Interacting roles of lateral and medial Orbitofrontal cortex in decision-making and learning : A system-level computational model.

Bhargav Teja Nallapu[1,2]❍, Frédéric Alexandre[1,2]

1 INRIA Bordeaux Sud-Ouest, Bordeaux, France
2 Institute of Neurodegenerative Diseases (IMN), Bordeaux, France

¤200 Avenue de la Vieille Tour, 33405 Talence, France
* frederic.alexandre@inria.fr (FA)
❍BTN is the graduate student who implemented the model, ran the simulations, and wrote the article. FA corrected the article. BTN and FA together conceived the idea and equally contributed to the conception and discussion.

Competing interests:

FA acts as a handling editor for PLOS ONE's call for papers on *Neuroscience of Reward and Decision Making.*

Abstract

In the context of flexible and adaptive animal behavior, the orbitofrontal cortex (OFC) is found to be one of the crucial regions in the prefrontal cortex (PFC) influencing the downstream processes of decision-making and learning in the sub-cortical regions. Although OFC has been implicated to be important in a variety of related behavioral processes, the exact mechanisms are unclear, through which the OFC encodes or processes information related to decision-making and learning. Here, we propose a systems-level view of the OFC, positioning it at the nexus of sub-cortical systems and other prefrontal regions. Particularly we focus on one of the most recent implications of neuroscientific evidences regarding the OFC - possible functional dissociation between two of its sub-regions : lateral and medial. We present a system-level computational model of decision-making and learning involving the two sub-regions taking into account their individual roles as commonly implicated in neuroscientific studies. We emphasize on the role of the interactions between the sub-regions within the OFC as well as the role of other sub-cortical structures which form a network with them. We leverage well-known computational architecture of thalamo-cortical basal ganglia loops, accounting for recent experimental findings on monkeys with lateral and medial OFC lesions, performing a 3-arm bandit task. First we replicate the seemingly dissociate effects of lesions to lateral and medial OFC during decision-making as a function of value-difference of the presented options. Further we demonstrate and argue that such an effect is not necessarily due to the dissociate roles of both the subregions, but rather a result of complex temporal dynamics between the interacting networks in which they are involved.

Author summary

We first highlight the role of the Orbitofrontal Cortex (OFC) in value-based decision making and goal-directed behavior in primates. We establish the position of OFC at the intersection of cortical mechanisms and *thalamo-basal ganglial* circuits. In order to understand possible mechanisms through which the OFC exerts emotional control over behavior, among several other possibilities, we consider the case of dissociate roles of two of its topographical subregions - *lateral* and *medial* parts of OFC. We gather predominant roles of each of these sub-regions as suggested by numerous experimental evidences in the form of a system-level computational model that is based on existing neuronal architectures. We argue that besides possible dissociation, there could be possible interaction of these sub-regions within themselves and through other sub-cortical structures, in distinct mechanisms of choice and learning. The computational framework described accounts for experimental data and can be extended to more comprehensive detail of representations required to understand the processes of decision-making, learning and the role of OFC and subsequently the regions of prefrontal cortex in general.

Introduction

Psychological and economic accounts of human and animal decision making have placed a great emphasis in the concept of value [1–3]. Value-based decision making is a process of evaluation of alternatives in terms of subjective preference to their potential consequences and their relevance to internal motivations. Thus often in the context of decision-making of an animal, *value* remains a subjective, comparable quantity, that spans across multiple neural pathways [4]. The Orbitofrontal Cortex

(OFC) is one of the prominent regions of the Prefrontal Cortex (PFC) that is believed to play a crucial role in value-based decision making [5,6] and learning [7] where *value* changes over time because of changes in internal states or external contingencies. However, specific functional role of the OFC in value-based decision making is not clear. Moreover, to some degree or the other, the OFC has been implicated in almost all of the fundamental processes formally described to be involved in value-based decision making [8] - Representation, Valuation, Action selection, Outcome evaluation and Learning.

In the context of 'Representation' of a decision scenario, the OFC has been proposed to represent a cognitive map of task space [9]. Further in the context of 'Valuation', the role of OFC has been implied in encoding the *value* of the offered and chosen goods [10,11]. It might appear that there are other valuation systems in the brain [12–14], but what makes OFC unique in valuation is that it encodes the value irrespective of the visuo-spatial and motor aspects. For instance, the value representation in the OFC has been argued, not only to guide choices consistent with *Transitivity* [15], but also to represent largely varying subjective values in an adaptive manner [16,17]. These representations were proposed to be based on a *common currency* [1,6,18] that guide the comparison for the decision between different objects that are otherwise incomparable. Alternative to the theory of *common currency*, it was proposed that what the OFC facilitates is the process of *common scaling* [19–21] which is qualitatively distinct from that of converting different rewards into a *common currency*. Instead, *common scaling* corresponds to retaining the individual value of each reward, and converting them to a different scale that makes them comparable. Evidently, with complex possibilities in the process of valuation, arise different possibilities of action selection processes.

Furthermore, exact representations and mechanisms through which the OFC contributes to behavior are still up to active debate [22]. Moreover, several early implications of the OFC in the paradigms related to reversal learning [23], response inhibition [24,25], flexible stimulus-outcome associations [26,27] have been overturned using the same experimental techniques [7] or even more accurate ones [28], modified task structures [29,30] or pointing out the fact that the findings from other related brain regions explain certain implications better [31–35].

Dissociate roles of lateral and medial OFC

The evident underlying complexity of studying the role of the OFC in value-based decision making and learning, and goal-directed behavior is underlined by the large heterogeneity of the region, unlike the rest of PFC which is homogeneously granular. The heterogeneity is multi-fold : different groups of neurons that encode different aspects of choice process in a single task context [11], cyto-architecturally different areas (granular and agranular) and their remarkably distinct connectivity pathways through different brain structures [36–38]. The possibility of functionally dissociate roles of topologically different sub-regions of the OFC has been of wide interest recently. While there are other sub-divisions of the OFC reported to be playing functionally distinct roles in behavior [39–41], the distinction that is most extensively reported to imply strikingly different functional roles is the one between lateral and medial parts of OFC [28,42,43]. In the scope of this article, as referred in most of the related experimental works, ventromedial prefrontal cortex (vmPFC) is also considered under the purview of medial OFC [44,45]. It has also been observed that lateral and medial OFC have clear divergent connections to different networks [46]. Both in monkeys and humans lateral OFC is reported to receive extensive projections from diverse sensory modalities through the somatosensory and insular cortices, and also heavy projections from amygdala [47–49]. Whereas, the medial OFC has strong

projections from hippocampus, hypothalamus, ventral striatum (VS), relatively less projections from amygdala, and is strongly connected with the cingulate cortical areas [50, 51].

In this current work, we present a recurrent neural network model of decision making and learning involving the OFC. The OFC, together with some nuclei of the basal ganglia (BG)(especially VS) and the thalamus (Th), forms a closed loop whose dynamics leads to action selection by competition resolution. This loop is a part of several similar generic loops that are formed between different cortical regions and different nuclei of the BG. A generic loop will be referred hereafter as a CBG loop. Notably, we separate the part involving the OFC into two CBG loops involving lateral and medial OFC, accounting for the individual experimental implications of the lateral and medial sub-regions. The input to the lateral OFC loop is provided with the information that represents *exteroception* - the value information arising from external factors like visual cues; the input to the medial OFC loop is predominantly the *interoception* - value information more with respect to internal motivational processes like satiety levels and internal needs [42][1]. Across both these loops, we use the idea of the Current Subjective Value (CSV) in the model as an input to lOFC and mOFC. Besides the activation in the loops that represents the visual salience of the cues, CSV represents the value based on which the sub-regions of OFC contribute to the decision-making process. Such a subjective value is known to arise from a comprehensive relation of lateral OFC with basolateral amygdala and the ventral striatum in an ongoing task context [52–55] (see Materials, CSV).

We provide a plausible explanation for one of the prominent experimental observations regarding the dissociate roles of lateral and medial OFC, studied in individual lesions in monkeys [43]. We represent this proposed dissociation in terms of the representation and processing of the task information. Furthermore we argue that, in the context of *learning* vs *choice* notion of lateral and medial OFC [56], more than a clear dissociation, it is the temporal interaction of both the sub-regions that highlights their roles at different stages of a decision task.

Results

We first describe the performance of an existing model of decision-making and learning on a 2-arm bandit task with probabilistic reward. We take the advantage of generic nature of the task to highlight the fundamental dynamics of the model. We then show that the model presented here with the distinct description of lateral and medial OFC replicates the results of basic model, robustly and in more realistic timescales. We further present complementary findings of separate lesions (simulated) of the lateral and medial OFC components in the model. We discuss the effect of these findings on the performance in different task contingencies, replicating a neuroscientific evidence found in monkeys with lesions to different subregions of OFC.

2-Arm Bandit Task and Probabilistic Reward Learning

Multi-arm bandit task is a classic reinforcement learning problem that has been used in the study of decision-making in experimental [7, 43, 57] and computational neuroscience [58–60]. Typically, in an N-arm bandit task, there are N possible cues (bandits) each carrying a different probability of reward and requiring a particular action to do, in order to select the cue. Fig. 1A shows an example trial of a 2-arm bandit task that has been used to study the computational models of probabilistic reward-based learning involving the basal ganglia (BG) [60, 61]. In this case, cue is one

---

[1]What was referred to as OFC in this work actually recorded from the lateral areas of the OFC

of the four possible shapes. The reinforcement in the model during the task is driven   106
by the probabilistic reward offered at the end of each trial, with a different probability   107
for each cue. It has been shown that monkeys learn to perform the task [57], learning   108
the reward contingencies over time and choosing always the best rewarding option   109
after learning.   110

The basic model (referred hereafter as OFC model) is a set of inter-connected CBG   111
loops and an associative network (ASC), each network processing different information   112
and contributing for a decision within the network (Fig 1C). In each trial, the CBG_{cue} 113
labeled 'limbic' takes as the input, the activation for the shapes that are presented in   114
the trial. This activation represents a constant visual salience component, that in the   115
simplest case, is same for every stimulus (shape). Similarly the other CBG position   116
loops (CBG_{pos}) takes as the input, the activation of the positions where the shapes are   117
presented. Since the positions are chosen randomly and carry no significance in   118
obtaining reward, there is no value-learning in this CBG_{pos} loop. Hence the activation   119
of a position represents just the presence of a cue at that position. Finally, the ASC   120
network takes as the input, the combined information of binding specific shape to a   121
specific position. The ASC network represent the associative loop through lateral PFC   122
and the dorsomedial striatum (DMS) which is believed to represent a multi-modal   123
information of stimulus-vs-position mapping [62]. This is implemented in the form a 2   124
dimensional mapping for each shape against all possible position and each position   125
against all possible shapes (Fig 1B, blue squares). The networks are inter-connected in   126
such a way that while each of the CBG loops independently processes the information   127
that it is activated with, it also affects the activities in the other through the ASC   128
network. The network architecture within each CBG loop that guarantees the   129
resolution of competition between the options is based on classical BG pathways that   130
have been previously explained with computational accounts [59, 60, 63].   131

In each trial of the task, the model is presented with pseudo-randomized pairwise   132
presentation of the four possible shapes in any two of the four possible positions ('Cue   133
presentation' phase in Fig 1A and first 6 panels in Fig 1D). Although the performance   134
of the model is assessed in terms of the shape it chooses for optimal reward   135
probability, the choice is confirmed only if the corresponding position of the shape is   136
chosen as the 'motor' decision (Fig 1A, black + sign under 'Decision' phase, dashed   137
lines under 'CBG' in Fig 1D). Thus, after the 'Decision' phase of the trial, the shape   138
at the chosen position is considered as the choice of 'cue' and the reward is delivered   139
according to the predetermined probability associated to that cue.   140

Fig 1. 2-arm bandit task. A. Sample trial from a 2-arm bandit task. Out of four possible shapes (cues) are shown at two random positions (of the four cardinal positions). The position that is chosen implies the choice of the shape made. **B.** Basic model involving two CBG loops and an associative loop (ASC), one CBG loop leading to a choice between the two cues and the other between the two positions. The final output that is considered from the model within a trial is that of the decision of CBG position, the cue shown at the *chosen position* is considered as the *chosen cue*. Note the CBG Cue is labelled *limbic*, as it will be developed more into components representing sub-regions of the OFC. Blue arrow represents the connection that can be modified by learning. **C.** The proposed change in the original model which will be described in detail in the following section. **D.** Activation of each cue that is shown in a choice, its position and the combined information. Also, the evolution of activity in a CBG loop - solid lines for cue, dashed lines for positions.

