Neuroscience and Biobehavioral Reviews 118 (2020) 27–41

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Prefrontal regulation of behavioural control: Evidence from learning theory and translational approaches in rodents

Karly M. Turner[a,*], Shauna L. Parkes[b,c,*]

[a] Decision Neuroscience Laboratory, School of Psychology, University of New South Wales, Sydney, NSW, Australia
[b] CNRS, INCIA, UMR 5287, Bordeaux, France
[c] Université de Bordeaux, France

ARTICLE INFO

Keywords:
Habit
Goal-directed
Instrumental
Compulsive
Impulsive
Punishment
Rodent
Prefrontal cortex
Insular cortex
Orbitofrontal cortex
Decision-making

ABSTRACT

Everyday activities require adaptive decision-making and control over our actions to achieve our goals. Sub-regions within the cortex are widely reported to regulate these choices. Here we review rodent studies from two disparate fields of instrumental action control – goal-directed and habitual responding, and impulsive and compulsive behaviour. Our aim was to compare findings across the spectrum, from precision associative learning to translational studies of action control. The evidence suggests that each cortical sub-region performs different roles depending on task requirements and, within tasks, clear dissociations exist between regions. Rather than synthesizing a single role or function for a given region, we should consider regions to be capable of many different functions. Further investigation of cortico-cortical connections and the pattern of input and output circuitry within each region may be needed to identify unique process-specific pathways. Despite differences in the scope and purpose of these two fields, integrating evidence across tasks provides a broader context for testing hypotheses about the role of cortical regions in adaptive actions and decision-making.

1. Introduction

We are constantly faced with situations that require choices between different actions. These actions may be selected through costly, adaptive decision-making processes or more efficient, although inflexible, automatic responses. The study of volitional behaviour in recent decades has led to the operational definition of these actions as goal-directed and habitual. These definitions hinge on behavioural adaptation under conditions where there are changes to either the value of an action's outcome or the causal relationship between the action and its consequence (for a recent review see Balleine and Dezfouli, 2019). Action control has also been studied in the context of inappropriate responses, including impulsive and compulsive behaviours, where actions often do not lead to the desired outcome. These behaviours provide insight into how volitional behaviour and action control can become disrupted. While studies in these fields have evolved somewhat independently, both fields have found a central role for corticostriatal circuitry in these processes, with cortical inputs involved in adaptive choice, learning and decision-making. Given the close proximity, interconnectivity and similar but dissociable inputs and outputs of the different regions of the prefrontal cortex, it is unlikely that regions are acting alone in instigating choices. Here we review literature from rodent models across a range of tasks in an attempt to provide a clearer picture of how these cortical regions contribute to different cognitive tasks and processes, and how they work together to solve these challenges.

This article reviews two domains of action control and decision-making in rodents that have remained largely separate in the literature: (i) habits and goal-directed actions, and (ii) impulsive and compulsive actions. These domains have evolved from reductionist versus translational approaches in psychology, respectively. Indeed, paradigms developed to assess habits versus goal-directed actions are founded in animal learning and cognition and allow us to precisely assess the psychological processes underlying an animal's behaviour; that is, what is an animal doing and why is it doing it (Balleine, 2019). By contrast, paradigms used to measure impulsive versus compulsive actions have a more translational focus with the goal of understanding executive functions (Fineberg et al., 2010). While there is a strong theoretical background to the study of complex executive functioning; and conversely an increasing interest in applying learning theory to translational studies, the two fields have remained largely independent. Our goal is to present current findings from various behavioural paradigms

* Corresponding author at: Institut de Neurosciences Cognitives et Integratives d'Aquitaine (INCIA), Université de Bordeaux, 146 rue Leo Saignat, 33076, Bordeaux, France.
E-mail addresses: karly.turner@unsw.edu.au (K.M. Turner), shauna.parkes@u-bordeaux.fr (S.L. Parkes).

https://doi.org/10.1016/j.neubiorev.2020.07.010
Received 18 February 2020; Received in revised form 21 April 2020; Accepted 13 July 2020
Available online 21 July 2020

within these learning theory and translational domains to illustrate how cortical regions work as an integrated unit to help us make adaptive choices.

For both domains, we focus only on behaviours (or responses) that are driven by instrumental associations rather than Pavlovian associations. Indeed, while the predictive learning processes engaged by Pavlovian conditioning provide organisms with the capacity to elicit anticipatory responses as a result of learning about the associations between cues or events, instrumental responses, established via instrumental conditioning, allow organisms to acquire new behavioral strategies and to exert control over their environment. It should be acknowledged that many tasks measuring impulsive and compulsive behaviour include components that are likely to contribute to Pavlovian processes, such as the use of cue lights for time out or reward delivery, however task performance is dependent on instrumental conditioning. The studies reviewed here therefore use tasks that depend on a causal relationship between the performance of a response and the availability of its associated outcome.

In Section 2, we review the cortical contribution to habits and goal-directed actions. The included studies use contingency degradation and outcome devaluation to assess whether an action is elicited by a stimulus-response association (i.e., habitual) or if the action is driven by a response-outcome association and the current value of the outcome (i.e., goal-directed). Studies of impulsive and compulsive behaviours are reviewed in Section 3. These studies are typically less well defined in terms of isolating the psychological processes being measured, but provide translational insight into the neural substrates of impaired action control. We examine the contribution of various cortical regions to impulsive action (including waiting to perform or stopping an action) and impulsive choice (i.e., selecting an action that will maximise reward). To assess compulsivity and perseveration, we review studies using reversal and switching tasks as well as punished responding. Schematics of the behavioural tasks covered in Sections 2 and 3 are shown in Fig. 1. Overall, the aim of this review is to integrate recent findings with traditional views to shed new light on the cortical control of decision-making and action control.

We focus on the involvement of several rodent cortical regions in reward-based decision making. These regions include the medial prefrontal cortex (mPFC), comprising anterior cingulate cortex (ACC), prelimbic cortex (PL), infralimbic cortex (IL), and medial orbitofrontal cortex (mOFC) as well as lateral prefrontal areas, including ventral OFC, lateral OFC and the most anterior agranular region of the insular cortex (IC). We also include the more posterior, gustatory region of insular cortex, between bregma +2.5 mm and +0.2 mm in the rat (Cechetto and Saper, 1987; McDonald, 1998; Allen et al., 1991; Shi and Cassell, 1998). Studies using manipulations that extend across more than one region have also been included (for example, studies targeting both PL and IL). Research focused on specific neurotransmitter systems and pharmacology studies were not included due to the additional complexity that would be generated, however integration of these findings with the outcomes of this review will be important moving forward.

It should also be noted that regions of rodent prefrontal cortex were recently redefined using the Brodmann nomenclature scheme in an attempt to integrate rodent and primate research (Vogt and Paxinos, 2014; Paxinos and Watson, 2014). Most notable for the current review is that the terms “prelimbic” and “infralimbic” are no longer used. These terms are extensively used in the studies included in this review and, therefore, we retain their use here. The homology between the rodent and primate prefrontal cortex is ongoing topic of debate and is not discussed in this review. However, it is acknowledged that in the latest edition of the Paxinos and Watson rat brain atlas, based on the Brodmann nomenclature, prelimbic has been redefined as area 32 (with dorsal and ventral divisions) and infralimbic redefined as area 25. In addition, cingulate areas 1 and 2 are redefined as area 24b and 24a, respectively (Paxinos and Watson, 2014; Franklin and Paxinos, 2019).

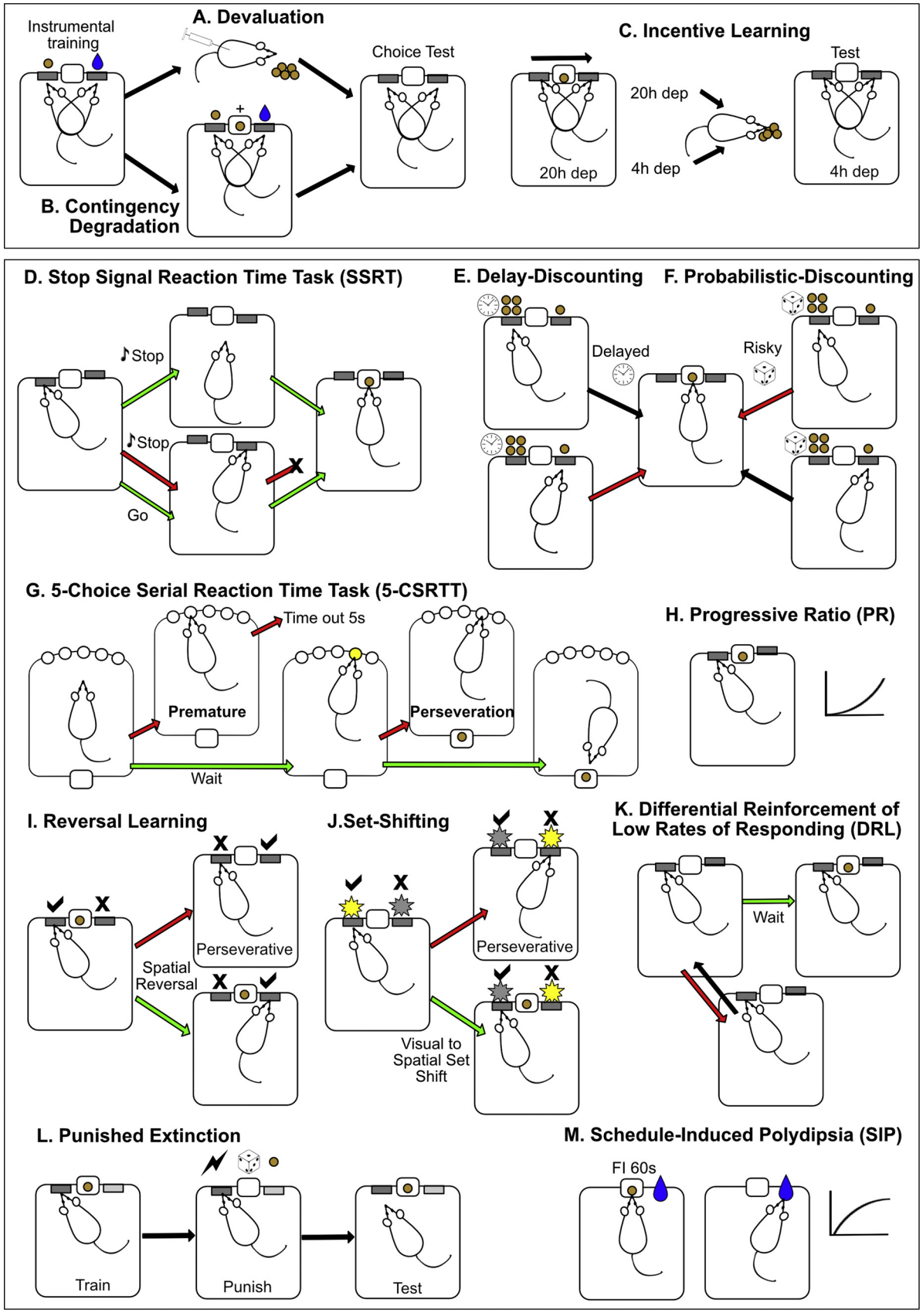
2. Learning theory domain: habits and goal-directed actions

Representations of value guide everyday choices. Faced with a decision between two courses of action we, and other animals, typically choose the action that delivers the more valuable consequence or outcome. Critically, outcome values fluctuate with our current needs and desires. Hence, the ability to update and encode value is a hallmark of flexible decision making. Such behaviour is defined as goal-directed (Balleine and Dickinson, 1998). However, action selection is not always guided by the value of their consequences. So-called “habits” are instead elicited by a stimulus-response association (Dickinson, 1985). Habits are an important component of decision making, allowing us to automatically generate appropriate behaviours with little cognitive effort.

Distinguishing between these two types of action control requires an understanding of their underlying associative processes. In order for a response to be goal-directed, it must be driven by a causal relationship between performance of the response and access to the goal and a representation that the outcome of that response is a desired goal (Dickinson and Balleine, 1994; Heyes and Dickinson, 1990). That is, a response is only goal-directed if its performance depends on *both* the contingency between the response and the outcome (i.e., the subject’s *belief* that performing the action will deliver the outcome) and the current motivational or incentive value of that outcome (i.e., the subject’s *desire* for the outcome) (Heyes and Dickinson, 1990; de Wit and Dickinson, 2009). Failure to meet these criteria indicates that the instrumental response is not goal-directed and may be under habitual control. Habitual responses are instead elicited by antecedent stimuli and are not performed to gain access to specific outcomes; that is, habits are performed independently of outcome value. Given these criteria, it is clear that Pavlovian responses (i.e., behaviours elicited by stimulus-outcome associations) are not, by definition, goal-directed. While Pavlovian responding may be sensitive to changes in outcome value, there is no contingent relationship between the performance of the response and the delivery of the outcome (Hershberger, 1986; Bussey et al., 1997). The *belief* criterion is therefore not fulfilled. However, Pavlovian influences on goal-directed and habitual behavior are well documented and, while beyond the scope of the current review, Pavlovian-instrumental interactions have been addressed in several other review papers (e.g., Holmes et al., 2010; Cardinal et al., 2002).

