Involvement of the rodent prelimbic and medial orbitofrontal cortices in goal-directed action: A brief review

Ellen P. Woon[1,2], **Michelle K. Sequeira**[1,2], **Britton R. Barbee**[2,3], **Shannon L. Gourley**[1,2,3,*]

[1]Graduate Program in Neuroscience

[2]Yerkes National Primate Research Center, Departments of Pediatrics and Psychiatry and Behavioral Sciences, Center for Translational and Social Neuroscience

[3]Graduate Program in Molecular and Systems Pharmacology Emory University, Atlanta, GA

Abstract

Goal-directed action refers to selecting behaviors based on the expectation that they will be reinforced with desirable outcomes. It is typically conceptualized as opposing habit-based behaviors, which are instead supported by stimulus-response associations and insensitive to consequences. The prelimbic prefrontal cortex (PL) is positioned along the medial wall of the rodent prefrontal cortex. It is indispensable for action-outcome-driven (goal-directed) behavior, consolidating action-outcome relationships and linking contextual information with instrumental behavior. In this brief review, we will discuss the growing list of molecular factors involved in PL function. Ventral to the PL is the medial orbitofrontal cortex (mOFC). We will also summarize emerging evidence from rodents (complementing existing literature describing humans) that it too is involved in action-outcome conditioning. We describe experiments using procedures that quantify responding based on reward value, the likelihood of reinforcement, or effort requirements, touching also on experiments assessing food consumption more generally. We synthesize these findings with the argument that the mOFC is essential to goal-directed action when outcome value information is not immediately observable and must be recalled and inferred.

Keywords

Part 1: Introduction

Goal-directed behavior refers to selecting actions based on desired outcomes. In contrast, habits are stimulus-elicited and insensitive to goals. Both goal-directed actions and habitual behaviors are important for survival, but maladaptive habits occurring at the expense of goal-sensitive actions are characteristic of many neuropsychiatric diseases (Griffiths et al., 2014;

*Contact Yerkes National Primate Research Center, 954 Gatewood Rd. NE, Atlanta, GA 30329, shannon.l.gourley@emory.edu, 404-727-2482.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Everitt and Robbins, 2016; Fettes et al., 2017) and may also contribute to compulsions and perseverative-like behaviors (Gillan et al., 2016).

Goal-directed actions and habits are commonly dissociated in rodents and primates using two tasks: reinforcer (or "outcome") devaluation and action-outcome (or "response-outcome") contingency degradation. Reinforcer devaluation assesses the ability of subjects to modify behaviors based on the value of expected outcomes. Rodents are typically trained to respond for two food reinforcers, one of which is then devalued in a separate environment in one of two ways: conditioned taste aversion or satiety-specific devaluation. In conditioned taste aversion procedures, one of the reinforcers is devalued by pairing it with lithium chloride (LiCl), which induces temporary malaise and conditioned taste aversion (fig.1a). In satiety-specific devaluation procedures, rodents are allowed unlimited access to one of the reinforcers, decreasing its value by virtue of satiety (fig.1b). When returned to the conditioning chambers, inhibiting responding associated with the now-devalued outcome is interpreted as evidence that animals modify their response strategies in reaction to the now-lower value of the reinforcer. Meanwhile, a failure to inhibit responding is considered habitual behavior. One "real-world" example of outcome devaluation is food poisoning. Even though one might enjoy hamburgers, negative experiences with hamburgers (such as food poisoning) typically result in other menu choices in the future, indicative of sensitivity to outcome value.

Action-outcome contingency degradation assesses an individual's ability to form and update the *association* between an action and its outcome. Animals are commonly trained to respond on two apertures for food reinforcers. Then, the association between one response and food delivery is disrupted, such that responding no longer reliably predicts pellet delivery. Thus, the action-outcome relationship associated with one of the responses is "degraded." A zero contingency is used, referring to a condition in which there is an equal likelihood of an outcome occurring following a response or no response at any given time (Hammond, 1980). Some investigations, particularly in mice, have alternatively taken the approach of delivering reinforcers noncontingently (*e.g.*, Gourley et al., 2012; Barker et al., 2018). In both procedures, inhibiting the response associated with the degraded contingency is considered goal-directed, while failing to modify responding is considered habitual (fig.2). As in reinforcer devaluation procedures, action-outcome memory formation and retrieval can be dissociated using a brief probe test following contingency degradation. At this time, the rodent presumably retrieves newly updated action-outcome memory in order to inhibit the response associated with the degraded contingency. A "real-world" example of sensitivity to instrumental contingency degradation is captured in one's typical reaction to a faulty vending machine – if inserting money is not reliably reinforced with our desired snack, and the machine randomly releases snacks, we stop feeding the machine. Meanwhile, we might fail to modify our familiar behaviors if we are instead relying on reflexive habits to navigate our worlds.

Importantly, reinforcer devaluation and instrumental contingency degradation measure two processes essential to goal-directed action (selecting actions based on outcome value and linking actions with valued outcomes, respectively). Nevertheless, these tasks are often erroneously regarded as interchangeable, unfortunate since some distinct mechanistic factors

have been identified and will be discussed in Part 2. Also notable, the relationship between goal-directed action and habitual behavior can be considered both antagonistic, as in our examples above, but also cooperative (Balleine and O'Doherty, 2010). Recent computational modeling suggests that goal-directed and habit-based behaviors are hierarchically organized, allowing for the dominance of one strategy over another under appropriate circumstances (Dezfouli and Balleine, 2013). These models emphasize that goal-directed action is not degraded when habits form. Action-outcome associations are not over-written or forgotten, but rather, the stimulus-response association is promoted when a familiar behavior is repeatedly reinforced (see for further discussion, Barker et al., 2014) or when a new action-outcome association fails to be consolidated or otherwise integrated into new future response strategies.

Reinforcer devaluation, action-outcome contingency degradation, and other tasks have been used to reveal that drugs of abuse, stressors, and stress hormones cause biases towards habit-based behaviors at the expense of goal-directed actions (Schwabe and Wolf, 2013; Barker and Taylor, 2014; DePoy and Gourley, 2015; Everitt and Robbins, 2016; Knowlton and Patterson, 2018). Developmental exposure to social adversity (see Hinton et al., 2019) and stressors or stress hormones (see Barfield and Gourley, 2019 and references therein) also cause habit biases evident in adulthood. These converging phenomena, observed across rodent and primate species, reinforce the utility of model organisms in understanding action/habit behavior. We will discuss evidence that specific structures within the rodent medial prefrontal cortex (mPFC) are essential for goal-directed action. In the interest of focus, we avoid investigations in which explicit Pavlovian cues, rather than action-outcome associations, are used to motivate responding (*i.e.*, studies of stimulus-outcome associations). We also do not discuss brain structures essential for habit formation, for instance, the infralimbic subregion of the mPFC and dorsolateral (sensorimotor) striatum (Killcross and Coutureau 2003; Coutureau and Killcross 2003; Yin et al., 2004). We refer interested readers to other excellent reviews on these structures (*e.g.*, Barker et al., 2014; Amaya and Smith, 2018), including reviews in this Special Issue.

