

从强化学习模型到精神病和神经疾病

From reinforcement learning models to psychiatric and neurological disorders

Tiago V Maia^{1,2} and Michael J Frank^{3,4}

¹*Department of Psychiatry, Columbia University, New York, New York, USA.*

²*New York State Psychiatric Institute, New York, New York, USA.*

³*Departments of Cognitive, Linguistic and Psychological Sciences, and Psychiatry and Human Behavior, Brown University, Providence, Rhode Island, USA.*

⁴*Brown Institute for Brain Science, Brown University, Providence, Rhode Island, USA.*

Correspondence should be addressed to T.V.M. (tmaia@columbia.edu) or M.J.F. (michael_frank@brown.edu).

Accepted: 2011 by Nature Neuroscience

(translated by zang jie)

摘要：在过去的十五年中，强化学习模型使人们对多巴胺和皮质—基底神经节—丘脑—皮质回路（CBGTC）的功能有了越来越复杂的了解。最近，这些模型及其所提供的见解已开始用于理解涉及多巴胺能系统和 CBGTC 回路障碍的几种精神病和神经病的重要方面。我们对这种方法及其在帕金森氏病，图雷特氏综合症，注意力不足/多动症，成瘾，精神分裂症和临床前动物模型（用于筛选新型抗精神病药物）中的现有和潜在应用进行了综述。该方法已被证明具有解释性和预测性，对于计算精神病学和计算神经病学的持续增长预示着良好的预兆。

1、引言

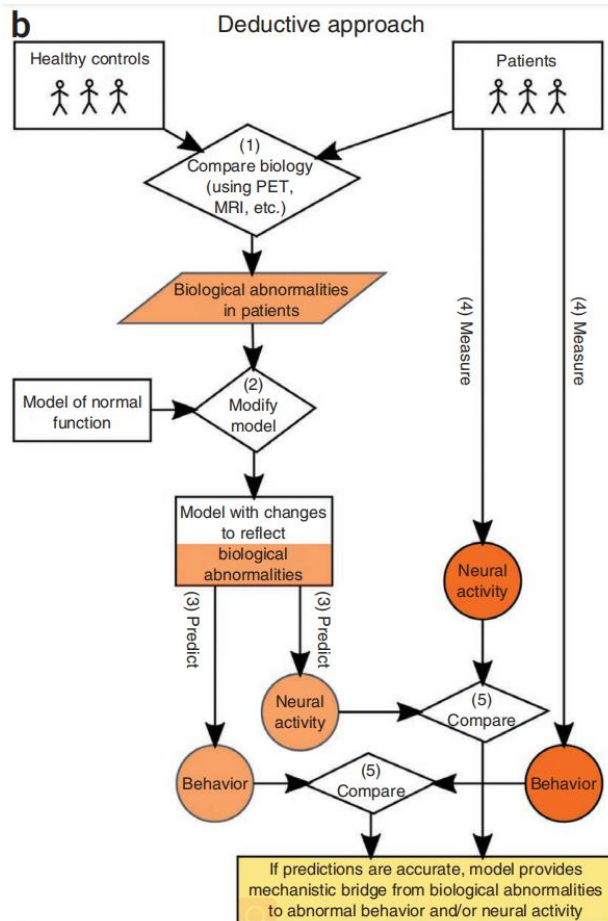
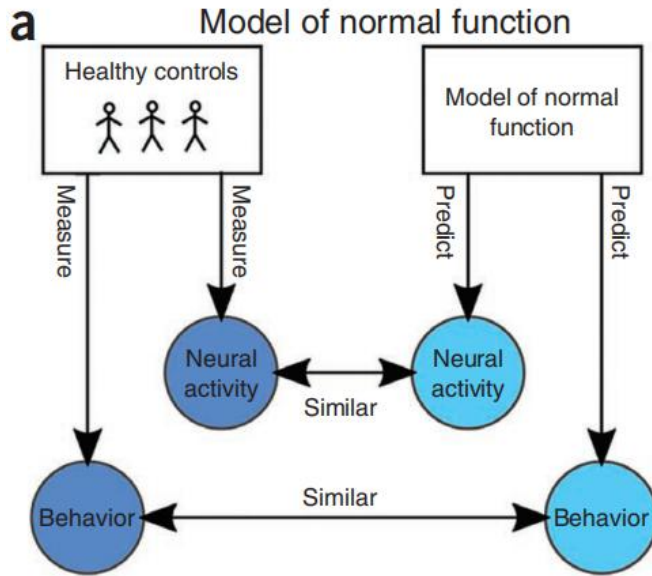
在新版《精神疾病诊断和统计手册》（DSM）的开发过程中，对精神病学领域最新技术的局限性进行了很多辩论。人们普遍认为，必须最终将基于症状的分类系统替换为基于病理生理学的系统¹。但是，目前对精神疾病的神经生物学和遗传学的理解仍然太有限，无法形成疾病学的支架²。这种有限的理解也反映在治疗的最新水平上，大多数精神科药物都是通过偶然发现的，而不是通过合理设计发现的。神经病学通常会用病因学更好地了解疾病（例如，帕金森氏病中的多巴胺能神经元缺失），但即使如此，通常仍不清楚这些病因过程如何产生复杂的症状模式，以及为什么治疗可以减轻某些缺陷，同时加重病情，甚至造成其他问题^{3,4}。解决这个问题的一部分是大脑和大脑的复杂性以及跨越这两者的许多分析层次。计算模型是解决这种复杂性的宝贵工具，因为它们促进了对多种分析水平的机械理解，并且可以解释系统某一组成部分的变化（例如，纹状体 D2 受体密度的增加）如何产生系统-转换为行为变化的级别变化。

强化学习是在神经科学的计算模型和实证研究相结合方面取得实质性进展的领域之一。这种方法已经建立了多巴胺和皮质基底神经节—丘脑—皮质（CBGTC）循环在学习强化阶段（奖励和惩罚）并指导行为以获取奖励和避免惩罚方面的作用的模型⁵。现有的模型解决了这些电路的各种功能，包括巴甫洛夫条件，工具条件及其相互作用。习惯，目标导向的行为及其相互作用；以及激励显着性，激励和活力的相互关联的问题⁵⁻⁹。

以能够获得适合当前动机状态的结果（例如，饥饿时获取食物）并避免有害结果的方式组织行为对于生存至关重要，因此是神经系统的主要组织原则。因此，毫不奇怪，多巴胺能系统和 CBGTC 回路的紊乱在几种精神病和神经病中起关键作用。强化学习模型最近已开始应用于这些疾病，并已显示具有实质性的解释和预测能力¹⁰⁻¹⁴。该方法基于对这些电路在健康个体中执行的计算的理解，并研究病理生理过程如何改变这些计算并产生症状。因此，我们首先回顾这些电路正常功能的计算神经生物学。然后，我们讨论使用强化学习模型已经受益或已经成熟的几种疾病。最后，我们讨论该工作体系对疾病学和治疗学的未来影响。

除了传达强化学习模型如何提供对精神病学和神经病学见解的细节外，我们希望这篇综述也将有助于促进新兴的计算精神病学和计算神经病学学科的发展。现在可以使用一组强大的计算技术来研究病理生理过程及其与行为的关系（图 1）。我

