3

Corticostriatal Intrusions

Bernard W. Balleine

Abstract

The loop-like circuits which link the cortex and basal ganglia have been implicated in a range of functions; most recently in the precursors to movement, including planning and decision making. Damage to these circuits induced by various disease states have, therefore, been heavily implicated in a range of symptoms, including intrusive involuntary thoughts and actions associated with, for example, neurodegenerative and psychiatric conditions as well as addictions of various kinds. This chapter focuses on recent evidence of parallel circuits that mediate the distinct forms of control associated with reflexive and volitional actions, and the interactions between these circuits in determining adaptive behavior. It discusses two kinds of interaction important for understanding intrusive actions and thoughts: competitive interactions, whereby circuits controlling volitional actions regulate reflexive or habitual responses, and cooperative processes that allow the simulation of specific actions to become manifest in performance. It then explores the role of information derived from predictive learning in action selection and choice. The influence of such information is conveyed through a specific corticobasal ganglia circuit, damage to which has been implicated in compulsive action. The evidence considered generally suggests that intrusive thoughts and actions are the product of an imbalance between corticobasal ganglia circuits rather than dysfunction in any one circuit or its related control process.

Introduction

Intrusive thoughts and actions are highly debilitating symptoms associated with a range of psychiatric disorders and addictions. There are numerous theories regarding the psychological, behavioral, and neural determinants of such intrusions; however, in recent years, a growing theme has been to link these conditions to the processes associated with involuntary or reflexive actions, commonly known as habits (Everitt and Robbins 2016; Robbins et al. 2019). The distinction between volitional, goal-directed actions and reflexes is one with which students of adaptive behavior have grappled for millennia. If, today, this issue feels more tractable, it is because of advances in our understanding of the behavioral processes and, consequently, the neural

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

networks that support these forms of action control. For the purposes of this chapter, I will focus on two aspects of these networks that appear to have important implications for future research into these issues. The first lies in the relationship between the neural networks that mediate goal-directed and habitual control: how they compete and cooperate to support our everyday activities, and the consequences of failures in these interactive processes. The second lies in the way an even more fundamental reflexive system mediating Pavlovian-conditioned reflexes interacts with goal-directed processes to provoke the initiation and performance of otherwise volitional actions in ways that can often appear maladaptive. Although both undoubtedly relate to the habitual control of intrusions, neither wholly depends on such factors. Both, however, highlight how actions and habits are integrated and, ultimately appear to suggest that failures, when they occur, are failures in that integrative process.

Actions and Habits

It is now well recognized that the performance of instrumental actions (i.e., those actions through which we manipulate the environment) is subject to two forms of control process, typically referred to as goal-directed and habitual control. Goal-directed actions are determined by their relationship to and the value of their consequences or outcome. They are rapidly acquired through a learning process sensitive to the causal relationship between action and outcome, and they uniquely engage a process of cognitive-emotional integration in linking these causal relations to goal values based on the motivational and emotional processes through which the incentive values of specific outcomes are established (Dickinson 1994; Dickinson and Balleine 1994; Balleine 2001). As such, goal-directed control is effortful and costly; it engages considerable cognitive and emotional resources but provides fast solutions that are highly flexible (i.e., they can be executed or withheld on demand). Habits, on the other hand, are determined by their antecedents rather than their consequences. They are acquired more slowly through a process of sensorimotor association with specific associations selected by reinforcement. They require few, if any, cognitive resources for their acquisition or performance and are emotionally subject to a reinforcement signal (during acquisition) and to the net level of arousal or drive (during performance) (Dickinson and Balleine 2002). Habits are released, rather than executed, by environmental events and are difficult to withhold once released. They are, however, highly organized actions that have a similar topography across repetitions; thus, they make an adaptive, low-cost solution to common or routine problems. Importantly, these two forms of action control have not only been found to be mediated by distinct learning rules, associative structures, and emotional feedback processes (cf. Dickinson 1994); they also engage distinct neural processes.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Common Tests of Goal-Directed Action

Much of the evidence for distinct forms of action control has been derived from the use of a battery of tests of goal-directed control through which the sensitivity of animals, whether rodent or human, to changes in the action–outcome relationship and in outcome value is evaluated (see Dickinson and Balleine 1994; Balleine and Dickinson 1998; Balleine 2005). One such test, the *contingency degradation test*, assesses sensitivity to the relationship between an action and its consequences by increasing the probability of access, p, to a specific outcome, O, in the absence of an action, A, or $p(O|noA)$. Given a specific probability that an action earns an outcome (i.e., $p(O|A)$), $p(O|noA)$ can be increased until the outcome is equally probable if the action is performed or not. At that point, the action is no longer causally related to the outcome and, if the actor is sensitive to this, then the actor should no longer perform the action (Hammond 1980). Increasing $p(O|noA)$ beyond that point means that the action prevents the outcome, which should reduce responding even faster, something demonstrated through the use of omission schedules (Davis and Bitterman 1971). Goal-directed action in humans and animals is exquisitely sensitive to these manipulations of the action–outcome relationship, largely because these manipulations alter the causal relationship between these events (Dickinson 2012). As a consequence, not only is performance affected, judgments as to how causal an action is, with respect to its specific outcome, are similarly modified by these contingency manipulations (Shanks and Dickinson 1991). Finally, these changes in contingency are highly selective; altering the relationship between one action and outcome does not affect performance based on the causal relationship between other actions and their outcomes (Balleine and Dickinson 1998).