Part 2: The prelimbic mPFC (PL) encodes action-outcome associations that support goal-directed action

As in humans, the rodent mPFC is involved in numerous aspects of complex decision making, including but not limited to: outcome-related learning, consolidation of memories, and forming associations between contexts and responses (Euston et al., 2012; Gourley and Taylor, 2016). The rodent mPFC can be subdivided into multiple regions with specific functions, and >2 decades of research indicate that the PL subregion is necessary for learning about relationships between actions and their outcomes (fig.3). PL inactivation in both mice and rats interferes with the ability of rodents to learn action-outcome associations (Balleine and Dickinson, 1998; Killcross and Coutureau, 2003; Corbit and Balleine, 2003; Ostlund and Balleine, 2005; Tran-Tu-Yen et al., 2009; Coutureau et al., 2009; Dutech et al., 2011; Coutureau et al., 2012; Swanson et al., 2017; Shipman et al., 2018). In instrumental reversal tasks (meaning, reversal tasks in which rodents must modify learned response strategies, rather than stimulus-outcome associations), PL inactivation can also delay

response acquisition (de Bruin et al., 2000; but see Gourley et al., 2010; Dalton et al., 2016), consistent with the notion that the PL is necessary for flexibly directing actions towards valued outcomes by encoding action-outcome associations. While the degree of homology between rodent and primate prefrontal cortex has long been a topic of contention (Preuss 1995; Carlén 2017), it has been argued that the rodent PL is functionally homologous to the ventromedial prefrontal cortex in humans (Balleine and O'Doherty, 2010), a brain region critical for goal-directed action selection (de Wit et al., 2009; Reber et al., 2017).

Additional investigations revealed that action-outcome contingency degradation induces immediate-early gene expression in the rodent PL (Fitoussi et al., 2018), and instrumental conditioning triggers the phosphorylation of Extracellular signal-Regulated Kinase (ERK1/2), a marker of activity-related synaptic plasticity, in the PL (Hart and Balleine, 2016). Action-outcome conditioning is associated with changes in neuronal excitability and covariations in firing rate between neurons in the PL (Singh et al., 2019). These modifications carry forward into periods of sleep following testing. With the caveat that the investigators used a maze task to probe action-outcome conditioning (rather than reinforcer devaluation or instrumental contingency degradation), these findings suggest that action-outcome conditioning leads to persistent changes in cortical networks, potentially associated with "replaying" newly learned information.

The PL consolidates specific action-outcome associations, at least in part, via glutamatergic projections to the posterior dorsomedial striatum (DMS) (Hart et al., 2018a). While the majority of fibers are ipsilateral in nature, the minority of direct bilateral projections is indispensable for behavioral sensitivity to reward value (Hart et al., 2018b). Meanwhile, connections with the ventral striatum are apparently dispensable (Hart et al., 2018b). PL-mediodorsal thalamic (MD) connections (both PL-to-MD and MD-to-PL) are also necessary for behavioral sensitivity to reward value (Bradfield et al., 2013; Alcaraz et al., 2018). Notably, sensitivity to the *link* between actions and their outcomes also requires MD-to-PL projections, but not the corresponding PL-to-MD projections (Alcaraz et al., 2018).

Interestingly, direct interactions between the PL and basolateral amygdala (BLA) are apparently not necessary for goal-directed action, even though both structures are individually essential for action-outcome conditioning (Coutureau et al., 2009). A possible intermediary structure is the ventrolateral orbitofrontal cortex (VLO), which is interconnected with the BLA and PL (Zimmermann et al., 2017; Vertes 2004). In a recent study, rats were trained to respond on two levers to earn grain pellets, then they underwent three sessions when one response still resulted in grain pellets, but the other now resulted in sugar pellets (Parkes et al., 2018). The rats then underwent satiety-specific devaluation. VLO inactivation blocked sensitivity to reinforcer value when contingencies had changed, suggesting that the VLO is necessary for integrating new action-outcome associations into prospective response strategies. In agreement, three independent investigations in our own laboratory revealed that VLO inactivation occludes sensitivity to instrumental contingency degradation (Whyte et al., 2019; Zimmerman et al. 2017, 2018), and dendritic spine plasticity in the VLO appears necessary for action-outcome response updating in this task (Whyte et al., 2019). Experiments that reduced levels of neurotrophic or cell adhesion factors necessary for VLO function and simultaneously inactivated the BLA or PL suggest

that the VLO interacts with these two structures to update outcome expectations (Zimmermann et al., 2017; DePoy et al., 2019). Thus, the VLO could conceivably serve as an intermediary between the PL and BLA and/or a site of integration; future investigations could explicitly test this hypothesis.

Instrumental responding can be context-dependent, such that changes in context can decrease instrumental responding (Thrailkill and Bouton, 2005). Some studies examining how context affects instrumental responding use a procedure termed "ABA renewal," in which animals are first trained to respond for reinforcement in Context A, then undergo extinction in Context B, and then tested back in Context A to assess "renewal" of responding (Bouton 2019). PL inactivation attenuates ABA renewal and increases responding in the extinction context, suggesting that the PL is necessary for associating contexts with instrumental behaviors (Eddy et al., 2016). In another study, rats were trained to lever press for a sucrose reinforcer in Context A, then the PL was inactivated in either Context A or a novel Context B. PL inactivation attenuated responding in Context A only, suggesting that the PL is necessary for detecting contexts in which responding had been previously reinforced (Trask et al., 2017). In a separate experiment, rats were allowed to reacquire the response in Context A, then the PL was inactivated in a novel extinction context (Context C) and assessed for renewal in a novel Context D. This "ACD" renewal procedure allowed the authors to further solidify the role of the PL in context-dependent responding – if the PL is specifically involved in context-dependent renewal, PL inactivation should produce no effect when renewal is assessed outside of the acquisition context (Context A). Indeed, there was no difference between PL-inactivated rats and their control counterparts in the renewal test. Thus, the PL appears to link instrumental response strategies with the contexts in which they are optimal or appropriate.

Neurobiological factors in PL-dependent action

Dopaminergic lesions and inactivation of dopamine D1/D2 receptors in the PL appear to occlude goal-sensitive action selection in a contingency degradation but not reinforcer devaluation procedure (Naneix et al., 2009; Lex and Hauber, 2010). This pattern suggests that dopamine in the PL is necessary for learning about action-outcome contingencies, but not necessarily outcome values *per se*. In related experiments, adolescent rats were less able than adults to optimize responding in a food-reinforced (though not ethanol-reinforced) contingency degradation task (Naneix et al., 2012; Serlin and Torregrossa, 2015). Age-related improvements in action-consequence conditioning were associated with the maturation of dopaminergic systems in the PL (Naneix et al., 2012). Subsequent investigations revealed that during adolescence, repeated stimulation of dopaminergic systems in rats derailed the typical maturation of PL dopamine systems; as adults, these rats were unable to associate actions with their consequences in an instrumental contingency degradation procedure (Naneix et al., 2013), again suggesting that dopamine signaling in the PL is necessary for learning about action-outcome associations. Notably, repeated stimulation of dopamine systems via experimenter-administered cocaine (Hinton et al., 2014; DePoy et al., 2014, 2017) and self-administered cocaine (DePoy et al., 2016) during adolescence also causes failures in action-outcome conditioning later in life.

Site-selective viral-mediated gene transfer allows for the modification of specific proteins in localized brain regions and has revealed multiple molecular factors in PL-dependent action selection. For instance, chronic loss of *Gabra1*, encoding GABA$_A$α1, during postnatal development causes response failures in a contingency degradation task (Butkovich et al., 2015). Butkovich et al. (2015) speculated that deficiencies might be attributable to a loss of synapses and dendritic spines that occurs with prolonged *Gabra1* deficiency during early postnatal development (see for example, Heinen et al., 2003). The prediction follows that the plasticity and stability of the actin cytoskeleton – the structural lattice that supports dendritic spines – should impact organisms' abilities to associate actions based on their outcomes. Consistent with this notion, inhibiting the cytoskeletal regulatory factor, Rho-kinase, enhances action-outcome conditioning, blocking habitual responding for both food and cocaine (Swanson et al., 2017). In the same report, successful action-outcome conditioning in a contingency degradation procedure was associated with dendritic spine loss on deep-layer PL neurons that was transient and tightly coupled with experiences that required mice to form new action-outcome associations (Swanson et al., 2017). These patterns suggest that some degree of dendritic spine pruning in the PL optimizes action-outcome conditioning.