The performance of the model is demonstrated under two conditions : EASY and   141
DIFFICULT. EASY is the condition where the reward probabilities related to each   142

shape are fairly separated and DIFFICULT is the condition where the reward   143
probabilities are either lower or closer, thus making the reinforcement difficult (Fig   144
2A). The effect of learning in the model after each trial can be observed in terms of   145
the decision times over the duration of the task. A decrease in decision times of both   146
cue and position is observed (Fig 2B, left). A running average over the choice of 10   147
trials is considered for the performance over 120 trials. The performance of the model   148
under the EASY condition replicates animals' behavior [57] (Fig 2D, blue). In the   149
DIFFICULT condition (Fig. 2A, right), the reward probabilities of both the shapes   150
are lower or closer. This should result in lower rate of reinforcement and thereby make   151
it difficult to make a correct choice. Animals however, with considerable amount of   152
training, were shown to identify the option with more chance of reward and thus make   153
correct choices [7,43]. We tested the same model as in the previous EASY case (Fig   154
2A, left), but the model couldn't learn the appropriate contingencies well. The   155
Decision Times (DTs) were longer compared to the previous case (Fig 2C) and the   156
overall performance was sub-optimal (Fig 2D, red).   157

Fig 2. N-arm bandit task. The task described in Fig 1**A** has two possible cues
(shapes) each with a predetermined probability of reward upon choice. **A.** Left and
right figures show two different reward probability schemes, in EASY and DIFFICULT
task scenarios. Each color represents a particular shape, as in the legend of right
sub-figure. **B, C.** Decision Times (DTs) in the model after cue presentation. 120 trials
are divided into 6 bins, with 20 trials per bin, and the DTs of both cue decisions and
the position decisions are averaged per bin. **B** shows the DTs in EASY condition of
the task and **C** in DIFFICULT condition. **D.** Performance of the model. Running
average of number of correct choices across 10 trials, averaged over 10 sessions.
Correct choice means the shape that rewards the most according to the predetermined
probabilities. Lighter color filling represents the standard deviation.

Precise Value Comparison   158

We then extend the 'limbic' CBG loop to individually describe two separate CBG   159
loops - one representing the lateral OFC and the other representing the medial OFC.   160
Here after this version of the model will be referred as *lmOFC* model. The CBG loop   161
involving lateral OFC builds on the top of the single limbic loop from the basic model   162
(described in Fig 1B). In addition to the activation (I_{ext}) to the network, a Current   163
Subjective Value (CSV) for each shape is also added to the input. CSV represents the   164
subjective value of a shape at any moment taking the externally learned reward   165
contingencies and internal bodily desire for the reward that the shape leads to (see   166
Materials, CSV). Another key aspect of lOFC is that it properly assigns the obtained   167
reward to the appropriate choice made in that trial (referred as *credit assignment*).   168
There has been evidence that neurons in lateral OFC are particularly active after the   169
reward delivery in a choice [42] and also the fact that medium spiny neurons neurons   170
which are extensively involved in decision-making are consistently active for a while   171
after reward delivery [55]. These evidences support the possibility that cortico-striatal   172
synaptic plasticity is a plausible phenomenon in the context of obtaining reward.   173
Similar arguments were made by other experimental findings [7].   174

The CBG loop with medial OFC receives input from the CSV layer. Medial OFC   175
has a separate value comparison mechanism implemented as a simple 'recurrent   176
excitation lateral inhibition' model, activated by the CSVs received. It was shown that   177
the activity in medial OFC correlated to the value difference between the options [64].   178
Supporting the view that the relative difference of the presented options is represented   179
in vmPFC, multiple value comparison mechanisms have been proposed. This value   180

difference signal further allows vmPFC to perform a value comparison to facilitate the $\quad$ 181
choice through principles of recurrent excitation and lateral inhibition [21,65–67]. The $\quad$ 182
output activities of mOFC are fed into its CBG loop. It has been shown that one of $\quad$ 183
the general function of populations in the PFC is to maintain history of decision $\quad$ 184
events such as previous action, previous reward etc [68]. Accordingly, we implemented $\quad$ 185
a simple history of rewards in mOFC, without cue-specific information. As the lOFC $\quad$ 186
maintains the current choice until the reward delivery and later [42], possibly a history $\quad$ 187
of choices is maintained in lOFC. It was shown that lesions to lOFC affect the $\quad$ 188
appropriate consolidation of the reward history with the choice history [7]. Hence, for $\quad$ 189
the sake of simplicity, both the histories in lOFC and mOFC are combined within the $\quad$ 190
lOFC to provide a combined choice-reward history up to one previous choice and $\quad$ 191
reward, and it is fed into the CBG loop of lOFC along with the activation of the cue. $\quad$ 192
In addition, a synaptic connection is added to the ASC layer outside the limbic $\quad$ 193
network, from each cue population in lOFC to all the possible position populations in $\quad$ 194
the 2-D mapping of ASC network. However the learning in these connections would be $\quad$ 195
less influential on the decision, compared to that in the lOFC network, since the $\quad$ 196
learning happens with respect to all four possible positions corresponding to the cue in $\quad$ 197
the 2-D mapping of ASC (Fig 1B, input to ASC). Although it has been shown that $\quad$ 198
mOFC / vmPFC encodes action-outcome associations in several task settings [69–72], $\quad$ 199
the design of the n-arm bandit task setting does not allow much of learning $\quad$ 200
action-values. The reason is that the task randomizes the positions where the cues are $\quad$ 201
present and hence the action required to chose a cue. $\quad$ 202

We then tested the lmOFC model on the DIFFICULT condition as in the previous $\quad$ 203
task. The model performed considerably well compared to the previous OFC model, $\quad$ 204
with much faster DTs. Both the models have an estimated value difference for the $\quad$ 205
ongoing task, across all the trials. Interestingly, the precise value comparison in $\quad$ 206
mOFC estimates the value difference across all the trials better than that estimated by $\quad$ 207
the OFC model under DIFFICULT condition (Fig 3D). $\quad$ 208

Fig 3. lmOFC Model : CBG loops with lateral and medial OFC. A. lmOFC
Model. Changes in the 'limbic' CBG loop, compared to the basic model. Lateral OFC
(lOFC) has access to cue identity (shape), hence drives learning the connections to its
CBG loop. lOFC also activates the Current Subjective Value (CSV) for each of the
presented cues from elsewhere. Medial OFC (mOFC) has a value comparison
mechanism to compare the CSVs of the presented cues it receives. mOFC further
drives its CBG loop with the ongoing value comparison outputs. Both lOFC and
mOFC also maintain general history of chosen cue-reward association and reward
respectively. This input is also used in the activation to their respective loops. **B.** The
average DTs of decisions choosing cue and position, across 120 trials binned every 20
trials. **C.** The performance of the lmOFC model (green) in comparison with the
performance of the basic model in DIFFICULT condition (red). **D.** Average value
difference of the presented options estimated in lmOFC model (green) and basic model
under DIFFICULT condition (red).

Proximity of Values and Decision Making $\quad$ 209

We tested the lmOFC model on a 3-arm bandit task (Fig 4). Each of the three cues $\quad$ 210
that are shown in every trial has a reward probability upon its choice. As shown in $\quad$ 211
Fig 4, V1, V2 and V3 are the reward probabilities associated to the cues *plus*, *delta* $\quad$ 212
and *star* respectively in a given experimental session. The task is carried out under $\quad$ 213
three different reward schedules (Fig 5A-C). In all the sessions, V1 and V3 are fixed to $\quad$ 214
be .7 and 0.05. V2 value is changed across three types of sessions : V2_HIGH, $\quad$ 215

V2_MID and V2_LOW where V2 is set to 0.6, 0.3 and 0.1. Similar task schedule was ₂₁₆
used on animals to test the effects of lesions of lateral and medial OFC separately [43]. ₂₁₇

Fig 4. 3-arm bandit task A sample trial from the 3-arm bandit task. Three
possible shapes (cues) are shown in three random positions (of the four cardinal
positions). The position that is chosen implies the choice of the shape made. Upon
selection of a shape, a reward is delivered with a probability (p), which is different for
each of the shapes (V1, V2 and V3).

The lmOFC model is used without any changes. At the time of presentation in ₂₁₈
every trial, 3 cues are activated simultaneously (along with their positions in the ₂₁₉
CBG$_{pos}$ loop and in the ASC network). In terms of the model parameters, just the ₂₂₀
input activation which represents the cue salience had to be increased as compared to ₂₂₁
when the choice was between 2 options (as in the previous tasks). In this task, a ₂₂₂
correct choice or a good choice is a V1 choice. The model reached optimal ₂₂₃
performance (more than 80% V1 choices) in less than 150 trials in each session, in all ₂₂₄
three reward schedules (V2_HIGH, V2_MID and V2_LOW). This is referred as the ₂₂₅
'Control' condition, green in Fig 4E-J. ₂₂₆

Fig 5. Effects of lateral and medial OFC lesions in the model. A-C. 3
conditions of the task (each column). Of the reward probabilities V1, V2 and V3
described in Fig 4, V1 and V3 are fixed in all 3 task conditions (A-C, red and green
respectively). The 3 conditions depend on the value V2 (A-C, blue) : V2_HIGH,
V2_MID and V2_LOW. **D-F.** Lesion of mOFC under each task condition. The
average performance in each condition with a lesion to mOFC (blue) is compared to
the control performance (green). **G-I.** Lesion of lOFC under each task condition. The
average performance in each condition with a lesion to lOFC (pink) is compared to the
control performance (green).

