

# 《意识》

---

# 意识

---

## 意识

---

### 一个浪漫还原主义者的忏悔录

Christof Koch

The MIT Press  
Cambridge, Massachusetts  
London, England

© 2012 Massachusetts Institute of Technology

版权所有。未经出版商书面许可，本书的任何部分均不得以任何电子或机械方式（包括影印、录音或信息存储和检索）复制。

有关特殊数量折扣的信息，请发送邮件至 [special\\_sales@mitpress.mit.edu](mailto:special_sales@mitpress.mit.edu)。

国会图书馆出版物编目数据

Koch, Christof, 1956 –

Consciousness : confessions of a romantic reductionist / Christof Koch.

p. cm.

包含书目参考文献和索引。

ISBN 978-0-262-01749-7 (硬皮书 : alk. paper)

ISBN 978-0-262-30103-9 (零售电子书)

1. 意识。2. 心灵与身体。3. 自由意志和决定论。4. 生命。I. 标题。

B808.9.K63 2012

153—dc23

10 9 8 7 6 5 4 3 2 1

献给Hannele

# 目录

---

前言

[致谢]

[第1章]: 我将介绍古老的心身问题，解释为什么我致力于用理性和实证研究来解决它，让你认识Francis Crick，解释他与这一探索的关系，做出一个忏悔，并以悲伤的音符结束

[第2章]: 我将写到我内心宗教与理性冲突的源泉，为什么我从小就想成为一名科学家，为什么我佩戴Calculus教授的胸针，以及我如何在晚年获得了第二位导师

[第3章]: 我将解释为什么意识挑战了科学的世界观，如何双脚坚实着地地对意识进行实证研究，为什么动物与人类共享意识，以及为什么自我意识并不像许多人认为的那么重要

[第4章]: 你将听到科学家-魔术师的故事，他们让你看而不见，他们如何通过窥视你的头骨来追踪意识的足迹，为什么你不是用眼睛看，以及为什么注意力和意识不是同一回事

第5章：你将从神经学家和神经外科医生那里了解到，有些神经元非常关心名人，将大脑皮层一分为二并不会让意识减半，颜色会因小块皮层区域的丧失而从世界中消失，以及方糖大小的脑干或丘脑组织的破坏会让你变成活死人

第6章：我将为两个我年轻时认为是胡说八道的命题进行辩护——你对头脑中发生的大多数事情都没有意识，僵尸代理(zombie agents)控制着你生活的大部分，尽管你自信地相信自己在掌控一切

第7章：我将抛开谨慎，提出自由意志、《尼伯龙根的指环》以及物理学对决定论的看法，解释你心智选择能力的匮乏，表明你的意志滞后于大脑的决定，而自由只是感觉的另一个词

[第8章]: 我将论证意识是复杂事物的基本属性，对integrated information theory(整合信息理论)赞不绝口，解释它如何解释关于意识的许多令人困惑的事实，并提供构建有感知能力机器的蓝图

[第9章]: 我将概述测量意识的电磁设备，描述利用基因工程力量追踪小鼠意识的努力，并发现自己在构建皮层观测站

[第10章]: 我将思考被礼貌科学话语视为禁区的终极问题：即科学与宗教的关系、上帝的存在、这位上帝是否能干预宇宙、我导师的去世，以及我最近的苦难

注释

参考文献

索引

# 前言

---

你手中拿着的是一本关于现代意识科学的简明阐述。在几个小时的时间里，你可以了解我们自然科学家在揭开我们存在的核心问题之一方面的立场——即主观感受、意识是如何进入世界的？“通过你的头脑”是显而易见的答案。但这个答案并不是很有用。你头脑中的大脑到底有什么特别之处，让你意识到颜色、疼痛和快乐、过去和未来、自己和他人？是否任何大脑都足够？昏迷患者、胎儿、狗、老鼠或苍蝇的大脑呢？计算机的“大脑”呢？它们能有意识吗？我将处理这些问题以及更多问题，包括自由意志、意识理论，以及我研究中的噩梦：quantum mechanics(量子力学)与理解意识的相关程度。

然而，这本书不仅仅是关于科学的。它也是一本忏悔录和回忆录。我不仅是一个冷静的物理学家和生物学家，也是一个只有几年时间来理解存在之谜的人。在过去的几年里，我了解到我的无意识倾向、我的信念以及我的个人优点和缺点如何强有力地影响了我的生活和对毕生事业的追求。我体验了小说家村上春树在一次引人注目的采访中所描述的：“我们内心有房间。其中大部分我们还没有探访过。被遗忘的房间。我们时不时能找到通道。我们发现奇怪的东西……旧留声机、图片、书籍……它们属于我们，但这是我们第一次发现它们。”当这些失落的房间与我的探索——揭开意识的根源——相关时，你将了解其中的一些。

加州福尼亚州帕萨迪纳

2011年5月

致谢

撰写、编辑和出版一本书需要许多人的积极配合。书籍证明了人类的美好本性——乐于为遥远的目标而努力，主要的回报就是工作出色完成后的满足感。

Blair Potter接手了我的散文并对其进行编辑。她识别出我写作中的三条不同线索，先将其拆分再重新编织在一起。如果最终成果读起来接近一个整体性的文本，那都要感谢她。John Murdzek和Katherine Almeida仔细校对了所有内容，Sara Ball、Amy Chung-Yu Chou和Kelly Overly提供了更多的编辑建议。

范德堡大学精神病学和宗教研究教授Volney Gay在纳什维尔邀请我于2007年春季就“哲学、宗教和科学中的意识问题”进行三场Templeton研究讲座。正是在那里，这本书的构思诞生了。我感谢John Templeton基金会对这些公开讲座的慷慨资助。

我要感谢许多学生、朋友和同事，他们阅读了本书的章节并指出了其中的诸多不当和不一致之处——Ralph Adolphs、Ned Block、Bruce Bridgeman、McKell Ronald Carter、Moran Cerf、David Chalmers、Michael Hawrylycz、Constanze Hipp、Fatma Imamoglu、Michael Koch、Gabriel Kreiman、Uri Maoz、Leonard Mlodinow、Joel Padowitz、Anil Seth、Adam Shai、Giulio Tononi和Gideon Yaffe。Heather Berlin建议了这个书名。Bruce Bridgeman、McKell Carter和Giulio Tononi花时间仔细通读了整个文本并对其进行修改。感谢他们的集体努力，他们含蓄或直接的批评，这本书读起来更加流畅，减少了突出、令人烦恼或分散注意力的粗糙之处。

感谢为我提供安静港湾的众多机构。首先是加州理工学院，我四分之一世纪以来的精神家园。在过去这些年里，也是我人生最艰难的时期，加州理工学院和它的人们是我唯一稳定的支撑。他们以各种大大小小的方式帮助我应对困难。首尔的高丽大学为我在远东提供了避难所，给我空间去写作、思考和反思这些页面中讨论的所有问题。西雅图的艾伦脑科学研究所慷慨地给了我时间来完成这本书。

我实验室的研究得到了国家科学基金会、国立卫生研究院、海军研究办公室、国防高级研究计划局、G. Harold & Leila Y. Mathers基金会、Swartz基金会、Paul G. Allen家族基金会以及韩国国家研究基金会世界一流大学项目的支持。我对所有这些机构深表感激。

# 第1章：

---

我介绍古老的心-身问题，解释为什么我要探索用理性和实证调查来解决它，让你认识Francis Crick，解释他与这个探索的关系，坦白一些事情，并以悲伤的音符结束

我无法告诉你它真正是什么，我只能告诉你它感觉起来是什么样的。

——Eminem，《Love the Way You Lie》（2010）

一个日常事件让我的人生走上了新的道路。我已经吃了阿司匹林，但牙痛依然存在。躺在床上，我因为下磨牙的阵阵疼痛而无法入睡。

为了让自己从这种痛苦的感觉中分心，我想知道它为什么会这么疼。我知道牙髓的炎症会沿着三叉神经的一个分支向上发送电活动，最终到达脑干。经过进一步的中继后，颅骨下方新皮质灰质部分的神经细胞变得活跃并释放电脉冲。大脑这一部分的生物电活动与疼痛意识密切相关，包括其可怕的、酸痛的感觉。

但是等等。刚刚发生了一些极其无法解释的事情。大脑中的活动是如何触发感觉的？那只是一些软糊糊的东西。正如赛博朋克小说轻蔑地称呼身体的“肉”，怎么能产生知觉？更一般地说，物理的东西怎么能产生非物理的东西，产生主观状态？无论是在遥远夏日大西洋岸边经历的疼痛，看到我的孩子时感受到的喜悦，还是品尝起泡Vouvray酒的味道，都有着同样成问题的起源——神经物质的激动。

这是有问题的，因为神经系统和其内在视角（它产生的感觉）之间似乎存在不可逾越的鸿沟。一方面是大脑，已知宇宙中最复杂的物体，一个受物理定律支配的物质事物。另一方面是意识的世界，生活的景象和声音，恐惧和愤怒，欲望、爱情和倦怠的世界。

这两个世界密切相关——中风或头部的重击都戏剧性地证明了这一点。Oscar Wilde诗意地表达道：“正是在大脑中，罂粟是红色的，苹果是芳香的，云雀在歌唱。”但这种转换到底是如何发生的？大脑是如何构建conscious experience的？通过其形状、大小、活动还是复杂性？

Consciousness并不出现在构成物理学基础的方程中，也不在化学元素周期表中，也不在我们基因无尽的ATGC分子序列中。然而我们两人——我，这些文字的作者，和你，读者——都是有知觉的。这就是我们发现自己所处的宇宙，一个高度有序物质的特定振动触发conscious feelings的宇宙。这看起来就像摩擦一盏黄铜灯然后有精灵出现并允许三个愿望一样神奇。

我是一个书呆子。小时候，我自己制造了家用电脑来实现布尔逻辑定律。我会在床上躺着不睡，在脑海中设计复杂的隧道挖掘机。所以，当我牙疼的时候，我很自然地想知道计算机是否能体验到疼痛。假设我将温度传感器连接到笔记本电脑上，并给它编程，如果房间变得太热，屏幕上就会出现红色大字“疼痛”。但是“疼痛”对我的Mac来说会有任何感觉吗？我愿意承认苹果产品的许多优点，尤其是酷炫，但不包括感知力。

但为什么不是呢？是因为我的笔记本电脑运行在不同的物理原理上吗？不是带正负电荷的钠离子、钾离子、钙离子和氯离子在神经细胞内外流动，而是电子流向晶体管的栅极，使它们切换。这就是关键差异吗？我不这么认为，因为在我看来，最终重要的一定是大脑不同部分之间的功能关系。而这些关系至少在原理上可以在计算机上模拟。是

因为人类是有机的，由骨骼、肌肉和神经构成，而计算机是合成的，由钛、铜线和硅构成吗？这似乎也不关键。那么也许是因为人类通过偶然和必然进化而来，而机器是明确设计出来的？动物的进化历史与数字机器截然不同，这种差异反映在它们不同的架构中。但我不明白这如何影响其中一个或另一个是否有意识。一定是系统现在的物理状态起决定作用，而不是它如何变成现在这样。

什么是产生差异的差异？

在哲学中，解释为什么有人能感受到任何东西的困难通常被称为困难问题。这个术语是哲学家David Chalmers创造的。他在1990年代早期通过一系列严密论证的推理链建立了声誉，得出结论认为意识体验不是从支配宇宙的物理定律中产生的。这些定律同样适用于一个没有意识的世界或具有不同意识形式的世界。永远不会有一个人还原论的、机械论的解释来说明客观世界如何与主观世界联系。困难问题这个术语，带着大写的H，如“不可能困难”，随后变得病毒式传播。没有人争议物理世界和现象世界在数十亿人的日常生活中密切相关，但为什么会这样一个谜。

Dave给我上了关于哲学家的重要一课。我邀请他在神经生物学和工程学听众面前演讲。之后，喝着葡萄酒，我震惊地发现他坚持认为没有任何经验事实、生物学发现或数学概念进步能让他相信这两个世界之间不可逾越的鸿沟是可以消除的。困难问题不受任何这样的进步影响。我感到震惊。仅仅是文字，没有数学或物理-经验框架的支持，如何能以如此确定的程度确立任何东西？对我来说，他有一个有力的论证，但肯定没有证明。

从那时起，我遇到了许多完全相信自己观点真理性的哲学家。对自己观点的这种信心——不被其他人无数竞争观点所困扰，而他们不可能都是对的——在自然科学家们中是罕见的。由于我们与混乱的大自然母亲的持续实验对抗，她迫使我们修改最聪明和最美观的理论，我们已经痛苦地学会了不要过分信任任何一个想法，直到它被确立得毫无合理怀疑为止。

然而，在某种无意识的层面上，这些论证对我产生了影响。它们表明，在寻求理解现象世界时，科学终于遇到了对手，意识抵抗理性解释，免疫于科学分析，超出了经验验证的理解范围。这为宗教提供了切入点。宗教对心灵现象有一个直观的、可信的解释：我们有意识是因为我们有一个非物质的灵魂，我们真正的、内在的自我。灵魂是超越性现实的组成部分，超越了空间、时间和因果关系的范畴。这个灵魂努力在时间的终点与上帝合一。这些是我在虔诚的罗马天主教家庭中成长时被教导要相信的传统答案。

宗教和科学是理解世界、其起源及其意义的两种模式。历史上，它们彼此对立。自启蒙时代开始以来，西方的宗教一直在退缩，一败再败。一个挫折是哥白尼革命，它将地球从宇宙中心移到了包含千亿颗恒星的银河系边缘。但最沉重的打击是达尔文自然选择进化论。它将人类从上帝赐予的对地球的统治地位中移除，用一个跨越永恒的故事取代了《创世纪》的史诗故事，充满了喧嚣和愤怒，却毫无意义。进化将人类从崇高地位上推翻；我们只是无数物种中的一个。在我们基因的分子特征中，我们可以追溯到我们从灵长类谱系的血统，在遥远的深时中，从池塘浮渣开始。

因此，许多宗教教义与现代世界观不相容。这种不匹配并不令人惊讶，因为支撑伟大一神教宗教的神话和教义形成于对地球或栖息在其上的生物的大小、年龄和进化知之甚少的时代。

许多人认为科学会从人类的行为、希望和梦想中抽走意义，留下荒凉和空虚。分子生物学先驱雅克·莫诺德(Jacques Monod)令人心寒地表达了这种凄凉的情感：

人类最终必须从千年的梦想中醒来，发现自己完全孤独，根本上是孤立的。他必须意识到，就像一个吉普赛人，他生活在一个异域世界的边界上；这个世界对他的音乐充耳不闻，对他的希望漠不关心，就像对他的痛苦和罪恶漠不关心一样。

在我的大学时代，这句警句与弗里德里希·尼采(Friedrich Nietzsche)和其他人同样冰冷的片段一起，装饰着我宿舍房间的墙壁。最终，我反抗了这种存在主义的、宇宙冷漠的表达。

此时我必须坦白一件事。凭借完美的后见之明，我现在意识到吸引我研究意识的是一种强烈而完全潜意识的渴望，想要证明我本能相信生活是有意义的这一信念是正确的。我认为科学无法解释感觉是如何进入世界的。通过全身心投入意识研究并在这一努力中失败，我打算向自己证明科学不足以完全理解心-身分离的本质，它无法解释现象存在核心的根本奥秘，莫诺德的凄凉情感是错误的。最终，结果并非如此。因此，我的牙痛让我踏上了探索意识海洋的航程，以困难问题作为我的指路星。

我开始与弗朗西斯·克里克(Francis Crick)一起研究心-身问题，这位物理化学家与詹姆斯·沃森(James Watson)一起在1953年发现了DNA的双螺旋结构，这是遗传分子。这一独特事件开创了分子生物学时代，是革命性科学发现最有据可查和最著名的例子。它在1962年获得了诺贝尔奖的赞誉。

正如霍勒斯·弗里兰·贾德森(Horace Freeland Judson)在其分子生物学精彩历史《创世第八日》中所记述的，弗朗西斯随后确立了自己作为该领域首席智者的地位。在破译生命通用密码的狂热而令人眩晕的竞争中，其他人都向他和他的理论思想寻求指导。当这一目标实现时，弗朗西斯的兴趣从分子生物学转向了神经生物学。1976年，在60岁时，他投入这一新领域，同时从旧世界的剑桥迁移到新世界的加利福尼亚。

在十六年中，弗朗西斯和我一起撰写了二十多篇科学论文和论述。它们都专注于灵长类大脑的解剖学和生理学及其与意识的联系。当我们在1980年代末开始这项热爱的工作时，撰写有关意识的文章被视为认知衰退的迹象。退休的诺贝尔奖获得者可以这样做，神秘主义者和哲学家也可以，但自然科学领域的严肃学者不行。对于一位年轻教授来说，特别是还未获得终身职位的教授，表现出对心-身问题超出业余爱好的兴趣是不明智的。意识是一个边缘话题：研究生总是敏锐地察觉到长辈的道德观念和态度，当这个话题出现时，他们会转动眼珠，宽容地微笑。

但这些态度发生了变化。我们与少数几位同事——伯尼·巴尔斯(Bernie Baars)、内德·布洛克(Ned Block)、大卫·查默斯(David Chalmers)、让-皮埃尔·尚热(Jean-Pierre Changeux)、斯坦尼斯拉斯·德阿内(Stanislas Dehaene)、杰拉德·埃德尔曼(Gerald Edelman)、史蒂文·劳里斯(Steven Laureys)、杰兰特·里斯(Geraint Rees)、约翰·塞尔(John Searle)、沃尔夫·辛格(Wolf Singer)和朱利奥·托诺尼(Giulio Tononi)等——共同催生了意识科学。尽管仍不成熟，但这门新科学代表了真正的范式转变和一种共识，即意识是科学的研究的合法主题。

助产这门学科诞生的是脑成像技术的幸运发展，使人类大脑能够安全而常规地被可视化观察其活动。这些技术对流行文化产生了令人振奋的影响：磁共振成像(MRI)的大脑图像，带有其特有的热点，已成为标志性符号。它们可以在杂志封面、T恤衫和电影中找到。

研究意识的生物学基础已经变成了一个主流的、合法的研究课题。

在过去的二十五年半中，我在加州理工学院(Caltech)指导了一个由二十多名学生、博士后研究员和工作人员组成的团队，他们专注于此类研究。我与物理学家、生物学家、心理学家、精神病学家、麻醉师、神经外科医生、工程师和哲学家合作过。我参与了无数心理学测试。我让强磁场脉冲和微弱电流刺激过我的大脑，把头伸进MRI扫描仪观察我颅内的状况，并在睡眠时记录我的脑电波。

在这本书中，我重点讲述来自意识神经生物学现代研究前沿的故事。正如光预设了它的缺失——黑暗，意识也预设了无意识。正如西格蒙德·弗洛伊德(Sigmund Freud)、皮埃尔·雅内(Pierre Janet)和其他人在十九世纪末意识到的，我们头脑中发生的很多事情是我们的心智无法接触的——是无意识的。确实，当我们反省时，我们经常欺骗自己，因为我们只能接触到头脑中发生的事情的一小部分。这种欺骗解释了为什么关于自我、意志和我们心智其他方面的哲学在两千多年来一直如此贫瘠。然而，正如我将描述的，无意识可以深刻地影响我们的行为。我还详细讨论了自由

意志的相关问题，即发起行动的感觉，以及物理学、心理学和神经外科如何解开这个形而上学的结。在没有太多宣传的情况下，这些领域的发现已经解决了自由意志问题的一个关键方面。

最后，我描述了一个合理的意识定量理论，它解释了为什么某些类型的高度组织化物质，特别是大脑，能够产生意识。这个由神经科学家兼精神病学家Giulio Tononi发展的整合信息理论，从两个基本公理出发，继而解释世界中的现象学。这不仅仅是投机的哲学，而是导向具体的神经生物学洞察，构建一个能够评估动物、婴儿、睡眠者、患者以及其他无法表达其体验的生物意识程度的意识测量仪。这个理论具有深远的后果，与Pierre Teilhard de Chardin的预言性思想有些相似（稍后会更多提及他）。

天文学和宇宙学的发现揭示了物理定律有利于氢和氦之外稳定重元素的形成。这些定律惊人地精细调节，需要四种基本物理力的精确平衡。否则，我们的宇宙永远不会发展到氢和氦聚集成巨大燃烧团块的阶段——这些长寿的恒星为围绕它们运行的岩石行星提供无尽的能量流。构成这些行星及其土壤、岩石和空气表面的物质——硅、氧等等——都是在第一代恒星的核熔炉中创造的，并在它们爆炸性死亡痛苦中散布到周围空间。我们，从字面意义上说，就是星尘。这个动态宇宙受热力学第二定律支配：任何封闭系统的熵永远不会减少；或者换句话说，宇宙正在展开以达到最大的无序和均匀。但这并不排除稳定有序岛屿的形成，这些岛屿以周围自由能海洋为食。这个定律的无情运作创造了统计确定性，即在宇宙中的某些这样的小岛上，长链复杂分子最终会出现。

一旦这个关键步骤发生，下一步也很可能发生：创世纪——在原始地球的洞穴或池塘中以及其他地方，在外星天空下生命的创造。生物体日益增加的复杂性，在化石记录中显而易见，是推动进化的无情生存竞争的结果。

它伴随着神经系统的出现和感知的最初征象。大脑的持续复杂化，用Teilhard de Chardin的术语，增强了意识直到自我意识出现：意识反思自身。这个递归过程在数百万年前就开始于一些高度发达的哺乳动物中。在智人中，它达到了暂时的顶峰。

但复杂化并不止步于个体自我意识。它在持续进行，实际上，正在加速。在今天技术复杂且相互交织的社会中，复杂化正呈现出超个体的、跨大陆的特征。通过手机、电子邮件和社交网络提供的即时全球通信，我预见到一个时代，人类数十亿的人口及其计算机将在一个巨大的矩阵中相互连接——一个行星级的超心智。只要人类避免末日——热核世界末日或完全的环境崩溃——这个意识过度发达的网络没有理由不能扩展到行星，并最终超越星夜延伸到整个银河系。

现在你知道为什么神经心理学家Marcel Kinsbourne称我为浪漫的还原主义者：还原主义者，因为我在数十亿微小神经细胞及其数万个突触的无休止且千变万化的活动中寻求意识的定量解释；浪漫，因为我坚持认为宇宙有意义的轨迹，可以在我们头顶的天空中和我们内心深处被解读。意义在于其宇宙演化的范围，而不一定在其中个体生物的生命中。有一个天体音乐，如果我们仔细聆听，我们可以听到它的片段，甚至可能是它整体形式的暗示。

这本书的副标题包含了承诺性的词语“忏悔录”。在这一体裁的演进中，从在罗马帝国黄昏年代发明它的圣奥古斯丁，到今天的脱口秀和真人秀节目，总是有一丝，如果不是一股臭味的话，展示癖、自私自利和虚假的味道。我打算远离那些恶臭的腐败。我也在一个强大的专业禁令面前写作，该禁令反对引入主观的、个人的因素。这个禁忌就是为什么科学论文用干燥的第三人称写成：“已经证明……”任何事情都要避免暗示研究是由有血有肉的生物完成的，他们有着不那么纯洁的动机和欲望。

在接下来的页面中，我将告诉你关于我自己的情况，只要我的生活与这些问题相关：为什么我被激励——有意识或无意识地——追求某些问题？为什么我采取了特定的科学立场？毕竟，正是在我们选择从事什么工作中，我们揭示了很多关于我们内在驱动力和动机的东西。

在过去几年中，随着我人生轨迹开始不可避免的下降，我迷失了方向。我无法或不愿控制的激情导致了一个深刻的危机，迫使我面对我的信念和内心的恶魔。但丁在其地狱篇中的开篇诗句完美地描述了这一点：

在我们人生旅程的中途

我发现自己处在一片黑暗的森林中，

因为正确的道路已经迷失。

但在我过多地涉及这些深夜事务之前，让我告诉你一些关于我早期生活的情况，这些与我的科学和我看待大脑的方式相关。

## 第二章：

---

在此我写下了内心宗教与理性冲突的源泉，为什么我从小就想成为科学家，为什么我佩戴卡尔库鲁斯教授的胸针，以及我如何在人生后期获得第二位导师

---

想想你们所出身的种族；

你们生来不是为了像野兽一样生活，

而是为了追求美德和知识。

我用这简短的劝告，

让我的同伴们如此渴望这次航行，

以至于我几乎无法阻止他们。

我们把船尾转向黎明，

用桨作翼进行疯狂的飞行，

不断在左舷一侧前进。

—但丁·阿利吉耶里，《地狱篇》(1531)

## 一个崭露头角的科学家的无忧童年

---

我在快乐中长大，痴迷于知识、结构和秩序。我和两个兄弟在父母最好的自由天主教传统中成长，在这种传统中，科学——包括自然选择进化论——在很大程度上被接受为解释物质世界的理论。我是个祭坛侍童，用拉丁语背诵祈祷词，聆听格里高利圣咏以及奥兰德·德·拉叙斯、巴赫、维瓦尔第、海顿、莫扎特、勃拉姆斯和布鲁克纳的弥撒曲、受难曲和安魂曲。暑假期间，我们全家旅行参观无数博物馆、城堡以及巴洛克和洛可可教堂。父母和哥哥赞叹地凝视着天花板、彩色玻璃窗、雕像和描绘宗教形象的壁画，而母亲为了我们所有人的利益，大声朗读每个物品的详细历史。尽管我觉得这种强制性的艺术饮食极其无聊，即使今天看到母亲书架上的三卷本艺术指南也不禁打个寒颤，但我却爱上了几世纪前罗马祈祷词的神奇吟诵和这些作曲家的神圣与世俗音乐。

天主教会是一个博学的、遍及全球的、文化丰富的、道德上无可指摘的机构，拥有跨越两千年延续到罗马和耶路撒冷的不同断血统。它的教义问答为生活提供了一个历史悠久且令人安心的解释，这对我来说是有意义的。宗教提供的慰藉如此强烈，以至于我传承了下去。我和妻子在信仰中抚养孩子们，为他们施洗，饭前祈祷，周日上教堂，带他们参加初领圣体仪式。

然而多年来，我开始越来越多地拒绝教会的教导。我得到的传统答案与科学世界观不相容。我的父母以及耶稣会和奉献会老师教给我一套价值观，但在书本、讲座和实验室里我听到了不同的鼓点。这种紧张关系让我对现实有了分裂的看法。在弥撒之外，我不太思考罪恶、牺牲、救赎和来世的问题。我完全用自然的术语来推理这个世界、其中的人以及我自己。这两个框架，一个神圣一个世俗，一个用于周日一个用于一周的其余时间，彼此不相交。教会通过将我微不足道的生命置于上帝创造的浩瀚和他的圣子为人类牺牲的背景中来提供意义。科学解释了我发现自己所处的实际宇宙的事实以及它是如何形成的。

拥有两个不同的解释，用美丽的亚里士多德意象来说，一个用于月上界一个用于月下界，这不是一个严肃的智识立场。我必须解决这两种解释类型之间的冲突。由此产生的冲突是我几十年来的固定伴侣。然而我始终知道那里只有一个单一的现实，而科学在描述它方面正变得越来越好。人类不注定永远在认识论的迷雾中徘徊，只知道事物的表面现象而永远不知道它们的真实本质。我们可以看到一些东西；我们凝视得越久，理解就越好。

直到最近几年我才设法解决了这个冲突。我缓慢但确实地失去了对人格化上帝的信仰。我不再相信有人看顾我，在世界上为我干预，并将在历史之外、在末世中复活我的灵魂。我失去了童年的信仰，但我从未失去一切都应如此的持久信念！我在骨子里深深感到宇宙有我们能够实现的意义。

我父亲学过法律，加入了德国外交部，成为一名外交官。我母亲是医生，在医院工作了几年。她为了丈夫放弃了职业生涯，将相当大的抱负倾注在我们孩子身上。

我1956年出生在密苏里州堪萨斯城，比我哥哥迈克尔晚一年。今天，你看不出我的中西部出身，因为我保留着相当浓重的德国口音。我们两年后离开，开始了流浪的生活，在阿姆斯特丹待了四年，我弟弟安德烈亚斯就在那里出生。随后，我们家住在波恩，当时的西德首都。在公立小学和耶稣会Gymnasium读了两年后，是时候回到大西洋彼岸的渥太华了。我在一所天主教修会办的学校学英语。但时间不长，三年后我们要去摩洛哥的拉巴特。我在那个北非国家的法国学校——完全世俗化的笛卡尔中学注册入学(这也许可以解释我对这位特定哲学家的持久喜爱)。尽管不断变换地方、学校和朋友，需要掌握第三种语言，我表现很好，1974年以数学和科学学士学位毕业。

我很幸运，从小就知道长大后想成为什么样的人。小时候我梦想成为一名自然学家和动物园园长，在塞伦盖蒂研究动物行为。青春期开始时，我的兴趣转向了物理学和数学。我大量阅读关于太空旅行、量子力学和宇宙学的科普书籍。我热爱相对论旅行的悖论、跌落时间视界进入黑洞、太空电梯等等。我至今还记得读乔治·伽莫夫的奇幻小说

《汤普金斯先生在仙境》时的美好时光，书中的主人公探索着一个超现实的世界，在那里骑自行车就能达到光速。还有《汤普金斯先生探索原子》，在那个世界里，普朗克常数（描述作用量子大小的数字）如此之大，以至于台球都表现出量子行为。这些书塑造了我的青少年思维。每次用我的零花钱买一本科学平装书时，我都会满怀爱意地在上面写下我的名字并珍藏它，随身携带，有机会就阅读。

我的父母通过给迈克尔和我购买德国品牌*Kosmos*的实验设计套装来进一步培养我的科学好奇心。这些是精致的玩具，通过一系列动手实验来教授物理学、化学、电子学或天文学。其中一套从电学基本定律开始，然后组装电磁继电器和感应电机，最后我制作了AM和FM无线电接收器。我花费大量时间摆弄这些电子设备，这种硬件破解是今天的孩子很少体验到的。另一套教授无机化学原理。我用新掌握的知识配制黑火药。当我制造火箭筒时，本应引导火箭的金属杆熔化了（推进剂点燃速度不够快），我父亲介入了，终结了我短暂的武器设计师生涯。在这个过程中，他可能救了我的四肢和眼睛。

我父亲给我们买了一架5英寸反射式望远镜，这是一台令人敬畏的仪器。我清楚地记得在拉巴特我们家屋顶上的一个夜晚，迈克尔和我在瓦格纳《漂泊的荷兰人》的背景音乐中计算天王星在星图上的位置。当我将望远镜指向天空中估算的方位角和仰角时，那颗闪烁的行星缓缓飘入视野，我感到多么兴奋！这是对宇宙秩序的绝佳确认！

在北非逗留期间，我永远被丁丁历险记的冒险故事所吸引——这个比利时男孩的官方职业是记者，但实际上他是探险家、侦探和全能英雄；他的白色狐狸梗小狗雪花（法语原版中叫Milou）；他豪放的朋友阿道克船长；还有疯狂的科学家、常驻天才、聪明但心不在焉、几乎耳聋的卡尔库路斯教授。这是我遇到的第一批卡通人物，因为我的父母不赞成漫画，认为它们要么太幼稚要么太粗俗。我把全部二十四本丁丁书都给了我的孩子们，他们也很喜欢，显然没有任何不良影响。丁丁海报甚至装饰着我家的走廊。卡尔库路斯教授是那种超脱世俗的学者的原型，他理解维系宇宙的秘密，但在处理日常世界时却是个笨手笨脚的人。他对我年轻的心灵产生了如此深刻的影响，以至于自1987年4月我作为教授发表就职演讲那天起，我就一直在夹克翻领上佩戴着他的头像徽章。

在不同国家长大、就读不同学校、学习不同语言，让我比那些不太流动的朋友更能看透任何一种文化的特殊性和独特性，并欣赏潜在的普遍特征。这是让我在离开家时想成为物理学家的许多形成性原因之一。

1974年，我进入德国西南部的图宾根大学。图宾根是一个围绕城堡建造的古朴小学术城市，很像它更著名的竞争对手海德堡。在大学里，我在一个击剑兄弟会中体验到了兄弟情谊。如果你不熟悉条顿学术传统，可以想象童子军移植到一个浪漫的500年历史大学，你就能大概理解我的意思。我也开始接触并沉迷于酒精、女性、舞蹈、弗里德里希·尼采和理查德·瓦格纳的快乐和危险，有时过度如此。我度过了离家后的第一个圣诞节——一个朋友和我把自己隔离在一个偏远的村庄里，沉醉于阅读《查拉图斯特拉如是说》以及《特里斯坦与伊索尔德》和《尼伯龙根的指环》的歌词和音乐。我年轻、不成熟、书呆子气，我需要通过生活中喧闹而光荣的混乱来进行这种自我发现之旅。

1979年，我从图宾根大学毕业，获得物理学硕士学位。在这个过程中，我还获得了哲学辅修学位，这让我接触到理想主义，这是一种一元论形式，教导宇宙只是心灵的显现。

到那时我已经意识到，我没有成为世界级宇宙学家所必需的数学技能。幸运的是，大约在那个时候，我被计算机迷住了。吸引我的是它们创造一个完全在我控制下的自包含虚拟世界的承诺。在它们简化的环境中，所有事件都遵循程序员制定的规则——算法。任何偏差都可以追溯到错误的推理或不完整的假设。如果程序不工作，你只能怪自己，只能怪自己。我编写程序，最初在穿孔卡片上提交给大学的集中式计算机，用Algol和汇编语言为天体物理学家和核物理学家编程。

## 研究神经细胞的生物物理学

---

我也开始对大脑是一种处理信息的计算机这一概念产生了强烈的迷恋。这种痴迷源于一本启发性的著作《大脑的质地：给控制论思维者的神经解剖学》，作者是德意志-意大利解剖学家Valentino Braatenberg。Valentino有着超凡的个性，他是活生生的证明，证明一个人可以同时成为伟大的科学家、美学家、音乐家、享乐主义者和好人。

Valentino是图宾根马克斯·普朗克生物控制论研究所的所长。通过他，我在该研究所找到了为意大利物理学家Tomaso Poggio编写代码的工作。Tommy，大家都这样称呼他，是世界上信息处理理论的伟大理论家之一。他发明了从同一场景的两个不同视角提取立体深度的第一个功能公式。在他的指导下，我完成了论文工作，在计算机上建模研究单个神经细胞上的兴奋性和抑制性synapses如何相互作用。

让我简要离题解释一下贯穿本书的几个基本概念。像所有器官一样，神经系统由数十亿个网络化的细胞组成，其中最重要的是neurons。就像肾细胞与血液或心脏细胞截然不同一样，也有不同类型的neurons，可能多达一千种。它们之间最重要的区别是它们是兴奋还是抑制所连接的neurons。Neurons是高度多样化和复杂的处理器，通过synapses或与其他神经细胞的接触点收集、处理和广播数据。它们通过精细分支的dendrites接收输入，dendrites上布满了数千个synapses。每个synapse都会短暂地增加或减少膜的电导率。由此产生的电活动通过dendrites和细胞体中复杂的膜结合机制转化为一个或多个全有或全无的脉冲，即著名的action potentials或spikes。每个脉冲的振幅约为十分之一伏特，持续时间不到千分之一秒。这些脉冲沿着neuron的输出线axon发送，axon通过synapses连接到其他neurons。（一些特化的neurons将输出发送到肌肉。）因此，循环闭合。Neurons通过synapses与其他neurons交流。这就是意识的栖息地。

神经系统的力量不在于其组件蜗牛般的速度，而在于其大规模并行通信和计算能力：它能够以非常特定的synaptic模式将非常庞大且高度异质的neurons联盟在很长距离上连接起来。正如我三十年后将要证明的，正是从这些模式中产生了我们的思想。Synapses类似于晶体管。我们的神经系统大约有1000万亿个synapses连接着约860亿个neurons。

在Tommy的指导下，我解决了描述神经细胞周围膜内外电荷如何通过其dendrites的分支模式和synapses的架构进行转换的微分方程。今天，这种建模是常规且备受尊重的，但当时生物学家对使用物理学描述大脑事件感到困惑。在我第一次以海报形式向其他科学家展示我的研究的全国会议上，我被隔离到会议厅的后面。只有两位访客过来，其中一位在寻找洗手间，但足够礼貌地留下来与我交谈。那天晚上我喝醉了，怀疑自己是否选择了正确的领域。尽管遭遇这些挫折，我还是在1982年获得了生物物理学博士学位。