Experimental procedures exist to assess whether responding is under goal-directed or habitual control. Studies employing these procedures have demonstrated that instrumental responding is supported by several regions of the mammalian cortex. Current evidence indicates that, in the rodent, these regions principally involve medial prefrontal, orbitofrontal, and insular cortices. In this section, we focus on two tasks: contingency degradation and outcome devaluation. These tasks explicitly test whether a response is sensitive to (a) the contingent relationship between the response and the outcome (i.e. belief criterion) and/or (b) the current incentive value of the outcome (i.e. desire criterion). Decades of behavioural research have been dedicated to understanding the associative learning processes involved in these tasks making them particularly amenable to stringent and coherent neurobiological analyses.

Contingency degradation assesses the subject’s understanding of the causal contingency between a response and its outcome. Indeed, performance of an instrumental response per se is not sufficient to determine whether that performance is driven by knowledge of the specific action-outcome contingency. Explicit tests of contingency perception are required (Dickinson and Balleine, 1994; Dickinson, 1994). In contingency degradation, rodents are typically trained to perform two actions for two different outcomes, then, rodents receive non-contingent deliveries of one of the outcomes. This weakens or degrades the causal relationship between that outcome and its associated action. If responding is under goal-directed control, rodents will stop responding on the degraded action. If responding is habitual, rodents



(caption on next page)

Fig. 1. Behavioural tasks used to assess habits and goal-directed actions versus compulsive and impulsive actions.
A) Devaluation. Rodents are first trained to perform two different actions (e.g., a left and right lever press) to earn two different food outcomes (e.g., a grain pellet or sucrose solution). One of the outcomes is then devalued either by sensory-specific satiety (1 h ad libitum access to the outcome) or by pairing the consumption of the outcome with a malaise-inducing injection (e.g., injection of lithium chloride). Finally, the two levers are presented and rodents can freely choose to respond on the levers. This test is unrewarded. Note that devaluation can also be run with a single action-outcome association with a sequential test of devalued versus non-devalued conditions. **B) Contingency degradation.** Rodents are trained to perform two actions for two distinct food outcomes. Then, one of the outcomes is delivered non-contingently, which weakens the causal relationship between that outcome and its associated action, while the other action-outcome association remains intact. Rodents are then given a choice test under unrewarded conditions. **C) Incentive learning.** In this example, rodents learn to perform a seeking-taking chain of actions for a food outcome while hungry [e.g., 20 h deprived of maintenance chow (20 h dep)]. Then, the rodent is either maintained in the hungry state (20 h dep) or shifted to a sated state [e.g., 4 h deprived of maintenance chow (4 h dep)] and is allowed to consume the outcome. All rodents are then tested in the sated state (4 h dep) with the seeking-taking chain of actions under unrewarded conditions. **D) StopSignal ReactionTime Task (SSRT).** Rodents are trained to respond in a rapid two-lever sequence within a limited hold period to receive a food outcome. If a tone is played after the first response, the rat must withhold the second press to receive an outcome. The stop-signal reaction-time is calculated as the time the stop cue must be played prior to the second response in order to stop on 50 % of trials. Longer stop times indicate poorer response control. **E) DelayDiscounting.** Rodents learn that one lever provides a large reward (left) and the other a small reward (right). Then, an increasing delay is incorporated between the response and reward delivery on the large reward lever and the shift towards choosing the small lever is recorded. Faster switching on an ascending delay curve indicates impulsive choice through higher sensitivity to delayed gratification. **F) ProbabilisticDiscounting.** Similar to Delay-Discounting, two levers provide small and large rewards. However, the large reward now becomes risky with an increasing probability of no reward. Staying on the risky, large reward lever too long indicates impulsive/risky choice. **G) 5-Choice Serial Reaction Time Task (5-CSRTT).** Rodents attend to a spatial array of five apertures and respond when one is illuminated to receive a reward. A premature response occurs when a nose poke is made during the inter-trial interval prior to any hole illuminating and results in a time out; indicating impulsive action. In contrast, if the rodent continues to nose poke after a correct response, these are counted as perseverative responses and suggest a compulsive phenotype. **H) Progressive Ratio (PR).** Rodents are trained to lever press to receive a food reward. Then, the number of presses to receive a reward escalates exponentially until the subject stops responding, deemed the break point. This task is commonly used to measure motivation but perseveration also resembles compulsive responding. **I) Reversal Learning.** Rodents are trained to make two responses and the contingencies are then reversed. Here, only one spatial response is rewarded, however in contingency reversal both are rewarded and the action-outcome relationships are reversed. Perseverating on the previously rewarded response indicates a more compulsive response. **J) Set-shifting.** In the version here, rodents learn to press the lever associated with the cue light. Then the rule changes through a set-shift and the rat should now ignore the cue light and press only at one spatial location. **K) Differential Reinforcement of Low Rates of Responding (DRL).** Rodents must wait an inter-response period (e.g. 20 s) between presses to receive a reward and earlier presses reset the timer, providing a measure of impulsive action. **L) Punished Extinction.** After training to lever press for a food reward, the rodent is exposed to a punishment session where presses result in either a reward or foot shock. They are then tested in the absence of shock to see if responding persists despite negative consequences as a measure of compulsive behaviour. **M) Schedule-Induced Polydipsia (SIP).** The infrequent delivery of food rewards (e.g., 60 s) leads to drinking behaviour that escalates beyond homeostatic requirements with further training. This excessive, repetitive behaviour provides a measure of compulsivity. *Note that these are simplified versions of the tasks to highlight key stages. Full methods are available in the relevant empirical articles cited in text.

will continue to perform the degraded action. Contingency degradation therefore measures the extent to which responding is driven by knowledge of the action-outcome association.

The instrumental outcome devaluation paradigm can also be used to evaluate the rodent's knowledge of the specific action-outcome contingencies but, in addition, it also measures the degree to which response performance is dependent on the incentive value of its outcome (e.g., Adams and Dickinson, 1981; Colwill and Rescorla, 1985). In this task, we can assess the subject's ability to update and encode changes in outcome value (incentive learning) and their ability to retrieve the current outcome value and use it to guide action selection (incentive memory). In a typical experiment, the rodent learns to perform two actions for two different yet equally desirable outcomes. Rodents readily learn to perform the actions but, as noted above, it is not yet apparent if responding is under habitual or goal-directed control. To distinguish between these two possibilities, one of the instrumental outcomes is devalued, often via sensory-specific satiety or conditioned taste aversion. If the response is habitual, devaluation of the outcome will be without effect and rodents will continue to select the lever associated with the devalued outcome. Although, it should be noted that a response that is resistant to selective devaluation is not necessarily habitual, particularly when choice is involved. By contrast, if the response is goal-directed, devaluation will decrease instrumental responding for the devalued outcome more than responding for the non-devalued outcome. Note that outcome devaluation experiments can also be performed with a single action-outcome association.

Incentive learning and memory can also be measured using a task that relies on shifts in primary motivational state (Balleine et al., 1995). In this so-called incentive learning paradigm, the rodent learns to perform a seeking-taking chain of actions for a food outcome while hungry. Then, the rodent is shifted to a sated state and is given the opportunity to consume the outcome in this new motivational state (incentive learning) before being given an unrewarded test in that state. If the rodent has learned the new incentive value of the outcome (i.e., that the food is less desirable when sated) and is able to recall this new value (incentive memory), it will perform the action less than a rodent that has not had the opportunity to learn the new value of the food outcome in the sated state. It is also possible to render the food outcome more valuable by training the rodent in a sated state and then giving access to the food in a hungry state. In this case, the value of the outcome is increased and the rodent will perform the associated action more when tested hungry.

2.1. Contingency degradation

Degradation measures whether a response is sensitive to the contingency between performance of the action and delivery of the outcome. If a response is goal-directed, its performance should adapt to changes in instrumental contingencies. If the response is habitual, changes in the response-outcome contingency should be without effect. Successful adaptation to changes in instrumental contingencies requires the PL (Balleine and Dickinson, 1998; Coutureau et al., 2012; Corbit and Balleine, 2003) but not the IL (Naneix et al., 2009), gustatory insular (Balleine and Dickinson, 2000) or mOFC (Bradfield et al., 2015). Pre-training lesions of the PL render rats insensitive to contingency degradation (Balleine and Dickinson, 1998; Corbit and Balleine, 2003) but only when the rat must evaluate the balance between contingent and non-contingent reinforcement. Indeed, rats with lesions of mPFC (including both PL and IL) remain able to learn a shift to a negative contingency (i.e., an omission schedule) but are unable to correctly detect changes when the contingency is shifted to a null contingency (Coutureau et al., 2012) (i.e., no clear relationship between the action and the outcome), as is the case in classic contingency degradation tasks.

Dopaminergic signaling in the PL is also required for adaptation to shifts in contingency (Naneix et al., 2009; Lex and Hauber, 2010; Naneix et al., 2013). Rats with lesions of dopaminergic terminals in the PL show normal acquisition of an instrumental response but fail to adapt their responding to contingency changes. The response, however, does remain sensitive to outcome devaluation [Naneix et al., 2009 but

see Lex and Hauber, 2010]. A similar impairment was observed when a dopamine antagonist was infused into the PL only during the degradation session (Naneix et al., 2009). Thalamic inputs to PL are also required for adapting to changes in instrumental contingencies. Chemogenetic inhibition of PL-projecting neurons in the mediodorsal thalamus (MD) impaired contingency degradation but left outcome devaluation intact (Alcaraz et al., 2018). Notably, this impairment was selective to the MD-to-PL pathway as inhibition of MD-projecting neurons in PL (i.e., the PL-to-MD pathway) left contingency degradation intact (Alcaraz et al., 2018). In contrast, habitual responding is abolished, and sensitivity to changes in the action-outcome relationship is restored, following D1 receptor antagonism or D2 receptor agonism in the IL (Barker et al., 2013). The broader circuitry with which the IL connects to mediate this effect is unknown but may involve direct cortico-cortical interactions with PL or an indirect pathway between IL and dorsolateral striatum (DLS), given that direct connections from IL to dorsolateral striatum are sparse (Gabbott et al., 2005).

While mOFC does not appear necessary for contingency degradation (Bradfield et al., 2015), studies using lesions and temporary chemogenetic inactivation suggest that the ventral and lateral OFC regions (vlOFC) could be required for adapting to changes in instrumental contingencies. In a series of papers, Gourley and colleagues have demonstrated that, in mice, inhibition of vlOFC or selective knockdown of vlOFC brain derived neurotrophic factor leaves acquisition of a nose-poke response intact but renders that response insensitive to contingency degradation (Zimmermann et al., 2017a, b; Gourley et al., 2013; Whyte et al., 2019) although this impairment can be overcome with additional degradation training (Zimmermann et al., 2017b). vlOFC has been proposed to exert top-down control over the DLS to ensure the successful adaptation to contingency changes (Gourley et al., 2013) and inputs to vlOFC from ventral hippocampus may also support this adaptation (Barfield and Gourley, 2019). However, it has been noted that the impairment observed in some of these studies is often driven by a decrease in responding on the non-degraded response and, given the design of these experiments and the demonstrated role for vlOFC in Pavlovian learning (e.g., Ostlund and Balleine, 2007), the contribution of stimulus-outcome associations to responding should be considered (see Robbins, 2017 for the full commentary on these studies).

There is evidence that ACC is required for instrumental responding. Minimal instrumental training has been associated with increased immediate early gene expression in rodent ACC and mPFC (PL, IL and mOFC) whereas expression in these regions was reduced (compared to non-contingent controls) following over-training (Hernandez et al., 2006). However, it was not clear in this study whether responding was goal-directed or habitual or whether rats had learned the specific action-outcome contingency at all. Acquisition of an instrumental response has been shown to depend on NMDA receptor activity in ACC (McKee et al., 2010), but not protein synthesis (Jonkman and Everitt, 2009), and lesions of the perigenual (area 24) cortex in marmosets impairs instrumental contingency degradation (Jackson et al., 2016). Others have argued that rodent ACC may regulate willingness to expend effort to obtain a reward (Schweimer and Hauber, 2005; Aly-Mahmoud et al., 2017), although such a role has not been observed for all instrumental responses, (Schweimer and Hauber, 2005) or response conflict resolution in the control of goal-directed behaviour (Jackson et al., 2016; de Wit et al., 2006).