们希望这里回顾的工作可以作为一个示例，促使这些技术在多种模型类型和疾病中得到一致和广泛的使用。



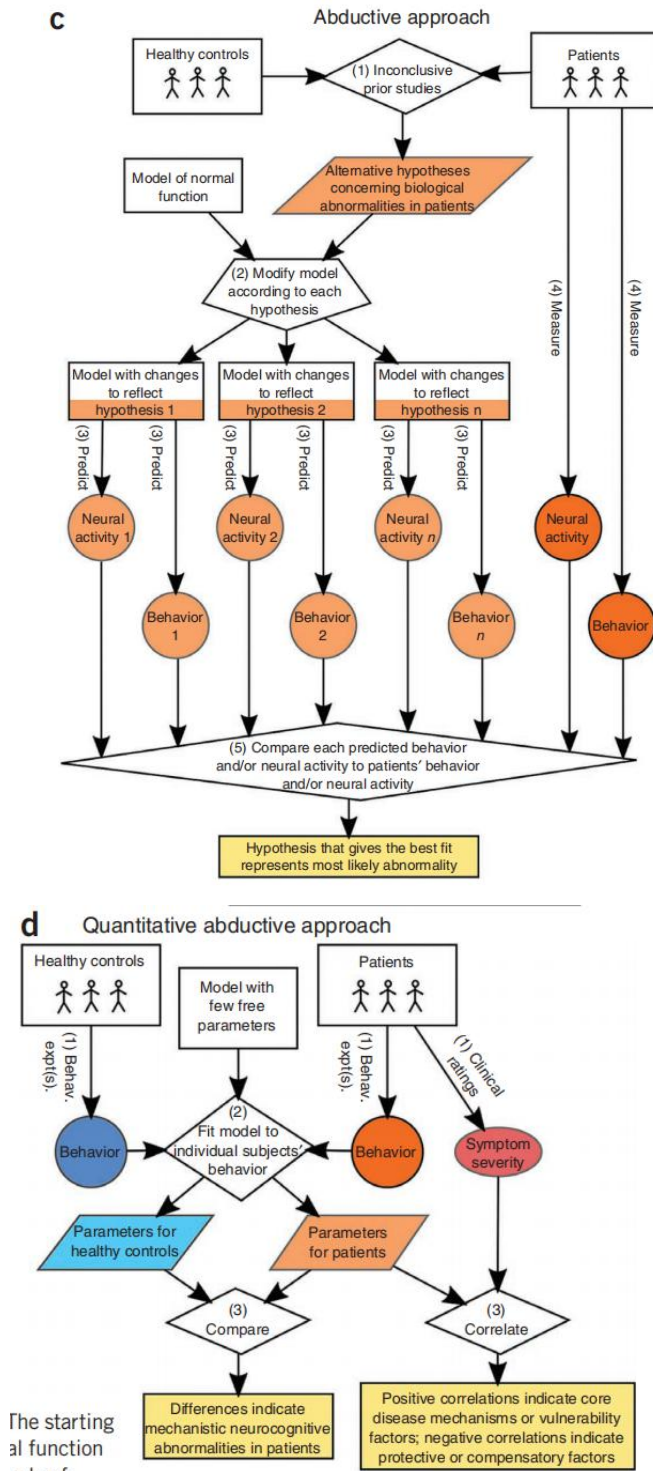


图 1 计算精神病学和计算神经病学原理。(a) 计算精神病学和计算神经病学的起点是正常功能的模型，该模型捕获行为，神经活动或两者的关键方面。各种抽象级别的模型都可能有用（例如，来自机器学习的算法模型或来自计算认知神经科学的神经模型）。然后可以采用几种

方法。(b) 使用详细的神经模型, 可以通过对模型进行原则上的改变来模拟病理生理过程, 该模型模仿所考虑的疾病的生物学变化(例如, 纹状体多巴胺能神经支配的变化)。然后可以探索这些更改在系统级别和行为方面的含义, 从而得出可检验的预测。我们将这种方法称为“演绎”, 因为模型用于重建原因(生物学异常)与其后果(系统级神经活动和行为异常)之间的机械联系。(c) 第二种方法涉及使用模型来尝试推断观察到的神经活动或行为异常的原因。我们将这种方法称为“归纳法”, 因为它涉及从后果(行为或系统级神经活动)到其可能原因(潜在的生物学异常)的推理。在这种方法中, 可以比较与给定疾病中可能的生物学异常有关的替代先验假设, 以确定哪些疾病(如存在)产生与该疾病相同的行为和神经活动异常(T. V. M. 和 B. S. Peterson, 未出版)。(d) 第三种方法是算法使用的次数多于神经模型的使用次数(主要是因为前者倾向于具有较少的参数), 该方法涉及将模型的参数拟合为适合个体任务或一组任务的行为, 然后确定患病和健康受试者组之间是否存在参数差异, 或者参数与疾病严重程度之间是否存在相关性。我们将此方法称为“定量归纳法”, 因为它还涉及从行为到其机械原因的推理。第四种方法(未显示)还涉及使模型适合受试者的行为, 但目标是在逐项试验的基础上估计每个受试者在模型中嵌入的数量的假定内部表示(例如状态值或预测错误)。然后将这些预测的内部表示用作功能成像(例如功能磁共振成像, 脑电图)中的回归指标, 以找到它们的神经相关性, 然后在患病和健康人群中进行比较。这四种方法中的每一种也可以适用于研究治疗效果(例如, 药物或神经外科手术)。此外, 有时可以通过在不同抽象级别上协同使用不同方法或模型来获得额外的杠杆作用。

2、大脑中的强化学习

多巴胺和预测误差。多巴胺神经元编码强化预测错误⁵, 这是许多强化学习模型中的关键信号¹⁵。预测误差表示观察到的结果与预期结果之间的差异: 正的预测误差表示结果要好于预期, 而负的预测误差表示结果要差于预期。阶段性多巴胺-神经元爆发的幅度定量表示正预测误差¹⁶。多巴胺能神经元放电的阶段性减少是否定量地代表负的预测误差存在更大的争议, 因为多巴胺能神经元的低调性发射率意味着这种减少的变化受到一定程度的限制。然而, 多巴胺能神经元放电停顿的持续时间似乎在定量上代表了负面的预测误差。正负预测误差之间的这种不对称编码(正负值的爆发幅度和负数的停顿持续时间)在生物学上是合理的, 这既是因为它允许较低的音调发射, 这在代谢上也是有利的, 而且在这些信号的突触后效应方面也是如此。在 D1 和 D2 受体上的作用, 如下所述。另一种假设或补充假设是, 阴性预测错误可能由血清素编码¹⁸。但是, 这一假设尚未得到充分验证⁵。尽管大多数多巴胺神经元会突然出现阳性预测错误, 但在其他情况下这些神经元中也会有较小比例的

破裂¹⁹。例如，一些多巴胺神经元会突然爆发出阳性预测错误，并爆发出预测阴性事件的阴性事件和刺激⁵。尽管有人尝试通过强化学习来调和这些发现²⁰，但在这一领域仍需要进一步的研究。然而，多巴胺爆发的主要功能是编码阳性预测错误¹⁹。确实，多巴胺神经元的阶段性光遗传学刺激引起了随后对这种刺激发生的位置的偏爱²¹，就像在那个地方已经获得了奖励一样。

基底神经节和动作选择。预测错误用于了解状态（刺激或情况），状态/动作对或两者的值（或“善”）。这些值然后用于选择最佳动作¹⁵。用来批评许多行为和神经发现的这一过程的模型是行为者评论家^{5, 22, 23}。行为者对大脑中动作选择的批评认为，皮质代表了当前状态，基底神经节实现了两个计算模块：评论家，它学习状态值，并且可以在腹侧纹状体中实现^{5,23}，甚至可能在杏仁核中实现。眼眶额叶皮质（OFC）⁵，以及学习刺激与反应关联的演员，可以在背侧纹状体中实现^[5,23]。评论家和演员都使用多巴胺发出的预测误差来更新其估计值（分别是状态值和刺激反应强度）。评论家到基底神经节的映射与以下观点一致：基底神经节对于刺激反应学习至关重要，但纹状体的不同部分（涉及不同的，平行的 CBGTC 回路²⁴）具有不同的特征²⁵。