Furthermore we simulated lesions of lateral and medial OFC in the model. Since ₂₂₇
the model generates a decision through at least one of the 'limbic' CBG loops and the ₂₂₈
other ASC and CBG$_{pos}$ loop, even in case of a lesion to lOFC or mOFC, a valid ₂₂₉
decision should be made. We first describe the changes in the model with respect to ₂₃₀
each of the lesions and the corresponding results. ₂₃₁

Medial OFC Lesion ₂₃₂

A lesion of mOFC to the lmOFC model shown in Fig 3A makes the model slightly ₂₃₃
similar to the basic OFC model described in Fig 1B. The credit assignment still works ₂₃₄
by lOFC during the period after reward (Fig 1D, CBG, after 'Reward' phase), because ₂₃₅
the identity of the chosen cue maintained in lOFC is available for learning after the ₂₃₆
reward. However, in the absence of mOFC, the input to lOFC from the ongoing ₂₃₇
precise value comparison in mOFC is absent. ₂₃₈
In all the control experiments, the model reached optimal performance within 150 ₂₃₉
trials. In the case of medial OFC lesions however, the performance was significantly ₂₄₀
impaired in the case of V2_HIGH scenarios, when V1 and V2 values were proximate. ₂₄₁
In the case of V2_MID and V2_LOW, the performance was observed to be similar to ₂₄₂
that of controls, except for a slight delay in reaching better performance. Such a ₂₄₃
normal performance in the case of V2_MID and V2_LOW can be attributed to the ₂₄₄
appropriate *credit assignment* by lateral OFC happening during learning. When the ₂₄₅
value difference is sufficiently large, and the credit correctly assigned to the correct ₂₄₆
choice, as the V2 anyway does not reward as much as V1, it is easily learned between ₂₄₇

the lOFC-CBG$_{cue}$ synaptic connections and they can drive the decision without a precise comparison (Fig. 4 D-F). 248 249

Lateral OFC Lesion 250

One of the major changes in case of the lateral lesion is the credit assignment. In the control condition, when there is a reward delivered, the activation of the chosen cue in lOFC is active (Fig 1D, CBG, after 'Reward' until 2500ms). When there is no lOFC in the network, the association of current reward to only current choice can no longer be done. In this case, we still consider that the CSV for each cue is sent as an input to mOFC, because mOFC/vmPFC has been shown to receive projections from the ventral striatum [50, 51, 73], which is a crucial component of the CSV layer. 251 252 253 254 255 256 257

Striking of the observations, in the case when the difference between V1 and V2 was close, monkeys with mOFC lesions showed impaired performance while the controls and animals with lOFC lesions fairly performed well as V1 and V2 were distinct. Such an impairment argued for the role of mOFC to be more sensitive to the value difference between the options. Conversely, when the difference between the values was more, quite surprisingly animals with lOFC lesions were impaired whereas the controls and animals with mOFC lesions could steadily perform optimal choices. While it was an interesting observation to see how mOFC could not compensate even when the difference between the values is high (meaning it is an easy choice), it highlighted the role of lOFC in appropriate credit assignment, i.e. assigning the reward to the appropriate choice made in the current trial rather than to the previous or even the succeeding choice or even to the choice that rewarded the most historically. 258 259 260 261 262 263 264 265 266 267 268 269

In the case of lOFC lesions, the performance was affected in rather contrasting manner. Although eventually the performances reached near-optimal in all three cases of V2_HIGH, V2_MID and V2_LOW, the performance was sub-optimal for most of the earlier part of the sessions especially in the cases where the value difference was larger. This highlights the importance of lOFC in appropriately assigning the credit of reward to the correct option. Impairment of performance in the absence of lateral OFC in the case of V2_MID and V2_LOW was observed for the initial part of the session. This may be due to partial learning in the form of reward-based history maintained in medial OFC (as it was maintained when lateral OFC in intact) (Fig. 4 H-I). 270 271 272 273 274 275 276 277 278

Discussion 279

We demonstrated the OFC on top of classical sub-cortical decision-making systems, with the descriptions of experimentally observed roles of its individual sub-regions. We explain the seemingly dissociated yet more complicated effects of the sub-regions of the OFC on the task performance depending on the task structure (value difference between the options). The OFC is clearly a crucial prefrontal region with heterogeneous representations and dynamics that result in complex behavior. Therefore clearly it is not a feasible idea to attempt a simplistic representation that relies on a unique way of information processing within the OFC, without implying several other brain regions that closely interact with the OFC during the behavior. Instead, we acknowledge the positioning of the OFC in the grand picture of several prefrontal and sub-cortical brain regions, as well as the heterogeneity within itself. Before attempting to model the possible mechanisms within the sub-regions of OFC in detail, it is crucial to build a framework that embeds a representation of environment in an embodied manner (with bodily needs and relevant behavior protocols for testing). Hence this work points towards the interest for modeling the dynamics of OFC as a part of a larger framework of related brain systems that the OFC interacts 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295

with, and the valuation systems it employs to guide decisions and learning. We chose ₂₉₆
one of the well-accounted frameworks of decision-making and reinforcement learning ₂₉₇
involving the BG (CBG loops) and complemented it with the specialized ₂₉₈
representations of the sub-regions of the OFC. We show that a simplistic model as in ₂₉₉
Fig 1B is sufficient for simple tasks. We further demonstrate that a more informed ₃₀₀
decomposed model (*lmOFC* model, Fig 3), while performing equally well on the ₃₀₁
simple tasks, also allows to study performance on more complex tasks in which the ₃₀₂
basic model cannot perform well. ₃₀₃

Choice ₃₀₄

Although the primary role of lateral OFC has been implied to be appropriate credit ₃₀₅
assignment, here we use the fact that lateral OFC still plays an important role in ₃₀₆
driving the activities in the downstream BG loops with the dynamic subjective values ₃₀₇
added to the visual salience. Since the synaptic weights that are changed through ₃₀₈
learning are the ones connecting lOFC to the CBG_{cue} loop, the dynamics would ₃₀₉
passively favor a choice whose connection weights have been sufficiently learned. For ₃₁₀
instance, in the case of a lesion to mOFC (Fig 5), the initial decisions before any ₃₁₁
learning may be guided by lOFC through the CBG loop, randomly with the help of ₃₁₂
intrinsic noise. However, as in the case of Fig 5E,F where only cue rewards ₃₁₃
significantly more, the learning can rapidly increase the synaptic weights in the ₃₁₄
network corresponding to that cue and thus guide the subsequent decisions to that cue. ₃₁₅
This could be one reason why the performance slowly picked up towards the latter end ₃₁₆
of the trials, in the cases of V2_MID and V2_LOW in case of medial OFC lesions. ₃₁₇
Whereas in the case of mOFC lesion under V2_HIGH condition, since V1 and V2 ₃₁₈
almost similarly reward, even if appropriate credit assignment is done between the ₃₁₉
cues corresponding to V1 and V2, the network may not be able to definitively guide ₃₂₀
the decision necessarily towards V1. ₃₂₁

Learning ₃₂₂

Learning in the system occurs at the level of both CSV (for expected values) and ₃₂₃
cortico-striatal synapses. The learning that occurs at the level of cortico-striatal ₃₂₄
synapses indirectly represents the reward contingencies of stimuli in terms of their ₃₂₅
probability. One of the possible motivations behind multiple learning mechanisms in ₃₂₆
the system is the feasibility of a shift of control from the value-comparison based ₃₂₇
processes in the mOFC/vmPFC at the beginning of the trials to a faster, network ₃₂₈
strength based decision through the lOFC-BG loops driven by the learned connection ₃₂₉
weights, in the trials after substantial learning. Such a distinction was reported where ₃₃₀
activities in vmPFC were more remarkably distinct between more deliberative ₃₃₁
situations with slower reaction times as opposed to trials towards the end of the ₃₃₂
experiment or even no-brainer trials (highly probable high reward versus the opposite). ₃₃₃
Moreover, the involvement of value-difference signal in vmPFC consistently decreased ₃₃₄
towards the later trials of the task [74]. However, it is important to note that, in a ₃₃₅
different formal description, it has been highlighted that ventrolateral PFC (vlPFC) ₃₃₆
encodes the *Availability* (probability) of rewards whereas the OFC was shown to ₃₃₇
encode the *Desirability* (palatability) of rewards [75]. However it was shown activity in ₃₃₈
medial and lateral orbitofrontal cortex, extending into vmPFC, was correlated with ₃₃₉
the probability assigned to the action actually chosen on a given trial [70]. ₃₄₀

Lateral and Medial in Learning and Choice : Dissociation or Interaction ? ₃₄₁

What is important to note here is that the dissociate effect observed does not ₃₄₂ necessarily imply that both the sub-regions have dissociate roles in decision-making ₃₄₃ and learning, as recently suggested [56]. The dissociation might be observed in terms ₃₄₄ of the anatomical connectivity in the sense that lateral OFC predominantly receives ₃₄₅ inputs from sensory regions about external environment whereas medial OFC has more ₃₄₆ inputs from internal bodily states and visceral responses. But we argue that seeming ₃₄₇ dissociation in terms of internal processes within these sub-regions is a possible ₃₄₈ network effect as a result of their temporal dynamics. Because, as we have shown here, ₃₄₉ both the sub-regions are involved in the circuitry that is capable of both guiding ₃₅₀ decisions and learning from outcomes. Albeit, the dissociation might be apparent ₃₅₁ because of the fact that each of them might have access to different information about ₃₅₂ the state of the environment and exert control at a different stage of the behavior. ₃₅₃

Extensions ₃₅₄

The task structures used in this work are related to only one type of outcome and ₃₅₅ assuming the motivational value of the outcome is always non-zero. And ventral ₃₅₆ striatum particularly plays an important role in learning action-values as well [4]. By ₃₅₇ incorporating computational accounts of emotional value learning in BLA, and taking ₃₅₈ the internal motivation into account through ventral striatum and lateral ₃₅₉ hypothalamus, experimental results of the role of OFC in paradigms like Reinforcer ₃₆₀ Devaluation [28, 42, 52] can be explained. Possibly, interesting findings like the one in ₃₆₁ Reversal Learning paradigm can be explored where neither lateral nor medial lesions ₃₆₂ of OFC do not affect the behavior whereas the lesion of OFC as a whole affects. As ₃₆₃ mentioned earlier, the OFC has been proposed to represent a cognitive map of task ₃₆₄ space [9] and to encode the *value* of the offered and chosen goods [10, 11] ₃₆₅

State/Task space representation in the OFC ₃₆₆

The OFC has been proposed to encode the task states and represent a cognitive map ₃₆₇ of this task space [9]. It has also been shown that OFC lesions in animals cause ₃₆₈ deficits in acquiring information about the task [27, 76]. In this work, related to the ₃₆₉ simple 2-arm bandit task done under a DIFFICULT condition (Fig 2D, red), the ₃₇₀ model fails to perform as good as in the EASY condition. However, with a slight ₃₇₁ change in the task structure, the model can be shown to perform better. That is, ₃₇₂ instead of presenting the same pair of shapes, imagine there are four possible shapes ₃₇₃ and the 6 possible pairs are presented for choice pseudo-randomly. Even if the reward ₃₇₄ probabilities of both the shapes used in Fig. 2A remain the same, besides two other ₃₇₅ possible options, the options with which these cues are presented change. This can be ₃₇₆ due to the modified learning rates because of the state-change across each trial ₃₇₇ (because no pair is presented consecutively). Alternatively it can also be remarked ₃₇₈ that, even though the value of the first two shapes did not change, the overall value of ₃₇₉ each trial or that of the entire task has changed in the presence of other options. ₃₈₀ These two factors can be hypothesized to be represented in terms of state prediction ₃₈₁ errors and value difference signals in lateral and medial OFC respectively. ₃₈₂

Temporal Dynamics : Delayed presentations, Opportunistic behaviors ₃₈₃

One of the clear limitations of the model presented in this work is that the temporal ₃₈₄ dynamics at various stages of decision-making processes with respect to the discussed ₃₈₅ sub-regions of OFC is not entirely accounted for. Particularly an intracranial EEG ₃₈₆ recordings of OFC in humans showed that lOFC was encoding experienced value in ₃₈₇

the reward delivery phase [77]. This model needs to extended further by incorporating the cortico-cortical interactions, within the subregions of OFC as well as other related prefrontal regions like the ACC which are believed to be rather interactive in their roles in behavior [78]. Furthermore, the role of any possible interaction between both the sub-regions of OFC, given their connectivity through the medial orbital sulci [46], has not been explored much. However, even if it is the case that the lateral OFC represents identity specific rewards and vmPFC represents general, scaled reward signals, it is unclear how these two signals could be linked to sub-serve goal-directed behavior. To this extent, there is not much evidence except one study that showed the functional connectivity was predictive of satiety related changes in choice behaviour [79]. Similarly, the role of VS also becomes crucial in serving such value signals that combine both value and internal motivation before the comparison processes in the mOFC/vmPFC.