A series of recent studies focused on a protein termed "p110β," a class 1A catalytic subunit of PI3-kinase. These investigations were initially motivated by the discovery that p110β is elevated in mouse models of fragile X syndrome (*Fmr1* knockout mice) (Gross et al., 2010). Reducing p110β in the PFC, including PL, of *Fmr1* knockout mice restored behavioral flexibility in tasks requiring animals to learn and update action-outcome associations, including instrumental contingency degradation (Gross et al., 2015). Notably, genetic reduction of p110β also normalized dendritic spine densities otherwise elevated with *Fmr1* deficiency (Gross et al., 2015), consistent with our argument above that dendritic spine plasticity – including pruning – is necessary for optimal action-outcome conditioning. In separate experiments, the same viral vector strategies rescued decision-making abnormalities following local *Fmr1* silencing (Gross et al., 2015), and systemic administration of a p110β-inhibiting drug also improves goal-sensitive action selection (Gross et al., 2019). Taken together, these findings suggest that an optimal balance in PI3-kinase activity in the PL is necessary for modifying behaviors based on action-outcome contingencies, and they again point to the likely importance of healthy dendritic spine plasticity, in that dendritic spine excess is linked with poor action-outcome conditioning. Similar associations were recently verified in the adjacent VLO (Whyte et al., 2019), where chemogenetic inactivation of excitatory neurons blocked both action-outcome contingency updating and learning-related dendritic spine elimination.

Other investigations focused on the neurotrophin Brain-derived neurotrophic factor (BDNF). BDNF is linked to a number of neuropsychiatric diseases, including depression, anxiety, schizophrenia, and addiction (Autry and Monteggia, 2012), in which complex decision making is impaired. In rats, mPFC *Bdnf* increases during the initial acquisition of a food-reinforced instrumental response, and then decreases with proficiency (Rapanelli et al., 2010), suggesting that it is involved in the initial phases of action-outcome conditioning – initially learning that an action produces specific consequences. Supporting this notion, substitution of a methionine allele for valine at codon 66 of the *BDNF* gene, which decreases activity-dependent BDNF release, increases the likelihood that humans will rely

on habit-based strategies (rather than goal-directed strategies) in spatial navigation tasks (Banner et al., 2011). Meanwhile, systemic administration of a bioactive, high-affinity tyrosine/tropomyosin receptor kinase B (trkB) agonist, 7,8-dihydroxyflavone, enhances action-outcome conditioning, blocking habits induced by response over-training in mice (Zimmermann et al., 2017). Subsequent studies confirmed that trkB stimulation enhances the formation of long-term action-outcome memory (Pitts et al., 2019).

Given patterns described above, we previously hypothesized that BDNF in the PL would be essential to goal-directed action. Thus, it was unexpected when bilateral *Bdnf* silencing in the PL *facilitated* action-outcome responding in mice bred on a BALB/c background (Hinton et al., 2014) and had no obvious effects in mice bred on a C57BL/6 background (Gourley et al., 2012)[1]. In C57BL/6 mice, PL-specific *Bdnf* silencing did, however, sensitize mice to failures in action-outcome conditioning when coupled with glucocorticoid receptor (GR) inhibition (Gourley et al., 2012). BDNF and GR systems coordinate dendritic spine plasticity (*e.g.*, Arango-Lievano et al., 2015; for review, Barfield and Gourley, 2018). Thus, one possibility is that BDNF and GR interactions stabilize synaptic contacts or plasticity necessary for optimal PL function, such that *Bdnf* silencing in the PL allows for the dominance of other brain regions, such as the infralimbic mPFC, during reward-related decision-making tasks. This possibility is supported by evidence that PL-selective *Bdnf* knockdown facilitates extinction conditioning (an IL-dependent form of learning and memory) (Gourley et al., 2009), but direct evidence is, to the best of our knowledge, not yet published. For further discussion of the behavioral functions of BDNF in the PL, we refer the reader to Pitts et al., 2016.

While the functions of BDNF in the PL in the context of action-outcome conditioning remain somewhat opaque, the functions of a primary downstream partner are clearer: ERK1/2 is a site of convergence of multiple signaling factors, and phosphorylated ERK1/2 (p-ERK1/2) is considered a marker of activity-related synaptic plasticity. Given the importance of the PL in goal-directed action, Hart and Balleine (2016) hypothesized that ERK1/2 in the PL could be a key molecular mechanism. They trained rats to respond for food, while others received pellets noncontingently. Instrumental conditioning increased p-ERK1/2 in layers 5 and 6 of the posterior PL 5 minutes after training, and anterior PL layers 2 and 3 60 minutes after training. The researchers then found that inhibiting p-ERK1/2 blocked the ability of rats to distinguish between the devalued and the non-devalued outcomes, suggesting that p-ERK1/2 is necessary for PL function. Experiments using post-training infusions allowed the investigators to conclude that ERK1/2 in the PL is involved in the *consolidation* of action-outcome memory. This process appears to involve a prolonged wave of ERK1/2 phosphorylation throughout the cell layers of the PL in the minutes-to-hours following the acquisition of new outcome value information.

Part 3: Functions of the medial OFC (mOFC) in action selection

The mOFC is positioned ventral to the PL at the base of the mPFC (fig.3). In humans, neuroimaging studies reveal that it is activated when making preference judgments (Paulus

---

[1]These strains also respond differently to action-outcome conditioning procedures in general – see Zimmermann et al., 2016

and Frank, 2003) and when the value of an outcome informs goal-directed behavior (Arana et al., 2003, Plassmann et al., 2007). These functions have been similarly identified in non-human primates (Wallis and Miller, 2003). Despite an explosion in recent years in research on the OFC in rodents (Izquierdo 2017), most reports focus on lateral OFC subregions, neglecting the mOFC. We will discuss emerging evidence that, as in humans, the mOFC is a key brain structure in rodents coordinating actions and habits.

A recent investigation revealed that lesions and chemogenetic inactivation of the mOFC in rats induce failures in reinforcer devaluation tasks. The specific pattern of response failures suggests that the healthy mOFC retrieves memories regarding the value of outcomes in order to guide response selection *when outcomes are not immediately observable* (Bradfield et al., 2015). Conversely, chemogenetic stimulation of the mOFC *enhances* sensitivity to outcome devaluation (Gourley et al., 2016). The anterior, but not posterior, mOFC appears necessary for this function (Bradfield et al., 2018), which might account for instances in which mOFC inactivation did not affect sensitivity to reinforcer devaluation in earlier investigations (Gourley et al., 2010; Münster and Hauber, 2018).

How does the mOFC retrieve memories regarding outcome value? A likely anatomical partner is the BLA, which is bidirectionally connected with the mOFC across rodent and primate species (McDonald and Culberson, 1986; Kita and Kitai, 1990; Kringelbach and Rolls, 2004; Hoover and Vertes, 2011; Gourley et al., 2016; Ghashghaei and Barbas, 2002). In a recent report (Malvaez et al., 2019), rats were first trained to lever-press for sucrose in a modestly food-restricted state. Then, rats were given the sucrose noncontingently, or "for free," when they were either sated or food-restricted, thereby increasing the value of the sucrose in the hungry rats and prompting memory encoding of the new value of the sucrose. The next day, rats, again food-restricted, lever-pressed during a brief probe test conducted in extinction. Rats that had undergone lever-pressing for sucrose in the food-restricted state generated higher response rates, indicating that they could retrieve updated value information to increase responding. Inactivating mOFC-to-BLA connections attenuated lever-pressing activity, however, indicating that mOFC-to-BLA connections are necessary for retrieving value memory. Notably, the investigators also discovered that if their rats had access to the sucrose during the probe tests, mOFC-to-BLA connections were unnecessary, presumably because memory retrieval was unnecessary. Thus, the mOFC appears necessary for optimal goal-oriented responding when the value of outcomes changes, requiring memory formation and retrieval to guide optimal responding, rather than situations in which animals can optimize their responding based on information held in working memory.

