

循环神经回路的感知决策

决策一直是认知神经生理学研究的一个中心课题，而实验和计算的研究引起了关于基本决策计算的皮层回路机制的提出。这种机制依赖于慢的突触循环兴奋和快的反馈抑制相平衡，这种平衡不仅为形成类别选择提供了吸引子状态的实例，而且为支持或反对其他选项的证据累积提供了长瞬态。这种具有奖赏依赖突触可塑性的回路能够产生适应性选择行为。而决策阈值是反应时间任务的核心概念，它可以从一般的决策规则中分离出来。此外，感知决策和基于价值的经济选择是在一个统一的框架内进行描述的，其中概率选择是由不规律的神经活动以及决策者与不确定环境或社会群体中其他不可预测的决策者之间的反复交互作用所导致的。

引言

决策是一种认知过程，是在两种或两种以上的选择方案中选择一种意见或一种行动，具有几个明确的特征。首先，可选的选项不仅仅是反射响应，还包括目标导向的行动，可以对预期的结果进行某种程度的评估，并在决策过程中加以考虑。其次，受控决策的一个特点是它是信息积累和深思熟虑的过程。第三，几乎所有感兴趣的决定都蕴含着风险；事实上，我们可以说，决策的本质是在面对长期后果的不确定性时做出正确的选择。

除了诸如战争与和平、婚姻或司法判决等重大决策之外，决策贯穿于我们日常生活中灵活行为的各个方面。我们制定一个目标，然后做出一系列的选择来实现这个目标。自愿选择性注意，在有意识地引导感官处理的意义上，依赖于决定外部世界中什么是在行为上最相关的。感知依赖于对感觉场景的判断，在这个场景中，需要检测、识别和分辨相互冲突和模糊的输入信号。给定了这些感官信息，生物体根据预期结果和这些行为的相关风险，在可用的选项中选择行动。当环境发生变化或结果取决于社会环境中相互作用的决策者的所有选择时，选择偏好和策略必须具有灵活的适应性。

尽管决策在认知中起着核心作用，但直到最近，人们对其神经基础知之甚少。由于动物行为生理学、人类脑成像、理论和神经回路建模的汇合，最近十年来，人们对这一领域的兴趣和研究活动激增。特别是，神经生理学家已经开始研究非人灵长类动物在各种决策任务中的行为，包括感知辨别(Shadlen and Newsome, 1996, 2001; Romo and Salinas, 2001; Roitman and Shadlen, 2002; Romo et al., 2004; Heekeren et al., 2008)，目标选择(Hanes and Schall, 1996; Schall, 2001, 2004; Cisek and Kalaska, 2005; Scherberger and Andersen, 2007)，经济的选择行为(Platt and Glimcher, 1999; Sugrue et al., 2004, 2005; Padoa-Schioppa and Assad, 2006)，和竞争游戏(Barraclough et al., 2004; Dorris and Glimcher, 2004; Glimcher, 2003; Lee, 2008)。这些实验揭示了单细胞水平上与决策计算的特定方面相关的神经信号。然而，在哺乳动物的大脑中，一个决定不是由单个细胞做出的，而是由神经回路的集体动力学行为做出的。观察到的神经信号是如何产生的？与初级感觉区的早期处理相比，使局部皮层区域(如前额叶或后顶叶皮层)能够辅助决策计算的特性是什么？怎样才能建立起连接分子、回路和决策行为的因果链呢？

在与实验的密切互动中，真实的神经回路建模为解决这些基本问题提供了一个有价值的工具。基于生物物理的模型可以将不同层次的描述、探测细胞和网络机制联系在一起，这些机制一方面构成了所观察到的神经尖峰活动的基础，另一方面它在行为层面上的解释了性能。此外，决策计算依赖于被赋予大量正反馈和负反馈回路的皮层回路，其行为仅凭直觉是难以预测的。非线性动力系统理论为描述和预测这类强循环神经系统的行为提供了一个数学框架。

细胞水平的建模已被证明对理解单个突触、单个神经元的行为和感觉处理(如初级视觉皮层的定向选择性机制)非常有用。另一方面，像决策这样的认知过程在很大程度上是由抽象的数学模型来描述的。基于生物物理的尖峰网络模型已经被用在各种实验范式里，包括涉及决策和工作记忆的感知任务，行动选择和准备，学习灵活的感觉运动关联，以及基于奖励的经济选择行为，如觅食或互动游戏。这些模型的基本假设是相似的。循环突触兴奋足够强就能产生多个自我维持稳定

的神经群体状态，这在数学上被称为“吸引态”。反射的兴奋性是由一个慢细胞过程来实例化的，随着时间的推移，会引起神经活动的长时间增加。因此，网络的行为不一定由稳态(代表类别选择)主导，而缓慢的瞬态动力学为信息的时间积累提供了一种神经机制。另一方面，反馈抑制实现了构成类别选择的竞争动力。此外，神经元高度不规则的尖峰活动在产生随机选择行为中起着关键作用。最后，基于奖赏的突触可塑性实现了反映过去选择历史结果的学习，导致神经元活动在不断变化的环境中或在社会环境中与其他决策者的互动中进行适应性计算。由于它们的共性，这些模型将被统称为“循环神经回路模型”。

本文从计算的角度综述了电生理学的最新发现。重点将放在基本计算上：(1) 证据的积累(信息随时间积累的细胞基础是什么?)，(2) 类别选择的形成(神经元考虑过程的终止规则是什么?)；(3) 基于奖励的适应(是通过神经元或突触学习到的选择响应的值；潜在的可塑性过程可能是什么?)；(4) 选择行为固有的随机性(关于世界的不确定性在大脑中如何表现？在行为选择中的随机性的内在神经来源是什么?)。这些计算是许多决策过程的核心；因此，了解它们的神经基础对于决策的生物学基础是至关重要的。

额顶叶回路中的神经元过程，是信息积累和分类选择的基础

深思熟虑的决策的一个标志是时间整合，这个过程使我们能够积累支持或反对备选方案的证据，并仔细考虑选择方案。虽然我们能够迅速作出反应，但仓促作出的决定可能产生不良影响。通常需要在速度和准确性之间进行权衡：响应时间越慢，性能越好(Wickelgren, 1977)。此外，当外部世界提供的信息相互冲突或有许多选择可供考虑时，我们通常会花更长的时间来考虑更困难的决策(Hick, 1952; Vickers, 1970)。

在行为层面，反应时间(RT)测量为探索在感知、记忆和认知过程中的时间整合提供了一个强大的工具。(Donders, 1969; Posner, 1978; Luce, 1986; Meyer et al., 1988)。RT 研究已经导致了累积模型的发展，它以各种方式实现了输入

信号到固定决策阈值的随机整合。在竞争模型中，表示不同选择选项的累加器构建了它们的活动，最先到达预定阈值的选项将生成选择 (Logan and Cowan, 1984)。在两种选择中强迫选择的漂移扩散模型中，累加器增加支持一种选择的证据，减去支持另一种选择的证据；当它达到一个正阈值或一个负阈值时，就会做出决策 (Stone, 1960; Laming, 1968; Ratcliff, 1978; Smith and Ratcliff, 2004)。一个线性泄漏竞争积分器 (LCA) 模型，它模拟了一个神经网络，考虑了积分的泄漏，并假设对不同选项的积分之间的竞争抑制 (Usher and McClelland, 2001)。这个模型很容易扩展到具有多个备选方案的决策。(Usher and McClelland, 2001; McMillen and Holmes, 2006; Bogacz et al., 2007)，但是这个模型对于扩散模型来说不那么明确对于有两个可选项的任务，在泄漏和抑制相互抵消的特殊情况下，LCA 模型可以简化为扩散模型 (Usher and McClelland, 2001)。扩散模型因其简单易行而广受欢迎，且在大量的人体研究中，它已经成功地拟合了行为数据 (Ratcliff, 1978; Busemeyer and Townsend, 1993; Smith and Ratcliff, 2004)。

尽管时间整合的概念很吸引人，但是哪种类型的选择行为涉及这种积累过程 (深思熟虑的决策的一个特征) 以及在什么时间尺度上参与这种积累过程并不明显 (Uchida et al., 2006)。当速度非常重要，在一组可能的动作中进行的是一种可以迅速发生的选择。简单生物体的例子说明了这一点 (Real, 1991)：蟾蜍会产生捕食或躲避行为，这取决于一个模糊的移动物体是被视为猎物还是掠食者 (Ewert, 1997)；或者是箭鱼，通过观察猎物的初始状态，快速的 (100 毫秒内) 决定捕捉猎物的行动路线 (Schlegel and Schuster, 2008)。在啮齿类动物中，一项研究报告说，嗅觉辨别在大鼠中是快速的 (300 毫秒)。通过调整二元混合物中两种气味的相对比例，任务难度发生了变化，但 RTs 只发生了轻微的变化 (变化了 30 毫秒)，因此，工作性能从随机水平变化到接近 100% 的正确率 (Uchida and Mainen, 2003)。相比之下，另一项研究发现，在老鼠中，无论辨别难度如何，嗅觉辨别能力都很高 (95% 正确率)，而 RTs 则从最简单到最困难增加了 80 ms (Abraham et al., 2004)，这表明在这些任务中，啮齿动物在不到 100 ms 的时间尺度上表现出速度-准确度的权衡。在人类研究中，在简单的感知任务中，

RTs 的平均时间通常在几十毫秒到大约一秒之间 (Luce, 1986; Usher and McClelland, 2001)。

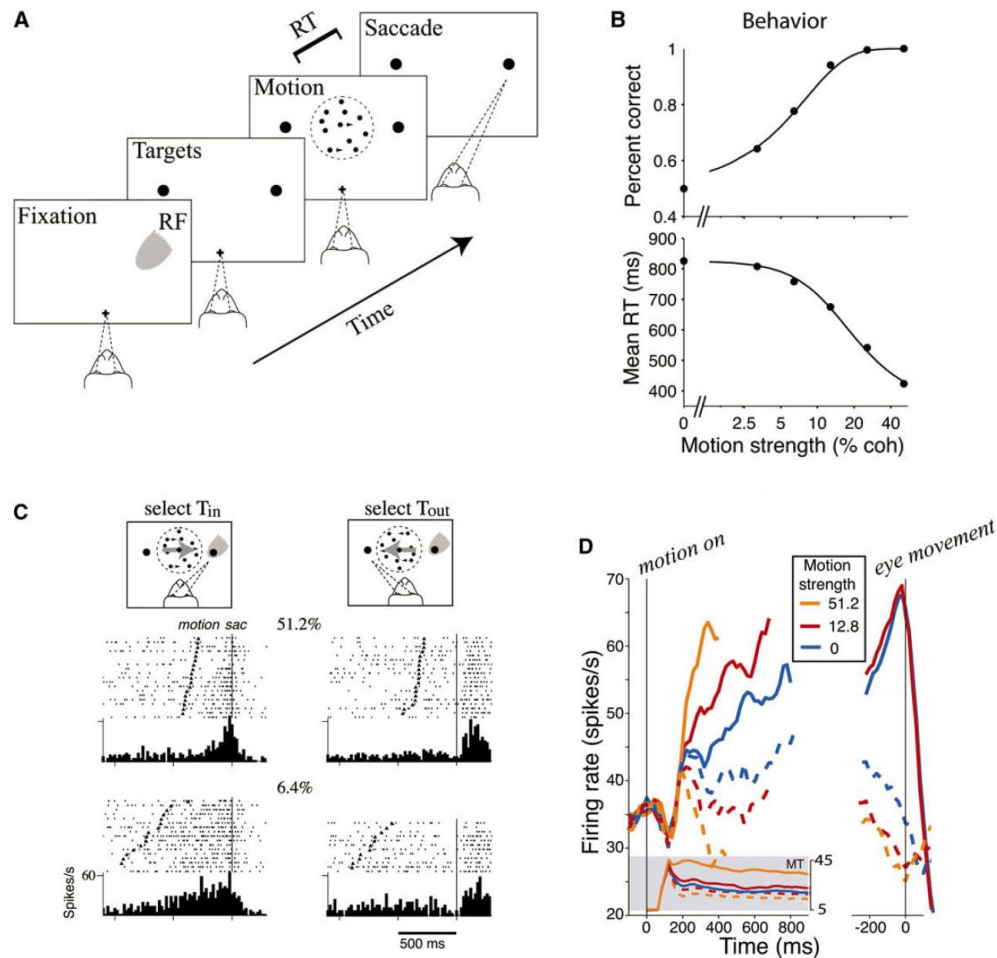


Figure 1. 猴子随机点运动方向识别实验中决策的神经机制。 (A) 任务的反应时间 (RT) 版本。受试者观察一组动态随机点，并决定运动的净方向。这个决定是通过眼睛移动到两个外围目标之一 (代表两个被迫的选择) 来表示的。在 RT 任务中，受试者控制观看时间，只要准备好，就用眼动结束每次试验。灰色斑块显示 LIP 神经元的响应场 (RF) 的位置。(B) 猴子的表现 (上) 和平均 RT (下) 作为运动强度的函数。(C) 单个 LIP 神经元的响应。只显示了在两种运动强度 (6.4% 和 51.2%) 下的正确选择。发放光栅图和反应直方图与猴子眼动反应的开始 (垂直线) 一致。符号 (^) 表示随机点运动的开始。试验栅格按 RT 排序。(D) LIP 神经元在决策形成过程中的平均反应，分为三个难度等级。如图所示，响应按运动强度和选择方向

进行分组。(左)响应与随机点运动的开始对齐。在决策形成过程中显示平均值(在眼动前的中位 RT 或 100ms 截断的曲线)。瞬态过后, MT 的响应速度几乎是恒定的。(右)LIP 的神经反应与眼睛的运动是一致的。(A)、(B)及(D)经 Gold and Shadlen (2007) 的许可而复制(从 Britten et al. [1992]使用的在线数据库插入); (C)转载自 Roitman and Shadlen (2002)。

