

# 基底神经节功能在学习、记忆和选择方面的神经计算模型

Neurocomputational models of basal ganglia function in learning, memory and choice

Michael X. Cohen, Michael J. Frank\*

Department of Psychology, Program in Neuroscience, University of Arizona, 1503 E University Blvd, Tucson, AZ 85721, United States

Article history: Received 2 June 2008

Received in revised form 24 September 2008

Accepted 24 September 2008

Available online 4 October 2008

(Song Jian, translate)

**摘要:** 基底神经节 (BG) 对运动、认知和情感功能的协调至关重要, 在从帕金森病到精神分裂症的几个病理状态下, 基底神经节都会出现功能障碍。在这里, 我们回顾了神经计算框架下发展的 BG 和相关回路的原理, 提供了洞察他们在行为中的功能作用。我们关注两类模型: 包括生物实在论的各个方面, 并受功能原理的约束, 以及更抽象的数学模型, 这些模型侧重于 BG 更高层次的计算目标。虽然前者可以说更“现实”, 但后者在能够以相对简单的方程组描述系统如何工作的功能原理方面具有互补优势, 但不太适合对特定核和神经细胞的作用进行特定假设。生理过程。我们回顾了这些模型的基本架构和假设, 它们与我们对 BG 神经生物学和认知功能的理解的相关性, 并提供了对现有模型中未明确纳入的生物学细节的潜在作用的更新。从转基因小鼠到多巴胺操纵、脑深部刺激和人类遗传学的经验研究, 在很大程度上支持模型预测, 并为进一步完善提供了基础。最后, 我们讨论了可能的未来方向和整合不同类型模型的可能方法。

**关键词:** 基底神经节; 计算模型; 多巴胺; 强化学习; 决策

## 1 引言

“基底神经节”(BG)一词是指在解剖学、神经化学和功能上相互连接的皮质下结构的集合[110]。基底神经节对几种认知、运动和情感功能至关重要, 是复杂功能/解剖回路的组成部分[73,72]。基底神经节的复杂程度可以从多个层面上看到, 从无数的皮质基底神经节丘脑皮质环路, 到多巴胺、血清素和乙酰胆碱等神经化学物质的调节, 再到不同受体亚型(例如, D1 和 D2 多巴胺受体)的作用差异和位置(例如, 突触前自动受体和异质受体与突触后受体)。对基底神经节功能组织的研究涉及许多物种、实验设计、理论框架和分析水平(例如, 从功能性神经影像学、非人类遗传操作到小鼠切片准备)。为了使问题更加复杂, 大多数个体实验只关注一个物种的一个分析水平, 而每种方法都有其自身的解释风险, 因此将研究结果和方法相结合, 使单一操作的影响可预测直接和间接影响变量的级联。

生物约束的计算模型提供了一个有用的框架, 在这个框架内, (1) 在更大的理论方法的背景下解释看似不同的经验研究的结果, (2) 生成新颖的、可测试的、有时是反直觉的假设, 评估这可以用来完善我们对基底神经节的理解。此外, 计算模型的数学基础消除了语义模糊和模糊术语, 允许对不同实验、物种和分析水平的结果进行直接比较, 并允许我们探索基底神经节错综的复杂性。同时将回路的功能与行为联系起来。

虽然也有一些需要注意的地方, 但是计算模型提供了一个工具来探索经典的“盒箭”图(无“盒”中是否包含解剖的大脑区域、认知/功能过程, 或者两者都包含)所不能实现的认知和大脑过程。“盒箭”图解释起来很混乱, 为语义不明确的解释提供了太多的余地, 并且不允许检查子系统之间丰富的时间动态交互, 更不用说这些动态是如何随着学习而随时间演变

的。计算模型是动态的，易于定量分析，并且可以通过直观检查“盒箭”图来预测或激发新的经验工作，这可能很难直观地理解。

在一些例子中，对基底神经节功能的新理解是由于在多个层次的分析上运行的计算模型而产生的。建立一些模型有助于理解控制神经功能的精确生物物理过程，如胆碱能神经元内部的离子通道门控；另一些是为了帮助理解可能导致认知过程（如学习、行为选择、甚至认知控制）的各种计算。每个模型和模型类都有其自身的优点和局限性，并且每个模型都适用于不同的应用程序。考虑到没有一个模型是完整的（即，无论生物物理或功能/行为受到怎样的约束，每一个模型一定是省略了一些分子和系统水平的影响无疑是相关），模型不应该仅仅由这些因素中的任何一个来判断，而应该由它们的能力来判断。捕捉有趣的现象并做出新的预测，这可能导致对其潜在机制的洞察。

本文综述了两类神经网络模型和更多抽象的数学模型，这些模型被反复用于理解基底神经节和相关回路的行为功能。神经网络模型使用简化的神经元单元和神经动力学来帮助理解回路多个部分之间的相互作用，以及多巴胺和其他神经化学物质的调节作用，如何支持认知和行为现象，如行为选择、学习和工作记忆。相比之下，更抽象的模型包括数学方程，其中许多建立在机器学习和人工智能研究的基础上。这些不一定在实施层面上受到生物架构的约束，但仍然与这些数据联系，并设计为使用较少的假设和参数来解释大范围的行为现象。鉴于对行为的关注，我们不讨论高度关注于理解单个神经元内更详细的生物物理过程的模型（例如[177,179,181,189,102]）。这一舍弃并不意味着对这些模型缺乏兴趣或兴奋。事实上，任何依赖于实现机制的抽象或系统级神经解释最终都需要使用更现实的模型神经元来测试其合理性，并且预计一些更高的水平解释可能会被这种努力所改变。然而，目前，使用高度详细的生物物理模型来开发一个需要系统级分析的认知和行为现象模型是很难的。下面回顾的经验数据证实，尽管在神经元层面上进行了一些简化，但模型做出了具体的预测，这些预测已经在多个涉及病灶、疾病、神经成像、药理学、遗传学和认知过程中的深部脑刺激的实验中得到证实。

在下面的章节中，我们将概述前两类模型，它们的基本结构和数学基础，以及他们对基底神经节的功能和相关回路提供了新颖的见解，包括实验测试特定的模型预测。根据这些概述，我们讨论了这些模型类如何在理论和实践方面相互关联。最后讨论了计算模型在理解基底神经节及其相关回路的功能组织方面的应用前景。

## 2.基底神经节神经网络模型

通过神经网络模型，我们提到了一类模型，其中提取了神经元功能的详细方面，如轴突的几何结构，而其他过程，如膜电位随时间的波动和包括活动依赖性通道在内的动态离子电导，则是相同的。由耦合微分方程[24,58,54,55,126,83,81]计算。因此，这些模型比简单的连接主义模型在生物学上更受限制，但比详细的生物物理模型要少得多。这种方法在捕获基础神经生物学的核心方面之间提供了平衡，同时允许网络扩展到与全局信息处理和行为相关的水平。不同的模型神经元被用来模拟具有不同发放特性的神经元、兴奋性和抑制性神经元以及一些基本的神经调节剂，如多巴胺及其对不同受体亚型的突触后效应。这些过程的参数可以被修改，以捕捉大脑不同区域的不同神经元特性（例如纹状体、苍白球和丘脑[55]）。突触效能通常被简化为一个单一的可修改的权重，这反映了突触前神经元对突触后神经元活动的影响程度。这种建模方法的数学和实现细节不在本综述的范围内；感兴趣的读者可参阅专门的教科书[128,49]，以及上文所引用的特定基底神经节模型参考。

## 2.1. 基底神经节模型的构建

大体上，我们将基底神经节概念化为一个系统，它动态地、自适应地控制额叶皮质以及从额叶皮质到运动系统的信息流（模型的图形概述见图1）。基底神经节通过几个明显但部分重叠的环路与额叶皮质和丘脑皮质运动系统进行了丰富的解剖学连接[68,116,72]。这种回路可以促进或抑制额叶皮质的动作表现[110,58,54,24,4]。这些表示可以从简单的动作到复杂的行为，再到认知操作，比如工作记忆更新。与目标相关度更高或获得正确或奖励的概率更高的表示将得到加强，而与目标相关度更低或奖励概率更低的表示将被削弱。多巴胺在这一过程中起着关键的作用，它通过互补的方式调节兴奋和抑制信号，从而具有调节信噪比（signal-to-noise ratio）的效果[180119,54]。

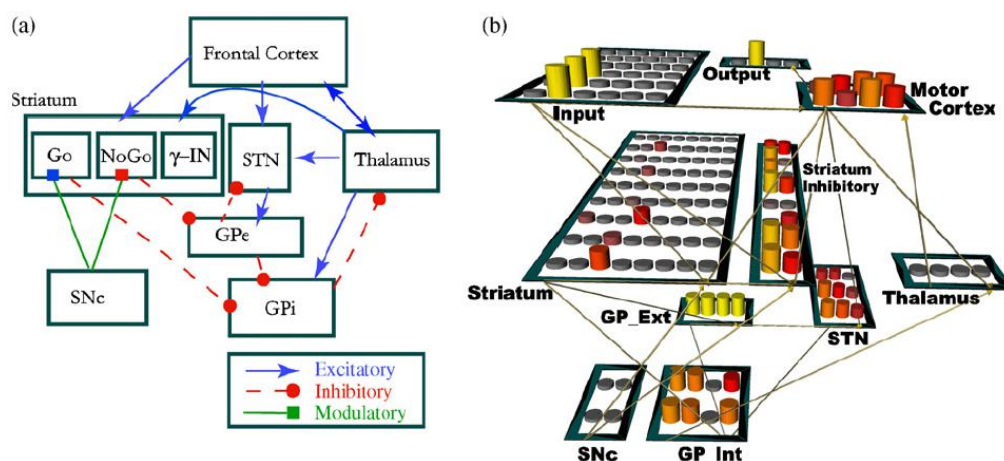


图1：左图:基底神经节回路的功能解剖图，显示了初级投射的更新模型。除了经典的直接和间接途径从纹状体到BG输出核分别起源于纹状体(Go)和纹状体(NoGo)细胞外，修正后的结构还具有从NoGo单元到GPe的聚焦投影和从皮质到丘脑的强自顶向下投影。此外，STN被合并为新发现的超直接通路的一部分(而不是最初设想的间接通路的一部分)，接收来自前额皮质的输入，并直接投射到GPe和GPi。右图:该回路的神经网络模型，四个不同的响应由四列运动单元表示，纹状体内Go和NoGo各四列，GPi、GPe和丘脑内对应的四列。快速尖峰的古ABA-ergic间神经元( $\gamma$ -IN)通过抑制投射调节纹状体活动。实现细节见[54,55]。

该系统的模型包括基底神经节的主要结构：纹状体、苍白球内外节段（GPi和GPe）、黑质致密部（SNc）、丘脑和底丘脑核（STN）。这包括向额叶皮质发送Go信号的经典直接通路和向额叶皮质发送NoGo信号的间接通路[2,110,68,54]。然而，正如我们将看到的，我们的计算模型超越了经典的直接/间接模型，（a）随着活动在整个系统中传播，以及作为突触可塑性的一个函数时，探索这个系统的动力学，这两者在静态模型中都不明显，（b）包含更多的不在原始模型中但对其功能选择至关重要的解剖学和生理学证据。直接通路来源于纹状体神经元，主要表达D1受体，对GPi和SNr有直接抑制作用。我们将此途径中的活动称为Go信号，因为当纹状体细胞处于活动状态时，它们抑制GPi，进而抑制丘脑[33]，并允许额叶皮质表现被自下而上的丘脑皮质驱动放大。注意，这种去抑制过程只有在同样的丘脑柱也受到自上而下的皮质-丘脑兴奋时，才会使相应的丘脑柱变得活跃。这意味着基底神经节系统不会直接选择要考虑的动作，而是调节皮质中已经活跃的表达的活动。这一功能使大脑皮层能够微弱地并行地表示多个潜在的动作；首先从纹状体输出接收到Go信号的动作，然后提供足够的额外激励来执行。一旦BG回路选择了获胜的反应，丘脑和皮层内的横向抑制会抑制竞争反应。

与直接途径互补，间接途径起源于纹状体的纹状体黑质细胞，主要表达D2受体，并向

GPe 提供直接抑制输入。我们将此途径中的活动称为发送 NoGo 信号以抑制特定的不需要的反应[2,54,110]。由于 GPe 通过直接聚焦投射来抑制 GPi，纹状体黑质 NoGo 活性消除了这种强直抑制，从而抑制 GPi，使其进一步抑制丘脑，并阻止特定的皮质动作被促进。这样，模型基底神经节可以促进（Go）或抑制（NoGo）在额叶皮质的表现。根据该模型，Go 和 NoGo 单元的特定亚群可以专门表示特定感官输入背景下的促进/抑制作用，以便不同的细胞被激活，例如在两个不同环境背景下促进相同的作用时[54]。还要注意，给定的动作可以同时具有 Go 和 NoGo 表示，选择它的概率是相对 Go-NoGo 激活差异的函数[54]。这是由于观察到 Go 和 NoGo 细胞从一个给定的皮质区域接收（从而编码一个给定的动作），它们起源于同一个纹状体区域，并终止于 GPi 内的同一区域，在该区域内保持躯体表示（例如[53,110,190]）。后一种结构中的神经元可以反映两个纹状体群体的相对差异，从而影响到抑制丘脑的可能性，进而选择动作。

