

25<br>26 **Acknowledgments:** A.F., C.E., B.E.R., and P.H.R. are supported by grants from the BRAIN<br>initiative (R01MH117040). B.E.R. is supported by grants from NIMH (R01MH111439) and<br>28 NINDS (R01NS109498). A.F. is supported by Over 28 NINDS (R01NS109498). A.F. is supported by Overseas Research Fellowship from Takeda<br>29 Science Foundation and a Brain & Behavior Research Foundation Young Investigator grant 29 Science Foundation and a Brain & Behavior Research Foundation Young Investigator grant<br>20 (#28979). We would like to thank Dr. Paula Croxson for providing the foundation on which this 29 Science Foundation and a Brain & Behavior Research Foundation Young Investigator grant<br>20 (#28979). We would like to thank Dr. Paula Croxson for providing the foundation on which this<br>21 work was built. Dr. Validae Hari (#28979). We would like to thank Dr. Paula Croxson for providing the foundation on which this<br>31 work was built, Dr. Yukiko Hori for advising on drug preparation, and Jairo Munoz and<br>22. Niveniane Biopkowske for essistance 31 work was built, Dr. Yukiko Hori for advising on drug preparation, and Jairo Munoz and<br>32 Niranjana Bienkowska for assistance with data acquisition. For help with fMRI data pre-<br>22 Precessing and analysis we thank Drs Ba 32 Niranjana Bienkowska for assistance with data acquisition. For help with fMRI data pre-<br>33 processing and analysis we thank Drs Paul Taylor, Alex Franco, and Vincent Costa. 33 processing and analysis we thank Drs Paul Taylor, Alex Franco, and Vincent Costa.

## 35

**Abstract**<br>36 Flexibly adjusting our behavioral strategies based on the environmental context is critical to 37 maximize rewards. Ventrolateral prefrontal cortex (vlPFC) has been implicated in both learning<br>38 and decision-making for probabilistic rewards, although how context influences these processes 38 and decision-making for probabilistic rewards, although how context influences these processes<br>39 remains unclear. We collected functional neuroimaging data while rhesus macaques performed a 39 remains unclear. We collected functional neuroimaging data while rhesus macaques performed a<br>30 probabilistic learning task in two contexts: one with novel and another with familiar visual 39 remains unclear. We collected functional neuroimaging data while rhesus macaques performed a<br>30 probabilistic learning task in two contexts: one with novel and another with familiar visual<br>41 stimuli We found that activ 40 probabilistic learning task in two contexts: one with novel and another with familiar visual<br>41 stimuli. We found that activity in vIPFC encoded rewards irrespective of the context but encoded<br>42 behavioral strategies t 41 stimuli. We found that activity in vIPFC encoded rewards irrespective of the context but encoded<br>42 behavioral strategies that depend on reward outcome (win-stay/lose-shift) preferentially in novel<br>43 contexts. Function behavioral strategies that depend on reward outcome (win-stay/lose-shift) preferentially in novel<br>43 contexts. Functional connectivity between vlPFC and anterior cingulate cortex varied with<br>44 behavioral strategy in novel 43 contexts. Functional connectivity between vlPFC and anterior cingulate cortex varied with<br>44 behavioral strategy in novel learning blocks. By contrast, connectivity between vlPFC and<br>45 mediadereal thelemus was bighet w 44 behavioral strategy in novel learning blocks. By contrast, connectivity between vlPFC and<br>45 mediodorsal thalamus was highest when subjects repeated a prior choice. Furthermore,<br>46 pharmacological D2-receptor blockade a 46 pharmacological D2-receptor blockade altered behavioral strategies during learning and resting-<br>47 state vlPFC activity. Taken together, our results suggest that multiple vlPFC-linked circuits pharmacological D2-receptor blockade altered behavioral strategies during learning and resting-<br>47 state vlPFC activity. Taken together, our results suggest that multiple vlPFC-linked circuits<br>48 septimbre to adoptive deci 47 state vlPFC activity. Taken together, our results suggest that multiple vlPFC-linked circuits<br>48 contribute to adaptive decision-making in different contexts. 48 contribute to adaptive decision-making in different contexts.

## 50

50 **Introduction**<br>51 In order to obtain the best possible outcome, organisms must flexibly adjust their behavior 51 In order to obtain the best possible outcome, organisms must flexibly adjust their behavior<br>52 depending on environmental context<sup>1</sup>. For example, you have probably already learned through depending on environmental context<sup>1</sup>. For example, you have probably already learned through  $\frac{2}{\sqrt{3}}$ 

53 trial and error which transportation method, whether it be taxi, bus, or subway is best to take to get you to the airport in your hometown. However, you would need to reassess your strategy and learn which option would 55 learn which option would be most reliable to take when faced with the same task in an unfamiliar<br>56 context like Tokyo or London. Such flexible adjustment of behavior depending on the context is 56 context like Tokyo or London. Such flexible adjustment of behavior depending on the context is<br>57 vital for optimal behavior in an uncertain world. Its failure can be catastrophic, such as in 57 vital for optimal behavior in an uncertain world. Its failure can be catastrophic, such as in<br>58 gambling disorder, in which patients often lack cognitive flexibility and engage in risky 58 gambling disorder, in which patients often lack cognitive flexibility and engage in risky<br>59 behaviors despite significant economic losses<sup>2-4</sup>, or in schizophrenia, which is known to be 58 gambling disorder, in which patients often lack cognitive flexibility and engage in risky<br>59 behaviors despite significant economic losses<sup>2-4</sup>, or in schizophrenia, which is known to be<br>60 accessited with inflavible d behaviors despite significant economic losses<sup>2-4</sup>, or in schizophrenia, which is known to be associated with inflexible decisions based on delusional beliefs<sup>5, 6</sup>. associated with inflexible decisions based on delusional beliefs<sup>5, 6</sup>. associated with inflexible decisions based on delusional beliefs<sup>3, o</sup>.<br>61 In non-human primates, the ventrolateral part of prefrontal cortex (vlPFC), especially Walker's

62 area 12, is critical for this type of learning and decision-making in uncertain environments.<br>63 Specifically, fMRI and single neuron recording studies show that activity in vlPFC represents 63 Specifically, fMRI and single neuron recording studies show that activity in vIPFC represents<br>64 outcome probability and integrates this information into subjective value computations<sup>7-11</sup>. 63 Specifically, fMRI and single neuron recording studies show that activity in vlPFC represents<br>64 outcome probability and integrates this information into subjective value computations<sup>7-11</sup>. outcome probability and integrates this information into subjective value computations<sup>7-11</sup>.<br>65 Studies that have either lesioned or inactivated vIPFC have reported that this area causally<br>66 . contributes to probabilist 65 Studies that have either lesioned or inactivated vlPFC have reported that this area causally<br>66 contributes to probabilistic learning and decision-making<sup>8, 12</sup>, but is not required when the<br>67 consolidations between s contributes to probabilistic learning and decision-making  $8, 12$ 66 contributes to probabilistic learning and decision-making<sup>8, 12</sup>, but is not required when the<br>67 associations between stimuli and rewards are deterministic<sup>13</sup>. A recent study also highlighted that<br>68 deceminarie proj associations between stimuli and rewards are deterministic<sup>13</sup> 67 associations between stimuli and rewards are deterministic<sup>13</sup>. A recent study also highlighted that<br>68 dopaminergic projections to vlPFC and premotor cortex are critical for making choices in<br>69 stochastic environment 68 dopaminergic projections to vlPFC and premotor cortex are critical for making choices in<br>69 stochastic environments<sup>14</sup> again indicating that a properly functioning vlPFC is essential for<br>70 stochastic docision making stochastic environments<sup>14</sup> again indicating that a properly functioning vlPFC is essential for probabilistic decision-making. The probabilistic decision-making.<br>
The situations where people or animals have to adaptively determine the best course of action

72 based on probabilistic feedback, they often use 'win-stay/lose-shift' behavioral strategies. Put<br>73 simply, subjects pursue previously-rewarded choices and avoid previously-unrewarded choices<sup>15</sup>. based on probabilistic feedback, they often use 'win-stay/lose-shift' behavioral strategies. Put<br>
73 simply, subjects pursue previously-rewarded choices and avoid previously-unrewarded choices<sup>15</sup>. simply, subjects pursue previously-rewarded choices and avoid previously-unrewarded choices<sup>15</sup> 74 Correctly applying such strategies to guide behavior in a given context can increase the rate of reward, and their use has been linked to the integrity of vlPFC<sup>7, 8, 12</sup>. A role for vlPFC in the use 74 Correctly applying such strategies to guide behavior in a given context can increase the rate of<br>75 reward, and their use has been linked to the integrity of vlPFC<sup>7, 8, 12</sup>. A role for vlPFC in the use<br>76 refersed to reward, and their use has been linked to the integrity of vlPFC<sup>7, 8, 12</sup> 76 of such strategies, however, may run counter to the view that vlPFC is essential for learning from<br>77 probabilistic feedback. One possibility is that the role of vlPFC in learning and decision-making 76 of such strategies, however, may run counter to the view that vlPFC is essential for learning from<br>77 probabilistic feedback. One possibility is that the role of vlPFC in learning and decision-making<br>78 vertice dependin probabilistic feedback. One possibility is that the role of vlPFC in learning and decision-making<br>78 varies depending on the learning context, that is, whether the stimulus-reward associations are<br>79 learning on must be le varies depending on the learning context, that is, whether the stimulus-reward associations are<br>79 known or must be learned. vlPFC may achieve this by differentially interacting with other<br>80 contigal and subsertial areas 79 known or must be learned. vlPFC may achieve this by differentially interacting with other cortical and subcortical areas, but the specific circuits are unclear<sup>16-18</sup>. cortical and subcortical areas, but the specific circuits are unclear  $16-18$ . cortical and subcortical areas, but the specific circuits are unclear<sup>16-18</sup>.<br>81 To test the role of vlPFC in learning and the use of strategies to guide behavior in different

81 To test the role of vlPFC in learning and the use of strategies to guide behavior in different<br>82 contexts, we first conducted a functional MRI experiment in awake macaque monkeys, while 82 contexts, we first conducted a functional MRI experiment in awake macaque monkeys, while<br>3 83 they participated in a probabilistic learning task. During the task, subjects completed blocks of<br>84 trials under different learning contexts; either new stimulus-reward associations had to be learned<br>85 (novel context) 85 (novel context) or knowledge about previously-learned associations could be used to guide<br>86 choices (familiar context). In a second experiment, we examined the role of dopamine in 85 (novel context) or knowledge about previously-learned associations could be used to guide<br>86 choices (familiar context). In a second experiment, we examined the role of dopamine in<br>87 influencing habouter career learnin 86 choices (familiar context). In a second experiment, we examined the role of dopamine in<br>87 influencing behavior across learning contexts through systemic injection of selective dopamine<br>88 measures attacamints. Finally, 87 influencing behavior across learning contexts through systemic injection of selective dopamine<br>88 receptor antagonists. Finally, we then conducted anesthetized functional MRI under the same<br>89 pharmacological challenge receptor antagonists. Finally, we then conducted anesthetized functional MRI under the same<br>89 pharmacological challenge and specifically looked at the effect of dopaminergic modulation on<br>80 viDEC estimity. Thus, this ser 89 pharmacological challenge and specifically looked at the effect of dopaminergic modulation on<br>90 vlPFC activity. Thus, this series of experiments allowed us to test how the role of vlPFC, and its<br>91 interactions with ce 90 vlPFC activity. Thus, this series of experiments allowed us to test how the role of vlPFC, and its<br>91 interactions with cortical and subcortical areas, varies depending on the contextual modulation of<br>92 learning and st 91 interactions with cortical and subcortical areas, varies depending on the contextual modulation of<br>92 learning and strategy use. The data indicate that vlPFC-linked pathways make distinct<br>92 contributions to decision mo 92 learning and strategy use. The data indicate that vlPFC-linked pathways make distinct<br>93 contributions to decision-making under different learning contexts. 93 contributions to decision-making under different learning contexts.<br>94

### 95

# 95 **Results**<br>96 Animals exhibit context-dependent behavioral adaptation in a probabilistic learning task

97 Monkeys ( $N = 4$ ) performed a probabilistic learning task for fluid rewards while they underwent 97 Monkeys ( $N = 4$ ) performed a probabilistic learning task for fluid rewards while they underwent<br>98 whole-brain functional neuroimaging (**Fig. 1A-C**). On each trial, subjects chose between two<br>99 viewel stimuli presente whole-brain functional neuroimaging (**Fig. 1A-C**). On each trial, subjects chose between two<br>99 visual stimuli presented on a monitor that were randomly selected from a larger pool of three<br>99 stimuli Each stimulus was ass 99 visual stimuli presented on a monitor that were randomly selected from a larger pool of three<br>100 stimuli. Each stimulus was associated with either 0.9, 0.5, and 0.3 probability of receiving a<br>101 reward. Subjects compl 100 stimuli. Each stimulus was associated with either 0.9, 0.5, and 0.3 probability of receiving a<br>101 reward. Subjects completed trials in two different task contexts or blocks. In novel blocks, visual<br>102 retimuli that w 101 reward. Subjects completed trials in two different task contexts or blocks. In novel blocks, visual<br>102 stimuli that subjects had never seen before were presented, whereas in familiar blocks stimuli<br>102 sthet subjects 102 stimuli that subjects had never seen before were presented, whereas in familiar blocks stimuli<br>103 that subjects had previously learned about were presented. Novel and familiar blocks were<br>104 pseudorandomly intermingl that subjects had previously learned about were presented. Novel and familiar blocks were<br>104 pseudorandomly intermingled in each session and neural and behavioral data were collected and<br>105 analyzed in an event related m 104 pseudorandomly intermingled in each session and neural and behavioral data were collected and<br>105 analyzed in an event-related manner (**Fig. 1D**). 105 analyzed in an event-related manner (**Fig. 1D**).<br>106 Subjects demonstrated distinct patterns of behavior across novel and familiar blocks. In novel

106 Subjects demonstrated distinct patterns of behavior across novel and familiar blocks. In novel<br>107 blocks, subjects' performance gradually improved as animals learned which option was<br>108 accordisted with the highest p 107 blocks, subjects' performance gradually improved as animals learned which option was<br>108 associated with the highest probability of reward, whereas in familiar blocks performance was<br>109 associated with the highest lev 108 associated with the highest probability of reward, whereas in familiar blocks performance was<br>109 consistently at a high level (**Fig. 1E, F**). We split each block into equal 25-trial bins and found<br>110 that performance 109 consistently at a high level (**Fig. 1E, F**). We split each block into equal 25-trial bins and found<br>110 that performance in early trial bins was different between novel and familiar blocks, whereas<br>111 later hine was n 110 that performance in early trial bins was different between novel and familiar blocks, whereas<br>111 later bins were not  $(2$ -way repeated-measures ANOVA, interaction of block type by trial bin,  $p <$ 111 later bins were not (2-way repeated-measures ANOVA, interaction of block type by trial bin,  $p <$ 4 112 0.01,  $F_{(1,178)} = 18.1$ ). This indicates that the animals successfully learned new stimulus-reward<br>113 associations in novel blocks while they maintained high performance across familiar blocks.<br>114 Within both learni 114 Within both learning contexts (i.e., novel and familiar blocks) we found that correct performance<br>115 in later bins reflected the relative reward probability associated with stimuli available on each 115 in later bins reflected the relative reward probability associated with stimuli available on each trial (2-way repeated-measures ANOVA, main effect of stimulus pair,  $p < 0.01$ ,  $F_{(2,267)} = 12.6$ ) 116 trial (2-way repeated-measures ANOVA, main effect of stimulus pair,  $p < 0.01$ ,  $F_{(2,267)} = 12.6$ )<br>117 **(Fig. 1G, H)**. This meant that the macaques were not always choosing the available option with 117 **(Fig. 1G, H)**. This meant that the macaques were not always choosing the available option with the highest probability of reward but were distributing the frequency of their choices to match 118 the highest probability of reward but were distributing the frequency of their choices to match<br>119 the relative option value. The response time (RT) also reflected the relative reward probability at 119 the relative option value. The response time (RT) also reflected the relative reward probability at<br>120 each stimulus pair in both block types (2-way repeated-measures ANOVA, main effect of 120 each stimulus pair in both block types (2-way repeated-measures ANOVA, main effect of stimulus pair,  $p = 0.016$ ,  $F_{(2,267)} = 4.2$ ) (**Supplementary Fig. 1**). Such a pattern of responding is 121 stimulus pair,  $p = 0.016$ ,  $F_{(2,267)} = 4.2$ ) (**Supplementary Fig. 1**). Such a pattern of responding is consistent with matching, a behavior whereby subjects distribute their responding to the 121 stimulus pair, p = 0.016,  $F_{(2,267)} = 4.2$ ) (**Supplementary Fig. 1**). Such a pattern of responding is consistent with matching, a behavior whereby subjects distribute their responding to the available ortions<sup>19, 20</sup> 122 consistent with matching, a behavior whereby subjects distribute their responding to the available options<sup>19, 20</sup>. available options $19, 20$ .