在我作为博士候选人的那些年里，我爱上了Edith Herbst并与她结婚。Edith是一名护士，在图宾根出生长大。在怀我们的儿子Alexander时，她将我的论文输入到研究所的主机中（只有128千字节的核心内存！）。当我的论文导师，我们德语中亲切地称为Doktorvater的人，成为麻省理工学院（MIT）的教授时，我们跟随他来到剑桥。25岁的我，作为博士后研究员在异国开始了新的生活。

MIT是一次智力爆发。我在心理学系和人工智能实验室待了四年。我可以自由地纯粹追求科学。与同一位导师待这么长时间是不寻常的，但对我的职业生涯很有帮助。Tommy和我至今仍有互动，这证明了博士父亲和儿子之间联系的持久性。

## Caltech、教学、研究和物理学家眼中的大脑

---

1986年秋，我带着家人——增加了女儿Gabriele——继续向西迁移，作为生物学和工程学助理教授来到加州理工学院。Caltech是美国最具选择性、最硬核的科学和工程大学之一，位于洛杉矶郊区帕萨迪纳。这里有宽阔的大道，两旁种植着棕榈树、橙树和橡树，坐落在圣加布里埃尔山脚下。我为能加入Caltech的教职员队伍感到无比自豪。

Caltech是一所灵活的私立大学——约280名教授和2000名本科生和研究生——致力于在逻辑、数学以及如何对自然世界进行推理方面教育最优秀和最聪明的人才。Caltech及其人员体现了大学——这些已经存在了八百年的机构——所有伟大和崇高的品质。它是最好意义上的象牙塔，为追求意识和大脑的本质提供了充足的自由和资源。

当人们发现我是一名教授时，他们问的第一个问题总是：“你教什么？”在公众眼中，教授完全由其传统的教师角色所定义。我享受“传授知识”的过程，教授各种课程。指导那些聪明而有动力的学生——他们毫不犹豫地指出错误或不一致之处——是最高层次的智力挑战，在情感上也很有成就感。在准备讲座或回答课堂问题时，我一次又一次地获得洞察，从意想不到的角度阐明了一个老生常谈的问题。

然而，我们这个群体的成员主要从他们的研究中获得自尊和价值感。我们在部落等级制度中的地位是由我们的研究成功程度决定的。研究是驱动我们的动力，也是我们最大的快乐源泉。衡量成功的标准是我们在高影响力、同行评议和竞争激烈的科学期刊上发表的论文数量和质量。

我们的发现在这个精英世界中产生的影响越大，我们的声誉就越大。教学在该领域的集体自我形象中只起次要作用。我们教授把大部分时间花在研究上——思考、推理和理论化，计算和编程，与同事和同事讨论想法，阅读大量文献，自己也为贡献内容，在研讨会和会议上发言，撰写无数的资助提案来推动研究机器运转并保持其润滑，当然，还要监督和指导那些设计、制造、测量、摇动、搅拌、成像、扫描、记录、分析、编程、调试和计算的学生和博士后研究员。我是一个约二十多名这样的研究人员团队的首领。

除了研究选择性视觉注意和视觉意识——稍后会详细介绍——我们还继续研究神经元的生物物理学。大脑是一个高度进化的器官，但它也是一个遵循能量守恒和电荷守恒等铁律的物理系统。高斯定律和欧姆定律调节着神经细胞内外的电荷分布及其相关电场。上述所有突触和放电过程都对插入大脑灰质的电极所检测到的电位有贡献。如果数万个神经元及其数百万个突触都很活跃，它们的贡献会累积成所谓的局部场电位(local field potential)。这种电活动的遥远回声在脑电图(EEG)记录的颅骨外永不停息的峰值和波谷中可见。局部场电位反过来又对单个神经元产生反馈。我们现在了解到，这种反馈迫使神经元同步其活动。

神经元的局部作用与它们集体产生并被包围的全局场之间的这种相互作用与电子硅电路非常不同，硅电路的设计者布置导线、晶体管和电容以避免相互干扰，将“寄生”串扰保持在最低水平。我对大脑的电场、这个场携带多少信息以及它在意识中的可能作用非常感兴趣。

对于研究大脑和心智的物理学家来说，令人震惊的是缺乏任何守恒定律：突触、动作电位、神经元、注意力、记忆和意识在任何有意义的意义上都不守恒。相反，生物学和心理学拥有大量的经验观察——事实。没有统一的理论，唯一的例外是达尔文的自然选择进化论——尽管它是一个极其强大的解释框架，但进化论是开放式的，不具有预测性。相反，生命科学有很多启发式的、半精确的规则，它们在特定的有机体尺度上捕获和量化现象——比如我在论文中研究的生物物理建模——而不追求普遍性。这使得这些领域的研究与物理学截然不同。

## 再次投入意识研究

---

当我在加利福尼亚开始工作时，我重新与Francis Crick相识。我第一次遇到他是在1980年夏天，他躺在图宾根郊外果园的一棵苹果树下。Francis来到镇上与Tommy谈话——这是他最喜欢的活动——讨论我们正在进行的树突和突触建模工作。

四年后，在另一个大陆上，Francis邀请我和MIT人工智能实验室的计算机科学家Shimon Ullman到Salk研究所待五天。Francis想了解关于Shimon和我刚刚发表的选择性视觉注意模型的一切。为什么是这种特殊的连接方案？涉及多少神经元？它们的平均放电率是多少？它们形成多少突触？它们的时间常数是什么？它们的轴突投射到丘脑的哪个部分？这能解释行为反应的快速性吗？这样的问题从早餐一直持续到下午晚些时候。休息后是晚餐和更多专注于大脑的对话。Francis不会闲聊。这让我喘不过气来。我钦佩他的妻子Odile如何数十年来应对这种强度。

几年后，Francis和我开始合作，每天通电话、写信、发邮件，每月我都会在他位于拉霍亚山坡上的家中长期居住，那里距离加州理工学院向南两小时车程。我们工作的重点是意识。尽管几代哲学家和学者都徒劳地试图解决心-身问题，但我们相信从神经科学角度的全新视角可能有助于解开这个戈尔迪亚斯之结。作为理论家，Francis的研究方法是安静思考、每日阅读相关文献——他能吸收大量文献——以及苏格拉底式对话。他对细节、数字和事实有着永不满足的渴求。他会不断地将假设组合起来解释某些事物，然后自己否定其中大部分。在早晨，他通常会用一些大胆的新假设轰炸我，这些假设是他在夜里失眠时想到的。我睡得更香，因此缺乏这样的夜间洞察。

在一生的教学、工作和与地球上一些最聪明的人辩论的过程中，我遇到过才华横溢和成就卓著的人，但很少遇到真正的天才。Francis是一个智力巨人，拥有我见过的最清晰、最深刻的头脑。他能够获取与其他人相同的信息，阅读相同的论文，却能提出全新的问题或推论。神经学家兼作家Oliver Sacks，我们两人的好朋友，回忆说遇见Francis的经历“有点像坐在智力核反应堆旁边……我从未有过如此白炽的感觉。”“有人说过，阿诺德·施瓦辛格在他作为宇宙先生的鼎盛时期，在别人甚至没有地方的地方都有肌肉。这句妙语，转换到理性思维上，正适用于Francis。

同样令人瞩目的是Francis是多么平易近人。他没有名人的态度。像James Watson一样，我也从未见过Francis表现得谦逊。但我也从未见过他傲慢。他愿意与任何人交谈，从地位卑微的本科生到诺贝尔奖得主同行，只要对话者能带给他有趣的事实和观察、令人惊讶的提议，或者他以前从未考虑过的问题。确实，他会很快对那些胡说八道或不理解自己推理错误的人失去耐心，但他是我见过的最开明的学者之一。

Francis是一个彻底的还原论者。他激烈反对任何哪怕有一丝宗教色彩或模糊思维的解释，这是他喜欢使用的表达。然而，无论是我的形而上学情感还是我们四十岁的年龄差距，都没有阻止我们发展深厚而持久的师生关系。他很享受与一个精力充沛、具有领域专门知识、热爱推测，并且有时敢于与他激烈争论的年轻人无休止地交流想法的机会。我非常幸运他喜欢我，基本上收养我为他的智力儿子。

现在让我来定义意识问题，并描述Francis和我探索其本质所采取的方法。

# 第三章：解释意识为何挑战科学世界观，如何以扎实的基础对意识进行实证研究，为何动物与人类共享意识，以及为何自我意识并不像许多人认为的那样重要

---

进化是如何将生物组织的水转化为意识的酒的？

—Colin McGinn，《神秘的火焰》(1999)

没有意识就什么都没有。你体验身体和世界——山脉和人们、树木和狗、星星和音乐——的唯一方式是通过你的主观体验、思想和记忆。你行动和移动，看和听，爱和恨，记住过去并想象未来。但最终，你只能通过意识以其所有表现形式遭遇世界。当意识停止时，这个世界也随之停止。

许多传统认为人类拥有心灵（或精神）、身体和超越性的灵魂。其他传统拒绝这种三分法，支持心-身二元论。古埃及人和希伯来人将精神置于心脏，而玛雅人将其定位在肝脏。我们现代人知道意识心灵是大脑的产物。要理解意识，我们必须理解大脑。

但问题就在这里。大脑如何将生物电活动转化为主观状态，光子从水面反射如何神奇地转化为对彩虹色水绿色山间小湖的感知，这是一个谜题。神经系统与意识之间关系的本质仍然难以捉摸，是激烈而无休止辩论的主题。

十七世纪法国物理学家、数学家和哲学家勒内·笛卡尔在他的《正确引导理性和在科学中寻求真理的方法论》中寻求终极确定性。他推理认为一切都可以怀疑，包括外部世界是否存在或他是否有身体。然而，他正在体验某些东西，即使他所体验的确切特征是妄想的，这是确定的。笛卡尔得出结论，因为他是有意识的，所以他存在：*Je pense, donc je suis*，后来翻译为*cogito, ergo sum*，或我思故我在。这句话点出了意识的根本重要性：它不是只有当你坐在山顶上的莲花座位上沉思肚脐、通过鼻孔哼着“Om”时才会降临的罕见状态。除非你在深度睡眠或昏迷中，否则你总是对某些东西有意识。意识是你生活的核心事实。

有意识的、体验中的观察者的单一视角被称为第一人称视角。解释一块高度组织化的物质如何能够拥有内在视角，这个问题一直困扰着科学方法，而科学方法在许多其他领域都被证明是极其富有成效的。

考虑一下NASA宇宙背景探测器(COBE)卫星所做的测量。1994年，COBE凭借其椭圆形、蓝绿色的全天空照片登上了头版新闻，照片上有黄色和红色的斑点。暖色调表示宇宙背景辐射温度的微小变化，这是产生我们宇宙的大爆炸的遗迹。宇宙学家可以聆听这次空间本身的巨大爆炸的回声，并推断早期宇宙的形状。COBE数据证实了他们的期望；也就是说，天文学可以对137亿年前发生的事件做出可检验的陈述！然而，像牙痛这样平凡的事情，就在此时此刻，仍然令人困惑。

生物学家还不能详细说明单个受精哺乳动物卵子内部的分子程序，这些程序将其转变为构成完全形成个体的肝脏、肌肉、大脑和其他器官的万亿个细胞。但毫无疑问，实现这样壮举的基本工具就在手边。在科学家兼企业家Craig Venter的领导下，分子工程师在2010年创造了一个新物种，实现了一个里程碑。他们对细菌的基因组（一段长达一百万个字母的单一DNA链）进行了测序，向其中添加了识别性水印，合成了构成基因组的基因（使用构成DNA的四种化学物质），并将它们组装成一条链，然后植入到已移除自身DNA的供体细菌的细胞体中。人工基因组成功地指挥了供体细胞的蛋白质制造机制，这个新生物被命名为*Mycoplasma mycoides JCVI-syn1.0*，一代又一代地复制下去。虽然创造一个新的细菌物种并不能制造出*golem*，但这仍然是一个惊人的行为，是历史上的分水岭时刻。以这种方式编程简单的多细胞植物和动物在理论上没有障碍，只有巨大的实践障碍。但这些最终会被克服。炼金术士的古老梦想——在实验室中创造生命——已经触手可及。

2009年，我参加了在加利福尼亚州山景城Googleplex举办的Science Foo (SciFoo) Camp。几百名超级极客、技术专家、科学家、太空爱好者、记者和数字精英聚集在一起，进行了一个周末的即兴研讨会和交流。关于人工智能的未来有很多讨论，一些人认为对真正人工智能的追求，比如达到六岁儿童的水平，已经被放弃了。然而，没有人怀疑那些智能将与我们匹敌，并最终超越我们的软件构造将会出现。虽然计算机科学家和程序员可能还需要几十年才能实现人类智能，但原则上，这样做并没有困难。SciFoo的与会者没有人认为达到这个目标是不可想象的。关于最佳实现方式存在争议，学习问题如何解决，人工智能对社会是好是坏等等都有讨论。但没有人质疑这个目标是可以实现的。

对于consciousness(意识)，没有类似的共识，对于理解其物理基础的可能性也没有共识。爱尔兰出生的物理学家John Tyndall发现了天空为什么是蓝色的，以及水蒸气和二氧化碳是地球大气中两个最大的吸热温室气体，他早在1868年就很好地阐述了将consciousness(意识)与大脑联系起来的困难：

从大脑的物理学到consciousness(意识)的相应事实的转变，作为力学的结果是不可思议的。假设一个明确的思想和大脑中一个明确的分子作用同时发生；我们不具备智力器官，也显然没有任何器官的雏形，能够使我们通过推理过程从一个现象转到另一个现象。它们一起出现，但我们不知道为什么。即使我们的思维和感官得到了扩展、强化和启发，能够看到和感受大脑的分子；即使我们能够跟踪它们所有的运动、所有的组合、所有的电放电(如果有的话)；即使我们对相应的思想和感觉状态非常熟悉，我们仍然远离问题的解决：“这些物理过程是如何与consciousness(意识)事实联系起来的？”两类现象之间的鸿沟在智力上仍然是不可逾越的。让love(爱)的consciousness(意识)与大脑分子的右旋螺旋运动相关联，hate(恨)的consciousness(意识)与左旋螺旋运动相关联。这样我们就知道，当我们love(爱)时，运动是朝一个方向的，当我们hate(恨)时，运动是朝另一个方向的；但“为什么？”这个问题仍然和以前一样无法回答。

这个挑战就是Chalmers所说的Hard Problem(难题)。

神经科学家通过显微镜和磁共振扫描仪观察神经系统，绘制其物理结构的细节图，用彩虹般的各种颜色染色神经元，并倾听猴子或人类观看图片或玩视频游戏时大脑内神经元的微弱窃窃私语。最新的技术热潮——很可能为开发这一技术的先驱们赢得诺贝尔奖——是*optogenetics*。这种方法针对动物大脑深处被改造病毒感染的特定神经细胞群。病毒使神经元长出只对特定波长光线响应的photoreceptors。然后可以用短暂的蓝光脉冲激活神经元，用同样短暂的黄光脉冲关闭神经元。演奏大脑的光之管风琴！Optogenetics之所以备受瞩目，是因为它允许研究人员在大脑紧密编织的网络中的任意点进行有意干预，从观察转向操控，从相关性转向因果性。任何具有独特基因条码的神经元群都可以以无与伦比的精确度被关闭或激活。我将在倒数第二章回到这项极具前景的技术。

所有这些技术都是从第三人称视角测量和干扰神经系统。但神经组织是如何获得内在的第一人称视角的？让我们回到150年前，回到Gottfried Leibniz的Monadology。他是被称为“最后一个什么都知道的人”的德国数学家、科学家和哲学家（他与他人共同发明了微积分和现代二进制数系统）。Leibniz在1714年写道：

此外，我们必须承认，感知以及依赖于它的的东西，无法用机械原理，即通过形状和运动来解释。如果我们想象有一台机器，其结构使它能够思考、感觉和产生感知，我们可以设想它被放大，保持相同的比例，这样人就可以进入其中，就像进入磨坊一样。假设当我们检查其内部时，我们只会发现相互推动的部件，而永远找不到任何可以解释感知的东西。

许多学者认为大脑机制与其散发的意识之间的这一鸿沟是无法跨越的。如果意识仍然无法得到解释，科学的解释领域就远比其从业者想要相信的要有限得多，也远比其宣传者从屋顶高声喧嚷的要有限得多。无法在定量的、经验可及的框架内解释知觉会是令人震惊的。

我不赞同这种失败主义态度。尽管解构主义者、“批判性”学者和社会学家的咄咄逼人，科学仍然是人类理解现实最可靠、累积性最强、最客观的方法。它远非万无一失；它充满了许多错误的结论、挫折、欺诈、从业者之间的权

力斗争以及其他人类的弱点。但在理解、预测和操控现实的能力方面，它远胜于任何替代方案。正因为科学在理解我们周围的世界方面如此出色，它也应该帮助我们解释我们内在的世界。

学者们不知道我们内在的精神世界为什么存在，更不用说它是由什么构成的。这个持续的谜题使意识令我的一些同事感到烦恼，令许多人厌恶。然而，意识对还原论理解的抗拒却令大众感到愉悦。他们贬低理性和那些服务于理性召唤的人，因为对意识的完全阐明威胁着关于灵魂、人类例外论、有机物优于无机物的长期而珍贵的信念。Dostoyevsky笔下的Grand Inquisitor深谙这种心态：“有三种力量，只有三种力量能够永远征服并俘获这些为了幸福而无力反抗者的良心——这些力量就是奇迹、神秘和权威。”

## Qualia和自然界

---

在这一点上，我需要介绍 *qualia*，这是心灵哲学家钟爱的概念。Qualia是 *quale* 的复数形式。拥有特定体验的感觉就是该体验的 *quale*：红色的 *quale* 是看到红色日落、中国红旗、动脉血、红宝石和 Homer 笔下酒色深海等如此不同感知的共同点。所有这些主观状态的共同分母是“红色性”。Qualia 是原始的感觉，是构成任何一种意识体验的元素。

一些 *qualia* 是基本的——黄色、腰部肌肉痉挛的突然而压倒性的疼痛，或 *déjà vu* 中的熟悉感。其他是复合的——我的狗依偎在我身边的气味和感觉、突然理解的“啊哈！”“，或当我第一次听到那些不朽台词时完全着迷的独特记忆：“我见过你们人类无法置信的事物。攻击舰在猎户座肩部附近燃烧。我看过了 c-beams 在 Tannhäuser Gate 附近的黑暗中闪闪发光。所有这些时刻都将在时间中消失，如同雨中的眼泪。该死去了。”“拥有体验意味着拥有 *qualia*，体验的 *qualia* 规定了该体验并使其区别于其他体验。

我相信 *qualia* 是自然世界的属性。它们没有神圣或超自然的起源。相反，它们是我想要揭示的未知定律的结果。

许多问题随之而来：感受质是物质本身的基本特征，还是只在极其复杂的有机系统中才会出现？换言之，基本粒子是否具有感受质，还是只有大脑才有？单细胞细菌是否体验某种形式的原始意识？那蠕虫或苍蝇呢？产生感受质需要最少多少个神经元？还是神经元连接的方式最重要？由硅晶体管和铜线组成的计算机能够有意识吗？正如菲利普·迪克修辞性地问道，机器人会梦见电子羊吗？我的 Mac 是否享受其内在的优雅，而我会计师的那台非 Mac 机器是否因其灰色外观和笨拙软件而痛苦？拥有数十亿节点的互联网是否具有感知能力？

在我的探索中，我不必从零开始。我们对意识和感受质有大量事实。最重要的是，我们知道感受质出现在某些高度网络化的生物结构中，包括专注观察者的中枢神经系统（比如你的）。因此，人类大脑必须是探索意识物理基础的起点。

不过，并非所有生物的、适应性的、复杂的系统都符合条件。你的免疫系统没有表现出意识的迹象。它日复一日地默默检测和消除各种病原体。你的身体防御系统可能现在就在对抗病毒感染，而你却没有意识到这一点。免疫系统会记住这个入侵者，如果同样的病毒再次袭击，它会产生抗体，给你终身免疫。然而这种记忆不是有意识的。

你肠道内壁纵横交错的 1 亿个神经元——肠神经系统，有时被称为第二大脑——情况也是如此。这些神经元默默地工作，负责胃肠道的营养提取和废物处理。这些你宁愿不知道的事情。有时，肠神经系统会发生异常——在重要的工作面试前你会感到胃里有蝴蝶在飞，或者在大餐后感到恶心。这些信息通过迷走神经传递到大脑皮层，然后产生紧张或沉重的感觉。你肠道中的第二大脑并不直接产生意识。

你的免疫系统和肠神经系统可能确实有自己的意识。你的心智，居住在头骨楼上，对肠道的感受一无所知，因为肠神经系统和中枢神经系统之间的交流有限。你的身体可能藏着几个自主的心智，永远孤立，如月球背面般遥远。目前，这种可能性不能完全排除；然而，鉴于肠神经系统有限的刻板行为，它似乎从属于大脑本体，没有独立体验的能力。

理解感受质如何产生只是消除心身问题中“问题”的第一步。接下来是理解特定感受质的特殊感受方式。为什么红色的感觉是那样的，与蓝色截然不同？颜色不是抽象的、任意的符号：它们代表某种有意义的东西。如果你问人们橙色是位于红色和黄色之间，还是蓝色和紫色之间，视力正常的人会选择前者。颜色感受质有着天生的组织结构。事实上，颜色可以排列成一个圆圈，即色轮。这种排列不同于其他感觉，如深度感或音调感，它们排列成线性序列。为什么？作为一个群体，颜色知觉具有某些共同特征，使它们不同于其他知觉，如看到运动或闻到玫瑰花香。为什么？

我寻求一种基于物理原理回答这些问题的方法，通过观察大脑的实际线路图，从其回路中推断出大脑能够体验的感觉。不仅是意识状态的存在，还有它们的详细特征。也许你认为这一壮举超出了科学的范围？不要这样想。诗人、词曲作者和安全官员都感叹不可能了解他人的内心。虽然从外部观察确实如此，但如果我能接触到一个人的整个大脑及其所有元素，情况就不同了。有了正确的数学框架，我应该能够准确判断她或他正在体验什么。无论喜欢与否，真正的读心术至少在原理上是可能的。第5章和第9章将详细讨论这一点。

## 意识的功能是什么？

---

另一个持续困扰的谜题是我们为什么要有任何体验。我们难道不能在没有意识的情况下生活、繁衍和抚养孩子吗？这样的僵尸式存在并不违反任何已知的自然法则。不过，从主观角度来看，这感觉就像在梦游中度过一生，或者像《活死人之夜》中那些可怕的不死族。没有知识。没有自我。没有存在。

我们心理体验的内在屏幕具有什么生存价值？意识的功能是什么，感受质的功能是什么？

随着意识到日常生活的起伏变化确实大部分发生在意识范围之外，这个谜团变得更加深刻。对于构成我们日常例行公事的大多数感觉-运动行为来说，这显然是真实的：系鞋带、在电脑键盘上打字、开车、回击网球发球、在岩石小径上跑步、跳华尔兹。这些行为在自动导航模式下运行，很少或没有意识反省。实际上，这些任务的流畅执行要求你不要过分专注于任何一个组成部分。虽然学习这些技能需要意识，但训练的目的是让你不再需要思考它们；你相信身体的智慧并让它接管。用耐克广告词来说，你“想做就做”。

Francis Crick和我假设每个人内部都存在一支由头脑简单的僵尸代理(*zombie agents*)组成的军队。这些代理专门负责可以自动化并在没有意识监督的情况下执行的刻板任务。你将在第6章读到关于它们的内容。这种无意识行为迫使我们审视意识的益处。为什么大脑不只是一堆专门的僵尸代理呢？既然它们工作毫不费力且迅速，为什么还需要意识呢？

因为生活有时会给你一个曲球！意外发生了，你必须在行动前思考。你上班的常规路线被拥堵的交通阻断，你考虑替代方案。Francis和我争论说，意识在这种情况下对规划很重要：在许多竞争的选择中思考和选择未来的行动方案。你应该等待交通堵塞自行解决，向左转走那条长长的、没有交叉路口的绕行路线，还是向右拐进内城区街道，面对无数红绿灯？这样的考虑需要将所有相关事实总结并呈现给心智的决策阶段。

这个论证并不意味着与规划相关的所有大脑活动都是意识可及的。很可能无意识过程也能规划，但比意识过程慢得多，或者无法看得像意识过程那样远。生物系统不像人工系统（想想你家里的电线），过于冗余、过于交织，当一个处理模块被击倒时不会完全失效。相反，Francis和我推测意识允许我们比僵尸代理更灵活地规划并看得更远。

功能问题是出了名的难以回答：为什么我们有两只眼睛而不是像蜘蛛那样有八只？蚓状阑尾的功能是什么？严格证明任何一个特定身体特征或行为在进化过程中为何出现是具有挑战性的，除了一般性术语，比如眼睛需要从远处探测猎物和捕食者。

一些学者否认意识有任何因果作用。他们接受意识的现实性，但争论主观感受没有功能——它们是存在海洋上的泡沫，对世界没有后果。技术术语是副现象的(*epiphenomenal*)。心脏跳动时发出的声音是一种副现象：它对诊断病人的心脏病专家有用，但对身体没有直接后果。Thomas Huxley，这位博物学家和达尔文的公共捍卫者，在1884年这样说过：

野兽的意识似乎与其身体机制的关系仅仅是其运作的附带产物，完全没有任何修改该运作的力量，就像伴随火车头工作的汽笛对其机械装置没有影响一样。

我发现这种论证路线不可信（但目前无法反驳它）。意识充满了有意义的知觉和有时难以承受强度的记忆。如果这种伙伴关系的感受部分对生物体的生存没有任何后果，为什么进化会偏爱神经活动与意识之间紧密而一致的联系，这种联系贯穿一生？大脑是一个选择过程的产物，该过程已经运行了数亿次迭代；如果qualia没有任何功能，它们就不会在这个无情的审查过程中幸存下来。

虽然关于意识功能的推测激发了哲学家、心理学家和工程师的想象力，但对其物质基础的实证探索以令人眩晕的速度进展。科学最擅长回答机制性的”如何？“问题，而不是终极的”为什么？“问题。沉迷于意识的效用作为研究途径，不如研究大脑的哪些部分对意识重要更有成果。

## 定义意识的困难

---

在关于动物意识的讲座后，一位女士走近我，强调地说：“你永远无法让我相信猴子是有意识的！”“而你也永远无法让我相信你是有意识的，”我反驳道。她被我的回应吓了一跳，但随后理解之光在她脸上闪现：无法感受猴子、狗或鸟的感受同样适用于人类，尽管程度较轻。考虑一个卧底间谍、欺骗的恋人或职业演员。他们伪造信任、爱国主义、爱情或友谊的感情。你永远无法对任何人的感受绝对确信！你可以观察他们的眼睛，分析他们说的话，但最终你无法通过观察真正了解他们的内心。

准确定义什么构成爵士乐是出了名的困难。因此有句话说，“兄弟，如果你还要问，你永远不会知道。”意识也是如此。要解释有意识的感觉而不涉及其他感觉、qualia或感知状态是不可能的。这种困难反映在意识的许多循环定义中。例如，《牛津英语词典》将意识定义为“有意识的状态，被视为健康清醒生活的正常状态。”“考虑一个相关但更简单的问题：向一个先天失明的人解释什么是颜色。要定义红色的感知而不涉及其他红色物体是不可能的。而且物体的颜色不同于其形状、重量或气味。它的颜色只与，嗯，其他颜色相似。

如果在达到透彻理解之前就试图下定义，定义可能是危险的：过早的严格性可能成为Procrustean床，限制进步。如果我看起来回避，试着定义一个基因。它是遗传传递的稳定单位吗？它必须编码单一酶吗？那么结构基因和调节基因呢？基因对应于核酸的一个连续片段吗？那么内含子、剪接和翻译后编辑呢？或者尝试更简单的东西，比如定义一颗行星。我从小学习九颗行星。几年前，我的加州理工学院同事Michael Brown领导的团队发现了Eris，一颗位于太阳系边缘的行星。Eris有自己的卫星，质量比冥王星更大。因为它也是众多所谓的海王星外天体之一，天文学家开始挠头。这意味着会有无穷无尽的新行星吗？为了避免这种情况，他们重新定义了什么是行星，并将冥王星降级为矮行星或小行星的地位。如果天文学家观测到围绕双星系统运行的行星，或者在太空中漫游而没有伴星照亮其天空的孤独行星，行星的定义将不得不再次适应。

一个习惯性的误解是，科学首先严格定义它所研究的现象，然后揭示支配它们的原理。从历史上看，科学的进步是在没有精确、公理化表述的情况下取得的。科学家使用可塑的、临时定义，随着更好的知识变得可用而适应它们。这样的工作定义指导讨论和实验，允许不同的研究社区互动，促进进步。

本着这种精神，让我提供四种不同的意识定义。就像佛教寓言中的盲人一样，每个人都在描述同一头大象的不同方面，每个定义都捕捉到了意识的一个重要方面，但没有一个能够描绘出完整的画面。

常识定义将意识等同于我们的内在精神生活。意识始于我们早上醒来时，持续整天直到我们进入无梦的睡眠。意识在我们做梦时存在，但在深度睡眠、麻醉和昏迷期间被放逐。在死亡时它永久消失。《传道书》说得对：“活着的人知道必死，死了的人毫无所知。”

意识的行为定义是一系列行动或行为的清单，任何能够做到其中一个或多个的有机体都可以被认定为有意识的。急诊室人员使用格拉斯哥昏迷评分快速评估头部外伤的严重程度。它根据患者控制眼睛、四肢和声音的能力给出分数。综合分数3对应昏迷，15表示患者完全清醒。中间值对应部分损伤：例如，“对询问感到困惑和迷失方向，但会因疼痛刺激而退缩。”虽然这种意识测量在处理儿童或成人时是完全合理的，但挑战在于提出适合那些不能轻易被要求做某事的有机体的行为标准，比如婴儿、狗、老鼠，甚至苍蝇。

进一步的困难在于区分刻板的、自动的反射与需要意识的更复杂行为（我将在第6章回到这些zombie行为）。在实践中，如果一个主体反复表现出有目的的、非常规的行为，涉及信息的短暂保持，她、他或它就被假定为有意识的。如果我对一个婴儿做鬼脸并伸出舌头，婴儿稍后模仿我，那么可以合理假设婴儿至少对其周围环境有一些基本的意

识。同样，如果要求患者将眼睛向左、向右移动，然后向上和向下移动，而她照做了，那么她是有意识的。同样，这样的任务必须适应被测试物种的手、爪、脚垫、爪子、鳍状肢或鼻子。

意识的神经元定义明确了任何一种意识感觉所需的最小生理机制。例如，临床医生知道，如果脑干受损，意识会急剧减少，甚至可能完全缺失，导致植物人状态。任何一种特定意识感觉的另一个必要条件是活跃且功能正常的皮质-丘脑复合体(cortico-thalamic complex)。这个复合体首先包括新皮质和其下方紧密相关的丘脑。新皮质是大脑皮质最新的部分，由构成著名灰质的神经元折叠层组成。它占据前脑的大部分，是哺乳动物独有的标志。丘脑是位于大脑中部一个鹌鹑蛋大小的结构，调节进入新皮质的所有输入并接收来自它的大量反馈。几乎大脑皮质的每个区域都从丘脑的特定区域接收输入，并将信息回传给它。与皮质-丘脑二元体归为一类的其他结构——因此称为“复合体”——包括海马体、杏仁核、基底神经节和屏状核。意识科学的大部分进展都发生在神经元领域。我将在第4章和第5章中详细阐述。

如果这三个定义仍不能解决问题，请咨询哲学家。她会给你第四个定义：“意识就是感受某种东西时的感觉”。特定体验的感觉只有经历这种体验的生物体才能知道。这种从内部感受的视角表达了现象意识的主要不可约特征——体验某种东西，任何东西。

这些定义都不是基础性的。没有一个用明确的术语描述了任何系统具有意识所需的条件。但就实用目的而言，行为和神经元定义最为有用。

## 关于动物的意识

---

为了在这些困难问题上取得进展而不陷入转移注意力的小冲突和激烈辩论，我必须做出一些假设，虽然现阶段还无法完全证明它们。这些工作假设在以后可能需要修正，甚至被拒绝。

作为起点，我假设意识的物理基础与神经元及其元素之间的特定相互作用密切相关。尽管意识完全符合物理学定律，但仅从这些定律预测或理解意识是不可行的。

我进一步假设许多动物，特别是哺乳动物，具有意识的某些特征：它们能看、听、嗅和其他方式体验世界。当然，每个物种都有其独特的感觉系统，与其生态位相匹配。但它们都体验着某些东西。相信其他情况是自大的，并且违背了动物与人类之间结构和行为连续性的所有证据。生物学家通过区分非人类和人类动物来强调这种连续性。我们都是大自然的孩子。我对此充满信心有三个原因。

首先，许多哺乳动物的行为是相似的，尽管不完全相同。以我的狗为例，当她叫喊、呜咽、啃咬爪子、跛行，然后来找我寻求帮助时：我推断她在疼痛，因为在类似条件下我也会有类似行为（除了啃咬）。疼痛的生理指标证实了这个推断——狗就像人一样，心率和血压升高，并在血液中释放应激激素。我们与动物共享的不仅是身体疼痛，还有痛苦。当动物被系统性虐待或年长宠物与同窝伙伴或人类伴侣分离时，就会发生痛苦。我并不是说狗的疼痛与人的疼痛完全一样，而是说狗——以及其他动物——不仅对有害刺激做出反应，而且有意识地体验疼痛。

其次，哺乳动物神经系统的结构是可比较的：需要专业的神经解剖学家才能区分取自小鼠、猴子和人的豌豆大小的大脑皮质块。我们的大脑很大，但其他生物——大象、海豚和鲸鱼——有更大的大脑。无论在基因组、突触、细胞还是连接层面，小鼠、猴子和人之间都没有质的差异。介导疼痛的受体和通路在物种间是相似的。经过长时间跑步后，我给我的狗服用与我相同的止痛药（剂量适合她的体型），她的跛行就消失了。因此，我认为狗的疼痛感觉很糟糕，就像人的疼痛一样。

尽管有这些相似之处，在硬件层面存在无数量化差异。这些差异累积起来的事实是，作为一个物种，智人(*Homo sapiens*)能够建造覆盖全球的互联网，计划核战争，等待戈多，这些是其他动物无法做到的壮举。但没有发现质的大脑差异。

第三，所有现存的哺乳动物彼此密切相关。胎盘哺乳动物在约6500万年前小行星撞击尤卡坦导致恐龙灭绝后，进化成今天看到的多样形式。类人猿、黑猩猩和大猩猩在最近600万年前与人类共享共同祖先。智人是进化连续体的一部分，而不是从天而降、完全有感知能力的独特生物。

意识可能是所有多细胞动物的共同特征。乌鸦、喜鹊、鹦鹉和其他鸟类；金枪鱼、慈鲷和其他鱼类；乌贼；以及蜜蜂都能够表现出复杂的行为。它们很可能也拥有某种觉知的萌芽，会感受痛苦和享受快乐。不同物种之间甚至同一物种的不同个体之间的差异在于，这些意识状态能够有多么分化、多么错综复杂。它们意识到的内容——它们觉知的内容——与它们的感官和生态位密切相关。各有各的特点。

意识状态的范围必定会随着有机体神经系统复杂性的降低而缩小。生物实验室中最受欢迎的两个物种——拥有302个神经细胞的线虫 *Caenorhabditis elegans* 和拥有10万个神经元的果蝇 *Drosophila melanogaster*——是否具有任何现象性状态 (phenomenal states)，目前还很难确定。在没有充分理解支持意识所需的神经元架构的情况下，我们无法知道动物王国中是否存在一条卢比孔河，将有感知能力的生物与那些没有任何感觉的生物分开。