时间整合背后的神经过程是什么? 最近,对行为正常的猴子进行的电生理学研究表明,反应时间与单细胞水平的神经活动有关。在二择一的视觉随机点运动方向(RDM)辨别任务中,训练猴子对接近阈值的随机点视觉运动刺激的运动方向做出二值判断,并以眼球运动的形式报告感知的运动方向。任务的难易程度可以通过在同一方向上相干运动的点的百分比来参数化,称之为运动强度或者一致性百分比 c' 。大量的生理和微刺激研究表明,在 MT 区域的方向敏感神经元编码运动刺激(Newsome et al., 1989; Britten et al., 1992, 1993, 1996),决策过程本身发生在大脑皮层的下游,可能位于后顶叶皮层和/或前额叶皮层。Shadlen 和 Newsome 发现,在正确和错误的实验中,顶叶外侧区域(LIP)神经元的活动与猴子的知觉选择有关(Shadlen and Newsome, 1996, 2001)。此外,在任务的反应时间版本中(Roitman and Shadlen, 2002; Huk and Shadlen, 2005),受试者的响应时间从最简单($c' = 100\%$)到最困难($c' = 3.2\%$)的任务中,增加了 400 ms (Figures 1A and 1B)。LIP 细胞表现出特有的放电时间历程,反映了猴子的反应时间和感知选择(Figures 1C and 1D)。从随机点运动刺激开始到猴子通过快速的眼球运动产生选择反应,对特定的扫视目标有选择性的 LIP 神经元的峰值活动增加了数百毫秒。 c' 越高,坡度越大(感官信息质量越高)。进一步地,我们观察到,当 LIP 神经元的放电速率(对该选择反应的选择性)达到一个不依赖于 c' 和响应时间的阈值时,就会做出决策选择(如扫视)。因此,正如神经积分器所预期的那样,这些 LIP 神经细胞显示出随机上升到一个设定的水平。

另一项生理学研究报告说,当猴子执行一项任务来检测视觉运动刺激的存在时,后顶叶皮层 (VIP)区域的神经元表现出上升的活动,这与主观判断(命中和误报

试验的活动高于未命中和正确拒绝试验)和响应时间(范围为 400 - 700 毫秒)相关(Cook and Maunsell, 2002)。MT 神经的响应在成功检测和失败检测中更大,这意味着由 MT 神经元向更高区域提供的信号强度在最终成功探测的试验中比在失败探测的试验中的更强。在 VIP 区域而非 MT 区域中,假警示试验中神经元活动明显大于基线,表明主观判断是在 MT 下游计算出来的。因此,这项研究支持了顶叶皮层是大脑系统的一部分,是视觉感知中信息积累和主观判断的基础。

关于上升的神经活动,到目前为止,大多数记录都局限于一次一个细胞,上升活动通常被报告为试验平均的神经活动 (Roitman and Shadlen, 2002)。这就留下了一个问题,即单次试验中的单峰放电是否确实显示了发放率的准线性上升。或者,神经元实际上可以经历一个速率到另一个发放率的突然跳跃,但是跳跃时间在不同的试验之间是不同的,因此试验的平均值显示出一个平滑的斜坡 (Okamoto et al., 2007)。附加的实验,也许是多个单元记录,将有助于解决这个问题。此外,目前尚不清楚顶叶皮质中观察到的神经活动是局部产生的,还是反映了来自其他地方的输入。后顶叶皮层在选择性视觉注意中起着重要作用 (Colby and Godberg, 1999; Corbetta and Shulman, 2002; Sereno and Amador, 2006; Ganguli et al., 2008),但在这些知觉决策实验中观察到的神经信号似乎不太可能完全归因于注意力效应。在做出选择之前,似乎应该把注意力集中在 RDM 刺激方案上,而不是目标上。即使在运动刺激和目标分散注意力的情况下,注意力如何解释神经动力学的时间过程也不清楚,即目标特定的上升活动,反映了决策的逐渐形成,并标志着这一过程的结束。另一种对上升活动的解释是运动意图,因为 LIP 特定地涉及眼跳的规划 (Andersen and Buneo, 2002)。检验这种可能性的一种方法是确定不管最终行为反应的形式如何,顶叶神经元的决策相关活动是否保持不变。

与决策相关的活动也出现在前额叶皮层,这是一个高级认知区域,在时间整合和其他认知功能中扮演着重要角色 (Fuster, 2008; Miller and Cohen, 2001; Wang, 2006a)。在 RDM 方向辨别任务的固定持续时间版本中,要求猴子在 1 s 的注视期内保持注视,刺激消失后是一段延迟期,然后猴子通过一个眼动扫视发出

自己选择的信号。与 RT 版本相反，现在记录的神经元的活动与决策(在观看运动时)相关，与运动反应(在延迟后)相关，可以暂时分离。前额叶皮层(Kim and Shadlen, 1999)和 LIP(Shadlen and Newsome, 1996, 2001; Roitman and Shadlen, 2002)区域的神经元显示出相似地活动模式：他们的活动反映了猴子的选择，但是，虽然它在运动观测期间是运动强度 c' 的一个分级函数，但在延迟期间的持续活动对 c' 变得不敏感。其含义是，受试者在观看刺激时做出决定，并在延迟期间主动维持二元选择，以指导稍后的行为反应。这些结果表明，决策和工作记忆可以由相同的皮层机制提供服务，可能位于大脑的顶额回路。值得注意的是，在 RDM 识别任务的反应时间版本中，前额叶神经元的记录尚未完成。在记忆延迟期，前额叶神经元通常表现上升活动，但这可能反映的是对即将到来的反应的预期和时间，而不是决策计算(Fuster, 2008; Chafee and Goldman-Rakic, 1998; Quintana and Fuster, 1999; Brody et al., 2003; Miller et al., 2003; Watanabe and Funahashi, 2007)。需要使用 RT 范式从前额叶皮层获取生理记录，以评估前额叶神经元是否以类似于顶叶神经元和顶叶神经元所观察到的方式显示出上升活动。如果是这样，上升活动是在一个区域本地产生(并在另一个区域反映)，还是通过两个区域之间的交互循环产生。

Romo 和他的合作者进行了一系列的实验，使用了不同的任务范式和感觉模式，这些实验为前额皮质参与感知决策提供了充足的证据。在体感延迟识别任务中，猴子报告了一个基于 f_1 和 f_2 两个机械振动频率比较的决定，这两个频率依次应用于指尖，间隔时间为 3-6 秒。因此，行为反应(信号 f_1 被认为是大于或小于 f_2)要求动物在延迟期间的工作记忆中保持第一个刺激的频率 (Figure 2A)。研究发现，在延迟过程中，前额叶皮层和运动前皮层的下凸性神经元表现出持续的活动，记忆活动的发放率随刺激频率单调增加(“正细胞”)或减少(“负细胞”) (Figures 2C and 2D)。在比较决策过程中，神经活性变为二元：当猴子选择 $f_1 > f_2$ 时，一个“正神经元”在实验中表现出高的放电，当猴子选择 $f_1 < f_2$ 时，表现出低的放电；而“负神经元”则表现出相反的趋势。在这个任务中，工作记忆先于决策，但是同样的回路参与两个过程。在任务的修改版本中，决策报告在比较期之后延迟几秒钟，研究发现，内侧前运动神经元以持续活动的形式保留了

猴子的选择和过去的感觉信息(Lemus et al., 2007), 强调了同一个回路同时参与工作记忆和决策的观点。Romo 的工作的一个显著成就是系统地探索了在同一任务中多个皮层区域(主要和次要躯体感觉区域, 运动前区和前额叶外侧区)的神经活动, 随着时间的推移, 这些区域的神经动力学产生了丰富的图像(Herna' ndez et al., 2002; Romo et al., 2004)。在一项使用近阈值振动触觉刺激的检测任务中, 报告了额叶在主观决策中发挥作用的额外证据。de Lafuente 和 Romo 的研究表明, 大脑额叶的前运动神经元的活动, 而不是主要的躯体感觉神经元的活动, 会随着逐次试验的主观报告(无论刺激是否存在)而共同变化(de Lafuente and Romo, 2005)。类似的探测心理测量函数是通过运动前皮层的误刺激或机械刺激获得的, 这表明受刺激的额叶部位可能与这种决策行为有因果关系。

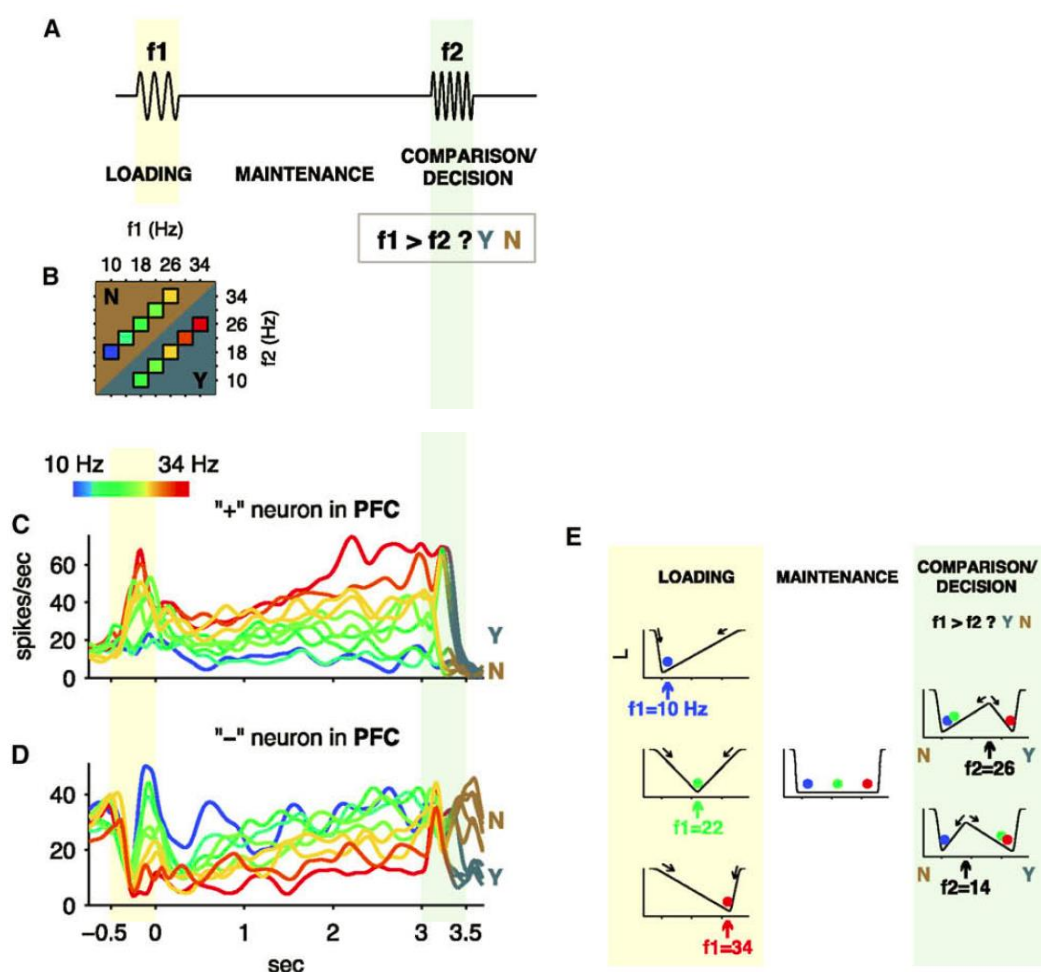


Figure 2. 延迟的触觉辨别任务及前额皮层的响应。

(A) 任务示意图，将两个频率分别为 f_1 和 f_2 的机械振动刺激依次(延时)作用于猴子指尖，受试者需判断 f_1 是否大于 f_2 。(B) 神经生理学研究中的典型刺激集。每个彩色框表示一个 (f_1, f_2) 刺激对。对于每对猴子，它们做出正确反应的几率都超过 91%。(C 和 D) 神经元反应。左上角的彩虹色代码表示在每种类型的试验中应用的 f_1 值。Y/N 色标表示每次试验结束时猴子按下的按钮。(C) 和 (D) 显示在许多试验中记录的两个不同 PFC 神经元的平滑放电率。(C) 为正单调(加)神经元，(D) 为负单调(减)神经元。(E) 二刺激区间辨别的一维动态算法。(E) 二刺激区间辨别的一维动态算法。横坐标: 状态变量(如 [C] 和 [D] 所示的正负神经元的放电率的差异)。纵坐标: 一个计算能量函数(与稳定吸引子状态对应的最小值)。在加载期间，第一个刺激产生一个独特的吸引子状态，位于水平轴上的一点，编码 f_1 值。在延迟期间，能量全局是平坦的，因此 f_1 的记忆在内部以参数持续活动的形式维持。在比较期间，再次对系统进行重构，第二个刺激 f_2 与内部记忆状态相互作用，产生分类决策 $f_1 > f_2$ 或 $f_1 < f_2$ 。转载自 Machens et al. (2005)。

随着事件相关电位测量的发展，人类对感知识别任务中 RTs 生理相关性的研究始于上世纪 80 年代(Gratton et al., 1988)。有趣的是，我们发现顶叶皮层区域记录的 P300 成分的峰间时间随着刺激识别性和 RT 的困难而增加，但与刺激-反应的兼容性无关(Kutas et al., 1977; McCarthy and Donchin, 1981)。然而，脑电图记录缺乏足够的空间分辨率来定位与时间整合密切相关的大脑区域。然而，脑电图记录缺乏足够的空间分辨率来定位与时间整合密切相关的大脑区域。在一项研究中(Binder et al., 2004)，研究人员要求受试者在嘈杂的听觉刺激中识别语音，发现听觉皮层的血氧水平依赖性(BOLD)功能磁共振成像(fMRI)能够反映出信号噪声比(感觉信息的质量)。而额叶皮层随反应时间(决策过程指标)呈线性增加。在另一项研究中(Ploran et al., 2007)，研究对象通过 8 个步骤(每个步骤 2 秒)逐步观看空白屏幕上显示的图片，随着识别时间的增加，活动逐渐增加的区域是顶叶、额叶和颞下叶区域。此外，最近的一项研究表明，在大脑额叶和顶叶皮质进入意识状态之前的很长一段时间(秒)内，一个自由的运动决策可以通过 BOLD 信号进行解码(Soon et al., 2008)。