123 available options<sup>19, 20</sup>.<br>124 Given that subjects  $\epsilon$ 124 Given that subjects exhibited aspects of matching behavior, which takes into account the<br>125 outcome of the previous trial, we next looked at subjects' use of reward delivery-based<br>126 behavioral strategies. Here we fo 125 outcome of the previous trial, we next looked at subjects' use of reward delivery-based<br>126 behavioral strategies. Here we found that animals demonstrated distinct behavioral strategies<br>127 depending on the learning co 127 depending on the learning context. Overall, they tended to switch their choices more frequently<br>128 following a 'loss' (unrewarded) trial compared to a 'win' (rewarded) trial, manifesting a win-128 following a 'loss' (unrewarded) trial compared to a 'win' (rewarded) trial, manifesting a win-<br>129 stay/lose-shift (WSLS) pattern (Fig. 11). This tendency was more pronounced in familiar blocks 128 following a 'loss' (unrewarded) trial compared to a 'win' (rewarded) trial, manifesting a win-<br>129 stay/lose-shift (WSLS) pattern (**Fig. 1I**). This tendency was more pronounced in familiar blocks<br>120 step novel blocks 129 stay/lose-shift (WSLS) pattern (**Fig. 1I**). This tendency was more pronounced in familiar blocks<br>130 than novel blocks (2-way repeated-measures ANOVA, interaction of outcome by block type,<br>121 E = 47.8 n < 0.01) indica 130 than novel blocks (2-way repeated-measures ANOVA, interaction of outcome by block type,<br>131  $F_{(1,178)} = 47.8$ ,  $p < 0.01$ ), indicating that the monkeys were more likely to apply WSLS strategy<br>132 when learning was not 132 when learning was not required. Specifically, the proportion of WSLS trials was at chance in the early phase of novel blocks, but gradually increased toward the end of blocks, while it was 133 early phase of novel blocks, but gradually increased toward the end of blocks, while it was not require maintained at a high level throughout familiar blocks (2-way repeated-measures ANOVA, 134 maintained at a high level throughout familiar blocks (2-way repeated-measures ANOVA,<br>135 interaction of trial bin by block type,  $F_{(3,350)} = 4.0$ ,  $p < 0.01$ ) (**Fig. 1J**). In addition, the proportion maintained at a high level throughout familiar blocks (2-way repeated-measures ANOVA,<br>135 interaction of trial bin by block type,  $F_{(3,356)} = 4.0$ ,  $p < 0.01$ ) (**Fig. 1J**). In addition, the proportion<br>126 of WSLS trials wa 135 interaction of trial bin by block type,  $F_{(3,356)} = 4.0$ ,  $p < 0.01$ ) (**Fig. 1J**). In addition, the proportion<br>136 of WSLS trials was positively correlated with choice performance in both novel and familiar<br>127 blocks 136 of WSLS trials was positively correlated with choice performance in both novel and familiar<br>137 blocks (Pearson's correlation,  $n = 55$  and 42 for novel and familiar blocks, respectively,  $p < 0.01$ )<br>128 (Fig. 1K) sugge 137 blocks (Pearson's correlation,  $n = 55$  and 42 for novel and familiar blocks, respectively,  $p < 0.01$ )<br>138 (Fig. 1K), suggesting that the use of these strategies based on the learning context was<br>120 educates for test 138 (**Fig. 1K**), suggesting that the use of these strategies based on the learning context was<br>139 advantageous for task performance. Taken together, these analyses show that the subjects<br>140 adentively used behavioral str 139 advantageous for task performance. Taken together, these analyses show that the subjects<br>140 adaptively used behavioral strategies to improve their task performance across the different<br>141 alsoming contexts 140 adaptively used behavioral strategies to improve their task performance across the different<br>141 learning contexts. learning contexts.



144 **Figure 1. Probabilistic learning task and behaviors.** (A) Trial sequence in a probabilistic 145 learning task On each trial animals make a choice between two visual stimuli by eye movement 145 learning task. On each trial, animals make a choice between two visual stimuli by eye movement<br>146 to earn juice reward (R) Stimulus sets in novel and familiar blocks. Each stimulus is associated 146 to earn juice reward. (**B**) Stimulus sets in novel and familiar blocks. Each stimulus is associated 147 with a reward probability of 0.9, 0.5, or 0.3. Different set of stimuli (Set A or B) are used by 147 with a reward probability of 0.9, 0.5, or 0.3. Different set of stimuli (Set A or B) are used by  $148$  subject in familiar blocks (C) Awake-fMRI setup. Subjects are placed in the sphyny position in 148 subject in familiar blocks. (**C**) Awake-fMRI setup. Subjects are placed in the sphynx position in 149 the 3T MRI scanner in front of a display screen with an eve-tracking system allowing them to 149 the 3T MRI scanner in front of a display screen with an eye-tracking system, allowing them to  $150$  perform tasks during functional scans (D) Analysis pineline. Neural and behavioral data are 150 perform tasks during functional scans. (**D**) Analysis pipeline. Neural and behavioral data are<br>151 collected simultaneously and separately preprocessed offline for subsequent event-related 151 collected simultaneously and separately preprocessed offline for subsequent event-related<br>152 analyses (E. F.) Choice performance in novel blocks (F.) and familiar blocks (F.) Average and 152 analyses. (**E, F**) Choice performance in novel blocks (E) and familiar blocks (F). Average and<br>153 SEM of choice performance (proportion of high-value option choice) of all monkeys ( $N - 4$ ) are 153 SEM of choice performance (proportion of high-value option choice) of all monkeys ( $N = 4$ ) are<br>154 polotted. Asterisk indicates significant interaction of trial bin by block type (\*\*p < 0.01 2-way 154 plotted. Asterisk indicates significant interaction of trial bin by block type  $(*p < 0.01, 2$ -way<br>155 prepeated-measures ANOVA). Dotted line indicates chance level. Green lines are individual 155 repeated-measures ANOVA). Dotted line indicates chance level. Green lines are individual 156 reformance  $(G, H)$  Performance for each stimulus pair in novel blocks  $(G)$  and familiar blocks 156 performance. (**G, H**) Performance for each stimulus pair in novel blocks (G) and familiar blocks 157 (H) Plots indicate performance in binned trials (left) where colors represent stimulus pair. Bar 157 (H). Plots indicate performance in binned trials (left) where colors represent stimulus pair. Bar

graph (right) indicates average performance for each stimulus pair in  $4<sup>th</sup>$  quartile. Asterisks 158 graph (right) indicates average performance for each stimulus pair in  $4^{\text{th}}$  quartile. Asterisks<br>159 indicate significant main effect of stimulus pair (\*\*p < 0.01, 2-way repeated-measures ANOVA).<br>160 Blue dotted li 160 Blue dotted line on the bar graph indicates the relative probability of a higher value option in<br>161 each pair. Symbols represent individual animals. (I) Proportion of switching choices. Bars<br>162 indicate average and S 160 Blue dotted line on the bar graph indicates the relative probability of a higher value option in<br>161 each pair. Symbols represent individual animals. (I) Proportion of switching choices. Bars<br>162 indicate average and S each pair. Symbols represent individual animals. (**I**) Proportion of switching choices. Bars<br>162 indicate average and SEM of switching probability for post-win trials and post-loss trials in<br>163 novel and familiar blocks, 163 novel and familiar blocks, respectively. Symbols represent each animal. Asterisk indicates significant interaction of block type by reward outcome  $(*p < 0.01, 2$ -way repeated-measures ANOVA). (J) Proportion of win-stay 163 novel and familiar blocks, respectively. Symbols represent each animal. Asterisk indicates<br>
164 significant interaction of block type by reward outcome  $(*p < 0.01, 2$ -way repeated-measures<br>
165 ANOVA). (J) Proportion o 164 significant interaction of block type by reward outcome (\*\*p < 0.01, 2-way repeated-measures<br>165 ANOVA). (J) Proportion of win-stay/lose-shift choices for novel (red) and familiar (blue) blocks<br>166 in each quartile bl 166 in each quartile block (average and SEM). Asterisks indicate significant interaction of trial bin by<br>167 block type  $(*\gamma > 0.01, 2$ -way repeated-measures ANOVA). (**K**) Correlation between the<br>168 proportion of WSLS and 167 block type (\*\*p < 0.01, 2-way repeated-measures ANOVA). (K) Correlation between the proportion of WSLS and choice performance in novel (left) and familiar (right) blocks. Each dot represents individual blocks and line 167 block type (\*\*p < 0.01, 2-way repeated-measures ANOVA). (**K**) Correlation between the proportion of WSLS and choice performance in novel (left) and familiar (right) blocks. Each dot represents individual blocks and lin 168 proportion of WSLS and choice performance in novel (left) and familiar (right) blocks. Each dot<br>169 represents individual blocks and lines indicate linear fitting of the data.<br>170 169 represents individual blocks and lines indicate linear fitting of the data.<br>170

## 171

171 *Whole-brain encoding of outcome and learning context*<br>172 Our behavioral analysis demonstrated that subjects adjusted their behavioral strategies between 173 learning contexts, altering their decisions to stay or shift from previous choices depending upon<br>174 the outcome. Consequently, we next set out to determine the network of brain areas that 173 learning contexts, altering their decisions to stay or shift from previous choices depending upon<br>174 the outcome. Consequently, we next set out to determine the network of brain areas that<br>175 exhibited neural estimit 174 the outcome. Consequently, we next set out to determine the network of brain areas that<br>175 exhibited neural activity associated with the task. First, we analyzed whole brain functional<br>176 examples also called from wh exhibited neural activity associated with the task. First, we analyzed whole brain functional<br>176 neuroimaging data collected from subjects while they performed the task (Fig. 1C, D), looking<br>177 for signals that were modu 177 for signals that were modulated either by learning context (Fig. 2A) or reward outcome (Fig. 2B).<br>178 These analyses revealed that bilateral lateral and ventral frontal areas as well as the ventral 177 for signals that were modulated either by learning context (**Fig. 2A**) or reward outcome (**Fig. 2B**).<br>178 These analyses revealed that bilateral lateral and ventral frontal areas as well as the ventral<br>179 temperal lab 178 These analyses revealed that bilateral lateral and ventral frontal areas as well as the ventral<br>179 temporal lobes were more active in the novel versus familiar blocks, while medial frontal areas<br>180 showed greater act 179 temporal lobes were more active in the novel versus familiar blocks, while medial frontal areas<br>180 showed greater activity in the familiar blocks (Fig. 2A). By contrast, reward receipt was<br>181 associated with increase 180 showed greater activity in the familiar blocks (**Fig. 2A**). By contrast, reward receipt was<br>181 associated with increased activity in ventrolateral frontal cortex as well as parts of sensorimotor<br>182 as external untral 181 associated with increased activity in ventrolateral frontal cortex as well as parts of sensorimotor<br>182 cortex and ventral striatum, and decreased activity in dorsolateral prefrontal cortex (**Fig. 2B**). 182 cortex and ventral striatum, and decreased activity in dorsolateral prefrontal cortex (**Fig. 2B**).<br>183 Next, we conducted a conjunction analysis looking for areas showing activations that varied

184 based on context and reward outcome. This analysis revealed clusters that encoded both context<br>185 and reward outcome, in a distinct network of areas including vIPFC, dorsal anterior cingulate based on context and reward outcome. This analysis revealed clusters that encoded both context<br>185 and reward outcome, in a distinct network of areas including vIPFC, dorsal anterior cingulate<br>186 acrts: (AACC), dereglater 2185 and reward outcome, in a distinct network of areas including vlPFC, dorsal anterior cingulate<br>186 cortex (dACC), dorsolateral PFC, supplemental motor area (SMA), and inferior temporal cortex<br>187 (TE) (2 year ANOVA, ma 186 cortex (dACC), dorsolateral PFC, supplemental motor area (SMA), and inferior temporal cortex<br>187 (TE) (2-way ANOVA, main effect of block type or outcome,  $p < 0.05$  with cluster-correction, **Fig.**<br>188 **2C**). We then pr 188 (2C). We then projected the effect of the novel versus familiar comparison back onto the areas<br>189 that showed an interaction between the effects of context and reward outcome to visualize the 189 that showed an interaction between the effects of context and reward outcome to visualize the strength of context encoding (**Fig. 2D**). This analysis showed that the vlPFC and dACC, the 190 strength of context encoding (**Fig. 2D**). This analysis showed that the vlPFC and dACC, the areas previously highlighted based on their potential role in probabilistic learning<sup>18, 21</sup>, are indeed 190 strength of context encoding (**Fig. 2D**). This analysis showed that the vlPFC and dACC, the areas previously highlighted based on their potential role in probabilistic learning<sup>18, 21</sup>, are indeed  $\overline{a}$ areas previously highlighted based on their potential role in probabilistic learning<sup>18, 21</sup>, are indeed<br> $\frac{7}{100}$ 

192 associated with learning context and reward, and respond more strongly in the novel learning 193 context compared to the familiar context. Taken together, these whole-brain analyses suggest that 194 the activity in vlPFC varies based on the ongoing learning context and the preceding outcome. A 195 full table of statistically significant clusters for this analysis can be found in **Supplementary** <sup>196</sup>**Table 1**.





199 **Figure 2. Whole-brain representations of learning context and outcome.** (A) Whole-brain<br>200 **Frances** representations of learning context. Coronal slices (2.5 mm apart) are shown from anterior (top) 200 representations of learning context. Coronal slices  $(2.5 \text{ mm apart})$  are shown from anterior (top 201 left) to posterior (bottom right) planes. Thresholded E-stat maps (p < 0.05 cluster-corrected) are 201 left) to posterior (bottom right) planes. Thresholded F-stat maps ( $p < 0.05$ , cluster-corrected) are<br>202 superimposed on a standard anatomical template. Positive and negative E-stats (warmer and 202 superimposed on a standard anatomical template. Positive and negative F-stats (warmer and 203 cooler colors) indicate more activity in novel blocks and in familiar blocks respectively (R) 203 cooler colors) indicate more activity in novel blocks and in familiar blocks, respectively. (**B**)<br>204 Whole-brain representations of reward outcome Larger E-stats indicate more activity in 204 Whole-brain representations of reward outcome. Larger F-stats indicate more activity in  $205$  rewarded than no reward trials. Data are displayed in the same manner as  $(4)$  (C) Conjunction 205 rewarded than no reward trials. Data are displayed in the same manner as  $(A)$ .  $(C)$  Conjunction

206 analysis result. Clusters highlighted (yellow) significantly encoded both learning context (novel vs. familiar) and reward outcome (rewarded vs. no reward) at cluster-level correction ( $p < 0.05$ ).<br>208 **(D)** F-stats ma 207 vs. familiar) and reward outcome (rewarded vs. no reward) at cluster-level correction (p < 0.05).<br>
208 (D) F-stats map of context coding (novel vs. familiar; A) masked for the clusters identified in the<br>
209 conjuncti 208 (D) F-stats map of context coding (novel vs. familiar; A) masked for the clusters identified in the conjunction analysis (C). dlPFC: dorsolateral prefrontal cortex, dACC: dorsal anterior cingulate cortex, pre-SMA: pre-209 conjunction analysis (C). dlPFC: dorsolateral prefrontal cortex, dACC: dorsal anterior cingulate<br>210 cortex, pre-SMA: pre-supplementary motor area, vlPFC: ventrolateral prefrontal cortex, AIns:<br>211 anterior insula, TE: 210 cortex, pre-SMA: pre-supplementary motor area, vlPFC: ventrolateral prefrontal cortex, AIns:<br>211 anterior insula, TE: inferior temporal cortex.<br>212 211 anterior insula, TE: inferior temporal cortex.<br>212

## 212

213 *vlPFC activity tracks outcome and behavioral strategy*  214 Previous work has emphasized the critical role of neural activity in vlPFC in probabilistic<br>215 learning<sup>7, 8, 12</sup>. Here we found that across all subjects, activity in bilateral vlPFC consistently<br>216 versied with lear learning<sup>7, 8, 12</sup> 215 learning<sup>7, 8, 12</sup>. Here we found that across all subjects, activity in bilateral vlPFC consistently<br>216 varied with learning context (novel > familiar) and this effect was most clearly differentiated in<br>217 the right 217 the right vIPFC (**Fig. 2** and **Supplementary Fig. 2**). Consequently, we chose right vIPFC as a<br>218 region of interest (ROI) for further analyses (**Fig. 3A**). In each novel and familiar block, the the right vlPFC (**Fig. 2** and **Supplementary Fig. 2**). Consequently, we chose right vlPFC as a<br>218 region of interest (ROI) for further analyses (**Fig. 3A**). In each novel and familiar block, the<br>219 regional in vlPEC worl region of interest (ROI) for further analyses (**Fig. 3A**). In each novel and familiar block, the<br>219 signal in vlPFC varied depending on the trial-by-trial outcome and whether subjects<br>220 subsequently stayed with their pr 219 signal in vlPFC varied depending on the trial-by-trial outcome and whether subjects<br>220 subsequently stayed with their prior choice or shifted to a different option (Fig. 3B). The average<br>221 signal in this area around 220 subsequently stayed with their prior choice or shifted to a different option (**Fig. 3B**). The average signal in this area around reward delivery (0-4 s after outcome) was marginally and negatively are semi-leader with 221 signal in this area around reward delivery (0-4 s after outcome) was marginally and negatively<br>222 correlated with task performance in novel blocks (Pearson's correlation,  $n = 51$ ,  $p = 0.056$ ), while<br>222 there was no 222 correlated with task performance in novel blocks (Pearson's correlation,  $n = 51$ ,  $p = 0.056$ ), while<br>223 there was no consistent relationship between performance and vlPFC activity in familiar blocks<br>224 ( $n = 42$ ,  $p =$ 224 (n = 42, p = 0.57) (**Fig. 3C**). Such a pattern of effects potentially indicates that activity in vlPFC is higher during explorative behavioral adaptation when the macaques are learning new stimulus-224 (n = 42, p = 0.57) (**Fig. 3C**). Such a pattern of effects potentially indicates that activity in vlPFC is higher during explorative behavioral adaptation when the macaques are learning new stimulus-<br>226 merced associa 225 is higher during explorative behavioral adaptation when the macaques are learning new stimulus-<br>226 reward associations. The 226 reward associations.<br>227 To more formally assess this relationship, we compared activity in vIPFC between learning

228 contexts. A multiple-regression analysis was performed on the ROI time-series for each block<br>229 type. To investigate the effects of different factors on signals in vlPFC, this analysis included the 229 type. To investigate the effects of different factors on signals in vlPFC, this analysis included the<br>230 following factors as regressors: reward outcome of the present trial, stay/shift decision in a 220 following factors as regressors: reward outcome of the present trial, stay/shift decision in a<br>231 subsequent trial, and the interaction between these two, i.e. whether subjects were using a win-230 following factors as regressors: reward outcome of the present trial, stay/shift decision in a<br>231 subsequent trial, and the interaction between these two, i.e. whether subjects were using a win-<br>222 stay/less shift (W 231 subsequent trial, and the interaction between these two, i.e. whether subjects were using a win-<br>232 stay/lose-shift (WSLS) strategy. This analysis revealed that vlPFC encoded whether reward was 232 stay/lose-shift (WSLS) strategy. This analysis revealed that vlPFC encoded whether reward was<br>233 delivered or not in a similar manner across both novel and familiar blocks (**Fig. 3D, top panels**).<br>234 By contrast, onl 234 By contrast, only activity within vlPFC during novel blocks was related to subjects' decision to<br>235 stay or shift and their use of WSLS strategies (compare left and right side of **Fig. 3D, middle** 234 By contrast, only activity within vlPFC during novel blocks was related to subjects' decision to<br>235 stay or shift and their use of WSLS strategies (compare left and right side of **Fig. 3D, middle**<br>226 and hetter panel 235 stay or shift and their use of WSLS strategies (compare left and right side of **Fig. 3D, middle**<br>236 **and bottom panels**). Thus, when subjects are actively learning stimulus-reward associations in 236 **and bottom panels**). Thus, when subjects are actively learning stimulus-reward associations in<br>9

237 novel blocks, activity within vlPFC is driven by both reward delivery as well as the behavioral<br>238 strategy that the subjects are using.