The second commonly used test is called *outcome devaluation*, which assesses sensitivity of performance to changes in the value of the outcome of an action and is usually conducted without feedback (i.e., in extinction). Hence, having learned a number of specific action–outcome relationships (A1 → O1, A2 → O2, etc.), the value of one outcome can be altered, after which the subject is given a choice between the various actions. Goal-directed control reflects the ability to integrate the action–outcome relationship with the altered value of the outcome to modify performance of the action. Again, considerable evidence has demonstrated that goal-directed actions in both humans and rodents are sensitive to these kinds of treatment (e.g., Balleine and Dickinson 1998; Balleine and O'Doherty 2010).

Neural Bases

Studies of goal-directed action have found evidence that the acquisition and performance of such actions depend on the rich connections of the prefrontal cortex (PFC), including the human ventromedial PFC and medial orbitofrontal

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

cortices, with the striatum, specifically the caudate nucleus (Balleine and O'Doherty 2010). The greater resolution allowed by rodent studies has significantly refined this general picture. Thus, it is now clear that the homologous region of the medial PFC in rats, the prelimbic cortex (area 32), is engaged during the early acquisition of goal-directed actions (Hart and Balleine 2016). Furthermore, whereas the input layers 2/3 contribute significantly to the consolidation of goal-directed learning, the output layers (particularly the intratelencephalic neurons in layer 5) are critical to this process. Activity in their bilateral projection is necessary for learning-related plasticity in the posterior dorsomedial striatum (one of their main targets) and thus for the acquisition of goal-directed actions (Hart et al. 2018a, b).

Using the above tests, we demonstrated, some time ago, that rodents with damage to the prelimbic cortex, the posterior dorsomedial, or the mediodorsal thalamus show deficits in their sensitivity to contingency degradation and outcome devaluation, and continued to respond as if neither the contingency nor outcome value had changed. Evidence that these structures share a high degree of interconnectivity led to the claim that they constitute the critical corticostriatal-thalamo-cortical loop through which goal-directed actions are encoded (see Figure 3.1a; Balleine 2005). Importantly, this same insensitivity to shifts in the action–outcome relationship and in outcome value is found in habitual actions. Early in training, actions are sensitive to changes in contingency and value, whereas with continued practice, performance naturally becomes insensitive to these changes as action control shifts from the consequences of an action to antecedent environmental stimuli with which the action becomes associated (Dickinson et al. 1995, 1998). Chief among these stimuli is the context in which the action is trained (Thrailkill and Bouton 2015).

At a neural level, the acquisition and performance of habits depends on a corticostriatal circuit linking the sensorimotor regions of frontal cortex with the putamen or dorsolateral striatum (Figure 3.1a). Again, treatments causing degeneration to, or temporary inactivation of, structures in this circuit block the acquisition of habit learning and render habitual actions goal directed; that is, despite extensive training, they continue to show sensitivity both to shifts in the action–outcome relationship and outcome value (Yin et al. 2004, 2006). Generally, therefore, these findings have been interpreted as suggesting that the goal-directed and habitual control of instrumental actions is a competitive process: any reduction in goal-directed control increases the likelihood actions will be habitual, whereas any reduction in habitual control increases the likelihood actions will be goal directed. Thus, inactivation of the dorsomedial striatum immediately places actions under habitual control (Yin et al. 2005a), whereas inactivation of the dorsolateral striatum appears immediately to render actions goal directed (Yin et al. 2006), as if these two processes are always active and merely compete for control over performance (Balleine et al. 2009).

In fact, even under normal circumstances the goal-directed process can rapidly inhibit habitual control. This can be detected in our everyday activities. A