2.2. Outcome devaluation

2.2.1. Prelimbic versus infralimbic cortices

Outcome devaluation studies have shown that the involvement of prelimbic (PL) versus infralimbic (IL) cortices in goal-directed versus habitual responding is doubly dissociable (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Coutureau and Killcross, 2003; Killcross and Coutureau, 2003). In sham-lesion rats, Killcross and Coutureau (2003) demonstrated the shift from goal-directed to habitual responding that emerges across extended training; after low amounts of training, sham rats showed sensitivity to outcome devaluation but sensitivity was lost after high amounts of training (Adams and Dickinson, 1981; Adams, 1982). By contrast, rats with PL lesions failed to show sensitivity after both limited or extended training, whereas IL lesioned rats showed persistent sensitivity to changes in outcome value even after extended training (Killcross and Coutureau, 2003). PL lesions therefore disrupt goal-directed actions but leave habits intact whereas the reverse is true for IL lesions. Temporary inactivation of the PL during acquisition, but not test, also impairs outcome devaluation (Tran-Tu-Yen et al., 2009; Ostlund and Balleine, 2005) and, following extended training, IL inactivation or infusion of dopamine into the IL at test abolishes habitual responding and restores sensitivity to outcome value (Coutureau and Killcross, 2003; Hitchcott et al., 2007). Similar results of IL involvement in habits have also been reported after over-training in a T-maze and indicate that the IL may be required for the acquisition of habits and not just their expression (Smith and Graybiel, 2013).

However, a recent study from Shipman and colleagues (Shipman et al., 2018) challenges this strict functional dichotomy. The authors trained two instrumental responses, one with minimal training and the other with extensive training. In vehicle control rats, both these responses showed sensitivity to outcome devaluation i.e., both responses were goal-directed. PL inactivation at test reduced the outcome devaluation effect for the minimally trained action but had no effect on the extensively trained action. The opposite pattern was observed for the IL; inactivation of IL at test reduced the outcome devaluation effect for the extensively trained, but not minimally trained, response. Previous studies reporting no effect of PL inactivation at test may have therefore been studying a more extensively trained goal-directed response (Shipman et al., 2018). The authors argue that goal-directed responding may be initially controlled by the PL but this control gradually shifts to the IL before the response becomes habitual (Shipman et al., 2018).

Overall, current evidence supports a role for PL in the acquisition of action-outcome associations, although outcome-action associations, which can also support goal-directed responding, are likely acquired elsewhere (Corbit and Balleine, 2003), whereas IL is required for the expression of extensively trained actions (Shipman et al., 2018). IL is also likely implicated in the process whereby the habitual system overrides the goal-directed system (by inhibiting the PL, for example) but perhaps not in the associative processes underpinning habit formation, which may occur in the DLS (Killcross and Coutureau, 2003; Yin et al., 2004). Direct projections from the PL or indirect projections from the IL to dorsal striatum may support the long-term storage of goal-directed and habitual responses. While the downstream targets of IL have received little attention, a series of papers from Hart and colleagues has revealed a role for the PL to posterior dorsomedial striatum (pDMS) pathway in the acquisition and consolidation of goal-directed learning. Increased MAPK/ERK phosphorylation (pERK) is observed in PL neurons projecting to pDMS shortly after an instrumental training session (Hart and Balleine, 2016) and inactivation of this pathway abolishes outcome devaluation (Hart et al., 2018a). This latter effect was shown to be driven by the bilateral and contralateral projecting intratelencephalic neurons in the PL (Hart et al., 2018b).

In addition to corticostriatal pathways, bidirectional information flow between the prefrontal cortex and the thalamus is required to show sensitivity to changes in outcome value. Chemogenetic inactivation of the PL to mediodorsal thalamus (MD) or MD to PL pathway both impaired outcome devaluation. However, the inactivation occurred throughout the behavioural task so it is unclear exactly when during behaviour these pathways are required. By contrast, it appears that projections from the PL to the basolateral amygdala (BLA) are not required for goal-directed behaviour (Coutureau et al., 2009); however, this remains to be confirmed with inactivation procedures that eliminate both ipsilateral and contralateral projecting PL neurons.

2.2.2. Orbitofrontal and insular cortices

The expression or retrieval of goal-directed actions requires the medial orbitofrontal cortex (mOFC) and the gustatory portion of the insular cortex (IC). The involvement of these regions in instrumental outcome devaluation was first established using pre-training excitotoxic lesions (Balleine and Dickinson, 2000; Bradfield et al., 2015). Their specific involvement was then revealed using temporary inactivation procedures which showed that outcome devaluation is impaired when perturbation of mOFC (Bradfield et al., 2015; Gourley et al., 2016) or IC (Parkes and Balleine, 2013; Parkes et al., 2015, 2018) activity was restricted to the choice test and the devaluation effect is facilitated when mOFC is chemogenetically activated during the test (Gourley et al., 2016). Notably, inhibition of IC during acquisition or satiety-induced devaluation does not impair goal-directed behaviour (Parkes et al., 2018, 2016). Both mOFC and IC are therefore necessary for retrieving a mental representation of the current outcome value (i.e., the value of an absent outcome) to guide action selection but not for the acquisition of instrumental actions or the encoding of outcome value.

Moreover, mOFC and gustatory IC may exert this function via similar neural circuits. Using the incentive learning paradigm, Malvarez et al. (Malvaez et al., 2019) demonstrated that inhibition of BLA inputs to the mOFC during retrieval of outcome value rendered rats unable to adjust their instrumental behaviour according to the updated outcome value. Similarly, the BLA to gustatory IC pathway is required for retrieving current outcome values in an outcome devaluation paradigm (Parkes and Balleine, 2013). However, it should be noted that, in the latter case, a pharmacological disconnection procedure was used so it is not clear if it is a direct or indirect pathway between BLA and gustatory IC that underlies this effect. Efferent projections from gustatory IC to NAc core [but not to BLA (Parkes and Balleine, 2013)] are also required to mediate the effect of outcome value on action selection (Parkes et al., 2015), suggesting a three-node serial circuit (BLA-IC-NAc core) for the encoding and retrieval of outcome value and subsequent instrumental performance based on that value. Additional work is required to delineate the roles of mOFC versus IC in goal-directed behaviour. It is unknown, for instance, whether communication between these cortical regions is required for successful retrieval of the outcome representation.

Functional heterogeneity may also exist within mOFC and IC. Recent work indicates that it is the anterior mOFC and not the posterior mOFC that is required to retrieve the value of an absent outcome (Bradfield et al., 2018; Munster and Hauber, 2017). An anteroposterior analysis of the IC in goal-directed behaviour has not yet been conducted but many studies investigating the lateral and ventral regions of OFC (vlOFC) also target the most anterior, primarily agranular, IC region. These studies indicate that the role of this more anterior IC region may be dissociable from that of the gustatory IC. However, a direct comparison of IC sub-regions is required to determine whether there is indeed a functional heterogeneity within rodent IC in goal-directed behaviour and if the role of anterior IC can be dissociated from that of vlOFC. The vlOFC has been shown to be required for encoding changes (either increases or decreases) in outcome value in an incentive learning paradigm (Baltz et al., 2018). Inhibition of vlOFC to BLA inputs during the encoding, but not retrieval, of incentive value also abolishes goal-directed behaviour (Malvaez et al., 2019). Based on these data, it appears that the more lateral regions of OFC (and, perhaps, anterior insular) are required for updating and encoding changes in outcome value whereas the anterior mOFC and gustatory insular are required for retrieving the value of an outcome that is unobservable (Bradfield et al., 2015; Parkes and Balleine, 2013; Malvaez et al., 2019; Bradfield et al., 2018).

However, studies using the instrumental outcome devaluation task have revealed a different role for vlOFC in goal-directed behaviour. Pre- or post-training lesions of vlOFC leave instrumental outcome devaluation intact (Ostlund and Balleine, 2007; Panayi and Killcross, 2018) and there is evidence to suggest that vlOFC is only required for instrumental outcome devaluation when the task requires the partitioning of two distinct action-outcome contingencies (Bradfield and Hart, 2020; Wilson et al., 2014). Parkes et al. (2017) trained rats to perform two different actions for two distinct food outcomes (e.g., A1-O1; A2-O2). One of these outcomes was then devalued via specific satiety and rats were then given a choice between the two actions. Consistent with the lesion studies, chemogenetic inhibition of vlOFC during either acquisition or the choice test did not impair the rats' ability to select the action associated with the non-devalued outcome. However, outcome devaluation was disrupted when the outcome identities were reversed. That is, following reversal of the action-outcome contingencies (A1-O2; A2-O1) rats were unable to bias their choice towards the action associated with the non-devalued outcome. This effect was not specific to reversal as rats also showed impaired outcome devaluation when they were first trained to perform actions for a common outcome (i.e., A1-Oc; A2-Oc) and then to perform these same actions for novel outcomes (A1-O1; A2-O2). Similar results have been reported when rats are trained with two different schedules of reinforcement on a single lever, likely leading to two different A-O associations for a single action (Gremel and Costa, 2013; Gremel et al., 2016).

Overall, data from the outcome devaluation paradigm suggests that rats with vlOFC inhibition show a specific deficit in action selection when the outcomes associated with that action are uncertain or changing (Bradfield and Hart, 2020; Wilson et al., 2014). The broader circuit mediating this segregation may involve interactions between the vlOFC and the dorsal striatum (Gremel et al., 2016) as well as interactions between the vlOFC and the thalamus (Fresno et al., 2019). Indeed, pre-training lesion disconnection of the vlOFC and the submedius thalamic nucleus impairs instrumental outcome devaluation but only following a reversal of outcome identities (Fresno et al., 2019). Interestingly, disconnecting the vlOFC from its other major thalamic partner, the mediodorsal thalamus, had no effect on outcome devaluation even following reversal based on outcome identity.

2.3. Summary

In this section, we have reviewed evidence from studies using tasks specifically designed to measure whether the performance of an action is driven by a response-outcome association or a stimulus-response association. These tasks have revealed essential information regarding the contribution of cortical subregions to habits and goal-directed actions (see Table 1). The current evidence indicates that the involvement of the ACC in these behaviours is limited with perhaps some contribution to instrumental responding per se but there is no evidence that this region is required for the acquisition or expression of goal-directed or habitual responding.

By contrast, the PL is clearly required for the acquisition of response-outcome contingencies and, thus, is necessary for goal-directed learning. It is not required for the acquisition or expression of habits but may be needed for the expression of a minimally, but not extensively, trained action. Incentive learning and memory do not require the PL indicating that this region contributes to the *'belief'* criterion but not the *'desire'* criterion of goal-directed behaviour. In direct contrast to PL, IL is necessary for the expression of stimulus-response (habit) associations but not response-outcome (goal-directed) associations. Although, under some circumstances, the expression of an extensively trained goal-directed response may come under IL control before that response becomes habitual. The IL may also be required for the acquisition of habits however, S-R encoding may actually reside in the DLS. Like the PL, IL is not required for incentive learning or memory.

Both mOFC and gustatory IC are necessary for goal-directed behaviour. These regions are specifically required for retrieving a representation of current outcome value (incentive memory) and not for encoding the motivational value of the outcome (incentive learning). Neither of these regions is involved in the acquisition of goal-directed behaviour nor in the acquisition or expression of habits.

Table 1
Summary of behavioural findings within the learning theory domain following loss of function in various rodent cortical regions.

	ACC	mPFC	PL	IL	OFC	mOFC	lOFC	Ant. IC	Gust. IC
Action-outcome acquisition									
Contingency degradation			↓	–		–	↓		–
Outcome devaluation			↓	–			–		–
Incentive learning									
Outcome devaluation			–	–			–		–
Incentive learning paradigm						–	↓		
Incentive memory									
Outcome devaluation			–	–		↓	–		↓
Incentive learning paradigm						↓	–		
Action-outcome updating									
Contingency degradation									
Outcome devaluation							↓		–
Habit acquisition									
Contingency degradation			–	↓					
Outcome devaluation			–	↓					
Habit expression									
Contingency degradation			–	↓					
Outcome devaluation			–	↓					

ACC: anterior cingulate cortex; mPFC: medial prefrontal cortex (encompassing prelimbic and infralimbic cortices); PL: prelimbic cortex; IL: infralimbic cortex, OFC: orbitofrontal cortex; mOFC: medial orbitofrontal cortex; lOFC: lateral regions of orbitofrontal cortex; Ant. IC: anterior insular cortex; Gust. IC: gustatory insular cortex. Arrows indicate the direction of the behavioural change following loss of function and - indicates no change following loss of function.