Consistent with the notion that the mOFC is involved in retrieving memories necessary for optimally calculating the likely consequences of one's behaviors, mOFC damage causes suboptimal responding in situations of uncertainty. In instrumental reversal procedures (referring to tasks in which rodents must modify response strategies based on reinforcement likelihood, rather than Pavlovian associations), mOFC inactivation impedes performance, causing mice to continue responding even when a given behavior is not reinforced (Gourley et al., 2010). The use of sophisticated probabilistic reversal tasks revealed that mOFC inactivation causes rats to err early in the task, in a manner suggesting that they struggle to differentiate between behaviors that yield high or low probabilities of outcome (Dalton et al.,

2016). mOFC inactivation also causes rats to favor win-stay strategies in tasks that assess "risky" decision making – meaning, they favor behaviors that were previously reinforced, even at the expense of utilizing new, potentially more favorable response strategies (Stopper et al., 2014). Together, these findings are consistent with arguments that the mOFC facilitates goal-directed response shifting under circumstances that require adapting to uncertain conditions (Gourley et al., 2010), potentially via memory retrieval processes (Bradfield et al., 2015; see for review, Bradfield and Hart, 2019).

A handful of studies used progressive ratio schedules of reinforcement to understand mOFC function in the context of uncertainty. In these experiments, organisms are typically trained to perform an operant response under a rich reinforcement schedule (such as a fixed ratio 1 schedule), then the schedule changes such that each reinforcer requires a progressively increasing number of responses. For instance, the first pellet might be delivered after 1 response, but the next requires 5 responses, the next 9, etc. Several measures can be collected, but a common one is the break point ratio, referring to the highest number of responses the animal exerts for a single reinforcer. This procedure is quite established (Hodos 1961), so a rich literature exists, and it has multiple other advantages. It can be conducted such that the organism is required to make minimal movement, advantageous for certain procedural and data interpretation considerations (discussed by Swanson et al., 2019). Further, a progressive ratio task can be devoid of explicit Pavlovian stimuli, encouraging rodents to utilize action-outcome strategies. Finally, it can also be quite simplistic, requiring organisms to develop only a single operant behavior.

To summarize current findings, the mOFC *inhibits* break point ratios when rodents are initially familiarizing themselves with the task, such that mOFC inhibition *elevates* break points (Gourley et al., 2010; Münster and Hauber, 2018) and stimulation *reduces* break points (Gourley et al., 2016; Münster and Hauber, 2018). Progressive ratio training also induces immediate-early gene expression in the mOFC (Münster and Hauber, 2018). One interpretation is that the mOFC is important for adapting to the demands of the task – namely, that for each reinforced action, food availability decreases because the response demand increases. This interpretation is compatible with the evidence that the mOFC retrieves value memory (Bradfield et al., 2015). For example, when rodents are confronted with the progressive ratio schedule of reinforcement, mOFC inactivation could prevent the retrieval and then integration of known value information into the development of new response strategies. Without the retrieval of value information, mOFC-inactivated rodents might expend inappropriate effort relative to the value of the (largely unseen) outcome.

Another interpretation is that the mOFC contributes to the extinction of action-outcome associations as the reinforcer becomes less and less available. One caveat, however, is that "extinction" in this interpretation would only apply to within-session extinction, given that progressive ratio schedules of reinforcement do not necessarily cause between-sessions extinction. Rodents will readily respond for food on progressive ratio schedules across several sessions, their response rates stabilizing, not extinguishing (see examples in Gourley et al., 2016).

Notably, mOFC inactivation does not seem to impact sensitivity to instrumental contingency degradation. Specifically, rats with mOFC lesions inhibit responding when a familiar action-outcome contingency is degraded via noncontingent pellet delivery, just like control rats (Bradfield et al., 2015). Why might the mOFC be necessary for typical responding in a progressive ratio task, but not in contingency degradation? Progressive ratio tasks presumably require animals to continually retrieve value representations of the outcome, given that the actual delivery of the reinforcer is infrequent. In contingency degradation, pellets are regularly delivered (contingently and noncontingently), meaning, they remain readily observable. Thus, the animal does not need to retrieve value representations to guide responding. In short, the mOFC appears to support goal-directed responding when reward value or response requirements change, and particularly when outcomes are not immediately observable.

It is important to note that in rodents familiar with the progressive ratio task or extensively trained in other food-reinforced procedures, mOFC inactivation has the opposite effect, decreasing responding (Swanson et al., 2019; Gardner et al., 2018). Whether this outcome is attributable to disruption in value memory retrieval is unclear, and potentially instead explained by the notion that another function of the mPFC – including the mOFC – is to help keep organisms "on task," maintaining responding over long delay periods (discussed in Swanson et al., 2019). The mOFC may help to keep organisms "on task" via connections with the ventral striatum and parts of the hypothalamus that control behavioral activation and autonomic and homeostatic processes (Swanson et al., 2019). For instance, the mOFC modulates nucleus accumbens shell-elicited feeding, greatly potentiating food intake (Richard and Berridge, 2013). The mOFC also supports *consistent* food intake, such that inactivating it disrupts typical patterns of sucrose consumption, causing fragmented intake and insensitivity to contrast effects between low and high concentrations of sucrose (Parent et al., 2015). Thus, one could imagine that the mOFC helps to sustain instrumental responding for food during periods of uncertainty or low reinforcement probability.

BDNF: A mechanistic factor in mOFC-dependent action selection

One molecular candidate likely involved in the ability of the mOFC to sustain goal-sensitive action is BDNF. Using *Bdnf*+/− mutant mice and viral-mediated mOFC-selective *Bdnf* knockdown, we demonstrated that regional loss of BDNF decreases behavioral sensitivity to reinforcer value (Gourley et al., 2016). *Bdnf*+/− mice also fail to habituate to a progressive ratio schedule of reinforcement, expending excessive effort relative to wildtype littermates (Gourley et al., 2016). In other words, *Bdnf*+/− mice, similarly to rodents with mOFC inhibition, fail to calculate the optimal effort expenditure relative to the value of the outcome that they can acquire. Importantly, BDNF infusion into the mOFC fully normalizes responding, indicating that BDNF in the mOFC is *sufficient* to support value-based responding, this normalization occurring likely in part by normalizing ERK1/2 activation (Gourley et al., 2016).

BDNF is subject to anterograde and retrograde transport (Conner et al., 1997; Sobreviela et al., 1996), such that OFC-selective *Bdnf* knockdown deprives interconnected regions, including the BLA and DMS, of BDNF (Gourley et al., 2013; Zimmermann et al., 2017).

Thus, *where* BDNF binding is necessary for prospective value-based action selection remains unclear. Dorsal striatal BDNF is a poor predictor of responding in a progressive ratio task, while BDNF in the mOFC is a strong predictor (Gourley et al., 2016). For these and other reasons, we think it likely that local mOFC BDNF binding to its high-affinity trkB receptor is essential for value-based action selection strategies, but this possibility needs to be empirically tested.