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

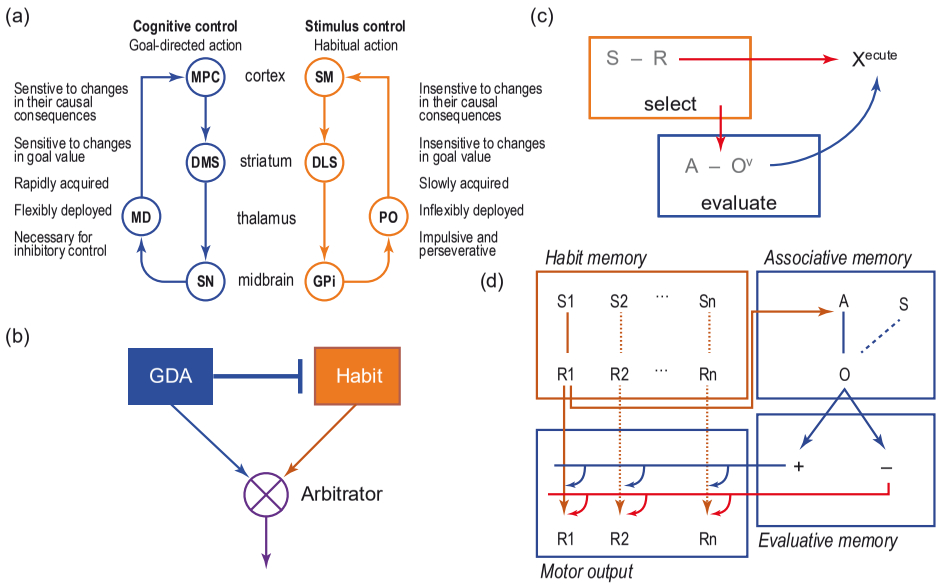

Figure 3.1 Competition and cooperation between actions and habits. (a) Schematic of neural circuits mediating goal-directed and habitual action control, as well as their primary characteristics: A loss of goal-directed control tends to yield dysregulated habitual actions, whereas a loss of habitual control yields dysregulated goal-directed actions. (b) One view of the competitive processes in goal-directed action (GDA) and habitual action control in which both compete for control of performance via some form of arbitration; habits emerge via arbitration but are heavily regulated. (c) Illustration of the cooperation between the stimulus-mediated selection and action-mediated evaluation processes necessary to generate action execution. Generally, stimuli (S) generate urges to respond (R) that initiate the retrieval of specific actions (A), their outcomes (O), and the value (v) of those consequences. If negative, this evaluation would check that urge; if positive, the urge would translate into an executed action. (d) This descriptive model can be further elaborated into an associative cybernetic model which views instrumental performance as the direct outcome of cooperation between S-R and A-O associative processes, whose joint influence sums at the motor system to drive motor output.

useful example is to consider the effect of seeing a police car in the rearview mirror while driving on the freeway: After a period of carefree, apparently cognitively disconnected driving, the sight of the highway patrol causes a very distinct change in our behavior. Do we carry on driving in such a carefree manner? Not likely! Even if we are within the speed limit and are generally obeying the rules of the road, our vigilance would dramatically increase and our driving would become far more deliberate. Thus, habitual control has been suppressed. Likewise, rats behaving habitually will rapidly transition to goal-directed control when the lever response is punished by the actual delivery of an aversive or noxious outcome (Dickinson et al. 1983, 1995). The rapidity of this adjustment is, however, severely curtailed by damage to or inactivation of the dorsomedial striatum (Yin et al. 2005b; Furlong et al. 2017), a finding that is consistent with the argument that the return of control to the goal-directed system is the source of the rapid suppression of habits in a punishment situation (see Figure 3.1b).

Loss of Control and Intrusive Action

Some forms of psychopathology that result in intrusive or compulsive actions find their source in a defective ability to suppress habitual actions; most notably in drug addiction. During the development of addiction, the pursuit of drugs of abuse rapidly becomes habitual, coming under the control of internal and external states and stimuli rather than the consequences of acting (Corbit et al. 2014; Furlong et al. 2016, 2017, 2018). It is important, however, to distinguish habitual drug seeking from other forms of habitual behavior. Under normal conditions, habit learning can be highly adaptive: habits allow us and other animals to relegate the control of routine behavioral responses to a system that uses few cognitive resources, freeing up a limited executive capacity. In contrast, habitual drug seeking is pathological: drug exposure increases the rate of acquisition of habitual actions and the influence of drug-associated contexts (Ostlund et al. 2010) and cues (Glasner et al. 2005) on their performance. Furthermore, a distinguishing feature of habitual drug seeking is the addict's loss of executive control over the habit. As is commonly noted, drug pursuit persists in the face of often severe, negative consequences. The compulsive pursuit of drugs can be viewed, therefore, as the product of (a) a drug-induced increment in habit and (b) a decrement in the addict's ability to exert control over these actions in the face of persistent, negative feedback (Ostlund and Balleine 2008). The need for these two processes to interact may help clarify why relatively few people who take drugs ultimately become addicted; although taking drugs may have some of the hallmarks of a habit, resilience in their goal-directed system to the effects of drug exposure may help to reduce the risk of those habits ultimately becoming dysregulated (Burt et al. 2016).

Consistent with this argument, we have found that exposure either to drugs themselves, such as cocaine or methamphetamine, or to contexts previously

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

paired with drugs can render goal-directed action habitual in tests of outcome devaluation (Ostlund et al. 2010; Corbit et al. 2014; Furlong et al. 2017). Furthermore, actions in these contexts can persist when punished by the delivery of an aversive outcome (Furlong et al. 2018), something that we found was linked to reduced activity in D1-expressing spiny projection neurons (SPNs) in the posterior dorsomedial striatum (Furlong et al. 2017). There are two populations of SPNs in the striatum that account for roughly 95% of the neurons: direct path D1-expressing (dSPNs) and indirect path D2-expressing SPNs (iSPNs). Whereas dSPNs tend to increase functional output, and hence increase goal-directed control, iSPNs inhibit it (Gerfen and Surmeier 2011). As such, a loss of dSPN activity should be expected to have the effects we observed. In an attempt to rescue normal function, therefore, we attempted to redress the relative balance between the activity of dSPNs and iSPNs by inhibiting iSPNs. How we achieved this was complex. The binding of dopamine at D2 receptors on iSPNS inhibits the activity of these neurons and, as such, a D2 antagonist might be expected to be sufficient. However, D2 receptors are also expressed on multiple interneurons as well as iSPNs in the dorsal striatum, reducing the selectivity of this manipulation. Importantly, however, adenosine A2A receptors are only expressed on iSPNs in this region, and the inhibition of these receptors increases the effects of D2 binding only on iSPNs (Tozzi et al. 2007). We hypothesized, therefore, that local infusion of an A2A antagonist would (a) increase D2 receptor activity, (b) reduce iSPN activity, (c) restore the balance between dSPNs and iSPNs, and thus (d) allow the animals to exert behavioral control over their instrumental performance. This is indeed what we found: rats were now able to exert behavioral control over their habits and reduced performance in a punishment test to a similar degree whether they were tested in a drug-paired or unpaired context (Furlong et al. 2017).