239 To explore the time course of the differences in vIPFC signal between the novel and familiar 240 blocks, we conducted a multidimensional analysis of activity related to reward outcome and<br>241 whether subjects chose to stay or shift their behavior. To do this we projected the regression 241 whether subjects chose to stay or shift their behavior. To do this we projected the regression<br>242 coefficients (beta values) from these two variables from 4 seconds before to 8 seconds after the whether subjects chose to stay or shift their behavior. To do this we projected the regression<br>242 coefficients (beta values) from these two variables from 4 seconds before to 8 seconds after the<br>242 containing anta a 2.D coefficients (beta values) from these two variables from 4 seconds before to 8 seconds after the<br>243 outcome onto a 2-D space for each block type (**Fig. 3E**). We then measured the Euclidean<br>244 distance between the project 243 outcome onto a 2-D space for each block type (**Fig. 3E**). We then measured the Euclidean<br>244 distance between the projected beta values from the novel and familiar blocks at each time point,<br>245 as a proxy of neural re 245 as a proxy of neural representational difference between the two contexts, that were plotted against the time relative to the outcome (Fig. 3F). This analysis revealed that the neural encoding 246 against the time relative to the outcome (Fig. 3F). This analysis revealed that the neural encoding<br>247 of reward outcome and decisions to stay or shift in vlPFC in the two contexts most prominently 246 against the time relative to the outcome (**Fig. 3F**). This analysis revealed that the neural encoding<br>247 of reward outcome and decisions to stay or shift in vlPFC in the two contexts most prominently<br>248 diversed arou 247 of reward outcome and decisions to stay or shift in vlPFC in the two contexts most prominently<br>248 diverged around the timing of the outcome (permutation test,  $p < 0.05$ ). This indicates that the<br>249 decisitive within 248 diverged around the timing of the outcome (permutation test,  $p < 0.05$ ). This indicates that the activity within vIPFC diverges at the timing when animals are adjusting their use of behavioral 250 strategies depending 249 activity within vlPFC diverges at the timing when animals are adjusting their use of behavioral<br>250 strategies depending on the context. 250 strategies depending on the context.



252

253 **Figure 3. vlPFC signal encodes behavioral strategy during learning.** (A) vlPFC ROI for<br>254 time-series analysis. The man of E-stats of context coding (novel ys. familiar) are shown on 254 time-series analysis. The map of F-stats of context coding (novel vs. familiar) are shown on<br>255 coronal (left) and sagittal (right) planes of a standard anatomical template. A spherical ROI is 255 coronal (left) and sagittal (right) planes of a standard anatomical template. A spherical ROI is<br>256 defined based on the peak coordinates of context coding in the right vIPEC cluster (R) ROI time-256 defined based on the peak coordinates of context coding in the right vlPFC cluster. (**B**) ROI time-<br>257 series around the outcome timing during novel (left) and familiar (right) blocks. Average and 257 series around the outcome timing during novel (left) and familiar (right) blocks. Average and<br>258 SEM of ROL time-series are plotted for win-stay win-shift lose-stay lose-shift trials 258 SEM of ROI time-series are plotted for win-stay, win-shift, lose-stay, lose-shift trials,  $259$  respectively  $(C)$  Correlation between vIPEC activity and choice performance in novel (left) and respectively. (C) Correlation between vlPFC activity and choice performance in novel (left) and<br>260 familiar (right) blocks. Each dot indicates a block and the line indicates a linear fitting of the 260 familiar (right) blocks. Each dot indicates a block and the line indicates a linear fitting of the 261 scatter plot (D) Regression analysis result. Beta coefficients for outcome coding (top) stay/shift 261 scatter plot. **(D)** Regression analysis result. Beta coefficients for outcome coding (top), stay/shift<br>262 decision coding (middle), and the interaction of outcome by stay/shift decision (i.e., WSI.S. 262 decision coding (middle), and the interaction of outcome by stay/shift decision (i.e., WSLS<br>263 behavioral strategy) coding (bottom) were computed using a sliding window analysis. The time-263 behavioral strategy) coding (bottom) were computed using a sliding window analysis. The time-<br>264 course of the beta coefficients were plotted around the timing of outcome (vertical dotted line) 264 course of the beta coefficients were plotted around the timing of outcome (vertical dotted line)<br>265 for each of novel (left) and familiar (right) blocks. Thick lines on the top of each panel indicate 265 for each of novel (left) and familiar (right) blocks. Thick lines on the top of each panel indicate<br>266 significant encoding compared to zero ( $p < 0.05$  at 3 consecutive bins, rank-sum test) (**F**, **F**) 266 significant encoding compared to zero  $(p < 0.05$  at 3 consecutive bins, rank-sum test). (**E, F**)<br>267 Multidimensional analysis result (**E**) Beta coefficients for outcome and stay/shift decision 267 Multidimensional analysis result. (**E**) Beta coefficients for outcome and stay/shift decision<br>268 coding are plotted at each time point of novel (warmer colors) and familiar (cooler colors) blocks 268 coding are plotted at each time point of novel (warmer colors) and familiar (cooler colors) blocks, LS<br>ne-

269 with the passage of time represented as a gradient of colors. The dotted line and squares indicate<br>270 the timing of outcome, and downward arrow and upward arrow indicate the start and end of the<br>271 analysis window (f 271 analysis window (from -4 to 8 seconds after the outcome), respectively. (F) The Euclidian distance between novel and familiar blocks was computed at each time point and plotted against the time. The shaded area indicat 272 distance between novel and familiar blocks was computed at each time point and plotted against<br>273 the time. The shaded area indicates the 95% confidence interval of the shuffled data. The data<br>274 that exceeded the 95 272 distance between novel and familiar blocks was computed at each time point and plotted against<br>273 the time. The shaded area indicates the 95% confidence interval of the shuffled data. The data<br>274 that exceeded the 95 273 the time. The shaded area indicates the 95% confidence interval of the shuffled data. The data<br>274 that exceeded the 95% CI are represented by thick lines.<br>275 274 that exceeded the 95% CI are represented by thick lines.<br>275

## 275

276 *vlPFC-ACC functional connectivity encoded behavioral strategy during learning*<br>277 vlPFC is a hub of the frontal attention network and the salience network<sup>16, 17</sup>. The 277 vlPFC is a hub of the frontal attention network and the salience network<sup>16, 17</sup>. Therefore, we next<br>278 asked how functional networks centered on the vlPFC are associated with reward outcome and<br>279 learning context 278 asked how functional networks centered on the vlPFC are associated with reward outcome and<br>279 learning context using a generalized psycho-physiological interaction (gPPI) analysis<sup>22</sup>. We first<br>280 menned out voyals i learning context using a generalized psycho-physiological interaction (gPPI) analysis<sup>22</sup> 280 mapped out voxels in the brain whose time-series showed interaction with the vlPFC seed time-<br>281 series and reward outcome or context. Based on this, we then identified significant clusters that 281 series and reward outcome or context. Based on this, we then identified significant clusters that<br>282 were modulated by the context that subjects were in (2-way ANOVA, main effect of block type, 282 were modulated by the context that subjects were in (2-way ANOVA, main effect of block type,<br>283 p < 0.05 with cluster-correction) (**Supplementary Fig. 3**). This analysis showed that the 283  $p < 0.05$  with cluster-correction) (**Supplementary Fig. 3**). This analysis showed that the functional connections between vIPFC and ACC, mediodorsal thalamus (MD), dIPFC, and pre-284 functional connections between vlPFC and ACC, mediodorsal thalamus (MD), dlPFC, and pre-<br>285 motor areas were modulated by the learning context. A full table of statistically significant 284 functional connections between vlPFC and ACC, mediodorsal thalamus (MD), dlPFC, and pre-<br>285 motor areas were modulated by the learning context. A full table of statistically significant<br>286 selectors is in Sunnlamenta 285 motor areas were modulated by the learning context. A full table of statistically significant clusters is in **Supplementary Table 2**. 286 clusters is in **Supplementary Table 2**.<br>287 Among these, we first focused on the vIPFC-ACC functional connection (**Fig. 4A**), as both

287 Among these, we first focused on the vlPFC-ACC functional connection (**Fig. 4A**), as both<br>288 vlPFC and ACC are implicated in adaptive behavior<sup>18</sup> and are known to be anatomically and<br>280 functionally connected<sup>16</sup>. H vlPFC and ACC are implicated in adaptive behavior $18$ 288 vIPFC and ACC are implicated in adaptive behavior<sup>18</sup> and are known to be anatomically and<br>289 functionally connected<sup>16</sup>. Here we found that vIPFC-ACC functional connectivity (FC) varied<br>200 with the habeviaral strat functionally connected<sup>16</sup>. Here we found that vlPFC-ACC functional connectivity (FC) varied<br>290 with the behavioral strategies used by the subjects after reward feedback (**Fig. 4B**). Specifically,<br>201 EQ between vlDEC an 290 with the behavioral strategies used by the subjects after reward feedback (**Fig. 4B**). Specifically,<br>291 FC between vIPFC and ACC increased around the time of outcome (rank-sum test,  $p < 0.05$  at 3<br>292 consecutive bin 292 consecutive bins) when the animals received reward and repeated the same choice (i.e., win-stay)<br>293 or when the animals received no reward and subsequently changed their choice (i.e., lose-shift), 293 or when the animals received no reward and subsequently changed their choice (i.e., lose-shift),<br>294 representing a WSLS pattern in novel blocks (Fig. 4B, left panels). By contrast, in familiar 293 or when the animals received no reward and subsequently changed their choice (i.e., lose-shift),<br>294 representing a WSLS pattern in novel blocks (**Fig. 4B, left panels**). By contrast, in familiar<br>205 blocks, the EG sha 294 representing a WSLS pattern in novel blocks (**Fig. 4B, left panels**). By contrast, in familiar<br>295 blocks, the FC changes did not follow the WSLS pattern although some significant modulation<br>206 was observed mainly bef 295 blocks, the FC changes did not follow the WSLS pattern although some significant modulation<br>296 was observed mainly before outcome period (**Fig. 4B, right panels**). Three-way ANOVA<br>297 confirmed these effects as there 297 confirmed these effects as there was a significant interaction of block type by stay/shift decision<br>298 and reward outcome, indicating that FC changes reflected WSLS pattern exclusively in novel 298 and reward outcome, indicating that FC changes reflected WSLS pattern exclusively in novel<br>299 blocks (**Supplementary Fig. 4**;  $F_{(1,368)} = 4.0$ ,  $p = 0.046$ ). This pattern of effects indicates that 299 blocks (**Supplementary Fig. 4**;  $F_{(1,368)} = 4.0$ ,  $p = 0.046$ ). This pattern of effects indicates that such connectivity was related to task performance, and indeed the functional connectivity 299 blocks (**Supplementary Fig. 4**;  $F_{(1,368)} = 4.0$ ,  $p = 0.046$ ). This pattern of effects indicates that such connectivity was related to task performance, and indeed the functional connectivity 300 such connectivity was related to task performance, and indeed the functional connectivity<br>12

301 between vlPFC and ACC in loss trials was marginally and negatively correlated with task 302 performance in the novel context (Pearson's correlation, ,  $n = 51$ ,  $r = -0.26$ ,  $p = 0.056$ ). No such 303 correlation was observed in the familiar blocks (n = 42, r = -0.061, p = 0.70) or for performance<br>304 and FC on win trials (p > 0.30) (**Fig. 4C**). A sliding-window regression analysis showed that the 304 and FC on win trials ( $p > 0.30$ ) (**Fig. 4C**). A sliding-window regression analysis showed that the 305 FC between vIPFC and ACC was associated with the WSLS strategy around the time of outcome 306 in novel blocks (permutation test,  $p < 0.05$ ), while they were anti-correlated around the reward 307 timing in the familiar context (**Fig. 4D**). This further suggests that the functional interaction 308 between vlPFC and ACC is related to the context-dependent use of behavioral strategies. ce<br>he



309

310 **Figure 4. vlPFC-ACC functional connection encodes behavioral strategy during learning.<br>311 (A) vlPEC seed and ACC ROL for functional connectivity analysis. ACC ROL (sagittal plane on** 311 (A) vlPFC seed and ACC ROI for functional connectivity analysis. ACC ROI (sagittal plane on<br>312 the right) was defined based on generalized PPI analysis using right vlPFC seed (R) FC time 312 the right) was defined based on generalized PPI analysis using right vlPFC seed. (**B**) FC time<br>313 course around the outcome timing vlPFC-ACC FC during novel (left) and familiar (right) blocks 313 course around the outcome timing. vlPFC-ACC FC during novel (left) and familiar (right) blocks<br>314 were computed using sliding window analysis and visualized for win-stay and win-shift trials 314 were computed using sliding window analysis and visualized for win-stay and win-shift trials<br>315 (top) and lose-stay and lose-shift trials (bottom) separately. The plots are made around the 315 (top) and lose-stay and lose-shift trials (bottom) separately. The plots are made around the<br>316 outcome timing (vertical dotted lines). The thick lines on the top of each panel indicate 316 outcome timing (vertical dotted lines). The thick lines on the top of each panel indicate<br>317 significant EC compared to zero for color matched trials (p < 0.05 with rank-sum test at 3 317 significant FC compared to zero for color matched trials ( $p < 0.05$  with rank-sum test at 3

318 consecutive bins). (**C**) Correlation between vlPFC-ACC FC and choice performance. The correlations were computed for win (rewarded) trials (top) and loss (unrewarded) trials (bottom), and for novel (left) and familiar 320 and for novel (left) and familiar (right) blocks separately. Each dot represents each block, and the lines are linear fitted to the data. (D) Time course of WSLS coding around the outcome. WSLS coding was computed as t 320 and for novel (left) and familiar (right) blocks separately. Each dot represents each block, and the<br>321 lines are linear fitted to the data. (D) Time course of WSLS coding around the outcome. WSLS<br>322 coding was compu 321 lines are linear fitted to the data. (**D**) Time course of WSLS coding around the outcome. WSLS<br>322 coding was computed as the interaction of outcome by stay/shift decision coding in a sliding<br>323 window multiple-regres 322 coding was computed as the interaction of outcome by stay/shift decision coding in a sliding<br>323 window multiple-regression analysis. Shaded areas (yellow) indicate 95% confidence interval of<br>324 the data, and the thic 323 window multiple-regression analysis. Shaded areas (yellow) indicate 95% confidence interval of the data, and the thick black lines indicate the significance of the data.<br>325 324 the data, and the thick black lines indicate the significance of the data.

## 325 *vlPFC-MD functional connectivity reflects decision to stay with a choice during learning*<br>327 Previous studies have shown that the fronto-thalamo pathway plays a critical role in learning and

327 Previous studies have shown that the fronto-thalamo pathway plays a critical role in learning and<br>328 decision-making<sup>23-25</sup>. Consequently, we next focused on the functional connection between<br>329 viDEG and MD (Fig. 54 decision-making $^{23-25}$ 328 decision-making<sup>23-23</sup>. Consequently, we next focused on the functional connection between vlPFC and MD (Fig. 5A). The functional connectivity between vlPFC and MD specifically increased in trials that were followed b 330 increased in trials that were followed by the repetition of the previous choice (i.e., stay decision)<br>331 regardless of whether reward was delivered or not in novel blocks (**Fig. 5B, left panels**). This 331 regardless of whether reward was delivered or not in novel blocks (Fig. 5B, left panels). This<br>332 pattern was less pronounced in familiar blocks (Fig. 5B, right panels and Supplementary Fig. 332 pattern was less pronounced in familiar blocks (**Fig. 5B, right panels** and **Supplementary Fig.**<br>333 5; three-way ANOVA, interaction of block type by stay/shift decision,  $F_{(1,368)} = 4.4$ ,  $p = 0.036$ ). 332 pattern was less pronounced in familiar blocks (**Fig. 5B, right panels** and **Supplementary Fig.**<br>333 5; three-way ANOVA, interaction of block type by stay/shift decision,  $F_{(1,368)} = 4.4$ ,  $p = 0.036$ ). 333 **5**; three-way ANOVA, interaction of block type by stay/shift decision,  $F_{(1,368)} = 4.4$ , p = 0.036).<br>334 Interestingly, the choice signal was not correlated to task performance in either block (**Fig. 5C**; p<br>325  $\rightarrow$  334 Interestingly, the choice signal was not correlated to task performance in either block (**Fig. 5C**; p<br>335 > 0.44), and behavioral strategy coding was primarily observed in familiar blocks (**Fig. 5D**). This<br>336 result s  $> 0.44$ ), and behavioral strategy coding was primarily observed in familiar blocks (**Fig. 5D**). This result suggests that this functional connection between vlPFC and MD encodes execution of the 227 result suggests that this functional connection between vlPFC and MD encodes execution of the decision to stay or switch *per se* that is not directly linked to the correct performance. decision to stay or switch *per se* that is not directly linked to the correct performance.<br>338 We additionally looked at other vIPFC functional connections that showed significant novel vs

339 familiar coding in the gPPI analysis (**Supplementary Fig. 3**). The functional connection between<br>340 vlPFC and supplementary motor area (SMA) increased in the 'stay' trials during novel blocks in 340 vlPFC and supplementary motor area (SMA) increased in the 'stay' trials during novel blocks in<br>341 a pattern similar to that was observed with vlPFC-MD FC, although the interaction of block type 341 a pattern similar to that was observed with vlPFC-MD FC, although the interaction of block type<br>342 by stay/shift decision was not significant ( $p = 0.30$ ). The functional connection between vlPFC 342 by stay/shift decision was not significant ( $p = 0.30$ ). The functional connection between vlPFC and dlPFC also showed changes depending on stay/shift decision in both block types, but again 343 and dlPFC also showed changes depending on stay/shift decision in both block types, but again<br>344 there was no significant interaction of block type by decision ( $p = 0.51$ ). A lack of clear 344 there was no significant interaction of block type by decision ( $p = 0.51$ ). A lack of clear relationship between connectivity in these pathways and the learning context suggests that they 344 there was no significant interaction of block type by decision ( $p = 0.51$ ). A lack of clear<br>345 relationship between connectivity in these pathways and the learning context suggests that they<br>346 relations associated 345 relationship between connectivity in these pathways and the learning context suggests that they<br>346 might be more associated with different aspects of the learning context such as attention. 346 might be more associated with different aspects of the learning context such as attention.<br>347



<sup>349</sup>**Figure 5. vlPFC-MD functional connection encodes decision to stay during learning.** (**A**) ) 350 vlPFC seed and MD thalamus ROI for functional connectivity analysis. (**B**) FC time course<br>351 around the outcome timing vlPFC-MD FC during novel (left) and familiar (right) blocks were 351 around the outcome timing. vlPFC-MD FC during novel (left) and familiar (right) blocks were<br>352 computed for win-stay win-shift lose-stay and lose-shift trials separately. The thick lines on the 352 computed for win-stay, win-shift, lose-stay, and lose-shift trials separately. The thick lines on the  $\frac{353}{12}$  top of each panel indicate significant EC compared to zero for color matched trials (p < 0.05 with 353 top of each panel indicate significant FC compared to zero for color matched trials (p < 0.05 with  $354$  rank-sum test at 3 consecutive bins) (C) Correlation between vIPEC-MD EC and choice 354 rank-sum test at 3 consecutive bins). (C) Correlation between vlPFC-MD FC and choice<br>355 performance for win trials (top) and loss trials (bottom) separately Dots and lines indicate 355 performance, for win trials (top) and loss trials (bottom) separately. Dots and lines indicate  $356$  blocks and linear fitted line respectively (D) Time course of WSI S coding around the outcome 356 blocks and linear fitted line, respectively. (D) Time course of WSLS coding around the outcome.<br>357 Shaded areas (vellow) indicate 95% confidence interval of the data and the thick black line 357 Shaded areas (yellow) indicate 95% confidence interval of the data, and the thick black line<br>358 indicate the significance of the data 358 indicate the significance of the data.