## 关于自我意识

---

如果问人们认为意识的定义性特征是什么，大多数人会指向自我觉知。意识到自己，担心孩子的病情，想知道为什么感到沮丧或她为什么激起你的嫉妒，这被认为是感知能力的巅峰。

幼儿对自己行为的洞察力非常有限。如果他们还不到18个月大，他们不会在镜子中认出自己。行为心理学家使用这种镜子测试作为自我识别的黄金标准。婴儿被偷偷在额头或脸上做上班点或色块标记。面对熟悉的镜子时，婴儿会与镜像玩耍，但不会抓挠或试图从脸上去除标记，这与那些可以无休止地霸占浴室、对外表挑三拣四的青少年不同。除了人类，还有许多物种能够通过镜子测试（经过适当修改）：类人猿、海豚、大象和喜鹊。猴子会露出牙齿或以其他方式与它们的倒影互动，但它们没有意识到“那里的图像对应于我这里的身体”。这并不是说它们没有自我感，但至少没有一个视觉表征来与镜子中的外部视图进行比较。

一些学者从这种缺乏自我觉知的情况下得出结论，认为动物王国的绝大多数都是无意识的。按照这种标准，只有人类，甚至不包括他们的幼儿，才是真正有意识的。

有一个观察结果使这个结论变得不太可信。当你真正投入到世界中时，你对自己只是模糊觉知的。当我攀登山峰、悬崖和沙漠岩塔时，我最强烈地感受到这一点。在高高的峭壁上，生命处于最激烈的状态。在好的日子里，我会体验到心理学家Mihaly Csikszentmihalyi所称的心流。这是一种强大的状态，在这种状态下，我对周围环境有着敏锐的意识——指尖下花岗岩的质地、吹拂头发的风、照射背部的阳光，以及，永远永远是距离我下方最后一个支点的距离。心流与平滑流畅的动作相伴，是感知和行动的无缝整合。所有注意力都集中在手头的任务上：时间的流逝变慢，自我感消失。那个内在的声音，我的个人批评者，总是准备提醒我的失败，变得沉默。心流是一种狂喜状态，与深度冥想中的佛教徒的心境相关。

作家兼攀岩者Jon Krakauer在《艾格之梦》(Eiger Dreams)中很好地表达了这一点：

渐渐地，你的注意力变得如此高度集中，以至于你不再注意到破皮的指关节、痉挛的大腿、保持不间断专注的紧张。一种恍惚状态笼罩着你的努力，攀登变成了一个清醒的梦。日常生活中积累的负罪感和杂乱，所有这些都被暂时遗忘，被压倒性的目标清晰度和手头任务的严肃性从你的思绪中挤走。

这种自我觉知的丧失不仅发生在登山运动中，也发生在做爱、参与激烈辩论、摇摆舞或赛车时。在这些情况下，你处于此时此地。你在世界中，属于世界，对自己几乎没有觉知。

以色列魏茨曼研究所的Rafael Malach付钱给志愿者，让他们躺在脑成像机器的狭窄空间内观看《好人、坏人与丑人》。虽然这不是你平均的电影体验，但观众们仍然觉得这部经典意大利式西部片很吸引人。在分析他们的大脑模式时，Rafi发现与内省、高级认知、规划和评估相关的大脑皮层区域相对不活跃，但涉及感觉、情感和记忆过程的区域却异常忙碌。更重要的是，脑扫描仪追踪到的皮层血流的起伏在志愿者之间是相似的。这两个观察结果都反映了天才意大利导演Sergio Leone控制观众的精湛技艺，让他们看到、感受到和回忆起他想让他们看到、感受到和回忆的东西。这就是我们喜欢看电影的一个原因——它们让我们从过度活跃的自我意识中分心，从我们日常的忧虑、焦虑、恐惧和疑虑的轰炸中分心。几个小时内，我们逃脱了我们颅骨大小的王国的暴政。我们高度意识着故事中的事件，但只是稍微觉知我们自己的内在状态。有时这是一个巨大的祝福。

如实际情况所示，大脑皮层前部广泛退化的人会有严重的认知、执行、情感和计划缺陷，同时对自己糟糕的状况缺乏洞察力。然而，他们的知觉能力通常是保持完好的。他们能看、能听、能闻，并且对自己的知觉有意识。

自我意识是意识的重要组成部分。这是一种特殊的觉知形式，不关注外在世界，而是专注于内在状态、对这些状态的反思，以及对这种反思的再反思。这种递归性使其成为一种特别强大的思维模式。

计算机科学家Doug Hofstadter推测，自我意识的核心是一种矛盾的、自我指涉的奇异循环，类似于Escher那幅一双手互相绘制对方的版画。如果真是如此——我对此持怀疑态度，因为心智很难进行自我指涉或递归，通常超不过“我在思考自己在思考”这个层面——那么这种奇异循环对意识本身来说是偶然的。相反，自我意识是身体和疼痛意识这些更古老形式的进化适应。

另一个独特的人类特征是言语。真正的语言使智人能够表征、操作和传播任意的符号和概念。语言产生了大教堂、慢食运动、广义相对论理论和《大师与玛格丽特》。这些都超出了我们动物朋友的能力范围。语言对文明生活大多数方面的首要地位，使得哲学家、语言学家和其他人产生了一种信念，即没有语言就不可能有意识，因此，只有人类才能感受和内省。

我反对这种观点。没有理由因为动物不会说话就否认它们有意识，也没有理由因为早产儿的大脑没有完全发育就否认他们有意识。更没有理由否认严重失语症患者的意识，这些患者在康复后能够清楚地描述他们在无法说话时的体验。长期的内省习惯使许多知识分子贬低生活中大部分无反思的、非语言的特征，并将语言提升到主宰者的地位。毕竟，语言是他们的主要工具。

那么情感呢？一个有机体必须感受到愤怒、恐惧、厌恶、惊讶、悲伤或兴奋才能有意识吗？尽管毫无疑问，这些强烈的情感对我们的生存至关重要，但没有令人信服的证据表明它们对意识是必不可少的。无论你是愤怒还是快乐，你都会继续看到面前燃烧的蜡烛，当你把手指伸进火焰时也会感到疼痛。

有些人由于严重抑郁症或大脑额叶损伤而情感平淡；他们的行为受到损害，判断力不健全。受伤的退伍军人可以冷漠地回忆起在他的悍马车下爆炸的地雷爆炸声，那次爆炸夺走了他的双腿。他似乎完全冷漠，疏离，对自己的状况漠不关心，但毫无疑问他正在体验着什么，哪怕只是来自伤口的极度不适。情感对于平衡而成功的生活不可或缺，但对意识来说并非必需。

我们已经完成了清理工作，这为解决心身问题的真正战场做好了准备。我刚才提到的那个病人提醒我们大脑的核心地位。神经科学教科书详细描述了这个器官，令人麻木，却遗漏了拥有大脑的感受。让我通过将体验主体的内在视角与脑科学家的外在视角联系起来，来弥补这个显著的遗漏。

## 第4章：

---

在这一章中，你会听到科学家魔术师的故事，他们让你看但看不见，他们如何通过窥视你的头骨来追踪意识的足迹，为什么你不是用眼睛看的，以及为什么注意力和意识不是同一回事

---

“数据！数据！数据！”他不耐烦地喊道。“没有粘土我造不出砖头。”

——《铜山毛榉案》中的夏洛克·福尔摩斯，阿瑟·柯南·道尔著（1892年）

人们愿意承认，在核物理学或肾透析方面，专业知识是必需的。但当话题转向意识时，每个人都会插嘴，假设在缺乏相关事实的情况下，他们都有权拥有自己偏爱的理论。没有什么比这更错误的了。

关于大脑和心智已经积累了大量的心理学、神经生物学和医学知识。全世界五万多名大脑和认知科学家的努力每年都为这个庞大的集合增添数千项新研究。

但别担心。我不会对这些材料的一小部分进行说教。相反，我将专注于实验室中的几个小插曲，这些插曲体现了现代寻找意识根源的探索。

## 在大脑中寻找意识

---

在上个千年的最后十年，一些勇敢的科学探索者厌倦了无休止的、争辩性的形而上学论证——意识真的存在吗？它独立于物理定律吗？心理状态的意向性是如何产生的？哲学家们以其无忧无虑的方式发明的众多”-主义”中，哪一种最能描述心智与身体之间的关系？——他们开始在大脑中寻找意识的足迹。这种实干的态度带来了巨大的概念进步——意识开始被设想为特定大脑机制的产物。

在1990年代早期，Francis Crick和我专注于我们所称的意识的神经（或神经元）*correlates* (NCC)。我们将其定义为共同足以产生任何一个特定有意识知觉的最小神经机制。（我们的定义没有涉及在机器或软件中创造意识的可能性，这是一个故意的遗漏，我稍后会回到这个问题。）

想象一下，你正在看一个红色立方体，神秘地留在沙漠的沙子里，有一只蝴蝶在它上方翩翩飞舞。你的大脑在瞬间理解了这个立方体。它能完成这一壮举是因为大脑激活了代表颜色的专门皮质神经元，并将它们与编码深度知觉的神经元结合起来，以及编码构成立方体的各种线条方向的神经元。引起有意识知觉的这些神经元的最小集合就是感知这个外来物体的意识神经*correlate*。

强调”最小”这一点很重要。如果没有这个限定词，整个大脑都可以被认为是一个*correlate*：毕竟，大脑确实日复一日地产生意识。但Francis和我追求的是更大的目标——产生、引起意识的特定突触、神经元和回路。作为严谨的科学家，我们使用更谨慎的”*correlates*”而不是更明确的“意识”原因”。

大脑的某些区域与意识内容有着比其他区域更亲密、更特权的关系。大脑不像全息图，在全息图中每个部分都对图像有相等的贡献。有些区域贡献很少，如果有的话，可以在不损失现象体验的情况下被损坏，而其他区域对意识至关重要。

你真的需要脊髓来有意识地看东西吗？偏瘫和四肢瘫痪患者失去了脊髓损伤以下身体的感觉和控制，但他们仍然清楚地意识到世界，并且可以过着与任何人一样丰富和有意义的生活。想想扮演超人的演员Christopher Reeve。在一次骑马事故后从颈部以下瘫痪，他创立了一个医学基金会，并成为干细胞研究和其他康复技术的雄辩发言人，这些技术将使像他这样的四肢瘫痪患者有一天能够重新获得活动能力。

那么小脑呢，头部后面、大脑皮质下面的小脑？“小”这个绰号很讽刺，因为小脑有690亿个神经细胞，是占尽风头的著名大脑皮质神经细胞数量的四倍多。如果中风或肿瘤袭击你的小脑，你的平衡和协调能力会受损。你的步态笨拙，站姿宽阔，你拖着脚走路，眼球运动不稳定，言语含糊不清。你曾经视为理所当然的流畅精细动作现在变得生硬，必须费力地有意识地控制。弹钢琴或打网球成了遥远的记忆。然而你的知觉和记忆没有太大变化，如果说的话。声音、视觉、触觉和嗅觉保持不受影响。

将脊髓或小脑损伤与大脑皮质或海马体lesion——被某种过程破坏的局限性组织——的影响进行对比。根据破坏发生的位置，你可能无法回忆起任何东西，甚至是你的孩子的名字；颜色可能从世界中消失；或者你可能无法识别熟悉的面孔。从这些临床观察中，在下一章中会更充分地描述，科学家们得出结论，大脑皮质及其卫星区域中离散区域的生物电活动对有意识体验的内容至关重要。

哪些大脑区域对意识至关重要的问题正在激烈辩论中。南加州大学的神经学家Antonio Damasio认为大脑皮质后部的顶叶部分是必不可少的。其他人指向”前脑岛”、“上颞叶多感觉”皮质，或大脑的其他区域。随着时间的推移，需要一个既必要又充分的意识区域列表，但现在这仍然是一个遥远的目标。

识别大脑的特定区域为关键只是开始。这就像说谋杀嫌疑人住在东北部的某个地方。这还不够好。我想要定位到那些大脑区域中介导特定体验的关键特定回路、细胞类型和突触。Francis和我提出，任何意识神经correlate的一个关键组成部分是位于大脑皮质后部的高阶感觉区域与位于前部的前额叶皮质的规划和决策区域之间的长距离、相互连接。我稍后会回到这个想法。

每种现象的、主观的状态都是由大脑中特定的物理机制引起的。有一个回路用于在照片中或生活中看到你的祖母，另一个用于听到风在山顶松树间低语的声音，第三个用于在城市交通中骑自行车快速穿梭时的替代性兴奋感。

这些不同qualia的神经correlates之间有什么共同点？它们都涉及相同的回路模式吗？或者，它们都涉及X区域的神经元？这三种感觉都是由前额叶皮质中活跃的锥体神经元介导的，这些神经元将其输出线索蛇行回相关的感觉区域？介导现象内容的神经元是否以节律性和高度协调的方式放电？这些是Francis和我一直在考虑的一些想法。

扰动任何特定意识体验的神经关联物将改变知觉。破坏或以其他方式使相关神经元静止将使知觉消失，尽管大脑，特别是年轻的大脑，可以在几周内补偿有限的损伤。

人工诱导意识的神经关联物将触发相关的知觉。这在神经外科中是常规操作：外科医生将电极放在大脑表面并通过它发送电流。根据位置和强度，这种外部刺激可以触发深刻的记忆、多年前听过的歌曲、想要移动肢体的感觉或运动的感觉。

可能性的极限在”构建边缘”电影《黑客帝国》中得到了探索。在通过头部和脊髓后部插孔连接的电子有机链接的帮助下，机器通过刺激适当的意识神经关联物在Neo的大脑中创造了一个完全虚构而完整的世界。只有通过拔掉这个刺激器的插头，Neo才能”觉醒”到他正在一个巨大的笼子堆中虚度生命的事，被类似昆虫的机器人照料着，这些机器人饲养人类以从他们的身体中获取能量。

你也会在头脑的私密空间中每夜产生幻觉。在睡眠中，你会有生动的、有时在情感上令人痛苦的现象体验，即使你不记得其中的大部分。你的眼睛是闭着的，然而做梦的大脑构建着它自己的现实。除了罕见的”清醒”梦境，你无法区分做梦和清醒意识之间的差别。梦境在持续时是真实的。你能对生活说出更多吗？

具有讽刺意味的是，当你在心中表演你的梦境时，你沉睡的身体很大程度上是瘫痪的：你的大脑限制身体运动以保护自己免受你梦中有时暴力运动的伤害。这种瘫痪暗示行为对意识来说并非真正必要。成年大脑，即使与大部分输入和输出切断，也足以产生那种神奇的东西——体验。古老的哲学家老生常谈——缸中之脑，在《黑客帝国》中看到了现代复兴——就足够了。

通过专注于相关神经回路来发现和表征意识的神经关联物是许多当代研究的主题，特别是视觉研究。

## 在明处隐藏物体

---

几年前，我与一位技艺高超的小偷Apollo Robbins在纽约市度过了几天。他住在拉斯维加斯，是一名专业魔术师，擅长所有常见的舞台戏法和魔术技巧。然而，当他坐在咖啡馆里我旁边时——没有雾气、聚光灯、比基尼助手和音乐，所有这些都用来分散观众注意力——他最具说服力。他从空气中变出硬币，向我扔纸球然后突然消失，还偷走了我的手表——所有这些都发生在我注视着他的每一个动作时。而我还自称是视知觉专家！Apollo最令人印象深刻的表演之一是当他从我儿子那里拿了一张卡片并把它贴在自己的额头上。卡片就在那里明摆着，然而我儿子对它去了哪里感到困惑，因为他的注意力被魔术师的手牢牢吸引住了。

我在与Apollo和其他类似人员共处期间学到的是，小偷和魔术师是操控观众注意力和期望的大师。如果Apollo能够误导你的目光或注意力到他的左手，那么你将对他右手的动作视而不见，即使你正在看着它。

你注意力的空间焦点被称为注意力聚光灯。任何被这个内在灯笼照亮的物体或事件都会得到优先处理，可以更快地被检测到且错误更少。当然，有光的地方也有阴影——那些没有被注意到的物体或事件往往只能勉强被感知。

让我们从繁忙的曼哈顿咖啡馆转移到幽闭恐惧症般的棺材状磁共振成像(MRI)扫描仪内部的嘈杂敲击声中，这是一台重达数吨的庞大机器。你躺在狭窄的圆筒内，拼命保持静止，甚至不敢眨眼，因为任何运动都会使信号模糊。通过一面镜子，你凝视着显示一副扑克牌中红心A的计算机显示器，同时机器监测你大脑中的血流。神经科学家不擅长戏法，所以他们通过向你的眼睛投射精确定时的第二个图像来操控你所看到的。如果操作得当，这种误导与魔术师的一样有效——你将看不到红心A。第二个图像掩蔽了第一个，使A变得不可见。你看，但看不见，颠覆了Yogi Berra的名言“通过观察你可以观察到很多”。

这项技术由我当时的研究生Naotsugu Tsuchiya完善，他称之为连续闪烁抑制。它通过将扑克牌的图像投射到你的一只眼睛中，同时持续向另一只眼睛闪烁大量明亮彩色的重叠矩形——就像荷兰画家Piet Mondrian的作品——来工作。如果你用这只眼睛眨眼，红心A就变得可见，但如果你保持双眼睁开，A会持续数分钟保持隐藏，被不断变化的彩色矩形显示伪装，分散你的注意力。

如此强大的掩蔽技术是视觉研究蓬勃发展的原因之一。有了它们，让人们看而不见、操控他们的感知变得相对容易——远比欺骗其他感官要容易。嗅觉和自我认知感更加稳固，不易被操控。我无法让你将玫瑰的香味与蓝色Stilton奶酪的气味混淆，或让你相信你一会儿是英格兰女王，下一刻又是Madonna。

最好的实验是将除了感兴趣的变量之外的所有内容都保持固定。这样，这一个因素对整个系统运作产生的差异就可以被分离出来。神经科学家使用磁性扫描仪来比较你看到红心A时的大脑活动与因为被遮挡而看不到红心A时的大脑活动。通过考虑这种差异，他们可以分离出主观体验看到红心A时独有的活动，通过追踪其足迹来追踪意识到其巢穴。

红心A的图像刺激你眼中的神经元，称为视网膜神经节细胞。它们的轴突构成视神经，将视觉信息发送到大脑本体。这些细胞对红心A的反应是爆发出动作电位，即第2章中描述的简短的全有或全无脉冲。眼睛的输出并不取决于眼球的主人是否有意识。只要眼睑是睁开的，视神经就忠实地发出外界存在什么的信号，并将其传递给大脑皮层的下游结构。这种活动最终触发形成一个稳定的活跃皮层神经元联盟，传达对红色A的有意识感知。这是如何发生的是第8章的主题。

然而，面对另一只眼睛的视神经上不断涌入的更加剧烈的尖峰波的竞争，没有联盟能够形成，另一只眼睛正受到不断变化的彩色矩形的刺激。这些尖峰触发它们自己的神经联盟，结果是你看到闪烁的彩色表面，而红色A仍然不可

见。

为了理解下一个实验，让我简要解释核磁共振成像的基本原理（由于负面含义，“核”这个词已被省略）。MRI扫描仪产生强大的磁场，大约比地球磁场强10万倍。某些元素的原子核，包括氢，表现得像微型条形磁铁。当你进入扫描仪的强磁场时，你体内的氢原子核与这个场对齐。你体重的一半以上是水，水由两个氢原子和一个氧原子组成。MRI扫描仪向你的头骨发送简短的无线电波脉冲，将原子核从对齐状态中敲出。当原子核放松回到其原始状态时，它们发出微弱的无线电信号，这些信号被检测并处理成数字图像。这样的图片揭示了软组织的结构。例如，它划定了大脑灰质和白质之间的边界。磁共振成像比X射线敏感得多。它通过允许定位和诊断从肿瘤到外伤的组织损伤，以微不足道的患者风险彻底革命了医学。

虽然MRI使器官的内部结构可见，功能性MRI (fMRI) 依赖于血液供应的一个巧合特性来揭示区域性大脑活动。活跃的突触和神经元消耗能量，因此需要更多氧气，氧气由循环红血球内的血红蛋白分子输送。当这些分子将氧气释放给周围组织时，它们不仅改变颜色——从动脉红变为静脉蓝——而且还变得轻微磁性。神经组织的活动导致新鲜血液的体积和流量增加。这种血液供应的变化，称为血流动力学信号，通过向头骨发送无线电波并仔细监听其返回回声来跟踪。请记住，fMRI不直接测量在毫秒过程中发生的突触和神经元活动。相反，它使用一个相对迟缓的代理——血液供应的变化——在几秒钟内上升和下降。fMRI的空间分辨率限制在豌豆大小的体积元素（体素），包含大约一百万个神经细胞。

那么，大脑如何对心智看不见的事物做出反应呢？

值得注意的是，不可见的图片可以在皮层中留下痕迹。无意识处理的丝缕可以在初级视觉皮层 (V1) 中被捕获。初级视觉皮层是从眼睛发送信息的终点。位于头部后面隆起的正上方，它是评估图像信息的第一个新皮层区域。大脑皮层的其他部分也对被抑制的图片做出反应——特别是延伸到初级视觉皮层之外的高阶视觉区域级联 (V2、V3等)，以及杏仁核，这个杏仁大小的结构处理情绪刺激，如恐惧或愤怒的面孔。

作为规则，视觉大脑的区域离视网膜越远，意识的影响就越强烈。随着期望、偏见和记忆在大脑高级区域发挥更大作用，外部世界的影响相应减弱。主观心智在皮层的上层表现得最为强烈。那是它的栖息地。

这暗示并非任何皮层活动都足以产生有意识的感觉。尽管初级视觉皮层中的一百万个神经元正在忙碌地放电，如果更高层的神经元没有反映这种活动，它们旺盛的尖峰可能不会产生体验。需要更多。也许这种活动必须超过某个阈值？也许皮层后部的一组特定的特殊神经元必须与额叶的神经元建立相互对话？也许这些神经元必须共同活跃，像同步游泳运动员一样一致地发射动作电位？我怀疑这三个条件都必须满足，信息才能被有意识地感知。这些结论激发了关于视觉处理层次早期神经元在多大程度上负责产生有意识感知的激烈辩论。

在明尼苏达大学Sheng He及其团队进行的一项值得注意的实验中，志愿者用一只眼睛看裸体男性或女性的照片，而彩色矩形不断在另一只眼睛中闪烁，使裸体图像变得不可见。尽管如此，敏感测试表明，不可见的裸体女性图片吸引了异性恋男性的注意力，而裸体男性的图像则让他们反感。然而这发生在*sub rosa*，在意识的雷达之下。志愿者看不到裸体，但仍然给予了关注。相反，异性恋女性和男同性恋者的注意力被不可见的裸体男性图片所吸引。从功能上讲，这是有道理的，因为你的大脑需要了解潜在的配偶。这也证实了一个关于欲望无意识本质的普遍陈词滥调。

## 并非所有神经元都参与意识

---

1995年，Francis和我在国际期刊《Nature》上发表了一份手稿，标题为“我们是否意识到初级视觉皮层中的神经活动？”（在《Nature》上发表文章就像在纽约或巴黎的顶级画廊展示你的艺术作品；这是一件大事。）我们对这个修辞性问题给出了否定的答案，论证视觉感知的神经关联不会在那里找到。我们基于猕猴的神经解剖学提出了这一假设。

猴子和人类都属于灵长类动物。我们最后的共同祖先生活在2400万到2800万年前。猴子的视觉系统与我们相似；它们很好地适应圈养；可以轻易训练；而且不是濒危物种。因此，猕猴是对高阶感知和认知感兴趣的科学家们的首选物种。出于明显的伦理原因，人类大脑不容易接受侵入性探测。因此，我们对猴子大脑中的线路了解比人类大脑多得多。

锥体神经元(pyramidal neurons)是大脑皮层的主力军。它们约占皮层神经元总数的五分之四，是唯一能将信息从一个区域传递到皮层内外其他区域（如丘脑、基底神经节或脊髓）的神经元。初级视觉皮层中的锥体细胞将信息发送到其他区域，如V2和V3，但它们都无法一直到达大脑前部。正是在那里，在前额叶皮层，特别是在其背外侧区域，高级智力功能——问题解决、推理和决策制定——位于此处。这一区域的损伤会使人的感觉形态和记忆保持完整，但做出理性决策的能力会受损：她会参与并持续进行极不合适的行为。

在前一章中，我提到了Francis和我的直觉，即意识的功能是规划。失去部分或全部前额叶皮层的患者在为近期或远期未来制定计划方面有困难。我们认为这意味着意识的神经关联必须包括前额叶皮层中的神经元。因为初级视觉皮层中的神经元不会将其轴突发送那么远，我们得出结论，位于初级视觉皮层中的神经元不是支撑视觉意识的神经元的组成部分。

我们的提议是违反直觉的，因为在许多条件下，初级视觉皮层中神经元的生物电活动确实反映了一个所看到的。以现在为例，当你正在阅读这些文字时：构成字母的笔画激发视网膜神经元，将其内容传递到初级视觉皮层。从那里，信息被传递到皮层视觉词形区域。你看到这些字母的能力似乎推翻了我们的假设——初级视觉皮层中神经元的活动可以与有意识的感知相关联。视网膜神经元也是如此——它们的活动可以镜像你所看到的。

尽管视网膜和初级视觉皮层中细胞的反应在某些时候确实与视觉意识共享某些属性，但在其他条件下它们的反应可能差异很大。让我给你举三个例子，说明为什么你不是用眼睛看东西的——画家们几个世纪以来就知道这一点。

首先，考虑你在浏览这些行时眼睛的不断运动。你每秒钟进行几次快速眼动，称为扫视(saccades)。尽管你的图像传感器几乎在不断运动，但页面看起来是稳定的。这应该让你感到惊讶。当你向右扫视时，世界应该向左移动——但它没有！考虑如果你以同样的节奏移动摄像机穿过书本会发生什么：观看由此产生的电影会让你恶心。为了避免这种情况，电视摄像机缓慢地扫过场景；它们的运动与你在捕捉图像突出方面时眼睛跳动的方式完全不同。如果视网膜神经元是传达静止世界感知的神经元，它们就必须专门信号外界的运动，但不对其所在眼睛的运动做出反应。然而，视网膜神经细胞，就像初级视觉皮层中的细胞一样，无法区分物体运动和眼球运动。它们对两者都有反应。与智能手机不同，它们没有附加的加速度计或GPS传感器来区分外部图像运动和内部眼部运动。是视觉皮层更高层的神经元产生了你对世界是静止的感知。

其次，考虑你视网膜上在盲点处的“空洞”：这是视神经离开视网膜的地方；组成神经纤维的轴突取代了光感受器，因此没有传入的光子在盲点处被捕获。如果你能意识到视网膜细胞的信息内容，你在那里什么都看不出来，就像你看不到视野边界在极左侧和极右侧一样。如果你手机的相机包含几个总是关闭的坏像素，你会很恼火：照片中的黑点会让你抓狂。但是你感觉不到眼睛中的洞，因为皮质细胞主动从边缘填充信息，补偿信息的缺失。

作为第三个证据，回想一下你梦境中的世界是彩色的、生动的和完全呈现的。因为你在黑暗中睡觉，闭着眼睛，你眼睛中的神经细胞不会发出关于外部世界的任何信号。是皮质-丘脑复合体提供了梦境的现象内容。

这只是视网膜神经元状态与你有意识看到的内容之间众多分离的三个例子。沿着你视神经传输的无数动作电位携带的数据在成为意识神经相关物的一部分之前经过了大量编辑。有时候视网膜信息完全被忽略，比如当你闭上眼睛想象小熊维尼，你形影不离的童年伙伴，或者当你梦见他时。

类似的反对意见可以针对初级视觉皮质细胞直接贡献于视觉感知的命题提出。以在猕猴大脑中用细电极记录的动作电位为例。一旦通过在动物麻醉下钻的小孔突破颅骨，微电极被慢慢插入灰质并连接到放大器。神经组织本身没有痛觉感受器，所以一旦电极就位，它们不会引起不适。（想想那50万名在大脑中植入电极以改善各种疾病的患者，比如帕金森病的震颤。）电极拾取神经细胞发出的微弱电脉冲。这些信号可以通过扬声器播放。你可以清楚地听到在神经系统自我嗡嗡声的背景嘘声之上的断续尖峰声。这样的记录证实初级视觉皮质中的神经元对猴子眼球运动和图像运动都有反应。当猴子的眼睛移动时，这些神经元一致发声，发出变化信号。然而正如我刚才讨论的，当你移动眼睛时，你的世界看起来是稳定的。

初级视觉皮质中的血流动力学活动通常反映一个人看到的内容，但有时也可能惊人地打断连接。伦敦大学学院的约翰-迪伦·海恩斯(John-Dylan Haynes)和杰兰特·里斯(Geraint Rees)向志愿者的眼睛短暂闪现向左或向右倾斜的条纹；这些条纹然后被遮挡，使观察者无法分辨它们的方向。他们所看到的只是由两条对角线组成的格子图案。然而对初级视觉皮质血流动力学反应的分析显示，这个区域区分了左边和右边的条纹。换句话说，初级视觉皮质——而不是更高的视觉区域如V2——看到了不可见条纹的方向，但该信息对大脑的所有者是不可获得的，这与我们的假设一致。

初级视觉皮质是通向大约三十几个以某种方式专门用于视觉处理的大脑皮质其他区域的门户。鉴于其战略位置，讽刺的是初级视觉皮质甚至对所有形式的视觉感知都不是必需的。对做梦志愿者的脑成像（考虑到扫描仪内狭小的空间和嘈杂的撞击声，这并非易事）表明，与很少做梦的非REM睡眠相比，在大多数梦境发生的快速眼动(REM)睡眠阶段，初级视觉皮质的活动受到限制。此外，初级视觉皮质受损的患者在做梦时不会伴随任何视觉内容的丢失。

其他初级感觉皮质——那些首先接收相关感觉数据流的皮质——也都不介导意识。对于因大脑大面积损伤而注定生活在植物状态、没有任何意识的人（我将在下一章更详细地讨论这些患者），巨大的噪声或痛苦的电击都未能引起任何有意义的反应。当扫描他们的皮质时，只有他们的初级听觉和躯体感觉皮质区域对这些强烈刺激显示显著活动。所有这些告诉我们的是，输入节点中的孤立活动不足以产生意识。需要更多。

## 新皮质高级区域的神经元与意识密切相关

---

遮挡和连续闪光抑制并不是视觉心理学家用来在意识雷达屏幕下向大脑输送载荷的唯一隐秘技术。双眼竞争是另一种。在双眼竞争中，一张小图片，比如一张脸，显示给你的左眼，另一张照片，比如旧日本帝国旗帜（从中央圆盘发出光线），显示给你的右眼。你可能认为你看到脸部叠加在旗帜上。不过，如果幻觉设置正确，你会感知到脸部与旗帜交替出现。你的大脑不会让你在同一地方同时看到两样东西。

起初，你清晰地看到那张脸，没有一丝阳光放射图案的痕迹；几秒钟后，旗帜的一块区域出现在你的视野某处，抹去了该位置的底层面部图案。从这个初始种子开始，图像逐渐扩散，直到脸部完全消失，只剩下旗帜。然后眼睛开始透出光芒。这种旗帜-面部图像的拼凑状态在几秒钟后解析为完整的脸部。接着旗帜的感知又重新占据主导地位。如此反复循环。这两个图像在意识中进进出出，进行着永无止境的舞蹈。你可以通过闭上一只眼睛来终止这种舞蹈——你会立即解决所有模糊性，感知到睁开眼睛看到的图像。

神经生理学家尼科斯·洛戈塞蒂斯(Nikos Logothetis)及其在图宾根马克斯·普朗克研究所的同事训练猴子在双眼竞争期间报告它们的感知。虽然人类志愿者因其劳动获得经济奖励——现金效果最好——但口渴的猴子会得到苹果汁作为奖励。猴子经过数月的学习，每当看到面部时拉一个拉杆，看到阳光放射图案时拉第二个拉杆，当看到其他任何东西时（比如两个图像的混合拼凑）则松开两个拉杆。主导时间的分布——阳光放射图案或面部一次被看到多长时间——以及改变图像对比度影响猴子报告的方式，都毫无疑问地表明猴子和人类有着性质相似的体验。

然后，洛戈塞蒂斯将细丝导入训练后的猴子大脑皮层，当时动物正处于双眼竞争设置中。在初级视觉皮层和附近区域，他发现只有少数神经元微弱地调节它们的反应以符合猴子的感知。大多数神经元的放电很少考虑动物看到的图像。当猴子发出一种感知信号时，初级视觉皮层中大批神经元对动物没有看到的被抑制图像产生强烈反应。这一结果完全符合弗朗西斯和我的假设，即初级视觉皮层无法被意识获取。

在高阶视觉区域，即下颞叶皮层中情况截然不同。该区域的细胞仅对猴子看到（并报告）的图片产生反应：它们都不对不可见图像产生反应。这样的神经元可能仅在动物指示它正在看面部时才发放动作电位。当猴子拉动另一个拉杆，表明它现在看到旗帜时，该细胞的放电活动会大幅下降——有时降至零——即使几秒钟前激活该细胞的面部图片仍然存在于视网膜上。细胞活动的增强和减弱期与动物感知报告之间这种紧密同步的双人舞，揭示了一组神经元与意识内容之间令人信服的联系。

正如我之前提到的，弗朗西斯和我关于意识神经相关物的推测集中在大脑皮层高阶感觉区域（在视觉情况下为下颞叶皮层）的神经元与它们在前额叶皮层目标之间建立直接回路。如果前额叶神经元通过轴突回延至下颞叶皮层，那么就建立了一个反响反馈回路并可以自我维持。放电活动然后可以扩散到支撑工作记忆、计划以及人类语言的区域。作为整体，这个神经元联盟介导对面部及其一些伴随属性的觉知，如凝视、表情、性别、年龄等。如果产生了代表旗帜的竞争回路，它会抑制面部回路的活动，意识内容就从面部转移到阳光放射图案。

最近，临床医生记录了两类严重脑损伤患者的脑电图，一类仍然昏迷，另一类至少恢复了某种程度的觉知。他们发现关键差异在于前额叶区域与后部颞叶、感觉皮层区域之间是否存在通信。如果存在这种反馈，意识得以保存。如果没有，意识就消失了。这是一个相当令人满意的发展。

现在还是早期阶段。我们还无法精确定位大脑中支撑意识的区域。但这是走错了方向——我们必须抵制脑扫描热点的催眠诱惑及其天真的颅相学解释：面部感知在这里计算，疼痛在那里，意识就在远处。意识不是从区域产生的，而是从区域内部和跨区域的高度网络化神经元产生的。

大脑的一个独特特征在过去二十年才变得明显，那就是神经元令人震惊的异质性。大约10万个神经元挤在皮层每平方毫米下面，这个区域大约相当于本页上字母“o”的大小，它们高度异质。可以根据它们的位置、树突的形状和形态、突触的结构、基因构成、电生理特性以及轴突投射的位置来区分它们。理解这种巨大多样性的参与者——也许多达一千种神经元细胞类型，它们是中枢神经系统的基本构建模块——如何对感受质(qualia)的产生做出贡献，这一点至关重要。

底线是这些生理实验正在稳步缩小心智与大脑之间的差距。可以提出假设，进行测试，然后拒绝或修改。在数千年的无效争论后，这是一个巨大的恩惠。

## 注意到某事，却看不见它

---

选择性注意和意识之间是什么关系？我们似乎对注意力聚光灯照亮的任何东西都会产生意识。当你努力在营地谈话声中倾听远处土狼的嚎叫时，你通过专注于声音而有意识地察觉到它们的嚎叫。由于注意力和意识之间的密切关系，许多学者将这两个过程混为一谈。实际上，在20世纪90年代初，当我公开出来举办关于心身问题的公开研讨会时，我的一些同事坚持让我用更中性的“注意力”来替代煽动性的“意识”，因为这两个概念无法区分，反正可能是同一回事。

我直觉地认为两者是不同的。注意力选择部分传入数据进行进一步仔细检查和检验，代价是忽视未被注意的部分。注意力是进化对信息过载的回应；这是没有大脑能处理所有传入信息这一事实的结果。离开眼球的视神经每秒传输几兆字节的信息，按照今天无线网络的标准来说非常少。这些信息不仅必须传送到皮层，还必须据此采取行动。大脑通过选择一小部分进行进一步处理来处理这种数据洪流——这些选择机制与意识是不同的。因此，注意力有着与意识使命不同的明确功能作用。