请注意，上述描述省略了 STN，传统上认为 STN 是连接 GPe 和 GPi 的间接通道内的关键中继站[2]。然而，最近的证据表明：（a）GPe 神经元向 GPi 发送直接抑制性投射，而不是必须通过 STN 间接施加其控制；（b）这些 GPe-GPi 投射更集中，允许抑制特定的反应，而从 STN 到 GPi 投射的反应是广泛的和扩散的[110,130]，可能提供更多的全局调节功能（见下文）。

这并不是为了削弱或降低 STN 的作用。相反，最近的证据表明 STN 应该被认为是第三个超直接通路的一部分（之所以这么命名是因为它完全绕过纹状体），而不仅仅是间接通路中的一个中继。事实上，STN 接收来自额叶皮质的直接兴奋性输入，并将弥漫性兴奋性投射发送给 GPi[117,118]，我们将 STN 中的活动称为发送一个全局 NoGo 信号，因为它对许多 GPi 神经元的弥漫性兴奋作用会阻止所有的反应，而不仅仅是一个反应[55]。模拟表明，这些信号是动态的：响应选择过程中早期观察到全局信号，防止过早选择任何响应，但随着 STN 活动减弱，响应就更有可能发生。STN 活性的短暂爆发与体内观察到的一致[173,105]。此外，在模型中，最初的全局 NoGo 信号被皮质“响应冲突”的程度自适应地调节：多个相互竞争的皮质激素指令的更大激活与更大的 STN 兴奋性驱动和一个明显的全局 NoGo 信号相关，使纹状体能够需要更多的时间来解决和整合过度噪声的内在活动，以选择最佳的响应[55,21]。如果没有这种 STN 功能，BG 网络更可能做出过早的响应，通常会解决次优的选择，特别是在存在高度响应冲突的情况下[55]在 STN 损伤的大鼠中观察到这种过早反应[11,10]。

多巴胺在基底神经节中起着特殊的调节作用。在纹状体中的 D1 受体处，多巴胺被认为是一种对比增强剂，增加高活性细胞的活性，同时降低低活性细胞的活性[77]。这有放大信号（高活性细胞）的效果，同时降低噪声（低活性细胞[54]）。在 D2 受体处，多巴胺是抑制性的，与细胞内的活动量无关[78]。由于 D1 受体在 Go 细胞中大量表达，而 D2 受体在 NoGo 细胞中大量表达[65]，多巴胺的升高有促进突触驱动的 Go 活性的净效应，同时抑制 NoGo 活性。相反，低水平的多巴胺会降低 Go 细胞的信噪比，同时释放 NoGo 细胞的抑制作用。这一概念解释了为什么在帕金森病中多巴胺水平的降低会导致 NoGo 通路的过度激活[158,150]和运动迟缓，类似于最初的提议[2]。此外，在我们的动态模型中，这些多巴胺对 Go 和 NoGo 活动的影响与强化学习特别相关[54,63,24]，并且对基底神经节系统如何学习哪些表征促进和哪些抑制具有重要意义，如下节所述。

## 2.2.基底神经节模型的强化学习

神经元之间的突触权重会随着时间和经验而动态变化，形成学习的基础。强烈和反复共同激活的单元之间的权重会变强（如长期增强；LTP），否则单元之间的权重不会改变或变

弱（如长期抑制；LTD）。多巴胺释放的存在和时间强烈调节纹状体中的这些效应[17,135,90,136,26,25,32]。事实上，基底神经节模型的主要学习机制依赖于已经被皮质纹状体谷氨酸输入激活的细胞的多巴胺调节。

具体来说，SNc 中的多巴胺能神经元在意外奖励期间以阶段性爆发的方式发放而闻名，当预期奖励但未收到奖励时，发放降低到强直基线水平以下[144,13]。在模型中，当模型选择正确的动作（取决于任务的性质）时，SNc 多巴胺迸发被模拟。结果，激活的 Go 单元进一步增强，使来自感觉皮层和运动前皮层的 Go 单元的权重增加。这意味着，下一次相同的刺激与相关的运动前皮质反应同时出现时，这些相同的 Go 单元可能会变得活跃并促进相同的奖励反应。相反，弱活动的 Go 单元被抑制。这些效应通过模拟的 D1 受体介导，与上述生理数据一致。

此外，当模型接收到多巴胺的下降（即当期望一个信号时缺乏奖励[144,13]），就会出现一个补充过程。在这种情况下，NoGo 单元，通常是由多巴胺通过模拟 D2 受体抑制，现在变得更活跃的皮质谷氨酸能输入。事实上，纹状体苍白球神经元从额叶皮质接收到更强的投射，对皮质刺激表现出特别增强的兴奋性[18,19,95,100]。关键的是，短暂增强的 NoGo 单元活性与长期增强有关（通过类似的 Hebb 学习原则），这样下次模型面临相同的感官刺激和潜在反应时，反应更可能被抑制。因此，多巴胺的阶段性下降诱导学习，以避免在特定刺激的情况下采取特定的行动[54]。最近的研究支持这一基本模型预测，表明尽管直接途径中的突触增强依赖于 D1 受体刺激，但间接途径中的增强依赖于缺乏 D2 受体刺激[149,191]。相反，当 DA 水平较高时，刺激 D2 受体会导致突触减弱[32,149]。因此，现有数据表明，DA 水平的升高（例如，在阶段性爆发期间）有可能促进学习，但削弱了 NoGo 学习，而低 DA 水平（在下降期间）则有相反的效果，正如模型所预测的那样[54]。

正是通过这种“推-挽”机制，基底神经节模型能够学会选择动作或加强额叶皮质表征，这更可能导致奖励或正确反馈，同时降低不正确或无奖励动作或表征的概率的这种情况不太可能发生。在这两种途径中学习的存在使模型能够增强不同“刺激-增强”概率之间的对比，从而更容易区分 60%和 40%奖励的选择。仅仅在 Go 通路中学习增加和减少突触权重的模型在复杂的概率环境中不太能够做出这些细微的区分[54]。在双通路模型中，Go 和 NoGo 两种通路都有 60%的响应，并且还记得 BG 输出（GPi）计算了每个反应的相对激活差异。因此，对 GPi 的净影响（为简单起见忽略非线性）为  $60-40=20\%$ 。同样地，40%的 GPi 响应将比 GO 响应更大，因此将表示为  $-20\%$ 。两种反应之间的净差异实际上是 20%，在 BG 输出时对比度提高到 40%。

### 2.2.1.多巴胺下降是否包含足够的学习信息？

多巴胺神经元的基线发放率一般在 5Hz 左右。因此，虽然发放率的增加可以随着预测误差的增大而增大，但它们不能随着预测误差的减小而减小（因为神经元不能有负的发放率）。这导致了一个问题，即是否需要大脑中单独的非多巴胺能机制来编码负性预测误差[46,12]。然而，最近的经验研究表明，在奖励缺失期间多巴胺神经元暂停的持续时间，可能包含负预测误差大小的信息，而不是发放率的变化[13]。有趣的是，纹状体中的 D2 受体对多巴胺的微小变化高度敏感，部分原因是大多数 D2 受体具有高亲和力[137]；因此，停顿时间的差异可能具有可检测的下游效应[56,60]。

回想一下，该模型需要缺乏 D2 受体刺激来增强 NoGo 单元的潜能并促进学习，最近的数据支持了这一点[149]。因此，较长的停顿时间为多巴胺转运体从突触中移除 DA 提供了更多的时间，增加了表达 D2 受体的神经元被解除抑制的可能性。这种解释在背侧纹状体特

别可信，那里有许多多巴胺转运体，突触中多巴胺的半衰期大约为 55-75ms[155,70,167]。这意味着长时间的停顿 (>200ms) 将给多巴胺几乎不存在的足够时间，并允许 NoGo 单元变得不受抑制（与腹侧纹状体形成对比，尤其是前额皮质，在这一过程中，恢复的过程可能过于缓慢，阶段性的下降不会产生任何功能影响）。此外，纹状体多巴胺水平的减少，如帕金森病，实际上会增强这种作用。尽管强直性多巴胺水平已经很低，但由此产生的 D2 受体超敏性[146]以及在 DA 耗尽状态下 NoGo 细胞的兴奋性增强[158,150]将有助于对 DA 暂停的突触后检测（因此，它们可能不需要持续很长时间）。事实上，最近的研究表明，在帕金森病的小鼠模型中，DA 缺失导致 NoGo 突触增强[149]。

### 2.2.2. 皮质系统的可塑性：从行为到习惯

最后，该模型还直接捕捉了从感觉皮层到运动前皮层的皮质通路的可塑性。当对特定刺激做出反应时，简单的 Hebb 学习发生，使得相同的运动前皮质单元在将来可能对相同的刺激做出反应，而与该反应是否得到奖励无关[54,56]。这种效应使皮质单元能够根据其先前选择的频率识别候选反应，从而对给定动作的适用性提供一个初步的最佳猜测，然后根据 Go/NoGo 强化值由 BG 促进或抑制。一旦这些皮质联系足够强，它们可能根本不需要 BG 来促进，与数据一致，这些数据表明纹状体多巴胺对于学习行为的初始获得是必要的，但对于他们的后期表达则更不需要如此[153,131]。同样地，背侧纹状体失活也会损害所学任务的执行，然而一旦行为根深蒂固，这种影响就微乎其微[7]。根据该模型，习惯学习依赖于纹状体多巴胺系统来获得导致奖励的反应，但它的表达是由更直接的皮质联系介导的（如果足够强烈，则不需要额外的纹状体增强）。注意，这种皮质学习意味着最终运动前皮质区域自身参与基于奖励的行动选择，这样，在 BG 进行任何促进之前，过去经常选择的反应立即优先于其他选项。

### 2.3. 局限性和与真实大脑解剖的比较

我们的模型远没有捕捉到与真正的基底神经节回路相关的所有有用的复杂性。事实上，基底神经节比上述段落中描述的要复杂得多。尽管我们已经模拟了不同的动态和解剖投影，这些投影不是经典模型的一部分，但我们的模型仍然高度简化，对于任何模型来说，质疑这些简化原则是否与真实系统相关始终是合理的。在这里，我们总结一些对框架的挑战。

#### 2.3.1. Go 和 NoGo 通道真的被隔离了吗？

尽管经典的 BG 模型成功地提供了一个预测框架，用于解释跨多个分析级别的多个数据模式，但对模型的基本原则仍有一些挑战。首先，该模型依赖于纹状体和纹状体黑质神经元中 D1 和 D2 受体的分离[65-67,20,97,84,8]。早期的挑战表明，事实上，D1 和 D2 受体共同定位在同一个神经元上，即使这种共定位相对于一种或另一种受体类型的整体表达来说很小[159,1]。最新的进展，尤其是转基因小鼠，已经完全停止了这种关注[158]。然而，还有一个关键的挑战是，纹状体和纹状体苍白球神经元的传出投射可能不像在模型中那样清晰地分离。事实上，虽然纹状体苍白球细胞只向 GPe (NoGo 细胞，用我们的模型的说法) 投射，但许多纹状体黑质细胞 (Go 细胞) 也有轴突侧支向 GPe 投射[88,101,183]。从表面上看，考虑到 Go 细胞也会投射到 GPe，这似乎对 Go 细胞的功能本身提出了挑战。然而，我们认为，这种设置实际上是有用的，以确保激活的 Go 通路保持瞬态，并使 GPi 计算 Go 信号的时间导数，而不是原始的 Go 信号[55]。也就是说，因为纹状体到 GPi 的直接投射是单突触的，而到 GPe 和 GPi 的直接投射是多突触的，Go 信号首先抑制丘脑，然后通过 GPe 路径延迟再抑制丘脑。这种类型的系统可以快速方便化并随后抑制表现，如果必须连续激活一系列运动