一种相关但略有不同的观点认为，基底神经节不负责产生动作，而是负责通过促进最适当的动作而抑制竞争动作来在皮质中正在考虑的动作之间进行仲裁²⁶⁻²⁸。换句话说，根据行为批评者的观点，基底神经节完全负责仅根据当前状态来选择动作，而这种替代观点认为，皮质本身最初会产生候选动作（例如，在大脑皮层上）。根据先前在当前状态下执行它们的频率²⁸），然后基底神经节在这些动作之间进行仲裁（可能基于其学习的增强概率^{28,29}）以促进（选通）最佳动作。但是，这些观点之间的共性远远超过了它们之间的分歧，因此，我们将不再进一步研究这种区别。

基底神经节解剖结构由皮质到基底神经节输出结构的直接，间接和超直接路径组成³⁰⁻³²（图 2a）。神经计算模型^[28,33-36]具有完善的理论^[26,30-32]，这些途径在行动选择中起着重要的作用。最初建议用来解释几种神经系统疾病的病理生理学的有影响力的说明^{30,31}表明，直接途径为当前状态的适当行动提供了集中的促进作用，而间接途径抑制了不适合该状态的行动。尽管该报告的原始版本基于对基底神经节的解剖结构和功能的某种过于简化的观点，但最近的发现强烈支持了直接（或 Go）和间接（或 NoGo）路径之间功能的这种分布³⁷⁻³⁹。这种功能分布在基底神经节 Go/NoGo（BG-GNG）模型²⁸中找到形式表达（图 2b），其中选择给定动作的概率与该动作的 Go 和 NoGo 活动之间的差异成比例在当前状态下。与该方案一致的是，电生理

学发现表明，动作的正值和负值在不同的纹状体人群中表现出来，神经元中的活性更高，代表动作的正值，预示着该动作的选择，而神经元的活动性更高。表示某个动作的负值的神经元，表示选择了另一个动作²⁹。BG-GNG 模型进一步显示了原始账户中未考虑的基底神经节解剖结构如何服务于基底神经节功能的其他方面。例如，该模型表明，超直接（或 GlobalNoGo）路径在处理的早期阶段，特别是在高冲突情况下（即，同时强烈激活多个动作时），提供了对所有动作的全局抑制。防止过早，欠佳的响应³⁵。

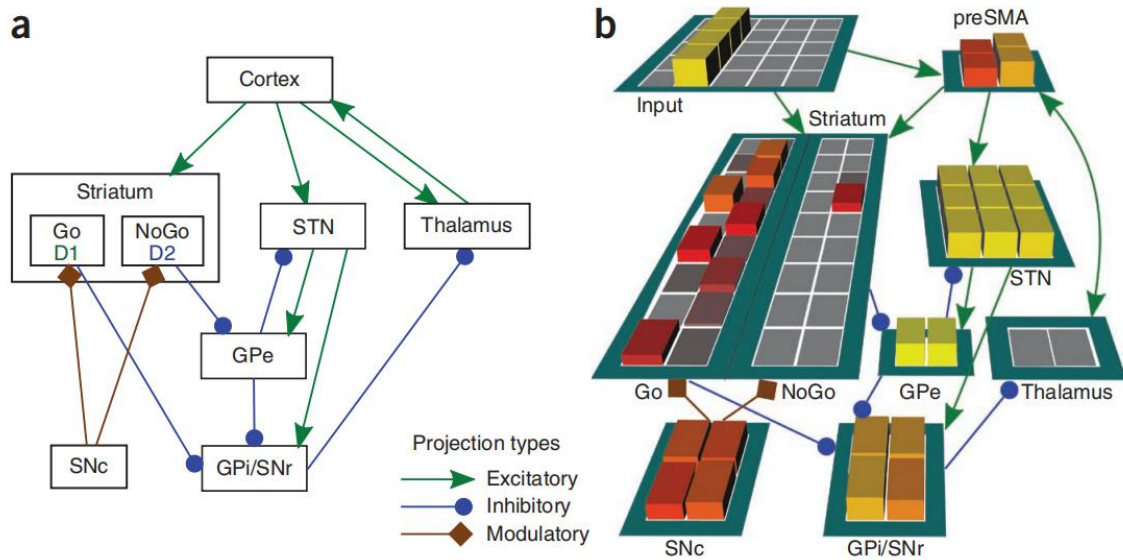


图 2CBGTC 回路的解剖和建模。(a) 解剖。直接途径中的纹状体中棘神经元 (Go 神经元) 主要表达 D1 受体⁴⁰，并直接投射至苍白球内部部分和黑质网状组织 (GPi/SNr)。运动神经元抑制 GPi/SNr，进而导致丘脑去抑制，从而促进相应动作的执行。间接途径的纹状体中棘神经元 (NoGo 神经元) 主要表达 D2 受体⁴⁰，并投射至苍白球外部节段 (GPe)，后者又投射至 GPi/SNr。NoGo 神经元可集中消除 GPi/SNr 对 GPe 的强直抑制作用，从而抑制 GPi/SNr，进而抑制丘脑中相应的作用。丘脑下核 (STN) 中的神经元直接接受从皮质在超直接通路中的投影，并投影到 GPe 和 GPi/SNr。从 STN 到 GPe 和 GPi/SNr 的投影是漫反射²⁶，因此据信它们可以调节所有动作，而不是特定动作。(b) 基底节 Go/NoGo 模型^{28,35}。模型中的连接与 a 中所示的解剖学连接一致。该模型学习将代表当前状态的输入映射到辅助运动前区域 (preSMA) (或 SMA) 中的动作。从输入层到 preSMA 的皮层皮质投影在适合当前状态的 preSMA 候选动作中激活。然后，基底神经节促进 (门控) 最佳动作，即当前状态下具有最佳强化历史的动作，同时抑制其他动作 (丘脑水平)。Go 和 NoGo 单位的分布人口分别代表当前状态下候选行动的正面和负面证据。Go 和 NoGo 途径之间的横向抑制作用确保选择给定动作的概率是该动作的正面和负面证据之间

差异的函数。根据过去的补强史，分别通过多巴胺对纹状体 Go 和 NoGo 单元中 D1 和 D2 受体的作用，了解每种状态下每种作用的正面和负面证据。从输入层到 preSMA 的权重本身是通过 Hebbian 机制学习的，从而使这些皮质预测与 preSMA 在给定状态下执行的概率成比例地激活 preSMA 中的候选动作。在高冲突情况下，STN 阻止了次优反应的促进³⁵。

BG-GNG 模型还显示了直接途径和间接途径如何利用多巴胺 28 传达的预测误差分别学习在每种状态下促进和抑制哪些动作。直接和间接途径分别主要表达 D1 和 D2 受体⁴⁰。在该模型中，当一个动作之后发生多巴胺爆发时，通过 D1 依赖性长期增强作用，可以将直接通路中的皮质口突触变为活跃的 Go 单元（编码状态与所执行动作之间的结合）（LTP），并通过依赖 D2 的长期抑郁症（LTD）削弱了进入有效 NoGo 单位以执行动作的间接途径中的皮层突触。当多巴胺浸入之后进行动作时，发生相反的情况。多巴胺对 D1 和 D2 介导的可塑性的双重影响得到了经验证据的支持⁴¹，模型预测也表明直接和间接途径分别介导了从积极和消极结果的学习³⁸。

多巴胺对 D1 和 D2 受体的动态影响也可以解释为什么爆发幅度和暂停持续时间分别代表正预测误差和负预测误差。D1 和 D2 受体对多巴胺的亲合力分别相对较低和较高⁴²。因此，假设 D1 刺激依赖于多巴胺的阶段爆发，更大的爆发产生更大的刺激。因此，由于正预测误差，爆发幅度对于 D1 介导的 LTP 至关重要。相反，假设 D2 受体被基线多巴胺水平调性刺激。因此，多巴胺能神经元放电中的停顿对 D2 受体的影响取决于多巴胺的重摄取，停顿时间越长，重摄取就越多，因此多巴胺浓度会降低。因此，由于负预测误差，暂停持续时间对于 D2 介导的 LTP 至关重要。LTD 机制在幅度和持续时间分别编码正负预测误差方面也起着关键作用。正预测误差会直接刺激 D2 受体，从而以间接途径产生 LTD。负的预测错误可能不会严重影响 D1 受体，因为强直性多巴胺可能基本上不会刺激 D1 受体。在 BG-GNG 模型中，负预测错误会通过其对 D2 受体的影响（取决于重摄取，因此取决于停顿持续时间），进而通过间接途径抑制直接途径，从而在直接途径中间接产生 LTD。直接途径依赖 LTD。

