

基底神经节做什么?建模的角度来看

What do the basal ganglia do? A modeling perspective

V. S. Chakravarthy · Denny Joseph · Raju S. Bapi

Received: 3 April 2009 / Accepted: 1 July 2010 / Published online: 20 July 2010

(Song Jian, translate)

摘要: 基底神经节(BG)是由7个大脑深核组成的网络,参与多种重要的大脑功能,包括:动作选择、动作门控、基于奖励的学习、运动准备、计时等。尽管今天有大量的数据可用,研究人员仍然想知道一个单一的大脑深层回路是如何执行如此令人困惑的一系列功能的。BG的计算模型侧重于单个功能,未能给出BG功能的综合描述。在我们对BG功能的理解上的一个重大突破可能是,中脑多巴胺能细胞的活动代表了对机体某种形式的奖励。这种洞察力使机器学习的一个分支——强化学习中的工具,在BG功能研究中的应用成为可能。然而,尽管有这些指明研究方向的点,我们离全面了解这些神秘核的目标还很远。对BG功能的全面了解有可能从根本上改变各种BG相关神经疾病(帕金森病、亨廷顿舞蹈症等)和神经精神疾病(精神分裂症、强迫症等)的治疗和管理。在这篇文章中,我们回顾了现有的关于BG的建模文献,并对这些核的功能提出了一个综合的假设。

关键词: 基底神经节; 多巴胺; 强化学习; Actor-Critic 模型; 动作选择; 探索

1. 引言

基底神经节(Basalganglia, BG)是大脑的深层核,涉及多种重要功能,如(1)基于奖励的学习 (*reward-based learning*)、(2)探索行为 (*exploratory behavior*)、(3)目标导向行为 (*goal-oriented behavior*)、(4)运动准备 (*motor preparation*)、(5)工作记忆 (*working memory*)、(6)计时 (*timing*)、(7)动作门控 (*action gating*)、(8)动作选择 (*action selection*)、(9)疲劳 (*fatigue*)和(10)冷漠 (*apathy*)。著名神经学家 C. Marsden 在 Robert Wartenberg 关于 BG 的演讲中,将这些核描述为神秘的(Marsden, 1982)。尽管自 Marsden 的讲座以来,我们已经做了很多工作,但我们对 BG 的相关论点充满了争议,比如,在治疗帕金森病等各种与 BG 相关的疾病时, BG 的知识缺乏确定性。一直困扰研究人员的问题是:这么多的功能是如何由一个皮层下回路支持的?

为理解 BG 功能打开大门的一个关键观点是, BG 中多巴胺能细胞的活动代表着奖赏信号 (*reward signaling*) (Schultz, 1998)。这种概念关联使得强化学习 (*Reinforcement learning, RL*) 概念 (Sutton 和 Barto, 1998) 在 BG 研究中的应用成为可能。RL 是机器学习的一个分支,它研究的是在没有明确“老师指导”的情况下,“智能体”如何学习对刺激做出最佳反应;“智能体”的学习过程是由环境对“智能体”行为的反应而产生的奖惩信号驱动的。今天,RL 在机器人领域有着出色的应用 (Dorigo 和 Colombetti, 1994)。计算 RL 实际上是受到“条件反射”作用的启发,条件反射作用是心理学的一个主题,在这个主题中,动物学习对感官刺激产生有益的反应 (Sutton 和 Barto, 1990)。导致奖励的反应得到加强,而惩罚的反应得到避免。这种对有回报的反应的搜索通常包括反复试验或探索过程。因此,如果假设 BG 辅助强化学习动态情况,则很容易预测其在探索行为中的作用。“奖励”是一个抽象的概念,在我们对 BG 功能提出一个观点之前,必须仔细定义它。虽然食物或果汁奖励是主要的奖励形式,但更抽象的(次要的)奖励形式可能被认为是成功地接近和达到目标状态的方法。大量的实验文献表明 BG 参与了目标导向行为。运动准备的定义通常很模糊,就像大脑在开始执行动作之前所做的那样。虽然 BG 一直被认为参与了运动的准备阶段,但对于 BG 在运动的准备

阶段中的作用还没有令人信服的理论。我们将运动准备描述为 BG 和运动皮质相互作用的结果。目标导向行为要求目标状态以适当的工作记忆表征状态保持。越来越多的人认识到多巴胺能投射在维持和更新前额皮质(PFC)工作记忆表征中的作用(Cohen 等人, 2002)。RL 的一个重要组成部分是预测未来的回报。预测与刺激-行动对相关的潜在奖励不仅包括预测奖励的存在/不存在, 甚至是奖励的大小, 还包括预测未来奖励发生的时间。因此, 辅助 RL 机制的结构自然适合学习奖励未来事件的时机。确定对给定刺激的最有益的反应包括比较给定背景事件中的许多潜在行为, 并选择最佳行为。BG 被称为脊椎动物大脑对行为选择问题的解决结构。因此, 大多数 BG 功能至少可以在 RL 框架内进行概念上的调整, 尽管在几个领域还需要更详细和更精确的知识。然而, BG 功能疲劳、冷漠等方面似乎并不容易在 RL 的网络中以任何明显的方式被捕获。这些为建模者在未来 BG 研究中提供了一些最困难的挑战。

在这篇文章中, 我们回顾了大量关于 BG 功能计算模型的代表性文献。尽管我们努力提供一个全面的文献回顾和总结在神经解剖学这个复杂的模块的当代思维, 我们真正的目标是收集许多建模工作链式线索, 把它们统一在 RL 的结构体系之下, 呈现出一个统一的, 综合的 BG 功能。

2. 基底神经节解剖

对 BG 的解剖有许多优秀的总结(Gerfen 和 Wilson, 1996; Mink, 1996; Smith 等人, 1998)。幸运的是, 虽然关于细胞水平连接模式的某些问题仍在研究中, 但关于 BG 的解剖学并没有太多争议。更困难的问题是 BG 功能的问题, 这也是本文研究的主题。

2.1 BG 核

BG 由 7 个核组成, 其中部分在正常的自主运动中起着重要作用(Gerfen 和 Wilson, 1996)。然而, 它们与脊髓没有直接的输入或输出连接。这些核从大脑皮层接收它们的主要输入, 并将它们的输出发送到脑干, 然后通过丘脑返回到前额叶、运动前皮质和运动皮质。因此, 基底神经节的运动功能在很大程度上是由额叶皮层的运动区域介导的。

BG 由 7 个广泛连接的皮质下核组成: 尾状核、壳核、苍白球(内部(GPi)和外部(GPe))、底丘脑核(STN)和黑质(致密部(SNc)和网状部(SNr))。尾状体和壳核共同构成纹状体, 纹状体是 BG 的输入核。苍白球可分为外部苍白球(GPe)和内部苍白球(GPi)两部分。SNc 将多巴胺能神经元的轴突投射到 STR 上, STR 的神经元投射到 GPi 和 GPe 上。STN 与 GPe 之间存在兴奋性和抑制性连接。STN 神经元投射到 GPi 上。GPi 核和 SNr 核构成 BG 的输出核, 后者将氨基酸能投射到丘脑。多巴胺能细胞也存在于腹侧被盖区(VTA), 即三叉神经中脑核(中脑核)。伏隔核(NAc)是另一个重要的 BG 核, 与嗅结节一起构成腹侧纹状体的一部分。它被认为在奖赏处理、快乐、攻击性和成瘾方面有作用。NAc 接收来自前额皮质区域和 VTA 的输入。VTA 神经元在 NAcc 上的末端是可卡因和安非他明等成瘾药物的作用区域(Di Chiara 等人, 2004)。背侧纹状体的 SNc 多巴胺能投射被认为参与了自主运动和习惯学习, 而腹侧纹状体的 VTA 投射有助于奖励和动机(Keath 等人, 2007)。多巴胺能投射到 PFC (Cohen 等人, 2002)和脊髓(Heise 和 Kayalioglu, 2008)。

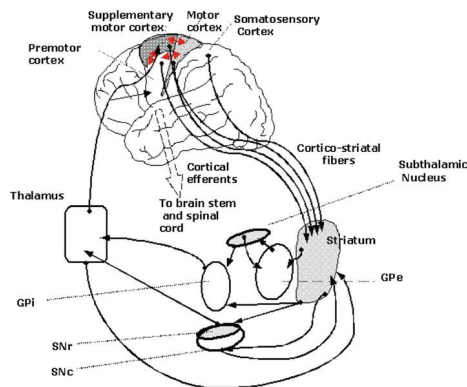


图 1: 基底神经节运动功能的解剖学基础

2.2 BG 的输入和输出

BG 接收来自大脑皮层大部分感觉运动区域的输入，包括初级和次级躯体感觉区、初级运动皮层(M1)和各种运动前区，包括补充区、背侧和腹侧运动前区。BG 运动功能的解剖学基础如图 1 所示。皮质部分负责运动，即补充运动区(SMA)、前运动区(PM)、初级运动区(M1)、体感皮质和顶上小叶，向壳核(BG 的输入核)的运动部分作出密集的、有拓扑组织的投射。这个通路的输出，称为 BG 的运动电路，主要返回到 SMA 和 PM 皮质。这些区域相互联系，并与运动皮层相互联系，这三个区域都有直接的下行投射到脑干运动中心和脊髓。

纹状体(STR)是皮层向 BG 输入信号的主要靶点。在 BG 的“皮质-纹状体-苍白球”通路中，纹状体输出预测形成两个截然不同的平行通道。因为它们对靶核，即丘脑的影响，这些投射被称为直接和间接的途径。直接通路由纹状体输出神经元对 GPi 神经元的抑制投射形成。纹状体神经元的激活抑制 GPi 中的神经元，GPi 抑制丘脑核。从概念上讲，直接通路可以看作是一个常闭的运动门。这个门是由纹状体活动打开，抑制苍白球的输出，允许出现运动。间接通路是通过对 GPe 的抑制投射而形成的，与直接通路中 GPi 神经元的抑制投射作用相反。间接通路的活性倾向于增加 GPi 细胞的活性，因此通过解除丘脑下核的抑制，关闭了“门”。这些通路可能参与调节运动参数(Contreras-Vidal 和 Stelmach, 1996)。