二十年后，我确信注意力和意识之间的区别是有效的。正如我刚才讨论的，连续闪光抑制会将图像遮蔽数分钟。在这种不可见性的掩护下，敏锐的实验者有足够的范围来操纵观看者的注意力。这就是盛和所做的，他证明了裸体男性的不可见图片吸引了女性的注意力，而裸体女性的不可见图片吸引了男性的注意力。毕竟，注意力是图像的选择性处理。大量其他实验也证实了这一点。例如，在遮蔽期间对初级视觉皮层的功能性脑成像显示，注意不可见物体会增强大脑对它们的反应。相反，操纵物体的可见性对V1的血流动力学反应几乎没有一致的影响。显然，大脑可以注意到它看不见的物体。

相反的情况，即缺乏注意力的意识，也会发生吗？当你注意某个特定位置或物体，专心审视它时，世界的其他部分并不会缩减为隧道，注意力焦点之外的一切都消失：你总是意识到周围世界的某些方面。你意识到你在读报纸或在有立交桥即将到来的高速公路上开车。

要旨(*Gist*)指的是场景的紧凑、高级摘要——高速公路上的交通堵塞、体育场的人群、持枪的人等等。要旨感知不需要注意处理：当大照片在你被告知要专注于视线中心某个细微细节时被短暂且意外地闪现到屏幕上时，你仍然能够理解照片的本质。只需要二十分之一秒的一瞥就足够了。在那短暂的时间里，注意选择并不起太大作用。

作为青少年，我儿子会在继续玩快节奏射击游戏的同时与我交谈。他的注意力显然没有完全被我们的谈话占据，为第二个更具挑战性且对他来说更重要的任务留下了足够的注意资源和带宽。

德国马格德堡大学的心理物理学家约亨·布劳恩完善了此类双重处理任务的实验室版本，用于测量某人在注意力聚光灯之外能看到多少。布劳恩的想法是通过给志愿者一个在凝视中心的困难工作（比如计算字母流中出现多少个X）和在计算机屏幕其他地方的次要任务来固定注意力。实验探究当这个任务与主要的需要注意力的工作同时进行时，表现如何恶化。

布劳恩发现，当注意力集中在凝视中心时，观察者可以用眼角余光区分出有动物的照片——有狮子的大草原、有鸟群的林冠、鱼群——与没有动物的图片。然而他们无法区分被分为红色和绿色两半的圆盘与其镜像——绿红圆盘（实验设置控制了视觉外围敏锐度的降低）。受试者可以判断在中央视觉外呈现的脸部性别，或者脸部是否著名，但对似乎更简单的工作感到挫败，比如区分旋转的字母“L”和旋转的“T”。布劳恩的实验表明，至少一些视觉行为可以在缺乏——或者采用更谨慎的立场——在几乎缺乏选择性注意的情况下完成。

最终分析中，心理学方法过于不足以完全解决这个问题。不精密地介入潜在的大脑回路，注意力和意识之间的区别将不会得到充分解决。这在小鼠或猴子等实验动物中现在正在迅速成为可能。最终的测试将是有条件地关闭——然后再次打开——大脑中注意力作用的控制线路，并观察动物仍然能够进行哪些视觉行为。我将在第9章回到这样的实验。

任何科学概念的历史——能量、原子、基因、癌症、记忆——都是一个不断分化和精细化的过程，直到能够在更低、更基本的层面上以定量和机械的方式进行解释。我刚才讨论的双向分离现象——无意识的注意和无注意的意识——彻底否定了两者相同的观念。它们并不相同。现有的实验文献中很大一部分都必须在注意和意识之间这一区别的指导下重新审视。这不容易。我所做的区分集中的神经生物学攻势清理了道路，以解决识别大脑中意识必要成因的核心问题。

## 第5章：

---

你将从神经科医生和神经外科医生那里了解到，某些神经元对名人极为关注，将大脑皮质一分为二并不会使意识减半，一个小皮质区域的缺失会使世界失去色彩，以及一块方糖大小的脑干或丘脑组织的破坏会让你变成活死人

换句话说，大脑中存在着不同的能力群体，而在大脑中，存在着不同的脑回群体；迄今为止科学获得的事实允许我们接受，正如我在其他地方所说的，大脑的大区域对应于大脑的大区域。正是在这个意义上，定位原理在我看来，如果不是严格得到证明，至少是极其可能的。但是要确定地知道每种特定能力是否在特定脑回中都有自己的位置，这是一个在科学的当前状态下几乎无法解决的问题。

——保罗·布洛卡，《解剖学会公报》(1861)

从历史上看，临床一直是关于大脑和心智洞察的最丰富源泉。自然和人类的变幻莫测，连同他们的汽车、子弹和刀具，造成了范围有限的破坏，这些破坏阐明了结构与功能之间的联系，并显现出在健康状态下几乎看不到的特征。让我告诉你患者及其主治医师、神经科医生和神经外科医生关于意识神经基础教给我们的四个重要课题。

我经常收到冗长的、主动寄来的思考文稿——密密麻麻写满的手稿，承诺会有更多内容，自费出版的书籍，或指向大量网页的链接——涉及生命和意识的终极答案。我对这些作品的态度是，除非它们尊重这些来之不易的神经学和科学知识，否则注定要进入我办公室尘封角落里不断增长的X档案。

## 小块灰质介导意识的特定内容

---

许多学者认为意识是大脑的整体性、完形特性。他们推理认为意识如此超世界，不可能由神经系统的任何一个特定特征引起。相反，意识只能归因于大脑整体。在我将在第8章中概述的技术意义上，他们是正确的：现象意识是因果相互作用部分的综合系统的特性。然而意识也有令人惊讶的局部方面。

中风、车祸、病毒感染、神经外科医生手术刀的控制性创伤都可能破坏脑组织。在它们的影响下，往往会留下永久性缺陷。对神经科学家来说非常感兴趣的是当损伤是有限和局部的时候。失去特定的一块神经组织会使世界变成灰色调，使熟悉的面孔变得陌生，这一事实表明该区域必须至少部分地负责产生颜色感或面部身份感。

以加州大学伯克利分校杰克·加兰特研究的患者A.R.为例。52岁时，A.R.遭受脑动脉梗塞，短暂失明。两年后进行的MRI扫描显示，在他的高级视觉中心右侧，即初级视觉皮质之外，有一个豌豆大小的病变。当加兰特和他的同事们在实验室测试A.R.时，他们发现他失去了色觉。不过，并非到处都失去，而只是在他视野的左上象限，正是他们根据MRI扫描所预期找到的地方。值得注意的是，这个人大多没有意识到他的部分世界已经变灰。

A.R.的低级视觉以及他的运动和深度感知都是正常的。唯一的其他缺陷是部分无法区分形状——他无法阅读文字——但这再次局限于他视野的左上象限。

纯粹的色觉缺失称为全色盲。这与主要影响男性的日常遗传性色盲完全不同。由于他们缺少眼睛中一种颜色色素的基因，这些二色视者不像具有三种视网膜光色素的正常视力人群(三色视者)那样感知到丰富的色彩调色板。相比之下，全色盲是视觉皮质中颜色中心被破坏后的结果。在其影响下，所有色调都从世界中消失了。不再有夕阳中壮丽的紫色和紫红色阿尔卑斯山余晖。取而代之的是，世界以明暗对比的方式被体验，就像彩色电视机退化为黑白色调一样。值得注意的是，颜色词汇和与颜色的关联(例如，“红色”与“消防车”)依然保留。

还有许多其他奇异的疾病。患有面盲症的人——医生称之为面容失认症(*prosopagnosia*)——在识别面孔方面存在困难。他们无法识别著名或熟悉的面孔。他们意识到自己正在看一张脸，只是无法识别出是谁。所有面孔看起来都一样，就像河床上的一堆鹅卵石一样难以区分。尽管面孔和石头都有许多不同的特征，但你很难区分这些光滑的石头，却能轻松识别数百张面孔。原因在于你大脑中相当多的回路专门用于处理面孔，但很少用于处理石头的外观（除非你是石头收藏家或地质学家）。由于大脑皮层高级区域中负责面孔身份识别的神经元被破坏或从一开始就不存在（相当多的人患有先天性面盲症），这些患者无法在机场的人群中认出他们的配偶。他们缺乏你我在看到所爱之人时所体验到的那种轻松而即时的识别能力。

面盲症导致社会隔离和羞怯，因为患者难以识别，更不用说说出与他们交谈的人的姓名。他们采用应对策略，专注于某些特殊标记，如痣或突出的鼻子、颜色鲜艳的衬衫，或声音特征。化妆和改变发型会阻碍识别，穿制服的人群也是如此。

在严重的面盲症中，患者甚至无法将面孔看作面孔。他们的视觉系统没有问题：他们能感知构成面孔的各个独立元素——眼睛、鼻子、耳朵和嘴巴——但无法将它们综合成一个统一的面孔感知。神经学家Oliver Sacks撰写的《错把太太当帽子的人》是一部观察精辟的病例研究集，书名来自这样一个患者：他误将时钟的表面认作人脸，试图与一个祖父钟握手。

有趣的是，这些患者在面对熟悉面孔时仍可能有自主神经反应：当他们看到知名政治家、电影明星、同事或家庭成员时，与看到陌生人的照片相比，会有增强的皮肤电反应——本质上是轻微出汗。然而，他们始终坚持认为自己不认识这些人。因此，无意识有其自己的方式来处理情感丰富的面孔。

面容失认症的反面是卡普格拉斯妄想症(*Capgras delusion*)。患有这种疾病的病人持续声称他的妻子已被外星人或冒名顶替者所替代，这个冒名顶替者看起来、说话和行动都与配偶一样，但在某种程度上是不同的。这种疾病可能相当局限，患者在其他方面并无异常。在这里，面孔识别是完整的，但熟悉感的自主神经反应缺失了。由于患者缺乏我们在遇到亲密之人时都理所当然的那种情感冲击，他感觉有什么地方严重不对劲。

运动盲症(*Akinetopsia*)是一种罕见且毁灭性的运动盲症。患有这种疾病的个体被放逐到一个只有闪光灯照明的世界，就像迪斯科舞厅或夜总会。舞者在每次闪光下都清晰可见，但他们看起来是静止的，没有运动。在镜子中看到自己以这种方式运动是迷人的，但我可以向你保证，兴奋感很快就会消失。运动盲患者通过比较物体在时间上的相对位置来推断物体已经移动，但她看不到它们移动。她能看到汽车改变位置，但看不到它向她驶来。视觉的其他方面，包括颜色、形状和检测闪烁光线的能力，都是完整的。

基于对这类局部损伤患者的仔细观察，伦敦大学学院的Semir Zeki创造了关键节点(*essential node*)这一术语，指负责特定意识属性的大脑部分。视觉皮层的一个区域包含颜色感知的关键节点；面孔感知和视觉运动感知中涉及几个这样的区域。杏仁核的部分区域对恐惧体验至关重要。任何一个节点的损伤都会导致相关感知属性的丧失，尽管其他意识属性保持不变。

临床数据的解释并不像我描述的那样直接，因为大脑，尤其是年轻的大脑，具有巨大的恢复能力。即使关键节点丢失，信息也可能被重新路由和重新表达，个体可能会慢慢恢复失去的功能。

要点是大脑皮层的小块区域负责特定的意识内容。皮层的这一部分赋予现象体验以面孔的生动性，那部分提供新奇感，而那边的部分调节声音的感知。皮层位置与功能之间的联系是神经系统的标志。将此与另一个重要器官——肝脏——进行对比。像大脑一样，肝脏重约三磅，有左叶和右叶。但肝组织比神经组织分化程度低得多，更加同质化。肝功能损害与损伤程度成正比，很少考虑破坏发生的位置。

## 概念神经元编码Homer Simpson和Jennifer Aniston

---

我生动地回忆起第一次坠落的经历。我最近开始攀岩，纯粹是出于绝望——我的儿子离家上了大学，女儿将在学年结束时离开。长期恐惧的空巢期到了！为了做点什么，任何事情，来消耗我多余的精力和热情，我开始了山地跑步和攀岩。

我当时正在加州沙漠的约书亚树国家公园攀登一条裂隙路线。我清楚地记得在那面橙棕色花岗岩近乎垂直的右弯墙壁上的确切位置，岩壁上嵌着晶体。花岗岩可能会撕裂攀岩者的手，但它也是攀岩者最好的朋友，因为它不平整的纹理提供了额外的摩擦力。我的左脚楔入裂隙，右脚在外侧踩踏，我用右手高高伸向头顶上方，将一个叫做cam的小型保护装备塞入裂隙。它进去得太顺利了，我担心受力时它可能会同样快速地弹出来。我重新调整了cam的位置，将它更深地塞入裂隙。就在这时我的左脚滑脱了，我从十到十二英尺高的地方坠落到地面，直接仰面着地，旁边有一块尖锐的岩石，这是一个我因愚蠢的运气而避免的潜在灾难。我背部的皮肤被磨破了，好几天都一瘸一拐，但这些轻微的伤害只是增强了攀岩的光环。这些事件不可磨灭地印在我的记忆中。

意识的内容怎么能如此充满具体而令人回味的细节？我的颅骨内没有我攀岩的图片，只有一个棕灰色的器官，具有煮过头的花椰菜的质地、大小和形状。这种豆腐般的组织，被血液和脑脊液缓冲，由神经细胞和胶质细胞组成。神经元(neurons)及其相互连接的突触(synapses)是感知、记忆、思维和行动的原子。如果科学要理解这些过程，它必须能够嵌入在深不可测的复杂网络中的大量神经元联盟的相互作用来解释它们。考虑一个类比：化学家如果不了解支配电子和离子相互作用的电磁力，就无望理解正常温度下物质的组成。

上段开头提出的问题是一个深刻的问题，目前还没有明确的答案。但我可以告诉你我密切参与的一个相关发现，它确实揭开了一些面纱。

癫痫发作——超同步化、自维持的神经放电，可以吞没整个大脑——是一种常见的神经系统疾病。对许多人来说，这些反复发作的阵发性脑痉挛通过抑制底层回路兴奋并增强抑制的药物得到控制。然而，药物治疗并不总是有效。当怀疑局部异常，如瘢痕组织或发育期错误连线，触发癫痫发作时，就需要进行神经外科手术来移除有问题的组织。尽管任何穿透颅骨的手术都有一些风险，但对于那些癫痫发作无法用其他方式控制的人来说，这种手术是有益的。

为了最大限度地减少副作用和术后生活质量的损失，准确定位大脑中癫痫发作的确切位置至关重要；这通过神经心理测试、脑扫描和脑电图(EEG)来完成。如果从外部看不到明显的结构性病变，神经外科医生可能会通过颅骨钻出的小孔将十几个电极插入大脑的软组织中，并让它们在那里停留一周左右。在此期间，患者在医院病房生活和睡觉，电线的信号被持续监测。当癫痫发作时，癫痫学专家和神经放射学专家三角测量异常电活动的起源。随后破坏或移除有问题的组织块可以减少癫痫发作的次数——有时完全消除它们。

加州大学洛杉矶分校医学院的神经外科医生和神经科学家Itzhak Fried是这个要求很高的行业中的世界大师之一，这需要极高的技术精密度。脑外科医生与攀岩者和登山者有许多共同特质，这也是我所向往的一套态度和行为。他们是技术极客——热衷于高科技和精密测量，但他们也很有见识和文学修养。他们对生活及其风险采取直截了当、不废话的方法；他们了解自己的限制，但对自己的技能有着极大的信心。（你不希望你的外科医生在即将钻入你的颅骨时犹豫不决。）而且他们可以专注数小时于手头的任务，排除其他一切。

Itzhak和他的外科同事完善了一种癫痫监测的变体，其中电极是空心的。这使他们能够将比头发还细的微小电线直接插入灰质。使用适当的电子设备和复杂的信号检测算法，微型电极束可以从大脑电活动的持续背景噪音中捕获十到五十个神经元群体的微弱聊天声。

在Itzhak的监督下，我实验室的一个小组，包括Rodrigo Quian Quiroga、Gabriel Kreiman和Leila Reddy，在内侧颞叶(medial temporal lobe)的丛林中发现了一组非凡的神经元。这个区域包括海马体(hippocampus)，将感知转化为记忆，但它也是许多癫痫发作的源头，这就是为什么Itzhak在这里放置电极的原因。

我们征求患者的帮助。虽然他们除了等待癫痫发作之外无事可做，我们向他们展示熟悉的人、动物、地标建筑和物体的图片。我们希望其中一张或多张照片能吸引被监测神经元的注意，促使它们发出一连串动作电位。大多数时候搜索都是徒劳的，尽管有时我们会发现一些对物体类别有反应的神经元，比如动物、户外场景或一般的面孔。但有些神经元要挑剔得多。当Gabriel向我展示第一批这样的细胞时，我非常兴奋。一个细胞只有在患者看时任总统Bill Clinton的照片时才会放电，对其他名人没有反应，另一个细胞则只对卡通人物Bart和Homer Simpson有反应。

尽管我们对这一发现最初相当怀疑——在单个神经细胞水平上如此惊人的选择性是闻所未闻的——内侧颞叶神经元确实对什么能激发它们极其挑剔。一个海马神经元只对电影明星Jennifer Aniston的七张不同照片有反应，但对其他金发女性或女演员的照片没有反应。海马中的另一个细胞只对女演员Halle Berry放电，包括她的卡通形象和拼出的姓名。我们发现了对Mother Teresa形象有反应的细胞，对可爱小动物有反应的细胞（“Peter Rabbit细胞”），对独裁者Saddam Hussein的形象以及他的语音和书面姓名有反应的细胞，还有对勾股定理 $a^2 + b^2 = c^2$ 有反应的细胞（这个在一位以数学为爱好的工程师大脑中）。

Itzhak将这些细胞称为概念神经元。我们尽量不将它们拟人化，避免称它们为“Jennifer Aniston细胞”的诱惑（细胞不喜欢你这样做！）。每个细胞连同它的姐妹细胞——因为在内侧颞叶中任何一个想法可能都有数千个这样的细胞——编码一个概念，比如Jennifer Aniston，无论患者是看到还是听到她的名字，看她的照片，还是想象她。可以把它想象为Jennifer Aniston柏拉图理想的细胞基础。无论这位女演员是坐着还是跑着，无论她的头发是盘起来还是放下，只要患者认出Jennifer Aniston，这些神经元就会活跃。

没有人生来就有对Jennifer Aniston选择性的细胞。就像雕塑家耐心地从一块大理石中雕刻出维纳斯或圣母怜子，大脑的学习算法雕刻出概念神经元嵌入其中的突触场。每次你遇到特定的人或物体时，在高阶皮质区域都会产生类似的尖峰神经元模式。内侧颞叶中的网络识别这种重复模式并将特定神经元专门分配给它们。你有编码家庭成员、宠物、朋友、同事、你在电视上看到的政治家、你的笔记本电脑、你喜爱的那幅画的概念神经元。我们推测概念细胞也代表更抽象但深度熟悉的想法，比如我们与记忆符号9/11、圆周率π或上帝概念相关联的一切。

相反，你对很少遇到的事物没有概念细胞，比如刚刚递给你脱脂柴茶拿铁的咖啡师。但如果你与她成为朋友，后来在酒吧遇到她，让她进入你的生活，内侧颞叶中的网络就会识别出相同的尖峰模式反复出现，并连接概念细胞来代表她。

视觉皮质中的许多神经元对特定方向的线条、灰色斑块或一般面孔以杂乱的热情做出反应，而内侧颞叶中的概念细胞则相当克制。任何一个个体或事物只会在很小一部分神经元中引起活动。这被称为稀疏表示。

概念细胞令人信服地证明了意识体验的特异性在细胞水平上有直接对应物。假设你正在回忆Marilyn Monroe站在地铁通风口上，阻止风吹起她裙子的标志性场景。通常认为大脑使用广泛的群体策略来表示这种知觉。当你看到Monroe时，几千万神经细胞以一种方式放电，当你看到Aniston、英国女王或你的祖母时以不同方式放电。总是相同的细胞群体在反应，但以不同的方式。然而我们的发现使这对于你深度熟悉的那些概念或个体来说不太可能。大多数细胞大多数时候都是沉默的，这是稀疏表示的本质。当Monroe出现时，少数细胞放电；不同的群体会对Aniston活跃，依此类推。任何一个有意识的知觉都是由可能数百或数千而不是数百万神经元联盟造成的。

最近，来自我实验室的Moran Cerf和其他人与Itzhak一起，将来自几个概念细胞的信号连接到外部显示器来可视化患者的思想。这个想法看似简单但执行起来极其困难。Moran——一位计算机安全专家和电影制片人转为Caltech研究生——花了三年时间才完成这一壮举。

让我为您介绍一个例子。Moran记录了一个神经元的活动，该神经元对演员Josh Brolin的图像产生反应（患者从她最喜欢的电影《七宝奇谋》中认识这位演员），另一个神经元则对我刚才提到的玛丽莲·梦露场景产生替代性反应。患者看着一个显示器，两张图像叠加在一起，两个细胞的放电活动控制着她在混合图像中看到Brolin或梦露的程度（通过患者大脑对显示器的反馈）。每当患者将注意力集中在Brolin身上时，相关的神经元就会更强烈地放电。Moran设置了反馈机制，使得这个细胞相对于另一个细胞放电越多，Brolin就越清晰可见，梦露的图像就越淡化，反之亦然。屏幕上的图像不断变化，直到只有Brolin或只有梦露可见，试验就结束了。患者很喜欢这个实验，因为她完全用思想控制着电影。当她专注于梦露时，相关神经元增加了放电频率，竞争概念Brolin的细胞同时抑制了活动，而绝大多数神经元保持不受影响。

按照我在这里讲述的方式，听起来有两个主角，就像电影《成为约翰·马尔科维奇》中木偶师Craig占据演员约翰·马尔科维奇头部的方式。一个是患者的心智，专注于梦露。另一个是患者的大脑——即内侧颞叶的神经细胞，它们根据心智的意愿上调和下调活动。但两者都是同一个人的一部分。那么谁控制谁呢？谁是木偶师，谁是木偶？

Itzhak的电极探测着意识神经相关物的核心地带。患者可以非常有意识地、非常有选择性地调高和调低内侧颞叶神经元的活动量。但大脑的许多区域将不受这种影响。例如，你不能意志性地让自己看到灰色阴影的世界。这很可能意味着你无法有意识地下调视觉皮层中的色彩神经元。尽管你有时可能想要这样做，但你无法关闭大脑中的疼痛中枢。

心智-身体关系的所有奇异性在这里都很明显。患者并不会在梦露神经元每次放电时感到瘙痒；她不会想“抑制，抑制”来将Brolin从屏幕上驱逐。她对头部内部发生的事情完全一无所知。然而，对梦露的想法转化为特定的神经元活动模式。她现象性心智中的事件在她的物质大脑中找到了对应。心智地震与大脑地震同时发生。

## 意识可以由任一大脑半球产生

---

大脑和身体的其他部分一样，具有显著的双侧对称性。将其视为放大的核桃会很有帮助。一侧并非另一侧的完全镜像，但大致如此。几乎每个大脑结构都有两个副本，一个在左侧，一个在右侧。视野的左侧由右半球的视觉皮层表示，而右侧则映射到左侧视觉皮层。当你观察世界时，你不会看到贯穿视野的细垂直线；两个半视野无缝整合。哲学家强调经验是统一的。你不会体验到两股意识流，每一侧一股，而只有一股。对视觉成立的，对触觉、听觉等同样适用。

大脑两半与一个心智之间的不协调被笛卡尔指出，他寻找反映经验统一性的单一器官。他错误地假设松果体没有左右两半，并著名地假设它是灵魂的所在地（用现代语言说，是意识的神经相关物）。当我在课堂上提到笛卡尔对松果体的选择时，一些学生会窃笑，“多么愚蠢。”事实上，笛卡尔领先他的时代几个世纪，寻找结构与功能之间的关系。在中世纪经院哲学末期尘土飞扬、陈腐不堪的氛围中，他是现代性和启蒙运动的清新空气。笛卡尔用机械性原因取代了陈旧的亚里士多德目的论最终原因——这些原因实际上什么都不能解释——木材燃烧是因为它具有寻求燃烧的内在形式。笛卡尔与Francis Crick和已故神经外科医生Joseph Bogen一起，居住在我的个人万神殿中。（说实话，少年记者丁丁和侦探夏洛克·福尔摩斯也是成员。）

胼胝体是大脑中最大的白质结构，主要负责这种整合。它是一个由约2亿根轴突组成的厚束，每根轴突从大脑一侧的锥体细胞延伸到另一侧。这些轴突连同一些次要的线束，紧密协调两个大脑半球的活动，使它们毫不费力地协同工作，产生对世界的单一视角。

如果这束轴突被切断会发生什么？如果小心地进行，不对其他结构造成损伤，患者应该保持有感知能力，尽管他或她的意识可能被一分为二，缩小到只包含左侧或右侧视野，另一半不可见。然而，这并不是实际发生的情况！

在某些顽固性癫痫发作的病例中，胼胝体的部分或全部会被切断，以防止起源于一侧半球的癫痫发作扩散到另一侧半球并引起全身性抽搐。这种手术最初于1940年代早期实施，能够缓解癫痫发作。值得注意的是，裂脑患者在手术恢复后，在日常生活中并不显眼。他们像以前一样看、听、闻，四处走动，说话，与他人进行适当的互动，智商也没有改变。他们保持着平常的自我意识，并报告在对世界的感知方面没有明显的改变——例如，他们的视野没有缩小。开创这种手术的外科医生，如南加州洛马林达大学的约瑟夫·博根(Joseph Bogen)，对这种缺乏明显症状的情况感到困惑。

然而，加州理工学院生物学家罗杰·斯佩里(Roger Sperry)对裂脑患者的仔细检查揭示了一种持续且深刻的断连综合征。如果特定数据被传给一个半球，这些信息不会与另一侧的双胞胎半球共享。此外，只有一个半球，通常是左半球，会说话。也就是说，如果右半球丢失或被麻醉剂抑制，患者仍然能够说话，这就是为什么左半球被称为优势半球的原因。右半球本身只有有限的语言理解能力，是哑的，尽管它可以咕哝和唱歌。所以，当与裂脑患者对话时，是这个人的左半球在做所有的谈话。他无法命名出现在左视野中的物体，因为该图像是由他哑的右视觉皮层处理的。但他可以用左手从托盘上的一组物体中挑出该物体，左手是由右运动皮层控制的。

如果一把钥匙放在他的右手中，在桌子下面，看不见的地方，患者会快速说出它的名称。来自他右手的触觉信息传递到他的左半球，在那里物体被识别，其标签被传递到语言中心。然而，如果钥匙放在这个人的左手中，他就无法说出它是什么，并且会胡言乱语。右半球很可能知道这个物体是钥匙，但它不能将这种知识传达给左侧的语言中心，因为通信链接已经被切断。

大脑的一半字面上不知道另一半在做什么，这导致了介于悲剧和闹剧之间的情况。北达科他大学的神经学家维克托·马克(Victor Mark)录制了与一位裂脑患者的访谈视频。当被问及手术后她有多少次癫痫发作时，她的右手举起了两

根手指。然后她的左手伸过来强行按下了右手的手指。在几次尝试计算她的癫痫发作次数后，她停了下来，然后同时用右手显示三根手指，用左手显示一根。当马克指出这种差异时，患者评论说她的左手经常自己做事情。两只手之间爆发了一场看起来像滑稽喜剧的打斗。只有当患者变得如此沮丧以至于泪如雨下时，我才想起了她的悲惨处境。

对裂脑患者的研究，斯佩里因此工作于1981年获得诺贝尔奖，告诉我们切断胼胝体会将皮质-丘脑复合体一分为二，但保持意识完整。两个半球都能独立地进行意识体验，其中一个比另一个更善于言语表达。无论意识的神经相关物是什么，它们必须独立存在于大脑皮质的两个半球中。一个头骨中的两个意识心智。我在第8章中会回到这个主题。

## 意识可能永久消失，留下僵尸

---

只要你醒着，你就对某些事物有意识——前方的道路，拉姆斯坦的重金属乐曲不断在你脑中回响，或者关于性的幻想。只有在某些冥想练习中，一个人才能在没有任何特定内容的情况下有意识，意识到而不意识到任何特定的东西。即使当你的身体睡着时，你也可以在梦中有生动的体验。相比之下，在深度睡眠、麻醉、昏厥、脑震荡和昏迷期间，根本没有任何体验。不是黑屏，而是虚无。

当严重的伤害袭击大脑时，意识可能不会回归。车祸、跌落、战斗伤口、药物或酒精过量、近乎溺水——任何这些都可能导致严重的无意识状态。由于救援直升机和急救医疗技术人员，他们迅速将受害者送到专业创伤护士和医生团队的照顾下，这些医生拥有先进的工具和药物宝库，许多人可以从死亡边缘被拯救回来。虽然这对大多数人来说是一种福祉，但对少数人来说却是一种诅咒。他们存活多年，永远无法恢复意识，成为不死之身。

这种全面的意识障碍发生在负责觉醒的大脑部分受损时。丘脑和大脑皮质中的神经元无法组装成介导任何一个意识内容的远程神经元联盟。意识受损状态包括昏迷、植物人状态和最小意识状态。总体觉醒从昏迷中的完全缺失，到植物人状态中的周期性睡眠-觉醒转换，以及最小意识状态中伴随有目的性运动的觉醒、梦游和某些部分性癫痫发作而波动。

仅在美国，就有多达25,000名患者在被称为持续性植物状态(PVS)的植物状态中徘徊数年，康复前景黯淡。让这种情况几乎无法忍受的是，与几乎没有反射的昏迷患者不同，处于这种边缘状态的患者有日常的睡眠-觉醒周期。当他们“清醒”时，眼睛是睁开的，可能会反射性地移动，四肢偶尔也会如此；他们可能会做鬼脸、转头、呻吟。对于床边天真的观察者来说，这些动作和声音暗示患者是清醒的，拼命想与她的亲人交流。毁掉的患者空白而空虚的生活悲剧，在安养院和疗养院中绝望地延续数十年，被她的家人为她的护理所投入的爱——和资源——所反映和放大，总是希望奇迹般的康复。

你可能还记得佛罗里达的Terri Schiavo，她在持续性植物状态中持续了十五年，直到2005年医学诱导死亡。由于她的丈夫（主张停止生命支持）和她的父母（相信他们的女儿有某种程度的意识）之间的激烈公开争斗，该案件引起了巨大轰动，在司法链上下诉讼，最终牵涉到时任总统乔治·W·布什。从医学角度来看，她的案例并无争议。她有短暂的自动症发作：转头、眼球运动等，但没有可重现的或一致的、有目的的行为。她的脑电图是平坦的，表明她的大脑皮层已经关闭。她的状况多年来没有改善。尸检显示她的皮层萎缩了一半，视觉中枢萎缩；因此与当时流传的公开报告相反，她不可能看到任何东西。

我要稍作题外话。美国目前的立法在撤回医疗护理和主动安乐死之间划出了明确的区别。在前一种情况下，绝症患者按照自己的时间表死亡。在后者中，医生通过阿片类药物或其他加速死亡到来的药物进行干预。我理解产生禁止安乐死法律的历史力量。但是允许患者，甚至像Schiavo这样的无意识患者，通过撤回所有液体或固体营养而饿死的方式死亡，在我看来是野蛮的。在第十二年，Trixie，我们心爱的家庭犬，患有心肌病。她停止进食，肚子充满水分，经常呕吐，并且难以控制排便。我和妻子最终带她进行最后一次兽医访问。当她信任地躺在我怀里，轻柔地舔我时，兽医给她注射了大剂量的巴比妥类药物，直到她勇敢的心脏停止跳动。这很快，无痛——尽管深深悲伤——这是正确的做法。我希望当我的时候到来时，有人会为我提供同样的服务。

不幸的是，区分处于持续性植物状态（有规律的睡眠-觉醒转换）的患者和处于最小意识状态（可以偶尔与周围人交流）的人通常是困难的。完成此任务的一个工具，一种意识测量仪，将在第9章中讨论。功能性脑成像可能是另一个工具。

英国剑桥大学的神经学家Adrian Owen将一名在交通事故中大脑严重受损的无反应女性放入MRI扫描仪。她的母亲为她读出指令，要求她想象打网球或参观她房子里的所有房间。患者没有显示出理解的迹象，更不用说反应了。然而，血流动力学脑活动模式与闭眼想象类似动作的健康志愿者相似。这种幻想是一种复杂而有目的的心理活动，持续数分钟：它不太可能在无意识中发生。受伤的女性，尽管无法用手、眼睛或声音发信号，至少偶尔是有意识的，能够遵循外部命令。大多数其他以这种方式测试的植物状态患者没有这样的大脑特征；他们似乎真的没有意识。MRI扫描仪可能是严重脑损伤患者的生命线，因为它开启了一种交流方式：“如果你感到疼痛，想象打网球。如果没有，想象在你的房子里走动。”

回到主要主题，值得注意的是，大脑皮层的大部分可以被破坏而在恢复后不会有整体功能损失。如上所述，患有局灶性皮层损伤的人缺陷有限。这种对损伤的抵抗力在额叶中尤其明显。用电流刺激它们不会产生任何抽搐的肢体——如刺激初级运动皮层那样——或闪光——如刺激视觉皮层那样。这就是为什么早期神经学家经常将额叶称为沉默区域。

经典精神外科的定义特征是有控制地破坏皮层额叶中的灰质（脑叶切除术）或切断连接前额皮层与丘脑和基底神经节的白质中的轴突（白质切开术）。这些手术臭名昭著地用通过眼窝插入的改良冰锥进行，造成人格改变和精神残疾。它们将“疯子变成白痴”，便于患者的监护护理，但它们不会造成意识的全面丧失。

然而，位于分离左右脑的假想中线附近的皮下结构的小而局限的损伤确实可以使人失去意识。我将这些中线结构视为意识的促成因子。它们控制意识所需的大脑唤醒程度。如果皮下区域的左右两个副本都被破坏，患者可能会永久失去意识。（一般来说，大脑能够承受身体一侧结构的损伤，但对两侧同时受损的抵抗力要弱得多。）其中一个这样的中线结构是网状激活系统(reticular activating system)，这是上脑干和下丘脑中核团的异质性集合。核团是具有独特细胞结构和神经化学特性的三维神经元集合。网状激活系统中的核团从其轴突向整个前脑释放调节性神经递质，如血清素、去甲肾上腺素、乙酰胆碱和多巴胺。

意识的另一个促成因子是丘脑的五个板内核团(intralaminar nuclei)，它们也聚集在中线周围。这些核团接受来自脑干核团和额叶的输入，并将输出发送到整个大脑皮层。在左右板内核团中不超过方糖大小的病变会使意识消失，很可能是永久性的。

丘脑和脑干中的大量核团保持前脑充分唤醒以使体验发生。这些具有不同化学特征的结构都不负责产生那种体验的内容，但它们使体验成为可能。它们努力的终点是大脑皮层中的160亿个神经元及其在丘脑、杏仁核、屏状核(claustrum)和基底神经节中的密切伙伴。通过控制神经递质混合物的释放，板内核团和大脑地下室中的其他核团调节突触和神经元活动的上下，使皮层-丘脑复合体能够形成和塑造紧密同步的神经元联盟，这是任何一次意识体验的核心。

总之，皮层及其附属结构的局部特性介导意识的具体内容，而全局特性对于维持意识本身至关重要。要使神经元的连贯联盟能够组装——并使意识出现——皮层-丘脑复合体需要被神经递质充满，这些化学物质由大脑更深层和更古老部分的神经元的长而蜿蜒的触手释放。局部和全局方面对意识都至关重要。

但神经解剖学和神经化学已经足够了。现在让我转向无意识。

# 第六章：我为两个我年轻时认为是胡说八道的命题辩护

## ——你对发生在你头脑中的大部分事情都没有意识，而且 僵尸代理控制着你生活的很大一部分，尽管你自信地相信 你在掌控一切

---

人究竟对自己了解什么？他真的能够完全感知自己吗，就像陈列在明亮展示柜中的物品一样？自然界难道没有对他隐瞒大多数事情——甚至关于他自己身体的事情——以便将他限制和锁定在一个骄傲的、欺骗性的意识中，远离肠道的盘绕、血流的快速流动和纤维的复杂颤动！她扔掉了钥匙。

——弗里德里希·尼采，《论超道德意义上的真理与谎言》(1873)

