Elsevier Editorial System(tm) for Biological Psychiatry

Manuscript Draft

Manuscript Number:

Title: Commentary on BDNF and the orbitofrontal regulation of behavior

Article Type: Commentary

Corresponding Author: Professor. Trevor Robbins, PhD

Corresponding Author's Institution: University of Cambridge

First Author: Trevor Robbins, PhD

Order of Authors: Trevor Robbins, PhD

Manuscript
Click here to view linked References

Commentary on BDNF and the orbitofrontal regulation of behavior (Zimmerman et al, this issue)

TW Robbins, Dept of Psychology and Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge U.K.

Whereas contemporary neuroimaging evidence supports a role for dysfunctioning prefrontal-amygdaloid-striatal circuitry in affective disorders, precise psychological functions of this interaction have still to be revealed, thus providing barriers to producing new treatments for these debilitating conditions. The great advantage of murine models of psychiatric disorders is the availability of the novel neuroscience tools including optogenetics and DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) which enable the molecular features of specific functional circuits identified in human neuroimaging studies to be further characterised and novel molecular targets for drug discovery to be identified. Possible drawbacks of mouse models, however, include unclear homology at both the anatomical and functional levels. Thus, can we be sure that the relevant circuits in the mouse correspond to those in the primate brain? And can we be sure that the rather crude models currently used in animal models of depression are of any use in modeling much more complex human psychopathology?

A welcome feature of recent work is that concepts originally derived from animal learning theory are gradually becoming recognized as relevant to human neuropsychiatric disorders, including depression. One historical example is that of learned helplessness, where exposure to uncontrollable and unpredictable events (generally aversive) leads to future impairments in voluntary behavior which may mimic some aspects of apathy and even anhedonia in humans. The advance of reinforcement learning theory [1] has further emphasized the distinction drawn between at least two forms of instrumental learning: goal-directed behavior and stimulus-response habit learning, which also engage distinct, though undoubtedly interacting, fronto-striatal systems [1, Figure 1]. A number of methods have evolved for ‘diagnosing’ imbalance in these systems; these include ‘devaluing the goal’ for example using poisoning of food or enhancing satiety to particular flavors, and also destroying the contingent relationship

between actions and their outcomes, so that the former no longer predict the latter; this is termed ‘contingency degradation’ [1]. Depression could be hypothesized to involve a loss of goal-directed behavior and a corresponding switch to habitual control. What the latter actually would mean for understanding the psychological nature of depression, as compared with other examples of loss of goal-directed behavior, such as addiction and OCD [2], is an interesting matter for conjecture. As parallel paradigms can model these forms of associative learning in rodents and humans, these methods can be used to enhance translation between the preclinical models and human clinical disorders. This is the commendable approach that has been deployed elegantly in the study by Zimmerman et al [7].

The neurobiological questions previously asked by this group have previously included the influence of key factors such as stress and neuronal plasticity- and their interactions [3]. Chronic stress has been shown to shift goal-directed behavior to habitual control, consistent with its role in depression. Neuronal plasticity has been established to underpin instrumental learning, although mechanisms underlying habit learning and the switch between the two are only now beginning to be elucidated [4]. Hence, this investigation [7] of BDNF and its biochemical pathways in the control of goal-directed behavior has a strong rationale. An innovative intervention here was to test effects of the tyrosine receptor kinase B agonist 7,8-dihydroxyflavone (DHF), which increases dendritic spine densities, on the goal-directed /habitual behavioral balance. The study thus comprised an exemplary multi-faceted investigation of neural circuitry underlying goal-directed learning that utilized anterograde tracing, viral-mediated gene silencing, functional disconnection strategies, pharmacological treatment, and DREADDs to determine anatomical and functional connectivity of the ventrolateral (vl-) orbitofrontal cortex (OFC) and basolateral amygdala (BLA).

OFC-amygdala disconnection produced by unilateral vl-OFC BDNF knockdown plus lesioning of the contralateral BLA robustly impaired discrimination between responding at an aperture for which responding is no longer necessary for food delivery and an alternative

aperture where the instrumental contingency had not been so degraded, in a manner similar to bilateral knockdown of BDNF in the vlOFC. DREADD mediated inhibition of the OFC produced similar effects. DHF administered to mice that had been over-trained according to a special schedule for producing habits exhibited rescue of their goal-directed behavior concomitant with enhanced dendritic spine density on excitatory neurons in the OFC.