因此，越来越多来自人类神经成像和猴子单神经元生理学的证据表明，顶叶和额叶皮质形成了一个核心的大脑系统，用于在感知决策任务中数据的暂时积累和分类选择的形成。这些区域也许提供了到感觉神经元的自上而下的信号，它的尖峰活动通常表现为与猴子的选择相关的弱反复试验的可变性(Britten et al., 1996; Shadlen and Newsome, 1996; Parker and Krug, 2003; Law and Gold, 2008)。

未来工作的一个挑战是准确阐明顶叶和额叶回路如何协同工作，在决策过程中可能扮演不同和互补的角色。一个决策可能涉及到跨大脑区域的并行处理，而不是严格按照串行阶段进行。在一个不同的大脑皮层区域同时进行行动选择和反应准备的模型中，已经提出了这种情况(Cisek, 2006)。然而，经验证据很少。解决这一突出问题的一个有希望的方法是，通过同时记录行为动物两个或两个以上区域的单个单元和局部场电位来检查区域间的相互作用(Pesaran et al., 2008)。

循环皮质回路机制

决策计算是如何在皮层电路中实例化的？皮质区域中的与决策相关的神经活动已经被报道，这些区域展现出在工作记忆维护中的记忆持续活动。例如，在一项眼动延迟反应任务中，LIP 和前额皮质的神经元都显示出定向调谐的持续性活动(Gnadt and Andersen, 1988; Funahashi et al., 1989)。基于这一观察，有人提出，这不仅仅是巧合，而是一种共同的回路机制，是决策和工作记忆的基础(Wang, 2002)。产生持续性活动的一个主要候选机制是在局部皮质回路中产生强烈的周期性兴奋，这导致了刺激选择性吸引子状态-神经网络的自我维持的群体活动模式(Amit, 1995; Goldman-Rakic, 1995; Wang, 2001)。这样的吸引子网络模型也能解释决策计算吗？

为了解决这个问题，一个最初为工作记忆开发的生物物理模型(Brunel and Wang, 2001) 应用到 RDM 辨别实验进行仿真(Wang, 2002)。值得强调的是，这个

局部回路模型强调的是前额叶和顶叶区域的共同特征,并没有涉及记忆或与决策相关的神经活动是在其中一个区域还是同时在这两个区域产生的问题。“基于生物物理的模型”一般是指具有解剖学上合理结构的模型,在这种模型中,不仅单个尖峰神经元以合理的精确度进行生物物理描述,突触的相互作用也通过定量神经生理学(结果证明这是非常重要的)进行校准。

图 3 说明了这种递归神经电路模型(Wang, 2002; Wong and Wang, 2006; Wong et al., 2007)。在该模型的两池版本中,尖峰神经元的亚群对两种选择方案具有选择性(例:A =左运动, B =右运动)。在每个锥体神经群中,强的周期性兴奋性连接可以维持由短暂的偏好刺激触发的持续性活动。这两个神经组通过神经元间的反馈抑制进行竞争。相互冲突的感官输入被传送到回路中的两个神经池中,运动强度 c' 用输入的相对差值来实现(图 3)。图 3B 为零运动强度仿真。刺激出现时,两个竞争神经群体发射率(r_A 和 r_B)最初在数百毫秒一起增加,然后分离:一个增加(由于在神经池中的循环性兴奋),而另一个下降(由反馈抑制导致的“winner-take-all”)。感知选择是根据两个神经群体中哪一个获胜来决定的。随着 c' 的变化,感官数据质量越高,神经元活动上升越快(图 3 C)。该模型捕捉了行为猴子 LIP 细胞活动的重要特征。首先,神经活动主要与决策选择相关(即使在错误试验或运动强度为零时)。其次,神经决策过程分为两个步骤:感官数据首先以一种分级的方式随着时间的推移进行整合,然后是 winner-take-all 的竞争引起二元选择。第三,在刺激退出后,网络将决策选择以对 c' 不敏感的持久化活动的形式存储在工作记忆中。

“吸引子场”可以在决策空间中进行说明, 绘制 r_A 与 r_B 的相空间(图 3D)。在这个例子中,感官证据支持选择 A,所以吸引子 A 比吸引子 B 有更大的吸引池(橙色)。系统最初在吸引 A 的吸引池中,并在正确的试验中向决策状态 A 演化(蓝色)。但在低 c' 时,偏差不强,噪声会诱导系统的轨迹穿越两个吸引池的边界,此时系统最终演化为错误试验(红色)中的决策状态 B。吸引池之间的边界穿越是缓慢的,这就解释了为什么在错误试验中反应时间比在正确试验中要长,就像在猴子实验中观察到的那样(Roitman and Shadlen, 2002)。这种决策空间分

析强调了一点，即系统不是刚性的，而是对外部信号作出灵活反应的。吸引子状态可以由输入产生或消除；因此，相同的网络可以提供不同的功能，例如，在固定持续时间的 RDM 任务中，在刺激呈现期间进行决策，然后在延迟期间使用活动记忆进行选择。这一结论得到了 RDM 实验最近其他建模研究的支持 (Roxin and Ledberg, 2008; Grossberg and Pilly, 2008)。

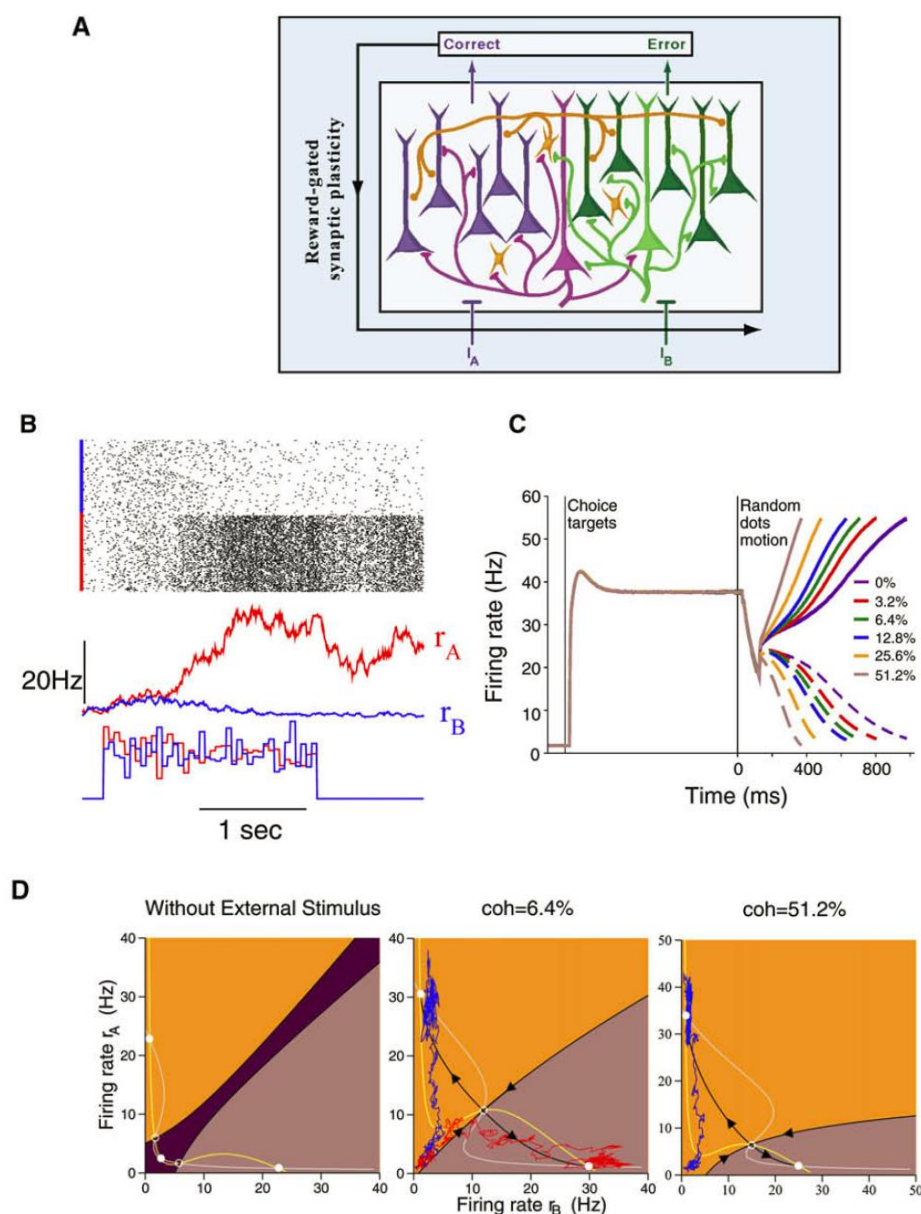


Figure 3. 一种大脑皮层回路模型，用于两种不同的强迫选择任务。(A) 模型方案。在视觉运动识别实验中，随机运动的点有两个(紫色和绿色)锥体细胞亚群，分别对两个方向(A 或 B)有选择性。第三个(橙色)神经亚群代表抑制性间神经元。

这三个亚群中的每一个都由几百个尖峰神经元组成。在每个选择性亚群的锥体细胞之间,该回路具有很强的周期性兴奋(由 AMPA 和 NMDA 受体介导)和竞争性抑制(GABA_A 受体介导)。运动一致性表示为 $c' = (I_A - I_B)/(I_A + I_B)$,其中 I_A 和 I_B 为平均输入。对于非零 c' ,一种选择是正确的,另一种是错误的,其结果可能导致某些输入突触的奖励依赖性可塑性改变。(B)一致性为零的尖峰网络模型的例子仿真。自顶向下:网络尖峰光栅,种群激发率 r_A 和 r_B ,随机输入。注意 r_A 和 r_B 最初的缓慢提升(时间整合)和最终的分歧(分类选择)。(C)不同运动强度的试验平均神经活动,包括目标输入。实体曲线:获胜的神经群体。虚线:失败的神经群体。注意 RDM 刺激开始时的短暂下降。(D)在一个固定持续时间版本的任务中,在 $c' = 6.4\%$ (中间组)或 51.2% (右边组)的运动刺激存在的情况下,不需要外部输入(左侧组),两个选择性神经种群的相平面图。在没有刺激的情况下(左面板),三个吸引子共存(白色圆圈):自发状态(当 r_A 和 r_B 都很低时),和两个持久的活动状态(高 r_A 和低 r_B 或相反)。当刺激的呈现时(中间面板 $c' = 6.4\%$),吸引子格局改变,自发的稳定状态消失,因此,系统被迫向代表知觉决策的两个活跃状态之一前进(A或B),如图所示的两个试验中的网络轨迹(蓝色和红色)。在刺激消失之后,系统的配置将恢复到左侧面板中的配置。因为持续的活动状态是自我维持的,所以知觉选择(A或B)可以储存在工作记忆中,以供以后使用,指导行为。彩色区域对应于共存吸引子状态的吸引池。在无噪声的情况下,从平面一个点出发的系统收敛到相应的吸引子状态。注意,正确选择状态的区域在高运动强度(右面板)比低(中面板)运动强度大得多。(A)及(B)经Wang (2002)允许转载,(C)来自Wong et al. (2007), (D)采用Wong and Wang (2006)的模型计算。

如上例所示,吸引子网络不仅限于稳态行为,还可以利用长瞬态来进行有趣的计算。正如最初提出的工作记忆,慢反射的一个候选细胞底物是局部循环性兴奋突触上的 NMDA 受体(Wang, 1999)。 $\tau_{network} = \tau_{syn}/(|1 - W_{rec}|)$ (其中 τ_{syn} 为突触时间常数, W_{rec} 为复发性连接强度)给出了网络时间常数的简单估计,当 W_{rec} 存在时, $\tau_{network}$ 的长度大于 τ_{syn} (Seung, 1996; Wang, 2001)。例如,如果 NMDA 受体介导的突触 $\tau_{syn} = 100 \text{ ms}$ (Hestrin et al., 1990), $W_{rec} = 0.9$,则 $\tau_{network}$

$= 1 \text{ s}$ 。因此，该网络在 1 秒的时间尺度上显示了提升活动的瞬时动态，而这种能力严格依赖于 NMDA 受体(Wang, 2002; Wong and Wang, 2006)。注意这个机制依赖于 NMDA 受体介导的通道的缓慢动力学并强调了 NMDA 受体对在线认知的重要性(而不是众所周知的在长期突触可塑性中的作用)。其他缓慢的生物物理机制，如短期突触促进(Abbott and Regehr, 2004)或单细胞钙依赖过程(Major and Tank, 2004)，也可能有助于时间整合。这些候选方案，所有的正反馈机制，都可以用行为正常的猴子(使用药理学手段)以及啮齿动物和其他更简单的动物系统进行实验测试。

循环神经回路模型仅仅是扩散模型的一种实现吗？有趣的是，与一维扩散模型相比，“决策空间”分析(图 3D)表明吸引子神经网络的动力学本质上是二维的(Wong and Wang, 2006)。这与 RDM 实验记录的 LIP 数据(Roitman and Shadlen, 2002)中发现的两个选择性神经池内的动力学均以慢模式为主的结论是一致的。因此，描述两个相互竞争的神经池需要两个动态变量。需要一个双变量模型来解释观察到的 LIP 神经活动呈现出一个双相的时间进程，即神经元对两个相反的目标的选择性在彼此分离之前首先一起加速(Roitman and Shadlen, 2002; Huk and Shadlen, 2005)。同样的行为也出现在一个自由运动决策任务中(Scherberger and Andersen, 2007)。

此外，重要的是，扩散模型和循环神经回路模型在行为层面上有不同的预测。首先，循环回路模型在错误试验中比在正确试验中产生更长的响应时间(Wong and Wang, 2006)，与猴子实验一致(Roitman and Shadlen, 2002)。相比之下，扩散模型的神经实现产生了相反的效果(Mazurek et al., 2003)。误差试验中较长的 RTs 可以在扩散模型中实现，额外的假设是起始点随试验随机变化(Ratcliff and Rouder, 1998)。其次，扩散模型从未达到稳定状态，它预测，随着刺激观察时间的延长，性能可能会无限期地提高，比如提高决策界限。在循环电路模型中，当达到吸引子状态时，上升活动最终停止(图 3D)。因此，在足够长的观看刺激的时间里，性能会停滞不前(Wang, 2002)。最近的猴子实验证实了这个模型的预测(Kiani et al., 2008)。第三，吸引子网络模型已经被证明能够