348

## <sup>360</sup>*Pharmacological manipulation of dopamine receptors affects vlPFC-mediated behavior*

361 The prior analyses indicate that the WSLS strategy is crucial for adaptive behavior depending on 362 the learning context, and this is in line with previous work showing that vlPFC plays a central<br>363 role in probabilistic learning and decision-making<sup>12, 14</sup>. The release of dopamine in frontal cortex 363 role in probabilistic learning and decision-making<sup>12, 14</sup>. The release of dopamine in frontal cortex 364 has been implicated in a variety of cognitive functions relevant to probabilistic learning, such as attention and working memory<sup>26, 27</sup>. Therefore, it is possible that the changes in vIPFC activity attention and working memory<sup>26, 27</sup>. Therefore, it is possible that the changes in vlPFC activity ral<br>tex<br>as<br>ity

366 that were associated with behavioral strategy are mediated by the action of dopamine through<br>367 cortical or subcortical dopamine receptors.

368 To address this, we conducted a pharmacological experiment with selective dopamine receptor 369 antagonists SCH-23390 (D1 antagonist) and haloperidol (D2 antagonist) and assessed their<br>370 effects on the performance of monkeys in the probabilistic learning task. The subject cohort in 370 effects on the performance of monkeys in the probabilistic learning task. The subject cohort in<br>371 this experiment ( $N = 4$ ) partially overlapped with the one used in our awake fMRI experiment 371 this experiment ( $N = 4$ ) partially overlapped with the one used in our awake fMRI experiment (see Table 1), and the effects of the drugs on the proportion of correct choice and reaction times 372 (see **Table 1**), and the effects of the drugs on the proportion of correct choice and reaction times<br>373 were analyzed in our previous paper in relation to resting-state functional connectivity<sup>28</sup>. Here, 372 (see **Table 1**), and the effects of the drugs on the proportion of correct choice and reaction times<br>373 were analyzed in our previous paper in relation to resting-state functional connectivity<sup>28</sup>. Here, were analyzed in our previous paper in relation to resting-state functional connectivity<sup>28</sup> 374 we specifically focused on the effects of systemically administered dopaminergic drugs on<br>375 WSLS behaviors. D2 antagonist haloperidol, but not D1 antagonist SCH-23390 or saline. 375 WSLS behaviors. D2 antagonist haloperidol, but not D1 antagonist SCH-23390 or saline,<br>376 increased the proportion of WSLS responses preferentially in novel blocks (Fig. 6A, B). A two-375 WSLS behaviors. D2 antagonist haloperidol, but not D1 antagonist SCH-23390 or saline,<br>376 increased the proportion of WSLS responses preferentially in novel blocks (Fig. 6A, B). A two-<br>277 way repeated magazines. ANOVA increased the proportion of WSLS responses preferentially in novel blocks (**Fig. 6A, B**). A two-<br>377 way repeated-measures ANOVA demonstrated a significant main effect of drug dose in novel<br>278 blocks but not in familiar b 377 way repeated-measures ANOVA demonstrated a significant main effect of drug dose in novel<br>378 blocks but not in familiar blocks after haloperidol administration (novel:  $F_{(2,440)} = 3.9$ ,  $p = 0.021$ ;<br>379 familiar:  $F_{(2$ 379 familiar:  $F_{(2,396)} = 0.072$ ,  $p = 0.93$ ), while there was no significant effect following SCH-23390 administration in either block type ( $p > 0.46$ ). This result suggests that dopamine D2 receptors 379 familiar:  $F_{(2,396)} = 0.072$ ,  $p = 0.93$ ), while there was no significant effect following SCH-23390 administration in either block type ( $p > 0.46$ ). This result suggests that dopamine D2 receptors 380 administration in either block type ( $p > 0.46$ ). This result suggests that dopamine D2 receptors play a key role in modulating behavioral strategy during learning. 381 play a key role in modulating behavioral strategy during learning.<br>382 Finally, to probe whether dopamine receptor-mediated manipulation of behavioral strategies also

383 influences neural activity in vlPFC, we performed a resting-state fMRI experiment with the same<br>384 dopamine antagonists ( $N = 7$ , see **Table 1**). In our previous study using the same dataset we 384 dopamine antagonists ( $N = 7$ , see **Table 1**). In our previous study using the same dataset we reported that the D1 and D2 receptor manipulation induced brain-wide functional connectivity 385 reported that the D1 and D2 receptor manipulation induced brain-wide functional connectivity<br>386 changes, most notably in the cortico-cortical and fronto-striatal FCs<sup>28</sup>. Here, we specifically 385 reported that the D1 and D2 receptor manipulation induced brain-wide functional connectivity<br>386 changes, most notably in the cortico-cortical and fronto-striatal  $FCs^{28}$ . Here, we specifically<br>287 forward an naural changes, most notably in the cortico-cortical and fronto-striatal  $FCs^{28}$ 387 focused on neural activity in the vlPFC region by analyzing regional homogeneity (ReHo). We chose this analysis approach as ReHo is sensitive to local changes in neural activity<sup>29, 30</sup>. The 387 focused on neural activity in the vlPFC region by analyzing regional homogeneity (ReHo). We<br>388 chose this analysis approach as ReHo is sensitive to local changes in neural activity<sup>29, 30</sup>. The<br>389 denominates drugs chose this analysis approach as ReHo is sensitive to local changes in neural activity<sup>29, 30</sup>. The<br>389 dopaminergic drugs induced different patterns of activity changes (ReHo signal) in bilateral<br>388 . The dopaminergic dr 389 dopaminergic drugs induced different patterns of activity changes (ReHo signal) in bilateral<br>390 vlPFC during resting-state neuroimaging (Fig. 6C). D2 receptor antagonist haloperidol, but not<br>201 DJ receptor antagonist vlPFC during resting-state neuroimaging (**Fig. 6C**). D2 receptor antagonist haloperidol, but not<br>391 D1 receptor antagonist SCH-23390 or saline, increased vlPFC activity (1-way repeated-measures<br>202 ANOVA, main effect of d 391 D1 receptor antagonist SCH-23390 or saline, increased vlPFC activity (1-way repeated-measures<br>392 ANOVA, main effect of drug,  $F_{(2,90)} = 3.4$ ,  $p = 0.036$ ) (**Fig. 6D**). While these data with<br>393 dopaminergic manipulati 393 dopaminergic manipulation were not acquired in the context of task performance, our results<br>394 suggest that the WSLS strategy that is associated with vIPFC activity is dependent on the 393 dopaminergic manipulation were not acquired in the context of task performance, our results<br>394 suggest that the WSLS strategy that is associated with vlPFC activity is dependent on the 394 suggest that the WSLS strategy that is associated with vlPFC activity is dependent on the

- 395 function of dopamine D2 receptors, and that pharmacological manipulations of dopamine via D2
- 396 receptors has a clear impact on the brain circuit that vlPFC is embedded within.



**Figure 6. D2 receptor blocker enhanced vlPFC activity and promoted adaptive behavior.**<br>399 (A) The effect of D1 receptor antagonism on WSLS behavior. The proportion of WSLS trials in 399 (A) The effect of D1 receptor antagonism on WSLS behavior. The proportion of WSLS trials in<br>400 quartile blocks (average and SEM) are plotted for each dose of SCH-23390 (0, 10, 30, 50 ug/kg) 400 quartile blocks (average and SEM) are plotted for each dose of SCH-23390  $(0, 10, 30, 50$  ug/kg)<br>401 for novel (left) and familiar (right) blocks, respectively (B) The effect of D2 receptor antagonism 401 for novel (left) and familiar (right) blocks, respectively. (**B**) The effect of D2 receptor antagonism (haloperidol: 0, 5, 10 ug/kg) on WSLS behavior. Plotted in same manner as (A). Asterisk 402 (haloperidol: 0, 5, 10 ug/kg) on WSLS behavior. Plotted in same manner as (A). Asterisk 403 indicates main effect of drug dose  $(*p < 0.05, 2$ -way ANOVA). (C) Regional homogeneity 403 indicates main effect of drug dose ( $p < 0.05$ , 2-way ANOVA). (**C**) Regional homogeneity<br>404 (ReHo) analysis of resting-state fMRI with pharmacological donamine receptor manipulation 404 (ReHo) analysis of resting-state fMRI with pharmacological dopamine receptor manipulation.<br>405 The clusters with significant ReHo values ( $p < 0.05$  cluster-corrected) are superimposed on a 405 The clusters with significant ReHo values ( $p < 0.05$ , cluster-corrected) are superimposed on a<br>406 coronal image from a standard anatomical template (D) The effect of donamine receptor 406 coronal image from a standard anatomical template. (**D**) The effect of dopamine receptor anti-<br>407 antagonists on ReHo value. Bar graph indicates average and SEM of ReHo value of the voxels in 407 antagonists on ReHo value. Bar graph indicates average and SEM of ReHo value of the voxels in<br>408 the vIPEC ROI for each drug condition with individual data points superimposed (\*p < 0.05.1-408 the vlPFC ROI for each drug condition with individual data points superimposed (\*p < 0.05, 1-<br>409 way ANOVA) 409 way ANOVA).

## <sup>411</sup>**Discussion**

412 We examined the brain-wide mechanisms underlying decision-making in different learning<br>413 contexts in macaques. In a probabilistic learning task animals' behavior was influenced by the<br>414 learning context and by the 414 learning context and by the preceding reward outcome. Specifically, animals used win-stay/lose-<br>415 shift strategies to different degrees depending on whether they were learning new stimulus 114 learning context and by the preceding reward outcome. Specifically, animals used win-stay/lose-<br>115 shift strategies to different degrees depending on whether they were learning new stimulus<br>116 separate associations o 415 shift strategies to different degrees depending on whether they were learning new stimulus<br>416 reward associations or exploiting known associations. When we analyzed the brain-wide neural<br>417 activity vIPEG stood out a 416 reward associations or exploiting known associations. When we analyzed the brain-wide neural<br>417 activity, vlPFC stood out as a key region where both behavioral strategies and reward outcomes<br>418 were encoded. Specific activity, vlPFC stood out as a key region where both behavioral strategies and reward outcomes<br>418 were encoded. Specifically, vlPFC encoded behavioral strategies during novel learning contexts.<br>419 Evastional connectivity were encoded. Specifically, vlPFC encoded behavioral strategies during novel learning contexts.<br>419 Functional connectivity in the pathways between vlPFC-ACC and vlPFC-MD was related to<br>420 distinct expects of the animals' 419 Functional connectivity in the pathways between vlPFC-ACC and vlPFC-MD was related to<br>420 distinct aspects of the animals' decisions that were dependent on the learning context.<br>421 Pharmacological experiments further 421 Pharmacological experiments further revealed that the manipulation of dopamine D2 receptors<br>422 influenced monkeys' behavioral strategy during learning as well as vlPFC neural activity at rest. 422 influenced monkeys' behavioral strategy during learning as well as vlPFC neural activity at rest.<br>423 Taken together, our results suggest a critical role of vlPFC and its associated neural networks in influenced monkeys' behavioral strategy during learning as well as vlPFC neural activity at rest.<br>423 Taken together, our results suggest a critical role of vlPFC and its associated neural networks in<br>424 Learning haborian Taken together, our results suggest a critical role of vlPFC and its associated neural networks in<br>424 adaptive behavior during probabilistic learning.

424 adaptive behavior during probabilistic learning.<br>425 The vlPFC has long been implicated in higher-cognitive function, however, the precise role of this area in learning and decision-making has only recently come into focus. Early lesion studies<br>427 highlighted that damage to this area leads to a deficit in the implementation of high-order 427 highlighted that damage to this area leads to a deficit in the implementation of high-order<br>428 decision-making strategy in non-human primates, especially lose-shift strategies<sup>13, 31, 32</sup>. More highlighted that damage to this area leads to a deficit in the implementation of high-order<br>428 decision-making strategy in non-human primates, especially lose-shift strategies<sup>13, 31, 32</sup>. More decision-making strategy in non-human primates, especially lose-shift strategies<sup>13, 31, 32</sup>. More<br>429 recent studies using chronic or transient lesions have demonstrated a causal link between the<br>420 **Euration of this mai** quared a causal link between the function of this region with associative learning processes in probabilistic settings where the history of reinforcement has to be used<sup>7, 8, 10, 12</sup>. This prior work did not, however, dir 430 function of this region with associative learning processes in probabilistic settings where the<br>431 history of reinforcement has to be used<sup>7, 8, 10, 12</sup>. This prior work did not, however, directly<br>422 compare the rol history of reinforcement has to be used<sup>7, 8, 10, 12</sup> 432 compare the role of vlPFC across multiple learning contexts. To address this question, we<br>433 designed a paradigm where subjects made choices between novel or familiar stimuli in separate 433 designed a paradigm where subjects made choices between novel or familiar stimuli in separate<br>434 blocks of trials and analyzed the pattern of choices as well as whole-brain neural activity across designed a paradigm where subjects made choices between novel or familiar stimuli in separate<br>434 blocks of trials and analyzed the pattern of choices as well as whole-brain neural activity across<br>425 different learning co 434 blocks of trials and analyzed the pattern of choices as well as whole-brain neural activity across<br>435 different learning contexts. Our behavioral analysis showed that the animals employed distinct<br>426 behavioral strat different learning contexts. Our behavioral analysis showed that the animals employed distinct<br>436 behavioral strategies depending on the context that they were in. In the blocks of familiar trials,<br>437 WSLS strategies bas 437 WSLS strategies based on monitoring of the preceding reward outcome and altering behavior<br>438 accordingly were more prominent<sup>19, 20</sup>, whereas in the novel context the use of WSLS gradually WSLS strategies based on monitoring of the preceding reward outcome and altering behavior<br>438 accordingly were more prominent<sup>19, 20</sup>, whereas in the novel context the use of WSLS gradually<br>420 increased as learning pregna accordingly were more prominent<sup>19, 20</sup>, whereas in the novel context the use of WSLS gradually<br>439 increased as learning progressed. By comparing whole-brain fMRI signal between novel and<br>440 familiar contexts we found t 439 increased as learning progressed. By comparing whole-brain fMRI signal between novel and<br>440 familiar contexts, we found that vlPFC activity encoded reward outcome in both contexts within<br>441 a similar time course, whi 441 a similar time course, while the same area encoded behavioral strategy preferentially in novel<br>442 contexts. This apparent disconnect between behavior and neural activity within vlPFC is notable 441 a similar time course, while the same area encoded behavioral strategy preferentially in novel<br>442 contexts. This apparent disconnect between behavior and neural activity within vlPFC is notable 442 contexts. This apparent disconnect between behavior and neural activity within vlPFC is notable

443 and may relate to the fact that the vlPFC is contributing to both the learning of stimulus-reward<br>444 associations and behavioral strategies in the novel context.

445 Notably, our vlPFC cluster that co-encoded reward outcome and ongoing learning context was 446 mainly localized to the ventral surface of the frontal cortex, within the areas 12o/l but also<br>447 extending into the anterior part of agranular insular cortex<sup>33</sup> (Fig. 2). The location of these mainly localized to the ventral surface of the frontal cortex, within the areas 12o/l but also<br>
447 extending into the anterior part of agranular insular cortex<sup>33</sup> (Fig. 2). The location of these<br>
448 extinctions were si extending into the anterior part of agranular insular cortex<sup>33</sup> activations was similar to the areas previously reported in neuroimaging studies using associative<br>449 learning tasks in macaques<sup>7, 10</sup>. This notion is also consistent with recent neural recording studies 448 activations was similar to the areas previously reported in neuroimaging studies using associative<br>449 learning tasks in macaques<sup>7, 10</sup>. This notion is also consistent with recent neural recording studies<br>450 that sh learning tasks in macaques<sup>7, 10</sup> 149 learning tasks in macaques<sup>7, 10</sup>. This notion is also consistent with recent neural recording studies<br>150 that showed a substantial reward probability or uncertainty coding in the ventral frontal cortex<sup>9, 11,</sup><br>151 <sup></sup> that showed a substantial reward probability or uncertainty coding in the ventral frontal cortex<sup>9, 11,</sup> 450<br>451<br>452 <sup>14</sup>, and a recent analysis of functional interactions showing a specific role for inputs from 451  $^{14}$ , and a recent analysis of functional interactions showing a specific role for inputs from<br>452 agranular insula to area 12o during feedback processing<sup>34</sup>. Taken together, our study reveals a<br>452 agray role for agranular insula to area 12o during feedback processing<sup>34</sup> agranular insula to area 12o during feedback processing<sup>34</sup>. Taken together, our study reveals a<br>453 new role for these parts of the ventral frontal cortex in adjusting behavior in uncertain<br>454 continuum 453 new role for these parts of the ventral frontal cortex in adjusting behavior in uncertain environments. 454 environments.<br>455 Beyond the vlPFC itself, we found that activity in this part of frontal cortex varied with other

456 parts of the brain during the different learning contexts. The FC between vlPFC and ACC tended<br>457 to encode WSLS behavior when outcomes were delivered in the novel but not familiar blocks of parts of the brain during the different learning contexts. The FC between vlPFC and ACC tended<br>457 to encode WSLS behavior when outcomes were delivered in the novel but not familiar blocks of<br>458 trials. Eurther, the creat to encode WSLS behavior when outcomes were delivered in the novel but not familiar blocks of<br>458 trials. Further, the greatest difference between the connectivity in this pathway between novel<br>459 and familiar blocks assum trials. Further, the greatest difference between the connectivity in this pathway between novel<br>459 and familiar blocks occurred when monkeys decided to switch to a different option after they<br>460 failed to receive a rewar 460 failed to receive a reward (Fig. 4B). Notably, the activity in this pathway was marginally related to better performance, indicating that dynamic interaction between vlPFC and ACC to guide 461 to better performance, indicating that dynamic interaction between vIPFC and ACC to guide<br>462 lose-shift strategies is potentially related to better behavioral performance. A specific role for this to better performance, indicating that dynamic interaction between vlPFC and ACC to guide<br>462 lose-shift strategies is potentially related to better behavioral performance. A specific role for this<br>463 a pathway in changin 162 lose-shift strategies is potentially related to better behavioral performance. A specific role for this<br>163 pathway in changing behaviors after failing to receive a reward agrees with reports that<br>164 expiration legion 463 pathway in changing behaviors after failing to receive a reward agrees with reports that<br>464 aspiration lesions of vlPFC result in a failure to use lose-shift strategies when learning novel<br>465 associations<sup>32</sup>. Furthe 464 aspiration lesions of vlPFC result in a failure to use lose-shift strategies when learning novel<br>465 associations<sup>32</sup>. Further, a number of prior investigations have highlighted a role for ACC in<br>466 diving originals associations<sup>32</sup> 465 associations<sup>32</sup>. Further, a number of prior investigations have highlighted a role for ACC in driving animals to switch to alternative options that are thought to be of higher value<sup>21, 35</sup>. Our driving animals to switch to alternative options that are thought to be of higher value<sup>21, 35</sup> 466 driving animals to switch to alternative options that are thought to be of higher value<sup>21, 35</sup>. Our data suggests that interaction between vlPFC and ACC is essential for guiding choices when the 468 visite of the par data suggests that interaction between vlPFC and ACC is essential for guiding choices when the<br>468 value of the perceived best option drops to a point where it is below the opportunity cost of<br>469 sharen's behaviors 468 value of the perceived best option drops to a point where it is below the opportunity cost of changing behaviors.