Conclusion

The anatomical description implicitly highlights the difference in the accessibility of lateral and medial OFC for experimentation procedures. The same applies to imaging studies, which happen to be a major contribution of studies in humans, that the regions highlighted by BOLD signals cannot be precise enough within the scope of subregions. Another challenge is the homologies of the OFC among humans, nonhuman primates and rodents. Since a good part of the literature on OFC is almost equally contributed by the studies in all three species, it would be major task at hand to be wary of the similarities and the differences among what is defined as OFC in each of these species. Functionally, it is also not straightforward to identify whether a difference in an ability of one species (say humans) to demonstrate a faculty and that of another (say rats) is a difference in kind or a difference in degree. Although, it would be fairly possible to extend the conclusions from one species to another depending on what is being studied (for example, the findings related to action-outcome contingencies or basic behavior in rats might extend well to primates beyond which more flexible representations might emerge).

Challenges in studying the representation and mechanisms in OFC

As far as the interest in the dissociate contribution of subregions of OFC is concerned, there are not so many experimental evidences that could establish a double dissociation between different subregions [28, 42, 43, 80]. Importantly, it has been found out that both lateral and medial regions of OFC represent the perceived value of task events, albeit with different levels of participation in different task settings [42]. It is generally single-neuron recording studies or lesion studies in macaques or rats [81] predominantly on the better accessible lateral OFC than the medial OFC, and BOLD signal correlation from fMRI studies in humans. Few behavioral studies on frontal damage patients have also discussed separate roles of OFC and vmPFC [4, 26, 82, 83]. However, owing to the different techniques and methodologies used, there are few inconsistencies where both lateral and medial regions were exclusively implied for signals during the anticipation of rewards [84]. Another interesting theory about the dissociating role of lOFC and mOFC that probably requires closer look is that while mOFC might represent the values of options in a context where there is no choice to be made, lOFC doesn't represent the value in a choice-free context [80].

Notwithstanding some complementary [85, 86] as well as contrasting [87, 88] findings, separable representations of absolute values and relative values seem to be a

key dissociation between the medial [5, 19, 89, 90] and lateral [91] OFC. Future studies that could effectively dissociate the factors such as salience of the stimuli, internal motivation and thereby investigate the difference between absolute and relative value encoding, should provide more insights into whether lateral and medial OFC play dissociate role to represent them. ₄₃₆ ₄₃₇ ₄₃₈ ₄₃₉ ₄₄₀

Well within one species, reported counter-intuitive findings also might encourage the need to dissociate the roles of lateral and medial OFC. Quite often the ambiguity arises from the task structure which doesn't sufficiently dissociate closely related aspects of a certain behavior. In the few studies that have described the nature of dissociation between lateral and medial OFC, certain results were difficult to explain and there appears to be several possibilities for such results. For instance, when it is observed that neither of the individual lesions of lateral and medial OFC impair the animal in Reversal Learning while the lesion of OFC as a whole does [83, 92], it could be possibly mediated by other sub regions of OFC (central, anterior or posterior) or there might very well be a possible mechanism through which they partially compensate for one another interacting with other parts of the brain. In consecutive reinforcer-devaluation tests, while monkeys with lOFC lesion showed significant impairment, the ones with mOFC lesion were less consistent in their choices across sessions contrasting overall performance similar to that of lateral lesions immediately after the lesion, and with that of the controls in the later sessions after the lesion [28].

Materials and methods

We use a neuro-computational connectionist modeling approach to highlight the organization of subsystems that drive decision-making and learning. The subsystems are built with simplified representations of the experimental findings related to the roles of lateral and medial OFC. The comprehensive model accommodates representations of a fairly complete set of phases involved in decision-making [8]. Besides the component of OFC, the information processing in the model for decision-making and learning also involves minimized yet biologically plausible representation of several sub-cortical mechanisms involved. However, we emphasize more on the contributions of the lateral and medial OFC, maintaining the generic nature of rest of the mechanisms. We first introduce the kind of tasks on which the model is tested, that would demonstrate both the dynamics of a decision as well as the progression of behavior through the task. Then the major computational aspects of the model will be presented, referring to the representations required for the tasks described.

Thalamo-Cortical Basal Ganglia (CBG) Loops

The fundamental decision making networks and learning in the system are implemented in the form of schematic thalamo-cortical basal ganglia (CBG) loops which process the information of the cues, estimated values of the cues and the actions required to select the cues. The core action selection mechanism between multiple options - cues or actions, is implemented using an architecture of the basal ganglia (BG) similar to that has been described in classical descriptions of pathways in BG (summarized well in [93]). Thus we define a CBG loop as a thalamo-cortical BG loop, an example of which is described in S1 Fig. The architecture presents a general idea of the connectivity between the input structures of BG - Subthalamic Nucleus (STN) and Striatum (STR) and the output structures - Globus Pallidus pars Interna (GPi) and Substantia Nigra pars Reticulata (SNr).

We implement computational model of parallel loops of three kinds, which were originally described as : *limbic*, *sensori-motor* and *associative* [94]. The *limbic* loops originate in the orbitomedial prefrontal cortex (generally comprised of OFC and ACC, through amygdala, hypothalamus and the subdivisions of VS (nucleus accumbens) and end back in the medial PFC. The *limbic* loops, besides processing external information, are based on interoceptive information. They are organized around the selection of the goal of the behavior, according to its motivational value, in response to perceived needs or according to its hedonic value. Individual sub-regions of medial prefrontal cortex form the feedback loop through different nuclei of VS [95, 96]. It is these *limbic* loops that we focus on in this work, specifically emphasizing the role of the OFC and possible dissociate roles of its sub-regions, since lateral and medial OFC also are part of this network of parallel loops. For example, in a 2-arm bandit task shown in Fig 1A, the cues (shapes) that are presented in each trial are represented within these *limbic* loops (CBG$_{cue}$ in Fig 1A). The information about the position of the cue (thus the required action to select the cue) is represented in the sensori-motor loops (CBG$_{pos}$ in Fig 1A), from the regions in Parietal Cortex to form the feedback loop through the dorsolateral striatum (DLS) [97, 98]. The lateral prefrontal cortex (lPFC) forms an associative loop with the dorsomedial striatum (DMS), receiving multimodal information from the associative regions of the posterior cortex (ASC in Fig 1A). The combined information of which cue is present in which position, which solves the binding-problem, is represented in the lPFC and DMS [62].

The population dynamics and the learning mechanisms described below have been adapted from similar works before on the thalamo-cortical BG loops [59, 60].

Population Dynamics

All the structures within a CBG loop are implemented as populations that are part of a recurrent neuronal network, with neuron units of similar dynamics. The dynamics of a neural population unit is described in equation 1, similar to previous computational accounts of decision-making in the BG [63]. Assuming each population unit represents an ensemble tuned towards a particular option : I_{ext} is the external input representing the salience of the option, I_s is the input to the unit from its connections (synaptic input) and τ is the decay time constant of the synaptic input and V is the resultant activity of the unit. External input, I_{Ext} is provided only in cortical structures (for the other structures $I_{Ext} = 0$), and T is the threshold of a neuron, depending on the population. Also, symmetry breaking is generated by Gaussian noise δ to the activity of each ensemble at each time step.

$$\tau \frac{dV}{dt} \quad = \quad \text{-}V + I_s + I_{ext} - T \tag{1}$$

$$U = fn(V + \delta) \tag{2}$$

The activation function fn in Eq. 2 is the same for all the structures within a CBG loop and it is a clamping function, except for the striatal structures. The activation of striatal populations, due to their neuronal properties [99–101], can be obtained by applying a sigmoidal transfer function to the activation of CTX-STR inputs in the form of the Boltzmann equation S1 Equation.

The synaptic input to a unit j, I_s^j, which is the input as a result of the connections from units of other structures (say i), depends on the connection weights (w_{ij}) between units i and j , as shown in the equation 3. Except the synaptic connections

that can be learned, the rest remains to be constant connection weights chosen at the beginning, within the range of 0.25 and 0.75, generally chosen around 0.5.

$$I_s^j = \Sigma_i w_{ij} * \hat{G}_{ij} * m_i \tag{3}$$

Also, there is a fixed gain parameter that characterizes the strength of interaction between the two populations to which i and j belong. For example, for any pair of connections ij between CTX(i) and STR(j), the gain $\hat{G}_{CTX\_STR}$ is fixed. A positive or negative $\hat{G}$ defines the connection as excitatory or inhibitory respectively. In the "direct" pathway, as a result of two inhibitory and one excitatory connection, it is referred as a positive feedback loop. In the "hyperdirect" pathway, as a result of two excitatory and one inhibitory connection, it is referred as a negative feedback loop [63].

Learning

The connections between the OFC and the CBG$_{cue}$ loop in the basic model (Fig. 1B) are modifiable. Similarly, after the model is changed to lmOFC model (Fig. 3), the connections between lOFC and its CBG loop, as well as mOFC and ASC are modifiable. The modifiable connections between lOFC and its CBG loop are One-To-One between cue populations in lOFC and cue specific populations in the structures of CBG loop. Whereas the modifiable connections between mOFC and the ASC network are One-To-All i.e, from each cue population in mOFC to all four position populations possible. After every decision and verifying the outcome, the weights are updated. Like in the previous models, all synaptic weights are initialized to 0.5 (SD, 0.005). The weight update term ΔW_t is calculated as a function of reward prediction error (RPE), which is believed to be signalled by dopamine at the level of cortico-striatal synapses. However, it was specifically found that striatal neurons involved in cortico-striatal synapses show long term potentiation (LTP) and long term depression (LTD) with respect to positive or negative prediction error, respectively ([102]). RPE precisely is the difference between the perceived reward value and the expected reward value. In the model, similar to a standard critic-learning RL framework, expected reward values of each stimulus population are maintained and updated.

$$\Delta W_{ij} = \alpha * \delta_t * U_j \tag{4}$$

$$\alpha = \begin{cases} \alpha_{LTP}, & \text{if } \delta_t > 0 \\ \alpha_{LTD}, & \text{otherwise} \end{cases} \tag{5}$$

The RPE, δ_t is calculated using a simple critic learning algorithm given below.

$$\delta_t = R - v_i \tag{6}$$

where R, the reward, is 0 or 1, depending on whether a reward was given or not on that trial. After the ΔW_t is calculated, the synaptic weights are updated according to S2 Equation. And upon weight changes, to make sure the weights stay within the initial bounds, every weight update is followed by a normalization of weights (S3 Equation). v_i is the CSV of the cue represented by neuron i in the CBG. The CSV of the chosen cue is then updated by :

$$v_i \leftarrow v_i + (\delta_t * \alpha_c) \tag{7}$$

where α_c is the critic learning rate and is set to 0.025 and α_{LTP} and α_{LTD} are set to ₅₆₁ 0.004 and 0.002 respectively. ₅₆₂