除了阶段性多巴胺在学习中的作用外，强直性多巴胺还增加了直接（Go）途径的兴奋性，降低了间接（NoGo）途径的兴奋性，因为 D1 受体具有兴奋性（至少对于接受强共价谷氨酸能输入的神经元而言）。D2 和 D2 受体具有抑制作用。因此，补品多巴胺的增加会产生 Go 偏倚，而减少会产生 NoGo 偏倚。使用 BG-GNG 模型进行的仿真表明，这些偏见对动作选择和反应时间有很大影响^[10,28,43]，而补药多巴

胺会促进动作的执行和速度（尤其是先前习得的 Go 和 NoGo 关联）。因此，多巴胺不仅调节学习，而且调节先验学习的表达。

尽管有多个平行的（尽管是相互作用的）⁴⁴CBGTC 循环穿过基底神经节²⁴，但常见的划分是感觉运动，缔合和边缘循环，它们分别连接到感觉运动皮层区，背外侧前额叶皮层，OFC 和前扣带回皮层⁴⁵。BG-GNG 模型不仅应用于涉及感觉运动回路的运动动作选择，而且还应用于认知“动作”的选择，尤其是工作记忆更新^{46,47}，很可能涉及联想回路。在后一种情况下，想法是 Go 信号有助于将刺激门控到工作记忆中，而 NoGo 信号阻止这种门控（例如，对于与任务无关的刺激）。与这个想法一致，基底神经节损伤会干扰仅将与任务相关的刺激选择性地进入工作记忆的能力⁴⁸。

试图推测直接途径和间接途径分别学习状态的正值和负值是很诱人的。与这一观点一致，在条件性位置偏好中，这涉及学习位置（状态）的值并取决于伏隔核⁴⁹，学习位置的正值取决于直接途径而非间接途径中的神经传递³⁸。此外，在学习过程中，可卡因诱导的条件性位置偏爱通过伏隔核中直接途径的光遗传激活而增加，而通过间接途径的激活而降低。

假设 OFC 会为状态增强值实现工作记忆，从而在其他情况下帮助指导这些值快速变化时的操作选择⁵¹。如果是这样，边缘循环还可以对“动作”实施 Go 和 NoGo，以决定何时应更新 OFC 中维护的值，以确保自适应，灵活的行为。

3、临床应用

这些思想和模型同时应用于运动领域和认知领域时，可以解释多种疾病的各种发现。鉴于多巴胺在强化学习中的核心作用，我们将重点研究多巴胺能参与性强的疾病：帕金森氏病，图雷特综合症，注意力缺乏/多动症（ADHD），药物成瘾和精神分裂症。当然，这些疾病通常还涉及非多巴胺能紊乱（例如，图雷特氏综合症涉及纹状体中间神经元异常⁵²，多动症涉及正肾上腺素能异常⁵³等）。此外，在某些情况下，多巴胺能障碍本身可能是上游异常引起的。例如，精神分裂症涉及中脑边缘的增加和中皮层的多巴胺的减少^{54, 55}，这两者可能至少部分是由于 NMDA 受体异常⁵⁶对多巴胺神经元的皮质控制失调引起的。（事实上，NMDA 异常可能通过破坏皮质吸引子的稳定性直接导致许多精神分裂症症状⁵⁷）这些疾病的综合模型最终将不得不整合这些各种异常。但是，对上述多巴胺和 CBGTC 电路的计算功能所

阐述的少数原理的更加有限的关注，为理解这些疾病的多个方面提供了重要的杠杆作用。

帕金森氏病。帕金森氏病中的多巴胺能细胞死亡导致纹状体多巴胺减少，从而导致 NoGo^{10,28} 的发展趋势。使用 BG-GNG 模型进行的模拟表明，这种趋势不仅可以解释帕金森氏病的运动不足症状，还可以解释这种疾病伴随的各种认知缺陷^{10, 28, 58}。例如，如上所述，该模型建议关联循环中的 Go 和 NoGo 信号分别促进和阻止工作存储器更新^[46,47]。因此，帕金森氏病中过度兴奋的 NoGo 途径应导致工作记忆更新不足，同时增加对于干扰物的抵抗力。此外，1-DOPA 和多巴胺激动剂应逆转这些作用。实证研究证实了这些预测^{59,60}。

受试者倾向于从积极或消极中学习的倾向可以使用概率选择任务来评估活动反馈（分别为 Go 和 NoGo 学习）⁵⁸（图 3）。健康的对照同样擅长学习以取得积极的成果（学习），避免负面的结果（NoGo 的学习）^[4,58]（尽管健康对照中 Go 和 NoGo 学习之间存在个体差异，这是由遗传预测的）直接和间接途径中影响多巴胺功能的变异³⁷）。与 BG-GNG 模型的预测一致，帕金森氏病的未经药物治疗的人在 NoGo 学习中胜于在 Go 学习中⁵⁸。药物逆转了这些偏见：与帕金森氏病或对照组相比，帕金森氏病的药物个体在 Go 上比在 NoGo 学习时更好，在 NoGo 学习时比在帕金森氏病或对照中未受治疗的个体更糟糕^{4,58}。在假设多巴胺能药物在阴性预测错误期间会减少多巴胺浸润的情况下，模型也可以预测这些药物作用（因为多巴胺能药物会导致多巴胺神经元放电暂停期间突触后多巴胺受体持续被占用）⁵⁸。负面预测错误的这种钝化会减少对负面结果的学习，从而导致 NoGo 学习的不足。对于帕金森氏病的非药物治疗和药物治疗患者，也使用其他一些任务获得了类似的发现，这些任务还评估了受试者从阳性或阴性结果中学到的知识¹⁰。从积极和消极结果学习中的这些不对称现象可能具有临床意义。例如，药物引起的从积极的结果中学习而不是从消极的结果中学习的趋势可能解释了为什么药物在帕金森氏病的一部分患者中诱发了病理性赌博。

BG-GNG 模型还表明，超直接途径的关键节点-丘脑下核（STN）提供了一个全局 NoGo 信号，该信号在动作选择期间会暂时抑制所有动作³⁵。STN 的激活以及全局 NoGo 信号的激活是通过响应冲突量进行动态调制的（另请参见参考资料³³）。因此，在发生高度冲突的情况下，GlobalNoGo 信号特别强，正在考虑采用相对相似的加固历史的行动。在这种情况下，全局 NoGo 抑制为赢得比赛的最佳动作提供了时间，从而阻止了过早的，次佳的动作。因此，该模型预测，STN 处理的中断（例

如，通过深部脑刺激（DBS）将破坏受试者在这种高冲突情况下的减速能力，从而导致更快但次优的响应⁴。该预测已在接受 STNDBS 的帕金森氏病患者中得到实验证实（并且当 DBS 关闭时，这一缺陷得以解决）⁴。此外，如模型所预测的那样，多巴胺能药物和 DBS 具有双重分解作用：药物会影响从积极和消极结果中学习的不对称性，但不会影响在高冲突情况下的减慢能力，而 DBS 具有相反的作用⁴。因此，药物和 DBS 对帕金森氏病个体的现实生活行为的不利影响可能是由不同的神经认知过程中断引起的。