只有在成年后，我才变得有死亡意识。十多年前，对我终结的内在洞察突然向我袭来。我浪费了整个晚上玩我十几岁儿子的一个令人上瘾的第一人称射击视频游戏——穿过诡异空旷的大厅、被淹没的走廊、噩梦般扭曲的隧道，以及异域阳光下的空旷广场，向无情追击我的外星人大军清空我的武器。我很晚才上床，像往常一样很容易入睡。几小时后我突然醒来。知识已经变成了确信——我要死了！不是在那里那时，而是某一天。

我没有对即将发生的事故、癌症等的任何预感——只是突然深深感受到我的生命将要结束的认识。死亡在十多年前进行了一次不速之访，当时我女儿Gabriele的同卵双胞胎妹妹Elisabeth在八周大时被婴儿猝死综合征(sudden infant death syndrome)夺走了生命。孩子不应该在父母之前死去；这严重违反了自然秩序。这可怕的经历污染了其后的一切，但奇怪的是，它并没有影响我对自己死亡的感觉。但这次夜间的洞察不同。我现在深深理解了，非常深刻地理解了，我也会死。对死亡的确定性一直伴随着我，使我更明智但不再快乐。

我对这个奇异事件的解释是，视频游戏中的所有杀戮触发了关于自我毁灭的无意识想法。这些过程产生了足够的焦虑，使我的皮层-丘脑复合体自己醒来，没有任何外部触发。在那一点上，自我意识点亮了，并面对了它的死亡性。这个独特而普遍的体验生动地让我认识到，在我头脑中发生的很多事情我都无法接触到。在大脑的某个地方，我的身体被监控着；爱、喜悦和恐惧诞生了；思想产生、被思考和丢弃；计划被制定；记忆被储存。有意识的我，Christof，对所有这些狂热的活动都毫无察觉。

抑制对等待我们所有人的必然厄运的了解，必定是进化过程中弗洛伊德所称的防御机制的一个主要因素（我们是唯一拥有这些机制的动物吗？黑猩猩能够压抑或抑制吗？）。这些是大脑将负面情绪、焦虑、内疚、不请自来的想法等从意识中清除的过程。如果没有这样的清理机制，早期人类可能会过度专注于自己的终极命运，无法成功主导自己的生存环境。也许临床抑郁症相当于这种防御机制的丧失。

但是在合适的触发条件下，无意识可以戏剧性地显现出来。每当我走在波士顿时，我都会尝试拜访伊丽莎白的墓地。在我深夜顿悟的几年后，我前往圣约瑟夫公墓朝圣。当我在温和的春雨中独自走过一排排墓碑时，我从远处注意到她的墓碑有些奇怪。走近后，我震惊地看到在刻有伊丽莎白名字的花岗岩块上有一个断翅的小陶制天使。在我女儿埋葬的地方，这个无助的小雕像立即唤起了几乎无法忍受的悲伤和失落感。我跪倒在地，在雨中哭泣。我给妻子打电话，她从远方安慰了我，但我在那天剩余的时间里仍然感到震惊。我从未弄清楚那个残损的雕像是如何到那里的。那天我学到的是，一个象征，在合适的语境中，可以突然释放长期休眠的记忆和情感。

在我的大学时代，两个亲密的朋友接受了原始疗法，这是一种由披头士约翰·列侬推广的心理疗法。我这个一向敏感的人，对他们坚持认为被压抑的记忆和本能欲望及需求在不为他们所知的情况下影响着他们这一点大加嘲笑。我激烈地断言，我完全掌控着我脑中发生的一切，谢谢。我坚决否认弗洛伊德无意识或我甚至不知道自己拥有的创伤

记忆，包括出生的痛苦——是的，这是原始疗法的重要组成部分（你不会惊讶地听到它起源于南加州）——正在影响我的行为。

三十年后，我变得更加审慎。我现在明白，“主权”“我”的行为是由很大程度上绕过意识检查的习惯、本能和冲动所决定的。我的神经系统如何让我的身体毫不费力地穿过充满匆忙行人的繁忙购物通道，它如何破译进入我耳朵的声音模式并将它们转化为某人正在提出的问题，我混乱的思想如何以合理的顺序从我的喉头和嘴巴中涌出成为言语，为什么我无法抗拒购买电紫色或帝王紫等艳丽颜色的衬衫——所有这些都超出了我意识自我的认知范围。这种缺乏意识的状态延伸到心智的最高层面。

如果你经历过情感充沛的时光，你会很熟悉怨恨和愤怒、恐惧和绝望、希望、悲伤和激情等强烈交叉暗流，它们是你的日常伴侣。有时，情感漩涡威胁着你的精神稳定。探索你自己潜在欲望、梦想和动机的洞穴，使它们变得有意识，从而可能使它们变得可理解，是非常困难的。精神分析和其他推理方法是不完美的；它们创造了一个新的虚构，一个基于关于人们为什么做他们所做的事情的直觉、民俗心理学概念的不同叙述。谈话疗法可能永远无法挖掘出一段关系破裂的真正原因：这些原因仍然被限制在大脑的黑暗地窖中，意识无法投射其窥探之光。

这些都不是新鲜事。亚精神的、非意识的或无意识的——我将其定义为任何不直接产生体验的处理——自19世纪后期以来一直是学术关注的话题。弗里德里希·尼采是第一个探索人类无意识中渴望支配他人并获得权力的黑暗角落的西方主要思想家，这种渴望经常被伪装成同情。在医学文学传统中，弗洛伊德论证童年经历，特别是那些具有性或创伤性质的经历，深刻决定了成人行为，而不被认识到它们的影响。这些弗洛伊德概念已经渗入日常语言，只是缓慢地被更多基于大脑的概念所取代。

让我从轶事性、自传性转向更客观的科学领域。我不会讨论神经质的上流社会患者躺在沙发上，以每小时200美元的价格滔滔不绝地谈论自己的案例研究。相反，我将叙述对一群每小时获得15美元报酬参与实验的大学生进行的实验。这些发现的压倒性结论是令人谦卑的：你的行为深受你无法察觉的无意识过程的塑造。

## 大脑中的僵尸

---

神经学和心理学侦查发现了各种专门的感觉-运动过程的动物园。这些伺服机制与传感器——眼睛、耳朵、平衡器官——相连，控制眼睛、颈部、躯干、手臂、手、手指、腿和脚，并服务于早晨的剃须、淋浴和穿衣；开车上班、在电脑键盘上打字和在手机上发短信；打篮球；晚上洗碗；等等。弗朗西斯·克里克和我称这些无意识机制为僵尸代理。总的来说，这支僵尸军队管理着肌肉和神经的流畅快速相互作用，这是所有技能的核心，构成了生活的体验。

僵尸智能体(zombie agents)类似于反射—眨眼、咳嗽、将手从炽热炉子上猛然缩回，或被突然的巨大噪音惊吓。经典反射是自动的、快速的，依赖于脊髓或脑干中的回路。僵尸行为可以被视为更加灵活和适应性的反射，涉及前脑。

僵尸智能体在意识雷达屏幕下方执行常规任务。你可以意识到僵尸智能体的行为，但只能在事后。我最近在山径跑步时，某种东西让我低头看。我的右腿立即加长了步幅，因为我的大脑探测到一条响尾蛇正在我即将踩到的石径上晒太阳。在我有意识地看到这条爬行动物并体验伴随的肾上腺素激增之前—在它发出不祥的警告嘶鸣之前—我已经避免踩到它并快速跑过。如果我依赖有意识的恐惧感来控制我的腿，我就会踩到蛇。

来自法国布隆认知科学研究所的Marc Jeannerod是动作神经心理学专家。他的实验，不涉及蛇，得出结论：动作确实可以比思维更快，纠正性运动动作的开始比有意识感知早约四分之一秒。为了将此放在透视中，考虑一个世界级短跑运动员在十秒内跑完一百米。当他有意识地听到手枪声时，跑者已经从起跑线冲出几步了。

无意识智能体通过训练而产生。反复重复相同序列强化各个组成部分，直到它们顺畅且自动地相互连接。你训练得越多，整体就变得越轻松和同步。训练给运动员和战士带来几分之一秒的优势，这决定了胜利与失败之间的差异—生与死之间的差异。

以扫视(saccades)为例，这是你不断扫描环境的快速眼动。你每秒将眼睛移动三到四次，一天100,000次—大约与你的心跳频率相同—然而你很少，如果有的话，意识到它们。你可以有意识地控制你的眼睛，比如当你将目光从你旁边那个人丑陋的裂唇上移开时，或当你避免与乞丐对视时，但这些都是例外。

加州大学圣塔克鲁兹分校的心理学家Bruce Bridgeman和其他人证明，你的眼睛看到你的头脑未意识到的细节。在一个实验中，一名志愿者坐在黑暗中，凝视一个发光二极管。当灯被关闭并在周边位置重新打开时，志愿者迅速将目光转向新位置。有时，“实验者”作弊，在志愿者的眼睛已经在途中执行扫视时，第二次移动灯光。她的眼睛没有错失节拍，准确落在灯光的新位置上，尽管它已经移动了。志愿者本人没有意识到灯光位置的第二次移动，因为视觉在扫视期间部分关闭。（这就是为什么你看不到自己眼睛的移动：试着在浴室镜子中来回移动眼睛。）确实，当灯光向内或向外稍微移动时，志愿者无法猜测灯光移动的方向，尽管她的眼睛仍然准确落在目标上。

扫视系统对它必须引导眼睛的位置极其敏感。鉴于其高度专业化，几乎不需要让意识参与其刻板的动作。如果你需要意识到并计划每一个眼动，你就无法做其他事情了。想象一下，如果你在追求思路时必须不断思考，“现在向左移，现在向下，现在到那里，然后再到这里！”，同时必须适当地控制十几块眼肌，会是什么样子。如果这些平凡的事情可以外包给专家，为什么要用它们堵塞体验呢？

僵尸智能体在此时此地运作。它们不为未来做计划。当你伸手去拿一杯热茶、在自行车上急转弯避开突然变道的汽车、回击网球截击或快速在键盘上打字时，你需要现在就行动，而不是几秒钟后。

获得新技能，如帆船或登山，需要大量的身心纪律。在攀岩中，你学习在哪里放置手、脚和身体来涂抹(smear)、支撑(stem)、后仰(lie back)、锁定手腕或手指在裂缝中。你注意将垂直花岗岩悬崖变成可攀爬墙壁的薄片和凹槽。一系

列不同的感觉-运动例程需要被缝合在一起，组装并编译成精细的运动程序。只有在数百小时的专注训练后，执行才变得自动，转变为通俗所说的肌肉记忆。持续重复招募一支僵尸智能体军队，使技能变得轻松，你身体的运动变得快速，没有浪费的努力。你从不考虑动作的细节，但它们需要瞬间协调和肌肉与神经的奇妙融合。

从需要注意力和占用意识的动作转向自动和无意识动作，是通过神经元资源从prefrontal cortex转移到basal ganglia和cerebellum来实现的。

矛盾的是，一旦学习完成，关注细节反而会干扰熟练的表现。专注于高度训练技能的任何一个方面，比如足球运球时右脚内侧接触球的确切时刻，会降低表现速度或使其更容易出错。当演奏一首很久没有弹过的熟练曲目时，最好让你的“手指去演奏”。意识到动作的流畅性或思考个别主题和音符序列会让你迷失方向。确实，夸奖你的网球对手反手动作姿势完美，下次他回球时就会关注他“完美”的姿势，结果球就会偏离目标。

范德比尔特大学的心理学家戈登·洛根和马修·克拉姆普在键盘打字这一我们时代的基本技能上证实了这个悖论。志愿者必须在三种条件下以正常打字速度输入屏幕上显示的单词。为了建立基准线，他们正常打字，双手并用。在左手试验中，他们必须抑制右手，只用键盘上传统分配给左手的字母打字。在右手试验中，右手只用来打字母。例如，如果在左手试验中显示器上出现单词“army”，受试者必须打出“a”和“r”，但不能打“m”或“y”，这些字母应该由右手来打。抑制这些自动的手部反应很困难，明显减慢了志愿者的打字速度，尽管只需要打一半的字母。然而，如果禁打的字母用绿色标出，允许打的字母用红色标出，受试者就能轻松阻止左手或右手打字，而且时间也很好。

哪只手打哪个字母是一个被委托给低级运动程序的细节。为了抑制动作而使这种知识变得有意识需要努力。然而，如果打字僵尸读取的输入已经标记为“不要打字”，就不需要挖掘这些信息，手可以及时停止，避免减慢整体打字速度。为了对抗这种“压力下的窒息”，教练和训练手册建议清空你的头脑。这释放了你内在的僵尸，带来了相当大的好处。

同样的思想在《箭术中的禅》这部沉思文学瑰宝中得到了表达。在这本书的结尾，作者欧根·赫里格尔对剑术有这样的描述：

学生必须培养一种新的感觉，或者更准确地说，培养所有感官的新警觉性，这将使他能够避开危险的刺击，仿佛他能感觉到它们的到来。一旦他掌握了躲避的艺术，他就不再需要全神贯注地观察对手的动作，甚至是多个对手的同时动作。相反，他看到并感觉到将要发生什么，在同一时刻他已经避开了其影响，感知和避开之间没有“毫发之差”。因此，重要的是：一种不再需要有意识观察的闪电般反应。至少在这个方面，学生使自己独立于所有有意识的目的。这是一个巨大的收获。

神经心理学家梅尔文·古达尔和大卫·米尔纳研究了一位在一次几乎致命的一氧化碳中毒事故中失去大部分视力的患者。她被诊断为视觉失认症(visual agnosia)，无法识别物体或看到它们的形状或方向（尽管她能看到颜色和纹理）。事故使她视觉皮层的部分区域缺氧，杀死了那里的神经元。因此，她无法区分水平线——比如邮箱中的槽——和垂直线。对她来说它们看起来是一样的。然而，如果她必须伸手将信件插入槽中，她会毫不犹豫地这样做，无论槽的倾斜度如何——水平的、垂直的或介于两者之间。尽管她没有意识到槽的方向，但她的视觉-运动系统可以获取这些信息，并顺利引导她的手，无需摸索。同样，当这位女士必须抓取放在她面前的物体时，她做得相当准确和自信，尽管她说不出自己拿的是细长的玻璃笛子还是大马克杯。她失去了负责识别物体的视觉皮层部分，但使用视网膜输入来塑造和引导她的手和手指的区域仍然功能正常。

早期的临床工作，加上猴子实验的支持，产生了两个不同视觉处理流的概念。两者都起源于初级视觉皮层，但随后分化，支配皮层中专门负责高级视觉和认知功能的不同区域。一个流经V2和V3进入下颞叶皮层和梭状回。这是感知、腹侧或什么通路。另一个，即动作、背侧或哪里通路，将数据传递到后顶叶皮层的视觉-运动区域。基于对他们

这位特殊患者的视觉-运动技能和缺陷的仔细分析，米尔纳和古达尔推断，她的什么通路（介导有意识的物体视觉）被近乎窒息的经历破坏了，而她的哪里通路（引导她的手和手指）基本上保持完整。其目的不是无休止的沉思——某些前额叶结构的特权——而是直接行动。功能良好的大脑紧密整合这两个通路，有许多交叉连接；大脑的拥有者甚至没有意识到他或她所体验的无缝界面是由两个或更多信息流组成的。

## 社会无意识

---

我们是极度社会化的动物。名人和八卦节目、杂志以及充满最新新闻聊、照片和关于聚会、婚外情、背后中伤和阴谋、非婚生子等传言的网站的流行，只是人们对彼此行为无穷好奇心的一种相当热情的表达。如果你认为人类是高尚的，只需查看谷歌的Zeitgeist档案中的前十大搜索词条。电影和流行明星、乐队、顶级运动员和时事政治是长期热门，其中几乎没有科学家或科学发现（这还是在最受欢迎的与性相关的搜索词条已经被筛选掉之后）。

没有人是孤岛。即使是隐居者也通过与他人的关系来定义自己，如果不是在现实生活中，那就通过书籍、电影或电视。

你可能像我年轻时一样坚信，你的有意识意图和deliberate choices控制着你与家庭成员、朋友和陌生人的互动。然而数十年的社会心理学研究清楚地表明并非如此。你的互动很大程度上受到你所不了解的力量支配，受到无意识的欲望、动机和恐惧支配。

美国心理学之父、小说家亨利·詹姆斯的兄弟威廉·詹姆斯认为，思考或观察某种行为的行为会增加参与该行为的倾向。这种*ideomotor action*（观念运动作用）原理是大脑皮层中感知和行动表征部分重叠的自然结果。在猴子大脑中发现的所谓*mirror neurons*（镜像神经元）证实了这样的观念：对行动的感知与执行该行动密切相关。

看着某人吃东西会激活你大脑中在你吃东西时被触发的区域，尽管很微弱。看到某人出丑，你会感到难受，因为你也可能会感到一点尴尬。当某人对你微笑时，你会感觉好一点。当你对某人有好感时，你会模仿那个人的行为和言语。下次你在咖啡馆见朋友时观察这种现象。你们两人可能会用同一只胳膊肘靠在桌子上，以同样的方式倾斜头部。当你低声说话时，你的朋友也会如此。当她挠头时，你也会这样做。当你打哈欠时，她也会如此。行为的reciprocity（互惠性）在讨好某人时也有帮助。

无穷无尽的多种因素影响着你与人的日常接触。他们的年龄、性别、种族、穿着、举止和情感表达会在你心中留下印象，并指导你接近、交谈和评判他们的方式。所有这些都逃脱了意识审查；因此“第一印象”的重要性。

一些人公开对特定群体持有强烈、通常是负面的观点：“自由主义者憎恨国家”、“基督徒是反科学的狂热分子”、“黑人男性具有攻击性”、“老年人很无聊”等等。这样的偏执者会基于他们的偏见做出选择。但即使你努力避免刻板印象，你仍然怀有无意识的偏见和倾向。你无法逃脱成为你的文化和成长环境之子的事实，从童话故事和神话、书籍、电影和游戏中，从你的父母、玩伴、老师和同龄人那里吸收*implicit judgments*（隐性判断）。如果你不相信我，只需做一个*implicit association test*（内隐联想测试）（我推荐哈佛的测试，你可以在线找到），在其中尽可能快地回答一系列问题。该测试以间接、斜向的方式测量你对特定宗教、性别、性取向或种族群体的偏见程度，这种方式抗manipulation（操控）或说谎。

无意识偏见可能比有意识偏见更有害，有两个原因。

首先，无意识倾向是广泛而自动的，在其trigger（触发器）出现时被激活。它们可能非常强大并被整个社区共享。

考虑1941年12月7日日本帝国海军对驻珍珠港美国太平洋舰队的突然袭击，以及2001年9月11日基地组织对世贸中心和五角大楼的袭击。两者都是致力于保护国家免受此类灾难的大型组织的巨大情报失败。学术和新闻调查发现了大量信息，这些信息在袭击发生前几天、几周和几个月就指向了即将发生的打击。在9·11事件的情况下，情报人员曾40次警告政府本·拉登构成的威胁；全都是徒劳。

为什么？无数政府小组和书籍得出了类似结论。存在多层次的无能。然而更加insidious（阴险）、比个人未能留意警告更广泛的，是那些本可以产生影响的人们心中明确和implicit（隐性）的种族傲慢和文化condescension（居高临下）态度。负责太平洋舰队的Admiral Kimmel（金梅尔上将）在众多国会调查中的一个不加掩饰的时刻完全清楚地表明了这一点：“我从未想过那些黄皮肤的狗崽子能在离日本如此远的地方发动这样的攻击。”50多年后，国防部副部长Wolfowitz对他的对手持同样的轻视态度，将本·拉登dismissing（斥为）“阿富汗的这个小恐怖分子”。广泛的stereotyping（刻板印象）蒙蔽了掌权者：“那些住在洞穴里头上缠着毛巾的未受教育的人怎么能威胁到我们这个地球上最强大的国家？”导致雷曼兄弟金融崩溃并几乎在2008年9月crashed（摧毁）市场的事件是这种思维pathology（病理学）的另一个例子。在这里，广泛持有的投资风险已得到控制并通过适当的金融工具leveraged away（杠杆化消除）的信念导致了全球recession（衰退）。

无意识倾向可以被有意地塑造以获取政治或经济利益。例如，在媒体饱和的民主国家中，每次重大选举前几个月充斥“新闻”的愚蠢内容令人极度厌烦，但它们确实产生了效果。我们如此沉浸在广告业的产品中，以至于不再注意到它们。然而，该行业在2010年全球投入约5000亿美元来影响消费者的购买决策是有原因的。它确实有效！

其次，虽然法律和公众意识活动可以根除公开的歧视，但无意识偏见却难以对抗。当你甚至没有意识到有需要改变的东西时，你如何能够改变它呢？

如果你还不相信，请考虑耶鲁大学John Bargh的实验，这些实验揭示了社会*priming*的力量。在*priming*中，图片、声音或文字影响对后续刺激的处理。如果你愿意配合并说出我将要说出的物体的颜色，我来演示*priming*。不要觉得尴尬。试着大声说出颜色，你会感到惊讶。开始：

空白纸张是什么颜色？婚纱是什么颜色？雪是什么颜色？蛋壳是什么颜色？

现在，不要思考，脱口而出这个问题的答案：“牛奶早餐喝什么？”

如果你和大多数人一样，你会想到或说出“牛奶”。只有过一会儿你才会意识到这完全是胡说八道。反复提到“白色”触发或*primed*了与其他白色事物相关的神经元和，可能最关键的是，突触活动。如果接下来要求你说出一种与牛奶和早餐都相关的液体，“牛奶”会自动浮现在脑海中。

Bargh想要为大学生*priming*“粗鲁”和“礼貌”的概念。他让学生从一列单词中生成句子，表面上是测试他们的语言技能。一组使用诸如*bold*、*rude*、*disturb*、*intrude*、*brazen*和*impolitely*等词，而第二组使用传达相反含义的词，如*respect*、*patiently*、*yield*和*courteous*。完成后，受试者被告知在走廊寻找实验者进行第二项测试。但实验者假装正忙着与同谋交谈，受试者不得不等待。

Bargh和他的同事秘密计时被*primed*的志愿者打断这次谈话需要多长时间。被礼貌词汇*primed*的学生表现出惊人的耐心，等待超过九分钟才被打断，而粗鲁组仅坚持了五分钟多一点就插话了。没有一个参与者怀疑单词测试影响了他们等待的时长。这个实验的结论——你听到或使用的词汇塑造你的行为——对我祖母来说并不是什么新闻，她总是说给小费、带小礼物和保持礼貌会以未知的方式得到回报。

这种技术的一个变体测量与老年人相关的偏见。Bargh让志愿者使用触发老年人刻板印象的词汇，如*old*, *lonely*, *forgetful*, *retired*, *wrinkled*, *ancient*, *helpless*和*Florida*，而对照组必须用中性词汇组成句子。他巧妙地挖掘老年偏见的方法是秘密计时志愿者从测试地点走到电梯所用的时间，距离约十米。被“*elderly*”*primed*的学生用了8.28秒，比接触与衰老无关词汇的组多约1秒。这是一个小但真实的效果。学生们完全没有察觉到这些词汇促使他们走得更慢。如果无关紧要地阅读与你没有直接关系的词汇都能让你放慢脚步，那么当朋友或配偶告诉你正在变老时，你的反应会多么强烈？所以要谨慎用词。

自助运动坚持认为拥有积极乐观的态度会产生影响。虽然积极思考不会治愈癌症，但它会塑造你的行为。这是我住在这里并从事现在工作的原因之一——美国人总体上，特别是自发选择移居西海岸的移民群体，相信只要有足够的渴望、汗水、奉献精神和明智地应用技术，几乎一切皆有可能。我认同这种能够做到的态度。当一个人付出了一切却失败，这是光荣的；因为害怕失败而不去尝试是一个重大的性格缺陷。

无意识处理的一个显著方面是，包括我年轻时的自己在内，许多人强烈否认它的存在。这种本能的防御反应在学者中尤其强烈，他们认为自己比其他人更客观、平衡和公正。作为一个群体，教授们竭尽全力在招聘和指导中补偿性别和种族偏见。但在政治和宗教方面，学者可能非常不宽容，将只有边缘人群才持有的观点归因于保守派或虔诚信徒。他们对大多数宗教，尤其是基督教，几乎难以掩饰的蔑视如此普遍，以至于许多大学生都不敢表达任何宗教情感。

你可能会因为两个原因反对潜意识的重要性。首先，因为接受它意味着失去控制。如果不是你在这里做决定，那么是谁呢？你的父母？你如此热切消费其产品的媒体？你的朋友和同龄人？其次，因为你没有意识到潜意识偏见（根据定义），你不知道自己有这些偏见。你不会回想起自己曾经秘密地基于某人的肤色、性别或年龄来判断他们的情况。当有人指出这样的情况时，你会想出许多看似合理的理由来解释为什么你会那样判断那个人——但你歧视了那个人的想法不会出现在你的脑海中。这很神秘，但大脑就是这样工作的。

为了强调这一点，让我告诉你关于选择盲视(choice blindness)的事。瑞典隆德大学的一百多名学生必须比较两张年轻女性的头像照片。实验者将每张照片拿在手中，并排放置，学生们必须在几秒钟内决定哪张看起来更有吸引力并指向她。然后照片被短暂地移出视线。紧接着，学生们再次看到他们认为更漂亮的面孔，并被要求解释他们的选择。在一些试验中，实验者在要求学生解释他们的选择之前，使用障眼法交换了照片。尽管两张照片中描绘的女性相当不同，大多数人都被愚弄了。只有四分之一的学生意识到照片已被调换，他们正在看的不是他们刚才选择的那张（在这些情况下，实验立即中断）。其他学生轻松地为他们的“选择”辩护，尽管它与他们几秒钟前做出的决定相矛盾：“她很有光彩。我宁愿在酒吧接近她而不是另一个。我喜欢耳环”，尽管他们选择的女性看起来严肃且没有戴耳环。

选择盲视不仅与约会相关，还与整个生活相关。你经常几乎不知道为什么要这样做你发现自己正在做的事情。然而解释的冲动如此强烈，以至于你当场编造一个故事，为你的选择辩护，在不自觉的情况下编造故事。

一些心理学家将拥有庞大处理资源的隐藏而强大的大脑这一概念推得更远。他们认为，当你必须做出需要平衡许多竞争因素的决定时，潜意识处理实际上优于有意识的思考。考虑选择租赁什么公寓。这取决于许多变量：月租金、可用日期、大小、位置、合同期限、房间状态等等。根据潜意识思维理论(unconscious thought theory)，在你获得了所有相关的住房信息后，你不应该试图内省，而应该分散注意力——比如解填字谜——然后再决定租哪个地方。不要为这个问题烦恼。想想别的事情，你的无形大脑会为你解决问题。

这些建议以其承诺获取潜意识心智的巨大力量而引起了大量公众关注。然而这些实验中的许多都相当“软”，统计显著性低且没有令人信服的对照。这些缺陷是研究人类固有的——他们的遗传学、环境、饮食、体育活动等等都很难得到适当的控制。对数据更谨慎的解读是，形成快速印象并做出有意识的决定可能比无休止地质疑第一眼评估更好。做出决定，相信自己，并坚持下去。

我同样对所声称的潜意识优越性持怀疑态度，无论是在方法学还是理论基础上。如上所述，历史上充满了广泛采用潜意识偏见导致灾难性后果的例子。现实生活中的决定总是涉及有意识和潜意识过程的混合，有些决定更多地依赖于其中一种或另一种。我还没有看到令人信服的证据表明，“如果——那么”类型的命题推理、复杂的符号操作，或处理计划外的突发情况可以在没有有意识的、深思熟虑的、耗时的、皱眉思考的情况下成功解决。否则，我们都会是爱因斯坦。这个结论也与几千年来传统教学的一致，即在做任何重要决定之前都需要自我检查和理性、冷静的思考。

## 潜意识的普遍影响对我的探索意味着什么？

---

醒来后，你的意识需要几秒钟来启动。一旦你确定了方向，它就为你提供了一个稳定的界面来应对令人眼花缭乱的丰富世界，没有冻结或恼人的蓝屏死机需要你的大脑重启。像任何好的界面一样，真正的工作发生在表面之下。这些是从你的眼睛捕捉反射光束到感知美丽女性的各种过程。

理解潜意识的重要性是心理学和神经科学持续面临的挑战。这对于理解你的生活也是必要的。因为如果没有自我检查，没有理解你的行为不仅仅是深思熟虑的、有意识选择的结果，你就无法改善自己。这里回顾的数据都没有表明潜意识具有可以被利用来解决你的爱情、家庭、金钱或职业问题的巨大而迄今未开发的力量。这些只能通过多年来培养的深思熟虑和有纪律的行为和习惯来解决，这是一个很少有人愿意听到的无聊信息。

我的追求是理解意识而非其缺失。无意识处理远不如意识神秘；毕竟，这正是计算机所做的事情。意识如何进入世界是令人困惑的。尽管如此，僵尸行为以及无意识欲望和恐惧的广泛存在对我的追求很重要，原因有三个。

首先，这引起了我可能本末倒置的担忧。如果潜意识、非意识的领域如此无所不包，那么大脑及其活动的大部分可能根本与意识无关。

这是真的！我一直在努力指出，视觉意识的神经相关物不在脊髓、小脑、视网膜或初级视觉皮层中。我怀疑皮层高级视觉区和前额叶区域的大多数神经元都是配角。也许只有一层薄薄的长程锥体神经元网络，它们相互连接前后区域，才负责意识内容。如果被迫给出一个数字，我会推测在任何给定时间，只有一小部分神经元活动直接参与构建有意识的知觉。健康、清醒大脑标志性的无休止神经元活动中的绝大部分都对意识起着从属作用。

为了让你不觉得这不可信，请考虑以下意识机制与遗传机制之间的类比(像任何类比一样，这个类比是不完美的)。将信息从一个细胞传递给其后代的分子过程，包括复制、转录和翻译，涉及数百种精巧的生化装置——DNA、tRNA、mRNA、核糖体、支架和中心体，仅举几例。然而，组装细胞的详细指令、蓝图，存在于单一的双链DNA分子中——在构成每个细胞的数百万分子中的一个非常长且稳定的分子。其序列中的一个拼写错误就可能导致严重的后果。意识的潜在机制可能同样特异。敲除皮质-丘脑神经元联盟中的任何一个成员都可能导致相关有意识知觉或思想的微小扰动。

其次，僵尸代理使意识研究者的生活变得更困难，因为它们迫使我们将行为与意识分开：有目的的、常规的、快速的行为本身并不意味着感知能力。仅仅因为一个重伤的人在有人进入病房时转动眼睛，并不意味着她有任何情境意识。这对早产儿、狗、小鼠和苍蝇同样成立。刻板行为不能保证主观状态。需要更多条件才能证明一个有机体具有现象体验。

第三，僵尸例程的广泛存在使意识神经相关物的问题变得突出。产生差异的差异在哪里？这仅仅是大脑皮层正确区域被激活的问题吗，正如Milner和Goodale所声称的？背侧区域用于无意识行为，腹侧区域用于有意识视觉？还是相同的回路可以参与两者，这取决于处理模式？

Francis和我认为，短暂的神经活动——它快速离开视网膜并迅速通过皮层的视觉-运动区域传递到运动神经元——对于意识来说是不充分的。这需要皮质-丘脑神经元的单一联盟建立优势并维持一段时间，这更类似于物理学中的驻波。我将在第9章中继续讨论改造小鼠以防止形成主导联盟的主题。

现在让我提出身心问题的一个关键方面，这个问题我在《意识的探索》(2004)中故意回避了——即大脑在其行为中有多少自由的问题。自由意志是哲学的卓越主题。它的根源可以追溯到古代。这是一个迟早会面对我们每个人的话

题。令人惊讶的是，这个问题的一个关键方面可以简化为知觉意识的问题。对我来说，这构成了形而上学中最具争议问题之一的重大进展。

## 第7章：我抛开谨慎，提及自由意志、《尼伯龙根的指环》以及物理学对决定论的看法，解释你心智选择的贫乏能力，表明你的意志滞后于你大脑的决定，而自由只是感觉的另一个词

---

你看，只有一个常数。一个普遍性。这是唯一的真理。因果关系。行动，反应。因果效应。

——《黑客帝国2：重装上阵》(2003)中的Merovingian

在宇宙的一个偏远角落，在一颗小蓝色星球上，它围绕着银河系不时尚的外围地区的一颗平凡太阳运转，有机体从原始泥浆和软泥中兴起，在一场跨越永恒的史诗般的生存斗争中。尽管所有证据都与此相反，这些双足生物认为自己异常特权，认为自己在拥有万亿颗恒星的宇宙中占据独特位置。他们如此自负，甚至相信他们，而且只有他们，能够逃脱支配一切的因果律的铁律。他们能够凭借他们称为自由意志的东西做到这一点，这使他们能够在没有任何物质原因的情况下做事。

你真的能自由行动吗？你能做出和说出那些不是直接源于你的性格倾向和环境条件的事情吗？你是出于自己的意愿选择阅读这本书的吗？你感觉自己是在面对其他竞争选择时自愿决定翻阅这些页面的，比如吃午饭或给朋友发短信。但这就是全部的真相吗？难道没有外在的原因影响了你——比如课堂作业或朋友称赞这本书流畅的文风？你可能会争辩说这些原因还不够充分，必须有其他因素介入：你的意志。然而，宿命论及其世俗化表亲决定论认为，你不可能以任何其他方式行动。在这件事上你没有真正的选择。你是绝对专制者的终身契约奴仆。你从来没有选择午饭的选项，而是从时间的开端就注定要阅读我的书。

自由意志问题不仅仅是哲学争论；它以一种很少有其他形而上学问题能做到的方式吸引着人们。它是社会责任观念、赞扬与谴责观念、因为你所做的事情而受到评判观念（无论好坏）的基石。归根结底，它关乎你对自己生活的控制程度。

你和一位可爱又深爱的配偶生活在一起。与陌生人的偶然相遇，仅仅持续几个小时，就把这种生活完全颠倒了。你质疑一切并搬了出去。你在电话里聊几个小时，你分享内心最深的秘密，你开始了一段恋情。你坠入爱河，这是一种令人陶醉而强烈的情感混合物，具有强迫症的许多特征。你非常清楚地意识到从道德角度来看这一切都是错误的；它会给许多人的生活带来破坏，而且不能保证有一个幸福而富有成效的未来。然而你内心的某些东西渴望改变。

这种令人痛苦的选择让你面对一个根本问题：你在这件事上到底有多少发言权？你是否只是在按照进化的指令行事，按照你DNA的古老舞蹈寻找繁衍自己的新途径？你的荷尔蒙，你的欲望，给你留下了任何自由吗？你当然觉得原则上你可以结束这段婚外情并回家。但尽管多次尝试，你却从未成功做到。你明知故犯地驶入完美风暴的风眼，撞毁了多年来安全载着你的婚姻之船。

自由意志是一个学术雷区。人们已经为任何可以想象的立场提出了深奥的论据，无论支持还是反对。在思考这些问题时，我忽略了数千年来博学的哲学辩论，而专注于物理学、神经生物学和心理学的观点，因为它们为这个古老的难题提供了部分答案。

## 强自由意志与实用主义自由的不同色调

---

让我先提供一个自由意志的定义：如果在完全相同的情况下，你本来可以采取不同的行动，那么你就是自由的。完全相同的情况不仅指相同的外部条件，还指相同的大脑状态。这是强、自由主义或笛卡尔立场，正如我们不断遇到的笛卡尔所阐述的那样。带有大写“W”的意志，真正的意志。

想想《黑客帝国》中的经典场景，Neo必须决定是否吞下墨菲斯提供给他的蓝色药丸（承诺幸福的无知）还是红色药丸（承诺痛苦地觉醒到严酷的现实）。Neo自由选择红色药丸意味着他本来可以同样容易地选择蓝色药丸，从而让我们失去近年来最引人入胜的电影之一。强意志意味着即使Neo的欲望、恐惧和想法，他大脑和环境中的一切，都与选择红色药丸时完全相同，他本来也可以选择蓝色药丸。

