.05
After	0.67 ± 0.10	0.00 ± 0.00	0.32 ± 0.05

The decision of interest was that at the beginning of each block, when participants choose between rooms that differed in terms of their divergence, expected monetary pay-offs and self- vs. auto-play. Model-derived choice probabilities and AICc scores for these decisions are listed in Table 2.

Table 2: Mean room-choice probabilities derived using the divergence-based and conventional models of expected value (EV), and associated AICc scores, for HDSP and HDAP choice scenarios, with standard deviations.

	Choice Probabilities		AICc Scores	
	HDSP	**HDAP**	**HDSP**	**HDAP**
Divergence EV	0.65 ± 0.13	0.58 ± 0.07	19.5 ± 2.7	38.3 ± 7.4
Conventional EV	0.55 ± 0.05	0.57 ± 0.07	21.6 ± 5.6	36.8 ± 7.5

A repeated measures analysis of variance (ANOVA) revealed that the model-derived probabilities of observed behavioral choice preferences were significantly greater for the divergence-based utility algorithm than for the conventional utility model, and this difference was significantly greater for HDSP choice scenarios than for HDAP choice scenarios, yielding a significant main effect of EV model, F(1,19)=12.40, $p<0.005$, as well as a model by choice scenario interaction, F(1,19)=9.52, $p<0.01$. Accordingly, there was also a significant interaction for the AICc scores, F(1,19)=7.71, $p<0.05$, such that scores were significantly lower, indicating a better fit, for the conventional than for the divergence-based utility model in HDAP blocks (t(19)=5.2, $p<0.001$) while being lower for the divergence-based utility model, albeit with only marginal significance (p=0.14) in HDSP blocks.

Neuroimaging Results

As with the behavioral data, the period of interest was the choice made at the beginning of each block, between two gambling rooms that differed in terms of divergence, monetary pay-offs and free choice. As illustrated in Figure 2, neural activity in the ventromedial prefrontal cortex (vmPFC) was parametrically modulated by the absolute difference in divergence-based EV between rooms, when at least one room option was both high-divergence and self-play (HDSP) but not when the high divergence room was auto-play, or both room options had zero divergence (HDAP). No significant effects of the difference between room options in expected monetary pay-offs emerged in this region. A similar pattern of results, with activity scaling selectively with the absolute difference in divergence-based EV between rooms options in HDSP choice scenarios, was found in the middle frontal gyrus, as well as the premotor cortex. Once a room had been selected, activity in a more dorsal aspect of the vmPFC, extending into the dorsal medial prefrontal cortex scaled with the expected monetary pay-off of the chosen slot machine, in self-play but not in auto-play rooms, as did activity in the lateral orbitofrontal cortex, posterior right middle temporal gyrus and right dorsolateral prefrontal cortex.

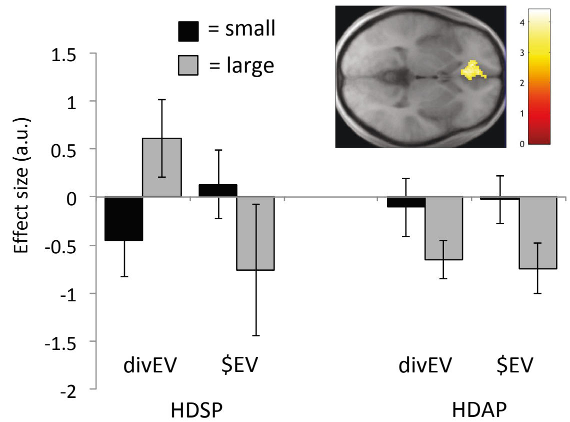

Figure 3: Map of the t-statistics for a test of differential parametric modulation by the difference across rooms in divergence-based expected value (divEV) for choice scenarios in which at least one high-divergence room option was self-play (HDSP) versus those in which the high-divergence room option was auto-play, or both options had zero-divergence (HDAP), showing significant effects in the vmPFC. Bar plots show effect sizes (y-axis) extracted from 4 mm spheres centered on the peak coordinate (x, y, z = -4, 34, -4), for small and large differences in divEV and monetary expected values ($EV), in HDSP and HDAP choice scenarios. Error bars=SEM.