减去消极信号以及增加关于选择备选方案的积极证据，但是随着时间的推移，随着网络收敛到代表备选方案的吸引子状态之一，新输入的影响会减弱(Wang, 2002)。猴子实验也证实了这一预测，实验表明，随着发病时间的提前，除了随机点刺激外，短暂的运动脉冲的影响更大(Huk and Shadlen, 2005; Wong et al., 2007)。这种违反时移不变性的行为不能用泄漏来解释。事实上，与循环回路模型相比，LCA 模型实际上预测的是更晚的信号，而不是更早的信号对最终决策的影响更大，因为更早的脉冲是逐渐形成的，由于泄漏而被“遗忘”，并且对稍后发生的决策没有显著影响(Wong et al., 2007)。

周期性的兴奋必须通过反馈抑制来平衡(Brunel and Wang, 2001)。扩散模型假设在输入水平上计算了关于冲突输入的差分信号(Ratcliff, 1978; Mazurek et al., 2003)。这一观点在人类 fMRI 实验中得到了严格的验证。在这项实验中，研究人员的任务是辨别一幅模糊的图像是一张脸还是一栋房子。研究发现，前额叶背外侧皮层的 BOLD 信号与腹侧颞叶皮层的面部和房屋选择区域的差异信号是共变的(Heekeren et al., 2004)。这一工作表明，大脑的某些区域可能对绝对不同的信号进行辨别时会编码不同的信号。在同一方面对不同的选择加以区别，情况可能会有所不同，比如左右运动方向，这很可能发生在一个局部网络中。在循环回路模型中，可选选项的神经池选择之间的竞争是通过侧突触抑制来体现的(Wang, 2002; Wong and Wang, 2006)。这个反馈机制不是前馈差值，而是通过观察 LIP 中一个神经池的微刺激来支撑的，这个反馈机制不仅加速了其偏好方向的决策，而且减缓了其反偏好方向的决策(Hanks et al., 2006)。在另一项相关分析中，Ditterich (2006)发现扩散模型产生了右尾较长的反应时间直方图(反映异常长的 RTs)，这与猴子实验中的观察不一致。横向抑制的加入加重了这一问题，导致右尾更长，尤其是在低一致性水平下。在循环神经回路模型中并非如此，该模型产生的决策时间分布不显示明显的右尾，与猴子实验中观察到的类似(X.-J. Wang, 2006, Soc. Neurosci., abstract)。非线性吸引子模型的一个显著特征是强重复兴奋，这是线性累加器模型所没有的。正反馈机制最终导致神经活动向决策边界加速，防止决策时间过长。Ditterich 的研究表明，利用累加器模型可以很好地拟合猴子的反应时间分布，并假设决策界随时间的增加而减小。

这在功能上类似于随时间增长的“紧急信号”(T.D. Hanks et al., 2007, Soc. Neurosci., abstract)。同样地,期望的效果可以通过一个暂时递增的上升坡度来实现,这在循环回路模型中自然发生,不需要额外的假设。另一方面,人类研究通常报告有右长尾的倾斜 RT 分布,这被扩散模型很好地捕捉到(Ratcliff, 1978; Luce, 1986; Ratcliff and Rouder, 1998; Usher and McClelland, 2001; Sigman and Dehaene, 2005),但在循环神经回路模型中不存在。在动物和人类的研究中,识别 RT 分布是否有突出尾巴的条件,并对这种现象提出神经机制的解释,将是非常有趣的。

在躯体感觉辨别实验中,也建立了循环回路模型(Romo et al., 2002, 2004)。Miller 等人(Miller et al., 2003)的研究表明,对该模型的连通性进行微调,可以得到一个具有参数化工作记忆能力的吸引子,类似于 gazecontrol 系统中的吸引子(Seung et al., 2000)。此外,两个相互抑制耦合的神经群体表现出持续的活动,分别为正的和负的,单调地调谐到第一频率 f_1 ,正如在这项任务中在前额神经元中观察到的那样(Romo et al., 1999; Brody et al., 2003; Miller et al., 2003)。因此,该回路能够在整个延迟期间存储 f_1 。Machens et al. (2005)等人结果表明,在比较期间(图 2E),该回路也可以进行辨别计算($f_1 > f_2$ 或 $f_1 < f_2$),为传入的刺激提供了一个转换机制,使得第二刺激到达记忆/决策回路,与第一刺激的调谐符号相反。在另一种情况下,积分反馈机制可以在不需要输入切换的情况下实例化比较计算(Miller and Wang, 2006a)。通过相平面分析, Machens 等人展示了吸引子曲面是如何通过外部信号为每一个任务的不同时期进行不同地重新配置的(提示加载、记忆维护、比较);因此,同一回路可以同时提供工作存储器和决策计算(Machens et al., 2005)。该模型预测了相同调谐(正单调还是负单调)神经对之间的正相关,相反调谐的神经对之间的负相关。从前额皮质记录的神经数据证实了这一预测。有趣的是,与 RDM 的固定持续时间版本的辨别相比,在这个任务中,决策要先于延迟响应(因为这个延迟响应需要二元选项的工作记忆),在振动触觉识别任务中,模拟量(f_1)的参数工作记忆先于双选择决策过程。然而,这两种行为任务的模型显示出惊人的相似性(包括相平面图)。近阈值检测实验(de Lafuente and Romo, 2005)也采用了相似的吸引子

网络模型 (Deco et al., 2007a)。综上所述，这些结果进一步支持慢反射动力学作为工作记忆和决策的一般机制。

决策过程的终止规则

我们如何准确地知道一个逐步积累过程何时结束，然后形成一个范畴决策？在 RT 任务中，可以从测量到的反应时间(减去运动反应潜伏期)来推断出决策时间，这已被证明与 LIP 神经活动的阈值交叉有关 (Roitman and Shadlen, 2002)。如果是这样，那么这种决定阈值的生物基础是什么？答案可能在下游神经元。一种可能的情况是，当决策神经元整合输入并达到特定的发放率水平时，该事件会在下游神经元中触发全或无响应，并导致行为输出的产生。在动眼神经任务中，一个自然的候选者是前额眼 (FEF) 和上丘 (SC) 中的运动神经元，是大脑中为选择，准备和初始化眼跳运动的必不可少的两个区域。这些神经元对眼跳的幅度和方向是有选择性的，在眼跳开始之前，它们会立即发出一串典型的尖峰脉冲 (Hanes and Schall, 1996; Munoz and Fecteau, 2002)。而在这里我们关注的是响应执行，FEF 和 SC，像 LIP 一样，也涉及动眼决策和响应选择的其他方面。

为了测试这个场景的决策阈值 (Lo and Wang, 2006)，我们考虑了一个扩展的多回路模型 (Lo and Wang, 2006)。皮层中的决策神经元(如上所述)投射到 SC 中的运动神经元 (Figure 4A)。该模型还包括基底神经节的一条直接通路，其中输入层(尾状核，CD)和输出层(黑质网状部，SNr)在控制自主运动中发挥着重要作用 (Hikosaka et al., 2000)。当皮层中的神经元随着时间上升时，到 SC 运动神经元池和 CD 神经元池中的突触输入也随之增加。在这个场景中，一个决策阈值(作为决策神经元放电率的一个界限)被突触输入的硬阈值实例化，用于触发下游运动神经元中的一个特殊事件。图 4B 为视觉运动方向辨别实验的模型仿真示例试验。由于大脑皮层的随机放电动力学，活动活动的速率在不同的试验中不断波动，并且在反复试验的基础上，爬坡活动的速率与决策时间(定义为 SC 中触发突发事件的时间)成反比(图 4C)。此外，当任务越困难(运动一致性越低)，上升活动越慢，导致反应时间越长。然而，被下游运动系统读出的皮质放电活动的阈值有相同的狭窄分布(图 4C 的插图)。不管上升速度或反应时间。因此，反应时间

的可变性主要归因于神经活动本身的不规则上升，而不是决策界的试验到试验的可变性。这个模型定量地再现了猴子的行为表现和反应时间(图 4D)。

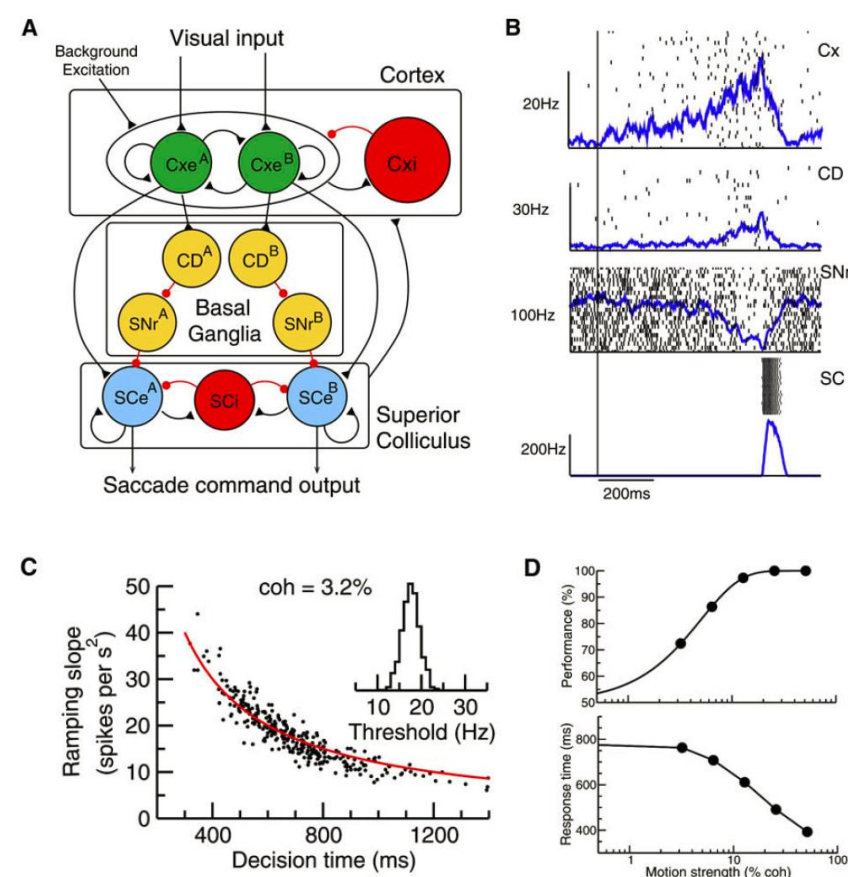


Figure 4. 用于两个可选的强制选择任务的机制的多模块网络。 (A) 示意图模型结构。皮层网络中的神经池整合了感觉信息，并相互竞争。它们投射到基底神经节的上丘(SC)和尾状核(CD)。CD向黑质网状部(SNr)发送抑制性投射，SNr通过抑制性突触与SC中的运动神经元连接。每个群体都由嘈杂的尖峰神经元组成。(B) 模型的单一试验模拟，显示单细胞的发放序列和Cxe、SNr、CD、SCe的群体放电率。当运动神经元的突触输入超过阈值水平时，就会触发运动神经元的簇放电，这是由皮质神经元的直接兴奋和皮质纹状体投射解除SNr抑制的结果。时间0对应于刺激开始。(C) Cxe发放率的上升与逐次试验的决策时间成反比(每个数据点对应一个单独的试验)，红色曲线是 $12,000 / (\text{决策时间})$ 。(D) 性能(正确选择的百分比)和平均响应时间作为运动相干性 c' 的函数。转载自 Lo and Wang (2006)。

该回路能自适应调整决策阈值吗？在速度-精度权衡中，阈值过低会导致更快的响应，但会导致更多的错误，而阈值过高会提高精度，但会延长响应时间。这些都不会产生最大的回报。一种普遍的观点是，通过自适应地调整决策阈值可以实现最优 (Gold and Shadlen, 2002; Bogacz et al., 2006)。由于在神经回路模型中，决策阈值被定义为引起下游 SC 神经元簇放电响应所需的最小皮质放电，因此可以预期，这个阈值可以通过皮质-丘体通路的可塑性变化来调整：随着突触强度的增强，同样水平的皮质输入到上丘可以用较少的皮质神经元的放电来实现。有趣的是，当系统受到基底神经节的控制时，情况并非如此。这是因为 SNr 中的神经元通常以高频率进行调理激发(图 4B)，并对 SC 运动神经元提供持续抑制(Hikosaka et al., 2000)。为了让 SC 神经元产生脉冲输出，这种抑制必须被释放(因为皮质的里的上升激活了CD神经元,反过来又抑制了SNr的活动)。这种高度非线性的去抑制机制意味着，通过调节皮质-纹状体通路的突触强度，比通过调节皮质-上丘突触强度的变化更容易调节决策阈值(Lo and Wang, 2006)。考虑到皮质-纹状体突触是多巴胺神经元支配神经的一个重要目标，这一发现尤其具有吸引力。考虑到皮质-纹状体突触是多巴胺神经元支配神经的一个重要目标，这一发现尤其具有吸引力。鉴于多巴胺神经元发出奖赏或奖赏预测误差的信号(Schultz, 1998)，皮质-纹状体突触的长期增强和抑制依赖于多巴胺信号(Reynolds et al., 2001; Shen et al., 2008)，我们的工作表明，皮质-纹状体突触依赖多巴胺的可塑性代表了大脑中决策阈值自适应调整的候选神经位点，这一预测在实验中是可以验证的。更一般地说，决定阈值(根据神经活动界限从神经生理学上定义的)或其他一些属性，如假定的决定神经元的开始时间或斜坡坡度，是否在速度和准确度之间进行了动态调整，还有待观察。

尽管决策阈值是 RT 任务中的一个关键元素，但它不应该等同于终止累积过程的一般决策规则。这一点在 RDM 任务的固定持续时间版本中得到了清晰的说明。在 RDM 任务中，观看时间由外部控制，在刺激消失(Britten et al., 1992; Kiani et al., 2008)或记忆延迟期结束之前(Shadlen and Newsome, 2001; Roitman and Shadlen, 2002)，受试者必须避免做出明显的反应。受试者的 LIP 神经元在整个刺激呈现期(只要足够长)似乎没有整合感官信息，这表明在这些情况下，阈