470 In contrast to the role of vIPFC-ACC interactions. FC between vIPFC and MD thalamus 471 increased when the subject decided to repeat their choice of a particular stimulus ('stay') even<br>472 when the preceding trial wasn't rewarded in the novel context. Such a pattern suggests that this 471 increased when the subject decided to repeat their choice of a particular stimulus ('stay') even<br>472 when the preceding trial wasn't rewarded in the novel context. Such a pattern suggests that this when the preceding trial wasn't rewarded in the novel context. Such a pattern suggests that this<br>19

173 connection encodes choice *per se* rather than a strategy to facilitate learning performance. MD<br>174 has been implicated in probabilistic learning<sup>23, 24, 36</sup>, and prior reports from lesion studies in<br>175 measures have has been implicated in probabilistic learning<sup>23, 24, 36</sup> 474 has been implicated in probabilistic learning<sup>23, 24, 36</sup>, and prior reports from lesion studies in<br>475 macaques have highlighted that MD is essential for promoting decision to stay with a particular<br>476 course of act macaques have highlighted that MD is essential for promoting decision to stay with a particular<br>476 course of action during learning<sup>37</sup>. Thus, our finding that there was higher functional connectivity<br>477 between vine con course of action during learning<sup>37</sup> course of action during learning<sup>37</sup>. Thus, our finding that there was higher functional connectivity<br>477 between vIPFC and MD on stay trials appears to indicate that such lesion effects are in part<br>478 across by disconnec 477 between vlPFC and MD on stay trials appears to indicate that such lesion effects are in part caused by disconnecting this area from vlPFC. 478 caused by disconnecting this area from vlPFC.<br>479 Our results highlight a set of circuits centered on vlPFC that coordinate the flexible adjustment

480 of behaviors in different learning contexts. It is reasonable to ask how and where the information<br>481 regarding outcome and learning context converge and transform into a behavioral strategy that 481 regarding outcome and learning context converge and transform into a behavioral strategy that<br>482 leads to a decision; addressing this question will require additional experiments using paired 482 leads to a decision; addressing this question will require additional experiments using paired<br>483 neurophysiology recordings and/or causal interrogation of specific neural circuits using viral 182 leads to a decision; addressing this question will require additional experiments using paired<br>183 neurophysiology recordings and/or causal interrogation of specific neural circuits using viral<br>184 toologisms<sup>38</sup> neurophysiology recordings and/or causal interrogation of specific neural circuits using viral<br>484 techniques<sup>38</sup>. techniques<sup>38</sup>.

484 techniques<sup>38</sup>.<br>485 We found tha We found that systemic administration of D2 antagonist haloperidol increased the use of WSLS<br>486 strategy exclusively in novel blocks, while D1 antagonist SCH-23390 did not (Fig. 6). Such a<br>487 pattern of effects systems a 486 strategy exclusively in novel blocks, while D1 antagonist SCH-23390 did not (**Fig. 6**). Such a<br>487 pattern of effects suggests a direct link between dopamine function via D2 receptors and the<br>488 beganing content depen pattern of effects suggests a direct link between dopamine function via D2 receptors and the<br>488 learning context-dependent behavioral strategy. This notion is in line with previous literature that<br>489 learning context dep 188 learning context-dependent behavioral strategy. This notion is in line with previous literature that<br>
189 has emphasized the role of dopamine in a wide variety of frontal-related cognitive functions,<br>
190 such as work thas emphasized the role of dopamine in a wide variety of frontal-related cognitive functions,<br>490 such as working memory, motivation, attention, and learning<sup>26, 27, 39</sup>. Specifically, a recent study<br>491 demonstrated a e such as working memory, motivation, attention, and learning<sup>26, 27, 39</sup> 491 demonstrated a critical role for the meso-vlPFC dopaminergic pathway in probabilistic decision-<br>492 making<sup>14</sup>, suggesting that dopaminergic inputs modulate vlPFC-centered functional circuits. Our demonstrated a critical role for the meso-vlPFC dopaminergic pathway in probabilistic decision-<br>492 making<sup>14</sup>, suggesting that dopaminergic inputs modulate vlPFC-centered functional circuits. Our<br>492 mating at the fMPL da making<sup>14</sup>, suggesting that dopaminergic inputs modulate vlPFC-centered functional circuits. Our<br>493 resting-state fMRI data analysis revealed that the administration of haloperidol, but not SCH-<br>494 residence and regional resting-state fMRI data analysis revealed that the administration of haloperidol, but not SCH-<br>494 23390, enhanced regional activity specifically within vlPFC. Thus, both dopaminergic<br>495 modulation via D2 receptors and pr 495 modulation via D2 receptors and probabilistic learning/use of behavioral strategies appear to<br>496 generally activate vlPFC. Taken together, our series of experiments reveal that a network of areas 496 generally activate vlPFC. Taken together, our series of experiments reveal that a network of areas<br>497 centered on vlPFC and including dopaminergic mechanisms underlie the flexible adjustment of querally activate vlPFC. Taken together, our series of experiments reveal that a network of areas<br>497 centered on vlPFC and including dopaminergic mechanisms underlie the flexible adjustment of<br>498 behavioral statesy dopen 497 centered on vlPFC and including dopaminergic mechanisms underlie the flexible adjustment of<br>498 behavioral strategy depending on the learning context. behavioral strategy depending on the learning context.<br>
499 A number of psychiatric conditions are characterized by maladaptive behaviors in uncertain

500 reward environments. Previous studies showed that the ability to associate stimuli to<br>501 probabilistic reward outcome or flexibly adapt risk tolerance in probabilistic paradigms was 501 probabilistic reward outcome or flexibly adapt risk tolerance in probabilistic paradigms was<br>502 impaired in human or non-human primate subjects with damage to prefrontal cortex<sup>13, 40, 41</sup> or 501 probabilistic reward outcome or flexibly adapt risk tolerance in probabilistic paradigms was<br>502 impaired in human or non-human primate subjects with damage to prefrontal cortex<sup>13, 40, 41</sup> or impaired in human or non-human primate subjects with damage to prefrontal cortex<sup>13, 40, 41</sup> or<br>20

patients with gambling disorders<sup>4, 42-44</sup>. Interestingly, a recent study showed that lesions of MD 503 patients with gambling disorders<sup>4, 42-44</sup>. Interestingly, a recent study showed that lesions of MD<br>504 thalamus, the area we found to interact with vlPFC during decisions to stay, are associated with<br>505 aberrant swi thalamus, the area we found to interact with vlPFC during decisions to stay, are associated with<br>505 aberrant switching choices akin to the behavioral patterns of human subjects with paranoia<sup>37</sup>.<br>506 This notatiolly impli aberrant switching choices akin to the behavioral patterns of human subjects with paranoia<sup>37</sup> 506 This potentially implicates dysfunction within this pathway as contributing to delusional beliefs<br>507 in disorders like schizophrenia. It is also noteworthy that dopamine function, particularly through 507 in disorders like schizophrenia. It is also noteworthy that dopamine function, particularly through<br>508 D2 receptors, has been implicated in schizophrenia, the behavioral pattern of which is also 508 D2 receptors, has been implicated in schizophrenia, the behavioral pattern of which is also characterized by impairment in flexible decision making based on probabilistic associations<sup>5, 6, 45,</sup> 508 D2 receptors, has been implicated in schizophrenia, the behavioral pattern of which is also<br>509 characterized by impairment in flexible decision making based on probabilistic associations<sup>5, 6, 45,</sup><br> $500<sup>46</sup>$  Inde characterized by impairment in flexible decision making based on probabilistic associations<sup>5, 6, 45,</sup> 510<br>511 <sup>46</sup>. Indeed, current theory posits that aberrant interactions between the salience and fronto-510 <sup>46</sup>. Indeed, current theory posits that aberrant interactions between the salience and fronto-<br>511 parietal networks, which include vIPFC as a main hub<sup>16</sup>, potentially underlie the biases in<br>512 decision making that parietal networks, which include vlPFC as a main hub<sup>16</sup> 511 parietal networks, which include vlPFC as a main hub<sup>16</sup>, potentially underlie the biases in<br>512 decision making that are observed in schizophrenia<sup>47, 48</sup>. Thus, the present study provides a new<br>512 direction incipal decision making that are observed in schizophrenia<sup>47, 48</sup> decision making that are observed in schizophrenia<sup>47, 48</sup>. Thus, the present study provides a new<br>513 insight regarding the neural mechanisms underlying the flexible adjustment of behavioral<br>514 strategy depending on the 513 insight regarding the neural mechanisms underlying the flexible adjustment of behavioral<br>514 strategy depending on the learning context that is potentially relevant to psychiatric disorders<br>515 with impaired aggrificit 514 strategy depending on the learning context that is potentially relevant to psychiatric disorders<br>515 with impaired cognitive flexibility. 515 with impaired cognitive flexibility.<br>516

## 517 517 **Methods**

519 Eight rhesus macaques (*Macaca mulatta*, 7-8 years old, 5 females) served as subjects. A subset 520 of four monkeys (monkeys Ee, Ge, Me, St) underwent awake-fMRI scans. Another subset of<br>521 seven monkeys (monkeys Bu, Cy, Ee, Me, Pi, St, Wo) underwent pharmacological experiments 520 of four monkeys (monkeys Ee, Ge, Me, St) underwent awake-fMRI scans. Another subset of<br>521 seven monkeys (monkeys Bu, Cy, Ee, Me, Pi, St, Wo) underwent pharmacological experiments<br>523 with denominaties drugs. The exper 521 seven monkeys (monkeys Bu, Cy, Ee, Me, Pi, St, Wo) underwent pharmacological experiments<br>522 with dopaminergic drugs. The experiments performed for each subject are summarized in **Table** with dopaminergic drugs. The experiments performed for each subject are summarized in **Table**<br>523 **1**. All procedures were reviewed and approved by the Icahn School of Medicine Animal Care and 1. All procedures were reviewed and approved by the Icahn School of Medicine Animal Care and<br>524 Use Committee. 524 Use Committee.<br>525 Surgery

526 Prior to training, an MRI compatible head-fixation device (Rogue research, Cambridge, MA) 527 was surgically implanted using dental acrylic (Lang Dental, Wheeling, IL) and ceramic screws<br>528 (Thomas Research Products, Elgin, IL) in the animals that underwent behavioral testing 528 (Thomas Research Products, Elgin, IL) in the animals that underwent behavioral testing<br>529 (monkeys Ee, Ge, Me, Pi, St). Briefly, following induction with ketamine (5 m/kg) and 529 (monkeys Ee, Ge, Me, Pi, St). Briefly, following induction with ketamine (5 m/kg) and<br>530 dexmedetomidine (0.0125 mg/kg), the animals were maintained on isoflurane (2-3%), and 8-10 530 dexmedetomidine (0.0125 mg/kg), the animals were maintained on isoflurane (2-3%), and 8-10 screws were implanted into the cranium and the head fixation device was bonded to the screws 530 dexmedetomidine (0.0125 mg/kg), the animals were maintained on isoflurane (2-3%), and 8-10<br>531 screws were implanted into the cranium and the head fixation device was bonded to the screws 531 screws were implanted into the cranium and the head fixation device was bonded to the screws<br>21

532 using dental acrylic. The animals were treated for discomfort and monitored by the researchers<br>533 and veterinary staff till fully recovered. The position of implant was determined based on a pre-<br>524 accuring T1 unigh 533 and veterinary staff till fully recovered. The position of implant was determined based on a pre-<br>534 acquired T1-weighted MR image. 534 acquired T1-weighted MR image.<br>535 Probabilistic learning task

536 We used a reward-based probabilistic learning task that we recently developed for macaque 536 We used a reward-based probabilistic learning task that we recently developed for macaque<br>537 monkeys<sup>28</sup>. The task was controlled by NIMH MonkeyLogic software<sup>49</sup> running on MATLAB<br>528 a 2010s (MathWarks, Natial: MA) monkeys<sup>28</sup>. The task was controlled by NIMH MonkeyLogic software<sup>49</sup> monkeys<sup>28</sup>. The task was controlled by NIMH MonkeyLogic software<sup>49</sup> running on MATLAB<br>538 2019a (MathWorks, Natick, MA) and presented on a monitor in front of the monkey. In this task,<br>539 animals were required to choose 2019a (MathWorks, Natick, MA) and presented on a monitor in front of the monkey. In this task,<br>539 animals were required to choose between two visual stimuli using a directed saccadic eye 539 animals were required to choose between two visual stimuli using a directed saccadic eye<br>540 movement. A trial began with the appearance of a fixation spot (white cross) at the center of the<br>541 second which the monkey 540 movement. A trial began with the appearance of a fixation spot (white cross) at the center of the screen, which the monkey had to maintain fixation on for 1-1.5 sec to initiate a trial. The fixation spot was then extin 541 screen, which the monkey had to maintain fixation on for 1-1.5 sec to initiate a trial. The fixation<br>542 spot was then extinguished and two stimuli were simultaneously presented to the right and left<br>542 section the se 542 spot was then extinguished and two stimuli were simultaneously presented to the right and left<br>543 on the screen. The two stimuli presented on each trial were randomly chosen from a larger pool<br>544 of three viewel stim 543 on the screen. The two stimuli presented on each trial were randomly chosen from a larger pool<br>544 of three visual stimuli that were associated with different reward probabilities (0.9, 0.5, and 0.3). 544 of three visual stimuli that were associated with different reward probabilities (0.9, 0.5, and 0.3).<br>545 Each trial fell into one of three categories based on the reward probabilities of the options<br>546 Each trial Li 545 Each trial fell into one of three categories based on the reward probabilities of the options<br>546 presented: High-Low (0.9-0.3), High-Mid (0.9-0.5), and Mid-Low (0.5-0.3). Stimuli were either<br>547 provel at the beginnin presented: High-Low (0.9-0.3), High-Mid (0.9-0.5), and Mid-Low (0.5-0.3). Stimuli were either<br>547 novel at the beginning of each block of 100 completed trials (novel block), or subjects had<br>548 novel at the prediction of r 547 novel at the beginning of each block of 100 completed trials (novel block), or subjects had<br>548 previously learned the probability of receiving a reward associated with each image, making<br>549 them highly familiar (fami 549 them highly familiar (familiar block). Once stimuli were presented, subjects were required to<br>550 move their gaze toward either the right or left stimulus option ('response') within 2 seconds. 550 move their gaze toward either the right or left stimulus option ('response') within 2 seconds.<br>551 Following a response, the chosen stimulus remained on screen for 0.3 sec, and then was removed 550 move their gaze toward either the right or left stimulus option ('response') within 2 seconds.<br>551 Following a response, the chosen stimulus remained on screen for 0.3 sec, and then was removed 551 Following a response, the chosen stimulus remained on screen for 0.3 sec, and then was removed<br>552 and a reward (1 drop of apple juice) was provided in accordance with the probability of the<br>552 absessmentian. Subseque 552 and a reward (1 drop of apple juice) was provided in accordance with the probability of the<br>553 chosen option. Subsequently an inter-trial interval (ITI, 3-3.5 sec) followed. A trial with a<br>554 fixation break during th 554 fixation break during the fixation period or with no response within the response window was<br>555 aborted; all stimuli were extinguished immediately and ITI started. The same trial was repeated 554 fixation break during the fixation period or with no response within the response window was<br>555 aborted; all stimuli were extinguished immediately and ITI started. The same trial was repeated<br>556 aborted trial 555 aborted; all stimuli were extinguished immediately and ITI started. The same trial was repeated following an aborted trial. 556 following an aborted trial.<br>557 The animals were trained in a mock MRI scanner for 3-6 months in advance of experiments. On

558 an experimental day, the animals performed 2-6 blocks, in which novel and familiar blocks (i.e.<br>559 learning context) were pseudorandomly interleaved. In awake-fMRI sessions, animals performed 559 learning context) were pseudorandomly interleaved. In awake-fMRI sessions, animals performed<br>560 the task in the MRI scanner during functional scans (see *Awake fMRI data acquisition* section). the task in the MRI scanner during functional scans (see *Awake fMRI data acquisition* section).<br>561 In pharmacology sessions, animals performed the task in the mock scanner. The injection of the task in the MRI scanner during functional scans (see *Awake fMRI data acquisition* section).<br>561 In pharmacology sessions, animals performed the task in the mock scanner. The injection of 561 In pharmacology sessions, animals performed the task in the mock scanner. The injection of<br>22

562 saline, SCH-23390, or haloperidol solution (i.m.) was performed 15 minutes prior to the task<br>563 start, and the order of drug treatment was randomized. The injections were at least a day apart<br>564 (SCH-23390) or a week 563 start, and the order of drug treatment was randomized. The injections were at least a day apart (SCH-23390) or a week apart (haloperidol) to avoid potential prolonged effects of the drug, in  $50$ 564 (SCH-23390) or a week apart (haloperidol) to avoid potential prolonged effects of the drug, in<br>565 accordance with known pharmacokinetics of the drugs in macaque monkeys<sup>50</sup>. accordance with known pharmacokinetics of the drugs in macaque monkeys<sup>50</sup>.<br>566 . Awake fMRI data acquisition