Finally, while there is now convincing evidence that the ventral and lateral regions of OFC are required for goal-directed responding the nature of this involvement remains somewhat unclear. Studies using shifts in primary motivational state suggest a role for vlOFC in outcome value encoding (incentive learning) whereas studies using the instrumental outcome devaluation task show that incentive learning is intact despite inhibition of vlOFC. It is somewhat perplexing that these two paradigms, which are both designed to test the degree to which response performance is dependent on the incentive value of its outcome, would produce conflicting results. Indeed, these two paradigms lead to a similar conclusion regarding the role of the mOFC in goal-directed behaviour. The reason for the inconsistency in vlOFC involvement warrants further attention and may require further dissection of the role for ventral *versus* lateral OFC involvement in goal-directed action. Outcome devaluation studies do, however, suggest an involvement for the ventral and lateral regions of OFC in updating response-outcome associations, particularly changes related to outcome identity, and segregating multiple response-outcome associations.

3. Translational domain: impulsive and compulsive actions

In contrast to the precise dissection of instrumental learning reviewed in Section 2, the study of impulsive and compulsive behaviours has developed from a translational perspective. Although impulsive and compulsive symptoms are common to many neuropsychiatric disorders, they are characteristic of obsessive-compulsive disorder (OCD), attention deficit hyperactivity disorder (ADHD) and substance abuse. Despite different diagnostic partitioning of these conditions, there is consensus that disorders of behavioural inhibition share alterations in corticostriatal circuitry. Rodent models have been widely used to study the neuroscience behind impulsive and compulsive symptoms, particularly in behavioural inhibition and action control. While rodent models do not encapsulate the human experience of impulsive and compulsive symptoms, behavioural studies based on tasks and observations in humans have been developed with minor adaptations. For this review, impulsivity is defined as rapid, inappropriate or risky actions with little regard for subsequent consequences, with a focus on impulsive actions and impulsive choices (Bari and Robbins, 2013; Winstanley et al., 2006a). Compulsive actions include persistent, repetitive or excessive actions that occur despite undesirable consequences (Robbins et al., 2012). The relationship between impulsive and compulsive behaviours is not entirely clear. More traditional views consider impulsive and compulsive actions as opposite ends of a spectrum (Fineberg et al., 2010) however, they have also been viewed as components within the same dimension with impulsivity leading to compulsivity (e.g., in addiction) (Belin et al., 2008) and, on the other hand, they can be grouped together as similar yet different forms of behavioural inhibition deficits (Fineberg et al., 2010; Dalley et al., 2011; Eagle and Baunez, 2010; Voon and Dalley, 2016). Deficits in response control occur in both impulsive and compulsive actions, but they appear to fail at different times during planning or execution with poor inter-correlation across measures validating the division of these constructs (Robbins et al., 2012; Dalley et al., 2011). While the nature of this interaction is still debated, cortical circuits are central to both impulsive and compulsive actions (Fineberg et al., 2010). In this section, we review studies examining the role of different cortical regions in tasks specifically assessing impulsive and compulsive actions in rodents.

With the goal of examining the neural systems underpinning human symptoms, many of the tasks designed for rodents were developed to mirror tasks used in humans. Deficits such as impulsivity and compulsivity are complex and multifaceted (even in rodents), so although these studies lack the specificity of those in Section 2, they provide a stepping-stone between understanding processes and understanding behavioural phenotypes or symptoms. We will focus on two forms of impulsivity - impulsive action and impulsive choice (Dalley et al., 2011; Evenden, 1999). *Impulsive actions* are characterised by motor disinhibition, which can be further broken down into actions that must be inhibited (waiting) and those that have already begun and must then be terminated (stopping) (Winstanley et al., 2006a; Eagle and Baunez, 2010). Waiting can be examined on the 5-choice serial reaction time task (5-CSRTT) by measuring premature responses that occur when the subject responds prior to stimulus onset during the inter-trial interval, triggering a time out period (Robbins, 2002). Another test of waiting is the differential reinforcement of low rates of responding (DRL) task where subjects must wait until a delay period has lapsed before responding and any response that occurs too soon leads to resetting of the delay period and, thereby, delays reward. In contrast, stopping can be measured using the stop signal reaction time task (SSRT), where a two-step response starts but must then be stopped if a cue is presented after the first response (Eagle and Robbins, 2003). The time required to stop on 50 % of trials is referred to as the 'stop signal reaction time' (SSRT), with a longer SSRT indicating more time is required to successfully stop the second response.

Impulsive choice describes the selection of likely or immediate small

rewards over unlikely or delayed large rewards. This behaviour can be measured with tasks such as delay-discounting and probabilistic-discounting. In these tasks, two responses are presented in fixed and free choice trials and the large reward option is discounted by altering the delay or probability of reinforcement, leading to gradual selection of the small reward option. In either case, the action can be made without delay, separating delayed gratification from impulsive motor action.

Compulsive behaviour is often more difficult to measure in rodents where knowledge of contingencies and consequences can be harder to determine. In animal models, perseveration may occur because of inadequate learning about contingency changes or an inability to switch actions. Evidence from human studies suggests that compulsive actions often occur despite insight of the negative consequences and this will be a crucial element in considering rodent studies. To explore action in the face of negative consequences, we can examine perseverative responding after outcome delivery on the 5-CSRTT (Robbins, 2002), during reversal learning, in schedule-induced polydipsia (SIP) where intermittent food delivery drives excessive drinking behaviour (Moreno and Flores, 2012), and punished extinction tasks where rewarded behaviour is extinguished using punishment (such as foot shock).

3.1. Impulsivity

3.1.1. Impulsive action

Premature responding on the 5-CSRTT has been used by a range of studies to examine the circuitry behind impulsive action and to phenotype individuals for baseline levels of impulsivity (Robbins, 2002; Muir et al., 1996; Jupp et al., 2013; Belin-Rauscent et al., 2016; Murphy et al., 2012; Chudasama et al., 2003; Feja and Koch, 2014; Paine et al., 2011). These studies have implicated the IL as the primary cortical region involved in impulsive action. Premature responses were substantially increased by lesions to the IL, with further evidence from IL infusion of an NMDA antagonist and muscimol ($GABA_a$ agonist), but not bicuculline ($GABA_a$ antagonist) also increasing premature responding (Murphy et al., 2012; Jupp et al., 2013; Belin-Rauscent et al., 2016; Murphy et al., 2012; Chudasama et al., 2003; Feja and Koch, 2014; Paine et al., 2011). Lesions encompassing the whole OFC (medial, ventral and lateral) increased premature responding and omissions on the 5-CSRTT (Chudasama et al., 2003). Parallel examination of the effects of lesions ACC did not increase premature responding demonstrating specificity of the role of the IL and OFC (Chudasama et al., 2003). Lesions to the post-genual cingulate cortex have been found to increase premature responses on the 5-CSRTT, but not the pre-genual cingulate (ACC) or PL cortex (Muir et al., 1996; Chudasama and Muir, 2001). However, a post-mortem study found reduced $GABA_a$ binding in ACC (but not PL, IL or vlOFC) of 'high impulsive' rats compared to rats with low levels of premature responding on the 5-CSRTT (Jupp et al., 2013). In contrast to the trend for increased impulsivity with IL lesions, gustatory insular cortex (IC) lesions reduced premature responses on the 5-CSRTT selectively in high impulsive rats where it was also found that the thickness of the gustatory IC correlated with the premature response rates (Belin-Rauscent et al., 2016). The role of the prefrontal insular portion, the anterior insular cortex, remains to be studied.

There have been far fewer studies examining the role of specific cortical regions in waiting using the DRL task. Lesions of the mPFC (PL and IL) impaired performance with mice showing a flattened response distribution around the delay time with training, rather than a distinct peak in responses. However, their response accuracy was enhanced with the provision of a cue indicating the correct response time, suggesting a deficit in temporal judgement rather than response inhibition per se (Cho and Jeantet, 2010). Depleting dopamine from the mPFC (PL and IL) also leads to more impulsive responding (Sokolowski and Salamone, 1994) and stimulating μ-opioid receptors in the IL increases impulsive responding on the DRL (Selleck et al., 2015). Overall, these results suggest that withholding a response until it is appropriate is impaired by IL loss of function and that gustatory IC function influences performance in a baseline-dependent manner.

The inability to stop an action that has been initiated can be measured using the SSRT. It should be noted that 'stopping' may share some features with compulsivity in that an initiated response must be terminated, however the time scales are likely to be quite different (Robbins et al., 2012). On the SSRT task, whole OFC (medial, ventral and lateral) lesions slow the stop reaction time (i.e. lesioned rats needed an earlier signal to be able to stop) without a change in stopping capacity when the stop signal was not delayed (effectively a no-go procedure) and no change to go response times (Eagle et al., 2008). Lesions to the lateral PFC (vlOFC and anterior IC) also did not impair performance on a go/no-go odour discrimination task, but did impair the reversal component (Schoenbaum et al., 2002). In contrast to lesion studies, vlOFC inactivation does not increase SSRT (Bari et al., 2011). Why OFC lesions impair stopping an action that has already started but not one that is waiting to be initiated is not entirely clear and further research is needed to address this question (Eagle and Baunez, 2010). Research using large OFC manipulations has suffered from contrasting effects in other behavioural paradigms, which have become clearer once the medial, ventral and lateral OFC sub-regions have been targeted specifically and this is likely to be another function that is not homogenously controlled (Izquierdo, 2017). This has not been examined to date and would help to clarify discrepancies.

There is no effect of mPFC lesions (ACC, PL and IL combined), or IL alone on SSRT or go/no-go tasks (Eagle and Robbins, 2003; Eagle et al., 2008). Yet, inactivation leads to an increased SSRT in the ACC and PL, but not IL cortex (Bari et al., 2011). Given the ACC is linked to error detection, involvement in SSRT may be related to stop cue detection and not necessarily response inhibition (Bari et al., 2011). Bari et al. (2011) also found that infusion of atomoxetine in the PL and vlOFC improved SSRT, suggesting noradrenergic tone within these regions is important for stopping responses (Bari et al., 2011). Although the IL does not play a role here, the conflicting data between lesion and inactivation studies in ACC, PL and OFC suggest that the effect of acquisition and performance of the SSRT may require different circuits or speak to the redundancy and/or compensation occurring in lesion studies. Using temporally specific manipulations and visualizing activity patterns across cortical regions could provide key information about how the stop signal is generated.

Overall, these results suggest that within the cortex, stopping is dependent on ACC, PL and OFC circuits. In addition, the IL and insular cortex have selective but opposing roles in the control of premature responses. However, there is still work to be done in further refining regional manipulations, which may clarify some inconsistencies between studies.

3.1.2. Impulsive choice

Lesions of the mPFC (PL and IL) and ACC were found not to alter impulsive choice on delay discounting (Cardinal et al., 2001) and neither does inactivation of the IL alone (Feja and Koch, 2014). While lesions of the PL and IL were found to impact delayed-reward choices, the impairment occurs on both ascending and descending discounting schedules (Cardinal et al., 2001). This indicates responding may be inflexible or less sensitive to change rather than more or less impulsive. This could occur because of impaired assessment of delay costs but not impulsive choice. A recent multiunit electrophysiological recording study in the PL found that high impulsive rats had a greater proportion of cue and press responsive neurons firing for the small/immediate reward compared to low impulsive rats (Sackett et al., 2019). This is consistent with suggestions that cortical regions hold information across delay periods and suggests altered PL functioning in impulsive animals. Infusion of the D2 receptor antagonist, raclopride, into the PL or vlOFC impaired shifting towards the smaller reward with increasing delays, however this may reflect the learning deficit detected in earlier mPFC lesion studies as this was only examined using a descending design (Pardey et al., 2013). There is mixed evidence for OFC

involvement when making choices that involve delays (Winstanley et al., 2006b; Zeeb et al., 2010; Mobini et al., 2002). The use of cues during the delay period may be an important part of this puzzle, as cues not only enhance learning by bridging the response-outcome gap but may also act as a conditioned reinforcer. Zeeb and colleagues found an interaction between baseline levels of impulsivity and the use of cues, such that lOFC inactivation increased impulsivity in low impulsive rats when the delay was cued but decreased impulsive choice in high impulsive rats when the delay was not cued (Zeeb et al., 2010). Conflicting evidence for both increasing and decreasing discounting in earlier OFC studies may also be due to functional heterogeneity within the OFC, with differences between mOFC and vlOFC emerging (Izquierdo, 2017; Winstanley et al., 2004).