值跨越可能仍然是终止决策过程的规则(Kiani et al., 2008)。这样的场景需要一个读出系统(检测阈值跨越事件)来发送反馈信号,以停止决策神经元中的集成过程。而在循环神经回路模型中,当系统达到稳态时积分自然停止。这在不触发明显的行为反应的情况下会发生,可能是因为下游运动神经元(在 FEF 和 SC 中)受到外部信号(例如注视信号)和/或内部产生的控制信号的抑制。从概念上讲,循环神经网络表明类别选择是由哪一个吸引子赢得竞争而自然决定的。响应时间是根据运动系统读取决策电路的神经信号的时间来解释的,在响应时间任务或固定持续时间任务中,它可以被灵活地调整 and 不同地控制。

这一普遍观点不仅适用于知觉决策,也适用于行动控制。事实上,一种循环电路方法已经被用来为动作选择和动作准备的建立模型(Wilimzig et al., 2006; Cisek, 2006; Heinzle et al., 2007)。运动的时间,甚至最终是否产生反应,都是由抑制过程有效控制的,比如通过控制神经元(Hanes and Schall, 1996; McPeck and Keller, 2002; Narayanan and Laubach, 2006; Boucher et al., 2007)及基底神经节(Hikosaka et al., 2000)来抑制运动神经元的门控。因此,至少在某些情况下,一个行为的选择可能不是偶然地与反应时间联系在一起的。在一种流行的行为抑制控制模型中,“走”过程和“停”过程之间存在着一种竞争,先跨越某个阈值的进程赢得竞争,并决定响应是否受到抑制(Logan and Cowan, 1984; Boucher et al., 2007)。在撤销行动的循环神经回路模型中,该决策被描述为双稳态动力学,类似于两种强迫选择感知决策,除了吸引子状态对应于被取消和非被取消的响应(C.C. Lo and X.-J. Wang, 2007, Soc. Neurosci., abstract)。这种观点认为,决策过程的结果在一定程度上对决策阈值的精确值不敏感,例如,将阈值提高到一定的水平并不一定会提高性能。因此,分类选择可以从决策阈值中分离出来。

基于价值的经济选择

做出“正确”的决定最终是为了实现一个行为目标。在实验室实验中,目标通常是获得最大的回报。这种类型的决策依赖于大脑评估可用选项的可取性的能力,这是选择的先决条件,并在选择结果不符合预期时自适应地改变决策策略。这个

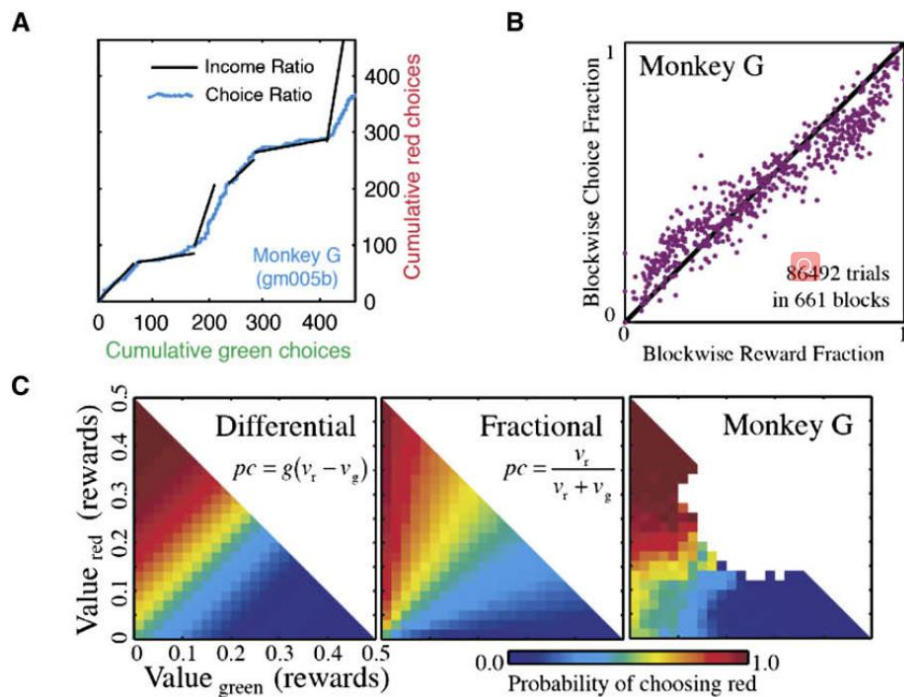
领域是强化学习理论和神经经济学的融合，是最近几篇综述的主题 (Sugrue et al., 2005; Rushworth and Behrens, 2008; Loewenstein et al., 2008; Soltani and Wang, 2008)。

在奖赏信号和奖赏期望的神经表征方面我们已经取得了重大进展。一项开创性的发现是，腹侧被盖区 (VTA) 多巴胺神经元的阶段性活动编码了一个奖赏预测错误 (预期奖赏与实际奖赏之间的差异) (Schultz et al., 1997; Schultz, 1998; Roesch et al., 2007)。与预测误差信号一致，多巴胺神经元的尖峰活动随着奖励幅度和概率的增加而增加 (Fiorillo et al., 2003; Tobler et al., 2005; Roesch et al., 2007)。此外，最近的一项研究发现，灵长类外侧系带的神经元反映奖励预测错误，其信号与多巴胺神经元相反 (当实际奖励小于预期结果时，表现出强烈的尖峰活动)。说明外侧系带是负预测误差信号的来源 (Matsumoto and Hikosaka, 2007)。另一方面，在记录动物行为的单个单元记录中，纹状体、杏仁核、眼窝额叶皮层 (OFC) 和前扣带皮层 (ACC) 一直观察到与奖赏期望相关的神经信号 (Rushworth and Behrens, 2008 中的综述)。期望值通常被描述为经验奖励的泄露积分因子。例如，对行为正常的猴子的 ACC 的神经刺激被描述为在几次试验 (或几十秒) 的时间尺度上对过去奖励的时间过滤器 (Kennerley et al., 2006; Seo and Lee, 2007)。在强化学习理论中，错误信号被假设为更新奖赏期望，而奖赏期望又被用来计算错误信号 (Sutton and Barto, 1998; Bayer and Glimcher, 2005; Rushworth and Behrens, 2008)。因此，奖赏期望和预测误差是相互依赖的，必须迭代计算，可能在不同的大脑区域以一个相互循环连接。例如，通过依赖多巴胺能输入的学习过程，包括纹状体和额叶皮质在内的回路可能会评估奖赏期望。然后将这个信号反馈给中脑多巴胺细胞，与实际奖励进行比较，得出一个预测错误。然而，奖赏期望和预测误差信号在大脑的多个区域是混合的，往往很难理清。

区分奖励评估的大脑系统和使用这些信息指导选择行为的神经回路是很有用的。在决策过程中被激活并被奖赏信号调节的大脑结构包括尾状核 (Samejima et al., 2005; Hikosaka et al., 2006)，外侧顶叶皮质 LIP (Platt and Glimcher,

1999; Sugrue et al., 2004)和前额皮质(Watanabe, 1996; Roesch and Olson, 2003)。强化学习模型表明, 行为值是在决策回路中的神经元突触上习得的, 从而影响选择行为(Seung, 2003; Worgotter and Porr, 2005)。为了说明这一点, 考虑一个如图2A所示的神经网络。回想一下, 网络行为是由softmax决策准则描述的, 也就是说, 选择A和B的概率是两个相互竞争的神经池的输入(ΔI)差的一个sigmoid函数(图4D, 上)。假设两个突触连接 c_A 和 c_B 的强度是可塑的, 那么突触的修饰会随着时间的推移改变网络的决策行为。具体地说, 我们使用了具有随机Hebbian学习规则的双突触, 即突触可塑性依赖于突触前和突触后神经元的共激活, 并且发生的概率性(Fusi, 2002; Fusi et al., 2007)。此外, 基于多巴胺信号可以启动纹状体(Wickens et al., 2003; Shen et al., 2008)和前额皮质(Otani et al., 2003; Matsuda et al., 2006)突触可塑性的观察, 我们认为突触学习依赖于奖赏信号。例如, 决定神经元输入的突触只有在选择得到奖励时才会增强, 反之则会减弱。因此, 在学习过程中, 突触获得了关于所选择反应的奖赏结果的信息, 即特定行为的价值观。由于突触修饰的结果, 决策网络中相互竞争的神经群的输入强度在不同的试验中有所不同, 从而导致选择行为的自适应动力学。该模型通过将其应用于觅食任务中进行测试, 在觅食任务中, 研究对象在不可预测的环境中自适应地做出一系列选择。在猴子实验中(Sugrue et al., 2004; Lau and Glimcher, 2005, 2008), 根据一个同步可变区间强化计划, 奖励分别以诱导率 λ_A 和 λ_B 随机分配给两个(A和B)响应选项。在该计划中, 选择的目标随机地被奖励诱导, 并保持诱导状态, 直到受试者选择目标并收集奖励。他研究发现, 猴子的行为符合匹配法则, 即被试按照与从这些选择中获得的相对强化相匹配的比例分配自己的选择(Herrnstein et al., 1997) (图5A和5B)。此外, 对眼跳反应具有选择性的LIP神经元的神经活动是通过结果值的表示来调节的, 结果值在行为上被定义为对该目标的过去奖励的泄漏整合(Sugrue et al., 2004)。有趣的是, 猴子的选择行为与两种收益差异的软最大值函数吻合得很好(Corrado et al., 2005) (图5)。这些行为和神经生理学观察到的结果在具有奖励依赖可塑性的决策神经回路模型中得到了重现(Soltani and Wang, 2006) (图5D–5G)。结果发现, 在模型中, 突触强度(c_A 和 c_B)与两个目标的回报(每个选择的平均回报)成比例, 而不是与收益(每个试验的平均回报)成比例(图5D)。请注意,

匹配意味着两个返回是相等的，因此根据返回对奖励值进行编码特别适合于匹配计算。此外，由于突触随时间随机增强或减弱，它们会健忘，就像过去选择结果的一个泄露的积分器，时间常数由学习速率和环境中的奖励统计数据决定 (Soltani and Wang, 2006)。因此，决策行为受到过去局部获得的奖励的影响，这与观察到的猴子的行为一致 (Sugrue et al., 2004; Lau and Glimcher, 2005)。正如LIP中所观察到的，在模型中，神经元被响应选项的值所调节 (图5E)，尽管它们并不直接负责评估本身。该模型再现了匹配行为：随着奖励率 λ_A/λ_B 从一个试验块到下一个试验块的变化，模型s的行为变化很快，以致于选择A和B的概率近似匹配 λ_A/λ_B (图5)。此外，该模型还解释了在猴子实验中的观察：匹配并不完美，选择更有回报的选项的相对概率略小于相对奖励率（“不匹配”） (图5G)。模型分析表明，不匹配是一种网络动力学波动的自然结果 (Soltani and Wang, 2006)。如果没有神经系统的可变性，决策行为往往会被更有益的选择所困；随机的尖峰活动使网络更具探索性，并导致不匹配。



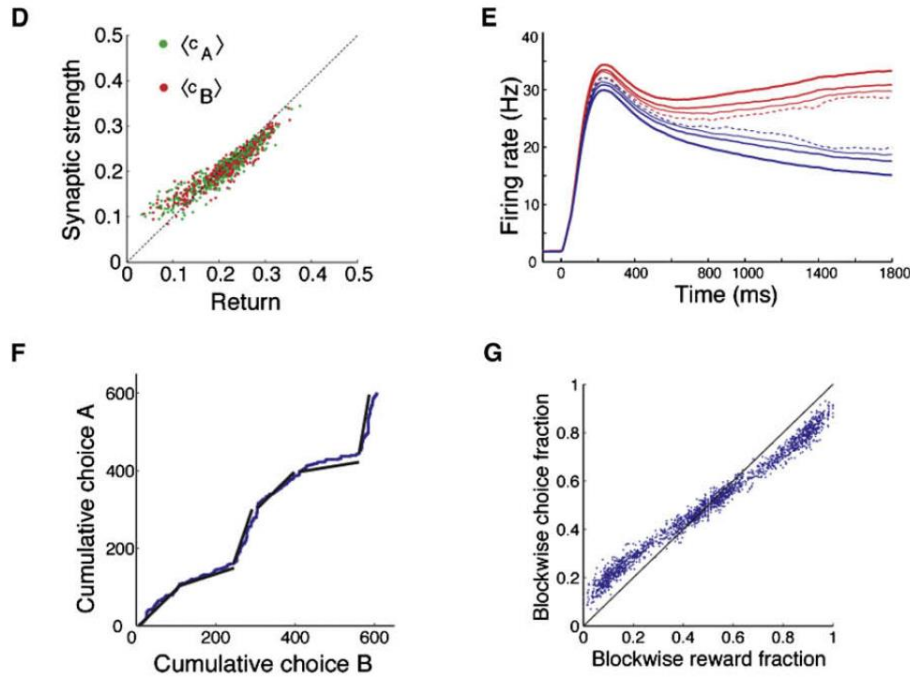


Figure 5. 觅食行为匹配规律的神经基础。 (A) 猴子在单一实验时段的动态匹配行为。连续的蓝色曲线显示了红色和绿色目标的累积选择。黑线表示每个区块的平均收入比率(红色:绿色) (这里, 1:1, 1:3, 3:1, 1:1, 1:6, 6:1)。匹配预示着蓝色和黑色曲线是平行的。(B) 块匹配行为。每个数据点代表一个试验块, 在这个试验块里每个目标的诱惑概率保持不变。红色目标的奖励和选择分数显示出来(绿色目标的奖励分数减去红色目标的奖励分数)。完全匹配对应于对角线上的数据点。当奖励概率大于 0.5 时, 选择概率低于奖励概率, 偏差(不匹配)明显。(C) 在线性-非线性模型中, 过去的奖励在以前的试验中加以整合, 过滤时间常数约为 5 至 10 次试验, 得到两个目标的估计值 v_r 和 v_g 。选择概率作为 v_r 和 v_g 的函数被建模为 softmax 规则(左面板)或分数规则(中面板)。利用 softmax (sigmoid) 决策准则(右面板)对猴子的行为数据拟合较好。(D) 在具有奖赏依赖性可塑性的循环神经回路模型中(图 3A), 平均突触强度是每个选择的回报的线性函数(每个选择对一个目标的回报概率)。红色和绿色数据点分别为突触强度 c_A (红色目标)和 c_B (绿色目标)。(E) 两种选择性神经群体神经元的分级活性。如果根据选择和突触强度的差异, 对单次试验的放电率进行排序和平均, 则决策神经元的活动呈现出一种分级模式。活动是由两个目标的开始排列的, 对于神经元的首选(红色)或非首选(蓝色)目标的选择, 活动是分开显示的。此外, 根据突触强度编码到两个相互竞争的神经群体上的值的差异, 将试验再分为四组 ($c_A - c_B = -0.05$ 到 -