Current Subjective Value (CSV) ₅₆₃

OFC is known to represent a current subjective value (CSV) of a stimulus with respect ₅₆₄ to the body's internal state (like satiety or desirability of the outcome the stimulus ₅₆₅ announces). Two primary brain structures that crucially involve with the OFC in this ₅₆₆ regard are : the amygdala and the ventral striatum (VS). The basolateral amygdala ₅₆₇ (BLA) has been shown to interact with the OFC and update its stimulus-outcome ₅₆₈ associations and hence the subjective value of a stimulus [53, 54]. On the other hand, ₅₆₉ the ventral striatum was found to represent a unified quantity as a combination of ₅₇₀ subjective value and internal motivation using different kind of neurons [55]. Several ₅₇₁ computational accounts have explained possible implementations of such ₅₇₂ representations [103–107]. A much detailed representation and role of ventral striatum ₅₇₃ and its distinct relation to lateral and medial OFC also could be a key factor to ₅₇₄ study [108–111]. ₅₇₅

Supporting information ₅₇₆

S1 Fig   High level architecture of Thalamo-Cortical-BG Loops in primates. ₅₇₇ Classic BG connectivity : STN and STR as inputs, GPi/SNr as outputs. GPi: Globus ₅₇₈ Pallidus pars Interna; SNr: Substantia Nigra pars Reticulata; STN: Subthalamic ₅₇₉ nucleus; STR: Striatum. The direct pathway from the prefrontal cortex (PFC, here ₅₈₀ used generally, including OFC) via STR to GPi, and the hyperdirect pathway from ₅₈₁ CTX via STN to GPi. It has to be noted that this is only one of several possible ₅₈₂ interpretations of action selection mechanism within BG, as it sufficiently explains ₅₈₃ using one excitatory and inhibitory pathway. Several other interpretations exist, for ₅₈₄ instance, an indirect pathway, which are not considered in this work. Indirect pathway, ₅₈₅ involving STN, GPe (Globus Pallidus pars externa) and STR, is also a part of ₅₈₆ "classical" view of CBG network [112–114]. Image re-illustrated, inspired from [93] ₅₈₇

S1 Appendix. ₅₈₈

S1 Equation.   Sigmoidal transfer function ₅₈₉

S2 Equation.   Oja weight update rule ₅₉₀

S3 Equation.   Weight normalization ₅₉₁

S1 Table.   Parameters of CBG loop structures ₅₉₂

References

1. Cabanac M. Pleasure: the common currency. J Theor Biol. 1992;doi:10.1016/S0022-5193(05)80594-6.

2. Sugden R, Kagel JH, Battalio RC, Green L. Economic Choice Theory: An Experimental Analysis of Animal Behaviour. Econ J. 1996;doi:10.2307/2235234.

3. Kahneman D, Tversky A. Choices, values, and frames. Am Psychol. 1984;doi:10.1037/0003-066X.39.4.341.

4. Fellows LK. Orbitofrontal contributions to value-based decision making: Evidence from humans with frontal lobe damage. Ann N Y Acad Sci. 2011;1239(1):51–58. doi:10.1111/j.1749-6632.2011.06229.x.

5. Grabenhorst F, Rolls ET, Parris BA. From affective value to decision-making in the prefrontal cortex. Eur J Neurosci. 2008;doi:10.1111/j.1460-9568.2008.06489.x.

6. Padoa-Schioppa C. Neurobiology of Economic Choice: A Good-Based Model. Annu Rev Neurosci. 2011;34(1):333–359. doi:10.1146/annurev-neuro-061010-113648.

7. Walton ME, Behrens TEJ, Buckley MJ, Rudebeck PH, Rushworth MFS. Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. Neuron. 2010;65(6):927–939. doi:10.1016/j.neuron.2010.02.027.

8. Rangel A, Camerer C, Montague PR. A framework for studying the neurobiology of value-based decision making; 2008.

9. Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. Orbitofrontal cortex as a cognitive map of task space. Neuron. 2014;81(2):267–279. doi:10.1016/j.neuron.2013.11.005.

10. Padoa-Schioppa C, Jandolo L, Visalberghi E. Multi-stage mental process for economic choice in capuchins. Cognition. 2006;99(1):B1–B13. doi:10.1016/j.cognition.2005.04.008.

11. Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. Nature. 2006;441(7090):223–226. doi:10.1038/nature04676.

12. Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci. 1998;doi:10.1038/1625.

13. Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. Nature. 1999;doi:10.1038/22268.

14. Shidara M, Richmond BJ. Anterior cingulate: Single neuronal signals related to degree of reward expectancy. Science (80-). 2002;doi:10.1126/science.1069504.

15. Padoa-Schioppa C, Assad JA. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat Neurosci. 2008;11(1):95–102. doi:10.1038/nn2020.

16. Padoa-Schioppa C. Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. J Neurosci. 2009;29(44):14004–14014. doi:10.1523/JNEUROSCI.3751-09.2009.

17. Conen KE, Padoa-Schioppa C. Partial Adaptation to the Value Range in the Macaque Orbitofrontal Cortex. J Neurosci. 2019; p. 2279–18. doi:10.1523/JNEUROSCI.2279-18.2019.

18. Montague PR, Berns GS. Neural economics and the biological substrates of valuation; 2002.

19. Rolls ET, McCabe C, Redoute J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. Cereb Cortex. 2008;doi:10.1093/cercor/bhm097.

20. Grabenhorst F, D'Souza AA, Parris BA, Rolls ET, Passingham RE. A common neural scale for the subjective pleasantness of different primary rewards. Neuroimage. 2010;51(3):1265–1274. doi:10.1016/j.neuroimage.2010.03.043.

21. Grabenhorst F, Rolls ET. Value, pleasure and choice in the ventral prefrontal cortex. Trends Cogn Sci. 2011;15(2):56–67. doi:10.1016/j.tics.2010.12.004.

22. Stalnaker TA, Cooch NK, Schoenbaum G. What the orbitofrontal cortex does not do. Nat Neurosci. 2015;18(5):620–627. doi:10.1038/nn.3982.

23. Roberts AC. Primate orbitofrontal cortex and adaptive behaviour. Trends Cogn Sci. 2006;10(2):83–90. doi:10.1016/j.tics.2005.12.002.

24. Kringelbach ML, Rolls ET. The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology; 2004.

25. Izquierdo A, Jentsch JD. Reversal learning as a measure of impulsive and compulsive behavior in addictions; 2012.

26. Fellows LK. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain. 2003;126(8):1830–1837. doi:10.1093/brain/awg180.

27. Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, et al. Reward-related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. J Cogn Neurosci. 2004;16(3):463–478. doi:10.1162/089892904322926791.

28. Rudebeck PH, Murray EA. Dissociable Effects of Subtotal Lesions within the Macaque Orbital Prefrontal Cortex on Reward-Guided Behavior. J Neurosci. 2011;31(29):10569–10578. doi:10.1523/jneurosci.0091-11.2011.

29. Schoenbaum G, Nugent SL, Saddoris MP, Setlow B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport. 2002;13(6):885–890. doi:10.1097/00001756-200205070-00030.

30. Schoenbaum G, Setlow B, Saddoris MP, Gallagher M. Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. Neuron. 2003;39(5):855–867. doi:10.1016/S0896-6273(03)00474-4.

31. Schoenbaum G, Chiba AA, Gallagher M. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Neurosci. 1998;doi:10.1038/407.

32. Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci. 1999;19(5):1876–1884.

33. Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature. 2006;439(7078):865–870. doi:10.1038/nature04490.

34. Morrison SE, Saez A, Lau B, Salzman CD. Different Time Courses for Learning-Related Changes in Amygdala and Orbitofrontal Cortex. Neuron. 2011;71(6):1127–1140. doi:10.1016/j.neuron.2011.07.016.

35. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area; 2012.

36. Carmichael ST, Price JL. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. J Comp Neurol. 1994;346(3):366–402. doi:10.1002/cne.903460305.

37. Öngür D, Ferry AT, Price JL. Architectonic subdivision of the human orbital and medial prefrontal cortex. J Comp Neurol. 2003;460(3):425–449. doi:10.1002/cne.10609.

38. Price JL. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. In: Ann. N. Y. Acad. Sci.; 2007.

39. Choi JS, Kang DH, Kim JJ, Ha TH, Lee JM, Youn T, et al. Left anterior subregion of orbitofrontal cortex volume reduction and impaired organizational strategies in obsessive-compulsive disorder. J Psychiatr Res. 2004;38(2):193–199. doi:10.1016/j.jpsychires.2003.08.001.

40. Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA. Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat Neurosci. 2013;16(8):1140–1145. doi:10.1038/nn.3440.

41. Hebscher M, Barkan-Abramski M, Goldsmith M, Aharon-Peretz J, Gilboa A. Memory, decision-making, and the ventromedial prefrontal cortex (vmpfc): The roles of subcallosal and posterior orbitofrontal cortices in monitoring and control processes. Cereb Cortex. 2016;26(12):4590–4601. doi:10.1093/cercor/bhv220.

42. Bouret S, Richmond BJ. Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys. J Neurosci. 2010;30(25):8591–8601. doi:10.1523/jneurosci.0049-10.2010.

43. Noonan MP, Walton ME, Behrens TEJ, Sallet J, Buckley MJ, Rushworth MFS. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc Natl Acad Sci. 2010;107(47):20547–20552. doi:10.1073/pnas.1012246107.

44. Cavada C. The Mysterious Orbitofrontal Cortex. Foreword. Cereb Cortex. 2000;doi:10.1093/cercor/10.3.205.

45. Mackey S, Petrides M. Quantitative demonstration of comparable architectonic areas within the ventromedial and lateral orbital frontal cortex in the human and the macaque monkey brains. Eur J Neurosci. 2010;doi:10.1111/j.1460-9568.2010.07465.x.

46. Zald DH, McHugo M, Ray KL, Glahn DC, Eickhoff SB, Laird AR. Meta-Analytic Connectivity Modeling Reveals Differential Functional Connectivity of the Medial and Lateral Orbitofrontal Cortex. Cereb Cortex. 2014;24(1):232–248. doi:10.1093/cercor/bhs308.

47. Carmichael ST, Price JL. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol. 1996;doi:10.1002/(SICI)1096-9861(19960722)371:2<179::AID-CNE1>3.0.CO;2-#.

48. Cavada C. The Anatomical Connections of the Macaque Monkey Orbitofrontal Cortex. A Review. Cereb Cortex. 2000;doi:10.1093/cercor/10.3.220.

49. Kahnt T, Chang LJ, Park SQ, Heinzle J, Haynes JD. Connectivity-based parcellation of the human orbitofrontal cortex. J Neurosci. 2012;doi:10.1523/JNEUROSCI.0257-12.2012.

50. Carmichael ST, Price JL. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J Comp Neurol. 1995;doi:10.1002/cne.903630408.

51. Carmichael ST, Price JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol. 1995;363(4):642–664. doi:10.1002/cne.903630409.

52. Málková L, Gaffan D, Murray EA. Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. J Neurosci. 1997;.

53. Elliott R. Dissociable Functions in the Medial and Lateral Orbitofrontal Cortex: Evidence from Human Neuroimaging Studies. Cereb Cortex. 2000;doi:10.1093/cercor/10.3.308.

54. Kennerley SW, Wallis JD. Evaluating Choices By Single Neurons in the Frontal Lobe. Neuroscience. 2009;29(10):2061–2073. doi:10.1111/j.1460-9568.2009.06743.x.Evaluating.

55. Bissonette GB, Burton AC, Gentry RN, Goldstein BL, Hearn TN, Barnett BR, et al. Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. PLoS One. 2013;8(5):1–10. doi:10.1371/journal.pone.0064673.

56. Miller KJ. Value Representations in Orbitofrontal Cortex Drive Learning , but not Choice. bioRXiv Prepr. 2018; p. 1–25. doi:10.1101/245720.

57. Pasquereau B, Nadjar A, Arkadir D, Bezard E, Goillandeau M, Bioulac B, et al. Shaping of Motor Responses by Incentive Values through the Basal Ganglia. J Neurosci. 2007;27(5):1176–1183. doi:10.1523/jneurosci.3745-06.2007.