The first evidence for dissociable roles of the medial and lateral OFC sub-regions in impulsive choice found that lesions of the mOFC *increased* preference for the large delayed reward whereas lesions of the lOFC *decreased* this preference (Mar et al., 2011). In contrast, when the response locations were reversed, lesions to the mOFC hastened switching while the lOFC lesions impaired reversal of contingencies. Another study found intact delay-discounting after mOFC inactivation but an increase in risky choice, with greater propensity to stay after an unlikely win, in a probabilistic-discounting task (Stopper et al., 2014). It was suggested that the mOFC is important for *sensitivity to relative value* as function of size, delay, uncertainty etc. Studies examining firing rates in the mOFC and lOFC during a task where odours indicated rewards with different delays and value and rats needed to choose between making a left and right response, found that both OFC regions were less active when reward delivery was delayed (Burton et al., 2014; Roesch et al., 2006). However, in contrast to most reward-related signals in the brain, the mOFC was found to increase activity during sampling when an odour was associated with a low value reward (Burton et al., 2014). These results suggest that the mOFC may aid decision-making by signalling the less desirable outcome as, without this information, decisions could be riskier, delay tolerant or imprecise. These findings highlight the importance of considering the role of sensitivity to reward value generally, and sensitivity to the magnitude of difference between choices, as potential drivers of impulsive responding. In contrast to the mOFC, the lOFC may contribute more to inhibitory control or to the updating of associative representations (Mar et al., 2011).

Larger lesions of the OFC have found that preference was shifted to small/immediate and small/certain rewards on delay- and probabilistic-discounting tasks (Mobini et al., 2002). However, when isolated to the mOFC, Stopper et al. (2014) found that inactivation increased risky choices via greater win-stay choices under a probabilistic schedule (irrespective of ascending or descending risk schedules) (Stopper et al., 2014). They found no effect on delay-discounting and suggested that the mOFC was required for explorative behaviour or negating the impact of unlikely wins (Stopper et al., 2014), consistent with the suggestion that this region provides a conservative signal to avoid risk. This is in contrast to the observed effects of ACC, IC (anterior and gustatory) and lOFC inactivation where there was no effect on probabilistic discounting (St Onge and Floresco, 2010). There was also no effect of lOFC pathway specific ablation on a three-choice probabilistic discounting task (Groman et al., 2019). Yet, recordings in the lOFC and anterior IC have demonstrated that firing patterns were responsive to cues conferring information about reward probability on a Pavlovian conditioning task (Jo and Jung, 2016). These results suggest that investigating the medial and lateral OFC sub-regions, as well as anterior and gustatory IC regions, will be required to clarify functionality and future studies will need to pay careful attention to protocol design.

In a T-maze discounting task, offering small and large rewards, rats were trained with the choice between a small immediate or large delayed reward in the first experiment or an easy small versus effortful large reward in the second experiment until they preferred the large reward in both cases (Rudebeck et al., 2006). Rats were then lesioned in either the OFC (medial, ventral and lateral) or ACC. OFC (but not ACC) lesioned rats stopped responding for the delayed large reward and ACC (but not OFC) lesioned rats stopped choosing the effortful arm with the large reward. These results demonstrate dissociable processing of delay and effortful costs within the OFC and ACC, respectively. Choice about economic value, without temporal cost or risk, does not appear to require either the mOFC or lOFC (Gardner et al., 2018, 2017). These studies highlight the importance of considering relative as well as absolute risk/value, different types of cost or risk, and also the importance of assessing ascending and descending discounting curves for more general learning or flexibility deficits (St Onge and Floresco, 2010).

3.2. Compulsivity

3.2.1. Perseveration

Inactivation of the mPFC (PL and IL) increased perseverative responding on the 5-CSRTT (Feja and Koch, 2014). Further studies found OFC (medial, ventral and lateral) and PL, but not IL, lesions also increased perseverative responding on the 5-CSRTT with a subtle effect in ACC lesioned rats (Chudasama et al., 2003; Chudasama and Muir, 2001). The OFC effects were consistent with the increased perseverative responding found on the reversal learning task (especially in the early stages of reversal) (Chudasama and Robbins, 2003). Further studies of the cingulate cortex found both the post-genual cingulate and medial frontal area (ACC and PL) increased perseverative responding, with the latter possibly driven by the PL as described above (Muir et al., 1996). Although not specifically a test of compulsive behaviour, the progressive ratio (PR) test involves the continuation of responding with diminishing returns and has the potential to detect perseverative responding. It has been shown that mOFC inactivation increased while stimulation decreased PR break point. This could suggest mOFC involvement in cost-benefit assessment of effort and reward in decision-making, providing further support for a role in conservative decision-making. There is limited causal evidence for the role of the cortex in schedule-induced polydipsia and this is an area for future research.

3.2.2. Reversal and switching

Reversal and switching tasks have found deficits associated with PL, IL and OFC with quite clear dissociations between these regions. PL lesions do not induce reversal deficits, but do lead to perseveration in set-shifting (see review (Brown and Tait (2016))). In contrast, the OFC has convincingly been implicated in reversal learning (Schoenbaum et al., 2002; McAlonan and Brown, 2003; Schoenbaum et al., 2003; Kim and Ragozzino, 2005), but does not impair attentional set-shifting (McAlonan and Brown, 2003) (but see Chase, Tait (Chase et al., 2012)). Lesions to the IL impaired performance in some reversal tasks and not others, however the late rather than early deficits observed are likely to reflect an impairment in learning rather than perseveration (Chudasama and Robbins, 2003; Boulougouris et al., 2007; Li and Shao, 1998) (see Izquierdo et al., 2017; Hamilton and Brigman, 2015; Ragozzino, 2007 for reviews of reversal studies across the cortex). However, there are some discrepancies that should be highlighted.

Although OFC lesions generally impair reversal learning, when reversals are repeated in serial reversal tasks, OFC (medial, ventral and lateral) lesioned rats have been found to do better on a subsequent reversal (Boulougouris et al., 2007). Using a probabilistic serial reversal task, Dalton et al. (2016) also found that inactivation of the PL cortex decreased the number of errors during initial learning and reversal, while the IL and ACC lesions had no effect (Dalton et al., 2016). Specific inactivation of the mOFC impaired performance, with more errors during initial discrimination and increased perseverative responding after reversal. While lOFC inactivation generally increased the number of errors made after reversal. In support of this finding, a recent study using a lOFC to nucleus accumbens pathway specific ablation found impaired updating after negative feedback on a probabilistic reversal task (Groman et al., 2019). Further support for the role of lOFC in

incorporating feedback in reversal was found by Piantadosi et al. (2018) using a three-choice visual reversal task. They showed that lOFC inactivation increased choice of a never-rewarded stimulus rather than increasing perseveration (Piantadosi et al., 2019). Inactivation of the mOFC and lOFC on a deterministic reversal task had no significant effect, suggesting the OFC aids the integration of information about the uncertainty or predictability of changing action-outcome contingencies. A recent study by Hervig et al. (2019) using a visual (rather than spatial) serial reversal task found that inactivation of the mOFC improved reversal learning by decreasing perseveration (Hervig et al., 2019). However, inactivating the lOFC impaired reversal learning by increasing early perseveration. The opposite was found when rats were then required to learn a new pair of stimuli such that mOFC inactivation impaired new learning and the lOFC inactivation improved new learning. Importantly, neither mOFC nor lOFC manipulations impacted learning of the first set (prior to reversal), suggesting a more nuanced role in integrating new learning rather than a general role in visual discrimination learning. Here, it was found that both PL and IL inactivation reduced the number of errors to reach the final reversal criteria rather than during any particular phase from early to late learning. As discussed in Section 2, using a two-outcome contingency reversal task it was found that inactivation of the vlOFC during training prevented appropriate response selection in a subsequent devaluation test (Parkes et al., 2018), indicating a role in updating two-way contingencies in deterministic paradigms. These studies are revealing the unique roles of cortical regions in flexibly updating contingencies, with probabilistic and deterministic paradigms resulting in contrasting findings. Across studies there is a variety of methods being used (e.g., spatial, visual, serial) offering insights into the specialized and generalised roles of the OFC in reversal learning.

3.2.3. Punished responding

To examine compulsive behaviour in rodents the once desirable intake of a rewarding substance can be paired with punishment, such as a foot shock, to observe persistent responding in the face of negative consequences. It has been found that inhibiting the IL, but not PL, prevents the reduction in responding that normally occurs after the introduction of punishment (Halladay et al., 2019). When a response paired with a food reward was punished, it was found that OFC inactivation increased responding on a punished lever, while inactivation of the anterior IC impaired choice when rats could choose between a punished and unpunished lever (Jean-Richard-Dit-Bressel and McNally, 2016). Here, the PL inactivation had little effect. In another study, reward pellets were delivered into the magazine, but a cue signified rats should inhibit consumption of the reward or a foot shock would be delivered if they were collected before a cue terminated (Verharen et al., 2019). Inactivation of the PL or IL impaired the rat's ability to collect rewards at the correct time, doubling the rates of entry during the shock period. Inactivation of the mOFC and ACC increased responses during the shock but not to the same extent as seen in the PL or IL. Inactivation of the lOFC reduced successful trials and increased omissions without increasing shocks. The effect in the lOFC was unexpected but may be due to impaired timing or credit assignment issues relating to the foot shock and hence a general extinction effect is observed. Using a platform-mediated signaled avoidance task, it was shown that lOFC inactivation prevented reinvigoration of avoidance behaviour after extinction training using exposure with response prevention (Rodriguez-Romaguera et al., 2016), again demonstrating that lOFC loss of function leads to a reduction in responding. These results again highlight the complex interplay of cognitive processes within translationally relevant tasks, with different combinations of behavioural effects across lesion groups on the same task. This provides a rich data set for exploring and comparing behavioural outcomes with relevance to complex decisions, yet it is more difficult to draw clear conclusions about the exact functions being manipulated.

3.3. Summary

We have reviewed a diverse array of literature examining the different forms of impulsive and compulsive behaviour in rodents. As shown in Table 2, it is clear that, across tasks, impulsive and compulsive behaviours rely on distinct but shared cortical circuits. Impulsive actions, which include waiting and stopping, show clear dissociation in circuitry. The IL and anterior IC were shown to be important for premature responses but not for stopping. Yet the ACC, PL and OFC are important for controlling stopping but not waiting. Lesions of ACC have also been suggested to cause general disinhibition and increase rates of non-specific responding (Hvoslef-Eide et al., 2018) which is not a primary measure of many of the tasks reviewed here, however there is evidence of increased incorrect responses on the 5CSRTT (Chudasama et al., 2003).

Impulsive choice appears to be dependent on mPFC function, with PL and IL leading to learning deficits rather than influencing impulsive responding. While evidence suggests mOFC loss of function increases preference for a delayed large or risky large reward, the opposite has been found for the lOFC. Given the contrast in roles for the OFC in impulsive action and choice, the OFC may be more involved in action control through the computation of state-dependent affective value rather than the decision to act or inhibit per se (Bari and Robbins, 2013). The ACC on the other hand appears to be more important for assessment of effortful responding. There are still many inconsistencies in the impulsive choice literature and further studies with discrete anatomical manipulations and comprehensive behavioural assessments are needed.

Compulsive behaviours involving excessive responding, such as perseveration on the 5CSRTT, are dependent on PL and OFC, but not IL function. Although superficially similar, perseveration on reversal is quite different as it measures the ongoing response to a previously rewarded stimulus when given a choice. Perseveration on reversal has classically been associated with OFC function, with the lOFC playing a key role. Interestingly, neither the PL nor IL is required for reversal, but lesioning the PL leads to perseveration in set-shifting. Given the often-opposing roles of the mOFC and vlOFC, it will be important for future studies to determine whether many of these behaviours are reliant on the vOFC, lOFC or mOFC. Further, the role of the anterior IC has not been as widely studied in impulsive and compulsive tasks and this should be addressed by future studies (Table 2).

4. Discussion

4.1. Synthesizing disparate literatures in decision-making

Studies on habits and goal-directed behaviour have remained largely detached from studies on compulsive and impulsive actions. While in both cases tasks were developed to understand action control, the tasks were designed to broadly address separate questions. In the former, tasks were designed to understand the fundamental processes involved in making choices and the control of actions using a learning theory framework. In the latter, tasks were driven by a desire to understand the neuroscience behind the regulation of learned behaviour and impairments in response control. As such, these domains of decision-making research have evolved independently with little overlap in the current literature.