图雷特氏综合症。图雷特氏综合症的特征是反复发作，刻板动作和发声，称为抽动症。假设抽动可反映出为 tic61 提供 Go 信号的纹状体神经元子集的异常激活。Tourette 的临床药理学和动物实验工作的证据表明，抽搐可能是由于相对于间接（NoGo）途径，直接（Go）的过度兴奋性或可塑性引起的。首先，D2 受体阻滞剂是 Tourette 的标准药物治疗方法，可增强间接途径（因为 D2 受体具有抑制作用）。其次，将多巴胺，苯丙胺或 D1 和 D2 激动剂的组合给药到纹状体中，它们同时增强直接途径（通过 D1 受体）和抑制间接途径（通过 D2 受体），引起动物刻板印象^{62,63}。一些发现表明，对这两种途径的影响协同作用以诱导刻板印象。例如，通过用 D1 或 D2 拮抗剂预处理可减少纹状体苯丙胺给药引起的刻板印象⁶²。

图雷特综合症中相对于 NoGo 活动而言，相对于 Go 而言过量的 Go 可能是纹状体中的多巴胺或多巴胺能受体敏感性过高的结果。确实，Tourette 综合征与纹状体中多巴胺的释放，多巴胺能神经支配和 D2 受体的增加有关（尽管在后两项研究中，后两种改变的证据不一致）^{55,64}。所有这些潜在的改变将导致相对于 NoGo 途径的 Go 增强。与可能过量的纹状体多巴胺相一致的是，患有图雷特氏综合症的未经药物治疗的人从奖励中学到的要比从惩罚中学到的更好。这些偏见与帕金森氏病的非医学个体所显示的偏见相反^[4,58,65]，反映出以下事实：在有限的程度上，这两种疾病的症状也相反：帕金森氏病是一种运动不足的疾病，而图雷特氏综合症是一种运动过度疾病。实际上，药物可以逆转这些学习偏见，使患有 L-DOPA 和多巴胺激动剂的帕金森氏病患者的表现与未接受 Tourette 综合征的非药物治疗的患者相似（从奖励中学习比从惩罚中学习更好），并使具有 D2 受体阻滞剂的 Tourette 综合征患者发挥作用。与患有帕金森氏病的未接受药物治疗的人类似（从惩罚中学到从奖励中学到）⁶⁵。对前一个发现的机理解释已经讨论过了。后一个发现与模拟结果相吻合，模拟结果表明，D2 阻滞增加了间接途径的兴奋性和可塑性，从而促进了 NoGo 学习。通过 D2 阻断对 NoGo 学习的这种增强表明，D2 拮抗剂的急性给药可能是行为疗法

的有效辅助手段，该行为疗法通过为抽动症赋予负面价值（例如，应急管理，抽动症后进行惩罚或抽搐的缺失会得到积极的加强，或者是大量的消极练习，在这种情况下，抽搐会由于疲劳而变得厌恶。尚不清楚急性 D2 阻滞是否也可以作为习惯性逆转训练的辅助手段，这是目前对 Tourette 综合征的最佳行为治疗方法，目前尚不清楚，因为该程序显然并不涉及厌恶性学习。

相对于运动 CBGTC 回路中的 NoGo 活动而言，过高的围棋动作也可能解释了图雷特氏综合症的一个显着特征就是事前的冲动。假设这些冲动是由辅助运动区域（SMA）的异常激活引起的，因为 SMA 的电刺激会引起类似的冲动⁶⁶。与这个想法一致，SMA 在 tic⁶⁷ 之前是活跃的，而 tic 的 SMA 激活要比健康对照（Z.Wang, TVM, R.Marsh, T）进行的模仿 tic（且在视觉上与 tic 无区别）的运动更大。Colibazzi, A. Gerber 等，未发表的数据。SMA 是电机 CBGTC 回路的主要目标²⁴，因此 SMA 异常激活可能是该回路中相对 Go 活动过多的结果。替代地或另外，如在 BG-GNG 模型中一样，异常 SMA 激活可能是由状态与 SMA（或 preSMA）之间的皮层投射引起的（参见图 2b）。在模型中，这些联系是通过 Hebbian 机制学习的，因此在一个或多个州（最初是由于相对于 NoGo 活动过多的围棋），基底神经节反复控制 tic 会加强这些州与 tic 的运动计划之间的联系在 SMA 中。然后，在这些状态下，SMA 运动计划往往会被激活，从而产生冲动（然后可以被基底神经节控制成实际的抽动发射，但也可以避免这种行为）。如果正确的话，这将解释图雷特氏综合症的重要临床特征。例如，它将解释抽动的状态依赖性（即抽动并非在所有情况下都频繁发生的事实）。这也将解释为什么随着时间的流逝，习惯性逆转训练等阻止抽动表现的治疗会导致冲动的减少：反复激活抽动引起状态的抽动而没有相应的抽动释放，会导致希伯来人不了解这些状态之间的联系和抽动。从更广泛的意义上讲，这表明图雷特氏综合症涉及一个恶性循环：进行抽动会增强执行抽动的冲动（通过皮层皮质希伯来语学习），从而增加抽动的倾向（通过皮层皮质激活抽动运动计划）在 SMA 中，这增加了该抽动的基底神经节门控的可能性）。

多动症。ADHD 的特征在于注意力不集中，活动过度和冲动的异常水平。ADHD 的经典理论是，它是由抑制性控制的主要缺陷导致的，该缺陷导致执行功能的一些缺陷⁶⁸。另一个著名的理论是，注意力缺陷多动症是由于延迟奖励的过度贴现而引起的⁶⁹。“多路径”说明表明行政功能障碍和过度的延误折扣都涉及⁷⁰。多动症似乎涉及功能低下的多巴胺能系统^{11,71}。

一种强化学习理论表明，腹侧纹状体中的多巴胺决定了一种可以向前看的强化学习系统的折扣因子（未来强化剂相对于直接强化剂的折扣程度），因为它包括环境内部模型⁷²。ADHD 腹侧纹状体中多巴胺的减少会产生较小的折现因子，从而导致延迟报酬的过度折现⁷²。这个想法似乎与一些环境证据相吻合：全身施用多巴胺阻滞剂可增加延迟贴现，而通过服用精神刺激药或选择性多巴胺再摄取抑制剂来增加多巴胺通常可减少延迟贴现⁷³。但是，伏隔核中的多巴胺消耗似乎不会产生过多的延迟贴现⁷⁴，而 OFC 中的多巴胺消耗确实会产生⁷⁵。因此，OFC 中的多巴胺含量低可能会导致 ADHD 的延误折扣过大。

生物物理现实的神经计算模型表明多巴胺可稳定 PFC 中的表达。鉴于 PFC 的自上而下的偏见在注意力，认知控制（包括抑制性控制）和工作记忆中的关键作用⁷⁷，假定的 ADHD 中低 PFC 多巴胺可能是所有这些执行功能不足的基础。这个想法与多动症的主要问题在于抑制控制，其余方面的问题是次要的⁶⁸。然而，ADHD 的认知功能障碍并不仅是由低 PFC 多巴胺引起的：关联环中的纹状体多巴胺含量低可能会导致例如工作记忆门控减少。相反，PFC 功能障碍不仅需要引起执行功能障碍：侧面 PFC 似乎参与了选择延迟奖励的能力⁷⁸（就像 OFC 一样），因此功能障碍的 PFC 也可能导致过度的延迟折扣。因此，神经生物学异常与认知和动机缺陷之间的关系可能不是一对一的，并且在多动症患者中可能会有所不同。

吸毒成瘾。纹状体多巴胺，类似阶段性变化在药物增强作用中的重要性⁷⁹使成瘾成为增强学习模型的自然候选者。有影响力的强化学习理论表明，这些快速增加的功能是每次服用药物时都会出现的积极预测错误¹⁴。药物的这种作用与自然奖励的作用形成对照，后者的预期误差在预期奖励后变为零。反复出现的由药物引起的阳性预测误差会导致导致药物接收的状态（或作用）值无限增加，从而促使人们强迫性使用药物。该理论解释了成瘾的重要特征¹⁴，但随后的研究驳斥了其成瘾性的一项主要预测。该理论预测，当将药物用作无条件刺激时，不应发生阻断作用。阻断是一种过程，其中刺激 A 首先与无条件刺激配对，然后，同时呈现 A 和另一刺激 B 与无条件刺激配对。通常，B 不会发生任何学习，因为 A 已经预测到了无条件的刺激，因此没有支持新学习的预测错误。但是，如果药物总是产生积极的预测误差，那么受试者应该学会将 B 与无条件刺激相关联，但是他们不会。这一发现促使人们寻找替代成瘾的强化学习方法。