566 *Awake fMRI data acquisition*  567 Animals sat in sphynx position in a custom-built MRI-compatible primate chair (Rogue Research,<br>568 Cambridge, MA) to perform a behavioral task in the MRI scanner (Siemens Skyra 3T). First the<br>569 Cambridge research, an 568 Cambridge, MA) to perform a behavioral task in the MRI scanner (Siemens Skyra 3T). First the<br>569 animals received an intravenous injection of a contrast agent, monocrystalline iron oxide<br>570 nanoparticle or MION (BIOPA 570 annoparticle or MION (BIOPAL, Worcester, MA), at a concentration of 10 mg/kg 30 minutes<br>571 prior to the scan<sup>51, 52</sup>. After head fixation, a custom-built 4-channel coil was placed around the manoparticle or MION (BIOPAL, Worcester, MA), at a concentration of 10 mg/kg 30 minutes<br>571 prior to the scan<sup>51, 52</sup>. After head fixation, a custom-built 4-channel coil was placed around the<br>573 bead. Five may many was ma prior to the scan<sup>51, 52</sup> 571 prior to the scan<sup>51, 52</sup>. After head fixation, a custom-built 4-channel coil was placed around the<br>572 head. Eye movement was monitored via infra-red camera and tracked using EyeLink 1000<br>573 as fitures (SB Bassarsh bead. Eye movement was monitored via infra-red camera and tracked using EyeLink 1000<br>573 software (SR Research, Ottawa, Canada). Juice reward was provided through pressurized tubing.<br>574 A secsion started with a set of set 573 software (SR Research, Ottawa, Canada). Juice reward was provided through pressurized tubing.<br>574 A session started with a set of setup scans which included shimming based on the acquired<br>575 Seldman, Fellowing, a 2D, 574 A session started with a set of setup scans which included shimming based on the acquired<br>575 fieldmap. Following a 3D T1-weighted MPRAGE image (0.5 mm isotropic, TR/TI/TE<br>576 3500/1200/2.81 ms flip angle 88) 2.6 mms o 575 fieldmap. Following a 3D T1-weighted MPRAGE image  $(0.5 \text{ mm}$  isotropic, TR/TI/TE<br>576 2500/1200/2.81 ms, flip angle 8°), 2-6 runs of echo planar image (EPI) functional scans  $(1.6 \text{ mm})$ <br>577 isotropic TR/TE 2120/16 ms, 576 2500/1200/2.81 ms, flip angle 8°), 2-6 runs of echo planar image (EPI) functional scans (1.6 mm<br>577 isotropic, TR/TE 2120/16 ms, flip angle 45°, 300-500 volumes per each run) were obtained, with<br>578 angele functional 577 isotropic, TR/TE 2120/16 ms, flip angle 45°, 300-500 volumes per each run) were obtained, with<br>578 each functional scan occurring in conjunction with a separate block of behavioral testing. Overall,<br>579 animals each co each functional scan occurring in conjunction with a separate block of behavioral testing. Overall,<br>579 animals each completed 4 to 7 scanning sessions, for a total of 23 scanning sessions (55 novel<br>580 and 42 familiar bla 579 animals each completed 4 to 7 scanning sessions, for a total of 23 scanning sessions (55 novel<br>580 and 42 familiar blocks). 580 and 42 familiar blocks).<br>581 Resting-state fMRI data acquisition

582 The scans were performed under the same protocol we previously developed for macaque 582 The scans were performed under the same protocol we previously developed for macaque<br>583 monkeys<sup>28, 53, 54</sup>. In brief, following sedation with ketamine (5mg/kg) and dexmedetomidine<br>584 (0.0125mg/kg) the animals were monkeys $^{28,\; 53,\; 54}$ 583 monkeys<sup>28, 53, 54</sup>. In brief, following sedation with ketamine (5mg/kg) and dexmedetomidine<br>584 (0.0125mg/kg) the animals were intubated. They were then administered MION (10 mg/kg, i.v.),<br>585 and three EPI functiona (0.0125mg/kg) the animals were intubated. They were then administered MION (10 mg/kg, i.v.),<br>585 and three EPI functional scans (300 volumes per each run) were obtained, along with a T1-<br>586 anisotical structurel seen (are 585 and three EPI functional scans (300 volumes per each run) were obtained, along with a T1-<br>586 weighted structural scan (pre-injection scans). Following drug i.v. injection (saline, SCH-23390,<br>587 ar halongridal) and 15 586 weighted structural scan (pre-injection scans). Following drug i.v. injection (saline, SCH-23390, or haloperidol) and 15 minutes waiting period, another set of three functional scans was acquired (post-injection scans) 588 (post-injection scans). Low-level isoflurane (0.7-0.9%) was used to maintain sedation through a<br>589 session so that neural activity was preserved while minimizing motion artifacts. The doses of 589 session so that neural activity was preserved while minimizing motion artifacts. The doses of<br>590 drugs used in the scans (50  $\mu$ g/kg and 10  $\mu$ g/kg for SCH and haloperidol, respectively) were pre-590 drugs used in the scans  $(50 \mu g/kg$  and  $10 \mu g/kg$  for SCH and haloperidol, respectively) were pre-<br>591 determined based on a prior PET study to achieve up to 70-80% occupancy of the DA receptors  $590$  drugs used in the scans (50  $\mu$ g/kg and 10  $\mu$ g/kg for SCH and haloperidol, respectively) were pre-<br>591 determined based on a prior PET study to achieve up to 70-80% occupancy of the DA receptors 591 determined based on a prior PET study to achieve up to 70-80% occupancy of the DA receptors<br>23

in macaques<sup>50</sup>. Vital signs (end-tidal  $CO<sub>2</sub>$ , body temperature, blood pressure, capnograph) were 592 in macaques<sup>50</sup>. Vital signs (end-tidal  $CO_2$ , body temperature, blood pressure, capnograph) were<br>593 continuously monitored and maintained as steadily as possible throughout an experimental<br>504 assains 593 continuously monitored and maintained as steadily as possible throughout an experimental<br>594 session. 594 session.<br>595 Drugs

596 SCH-23390 hydrochloride (Tocris Bioscience, Minneapolis, MN) and haloperidol (Sigma-596 SCH-23390 hydrochloride (Tocris Bioscience, Minneapolis, MN) and haloperidol (Sigma-<br>597 Aldrich, St. Louis, MO) were used as D1 and D2 receptor selective antagonists, respectively<sup>55</sup>. Aldrich, St. Louis, MO) were used as D1 and D2 receptor selective antagonists, respectively<sup>55</sup> 597 Aldrich, St. Louis, MO) were used as D1 and D2 receptor selective antagonists, respectively<sup>55</sup>.<br>598 Both SCH and haloperidol were dissolved and diluted in 0.9% saline to achieve target dose of 1<br>599 ml solution. 0.9% 598 Both SCH and haloperidol were dissolved and diluted in 0.9% saline to achieve target dose of 1 ml solution. 0.9% saline (1 ml) was also used as a control solution. The solution was prepared  $500$  fresh an avery experim 599 ml solution. 0.9% saline (1 ml) was also used as a control solution. The solution was prepared<br>600 fresh on every experimental day. 600 fresh on every experimental day.<br>601 Behavioral data analyses

601 *Behavioral data analyses*  602 All behavioral data was analyzed using MATLAB 2022b. Choice performance was defined as<br>603 the proportion of trials in a block (100 trials) in which monkeys chose an option associated with<br>604 bigher reward probability the proportion of trials in a block (100 trials) in which monkeys chose an option associated with<br>604 higher reward probability in the stimulus pair presented. Reaction time (RT) was defined as the<br>605 duration from the ti 605 duration from the timing of visual stimuli presentation to the timing of response initiation.<br>606 Choice performance was computed for bins of 10 trials at each block and averaged for each 606 Choice performance was computed for bins of 10 trials at each block and averaged for each<br>607 subject, then finally averaged across subjects for each context, novel and familiar. We also 606 Choice performance was computed for bins of 10 trials at each block and averaged for each<br>607 subject, then finally averaged across subjects for each context, novel and familiar. We also 607 subject, then finally averaged across subjects for each context, novel and familiar. We also<br>608 computed choice performance for each quartile and performed a two-way repeated-measures<br>609 ANOVA (trial bin: 1.4 block t 608 computed choice performance for each quartile and performed a two-way repeated-measures<br>609 ANOVA (trial bin: 1-4, block type: novel or familiar) for each block type. We reasoned that a<br>610 significant interaction of 609 ANOVA (trial bin: 1-4, block type: novel or familiar) for each block type. We reasoned that a<br>610 significant interaction of trial bin by block type  $(p < 0.05)$  indicates the improvement of<br>611 apartomenes through succ 610 significant interaction of trial bin by block type  $(p < 0.05)$  indicates the improvement of<br>611 performance through successful learning in novel blocks. Choice performance and RT on each<br>612 stimulus pain was assessed 611 performance through successful learning in novel blocks. Choice performance and RT on each<br>612 stimulus pair were assessed by two-way repeated-measures ANOVA (stimulus pair: 0.9-0.3, 0.9-<br>612 0.5.0.5.0.2 khoal two nove 612 stimulus pair were assessed by two-way repeated-measures ANOVA (stimulus pair: 0.9-0.3, 0.9-<br>613 0.5, 0.5-0.3, block type: novel or familiar). 613 0.5, 0.5-0.3, block type: novel or familiar).<br>614 Switching trials were defined as trials in which the monkeys chose a different stimulus although

615 the previously-chosen stimulus was available in the current trial, as opposed to stay trials in which the monkeys chose the same stimulus sequentially. Thus the proportion of switching trials the previously-chosen stimulus was available in the current trial, as opposed to stay trials in<br>616 which the monkeys chose the same stimulus sequentially. Thus the proportion of switching trials<br>617 was limited to these t 616 which the monkeys chose the same stimulus sequentially. Thus the proportion of switching trials<br>617 was limited to those trials in which the previously-chosen stimulus was available. The proportion<br>618 and stimulus res was limited to those trials in which the previously-chosen stimulus was available. The proportion<br>618 of switching trials regarding previous outcome and block type was assessed using two-way<br>619 repeated-measures ANOVA (ou 619 repeated-measures ANOVA (outcome: reward or no-reward, block type: novel or familiar). We interpreted a significant interaction of outcome by block type ( $p < 0.05$ ) to indicate that the 619 repeated-measures ANOVA (outcome: reward or no-reward, block type: novel or familiar). We<br>620 interpreted a significant interaction of outcome by block type ( $p < 0.05$ ) to indicate that the 620 interpreted a significant interaction of outcome by block type  $(p < 0.05)$  to indicate that the<br>24

621 proportion of win-stay/lose-shift (WSLS) trials was varied depending on the learning context.<br>622 We also performed a two-way repeated-measures ANOVA (trial bin: 1-4, block type: novel or<br>623 familiar) on the proportio 623 familiar) on the proportion of WSLS choices to assess the impact of learning on WSLS strategy.<br>624 The direct relationship between WSLS choices and choice performance was assessed by familiar) on the proportion of WSLS choices to assess the impact of learning on WSLS strategy.<br>624 The direct relationship between WSLS choices and choice performance was assessed by<br>625 according the Pearson's correlation 624 The direct relationship between WSLS choices and choice performance was assessed by<br>625 calculating the Pearson's correlation coefficient. To do this with all subjects combined, the<br>626 career at WSLS schooles and the 625 calculating the Pearson's correlation coefficient. To do this with all subjects combined, the<br>626 proportion of WSLS choices and the choice performance on each block were z-transformed for<br>627 a sock assign and sach bl 626 proportion of WSLS choices and the choice performance on each block were z-transformed for each session and each block type. 627 each session and each block type.<br>628 The effect of drug injection on the proportion of WSLS choices was assessed by two-way

629 repeated-measures ANOVA (trial bin: 1-4, dose: 0, 5, or 10  $\mu$ g/kg of haloperidol, or 0, 10, 30, or<br>630 50  $\mu$ g/kg of SCH-23390) for each block type. The significant main effect of drug dose (p < 0.05) 630 50  $\mu$ g/kg of SCH-23390) for each block type. The significant main effect of drug dose ( $p < 0.05$ ) indicates the impact of the drug on the proportion of WSLS trials in a specific learning context. 630 50  $\mu$ g/kg of SCH-23390) for each block type. The significant main effect of drug dose (p < 0.05) indicates the impact of the drug on the proportion of WSLS trials in a specific learning context. indicates the impact of the drug on the proportion of WSLS trials in a specific learning context.<br>632 *fMRI data analyses* 

632 *fMRI data analyses*  633 Imaging data were analyzed using a customized AFNI processing pipeline for non-human<br>634 primates<sup>56</sup> and the standard NMT atlas<sup>57</sup>. Following the preprocessing steps, whole-brain<br>625 analyzis was performed. For aval primates<sup>56</sup> and the standard NMT atlas<sup>57</sup> 634 primates<sup>56</sup> and the standard NMT atlas<sup>57</sup>. Following the preprocessing steps, whole-brain<br>635 analysis was performed. For awake-fMRI data, regression analysis was performed for each<br>636 session with the timing of th 636 session with the timing of the outcome as a regressor. One scanning session with only novel<br>637 blocks was excluded from this analysis, resulting in 22 sessions (51 novel and 42 familiar 636 session with the timing of the outcome as a regressor. One scanning session with only novel<br>637 blocks was excluded from this analysis, resulting in 22 sessions (51 novel and 42 familiar<br>628 blocks). The resulting corr blocks was excluded from this analysis, resulting in 22 sessions (51 novel and 42 familiar blocks). The resulting correlation coefficients for each voxel were submitted to a two-way 638 blocks). The resulting correlation coefficients for each voxel were submitted to a two-way<br>639 ANOVA (outcome: reward or no reward, block type: novel or familiar) with subject and session<br>640 as random effects<sup>10</sup>. Gr 639 ANOVA (outcome: reward or no reward, block type: novel or familiar) with subject and session<br>640 as random effects<sup>10</sup>. Group-level statistics were computed by 3dClustSim using initial<br>641 thresholding at  $n \le 0.05$  i as random effects<sup>10</sup>. Group-level statistics were computed by 3dClustSim using initial<br>641 thresholding at  $p < 0.05$  in the ANOVA and a cluster size of 14 voxels that is corrected for<br>642 multiple comparisons at  $p < 0.05$ 641 thresholding at  $p < 0.05$  in the ANOVA and a cluster size of 14 voxels that is corrected for multiple comparisons at  $p < 0.05$ . Subsequent conjunction analyses specified the areas that  $642$  survived aluster hased ser 642 multiple comparisons at  $p < 0.05$ . Subsequent conjunction analyses specified the areas that survived cluster-based correction in both outcome and block type (context) contrast.

643 survived cluster-based correction in both outcome and block type (context) contrast.<br>644 A region of interest (ROI) was chosen based on the result of the conjunction analys 644 A region of interest (ROI) was chosen based on the result of the conjunction analysis. The peak<br>645 voxel from right vlPFC and its adjacent voxels (faces touching) was used as the main ROI for<br>646 voxes were also used 645 voxel from right vlPFC and its adjacent voxels (faces touching) was used as the main ROI for<br>646 subsequent ROI-based analyses. First, the time series was extracted from the ROI and z-<br>647 transformed for seek block an 646 subsequent ROI-based analyses. First, the time series was extracted from the ROI and z-<br>647 transformed for each block, and the timing of the signal was aligned to the timing of outcome. transformed for each block, and the timing of the signal was aligned to the timing of outcome.<br>
648 Then, trials were divided based on outcome (reward or no reward) and subsequent decision (stay<br>
649 or shift), and the ave 649 or shift), and the averaged time series for each trial type with smoothing was computed to create<br>650 a peri-stimulus timing histogram for each task. A sliding-window multiple linear regression (500) 649 or shift), and the averaged time series for each trial type with smoothing was computed to create<br>650 a peri-stimulus timing histogram for each task. A sliding-window multiple linear regression (500 650 a peri-stimulus timing histogram for each task. A sliding-window multiple linear regression (500 651 msec window, 100 msec step) concerning outcome and stay/shift decision was performed and the resulting beta coefficients were plotted around the timing of outcome. Also, a correlation between ROI value (0-4 sec from th 653 between ROI value (0-4 sec from the timing of outcome) and normalized choice performance<br>654 was calculated. A multidimensional analysis was performed by plotting beta coefficients 653 between ROI value (0-4 sec from the timing of outcome) and normalized choice performance<br>654 was calculated. A multidimensional analysis was performed by plotting beta coefficients<br>655 (externe and star) which decision 654 was calculated. A multidimensional analysis was performed by plotting beta coefficients<br>655 (outcome and stay/shift decision coding) in 2-D space for each block type separately. Then, the<br>656 Evalidaen distance between 655 (outcome and stay/shift decision coding) in 2-D space for each block type separately. Then, the<br>656 Euclidean distance between novel and familiar blocks across the timing of the trial was<br>657 computed and plotted again 656 Euclidean distance between novel and familiar blocks across the timing of the trial was<br>657 computed and plotted against the trial time-course. The distance measure was compared to<br>658 shuffled data (05% CI) that was c 657 computed and plotted against the trial time-course. The distance measure was compared to<br>658 shuffled data (95% CI) that was computed by randomly assigning block types for each trial and<br>659 starsted 1.000 times. 658 shuffled data (95% CI) that was computed by randomly assigning block types for each trial and<br>659 iterated 1,000 times.