Indeed, most of the tasks used to study compulsive and impulsive actions, as described in Section 3, were designed to measure broader domains of cognitive control and, therefore, it is often difficult to delineate the precise learning processes that may contribute. This is particularly true for the segregation of instrumental and Pavlovian processes, which both rely on corticostriatal circuitry. But these tasks represent a link to studies conducted in humans and are important for improving our understanding of psychiatric conditions and symptoms, which are also unlikely to be dependent on a single cognitive process.

Table 2
Summary of behavioural findings within the translational domain following loss of function in various rodent cortical regions.

	ACC	mPFC	PL	IL	OFC	mOFC	lOFC	Ant. IC	Gust. IC
Impulsive Action									
5-CSRTT Premature	–	↑	–	↑	↑/-				↓
SSRT	↑	↑-	↑-	–	↑-				
Impulsive Choice									
Delay-discounting	–	–	–	–	↑↓	↓/-	↑↓		
Probabilistic-discounting	–	–				↑	–	–	
Perseveration									
5-CSRTT Perseveration	Subtle	↑	↑	–	↑				
Reversal Perseveration	–		–	–	↑	–	↑	↑	
Set shifting Perseveration		↑	↑		–				
PR responding						↑			
Punished extinction	↑		↑-	↑-	↑	↑	↓	↑	

ACC: anterior cingulate cortex; mPFC: medial prefrontal cortex (encompassing prelimbic and infralimbic cortices); PL: prelimbic cortex; IL: infralimbic cortex, OFC: orbitofrontal cortex; mOFC: medial orbitofrontal cortex; lOFC: lateral regions of orbitofrontal cortex; Ant. IC: anterior insular cortex; Gust. IC: gustatory insular cortex. Arrows indicate the direction of the behavioural change following loss of function with (-) indicating no change following loss of function.

By contrast, habits and goal-directed actions are studied using specific tasks that allow us to identify the learning and memory processes driving behaviour. And, conversely, it can be difficult to directly integrate tests for habits or for goal-directed actions within more complex translational tasks, although this is a growing area. Yet, we are studying the same neural circuits and, ultimately, overlapping behavioural systems. If we can bring together studies from across this spectrum, from precise function to symptom or phenotype level, then both fields have much to gain.

4.2. Cortical contributions to decision-making: similarities and differences between the domains

The traditional view of the prefrontal cortex as being central to response inhibition is evolving with greater appreciation for its role in decision-making and associative learning as well as an enhanced understanding of its functional heterogeneity (Bradfield and Hart, 2020; Wilson et al., 2014; Izquierdo, 2017; Sharpe et al., 2019; Lopatina et al., 2017; Stalnaker et al., 2015). Here, we have attempted to present the most consistent findings to date regarding the cortical bases of action control. As illustrated in Fig. 2, these findings show that each cortical sub-region is likely to be involved in multiple functions and an isolated region can play quite distinct roles depending on the task at hand. It can also be seen that the findings in Section 2 are more consistent than those in Section 3, indicative of the limited but consistent range of studies used in this field compared to the highly diverse range of tasks used to measure impulsive and compulsive behaviours. The different roles that regions play across tasks are likely to depend on the arrangement of cortical inputs and outputs, such that two different processes that co-localise in their dependence on a region may be reliant on different input and output circuitry. In an effort to reconcile findings from different platforms, it is often tempting to consolidate results to a single function, but this neglects to acknowledge the intrinsic complexity of circuitry into and out of each region.

Although we did not set out to review the involvement of the broader circuits of each cortical region, these will define the functional specialization and limitations of each region. For example, more lateral regions receive more sensory and motor feedback whereas medial regions have strong thalamic and visceral inputs. Output, particularly to the striatum in the context of the tasks reviewed, is also highly relevant to the regulation of behaviour with a topographical pattern of innervations that in many ways mirrors the layout of the cortex. Cortical regions also share patterns of input and output, so these alone do not define function. An understudied component of the cortex is the cortico-cortical connections that could allow parallel processing and integration of information from a variety of indirect sources. With the development of cell and circuit specific technologies, we now have the capacity to isolate and characterise these networks in ways that were not available when these tasks were originally developed. For example, cortico-cortical interactions have largely been ignored in comparison to studies on more distant circuits, including corticostriatal, corticolimbic, corticothalamic, or thalamocortical pathways. Advanced technologies, such as viral and optical tools, provide new opportunities to better understand cross-communication and synergistic activity *within* the cortex. We suggest that rather than isolating functions to regions (i.e. *region A does function X*), that we can develop a more nuanced description by looking deeper within and across networks (i.e. *under conditions A, specific cells/circuits within region B and C contribute to function X*). Nevertheless, reviewing a variety of behavioural paradigms across fields of decision-making does indicate some functional convergence within the cortical sub-regions.

4.3. Patterns of functional convergence

Within the limitations described above, here we attempt to draw some general conclusions and examples of cross-task evidence for functional convergence within the cortex. Data from both fields indicates a functional dichotomy between the PL and IL. Loss of function in the PL impairs goal-directed behaviour, but not habits. Within the translational literature, goal-directed impairments would be expected to lead to general learning deficits or impairments in rule-formation. This notion is strongly supported by the role of the PL in set-shifting (rule learning) but not reversal tasks. Given general learning deficits were found on delay discounting in mPFC (PL and IL) lesioned rats, with deficits on both ascending and descending curves, it is tempting to suggest that this effect would also be seen with isolated PL lesions. The perseveration observed in the 5-CSRTT with PL lesions is not observed on reversal tasks and, combined with evidence from the SSRT, suggests a deficit in stopping an ongoing action rather than choosing to continue the same response across trials. Overall, these findings are consistent with a role for the PL in learning the relationship between an action and its associated outcome, but also suggest a role in stopping actions that have been initiated. By contrast, IL loss of function is associated with impaired habitual, but not goal-directed control. The IL is not involved in stopping actions but is required to wait to respond. This convergent evidence suggests IL is involved in various forms of pre-potent responding. Whether the circuits supporting habits interfere with updating at *initiation* versus *execution* of an action has not been as well defined in literature. But based on comparisons in the translational literature, it could be inferred that IL-dependent habitual responding is limited to response initiation and not to ongoing responses. The anterior IC is also required to inhibit premature or punished responses but its role in compulsive and impulsive actions has not been as widely studied as the orbitofrontal and medial prefrontal cortices. In addition,

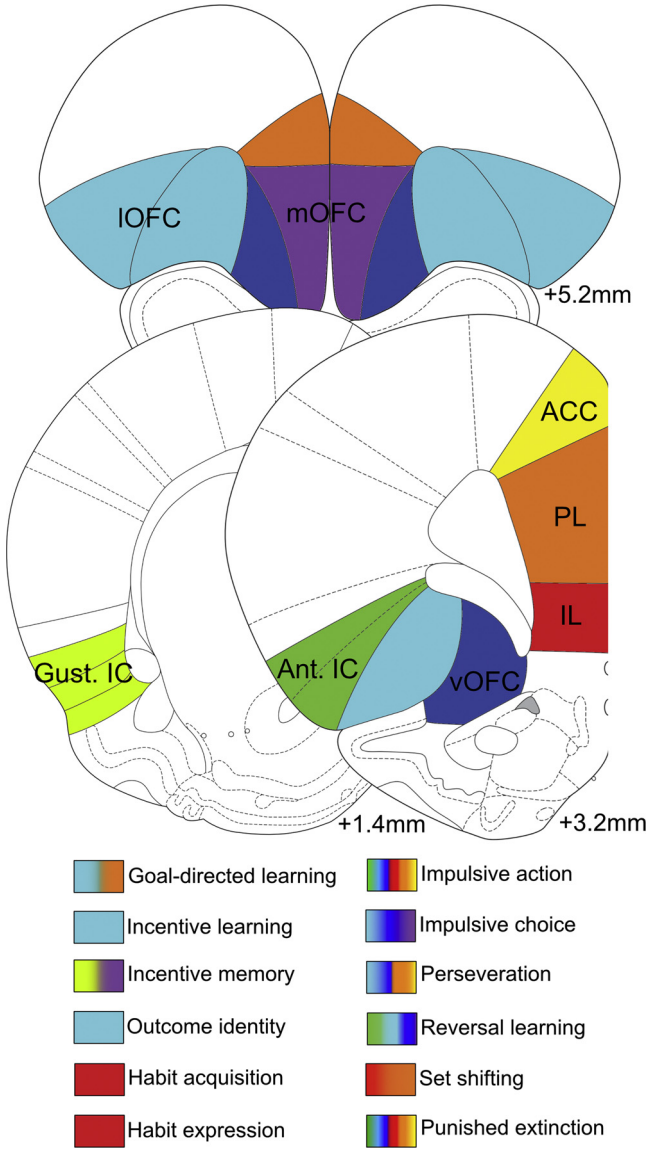


Fig. 2. Cortical regulation of actions. Graphical representation of the cortical regions involved in various behavioural functions. Atlas images taken from Paxinos and Watson 6th Edition The Rat Brain (Paxinos and Watson, 2007). ACC: anterior cingulate cortex; PL: prelimbic cortex; IL: infralimbic cortex, vOFC: ventral orbitofrontal cortex; mOFC: medial orbitofrontal cortex; lOFC: lateral orbitofrontal cortex; Ant. IC: anterior insular cortex; Gust. IC: gustatory insular cortex.

its specific role in goal-directed and habitual behaviour remains to be investigated.

Research in both domains also points to a role for OFC in the representation of outcome value and/or outcome identity. Indeed, there is no evidence that the OFC (medial or lateral regions) is required for instrumental learning per se but responses can become insensitive to manipulations of outcome value (either motivational value or reward magnitude) following loss of OFC function. Moreover, both domains agree that there exists a functional heterogeneity within OFC. For instance, mOFC inactivation enhances early reversal and impairs sensitivity to devaluation. Both these functions suggest a role in retrieving current reward value, where inactivation during early reversal may reduce interference by impairing retrieval of the previous association as is found in devaluation. By contrast, inactivation of lateral regions of OFC impairs reversal learning and the ability to encode the current reward value in an incentive learning paradigm. Understanding the distinct and overlapping roles of the OFC sub-regions in decision-making processes should remain a priority in future research (Izquierdo, 2017). One fruitful avenue of research will be to better understand the involvement of OFC in instrumental versus Pavlovian conditioning. For instance, there are clear demonstrations that vlOFC loss of function leaves instrumental outcome devaluation intact (when there is no change in outcome identity) but impairs Pavlovian reinforcer devaluation (Ostlund and Balleine, 2007; Parkes et al., 2018; Panayi and Killcross, 2018). Understanding this disparity may allow us to better grasp the function(s) of rodent ventral and lateral OFC in behavioural control.

4.4. Primate homology

It will be increasingly important to compare and contrast these findings with those from primate studies. Indeed, some clear similarities and dissociations have been demonstrated in studies that have compared functions across species (for example, in contingency degradation and reversal learning). We have restricted this review to evidence in rodents, however if these results are to contribute to our understanding of human cognition, then cross-species differences in functional and anatomical homology will need to be constantly intertwined. Substantial efforts are being made in this space across both the fields addressed in this review. There are many excellent reviews highlighting similarities and limitations in homology across species (Heilbronner et al., 2016; Balleine and O'Doherty, 2010; Laubach et al., 2018; Carlen, 2017; Wallis, 2011). As identified by others, consistency of nomenclature and atlas-based visual representation of manipulations will aid the integration of findings now and in the future as naming conventions and boundaries are refined. Despite the cortical expansion and differentiation seen in primates, rats and mice are certainly capable of solving simplified tasks and performing many of the cognitive functions of interest. Therefore, the more restricted rodent cortical circuitry provides an ideal reductionist model for understanding decision-making. As long as these limitations are addressed, then we have much to gain by taking the results from each species and comparing the circuits involved.