However, these exciting results also raise more questions than they address. One controversy emerging from the literature provoked further by this study is whether the rodent OFC is indeed implicated in instrumental (i.e. voluntary, goal-directed behavior) in the face of other evidence implicating this region in Pavlovian conditioning. There is evidence that the rodent vl-OFC mediates stimulus-outcome learning, but *not* action-outcome learning [6]. Thus, using the goal devaluation procedure to probe for instrumental habits, there was no effect of vl-OFC lesions but impairments in outcome specific pavlovian-to-instrumental transfer (PIT), where the motivational effects of presenting a Pavlovian CS are expressed by increases in instrumental responding in a food reinforcement procedure [6]. This conclusion is also compatible with evidence of effects of Pavlovian devaluation of stimuli in rhesus monkeys with lesions of specific regions (Brodmann areas (B.A.) 11 and 13) in the rhesus macaque [5]. Moreover, a considerable body of evidence in humans and rodents has apparently implicated areas of *ventromedial* (rather than ventrolateral) PFC structures in instrumental goal-directed behavior and instrumental contingencies, including such structures as the medial OFC and BA 14 and 32 [1]. Although there are major complications of homology imposed by comparisons between the rodent and primate brain – i.e. that specific regions of the rodent PFC correspond to specific regions of the human PFC, there does appear to be genuine discrepancy, perhaps related to the different procedures (devaluation versus contingency degradation) by which habits are detected. Other evidence [4] has supported a role for the vl-OFC in action-outcome learning using a reinforcement schedule training method for inducing habits in mice. Chemogenetic inhibition (via DREADDs) of OFC disrupted goal-directed actions, whereas optogenetic activation of OFC specifically increased goal-directed lever-

pressing (via OFC-striatal interactions). Moreover, the training methodology was validated by goal devaluation assessment of habit learning. In discussing the obvious discrepancy with the rat findings [6], Gremel and Costa [4] speculate their procedure to be more sensitive for detecting habits, or alternatively that "inhibiting a single action following devaluation recruits different neural mechanisms than the choice behaviour between two outcomes (albeit one devalued) observed following training with two actions and two outcomes" (p9). The issue, however, remains unresolved and there is a further complication posed by the data presented by Zimmerman et al in this, and in earlier, publications [3,7]. Whilst discrimination of contingency degradation is impaired in these mice, the prediction of a specific impairment in the latter is actually not upheld. This is because, as I see it, the deficit in discrimination clearly depicted in several of the experiments is not produced by an increase in responding in the degraded contingency condition, but by a *reduction* in responding in the non-degraded condition. There thus must be some other factor that causes a reduction in this responding not at all necessarily due to contingency degradation *per se*; perhaps it is motivation or extinction-related. Thus, further interpretation of these data is warranted. As ever, the behavioral or functional analysis is primary in understanding the significance of these elegant neurobiological manipulations for human mental health disorders, and should not be ignored.

One might question whether such controversies of interpretation matter in the face of otherwise clear findings regarding effects of manipulating important mechanisms of neuronal plasticity. I would argue that there are implications for understanding the psychological nature of depression, for example whether we need to impose control over our voluntary tendencies in depressive states and whether this can be modified by pharmacological treatments (perhaps including BDNF type mechanisms). It also matters in terms of how we try to understand how behavior is controlled by the brain and the current modal model of multiple control systems whose output has to be controlled and co-ordinated seamlessly together to produce the coherent expression of behavioral and cognitive output.

Acknowledgement My work is supported by the Wellcome Trust.