Conclusions

Goal-directed action refers to selecting behaviors based on: *1)* the value of anticipated outcomes, and *2)* the causal link between actions and outcomes. The PL subregion of the mPFC is essential for both processes via action-outcome memory consolidation, though molecular mechanisms are still being defined. The ventrally-situated mOFC also appears necessary for goal-directed action, particularly when outcome information is not immediately available and must be recalled and inferred and response strategies must be updated. Relatively few investigations in rodents have focused on this structure, compared to other sub-regions of the mPFC or OFC. As such, our understanding of this brain region will inevitably continue to evolve and refine as we better comprehend how organisms coordinate goal-directed action.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Works Cited

Alcaraz F, Fresno V, Marchand AR, Kremer EJ, Coutureau E, Wolff M (2018) Thalamocortical and corticothalamic pathways differentially contribute to goal-directed behaviors in the rat. Elife 7:e32517.

Amaya KA, Smith KS (2018) Neurobiology of habit formation. Curr Opin Behav Sci 20:145–152.

Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC (2003) Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. J Neurosci 23:9632–9638. [PubMed: 14573543]

Arango-Lievano M, Lambert WM, Bath KG, Garabedian MJ, Chao MV, Jenneteau F (2015) Neurotrophic-priming of glucocorticoid receptor signaling is essential for neuronal plasticity to stress and antidepressant treatment. Proc Natl Acad Sci USA 112:15737–15742. [PubMed: 26630005]

Autry AE, Monteggia LM (2012) Brain-derived neurotrophic factor and neuropsychiatric disorders. Pharmacol Rev 64:238–258. [PubMed: 22407616]

Balleine BW, Dezfouli A (2013) Actions, action sequences and habits: Evidence that goal-directed and habitual action control are hierarchically organized. PLoS Comput Biol 9:e1003364.

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419. [PubMed: 9704982]

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35: 48–69. [PubMed: 19776734]

Banner H, Bhat V, Etchamendy N, Joober R, Bohbot VD (2011) The brain-derived neurotrophic factor Val66Met polymorphism is associated with reduced functional magnetic resonance imaging activity in the hippocampus and increased use of caudate nucleus-dependent strategies in a human virtual navigation task. Eur J Neurosci 33:968–977. [PubMed: 21255124]

Barfield ET, Gourley SL (2018) Prefrontal cortical trkB, glucocorticoids, and their interactions in stress and developmental contexts. Neurosci Biobehav Rev 95:535–558. [PubMed: 30477984]

Barfield ET, Gourley SL (2019) Glucocorticoid-sensitive ventral hippocampal-orbitofrontal cortical connections support goal-directed action – Curt Richter Award Paper 2019. Psychoneuroendocrinology, epub ahead of print.

Barker JM, Taylor JR (2014) Habitual alcohol seeking: modeling the transition from casual drinking to addiction. Neurosci Biobehav Rev 47:281–294. [PubMed: 25193245]

Barker JM, Taylor JR, Chandler LJ (2014) A unifying model of the role of the infralimbic cortex in extinction and habits. Learn Mem 21:441–448. [PubMed: 25128534]

Barker JM, Bryant KG, Chandler LJ (2018) Inactivation of ventral hippocampus projections promotes sensitivity to changes in contingency. Learn Mem 26:1–8. [PubMed: 30559114]

Bouton ME (2019) Extinction of instrumental (operant) learning: interference, varieties of context, and mechanisms of contextual control. Psychopharmacology 236:7–19. [PubMed: 30350221]

Bradfield LA, Hart G, Balleine BW (2013) The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. Front Syst Neurosci 7:51. [PubMed: 24130522]

Bradfield LA, Dezfouli A, van Holstein M, Chieng B, Balleine BW (2015) Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron 88:1268–1280. [PubMed: 26627312]

Bradfield LA, Hart G, Balleine BW (2018) Inferring action-dependent outcome representations depends on anterior but not posterior medial orbitofrontal cortex. Neurobiol Learn Mem 155:463–473. [PubMed: 30243849]

Bradfield LA, Hart G (2019) Medial and lateral orbitofrontal cortices represent unique components of cognitive maps of task space. Available on PsyArXiv, accessed June 7, 2019.

Butkovich LM, DePoy LM, Allen AG, Shapiro LP, Swanson AM, Gourley SL (2015) Adolescent-onset GABAAα1 silencing regulates reward-related decision making. Eur J Neurosci 42:2114–2121. [PubMed: 26096050]

Carlén M (2017) What constitutes the prefrontal cortex? Science 358:478–482. [PubMed: 29074767]

Conner JM, Lauterborn J, Yan Q, Gall CM, Varon S (1997) Distribution of brain-derived neurotrophic factor (BDNF) protein and mRNA in the normal adult rat CNS: Evidence for anterograde axonal transport. J Neurosci 17:2295–2313. [PubMed: 9065491]

Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. Behav Brain Res 146:145–157. [PubMed: 14643467]

Coutureau E, Killcross S (2003) Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. Behav Brain Res 146:167–174. [PubMed: 14643469]

Coutureau E, Marchand AR, Di Scala G (2009) Goal-directed responding is sensitive to lesions to the prelimbic cortex or basolateral nucleus of the amygdala but not their disconnection. Behav Neurosci 123:443–448. [PubMed: 19331467]

Coutureau E, Esclassan F, Di Scala G, Marchand AR (2012) The role of the rat medial prefrontal cortex in adapting to changes in instrumental contingency. PLoS One 7:e33302.

Dalton GL, Wang NY, Phillips AG, Floresco SB (2016) Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. J Neurosci 36:1996–2006. [PubMed: 26865622]

de Bruin JP, Feenstra MG, Broersen LM, van Leeuwen M, Arens C, de Vries S, Joosten RN (2000) Role of the prefrontal cortex of the rat in learning and decision making: effects of transient inactivation. Prog Brain Res 126:103–113. [PubMed: 11105643]

de Wit S, Corlett PR, Aitken MR, Dickinson A, Fletcher PC (2009) Differential Engagement of the Ventromedial Prefrontal Cortex by Goal-Directed and Habitual Behavior toward Food Pictures in Humans. J Neurosci 29:11330–11338. [PubMed: 19741139]

DePoy LM, Perszyk RE, Zimmermann KS, Koleske AJ, Gourley SL (2014) Adolescent cocaine exposure simplifies orbitofrontal cortical dendritic arbors. Front Pharmacol 5:228. [PubMed: 25452728]

DePoy LM, Gourley SL (2015) Synaptic cytoskeletal plasticity in the prefrontal cortex following psychostimulant exposure. Traffic 16:919–940. [PubMed: 25951902]

DePoy LM, Allen AG, Gourley SL (2016) Adolescent cocaine self-administration induces habit behavior in adulthood: sex differences and structural consequences. Transl Psychiatry 6:e875.

DePoy LM, Zimmermann KS, Marvar PJ, Gourley SL (2017) Induction and blockade of adolescent cocaine-induced habits. Biol Psychiatry 81:595–605. [PubMed: 27871669]

DePoy LM, Shapiro LP, Kietzman HW, Roman KM, Gourley SL (2019) β1-Integrins in the Developing Orbitofrontal Cortex Are Necessary for Expectancy Updating in Mice. J Neurosci 39: 6644–6655. [PubMed: 31253753]

Dutech A, Coutureau E, Marchand AR (2011) A reinforcement learning approach to instrumental contingency degradation in rats. J Physiol Paris 105:36–44. [PubMed: 21907801]

Eddy MC, Todd TP, Bouton ME, Green JT (2016) Medial prefrontal cortex involvement in the expression of extinction and ABA renewal of instrumental behavior for a food reinforcer. Neurobiol Learn Mem 128:33–39. [PubMed: 26723281]

Euston DR, Gruber AJ, McNaughton BL (2012) The role of medial prefrontal cortex in memory and decision making. Neuron 76:1057–1070. [PubMed: 23259943]

Everitt BJ, Robbins TW (2016) Drug addiction: updating actions to habits to compulsions ten years on. Annu Rev Psychol 67:23–50. [PubMed: 26253543]

Fettes P, Schulze L, Downar J (2017) Cortico-striatal-thalamic loop circuits of the orbitofrontal cortex: promising therapeutic targets in psychiatric illness. Front Syst Neurosci 11:25. [PubMed: 28496402]

Fitoussi A, Renault P, Le Moine C, Coutureau E, Cador M, Dellu-Hagedorn F (2018) Inter-individual differences in decision-making, flexible and goal-directed behaviors: Novel insights within the prefronto-striatal networks. Brain Struct Func 223:897–912.