除了皮质，丘脑还向纹状体发送大量的反馈投射。丘脑尾端髓板内核、中心核(CM)和束旁核(Pf)是丘脑纹状体投射的重要来源。CM 主要作用于壳核，而 Pf 主要作用于尾状核和 NAc (Smith, 2008; Arbib 等人, 1998)。

纹状体神经元由两种多巴胺受体组成：D1 和 D2 (Clark 等人, 2005)。当 D1 受体被 SNc 的多巴胺信号激活时，直接通路被激活。同样，D2 受体被激活时，间接通路也被激活。此外，纹状体多巴胺的增加使平衡转向直接途径，从而增加整体运动活动。因此，间接途径是正常的活性途径。在运动开始前，多巴胺释放到纹状体激活了直接通路，平衡就被改变了 (Clark 等人, 2005)。图 2 显示了基底神经节的功能解剖。

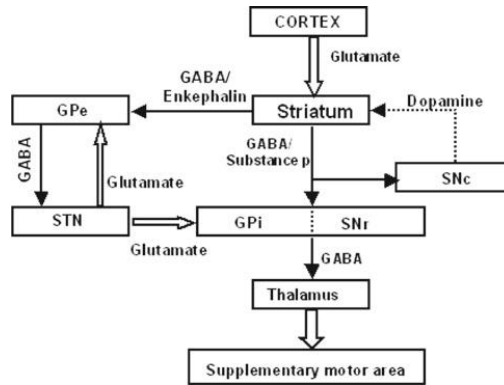


图 2: 基底神经节功能解剖

3.多巴胺、奖励信号和强化学习

多巴胺功能的经典理论是 Wise(1982)提出的“RL”。这一假设表明，多巴胺代表了负责行为控制的大脑系统的天然增强剂的“享乐品质”。这些理论的强大之处在于，成瘾药物主要通过影响多巴胺能系统来发挥作用。大脑刺激奖励(BSR)实验也支持这一观点。在 BSR 实验中，动物选择接受多巴胺相关回路周围的电刺激，动物甚至对这种刺激比食物或水等自然强化物更喜欢(Daw, 2003)。

中脑多巴胺神经元对自然奖励和新奇或意想不到的刺激有反应(Schultz, 1998)。然而，多巴胺能神经元对自然奖励的反应只是对其功能的一个近似描述。Schultz 等人(1997)的实验表明多巴胺反应与奖赏本身无关，而是与对未来奖赏的预测有关。多巴胺神经元被比预期更好的奖励事件激活，不受预期更好的事件影响，被比预期更糟的事件抑制。因此，更准确地说，多巴胺信号似乎代表了预测的未来奖励和实际奖励之间的误差(Schultz 等人, 1997)。

中脑多巴胺信号代表预测未来奖励和实际奖励之间的误差，这一发现标志着我们对 BG 的理解取得了重要进展。有趣的是，还观察到一个复杂的量，比如预测未来奖励和实际奖励之间的误差，与 RL 术语中称为时间差异(TD)误差的类似量非常相似(Sutton 和 Barto, 1998)。这提出了一个实现 RL 机制的可能的场景 BG，这需要 SNc 的多巴胺信号来处理它，并向感觉运动皮层通路发出适当的调节信号，形成 S-R 学习的神经底物(图 1)。

在我们继续进行之前，来描述 RL 的概念及其与 BG 功能的关系，我们必须强调，所有 BG 模型不是 RL 模型。此外，目前对于标准 RL 组件如何精确地映射到 BG 解剖结构上还没有共识。如果上一段所描述的可以被称为“严格的”RL(其中多巴胺信号用 TD 误差识别)，这将有助于提出一个更普遍、更广泛的 RL 概念，其中多巴胺只是携带某种形式的奖励相关信息。这一主题的变异将多巴胺解释为显著性、错误等。这个更普遍、更广泛的概念允许大量特定的数学实现。许多现有的 BG 模型可以适应这个更通用、更广泛的 RL 框架。在此前提下，我们讨论了一些关键的 RL 概念，并考虑了它们与 BG 功能的关系。

3.1 多巴胺和强化学习

BG 被认为参与了记忆的 S-R 刺激反应(S-R)形式的形成，这种记忆形式被称为“操作条件反射”(Squire 和 Zola, 1996)。在操作条件反射中，有益的反应得到加强，而无益的反应减弱。操作条件作用也与 RL 密切相关，RL 研究的是个体如何选择自己的行为，从而最大化回报(或最小化错误)(Sutton 和 Barto, 1998)。

如果我们将 RL 与另一种学习形式(如监督学习)进行比较，就会清楚地认识到将探索作

为实现 RL 的过程的必要性。在监督学习中，对刺激物的期望反应对智能体是显式可用的；将智能体的实际响应与期望的响应进行比较，并使用该误差纠正智能体的 S-R 映射。在 RL 的情况下，没有这样明确的期望反应(老师)。智能体必须探索对刺激可能作出反应的空间；从环境中获得奖励的反应得到选择性的加强。在 RL 框架中，这种探索是由一个称为“Explorer”的模块执行的，但是，它可以根据正在面临的问题实现多种实现。

如果不是为了探索“反应或行动”的空间，使用奖励信息学习 S-R 映射本来是很容易的。但是，在自然环境中，奖励通常是在动物完成一系列动作后才发放的，而这些动作对奖励发放的贡献很难确定。根据过去行为与未来奖励的相关性来评估过去行为的需要被称为“时间信用分配问题”。在 RL 框架中，“智能体”通过开发一个称为 Critic 的内部评估模型来解决这个问题，该模型估计了“智能体”当前状态下的总预期未来报酬，即“值函数” $V(t)$ 。

Critic 的评价和 Explorer 的干扰被用来训练“S-R 映射”，这是一个用 RL 术语称为 Actor 的模块。因此，Actor、Critic 和 Explorer 是 RL 机器人的关键执行组件。

建模 BG 的下一步自然是寻找 RL 的 Critic、Explorer、Actor 和其他关键组件的神经类似物，从而使我们对 BG 的理解更加完整。许多建模工作已经被用于使用 RL 概念描述 BG 功能(Montague 等人, 1996; Schultz 等人, 1997; Houk 等人, 1995; Suri 和 Schultz, 1999; Daw 和 Touretzky, 2000; Frank, 2005)。我们现在描述一些从这一研究中产生的关键思想。

4. 基底神经节计算模型

虽然早期的研究涉及到基底神经节在运动功能中的作用，但后来的工作也认识到它对认知、学习和记忆相关现象的重要性。此外，基底神经节的损伤与复杂的神经精神认知和行为障碍有关，反映了这些核在额叶不同功能中的广泛作用(Stout 和 Johnson, 2005)。关于 BG 的建模文献的一个主要部分涉及 BG 在运动功能和运动障碍中的作用(Chesselet 和 Delfs, 2003)。虽然在解剖学、病理学、电生理学和与 BG 相关的影像学研究方面已经取得了巨大的进展，但对这些核在行为控制方面的作用的全面理解仍然是难以实现的。关于 BG 的计算建模文献确实很多，但大多数模型往往只关注基底神经节的一个或两个功能。需要一个可以将 BG 的所有功能集成到一个框架中的统一功能模型。从 BG 计算模型综述(Prescott 等人, 2002; Houk 等人, 1995)可知，似乎大多数 BG 模型都是排他性的，并只实现特定的功能角色。例如，有一些模型描述了 BG 在“动作门控”中的作用(Contreras-Vidal 和 Stelmach, 1995); 竞争行动之间的“行动选择”(Redgrave 等人, 1999b; , 1995 年); 维持“工作记忆表征”(Houk 等人, 1995); “顺序学习”(Bapi 和 Doya, 1998; Berns 和 Sejnowski, 1998)，最重要的是 RL(Montague 等人, 1995; Schultz, 1998; Barto, 1995)。Buhusi 和 Meck(2005)也回顾了 BG 在“计时”方面的作用。BG 建模者的一个巨大的挑战，即未来要构造由以上提到的模型改编而成的具有独特的、有用的洞察力的一个集成框架，构建一个统一的 BG 功能构想，并显示 BG 核运用奖励信息引起的多巴胺信号和用它来进行多方面的操作。

BG 模型的一个重要类别被称为 Actor-Critic 模型或基于 RL 的模型。上文提到，中脑多巴胺能神经元的活动代表了预测奖励和实际奖励之间的差异，这个量在 RL 术语中被称为 TD 误差(Schultz, 1998)。Houk、Adams 和 Barto(参见 Houk 等人, 1995 年的著作)可能是最早提出类似于 Actor-Critic 学习系统的一些作者。Montague 及其同事(Montague 等人, 1996; Schultz 等人, 1997)提出了一个计算 RL 模型来解释在(Schultz, 1998)中的实验研究中观察到的多巴胺能神经活动模式。

一些作者认为纹状体在 BG 中起着 Critic 的作用(Houk 等人, 1995; Suri 和 Schultz, 1999; Contreras-Vidal 和 Schultz, 1999)。纹状体接受大量来自感觉运动皮层区域的输入。特别是, 每个运动皮质区域都投射到一个由尾状核和壳核构成的结构上截然不同的区域。因为纹状体也收到从 SNc 的多巴胺能预测, 它实现的计算是十分恰当的价值函数: 正如 TD 误差(δ)是用于更新价值函数的权重, 纹状体多巴胺预测调节“皮质-纹状体”连接的可塑性。

虽然 RL 结构成分在 BG 解剖结构上的映射研究已经取得了一定的成果, 但在确定一个重要成分的皮层下底物上仍有一定的困难。虽然在 RL 框架中, “运用”和“探索”是同等重要的过程, 但基于 RL 的 BG 观似乎主要集中在奖赏信号——它的化学信使, 它的解剖位置, 它在学习中的结果等等, 而只对探索进行了总结。Daw 等人(2006)的研究表明, 前额叶和顶叶位置与探索相对应, 但没有发现相应的皮质下位置。