58. Garenne A, Pasquereau B, Guthrie M, Bioulac B, Boraud T. Basal Ganglia Preferentially Encode Context Dependent Choice in a Two-Armed Bandit Task. Front Syst Neurosci. 2011;5(May):1–9. doi:10.3389/fnsys.2011.00023.

59. Guthrie M, Leblois A, Garenne A, Boraud T. Interaction between cognitive and motor cortico-basal ganglia loops during decision making: a computational study. J Neurophysiol. 2013;109(12):3025–3040. doi:10.1152/jn.00026.2013.

60. Topalidou M, Rougier NP, Topalidou M, Rougier NP, Interaction R, Topalidou M, et al. [Re] Interaction between cognitive and motor cortico-basal ganglia loops during decision making : a computational study To cite this version : HAL Id : hal-01201790 Re Science [Re] Interaction between cognitive and motor cortico-basal ganglia loops du. 2015; p. 0–6.

61. Nallapu BTBT, Rougier NPNP. Dynamics of reward based decision making: A computational study. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). 2016;9886 LNCS:322–329. doi:10.1007/978-3-319-44778-0_38.

62. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol. 1989;61(2):331–349. doi:10.1152/jn.1989.61.2.331.

63. Leblois A. Competition between Feedback Loops Underlies Normal and Pathological Dynamics in the Basal Ganglia. J Neurosci. 2006;26(13):3567–3583. doi:10.1523/JNEUROSCI.5050-05.2006.

64. FitzGerald THB, Seymour B, Dolan RJ. The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. J Neurosci. 2009;29(26):8388–8395. doi:10.1523/jneurosci.0717-09.2009.

65. Wang XJ. Decision Making in Recurrent Neuronal Circuits. Neuron. 2008;60(2):215–234. doi:10.1016/j.neuron.2008.09.034.

66. Rolls ET, Grabenhorst F, Deco G. Choice, difficulty, and confidence in the brain. Neuroimage. 2010;doi:10.1016/j.neuroimage.2010.06.073.

67. Strait CE, Blanchard TC, Hayden BY. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron. 2014;doi:10.1016/j.neuron.2014.04.032.

68. Tsutsui KI, Grabenhorst F, Kobayashi S, Schultz W. A dynamic code for economic object valuation in prefrontal cortex neurons. Nat Commun. 2016;7. doi:10.1038/ncomms12554.

69. Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. In: Behav. Econ. Prefer. Choices, Happiness; 2016. p. 593–616.

70. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature. 2006;doi:10.1038/nature04766.

71. Hampton AN, Bossaerts P, O'Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J Neurosci. 2006;doi:10.1523/JNEUROSCI.1010-06.2006.

72. Kim M, Ducros M, Carlson T, Ronen I, He S, Ugurbil K, et al. Anatomical correlates of the functional organization in the human occipitotemporal cortex. Magn Reson Imaging. 2006;doi:10.1016/j.mri.2005.12.005.

73. Haber SN, Knutson B. The reward circuit: Linking primate anatomy and human imaging; 2010.

74. Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MFSS, Behrens TEJJ. Mechanisms underlying cortical activity during value-guided choice. Nat Neurosci. 2012;15(3):470–476. doi:10.1038/nn.3017.

75. Rudebeck PH, Saunders RC, Lundgren DA, Murray EA. Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes. Neuron. 2017;95(5):1208–1220.e5. doi:10.1016/j.neuron.2017.07.042.

76. Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, et al. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci. 2011;doi:10.1038/nn.2957.

77. Li Y, Vanni-Mercier G, Isnard J, Mauguière F, Dreher JC. The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. Brain. 2016;139(4):1295–1309. doi:10.1093/brain/awv409.

78. Fs Rushworth M, Kolling N, Rô Me Sallet J, Mars RB, Rushworth MFS, Kolling N, et al. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr Opin Neurobiol. 2012;22(6):946–955. doi:10.1016/j.conb.2012.04.011.

79. Howard JD, Kahnt T. Identity-specific reward representations in orbitofrontal cortex are modulated by selective devaluation. J Neurosci. 2017;doi:10.1523/JNEUROSCI.3473-16.2017.

80. Nogueira R, Abolafia JM, Drugowitsch J, Balaguer-Ballester E, Sanchez-Vives MV, Moreno-Bote R. Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. Nat Commun. 2017;8. doi:10.1038/ncomms14823.

81. Murray EA, Izquierdo A. Orbitofrontal cortex and amygdala contributions to affect and action in primates. Ann N Y Acad Sci. 2007;1121:273–296. doi:10.1196/annals.1401.021.

82. Fellows LK, Farah MJ. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. Cereb Cortex. 2005;doi:10.1093/cercor/bhh108.

83. Fellows LK, Farah MJ. The role of ventromedial prefrontal cortex in decision making: Judgment under uncertainty or judgment per se? Cereb Cortex. 2007;17(11):2669–2674. doi:10.1093/cercor/bhl176.

84. O'Neill M, Schultz W. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. Neuron. 2010;doi:10.1016/j.neuron.2010.09.031.

85. Kringelbach ML, O'Doherty J, Rolls ET, Andrews C. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. Cereb Cortex. 2003;doi:10.1093/cercor/13.10.1064.

86. Abler B, Walter H, Erk S. Neural correlates of frustration. Neuroreport. 2005;16(7):669–672. doi:10.1097/00001756-200505120-00003.

87. Fujiwara J, Tobler PN, Taira M, Iijima T, Tsutsui KI. Personality-dependent dissociation of absolute and relative loss processing in orbitofrontal cortex. Eur J Neurosci. 2008;doi:10.1111/j.1460-9568.2008.06096.x.

88. Elliott R, Agnew Z, Deakin JFW. Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans. Eur J Neurosci. 2008;27(9):2213–2218. doi:10.1111/j.1460-9568.2008.06202.x.

89. O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci. 2001;doi:10.1038/82959.

90. O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ. Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. Neuropsychologia. 2003;41(2):147–155. doi:10.1016/S0028-3932(02)00145-8.

91. Grabenhorst F, Rolls ET. Different representations of relative and absolute subjective value in the human brain. Neuroimage. 2009;48(1):258–268. doi:10.1016/j.neuroimage.2009.06.045.

92. Clarke HF, Robbins TW, Roberts AC. Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. J Neurosci. 2008;28(43):10972–10982. doi:10.1523/JNEUROSCI.1521-08.2008.

93. Boraud T, Leblois A, Rougier NP. A natural history of skills. Prog Neurobiol. 2018;doi:10.1016/j.pneurobio.2018.08.003.

94. Krack P, Hariz MI, Baunez C, Guridi J, Obeso JA. Deep brain stimulation: From neurology to psychiatry? Trends Neurosci. 2010;33(10):474–484. doi:10.1016/j.tins.2010.07.002.

95. Kringelbach ML. The human orbitofrontal cortex: linking reward to hedonic experience. Nat Rev Neurosci. 2005;6(9):691–702. doi:10.1038/nrn1747.

96. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: Opportunity costs and the control of response vigor. Psychopharmacology (Berl). 2007;doi:10.1007/s00213-006-0502-4.

97. Alexander G. Parallel Organization of Functionally Segregated Circuits Linking Basal Ganglia and Cortex. Annu Rev Neurosci. 1986;doi:10.1146/annurev.neuro.9.1.357.

98. Sommer MA, Wurtz RH. What the Brain Stem Tells the Frontal Cortex. I. Oculomotor Signals Sent from Superior Colliculus to Frontal Eye Field Via Mediodorsal Thalamus. J Neurophysiol. 2004;doi:10.1152/jn.00738.2003.

99. Wilson CJ, Groves PM. Spontaneous firing patterns of identified spiny neurons in the rat neostriatum. Brain Res. 1981;doi:10.1016/0006-8993(81)90211-0.

100. Nisenbaum ES, Wilson CJ. Potassium currents responsible for inward and outward rectification in rat neostriatal spiny projection neurons. J Neurosci. 1995;.

101. Sandstrom MI, Rebec GV. Characterization of striatal activity in conscious rats: Contribution of NMDA and AMPA/kainate receptors to both spontaneous and glutamate-driven firing. Synapse. 2003;doi:10.1002/syn.10142.

102. Pawlak V, Kerr JND. Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. J Neurosci. 2008;28(10):2435–2446. doi:10.1523/JNEUROSCI.4402-07.2008.

103. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci. 1996;.

104. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science (80-). 1997;doi:10.1126/science.275.5306.1593.

105. Grossberg S, Bullock D, Dranias MR. Neural Dynamics Underlying Impaired Autonomic and Conditioned Responses Following Amygdala and Orbitofrontal Lesions. vol. 122; 2008.

106. Vitay J, Hamker FH. Timing and expectation of reward: A neuro-computational model of the afferents to the ventral tegmental area. Front Neurorobot. 2014;doi:10.3389/fnbot.2014.00004.

107. Kaushik PS, Carrere M, Alexandre F, Raju SB. A biologically inspired neuronal model of reward prediction error computation. In: Proc. Int. Jt. Conf. Neural Networks; 2017.

108. Corbit LH, Fischbach SC, Janak PH. Nucleus accumbens core and shell are differentially involved in general and outcome-specific forms of Pavlovian-instrumental transfer with alcohol and sucrose rewards. Eur J Neurosci. 2016;43(9):1229–1236. doi:10.1111/ejn.13235.

109. West EA, Carelli RM. Nucleus accumbens core and shell differentially encode Reward-Associated cues after reinforcer devaluation. J Neurosci. 2016;36(4):1128–1139. doi:10.1523/JNEUROSCI.2976-15.2016.

110. Saddoris MP, Cacciapaglia F, Wightman RM, Carelli RM. Differential dopamine release dynamics in the nucleus accumbens core and shell reveal complementary signals for error prediction and incentive motivation. J Neurosci. 2015;35(33):11572–11582. doi:10.1523/JNEUROSCI.2344-15.2015.

111. Bassareo V, Cucca F, Musio P, Lecca D, Frau R, Di Chiara G. Nucleus accumbens shell and core dopamine responsiveness to sucrose in rats: Role of response contingency and discriminative/conditioned cues. Eur J Neurosci. 2015;41(6):802–809. doi:10.1111/ejn.12839.

112. Gurney K, Prescott TJ, Redgrave P. Gurney Et Al. 2001a. 2001;410:1–10.

113. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. Biol Cybern. 2001;84(6):411–423. doi:10.1007/PL00007985.