Similar deficits in contingency degradation and outcome devaluation have been described in various psychiatric conditions linked to changes in the circuitry mediating goal-directed action control (Griffiths et al. 2014). We found, for example, that the causal judgements of a cohort of youths diagnosed with major depression were relatively insensitive to changes in the causal relationship between action and outcome (Griffiths et al. 2015, 2016a). Furthermore, reductions in causal awareness were correlated with size and shape changes of the globus pallidus (GP) and midline thalamic structures in the basal ganglia output circuit that feeds back to the cortex. Tractography confirmed the relationship between the dysfunctional area of GP and the striatum, on one hand, and the midline thalamus, on the other (Griffiths et al. 2015). This suggests, particularly given that subjects were at an early stage in illness progression, that such changes may reflect the precursors of later prefrontal cortical and corticostriatal deficits in depression, as has also been claimed by others with regard to schizophrenia (e.g., Simpson et al. 2010). Indeed, in a recent study, we found deficits in the flexibility of causal judgment in chronic schizophrenia similar to that observed in major depression (Morris et al. 2018). Furthermore,

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

we found that schizophrenic subjects had a severe deficit in the sensitivity of their instrumental performance to outcome devaluation. We did not, however, find deficits in the ability of schizophrenic subjects to describe the immediate relationship between actions and their consequences, nor did we find that schizophrenic subjects differed from healthy controls in their ratings of outcome value after devaluation. Rather, it appeared that the schizophrenic subjects had difficulty integrating their knowledge of the action–outcome relationship with the changes in outcome value to permit them to make accurate choices. As might be expected, this effect was related to reduced activity in the dorsolateral prefrontal cortex and in the head of the caudate. Furthermore, whereas the deficits in neural activity were related to the severity of negative symptoms, particularly avolition and alogia, devaluation sensitivity correlated with functional measures, such as hours of paid employment over the last six weeks (Morris et al. 2015). Generally, therefore, with respect to its role in controlling intrusive actions, each of these effects suggests that dysfunction in the prefrontal-caudate-globus pallidus-thalamic feedback circuit compromises goal-directed control.

The Dysregulation of Goal-Directed Action

It would be interesting to know whether the same kind of dysregulation occurs when habitual control is inhibited. Does reduced habitual control result in dysregulated goal-directed action? If so, what would that dysregulation look like? There are, in fact, layers of complexity in attempting to understand the interaction of these apparently distinct action controllers. At one level it is clear that goal-directed and habit processes compete for control of performance (as reviewed above). There is also evidence that these processes cooperate in the integration of stimulus-mediated action selection with action evaluation, particularly during the performance of goal-directed actions. Perhaps the strongest evidence comes from studies of instrumental reinstatement: after a period of extinction, we have found that free delivery of the instrumental outcome reinstates performance of its associated action (Ostlund and Balleine 2007b). Importantly, the ability of the outcome to generate this effect does not depend on its value: devaluation of the freely paired outcome does not affect its ability to select its associated action in reinstatement tests. Nevertheless, the subsequent level of performance of that action is affected by the devaluation treatment: devaluing the outcome that serves as a goal for the retrieved action reduces the vigor of its performance but not the ability of the outcome to reinstate performance (Balleine and Ostlund 2007; de Wit et al. 2009).

Thus, in the ordinary course of events, this evidence suggests that the outcome controls actions in two ways: (a) through a form of stimulus–response association, according to which the stimulus properties of the outcome select its associated action, and (b) through the action–outcome association, through which the action retrieves the outcome as a goal. This behavioral evidence suggests that a

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

selection–evaluation–execution sequence lies at the heart of goal-directed instrumental performance and that this control requires the cooperative integration of the goal-directed and habitual control processes (see Figure 3.1c; Balleine and Ostlund 2007; Balleine and O'Doherty 2010). From this perspective, two possibilities may arise. First, if maladaptive habitual control results in attenuated cooperation between the habitual and goal-directed control processes, then this would break the link between the urge to respond and the regulation of the effect of that urge on performance, potentially resulting in an urge to act but without the subsequent capacity to evaluate and, when necessary, to suppress that action. Such intrusive urges to act may be one source of obsessions and, if strong enough, of compulsions (as discussed above). Alternatively, an actual loss of habitual control could result in a loss of action selection and/or initiation leaving the actor cognitively able to retrieve specific action outcomes and their value but unable to implement those actions. Cognitively simulating actions and thinking through their consequences sounds like planning. However, repeatedly engaging in cognitive simulation because of an attenuated capacity to complete the retrieval–evaluation–execution sequence does not resemble adaptive planning but rather *obsessive* or *intrusive thinking*. It might, therefore, be argued that whereas attenuated goal-directed control results in compulsive or intrusive actions, a loss of habitual control results in obsessive or intrusive rumination on the potential consequences of actions that are, then, unable (or less able) to be performed. Such failures could result from dysregulation in the competition between these control processes (resulting in a loss in the ability to inhibit habits) or in the cooperation between them (resulting in an inability to execute actions and so bring planning to an end).