命令或工作记忆中的项目，这将是相关的。

### 2.3.2.纹状体中间神经元的作用

为了简单起见，我们的模型并没有明确地包含胆碱能（强直性）中间神经元的功能，我们只是开始探索 GABA 能（快-尖峰）中间神经元的作用（图 1），它们共同构成了 5% 的纹状体神经元[164,68]。这些细胞类型中相对较小的比例并不一定会降低其潜在的功能意义。例如，胆碱能中间神经元被认为深入参与基于奖励的学习，当它们成为奖励的预测因子时，它们对刺激作出动态反应[178,3]。胆碱能神经元也在纹状体的长期塑性变化中起着许可的作用[31]，并似乎间接介导了“多巴胺-诱导”塑性的某些效应[171]。最近的证据表明，乙酰胆碱和多巴胺可能在基于奖励的学习中起到协同作用：在显著事件中，中脑多巴胺细胞和纹状体胆碱能细胞在同一时间窗内响应，但只有多巴胺细胞按奖励概率的比例发放[112]。有人认为，胆碱能发放的暂停可以作为一个时间框架，根据多巴胺能信号的大小决定何时学习。此外，[44]认为胆碱能暂停提供了一种对比增强效果，区分了强直性和阶段性多巴胺能状态，由于多巴胺的爆发和下降，有效地提高了学习。这种效应的部分原因是乙酰胆碱通过烟碱受体对多巴胺释放的突触前效应[44]。因此，尽管多巴胺有助于学习什么，胆碱能中间神经元却有助于何时学习。

尽管在我们的模型中，这些效应都没有在生物物理学层面上进行模拟，但我们还是隐晦地包含了其中的一些效应。也就是说，在我们的模型中，控制学习的方程式相当于一种“对比的 Hebb 学习”形式，其中阶段性多巴胺信号对 Go/NoGo 活动的影响是相对于前一种状态（在此期间多巴胺能信号是强直的）的影响进行计算的。因此，这种机制自动确保在正确的时间窗口中进行学习，并提供强直状态和阶段性状态之间的对比；如上文所述，这两种功能都可以由胆碱能激发的暂停来支持[112,44]。然而，毫无疑问，这些相互作用是相当复杂的，并且可能受益于更明确的模拟。

### 2.3.3.丘脑“背部-投射”

除了丘脑和额叶皮质之间的反复投射，以及从 GPi 到丘脑的前馈投射外，也有经常被忽视的从束旁核丘脑到纹状体和底丘脑核的“背部-投射”（例如[113,30]）。考虑到丘脑纹状体投射突触主要位于胆碱能神经元间，调节胆碱能外排[96,188]，可能是束旁核丘脑提供了一个警报信号，在突出的事件中诱导暂停胆碱能发放和促进学习。此外，我们的模型中的初步（未发表）模拟表明，从丘脑到 STN[30]的“背部-投射”可能在运动反应被抑制后起到终止作用。

### 2.3.4.腹侧纹状体与背侧纹状体

虽然纹状体在我们的模型和其他几个模型中表现为单一的结构，但它实际上包括几个子区域。这些亚区遵循腹内侧到背外侧的梯度，从皮质大致平行的梯度传入[72,35]。虽然根据细胞构造性质很难确定子区域之间的精确边界[169,103]，但可以通过输入/输出纤维的模式[73]和某些情况下的功能分离[27,133,7,122]来划分子区域。背侧纹状体区域与背侧前额叶区域紧密相连，因此被认为在调节认知操作（如工作记忆更新）方面起着核心作用[58,38,138]。相反，腹内侧区域，包括伏隔核，更与强化引导的学习和成瘾相关过程有关[28,52,93]。在伏隔核内部、外壳和核心区域之间可以进一步区分。

关于强化学习领域腹侧/背侧功能分离的一个经典解释是“Critic”和“Actor”之间的解释[85,82]。由腹侧纹状体扮演的“Critic”，评估当前的环境状态是否是对奖励的预测，并通过在特定的状态中体验奖励来学会这样做。在意想不到的回报（或缺乏回报）期间，阶段性多

巴胺反应的变化被认为会推动“Critic”的学习，从而使其预测在未来更加准确。相反，由背侧纹状体扮演的“Actor”决定选择哪一个动作，并在执行特定动作后通过这些相同的阶段性多巴胺信号学习选择哪一个动作，从而形成动作特定的值表示。（请注意，一旦“Critic”了解到这一点，当遇到预测未来奖励的环境状态时，就会产生多巴胺爆发，这有助于训练“Actor”产生这种状态的行为，即使这些行为并没有立即先于奖励本身。）尽管有证据支持这一观点[85,122]，但这个模式可能更复杂[7]。

基于上述建模框架，我们认为纹状体的不同亚区与额叶皮质的计算和相互作用相似，但在不同区域处理的信息类型取决于额叶皮质的亚区，而额叶皮质的亚区与额叶皮质的亚区之间存在着相互作用。纹状体亚区相互作用（另见[174]）。例如，由于背侧纹状体受背侧前额叶和外侧前额叶区域的神经支配最为密集，它可能控制与背外侧前额叶皮质参与的过程有关的信息流，即工作记忆、计划、认知控制等等[58,126]。相反，腹侧纹状体与眶额皮质和腹内侧额前皮质紧密相连，可能会控制有关奖励和动机的信息[56]。伏隔核的其他部分可能参与学习哪些与奖励有关的环境状态（包括外部和内部），从而驱动多巴胺信号，训练“Actor”[127,22]。最近，O'Reilly 和他的同事们提出了多巴胺介导学习的神经生物学机制的扩展模型。在主值-习得值（Primary value-learned value, PVLV）模型[127]中，与 SNc 相关的单个节点现在是一个由腹侧纹状体、外侧下丘脑、杏仁核和 SNc 等区域组成的网络。“主值”（PV）系统由腹侧纹状体中的“斑片状”纹状体神经元介导，负责学习何时会出现无条件的奖赏，并在预期的情况下（由于纹状体对 SNc 和 VTA 的抑制性投射）采取行动来消除多巴胺迸发[86]。PV 系统产生的活动与动物学习预测初级奖励时多巴胺神经元活动的最初增加和随后的减少相匹配。

模型的“习得值”（LV）系统学习将奖励值分配给预测后期奖励（即条件刺激）的任意刺激。只有当存在外部奖励或 PV 系统期望主要奖励通过 PV 激活进行门控时，才会在该系统中进行学习。这样，在环境中没有奖励的时候，LV 系统可以表达一般的奖励值（相反，PV 系统总是了解奖励或奖励的缺失，因此不会在奖励值出现之前表示奖励值）。LV 由杏仁核的中央核表示，杏仁核与奖赏学习密切相关，并向中脑多巴胺神经元发送兴奋性投射。这一系统在生物学上比以往使用时间差异学习对中脑多巴胺系统功能的数学估计更可信，并且在某些情况下比这个系统更强大（如刺激-奖励时间变异性和对干扰性刺激的敏感性[127]）。

总之，尽管我们的计算模型不完整，大脑不仅仅是复杂的突触、神经和化学部分的总和：大脑可以学习并参与一系列令人印象深刻的认知和行为过程。从这个意义上说，上面描述的建模方法在生物学上是相关的，因为，正如下一节中详细介绍的，模型可以产生类似于生物有机体的输出，并且模型的行为是通过药物、疾病状态和遗传变异的模拟来调节的。因此，神经网络建模方法的目的是不是捕获基底神经节神经生物学的每个已知方面，而是将基底神经节神经生物学的关键要素与认知和行为过程联系起来。在下一节中，我们将描述一些已经由经验结果证实的模型预测。

## 2.4.基底神经节模型预测的经验证据

这个基底神经节模型对强化学习过程中的行为和神经反应做出了一些可测试和可证伪的预测，以及这些反应应如何受药物、疾病或遗传状态的调节。最初的模型被设计为受生理和解剖数据的限制，但也考虑到帕金森病和药物状态导致的认知变化，包括强化值和逆转[54]之间的复杂概率辨别，以及丘脑下核在高冲突决策中的作用[55]。

### 2.4.1.Go 和 NoGo 学习的多巴胺调节



在神经层面，该模型预测了不同的纹状体种群的存在，这些纹状体分别编码积极和消极的刺激反应动作值。尽管这些神经元是否与 Go 和 NoGo 的单位（即神经元）相对应还有待确定，但在猴子身上已经有了这类神经元的报道[139]。但突触可塑性研究支持了该模型的预测，即这些不同的群体如何通过 D1 和 D2 受体机制分化出来，从而增强 Go 和 NoGo 突触的功能[149]。

在行为层面，猴子为了获得较大的奖励而加快反应时间的能力（在我们的模型中需要 Go 学习）依赖于纹状体 D1 受体的刺激，而为了较小的奖励而放慢反应速度的倾向（NoGo 学习）依赖于 D2 受体的去抑制作用[115]。同样，我们的计算模型模拟了一系列关于 D2 受体拮抗作用对啮齿动物过氧化氢表达的影响的报道结果，作为 NoGo 学习的一种形式，包括敏化、背景事件依赖和消除[175]。

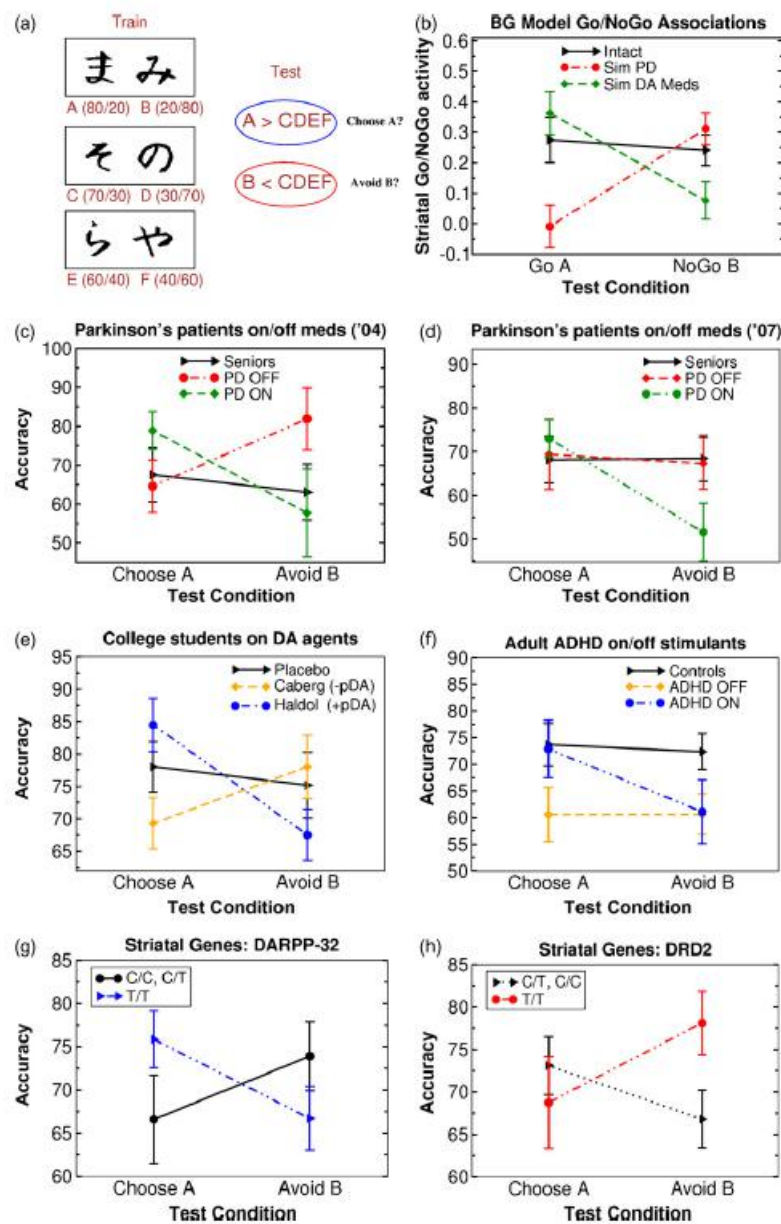


图 2: (a) 概率选择强化学习任务。在训练过程中，参与者在每一对刺激中进行选择。括号中表示每个刺激收到正/负反馈的概率。在测试阶段，所有的刺激组合都没有反馈。在这些新的配对中，Go 学习通过可靠地选择最积极的刺激 A 来建立索引，而 NoGo 学习则通过可靠地避免最消极的刺激 B 来建立索引。(b) 分别在输入刺激 A 和 B 时，纹状体 Go 和 NoGo

激活状态。模拟帕金森病(Sim PD)通过降低纹状体 DA 水平来实现, 而药物(Sim DA Meds)通过提高 DA 水平来实现, 并在负反馈时部分分流 DA 下降的影响。(c) PD 患者用药前后的行为研究结果支持模型预测[63]。(d)在另一组患者中进行复制, 其中在 NoGo 学习条件下观察到的效果最为显著[61]。(e)在服用调节突触前 DA (pDA)的多巴胺激动剂和拮抗剂的健康参与者和服用或停用刺激性药物的成人 ADHD 参与者中也有类似的结果。(g)和(h)通过控制纹状体 D1/D2 功能的基因预测大学生 Go/NoGo 学习的个体差异。