一种方法是从标准时差学习过渡到平均奖励强化学习⁸¹。在平均奖励强化学习中，相对于使用过去强化的缓慢变化的加权平均值计算出的平均强化值 R_t 评估强

化。因此，在时间 t 的“有效”强化是 $r_t - R_t$ ，其中 r_t 是收到的强化。这种上瘾方法的部分动机是要捕捉对长期吸毒引起的自然奖励的敏感性下降。凭直觉，如果毒品极度强化，长期吸毒会使 R_t 膨胀，从而使自然奖励的强化程度降低。但是，该模型通过在每次使用毒品时人为地夸大 R_t 进一步增加了这种效果。在长期吸毒之后，该模型显示出对自然奖励的敏感性降低，并且还显示出阻断作用和其他相关作用⁸¹。

毫无疑问，将继续寻找简单的单因素成瘾理论，但成瘾可能涉及强化学习的多个方面⁸²。例如，长期吸毒会导致重要的强化学习大脑区域（例如，OFC83）的功能和结构发生变化，从而进一步失调强化学习，并可能有助于维持或加剧成瘾。另一个例子是，小鼠的光遗传学发现表明，在给药过程中直接或间接的途径刺激分别增加或减少了药物的增强作用⁵⁰，这表明相对于直接途径的活性降低的间接作用可能是成瘾的危险因素。实际上，间接途径活动的减少也可以解释对成瘾的负面结果敏感性降低。

精神分裂症。精神分裂症的特征是阳性症状（例如，妄想和幻觉），阴性症状（例如，快感和无力）和认知症状（例如，注意力和认知控制障碍）。精神分裂症涉及纹状体中过量的多巴胺和 D2 受体，但 PFC 中的多巴胺减少^{54,55}。

一种理论基于多巴胺发出激励性显著性的观点⁸⁴，表明精神分裂症中多巴胺能放电失调会给人以异常显著性的感觉，思想和记忆，而这种异常显著性经历是幻觉和幻觉的基础⁸⁵。另一种理论认为，心理是由于异常的预测错误导致的，这些错误产生了不合适的联想，因果归因和注意力显著⁸⁶。患有精神病的个体在预测错误期间确实表现出异常的神经活动^[87,88]，但这些异常与精神病之间的因果关系仍然是假设的。

消极症状可能是反映奖励敏感性降低的原因，但是对此的证据是混杂的。与此假设相符，精神分裂症患者对阳性预测错误的神经反应减弱，与更强的意志相关的壳核反应较弱⁸⁹。精神分裂症患者的围棋学习也减少^[13,43,90]，但这些缺陷似乎与不良症状^[43,90]无关。在这些研究中，阴性症状与 PFC 功能障碍的指标相关^[13,43]。尤其是无语症，与减少不确定性驱动的探索有关（在这种探索中，替代性行动是根据其补强统计的不确定性相对于当前首选行动的补强统计的不确定性成比例地进行探索的）⁹⁰。这可能反映出这样一个事实，即使用的性欲低下的评估部分取决于个人从事愉悦活动的频率，这可能取决于诸如勘探之类的战略过程。

减少对阳性预测误差的纹状体反应⁸⁹和精神分裂症的围棋学习减少^{13,43,90}提示减少了相性纹状体多巴胺。与精神分裂症一样，强直性纹状体多巴胺的增加⁵⁴可能通过抑制性自体受体使多巴胺减少。实际上，精神分裂症患者表现出整体的Go偏向性（与强直性纹状体多巴胺增加相一致）以及Go的学习减少（与阶段性纹状体多巴胺减少相一致）⁴³。在关联循环中，强音偏音可能会产生带有无关信息的PFC表示的过度更新。减少PFC多巴胺可能会增加PFC表现的不稳定性⁷⁶。所导致的PFC表示形式的极度不稳定性可能是认知症状的基础，并有助于阳性症状。

临床前动物模型。强化学习模型还可以阐明用于测试新药物的临床前动物模型。一个例子是使用条件回避来筛查抗精神病药⁹¹。在有条件的回避中，除非动物在警告刺激发生后但在电击之前执行一定的回避反应，否则警告刺激后将进行电击。回避响应会导致从负值状态（预期会发生电击）过渡到中性值状态（会预期无电击），因此会引起正向预测误差²²。假设这些积极的预测错误可加强行为批评体系中警告刺激与避免反应之间的刺激-反应关联²²。与这个想法相一致，腹侧纹状体（预期会损害批评家）和黑纹状体多巴胺能投射（预期可防止将预测错误传递给演员）的病变破坏了回避学习^{92,93}。仅在训练后才破坏背侧纹状体中的多巴胺能信号传导（例如，通过黑纹状体病变或背侧纹状体D2阻滞）不会破坏回避性能^{92,94}。这些发现也与行动者评论的预测相一致，因为背侧纹状体中预测错误的多巴胺能信号对于刺激反应学习是必需的，但对于刺激反应表达不是必需的。

与背侧纹状体不同，训练后伏隔核中的D2阻断会破坏回避表现⁹⁴。实际上，有条件回避测试抗精神病药物的标准用途是在训练后对其进行管理。标准的发现是，低抗精神病药剂量会破坏回避行为，但不能避免持续的电击，伏击核是一种介导的作用⁹¹。伏隔核中的多巴胺可以调节仪器行为的激活^{95,96}，可能是通过纹状体—黑质—纹状体的螺旋结构使伏隔核影响背侧纹状体^[44,96]。因此，抗精神病药引起的回避反应的中断可能反映了工具行为的激活减少。逃生将受到较小的影响，因为疼痛的即时性会激活先天的飞行响应。在食欲方面存在类似的情况：全身或伏打小剂量多巴胺拮抗剂的使用会破坏对食物的工具压迫，而不会影响食物进食或食用⁹⁵。

the activation of instrumental behavior^{95,96}, possibly via the striato-nigro-striatal spirals that allow the nucleus accumbens to influence the dorsal striatum^{44,96}. The disruption of avoidance responding by antipsychotics may therefore reflect a decrease in the activation of instrumental behavior. Escapes would be less affected because the immediacy of pain would activate an innate flight response. An analogous situation is found in the appetitive domain: systemic or intra-accumbens administration of low doses of dopamine antagonists disrupts instrumental lever-pressing for food without affecting food approach or consumption⁹⁵.

Conclusions

We focused on a small number of reinforcement learning principles and examined how they can shed light on multiple disorders. Other aspects of reinforcement learning and related computational approaches also seem likely to be relevant for psychiatric and neurological disorders. For example, models of the role of the OFC in reinforcement learning⁵¹ may be relevant for obsessive-compulsive disorder, which involves prominent OFC disturbances⁹⁷; models of the role of serotonin in reinforcement learning¹⁸ may be relevant for disorders that involve serotonergic abnormalities; and models of Pavlovian conditioning and extinction^{98,99} may be relevant for some anxiety disorders (and have in fact already been shown to explain complex findings in fear conditioning in humans¹⁰⁰).