0.14 [虚线], 0 到 -0.05 [淡的], 0 到 0.05 [常规的], 0.05 到 0.14 [重的])。 (F) 对某一时段的模型模拟觅食实验, 绘制目标 A 的累积选择与目标 B 的累积选择 (蓝色) 的对比图。黑色直线表示各区块的诱饵概率比。在猴子的实验中使用了相同的诱饵概率比率 (A)。 (G) 每一点都显示出, 在保持诱饵概率不变的一组试验中, 区块选择分数是区块奖励分数的函数。该模型再现了匹配行为和欠匹配现象。 (A) 转载自 Sugrue et al. (2004), (B) 和 (C) 来自 Corrado et al. (2005), (D) - (G) 来自 Soltani and Wang (2006)。

同样的模型也适用于竞争游戏, 即多个决策者根据收益矩阵进行交互。具体来说, Barraclough 等人 Barraclough et al. (2004) 的实验进行了模型模拟, 在实验中, 猴子与电脑对手进行硬币匹配。该模型再现了显著的行为观察结果 (Soltani et al., 2006)。与猴子的行为类似, 当对手按照匹配硬币的规则进行完全交互时, 模型行为变为准随机。例如, 如果初始 c_A 大于 c_B , 系统更频繁地选择目标 A, 则会被对手利用。此外, 选择 A 而没有得到奖励的结果会导致神经池 A 的突触抑郁, 以致于随着时间的推移, 差值 $c_A - c_B$ 逐渐减小, 系统会越来越频繁地选择 B (Soltani et al., 2006; Lee and Wang, 2008)。因此, 虽然决策神经元的活动依赖于反应选项的值, 但评估可能发生在其他地方, 也许在突触水平。当结果 (奖励与否) 在行为反应之后很久才显现出来时, 这种学习规则是如何起作用的, 还有待观察, 它是通过整合连接两个事件之间的时间间隔的持续性神经活动, 还是通过突触中的“合格性追踪” (Sutton and Barto, 1998; Seung, 2003; Izhikevich, 2007)。在这个重要的问题上, 几乎没有任何经验可以证明, 在这个方向上的新实验是非常可取的。另一个关键因素是成本 (惩罚、损失和努力), 这是奖励的另一面, 在神经层面上没有被很好的理解。

候选位点包括基底神经节的皮质纹状体连接或眶额皮质内的突触通路, 这些已在连接主义神经网络模型中得到探索 (Cohen et al., 1996; Frank and Claus, 2006) 和强化学习模型 (Samejima and Doya, 2007)。因此, 奖励依赖性突触可塑性很可能发生在特定的大脑区域 (或这些区域的神经元亚群), 这些区域专门负责

发出行为价值的信号，而其他区域则更直接地参与行为选择的产生 (Samejima and Doya, 2007; Rushworth and Behrens, 2008)。阐明这种大规模决策回路的内部工作是该领域的一大挑战。

不确定性和随机神经动力学

决策往往充满风险，因为感知世界和选择结果，以及互动决策主体的意图，只有在不同的概率水平上才为人所知。以上猴子实验说明了这一点：在一项接近阈值的识别任务中，感官信息贫乏且相互冲突；而在觅食任务中，响应选项的可能结果由奖励概率给出，奖励概率随时间不可预测地变化。在匹配便士的游戏任务中，代理必须在不知道彼此的预期行为的情况下做出决定，但结果取决于所有代理的选择。不确定性被认为是解释经济选择行为的关键因素 (Kahneman, 2002)，而风险下的决策代表了经济学家和神经科学家的共同兴趣点。最近，利用赌博任务和功能神经成像以及神经生理学对行为动物的研究进行了人类研究，以检验不确定性的神经表征。一些研究的目的是识别不同类型的不确定性所招募的不同的脑系统 (Yu and Dayan, 2005; Hsu et al., 2005; Huettel et al., 2006; Behrens et al., 2007)。其他研究用概率量化不确定性相关的神经信号 (Fiorillo et al., 2003) 或概率分布的方差 (Tobler et al., 2007)。还有人研究了关于概率决策信心的估计 (Grinband et al., 2006; Kepecs et al., 2008)。

在决策过程中随机性的起源一直是一个争论的问题，争论的焦点是相同的核心机制是否能够构成感知决策和基于价值的选择行为的基础 (Glimcher, 2005; Gold and Shadlen, 2007)。Gold and Shadlen (2007) 将决策描述为一个过程，其中“决策变量 (DV) 表示所有先验、证据和价值来源的应计收益，并将其转换成一个由决策规则解释以产生选择的数量”。在感知决策的累加器模型中，随机性来源于外部输入的噪声，DV 是随机的，但决策规则 (界) 是固定的。时间积分的计算效益可以通过信噪比来理解，信噪比随时间的增加而增加 (在扩散模型中约为 $\sqrt{t}$)。另一方面，在奖励依赖选择的强化模型中 (Barraclough et al., 2004; Sugrue et al., 2004, 2005; Lau and Glimcher, 2005)，DV 由响应选项的值

定义，响应选项的值根据过去的奖励进行确定性更新，而选择是由基于 DV 的概率决策规则(例如 softmax 标准)生成的。因此，随机性的来源被解释为内在的。Glimcher (2005) 认为内在的不确定性对于不可预测的行为可能是必不可少的。例如，在像匹配便士或石头剪刀布这样的互动游戏中，任何偏离代理人随机选择的趋势都可能被他的对手利用。

循环神经回路模型提供了一种调和这两种看似对立的观点的方法。在这个模型中，DV 和决策规则之间没有根本的区别，因为相同的循环神经动力学实例化了收益过程和分类选择。奖赏或价值信号通过突触输入调节神经放电，就像感觉刺激一样，与 Gold and Shadlen (2007) 的观点一致。神经动力学产生了随机决策，其总体行为特征是竞争决策神经元输入的差分 ΔI 的 softmax 函数。softmax 仅仅是对行为统计的描述，而不是用于在单个试验中产生个体选择的决策标准。 ΔI 的绝对值越小，网络行为越随机。在觅食任务中，序列反应和结果导致突触强度的变化， ΔI 反映了选择选项值的差异。当奖励不确定性的量变化时， ΔI 通过突触可塑性进行调节，使系统的行为或多或少是随机的，符合匹配规律(Soltani and Wang, 2006)。在竞争游戏中，与对手的相互作用会诱导奖励依赖性突触可塑性，迫使 ΔI 接近于零，导致随机行为。因此，决策者并没有随机游戏的目标，而只是试图在互动游戏中，在给定的环境或其他决策代理的情况下，达到最佳状态(Soltani et al., 2006; Lee and Wang, 2008)。这一结论与行为研究一致，行为研究表明反馈在产生随机反应模式中起着不可或缺的作用(Rapoport and Budescu, 1992; Camerer, 2003; Glimcher, 2005)。决策回路产生随机选择行为，并不是因为系统依赖于“随机数发生器”，而是因为决策者与不稳定环境或其他决策者之间的相互作用会导致适应性的、看似随机的决策模式。

这种强调内在随机神经动力学的观点也适用于知觉决策， ΔI 测量相对输入强度(如 RDM 刺激的一致性)。为了理解这一点，值得注意的是，即使是决定性的选项也与某些行为的可变性有关，特别是反应时间在不同的试验中的可变性。考虑一下对突然出现的视觉目标快速扫视(一种弹道眼动)的简单动作。没有感觉上的不确定性，行为反应总是一样的。然而，眼跳反应时间(目标开始和眼球运动之间的时间)在不同的试验中波动很大(Carpenter, 1981)。在猴子的生理实验中，

Hanes 和 Schall 发现, 当放电活动达到 FEF 中运动神经元的阈值水平时, 就会启动眼跳(Hanes and Schall, 1996; Schall, 2001)。眼跳反应时间在不同试验上的变异性被证明与运动相关神经元累积活动的斜率成反比, 而阈值水平与反应时间无关, 保持恒定(Hanes and Schall, 1996)。也有证据表明, 反应潜伏期与皮质运动神经元在目标发病前、眼跳和其他感觉运动任务中的准备活动之间存在一种试对试的相关性(Dorris et al., 1997; Churchland et al., 2006; Nakahara et al., 2006)。在发生冲突的情况下, 例如, 当实验对象在目标开始后的一小段时间内必须通过引入的停止信号来抑制计划好的扫视时, 行为就变成了概率性的(在某些试验中, 扫视受到抑制, 但在其他试验中则没有)(Logan and Cowan, 1984; Boucher et al., 2007)。因此, 在某种意义上, 神经系统固有的随机性通过外部不确定性表现出来(Green and Swets, 1966)。

感知信号检测理论用噪声输入来解释行为的不确定性(Green and Swets, 1966)。然而, 在生理学上, 我们对导致决策过程随机性的各种来源的身份和相对权重知之甚少。在猴子的 RDM 识别实验中, 至少有三种噪声成分会影响像 LIP 这样的决策回路: 实验中呈现的随机时空点模式, 实验间刺激的变化, 以及神经系统固有的波动动力学。有趣的是, 不同试验的刺激变化对 MT 神经元的放电活动的在不同试验中里的方差没有明显的影响(Britten et al., 1993), 对 MT 神经反应和行为选择之间的关系也没有明显的影响(Britten et al., 1996)。在循环神经回路模型中, 当外部输入在试验间变化或在试验间保持不变时, 通过心理测量函数和反应时间变异性测量的概率决策行为在模拟中被发现是相同的, 这表明变化的主要来源可能不是感官刺激而是在神经系统内部的感觉 (Wang, 2002)。此外, Deco 等将该模型应用于猴子体感识别实验, 发现本质上的随机决策电路动力学可以解释韦伯定律, 即刚刚可察觉的输入差与绝对输入强度的比值是恒定的(Deco and Rolls, 2006; Deco et al., 2007b)。这些例子说明了统计行为规律最终可以用神经元术语来解释的潜力。

无论是对于知觉决策模型还是基于价值的选择行为, 内在随机性都源自高度不规则的神经放电(Amit and Brunel, 1997; Brunel and Wang, 2001), 皮质神经元放电的特征(Softky and Koch, 1993; Shadlen and Newsome, 1994)。有证据

表明,这种随机性是皮质回路固有的。例如,即使在工作记忆中没有外部刺激的情况下,从行为正常的猴子身上记录下来的前额叶神经元的延续性活动,也具有类似泊松统计的特点(Compte et al., 2003)。在理论上提出的机制,为高度不规则的神经活动的皮层假定复发的皮层回路动力学赋有平衡的突触兴奋和抑制(van Vreeswijk and Sompolinsky, 1996; Amit and Brunel, 1997; Mattia and Del Giudice, 2004; Renart et al., 2007; Barbieri and Brunel, 2008)。在扩散模型中,输入由 ΔI 给出并整合,而在循环回路中,突触输入的一个重要组成部分“来自内部”,并随着时间的推移与尖峰活动同时建立。因此,时间整合超越了全部(外部的和重复的)输入,而不仅仅是感官刺激。假设像 LIP 这样的决策回路中的神经元通过一个递增的时间过程表现出泊松统计,然后信噪比随着时间的推移而提高,这仅仅是平均射速增加的结果(对任何泊松过程都是如此),而不是因为外部刺激中的噪声被平均掉了。此外,对于决策神经元来说,信噪比确实会随着时间的推移而衰减,这并不是一个必然的结论,因为神经累加器由于没有明显的泄漏,预计会显示不寻常的波动,例如,放电序列的 Fano 因子(方差与峰放电数目的平均值的比值)实际上可能会随着时间的推移而增长(Miller and Wang, 2006b)。

相平面分析(图 3D)为信噪比问题提供了一个新的视角。相平面分析(图 3D)为信噪比问题提供了一个新的视角。相平面分析(图 3D)为信噪比问题提供了一个新的视角。在任意给定时刻,将信号定义为当前网络状态(由两个神经池激发率 r_A 和 r_B 给出)与两个选择吸引子的吸引池边界之间的距离 $d(t)$ 。噪声可以通过网络状态在不同试验中的标准差 $[\sigma(t)]$ 来量化。信噪比为 $d(t)/\sigma(t)$ 。RDM 刺激开始时,初始信号 $d(t = 0)$ 依赖于运动强度 c' ; 如果 $c' = 0$, 它是零,但对于非零 c' 是正的,因为网络已经开始在流域内正确选择吸引子(参考图 3D)。然而,只要 $d(t)/\sigma(t)$ 很小,网络就仍然是“未确定的”。随着网络进一步演化为吸引盆地之一, $d(t)$ 和 $\sigma(t)$ 可能会随着时间的推移而增加,但 $d(t)/\sigma(t)$ 的比值会增加,因此噪声越来越不可能将网络“带回”另一个可选的吸引盆地。从这个意义上说,当 $d(t)$ 远大于 $\sigma(t)$, 即使网络可能仍然远离实际的吸引子状态,或者两个激发率都没有达到规定的决策阈值。根据这种信噪比的状态动力学观点,