Discussion

Countless studies on motivated behavior have investigated the neural representation of primary and monetary rewards (Abe & Lee, 2011; Abler et al., 2009; Belova et al., 2007; Cador e al., 1989). Here, having previously demonstrated a behavioral preference for instrumental divergence – a formal index of flexible operant control – we explored neural substrates mediating the influence of this information theoretic variable on economic choice. Specifically, participants were scanned with fMRI as they chose between gambling rooms that differed with respect to both instrumental divergence and expected monetary pay-offs. Using a model-based analysis, we found that activity in the ventromedial prefrontal cortex (vmPFC) scaled with a divergence-based measure of expected utility that reflected both instrumental divergence and monetary pay-offs.

Considerable evidence from neurophysiological and neuroimaging studies suggest that the vmPFC encodes the subjective values of primary rewards, such as tastes and odors (Rolls et al., 2003; Anderson et al., 2003; Small et al., 2003), as well as visual stimuli, including the attractiveness of faces or pictorial scenes (O'Doherty et al., 2003; Kirk et al., 2009), and more abstract goods, like social praise (Elliot et al., 1997) and monetary gain (O'Doherty et al., 2001).

Two notable features of the vmPFC shed important light on the current results: First, value encoding in the vmPFC appears to be relative, such that the value signal for a particular stimulus depends on the values of other, proximal, stimuli (O'Doherty, 2011). One might expect, thus, that the vmPFC signal would respond most clearly to a *difference in value* between concurrently available stimuli. Second, recent findings suggest that the vmPFC encodes stimulus values that are independent of the particular stimulus category, essentially implementing a common neural value scale for different types of goods (Chib et al., 2009; McNamee et al., 2013). The currently demonstrated value signal in the vmPFC, corresponding to a difference between options in divergence-based utility, suggest that this common value scale can be extended to a relative analysis of exceedingly abstract concepts.

Our previous work has implicated the right supramarginal gyrus (rSMG) of the inferior parietal lobule in encoding instrumental divergence. Specifically, using a simple value-based decision-making task, Liljeholm et al. (2013) found that activity in the rSMG scaled parametrically with trial-by-trial estimates of instrumental divergence, and that this signal was dissociable from other information theoretic and motivational variables, including outcome entropy and expected utility. In a subsequent task, aimed at assessing neural substrates mediating the acquisition of goal-directed vs. habitual instrumental behavior, Liljeholm et al., (2015) found that activity in the rSMG increased across blocks of instrumental acquisition in a high-divergence, but not in a zero-divergence, condition. In contrast, we did not find any effects of instrumental divergence in the rSMG in the current study. There are several possible reasons for this discrepancy: First, none of the previous studies assessed the motivational significance of instrumental divergence, in terms of a behavioral preference for environments with relatively high divergence. Second, in the current study, outcome probability distributions were trained to criterion prior to scanning (eliminating acquisition effects), and instrumental divergence remained constant within a room (eliminating responses to trial-by-trial fluctuations in divergence). Further work is needed to determine how these differences may account for a differential engagement of the rSMG.

A fundamental property of stimuli that possess intrinsic value is their ability to transfer that valence to neutral stimuli with which they are paired – a phenomenon termed *conditioned reinforcement,* that has been studied extensively using a wide range of stimuli, species and procedures (e.g., Arroyo et al., 1998; Williams, 1994). This large body of research has demonstrated that conditioned reinforcers are powerful behavioral determinants, maintaining instrumental responding in the absence of primary rewards, such as food and sex, and even serving as goals in themselves. Moreover, once established, previously neutral conditioned reinforcers can pass on their motivational significance to other neutral stimuli; For example, casino chips maintain gambling based on their association with monetary reward, which in turn obtains valence from its usefulness in acquiring primary rewards. One might expect, therefore, that any sufficiently valuable stimulus, no matter how abstract, should be able to induce conditioned reinforcement in associated arbitrary, and initially neutral, stimuli. Another important question, thus, is whether the affective properties of instrumental divergence may transfer to concomitant stimuli, and what brain regions might mediate such a processes.

Formal theories of goal-directed decisions postulate that the agent generates a "cognitive map" of stochastic relationships between actions and states such that, for each action in a given state, a probability distribution is specified over possible outcome states. These transition probabilities are then combined with current estimates of outcome utilities in order to generate action values – the basis of goal-directed choice (Doya et al., 2002). Although computationally expensive (Otto et al., 2013), the dynamic binding of outcome probabilities with utilities offers adaptive advantage over more automatic action selection, which uses cached values based on reinforcement history. However, when instrumental divergence is zero, or very low, the processing cost of goal-directed computations does not yield the return of flexible control, suggesting that a less resource-intense automatic decision strategy might be optimal. As noted, in a previous study we found evidence implicating instrumental divergence in the deployment of goal-directed and habitual behavior, and this is an important avenue for future work.