我最近在洛杉矶联邦地区法院担任陪审员。被告是一个身上有大量纹身的街头帮派成员，从事毒品走私和销售。他被控谋杀一名帮派同伙，向其头部开了两枪。尽管这场审判迫使 I 突然搁置生活去法庭听取证据，但这次经历大大开阔了我的视野。它以残酷的细节揭示了一个围绕着枪支、毒品、现金、尊重和名誉的高度部落化社会的生活。街头帮派居住在一个与我和我的家人、朋友、同事所生活的特权和受保护世界平行的宇宙中。这些世界虽然只相距几英里，却很少相交。

当执法部门、亲属以及过去和现在的帮派成员——其中一些人在作证时被戴着手铐、脚镣，穿着鲜艳的橙色监狱制服——描述犯罪背景时，我思考着塑造被告的个人和社会力量。他曾经有过笛卡尔式的选择吗？他的暴力成长经历是否让他必然走向杀戮？幸运的是，陪审团不需要回答这些无法解决的问题或决定他的惩罚。我们只需要排除合理怀疑地决定他是否有罪，是否在特定地点和时间射杀了特定的人。我们做到了这一点。

强自由定义除了激烈而最终无益的辩论外，对任何事情都没有用处，因为在现实世界中你无法回头做不同的事情。正如古代哲人赫拉克利特所说：“你不能两次踏进同一条河流。”然而这种笛卡尔式的意志观点是大多数普通人相信的。它与灵魂概念密切相关。灵魂像格兰芬多学院的鬼魂差点没头的尼克一样悬浮在大脑上方，自由地选择这个或那个，让大脑按照其意愿行动，就像司机驾驶汽车选择走这条路或那条路一样。

将强烈的自由概念与一种更为实用的概念——相容论进行对比，这是生物学、心理学、法律和医学界的主流观点。如果你能够追随自己的欲望和偏好，你就是自由的。决定论和自由意志可以共存。它们彼此相容。只要你不受某种内在强迫的支配，也不在他人或外力的不当影响下行动，你就是自己命运的主宰。一个想要戒烟但一次又一次点燃香烟的长期吸烟者并不自由。他的欲望被成瘾所挫败。根据这个定义，我们中很少有人是完全自由的。相容论不依赖于灵魂等超世俗实体。它完全属于这个世界。

很少有人——我想到了圣雄甘地——能够坚定地让自己为了更高的伦理目的而连续数周拒绝食物。另一个极端的铁一般自制力案例是佛教僧人释广德于1963年的自焚，以抗议南越的专制政权。这一事件如此独特，被摄入令人难忘的照片中，这些照片至今仍是二十世纪最容易识别的图像之一，其特殊之处在于他英勇行为的平静和深思熟虑。在被火焚身而死的过程中，广德始终保持着冥想的莲花坐姿，没有移动一块肌肉或发出一声声响，任由火焰吞噬他。

对于我们其余那些努力避免吃甜点的人来说，自由始终是一个程度问题，而不是我们拥有或不拥有的绝对善。

刑法承认责任减轻的情况，即被告没有自由行动。当丈夫在当场捉奸时盛怒之下将妻子的情人打死，与他在数周后以冷静、有预谋的方式寻求报复相比，前者被认为罪行较轻。以冷血方式射杀六十多人的偏执型精神分裂症患者被认定为“因精神失常无罪”并被关押在精神病院。在没有此类减轻情节的情况下，被告被假定为具有出庭受审的能力。当代社会和司法制度建立在这样一种实用的、心理学的自由概念之上。

理查德·瓦格纳的不朽之作《尼伯龙根的指环》是一系列四部歌剧，围绕命运与自由之间的冲突展开。不受恐惧或社会道德约束的英雄齐格弗里德杀死了龙，穿过火环去追求布伦希尔德，粉碎了沃坦的长矛，导致了众神旧世界秩序的毁灭。齐格弗里德不遵循任何法律，只遵循内心的欲望和冲动。他是自由的，但他盲目行动，不理解自己行为的后果。（齐格弗里德很可能在amygdala有病变——他不知恐惧——以及ventromedial prefrontal cortex有损伤，剥夺了他的决策能力。遗传和发育因素导致了他的功能失调行为：他的父母是兄妹；他作为孤儿被一个唯一的照顾者抚养长大，这是一个爱争吵的侏儒，痴迷于一堆黄金；他在德国森林深处孤立地长大。这种社交技能的缺乏最终导致他被信任的朋友哈根谋杀。）歌剧的女主角布伦希尔德通过自我牺牲，自由而明知地迎来人类的新时代。这部戏剧配有一些有史以来最非凡、质感丰富、多层次且动人的音乐。从相容论的观点来看，齐格弗里德和布伦希尔德都是自由行动的。

但我想要深入挖掘。我想要揭示这种“自由”行为的根本原因。你每天的马拉松是一系列选择的考验——穿哪件衬衫、听哪个广播电台、点哪道菜等等。我在前一章报告的证据应该提醒你，你的大部分行为逃脱了意识的内省和控制。你的自由受到你过去形成的习惯和一贯选择的限制。承载和引导你意识流的河床本身就是由抚养你的家庭和文化塑造的。你“自由”表现出的欲望和偏好似乎是完全被决定的！

相容论的自由留下了不安的残余，一种挥之不去的疑虑。缺乏明显的内在或外在胁迫当然是自由的必要条件，但这并不能保证强意义上的自由。如果来自天性、后天培养以及环境中随机因素的所有影响都被考虑在内，还有任何操作空间吗？你难道不是这些约束的彻底奴隶吗？看起来相容论等同于自由的简化版！我们的概念挖掘工作是否已经触及决定论的基础岩石？

物理学对这个问题有什么看法？

## 经典物理学与决定论：发条宇宙

---

人类理解宇宙的持续过程中的一个高峰出现在1687年，当时艾萨克·牛顿发表了他的《原理》。它阐述了万有引力定律和三大运动定律。牛顿第二定律将作用在一个孤立系统——在绿色毡桌上滚动的台球——上的力与其加速度联系起来，即与其速度变化联系起来。这个定律具有深远的后果，因为它意味着组成一个实体的所有组件在任何特定时刻的位置和速度，连同它们之间的力，不可改变地决定了该实体的命运——即它未来的位置和速度。没有其他东西介入；不需要其他东西。系统的命运直到时间的尽头都已注定。

这一定律的效力延伸至整个大地——从树上掉落的苹果，到月球绕地球运行的轨道，或是数十亿颗恒星围绕银河系中心旋转——牛顿定律统治着它们所有。给我这些力以及系统的当前状态——物理学家用来精确指定系统所有组成部分的确切位置和速度的简略表达——我就能告诉你该系统在未来任何时间点的状态。

这就是决定论的精髓。行星在围绕太阳的轨道中运行时的质量、位置和速度，完全决定了它们在从今天起一千年、一百万年或十亿年后将处于何处，只要所有作用于它们的力都得到适当考虑。如果将整个宇宙作为一个整体来考虑，牛顿定律同样有效。这一概念性飞跃在法国数学家皮埃尔·西蒙·德·拉普拉斯那里找到了最雄辩的支持者，他在1814年写道：

我们可以将宇宙的现在状态视为其过去的结果和未来的原因。一个在某一时刻能够知晓推动自然运动的所有力，以及构成自然的所有物体的所有位置的智慧，如果这个智慧也足够强大，能够将这些数据进行分析，它将在一个单一的公式中包含宇宙最大天体和最微小原子的运动；对这样的智慧来说，没有什么是不确定的，未来就像过去一样会呈现在它眼前。

宇宙一旦启动，就会不可避免地运行其轨道，如同钟表机械。对于一台全知的超级计算机来说，未来是一本敞开的书。除了物理定律所规定的之外，没有自由可言。你与内心恶魔斗争的所有努力，无论好坏，都是徒劳的。你未来行动的结果在时间开始时宇宙启动时就已注定。

十一世纪的波斯天文学家、数学家和诗人欧玛尔·海亚姆在他的《鲁拜集》中明白地表达了这一点：

那倒扣的碗我们称之为天空，

在其下我们匍匐生死如笼中，

莫向其举手求助——因为它

无力地滚动着，如你我一般。

这台巨大机器并非如预期那般可预测的第一个暗示，来自十九世纪最后十年的法国数学家亨利·庞加莱。直到数字计算机的出现才揭示了决定论混沌的真面目——对未来可以精确预测这一观念的全面挫折。气象学家爱德华·洛伦兹在解决三个描述大气运动的简单数学方程的过程中发现了这一点。当他输入仅有微小差异的初始值时，计算机程序预测的解决方案差异很大。这就是混沌的标志：方程初始点的无限小扰动导致截然不同的结果。洛伦兹创造了蝴蝶效应一词来表示这种对初始条件的极度敏感性：蝴蝶翅膀的拍打在大气中产生几乎无法察觉的涟漪，最终改变了其他地方龙卷风的路径。

股票市场是混沌系统的一个好例子。微小的干扰——关于董事会冲突的谣言或遥远土地上的罢工——可能以不可预测的、不稳定的方式影响公司股票的命运。混沌也是精确的长期天气预报不可能实现的原因。

牛顿和拉普拉斯钟表宇宙的缩影是天体力学。行星沿着重力的geodesics(测地线)庄严地运行，由形成太阳系的云团的初始旋转推动。因此，当1990年代的计算机建模证明冥王星具有混沌轨道，发散时间为数百万年时，这是一个巨大的惊喜。天文学家无法确定一千万年后冥王星是在太阳的这一侧（相对于地球的位置）还是另一侧！

如果这种不确定性适用于一颗内部构造相对简单、在太空真空中仅受重力这一单一力作用而运动的行星，那么对于人类、微小昆虫或极小的神经细胞这些受到无数因素影响的个体的可预测性意味着什么呢？

考虑一群数百只基因相同的果蝇，它们一起孵化并在12小时明暗周期下的塑料管中饲养。即使在控制良好的实验室条件下，果蝇的行为也是反复无常的。当它们被释放到迷宫中并接近分叉处时，有些会选择左分支，有些选择右分支，有些会转身返回。其他果蝇会停留在原地，无法做出决定。未来的生物学家将能够预测果蝇群体在这种情况下的行为，然而预见任何单个果蝇的选择将被证明与预测一只股票的命运一样不可能——且出于同样的根本原因，决定论混沌。

然而，蝴蝶效应并不会使因果关系的自然定律失效。它继续占据主导地位。行星物理学家不太确定冥王星在永恒后会在何处，但他们相信其轨道将始终完全受重力支配。在混沌中破坏的不是行动与反应的链条，而是可预测性。宇宙仍然是一个巨大的钟表机械，尽管我们无法确定一周后分针和时针将指向何方。

这个观点在生物学中同样适用。任何细胞器，如突触或细胞核，都由悬浮在水溶液中的数量庞大的分子组成。这些分子不断相互碰撞，以一种无法精确捕捉的方式四处运动。为了驯服这种热运动，物理学家依靠概率论的工具。然而，分子过程的随机性并不是由于微观尺度上决定论的崩溃造成的。不是的，实际上是出于实用考虑，无法追踪无数分子的运动。在经典物理学定律下，不可否认的是，给定所有分子的作用力、初始位置和速度，它们的未来状态将不可避免地遵循确定的轨迹。

请记住我的话——如果物理决定论成立，就不存在笛卡尔式的自由。宇宙中将要发生的一切，包括你的所有行为，在宇宙诞生时就已经固有了。所有事件都是预先注定的。你注定要观看一部专门为你放映、持续一生的电影。导演——物理定律——对你改变任何一个场景的恳求都充耳不闻。

## 发条宇宙的消亡

---

这种宿命论的宇宙观随着20世纪20年代量子力学的诞生而发生了决定性的改变。量子力学是我们对原子、电子和光子在非相对论速度下最好的描述。其理论大厦在预测能力上令人惊叹，是人类最高的智力成就，无与伦比。

对牛顿-拉普拉斯梦想——或者在我看来是噩梦——的致命一击是著名的量子力学不确定性原理，由维尔纳·海森堡于1927年提出。这是对粒子位置和动量能够同时被精确测量的一个不可简化的限制（粒子的动量是其质量乘以速度）。在最常见的解释中，海森堡原理断言宇宙的构建方式使得任何粒子，比如光子或电子，都不能同时具有确定的位置和确定的动量。如果你准确知道它的速度，其位置就相应地不明确，反之亦然。这个原理代表的不是当今仪器的粗糙程度（这可以通过更好的技术来克服），而是现实的本质结构。像我的红色Mini敞篷车这样的宏观重物体，在高速公路上以明确定义的速度移动时，在空间中占据精确的位置。但像基本粒子、原子和分子这样的微观事物违背了常识：你越精确地确定它们在哪里，它们的速度就越不确定、越模糊，反之亦然。

海森堡的不确定性原理是对经典物理学的根本背离，其影响尚未完全明确。它用模糊性取代了教条式的确定性。一切的基础是一个叫做波函数的数学抽象，它以薛定谔定律决定的确定方式演化。从这个定律中，物理学家推导出任何给定事件的概率，比如电子占据氢离子特定原子轨道的机会。概率本身可以被精确计算到令人惊异的程度，但电子在任何特定时刻实际位于何处是不可能确定的。

考虑一个实验，结果是电子有90%的机会在这里，10%的机会在那里。如果实验重复一千次，大约九百次试验中，电子会在这里，误差不大；否则，它会在那里。然而，这种统计结果并不决定电子在下次试验中会在哪里。它更可能在这里而不是那里，但它实际最终在哪里是偶然的。阿尔伯特·爱因斯坦永远无法接受自然的这种随机性。正是在这种背景下，他著名地宣称：“Der Alte würfelt nicht”（老人（即上帝）不掷骰子）。

当你仰望天空时，有令人惊叹的证据证明这种随机性。星系在浩瀚太空中分布并不均匀。它们聚集成细长的丝状结构，排列在片状和壁状结构中，围绕着空旷的虚无，其巨大的空虚程度令人震撼。光束需要数百万年才能穿越这样的深渊！我们自己的银河系是室女座超星系团的一部分，包含数十万亿颗恒星。

根据宇宙学的暴胀理论，这些超星系团——宇宙中最大的结构——是由大爆炸后瞬间发生的随机量子涨落引起的，大爆炸形成了宇宙。最初，宇宙比针头还小得多，在时间开始时这种质量-能量混合物的紧密限制下，这里的密度稍高一些，那里稍薄一些。当这个婴儿宇宙膨胀创造出空间本身时，其量子印记被放大到今天观察到的星系的巨大且不均匀分布。

宇宙具有不可简化的随机特征。如果它是一个钟表机械，其齿轮、弹簧和杠杆不是瑞士制造的；它们不遵循预定的路径。物理决定论已被概率的决定论所取代。没有什么再是注定的了。量子力学定律决定不同未来发生的概率，但不决定哪一个会发生。

但是等等——我听到了一个严肃的反对意见。毫无疑问，人类体验的宏观世界建立在微观的量子世界之上。但这并不意味着汽车等日常物品会继承量子力学的所有奇异属性。当我停好我的Mini时，它相对于路面的速度为零。因为它比电子重得多，与其位置相关的模糊性实际上为零。假设我没有忘记停车的地方，车也没有被拖走或偷走，我会在我离开它的正确位置找到它。在我们生活的世界中，物体在短时间内表现得相当可靠。只是在较长的时间里，不可预测性才会悄然出现。

汽车具有相对简单的内部结构。蜜蜂、小狗和男孩的大脑要复杂得多，构成它们的组成部分具有噪声特性。随机性在它们的神经系统中随处可见，从感知光线和气味的感觉神经元到控制身体肌肉的运动神经元。

考虑一下我在第5章中写到的概念神经元之一。每当患者看到Jennifer Aniston的照片时，概念神经元就会变得兴奋，并在半秒钟内发射大约五个动作电位。然而，动作电位的确切数量在每次观看时都会变化——在一次试验中是六个脉冲，下一次是三个。这种变化性的部分原因是由于眼球颤动、心跳、呼吸等等。其余的不可预测性被认为来自水分子和其他分子的持续抖动，即我们称为温度的热运动，这完全受经典物理学支配。

生物物理学家——研究蛋白质和双分子膜层面细胞结构的专家——大体上看不到量子波动在神经元生命中起关键作用的证据。神经系统和其他任何事物一样，遵循量子力学定律；然而，所有这些分子疯狂运动的集体效应是平滑掉任何量子不确定性，这种效应被称为退相干(decoherence)。退相干意味着生命分子可以使用完全经典的、确定性的定律而不是量子力学的、概率性定律来处理。如果是这样，观察到的行为不确定性——预测蜜蜂、小狗和男孩狂野行为的实际不可能性——是由于我们能够多准确地跟踪事件过程的经典限制，这些限制是完全可以理解的。然而，我们不能排除量子不确定性同样导致行为不确定性的可能性。这种随机性可能发挥功能性作用。任何偶尔能够以不可预测方式行动的生物体，都比行动完全可预见的生物更可能找到猎物或逃脱捕食者。如果一只被捕食者追赶的家蝇在飞行中突然急转弯，它比更可预测的同伴更可能看到新一天的曙光。因此，进化可能偏好利用量子随机性进行某些行为或决策的回路。大脑深处的随机量子波动，其后果被确定性混沌放大，可能导致可测量的结果。蜜蜂、小狗和男孩做事不规律，没有任何明显的原因。如果你和它们生活在一起，你就知道这一点。量子力学和确定性混沌都导致不可预测的结果。

这种自发行为的一个令人寒心的例子记录在Truman Capote的《冷血》(In Cold Blood)中，这是一个真实犯罪记录，讲述两名前罪犯一夜闯入农夫家中抢劫，无理由屠杀了农夫、他的妻子和两个孩子。残忍杀害整个家庭的决定不是预谋的，而是当场做出的，没有任何令人信服的理由。就这样。罪犯本可以轻易逃离房屋而不犯下这一暴行，他们后来为此被绞死。人生中有多少关键决策是不请自来的、不经思考的、无法解释的行为，由投掷（量子）硬币这种比喻性的方式决定？

不确定性具有深远的后果。它意味着人类的行为无法预言。尽管宇宙和其中的一切都遵循自然法则，但未来世界的状态总是模糊的，我们试图向前看得越远，不确定性就越大。

就我个人而言，我觉得确定论令人憎恶。你此时此刻阅读我书籍这一想法内在于大爆炸中，这个观念让我感到完全无助。（当然，我在这个问题上的个人感受与世界的实际情况无关。）

尽管不确定性对于我是否能产生影响、是否能开始自己的因果链条没有什么可说的，但它至少确保了宇宙以不可预测的方式展开。

## 心灵在实现一个量子事件而非另一个量子事件方面的贫乏自由

罗马诗人Lucretius在《物性论》(De rerum natura)中假设了他著名的”偏转”——原子的随机急动运动，用他的话说，是为了保证”意志摆脱命运的束缚”。然而，不确定性并没有为真正的自由意志论者提供慰藉；它不能替代自由意志。我的行为当然应该是因为我想要它们发生而引起的，而不是偶然发生的。用随机性的模糊性来交换确定论的确定性，这并不是Descartes心中所想的。心灵的自由意志论概念要求心灵控制大脑，而不是大脑任意决定。

有一种经常被援引的解释，试图说明这种情况是如何产生的，这种解释可以追溯到量子力学的创立时期。它假设意识与众多量子事件中实际发生的那一个事件之间存在密切联系。这种观点认为，需要一个有感知能力的人类观察者（是否猴子也能胜任从未被考虑过）才能使量子力学所处理的概率坍缩为一个或另一个实际结果。这就是臭名昭著的测量过程，它引发了大量的文献讨论。

最近，争论的焦点集中在量子纠缠上，这是一个得到充分验证的观察结果，即某些经过精心准备的量子系统无论相距多远，都会保持神秘的联系。纠缠的量子系统，比如两个自旋相反、彼此远离的电子，或者两个偏振光子，无论它们相距多远都会始终保持关联（只要它们在此期间不与其他任何东西相互作用）。一旦测量了其中一个电子的自旋，另一个电子的自旋就会瞬间确定，即使它可能相距一光年。这很奇怪，但却是真实的。物理学家Roger Penrose、麻醉学家Stuart Hameroff和其他人推测，这种超世俗的非定域性与意识密切相关。佛教的某些流派，一个更古老的传统，同样认为客体和主体不可分割地联系在一起，意识是物理宇宙的基本特征。

但是，在生物系统中是否有这种量子力学效应的证据呢？直到最近，答案还是响亮的”否”。2010年，著名期刊《自然》发表了在室温下光合成蛋白质内量子力学电子相干性的测量结果。这种效应延伸超过五十亿分之一米，使得光合作用在将阳光转化为有用能量方面异常高效。相干性在概率中显现出来，这些概率决定了被捕获光子的能量如何在从一个分子移动到下一个分子时发生转化，遵循的是量子力学而非经典力学定律。相干性是否在大脑本身的核心操作中起作用还有待观察。至少目前还没有证据表明神经系统的任何分子组成部分——一个温暖潮湿、与环境强耦合的组织——显示出量子纠缠。

总的来说，生物物理学阻碍了稳定的量子纠缠。神经元信息处理的基础是两种操作：信息在突触处从一个神经元化学传递到另一个神经元，以及动作电位的产生。每种操作都会破坏相干量子态，因为这两种操作要么涉及数百个神经递质分子在突触间隙中扩散，要么涉及数百个沿着神经元膜分布的离子蛋白通道改变其构型。发射动作电位的神经元只能接收和发送经典信息，而不是量子信息；也就是说，在每个时刻，神经元要么产生一个二进制脉冲，要么不产生。因此，神经元从不处于叠加态：它不会同时既发射又不发射动作电位。

哲学家Karl Popper和神经生理学家John Eccles是两位现代的灵魂捍卫者。Popper是著名的科学和政治哲学家，Eccles是发现突触传递全或无特性的先驱，他因此于1963年获得诺贝尔奖。因此，他们不是那种对薛定谔的猫、纠缠和万物互联系胡言乱语的典型怪人。

根据Popper和Eccles的观点，有意识的心灵通过操纵与运动规划相关的皮质区域中神经元相互交流的方式，将其意志施加于大脑。通过促进这里的神经元之间的突触交通，并阻止那里的突触交通，心灵将其意志施加于物质世界。对于强烈相信意志力的人来说，Popper-Eccles理论很有吸引力，因为它似乎调和了宗教观点与科学立场。

但这个提议在物理学基础上是否合理呢？

不——如果它要求心灵强迫大脑执行某些物理行动的话。就像一个吵闹鬼，心灵必须撼动和拖拽突触。这是工作，工作需要消耗能量。即使是调节突触传递所需的微量能量消耗也必须显示在自然的平衡表上。物理学不允许任何例

外。能量守恒原理经过一次又一次的检验，总是胜出者。

如果心灵真的是短暂的、不可言喻的，像鬼魂或精神一样，它就无法与物理宇宙相互作用。它无法被看见、听见或感受到。它当然也无法让你的大脑做任何事情。

自由主义式自由选择的唯一真正可能性是让心灵实现一个而非另一个量子力学事件，正如薛定谔方程所规定的那样。假设在特定的时间点和特定的突触处，发生了两种量子力学态的叠加。突触有15%的概率会切换，发送化学信号穿过分离两个神经元的突触间隙，85%的概率不会切换。然而，这种概率计算不足以确定下次动作电位到达突触时会发生什么。所能说的只是可能不会发生释放。（神经科学家仍然不清楚这种极低的突触切换概率是神经系统的特征还是缺陷；也就是说，它是否起到某种功能作用，还是将大约十亿个突触装入一立方毫米皮质组织的不良后果？）

根据我们目前对量子力学的理解，Popper-Eccles心智可以利用这种特殊的自由。心智无法改变概率，但它可以决定任何一次试验中会发生什么。心智的行为将始终保持隐蔽，秘密进行，因为如果我们考虑许多次试验，不会发生任何异常：只有自然法则所预期的结果。意识意志将在物理学的束缚下在世界中行动。它将与偶然性无法区分。

如果这些推测是正确的，这将是意识心智所能获得的最大自由。在平衡于刀刃边缘的选择中，朝任何一个方向的微小推动都可能产生影响。但如果一个结果比另一个结果更有可能发生，那么意识心智的低语就太微不足道，无法对抗这种几率（假设从能量观点来看，不太可能的结果不太受青睐）。这是微薄、贫瘠的自由，因为心智的影响只有在结果大致同等可能时才有效。

外行人和神秘主义者都过分偏爱这样的假设：量子力学的奇异性必须以某种方式对意识负责。除了将宇宙奥秘的数量从两个减少到一个之外，完全不清楚能获得什么。即使我们接受纠缠以某种方式对意识至关重要，纠缠如何解释心身问题的任何一个具体方面？它如何解释可兴奋脑物质向现象体验的转化？

## 意志作为行动的事后想法

---

让我回到坚实的基础上，告诉你一个经典实验，它说服了许多人相信自由意志是虚幻的。这个实验由Benjamin Libet构思并执行，他是加州大学旧金山分校的神经心理学家，实验进行于1980年代初。

大脑和海洋有一个共同点——两者都在不断运动。可视化这一点的一种方法是记录头皮外侧微小的电位波动，大小约为几百万分之一伏特，使用脑电图(EEG)。就像地震仪的记录一样，EEG轨迹狂热地上下移动，记录着下方大脑皮层中看不见的震动。每当被测试的人即将移动肢体时，一个缓慢上升的电位就会积聚。被称为预备电位(readiness potential)，它比实际运动开始提前多达一秒。还有其他EEG特征，但我将通过专注于这一个来保持简单。

直觉上，导致自主行为的事件序列必须如下：你决定举起手；你的大脑将这个意图传达给负责计划和执行手部运动的神经元（它们电活动的微弱回声就是预备电位）；这些神经元将适当的指令传递给脊髓中收缩手臂肌肉的运动神经元。心智选择，大脑行动。当我内省时，这很有道理。我的心智决定去跑步，我的大脑发出适当的指令，然后我寻找我的运动鞋。但Libet并不确信。难道心智和大脑同时行动，甚至大脑在心智之前行动不是更有可能吗？

Libet着手确定心理事件（一个人的深思熟虑的决定）的时机，并将其与物理事件（预备电位的开始）的时机进行比较。多么令人欣慰——经过几千年令人厌倦的哲学辩论，终于有了一个可以用一种或另一种方式解决的问题。实验的棘手部分是确定心理行为发生的确切时刻。我挑战你推断出你第一次感到举起手的冲动的确切时刻；这不容易。

为了帮助他的受试者，Libet将一个明亮光点投射到老式绿色示波器屏幕上。光点绕圈移动，就像时钟分针的尖端。坐在椅子上，头上戴着EEG电极，每个志愿者必须自发但有意地弯曲他们的手腕。他们在注意到想要行动的愿望或冲动时记录光点的位置。为了确保志愿者对神经事件的主观时机判断是准确的，Libet让他们在单独的实验中记录他们的手腕实际开始弯曲的时间，这个时间点可以通过记录他们的肌肉活动来客观确认。受试者可以做得相当好，比实际开始时间仅提前八十毫秒。

结果讲述了一个明确的故事。预备电位的开始先于有意识的运动决定至少半秒，通常要长得多。大脑在心智决定之前就行动了！这完全颠覆了心理因果关系的根深蒂固的直觉——大脑和身体只有在心智意志之后才行动。这就是为什么这个实验过去是、现在仍然是有争议的。但它在这些年里被重复和完善——该实验的脑成像版本最近出现在新闻中——其基本结论依然成立。

在大脑深处的某个地方，可能是在basal ganglia中，一些钙离子聚集在突触前膜附近，一个单一的突触囊泡被释放，越过了阈值，一个action potential诞生了。这个孤立的脉冲级联成一股spike洪流，侵入已准备就绪、随时准备行动的 premotor cortex。在收到这个前进信号后，premotor cortex通知motor cortex，其锥体细胞将详细指令发送到脊髓和肌肉。所有这一切都是在precognitive状态下发生的。然后，一个调节agency感的皮层结构开始工作。它产生了“我刚刚决定移动”的意识感受。肌肉运动的时间和意愿它的感觉或多或少是同时的，但实际的移动决定发生得更早，在意识之前。

## Agency, 或意志的意识体验

---

为什么不立即重复这个实验，不用EEG电极的帮助。去吧，弯曲你的手腕。你体验到与你最初移动计划、意愿运动和实际运动相关的三种相关感受。每一种都有其独特的主观标签。首先是移动的*intention*。一旦你移动了手，你会感到*ownership*——是你的手移动了——和*agency*。你决定移动它。如果朋友拿着你的手并弯曲它，你会感到你的手以特定方式转动(*ownership*)，但你不会体验到*intention*。你也不会感到对手腕运动负责。如果你反射性地用手推桌子起身，你会感到*agency*但几乎没有或没有*intention*。

这是free will辩论中一个被忽视的观点——mind-body nexus创造了一种特定的、意识的voluntary movement感觉，一种引人注目的“我意愿了这个”或“我是这个行动的作者”体验。像其他主观体验一样，这种意愿感具有特定的现象学内容。它有一个*quale*，在本质上与品尝苦杏仁的*quale*没有不同。

就Libet的实验而言，你的大脑决定现在是弯曲手腕的好时机，readiness potential开始积累。稍后，*agency*的neural correlate变得活跃。正是对这个percept，你错误地归因了因果关系。由于这些事件在闪电般的瞬间发生，不到一秒钟，很难捕捉到它们。

*agency*感对实际决定的责任不比雷声对闪电的责任更大。雷声和闪电显然有潜在的因果关系——雨云与地面之间电荷的积累导致电荷均衡，触发声波冲击波——然而我们现代人不会混淆二者。想象你是一个Cro-Magnon人，当闪电击中附近的树时。你几乎被尖锐的雷声震聋了。你闻到臭氧和燃烧木材的味道。难道你不会，非常合理地，将这次撞击归因于雷声(“众神愤怒了”)吗？

但即使你意愿行动的感觉实际上并没有引起它，也不要忘记，采取行动的仍然是你的大脑，而不是别人的。只是不是你的conscious mind这样做的。

这个结论是否仅在Libet实验室的狭窄范围内成立？毕竟，志愿者拥有的唯一自由是决定何时移动他们的手腕，或者——在基本实验的变体中——是否移动左手腕或右手腕。这就像从两个相同的可乐罐中选择一个——谁在乎你选择哪一个？那些更重要的行为，那些涉及长期和深思熟虑的推理的行为呢？你应该养狗吗？你应该娶她吗？所有这些关键决定是否也伴随着先于选择的readiness potential？现在，我们不知道。

所有感官都可能被愚弄。科学家和艺术家称这样的错误为*illusions*。*agency*感也不能完美工作。它也会犯错误。因此，并非你所有的行为都伴随着*agency*。熟练的zombie actions——你的手指在键盘上敲击——唤起微弱的意志体验或根本没有。其他行动需要很多。你需要强烈地发挥你的意志，像内在的肌肉一样，克服爬过暴露crux的恐惧。但一旦越过这个部分，你的身体管理得很好，不需要任何进一步的conscious意志发挥。

在automatism中，*agency*感可能完全缺失。例子包括宗教仪式中的possession和trance、posthypnotic suggestion、Ouija board游戏、dowsing和其他pseudo-occult现象。参与者强烈否认他们导致了这些事情的发生。相反，他们将责任投射到遥远的神灵、精灵或催眠师身上。

在你的生活中，远离occult实践，你发现自己在做事情而没有处于行动中。当人们对他们想要什么深度矛盾时，这种情况尤其明显。强迫性赌徒突然发现自己在赌场，相信今晚他会赢大钱，尽管他在某种程度上“知道”，到夜晚结束时，他将失去一切。强大的psychodynamic力量在起作用，减少了责任感。

精神疾病可能导致明显的病理，阻碍意志体验。这个spectrum包括临床肥胖者，他们无法停止supersizing他们的食物；drug addicts，他们转向卖淫和犯罪来资助他们的习惯；患有Tourette’s syndrome的个体，他们的身体定期爆发出

疯狂的tics、jerks和grimaces的暴动；以及obsessive-compulsives，他们洗手如此频繁以至于变得粗糙和流血，或者被迫在去洗手间时执行奇怪的仪式。患者知道他们所做的是dysfunctional的，是“疯狂的”，但他们无法阻止自己。显然，他们并不总是自己行为的主人。

感染了弓形虫 *Toxoplasma gondii* 这种原虫的啮齿动物也不例外。健康的老鼠会小心避开有猫尿味的地方，但被这种单细胞寄生虫感染的老鼠却失去了对猫科动物气味的天然厌恶，甚至可能被这种气味吸引。失去对猫的恐惧对生病的老鼠来说是不幸的转折，因为它现在更容易被猫杀死。但对原虫来说这是一笔好交易。当猫吞食老鼠时，这个讨厌的搭便车者就进入了新宿主体内，完成了它的生命周期（它们只能在猫的肠道内进行有性繁殖）。这种行为操控非常特异：生病的老鼠总体上并不比健康老鼠更少焦虑，也不会失去对它们与疼痛足部电击相关联的音调的恐惧。*T. gondii*专门瞄准大脑中支撑特定恐惧的部分——它在杏仁核中的囊肿密度几乎是其他参与气味感知脑结构中密度的两倍。

将这个野外生活的小插曲提升到史诗级高度的是，10%的美国人口感染了 *T. gondii*。这被称为弓形虫病 (toxoplasmosis)。科学家早就知道精神分裂症患者更容易携带 *T. gondii* 抗体，甚至有人声称这种常见寄生虫在文化习惯进化中发挥作用。然而感染者大概觉得可以自由地做他们想做的事。就像在好莱坞恐怖片中一样，他们可能会执行这些脑寄生虫的无声命令。

Daniel Wegner是哈佛大学的心理学家，是现代意志研究的先驱之一。在他引人入胜的专著《意识意志的幻觉》(*The Illusion of Conscious Will*)中，他探索了agency sensation的本质以及它如何被操控。

在一个引人注目的实验中，Wegner让一名志愿者穿上黑色罩衫和白手套，站在镜子前，双臂垂在身体两侧。她身后直接站着一名实验室成员，穿着相同的服装。他将手臂从她腋下伸出，这样当女子看镜子时，他的两只戴手套的手看起来就像是她自己的（男子的头被屏幕遮住了）。两名志愿者都戴着耳机，Wegner通过耳机发出指令，比如“拍手”或“弹左手手指”。志愿者应该倾听并报告实验室成员手部动作在多大程度上是她自己的。当女子在男子手部执行动作之前听到Wegner的指令时，相比于Wegner的指令在男子已经移动手部之后才发出时，她报告感到更强烈的自己意志了该动作的感觉。当两人都被要求拍手三次时，相比于她没听到指令而看到手部拍手时，女子感到更强的引起手部鼓掌的感觉。记住，女子从未移动过她的手——移动的始终是她身后男子的手。

agency感是由大脑模块产生的，该模块基于简单规则将authorship分配给自主动作。如果你计划弹手指，然后你低头看到它们在这样做，agency模块就得出结论你启动了该动作。另一套规则涉及时机。想象独自走过黑暗森林时听到树枝断裂的声音。如果声音恰好在你踩到树枝之后发出，你会松一口气，因为你的agency模块得出结论是你发出了声音，一切安好。但如果尖锐的破裂声在你踩到树枝之前发生，可能有什么东西或某人在追赶你，你的所有感官都会高度警觉。

这些意图和agency感觉的真实性被神经外科医生强调了，他们有时必须切除脑组织，要么因为它是肿瘤性的，要么因为它可能剧烈放电，引发大发作(grand mal)癫痫发作。切除或烧灼多少组织是一种平衡行为，既要避免留下癌性或易发物质残留的Scylla，又要避免切除对言语或其他重要行为至关重要区域的Charybdis。为确定要切除多少物质，脑外科医生探测附近组织的功能。他通过用短脉冲电流刺激组织来做到这一点，同时患者用拇指依次触摸每根手指、数数或做一些其他简单任务。

在这样的探索过程中，Itzhak Fried——我在第5章首次提到的外科医生——刺激了前辅助运动区(presupplementary motor area)，这是位于初级运动皮层前方的大脑皮层广阔区域的一部分。他发现这种刺激可以引发移动肢体的冲动。患者报告感到需要移动腿部、肘部或手臂。法国Bron认知科学研究所(Institut des Sciences Cognitives)的Michel Desmurget和Angela Sirigu在刺激后顶叶皮层时发现了类似的情况，该区域负责将视觉信息转化为运动命令。激活这个灰质产生了纯粹意图的感觉。患者评论道：“感觉像我想移动我的脚。不确定如何解释，”“我有移动右手的欲