不需要单独处理外部噪声刺激和内部神经元的随机性。时间累积的计算效益是通过吸引子景观中的网络动力学来理解的，而不是传统的时域分析。注意到网络并不直接“看到”运动相干性 c' ，只看到 RDM 刺激，但是决策空间对反映 c' 的平均输入是敏感的。对于较高的 c' （图 3D， $c' = 51.2\%$ vs . 6.4% ），网络在正确选择吸引子的区域开始越深，刺激开始时 d 越大，分类选择的时间点越早。当 c' 足够大时，性能是 100%，因为刺激一出现， $d(t)/\sigma(t)$ 就已经非常大，不可能切换到另一个选择吸引子（代表一个错误决策的吸引子）。我们可以说，系统在刺激开始时就已经下定决心了，尽管神经活动需要一段时间才能达到阈值水平。注意，每个基于状态的描述对应一个（足够长的）固定输入。这并不意味着这个决定是不可逆转的；外部输入的改变（例如，反转运动刺激的方向）可以从根本上改变吸引子景观，导致不同的选择。

有人提出，在一次试验中，神经群体活动模式明确地表示概率密度函数 (Ma et al., 2006)。在 RDM 鉴别实验中，LIP 神经元的上升活动被解释为两个备选方案的似然比 (Gold and Shadlen, 2001; Jazayeri and Movshon, 2006) 或后验（结合证据和先验信息）对数的时间总和 (Ma et al., 2006; Jazayeri and Movshon, 2006)。这类模型需要贝叶斯解码器，它使用非线性过程读出分类选项。递归神经电路模型提供了一个不同的视角，同一决策电路通过吸引子动力学进行数据的时间累积和分类选择。此外，单次试验的决策是基于波动神经网络活动的随机抽样；概率分布只出现在试验的汇总统计数据中。未来的研究将帮助我们理解这些不同的操作模式，可能在不同的大脑区域有不同的部署。

结论

最近，决策不仅在认知的神经生物学研究中受到越来越多的关注，而且在精神病学中也受到越来越多的关注，因为人们认识到，决策能力受损与各种精神障碍有着显著的联系 (Fellows, 2004; Sachdev and Malhi, 2005)。在这篇综述中，我整理了实验结果，并在此基础上开发了用于决策的递归神经回路理论。因为它现在已经成为可以跨物种调查决策，从苍蝇，老鼠，人类被试和猴子，从单个神经

元的生物物理学 (Llinas, 1988; Magee et al., 1998), 突触连接动力学 (Abbott and Regehr, 2004), 以及微电路连接 (Somogyi et al., 1998; Douglas and Martin, 2004) 角度研究其潜在机制的时机已经成熟。从非线性动力系统的角度来看, 数量上的差异会产生性质上不同的函数。因此, 虽然后顶叶和前额叶皮质的结构布局可能与感觉皮质在质量上相同, 但如果突触复发率足够高(前提是它的复发率较低), 就可以自然地导致持续活动的产生, 以及适合于为认知型计算提供帮助的兴奋性活动 (Wang, 2006a)。相反, 周期性兴奋和抑制的相对适度的减少则会导致认知功能的显著损害 (Wang, 2006b; Durstewitz and Seamans, 2008; Rolls et al., 2008)。

工作记忆和决策中的关键神经计算可以被理解为输入的时间整合: 工作记忆依赖于神经元, 神经元将短暂的输入脉冲转化为自我持续的持续活动, 而决策则涉及到准线性的提升活动, 以应对不断输入的信息积累。知觉辨别 (Pasternak and Greenlee, 2005) 和行为选择 (Tanji and Hoshi, 2008) 任务通常也依赖于工作记忆, 以保留对未来决策有用的信息或记住以前做出的选择。然而, 目前尚不清楚, 在工作记忆和决策过程中证据积累的刺激选择持续性活动中, 潜在的回路机制是否必然相同。在 RDM 识别实验中, 记录的 LIP 神经细胞在延迟动眼神经反应任务中显示定向调谐的助记活性, 并以此为标准进行预选 (Shadlen and Newsome, 1996, 2001; Roitman and Shadlen, 2002; Huk and Shadlen, 2005)。有趣的是, 在 RDM 识别任务中, 其他没有表现出持续活动的神经元是否也表现出缓慢的启动活动。此外, 细胞和突触机制的神经上升活动仍有待阐明的实验。LIP 是否真的作为一个吸引子网络也受到了质疑, 因为在选择性注意时 LIP 神经反应的某些方面还没有被现有的吸引子模型复制 (Ganguli et al., 2008)。

选择行为通常用基于值的优化来表示。因此, 阐明大脑中不同的评估维度是如何表现的是一个挑战 (Sugrue et al., 2005; Rushworth and Behrens, 2008)。另一个是理解不确定性和风险的神经指标, 以及它如何影响决策。在一个不确定的世界中, 有时需要用探索性决策来平衡强化。根据行为需求, 研究大脑如何利

用已知环境，或在多变的世界中探索最佳选择，将是有趣的 (Daw et al., 2006; Behrens et al., 2007)。

虽然本文的重点是简单的行为任务，但决策计算的基本构建及其神经机制可能也与更复杂的认知过程相关。即使在口语处理中，也有证据表明，一个口语单词会引发多种词汇表征，口语单词识别是从信息源的实时集成到语音竞争者之间的分类选择 (Spivey et al., 2005)。因此，我们希望通过更简单的知觉和经济决策的详细的神经生理学研究来了解更高层次的决策。

参考文献

- Abbott, L.F., and Regehr, W.G. (2004). Synaptic computation. *Nature* 431, 796 – 803.
- Abraham, N., Spors, H., Carleton, A., Margrie, T., Kuner, T., and Schaefer, A. (2004). Maintaining accuracy at the expense of speed: stimulus similarity defines odor discrimination **time in mice**. *Neuron* 44, 865 – 876.
- Amit, D.J. (1995). The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behav. Brain Sci.* 18, 617 – 626.
- Amit, D.J., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237 – 252.
- Andersen, R.A., and Buneo, C.A. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189 – 220.
- Barbieri, F., and Brunel, N. (2008). Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex? *Frontiers in Neurosci.* 2, 114 – 122.
- Barracough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404 – 410.
- Bayer, H., and Glimcher, P. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129 – 141.
- Behrens, T., Woolrich, M., Walton, M., and Rushworth, M. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214 – 1221.
- Binder, J., Liebenthal, E., Possing, E., Medler, D., and Ward, B. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nat. Neurosci.* 7, 295 – 301.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two alternative forced-choice tasks. *Psychol. Rev.* 113, 700 – 765.
- Bogacz, R., Usher, M., Zhang, J., and McClelland, J. (2007). Extending a

biologically inspired model of choice: multi-alternatives, nonlinearity and valuebased multidimensional choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1655 – 1670.

Boucher, L., Palmeri, T., Logan, G., and Schall, J. (2007). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376 – 397.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12, 4745 – 4765.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157 – 1169.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis. Neurosci.* 13, 87 – 100.

Brody, C., Hernandez, A., Zainos, A., and Romo, R. (2003). Timing and neural encoding of somatosensory parametric working

memory in macaque prefrontal cortex. *Cereb. Cortex* 13, 1196 – 1207.

Brunel, N., and Wang, X.-J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63 – 85.

Busmeyer, J., and Townsend, J. (1993). Decision field theory: a dynamic cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* 100, 432 – 459.

Camerer, C.F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction* (Princeton, NJ: Princeton University Press).

Carpenter, R.H.S. (1981). Oculomotor procrastination. In *Eye Movements: Cognition and Visual Perception*, D.F. Fisher, R.A. Monty, and J.W. Senders, eds. (Hillsdale, NJ: Lawrence Erlbaum), pp. 237 – 246.

Chafee, M.V., and Goldman-Rakic, P.S. (1998). Neuronal activity in macaque prefrontal area 8a and posterior parietal area 7ip related to memory guided saccades. *J. Neurophysiol.* 79, 2919 – 2940.

Churchland, M., Afshar, A., and Shenoy, K. (2006). A central source of movement variability. *Neuron* 52, 1085 – 1096.

Churchland, A., Kiani, R., and Shadlen, M. (2008). Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693 – 702.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model. *J. Neurosci.* 26, 9761 – 9770.

Cisek, P., and Kalaska, J. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron* 45, 801 – 814.

Cohen, J., Braver, T., and O'Reilly, R. (1996). A computational approach to prefrontal cortex, cognitive control and schizophrenia: recent developments and

current challenges. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1515–1527.

Colby, C.L., and Godberg, M.E. (1999). Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349.

Compte, A., Constantinidis, C., Tegner, J., Raghavachari, S., Chafee, M.V., **Goldman-Rakic, P.S., and Wang, X.-J.** (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol.* 90, 3441–3454.

Cook, E.P., and Maunsell, J.H.R. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat. Neurosci.* 5, 985–994.

Corbetta, M., and Shulman, G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Corrado, G., Sugrue, L., Seung, H., and Newsome, W. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.

Daw, N., O’Doherty, J., Dayan, P., Seymour, B., and Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

de Lafuente, V., and Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* 8, 1698–1703.

Deco, G., and Rolls, E. (2006). Decision-making and Weber’s law: a neurophysiological model. *Eur. J. Neurosci.* 24, 901–916.

Deco, G., Prez-Sanagustn, M., de Lafuente, V., and Romo, R. (2007a). Perceptual detection as a dynamical bistability phenomenon: a neurocomputational correlate of sensation. *Proc. Natl. Acad. Sci. USA* 104, 20073–20077.

Deco, G., Scarano, L., and Soto-Faraco, S. (2007b). Weber’s law in decision making: integrating behavioral data in humans with a neurophysiological model. *J. Neurosci.* 27, 11192–11200.

Ditterich, J. (2006). Evidence for time-variant decision making. *Eur. J. Neurosci.* 24, 3628–3641.

Donders, F.C. (1969). On the speed of mental processes. [Translation of *Die Schnelligkeit Psychischer Prozesse*, first published in 1868]. *Acta Psychologica* 30, 412–431.

Dorris, M., Pare’, M., and Munoz, D. (1997). Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J. Neurosci.* 17, 8566–8579.

Dorris, M.C., and Glimcher, P.W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Douglas, R.J., and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451.

Durstewitz, D., and Seamans, J. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol omethyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 10.1016/j.biopsych.2008.05.015.

Ewert, J. (1997). Neural correlates of key stimulus and releasing mechanism: a case study and two concepts. *Trends Neurosci.* 20, 332–339.

Fellows, L.K. (2004). The cognitive neuroscience of human decision making: a

review and conceptual framework. *Behav. Cogn. Neurosci. Rev.* 3, 159 – 172.

Fiorillo, C., Tobler, P., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898 – 1902.

Frank, M., and Claus, E. (2006). Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol. Rev.* 113, 300 – 326.

Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331 – 349.

Fusi, S. (2002). Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.* 87, 459 – 470.

Fusi, S., Asaad, W.F., Miller, E.K., and Wang, X.-J. (2007). A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* 54, 319 – 333.

Fuster, J.M. (2008). *The Prefrontal Cortex, Fourth Edition* (New York: Academic Press).

Ganguli, S., Bisley, J.W., Roitman, J.D., Shadlen, M.N., Goldberg, M.E., and Miller, K.D. (2008). One-dimensional dynamics of attention and decision making in LIP. *Neuron* 58, 15 – 25.

Glimcher, P.W. (2003). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (Cambridge, MA: MIT Press).

Glimcher, P.W. (2005). Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25 – 56.

Gnadt, J.W., and Andersen, R.A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216 – 220.

Gold, J.I., and Shadlen, M.N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10 – 16.

Gold, J.I., and Shadlen, M.N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299 – 308.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535 – 574.

Goldman-Rakic, P.S. (1995). Cellular basis of working memory. *Neuron* 14, 477 – 485.

Gratton, G., Coles, M., Sirevaag, E., Eriksen, C., and Donchin, E. (1988). Preand poststimulus activation of response channels: a psychophysiological analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 331 – 344.

Green, D.M., and Swets, J.A. (1966). *Signal Detection Theory and Psychophysics* (New York: John Wiley and Sons).

Grinband, J., Hirsch, J., and Ferrera, V.P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron* 49, 757 – 763.

Grossberg, S., and Pilly, P. (2008). Temporal dynamics of decision-making during motion perception in the visual cortex. *Vision Res.* 48, 1345 – 1373.

Hanes, D.P., and Schall, J.D. (1996). Neural control of voluntary movement

initiation. *Science* 274, 427 – 430.

Hanks, T.D., Ditterich, J., and Shadlen, M.N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9, 682 – 689.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* 431, 859 – 862.

Heekeren, H., Marrett, S., and Ungerleider, L. (2008). The neural systems that mediate human perceptual decision making. *Nat. Rev. Neurosci.* 9, 467 – 479.

Heinzle, J., Hepp, K., and Martin, K. (2007). A microcircuit model of the frontal eye fields. *J. Neurosci.* 27, 9341 – 9353.

Hernández, A., Zainos, A., and Romo, R. (2002). Temporal evolution of a decision making process in medial premotor cortex. *Neuron* 33, 959 – 972.

Herrnstein, R.J., Rachlin, H., and Laibson, D.I. (1997). *The Matching Law: Papers in Psychology and Economics* (Cambridge: Harvard University Press).

Hestrin, S., Sah, P., and Nicoll, R. (1990). Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices. *Neuron* 5, 247 – 253.

Hick, W.E. (1952). On the rate of gain of information. *Q. J. Exp. Psychol.* 4, 11 – 26.

Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953 – 978.

Hikosaka, O., Nakamura, K., and Nakahara, H. (2006). Basal ganglia orient eyes to reward. *J. Neurophysiol.* 95, 567 – 584.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680 – 1683.

Huettel, S., Stowe, C., Gordon, E., Warner, B., and Platt, M. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765 – 775.

Huk, A.C., and Shadlen, M.N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420 – 10436.

Izhikevich, E. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* 17, 2443 – 2452.

Jazayeri, M., and Movshon, J.A. (2006). Optimal representation of sensory information by neural populations. *Nat. Neurosci.* 9, 690 – 696.