在人类中, 一个直接的模型预测是, 从正反馈和负反馈中学习的能力应该取决于 Go 和 NoGo 学习, 其平衡取决于多巴胺的水平。多巴胺的阶段性爆发促进从正反馈中获得 Go 学习, 而阶段性下降则促进从负反馈中获得 NoGo 学习[54]。如果这些多巴胺的阶段性水平被疾病或药理学调节或损害, 那么个体从正反馈和负反馈中学习的方式也应该被调节。帕金森病患者提供了一个测试这些假设的机会: 这些患者在停止服药时降低了多巴胺信号, 但在服药时提高了多巴胺水平。先前的研究发现, 帕金森病患者在强化学习方面受到损伤, 这与纹状体和前额皮质中多巴胺含量较低有关[161,152,39,41]。因此, 人们可能期望多巴胺药物能改善这些患者的表现。然而, 奇怪的是, 根据所使用的认知任务的不同, 表现可以提高, 也可以降低[42,151,54,63]。

计算模型可能有助于澄清这种明显的不一致性。具体来说, 该模型预测多巴胺水平应不同地影响从负反馈和正反馈的学习(图 2)。当患者停止服药时, 他们应该从负反馈中学习, 而不是从正反馈中学习, 因为低水平的多巴胺激活了 NoGo 通路(如[158]), 并且, 与 D2 受体超敏性一起, 可能有助于检测 DA 下降, 但是要防止在奖励过程中 Go 通路被充分激活。相反, 当患者在服药时, 突触前多巴胺合成增加[163,132]。此外, 长期服用左旋多巴(主要用于治疗 PD 的 DA 药物)可增加阶段性(尖峰依赖性)DA 暴增[75,176,89], 并表达 Zif-268, 一种与突触可塑性相关的早期基因[92], 在纹状体黑质(Go), 而不是纹状体苍白球(NoGo)神经元[29]。因此, 该模型预测药物可以改善 Go 路径中的正反馈学习。有趣的是, 同一个模型预测多巴胺药物会削弱从负反馈中学习的能力: 因为药物持续刺激 D2 受体(这一假设最明显地适用于大多数患者服用的 D2 激动剂药物, 但也可能是左旋多巴, 它可能增加强直性 DA 释放, 除了阶段性发作), 它们有效地阻止了当奖励被忽略时检测到的 DA 激活中的阶段性停顿[54]。

这种结果的模式最近在帕金森病患者中得到了证实, 他们在概率强化学习范式中药物进行了断断续续的测试, 其中一些选择与正反馈和负反馈相关的概率更大[63](图 2)。不服用药物的患者从负反馈中学习的比从正反馈中学习的好, 而服用药物的患者从正反馈中学习的比从负反馈中学习的好。当 DA 消耗和药物在模型中被模拟时, 这些效应也产生了[63,61], 并且在不同的实验室中使用不同的模式再现了这些效应[40]。此外, 它们与上述突触可塑性研究惊人的一致, 其中 DA 耗竭与 Go 突触 D1 相关电位的降低有关, 但与 NoGo 突触 D2 相关电位的增强有关, 而 D2 激动剂的使用逆转了 NoGo 突触的电位[149]。值得注意的是, 在多巴胺转运体基因敲除小鼠(纹状体多巴胺水平升高)中也发现了类似的行为结果模式(Go 增强但 NoGo 学习减少)[43]。所有这些发现都证实多巴胺不仅与积极的预测误差有关, 而且与消极的预测误差有关。

在服用小剂量多巴胺激动剂和拮抗剂的年轻健康大学生身上也观察到了同样的 Go 和 NoGo 学习概率的调节[60](图 2)。此外, 纹状体 DA 缺失和 DA 细胞完整性受损的老年人(70 岁以上)比年轻人(60-70 岁)有选择性地表现出更好的负反馈学习, 这与帕金森氏研究结果一致[57]。成人多动症参与者的结果呈现相反的模式, 他们在服用兴奋剂时表现出比负反馈学习更好的正反馈[62](图 2), 后者阻断多巴胺转运体并提升纹状体 DA[168,104]。总之, 在广泛的人群和操作中, 纹状体多巴胺的增加与相对较好的 Go 学习有关, 尤其是与

较差的 NoGo 学习有关，而纹状体多巴胺的减少与相反的模式有关。

从行为上讲，该模型表明，具有独立的 Go 和 NoGo 路径可以提高不同增强概率之间的概率辨别能力。也就是说，仅从正反馈或负反馈中学习的网络不会像那些同时接受正反馈和负反馈的网络那样产生强大的学习效果（即使反馈试验的数量是相等的）。最近在一项基底神经节依赖性概率学习任务[6]中发现了这种模式，在该任务中得出结论，需要双通道 go/NoGo 模型来捕捉基本的行为发现。

尽管我们在概率强化范式中发现，平均而言，健康个体从正反馈和负反馈中获得同样好的学习，但在这些衡量标准中，仍然存在着显著的个体差异，例如，一些参与者是积极的学习者，而一些则是消极的学习者[64]。我们假设至少部分这种变异可能是由控制纹状体多巴胺功能的遗传因素造成的。为了验证这一假设，我们从 69 名健康参与者身上采集了 DNA，并用相同的概率强化学习任务对他们进行了测试[59]。如果 Go-学习的个体差异归因于 D1 功能和 NoGo-学习归因于 D2 功能，那么控制纹状体 D1 和 D2 功效的遗传因素可能是这种学习的预测因素。因为还没有一个基因多态性显示优先影响纹状体 D1 受体，我们分析了一个多态性，该多态性控制了蛋白质 DARPP-32，该蛋白质在纹状体中高度集中，是动物 D1 依赖可塑性和奖赏学习所必需的[129,170,25,154]。此外，在人类中，根据 DARPP-32 基因型进行功能调控的唯一大脑区域是纹状体及其与额叶皮质的功能连接[109]，我们还分析了 DRD2 基因中编码突触后纹状体 D2 受体密度的多态性[79]。值得注意的是，我们发现 DARPP-32 基因功能的个体差异，作为纹状体 D1 依赖可塑性的替代测量，可以预测更好的正反馈学习，而 DRD2 功能的个体差异，作为纹状体 D2 受体密度的测量，可以预测更好的负反馈学习活动[59]（图 2）。另一组[91]也独立发现了后一种效应，他们分析了不同的 DRD2 多态性。此外，Go/NoGo 学习效应是特定于纹状体遗传功能的，因为第三个主要编码前额叶多巴胺能功能的基因[166]与 Go 或 NoGo 递增概率学习无关，相反，纹状体基因预测的是受试者工作记忆的强化效果最大[59]。这种工作记忆效应与其他详细的计算模型一致，这些模型表明，前额叶多巴胺对于活跃状态下信息的稳定维护至关重要[51]，并且部分前额叶皮层支持工作记忆以获取奖励值，这指导了试验的临床行为并补充渐进学习基底神经节系统[56]。

这种清晰的基因发现，即具有不同功能的大脑效应的不同多态性与可分离的认知功能相关，在文献中很少见。如果没有计算模型，这些特定的基因不太可能在这些特定类型决策的背景下进行分析。然而，前额多巴胺基因的影响的性质和方向，尽管与前额叶皮质在快速适应试验中的一般作用相一致，但与我们现有的 DA 在该系统中的作用模式不一致[56]，这将使我们重新审视和改进模型（即，因此前额叶 DA 的作用更接近于[51]的提议作用）。

#### 2.4.2. 在高冲突决策中的底丘脑核

基底神经节模型也对决策的其他非多巴胺和非学习方面做出预测。如前一节所述，STN 向 BG 输出核（GPi 和 GPe）扩散投射，并通过背内侧额叶皮质的超直接途径接收直接兴奋性输入[117,4]。该模型通过动态（和短暂）调整决策阈值，将 STN 纳入防止冲动决策的考虑范围[55]。在做出涉及高度应对冲突的决策时，这种作用是显而易见的（见图 3a）。神经影像学支持这一结论，即在高而非低冲突条件下，背内侧额叶皮层和 STN 之间的协同激活增加与反应时间越来越慢有关[4]。

为了证明 STN 在减缓冲突下的反应中起着关键（而不是相关）的作用，必须对其进行操纵。帕金森病患者的深部脑刺激（DBSS）植入 STN，为了解 STN 在人类冲突相关决策中的作用提供了一个独特的窗口。这些刺激器以异常高的频率和电压向 STN 提供电流，干扰

STN 的功能，有效地起到病变的作用（或像给系统增加噪声一样，阻止它对皮质输入作出自然反应）[15,16,108]。然而，这个虚拟损伤是暂时的，因为刺激器可以由医生打开或关闭。当刺激器打开时，许多与运动相关的帕金森氏病症状都会急剧减轻；刺激器关闭后几分钟到一小时内，症状就会恢复。在最近的一项研究中，Frank 和同事们在强化学习任务中对这些患者进行了开/关刺激测试，并将他们的表现与另一组开/关多巴胺药物的患者进行了比较[61]。高冲突决策是指两种选择之间的强化概率只有细微差别的选择（例如，一种选择有 80% 的机会获得奖励，而另一种选择有 70% 的机会），而低冲突决策的特征是包含不同强化概率的选择（例如，80%和 30%）。

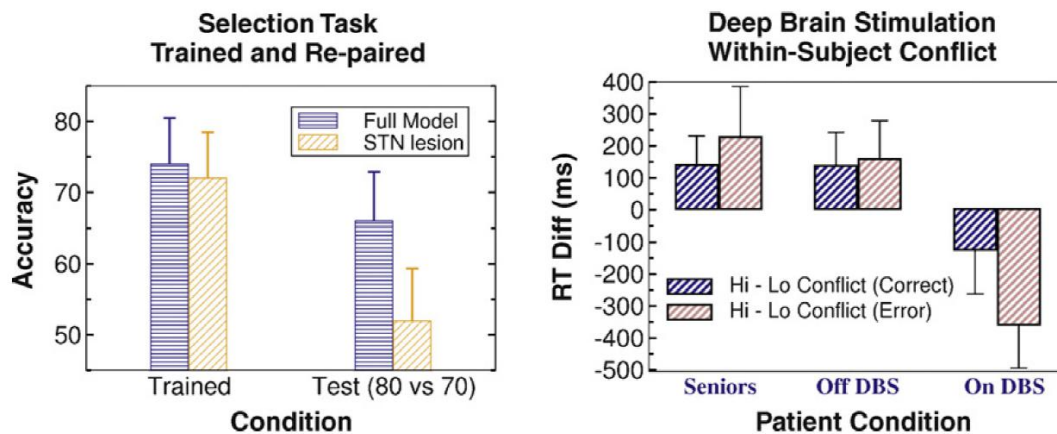


图 3.(a)丘脑下核在概率选择任务中对模型性能的贡献。在训练有素的低冲突识别中，STN 受损网络的选择与完整网络没有区别(80 和 20, 70 和 30)，但在高冲突选择中，当 STN 受损网络与 70% 的响应竞争时，其 80% 正强化响应的选择性受损。模型 STN 全局 NoGo 信号可以防止在多个响应可能有回报时过早做出响应，从而增加了准确选择[55]的可能性。(b)帕金森患者在 DBS 前后的行为结果，证实了模型预测。相对于低冲突试验，反应时间差异较大。健康对照组、开/停药组(未显示)和停药组在高冲突试验和低冲突试验中决策时间较慢，而停药组在这些试验中冲动反应较快(改编自[61])。

通常，当面对这些高冲突选择时，响应时间会减慢；在健康对照组、停药和停药患者以及 DBS 患者中观察到这种模式。值得注意的是，DBS 患者在决策冲突增加的情况下未能减缓反应时间（图 3b）。此外，患者对高冲突的反应比低冲突的反应更快。当患者选择了次优选项（强化概率较低的选择[61]）时，这些加速的高冲突决策时间甚至被夸大了，这表明刺激破坏了 STN 在高冲突决策期间提供全局 NoGo 信号的能力。此外，当模型被给予一个 STN 损伤或当模拟高频 DBS 应用时，它产生相同的结果模式。与上述药物效应（并在 2007 年的研究中重复）一起，这些发现揭示了治疗类型在帕金森病认知决策的两个方面的双重分离：多巴胺药物影响正/负学习偏差，而 DBS 影响冲突导致的学习偏差减缓，但不是正/负。

总之，尽管我们的神经模型相对于实际的基底神经节回路的复杂性进行了简化，并抽象出了许多生物物理和分子机制，但建模工作已被证明是开发明确的可测试和可证伪假设的一个有价值的工具，而这些假设可以通过一个简单的神经模型来实现。直接导致了经验实验，为这些预测提供了支持。尽管如此，我们承认，我们的模型所用的一些详细机制，虽然在神经学上是可信的，但可能过于简化。我们期待进一步的改进和具有挑战性的数据，这些数据将使我们重新审视一些基本机制。

### 3.行为选择和学习的抽象模型

与前一节描述的神经网络模型相比，抽象模型通常不捕获神经生物学或神经解剖过程，而是关注可能导致特定行为输出（如学习和决策）的认知操作的本质。尽管这些模型与神经

生物学事件有关，并且在某些情况下，包括特定的神经过程，如多巴胺的作用[182,34]，但这些模型通常不受已知的生物学限制（既不包括解剖学也不包括生理学）。尽管如此，通过采用自上而下的功能方法，这些模型在揭示奖励引导学习和决策的认知机制方面被证明是很有价值的，并在将这些机制与基底神经节、前额皮质和多巴胺系统的神经生物学联系起来方面取得了若干进展[34,111,47,122]。