4.5. Future directions

A few key points can be made from this collection of studies. First, these tasks allow us to isolate a range of cognitive processes and comparing performance across tasks will help to build a larger and more comprehensive model of instrumental behaviour. A benefit of this approach is the ability to compare manipulations across cognitive processes to confirm specific and generalised functions. Second, cortical regions do not operate single functions, they appear to work on different problems under different conditions and trying to condense functions to a single explanation is unnecessary and may not be accurate. Given the proximity and connectivity of cortical regions, it is likely that these regions cooperate as hubs that receive various inputs and then direct information to subcortical structures. Although understudied, we need to consider cortico-cortical connectivity as substantial pathways within the standard long-range circuits we commonly discuss. For example, given the ventral and lateral portions of the PFC receive more sensory input and the dorsal and medial portions have been linked to more goal or executive function roles, many cognitive functions likely require recruitment of multiple cortical sub-regions and cortico-cortical communication will be important for the transfer of information required to orchestrate decisions. This will require continued precise targeting of cortical regions and the implementation of circuit- and cell-specific tools. To build a more comprehensive model, it would be helpful to use more consistent and reproducible protocol methodology, particularly in the paradigms used to study compulsive and impulsive actions. Indeed, the studies described under Section 2 are an example of how defined protocols deliver more clarity on the

learning processes being studied. At the very least, we must have a greater appreciation for the diversity of processes being measured in our tasks and try to address this limitation when making conclusions about cortical functions.

Overall, this review demonstrates that comparing studies of specific processes in habits and goal-directed actions with the impaired control measured in impulsive and compulsive tasks gives a more general overview of how cortical regions operate under different conditions. This highlights the diversity of functions a region may be involved in, as well as the number of regions that may be involved in any particular task. Trying to label each region with a function or specific purpose may be futile given the diversity of pathway-specific circuits and cortical cross-talk that happens within regions. Future studies should examine these specific pathways within the cortex and consider clues from across the research divide. Despite differences in perspective and approach, similar tools are being used to study how the rodent cortex contributes to choices and actions. There is greater potential for insight and collaboration across these fields than is currently being utilised. Integration across these fields may be critical for understanding cortical regulation of adaptive choices and how it can go wrong.

Funding

This work was supported by research grants from the National Alliance for Research on Schizophrenia and Depression (27402) and the French National Agency for Scientific Research (ANR-19-CE37-0004-07) awarded to SP; and a National Health and Medical Research Council Early Career Fellowship (APP1122221) awarded to KT.

Acknowledgements

This manuscript is based on discussions with Andrew Holmes and Bernard Balleine during the symposium *Recent Perspectives on Action Control and Decision-Making* at the 2019 International Behavioral Neuroscience Society meeting. We thank Andrew Holmes and Bernard Balleine for these insightful discussions as well as their comments during the preparation of this manuscript. We also thank Etienne Coutureau and Trevor Robbins for their immensely helpful comments on an earlier version of this manuscript.

References

Adams, C.D., 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. Q. J. Exp. Psychol. (Hove) 34 (2), 77–98.
Adams, C.D., Dickinson, A., 1981. Instrumental responding following reinforcer devaluation. Q. J. Exp. Psychol. 1981 (33), 109–122.
Alcaraz, F., et al., 2018. Thalamocortical and corticothalamic pathways differentially contribute to goal-directed behaviors in the rat. Elife 7.
Allen, G.V., et al., 1991. Organization of visceral and limbic connections in the insular cortex of the rat. J. Comp. Neurol. 311 (1), 1–16.
Aly-Mahmoud, M., et al., 2017. Role of anterior cingulate cortex in instrumental learning: blockade of dopamine D1 receptors suppresses overt but not covert learning. Front. Behav. Neurosci. 11, 82.
Balleine, B.W., 2019. The meaning of behavior: discriminating reflex and volition in the brain. Neuron 104 (1), 47–62.
Balleine, B.W., Dezfouli, A., 2019. Hierarchical action control: adaptive collaboration between actions and habits. Front. Psychol. 10, 2735.
Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37 (4–5), 407–419.
Balleine, B.W., Dickinson, A., 2000. The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. J. Neurosci. 20 (23), 8954–8964.
Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35 (1), 48–69.
Balleine, B.W., et al., 1995. Motivational control of heterogeneous instrumental chains. J. Exp. Psychol. Anim. Behav. Process. 21, 203–217.
Baltz, E.T., et al., 2018. Orbital frontal cortex updates state-induced value change for decision-making. Elife 7.
Barfield, E.T., Gourley, S.L., 2019. Glucocorticoid-sensitive ventral hippocampal-orbitofrontal cortical connections support goal-directed action - curt richter award paper 2019. Psychoneuroendocrinology 110, 104436.
Bari, A., Robbins, T.W., 2013. Inhibition and impulsivity: behavioral and neural basis of response control. Prog. Neurobiol. 108, 44–79.
Bari, A., et al., 2011. Prefrontal and monoaminergic contributions to stop-signal task performance in rats. J. Neurosci. 31 (25), 9254–9263.
Barker, J.M., Torregrossa, M.M., Taylor, J.R., 2013. Bidirectional modulation of infralimbic dopamine D1 and D2 receptor activity regulates flexible reward seeking. Front. Neurosci. 7, 126.
Belin, D., et al., 2008. High impulsivity predicts the switch to compulsive cocaine-taking. Science 320 (5881), 1352–1355.
Belin-Rauscent, A., et al., 2016. From impulses to maladaptive actions: the insula is a neurobiological gate for the development of compulsive behavior. Mol. Psychiatry 21 (4), 491–499.
Boulougouris, V., Dalley, J.W., Robbins, T.W., 2007. Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. Behav. Brain Res. 179 (2), 219–228.
Bradfield, L.A., Hart, G., 2020. Rodent medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. Neurosci. Biobehav. Rev. 108, 287–294.
Bradfield, L.A., et al., 2015. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron 88 (6), 1268–1280.
Bradfield, L.A., Hart, G., Balleine, B.W., 2018. Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. Neurobiol. Learn. Mem. 155, 463–473.
Brown, V.J., Tait, D.S., 2016. Attentional set-shifting across species. Curr Top Behav Neurosci 28, 363–395.
Burton, A.C., et al., 2014. Increased firing to cues that predict low-value reward in the medial orbitofrontal cortex. Cereb. Cortex 24 (12), 3310–3321.
Bussey, T.J., Everitt, B.J., Robbins, T.W., 1997. Dissociable effects of cingulate and medial frontal cortex lesions on stimulus-reward learning using a novel Pavlovian autoshaping procedure for the rat: implications for the neurobiology of emotion. Behav. Neurosci. 111 (5), 908–919.
Cardinal, R.N., et al., 2001. Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science 292 (5526), 2499–2501.
Cardinal, R.N., et al., 2002. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci. Biobehav. Rev. 26 (3), 321–352.
Carlen, M., 2017. What constitutes the prefrontal cortex? Science 358 (6362), 478–482.
Cechetto, D.F., Saper, C.B., 1987. Evidence for a viscerotopic sensory representation in the cortex and thalamus in the rat. J. Comp. Neurol. 262 (1), 27–45.
Chase, E.A., Tait, D.S., Brown, V.J., 2012. Lesions of the orbital prefrontal cortex impair the formation of attentional set in rats. Eur. J. Neurosci. 36 (3), 2368–2375.
Cho, Y.H., Jeantet, Y., 2010. Differential involvement of prefrontal cortex, striatum, and hippocampus in DRL performance in mice. Neurobiol. Learn. Mem. 93 (1), 85–91.
Chudasama, Y., Muir, J.L., 2001. Visual attention in the rat: a role for the prelimbic cortex and thalamic nuclei? Behav. Neurosci. 115 (2), 417–428.
Chudasama, Y., Robbins, T.W., 2003. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. J. Neurosci. 23 (25), 8771–8780.
Chudasama, Y., et al., 2003. Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. Behav. Brain Res. 146 (1–2), 105–119.
Colwill, R.M., Rescorla, R.A., 1985. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. J. Exp. Psychol. Anim. Behav. Process. 11 (4), 520–536.
Corbit, L.H., Balleine, B.W., 2003. The role of prelimbic cortex in instrumental conditioning. Behav. Brain Res. 146 (1–2), 145–157.
Coutureau, E., Killcross, S., 2003. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behav. Brain Res. 146 (1–2), 167–174.
Coutureau, E., Marchand, A.R., Di Scala, G., 2009. Goal-directed responding is sensitive to lesions to the prelimbic cortex or basolateral nucleus of the amygdala but not to their disconnection. Behav. Neurosci. 123 (2), 443–448.
Coutureau, E., et al., 2012. The role of the rat medial prefrontal cortex in adapting to changes in instrumental contingency. PLoS One 7 (4), e33302.
Dalley, J.W., Everitt, B.J., Robbins, T.W., 2011. Impulsivity, compulsivity, and top-down cognitive control. Neuron 69 (4), 680–694.
Dalton, G.L., et al., 2016. Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. J. Neurosci. 36 (6), 1996–2006.
de Wit, S., Dickinson, A., 2009. Associative theories of goal-directed behaviour: a case for animal-human translational models. Psychol Res 73 (4), 463–476.
de Wit, S., et al., 2006. Dorsomedial prefrontal cortex resolves response conflict in rats. J. Neurosci. 26 (19), 5224–5229.
Dickinson, A., 1985. Actions and habits: the development of behavioral autonomy. Philos Trans R Soc Lond B Biol Sci 308 (1135), 67–78.
Dickinson, A., 1994. Instrumental conditioning. In: Mackintosh, N.J., Mackintosh, N.J. (Eds.), Animal Learning and Cognition. Academic Press, San Diego, CA, US, pp. 45–79.
Dickinson, A., Balleine, B.W., 1994. Motivational control of goal-directed action. Anim. Learn. Behav. 22 (1), 1–18.
Eagle, D.M., Baunez, C., 2010. Is there an inhibitory-response-control system in the rat? Evidence from anatomical and pharmacological studies of behavioral inhibition. Neurosci. Biobehav. Rev. 34 (1), 50–72.
Eagle, D.M., Robbins, T.W., 2003. Lesions of the medial prefrontal cortex or nucleus accumbens core do not impair inhibitory control in rats performing a stop-signal