因此, 尽管有大量的文献应用于 BG 功能的 RL 概念, 但没有共识, 如果, RL 机制可以被映射——组件到组件——BG 解剖, 并且映射关系如何解释 BG 功能的全部功能(Joel 等人, 2002; Worgotter 和 Porr, 2005)。

5.探索性行为

动物必须在不确定、未知的环境中觅食。有两种神经系统被认为可以调节探索性行为, 一种基于海马, 另一种基于 BG (Packard 和 Knowlton, 2002)。海马体在大鼠空间导航能力中的作用已被广泛研究。几项研究报告了大鼠海马中存在的位置细胞, 当动物位于熟悉的空间环境中的某个位置时, 这些细胞会被激活(Packard 等人, 1989)。在一项实验中, 老鼠被训练从一个固定的起始位置走到 T 形迷宫中有诱饵的端口处(例如指向西方的端口)(Packard 和 McGaugh, 1996)。经过几天的训练, 迷宫旋转 180°。目的是观察动物是否表现出位置学习的迹象, 并朝着相同的绝对方向(西方), 或表现出刺激反应(S-R)学习, 使相同的身体转向并朝着错误的绝对方向(东方)。换句话说, 在位置学习中, 动物在视觉空间坐标系中表示目标; 在 S-R 学习中的编码是在就身体位置而言的运动坐标。经过 8 天的训练, 大多数动物表现出位置学习, 但经过 16 天的训练, 大多数动物表现出 S-R 形式的学习。第 16 天背外侧纹状体功能受损, 使动物表现出与 S-R 学习相背的位置学习。因此, 即使保留了位置学习, 动物也会在长时间的训练中从位置学习策略切换到 S-R 策略。这些和其他实验表明, BG 参与了 S-R 形式的导航。

研究了多巴胺能活动对蜜蜂觅食行为的影响(Scheiner 等人, 2002)。Montague 等人(1995)提出了一种基于预测的 Hebb 学习和多巴胺神经元的蜜蜂觅食模型。在这个模型中, 模拟的蜜蜂遵循一个简单的觅食策略: 如果多巴胺信号足够高(高于基线), 有机体就会保持航向, 否则(多巴胺细胞活动低于基线), 有机体就会随机摇摆, 这是一种被称为“变向无定位运动”的运动形式。然而, 是什么神经机制驱动有机体的随机转向? 这个问题没有答案(Montague 等人, 1995)。虽然在 RL 框架中, “运用”和“探索”同样重要, 但是关于 BG 在 RL 中的作用的文献, 似乎主要集中在奖赏信号——它的化学信使, 它的解剖位置, 它在学习中的结果等等, 而只是对探索进行了总结。人们常说, SNc 和/或 VTA 的多巴胺能细胞的活动表明奖赏(Montague 等人, 1995)。但是 BG 的哪一部分产生探索所需的随机信号呢?

Daw 等人(2006)也注意到, 虽然对“运用”的解剖位置有相当清楚的认识, 但“探索”的神经底物相对不清楚。Daw 等人(2006)在一项基于成像的研究中, 研究对象玩四臂强盗游戏, 这是一种“探索”与“运用”很好地融合在一起的 RL 游戏, 他们观察到, 当受试者进行探索活动时, 额皮质和顶叶沟优先活跃。刚才提到的两个皮层区域可能是支持探索性决策的皮层区域。但是, 由于其他与 BG 功能相关的 RL 成分已经在 SNc 的 BG 电路——TD 误

差、纹状体的 Critic 等中被识别出来，因此没有相应的 BG 核(或细胞核)被提及作为探索行为的皮质下基础，这是令人惊讶的。

我们之前曾提出，BG 电路中的 STN-GPe 环路被正确地放置在支持探索行为的位置(Sridharan 等人，2006)。在 RL 模型中，“探索”总是由随机分量建模。从生物学的角度来说，只有能够产生复杂的尖峰活动的神经系统才有资格成为“Explorer”。STN-GPe 层形成了兴奋-抑制神经元层对，以其在完整 BG 中的复杂的尖峰活动而闻名，在多巴胺缺乏或帕金森条件下观察到这种复杂活动的丧失(Nini 等人，1995；Bergman 等人，1994；Brown 等人，2001)。

5.1 STN-GPe 环路层作为随机性的来源

Plenz 和 Kitai(1999)在对基底神经节器官型组织培养的研究基础上提出，STN 和 GPe 结构中都可能出现相关的活性，并且相关的活性是由这两个结构之间的相互作用而不是由外部源驱动的。Terman 等人(2002)提出 GPe 内连接减弱和纹状体输入增强的组合为同步 STN-GPe 振荡和相关节律性 STN 输出奠定了基础。关于 GPe-STN 和 STN-GPe 投射的空间组织，以及这两个核是否以“互反”或“失域”的方式相互投射，目前缺乏精确的解剖学信息。Hurtado 等人(1999)记录了清醒帕金森患者的神经元活动，并观察了 GPi 内的一些成对记录位点，它们显示出短暂的同步。Bergman 等(1998)报道了 MPTP 治疗猴子大脑皮层活动的动态同步。最近的实验研究显示，放电具有显著的低频周期性(4-30Hz)，并显著增加了 GPe 和 STN 神经元之间的相关性(Bergman 等人，1994；Nini 等人，1995；Magnin 等人，2000；Raz 等人，2000；Brown 等人，2001)。对 STN 和 GPe 活性的实验研究表明，在多巴胺耗尽的情况下(类似于帕金森条件)，这些核的活性虽然没有显著降低放电率，但神经元之间的相关性却显著增加(Bergman 等人，1994；Brown 等人，2001)。在建模研究中，STN-GPe 回路神经元的相关活动在功能上与帕金森震颤频率(Terman 等人，2002)和帕金森笔迹扭曲(Gangadhar 等人，2008)有关。

STN-GPe 环路在正常 BG 中的复杂活性及其在帕金森条件下的损失被认为具有深刻的功能意义，并被解释为被解释为 RL 所要求的随机勘探的来源(Sridharan 等人，2006)。Sridharan 等人(2006)的模型描述了一个模拟的莫里斯水池实验，其中一只虚拟老鼠在可见地标的帮助下探索一个隐藏的平台。当平台(即与之相关的地标集)不在“智能体”的视野中时，模型的 STN-GPe 部分表现出不相关的活动，反映了探索性行为。当平台(即与之相关的一组地标)不在智能体视野之内时，模型的 STN-GPe 部分表现出不相关的活动，反映了探索性行为。

Doya(2002)从 RL 的角度讨论了各种神经调节因子(多巴胺、血清素、去甲肾上腺素和乙酰胆碱)在脑功能中的可能作用，并假设：(i)多巴胺是预测奖励和强化行为的全局学习信号，(ii)血清素控制着对奖励的短期和长期预测之间的平衡，(iii)去甲肾上腺素控制着广泛探索和集中执行之间的平衡，(iv)乙酰胆碱控制着记忆存储和更新之间的平衡。特别地，为了支持去甲肾上腺素在探索中的假设作用，Doya 指出蓝斑(LC)中的去甲肾上腺素能神经元在紧急情况下被激活(Doya，2002)。此外，我们知道，在刺激呈现时，LC 神经元的相态反应与反应的高精度相关(Aston-Jones 等人，1994)。其他作者也提出去甲肾上腺素与“探索-运用”问题之间存在联系(Usher 等人，1999)。也有证据表明去甲肾上腺素与苍白球的活动水平有关(Russell 等人，1992)。这种感知与 STN-GPe 环路的观点非常一致，STN-GPe 环路在探索行为中扮演着重要角色(Sridharan 等人，2006)。

一个主要的理论问题涉及探索行为，以及探索获取新信息和运用现有知识以获得回报之

间的权衡。研究动物如何选择在“探索”和“运用”之间保持平衡是一个丰富的研究领域(Montague 和 Berns, 2002)。已有研究表明,多巴胺系统的活动与奖励预测误差无关,但可能与确保适当的探索有关(Suri 和 Schultz, 1999; Kakade 和 Dayan, 2002)。多巴胺与强化学习的时间差误差之间的密切关系,揭示了动物与人工决策系统之间新的、强有力的联系。

6.运动准备

Kornhuber 和 Deecke(1990)在他们关于自主运动的经典脑电图研究中发现,皮质电位在运动开始前就存在缓慢的负迁移。这种电位被称为“准备电位”(Bereitschafts potential, BP),被认为是运动皮层区域运动前准备的标志。仔细分析 BP 的电流偶极子源,可以发现 SMA 是一个关键因素(Lang 等人, 1991)。然而,与运动方向相对应的准备活动在许多其他脑区也被发现,包括 M1 (Georgopoulos 和 Grillner, 1989)、运动前皮质(Kubota 和 Hamada, 1978)、PFC (Kubota 和 Funahashi, 1982)、顶叶皮质(Crammond 和 Kalaska, 1989)和 BG (Alexander, 1987)。Churchland 等人(2006)在灵长类动物实验中提出了一个有趣的运动准备功能定义。这个小组假设,准备是一个过程,通过这个过程,运动皮层神经元的活动,随机的和可变的在准备的早期阶段,是逐步推进到一个有限的区域的状态空间,这是特定于给定的运动。来自灵长类动物运动前皮层神经元的活动似乎证实了他们的假设 Churchland 等人, 2006)。然而,与其他运动皮质区相比, SMA 在运动准备中起着主导作用。

6.1 SMA 和运动准备

SMA 作为运动准备信号的主要来源,似乎与其他几个相关运动领域竞争。灵长类动物的单细胞记录显示,与 M1 神经元相比,神经神经元的准备相关的变化更为明显(Tanji, 1994)。问题是 SMA 和 M1 哪个先出现?如果能证明 SMA 神经元的准备活动先于 M1 神经元的类似活动,就可以解决这个问题。已有研究表明,具有预备活动的 SMA 神经元可以投射到 M1 (Tanji, 1994)。相反,研究还发现,表现出准备活动的 M1 神经元接受来自 SMA 的输入,而不是来自丘脑或顶叶皮质(Aizawa 和 Tanji, 1994)。这些研究有力地说明 SMA 在运动准备中的作用。然而, SMA 可能并不完全负责运动准备工作。它的准备作用可能包括皮层下结构之间的相互作用,如基底神经节,这往往涉及运动定时功能: SMA 和基底神经节之间的相互作用被认为在学习的运动序列中起着至关重要的作用(Cunnington 等人, 1995; Bapi 等人, 2006)。有人认为基底神经节的阶段性活动可以作为 SMA 的复位信号,清除一次运动后的痕迹,为连续运动做好准备(Georgiou 等人, 1993)。