Gardner MP, Conroy JC, Styer CV, Huynh T, Whitaker LR, Schoenbaum G (2018) Medial orbitofrontal inactivation does not affect economic choice. Elife 7:e38963.

Ghashghaei HT, Barbas H (2002) Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. Neuroscience 115:1261–1279. [PubMed: 12453496]

Gillan CM, Robbins TW, Sahakian BJ, van den Heuvel OA, van Wingen G (2016) The role of habit in compulsivity. Eur Neuropsychopharmacol 26:828–840. [PubMed: 26774661]

Gourley SL, Howell JL, Rios M, DiLeone RJ, Taylor JR (2009) Prelimbic cortex bdnf knock-down reduces instrumental responding in extinction. Learn Mem 16:756–760. [PubMed: 19926781]

Gourley SL, Lee AS, Howell JL, Pittenger C, Taylor JR (2010) Dissociable regulation of instrumental action within mouse prefrontal cortex. Eur J Neurosci 32:1726–1734. [PubMed: 21044173]

Gourley SL, Swanson AM, Jacobs AM, Howell JL, Mo M, DiLeone RJ, Koleske AJ, Taylor JR (2012) Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. Proc Natl Acad Sci USA 109:20714–20719. [PubMed: 23185000]

Gourley SL, Olevska A, Zimmerman KS, Ressler KJ, DiLeone RJ, Taylor JR (2013) The orbitofrontal cortex regulates outcome-based decision-making via the lateral striatum. Eur J Neurosci 38:2382–2388. [PubMed: 23651226]

Gourley SL, Zimmermann KS, Allen AG, Taylor JR (2016) The medial orbitofrontal cortex regulates sensitivity to outcome value. J Neurosci 36:4600–4613. [PubMed: 27098701]

Gourley SL, Taylor JR (2016) Going and stopping: Dichotomies in behavioral control by the prefrontal cortex. Nat Neurosci 19:656–664. [PubMed: 29162973]

Griffiths KR, Morris RW, Balleine BW (2014) Translational studies of goal-directed action as a framework for classifying deficits across psychiatric disorders. Front Syst Neurosci 8:101. [PubMed: 24904322]

Gross C, Nakamoto M, Yao X, Chan CB, Yim SY, Ye K, Warren ST, Bassell GJ. (2010) Excess phosphoinositide 3-kinase subunit synthesis and activity as a novel therapeutic target in fragile X syndrome. J Neurosci 30:10624–10638. [PubMed: 20702695]

Gross C, Raj N, Molinaro G, Allen AG, Whyte AJ, Gibson JR, Huber KM, Gourley SL, Bassell GJ (2015) Selective role of the catalytic PI3K subunit p110beta in impaired higher order cognition in fragile x syndrome. Cell Rep 11:681–688. [PubMed: 25921527]

Gross C, Banerjee A, Tiwari D, Longo F, White AR, Allen AG, Schroeder-Carter LM, Krzeski JC, Elsayed NA, Puckett R, Klann E, Rivero RA, Gourley SL, Bassell GJ. (2019) Isoform-selective phosphoinositide 3-kinase inhibition ameliorates a broad range of fragile X syndrome-associated deficits in a mouse model. Neuropsychopharmacology 44:324–333. [PubMed: 30061744]

Hammond LJ (1980) The effect of contingency upon the appetitive conditioning of free-operant behavior. J Exp Anal Behav 34:297–304 [PubMed: 16812191]

Hart G, Balleine BW (2016) Consolidation of goal-directed action depends on MAPK/ERK signaling in rodent prelimbic cortex. J Neurosci 36:11974–11986. [PubMed: 27881782]

Hart G, Bradfield LA, Balleine BW (2018a) Prefrontal corticostriatal disconnection blocks the acquisition of goal-directed action. J Neurosci 38:1311–1322. [PubMed: 29301872]

Hart G, Bradfield LA, Fok SY, Chieng B, Balleine BW (2018b) The bilateral prefronto-striatal pathway is necessary for learning new goal-directed actions. Curr Biol 28:2218–2229. [PubMed: 30056856]

Heinen K, Baker RE, Spijker S, Rosahl T, van Pelt J, Brussaard AB (2003) Impaired dendritic spine maturation in GABAA receptor alpha1 subunit knock out mice. Neuroscience 122:699–705. [PubMed: 14622913]

Hinton EA, Wheeler MG, Gourley SL (2014) Early-life cocaine interferes with BDNF-mediated plasticity. Learn Mem 21:253–257. [PubMed: 24737916]

Hinton EA, Li DC, Allen AG, Gourley SL (2019) Social isolation in adolescence disrupts cortical development goal-dependent decision making in adulthood, despite social reintegration. eNeuro, epub ahead of print.

Hodos W (1961) Progressive ratio as a measure of reward strength. Science 134:943–944. [PubMed: 13714876]

Hoover WB, Vertes RP (2007) Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. Brain Struct Funct 212:149–179. [PubMed: 17717690]

Izquierdo A (2017) Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. J Neurosci 37:10529–10540. [PubMed: 29093055]

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex. Cereb Cortex 13:400–408. [PubMed: 12631569]

Kita H, Kitai S (1990) Amygdaloid projections to the frontal cortex and the striatum in the rat. J Comp Neurol 298:40–49. [PubMed: 1698828]

Knowlton BJ, Patterson TK (2018) Habit formation and the striatum. Curr Top Behav Neurosci 37:275–295. [PubMed: 27677776]

Kringelbach ML, Rolls ET (2004) The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. Prog Neurobiol 72:341–372. [PubMed: 15157726]

Lex B, Hauber W (2010) The role of dopamine in the prelimbic cortex and the dorsomedial striatum in instrumental conditioning. Cereb Cortex 20:873–883. [PubMed: 19605519]

Malvaez M, Shieh C, Murphy MD, Breenfield Wassum KM (2019) Distinct cortical-amygdala projections drive reward value encoding and retrieval. Nat Neurosci 22:762–769. [PubMed: 30962632]

McDonald AJ, Culberson J (1986) Efferent projections of the basolateral amygdala in the opposum, didelphis virginiana. Brain Res Bull 17:335–350. [PubMed: 3768738]

Münster A, Hauber W (2018) Medial orbitofrontal cortex mediates effort-related responding in rats. Cereb Cortex 28:4379–4389. [PubMed: 29161356]

Naneix F, Marchand AR, Di Scala G, Pape J-R, Coutureau E (2009) A role for medial prefrontal dopaminergic innervation in instrumental conditioning. J Neurosci 29:6599–6606. [PubMed: 19458230]

Naneix F, Marchand AR, Di Scala G, Pape J-R, Coutureau E (2012) Parallel maturation of goal-directed behavior and dopaminergic systems during adolescence. J Neurosci 32:16223–16232. [PubMed: 23152606]