reaction time task. Behav. Brain Res. 146 (1–2), 131–144.
Eagle, D.M., et al., 2008. Stop-signal reaction-time task performance: role of prefrontal cortex and subthalamic nucleus. Cereb. Cortex 18 (1), 178–188.
Evenden, J.L., 1999. Varieties of impulsivity. Psychopharmacology (Berl) 146 (4), 348–361.
Feja, M., Koch, M., 2014. Ventral medial prefrontal cortex inactivation impairs impulse control but does not affect delay-discounting in rats. Behav. Brain Res. 264, 230–239.
Fineberg, N.A., et al., 2010. Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. Neuropsychopharmacology 35 (3), 591–604.
Franklin, K., Paxinos, G., 2019. Paxinos and Franklin'S The Mouse Brain in Stereotaxic Coordinates. Elsevier: Academic Press.
Fresno, V., et al., 2019. A thalamocortical circuit for updating action-outcome associations. Elife 8.
Gabbott, P.L., et al., 2005. Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. J. Comp. Neurol. 492 (2), 145–177.
Gardner, M.P.H., et al., 2017. Lateral orbitofrontal inactivation dissociates devaluation-sensitive behavior and economic choice. Neuron 96 (5), 1192–1203 e4.
Gardner, M.P., et al., 2018. Medial orbitofrontal inactivation does not affect economic choice. Elife 7.
Gourley, S.L., et al., 2013. The orbitofrontal cortex regulates outcome-based decision-making via the lateral striatum. Eur. J. Neurosci. 38 (3), 2382–2388.
Gourley, S.L., et al., 2016. The medial orbitofrontal cortex regulates sensitivity to outcome value. J. Neurosci. 36 (16), 4600–4613.
Gremel, C.M., Costa, R.M., 2013. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. 4, 2264.
Gremel, C.M., et al., 2016. Endocannabinoid modulation of orbitostriatal circuits gates habit formation. Neuron 90 (6), 1312–1324.
Groman, S.M., et al., 2019. Orbitofrontal circuits control multiple reinforcement-learning processes. Neuron 103 (4), 734–746 e3.
Halladay, L.R., et al., 2019. Prefrontal regulation of punished ethanol self-administration. Biol. Psychiatry.
Hamilton, D.A., Brigman, J.L., 2015. Behavioral flexibility in rats and mice: contributions of distinct frontocortical regions. Genes Brain Behav. 14 (1), 4–21.
Hart, G., Balleine, B.W., 2016. Consolidation of goal-directed action depends on MAPK/ERK signaling in rodent prelimbic cortex. J. Neurosci. 36 (47), 11974–11986.
Hart, G., Bradfield, L.A., Balleine, B.W., 2018a. Prefrontal corticostriatal disconnection blocks the acquisition of goal-directed action. J. Neurosci. 38 (5), 1311–1322.
Hart, G., et al., 2018b. The bilateral prefronto-striatal pathway Is necessary for learning new goal-directed actions. Curr. Biol. 28 (14), 2218–2229 e7.
Heilbronner, S.R., et al., 2016. Circuit-based corticostriatal homologies between rat and primate. Biol. Psychiatry 80 (7), 509–521.
Hernandez, P.J., Schiltz, C.A., Kelley, A.E., 2006. Dynamic shifts in corticostriatal expression patterns of the immediate early genes Homer 1a and Zif268 during early and late phases of instrumental training. Learn. Mem. 13 (5), 599–608.
Hershberger, W.A., 1986. An approach through the looking glass. Anim. Learn. Behav. 14, 443–451.
Hervig, M.E., et al., 2019. Dissociable and paradoxical roles of rat medial and lateral orbitofrontal cortex in visual serial reversal learning. Cereb. Cortex.
Heyes, C., Dickinson, A., 1990. The intentionality of animal action. Mind Language 5, 87–104.
Hitchcott, P.K., Quinn, J.J., Taylor, J.R., 2007. Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. Cereb. Cortex 17 (12), 2820–2827.
Holmes, N.M., Marchand, A.R., Coutureau, E., 2010. Pavlovian to instrumental transfer: a neurobehavioural perspective. Neurosci. Biobehav. Rev. 34 (8), 1277–1295.
Hvoslef-Eide, M., et al., 2018. Effects of anterior cingulate cortex lesions on a continuous performance task for mice. Brain Neurosci. Adv. 2.
Izquierdo, A., 2017. Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. J. Neurosci. 37 (44), 10529–10540.
Izquierdo, A., et al., 2017. The neural basis of reversal learning: an updated perspective. Neuroscience 345, 12–26.
Jackson, S.A., et al., 2016. Role of the perigenual anterior cingulate and orbitofrontal cortex in contingency learning in the marmoset. Cereb. Cortex 26 (7), 3273–3284.
Jean-Richard-Dit-Bressel, P., McNally, G.P., 2016. Lateral, not medial, prefrontal cortex contributes to punishment and aversive instrumental learning. Learn. Mem. 23 (11), 607–617.
Jo, S., Jung, M.W., 2016. Differential coding of uncertain reward in rat insular and orbitofrontal cortex. Sci. Rep. 6, 24085.
Jonkman, S., Everitt, B.J., 2009. Post-learning infusion of anisomycin into the anterior cingulate cortex impairs instrumental acquisition through an effect on reinforcer valuation. Learn. Mem. 16 (11), 706–713.
Jupp, B., et al., 2013. Dopaminergic and GABA-ergic markers of impulsivity in rats: evidence for anatomical localisation in ventral striatum and prefrontal cortex. Eur. J. Neurosci. 37 (9), 1519–1528.
Killcross, S., Coutureau, E., 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb. Cortex 13 (4), 400–408.
Kim, J., Ragozzino, M.E., 2005. The involvement of the orbitofrontal cortex in learning under changing task contingencies. Neurobiol. Learn. Mem. 83 (2), 125–133.
Laubach, M., et al., 2018. What, If anything, Is rodent prefrontal cortex? eNeuro 5 (5).
Lex, B., Hauber, W., 2010. The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. Cereb. Cortex 20 (4), 873–883.
Li, L., Shao, J., 1998. Restricted lesions to ventral prefrontal subareas block reversal learning but not visual discrimination learning in rats. Physiol Behav 65 (2), 371–379.
Lopatina, N., et al., 2017. Ensembles in medial and lateral orbitofrontal cortex construct cognitive maps emphasizing different features of the behavioral landscape. Behav. Neurosci. 131 (3), 201–212.
Malvaez, M., et al., 2019. Distinct cortical-amygdala projections drive reward value encoding and retrieval. Nat. Neurosci. 22 (5), 762–769.
Mar, A.C., et al., 2011. Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. J. Neurosci. 31 (17), 6398–6404.
McAlonan, K., Brown, V.J., 2003. Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. Behav. Brain Res. 146 (1–2), 97–103.
McDonald, A.J., 1998. Cortical pathways to the mammalian amygdala. Prog. Neurobiol. 55 (3), 257–332.
McKee, B.L., et al., 2010. Operant learning requires NMDA-receptor activation in the anterior cingulate cortex and dorsomedial striatum, but not in the orbitofrontal cortex. Behav. Neurosci. 124 (4), 500–509.
Mobini, S., et al., 2002. Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology (Berl) 160 (3), 290–298.
Moreno, M., Flores, P., 2012. Schedule-induced polydipsia as a model of compulsive behavior: neuropharmacological and neuroendocrine bases. Psychopharmacology (Berl) 219 (2), 647–659.
Muir, J.L., Everitt, B.J., Robbins, T.W., 1996. The cerebral cortex of the rat and visual attentional function: dissociable effects of mediofrontal, cingulate, anterior dorsolateral, and parietal cortex lesions on a five-choice serial reaction time task. Cereb. Cortex 6 (3), 470–481.
Munster, A., Hauber, W., 2017. Medial orbitofrontal cortex mediates effort-related responding in rats. Cereb. Cortex 1–11.
Murphy, E.R., et al., 2012. Impulsive behaviour induced by both NMDA receptor antagonism and GABAA receptor activation in rat ventromedial prefrontal cortex. Psychopharmacology (Berl) 219 (2), 401–410.
Naneix, F., et al., 2009. A role for medial prefrontal dopaminergic innervation in instrumental conditioning. J. Neurosci. 29 (20), 6599–6606.
Naneix, F., et al., 2013. Adolescent stimulation of D2 receptors alters the maturation of dopamine-dependent goal-directed behavior. Neuropsychopharmacology 38 (8), 1566–1574.
Ostlund, S.B., Balleine, B.W., 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J. Neurosci. 25 (34), 7763–7770.
Ostlund, S.B., Balleine, B.W., 2007. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J. Neurosci. 27 (18), 4819–4825.
Paine, T.A., Slipp, L.E., Carlezon Jr., W.A., 2011. Schizophrenia-like attentional deficits following blockade of prefrontal cortex GABAA receptors. Neuropsychopharmacology 36 (8), 1703–1713.
Panayi, M.C., Killcross, S., 2018. Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. Elife 7.
Pardey, M.C., et al., 2013. Catecholamine receptors differentially mediate impulsive choice in the medial prefrontal and orbitofrontal cortex. J. Psychopharmacol. 27 (2), 203–212.
Parkes, S.L., Balleine, B.W., 2013. Incentive memory: evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. J. Neurosci. 33 (20), 8753–8763.
Parkes, S.L., Bradfield, L.A., Balleine, B.W., 2015. Interaction of insular cortex and ventral striatum mediates the effect of incentive memory on choice between goal-directed actions. J. Neurosci. 35 (16), 6464–6471.
Parkes, S.L., Ferreira, G., Coutureau, E., 2016. Acquisition of specific response-outcome associations requires NMDA receptor activation in the basolateral amygdala but not in the insular cortex. Neurobiol. Learn. Mem. 128, 40–45.
Parkes, S.L., et al., 2018. Insular and ventrolateral orbitofrontal cortices differentially contribute to goal-directed behavior in rodents. Cereb. Cortex 28 (7), 2313–2325.
Paxinos, G., Watson, C., 2007. The Rat Brain in Stereotaxic Coordinates, 6th ed. Academic Press, San Diego.
Paxinos, G., Watson, C., 2014. The Rat Brain in Stereotaxic Coordinates. Academic Press.
Piantadosi, P.T., et al., 2019. A novel multichoice touchscreen paradigm for assessing cognitive flexibility in mice. Learn. Mem. 26 (1), 24–30.
Ragozzino, M.E., 2007. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. Ann. N. Y. Acad. Sci. 1121, 355–375.
Robbins, T.W., 2002. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. Psychopharmacology (Berl) 163 (3-4), 362–380.
Robbins, T.W., 2017. Brain-derived neurotrophic factor and the orbitofrontal regulation of behavior. Biol. Psychiatry 81 (4), 282–284.
Robbins, T.W., et al., 2012. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. Trends Cognit. Sci. 16 (1), 81–91.
Rodriguez-Romaguera, J., et al., 2016. An avoidance-based rodent model of exposure with response prevention therapy for obsessive-compulsive disorder. Biol. Psychiatry 80 (7), 534–540.
Roesch, M.R., Taylor, A.R., Schoenbaum, G., 2006. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron 51 (4), 509–520.
Rudebeck, P.H., et al., 2006. Separate neural pathways process different decision costs. Nat. Neurosci. 9 (9), 1161–1168.
Sackett, D.A., Moschak, T.M., Carelli, R.M., 2019. Prelimbic cortical neurons track preferred reward value and reflect impulsive choice during delay discounting behavior. J. Neurosci. 39 (16), 3108–3118.
Schoenbaum, G., et al., 2002. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport 13 (6), 885–890.
Schoenbaum, G., et al., 2003. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn.

Mem. 10 (2), 129–140.
Schweimer, J., Hauber, W., 2005. Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. Learn. Mem. 12 (3), 334–342.
Selleck, R.A., et al., 2015. Endogenous opioid signaling in the medial prefrontal cortex is required for the expression of hunger-induced impulsive action. Neuropsychopharmacology 40 (10), 2464–2474.
Sharpe, M.J., et al., 2019. An integrated model of action selection: distinct modes of cortical control of striatal decision making. Annu. Rev. Psychol. 70, 53–76.
Shi, C.J., Cassell, M.D., 1998. Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. J. Comp. Neurol. 399 (4), 440–468.
Shipman, M.L., et al., 2018. Inactivation of prelimbic and infralimbic cortex respectively affects minimally-trained and extensively-trained goal-directed actions. Neurobiol. Learn. Mem. 155, 164–172.
Smith, K.S., Graybiel, A.M., 2013. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron 79 (2), 361–374.
Sokolowski, J.D., Salamone, J.D., 1994. Effects of dopamine depletions in the medial prefrontal cortex on DRL performance and motor activity in the rat. Brain Res. 642 (1–2), 20–28.
St Onge, J.R., Floresco, S.B., 2010. Prefrontal cortical contribution to risk-based decision making. Cereb. Cortex 20 (8), 1816–1828.
Stalnaker, T.A., Cooch, N.K., Schoenbaum, G., 2015. What the orbitofrontal cortex does not do. Nat. Neurosci. 18 (5), 620–627.
Stopper, C.M., Green, E.B., Floresco, S.B., 2014. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. Cereb. Cortex 24 (1), 154–162.
Tran-Tu-Yen, D.A.S., et al., 2009. Transient role of the rat prelimbic cortex in goal-directed behaviour. Eur. J. Neurosci. 30, 464–471.
Verharen, J.P.H., et al., 2019. Corticolimbic mechanisms of behavioral inhibition under threat of punishment. J. Neurosci. 39 (22), 4353–4364.
Vogt, B.A., Paxinos, G., 2014. Cytoarchitecture of mouse and rat cingulate cortex with human homologies. Brain Struct. Funct. 219 (1), 185–192.
Voon, V., Dalley, J.W., 2016. translatable and Back-translatable measurement of impulsivity and compulsivity: convergent and divergent processes. Curr. Top. Behav. Neurosci. 28, 53–91.
Wallis, J.D., 2011. Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci. 15 (1), 13–19.
Whyte, A.J., et al., 2019. Reward-related expectations trigger dendritic spine plasticity in the mouse ventrolateral orbitofrontal cortex. J. Neurosci. 39 (23), 4595–4605.
Wilson, R.C., et al., 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron 81 (2), 267–279.
Winstanley, C.A., et al., 2004. Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. J. Neurosci. 24 (20), 4718–4722.
Winstanley, C.A., Eagle, D.M., Robbins, T.W., 2006a. Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. Clin. Psychol. Rev. 26 (4), 379–395.
Winstanley, C.A., et al., 2006b. Double dissociation between serotonergic and dopaminergic modulation of medial prefrontal and orbitofrontal cortex during a test of impulsive choice. Cereb. Cortex 16 (1), 106–114.
Yin, H.H., Knowlton, B.J., Balleine, B.W., 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. 19 (1), 181–189.
Zeeb, F.D., Floresco, S.B., Winstanley, C.A., 2010. Contributions of the orbitofrontal cortex to impulsive choice: interactions with basal levels of impulsivity, dopamine signalling, and reward-related cues. Psychopharmacology (Berl) 211 (1), 87–98.
Zimmermann, K.S., et al., 2017a. Memory retention involves the ventrolateral orbitofrontal cortex: comparison with the basolateral amygdala. Neuropsychopharmacology.
Zimmermann, K.S., et al., 2017b. Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by brain-derived neurotrophic factor. Biol. Psychiatry 81 (4), 366–377.