从上面的叙述中,我们了解到运动准备是受 BG 和 SMA 双向作用的影响。类似的观点也出现在“笔迹生成”的计算神经运动模型中(Gangadhar 等人, 2007)。该模型的灵感来自于 Hollerbach(1981)的“手写的振子理论”,根据该理论,“手写”可以看作是两个振荡过程的叠加运动——一个是垂直的,另一个是水平的。笔迹的形状可以通过调制各种参数的振荡,如相位,孤立时刻的振幅。Gangadhar 等人(2007)描述了一个神经振荡器网络,该网络可以根据网络中神经元的振荡活动,以傅里叶式的方式解析笔画速度,从而学习任意的笔画。准确的行程执行不仅取决于网络的“权重”,而且关键取决于振荡器的初始相位。模型中的“基底神经节”模块向振荡神经元发送预备信号,并通过将它们从随机的初始阶段驱动到“标准阶段”来初始化它们的相位。这种运动准备的观点与 Churchland 等人(2006)的观点强烈呼应, Churchland 等人观察到背侧运动前神经元(PMd)的活动模式是从随机初始运动开始运动的状态到“最优子空间”,该子空间指的是一组受限制的状态(图 3)。

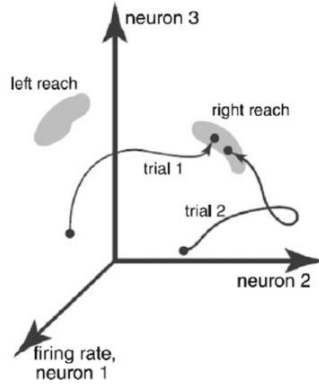


图 3: 显示了运动前活动如何随着时间的推移而稳定下来。活动在高维空间中表示, 每个神经元的放电率提供一个轴。活动最初是可变的, 接近基线(每次试验一个点)。在目标开始后, 活动将(曲线路径)确定为一个子空间状态, 该子空间适合所需的运动(阴影区域)(摘自 Churchland 等人, 2006 年的研究)。

作为(Gangadhar 等人, 2008)中提出的运动准备研究的延伸, 我们的小组最近研究了帕金森条件下的准备(未发表的结果)。在模型中, 我们发现从运动开始到运动结束的驱动力与阶段的多巴胺爆发有关, 而阶段的多巴胺爆发已知发生在运动开始或运动之前。由于多巴胺缺乏症是帕金森病(PD)的一个特征, 我们在模型中通过减弱准备神经振荡器的驱动信号来模拟 PD 的情况。驱动信号(缩放因子称为驱动因子, γ_{df})被削弱, 准备花费的时间更长; γ_{df} 低于临界值时, 准备永远不会终止, 动作永远不会启动。这种不能使运动皮层为运动做好准备的情况类似于 PD 患者在自主控制时所面临的困难。特别是对 PD 的冻结现象(Giladi 等人, 1992)提供了一个有趣的解释, 我们认为这可能是一个失败的运动准备的例子。

7.BG, PFC 和工作记忆

前额-基底神经节神经回路参与一系列神经精神疾病, 如精神分裂症、注意缺陷多动障碍等(Cohen 等人, 2002), 由此可见前额-基底神经节神经回路的 BG、PFC 和工作记忆的临床重要性。前额叶区域在认知、目标导向行为、工作记忆和注意力功能方面具有公认的作用(Gabrieli 等人, 1995; Knight 等人, 1995)。目标导向行为要求当个体从事有助于实现目标的行为时, 与目标相关的信息被稳定地保存在大脑中。在这个过程中, PFC 被认为是一个关键的执行者, 因为这个皮层区域在涉及目标寻找的任务中被优先激活。此外, 对 PFC 的损伤会导致目标寻找的障碍(Cohen 等人, 2002)。

虽然 PFC 满足了对目标相关信息维护机制的需求, 但仍有两个问题没有得到解答: (1) 目标相关信息在 PFC 中应该保存多久? (2) PFC 中的信息更新是通过什么机制产生的? 已有研究表明, 腹侧被盖区(VTA)对 PFC 的多巴胺能投射在 PFC 中执行这种信息门控(Cohen 等人, 2002)。根据这一建议, 在没有阶段性多巴胺释放的情况下, PFC 的活性对干扰传入的输入保持强劲。阶段性多巴胺爆发打开了 PFC 的大门, 允许更新目标状态(Montague 等人, 2004)。O Reilly 和 Frank(2006)提出了一个包含 PFC 和 BG 的工作记忆模型, 该模型在 RL 框架内描述了多巴胺能投射到 PFC 的工作记忆门控。多巴胺信号中的奖赏相关信息被用来决定哪些信息随着时间的推移而保持, 哪些信息可以在 PFC 中得到更新。

多巴胺能传递到 PFC 的干扰与精神分裂症的发病机制有关(Montague 等人, 2004)。近半个世纪以来, 精神分裂症一直被认为是一种高多巴胺能状态。阻断多巴胺能传递的药物改善了精神分裂症的妄想症状, 多巴胺激动剂重现了一些精神分裂症症状, 这一事实支持了这一观点。然而, 后来的研究表明多巴胺和精神分裂症之间的联系比这更复杂。最近的一项评

估表明，精神分裂症与 PFC 中的多巴胺减少和纹状体中多巴胺水平升高有关。通过使用计算模型模拟 PFC 中多巴胺能增益的降低，有可能重现精神分裂症的症状(Servan-Schreiber 等人，1990)。

8. 计时

协调多个处理流的相对时序对运动性能和感官感知都至关重要。生物系统的时间处理发生在一定的时间范围内，大致可分为三类：(1)昼夜时间，对应于时间的顺序，由大脑结构如坐骨上核处理，(2)间隔时间，这对应于时间的秒到几分钟，主要由“皮质-纹状体”和协调互动，和(3)毫秒时间，这显然与毫秒时间，由小脑控制(Buhusi 和 Meck, 2005)。除了时间方面，皮质纹状体和皮质小脑系统在运动技能学习方面的贡献已经被区分出来(Doyon 等人，2003)。

最近的数据表明，时间可能在大脑中以一种分布式的方式来表示，说明时间是一个检测不同神经群的巧合激活的问题。间隔时间取决于完整的纹状体(Malapani 等人，1998)。处理时间在秒到分钟范围内的能力受损的患者发现，影响多巴胺能通路的疾病，如帕金森氏病(PD)，亨廷顿舞蹈症和精神分裂症。PD 患者在服用 L-dopa 时表现出无向性，但在非药物治疗时则没有(Malapani 等人，1998)。他们无法独立计算两个时间段的时间；这两种标准持续时间的重复标准趋向于相互迁移。这种迁移被消除，并在刺激丘脑下核后恢复准确的时间。它们也显示出运动动作的时间不佳。小脑损伤后无向性的保留支持了纹状体和小脑参与不同方面的时间和时间知觉的观点。如前所述，小脑负责毫秒计时，BG 负责间隔计时。

基底神经节在“间隔计时”中的作用似乎来自于“丘脑皮质-纹状体”环路的动态情况。在纹状体搏动频率(SBF)模型中(Matell 和 Meck, 2004)，一个强调 BG 定时功能的模型，皮质振荡器被假定在运动开始前增加同步，并在整个表演过程中保持节奏。根据 SBF 模型(Buhusi 和 Meck, 2005)，实验开始时的多巴胺能脉冲可以触发皮质振荡器的同步。纹状体神经元被调节以响应特定的皮质振荡模式(Matell 和 Meck, 2004)。皮质纹状体传递的经验依赖性变化被认为使纹状体神经元更有可能通过皮质纹状体 LTP 和 LTD 在奖赏传递时检测皮质振荡器的特定激活模式(Buhusi 和 Meck, 2005)。

如果我们接受 BG 核是大脑 RL 机制底物的观点，那么 BG 参与与时间相关的功能是很自然的。对未来报酬的预测是 RL 的关键子问题之一，涉及到回答两个问题：“多少？”和“多久？”BG 在解决 RL 下的时间信用分配问题的过程中，很自然地处理了时间问题，学习了奖励事件、刺激和行为之间的时间关系。

9. BG 与动作选择

BG 解决竞争行动之间的联系的建议是基于一个日益增长的共识，即这些结构的一个关键功能是在竞争进入最终共同运动路径的传感器运动系统之间进行仲裁。从这个想法发展而来的一个计算假设依赖于这样一个前提：传入纹状体的信号编码了进入运动系统的请求的显著性(Redgrave 等人，1999b)。BG 中嵌入的多重选择机制可以解决竞争对手之间的冲突，并在赢家之间提供干净、快速的切换。首先，纹状体神经元的上/下状态可以作为第一遍过滤器来排除弱支持请求。其次，纹状体内的局部抑制可以选择性地增强最显著通道的活性。第三，纹状体的聚焦抑制与 STN 的扩散(发散)激发相结合，可以作为一个前馈的、离中心的/环绕的网络，作为一个整体贯穿 BG(参见 Mink, 1996)。最后，输出核内的局部相互抑制可以提高最终的选择。

Prescott 等人(2002)以动作选择假设为组织原则,提出了一种新的基于选择和控制电路的功能分组方法。该模型的强度已经通过将其嵌入一个移动机器人的控制体系结构中进行了测试,该机器人具有少量的类动物行为(参见 Prescott 等人, 2002)。这项工作证实了模拟的 BG 可以在需要适当和及时的行为切换的真实环境中提供有效的行动选择。该机器人模型还提供了一个洞察异常多巴胺调制行为选择的紧急后果。例如,让人想起帕金森病的一些运动症状的是,多巴胺的减少被发现会导致无法选择适当的行为或曾经选择完成的行为。