### 3.1.模型背后的数学

我们着重研究在理解基底神经节功能方面应用最广泛的模型。这些强化学习模式背后的基本学习机制可以用 Thorndike 效应定律[165]在语义上加以概括：在对同一情景所作的几种反应中，在其他条件相同的情况下，伴随或紧跟着对动物感到满意的反应将与情景更紧密地联系在一起，因此，当情景再次出现时，它们更有可能再次出现；在其他条件相同的情况下，那些伴随或紧跟着动物感到不适的动物，它们与这种情况的联系就会减弱，因此，当这种情况再次发生时，它们就不太可能发生了。满足感或不适感越强，这种联系就越强或越弱。

换句话说，与积极反馈相关的行动更有可能被重复，而与消极反馈相关的行动则不太可能被重复。在模型中，不同的动作可以用 Q 值表示；相对于其他动作，Q 值越大，模型越有可能选择该动作。然而，选择函数策略通常是概率的，因此有时会选择 Q 值较低的其他选择。这确保了模型偶尔会探索其他可选的行为，从而避免了其他决策选项提供更高的回报但没有被选中的情况，因为模型不断地选择一个决策选项（即，局部极小值）[160]。最常用的选择函数被称为 SoftMax，因为它为选择具有最大 Q 值的动作分配了更高的概率，但 Q 值之间的仲裁是软的，因此只有较小值的动作几乎同样可能被选择。SoftMax 函数的斜率决定了选择最大 Q 值的程度与进行探索性选择的概率[160,48,59]。

为了了解哪个 Q 值会带来最高的回报，Q 值会根据增强值进行调整。更新 Q 值最常用的方法是通过一个奖励预测误差，即期望和收到的奖励之间的差异： $\delta=r-Q$ ，其中  $\delta$  是预测误差， $r$  是奖励， $Q$  是与所选动作对应的权重值（在更复杂的算法中，预测误差不仅考虑当前的奖励，还考虑基于先前学习的对未来试验的预测奖励，以及对未来奖励的折现[172]。这一功能对于增强学习代理不仅能够了解哪些行为会导致立即的奖励，而且还能够了解哪些行为会在稍后发生其后果时加强，从而最大化未来的总奖励[160]。然而，我们在这里只讨论简单的情况，即简单的行为会导致立即的回报或回报的缺乏。）。此预测误差项可能反映了中脑多巴胺神经元的阶段性活动，详情见下文[157]。因此，当遵循特定行动的奖励大于该特定行动预期的奖励（即 Q 值）时，预测误差为正；当奖励与预期完全一致时，预测误差为零；当奖励低于预期时，预测误差为负。这些预测误差随后在后续试验中调整 Q 值： $Q(t+1)=Q(t)+\delta$ ，其中  $t$  指试验。导致奖励（惩罚）的 Q 值被加强（削弱），因此在随后的试验中被选择的可能性越来越小。因此，Q 值更新方程可以看作是部分 Thorndike 效应定律的简明数学表示。注意，预测误差项乘以一个学习率，该学习率衡量预测误差对后续 Q 值的影响： $Q(t+1)=Q(t)+\alpha*\delta$ 。

学习率描述了预测误差调整 Q 值的程度，并可能与单次学习经验激活的 AMPA 受体的相对数量相对应。这些学习率在不同的大脑区域也可能不同。例如，海马学习系统能够快速、单次试验学习（学习率高），而基底神经节学习系统通过随着时间的推移更慢地整合来学习，从而利用较低的学习率。因此，一个重要的计算问题是确定何时使用学习率高与低的系统。一些人认为不确定性的数量在决定学习率方面起着作用[14,47,187]。

学习 Q 值也可能有助于解释我们如何形成习惯，正如“优势学习”中的形式化一样[50]。“优势学习”理论指出，当与该行为相关的值超过该状态(例如，时间点)的整个可能行为集的

平均值时，选择该行为。随着时间的推移，随着智能体学习最优响应策略，特定操作的优势会下降，因为该操作状态的总体价值会增加。在这一点上，行为选择变得更加自动化，刺激反应习惯形成。我们的神经模型也显示了类似的转变，从根据奖励选择行为到根据习惯选择行为。在我们的神经模型中，这种转变是通过在皮质-皮质投射中缓慢的 Hebb 学习，在许多试验中逐渐发生的，如上所述。

注意，在神经网络和抽象模型之间，强化学习的基本原则是加强奖励行为的表示，同时削弱非奖励行为的表示。神经网络模型更关注假定的神经实现，而 Q 模型抽象神经实现，有利于集中于基本的计算实现。

许多用来理解基底神经节功能的模型比这些简单的方程更为复杂。例如，其他抽象模型讨论了动物如何在基于 BG 的习惯性系统与位于前额叶皮质的目标导向系统之间进行仲裁 [47]，何时在动态概率环境中进行探索 [48,107]，在可变的奖励计划中如何作出反应 [121]，以及何时补充多巴胺介导的基础强化学习，并明确检测环境何时发生变化 [74]。然而，上述基本方程在许多情况下都是稳健的，并继续构成这些更复杂模型的支柱。尽管神经生物学的许多方面没有纳入这些模型（例如膜电位动力学、D1 和 D2 受体的不同作用、不同 BG 核的作用），这些方程预测特定纹状体和前额叶区域的活动，证明了这些方程的优雅性。简单但强大的模型，用于解释基底神经节的计算，而无需对神经回路中的精确实现形式进行太多假设。

## 3.2.抽象模型的神经生物学关联

### 3.2.1.预测误差的神经生物学

奖励预测误差被认为是由中脑多巴胺细胞在 VTA 和 SNc 的阶段性爆发活动所发出的信号。多巴胺在纹状体和边缘系统的广泛区域迅速释放。就像上面描述的强化学习模型中的预测误差项一样，中脑多巴胺活动在接收到意想不到的奖励时阶段性地增加，在没有接收到预期奖励时阶段性地减少，在接收到预期奖励时与基线水平没有变化。对这一证据的详细审查可在其他地方找到 [143,142,145]。

多巴胺细胞活动与计算模型中的预测误差项之间的联系启发了许多研究人员，他们利用人类的非侵入性神经成像技术来研究预测误差与神经的关系。例如，在功能性核磁共振研究中，计算强化学习模型，类似于上面概述的，被用来在每次试验中产生奖励预测误差。然后，这些预测误差被用于回归，以确定大脑区域的活动与从模型中得到的预测误差相关的区域。这些相关性通常在纹状体和额叶皮层以及其他区域（下面将更深入地讨论）中被观察到，并被用来反映从中脑到纹状体回路的奖赏预测误差信号 [34,124,123,148]。

在另一项使用头皮记录的人类脑电图 (EEG) 的研究中，研究人员发现了一种称为误差-相关负性 (ERN) 和反馈-相关负性 (FRN) 的成分，这种成分可能反映了奖励预测误差信号 [185,80,37,64,120]。这些成分位于前额中央头皮部位，与正反馈相比，负反馈后约 200-400 ms，或与正确响应相比，误差后约 200-400 ms。有人提出，FRN 反映了来自中脑多巴胺系统的负奖励预测误差信号的影响，然后将其用于适应奖励寻求行为 [80,23]。这与中脑多巴胺神经元投射并能调节扣带回皮质锥体细胞的活动发现是一致的 [125]。然而，考虑到额叶皮质 DA 再摄取的缓慢过程，扣带回是否能检测 DA 下降尚不清楚（如上所述）。然而，有可能这些头皮脑电图记录实际上反映了 BG 中 DA 下降的影响，后者激活了 NoGo 通路，然后间接导致额叶皮质活动的改变（例如，通过增加反应后冲突 [186]）。

### 3.2.2. “行为(Q值)” 神经生物学

这些强化学习模型的另一个主要组成部分是 Q 值，它代表特定的行动或决定。尽管与预测误差的神经生物学相关因素相比，Q 值的可能神经生物学相关因素受到的关注较少，但有证据表明，模型中的 Q 值可能与负责计划和执行这些特定行为的大脑区域的活动相对应。例如，纹状体中代表特定动作的神经元的活动（例如向右或向左的扫视）受正确反应获得的奖励量的调节[139]。在本研究中，这些神经元的特性将通过 Q 学习算法来拟合。此外，与奖励相关的运动区域活动调制可能会使决策和行动选择过程产生偏差[69,140,156]。虽然这些发现并不总是以计算学习模型中的 Q 值来讨论，但是观察结果与模型中的 Q 值或权重与感觉运动系统中的活动相对应的观点是一致的。人类的初步证据表明，皮质运动区的活动可能与 Q 值相对应。例如，[37]报道，外侧额叶电极部位（位置 c3/4，通常用于指示运动皮层活动）上的脑电图活动类似于从计算模型中获得的 Q 值，而模型玩的是人类受试者玩的相同的战略游戏。

一个重要的问题是，Q 值在大脑中的意义是什么，它存储在哪里。如前一段所述，对于每个决策映射到特定动作或响应上的简单决策（例如，向左扫视，或按右手食指），Q 值可能与基底神经节和/或皮质运动区域的运动活动激活的强度相对应。但是我们面临的大多数决定都比较复杂，并且没有与之相关的具体的、离散的运动行为（例如，上哪所大学？晚饭吃什么？我应该嫁给这个人吗？）。与此相关的是，在一些实验中，相同的刺激在不同的实验中与不同的运动反应相关。这对于平衡运动反应的需求是有用的，但仍然存在这样一个问题：这些实验中的 Q 值是否与刺激表征相关联，或者它们是否仍然与更抽象的反应表征相关联，后者是灵活的，并根据任务需求进行变化。一种可能性是，多个 Q 型的表征由不同的大脑区域维持，并对应于不同类型信息的奖励调节权重。例如，眶额皮质或腹侧纹状体可能包含特定世界状态的基本值表示（与动作无关）；背侧纹状体和补充运动区域可能包含特定运动动作的 Q 型表示；以及背外侧或前额叶前部皮层可能包含更抽象的目标或计划的 Q 型表征。

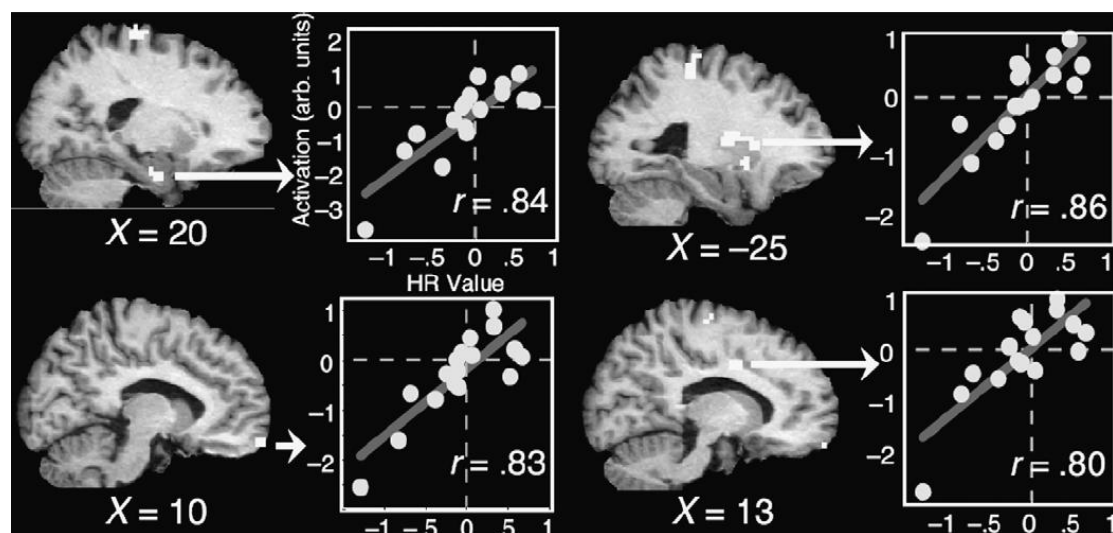


图 4：抽象强化学习模型可用于研究个体差异。在此，我们使用一个模型来评估强化(在赌博任务中赢钱或不赢钱)对后续试验中进行低风险或高风险赌博的可能性的影响。每个受试者的最佳拟合参数决定了高风险选项在获得高风险奖励后的权重变化幅度和符号。这个参数的个体差异与强化相关的大脑激活相关。结果表明，在包括侧纹状体(右上角)在内的区域网络中，这个权重更新参数(x 轴)预测大脑对大奖励的激活是否与随后的高风险(y 轴正值)或非高风险(负值)选择有关。在这种情况下，个体差异被证明对理解强化如何指导后续决策至关重要：对一些受试者来说，与奖励相关的活动预测了他