114. Calabresi P, Picconi B, Tozzi A, Ghiglieri V, Di Filippo M. Direct and indirect pathways of basal ganglia: A critical reappraisal; 2014.
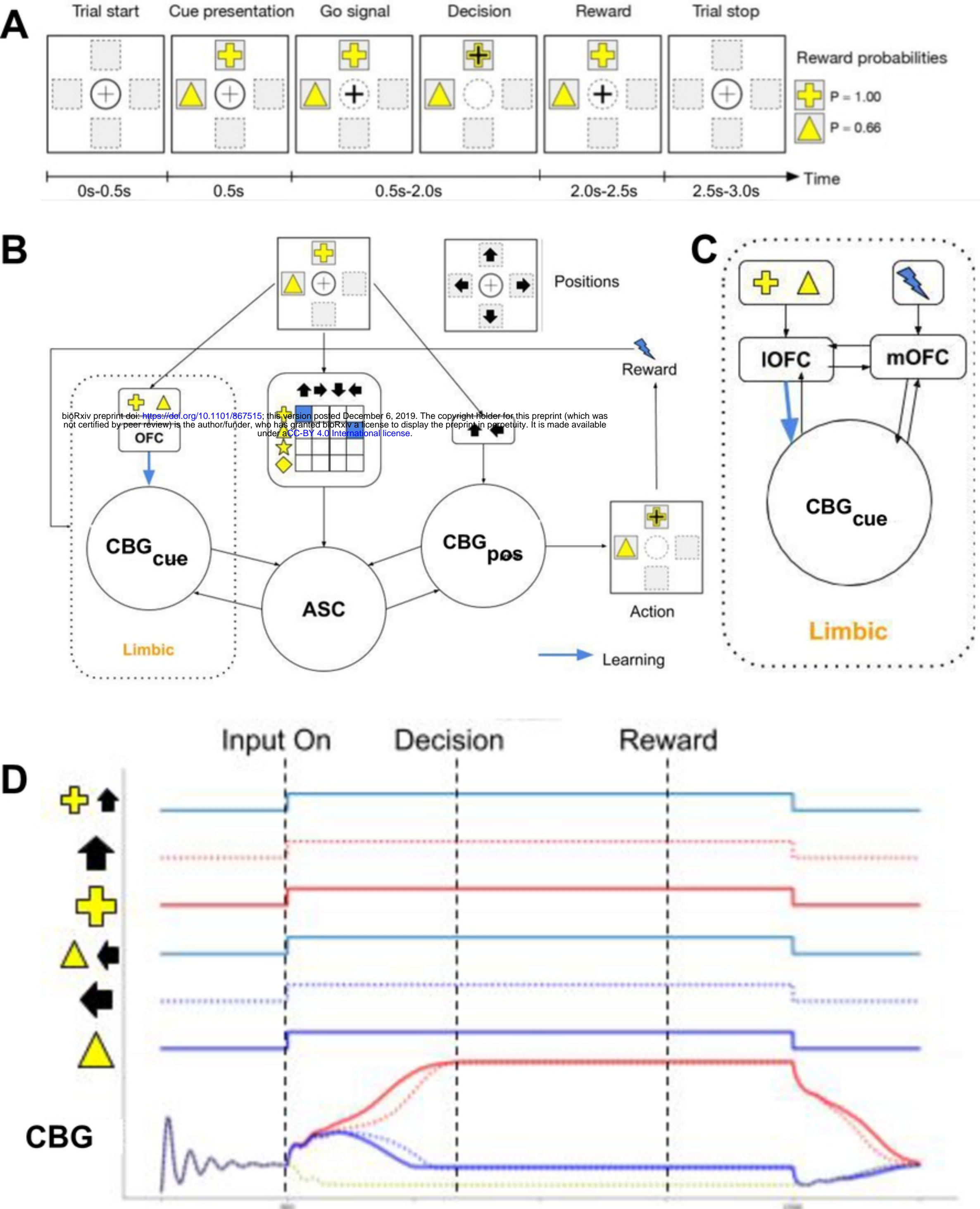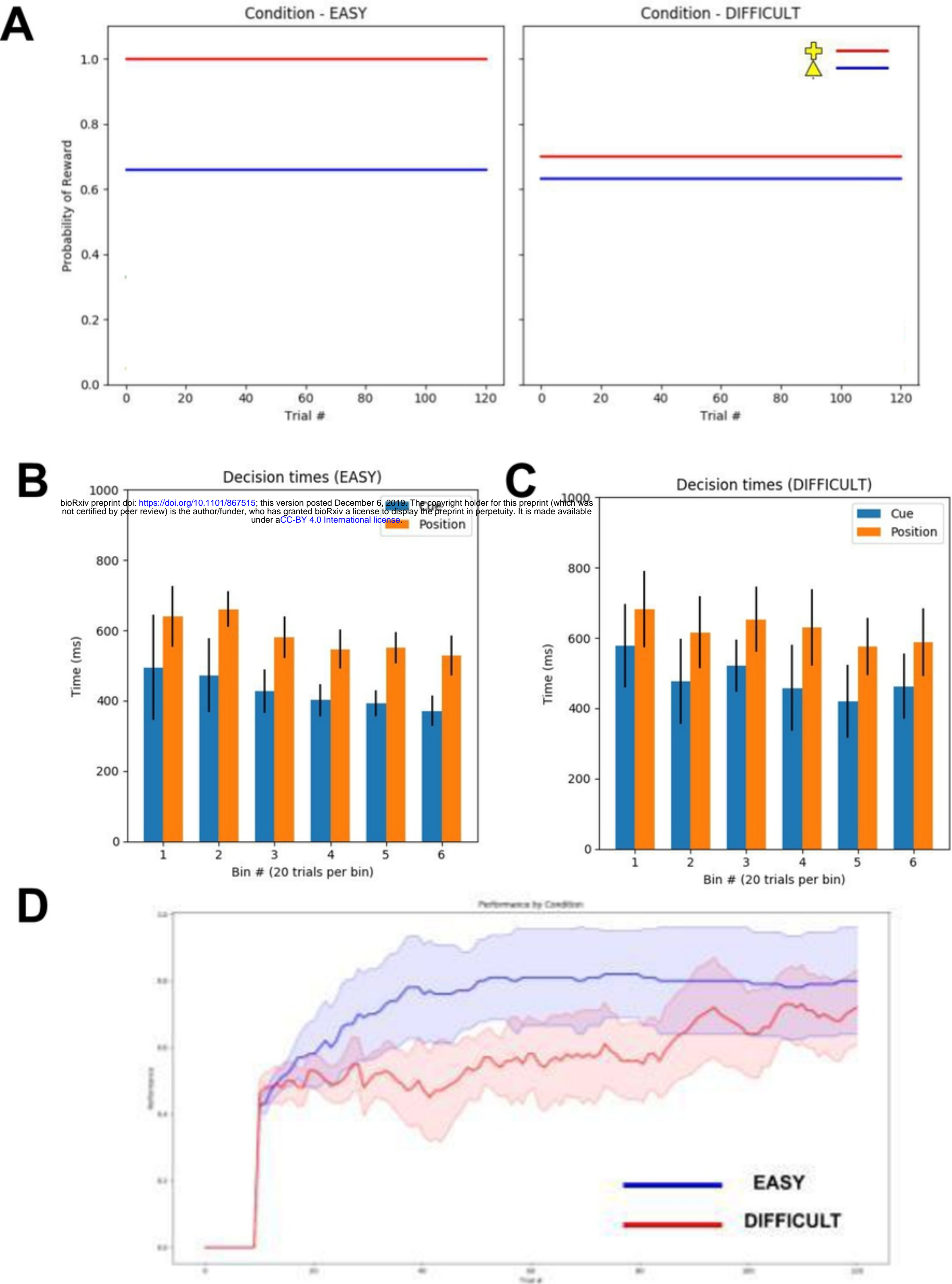
Fig1

A

Condition - EASY · Condition - DIFFICULT

B

Decision times (EASY)

C

Decision times (DIFFICULT)

D

Performance by Condition

EASY

DIFFICULT

Fig2

Fig3

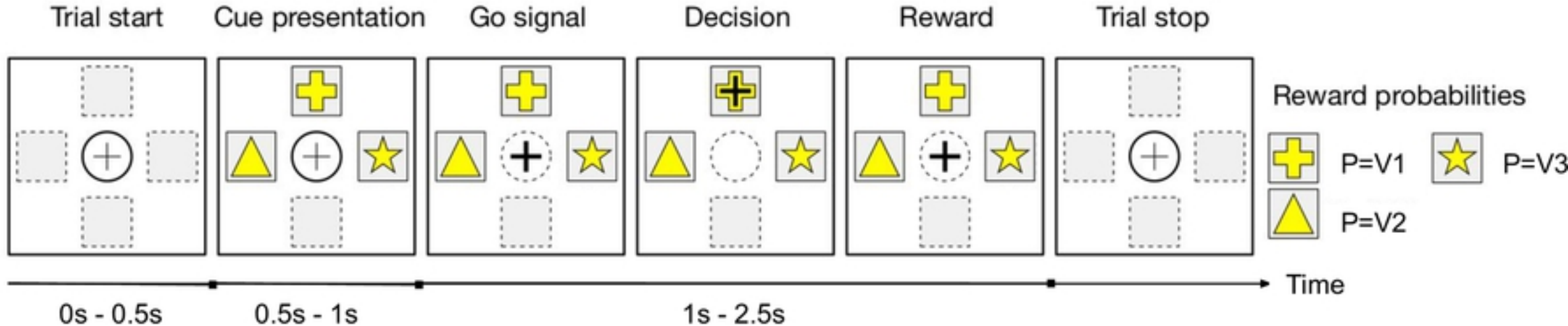| Trial start	Cue presentation	Go signal	Decision	Reward	Trial stop
0s - 0.5s	0.5s - 1s	1s - 2.5s			

Reward probabilities

$P=V1$ (cross) $P=V3$ (star)

$P=V2$ (triangle)