最近的研究已经为纹状体内的局部抑制提供了证据,这种抑制要么是通过局部间神经元介导的,要么是通过输出细胞之间的相互抑制网络介导的(Oorschot 等人, 2002)。Wickens(1997)运用网络模型研究了纹状体神经元这种局部邻域的动态。在拓扑结构和大小不同的假设下,观察到相互抑制通常会导致网络动态竞争,也就是说,最活跃的神经元往往会抑制其较不活跃的邻居的活动。这项研究还解释了模拟多巴胺输入的影响,表明多巴胺低的环境下,网络的动态变化与竞争“同激活”(活动均匀分布于神经元的局部种群中),这种模式可以为多巴胺缺乏的帕金森患者的肌肉僵硬提供一个模型。

Redgrave 等人(1999a)认为,多巴胺能细胞对意外出现的行为显著性刺激的短时间、短潜伏反应,是过程的一个重要组成部分,这些过程负责真实定位注意力和行为资源,使其有利于意外的显著性事件。据推测,这样的爆发产生了一种相对非分化的多巴胺能输入波到纹状体大面积区域。作者(Redgrave 等人, 1999a)指出,脊椎动物 BG 已经进化成一种中枢选择装置,专门用来解决多个子系统之间的冲突,这些子系统争夺有限的运动或认知资源。在这个框架下,选择操作有选择地解除对获胜竞争者“感觉-运动”连接的抑制,同时保持或增加对失败竞争者的抑制控制。多巴胺爆发的另一个隐含功能是,在奖励刺激被定位和消费之前,首先有必要中断正在进行的行为,并切换注意力和行为资源来处理奖励事件。人们发现,意想不到的新奇或强烈的刺激总是会引发短期的多巴胺爆发,从而证实了这种作用。

10. BG 和疲劳

疲劳是指一个人感到无力产生足够的力量来完成重要工作时的一种疲惫状态。然而,它必须与软弱相区别,在软弱中,个人实际上无法产生与手头任务相称的足够的肌肉力量。后一种形式的疲劳可称为肌肉疲劳,甚至是“外周疲劳”;它指的是肌肉无法产生力量。它可能与中枢疲劳(central fatigue)不同,中枢疲劳指的是疲惫的主观感觉。这种中枢疲劳在慢性疲劳综合症(CFS)中表现突出,其特征包括认知问题、睡眠障碍、过敏和敏感、头痛、低烧等。通常定义为持续性和复发的原因不明的疲劳持续至少 6 个月(Fukuda 等人, 1994)。让 CFS 患者感到困惑的是,他们并不一定是真的很虚弱,他们只是经常感到虚弱或疲劳。Chaudhuri 和 Behan (2000a)提到了与中枢疲劳相关的神经系统疾病(CFS 除外)。这些包括脊髓灰质炎后综合征、病毒后综合征、格林巴利综合征、多发性硬化症、帕金森病等。

CFS 症状学的一些重要定义特征是(Chaudhuri 和 Behan, 2000b)

- (1) CFS 患者表现出与多发性硬化症患者相似的延迟运动传导;
- (2) CFS 患者虽然拥有完整的肌肉系统,但仍不能完全恢复肌肉;
- (3) 他们对努力有一种夸大的感觉;
- (4) 与对照组相比,他们在运动过程中也表现出运动诱发电位的降低。

经磁刺激(TMS)对慢性疲劳综合症患者的研究表明,患者不仅没有最大程度地激活肌肉,而且看到的肌肉激活减少是由于中枢激活减少(Sacco 等人, 1999)。除了运动过程中运动皮层活动减少,CFS 患者的运动前或准备活动也减少(Starr 等人, 2000)。这些研究表明,

导致慢性疲劳综合症的一个关键因素是对运动皮层的驱动减少。

Chaudhuri 和 Behan (2000c)假设中心疲劳的起源在于 BG 电路的动态变化。基于神经病理学数据,这些作者认为中枢疲劳的产生是由于基底神经节内边缘输入和运动功能的整合失败,影响了纹状体-丘脑-额叶皮质系统(Chaudhuri 和 Behan, 2000c)。在其他地方,Chaudhuri 和 Behan (2000a)将自主工作描述为一个共同依赖于实际努力的变量,而实际努力又依赖于动机输入;以及感知运动,这取决于来自身体的感觉运动反馈。中枢性疲劳似乎是感知力受损的一种表现。

认为疲劳是由于边缘和运动输入在 BG 中的集成失败而引起的,这与本文中发展的 BG 功能的观点非常一致。到目前为止,在这篇综述中,我们主要关注 BG 功能的感知运动方面。然而,通过 BG 核的“皮质-皮质-苍白球-丘脑”通路被组织成平行的回路,并根据其功能进行分离(Alexander 等人, 1990; Hikosaka 等人, 1999)。根据 Alexander 等人(1990)的著名分类,这些回路包括:(1)骨骼运动回路;(2)动眼神经回路;(3)联想回路和(4)边缘回路。由于这些通道是按结构组织的,因此在整个 BG 电路中保持了回路的隔离。头两个循环,骨骼运动循环和眼动循环,顾名思义,分别控制肌肉骨骼功能和眼球运动。联想回路涉及认知、记忆和注意力。边缘回路与皮层和皮层下边缘区相互作用,如前丘状皮质、杏仁核等,并参与动机活动、体内平衡、内脏监控和控制。到目前为止,感知运动功能和后处理功能与前两个动作(骨骼运动和肌肉运动)最为相关。但是,对于理解 BG 和疲劳之间的关系,理解联想环和边缘环的参与似乎是关键。

为了理解 BG 在边缘区域的功能,我们假设了 BG 对奖赏处理的扩展图。到目前为止,BG 中奖励处理的基本情况如下:BG 结合代表自然奖励(食物、果汁等)的中性感官刺激信号,预测未来的奖励,并运用这些预测选择奖励行为。但是,仅仅把一个行动看作是有回报的还是没有回报的,是对行动及其含义的狭隘看法。这里似乎漏掉了一个重要的组成部分。为了完成一个动作,动物必须消耗能量;或者它可能不得不承担风险。因此,行动的选择不仅基于其内在的回报;只有当奖励超过了动物的代谢成本和/或选择所带来的风险时,这个行为才会被选择和执行。确切地说,由于动物在采取行动之前并不知道实际的奖励或实际的代谢成本,动物可能通过比较预期的奖励和/或预期的代谢成本和/或风险来做出判断。我们知道 BG 电机回路提供奖励预测;我们假设边缘回路提供了关于预期代谢成本的信息。

11. BG 和冷漠

冷漠通常被定义为缺乏感觉、情感、兴趣或关心(Levy 和 Dubois, 2005)。就可观察到的特征而言,它的显著特征是自愿或目标导向行为的急剧减少。因此,冷漠是神经回路改变的自然结果,最重要的是 PFC 和 BG,它们辅助自主行为。PFC 病变(Eslinger 和 Damasio, 1985; Fuster, 1997; Stuss 等人, 2000)以及 BG 的具体结构(Bhatia 和 Marsden, 1994; Engelborghs 等人, 2000; Ghika-Schmid 和 Bogousslavsky, 2000)经常导致冷漠。然而,冷漠并不是一个简单的单一过程:认知、情感等多种因素可以构成冷漠的基础。因此,Stuss 等人(2000)将群体冷漠分为三类:(1)认知冷漠、(2)情感冷漠和(3)行为冷漠。这些分类与 BG-PFC 电路中某些精确的解剖底物有进一步的联系。

冷漠的情感成分与眼眶和内侧 PFC 以及与之相连的腹侧纹状体有关。灵长类动物的电生理学研究和对人类的成像研究表明,眼眶和内侧 PFC 对于解释奖励的相对价值或背景事件价值并将其作为行为的推动力非常重要(Thut 等人, 1997; Tremblay 和 Schultz, 1999)。因此,眼眶和内侧 PFC 的损伤可能导致奖赏检测的缺陷,从而导致自主行为的减少,这可能被解释为冷漠。

同样，认知成分与外侧 PFC 和尾状核背侧相关(Levy 和 Dubois, 2005)。冷漠的这一方面，也可以称为认知惰性，与目标导向行为受损有关。目标导向行为包括许多组件功能，如计划、工作记忆、规则发现和设置转移；这些功能优先受到外侧 PFC 病变的损害(Goldman-Rakic, 1987; Fuster, 1997)。此外，尾状核背侧病变会产生行为缺陷，例如延迟交替任务，这与侧 PFC 病变类似(Dean 和 Davis, 1958; Iversen, 1979)。

因此，冷漠的情感成分似乎源于无法充分感知和回应奖励的存在。认知方面来自于目标导向行为的各种组成功能的缺陷。冷漠的第三个方面，行为方面，也被称为自激活缺陷(Levy 和 Dubois, 2005)，指的是自发激活的丧失。这种情况下的患者可能更喜欢长时间静静地坐着，不与任何人交谈，也不采取任何主动。然而，他们可能会对外界刺激做出适当的反应，并产生相应的答案和行动。

综上所述，冷漠的产生可能是由于额基底神经节回路的紊乱。此外，由于它与奖赏处理的缺陷有关，多巴胺能信号的参与可能是预期的。因此，不出所料，在 PD 患者中常常观察到冷漠，而且，在 PD 患者的开、关状态之间，甚至可以观察到明显的冷漠差异(Czernecki 等人, 2002)。因此，似乎冷漠和疲劳可能与动机和情绪信息从边缘通路到腹侧纹状体和 BG 中的纹状体模块的融合中断有关。另一方面，母细胞模块似乎正在接收来自大脑皮层的感觉运动信息。BG 被认为是将情感信息结合成目标导向行为的最佳位置(Houk 等人, 1995)。

12. 总结与结论

从现有的 BG 功能方面的文献中，我们试图对 BG 功能进行概念上的综合。从这个综合来看，BG 功能的本质可以表述为：BG 将机体的感觉运动环境与奖励信息相结合，并将这种结合后的信息传递到运动皮层和前额皮质区域进行决策。BG 的大部分功能似乎都起源于这一基本操作。虽然目前还没有针对 BG 功能所有方面的计算模型，但是初步的概念综合可以为正式的建模工作打下基础。