们做出后续风险选择的可能性增加，而对另一些受试者来说，它根据他们的估计参数预测了可能性降低。详见[36]。

### 3.3.个体差异

上面描述的强化学习方程式是规范的，因为它们规定了所有人应该如何行动和从强化中学习。然而，人类的决策是可变的；无数的个体差异影响着人们的决策方式，不同的个体可以采取不同的行动和学习，即使在相同的行动之后给予相同的强化。抽象模型的一个优点是，它们可以用来在数学上描述这种个体差异。这是通过将模型拟合到每个受试者的行为数据，并通过统计拟合程序估计一些模型变量来完成的。例如，我们可以估计每个受试者的独特学习率，它衡量预测误差对 Q 值调整的影响。该方法已成功地用于将行为任务表现和基底神经节和额叶皮质的大脑活动与决策个体差异联系起来[36,34,141,14]。Frank 及其同事最近证实，与人类纹状体和前额叶皮质多巴胺受体表达相关的遗传多态性与不同的学习率有关[59]。此外，通过 DARPP-32 和 DRD2 基因预测获得（Go）和损失（NoGo）的单独学习率，为神经网络模型提供了一个很好的映射。Lee 和他的同事在猴子身上发现，根据这些估计的模型参数可以预测前额叶皮质细胞的活动[98,99]。

当受试者在如何使用增强功能来调整决策（例如，在没有正确答案或学习策略的赌博研究中[36]）方面存在很大差异时，将模型参数与受试者数据相匹配对于阐明决策模型的神经计算机制至关重要（图 4）。在这些情况下，忽略个别差异（即规范性方法）可能导致错误的解释，即模型不能解释数据。

### 3.4.抽象模型与神经生物学联系的不确定性和不一致性

虽然已有文献表明，额纹状体回路的活动与抽象计算模型的某些方面有关，但在涉及到何种大脑系统以及大脑活动与抽象模型预测的一致性和不确定性仍然存在。这种不确定性的部分原因在于，模型比实际的基底神经节系统要简单得多。例如，上面详细描述方程不太可能描述实验学习任务中涉及的所有内部心理过程，即使在神经系统简单的物种中也是如此；人类和动物很可能与这些机制以及其他复杂而动态的高级过程（如假设测试）类似。

不确定性的一个方面涉及正预测误差和负预测误差。如前一节所述，最近的研究表明多巴胺细胞发放暂停的持续时间可能编码负预测误差[13]。血清素系统也被认为在负预测误差信号中起作用[46]。关于这个问题，功能性 MRI 的文献不太清楚：一些研究发现，正/负预测误差导致额纹状体回路活性增加/减少[34,106,124]，而另一些研究发现腹侧纹状体活性仅与正预测误差相关[184]。然而，也有人提出纹状体的不同亚区与积极和消极预测误差有关[147]。在中脑的研究报道激活，有些报道活动增加对奖励与惩罚[114]相比，其他人已经报道的中脑活动增加消极与积极的反馈[5]相比，和其他人已经报道增加但无意义的积极的预测误差减少消极预测误差[45]。

另一个不确定性领域涉及活动与预测误差相关的精确区域：尽管通常报告腹侧纹状体的预测误差相关活动，但不同的研究也显示，背侧纹状体的预测误差型反应，如前额皮质区域，包括眶额、腹外侧和背外侧、中脑和小脑[106,148,124,76,134]。根据当前的任务目标，大脑中的不同网络可能会利用预测误差，尽管据我们所知，这还没有被研究过。

对 BOLD 反应的解释还有另一个问题，尤其是对基底神经节的解释：BOLD 反应是一种时间和空间迟缓的信号，无法区分不同类型的神经元或不同功能网络的活动，尤其是当这些网络在空间上重叠时。例如，Go 和 NoGo 细胞的网络在空间上是重叠的，所以我们不能用功能性 MRI 来区分这些系统。同样，功能性 MRI 不能将中间神经元与中等多棘神经元分离，



也不能将不同频率或特定基底神经节结构内不同亚区的振荡分离。如果假设 Go 和 NoGo 细胞相互对立，这个问题可能会变得非常重要。在这种情况下，与低 Go 和高 NoGo 活性条件相比，高 Go 和低 NoGo 活性条件下纹状体的 BOLD 反应可能没有差异。

尽管这些研究之间的不一致性与它们的共性相比可能是相对较小的，但是中脑多巴胺神经元发出的奖励预测误差信号的理论已经证明是一个显著的理论，其简单、优雅，并且能够将极为不同的神经细胞连接在一起。从人工智能到细胞电生理学再到人类神经影像学的研究。它继续激发新的、创造性的和跨学科的研究，并为基底神经节回路在强化学习和决策中的作用提供了新的线索。

#### 4.神经网络与抽象模型的整合

这两种理解基底神经节计算功能的方法传统上是分开进行的，通常由不同的研究小组进行。如前几节所述，不同的模型有不同的优点和缺点。如果它们的优势和局限相匹配，那么结合这两种建模方法可能比单独使用两种方法更有成效。例如，抽象模型（而不是神经网络模型）可以评估学习率和其他参数中的个体差异，并将这些个体差异与表现或大脑活动联系起来；相比之下，神经网络模型，而不是抽象模型，对功能计算如何通过多个脑区之间的交互动力学产生做出了具体的预测，进而预测了局灶性脑损伤、药物操作和遗传学的影响。将这些建模方法结合起来的一种方法是使用抽象数学模型来估计。将这些建模方法结合起来的一种方法是使用抽象的数学模型来估计神经网络模型的学习参数，就像它是一门人类学科一样。也就是说，当评估个体的学习速度时，当面对特定的强化序列时，抽象模型通常适合考虑给定对象的实际逐项试验选择。我们可以应用相同的程序，将神经网络模型的输出视为行为选择，然后使用抽象模型来估计神经网络模型作为整个系统使用的学习率（这可能与给定突触的学习率有很大差异）。这可能有助于理解个体行为学习率差异的神经生物学。尽管一些研究已经调查了学习率的个体差异以及这些个体在行为和大脑活动方面的差异之间的关联，但仍不清楚什么样的神经生物学因素可能导致不同的个体有不同的学习率。是多巴胺系统反应幅度、多巴胺受体浓度，还是苍白球丘脑效应？根据经验，很难确定导致行为学习率差异的神经生物学机制。然而，这正是神经网络模型变得有用的地方：可以操纵神经网络模型中的各种参数，并且可以在虚拟实验中对模型进行测试。不同模型版本（例如，多巴胺系统功能完整或受损的模型，以模拟帕金森氏病）产生的学习率可与不同学习率的受试者组进行比较。如果来自不同模型版本的学习率与来自不同受试者组（例如，具有不同基因型的受试者）的学习率相匹配，我们可以得出结论，对模型所做的更改代表了一种生物学上可信的机制，通过这种机制可以实现不同的学习率。当然，对神经网络模型所做的改变应该由先验假设驱动，并受生理证据的约束。这将为抽象模型提供一个重要的验证，因为尽管大脑不同区域的激活与学习率个体差异产生的模型参数相关[34,162]，但尚不清楚哪种生物过程可以解释这些差异。

#### 5.结论和未来方向

像这里讨论的那些计算模型是理论，和所有理论一样，它们被简化了，范围有限，并且可能会随着新的经验数据改进我们的理解而经历重大的修改。许多实证论文往往依赖于概念模型，但这些往往是静态解剖图，缺乏这里所述模型的数学精度，而且往往是相对简单的。这里讨论的计算模型在某种意义上是相似的，它们都是省略了许多细节的简化版本。然而，与数学基础不那么扎实的理论相比，计算模型有明显的优势：它们可以进一步考虑大脑试图解决的计算功能、这些计算的实现以及基底神经节回路的丰富动态。最终，由特定模型捕获的数据模式应该由上面一层的模型（为了优雅、分析的易处理性和简洁）和下面一层的模型

(为了探索更详细的生物物理约束并相应地调整模型) 再现。

在过去的几十年中, 计算建模领域, 特别是基底神经节系统的建模领域已经有了很大的发展, 我们设想未来利用计算建模来理解基底神经节及其相关回路的几个平行方向。由于新软件的出现, 使得进入这一领域变得容易, 以及计算机硬件速度和效率的提高, 很可能会有更多的研究人员使用更加复杂和生物细节化的模型。随着计算机速度的提高, 并行处理越来越普遍, 高度详细的神经模型可以被放大到一个水平, 在这个水平上它们可以产生行为和认知上有意义的输出。我们还设想计算模型将更多地与实证研究相结合。在本综述中讨论了预测误差和相关强化学习变量的假定神经机制。最后, 从神经生物学上可信的计算模型得出的见解可能会越来越多地进入神经科学以外的领域, 如人工智能和机器人学[71]。

## 参考文献

- [1] Aizman O, Brismar H, Uhlen P, Zettergren E, Levet AI, Forsberg H, et al. An anatomical and physiological evidence for D1 and D2 dopamine receptor colocalization in neostriatal neurons. *Nat Neurosci* 2000;3:226–30.
- [2] Albin RL, Young AB, Penney JB. The functional anatomy of basal ganglia disorders. *Trends Neurosci* 1989; 12:366–75.
- [3] Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M. Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J Neurosci* 1994;14:3969–84.
- [4] Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging(mri) and functional mri. *J Neurosci* 2007;27(14):3743–52.
- [5] Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J Neurophysiol* 2004;92(2):1144–52.
- [6] Ashby FG, O'Brien JB. The effects of positive versus negative feedback on information-integration category learning. *Percept Psychophys* 2007;69(6):865–78.
- [7] Atallah HE, Lopez-Paniagua D, Rudy JW, O'Reilly RC. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat Neurosci* 2007;10:126–31.
- [8] Aubert I, Ghorayeb I, Normand E, Bloch B. Phenotypical characterization of the neurons expressing the D1 and D2 dopamine receptors in the monkey striatum. *J Comp Neurol* 2000;418:22–32.
- [9] Bäckman L, Ginovart N, Dixon RA, Wahlin TR, Wahlin A, Halldin C, et al. Age-related cognitive deficits mediated by changes in the striatal dopamine system. *Am J Psychiatry* 2000;157(4):635–7.
- [10] Baunez C, Christakou A, Chudasama Y, Forni C, Robbins TW. Bilateral high-frequency stimulation of the subthalamic nucleus on attentional performance: transient deleterious effects and enhanced motivation in both intact and parkinsonian rats. *Eur J Neurosci* 2007;25(4):1187–94.
- [11] Baunez C, Robbins TW. Bilateral lesions of the subthalamic nucleus induce multiple deficits in an attentional task in rats. *Eur J Neurosci* 1997;9(10):2086–99.
- [12] Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 2005;47(1):129–41.
- [13] Bayer HM, Lau B, Glimcher PW. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J Neurophysiol* 2007;98(3):1428–39.
- [14] Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nat Neurosci* 2007;10(9):1214–21.
- [15] Benabid AL. Deep brain stimulation for Parkinson's disease. *Curr Opin Neurobiol* 2003;13(6):696–706.
- [16] Benazzouz A, Hallett M. Mechanism of action of deep brain stimulation. *Neurology* 2000;55:13–S16.
- [17] Berke JD, Hyman SE. Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 2000;25(3):515–32.
- [18] Berretta S, Parthasarathy HB, Graybiel AM. Local release of GABAergic inhibition in the motor cortex induces immediate-early gene expression in indirect pathway neurons of the striatum. *J Neurosci* 1997;17(12):4752–63.
- [19] Berretta S, Sachs Z, Graybiel AM. Cortically driven Fos induction in the striatum is amplified by local dopamine D2-class receptor blockade. *Eur J Neurosci* 1999;11(12):4309–19.
- [20] Bloch B, LeMoine C. Neostriatal dopamine receptors. *Trends Neurosci* 1994;17:3–4.
- [21] Bogacz R, Gurney K. The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Comput* 2007;19(2):442–77.
- [22] Brown J, Bullock D, Grossberg S. How the basal ganglia use parallel excitatory and inhibitory learning pathways to selectively respond to unexpected rewarding cues. *J Neurosci* 1999;19:10502–11.
- [23] Brown JW, Braver TS. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 2005;307(5712):1118–21.
- [24] Brown JW, Bullock D, Grossberg S. How laminar frontal cortex and basal ganglia circuits interact to control planned and reactive saccades. *Neural Networks* 2004;17:471–510.