660 Subsequently, functional connectivity (FC) analysis was performed using the right vIPFC seed 661 that was used in the main ROI analysis. We performed a generalized form of context-dependent<br>662 syven optional interactions (gPPI)<sup>22</sup> to compute a vlPFC-derived network modulated by 661 that was used in the main ROI analysis. We performed a generalized form of context-dependent<br>662 psychophysiological interactions  $(gPPI)^{22}$  to compute a vIPFC-derived network modulated by<br>662 begaining sentext (nevel psychophysiological interactions  $(gPPI)^{22}$ 662 psychophysiological interactions  $(gPPI)^{22}$  to compute a vlPFC-derived network modulated by<br>663 learning context (novel or familiar) at the whole-brain level. Then, a secondary ROI was defined<br>664 learning on pack wou 1663 learning context (novel or familiar) at the whole-brain level. Then, a secondary ROI was defined<br>664 based on peak voxels in the brain map. The ROI time-course of FC was computed by calculating<br>665 the seed BOI correl 664 based on peak voxels in the brain map. The ROI time-course of FC was computed by calculating<br>665 the seed-ROI correlation (Pearson's r) over time using a sliding window analysis. The trials were<br>666 divided based on ou 666 divided based on outcome and subsequent choice, and the FC time course for each trial type was<br>667 calculated around the timing of outcome, for each block type. The significant FC change 667 calculated around the timing of outcome, for each block type. The significant FC change<br>668 between -4 to 8 seconds after the outcome timing was detected when 3 consecutive bins reached 667 calculated around the timing of outcome, for each block type. The significant FC change<br>668 between -4 to 8 seconds after the outcome timing was detected when 3 consecutive bins reached<br>669  $R \le 0.05$  in a Wilayam's r 668 between -4 to 8 seconds after the outcome timing was detected when 3 consecutive bins reached<br>669  $p < 0.05$  in a Wilcoxon's rank-sum test. We further performed a sliding-window multiple<br>670 measures and the subset of 669  $p < 0.05$  in a Wilcoxon's rank-sum test. We further performed a sliding-window multiple<br>670 regression analysis concerning outcome and stay/shift decision in the FC. The beta coefficients<br>671 of the interaction of out 670 regression analysis concerning outcome and stay/shift decision in the FC. The beta coefficients<br>671 of the interaction of outcome by stay/shift decision were plotted as WSLS coding of the FC and<br>672 accuraced to the 05 671 of the interaction of outcome by stay/shift decision were plotted as WSLS coding of the FC and<br>672 compared to the 95% CI of the shuffled data. 672 compared to the 95% CI of the shuffled data.<br>673 Resting-state fMRI data were preprocessed in the same manner as the awake scans. The residual

674 error file for each run was split in half, resulting in 6 runs for pre- and post-injection scans,<br>675 espectively, for each of the drug injection sessions. Regional homogeneity (ReHo) analysis was 675 respectively, for each of the drug injection sessions. Regional homogeneity (ReHo) analysis was performed using the function 3dReHO in the AFNI FATCAT toolbox<sup>30</sup>. ROI analysis was 675 respectively, for each of the drug injection sessions. Regional homogeneity (ReHo) analysis was<br>676 performed using the function 3dReHO in the AFNI FATCAT toolbox<sup>30</sup>. ROI analysis was<br>677 performed using the right ul performed using the function 3dReHO in the AFNI FATCAT toolbox<sup>30</sup>. ROI analysis was<br>677 performed using the right vIPFC ROI that was used in the main ROI analyses. Two-way ANOVA<br>678 (drugs seline, SGU 22200) or helopogid 677 performed using the right vlPFC ROI that was used in the main ROI analyses. Two-way ANOVA<br>678 (drug: saline, SCH-23390, or haloperidol, injection: pre or post) was performed to assess the<br>679 effect of drug injection o 678 (drug: saline, SCH-23390, or haloperidol, injection: pre or post) was performed to assess the effect of drug injection on vlPFC activity. The results were superimposed on the NMTv2.0 679 effect of drug injection on vlPFC activity. The results were superimposed on the NMTv2.0<br>680 template for visualization purpose<sup>57</sup>. 680 template for visualization purpose<sup>57</sup>.

# 681<br>682 682 **Conflict of interest:** The authors declare no competing financial interest.

- 683
- **Author contributions:** A.F., C.E., B.E.R., and P.H.R. designed the study. A.F., C.E., S.H.F., and L.F. performed the study. A.F. analyzed the data. A.F., C.E., B.E.R., and P.H.R. wrote the original data data data and  $\frac{$ 685 L.F. performed the study. A.F. analyzed the data. A.F., C.E., B.E.R., and P.H.R. wrote the original draft. All authors edited the paper.
- 686 draft. All authors edited the paper.<br>687

688 **Data Availability:** The data that support the findings of this study are available from the corresponding authors upon reasonable request. 689 corresponding authors upon reasonable request.<br>690

## 691

 $(201)$ 

**References**<br> **692** 1. Stephens, D.W. & Krebs, J.R. Foraging theory (Princeton university press, 1986).<br> **693** 2. Hodgins, D.C., Stea, J.N. & Grant, J.E. Gambling disorders. *Lancet (London, Engl*)

693 2. Hodgins, D.C., Stea, J.N. & Grant, J.E. Gambling disorders. *Lancet (London, England)* 378,<br>694 1874-1884 (2011).<br>695 3. APA. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®) (American* 

695 3. APA. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5<sup>®</sup>) (American<br>696 Psychiatric Association Publishing Arlington, VA, 2013).* 696 Psychiatric Association Publishing, Arlington, VA, 2013).<br>697 4. Goudriaan, A.E., Oosterlaan, J., de Beurs, E.

697 4. Goudriaan, A.E., Oosterlaan, J., de Beurs, E. & van den Brink, W. Neurocognitive<br>698 functions in pathological gambling: a comparison with alcohol dependence. Tourette syndrome 697 4. Goudriaan, A.E., Oosterlaan, J., de Beurs, E.<br>698 functions in pathological gambling: a comparison with a 698 functions in pathological gambling: a comparison with alcohol dependence, Tourette syndrome<br>699 and normal controls. Addiction (Abingdon, England) 101, 534-547 (2006).

For the matrices in pathological gambling: a comparison with alcohol dependence, Tourette syndrome<br>699 and normal controls. Addiction (Abingdon, England) 101, 534-547 (2006).<br>700 5. Huq, S.F., Garety, P.A. & Hemsley, D.R. 699 and normal controls. Addiction (Abingdon, England) 101, 534-547 (2000).<br>700 5. Huq, S.F., Garety, P.A. & Hemsley, D.R. Probabilistic judgements<br>701 deluded subjects. The Quarterly journal of experimental psychology. A, 702 5. Huq, S.F., B.F., Garety, P.A. Probabilistic judgements in deluded and non-resource in deluded and non-<br>The non-second and n For deluded subjects. The Quarterly Journal of experimental psychology. A, Human experimental<br>702 psychology 40, 801-812 (1988).<br>703 6. Garety, P.A. & Freeman, D. Cognitive approaches to delusions: a critical review of

psychology 40, 801-812 (1988).<br>1903 6. Garety, P.A. & Freeman, D. Cognitive approaches to delusions: a c<br>1999). To theories and evidence. *British journal of clinical psychology* 38, 113-154 (1999).<br>1905 1. Chau, B.K., et

2022 6. Freeman entity, President approaches approaches to the oriental review of<br>1993 6. Garety, P.A. Garety, P. Contrasting Roles for Orbitofrontal Cortex and Amygdala in Credit 704 theories and evidence. British journal of clinical psychology 38, 113-154 (1999).<br>705 7. Chau, B.K., et al. Contrasting Roles for Orbitofrontal Cortex and Am<br>706 Assignment and Learning in Macaques. Neuron 87, 1106-111

706 Assignment and Learning in Macaques. *Neuron* **87**, 1106-1118 (2015).<br>707 8. Folloni, D.*, et al*. Ultrasound modulation of macaque prefrontal cortex selectively alters<br>708 credit assignment-related activity 708 credit assignment-related activity and behavior. Science advances 7, eabg7700 (2021).<br>709 9. Jezzini, A., Bromberg-Martin, E.S., Trambaiolli, L.R., Haber, S.N. & Monos

709 9. Jezzini, A., Bromberg-Martin, E.S., Trambaiolli, L.R., Haber, S.N. & Monosov, I.E. A<br>710 prefrontal network integrates preferences for advance information about uncertain rewards and 708 credit assignment-related activity and behavior. Science duvances 7, eabg7700 (2021).<br>709 9. Lezzini, A., Bromberg-Martin, E.S., Trambaiolli, L.R., Haber, S.N. & Monos<br>710 prefrontal network integrates preferences for 710 prefrontal network integrates preferences for advance information about uncertain rewards and<br>711 punishments. Neuron 109, 2339-2352.e2335 (2021).

711 punishments. *Neuron* **109**, 2339-2352.e2335 (2021).<br>712 10. Kaskan, P.M.*, et al*. Learned Value Shapes Responses to Objects in Frontal and Ventral<br>713 Stream Networks in Macague Monkeys. *Cerebral cortex (New York. N* 711 punishments. Neuron 109, 2339 2332.e2333 (2021).<br>712 10. Kaskan, P.M., et al. Learned Value Shapes Re<br>713 Stream Networks in Macaque Monkeys. Cerebral co 712 110. Kaskan, P.M., et al. Learned Value Shapes Responses to Objects in Frontal and Ventral<br>713 Stream Networks in Macaque Monkeys. *Cerebral cortex (New York, N.Y. : 1991)* 27, 2739-2757<br>714 (2017). 713 Stream Networks in Macaque Monkeys. Cerebral cortex (New York, N.Y. : 1991) 27, 2739-2757<br>714 (2017).

11. Stoll, F.M. & Rudebeck, P.H. Dissociable Representations of Decision Variables within<br>16 Subdivisions of the Macaque Orbital and Ventrolateral Frontal Cortex. The Journal of<br>17 Neuroscience 44, e0464242024 (2024).<br>18 1 11.

718 12. Rudebeck, P.H., Saunders, R.C., Lundgren, D.A. & Murray, E.A. Specialized<br>719 Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus 717 Neuroscience 44, e0404242024 (2024).<br>718 12. Rudebeck, P.H., Saunders, 1<br>719 Representations of Value in the Orbita 719 Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus<br>720 Availability of Outcomes. Neuron 95, 1208-1220.e1205 (2017).

720 Availability of Outcomes. *Neuron* **95**, 1208-1220.e1205 (2017).<br>721 13. Rygula, R., Walker, S.C., Clarke, H.F., Robbins, T.W. & Roberts, A.C. Differential<br>722 contributions of 720 Availability of Outcomes. Neuron 95, 1208-1220.e1205 (2017).<br>721 13. Rygula, R., Walker, S.C., Clarke, H.F., Robbins, T.V<br>722 contributions of the primate ventrolateral prefrontal and orbi 722 contributions of the primate ventrolateral prefrontal and orbitofrontal cortex to serial reversal<br>723 learning. The Journal of neuroscience : the official journal of the Society for Neuroscience 30, 723 learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**,<br>724 14552-14559 (2010).<br>725 14. Sasaki, R., *et al.* Balancing risk-return decisions by manipulating the mesofrontal

- 725 14. Sasaki, R., et al. Balancing risk-return decisions by manipulating the mesofrontal circuits 726 in primates. Science (New York, N.Y.) 383, 55-61 (2024). 726 in primates. Science (New York, N.Y.) 383, 55-61 (2024).<br>727 15. Evenden, J.L. & Robbins, T.W. Increased re
- 727 15. Evenden, J.L. & Robbins, T.W. Increased response switching, perseveration and<br>728 perseverative switching following d-amphetamine in the rat. Psychopharmacology 80, 67-73 226 In primates. Science (New York, N.Y.) 383, 35-61 (2024).<br>727 15. Evenden, J.L. & Robbins, T.W. Increased response perseverative switching following d-amphetamine in 728 perseverative switching following d-amphetamine in the rat. Psychopharmacology 80, 67-73<br>729 (1983).
- 729 (1983).<br>730 16. Trambaiolli, L.R.*, et al.* Anatomical and functional connectivity support the existence of a<br>731 salience network node within the caudal ventrolateral prefrontal cortex. *eLife* 11 (2022). 731 salience network node within the caudal ventrolateral prefrontal cortex. *eLife* 11 (2022).<br>732 17. Rapan, L., et al. Cytoarchitectonic, receptor distribution and functional conr 730 16. Trambaiolii, L.R., et al. Anatomical and functional comectivity support the existence of a<br>731 salience network node within the caudal ventrolateral prefrontal cortex. *eLife* 11 (2022).<br>732 17. Rapan, L., *et al.*

731 Salience network hode within the caudal ventrolateral prefrontal cortex. *ELife* 11 (2022).<br>732 17. Rapan, L., *et al.* Cytoarchitectonic, receptor distribution and functional conr<br>733 analyses of the macaque frontal

- 733 analyses of the macaque frontal lobe. *eLife* 12 (2023).<br>734 18. Monosov, I.E. & Rushworth, M.F.S. Interactions between ventrolateral prefrontal and<br>735 anterior cingulate cortex during learning and behavioural change. 733 analyses of the macaque frontal lobe. eLife 12 (2023).<br>734 18. Monosov, I.E. & Rushworth, M.F.S. Interactio<br>735 anterior cingulate cortex during learning and behavio 735 anterior cingulate cortex during learning and behavioural change. Neuropsychopharmacology:<br>736 official publication of the American College of Neuropsychopharmacology 47, 196-210 (2022).
- 736 official publication of the American College of Neuropsychopharmacology **47**, 196-210 (2022).<br>737 19. Soltani, A. & Wang, X.-J. A Biophysically Based Neural Model of Matching Law Behavior:<br>738 Melioration by Stochastic 738 Melioration by Stochastic Synapses. *The Journal of Neuroscience* 26, 3731-3744 (2006).<br>739 20. Trepka. E., *et al.* Entropy-based metrics for predicting choice behavior based
- 738 Melioration by Stochastic Synapses. The Journal of Neuroscience 26, 3731-3744 (2006).<br>739 20. Trepka, E., et al. Entropy-based metrics for predicting choice behavior based on local 738 Menoration by Stochastic Synapses. The Journal by Neuroscience 20, 3731-3744 (2000).<br>739 20. Trepka, E., et al. Entropy-based metrics for predicting choice behavior based<br>740 response to reward. Nature communications 1
- 740 response to reward. *Nature communications* 12, 6567 (2021).<br>741 21. Fouragnan, E.F., *et al*. The macaque anterior cingulate cortex translates counterfactual<br>742 choice value into actual behavioral change. Nature neur
- 742 be choice value into actual behavioral change. *Nature neuroscience* 22, 797-808 (2019).<br>743 122. McLaren, D.G., Ries, M.L., Xu, G. & Johnson, S.C. A generalized form 743 22. McLaren, D.G., Ries, M.L., Xu, G. & Johnson, S.C. A generalized form of context-<br>744 dependent psychophysiological interactions (gPPI): a comparison to standard approaches. 744 dependent psychophysiological interactions (gPPI): a comparison to standard approaches.<br>745 Neurolmage **61**, 1277-1286 (2012). 745 22. Meuroliniage **61,** 1277-1286 (2012).<br>The Context-Million of The Million of Context-Context-Context-Context-Context-Context-Context-Context-Context-
- 745 Neurolmage 61, 1277-1286 (2012).<br>746 23. Mitchell, A.S. The mediodorsal thalamus as a higher order thalamic relay nucleus. 745 Neurolinge 61, 1277-1288 (2012).<br>746 23. Mitchell, A.S. The medioder<br>747 important for learning and decision 747 important for learning and decision-making. Neuroscience and biobehavioral reviews 54, 76-88<br>748 (2015).
- 747 important for learning and decision-making. Neuroscience and biobehavioral reviews 54, 76-88<br>748 (2015).<br>749 24. Mitchell, A.S., Baxter, M.G. & Gaffan, D. Dissociable performance on scene learning and<br>750 strategy impl 749 24.<br>750 strategy<br>--750 strategy implementation after lesions to magnocellular mediodorsal thalamic nucleus. The<br>751 *Journal of neuroscience : the official journal of the Society for Neuroscience* 27, 11888-11895 751 *Journal of neuroscience : the official journal of the Society for Neuroscience* 27, 11888-11895<br>752 (2007).<br>753 25. Oyama, K., *et al.* Distinct roles of monkey OFC-subcortical pathways in adaptive behavior.
- 753 25. Oyama, K., et al. Distinct roles of monkey OFC-subcortical pathways in adaptive behavior.<br>754 Nature communications 15, 6487 (2024). 754 *Nature communications* **15**, 6487 (2024).<br>755 26. Sawaguchi, T. & Goldman-Rakic
- 755 26. Sawaguchi, T. & Goldman-Rakic, P.S. D1 dopamine receptors in prefrontal cortex:<br>756 involvement in working memory. Science (New York, N.Y.) 251, 947-950 (1991). 756 involvement in working memory. Science (New York, N.Y.) 251, 947-950 (1991).<br>757 27. Froudist-Walsh. S., et al. A dopamine gradient controls access to dis
- 757 27. Froudist-Walsh, S., et al. A dopamine gradient controls access to distributed working<br>758 memory in the large-scale monkey cortex. Neuron 109, 3500-3520.e3513 (2021). 756 involvement in working memory. Science (New York, N.H.) 251, 347-950 (1991).<br>757 27. Froudist-Walsh, S., et al. A dopamine gradient controls access to dis<br>758 memory in the large-scale monkey cortex. Neuron 109, 3500-3  $757$  27. Froudist-Walsh, S., et al. A dopalifie gradient controls access to distributed working<br>758 memory in the large-scale monkey cortex. Neuron **109**, 3500-3520.e3513 (2021).  $758$  memory in the large-scale monkey cortex. Neuron 109, 3500-3520.e3513 (2021).

28. Fujimoto, A., et al. Pharmacological modulation of dopamine D1 and D2 receptors<br>760 reveals distinct neural networks related to probabilistic learning in non-human primates. *bioRxiv*<br>761 (2023). Zang, Y., Jiang, T., L

762 29. Zang, Y., Jiang, T., Lu, Y., He, Y. & Tian, L. Regional homogeneity approach to fMRI data<br>763 analysis. Neurolmage 22. 394-400 (2004). 763 - analysis. *Neurolmage* 22, 394-400 (2004).<br>764 - 30. - Taylor, P.A. & Saad, Z.S. FATCAT: (a

764 30. Taylor, P.A. & Saad, Z.S. FATCAT: (an efficient) Functional and Tractographic Connectivity 765 Analysis Toolbox. *Brain connectivity* 3, 523-535 (2013). 765 Analysis Toolbox. *Brain connectivity* **3**, 523-535 (2013).<br>766 31. Baxter, M.G., Gaffan, D., Kyriazis, D.A. & Mitch

766 31. Baxter, M.G., Gaffan, D., Kyriazis, D.A. & Mitchell, A.S. Ventrolateral prefrontal cortex is<br>767 Prequired for performance of a strategy implementation task but not reinforcer devaluation 765 Analysis Toolbox. Brain connectivity 3, 523-535 (2013).<br>766 31. Baxter, M.G., Gaffan, D., Kyriazis, D.A. & Mitch<br>767 required for performance of a strategy implementat reception of the matter, M.G., Gammington, M.G., Gammington, M.G., Commission, Premium Premium Preference<br>768 effects in rhesus monkeys. The European journal of neuroscience 29, 2049-2059 (2009).