Some years ago, we implemented this general scheme within an associative cybernetic model of instrumental performance (see Figure 3.1d). Briefly, a response tendency or urge is generated in a stimulus-response memory which then brings to mind the retrieval of a specific action and its consequences in an associative memory. Outcome retrieval in the associative memory activates an incentive memory of that outcome which, by marshalling specific motivational and emotional processes, determines its value, which acts to potentiate (or depotentiate) the motor effects of the stimulus-response urge and thus increase (or decrease) the probability that the action will be executed. It is this latter process that constitutes the cybernetic or feedback component of the model without which either response urges could run off unchecked or not run off at all (depending on their strength). Likewise, a significantly attenuated stimulus-response memory would result in the failure of activity in the associative memory to result in an actual response. One question that arises in this model is why there is any activity at all in the associative memory when the habit memory is attenuated. To understand this issue, it is necessary to introduce perhaps the single most important influence on the performance of goal-directed and habitual actions in the context of intrusive thoughts and actions: the effect of predictive stimuli on action selection.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Cognitive Control of Action

Although both prediction and control are necessary for adaptive behavior, discussions of the importance of predictive learning are usually confined to its role in determining conditioned reflexes of one kind or another. There are, however, multiple ways in which predictive learning can influence instrumental actions. Pavlovian predictors of important environmental events could elicit conditioned reflexes congruent or incongruent with our actions; for example, a stimulus paired with an aversive event could generate freezing, which would be incongruent with active avoidance responses. Furthermore, there is evidence that predictive learning can influence actions independently of the conditioned reflexes predictors produce. Indeed, there is good evidence that a stimulus that has become a reliable predictor of a valued outcome can enhance the performance of actions associated with that outcome while leaving actions associated with other rewarding outcomes unaffected. Furthermore, treatments that abolish the predictive validity of such stimuli can abolish these effects on instrumental performance without affecting their ability to evoke a conditioned reflex (Delamater 1995). Over and above conditioned reflexes, therefore, predictive learning clearly provides *information* regarding forthcoming rewards and punishers; if that information is important for adaptive behavior, its effects will then be mediated by changes in instrumental action, via its effects on actions sufficiently flexible to be modified in response to that information and to do so rapidly.

The influence of the information provided by predictive learning on instrumental action is typically studied in the laboratory using the *Pavlovian-instrumental transfer* paradigm (for a recent review, see Cartoni et al. 2016). In these experiments, subjects are exposed to a period of Pavlovian conditioning, during which cues of various kinds are paired with specific, usually rewarding, events (e.g., specific foods or fluids), after which they are trained to perform distinct actions to earn those same food or fluid outcomes (Figure 3.2a). In a typical rodent experiment, rats will be first given a period of predictive learning, during which they learn that stimuli S1 and S2 (e.g., tones or clickers) predict the delivery of distinct outcomes, O1 and O2 (e.g., dry food pellets or liquid sucrose): S1–O1 and S2–O2. Subsequently, they are trained to perform two novel instrumental actions (A1 and A2) for these same outcomes. They might be trained, for instance, to press one lever for the pellets and a second lever for the sucrose: A1 → O1 and A2 → O2. In a final test, the previous Pavlovian and instrumental phases are brought together to examine the influence of the former on the latter: rats must choose between A1 and A2 in the presence of S1 and S2 (i.e., S1: R1 vs. R2, and S2: R1 vs. R2). Importantly, no outcomes are delivered in this test phase, either after the stimuli or the actions. Thus, the test provides an opportunity to observe how predictive learning influences instrumental performance directly. Typically, the stimuli cause the rats to select and perform more vigorously the response previously associated with

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

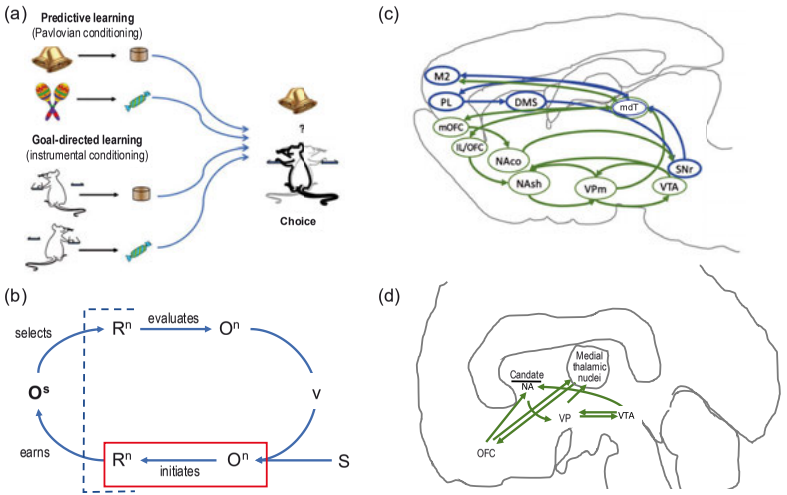

Figure 3.2 The influence of predictive learning on choice. (a) Design of a typical experiment used to study the influence of predictive learning (Pavlovian conditioning) on instrumental choice performance, here in rodents. Rats are first exposed to two stimuli paired with distinct outcomes and then trained to perform two actions for those outcomes before the influence of predictive learning on choice is assessed in an extinction test. (b) Evidence (see text) suggests that the influence of predictive learning is mediated by the ability of a stimulus, S, to retrieve a specific outcome, O ($S - O^n$), and so the specific action associated with that O ($O^n - R^n$). The outcome delivered as a consequence of R^n serves both as the goal of the action and as a stimulus that selects the next action. (c) Schematic of the cortical-striatal-pallidal-thalamo-cortical circuits underlying goal-directed action and the way predictive learning influences these actions. Note that the influence of stimuli on infralimbic and orbitofrontal cortex (IL/OFC) activity is driven by the ability of nucleus accumbens shell (NAsh) and medial ventral pallidum (VPm) to inhibit mediodorsal thalamus (mdT), and so disinhibit this input. The involvement of medial OFC and its connections with nucleus accumbens core (NAco) is crucial to the integration of this circuit with the larger goal-directed circuit and its influence on performance through medial agranular cortex (M2). (d) Dysfunction in the same cortical-striatal-pallidal-thalamo-cortical circuit in humans has been argued to underlie the compulsive actions associated with obsessive-compulsive disorder; ventral tegmental area (VTA), caudate nucleus (CAUD) (after Modell et al. 1989).