- [25] Calabresi P, Gubellini P, Centonze D, Picconi B, Bernardi G, Chergui K, et al. Dopamine and camp-regulated phosphoprotein 32 kda controls both striatal long-term depression and long-term potentiation, opposing forms of synaptic plasticity. *J Neurosci* 2000;20(22):8443–51.
- [26] Calabresi P, Pisani A, Centonze D, Bernardi G. Synaptic plasticity and physiological interactions between dopamine and glutamate in the striatum. *Neurosci Biobehav Rev* 1997;21:519–23.
- [27] Cardinal RN. Neural systems implicated in delayed and probabilistic reinforcement. *Neural Networks* 2006;19:1277–301.
- [28] Cardinal RN, Parkinson JA, Hall J, Everitt BJ. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci Biobehav Rev* 2002;26:321–52.
- [29] Carta AR, Tronci E, Pinna A, Morelli M. Different responsiveness of striatonigral and striatopallidal neurons to DOPA after a subchronic intermittent-DOPA treatment. *Eur J Neurosci* 2005;21(5):1196–204.
- [30] Castle M, Aymerich MS, Sanchez-Escobar C, Gonzalo N, Obeso J A, Lanciego JL. Thalamic innervation of the direct and indirect basal ganglia pathways in the rat: ipsi- and contralateral projections. *J Comp Neurol* 2005;483(2):143–53.
- [31] Centonze D, Gubellini P, Bernardi G, Calabresi P. Permissive role of interneurons in corticostriatal synaptic plasticity. *Brain Res Brain Res Rev* 1999;31(1):1–5.
- [32] Centonze D, Picconi B, Gubellini P, Bernardi G, Calabresi P. Dopaminergic control of synaptic plasticity in the dorsal striatum. *Eur J Neurosci* 2001;13:1071–7.
- [33] Chevalier G, Deniau JM. Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci* 1990;13:277–80.
- [34] Cohen MX. Individual differences and the neural representations of reward expectation and reward prediction error. *Soc Cogn Affect Neurosci* 2007;2(1):20–30.
- [35] Cohen MX, Lombardo MV, Blumenfeld RS. Covariance-based subdivision of the human striatum using t1-weighted mri. *Eur J Neurosci* 2008;27:1446–534.
- [36] Cohen MX, Ranganath C. Behavioral and neural predictors of upcoming decisions. *Cogn Affect Behav Neurosci* 2005;5(2):117–26.
- [37] Cohen MX, Ranganath C. Reinforcement learning signals predict future decisions. *J Neurosci* 2007;27(2):371–8.
- [38] Collins P, Wilkinson LS, Everitt BJ, Robbins TW, Roberts AC. The effect of dopamine depletion from the caudate nucleus of the common marmoset (*Callithrix jacchus*) on tests of prefrontal cognitive function. *Behav Neurosci* 2000;114:3–17.
- [39] Cools R. Dopaminergic modulation of cognitive function – implications for L-DOPA treatment in Parkinson's disease. *Neurosci Biobehav Rev* 2006;30(1):1–23.
- [40] Cools R, Altamirano L, D'Esposito M. Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia* 2006;44:1663–73.
- [41] Cools R, Barker RA, Sahakian BJ, Robbins TW. Enhance dopamine impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cerebral Cortex* 2001;11:1136–43.
- [42] Cools R, Barker RA, Sahakian BJ, Robbins TW. Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain* 2001;124:2503–12.
- [43] Costa RM, Gutierrez R, de Araujo IE, Coelho MRP, Kloth AD, Gainetdinov RR, et al. Dopamine levels modulate the updating of task values. *Genes Brain Behav* 2007;6(4):314–20.
- [44] Cragg SJ. Meaningful silences: how dopamine listens to the a pause. *Trends Neurosci* 2006;29(3):125–31.
- [45] D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. Bold responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 2008;319(5867):1264–7.
- [46] Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. *Neural Networks* 2002;15:603–16.
- [47] Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 2005;8(12):1704–11.
- [48] Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature* 2006;441(7095):876–9.
- [49] Dayan P, Abbott L. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press; 1999.
- [50] Dayan P, Balleine BW. Reward, motivation, and reinforcement learning. *Neuron* 2002;36:285–98.
- [51] Durstewitz D, Seamans JK, Sejnowski TJ. Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J Neurophysiol* 2000;83:1733–50.
- [52] Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* 2005;8(11):1481–9.
- [53] Féger J, Crossman AR. Identification of different subpopulations of neostriatal neurons projecting to globus pallidus or substantia nigra in the monkey: a retrograde fluorescence double-labelling study. *Neurosci Lett* 1984;49(1–2):7–12.
- [54] Frank MJ. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated parkinsonism. *J Cognit Neurosci* 2005;17:51–72.
- [55] Frank MJ. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Networks* 2006;19:1120–36.
- [56] Frank MJ, Claus ED. Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 2006;113(2):300–26.
- [57] Frank MJ, Kong L. Learning to avoid in older age. *Psychol Aging* 2008;23: 392–8.
- [58] Frank MJ, Loughry B, O'Reilly RC. Interactions between the frontal cortex and basal ganglia in working memory: a computational model. *Cognit Affect Behav Neurosci* 2001;1:137–60.
- [59] Frank MJ, Moustafa AA, Haughey H, Curran T, Hutchison K. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci* 2007;104:16311–6.
- [60] Frank MJ, O'Reilly RC. A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and

- haloperidol. *Behav Neurosci* 2006;120:497–517.
- [61] Frank MJ, Samanta J, Moustafa AA, Sherman SJ. Hold your horses: impulsivity, deep brain stimulation and medication in parkinsonism. *Science* 2007;318:1309–12.
- [62] Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* 2007;32:1583–99.
- [63] Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 2004;306:1940–3.
- [64] Frank MJ, Woroch BS, Curran T. Error-related activity predicts reinforcement learning and conflict biases. *Neuron* 2005;47:495–501.
- [65] Gerfen CR. The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu Rev Neurosci* 1992;15:285–320.
- [66] Gerfen CR, Keefe KA. Neostriatal dopamine receptors. *Trends Neurosci* 1994;17:2–3.
- [67] Gerfen CR, Keefe KA, Gauda EB. D1 and D2 dopamine receptor function in the striatum: coactivation of D1- and D2-dopamine receptors on separate populations of neurons results in potentiated immediate early gene response in D1-containing neurons. *J Neurosci* 1995;15:8167–76.
- [68] Gerfen CR, Wilson C. The basal ganglia. In: Swanson L, Bjorkland A, Hokfelt T, editors. *Handbook of chemical neuroanatomy*. vol 12: integrated systems of the CNS. Amsterdam: Elsevier; 1996. p. 371–468.
- [69] Gold JI, Shadlen MN. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 2002;36:299–308.
- [70] Gonon FJ. Prolonged and extrasynaptic excitatory action of dopamine mediated by D1 receptors in the rat striatum in vivo. *J Neurosci* 1997;17:5972–8.
- [71] Gurney K, Prescott TJ, Wickens JR, Redgrave P. Computational models of the basal ganglia: from robots to membranes. *Trends Neurosci* 2004;27(8):453–9. [72] Haber SN. The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat* 2003;26(4):317–30.
- [73] Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 2000;20:2369–82.
- [74] Hampton AN, Bossaerts P, O'Doherty JP. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J Neurosci* 2006;26(32):8360–7.
- [75] Harden DG, Grace AA. Activation of dopamine cell firing by repeated l-DOPA administration to dopamine-depleted rats: its potential role in mediating the therapeutic response to l-DOPA treatment. *J Neurosci* 1995;15(9): 6157–66.
- [76] Haruno M, Kawato M. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus–action–reward association learning. *Neural Network* 2006;19(8): 1242–54.
- [77] Hernandez-Lopez S, Vargas J, Surmeier DJ, Reyes A, Galarraga E. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca<sup>2+</sup> conductance. *J Neurosci* 1997;17:3334–42.
- [78] Hernandez-Lopez S, Tkatch T, Perez-Garci E, Galarraga E, Vargas J, Hamm H, et al. D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca<sup>2+</sup> currents and excitability via an ovelplc[beta]1-ip3-calcineurin-signaling cascade. *J Neurosci* 2000;20(24):8987–95.
- [79] Hirvonen M, Laakso A, Abd Nagren K, Rinne J, Pohjalainen T, Hietala J. C957t polymorphism of the dopamine d2 receptor (drd2) gene affects striatal drd2 availability in vivo (corrigendum). *Mol Psychiatry* 2005;10:889.
- [80] Holroyd CB, Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 2002;109:679–709.
- [81] Houk JC. Agents of the mind. *Biol Cybern* 2005;92(6):427–37.
- [82] Houk JC, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk JC, Davis JL, Beiser DG, editors. *Models of information processing in the basal ganglia*. Cambridge, MA: MIT Press; 1995. p. 233–48.
- [83] Humphries MD, Stewart RD, Gurney KN. A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J Neurosci* 2006;26(50):12921–42.
- [84] Ince E, Ciliax BJ, Levey AI. Differential expression of D1 and D2 dopamine and m4 muscarinic acetylcholine receptor proteins in identified striatonigral neurons. *Synapse* 1997;27:257–366.
- [85] Joel D, Niv Y, Ruppin E. Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks* 2002;15:535–47.
- [86] Joel D, Weiner I. The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* 2000;96:451.
- [87] Kaasinen V, Rinne JO. Functional imaging studies of dopamine system and cognition in normal aging and Parkinson's disease. *Neurosci Biobehav Rev* 2002;26(7):785–93.
- [88] Kawaguchi Y, Wilson CJ, Emson PC. Projection subtypes of rat neostriatal matrix cells revealed by intracellular injection of biocytin. *J Neurosci* 1990;10(10):3421–38.
- [89] Keller RW, Kuhr WG, Wightman RM, Zigmond MJ. The effect of l-dopa on in vivo dopamine release from nigrostriatal bundle neurons. *Brain Res* 1988;447(1):191–4.
- [90] Kerr JN, Wickens JR. Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. *J Neurophysiol* 2001;85:117–24.
- [91] Klein TA, Neumann J, Reuter M, Hennig J, von Cramon DY, Ullsperger M. Genetically determined differences in learning from errors. *Science* 2007;318(5856):1642–5.
- [92] Knapka E, Kaczmarek L. A gene for neuronal plasticity in the mammalian brain: Zif268/egr-1/ngfi-a/krox-24/tis8/zenk? *Prog Neurobiol* 2004;74(4):183–211.
- [93] Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. *Science* 1997;278:52–8.
- [94] Kravtsov Y, Kudryavtseva E, McKee AC, Geula C, Kowall NW, Khrapko K. Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat Genet* 2006;38(5): 518–20.

- [95] Kreitzer AC, Malenka RC. Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* 2007;445(7128):643–7.
- [96] Lapper SR, Bolam JP. Input from the frontal cortex and the parafascicular nucleus to cholinergic interneurons in the dorsal striatum of the rat. *Neuroscience* 1992;51(3):533–45.
- [97] Le Moine C, Bloch B. D1 and D2 dopamine receptor gene expression in the rat striatum: sensitive cRNA probes demonstrate prominent segregation of D1 and D2 mRNAs in distinct neuronal populations of the dorsal and ventral striatum. *J Comp Neurol* 1995;355:418–26.
- [98] Lee D, Conroy ML, McGreevy BP, Barraclough DJ. Reinforcement learning and decision making in monkeys during a competitive game. *Brain Res Cogn Brain Res* 2004;22(1):45–58.
- [99] Lee D, McGreevy BP, Barraclough DJ. Learning and decision making in monkeys during a rock–paper–scissors game. *Brain Res Cogn Brain Res* 2005;25(2):416–30.
- [100] Lei W, Jiao Y, Del Mar N, Reiner A. Evidence for differential cortical input to direct pathway versus indirect pathway striatal projection neurons in rats. *J Neurosci* 2004;24(38):8289–99.
- [101] Lèvesque M, Parent A. The striatofugal fiber system in primates: an evaluation of its organization based on single-axon tracing studies. *Proc Natl Acad Sci USA* 2005;102(33):11888–93.
- [102] Lindskog M, Kim M, Wikström MA, Blackwell KT, Kotaleski JH. Transient calcium and dopamine increase pka activity and darpp-32 phosphorylation. *PLoS Comput Biol* 2006;2(9):e119.
- [103] Liu FC, Graybiel AM. Region-dependent dynamics of cAMP response element binding protein phosphorylation in the basal ganglia. *Proc Natl Acad Sci USA* 1998;95:4708–13.
- [104] Madras BK, Miller GM, Fischman AJ. The dopamine transporter and attention deficit/hyperactivity disorder. *Biol Psychiatry* 2005;57(11):1397–409.
- [105] Magill PJ, Sharott A, Bevan MD, Brown P, Bolam JP. Synchronous unit activity and local field potentials evoked in the subthalamic nucleus by cortical stimulation. *J Neurophysiol* 2004;92(2):700–14.
- [106] McClure SM, Berns GS, Montague PR. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 2003;38:339–46.
- [107] McClure SM, Gilzenrat MS, Cohen JD. An exploration-exploitation model based on norepinephrine and dopamine activity. In: *Advances in neural information processing systems vol 18*. MIT Press; 2006. p. 867–74.
- [108] Meissner W, Leblois A, Hansel D, Bioulac B, Gross CE, Benazzouz A, et al. Subthalamic high frequency stimulation resets subthalamic firing and reduces abnormal oscillations. *Brain* 2005;128(Pt 10):2372–82.
- [109] Meyer-Lindenberg A, Straub RE, Lipska BK, Verchinski BA, Goldberg T, Callicott JH, et al. Genetic evidence implicating darpp-32 in human frontostriatal structure, function, and cognition. *J Clin Invest* 2007;117(3):672–82.
- [110] Mink JW. The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 1996;50:381–425.
- [111] Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 1996;16:1936–47.
- [112] Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* 2004;43:133–43.
- [113] Mouroux M, Féger J. Evidence that the parafascicular projection to the subthalamic nucleus is glutamatergic. *Neuroreport* 1993;4(6):613–5.
- [114] Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, et al. Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Mol Psychiatry* 2008;13(3):239, 267–239, 276.
- [115] Nakamura K, Hikosaka O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci* 2006;26(20):5360–9.
- [116] Nakano K, Kayahara T, Tsutsumi T, Ushiro H. Neural circuits and functional organization of the striatum. *J Neurol* 2000;247(Suppl. 5):1–15.
- [117] Nambu A, Tokuno H, Hamada I, Kita H, Imanishi M, Akazawa T, et al. Excitatory cortical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *J Neurophysiol* 2000;84:289–300.
- [118] Nambu A, Tokuno H, Takada M. Functional significance of the cortico-subthalamo-pallidal ‘hyperdirect’ pathway. *Neurosci Res* 2002;43:111–7.
- [119] Nicola SM, Surmeier J, Malenka RC. Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu Rev Neurosci* 2000;23:185–215.
- [120] Nieuwenhuis S, Holroyd CB, Mol N, Coles MGH. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neurosci Biobehav Rev* 2004;28:441–8.
- [121] Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 2007;191(3):507–20.
- [122] O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 2004;304(5669):452–4.
- [123] O’Doherty JP. Lights, camera, action! the role of human orbitofrontal cortex in encoding stimuli, rewards, and choices. *Ann N Y Acad Sci* 2007;1121:254–72.
- [124] O’Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ. Temporal difference models and reward-related learning in the human brain. *Neuron* 2003;38:329–37.
- [125] Onn SP, Wang XB. Differential modulation of anterior cingulate cortical activity by afferents from ventral tegmental area and medial dorsal thalamus. *Eur J Neurosci* 2005;21:2975–92.
- [126] O’Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput* 2006;18:283–328.
- [127] O’Reilly RC, Frank MJ, Hazy TE, Watz B. PVLV: the primary value and learned value pavlovian learning algorithm. *Behav Neurosci* 2007;121:31–49.