望，”或”我有在嘴里卷舌头的欲望。“值得注意的是，他们从未对电极诱发的这些特定冲动采取行动。他们的感觉明显来自内部，没有检查者的任何提示。

## 评估现状

---

让我总结一下。经典决定论已经过时：未来不是由当前事实完全决定的。毫无疑问，量子力学随机性是物质基本结构的固有特征。哪个未来会发生并不是完全确定的。你的行动不是注定的。Omar Khayyam诗句中的无奈感叹

运动的手指在书写；写完后，

继续前行：无论多少虔诚与智慧

也不能诱使它回头取消半行字，

所有眼泪也洗不掉其中一个词

这并不适用于未来。你展开的人生是一本未写完的书。你的命运掌握在你的手中，也掌握在宇宙其余部分好管闲事的手中。大脑的复杂特性和确定性混沌限制了即使是未来最博学的科学家也无法准确预测行为。有些行为将永远显得自发、无法解释。量子力学不确定性(quantum mechanical indeterminacy)在其生成中发挥何种作用仍然未知。

强烈的、笛卡尔式的自由意志观念——相信如果你再次被置于完全相同的环境中，包括与之前完全相同的大脑状态，你本可以通过意志让自己采取不同的行动——这种观念无法与自然法则调和。意识心智，这个灵魂的避难所，不可能在不留下明显痕迹的情况下影响大脑。物理学不允许这样的幽灵般的相互作用。世界上发生的任何事情都有一个或多个也属于世界的原因；宇宙在因果关系上是封闭的。

至少在实验室中，大脑的决定远早于心智的决定；意志支配简单行为的意识体验——能动性(agency)或作者身份的感觉——是次于实际原因的。能动性具有现象学内容，或感质(qualia)，就像意识体验的感官形式一样，由皮质-丘脑回路触发。心理学实验、精神病患者和神经外科干预揭示了自主行动这一方面的现实。决定如何形成仍然是无意识的。你为什么以那种方式选择在很大程度上对你来说是不透明的。

我从这些洞察中汲取了两个教训。首先，我采用了一种更加实用的、相容主义的自由意志概念。我努力生活得尽可能摆脱内在和外在的约束。唯一的例外应该是我故意和有意识地强加给自己的限制，其中最主要的是出于伦理考虑的约束；无论你做什么，都不要伤害他人，并努力让这个星球变得比你发现它时更美好。其他考虑因素包括家庭生活、健康、财务稳定和正念。其次，我试图更好地理解我的无意识动机、欲望和恐惧。我比年轻时的自己更深入地反思自己的行为和情感。

我在这里并没有开创新的领域——这些是智者几千年来一直在教授的教训。古希腊人在德尔斐阿波罗神庙的入口上方刻着”gnothi seauton”(认识你自己)，拉丁文版本装饰着《黑客帝国》中先知厨房的墙壁。耶稣会士有一个近500年的精神传统，强调每日两次的良心省察。这是一种自我意识的练习：持续的内在质疑会提高你对自己行为、欲望和动机的敏感性。你真诚地试图识别自己的错误并努力消除它们。你寻求将无意识的动机带入意识中。这不仅能让你更好地了解自己，还能让你过上更符合自己性格和长期目标的生活。

仍然未解决的是能动性的现象学感受如何从神经活动中产生。又是意识！我们回到了心身问题的内圣殿。我在下一章中概述了这个难题的信息理论解决方案，这是本书中最具推测性的一章。

## 第8章：我论证意识是复杂事物的基本属性，热情赞美整合信息理论，它如何解释关于意识的许多令人困惑的事实，并提供构建有感知能力机器的蓝图

---

哲学写在这本宏伟的书中——我说的是宇宙——它在我们眼前敞开着。但除非我们首先学会理解其中的语言，并知道其中的字符，否则无法理解这本书。它是用数学语言写成的。

——伽利略·伽利雷，《试金者》(1623)

对意识物理基础的追求是我智力生活的焦点——在过去二十多年里一直如此。

弗朗西斯·克里克(Francis Crick)和我花了好几天坐在他工作室的藤椅上，辩论活体物质如何产生主观感受。我们写了两本书和二十多篇学术文章，阐述了将意识的不同方面与特定大脑机制和区域联系起来的必要性。我们假设了觉知与皮质神经元有节律放电之间的关系，每二十到三十毫秒放电一次。(我们所谓的40赫兹假说目前在选择性注意的背景下正在复兴。)我们痴迷于神经元同步发放动作电位的需要。我们为第5层新皮质锥体细胞作为意识内容传达者的关键作用进行辩护。我们推理大脑皮层下面一层神经元的神秘薄片，被称为屏状核(claustrum)，对于跨越视觉和听觉或视觉和触觉的知觉觉知是必不可少的。我读了无数通常是可有可无的手稿和书籍；参加(偶尔在其中打盹)了数百场研讨会。我与学者、朋友和来自各行各业的人们讨论过意识和大脑。我甚至在《花花公子》杂志上发表过一封关于这个话题的信。

对我来说越来越清楚的是，无论关键的神经回路是什么，它们的识别将引发一个我在1992年首次遇到的根本问题。那是我游学漫游的早期，被传播好消息的冲动驱动着，从现在开始，意识将完全落在经验的领域内，它将能够接受科学分析。

在一次这样的研讨会后，已故的苏黎世神经学家Volker Henn提出了一个简单的问题：假设Crick和你的所有想法都得到证实，视觉皮层第5层皮质神经元有节律地放电并将其输出发送到大脑前部，这些就是意识的关键神经关联物。那么这些细胞的什么特性产生了觉知？从原则上讲，你的假设与笛卡尔关于松果体是灵魂所在地的提议有何不同？声称以节律方式放电的神经元产生看到红色的感觉，这与假设松果体中动物精神的扰动引起灵魂的激情一样神秘。你的语言比笛卡尔的更机械化——毕竟，三个半世纪已经过去了——但基本的困境仍然和以往一样尖锐。在这两种情况下，我们都必须凭信念接受某种类型的物理活动能够产生现象性感受。

我用一张期票回应了Henn：随着时间的推移，科学将回答这个问题，但现在，神经科学应该继续推进，寻找意识的关联物。否则，对意识根本原因的探索将被不必要地延迟。

Henn的问题可以推广。全局可用性、奇异环、吸引子网络、这种神经递质或那个大脑区域都被提名为意识的本质。更非常规的提议涉及量子力学纠缠或其他奇异物理学。但无论什么特征被证明是关键的，这些特定特征的什么解释了主观性？Francis和我曾考虑过意识必须涉及皮层内的反馈回路的想法，但反馈的什么产生了现象学、产生了感受？房间恒温器也有反馈：当环境空气温度达到预定值时，冷却就会关闭。它有一点意识吗？这与相信摩擦黄铜灯会让精灵出现有什么根本不同？

多年来，我将Henn的问题推到一边，认为它没有成效。我想推进意识项目。我想说服分子生物学家和神经科学家加入进来，将他们不断增长的工具集用于干预心智的关键回路。

然而Henn的挑战必须得到回答。我探索的终点必须是一个理论，解释物理世界如何以及为什么能够产生现象性体验。这样的理论不能只是模糊、空洞的，而必须是具体的、可量化的、可测试的。

我相信信息论，如果得到适当的表述和完善，能够完成如此巨大的壮举，分析任何生物的神经元连线图并预测该生物体将体验的意识形式。它可以为有意识人工制品的设计绘制蓝图。而且，令人惊讶的是，它提供了宇宙中意识进化的宏伟观点。

这些都是大胆、雄心勃勃和夸张的声明。请耐心听我为它们辩护。

## 狗，或者意识是从大脑中涌现出来的吗？

---

如果你与狗共同生活——我对这些生物怀有极大的情感——你会知道它们不仅在意想不到的方面聪明，还表现出各种情绪。我的黑色德国牧羊犬Nosy第一次在洛矶山脉高处遇到雪时，她把鼻子伸进这种奇怪的白色物质中，把一些抛向空中，然后再接住它。她咬了咬冰冷的表层，对雪堆吠叫，最后她仰面躺下，来回滑动，将冰冷的晶体摩擦进她的毛皮。她是快乐的活体体现！当一只小狗加入我们家庭并且生活围绕着这个新家庭成员转动时，Nosy沉郁了几个星期；捡网球时兴奋；当另一只狗挑战她时变得好斗；犯了错误时羞愧，夹着尾巴；在烟花时恐惧并需要Prozac；当我整天工作忽略她时感到无聊；一有车进入车道就警觉；当她在做饭时等食物掉落而其中一个孩子戳她时感到恼怒；当我们购物回来时好奇，把鼻子伸进每个袋子检查里面的内容。

作为群居动物，狗进化出了广泛的精巧交流技能。Charles Darwin这位敏锐的动物行为观察者，同时也是狗的爱好者，在《人类的由来》中写道关于犬类声音的以下内容：

我们有急切的吠叫，如在追逐中；愤怒的吠叫，以及咆哮；绝望的呜咽或嚎叫，如被关起来时；夜晚的吠叫；快乐的吠叫，如与主人开始散步时；以及非常明显的要求或恳求的叫声，如希望门或窗被打开时。

狗的尾巴、鼻子、爪子、身体、耳朵和舌头表达它的内在状态，它的感受。狗不会——它们不能——伪装。

这种丰富的行为表现以及犬类和人类大脑之间众多的结构和分子相似性让我得出结论：狗有现象性感受。任何否认它们有感知能力的哲学或神学都是严重缺陷的。（我在童年时直觉地感到这一点；我无法理解为什么上帝会在审判日复活人类但不复活狗。这没有任何意义。）对狗来说是真的，对猴子、老鼠、海豚、鱿鱼，以及可能还有蜜蜂来说也是真的。我们都是自然的孩子；我们所有人都体验生活。

尽管这一论点在信奉一神教、拒绝承认动物有灵魂的西方国家中说服力较弱，但东方宗教则更加宽容。印度教、佛教、锡克教和耆那教都承认所有生物都是同类的、有感知能力的存在。美洲原住民也没有人类例外论的信念，而这种信念在犹太教-基督教的世界观中根深蒂固。

实际上，我经常认为狗比人更接近真正的佛性。它们本能地知道生命中什么是重要的。它们无法心怀恶意或怨恨。它们对生存的喜悦、渴望取悦的热切、直到生命终结的简单而纯真的忠诚，这些都是人类只能向往的品质。

狗和人类在数万年前的热带草原、草原和森林中缔结了盟约，当时狼和人开始在近距离生活。这种有利的关系，两个物种在其中共同进化并相互驯化，一直延续至今。

然而，毫无疑问，犬类意识的范围和深度不如我们的。狗不会反思自己，也不会担心为什么它们的尾巴摆动方式很奇怪。它们的自我意识是有限的。它们没有受到亚当诅咒——对自己死亡的认知——的折磨。它们不会分享人类的愚蠢行为，从存在主义恐惧到大屠杀和自杀式炸弹袭击。

考虑更简单的动物——简单性以神经元数量及其相互连接来衡量——比如小鼠、鲱鱼或苍蝇。它们的行为分化程度较低，比狗的行为更加刻板化。因此，假设这些动物的意识状态不如犬类意识那样丰富，其中包含的联想和意义要少得多，这并非不合理。

基于这样的推理，学者们认为意识是大脑的一种涌现特性(emergent property)。这种观点在生物学家们中广泛共享。这究竟是什么意思？涌现特性是指整体所表现出的、但其个别部分不一定具备的属性。系统拥有其组成部分所不具备的特性。

涌现不必具有神秘的、新时代的含义。考虑水的湿润性，即它保持与表面接触的能力。这是分子间相互作用的结果，特别是相邻水分子之间氢键的结果。一两个H<sub>2</sub>O分子并不湿润。但在适当的温度和压力下将数以亿计的分子聚集在一起，湿润性就会涌现。遗传定律从DNA和其他大分子的分子特性中涌现出来。当太多汽车在过于狭窄的空间内朝不同方向行驶时，交通堵塞就会涌现。你明白这个道理。

以这种尚待定义的方式，意识不会在少数几个神经元连接在一起时显现；它从大型细胞网络中涌现出来。这些组合越大，网络可用的意识状态repertoire就越大。

理解意识的物质基础需要深刻理解这些由数百万异质神经细胞紧密连接的网络是如何编织我们精神生活的挂毯的。要想象大脑惊人的复杂性，回想一下那些自然纪录片，其中单螺旋桨飞机通过在丛林上空飞行数小时来捕捉亚马逊的广袤。这片雨林中的树木数量大约相当于你大脑中的神经元数量（如果森林继续以目前的速度被砍伐，这种说法在几年内将不再成立）。这些树木的形态多样性，它们独特的根部、枝干和覆盖着藤蔓和爬行动物的叶子，与神经细胞的多样性也相当。想想这一点。你的大脑就像整个亚马逊雨林。

神经元联盟从与环境和彼此的相互作用中学习的能力经常被低估。单个神经元是深奥复杂的信息处理器；每个神经元树突（处理突触输入）和轴突（分发其输出）的配置都是独特的。反过来，突触是纳米机器，配备了修改这些神经连接强度和动力学的学习算法，时间尺度从几秒到一生不等。人类对如此庞大、复杂和自适应的网络缺乏直接经验。

理解意识如何从大脑中涌现的概念困难，在历史上与十九世纪和二十世纪初关于活力论(vitalism)和遗传机制的争论相似。遗传背后的化学定律深深令人困惑。指定独特个体的所有信息是如何储存在细胞中的？这些信息是如何被复制并传递给细胞后代的？当时已知的简单分子如何使卵子发育成成体？

这种困惑由英国顶尖遗传学家威廉·贝特森(William Bateson)在1916年很好地表达了：

生物的特性在某种程度上附着于物质基础，也许在某种特殊程度上附着于核染色质；然而，染色质颗粒或任何其他物质的颗粒，无论多么复杂，都不可能拥有必须分配给我们的因子或基因[原文如此]的那些能力。假设染色质颗粒彼此无法区分，实际上在任何已知测试下几乎是均质的，却能通过其物质性质赋予生命的所有特性，这超出了即使最坚定的唯物主义的范围。

为了解释生命，学者们求助于神秘的vitalistic力量，亚里士多德的entelechy，叔本华的phenomenal will，或柏格森的élan vital。其他人，包括以其同名方程而闻名的物理学家薛定谔，假设了新的物理定律。化学家无法想象，串状分子中四种核苷酸类型的精确序列竟然掌握着关键。遗传学家低估了大分子储存大量信息的能力。他们未能理解经过数十亿年自然选择作用塑造的蛋白质所具有的惊人特异性。但这个特殊的谜题最终得到了解决。我们现在知道生命是一种涌现现象，最终可以归结为化学和物理学。没有vitalistic力量或能量将无机的死亡世界与有机的生命世界分开。

这种缺乏明确分界线的情况对于涌现来说是典型的。像H<sub>2</sub>O这样的简单分子显然没有生命，而细菌则是有生命的。但是引起疯牛病的朊病毒蛋白呢？病毒呢？它们是死的还是活的？

如果意识是一种涌现现象，最终可归结为神经细胞的相互作用，那么一些动物会有意识，而其他动物则没有。微小的大脑——想想著名的线虫*Caenorhabditis elegans*，不比字母I大，其大脑恰好有302个神经元——可能没有心智。大脑袋——人类的160亿神经元——有心智。这种涌现与物理思维的基本原则相矛盾——*ex nihilo nihil fit*，或者无中不能生有。这是一种原始守恒定律的形式。如果一开始什么都没有，再增加一点点也不会有什么不同。

我曾经是意识从复杂神经网络中涌现这一观念的支持者。只需读读我早期的Quest就知道。但多年来，我的想法发生了改变。主观性与任何物理现象都有根本性的不同，不可能是一种涌现现象。某种蓝色与眼睛锥体光感受器的电活动根本不同，尽管我完全认识到后者对于前者是必要的。一个是我大脑固有的，无法从外部推断，而另一个具有可

被外部观察者获取的客观属性。现象界来自与物理界不同的王国，服从不同的定律。我看不出无意识生物与有意识生物之间的鸿沟如何能够通过更多神经元来弥合。

有一个明确的涌现和还原论的替代方案，对像我这样的隐秘柏拉图主义者来说很有吸引力。莱布尼茨在十八世纪初在他的单子论开篇声明中阐述了这一点：

1. 我们将在此讨论的单子，只不过是进入复合体的简单实体——简单的，也就是说，没有部分的。
2. 必须有简单实体，因为有复合体；因为复合体只不过是简单体的集合或聚合。

这种观点确实带来了许多人不愿意承担的形而上学代价——承认体验，即功能性大脑的内在视角，是与引起它的物质事物根本不同的东西，永远无法完全归结为大脑的物理属性。

## 意识内在于复杂性

---

我相信意识是生命物质的基本属性，是基本属性。它不能从任何其他东西推导出来；用莱布尼茨的话说，它是一种简单实体。

我的推理类似于研究电荷的学者们提出的论证。电荷不是生物的涌现属性，正如最初在青蛙抽搐肌肉中发现电时所认为的那样。没有无电荷粒子在聚合时产生电荷。电子有一个负电荷，质子——氢离子——有一个正电荷。与分子或离子相关的总电荷简单地等于所有单个电子和质子电荷的总和，无论它们彼此的关系如何。就化学和生物学而言，电荷是这些粒子的固有属性。电荷不是从物质中涌现的。

意识也是如此。意识伴随着有组织的物质块。它内在于系统的组织中。它是复杂实体的属性，不能进一步归结为更基本属性的作用。我们已经到达了还原论的底层（这就是为什么本书副标题中的还原论者被浪漫主义者所调和）。

你和我发现自己身处这样一个宇宙中：任何以及所有相互作用部分的系统都具有某种程度的感知力。系统越大、网络化程度越高，意识程度就越高。人类意识比犬类意识更加精细，因为人脑比狗脑多二十倍的神经元，并且具有更密集的网络连接。

注意我省略了什么。我写的是“相互作用部件的系统”而不是“相互作用部件的有机系统”。我没有特别指出生物系统，因为在心理学家和工程师中有一个广泛的信念，称为功能主义。要理解功能主义，请想想数字的乘法或除法。你可以用铅笔在纸上书写，可以在计算尺上滑动两块有刻度的木片，可以在算盘上移动珠子，或者可以按袖珍计算器上的按钮。所有这些方法都实现了相同的代数规则；它们在功能上是等价的。它们在灵活性、优雅性、价格等方面有所不同，但都做同样的事情。对人工智能的追求基于对功能主义的强烈信念——智能可以有非常不同的载体，无论是头骨、外骨骼还是铝制盒子。

应用于意识的功能主义意味着任何内部结构在功能上等同于人脑的系统都拥有相同的心智。如果我大脑中的每一个axon(轴突)、synapse(突触)和神经细胞都被执行完全相同功能的导线、晶体管和电子电路所取代，我的心智将保持不变。我大脑的电子版本可能更笨重、更大，但只要每个神经元组件都有一个忠实的硅制替身，意识就会保持。

对心智来说，重要的不是大脑所由之构成的物质的性质，而是那种物质的组织——系统各部分的连接方式，它们的因果相互作用。更精确的表述是：“意识是substrate-independent(基质独立的)。”功能主义在理解和模仿自然方面为生物学家和工程师提供了很好的服务，那么在涉及意识时为什么要放弃它呢？

## 意识与信息论

---

理解心智-大脑界面的发现——如第4、5、6章所描述的——需要一个大规模、逻辑一致的框架，即意识理论。这样一个理论体系需要将awareness(觉知)与synapses(突触)和neurons(神经元)联系起来——这是意识科学的圣杯。它不能仅仅是描述性的(即，意识涉及大脑的这一部分和那些连接)；它必须是规范性的(即，它必须给出意识发生的必要和充分条件)。这个理论必须基于第一性原理，将现象学经验建立在宇宙的某个基本方面上。这样的理论必须是精确和严格的，不只是形而上学断言的集合。

任何科学理论的一个基本要求是它必须处理可测量的事物。伽利略表达为“测量可测量的，使不可测量的变得可测量。”“意识理论必须量化意识，将neuroanatomy(神经解剖学)和physiology(生理学)的特定方面与qualia(感性质)联系起来，并解释为什么意识在麻醉和睡眠期间会减弱。它必须解释意识对生物体有什么用处，如果有的话。它应该从少数几个公理开始，通过诉诸我们自己的phenomenal(现象学的)、conscious(有意识的)经验来证明它们。这些公理将产生某些后果，这些后果应该能以通常的经验方式得到验证。

除了一个值得注意的例外(我稍后会谈到)，关于意识理论的基础性持续工作很少。有一些模型将心智描述为许多功能盒子，箭头进进出出并连接这些盒子：一个用于早期视觉的盒子，一个用于物体识别的盒子，一个用于工作记忆的盒子，等等。这些盒子与大脑中的特定处理阶段相对应。这种方法的支持者然后指向其中一个盒子并宣称，每当信息进入这个盒子时，它就神奇地被赋予了phenomenal awareness(现象学觉知)。

我也犯过这样的错误。Francis和我声称“在视觉皮层的高级区域和前额叶皮层的规划阶段之间来回穿梭的信息将被有意识地体验到”，就是这种类型。从经验上看，在neocortex(新皮层)的后部和前部之间建立双向对话确实可能产生主观感受，但为什么会这样并不明显。

同一类别中还有cognitive psychologist(认知心理学家) Bernie Baars的*global workspace*(全局工作空间)模型。它的谱系可以追溯到人工智能早期的*blackboard architecture*(黑板架构)，在这种架构中，专门化程序访问一个共享的信息存储库——黑板。Baars假设人类心智中存在这样一个共同的处理资源。写入工作空间的任何数据都可供一系列辅助过程使用——工作记忆、语言、规划模块等等。全局广播信息的行为使我们意识到它。不过，工作空间非常小，所以一次只能表示一个percept(知觉)、思想或记忆。新信息与旧信息竞争并取代它。

*global workspace*模型背后的核心直觉是有效的。有意识的信息可以被整个系统全局访问且是有限的。相比之下，zombie agents(僵尸代理)则把它们的知识留给自己。它们在信息上是encapsulated(封装的)，超出了意识的范围。

著名的分子生物学家让-皮埃尔·尚热 (Jean-Pierre Changeux) 和他的年轻同事、数学家兼认知神经科学家斯坦尼斯拉斯·德阿纳 (Stanislas Dehaene) 在巴黎法兰西学院，将这一模型用神经科学的术语进行了阐述。他们认为前额皮质中的长程锥体神经元实例化了巴尔斯的全局工作空间。德阿纳的研究小组正在领导一项协调努力来阐明这个神经元工作空间，使用创新的心理物理学程序、fMRI扫描和手术患者的脑电图记录。他们的模型很好地捕捉了非意识的局部处理与意识的全局处理和内容可达性之间的突然转换。

描述性模型对于制定可检验假设至关重要。它们推动任何科学的早期阶段。但不应将它们与规范性理论混淆，因为它们未能回答亨恩的问题：为什么大脑皮质前后之间的回响性、整合性神经活动会被有意识地体验？为什么通过长程皮质纤维的扩音器广播信息会产生感觉？这些模型只是断言这就是发生的事情；它们没有解释如何发生。

很长时间以来，弗朗西斯和我都抗拒用数学方法正式描述意识的尝试。散布在知识景观中的许多沉没心身模型的残骸加剧了我们的怀疑，即扶手椅式的理论化，即使有数学和计算机模拟的支持，也无法带来进步。弗朗西斯在分子

生物学方面的经验加强了这种反理论偏见：数学模型——包括他自己利用编码理论的徒劳尝试——在分子生物学的惊人成功中最多只起到了次要作用。这就是为什么弗朗西斯和我在我们的著作和演讲中强调了一个严格的实验计划来发现和探索意识的生物学基础。

在他生命的最后十年里，弗朗西斯总是愿意根据新的证据和思维方式改变自己的观点，他开始认可信息论作为意识理论的合适语言。为什么？嗯，在没有某种特殊物质的情况下，比如笛卡尔的思维物质能够神奇地赋予有机体主观性，意识必须从超连接脑细胞之间的因果互动中产生。在这种情况下，因果意味着神经元A中的活动，直接或间接地影响神经元B在即时或更远的未来发生活动的可能性。

我也想要更一般的东西。我想要回答银河系、蚁丘、蜜蜂或iPhone是否有意识的问题。为此，我需要一个超越宇宙学、行为生物学、神经生物学和电路分析细节的理论。适当地表述，信息论是一种数学形式主义，可以量化任何系统组件的因果互动。它形式化了这里这一部分的状态——一颗恒星、一只蚂蚁、一个神经元或一个晶体管——对那里那一部分的影响程度，以及这种影响如何随时间演化。信息论详尽地编目和表征了任何复合实体所有部分之间的互动。

信息是21世纪的通用通用语。股票和债券价格、书籍、照片、电影、音乐和我们的基因构成都可以转换成无穷无尽的0和1数据流的想法是我们熟悉的。一个简单的电灯开关可以处于两个位置之一，开或关；知道它处于哪个位置或状态对应一比特信息。需要几个比特来指定一个突触对其连接的神经元的影响大小。比特是数据的原子。它们通过以太网电缆或无线传输、存储、重放、复制，并组装成巨大的知识库。这种外在的信息概念，即产生差异的差异，是通信工程师和计算机科学家所熟悉的。

大卫·查尔默斯（David Chalmers）是信息论理解意识潜力的哲学辩护者。他的意识双重属性草图假设信息具有两种不同的、固有的、基本的属性：外在属性和内在属性。信息的隐藏的内在属性就是成为这样一个系统的感觉；某种程度的意识，某种最小的感质（ quale）与作为信息处理系统相关联。这就是宇宙的本来面目。任何具有可区分物理状态的东西——无论是两种状态，如开关开关，还是数十亿种状态，如硬盘或神经系统——都具有主观的、短暂的、有意识的状态。可区分状态的数量越大，意识体验的 repertoire 就越大。

查尔默斯的双重属性理论表述是粗糙的。它只考虑信息的总量，而意识并不随着比特的单纯积累而增加。在什么有意义的方式下，一个具有1千兆字节存储容量的硬盘比一个具有128千兆字节的硬盘感知能力更低？当然，重要的不仅仅是越来越多数据的积累，而是各个数据比特之间的关系。系统的架构，其内部组织，对意识至关重要。但查尔默斯的思考并不关心架构，系统的内部组织。因此它们未能解释为什么大脑的某些区域对意识比其他区域重要得多，无意识和有意识行为之间的差异等等。

在我们不懈的探索中，Francis和我遇到了一个更加精密的双面向理论版本。其核心是由Giulio Tononi提出的整合信息概念，他当时在加利福尼亚拉霍亚的Gerald Edelman的神经科学研究所工作，现在是威斯康星大学麦迪逊分校的教授。Edelman是帮助破译抗体化学结构的免疫学家，这项工作为他赢得了诺贝尔奖。

Giulio安排我们四人在Edelman神经科学研究所美丽的园区里共进午餐。我们在生物学两位老前辈的竞争氛围中相遇。然而，本可能紧张的聚会却变得和谐，美食佳肴和Edelman喜欢讲述的无穷无尽的笑话轶事为这次聚会增色不少。我们年轻人彼此产生了好感，这种友谊随着时间的推移只会更加深厚。

那个下午我们四人相互学习。Francis和我更好地理解了他们对皮质-丘脑复合体广阔领域的整体、全局特性的强调，以及静默的重要性——那些没有演奏的乐器的重要性。（我将在几页后澄清这个神秘的评论。）反过来，他们也开始欣赏我们的坚持，即寻找意识的神经相关物可能取决于神经元及其连接的局部、特定属性。

让我向你介绍Giulio的思想。

## 整合信息理论

---

任何意识状态都蕴含着丰富信息，这是一个普遍的观察结果。事实上，它是如此具体，以至于你永远不会重新体验到完全相同的感受——永远不会！不仅如此，任何意识状态都排除了无数种替代体验。当你在漆黑的房间里睁开眼睛时，你什么也看不见。纯粹的黑暗似乎是你能拥有的最简单的视觉体验。事实上，你可能认为它几乎不传达任何信息。然而，这种漆黑的感知暗示着你没有看到安静、光线充足的客厅，或者优胜美地半穹顶的花岗岩表面，或者过去或未来拍摄的任何电影的任何画面。你的主观体验隐含地排除了所有这些你本可以看到的、想象到的、听到的、闻到的其他事物。这种不确定性的减少（也称为熵）正是信息论之父、电气工程师Claude Shannon定义信息的方式。也就是说：每一个意识体验都蕴含着丰富的信息，都是极其分化的。

意识状态还具有第二个特性：它们是高度整合的。任何意识状态都是一个单子(monad)、一个单位——它不能被细分为独立体验的组件。无论你如何努力，你都无法以黑白的方式看世界；也无法只看到你视野的左半部分（或右半部分）（除非闭上一只眼睛或采取其他类似的技巧）。当我写下这些字句时，我正在聆听神秘主义极简派作曲家Arvo Pärt的《纪念Benjamin Britten的圣歌》中的哀歌。我意识到整个声音景观。我无法不听到管钟声或最后渐入静寂的下降。这是一个单一的感知。

我所意识到的任何信息都完整地呈现在我的心中。意识统一性的基础是我大脑相关部分之间的众多因果相互作用。如果大脑区域变得分裂、断开和分化，就像麻醉时发生的情况一样，意识就会消退。相反，当许多区域同步激活时——在EEG信号的共同起伏中显现，如在深度睡眠中——整合度很高，但传达的特定信息却很少。

Giulio的整合信息理论从这两个公理性前提推断出，任何意识系统都必须是一个单一的、整合的实体，具有大量高度分化状态的储备。这就是他的处方——整合和分化。这构成了他的单子。不多不少。

我那台时尚Apple笔记本电脑的磁盘存储容量远远超过了我记忆事物的能力。然而磁盘上的信息并非整合的。我Mac上的家庭照片彼此之间没有关联。笔记本电脑不知道那些照片中的女孩是我的女儿，从可爱的小女孩成长为瘦长的少女再到优雅的成年人，也不知道我日历中的”Gabi”条目指的是与那些图像中的人的会面。对计算机而言，所有这些数据都是一样的，是零和一组成巨大随机图案。我从这些图像中获得意义，因为我的记忆建立在成千上万的其他事实和记忆的基础上，并与它们交叉关联。它们越是相互关联，就变得越有意义。

Giulio将这些公理正式化为他的整合信息理论的两大支柱。他假设，任何物理系统在特定状态下产生的意识体验的量等于该系统在该状态下产生的整合信息量，超出其各部分产生的信息量。系统必须在大量状态储备中进行区分（分化），并且必须作为统一整体的一部分来做到这一点，这个整体不能被分解为一系列因果独立的部分（整合）。

考虑一个进入特定状态的神经系统，其中一些神经元正在放电，而其他神经元保持沉默。假设在这种状态下，大脑体验到红色。大脑之所以能做到这一点，是因为它能够在其分布广泛的神经元领域之间整合信息，这种信息无法通过将大脑分解为更小的独立组件来产生。如前所述，当连接左右半球的胼胝体完全切断时，大脑半球不再整合。从信息学角度来说，整个大脑的熵现在变成左右半球两个独立熵的总和，而整合信息变为零。大脑作为一个整体不再有conscious experiences。相反，每个半球都有自己整合信息，尽管比整个大脑的整合程度要低。在裂脑患者的颅骨内生活着两个断开连接的大脑和两个conscious minds，每个都拥有另一个半球无法获得的信息。这引发了关于自我连续性的迷人而诱人的问题。自我意识、人格感是传递给两个半球，还是仅与占主导地位的、能说话的半球相关？这些问题从未得到适当考虑。

在极少数情况下，相反的情况可能发生，当双胞胎出生时颅骨相连。在最近的一个案例中，有可信证据表明两个年轻女孩的大脑在丘脑水平相连。每个人似乎都能接触到另一个人看到的东西。如果这是真的，这将是大脑和minds的

非凡融合，远远超越了瓦格纳同名歌剧中特里斯坦和伊索尔德身份的狂欢般的orgasmic dissolution。

整合信息理论引入了一个精确的测量方法来捕捉consciousness的程度，称为 $\Phi$ 或phi（发音为“fī”）。以比特为单位表示， $\Phi$ 量化了系统进入特定状态时发生的不确定性减少，这超出了其部分独立产生的信息。（记住，信息是不确定性的减少。）系统的部分——modules——尽可能多地解释非整合的独立信息。因此，如果大脑的所有独立块单独考虑时已经解释了大部分信息，那么进一步的整合就很少发生。 $\Phi$ 测量网络在其当前状态下是多么协同的，即系统超越其部分总和的程度。因此， $\Phi$ 也可以被视为网络整体性的测量。

整合信息理论做出了许多预测。其中一个更违反直觉，因此更有力的预测是，整合信息来自系统内的causal interactions。当这些相互作用不再能发生时，即使系统的实际状态保持不变， $\Phi$ 也会收缩。

像你一样，我也许会被迪拜的哈利法塔所震撼，它几乎伸展一公里高进入蔚蓝的沙漠天空。当我在电脑屏幕上看到这座摩天大楼时，我视觉皮层中代表其形状的神经元放电，而我的auditory cortex或多或少处于静止状态。假设我auditory brain中的所有神经元都被短效巴比妥酸盐沉默，而我的形状神经元继续对这个priapic结构作出反应。我将无法听到任何声音。直觉上，由于本来就几乎没有声音，这应该没有任何区别。然而，整合信息理论预测，即使我大脑中的活动在两种情况下都相同（视觉形状中心有活动，但auditory区域没有）， $\Phi$ ——因此perceptual experience——会有所不同。神经元可以放电但不放电这一事实是有意义的，与神经元因为被人为沉默而无法放电的情况截然不同。

《银色白马》是夏洛克·福尔摩斯短篇小说中较为著名的一篇。其情节的关键在于“夜晚狗的奇异事件”，侦探向愚钝的警官指出狗没有吠叫这一事实。如果狗因为被套上口套而无法吠叫，这就没有意义了。但它没有被套口套，它没有吠叫是因为它认识入侵者。大脑也是如此。cortico-thalamic管弦乐队中的所有乐器都很重要，演奏的和不演奏的都是。虽然在这个例子中qualia的实际差异是微小的，但敏感的psychophysical技术应该能够检测到它。

Giulio整合信息理论的这种整体性方面并不否定大脑的某些片段对某些类别的qualia比其他部分更重要这一概念。关闭皮层中的visual form center几乎会消除摩天大楼的percept，而几乎不影响世界听起来的方式。相反，关闭auditory cortex几乎不会减少世界最高建筑的视觉，但会让我变聋。因此，寻找颜色、声音和agency的neural correlates仍然是有意义的。

计算 $\Phi$ 相当困难，因为必须考虑系统可以被分割的所有可能方式——即将网络切割成两部分的每种方式，将其切割成三部分的所有方式，以此类推，直到到达atomic partition，其中构成网络的所有单元都被单独考虑。在组合数学中，所有这些分割的数量是Bell's number。这个数字很大。对于构成*C. elegans*神经系统的302个神经元，这个网络可以被切割成部分的方式数量是超天文数字10后面跟467个零。因此，为任何神经系统计算 $\Phi$ 都是极其困难的，需要启发式方法、捷径和近似。

计算机对小型网络的仿真显示，实现高 $\Phi$ 值是困难的。通常，这样的电路只拥有几比特的integrated information。高 $\Phi$ 网络需要专门化和整合两者兼备，这是皮质-丘脑复合体中神经回路的一个标志。 $\Phi$ 表示与任何因果相互作用部分的网络相关联的意识repertoire的大小。系统越整合且分化，它就越有意识。

神经元间动作电位的同步放电是整合的另一种方式。请注意，如果大脑的所有神经元都以同步方式放电，如在大发作癫痫发作中，整合将是最大的，但分化将是最小的。最大化 $\Phi$ 就是在这两种对立倾向之间找到最佳平衡点。

就神经连接而言，大脑皮质中pyramidal neurons的主要特征是它们丰富的局部兴奋性连接，辅以较少的与远距离神经元的连接。由这样的组件构成的网络在数学上被称为小世界图。这些网络中的任意两个单元，任意两个皮质神经元，相距不超过几个突触。这个特性倾向于最大化 $\Phi$ 。