the outcome predicted by the stimulus: given the scenarios described above, S1: R1 > R2 and S2: R1 < R2. Very similar effects are observed in mouse and human subjects to those observed in rats (Cartoni et al. 2016).

As intimated above, the influence of predictive stimuli on choice in the transfer paradigm depends on their predictive status; degrading the contingency between a stimulus and its specific outcome does not affect its ability to elicit a conditioned response but completely abolishes its effects on choice between actions. It is clear, therefore, that it is the information that the stimulus provides with respect to a predicted outcome that is critical to the ability of such stimuli to affect performance, rather than its ability to evoke the conditioned response. Nor, as it turns out, is this effect due to the ability of the stimulus to retrieve and activate goal-directed control generally via retrieval of the action–outcome relationship and, subsequently, the value of the outcome. One of the more interesting and telling effects in this literature is the finding that devaluing the outcome predicted by the stimulus does not affect the ability of that stimulus to influence choice (Rescorla 1994; Holland 2004). Although the predictive validity of the stimulus is critical, the value of the outcome predicted by that stimulus is not.

Finally, it is important to note that these effects of predictive learning on choice between actions are mediated by both excitatory and inhibitory action–outcome associations and their effects on performance. When a stimulus predicting a particular outcome elevates the performance of an action associated with that outcome, it does so without affecting the performance of other actions (Laurent and Balleine 2015). We have hypothesized that this is at least partly due to the fact that, in goal-directed learning situations, actions can become associated with the outcome that they deliver (e.g., A1 → O1, A2 → O2) as well as with the absence of outcomes that they do *not* deliver (i.e., A1 → no O2, A2 → no O1). Hence, a stimulus that predicts a particular outcome is likely to elevate the performance of actions associated with the predicted outcome, but not the responses associated with the absence of that outcome; information (e.g., S1 predicts O1) can elevate an action, R1, that delivers O1 but not another action, R2, that predicts no O1 (Laurent and Balleine 2015). In summary, Pavlovian-instrumental transfer provides insight into the way predictive learning affects instrumental performance. As a phenomenon, it reveals the following:

1. Presenting cues that predict specific outcomes elevates the performance of actions associated with those outcomes (without affecting those that are not) by selecting those actions and increasing (or inhibiting) their execution in an ongoing manner.
2. The ability of stimuli to produce these effects depends on how specifically they provide information about those outcomes.
3. The ability of these cues to provoke the performance of actions does not depend on the value of the outcome with which they are associated.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Of the theories advanced thus far to explain the effects of predictive learning on instrumental performance, the one best supported suggests that the association between the stimulus and outcome, acquired during Pavlovian conditioning, allows the stimulus to subsequently retrieve the outcome, thereby retrieving (or inhibiting) the action associated with that outcome (Figure 3.2b; Balleine and Ostlund 2007). Any failure to inhibit competing actions in the presence of a specific predictive stimulus would result in an unwanted, intrusive action. This is strongly reminiscent of the occurrence of intrusive, compulsive actions in various types of addiction as well as in multiple, severe neuropsychiatric disorders, including Tourette syndrome (Leckman et al. 2010), grooming disorders (e.g., skin picking, trichotillomania; Chamberlain et al. 2009), and obsessive-compulsive disorder (OCD) (Robbins et al. 2019). Furthermore, not only does the superficial similarity of the behavioral influence of predictive learning on instrumental performance suggest this, there is a very close similarity between the neural bases of Pavlovian-instrumental transfer and the circuitry previously implicated in the compulsive actions associated with these conditions.

Neural Bases of Pavlovian-Instrumental Transfer

A number of structures have been found to be involved in the way predictive learning affects action selection in transfer situations; for a summary, see Figure 3.2c. Many of these have also been implicated in either Pavlovian conditioning and/or instrumental conditioning, as would be anticipated. Treatments, for example, that lesion or block the activity of the basolateral amygdala (BLA) are particularly effective in blocking transfer effects, but they also abolish the ability of animals to encode the association of Pavlovian cues and instrumental actions with specific outcomes. Similarly, lesions to or inactivation of the mediodorsal thalamus (mdT) or the ventrolateral orbitofrontal cortex (vlOFC) abolish the selectivity of Pavlovian predictions with respect to specific outcomes, thereby abolishing the selectivity of transfer effects with respect to those outcomes (reviewed in Cartoni et al. 2016). However, the projection from the BLA to the nucleus accumbens shell is different in this regard: Disconnection of the BLA from the shell abolishes Pavlovian-instrumental transfer but has no detectable effect on either instrumental- or Pavlovian-conditioning processes. This suggests that this circuit is critical for the way information derived from predictive learning informs choice between instrumental actions (Laurent et al. 2015a).