Time
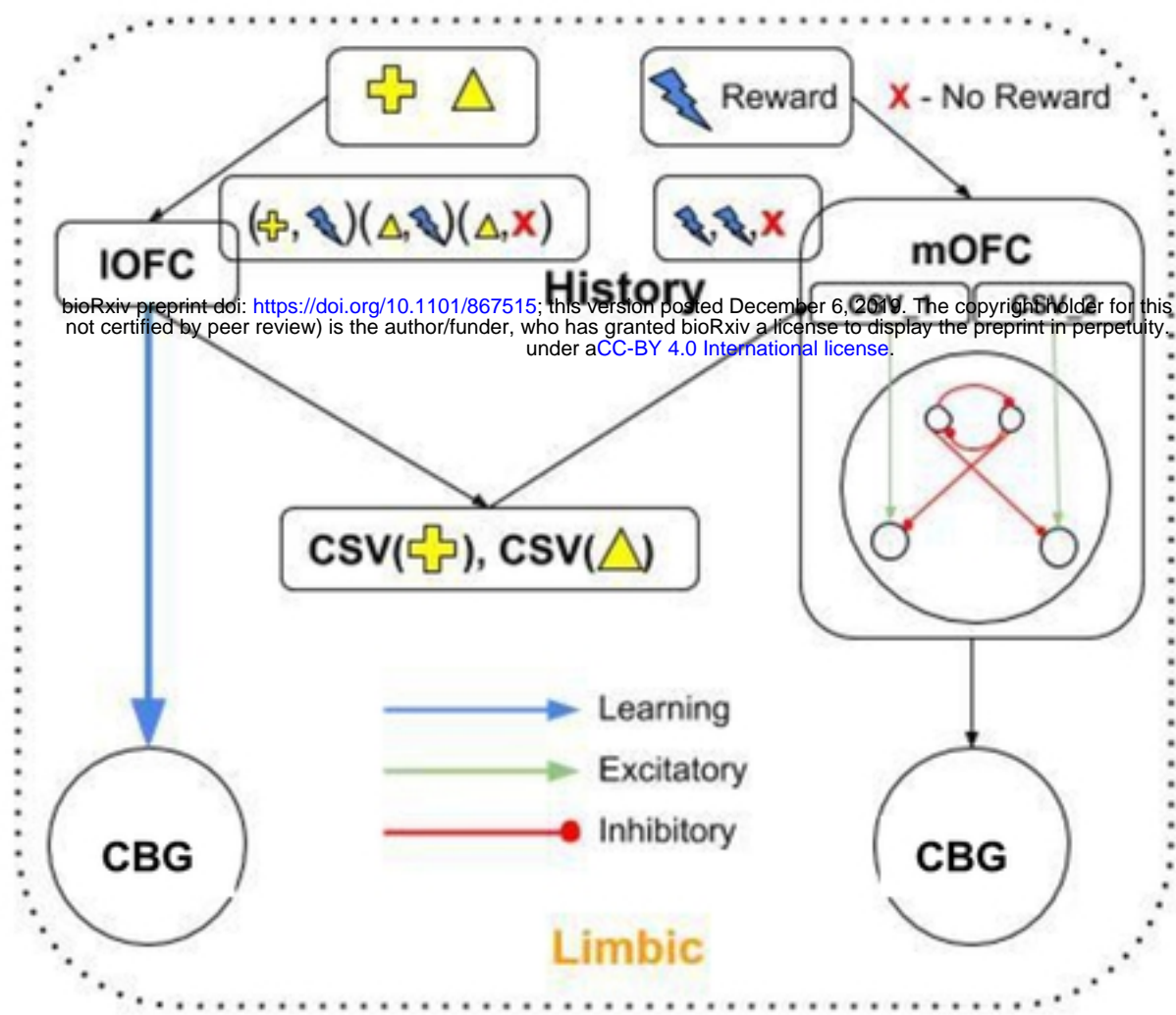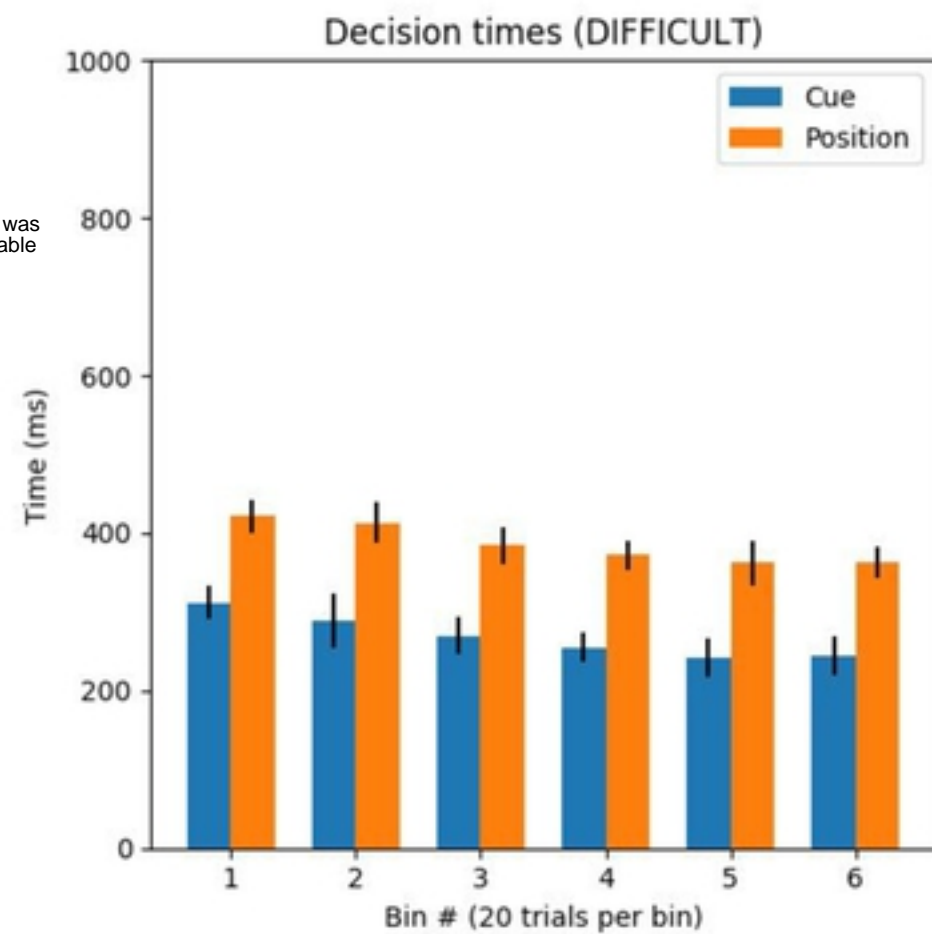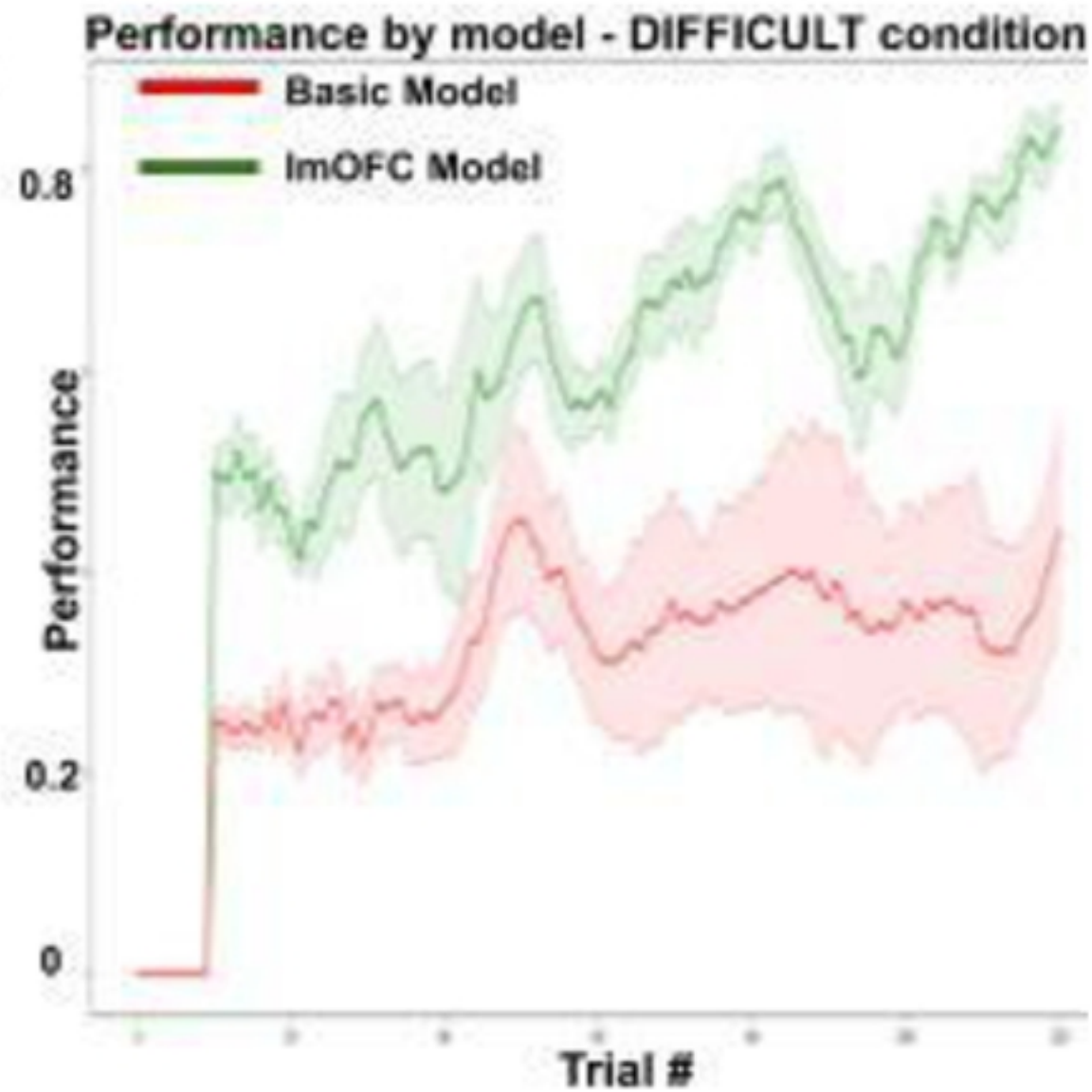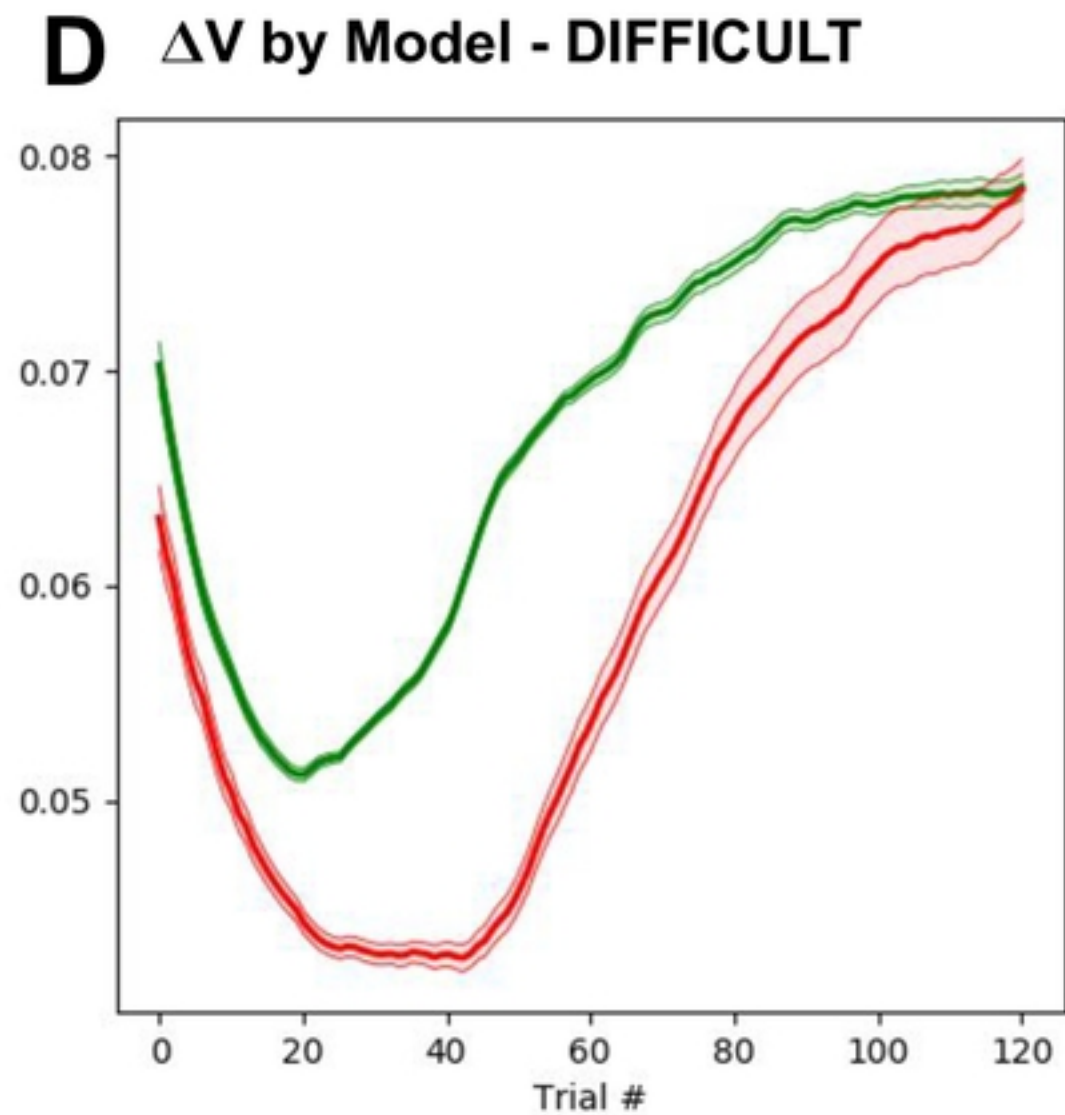
Fig4

Effects of lesions on 3-armed bandit task

Fig5