Naneix F, Marchand AR, Pichon A, Pape J-R, Coutureau E (2013) Adolescent stimulation of D2 receptors alters maturation of the dopamine-dependent goal-directed behavior. Neuropsychopharmacology 38:1566–1574. [PubMed: 23443719]

Ostlund SB, Balleine BW (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J Neurosci 25:7763–7770. [PubMed: 16120777]

Parent MA, Amarante LM, Liu B, Weikum D, Laubach M (2015) The medial prefrontal cortex is crucial for the maintanance of persistent licking and the expression of incentive contrast. Front Integr Neurosci 9:1–14. [PubMed: 25698945]

Parkes SL, Ravassard PM, Cerpa JC, Wolff M, Ferreira G, Coutureau E (2018) Insular and ventrolateral orbitofrontal cortices differentially contribute to goal-directed behavior in rodents. Cereb Cortex 28:2313–2325. [PubMed: 28541407]

Paulus MP, Frank LR (2003) Ventromedial prefrontal cortex activation is critical for preference judgments. Neuroreport 14:1311–1315. [PubMed: 12876463]

Pitts EG, Taylor JR, Gourley SL (2016) Prefrontal cortical BDNF: A regulatory key in cocaine- and food-reinforced behaviors. Neurobiol Dis 91:326–335. [PubMed: 26923993]

Pitts EG, Barfield ET, Woon EP, Gourley SL (2019) Action-outcome expectancies require orbitofrontal neurotrophin systems in naïve and cocaine-exposed mice. Neurotherapeutics, epub ahead of print.

Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday econonmic transactions. J Neurosci 27:9984–9988. [PubMed: 17855612]

Preuss TM (1995) Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. J Cogn Neurosci 7:1–24. [PubMed: 23961750]

Rapanelli M, Lew SE, Frick LR, Zanutto BS (2010) Plasticity in the rat prefrontal cortex: linking gene expression and an operant learning with a computational theory. PLoS One 5:e8656.

Reber J, Feinstein JS, O'Doherty JP, Liljeholm M, Adolphs R, Tranel D (2017) Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. Brain 6:1743–1756.

Richard JM, Berridge KC (2013) Prefrontal cortex modulates desire and dread generated by nucleus accumbens glutamate disruption. Biol Psychiatry 73:360–370. [PubMed: 22981656]

Schwabe L, Wolf OT (2013) Stress and multiple memory systems: from 'thinking' to 'doing.' Trends Cogn Sci 17:60–68. [PubMed: 23290054]

Serlin H, Torregrossa MM (2015) Adolescent rats are resistant to forming ethanol seeking habits. Dev Cogn Neurosci 16:183–190. [PubMed: 25575668]

Shipman ML, Trask S, Bouton ME, Green JT (2018) Inactivation of prelimbic and infralimbic cortex respectively affects minimally-trained and extensively-trained goal-driected actions. Neurobiol Learn Mem 155:164–172. [PubMed: 30053577]

Singh A, Peyrache A, Humphries MD (2019) Medial prefrontal cortex population activity is plastic irrespective of learning. J Neurosci 39:3470–3483. [PubMed: 30814311]

Sobreviela T, Pagcatipunan M, Kroin JS, Mufson EJ (1996) Retrograde transport of Brain-Derived Neurotrophic Factor (BDNF) following infusino in the neo- and limbic cortex in rat: Relationship to BDNF mRNA expressing neurons. J Comp Neurol 375:417–444. [PubMed: 8915840]

Stopper CM, Green EB, Floresco SB (2014) Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. Cereb Cortex 24:154–162. [PubMed: 23042736]

Swanson AM, DePoy LM, Gourley SL (2017) Inhibitng Rho kinase promotes goal-directed decision making and blocks habitual responding for cocaine. Nat Comm 8:1861.

Swanson K, Goldback HC, Laubach M (2019) The rat medial frontal cortex controls pace, but not breakpoint, in a progressive ratio licking task. Available on PsyArXiv Preprints, accessed May 3, 2019.

Thraikill EA, Bouton ME (2015) Contextual control of instrumental actions and habits. J Exp Psychol Anim Learn Cogn. 41:69–80. [PubMed: 25706547]

Tran-Tu-Yen DA, Marchand AR, Pape JR, Di Scala G, Coutureau E (2009) Transient role of the rat prelimbic cortex in goal-directed behaviour. Eur J Neurosci 30:464–471. [PubMed: 19614748]

Trask S, Shipman ML, Green JT, Bouton ME (2017) Inactivation of the prelimbic cortex attenuates context-dependent operant responding. J Neurosci 37:2317–2324. [PubMed: 28137970]

Vertes RP (2004) Differential projections of the infralimbic and prelimbic cortex in the rat. Synapse 51:32–58. [PubMed: 14579424]

Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. Euro J Neurosci 18:2069–2081.

Whyte AJ, Kietzman HW, Swanson AM, Butkovich LM, Barbee BR, Bassell GJ, Gross C, Gourley SL (2019) Reward-related expectations trigger dendritic spine plasticity in the mouse ventrolateral orbitofrontal cortex. J Neurosci 39:4595–4605. [PubMed: 30940719]

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of the dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189. [PubMed: 14750976]

Zimmermann KS, Hsu CC, Gourley SL (2016) Strain commonalities and differences in response-outcome decision making in mice. Neurobiol Learn Mem 131:101–108. [PubMed: 27003118]

Zimmermann KS, Yamin JA, Rainnie DG, Kessler KJ, Gourley SL (2017) Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by BDNF-trkB. Biol Psychiatry 81:366–377. [PubMed: 26786312]

Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, Gourley SL (2018) Memory retention involves the ventrolateral orbitofrontal cortex: Comparison with the basolateral amygdala. Neuropsychopharmacology 43:373–383. [PubMed: 28664926]