In recent years, considerable detail has been added to this basic circuit in understanding how exactly the shell is involved in transfer. Briefly, what has become clear is that inputs to the shell from the ventral tegmental dopamine neurons are critically involved (Laurent et al. 2014) as are two other modulatory processes that are localized to the shell itself, involving the interaction of opioid and cholinergic processes (Laurent et al. 2012, 2014, 2015b). The

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

latter process is controlled by the giant aspiny cholinergic interneurons that make up about 3% of the neurons in the shell. These neurons express delta opioid receptors on their membrane which bind enkephalin that inhibits the release of acetylcholine, thereby releasing its inhibitory effects on dopamine D1 receptor-expressing SPNs in the shell (Bertran-Gonzalez et al. 2012). Importantly, the degree to which delta opioid receptors are expressed on cholinergic interneurons is related to the effectiveness of predictive learning. Mice that show strong Pavlovian learning also show strong delta receptor expression, and this is true whether one assesses excitatory (Bertran-Gonzalez et al. 2012) or inhibitory associations (Laurent et al. 2015b).

The shell projects to the broader basal ganglia circuitry through several pathways, but particularly to the medial and lateral segments of the ventral pallidum. In a number of studies, we found that the more critical projection for transfer effects is to the ventral segment of the medial ventral pallidum (VPm) based on the effects of disconnecting the shell from the VPm versus the VP-l prior to the Pavlovian-instrumental transfer test (Leung and Balleine 2013). Shell inputs to the VPm drive the output neurons from this structure, which project to multiple efferent structures including the ventral tegmental area (VTA) and the mdT (Root et al. 2015). Again, using disconnection procedures, we found that the critical output for the transfer effect is to the mdT, whereas the projection to the VTA appears to be important for response vigor irrespective of the direction of that response (Leung and Balleine 2015). Thus, although damage to both VTA and mdT projections completely removes the transfer effect, damage to the VTA projection reduces the overall level of performance, whereas damage to the mdT projection blocks the bias in performance typically induced during the transfer test.

The VPm to mdT projection is GABAergic and its activity inhibits thalamic output neurons (Lavín and Grace 1994). These mdT output neurons inhibit the vlOFC (Alcaraz et al. 2016); thus, turning off this mdT projection disinhibits the vlOFC and alters the activity of this corticostriatal projection and its effects on performance. As such, this loop serves the function of maintaining the selective drive exerted by predictive learning on action selection, although precisely how it influences motor performance is currently unknown. One possibility is that, like other cortical areas that receive input from the thalamus, the OFC provides a return projection that, ultimately, excites motor output through premotor cortex (Hart et al. 2014). As this output is also likely to be engaged by the dorsomedial striatal goal-directed circuit, it is possible that the corticothalamic circuit is generally functioning to allow predictive cues to potentiate the performance of actions selected and evaluated by the goal-directed circuit (Hart et al. 2014). Whatever the merits of these speculations, it is important to recognize that any failure of the VPm to inhibit the mdT will result in a failure to disinhibit the OFC, and that the effect of this failure is not to block transfer effects per se but to render them nonselective; that is, the action typically inhibited in the transfer test via the

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

inhibitory response-outcome association will be disinhibited and released as an intrusive, involuntary, or compulsive action.

Finally, one other structure implicated in the changes in instrumental performance induced by predictive learning is the medial OFC (mOFC). Damage to or inactivation of the mOFC also abolishes transfer effects without generally elevating performance; rather, Pavlovian cues have no influence on instrumental performance (Bradfield et al. 2015). One explanation for this effect is that the mOFC critically mediates instrumental performance in extinction; indeed we have recently argued that the mOFC functions to retrieve action outcomes when they are freely recalled from memory as opposed to when they are directly observable in the environment. We reached this conclusion from the finding that mOFC manipulations altered sensitivity to outcome devaluation, but only in tests conducted in extinction: when the devalued outcome was delivered contingent on performance, the response was adjusted appropriately. Importantly, these manipulations also altered the animals' sensitivity to Pavlovian-instrumental transfer (Bradfield et al. 2015). Thus, unlike the vlOFC, the mOFC is not just involved in influencing predictive stimuli on instrumental performance; it also affects experienced changes in reward on that performance. This suggests that its effects on transfer might be related to its involvement in retrieving the outcomes of goal-directed actions more generally when these outcomes are unobservable.

Neural Bases of Intrusive Thoughts and Actions

Over and above the relationship between the altered activity in various cortical-dorsal-striatal circuits and the disease states described above, theories related to the neural bases of addictive conditions and relapse as well as of OCD and related conditions have implicated changes in the OFC-shell-VPm-mdT-OFC loop mediating Pavlovian-instrumental transfer in those conditions (Fettes et al. 2017). The role of this circuitry in OCD is particularly well documented, especially the role of the vlOFC and mOFC and its reciprocal connections with the mdT in this condition. Thus, for example, it has long been known that the symptoms associated with OCD are ameliorated by surgical ablation of either the OFC or of the midline and intralaminar thalamus, including the mdT. Importantly, they are similarly ameliorated by thermal or gamma radio lesions of the ventral anterior limb of the internal capsule (also called anterior capsulotomy)—the region through which the white matter tracks pass containing the bidirectional fibers which connect the thalamus and the OFC (Pepper et al. 2015). More recently, deep brain stimulation of these fiber pathways has been reported to have a similar effect (reviewed in Greenberg et al. 2010).