To conclude, reinforcement learning models have been used to explain a wealth of findings across several psychiatric and neurological disorders. Although disorders as seemingly disparate as Parkinson's disease, Tourette's syndrome, ADHD, schizophrenia and addiction might seem to have little in common, they all involve disturbances in dopamine and CBGTC loops. The work reviewed above demonstrates that a mechanistic, computationally grounded understanding of the functions of these circuits sheds important light on all of these disorders. This approach relates to the new Research Domain Criteria initiative from the US National Institute of Mental Health, which calls for research that cuts across diagnostic criteria and focuses instead on neurocognitive domains and how they go awry in a variety of DSM-defined conditions. The work reviewed above exemplifies this strategy. This work also demonstrates the new level of theoretical sophistication that computational psychiatry and computational neurology bring to the venerable disciplines of psychiatry and neurology. Such theoretical sophistication and depth is essential if we are to fulfill the promise of a neuroscience-based, mechanistically detailed approach to diagnosis and treatment, which many agree should characterize the psychiatry and neurology of the future.

ACKNOWLEDGMENTS

Preparation of this article was funded by a Research Associate Award from the New York State Psychiatric Institute and the Research Foundation for Mental Hygiene, by National Institute of Mental Health grant R01 MH080066 and by a grant from the Michael J. Fox Foundation for Parkinson's Research.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/natureneuroscience/>.

Reprints and permissions information is available online at <http://www.nature.com/reprintsandpermissions/>.

- Charney, D.S. *et al.* Neuroscience research agenda to guide development of a pathophysiologically based classification system. in *A Research Agenda for DSM-V* (eds. Kupfer, D.J., First, M.B. & Regier, D.A.) 31–83 (American Psychiatric Association, Washington, D.C., 2002).
- Hyman, S.E. Can neuroscience be integrated into the DSM-V? *Nat. Rev. Neurosci.* **8**, 725–732 (2007).
- Cools, R. Dopaminergic modulation of cognitive function—implications for L-DOPA treatment in Parkinson's disease. *Neurosci. Biobehav. Rev.* **30**, 1–23 (2006).
- Frank, M.J., Samanta, J., Moustafa, A.A. & Sherman, S.J. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. *Science* **318**, 1309–1312 (2007).
- Maia, T.V. Reinforcement learning, conditioning, and the brain: successes and challenges. *Cogn. Affect. Behav. Neurosci.* **9**, 343–364 (2009).
- Dayan, P., Niv, Y., Seymour, B. & Daw, N.D. The misbehavior of value and the discipline of the will. *Neural Netw.* **19**, 1153–1160 (2006).
- McClure, S.M., Daw, N.D. & Montague, P.R. A computational substrate for incentive salience. *Trends Neurosci.* **26**, 423–428 (2003).
- Dayan, P. & Balleine, B.W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).
- Niv, Y., Daw, N.D., Joel, D. & Dayan, P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl.)* **191**, 507–520 (2007).
- Wiecki, T.V. & Frank, M.J. Neurocomputational models of motor and cognitive deficits in Parkinson's disease. *Prog. Brain Res.* **183**, 275–297 (2010).
- Frank, M.J., Santamaria, A., O'Reilly, R.C. & Willcutt, E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* **32**, 1583–1599 (2007).
- Frank, M.J., Scheres, A. & Sherman, S.J. Understanding decision-making deficits in neurological conditions: insights from models of natural action selection. *Phil. Trans. R. Soc. Lond. B* **362**, 1641–1654 (2007).
- Waltz, J.A., Frank, M.J., Robinson, B.M. & Gold, J.M. Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol. Psychiatry* **62**, 756–764 (2007).
- Redish, A.D. Addiction as a computational process gone awry. *Science* **306**, 1944–1947 (2004).
- Sutton, R.S. & Barto, A.G. *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
- Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
- Bayer, H.M., Lau, B. & Glimcher, P.W. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* **98**, 1428–1439 (2007).
- Daw, N.D., Kakade, S. & Dayan, P. Opponent interactions between serotonin and dopamine. *Neural Netw.* **15**, 603–616 (2002).
- Schultz, W. Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* **6**, 24 (2010).
- Morris, G., Schmidt, R. & Bergman, H. Striatal action-learning based on dopamine concentration. *Exp. Brain Res.* **200**, 307–317 (2010).
- Tsai, H.C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324**, 1080–1084 (2009).
- Maia, T.V. Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn. Behav.* **38**, 50–67 (2010).
- O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
- Alexander, G.E., DeLong, M.R. & Strick, P.L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
- Yin, H.H. & Knowlton, B.J. The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* **7**, 464–476 (2006).
- Mink, J.W. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**, 381–425 (1996).
- Prescott, T.J., Gurney, K. & Redgrave, P. Basal ganglia. in *The Handbook of Brain Theory and Neural Networks* (ed. Arbib, M.A.) 147–151 (MIT Press, Cambridge, Massachusetts, 2003).
- Frank, M.J. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J. Cogn. Neurosci.* **17**, 51–72 (2005).
- Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
- Albin, R.L., Young, A.B. & Penney, J.B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12**, 366–375 (1989).
- DeLong, M.R. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* **13**, 281–285 (1990).
- Nambu, A., Tokuno, H. & Takada, M. Functional significance of the cortico-subthalamo-pallidal 'hyperdirect' pathway. *Neurosci. Res.* **43**, 111–117 (2002).
- Bogacz, R. & Gurney, K. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput.* **19**, 442–477 (2007).
- Brown, J.W., Bullock, D. & Grossberg, S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Netw.* **17**, 471–510 (2004).
- Frank, M.J. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw.* **19**, 1120–1136 (2006).
- Gurney, K., Prescott, T.J. & Redgrave, P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol. Cybern.* **84**, 411–423 (2001).
- Frank, M.J. & Fossella, J.A. Neurogenetics and pharmacology of learning, motivation, and cognition. *Neuropsychopharmacology* **36**, 133–152 (2010).
- Hikida, T., Kimura, K., Wada, N., Funabiki, K. & Nakanishi, S. Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. *Neuron* **66**, 896–907 (2010).