- 1978 republic for performance of a strategy implementation task but not republic effects in rhesus monkeys. The European journal of neuroscience 29, 2049-2059 (2009).<br>769 32. Bussey, T.J., Wise, S.P. & Murray, E.A. The rol 768 effects in filesus monkeys. *The European Journal of neuroscience* 23, 2043-2033 (2003).<br>769 32. Bussey, T.J., Wise, S.P. & Murray, E.A. The role of ventral and orbital prefrontal of<br>770 conditional visuomotor learning 32. Bussey, T.J., Wise, S.P. & Murray, E.A. The role of ventral and orbital prefrontal cortex in<br>770 conditional visuomotor learning and strategy use in rhesus monkeys (Macaca mulatta).<br>771 Behavioral neuroscience 115, 971
- 772 33. Carmichael, S. & Price, J.L. Architectonic subdivision of the orbital and medial prefrontal<br>773 cortex in the macague monkey. Journal of Comparative Neurology 346, 366-402 (1994). 771 Behavioral neuroscience 115, 971-982 (2001).<br>772 33. Carmichael, S. & Price, J.L. Architecton<br>773 cortex in the macaque monkey. Journal of Com
- 773 cortex in the macaque monkey. *Journal of Comparative Neurology* 346, 366-402 (1994).<br>774 34. Stoll, F.M. & Rudebeck, P.H. Decision-making shapes dynamic inter-areal communication<br>---774 34. Stoll, F.M. & Rudebeck, P.H. Decision-making shapes dynamic inter-areal communication<br>775 vithin macaque ventral frontal cortex. *Current Biology*.<br>776 35. Hayden, B.Y., Pearson, J.M. & Platt, M.L. Neuronal basis o
- 776 35. Hayden, B.Y., Pearson, J.M. & Platt, M.L. Neuronal basis of sequential foraging decisions<br>777 in a patchy environment. Nature neuroscience 14, 933-939 (2011). 777 in a patchy environment. Nature neuroscience 14, 933-939 (2011).<br>778 36. Chakraborty, S., Kolling, N., Walton, M.E. & Mitchell, A.S. Cr
- 778 36. Chakraborty, S., Kolling, N., Walton, M.E. & Mitchell, A.S. Critical role for the mediodorsal<br>779 thalamus in permitting rapid reward-guided updating in stochastic reward environments. *eLife* 5 777 In a patchy environment. Nature neuroscience 14, 999 999 (2011).<br>778 36. Chakraborty, S., Kolling, N., Walton, M.E. & Mitchell, A.S. Cr<br>779 thalamus in permitting rapid reward-guided updating in stochastic 779 thalamus in permitting rapid reward-guided updating in stochastic reward environments. *eLife* 5<br>780 (2016).
- 780 (2016).<br>781 37. Suthaharan, P., *et al.* Lesions to the mediodorsal thalamus, but not orbitofrontal cortex,<br>782 enhance volatility beliefs linked to paranoia. *Cell reports* 43, 114355 (2024). 780 (2016). 781 37. Suthaharan, P., et al. Lesions to the mediodorsal thalamus, but not orbitofrontal cortex,<br>782 enhance volatility beliefs linked to paranoia. *Cell reports* 43, 114355 (2024).<br>783 38. Roth, B.L. DREADDs for Neurosci
- 
- 783 38. Roth, B.L. DREADDs for Neuroscientists. *Neuron 89, 683-694 (2016)*.<br>784 39. Jocham, G., Klein, T.A. & Ullsperger, M. Dopamine-Mediated Reinforcement Learning<br>785 Signals in the Striatum and Ventromedial Prefrontal 783 38. Roth, B.L. BREADDS for Neuroscientists. Neuron 89, 683-694 (2016).<br>784 39. Jocham, G., Klein, T.A. & Ullsperger, M. Dopamine-Mediated Reir<br>785 Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Val 785 Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Value-Based Choices. The<br>786 *Journal of Neuroscience* 31, 1606-1613 (2011).
- 786 *Journal of Neuroscience* **31**, 1606-1613 (2011).<br>787 40. Bechara, A., Damasio, H., Tranel, D. & Damasio, A.R. Deciding advantageously before<br>788 Knowing the advantageous strategy. *Science (New York, N.Y.)* 275, 1 788 knowing the advantageous strategy. Science (New York, N.Y.) 275, 1293-1295 (1997).<br>789 41. Axelsson. S.F.A.. Horst. N.K.. Horiguchi. N.. Roberts. A.C. & Robbins. T.W. Fle

789 41. Axelsson, S.F.A., Horst, N.K., Horiguchi, N., Roberts, A.C. & Robbins, T.W. Flexible versus<br>790 Fixed Spatial Self-Ordered Response Sequencing: Effects of Inactivation and Neurochemical 790 Fixed Spatial Self-Ordered Response Sequencing: Effects of Inactivation and Neurochemical<br>791 Modulation of Ventrolateral Prefrontal Cortex. The Journal of neuroscience : the official journal Frame Constitution, States, May 1998, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1999, 1<br>791 Modulation of Ventrolateral Prefrontal Cortex. The Journal of neuroscience : the official journal 791 Modulation of Ventrolateral Prefrontal Cortex. The Journal of neuroscience: the official journal<br>792 of the Society for Neuroscience 41, 7246-7258 (2021).

792 of the Society for Neuroscience **41**, 7246-7258 (2021).<br>793 – 42. Fujimoto, A., et al. Deficit of state-dependent risk attitude modulation in gambling<br>794 – disorder. Translational psychiatry 7, e1085 (2017). 794 disorder. Translational psychiatry 7, e1085 (2017).<br>795 43. Leppink, E.W., Redden, S.A., Chamberlain, S.

- 795 43. Leppink, E.W., Redden, S.A., Chamberlain, S.R. & Grant, J.E. Cognitive flexibility correlates<br>796 with gambling severity in young adults. *Journal of psychiatric research* 81, 9-15 (2016). 796 vith gambling severity in young adults. *Journal of psychiatric research* 81, 9-15 (2016).<br>797 144. Miehler. A., Chakroun. K. & Peters. J. Attenuated Directed Exploratio
- 196 with gambling severity in young adults. Journal of psychiatric research 81, 9-15 (2016).<br>797 44. Wiehler, A., Chakroun, K. & Peters, J. Attenuated Directed Exploration during 796 With gambing severity in young adults. Journal by psychiatric research 81, 5-15 (2016).<br>797 44. Wiehler, A., Chakroun, K. & Peters, J. Attenuated Directed Exploratio<br>798 Reinforcement Learning in Gambling Disorder. The 798 Reinforcement Learning in Gambling Disorder. *The Journal of Neuroscience* 41, 2512-2522<br>799 (2021).<br>800 45. Suhara, T., *et al*. Decreased dopamine D2 receptor binding in the anterior cingulate
- 800 45. Suhara, T., et al. Decreased dopamine D2 receptor binding in the anterior cingulate 801 cortex in schizophrenia. Archives of aeneral osvchiatry 59. 25-30 (2002).  $800$  45. 800 45. Suhara, T., et al. Decreased dopartile D2 receptor binding in the anterior cingulate<br>801 cortex in schizophrenia. Archives of general psychiatry 59, 25-30 (2002). 801 cortex in schizophrenia. Archives of general psychiatry 59, 25-30 (2002).

802 46. Seeman, P. Dopamine receptors and the dopamine hypothesis of schizophrenia. Synapse<br>803 (New York, N.Y.) 1, 133-152 (1987).<br>804 47. Kapur, S. Psychosis as a state of aberrant salience: a framework linking biology,<br>

803 (*New York, N.Y.*) 1, 133-132 (1987).<br>804 47. Kapur, S. Psychosis as a<br>805 phenomenology, and pharmacolog 805 phenomenology, and pharmacology in schizophrenia. The American journal of psychiatry 160,<br>806 13-23 (2003).

806 13-23 (2003).<br>807 48. Menon, V., Palaniyappan, L. & Supekar, K. Integrative Brain Network and Salience Models<br>808 of Psychopathology and Cognitive Dysfunction in Schizophrenia. *Biological psych* end<br>807 48. Menor<br>808 of Psychopath 808 of Psychopathology and Cognitive Dysfunction in Schizophrenia. *Biological psychiatry* 94, 108-<br>809 120 (2023).

809 120 (2023).<br>810 49. Hwang, J., Mitz, A.R. & Murray, E.A. NIMH MonkeyLogic: Behavioral control and data<br>811 acquisition in MATLAB. *Journal of neuroscience methods* **323**, 13-21 (2019). 811 acquisition in MATLAB. Journal of neuroscience methods 323, 13-21 (2019).<br>812 50. Hori, Y., et al. D1- and D2-like receptors differentially med

812 50. Hori, Y., et al. D1- and D2-like receptors differentially mediate the effects of<br>813 dopaminergic-transmission-on-cost-benefit-evaluation-and-motivation-in-monkeys. PLoS-biology 811 acquisition in MATLAB. Journal by hearbscreme methods 323, 13-21 (2015).<br>812 50. Hori, Y., et al. D1- and D2-like receptors differentially medi<br>813 dopaminergic transmission on cost-benefit evaluation and motivation in  $812$  50. Hori, H, et al. D1- and D2-like receptors differentially inequate the effects of<br>813 dopaminergic transmission on cost-benefit evaluation and motivation in monkeys. PLoS biology<br>814 **19**, e3001055 (2021).

814 **19**, e3001055 (2021).<br>815 dopaminer 51. dopaminergic transmission and mother est-band monetarine and monetaring<br>816 macagues at 3 Tesla. Neurolmage 16, 283-294 (2002). 814 19, e3001055 (2021).

817 52. Russ, B.E., et al. Common functional localizers to enhance NHP & cross-species 818<br>818 neuroscience imaging research. Neurolmage 237, 118203 (2021). 818 neuroscience imaging research. NeuroImage 237, 118203 (2021).<br>819 53. Fujimoto, A., et al. Resting-State fMRI-Based Screening of

819 53. Fujimoto, A., et al. Resting-State fMRI-Based Screening of Deschloroclozapine in Rhesus<br>820 Macaques Predicts Dosage-Dependent Behavioral Effects. *The Journal of neuroscience : the* 818 1. Incurscience imaging research. *Neuroimage 237*, 118203 (2021).<br>819 53. Fujimoto, A., *et al.* Resting-State fMRI-Based Screening of<br>820 Macaques Predicts Dosage-Dependent Behavioral Effects. The 819 53. Frujinoto, A., et al. Resting-State film Based Screening of Deschloroclozapine in Rhesus<br>820 Macaques Predicts Dosage-Dependent Behavioral Effects. The Journal of neuroscience: the<br>821 official journal of the Socie

821 *official journal of the Society for Neuroscience* **42**, 5705-5716 (2022).<br>822 54. Elorette, C., *et al*. The neural basis of resting-state fMRI functional connectivity in fronto-<br>823 limbic circuits revealed by chemog 821 official journal of the Society for Neuroscience 42, 5705-5716 (2022).<br>822 54. Elorette, C., et al. The neural basis of resting-state fMRI functi<br>823 limbic circuits revealed by chemogenetic manipulation. Nature comm 822 54. Elorette, C., et al. The neural basis of resting-state finith functional connectivity in fronto-<br>823 Ilmbic circuits revealed by chemogenetic manipulation. Nature communications 15, 4669 (2024)<br>824 55. Beaulieu, J.

824 55. Beaulieu, J.M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of<br>825 dopamine receptors. *Pharmacol Rev* 63, 182-217 (2011).<br>826 56. Cox, R.W. AFNI: software for analysis and visualization of funct

826 56. Cox, R.W. AFNI: software for analysis and visualization of functional magnetic resonance<br>827 neuroimages. Computers and biomedical research, an international journal 29, 162-173 (1996). 825 dopartme receptors. *Pharmacol Rev 63*, 182-217 (2011).<br>826 56. Cox, R.W. AFNI: software for analysis and visualization.<br>827 neuroimages. Computers and biomedical research, an int

827 neuroimages. Computers and biomedical research, an international journal 29, 162-173 (1996).<br>828 57. Seidlitz, J., et al. A population MRI brain template and analysis tools for the macaque. 828 57. Seidlitz, J., et al. A population MRI brain template and analysis tools for the macaque.<br>829 Neurolmage 170, 121-131 (2018). 828 57. Seidlitz, J., et al. A population MRI brain template and analysis tools for the macaque.<br>829 Neurolmage 170, 121-131 (2018).<br>830 829 Neurolinge 170, 121-131 (2018).<br>830

	<b>Behavioral</b>				
<b>Subject</b>	Awake fMRI	pharmacology	<b>SCH-rsMRI</b>	<b>HAL-rsMRI</b>	Saline-rsMRI
Ee	Y	Y	Y	N	Y
<b>Ge</b>	Y	N	N	N	N
Me	Y	Y	N	Y	N
St	Y	Y	Y	Y	N
Bu	N	N	N	N	Y
Cy	N	$\mathbf N$	N	N	Y
$\mathbf{P}$ i	N	Y	Y	Y	Y
<b>Wo</b>	N	N	N	N	v

831<br>832<br>833 **Table 1. Assignment of experiments for each subject.** Y and N indicate the condition that the data was collected and not collected, respectively. SCH: SCH-23390 (10 μg/kg), HAL: haloperidol (50 μg/kg). rsMRI: resting-sta data was collected and not collected, respectively. SCH: SCH-23390 (10 μg/kg), HAL:<br>834 haloperidol (50 μg/kg). rsMRI: resting-state fMRI. Note that animals assigned to behavioral<br>835 pharmacology experiments went through 834 haloperidol (50  $\mu$ g/kg). rsMRI: resting-state fMRI. Note that animals assigned to behavioral pharmacology experiments went through all drug treatment conditions. pharmacology experiments went through all drug treatment conditions.



860 atsushi.fujimoto@mssm.edu or peter.rudebeck@mssm.edu

- 862
- 862 **Supplementary Figures 1-5**  863 **Supplementary Tables 1-2**



866 **Supplementary Figure 1. Response time (RT) of monkeys.** Bar graphs show average and SEM<br>867 **of RT** for each stimulus pair in novel (left) and familiar (right) blocks. Symbols represent each 867 of RT for each stimulus pair in novel (left) and familiar (right) blocks. Symbols represent each<br>868 animal Asterisks indicate significant main effect of stimulus pair (\*p<0.05, 2-way repeated-868 animal. Asterisks indicate significant main effect of stimulus pair (\*p<0.05, 2-way repeated-- 869 measures ANOVA).<br>870



872 **Supplementary Figure 2. Context coding in individual monkeys. (A) Anatomical templates<br>873 <b>Supplementary Figure 2. Context coding in individual monkeys. (A)** Anatomical templates 873 showing coronal slices around vlPFC ROI. (**B**) Unthresholded map of F-stats superimposed on<br>874 the anatomical templates in (A) The data for each animal is shown in each row Crosshairs and 874 the anatomical templates in  $(A)$ . The data for each animal is shown in each row. Crosshairs and arrowheads indicate the neak coordinates of vIPEC ROI used in time-course analyses 875 arrowheads indicate the peak coordinates of vlPFC ROI used in time-course analyses.

vIPFC-FC analysis, Context (Novel > Familiar)



 $F > 3.0$  and cluster size  $>14$  voxels

## 877

878 Supplementary Figure 3. Functional connectivity analysis using vlPFC seed. Whole-brain<br>879 **man of E-stats in significant clusters** (p. < 0.05 cluster-corrected generalized psycho-879 map of F-stats in significant clusters ( $p < 0.05$ , cluster-corrected, generalized psycho-<br>880 polygical interaction or gPPD superimposed onto an anatomical template Coronal slices (4.0) 880 physiological interaction or gPPI) superimposed onto an anatomical template. Coronal slices (4.0<br>881 mm apart) are shown from anterior (top left) to posterior (bottom right) planes 881 mm apart) are shown from anterior (top left) to posterior (bottom right) planes.



<sup>884</sup>**Supplementary Figure 4. The impact of outcome and stay/shift decision on vlPFC-ACC C**  885 **functional connection.** The average FC between vlPFC and ACC around the timing of outcome<br>886  $\left($ -2 to +2 seconds after outcome) are plotted for win-stay and win-shift trials (A) and lose-stay 886 (-2 to +2 seconds after outcome) are plotted for win-stay and win-shift trials (A) and lose-stay<br>887 and lose-shift trials (B) for novel (left) and familiar (right) blocks, respectively. Frror hars 887 and lose-shift trials (B) for novel (left) and familiar (right) blocks, respectively. Error bars<br>888 indicate SEM, Asterisks indicate significant EC changes from zero (\*\*n < 0.01 or \*n < 0.05 888 indicate SEM. Asterisks indicate significant FC changes from zero (\*\*p < 0.01 or \*p < 0.05, rank-sum test) 889 rank-sum test).



## vIPFC-MD functional connection

891

**Supplementary Figure 5. Functional connection between vIPFC and MD thalamus around<br>893 the outcome timing (A, B) The average EC between vIPEC and MD around the timing of** 

893 **the outcome timing.**  $(A, B)$  The average FC between vlPFC and MD around the timing of outcome are plotted for novel (left) and familiar (right) blocks respectively. The conventions are

894 outcome are plotted for novel (left) and familiar (right) blocks, respectively. The conventions are<br>895 the same as Figure S4. Asterisks indicate significant FC changes from zero (\*\*n < 0.01 or \*n < 895 the same as Figure S4. Asterisks indicate significant FC changes from zero (\*\*p < 0.01 or \*p <  $\alpha$  0.05 rank-sum test) 896 0.05, rank-sum test).

Peak



Цý.

898<br>899<br>900 899 **Supplementary Table 1. Full list of clusters in the whole-brain analysis that encoded learning context and reward outcome. dlPFC: dorsolateral prefrontal cortex, vlPFC:<br>39** 900 **learning context and reward outcome.** dlPFC: dorsolateral prefrontal cortex, vlPFC:

901 ventrolateral prefrontal cortex, V1: primary visual cortex, V2: secondary visual cortex, V3: third<br>902 visual cortex, V4: fourth visual cortex, TE: anterior inferotemporal cortex, TEO: posterior<br>903 inferotemporal cort

902 visual cortex, V4: fourth visual cortex, TE: anterior inferotemporal cortex, TEO: posterior inferotemporal cortex, dACC: dorsal anterior cingulate cortex, PCC: posterior cingulate cortex.<br>904

903 inferotemporal cortex, dACC: dorsal anterior cingulate cortex, PCC: posterior cingulate cortex.<br>904

Peak



905<br>906<br>907

906 **Supplementary Table 2. Full list of clusters in the functional connectivity (gPPI) analysis<br>907 <b>that encoded learning context in relation to right vlPFC seed timeseries.** dlPFC: dorso-lateral<br>908 **example:** external 907 **that encoded learning context in relation to right vlPFC seed timeseries.** dlPFC: dorso-lateral prefrontal cortex, V1: primary visual cortex, V2: secondary visual cortex, dACC: dorsal anterior cingulate cortex. 908 prefrontal cortex, V1: primary visual cortex, V2: secondary visual cortex, dACC: dorsal anterior<br>909 cingulate cortex.

909 cingulate cortex.