Kahneman, D. (2002). Nobel prize lecture: Maps of bounded rationality: a perspective on intuitive judgment and choice. In *Nobel Prizes 2002: Nobel Prizes, Presentations, Biographies, & Lectures*, T. Frangsmyr, ed. (Stockholm: Almqvist & Wiksell Int), pp. 416 – 499.

Kennerley, S., Walton, M., Behrens, T., Buckley, M., and Rushworth, M. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940 – 947.

Kepecs, A., Uchida, N., Zariwala, H., and Mainen, Z.F. (2008). Neural

correlations, computation and behavioral impact of decision confidence. *Nature* 455, 227 – 231.

Kiani, R., Hanks, T., and Shadlen, M. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017 – 3029.

Kim, J.-N., and Shadlen, M.N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2, 176 – 183.

Kutas, M., McCarthy, G., and Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science* 197, 792 – 795.

Laming, D. (1968). *Information Theory of Choice Reaction Times* (New York: Academic Press).

Lau, B., and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555 – 579.

Lau, B., and Glimcher, P. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451 – 463.

Law, C., and Gold, J. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11, 505 – 513.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404 – 409.

Lee, D., and Wang, X.-J. (2008). Mechanisms for stochastic decision making in the primate frontal cortex: Single-neuron recording and circuit modeling. In *Neuroeconomics: Decision Making and the Brain*, E.F.P.W. Glimcher, C.F. Camerer, and R.A. Poldrack, eds. (New York: Academic Press).

Lemus, L., Hernández, A., Luna, R., Zainos, A., Ncher, V., and Romo, R. (2007). Neural correlates of a postponed decision report. *Proc. Natl. Acad. Sci. USA* 104, 17174 – 17179.

Llinás, R. (1988). The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function. *Science* 242, 1654 – 1664.

Lo, C.C., and Wang, X.-J. (2006). Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956 – 963.

Loewenstein, G., Rick, S., and Cohen, J. (2008). *Neuroeconomics*. *Annu. Rev. Psychol.* 59, 647 – 672.

Logan, G.D., and Cowan, W.B. (1984). On the ability to inhibit thought and action: a theory of an act of control. *Psychol. Rev.* 91, 295 – 327.

Luce, R.D. (1986). *Response Time: Their Role in Inferring Elementary Mental Organization* (New York: Oxford University Press).

Ma, W., Beck, J., Latham, P., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432 – 1438.

Machens, C.K., Romo, R., and Brody, C.D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 307, 1121 – 1124.

Magee, J., Hoffman, D., Colbert, C., and Johnston, D. (1998). Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annu. Rev. Physiol.* 60, 327 – 346.

- Major, G., and Tank, D.** (2004). Persistent neural activity: prevalence and mechanisms. *Curr. Opin. Neurobiol.* 14, 675 – 684.
- Matsuda, Y., Marzo, A., and Otani, S.** (2006). The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J. Neurosci.* 26, 4803 – 4810.
- Matsumoto, M., and Hikosaka, O.** (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* 447, 1111 – 1115.
- Mattia, M., and Del Giudice, P.** (2004). Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Phys. Rev. E Stat. Nonlin. Soft Matter Physiol.* 70, 052903.
- Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N.** (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257 – 1269.
- McCarthy, G., and Donchin, E.** (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science* 211, 77 – 80.
- McMillen, T., and Holmes, P.** (2006). The dynamics of choice among multiple alternatives. *J. Math. Psychol.* 50, 30 – 57.
- McPeck, R., and Keller, E.** (2002). Saccade target selection in the superior colliculus during a visual search task. *J. Neurophysiol.* 88, 2019 – 2034.
- Meyer, D.E., Osman, A.M., Irwin, D.E., and Yantis, S.** (1988). Modern mental chronometry. *Biol. Psychol.* 26, 3 – 67.
- Miller, E.K., and Cohen, J.D.** (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167 – 202.
- Miller, P., and Wang, X.-J.** (2006a). Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. *Proc. Natl. Acad. Sci. USA* 103, 201 – 206.
- Miller, P., and Wang, X.-J.** (2006b). Power-law neuronal fluctuations in a recurrent network model of parametric working memory. *J. Neurophysiol.* 95, 1099 – 1114.
- Miller, P., Brody, C.D., Romo, R., and Wang, X.-J.** (2003). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb. Cortex* 13, 1208 – 1218.
- Munoz, D.P., and Fecteau, J.H.** (2002). Vying for dominance: dynamic interactions control visual fixation and saccadic initiation in the superior colliculus. *Prog. Brain Res.* 140, 3 – 19.
- Nakahara, H., Nakamura, K., and Hikosaka, O.** (2006). Extended LATER model can account for trial-by-trial variability of both pre- and post-processes. *Neural Netw.* 19, 1027 – 1046.
- Narayanan, N., and Laubach, M.** (2006). Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron* 52, 921 – 931.
- Newsome, W.T., Britten, K.H., and Movshon, J.A.** (1989). Neuronal correlates of a perceptual decision. *Nature* 341, 52 – 54.
- Niwa, M., and Ditterich, J.** (2008). Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435 – 4445.

- Okamoto, H., Isomura, Y., Takada, M., and Fukai, T. (2007). Temporal integration by stochastic recurrent network dynamics with bimodal neurons. *J. Neurophysiol.* 97, 3859 – 3867.
- Otani, S., Daniel, H., Roisin, M., and Crepel, F. (2003). Dopaminergic modulation of long-term synaptic plasticity in rat prefrontal neurons. *Cereb. Cortex* 13, 1251 – 1256.
- Padoa-Schioppa, C., and Assad, J. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223 – 226.
- Parker, A., and Krug, K. (2003). Neuronal mechanisms for the perception of ambiguous stimuli. *Curr. Opin. Neurobiol.* 13, 433 – 439.
- Pasternak, T., and Greenlee, M. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97 – 107.
- Pesaran, B., Nelson, M., and Andersen, R. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* 453, 406 – 409.
- Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233 – 238.
- Ploran, E., Nelson, S., Velanova, K., Donaldson, D., Petersen, S., and Wheeler, M. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. *J. Neurosci.* 27, 11912 – 11924.
- Posner, M.I. (1978). *Chronometric Exploration of Mind* (Hillsdale, NJ: Lawrence Erlbaum Associates).
- Quintana, J., and Fuster, J.M. (1999). From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cereb. Cortex* 9, 213 – 221.
- Rapoport, A., and Budescu, D.V. (1992). Generation of random series in twoperson strictly competitive games. *J. Exp. Psychol. Gen.* 121, 352 – 363.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychol. Rev.* 85, 59 – 108.
- Ratcliff, R., and Rouder, J.F. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347 – 356.
- Real, L. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science* 253, 980 – 986.
- Renart, A., Moreno-Bote, R., Wang, X.-J., and Parga, N. (2007). Mean-driven and fluctuation-driven persistent activity in recurrent networks. *Neural Comput.* 19, 1 – 46.
- Reynolds, J.N.J., Hyland, B.I., and Wickens, J.R. (2001). A cellular mechanism of reward-related learning. *Nature* 413, 67 – 70.
- Roesch, M., and Olson, C. (2003). Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90, 1766 – 1789.
- Roesch, M., Calu, D., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615 – 1624.
- Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task.

- J. Neurosci. 22, 9475 - 9489.
- Rolls, E., Loh, M., Deco, G., and Winterer, G.** (2008). Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nat. Rev. Neurosci.* 9, 696 - 709.
- Romo, R., and Salinas, E.** (2001). Touch and go: Decision-making mechanisms in somatosensation. *Annu. Rev. Neurosci.* 24, 107 - 137.
- Romo, R., Brody, C.D., Herná' ndez, A., and Lemus, L.** (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470 - 474.
- Romo, R., Herná' ndez, A., Zainos, A., Lemus, L., and Brody, C.** (2002). Neuronal correlates of decision-making in secondary somatosensory cortex. *Nat. Neurosci.* 5, 1217 - 1225.
- Romo, R., Herná' ndez, A., and Zainos, A.** (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* 41, 165 - 173.
- Roxin, A., and Ledberg, A.** (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046.
- Rushworth, M., and Behrens, T.** (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389 - 397.
- Sachdev, P., and Malhi, G.** (2005). Obsessive-compulsive behaviour: a disorder of decision-making. *Aust. N. Z. J. Psychiatry* 39, 757 - 763.
- Samejima, K., and Doya, K.** (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. *Ann. N Y Acad. Sci.* 1104, 213 - 228.
- Samejima, K., Ueda, Y., Doya, K., and Kimura, M.** (2005). Representation of actionspecific reward values in the striatum. *Science* 310, 1337 - 1340.
- Schall, J.D.** (2001). Neural basis of deciding, choosing and acting. *Nat. Neurosci.* 2, 33 - 42.
- Schall, J.** (2004). On building a bridge between brain and behavior. *Annu. Rev. Psychol.* 55, 23 - 50.
- Scherberger, H., and Andersen, R.** (2007). Target selection signals for arm reaching in the posterior parietal cortex. *J. Neurosci.* 27, 2001 - 2012.
- Schlegel, T., and Schuster, S.** (2008). Small circuits for large tasks: high-speed decision-making in archerfish. *Science* 319, 104 - 106.
- Schultz, W.** (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1 - 27.
- Schultz, W., Dayan, P., and Montague, P.R.** (1997). A neural substrate of prediction and reward. *Science* 275, 1593 - 1599.
- Seo, H., and Lee, D.** (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366 - 8377.
- Sereno, A., and Amador, S.** (2006). Attention and memory-related responses of neurons in the lateral intraparietal area during spatial and shape-delayed matchto- sample tasks. *J. Neurophysiol.* 95, 1078 - 1098.
- Seung, H.S.** (1996). How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. USA* 93, 13339 - 13344.

- Seung, H. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron* 40, 1063 – 1073.
- Seung, H.S., Lee, D.D., Reis, B.Y., and Tank, D.W. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26, 259 – 271.
- Shadlen, M.N., and Newsome, W.T. (1994). Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4, 569 – 579.
- Shadlen, M.N., and Newsome, W.T. (1996). Motion perception: seeing and deciding. *Proc. Natl. Acad. Sci. USA* 93, 628 – 633.
- Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916 – 1936.
- Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848 – 851.
- Sigman, M., and Dehaene, S. (2005). Parsing a cognitive task: a characterization of the mind's bottleneck. *PLoS Biol.* 3, e37.
- Smith, P.L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161 – 168.
- Softky, W.R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* 13, 334 – 350.
- Soltani, A., and Wang, X.-J. (2006). A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* 26, 3731 – 3744.
- Soltani, A., and Wang, X.-J. (2008). From biophysics to cognition: rewarddependent adaptive choice behavior. *Curr. Opin. Neurobiol.* 18, 209 – 216.
- Soltani, A., Lee, D., and Wang, X.-J. (2006). Neural mechanism for stochastic behavior during a competitive game. *Neural Netw.* 19, 1075 – 1090.
- Somogyi, P., Tamás, G., Lujan, R., and Buhl, E.H. (1998). Salient features of synaptic organization in the cerebral cortex. *Brain Res. Brain Res. Rev.* 26, 113 – 135.
- Soon, C., Brass, M., Heinze, H., and Haynes, J. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* 11, 543 – 545.
- Spivey, M., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. USA* 102, 10393 – 10398.
- Stone, M. (1960). Models for choice reaction time. *Psychometrika* 25, 251 – 260.
- Sugrue, L.P., Corrado, G.C., and Newsome, W.T. (2004). Matching behavior and representation of value in parietal cortex. *Science* 304, 1782 – 1787.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363 – 375.
- Sutton, R.S., and Barto, A.G. (1998). *Reinforcement Learning: An Introduction* (Cambridge, MA: MIT Press).
- Tanji, J., and Hoshi, E. (2008). Role of the lateral prefrontal cortex in

executive behavioral control. *Physiol. Rev.* 88, 37 – 57.

Tobler, P., Fiorillo, C., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642 – 1645.

Tobler, P., O’ Doherty, J., Dolan, R., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621 – 1632.

Usher, M., and McClelland, J. (2001). On the time course of perceptual choice: the leaky competing accumulator model. *Psychol. Rev.* 108, 550 – 592.

Uchida, N., and Mainen, Z. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* 6, 1224 – 1229.

Uchida, A., Kepecs, A., and Mainen, Z.F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nat. Rev. Neurosci.* 7, 485 – 491.

van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274, 1724 – 1726.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. In *Contemporary Problems in Perception: Ergonomics*, A.T. Welford and L. Houssiadass, eds. (London: Taylor & Francis), pp. 37 – 58.

Wang, X.-J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19, 9587 – 9603.

Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455 – 463.

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955 – 968.

Wang, X.-J. (2006a). A microcircuit model of prefrontal functions: ying and yang of reverberatory neurodynamics in cognition. In *The Prefrontal Lobes: Development, Function and Pathology*, J. Risberg, J. Grafman, and F. Boller, eds. (New York: Cambridge University Press), pp. 92 – 127.

Wang, X.-J. (2006b). Toward a prefrontal microcircuit model for cognitive deficits in schizophrenia. *Pharmacopsychiatry* 39 (Suppl 1), 80 – 87.

Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629 – 632.

Watanabe, K., and Funahashi, S. (2007). Prefrontal delay-period activity reflects the decision process of a saccade direction during a free-choice ODR task. *Cereb. Cortex* 17 (Suppl 1), 88 – 100.

Wickelgren, W.A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol. (Amst.)* 41, 67 – 85.

Wickens, J.R., Reynolds, J.N.J., and Hyland, B.I. (2003). Neural mechanisms of reward-related motor learning. *Curr. Opin. Neurobiol.* 13, 685 – 690.

Wilimzig, C., Schneider, S., and Scho¨ner, G. (2006). The time course of saccadic decision making: dynamic field theory. *Neural Netw.* 19, 1059 – 1074.

Wong, K.F., and Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314 – 1328.

Wong, K.F., Huk, A.C., Shadlen, M.N., and Wang, X.-J. (2007). Neural circuit

dynamics underlying accumulation of time-varying evidence during perceptual decision-making. *Frontiers in Comput. Neurosci.* 1, 6. 10.3389/neuro.10/006. 2007.

Worgotter, F., and 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.

Yu, A. J., and Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron* 46, 681 – 692.