In summary, we have used model-based fMRI to investigate the neural computations mediating a behavioral preference for instrumental divergence. We found that activity in the vmPFC was significantly modulated by a variant of expected value that reflected both instrumental divergence an monetary pay-offs, but not by a conventional model of expected value, based solely on monetary gain. Our results complement previous work on the role of the vmPFC in value-based choice.

Acknowledgements

This work was supported by a CAREER grant from the National Science Foundation (1654187) awarded to Mimi Liljeholm.

References

Abe, H. & Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**, 731-741, doi:10.1016/j.neuron.2011.03.026 (2011).

Abler, B., Herrnberger, B., Gron, G. & Spitzer, M. From uncertainty to reward: BOLD characteristics differentiate signaling pathways. *BMC Neurosci* **10**, 154, doi:10.1186/1471-2202-10-154 (2009).

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., ... & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature neuroscience*, *6*(2), 196.

Arroyo, M., Markou, A., Robbins, T. W. & Everitt, B. J. Acquisition, maintenance and reinstatement of intravenous cocaine self-administration under a second-order schedule of reinforcement in rats: effects of conditioned cues and continuous access to cocaine. *Psychopharmacology* **140**, 331-344 (1998).

Ayal S, Zakay D (2009) The perceived diversity heuristic: the case of pseudodiversity. Journal of personality and social psychology 96:559-573.

Belova, M. A., Paton, J. J., Morrison, S. E. & Salzman, C. D. Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* **55**, 970-984 (2007).

Cador, M., Robbins, T. W. & Everitt, B. J. Involvement of the amygdala in stimulus-reward associations: interaction with the ventral striatum. *Neuroscience* **30**, 77-86 (1989).

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, *29*(39), 12315-12320.

den Ouden, H. M., Frith, U., Frith, C. & S.J., B. Thinking about intentions. *NeuroImage* **28**, 787-796 (2005).

Doya K, Samejima K, Katagiri K, Kawato M (2002) Multiple model-based reinforcement learning. Neural computation 14:1347-1369.

Elliott, R., Frith, C. D., & Dolan, R. J. (1997). Differential neural response to positive and negative feedback in planning and guessing tasks. *Neuropsychologia*, *35*(10), 1395-1404.

Ernst, M. *et al.* Choice selection and reward anticipation: an fMRI study. *Neuropsychologia* **42**, 1585-1597, doi:10.1016/j.neuropsychologia.2004.05.011 (2004).

Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *Neuroimage*, *44*(3), 1125-1132.

Leotti, L. A., & Delgado, M. R. (2011). The inherent reward of choice. *Psychological science*, *22*(10), 1310-1318.

Liljeholm, M., Wang, S., Zhang, J., & O'Doherty, J. P. (2013). Neural correlates of the divergence of instrumental probability distributions. *Journal of Neuroscience*, *33*(30), 12519-12527.

Liljeholm, M., Dunne, S., & O'doherty, J. P. (2015). Differentiating neural systems mediating the acquisition vs. expression of goal-directed and habitual behavioral control. *European Journal of Neuroscience*, *41*(10), 1358-1371.

Liljeholm, M., Mistry, P., & Koh, S. (2018). The Influence of Schizotypal Traits on the Preference for High Instrumental Divergence. *Cog Sci 2018 Proceedings*.

Liljeholm, M. (2018). Instrumental Divergence and Goal-Directed Choice. In *Goal-Directed Decision Making* (pp. 27-48). Academic Press.

McNamee, D., Rangel, A., & O'doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature neuroscience*, *16*(4), 479.

Mistry, P., & Liljeholm, M. (2016). Instrumental Divergence and the Value of Control. *Scientific reports*, *6*, 36295.

O'Doherty, J. P. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Annals of the New York Academy of Sciences* **1239**, 118-129, doi:10.1111/j.1749-6632.2011.06290.x (2011).

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature neuroscience*, *4*(1), 95.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, *41*(2), 147-155.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A. & Daw, N. D. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A* **110**, 20941-20946, doi:10.1073/pnas.1312011110 (2013).

Rolls, E. T., Kringelbach, M. L., & De Araujo, I. E. (2003). Different representations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience*, *18*(3), 695-703.

Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, *39*(4), 701-711.

Williams, B. A. Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic bulletin & review* **1**, 457-475, doi:10.3758/BF03210