References

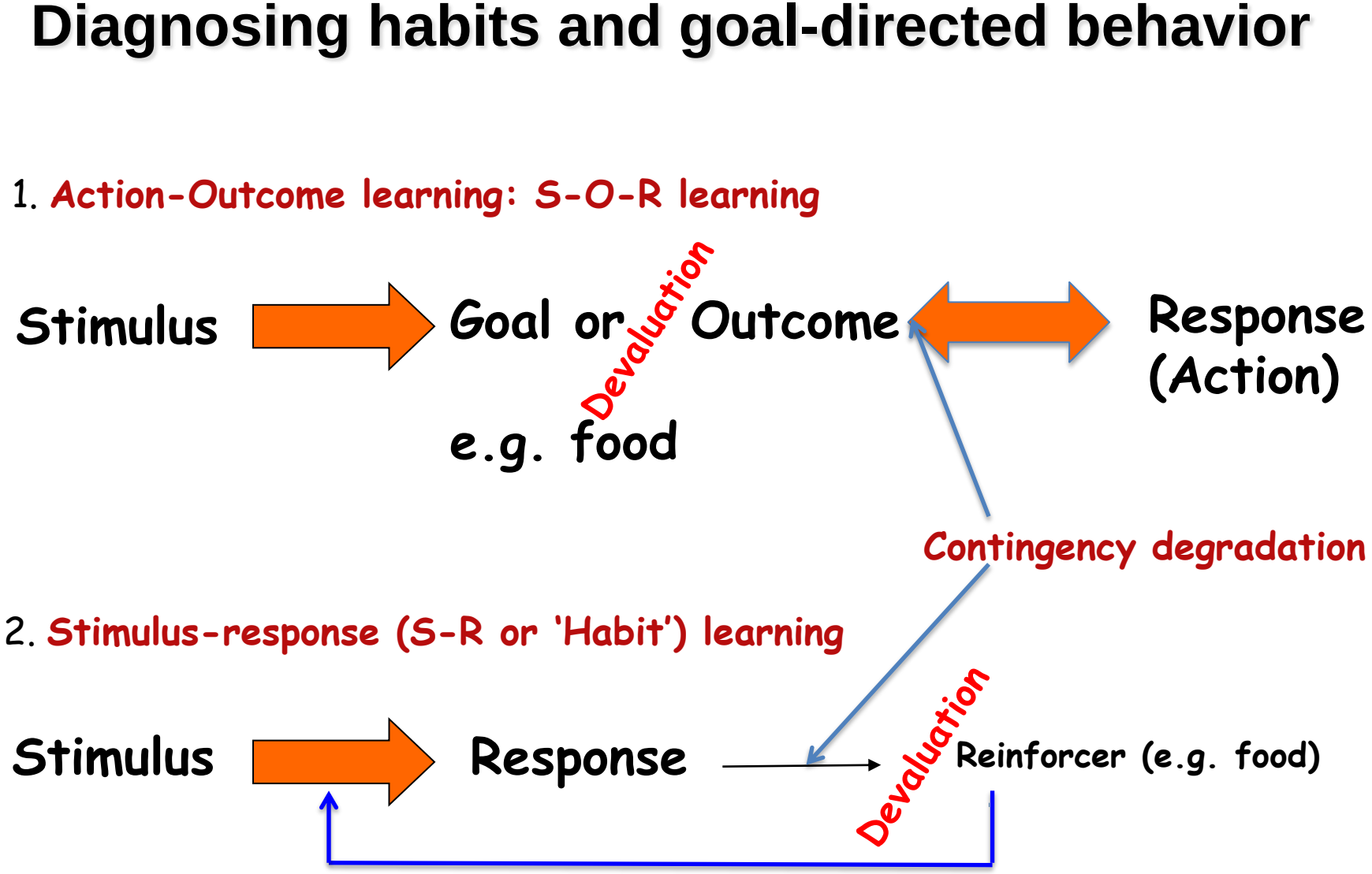
1. Balleine BW, O'Doherty (2011) Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacol 5: 48-69
2. Gillan, CM, Robbins, TW. (2014). Goal-directed learning and obsessive-compulsive disorder. Philos Trans R Soc Lond B Biol Sci 369
3. Gourley SL, Swanson AM, Jacobs AM, Howell JL, Mo M, Dileone RJ, et al. (2012): Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. Proc Natl Acad Sci U S A 109: 20714–20719.
4. Gremel CM, Costa RM (2013) Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nature Communications 4: 2264
5. Murray EA, Wise SP (2010) Interactions between orbital prefrontal cortex and amygdala: Advanced cognition, learned responses and instinctive behaviors. Curr Opinion Neurobiol 20:2012-2020.
6. Ostlund SB, Balleine BW, Oribitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J. Neurosci 27: 4819-4825.
7. Zimmerman KS, Yamin JA, Rainnie DG, Ressler KJ, Gourley SL (2016) Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by brain-derived neurotrophic factor. Biological Psychiatry, http://dx.doi.org/10.1016/j.biopsych.2015.10.026

Main text plus references 1491 words

Figure legend

The two main forms of instrumental behavior; action-outcome learning (or 'control') and stimulus-response habit learning (or 'control') which may proceed in parallel and via distinct neural systems jointly determine behavioral output [1]. Action-outcome learning depends on two main associations: between environmental stimulus and goal (or the currently valued outcome or 'reward') and between the valued outcome and the action with which it is associated. This enables voluntary behavior, such as working for the goal or outcome of food. In habit learning, by contrast, stimuli directly elicit the response without the need for a representation of the goal. In this case the food outcome functions as a reinforcer that promotes stimulus-response learning. The relative balance between goal-directed behavior and habits can be determined by two distinct probe tests; *goal devaluation* and *contingency degradation*. In the former case, the goal is devalued and so weakens the link between outcome and action. However, such devaluation would have no effect on habitual performance as the currently valued outcome does not participate in these associations. Similarly, instrumental contingency degradation or the weakening of the correlation between action (A) and outcome (O), i.e. towards the state in which P(O/A) = P(O/no A), also reduces goal-directed behavior for the degraded contingency and if instrumental responding persists, is assumed to be habitual.

Figure


Diagnosing habits and goal-directed behavior
1. Action-Outcome learning: S-O-R learning
Stimulus
Goal or Outcome
e.g. food
Devaluation
Response (Action)
Contingency degradation
2. Stimulus-response (S-R or 'Habit') learning
Stimulus
Response
Reinforcer (e.g. food)
Devaluation