Significance Statement

Goal-directed action refers to selecting behaviors based on their likely outcomes. It requires structures along the medial wall of the prefrontal cortex in rodents, but mechanistic factors and the functions of specific subregions are still being defined. We will discuss molecular factors involved in the ability of the prelimbic prefrontal cortex to form action-outcome associations. Then, we will summarize evidence that the medial orbitofrontal cortex is also involved in action-outcome conditioning.
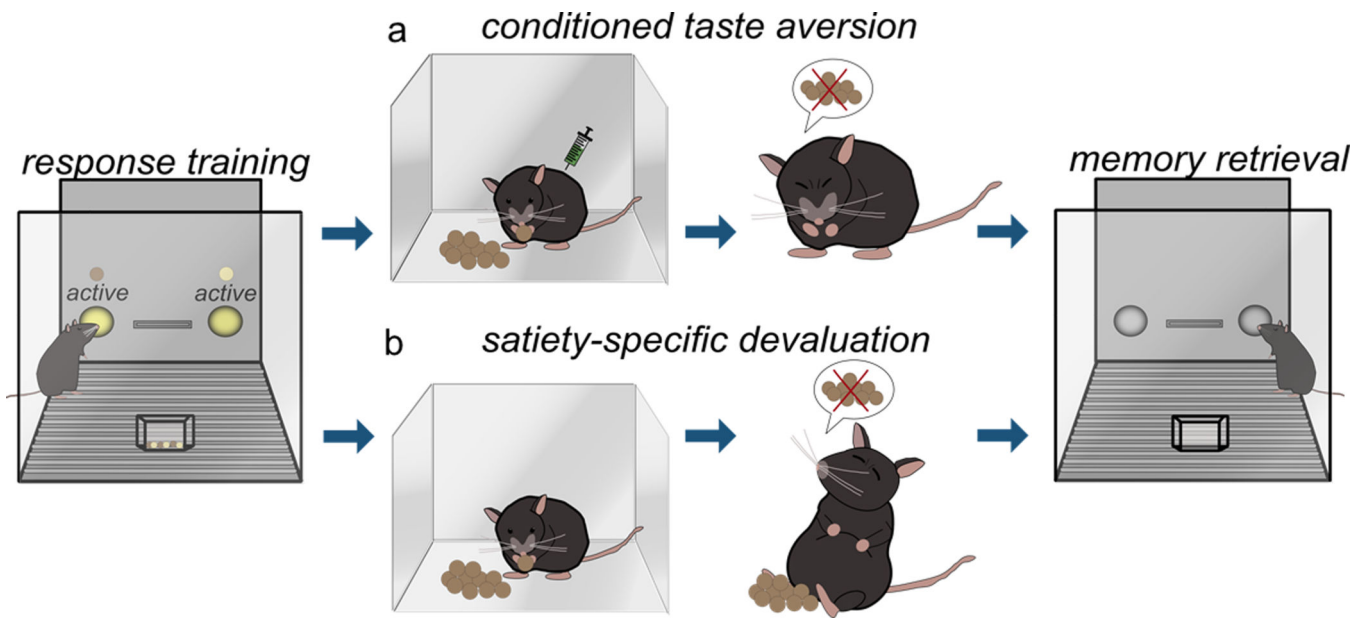
Figure 1. Schematic of (a) conditioned taste aversion and (b) satiety-specific prefeeding devaluation.
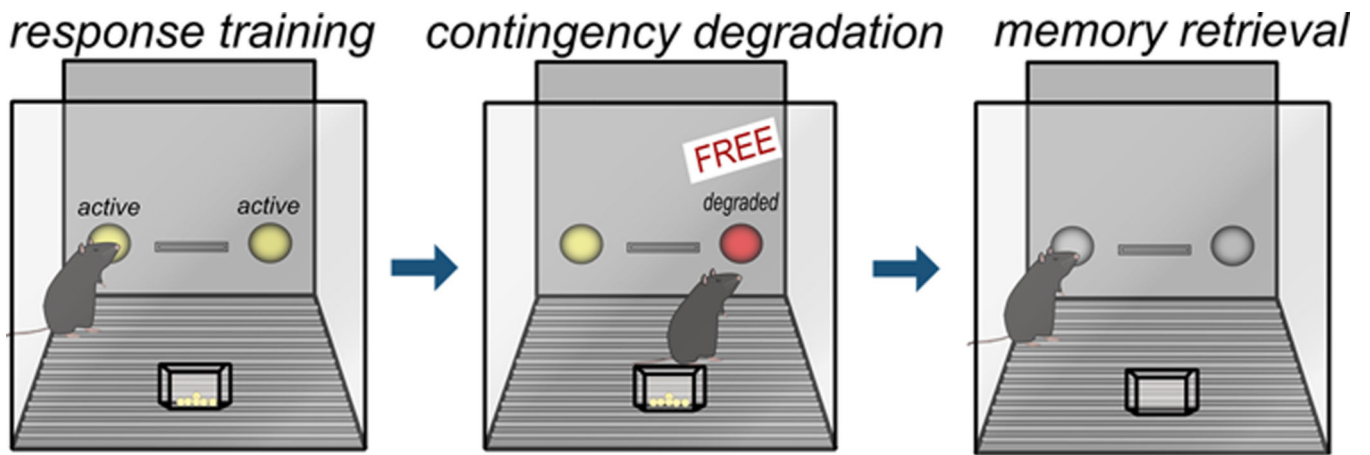See text in Part 1 for description of behavioral procedures.

Figure 2. Schematic of action-outcome contingency degradation.
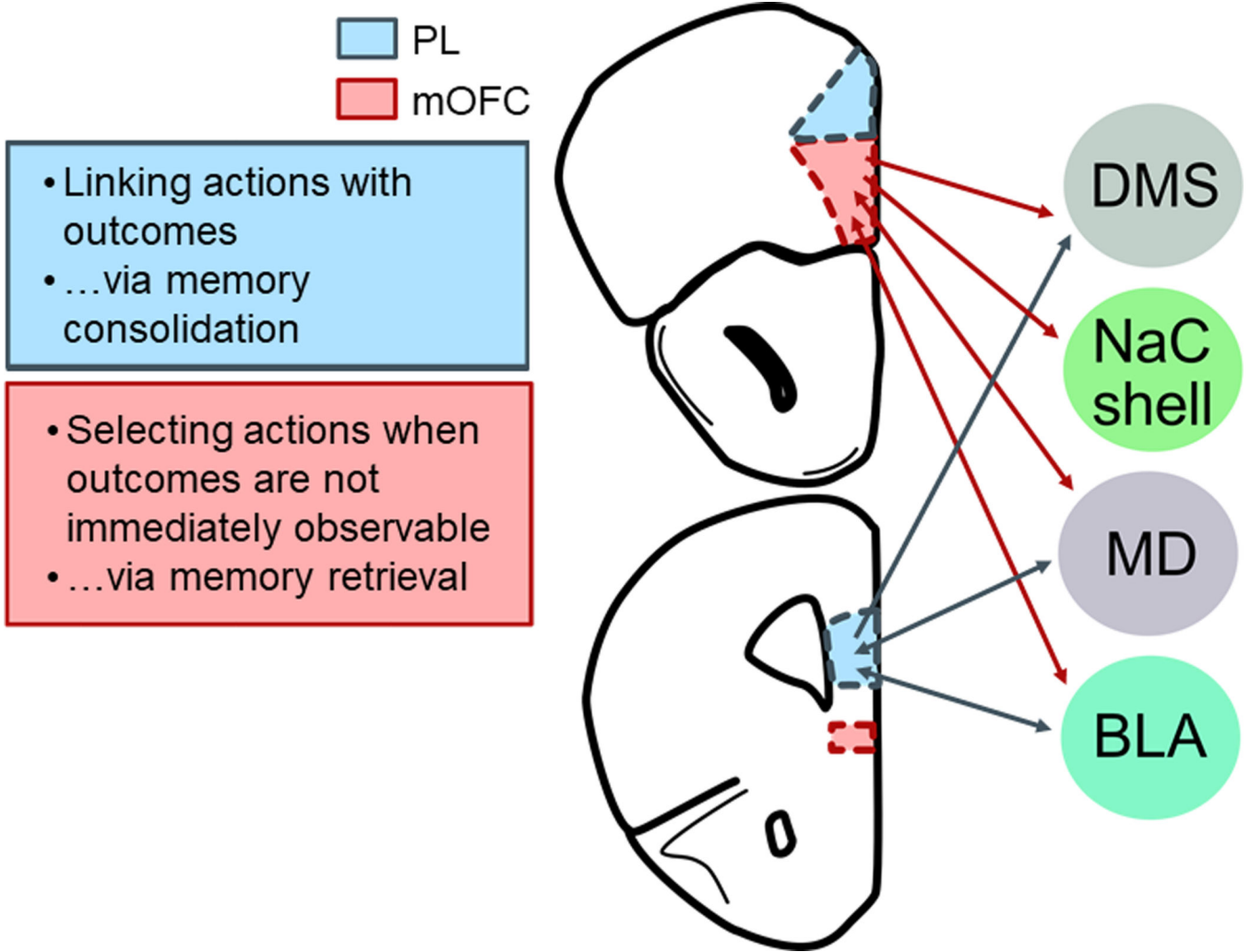See text in Part 1 for description of behavioral procedures.

Figure 3. Differential contributions of the mOFC (top) and PL (bottom) to goal-directed action selection.

Connections discussed in this mini-review are highlighted.