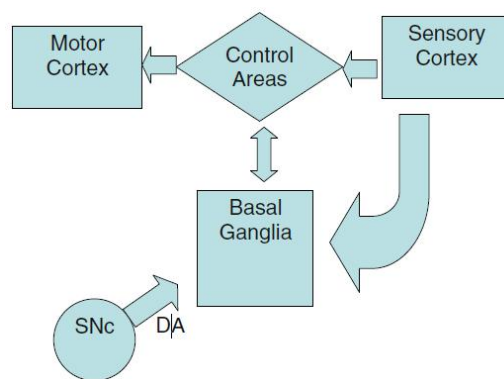


图 4: BG 门控函数的简化示意图。显示了 BG 相对于感觉运动皮层区域的位置。控制区包括前额区和辅助运动区。虽然 SNc 通常被认为是 BG 的一部分，但它显示在 BG 之外以显示 DA 信号。并不是所有的连接都显示为清晰的表示。

BG 皮质下核相对于皮质的功能位置如图 4 所示。BG 提供综合奖励和感觉运动信号的结果来控制或决策皮层区域(前额叶区域和补充运动区域)。BG 对这些领域的输入作为一个“门”，通常允许有回报的行为或不允许无回报的行为。此外，当多个动作同时竞争执行时，BG 系统作为仲裁器，运用多巴胺能奖励信息作为显著性度量来解决竞争。

“门控”这个术语可以用来粗略地描述 BG 功能。更准确的描述应该是 BG 影响或调节控制区域(PFC 区域和 SMA)。这种影响或调制的另一个例子是运动准备，BG 监控 SMA 的

活动,并在对下一个动作排序之前将准备驱动发送给 SMA。图 4 所示的控制区(此处为 SMA)与 BG 的双向相互作用,为这种准备过程提供了解剖学基础。

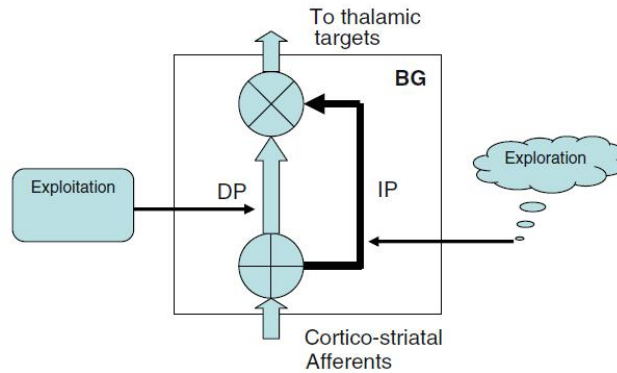


图 5: BG 的直接和间接途径,假设分别为辅助“运用”和“探索”功能

从事强化学习的“智能体”需要在开发和探索之间进行切换。结合由 STN 和 GPe 组成的间接通路(IP)具有产生驱动探索所需随机信号的机制,我们现在有了 BG 功能的图像,其中直接通路和间接通路分别辅助“探索”和“运用”(图 5)。因此, BG 电路可能被认为“进化引擎”, DP 的作用类似于自然选择, IP 对应于随机突变的过程。关键是,虽然真正的进化发生在数百万年之后,但 BG 驱动的进化发生在行为的时间尺度上。因此, Yin 和 Knowlton(2006)认为 BG 可以被认为是一个遵循经典选择原则运作的生物系统,具有多样性的产生机制和选择和差异放大,他们呼应了 BG 功能的类似观点。我们进一步提出, BG 中多样性的产生是在 IP 的复杂动态过程中产生的。已有几项研究表明,间接途径中复杂活动的丧失可能是帕金森震颤的原因。因此,不像之前的一些理论建议,认为间接途径发挥了一些较弱的作用,如提供背景事件信息,或关注和排序的 GPi 输出(Hikosaka 等人, 2000),切换(Isoda 和 Hikosaka, 2008), IP 角色扮演相当重要的角色——从本综述中开发的 BG 功能的角度来看, DP 的作用是完全互补的,而不仅仅是相反的(Frank 2005)(见附录)。

虽然我们努力提出一个连贯的 BG 功能的概念综合,将其众多的功能与其通过强化学习的本质作用联系起来,但我们承认,将这一概念图转化为一个详细而全面的计算模型的时机尚未成熟。将现在在概念级别上明显的一致性转换为一个全面的、包含所有内容的模型需要大量的建模工作。能够用完美的计算术语描述这个神秘的(套用 Marsden 对 BG 的描述)系统,确实是建模者的乐趣。

BG 功能的非 RL 观点

尽管图 4 和图 5 所示的 BG 的功能结构与大脑其他关键区域的关系本身就很有趣,也很有说服力,但必须指出的是,当代文献对 BG 功能的观点存在着根本的不同。随着新的实验数据的获得,有人试图添加新的连接,甚至模块到现有的 BG 架构。这样的添加挑战了 BG 功能的早期概念。他们似乎要求彻底修改这个问题,如果可能的话,将新数据纳入现有的框架。

将 BG 模型分为严格 RL 模型和广义 RL 模型两大类是有帮助的。Schultz 等人(1997)的模型属于严格的 RL,其中多巴胺信号被识别为 TD 误差。在基于更普遍、更广泛的 RL 概念的模型中,多巴胺只是表示某种形式的奖励、显著性或错误信息。甚至 Prescott、Redgrave、Gurney 等人的模型(Redgrave 等人, 1999a, b; Prescott 等人, 2002),其中多巴胺信号表示动作的显著性,并用于动作选择,可以解释为属于这一广泛的 RL 范畴。事实上,从目前的

情况来看, 这种对 BG 功能的更广泛看法似乎更接近事实真相。现在假设严格的基于 RL 的方法可以解决所有问题还为时过早, 因为这种解释并非没有缺点。

虽然 Schultz(1998)最初的电生理实验已经被重复了几次, 但它们只在一个实验室中进行过。很明显, 其他实验室需要重复这些实验。甚至在建模领域, 也有几次尝试扩展多巴胺信号与 TD 误差相等的主题。最近的一个模型提出了强直性 DA 活动和阶段性 DA 活动的不同作用, 前者与长期奖励预测有关, 后者与试验特异性预测有关(Daw 和 Touretzky, 2000)。另一种观点认为, 相位 DA 激发与预测误差有关, 紧张性 DA 活性成分与预测的不确定性有关(Fiorillo 等人, 2003)。

因此, 可以对 BG 功能给出一个更具包容性的概念总结如下:

BG 将多巴胺信号所携带的奖励或显著性或误差相关信息与感觉运动皮层状态相结合, 为运动皮层提供适当的纠正。

还有其他一些 BG 模型甚至不属于这个更广泛的 RL 的范畴。例如, Leblois 等人(2006)将 BG 及其相关结构组织成两个回路: 第一个回路由“纹状体-皮质-GPi-丘脑-皮质”组成, 他们将其称为传统的直接通路; 第二个回路由“皮质-STN-GPi-丘脑-皮质”组成, 他们将其称为超直接通路。作者认为这两个环路之间的竞争负责行动选择。它们还显示了多巴胺的适度减少是如何导致单配对行为选择的, 而更大的多巴胺缺乏会产生振荡。在另一项激进的提议中, Obeso 等人(2006)将 GPe 置于中心位置, 皮质-STR-GPi 和皮质-STN-GPi 为两条环绕 GPe 的平行通路(图 6)。作者主要通过实验研究来支持他们的架构, 没有提到强化学习。多巴胺信号在 BG 功能中的作用也没有明确的描述。看看这样的架构是否可以解释 BG 的 RL 功能将是一个不小的挑战。

在对 BG 功能进行了如此多的重新思考之后, 即使经过近 30 年对建模的深入研究, 这一课题似乎比以往任何时候都更具挑战性。然而, 在这些相互矛盾的建议中, 一个基本的观点是 BG 是一个皮层下结构的网络, 它通过处理中脑多巴胺能信号中与奖赏相关的信息, 对皮层控制结构中的决策产生强烈影响。所有激烈的争论都是关于在特定的 BG 核中发生的详细信息处理的性质。但是, 俗话说, 细节决定成败。

致谢

作者致谢他们与 Gangadhar Garipelli 和 Sridharan Devarajan 就基底神经节建模的各种主题进行了许多有益的讨论。这项工作由印度科学技术部的一项研究资助。

附录

An alternative view of DP and IP function

The proposed complementary function of DP and IP was recently developed in a model that highlights the role of BG

in reaching movements (Mohamed et al. 2009). In this model, the motor cortex (MC) is trained to reach one of four targets on command. It is modeled as a perceptron which takes the target selection vector (TSV), ξ , as input and produce, g_m , as output. The TSV of the i th target is represented as $\xi_i(t) = 1$, and $\xi_i(j) = 0$ for $j \neq i$. The output of MC, g_m , and the output of BG (g_{bg}) are combined linearly to produce:

$$g = \alpha g_m + \beta g_{bg}$$

where g represents activations to be given to the four muscles of a two-link arm. BG output, g_{bg} , is a time-varying quantity driven by a chaotic source located in the IP. This variable g_{bg} induces random, wandering movements of the arm. The value function, $V(t)$, evaluates the nearness of the arm's end-effector to the target and is given as:

$$V(X; X_i^{\text{tar}}) = A \left(1 - \frac{d^2}{r^2}\right) \text{ for } d < r \\ = 0 \text{ otherwise}$$

where $d = \|X - X_i^{\text{tar}}\|$.

Value is thought to be computed within the Striatum. The TD error, δ , which represents the phasic activity of DA cells of SNc, as follows:

$$\delta(t) = r(t) + \gamma V(t) - V(t-1)$$

The reward, $r(t) = A$, when $d < R_{\text{small}}$, otherwise $r(t) = 0$, where R_{small} is a small positive quantity.