Neuroimaging studies of OCD patients have implicated aspects of this circuitry in OCD, particularly the subregions of the OFC. Generally, the vlOFC has been reported to be hypoactive (Remijnse et al. 2006) and the mOFC to be hyperactive (Gillan et al. 2015) during tests that engage these regions; these

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

findings are broadly consistent with other studies (e.g., Kahnt et al. 2012; Zald et al. 2014). Two important findings from the animal literature suggest that these differences in activity may be important for aspects of symptom formation. For example, a study by Burguiere et al. (2013), which uses the SAPAP3 mouse model of OCD, found that these mice show a persistent conditioned grooming response when it was elicited by a cue that had been paired with a water drop to the head (Burguiere et al. 2013). The authors hypothesized that this effect was generated by a hypoactive vlOFC and thus sought to activate this region in an attempt to reduce the excessive grooming response. They achieved this using optogenetic stimulation of the vlOFC and found that the persistent grooming was, indeed, inhibited by optogenetic stimulation of the lateral OFC neurons projecting to the striatum. A similar study by Ahmari et al. (2013), optically stimulated the terminals of mOFC cells expressing channelrhodopsin in the ventromedial striatum, centering on the junction between the nucleus accumbens core and shell. Stimulation of this pathway progressively increased grooming, which persisted after the stimulation (Ahmari et al. 2013). As such it appears that enhancing activation of accumbens projecting neurons in the vlOFC reduces compulsive grooming, whereas stimulating the mOFC projections to the ventral striatum enhances compulsive grooming in line with the hypo- and hyperactive phenotypes of the vlOFC and mOFC in OCD, respectively.

Sometime ago, based on similar kinds of observations in OCD patients, Modell et al. (1989) developed a circuitry model of OCD with striking similarities to the circuitry involved in Pavlovian-instrumental transfer (Figure 3.2d). They argued that "the primary pathogenetic mechanism of OCD lies in dysregulation of a basal ganglia/limbic striatal circuit that modulates neuronal activity in and between posterior portions of the orbitofrontal cortex and the associated medial thalamic nuclei" (Modell et al. 1989:32). Furthermore, they proposed that specifically the compulsive symptoms of OCD are associated with aberrant activity in a positive feedback loop in the reciprocally excitatory frontothalamic neurons, due to the loss of inhibitory input from the ventromedial portions of the striatum. Further, they postulated that: "the net result of ventral striatal activation is increased (inhibitory) pallidothalamic output: essentially the loop transduces excitatory input from the vlOFC into inhibitory output to the thalamus which serves to modulate activity of the frontothalamic circuit by means of an interposed negative feedback loop" (Modell et al. 1989:32).

The same circuit is crucial to normal Pavlovian-instrumental transfer. Damage to this circuit results in aberrant transfer; that is, it results in predictive learning eliciting *both* the performance of actions associated with the outcomes predicted by those stimuli as well as actions that predict the absence of those outcomes, which is the definition of an intrusive involuntary and maladaptive movement. This hypothesis suggests, therefore, that the transfer effect and its attendant circuitry, which provides such an important translational model of the cognitive control of actions, also provides a model of OCD in

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

which abnormality in this circuit results in the inhibition of the vlOFC and, by extension, excitation of the mOFC.

Summary and Conclusions

Corticostriatal circuits play a central role in the induction of intrusive thoughts and actions that are associated with major psychiatric disorders and addictions of various kinds. Here, I point to two types of involvement based on an assessment of (a) the corticostriatal circuits engaged during the acquisition and performance of goal-directed and habitual actions and (b) the modulation of such actions by predictive learning. The first indicates deviations in the balance between the dual corticostriatal circuits that mediate goal-directed and habitual action control and the competitive and cooperative relationship between them that is required to generate adaptive behavior. Changes to the goal-directed circuit reduces its regulation of habits and allows the intrusion of stimulus-driven actions without regard to their consequences. Alternatively, changes to the habitual system could result in a loss of the cooperative processes through which stimulus-driven urges provide the basis for the retrieval and execution of goal-directed action. As a consequence, the operation of a planning process that retrieves action–outcome associations could result in the mental simulation of their performance without the ability to initiate those actions. This simulation has much in common with intrusive thinking.

One initiator of this planning process is predictive learning. Stimuli associated with the outcomes with which actions are associated can directly elicit those actions. Predictive learning can engage both excitatory and inhibitory aspects of action control: excitatory in the sense that actions associated with predicted outcomes are released, and inhibitory in the sense that such stimuli can block this influence from extending to actions associated with unpredicted outcomes. A circuit involving the orbitofrontal cortex, nucleus accumbens shell, ventral pallidum, and mediodorsal thalamus controls this balance between the excitatory and inhibitory effects of predictive learning. Interference with this circuit results in the release of inhibitory control, causing predictive cues to elicit actions associated with unpredicted outcomes, which, to a first approximation, is the definition of intrusive or compulsive action. Generally, therefore, the arguments presented here suggest that it is not a failure of habitual or goal-directed control processes per se, but rather the failure of the cooperative and competitive interactions between these processes that forms the basis of intrusive thoughts and actions.

Acknowledgments

Support for this work was provided by a Senior Principal Research Fellowship from the National Health and Medical Research Council of Australia (GNT#1079561). The

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

author thanks members of the Decision Neuroscience Lab at UNSW and attendees at the ES Forum on Intrusive Thinking for the many discussions that have helped to shape this work.

From “Intrusive Thinking: From Molecules to Free Will,” edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1