39. Kravitz, A.V. *et al.* Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* **466**, 622–626 (2010).
40. Gerfen, C.R. Molecular effects of dopamine on striatal-projection pathways. *Trends Neurosci.* **23**, S64–S70 (2000).
41. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D.J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321**, 848–851 (2008).
42. Creese, I., Sibley, D.R., Hamblin, M.W. & Leff, S.E. The classification of dopamine receptors: relationship to radioligand binding. *Annu. Rev. Neurosci.* **6**, 43–71 (1983).
43. Waltz, J.A., Frank, M.J., Wiecki, T.V. & Gold, J.M. Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology* published online, doi:10.1037/a0020882 (22 November 2010).
44. Haber, S.N. The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* **26**, 317–330 (2003).
45. Postuma, R.B. & Dagher, A. Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cereb. Cortex* **16**, 1508–1521 (2006).
46. Frank, M.J., Loughry, B. & O'Reilly, R.C. Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn. Affect. Behav. Neurosci.* **1**, 137–160 (2001).
47. O'Reilly, R.C. & Frank, M.J. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* **18**, 283–328 (2006).
48. Baier, B. *et al.* Keeping memory clear and stable: the contribution of human basal ganglia and prefrontal cortex to working memory. *J. Neurosci.* **30**, 9788–9792 (2010).
49. Tzschentke, T.M. Measuring reward with the conditioned place preference paradigm: a comprehensive review of drug effects, recent progress and new issues. *Prog. Neurobiol.* **56**, 613–672 (1998).
50. Lobo, M.K. *et al.* Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* **330**, 385–390 (2010).
51. Frank, M.J. & Claus, E.D. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* **113**, 300–326 (2006).
52. Kalanithi, P.S. *et al.* Altered parvalbumin-positive neuron distribution in basal ganglia of individuals with Tourette syndrome. *Proc. Natl. Acad. Sci. USA* **102**, 13307–13312 (2005).
53. Biederman, J. & Spencer, T. Attention-deficit/hyperactivity disorder (ADHD) as a noradrenergic disorder. *Biol. Psychiatry* **46**, 1234–1242 (1999).
54. Guillin, O., Abi-Dargham, A. & Laruelle, M. Neurobiology of dopamine in schizophrenia. *Int. Rev. Neurobiol.* **78**, 1–39 (2007).
55. Nikolaus, S., Antke, C. & Müller, H.W. *In vivo* imaging of synaptic function in the central nervous system. II. Mental and affective disorders. *Behav. Brain Res.* **204**, 32–66 (2009).
56. Stahl, S.M. Beyond the dopamine hypothesis to the NMDA glutamate receptor hypofunction hypothesis of schizophrenia. *CNS Spectr.* **12**, 265–268 (2007).
57. Rolls, E.T., Loh, M., Deco, G. & Winterer, G. Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat. Rev. Neurosci.* **9**, 696–709 (2008).
58. Frank, M.J., Seeberger, L.C. & O'Reilly, R.C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**, 1940–1943 (2004).
59. Moustafa, A.A., Sherman, S.J. & Frank, M.J. A dopaminergic basis for working memory, learning and attentional shifting in Parkinsonism. *Neuropsychologia* **46**, 3144–3156 (2008).
60. Cools, R., Miyakawa, A., Sheridan, M. & D'Esposito, M. Enhanced frontal function in Parkinson's disease. *Brain* **133**, 225–233 (2010).
61. Mink, J.W. Basal ganglia dysfunction in Tourette's syndrome: a new hypothesis. *Pediatr. Neurol.* **25**, 190–198 (2001).
62. Delfs, J.M. & Kelley, A.E. The role of D1 and D2 dopamine receptors in oral stereotypy induced by dopaminergic stimulation of the ventrolateral striatum. *Neuroscience* **39**, 59–67 (1990).
63. Walters, J.R., Bergstrom, D.A., Carlson, J.H., Chase, T.N. & Braun, A.R. D1 dopamine receptor activation required for postsynaptic expression of D2 agonist effects. *Science* **236**, 719–722 (1987).
64. Steeves, T.D. & Fox, S.H. Neurobiological basis of serotonin-dopamine antagonists in the treatment of Gilles de la Tourette syndrome. *Prog. Brain Res.* **172**, 495–513 (2008).
65. Palminteri, S. *et al.* Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proc. Natl. Acad. Sci. USA* **106**, 19179–19184 (2009).
66. Peterson, B.S. *et al.* Neuroanatomical circuitry in Tourette's Syndrome: Tics, Obsessions, Compulsions. *Developmental Psychopathology and Clinical Care* (eds. Leckman, J.F. & Cohen, D.J.) 230–259 (John Wiley & Sons, New York, 1999).
67. Bohlhalter, S. *et al.* Neural correlates of tic generation in Tourette syndrome: an event-related functional MRI study. *Brain* **129**, 2029–2037 (2006).
68. Barkley, R.A. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* **121**, 65–94 (1997).
69. Sagvolden, T. & Sergeant, J.A. Attention deficit/hyperactivity disorder—from brain dysfunctions to behaviour. *Behav. Brain Res.* **94**, 1–10 (1998).
70. Sonuga-Barke, E.J. Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol. Psychiatry* **57**, 1231–1238 (2005).
71. Sagvolden, T., Johansen, E.B., Aase, H. & Russell, V.A. A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav. Brain Sci.* **28**, 397–419 (2005).
72. Smith, A.J., Becker, S. & Kapur, S. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural Comput.* **17**, 361–395 (2005).
73. Pattij, T. & Vanderschuren, L.J. The neuropharmacology of impulsive behaviour. *Trends Pharmacol. Sci.* **29**, 192–199 (2008).
74. Winstanley, C.A., Theobald, D.E., Dalley, J.W. & Robbins, T.W. Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* **30**, 669–682 (2005).
75. Kheramin, S. *et al.* Effects of orbital prefrontal cortex dopamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl.)* **175**, 206–214 (2004).
76. Durstewitz, D., Seamans, J.K. & Sejnowski, T.J. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* **83**, 1733–1750 (2000).
77. Maia, T.V. & Cleeremans, A. Consciousness: converging insights from connectionist modeling and neuroscience. *Trends Cogn. Sci.* **9**, 397–404 (2005).
78. Figner, B. *et al.* Lateral prefrontal cortex and self-control in intertemporal choice. *Nat. Neurosci.* **13**, 538–539 (2010).
79. Volkow, N.D., Fowler, J.S., Wang, G.J., Swanson, J.M. & Telang, F. Dopamine in drug abuse and addiction: results of imaging studies and treatment implications. *Arch. Neurol.* **64**, 1575–1579 (2007).
80. Panlilio, L.V., Thorndike, E.B. & Schindler, C.W. Blocking of conditioning to a cocaine-paired stimulus: testing the hypothesis that cocaine perpetually produces a signal of larger-than-expected reward. *Pharmacol. Biochem. Behav.* **86**, 774–777 (2007).
81. Dezfouli, A. *et al.* A neurocomputational model for cocaine addiction. *Neural Comput.* **21**, 2869–2893 (2009).
82. Redish, A.D., Jensen, S. & Johnson, A. A unified framework for addiction: vulnerabilities in the decision process. *Behav. Brain Sci.* **31**, 415–437, discussion 437–487 (2008).
83. Schoenbaum, G., Roesch, M.R. & Stalnaker, T.A. Orbitofrontal cortex, decision-making and drug addiction. *Trends Neurosci.* **29**, 116–124 (2006).
84. Berridge, K.C. & Robinson, T.E. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Brain Res. Rev.* **28**, 309–369 (1998).
85. Kapur, S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am. J. Psychiatry* **160**, 13–23 (2003).
86. Corlett, P.R., Honey, G.D. & Fletcher, P.C. From prediction error to psychosis: ketamine as a pharmacological model of delusions. *J. Psychopharmacol.* **21**, 238–252 (2007).
87. Corlett, P.R. *et al.* Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. *Brain* **130**, 2387–2400 (2007).
88. Murray, G.K. *et al.* Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol. Psychiatry* **13**, 267–276 (2008).
89. Waltz, J.A. *et al.* Patients with schizophrenia have a reduced neural response to both unpredictable and predictable primary reinforcers. *Neuropsychopharmacology* **34**, 1567–1577 (2009).
90. Strauss, G.P. *et al.* Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol. Psychiatry* (in the press).
91. Wadenberg, M.L. & Hicks, P.B. The conditioned avoidance response test re-evaluated: is it a sensitive test for the detection of potentially atypical antipsychotics? *Neurosci. Biobehav. Rev.* **23**, 851–862 (1999).
92. Fibiger, H.C., Phillips, A.G. & Zis, A.P. Deficits in instrumental responding after 6-hydroxydopamine lesions of the nigro-neostriatal dopaminergic projection. *Pharmacol. Biochem. Behav.* **2**, 87–96 (1974).
93. Fantin, G. & Botteccchia, D. Effect of nucleus accumbens destruction in rat. *Experientia* **40**, 573–575 (1984).
94. Wadenberg, M.L., Ericson, E., Magnusson, O. & Ahlenius, S. Suppression of conditioned avoidance behavior by the local application of (–)sulpiride into the ventral, but not the dorsal, striatum of the rat. *Biol. Psychiatry* **28**, 297–307 (1990).
95. Salamone, J.D. & Correa, M. Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* **137**, 3–25 (2002).
96. Yin, H.H., Ostlund, S.B. & Balleine, B.W. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* **28**, 1437–1448 (2008).
97. Maia, T.V., Cooney, R.E. & Peterson, B.S. The neural bases of obsessive-compulsive disorder in children and adults. *Dev. Psychopathol.* **20**, 1251–1283 (2008).
98. Dayan, P., Kakade, S. & Montague, P.R. Learning and selective attention. *Nat. Neurosci.* **3** Suppl: 1218–1223 (2000).
99. Gershman, S.J., Blei, D.M. & Niv, Y. Context, learning, and extinction. *Psychol. Rev.* **117**, 197–209 (2010).
100. Maia, T.V. Fear conditioning and social groups: statistics, not genetics. *Cogn. Sci.* **33**, 1232–1251 (2009).