Accidental target-wards movement of the arm, results in positive $\delta(t)$, while negative shifts in $\delta(t)$ accompany movements away from the target. Fluctuations in $\delta(t)$ in turn control the BG output, g_{bg} , according to the following rule:

$$\begin{aligned} \text{if } (\delta(t) > DA_{hi}) & \\ \Delta g_{bg}(t) &= +\Delta g_{bg}(t-1) & \text{--"Go"} \\ \text{else if } (\delta(t) > DA_{lo} \text{ and } \delta(t) \leq DA_{hi}) & \\ \Delta g_{bg}(t) &= \varphi & \text{--"Explore"} \\ \text{else // } (\delta(t) \leq DA_{lo}) & \\ \Delta g_{bg}(t) &= -\Delta g_{bg}(t-1) & \text{--"NoGo"} \end{aligned}$$

where φ is a random two-dimensional vector such that norm (φ) = η , and g_{bg} is updated such that $g_{bg}(t+1) = g_{bg}(t) + \Delta g_{bg}$.

A large positive fluctuation in $\delta(t)$ ($> DA_{hi}$) indicates that the last move is in the right direction, resulting in continuation in the same direction: this regime corresponds to the Go case of classical Go/NoGo interpretation of DP/IP function. Similarly, a large negative fluctuation in $\delta(t)$ ($< DA_{lo}$) indicates that the last move is unacceptable, in response to which the arm is commanded to move in exactly opposite direction. This regime corresponds to the NoGo case. A novel aspect of the proposed model consists in positing a third regime—the Explore regime—between Go and NoGo regimes. For $DA_{lo} < \delta(t) < DA_{hi}$, the last move is “neither good nor

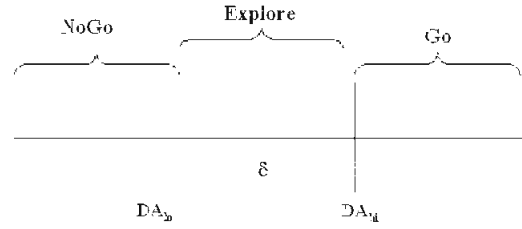


Fig. A1 The three-regime scheme involving Go-Explore-NoGo in which each regime corresponds to a range of values of δ

bad.” In this case, the arm randomly explores a new direction. Driven by such wandering movements, when the arm accidentally arrives at the target, it receives a reward. The muscle activations that the arm receives at that instant are used to train the MC. With such a training scheme, the arm negotiates long, tortuous movements to reach the target in the early stages. This variability in reaching trajectories reduces as training progresses.

Thus as a radical departure from the classical Go-NoGo picture of DP/IP function, we propose a Go-Explore-NoGo scenario (Fig. A1), where the site of “Go” is DP, while both “Explore” and “NoGo” are subserved by IP.

References

- Aizawa H, Tanji J (1994) Corticocortical and thalamocortical responses of neurons in the monkey primary motor cortex and their relation to a trained motor task. *J Neurophysiol* 71:550–560
- Alexander GE (1987) Selective neuronal discharge in monkey Putamen reflects intended direction of planned limb movements. *Exp Brain Res* 67:623–634
- Alexander GE, Crutcher MD, DeLong MR (1990) Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, prefrontal and limbic functions. *Prog Brain Res* 85:119–146
- Arbib MA, Erdi P, Szentagothai J (1998) Neural organization: structure, function and dynamics. MIT Press, Cambridge, pp 203
- Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *J Neurosci* 14:4467–4480
- Bapi RS, Doya K (1998) A sequence learning architecture based on cortico-basal ganglionic loops and reinforcement learning. In: *ICONIP'98, Kita-Kyushu, Japan, 21–26 Oct 1998*
- Bapi RS, Miyapuram KP, Graydon FX, Doya K (2006) fMRI investigation of cortical and subcortical networks in the learning of abstract and effector-specific representations of motor sequences. *Neuroimage* 32(2):714–727
- Barto AG (1995) Adaptive critics and the basal ganglia. In: Houk JC, Davis J, Beiser D (eds) *Models of information processing in the basal ganglia*. MIT Press, Cambridge, pp 215–232
- Bergman H, Wichmann T, Karmon B, DeLong MR (1994) The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of Parkinsonism. *J Neurophysiol* 72:507–520
- Bergman H, Feingold A, Nini A, Raz A, Slovlin H, Abeles M, Vaadia E (1998) Physiological aspects of information processing in the basal ganglia of normal and parkinsonian primates. *Trends Neurosci* 21:32–38

- Berns GS, Sejnowski TJ (1995) A computational model of local memory in the primate pallidal-subthalamic circuit. *Soc Neurosci Abstr* 21:678
- Berns GS, Sejnowski TJ (1998) A computational model of how the Basal ganglia produce sequences. *J Cognit Neurosci* 10:108–121
- Bhatia KP, Marsden CD (1994) The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain* 117:859–876
- Brown PAO, Mazzone P, Insola A, Tonali P, Di Lazzaro V (2001) Dopamine dependency of oscillations between subthalamic nucleus and pallidum in Parkinsons disease. *J Neurosci* 21:1033–1038
- Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6:755–756
- Chaudhuri A, Behan PO (2000a) Fatigue in neurological disorders. *Lancet* 179:34–42
- Chaudhuri A, Behan P (2000b) Neurological dysfunction in chronic fatigue syndrome. *J Chron Fatigue Syndr* 6:51–68
- Chaudhuri A, Behan P (2000c) Fatigue and basal ganglia. *J Neurol Sci* 179:34–42
- Chesselet M, Delfs JM (2003) Basal ganglia and movement disorders: an update. *Trends Neurosci* 19(10):417–422
- Churchland MM, Yu BM, Ryu SI, Santhanam G, Shenoy KV (2006) Neural variability in premotor cortex provides a signature of motor preparation. *J Neurosci* 26:3697–3712
- Clark D, Boutros N, Mendez M (2005) The brain and behavior. Cambridge University Press, Cambridge
- Cohen JD, Braver TS, Brown JW (2002) Computational perspectives on dopamine function in prefrontal cortex. *Curr Opin Neurobiol* 12:223–229
- Contreras-Vidal JL, Schultz W (1999) A predictive reinforcement model of dopamine neurons for learning approach behavior. *J Comput Neurosci* 6(3):191–214
- Contreras-Vidal JL, Stelmach GE (1995) A neural model of basal ganglia-thalamocortical relations in normal and parkinsonian movement. *Biol Cybern* 73(5):467–476
- Contreras-Vidal JL, Stelmach GE (1996) Effects of Parkinsonism on motor control. *Life Sci* 58:165–176
- Crammond DJ, Kalaska JF (1989) Neuronal activity in primate parietal cortex area 5 varies with intended movement direction during an instructed-delay period. *Exp Brain Res* 76:458–462
- Cunnington R, Iansek R, Bradshaw JL, Phillips JG (1995) Movement-related potentials in Parkinsons-disease presence and predictability of temporal and spatial cues. *Brain* 118:935–950
- Czernecki V, Pillon B, Houeto JL, Pochon JB, Levy R, Dubois B (2002) Motivation, reward, and Parkinson's disease: influence of dopatherapy. *Neuropsychologia* 40:2257–2267
- Daw ND (2003) Reinforcement Learning models of the dopamine system and their behavioral implications. Ph.D. thesis. Carnegie Mellon University, Pittsburgh
- Daw ND, Touretzky DS (2000) Behavior results suggest an average reward TD model of dopamine function. *Neurocomputing* 32:679–684
- Daw ND, O'doherty JP, Seymour B, Dayan P, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879
- Dean WH, Davis GD (1958) Behavior changes following caudate lesions in rhesus monkeys. *J Neurophysiol* 22:165–187
- Di Chiara G, Bassareo V, Fenu S, De Luca MA, Spina L, Cadoni C, Acquas E, Carboni E, Valentini V, Lecca D (2004) Dopamine and drug addiction: the nucleus accumbens shell connection. *Neuropharmacology* 47(Suppl 1):227–241
- Dorigo M, Colombetti M (1994) Robot shaping: developing autonomous agents through learning. *Artif Intell* 71(2):321–370
- Doya K (2002) Metalearning and neuromodulation. *Neural Netw* 15(4):495–506
- Doyon J, Penhune V, Ungerleider LG (2003) Distinct contribution of the corticostriatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41(3):252–262
- Engelborghs S, Marien P, Pickut BA, Verstraeten S, De Deyn PP (2000) Loss of psychic self-activation after paramedian bithalamic infarction. *Stroke* 31:1762–1765
- Eslinger PJ, Damasio AR (1985) Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 35:1731–1741
- Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299(5614):1898–1902
- Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J Cognit Neurosci* 17:51–72
- Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A (1994) The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. *Ann Intern Med* 121:953–959
- Fuster JM (1997) The prefrontal cortex. Raven Press, New York
- Gabrieli JDE, Poldrack RA, Desmond JE (1995) The role of left prefrontal cortex in language and memory. *Adv Neurol* 66:21–34
- Gangadhar G, Joseph D, Chakravarthy VS (2007) An oscillatory neuro-motor model of handwriting generation. *Int J Doc Anal Recognit* 10(2):69–84
- Gangadhar G, Joseph D, Chakravarthy VS (2008) Understanding Parkinsonian Handwriting using a computational model of basal ganglia. *Neural Comput* 20:1–35
- Georgiou N, Iansek R, Bradshaw JL, Phillips JG, Mattingley JB (1993) An evaluation of the role of internal cues in the pathogenesis of Parkinsonian hypokinesia. *Brain* 116:1575–1587
- Georgopoulos AP, Grillner S (1989) Visuomotor coordination in reaching and locomotion. *Science* 245:1209–1210
- Gerfen CR, Wilson CJ (1996) The basal ganglia. In: Swanson LW, Bjorklund A, Hokfelt T (eds) *Handbook of chemical neuroanatomy*. Elsevier, Amsterdam, pp 371–468
- Ghika-Schmid F, Bogousslavsky J (2000) The acute behavioral syndrome of anterior thalamic infarction: a prospective study of 12 cases. *Ann Neurol* 48:220–227
- Giladi N, McMahon D, Przedborski S, Flaster E, Guillory S, Kostic V, Fahn S (1992) Motor blocks in Parkinson's disease. *Neurology* 42:333–339
- Goldman-Rakic PS (1987) Circuitry of primate prefrontal cortex and regulation of behaviour by representational memory. In: Plum F, Mountcastle U (eds) *Handbook of physiology*. The American Physiological Society, Washington, pp 373–417
- Heise C, Kayalioglu G (2008) Cytoarchitecture of the spinal cord. In: Watson C, Paxinos G, Kayalioglu G (eds) *The spinal cord*. Elsevier, London
- Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Miyachi S, Doya K (1999) Parallel neural networks for learning sequential procedures. *Trends Neurosci* 22:464–471
- Hikosaka O, Takikawa Y, Kawagoe R (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol Rev* 80:953–978
- Hollerbach JM (1981) An oscillation theory of handwriting. *Biol Cybern* 39:139–156
- Houk JC, Davis JL, Beiser DG (1995) Models of information processing in the basal ganglia. MIT Press, Cambridge
- Hurtado JM, Graym CM, Tamas LB, Sigvardt KA (1999) Dynamics of tremor-related oscillations in the human globus Pallidus: a single case study. *Neurobiology* 96:1674–1679