- [128] O'Reilly RC, Munakata Y. Computational explorations in cognitive neuroscience: understanding the mind by simulating the brain. Cambridge, MA: The MIT Press; 2000.
- [129] Ouimet CC, Miller PE, Hemmings HCJ, Walaass SI, Greengard P. DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein enriched in dopamine-innervated brain regions. III. Immunocytochemical localization. *J Neurosci* 1984;4:111–24.
- [130] Parent A, Hazrati L. Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Res Rev* 1995;20:128–54.
- [131] Parkinson JA, Dalley JW, Cardinal RN, Bamford A, Fehner B, Lachenal G, et al. Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behav Brain Res* 2002;137:149–63.
- [132] Pavese N, Evans AH, Tai YF, Hotton G, Brooks DJ, Lees AJ, et al. Clinical correlates of levodopa-induced dopamine release in Parkinson disease: apertstudy. *Neurology* 2006;67(9):1612–7.
- [133] Pothuizen HHJ, Jongen-Rêlo AL, Feldon J, Yee BK. Double dissociation of the effects of selective nucleus accumbens core and shell lesions on impulsive-choice behaviour and salience learning in rats. *Eur J Neurosci* 2005;22:2605–16.
- [134] Ramnani N, Elliott R, Athwal BS, Passingham RE. Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage* 2004;23(3):777–86.
- [135] Reynolds JN, Wickens JR. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks* 2002;15:507–21. [136] Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. *Nature* 2001;412:67–9.
- [137] Richfield EK, Penney JB, Young AB. Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience* 1989;30:767–77.
- [138] Saint-Cyr JA, Taylor AE, Lang AE. Procedural learning and neostriatal dysfunction in man. *Brain* 1988;111(Pt 4):941–59.
- [139] Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science* 2005;310(5752):1337–40.
- [140] Schall JD. Neural correlates of decision processes: neural and mental chronometry. *Curr Opin Neurobiol* 2003;13:182–6.
- [141] Schönberg T, Daw ND, Joel D, O'Doherty JP. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward based decision making. *J Neurosci* 2007;27(47):12860–7.
- [142] Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol* 1998;80:1.
- [143] Schultz W. Getting formal with dopamine and reward. *Neuron* 2002;36:241–63.
- [144] Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science* 1997;275:1593.
- [145] Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci* 2000;23:473–500.
- [146] Seeman P. Dopamine d2(high) receptors on intact cells. *Synapse* 2008;62(4):314–8. [147] Seymour B, Daw N, Dayan P, Singer T, Dolan R. Differential encoding of losses and gains in the human striatum. *J Neurosci* 2007;27(18):4826–31.
- [148] Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, et al. Temporal difference models describe higher-order learning in humans. *Nature* 2004;429(6992):664–7.
- [149] Shen W, Flajolet M, Greengard P, Surmeier DJ. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 2008;321(5890):848–51.
- [150] Shen W, Tian X, Day M, Ulrich S, Tkatch T, Nathanson NM, et al. Cholinergic modulation of kir2 channels selectively elevates dendritic excitability in striatopallidal neurons. *Nat Neurosci* 2007;10(11):1458–66.
- [151] Shohamy D, Myers CE, Gekhman KD, Sage J, Gluck MA. l-Dopa impairs learning, but spares generalization, in Parkinson's disease. *Neuropsychologia* 2006;44(5):774–84.
- [152] Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA, Poldrack RA. Corticostriatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain* 2004;127:851–9.
- [153] Smith-Roe SL, Kelley AE. Coincident activation of NMDA and dopamine D1 receptors within the nucleus accumbens core is required for appetitive instrumental learning. *J Neurosci* 2000;22:7737–42.
- [154] Stipanovich A, Valjent E, Matamales M, Nishi A, Ahn J-H, Maroteaux M, et al. A phosphatase cascade by which rewarding stimuli control nucleosomal response. *Nature* 2008;453(7197):879–84.
- [155] Suaud-Chagny MF, Dugast C, Chergui K, Msghina M, Gonon F. Uptake of dopamine released by impulse flow in the rat meso limbic and striatal systems in vivo. *J Neurochem* 1995;65(6):2603–11.
- [156] Sugrue LP, Corrado GS, Newsome WT. Neuroscience: matching behavior and the representation of value in the parietal cortex. *Science* 2004;304(5678):1782–6.
- [157] Suri RE, Schultz W. Learning of sequential movements by neural network model with dopamine-like reinforcement signal. *Exp Brain Res* 1998;121:350.
- [158] Surmeier DJ, Ding J, Day M, Wang Z, Shen W. D1 and d2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci* 2007;30(5):228–35.
- [159] Surmeier DJ, Song WJ, Yan Z. Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J Neurosci* 1996;16:6579–91.
- [160] Sutton RS, Barto AG. Reinforcement learning: an introduction. Cambridge, MA: MIT Press; 1998.
- [161] Swanson R, Rogers RD, Sahakian BJ, Summers BA, Polkey CE, Robbins TW. Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia* 2000;38:596–612.
- [162] Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* 2004;7:887–93.
- [163] Tedroff J, Pedersen M, Aquilonius SM, Hartvig P, Jacobsson G, Långström B. Levodopa-induced changes in synaptic dopamine in patients with Parkinson's disease as measured by [<sup>11</sup>C] raclopride displacement and PET. *Neurology* 1996;46(5):1430–6.
- [164] Tepper JM, Bolam JP. Functional diversity and specificity of neostriatal interneurons. *Curr Opin Neurobiol* 2004;14(6):685–92.

- [165] Thorndike EL. *Animal intelligence: experimental studies*. MacMillan Press; 1911.
- [166] Tunbridge EM, Bannerman DM, Sharp T, Harrison PJ. Catechol-omethyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. *J Neurosci* 2004;24(23):5331–5.
- [167] Venton BJ, Zhang H, Garris PA, Phillips PEM, Sulzer D, Wightman RM. Realtime decoding of dopamine concentration changes in the caudate-putamen during tonic and phasic firing. *J Neurochem* 2003;87(5):1284–95.
- [168] Volkow ND, Wang GJ, Fowler JS, Logan J, Gerasimov M, Maynard L, et al. Therapeutic doses of methylphenidate significantly increase extracellular dopamine in the human brain. *J Neurosci* 2001;21:RC121.
- [169] Voorn P, Vanderschuren LJMJ, Groenewegen HJ, Robbins TW, Pennartz CMA. Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci* 2004;27:468–74.
- [170] Walaas SI, Aswad D, Greengard P. A dopamine- and cyclic AMP-regulated phosphoprotein enriched in dopamine-innervated brain. *Nature* 1983;301:69–71.
- [171] Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J, et al. Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* 2006;50(3):443–52.
- [172] Watkins CJCH, Dayan P. Technical note: Q-learning. *Mach Learn* 1992;8:279.
- [173] Wichmann T, Bergman H, DeLong MR. The primate subthalamic nucleus. I. Functional properties in intact animals. *J Neurophysiol* 1994;72: 494–506.
- [174] Wickens JR, Budd CS, Hyland BI, Arbuthnott GW. Striatal contributions to reward and decision making: making sense of regional variations in a reiterated processing matrix. *Ann N Y Acad Sci* 2007;1104:192–212.
- [175] Wiecki TV, Riedinger K, Meyerhofer A, Schmidt WJ, Frank MJ. A neurocomputational account of catalepsy sensitization induced by D2-receptor blockade: Context dependency, extinction and renewal; submitted for publication.
- [176] Wightman RM, Amatore C, Engstrom RC, Hale PD, Kristensen EW, Kuhr WG, et al. Real-time characterization of dopamine overflow and uptake in the rat striatum. *Neuroscience* 1988;25(2):513–23.
- [177] Wilson CJ, Callaway JC. Coupled oscillator model of the dopaminergic neuron of the substantia nigra. *J Neurophysiol* 2000;83:3084.
- [178] Wilson CJ, Chang HT, Kitai ST. Firing patterns and synaptic potentials of identified giant aspiny interneurons in the rat neostriatum. *J Neurosci* 1990;10(2):508–19.
- [179] Wilson CJ, Weyrick A, Terman D, Hallworth NE, Bevan MD. A model of reverse spike frequency adaptation and repetitive firing of subthalamic nucleus neurons. *J Neurophysiol* 1963–1980;91(5).
- [180] Winterer G, Weinberger DR. Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends Neurosci* 2004;27(11):683–90.
- [181] Wolf JA, Moyer JT, Lazarewicz MT, Contreras D, Benoit-Marand M, O'Donnell P, et al. NMDA/AMPA ratio impacts state transitions and entrainment to oscillations in a computational model of the nucleus accumbens medium spiny projection neuron. *J Neurosci* 2005;25(40):9080–95.
- [182] Wörgötter F, Porr B. Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Comput* 2005;17(2):245–319.
- [183] Wu Y, Richard S, Parent A. The organization of the striatal output system: a single-cell juxtacellular labeling study in the rat. *Neurosci Res* 2000;38(1):49–62.
- [184] Yacubian J, Sommer T, Schroeder K, Gläscher J, Kalisch R, Leuenberger B, et al. Gene–gene interaction associated with neural reward sensitivity. *Proc Natl Acad Sci USA* 2007;104(19):8125–30.
- [185] Yasuda A, Sato A, Miyawaki K, Kumano H, Kuboki T. Error-related negativity reflects detection of negative reward prediction error. *Neuroreport* 2004;15:2561–5. [186] Yeung N, Botvinick M M, Cohen J D. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev* 2004;111(4):931–59.
- [187] Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. *Neuron* 2005;46(4):681–92.
- [188] Zackheim J, Abercrombie ED. Thalamic regulation of striatal acetylcholine efflux is both direct and indirect and qualitatively altered in the dopamine-depleted striatum. *Neuroscience* 2005;131(2):423–36. [189] Zador A, Koch C. Linearized models of calcium dynamics: formal equivalence to the cable equation. *J Neurosci* 1994;14:4705–15.
- [190] Deniau JM, Menetrey A, Charpier S. The lamellar organization of the rat substantia nigra pars reticulata: segregated patterns of striatal afferents and relationship to the topography of corticostriatal projections. *Neuroscience* 1996;73:761–81.
- [191] Centonze D, Usiello A, Costa C, Picconi B, Erbs E, Bernardi G, Borrelli E, Calabresi P. Chronic haloperidol promotes corticostriatal long-term potentiation by targeting dopamine D2L receptors. *Journal of Neuroscience* 2004;24:8214–22.