- Isoda M, Hikosaka O (2008) Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. *J Neurosci* 28(28):7209–7218
- Iversen SD (1979) Behaviour after neostriatal lesions in animals. In: Divac I, Oberg RGB (eds) *The neostriatum*. Pergamon, Oxford pp 195–210
- Joel D, Niv Y, Ruppel E (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw* 15:535–547
- Kakade S, Dayan P (2002) Dopamine: generalization and Bonuses. *Neural Netw* 14(4):549–559
- Keath JR, Iacoviello MP, Barrett LE, Mansvelter HD, McGehee DS (2007) Differential modulation by nicotine of substantia nigra versus ventral tegmental area dopamine neurons. *J Neurophysiol* 98:3388–3396
- Knight RT, Grabowecy MF, Scabini D (1995) Role of human prefrontal cortex in attention control. *Adv Neurol* 66:21–36
- Kornhuber HH, Deecke L (1990) Readiness for movement—the Bereitschafts potential-story. *Curr Contents Life Sci* 33(22)
- Kubota K, Hamada I (1978) Visual tracking and neuron activity in the post-arcuate area in monkeys. *J Physiol (Paris)* 74:297–312
- Kubota K, Funahashi S (1982) Direction-specific activities of dorsolateral prefrontal and motor cortex pyramidal tract neurons during visual tracking. *J Neurophysiol* 47:362–376
- Lang W, Cheyne D, Kristeva R, Beisteiner R, Lindinger G, Deecke L (1991) Three-dimensional localization of SMA activity preceding voluntary movement. *Exp Brain Res* 87:688–695
- Leblois A, Boraud T, Meissner W, Bergman H, Hansel D (2006) Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *J Neurosci* 26:3567–3583
- Levy R, Dubois B (2005) Apathy and the functional anatomy of the prefrontal cortex—basal ganglia circuits. *Cereb Cortex* 16:916–928
- Magnin M, Morel A, Jeanmonod D (2000) Single-unit analysis of the pallidum, thalamus and subthalamic nucleus in parkinsonian patients. *Neuroscience* 96:549–564
- Malapani C, Rakitin B, Levy R, Meck WH, Deweer B, Dubois B, Gibbon J (1998) Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. *J Cognit Neurosci* 10:316–331
- Marsden CD (1982) The mysterious motor function of the basal ganglia: the Robert Wartenberg Lecture. *Neurology* 32(5):514–539
- Matell MS, Meck WH (2004) Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognit Brain Res* 21:139–170
- Mink JW (1996) The basal ganglia: focused selection and inhibition of competing motor programs. *Prog Neurobiol* 50:381–425
- Mohamed M, Chakravarthy VS, Subramanian D, Ravindran B (2009) The role of basal ganglia in performing simple reaching movements: a computational model. In: *Proceedings of 14th conference of the international graphonomics society*, Dijon, France, 13–16 September
- Montague PR, Berns G (2002) Neural economics and the biological substrates of valuation. *Neuron* 36:265–284
- Montague PR, Dayan P, Person C, Sejnowski TJ (1995) Bee foraging in uncertain environments using predictive Hebbian learning. *Nature* 377:725–728
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947
- Montague PR, Hyman SE, Cohen JD (2004) Computational roles for dopamine in behavioral control. *Nature* 431:760–767
- Nini A, Feingold A, Slovlin H, Bergman H (1995) Neurons in the globus pallidus do not show correlated activity in the normal monkey, but phase-locked oscillations appear in the MPTP model of Parkinsonism. *J Neurophysiol* 74:1800–1805
- Obeso JA, Rodriguez-Oroz MC, Blesa FJ, Guridi J (2006) The globus pallidus pars externa and Parkinson's disease. Ready for prime time?. *Exp Neurol* 202:1–7
- Oorschot DE, Tunstall MJ, Wickens JR (2002) Local connectivity between striatal spiny projection neurons: a re-evaluation. In: Nicholson L, Faull R (eds) *Basal ganglia VII*. Plenum Press, New York
- O'Reilly RC, Frank MJ (2006) Making working memory work: a computational model of learning in the frontal cortex and basal ganglia. *Neural Comput* 18:283–328
- Packard MG, Knowlton BJ (2002) Learning and memory functions of the basal ganglia. *Annu Rev Neurosci* 25:563–593
- Packard MG, McGaugh JL (1996) Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Memory* 65:65–72
- Packard MG, Hirsh R, White NM (1989) Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *J Neurosci* 9:1465–1472
- Plenz D, Kitai ST (1999) A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus. *Nature* 400:677–682
- Prescott TJ, Gurney K, Montes-Gonzalez F, Humphries M, Redgrave P (2002) The robot basal ganglia: action selection by an embedded model of the basal ganglia. In: Nicholson L, Faull R (eds) *Basal ganglia VII*. Plenum Press, New York, pp 349–356
- Raz A, Vaadia E, Bergman H (2000) Firing patterns of spontaneous discharge of pallidal neurons in the model of Parkinsonism. *J Neurosci* 20:8559–8571
- Redgrave P, Prescott TJ, Gurney K (1999a) Is the short-latency dopamine response too short to signal reward error. *Trends Neurosci* 22:146–151
- Redgrave P, Prescott TJ, Gurney K (1999b) The basal ganglia: a vertebrate solution to the selection problem?. *Neuroscience* 89:1009–1023
- Russell VA, Allim R, Lamm MC, Taljaard JJ (1992) Regional distribution of monoamines and dopamine D1- and D2-receptors in the striatum of the rat. *Neurochem Res* 17:387–395
- Sacco P, Hope PAJ, Thickbroom GW, Byrnes ML, Mastaglia FL (1999) Corticomotor excitability and perception of effort during sustained exercise in the chronic fatigue syndrome. *Clin Neurophysiol* 110:1883–1891
- Scheiner R, Pluckhahn S, Oney B, Blenau W, Erber J (2002) Behavioral pharmacology of octopamine, tyramine, and dopamine in honey bees. *Neurology* 42:333
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599
- Servan-Schreiber D, Printz H, Cohen JD (1990) A network model of catecholamine effects: gain, signal-to-noise ratio and behavior. *Science* 249:892–895
- Smith Y (2008) The thalamus. In: Michael Conn P (ed) *Neuroscience in medicine*. Humana Press, New York
- Smith Y, Bevan MD, Shink E, Bolam JP (1998) Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* 86:353–387
- Squire LR, Zola SM (1996) Structure and function of declarative and nondeclarative memory systems. *Proc Natl Acad Sci USA* 93:13515–13522
- Sridharan D, Prashanth PS, Chakravarthy VS (2006) The role of the basal ganglia in exploration in a neural model based on reinforcement learning. *Int J Neural Syst* 16:111–124
- Starr A, Scalise A, Gordon R, Michalewski HJ, Caramia MD (2000) Motor cortex excitability in chronic fatigue syndrome. *Clin Neurophysiol* 111:2025–2031

-
- Stout JC, Johnson SA (2005) Cognitive impairment and dementia in basal ganglia disorders. *Curr Neurol Neurosci Rep* 5:355–363
- Stuss DT, Van Reekum R, Murphy KJ (2000) Differentiation of states and causes of apathy. In: Borod JC (ed) *The Neuropsychology of emotion*. Oxford University Press, Oxford, pp. 340–363
- Suri RE, Schultz W (1999) A neural network model with dopamine-like reinforcement signal that learns a spatial delayed response task. *Neuroscience* 91:871–890
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. MIT Press, Cambridge
- Sutton RS, Barto AG (1990) Time-derivative models of Pavlovian reinforcement. In: Gabriel M, Moore J (eds) *Learning and computational neuroscience*. MIT Press, Cambridge
- Tanji J (1994) The supplementary motor area in the cerebral cortex. *Neurosci Res* 19:251–268
- Terman D, Rubin JE, Yew AC, Wilson CJ (2002) Activity patterns in a model for the subthalamopallidal network of the basal ganglia. *J Neurosci* 22:2963–2976
- Thut G, Schultz W, Roelcke U, Nienhusmeier M, Missimer J, Maguire RP, Leenders KL (1997) Activation of the human brain by monetary reward. *Neuroreport* 8:1225–1228
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708
- Usher M, Cohen JD, Servan-Schreiber D, Rajkowski J, Aston-Jones G (1999) The role of locus coeruleus in the regulation of cognitive performance. *Science* 22:549–554
- Wickens J (1997) Basal ganglia: structure and computations. *Netw Comput Neural Syst* 8(4):R77–R109
- Wise PM (1982) Norepinephrine and dopamine activity in microdissected brain areas of the middle-aged and young rat on Proestrus. *Biol Reprod* 27:562–574
- Worgotter F, Porr B (2005) Temporal sequence learning, prediction, and control: a review of different models and their relation to biological mechanisms. *Neural Comput* 17:245–319
- Yin HH, Knowlton BJ (2006) The role of basal ganglia in habit formation. *Nat Neurosci Rev* 7:464–476