相反，对于由众多小的、准独立模块组成的网络， $\Phi$  是低的。这可能解释了为什么小脑尽管有大量神经元，但对意识贡献不大：它的类晶体突触组织使得其模块彼此独立运作，远距离模块间几乎没有相互作用。

为什么进化会偏好具有高  $\Phi$  的系统并不立即显而易见。它们获得了什么行为优势？

一个好处是能够结合不同传感器的数据来思考和规划未来的行动方案。具有高  $\Phi$  值的通用神经网络，如皮质-丘脑复合体，应该能够比更专门化的网络更好地处理意外和新颖情况——Black Swan事件。具有高  $\Phi$  大脑的生物应该比拥有相同数量神经元但整合性较差的大脑的生物更好地适应一个有许多独立actor在各种不同时间尺度上运作的世界。

重要的是要确立具有高  $\Phi$  的大脑在某种程度上优于整合性较差的大脑；也就是说，它们更好地适应复杂的自然环境。这将证明意识对生存是有益的，因为生存优势会给予具有高  $\Phi$  的生物。即使没有这样的证明，该理论也不会被否定，但这会使体验成为epiphenomenal。

integrated information theory面临的另一个重大挑战是解释无意识。它的假设意味着无意识过程比意识过程依赖较少的整合。然而，传统上归因于无意识的许多属性看起来相当复杂。如果用算法术语适当分析，它们是否比依赖意识心智的任务更少协同但更耗费劳动？你在第6章读到的sensory-motor zombie agents确实符合这个描述。它们表现出高度适应性但刻板的行为，这些行为依赖于专门化信息。需要记录技术来追踪清醒大脑内高度整合的主要复合体，并将其与大脑中同样活跃但整合性较差且调节无意识的那些部分区分开来。

Integrated information theory不仅指定了与系统每个状态相关联的意识量  $\Phi$ 。它还捕获了该体验的独特质量。它通过考虑底层物理系统能够产生的所有信息关系的集合来做到这一点。也就是说，integrated information产生的方式不仅决定了系统有多少意识，还决定了它有什么样的意识。Giulio的理论通过引入qualia space的概念来做到这一点，其维数与系统可以占据的不同状态数相同。对于具有  $n$  个二进制开关元素的简单网络，qualia space有  $2^n$  维，每个可能状态对应一维。每个轴表示系统处于该状态的概率。

任何物理系统的状态都可以映射到这个极其多维的qualia space中的一个形状上。它的表面是facets。这个形状的技术术语是*polytope*，但我更喜欢更诗意的*crystal*。处于任一特定状态的神经网络在qualia space中有一个相关的形状；它由信息关系构成。如果网络转变到不同状态，*crystal*会改变，反映网络各部分之间的信息关系。每个conscious experience都完全由其相关的*crystal*描述，每个状态感觉不同是因为每个*crystal*都是完全独特的。看到红色的*crystal*在某种独特的几何方式上不同于与看到绿色相关的那个。颜色体验的topology将不同于看到运动或闻到鱼的topology。

*Crystal*不同于底层的机制性、因果相互作用网络，因为前者是phenomenal experience，而后者是物质事物。该理论假设宇宙中存在两种不能相互还原的属性——心理的和物理的。它们通过一个简单而精密的定律联系起来，即integrated information的数学。

水晶是从内部观察的系统。它是头脑中的声音，颅骨内的光芒。它是你将永远了解世界的一切。它是你唯一的现实。它是体验的本质。食莲者的梦境、冥想僧侣的正念，以及癌症患者的痛苦之所以有这样的感受，是因为在万亿维度空间中不同水晶的形状——真正是一种至福的景象。整合信息的代数被转化为体验的几何学，验证了毕达哥拉斯的信念，即数学是终极现实：

数是形式和理念的统治者，是神灵和魔鬼的起因。

莱布尼茨会对整合信息理论非常认同。

这个理论可以用来构建一个“意识测量仪”。这个装置获取任何相互作用组件系统的连接图，无论是湿润的生物回路还是蚀刻在硅片上的回路，来评估该系统的意识repertoire的大小。意识测量仪扫描网络的物理回路，读出其活动水

平来计算 $\Phi$ 和网络momentarily体验的感受质(qualia)的水晶形状。需要开发几何微积分来确定这个水晶是否具有痛苦撞击脚趾的形态，还是满月下玫瑰香味的形态。

## 泛心论和德日进

---

我一直谨慎地强调，任何网络都拥有整合信息。这个理论在这一点上非常明确：任何功能连接性和架构产生大于零的 $\Phi$ 值的系统都至少有一丝体验。这包括地球上每个活细胞中发现的各种生化和分子调节网络。它也包括由固态器件和铜线制成的电子电路。实际上，在一篇为计算机科学家写的文章中，Giulio和我论证了人工智能的难以捉摸的目标——模拟人类智能——最终将通过关联和整合关于世界的大量信息的机器来实现。它们的处理器将具有高 $\Phi$ 。

无论生物体或人工制品是来自古老的动物界，还是来自其最近的硅基后代，无论这个东西有腿可以行走、有翅膀可以飞翔，还是有轮子可以滚动——如果它既有分化又有整合的信息状态，那么成为这样一个系统就会有某种感觉；它有一个内在的视角。它们相关的现象体验的复杂性和维度可能差异巨大，但每一个都有自己的水晶形状。

按照这个标准，在21世纪，填满地球的无数万亿有意识生物体被数十亿有感知的人工制品所加入——个人计算机、嵌入式处理器和智能手机。单独来看，这些人工制品可能只是最低限度的有意识，是照亮黑暗的火花。但综合起来，它们照亮了现象空间。

考虑地球表面所有通过互联网相互连接的计算机——几十亿台。每一台都由数亿个晶体管构建。整个网络中晶体管的总数（约 $10^{18}$ ）比单个人脑中的突触数量（高达 $10^{15}$ ）大一千倍。中央处理器中晶体管的典型门极只连接到少数几个其他门极，而单个皮质神经元则与数万个其他神经元相连。换句话说，神经组织实现了在二维硅技术中难以模仿的信息整合程度。

尽管如此，网络可能已经具有感知能力了。我们将通过什么迹象来识别它的意识？它会在不久的将来开始独立行动，以其自主性以令人担忧的方式让我们感到惊讶吗？

其含义还不止于此。即使是简单的物质也有一点 $\Phi$ 。质子和中子由从未单独观察到的三个夸克组成。它们构成了一个无穷小的整合系统。

通过假设意识是宇宙的一个基本特征，而不是从更简单的元素中涌现出来的，整合信息理论是泛心论的一个精致版本。所有物质在某种程度上都具有感知能力的假设因其优雅、简洁和逻辑一致性而极具吸引力。一旦你假设意识是真实的，在本体论上与其物理基质不同，那么得出整个宇宙都充满感知能力的结论就是简单的一步。我们被意识所包围和沉浸其中；它存在于我们呼吸的空气中、我们踩踏的土壤中、定居在我们肠道中的细菌中，以及使我们能够思考的大脑中。

苍蝇的 $\Phi$ ，更不用说细菌或粒子的 $\Phi$ ，在所有实际目的上都远低于我们在深度睡眠时体验到的 $\Phi$ 。充其量是对某种事物的模糊和未分化的感觉。因此，按照这个标准，苍蝇的意识水平会低于你在深度睡眠中的意识水平。但仍然如此。

当我谈论泛心论时，我经常遇到茫然不解的凝视。这样的信念违背了人们强烈持有的直觉，即感知能力只有人类和少数密切相关的物种才拥有。但当我们小时候第一次被告知鲸鱼不是鱼而是哺乳动物时，我们的直觉也会失效。我们必须习惯这一点。除非有理论在我们这边，否则我们可能无法认识到原子意识的本质。

泛心论有着古老而著名的谱系，不仅在佛教中，也在西方哲学中：从前苏格拉底思想家米利都的泰勒斯，到希腊时期的柏拉图和伊壁鸠鲁，启蒙时代的斯宾诺莎和莱布尼茨，浪漫主义时期的叔本华和歌德，一直延续到20世纪。

这让我想到了耶稣会神父兼古生物学家皮埃尔·德日进(Pierre Teilhard de Chardin)。他参与了北京人化石的发现，北京人是直立人(Homo erectus)的化石成员。他最著名的作品《人的现象》(The Phenomenon of Man)直到他死后才出版，因为罗马天主教会禁止其发表。在这本书中，德日进生动地描述了宇宙中精神通过达尔文进化论的兴起。他的复杂化定律(law of complexification)断言，物质具有内在的冲动，会组装成越来越复杂的群体。而复杂性孕育意识。德日进对他的泛心论立场非常明确：

我们在逻辑上被迫假设在每一个微粒中都存在某种原始形式的心理，甚至在那些（大分子及以下）复杂性如此低或适度以至于使其（心理）无法察觉的微粒中也是如此。

德日进并不止步于分子。不，精神的上升还在继续。这种原始的意识形式在动物中通过自然选择的力量变得更加高度发达。在人类中，意识转向内在，产生了自我意识。正是在这种背景下，朱利安·赫胥黎(Julian Huxley)说：“进化无非是物质意识到了自己。”复杂化是一个持续进行的过程，现在参与到智慧圈(noosphere)中，即无数人类心智的相互作用，这在当代城市社会中得以体现：

一束光芒从意识反思的第一个火花向外涟漪扩散。点燃点变得更大。火焰以不断扩大的圆圈蔓延，直到最终整个星球都被白炽光覆盖。这真的是一个新的层次，思维层，它……已经在植物和动物世界之上扩散开来。换句话说，在生物圈之外和之上存在着智慧圈。

如果互联网有守护神的话，那应该是德日进。

没有理由认为复杂化应该在我们蓝色星球与行星际空间的边界处停止。德日进相信整个宇宙朝着他所称的欧米伽点(Omega point)进化，届时宇宙通过最大化其复杂性和协同性而意识到自己。德日进之所以吸引人，是因为他的基本洞察与观察到的生物多样性（以变异量衡量）和复杂性在进化过程中增加的趋势相兼容，也与我在这里概述的关于整合信息和意识的观点相符。

不过，让我们不要过于兴奋。朱利奥的整合信息理论(integrated information theory)详细说明了蜜蜂的意识与大脑袋的两足动物的意识有何不同，它做出预测并提供如何构建有感知能力机器的蓝图。泛心论两者都做不到。

整合信息理论还处于起步阶段。它没有说明系统输入和输出之间的关系（不像著名的图灵测试那样测试智能）。整合信息关注的是系统内部发生的因果相互作用，而不是它与周围环境的关系（尽管外部世界会通过进化深刻地塑造系统的构成）。而且该理论尚未解释记忆或规划。我并不声称它是意识的终极理论，但它是朝着正确方向迈出的良好一步。如果它被证明是错误的，它将以有趣的方式错误，从而阐明问题。

## 令人谦卑的离别思考

---

弗朗西斯·培根(Francis Bacon)与笛卡尔一起是科学方法之父。培根比笛卡尔早生活和死亡二十年，在许多方面是他的英国对应者。笛卡尔是演绎理论家的原型，在总体原则驱动下寻找一般规律，而培根是完美的经验主义者，以归纳方式研究自然现象并跟随数据的指引。科学在自下而上的培根式分析和自上而下的笛卡尔式分析的相互作用中表现极其出色。尽管有反对者，科学最终将通过结合经验和临床研究与数学理论，以及日益增长的有意识人工制品工程学来理解意识。

让我以谦逊的请求作为结尾。宇宙是一个奇怪的地方，我们对它仍然知之甚少。仅仅二十年前，科学家发现宇宙质能中只有4%是构成恒星、行星、树木、你和我的那种材料。四分之一是冷暗物质，其余是被称为暗能量的奇异东西。宇宙学家不知道暗能量是什么或它遵循什么定律。这种诡异的东西与意识之间是否存在某种短暂的联系，正如小说家菲利普·普尔曼(Philip Pullman)在他的三部曲《黑暗物质》(His Dark Materials)中所暗示的？极不可能，但仍然……我们的知识只不过是照亮我们周围广阔黑暗的一团火，在风中摇曳。所以，让我们对探索意识源泉的理性替代解释保持开放。

## 第九章：

---

### 我概述了一种测量意识的电磁装置，描述了利用基因工程力量在小鼠中追踪意识的努力，并发现自己在建造皮质观测台

---

关于恒星的主题，所有最终无法归结为简单视觉观察的研究都……必然被我们拒之门外……我们永远无法通过任何手段研究它们的化学组成。

—奥古斯特·孔德(August Comte)，《实证哲学教程》(Cours de Philosophie Positive) (1830-1842)

意识是现实的基本的、不可简化的方面吗？还是它从有序物质中产生，正如大多数科学家和哲学家所相信的那样？我想在死前知道答案；所以我不能永远等下去。诡辩式的哲学辩论很有趣，甚至可能有所帮助，但它们无法解决根本问题。发现物质之水如何变成意识之酒的最佳方式是通过实验结合理论的发展。

目前，我忽略了关于意识确切定义的琐碎辩论，比如它是否是一种现象伴随物(epiphenomenon)，无力影响世界，或者我的直觉是否有意识但没有告诉我。这些问题最终都需要解决，但今天担心它们只会阻碍进步。不要被哲学上的哗众取宠和声称意识难题(Hard Problem)将永远伴随我们的宣言所迷惑。哲学家处理的是信念体系、简单逻辑和观点，而不是自然法则和事实。他们提出有趣的问题，提出迷人且具有挑战性的困境，但他们在预测方面的历史记录平平。考虑一下法国哲学家、实证主义之父奥古斯特·孔德(August Comte)的章节引言。在他自信地宣称我们永远不了解恒星物质的几十年后，通过光谱分析恒星的光就推断出了它们的化学成分，这直接导致了气体氦的发现。相反，听听预测记录要好得多的学者弗朗西斯·克里克(Francis Crick)的话：“说事情超出了科学的范围是非常鲁莽的。”我们没有理由不能最终理解现象心智如何融入物理世界。

我的方法是直接的，我的许多同事认为这是不明智和幼稚的。我将主观体验视为既定事实，并假设大脑活动足以体验任何事物。虽然内省和语言对社会生活至关重要，是文化和文明的基础，但它们对体验某些事物来说并非必需。这些假设使我们能够在人类和动物身上以前所未有的精度研究意识的大脑基础。让我给你举两个例子来说明我的意思。

## 重伤患者的意识测量仪

---

当你从无梦的深度睡眠中被唤醒时，你记得任何事情。前一刻你还在回顾一天的事件，下一刻你知道的就是早上醒来了。与伴有生动且常常奇异梦境体验的快速眼动(REM)睡眠不同，在非REM睡眠期间意识处于低潮。然而，当身体睡觉时，大脑仍在活跃。只要看看睡眠大脑的脑电图(EEG)轨迹，其特征是缓慢、深沉且规律的波形。此外，皮质神经元的平均活动与安静清醒时大致相同。那么为什么意识会消退呢？朱利奥·托诺尼(Giulio Tononi)在前一章讨论的理论预测，如果深度睡眠时的整合度比清醒时低，就会出现这种情况。

朱利奥和他的年轻同事、现为意大利米兰教授的马尔切洛·马西米尼(Marcello Massimini)着手证明这一点。他们通过一种称为经颅磁刺激(TMS)的技术，向志愿者的大脑发送单次高磁场脉冲。放电的塑料包裹线圈贴住头皮，最终在颅骨下方的灰质中诱导出短暂的电流(受试者由于皮肤刺激会感到轻微刺痛)。这个脉冲激发脑细胞和附近的通过纤维，进而通过突触连接的神经元形成级联活动，在头部内产生回响。在不到一秒钟的时间内，这种兴奋消退。

朱利奥和马尔切洛在受试者头皮上安装了64个电极，测试对象要么安静休息，要么睡眠。清醒时，TMS脉冲后的脑电图显示典型的快速、反复波形的增强和衰减模式，持续约三分之一秒。对脑电图信号的数学分析显示，高振幅电位的热点从TMS线圈所在位置上方的前运动皮质，传播到另一半球匹配的前运动皮质，再到运动皮质和后方的顶叶皮质。把大脑想象成一个大教堂钟，TMS设备就是钟锤。一旦敲击，铸造良好的钟会以其特有的音调响起相当长的时间。清醒的皮质也是如此，以每秒十到四十次的频率嗡嗡作响。

相反，睡眠受试者的大脑表现得像一个发育不良、调音不佳的钟。虽然脑电图的初始振幅比受试者清醒时更大，但其持续时间要短得多，并且不会在皮质间向其他连接区域产生回响。虽然神经元仍然活跃，正如强烈的局部反应所显示的，但整合已经崩溃。很少有清醒大脑典型的空间分化和时间多样化的电活动序列存在，正如预测的那样。接受全身麻醉的志愿受试者也是如此。TMS脉冲总是产生简单的局部反应，表明皮质-皮质相互作用的崩溃和整合度的降低，符合朱利奥的理论。到目前为止，理论得分1比0。但情况变得更好。

我在第5章介绍了植物状态患者。大脑的重大创伤使他们活着，保持了觉醒，即睡眠-清醒循环，但严重残疾，卧床不起，没有任何目的的行为。相比之下，微意识状态(MCS)患者有波动的非反射性反应迹象，如用眼睛追踪目标或对简单命令的言语或手部反应。虽然意识已经离开了植物状态患者，但在MCS患者中得到了部分保留。

神经学家Steven Laureys、Marcello Massimini、Giulio和他们的同事测量了此类患者的大脑整合范围。他们对睁眼患者的顶叶或额叶施加TMS脉冲。结果是明确的。植物状态患者的EEG反应简单而局部—通常是缓慢的单一正负波(当他们有任何反应时)，与深度睡眠和麻醉反应非常相似。相反，在最小意识状态(MCS)患者中，磁脉冲触发了健康清醒受试者预期的复杂电反应，多个焦点在不同皮质部位转移。五名患者在从昏迷中苏醒后立即从重症监护室招募。三名最终恢复了意识，两名没有。那些最终恢复的患者意识的出现之前，对磁脉冲的EEG反应出现了延长和复杂化—他们从单一的局部波进展为更丰富的时空模式。换句话说，Massimini-Tononi评估整合的方法可以作为一个粗略的意识测量仪，评估严重受损患者的意识水平。小型化的TMS线圈与几个电极的EEG设备相结合，可以轻松组装成临床实践的仪器。这将提高区分真正无意识患者与部分或完全有意识患者的精确度，基于皮质-丘脑复合体在意识状态下比在植物状态、无意识状态下具有更大的整合性。

## 使用光遗传学追踪意识的足迹

---

当你凝视一只狗的眼睛时，一个在两端都被永恒所包围的旅程上的同伴凝视着你。它的心智与你的不同，但与你的心智相关。狗和人都体验生活。人类是特殊的，他们在所有其他生物之上被意识的天赋所选中，这种想法源于深深持有的犹太-基督教信仰，即我们在事物秩序中占据特权地位，这是一种有圣经但无经验基础的信仰。我们并不特殊。我们只是无数物种中的一种。我们是不同的。但每个其他物种也是如此。从科学角度来看，这意味着我们可以在其他有情生物中研究意识。

但在我们这样做之前，我们需要解决一个紧迫的伦理问题。人类有什么权利将其他物种置于自己的欲望之下？这当然是一个复杂的话题。但简单和详细的答案是，唯一可能的理由是减少那些由于内省习惯而特别容易遭受此类痛苦的生物的可预防痛苦，也就是人类。

我遇到过一只在肇事逃逸事故中后腿被压碎的狗。兽医将她安装到一种战车中，使她能够用两条腿和两个轮子奔跑。她是运动和活动的模糊影像，是我见过的最快乐的狗之一，似乎完全没有意识到她的缺陷。看着她让我哭泣。她没有认知装置来思考可能的情况，如果那辆车那天没有撞到她，她如何能够四处奔跑。她活在当下。另一方面，我们人类被“祝福”拥有前额皮质，使我们能够将自己投射到不同的未来，想象替代生活，可能的情况。这使得类似的人类残疾—想想被路边炸弹炸掉一只或多只肢体的归来老兵—更难以承受。

减轻人类痛苦是以侵入性方式研究动物的唯一伦理上值得的理由。我的一个女儿死于婴儿猝死综合征；我的父亲被帕金森病摧残；一个朋友在精神分裂症的华丽发作中自杀；阿尔茨海默病等待着我们许多人的生命终点。消除这些和其他折磨大脑的疾病需要动物实验—以关怀和同情进行，并尽可能在他们的合作下进行。

通过这种从人类转向动物的焦点转移，我们获得的是直接探测他们大脑的潜力，这是我们无法在人类身上做到的。我们失去的是受试者向我们讲述他们体验的可能性。但婴儿和严重残疾患者也不能这样做。所以我们必须想出巧妙的方法，通过观察动物的行为来推断它正在经历什么，就像每个父母对新生儿所做的一样。

研究感知和认知的心理学家和神经科学家的选择生物是旧世界猴。它们没有濒危，它们的大脑皮质在许多凹陷和褶皱方面与我们的相似。人类大脑重3磅(1,500克)，有860亿个神经元，而猴子的大脑要轻得多，重3盎司(86克)，包含60亿个神经细胞。正如我在第4章中讨论的，猴子感知许多与人类相同的视觉错觉。这意味着可以使用倾听的微电极和观察单个神经细胞工作的显微镜来探索视觉感知的机制基础。

然而，我已经提到的一个惊人的技术突破使得谦逊的小鼠—其大脑重量不到0.5克，只有7100万个神经元—成为科学家最有可能首先识别意识细胞足迹的生物。

每一代天文学家都发现宇宙比他们的前辈想象的要大得多。大脑复杂性也是如此。每个时代最先进的技术，当应用于大脑研究时，不断揭示更多层次的嵌套复杂性，像一套永无止境的俄罗斯套娃。

动物由大量不同的细胞类型组装而成：血细胞、心脏细胞、肾脏细胞等。同样的原理也适用于中枢神经系统。在神经系统内可能有多达一千种不同的神经细胞亚型和支持细胞——胶质细胞和星形胶质细胞。每种细胞类型都由特定的分子标记、神经元形态、位置、突触结构和输入-输出处理来定义。在视网膜中，大约有六十种神经元细胞类型，每一种都完全覆盖视觉空间（意味着视觉空间中的每一个点都至少由每种类型的一个细胞处理）。这个数字对于任何一个脑区来说可能都具有代表性。

不同的细胞类型以特定的方式连接。新皮质深层第5层锥体神经元将其纤细如丝的输出导线——轴突——蜿蜒延伸到远处中脑的上丘，而附近锥体细胞的轴突在其邻近区域产生侧支，然后将其脉冲通过胼胝体发送到对侧大脑皮质半球；第三个锥体细胞将其信息传回丘脑，并通过分支轴突将副本发送到网状核。可以合理地假设每个细胞类别都向其目标传达独特的信息（否则，一条分支到不同目标的单一轴突就足够了）。然后还有许多局部的抑制性中间神经元，每个都有其独特的支配目标的方式。所有这些构成了细胞间相互作用的丰富基质，具有组合式的大量回路模式。想象一个由一千种不同类型的LEGO积木组成的建筑套装，具有不同的颜色、形状和大小。人类大脑皮质有160亿块从这些类型中选择的积木，按照极其复杂的规则组装，比如红色 $2\times 4$ 积木连接到蓝色 $2\times 4$ 积木，但只有当它靠近黄色 $2\times 2$ 屋顶瓦片和绿色 $2\times 6$ 块时才可以。从这里诞生了大脑巨大的相互连接性。

批量组织技术如fMRI可靠地识别哪些脑区与视觉、想象、疼痛或记忆相关，这是颅相学思想的重生。脑成像追踪一百万个神经元的能耗，无论它们是兴奋性还是抑制性的，是局部投射还是全局投射，是锥体神经元还是星形细胞。由于无法解析至关重要的回路水平的细节，它们不足以胜任当前的任务。

随着我们对大脑理解的成熟，我们干预的愿望，帮助改善心智容易患上的许多病理状况的愿望也相应增长。然而今天的工具——药物、深部脑电刺激和经颅磁刺激——都很粗糙、不精确，有许多不良副作用。我在加州理工学院的同事David Anderson将使用它们比作通过将机油倒在发动机上来换汽车机油：一些机油最终会渗透到正确的位置，但大部分会流到造成更多伤害的地方。

技术突破前来救援，这是分子生物学、激光和光纤的融合，被称为光遗传学(optogenetics)。它基于三位德国生物物理学家——Peter Hegemann、Ernst Bamberg和Georg Nagel——在单细胞绿藻光感受器上的基础发现。这些光感受器直接（而不是像你眼中的那些间接地）将入射蓝光转换为兴奋性的正电信号。这个三人组分离出了这种蛋白质的基因，一个跨越神经元膜的光门控离子通道，称为通道视紫红质-2(channelrhodopsin-2, ChR2)。Bamberg和Nagel随后与斯坦福大学的精神病学家和神经生物学家Karl Deisseroth，以及现在麻省理工学院的神经工程师Ed Boyden进行了富有成效的合作。

该小组取得了ChR2基因，将其插入小病毒中，并用这种病毒感染神经元。许多神经元接受了这些外来指令，合成了ChR2蛋白，并将这些异位的光感受器整合到其膜中。在黑暗中，这些受体安静地停留在那里，对其宿主细胞没有明显影响。但蓝光的短暂闪烁会导致这些细菌光感受器轻微震动其宿主细胞。它们的集体行动产生动作电位。每次打开光时，细胞可靠地脉冲，正好一次。因此，神经元可以通过精确定时的蓝光刺激被驱动产生脉冲。

生物物理学家在他们的工具包中添加了另一种天然存在的光敏蛋白质。它来源于生活在撒哈拉沙漠干燥盐湖中的古细菌。向其照射黄光会产生抑制性的负信号。使用相同的病毒策略，可以使神经元稳定地在其膜中整合任一类型的蛋白质，这样它就可以被蓝光兴奋或被黄光抑制。每次蓝光闪烁都会引起脉冲，就像按下钢琴键时发出的音符。同时的黄光闪烁可以阻止脉冲。这种以毫秒精度在单个神经元水平控制电活动的能力是前所未有的。

这项技术的好处更加深远，因为携带光感受器基因的病毒可以被改造携带一个载荷（启动子），只在具有适当分子标签的细胞中激活病毒遗传指令。因此，不是激发特定区域的所有神经元，激发将被限制在合成特定神经递质或将输出发送到特定位置的神经元上。所需要的只是该特定细胞类型的分子邮政编码；例如，所有表达生长抑素激素的皮层抑制性中间神经元。它们为什么合成这种物质并不重要，重要的是这种蛋白质可以用作独特的分子标签来标记细胞，使它们容易被激光激发或抑制。

Deisseroth的团队利用这种能力，将ChR2引入位于小鼠大脑深处侧下丘脑的一个神经元亚群。在这里，不到一千个细胞产生orexin（也称为hypocretin），这是一种促进清醒的激素。orexin受体的突变与narcolepsy（一种慢性睡眠障碍）有关。经过操作后，几乎所有orexin神经元，但没有其他混合的神经元，携带ChR2光感受器。此外，通过光纤传递的蓝光精确可靠地在orexin细胞中产生尖峰波。

如果在睡眠中的小鼠身上进行这个实验会发生什么？在没有这种特定基因操作的对照动物中，几百次蓝色闪光在约一分钟后唤醒了啮齿动物。这是判断手术效果的基线条件，包括插入光纤等。当同样的光照射到携带ChR2通道的动物时，动物醒来的时间缩短了一半。也就是说，照亮大脑深处并导致一个已知身份和位置的微小神经元亚群产生电尖峰的光线唤醒了动物。是来自侧下丘脑的orexin释放驱动了这种行为。这个典型研究在大脑神经元亚群的电活动与睡眠-觉醒转换之间建立了令人信服的因果关系。

过去几年中数十个这样精美的干预性小鼠实验教会了我们关于参与厌恶条件反射、Parkinson's disease、交配、雄性间攻击和其他社会互动、视觉辨别和焦虑等的电路元素，仅举几例。它们甚至帮助恢复了因视网膜退化而失明的小鼠的视力。

使用基因工程已经开发出各种变体。在某些变体中，一次蓝光脉冲可以将神经元开启数分钟，而一次黄光脉冲会再次关闭它们，类似于电灯开关。在pharmacogenetics中，向大脑区域注射一种本来无害的化合物，可以开启或关闭基因识别的细胞亚群，从而实现对神经元群体的长期控制。分子导向神经工程师的工具集正在不断扩展。

## 迈向新天地

---

2011年，我加入西雅图的Allen Institute for Brain Science，担任首席科学官。这家非营利医学研究机构成立于2003年，由Microsoft创始人和慈善家Paul G. Allen慷慨的种子资助启动，旨在推动神经科学的研究进步（“推动发现”是他们的座右铭）。为此，Allen Institute开展了一种独特的高通量神经科学研究，这在学术环境中无法完成。其旗舰产品是在线Allen Mouse Brain Atlas，这是一个高度标准化、细胞分辨率、公开可访问的数字图谱，展示了小鼠基因组中所有两万个基因在动物大脑中的表达模式。对于任何特定基因，你可以在线查找其相关RNA在大脑中的表达位置（通过原位杂交协议映射）。这项大规模的工作是人类对哺乳动物大脑回路如何构建的缓慢增长理解中的一个重要里程碑。其他在线公共资源包括人脑图谱和小鼠大脑中神经投射的图谱。

该研究所现在正在寻找理解神经元信息如何编码和转换的方法。天文学家、物理学家和工程师建造太空和地面望远镜，凝视时空的遥远处，见证我们宇宙及其恒星组成部分的起源。这些观测站需要十年或更长时间来组装，需要数百名或更多技术专家和科学家的专业知识。我们正在规划阶段构建观测站来观察颅骨下大脑中运行的心智。我称之为“Project Mindscope”。实验挑战是装配光学、电子和计算机设备，在小鼠从事某种视觉行为时观察数万个基因可识别回路元素的同时放电活动。

当我渴望理解意识时，为什么要研究小鼠而不是在进化上更接近人类的猴子呢？首先，小鼠大脑与我们的大脑在遗传和神经解剖学上有很多相似之处：它有一个小而光滑的新皮层，包含1400万个神经元——比我们的少一千倍。小鼠皮层的一小块与人类灰质的一小块并没有太大差异。但最重要的是，它非常适合基因操作。在所有脊椎动物中，对小鼠分子生物学的理解——即其DNA如何转录为RNA并翻译为蛋白质——是迄今为止最先进的。小鼠重组DNA技术始于1970年代中期，转基因小鼠的创造已经是一项成熟的技术。对于我的研究至关重要的是，主要神经元细胞类型的独特分子邮政编码及其投射位置正在被破译。Allen Institute的Hongkui Zeng是利用这种遗传地址簿来工程化小鼠的大师——健康的动物，其神经元表达ChR2，这样将光照射到大脑中会在易受影响的神经元中引发一连串尖峰放电；或其他动物，其抑制性中间神经元在适当波长的光照射下会发出荧光，呈现诡异的绿色或番茄红色。

这种光学与遗传学的奇妙结合的重要性在于，它允许测试关于心智回路的非常具体的想法。考虑将图像闪现到小鼠眼中时触发的尖峰波。它沿着视神经上行，进入初级视觉皮层，然后到达其他八个视觉区域，穿过运动皮层，然后下行到控制头部、前爪或其他肢体的运动神经元。第4章概述了Francis Crick和我的推测，即单一的这种尖峰波可以在几百毫秒内触发一些简单行为，比如推杠杆，但不会产生有意识的感觉。第6章提到了我们所有人在一天中都会进行的许多这种zombie行为。我们假设，一旦皮层-皮层或皮层-丘脑反馈通路参与进来，建立起表现为神经元联盟强烈共同放电的reverberatory activity时，意识就会产生，这在某种程度上与物理学中驻波的概念相关。当神经活动从高级视觉皮层向下传递到低级区域，或从大脑前部传递到后部时，这个神经元联盟所代表的整合信息会激增，产生有意识的感觉或思想。

这样的假设现在可以在适当工程化的小鼠中进行测试：训练它们进行任何一种视觉辨别行为，然后暂时关闭从高级皮层区域反馈到低级区域的连线。如果Francis和我的观点正确，先天的、刻板的或高度排练的视觉-运动行为只会受到最小影响。但依赖于小鼠意识的复杂行为会失败。测试我们的假设需要小鼠能够被训练对光学双稳态错觉作出反应，区分图形与其背景，或学会将视觉地标与美味食物关联起来。如果皮层-皮层反馈在整个大脑中被关闭，我们将创造出真正的zombie小鼠，无法进行现象体验！如果反馈重新激活，有意识的感觉就会恢复。

Stanislas Dehaene的一项经典fMRI实验对比了人们对短暂闪现的可见单词和相同持续时间但被掩蔽而变得不可见的同一单词之间的差异（见第4章）。当单词不可见时，血流动力学活动仅限于一小组区域，而对单词的有意识觉察激活了大脑前部和后部更大范围的皮层区域。当另一组研究人员使用掩蔽音调而不是图像时，也发现了同样的结果。阈

下刺激只引发微弱活动，而有意识感知的刺激被放大了许多倍。没有理由不能在小鼠中重复这种实验的变体，但现在使用微电极阵列或共聚焦双光子显微镜来观察支撑与有意识感知相关的广泛激活的所有神经元。

对构成丘脑-皮层复合体的庞大、异质和纠缠网络进行系统和全面的结构和功能探索的重要性怎么强调都不过分。几年内，Allen Institute将拥有构成小鼠皮层及其输入的所有细胞类型的完整分类学。确实，解剖学如此重要，以至于这位大师级神经科学家、西班牙人Santiago Ramón y Cajal绘制的啮齿动物皮层微回路图现在纹在我的左上臂上，这是我所从事探索的无声见证。

这是令人兴奋的时代。我享受在西雅图的生活，那里有壮观的景色、户外文化和自行车道。有时有点困难，因为我仍然是加州理工学院的教授，有一大群学生和博士后研究员需要指导。但我一直信奉“死后再睡”的理念。

生物学涉及分子-细胞水平上闻所未闻的复杂性和特异性。当物质被认为是希腊古典四元素（土、水、气、火）的混合物时，化学没有取得进展。意识也是如此。现象体验不是从活跃或沉默的大脑区域中产生的，而是从神经元联盟的不断形成和解散中产生的，其复杂性和表征能力是我们最私密思想的终极基质。

## 第10章：

---

### 关于最终问题的思考

---

在此我思考那些在礼貌的科学讨论中被视为禁区的最终问题：即科学与宗教的关系，上帝的存在，这个上帝是否能干预宇宙，我的导师的去世，以及我最近的苦难

当我思考我生命的短暂，被前后的永恒所吞没，我所占据的微小空间，甚至我所能看到的，都被我无知且不知我的无限广袤空间所吞没时，我感到恐惧，并惊讶于我在这里而不是在那里；因为没有理由说为什么在这里而不是在那里，为什么是现在而不是那时。是谁把我放在这里？通过谁的命令和指导，这个地方和时间被分配给了我？……这些无限空间的永恒沉默让我恐惧。

——布莱兹·帕斯卡尔，《思想录》(1670)

保罗·高更令人难忘的杰作《我们从哪里来？我们是谁？我们要到哪里去？》，创作于他生命最后几年在大溪地的时光，完美地概括了我痴迷的三个问题：我们——人类、狗和其他有感知的生物——从哪里来？我们是谁？我们要到哪里去？我是一名自然科学家。我有一种根深蒂固的愿望，要找到这些问题的答案，理解物理宇宙以及意识。我寻求理解整个存在——不是像神秘主义者那样，通过我有时在圣盖博山脉高海拔地区跑步数小时体验到的那种狂喜体验，而是以理性、智慧的方式。

这些最后的篇章呈现了我对科学与宗教的一些个人思考，一个迟来的成年过程，迫使 I 重新评估我童年的信仰，以及一些自传片段，阐明了我为什么关心自由意志的问题。我通过与学生和同事的接触知道，不少人夜不能寐，思考着这些问题。这最后一章是为你们而写的。

## 二元论、灵魂与科学

---

柏拉图，西方哲学的鼻祖，将个体设想为一个非物质的、不朽的灵魂，被监禁在一个物质的、凡人的身体中。这个概念正是二元论的体现，即现实由两种根本不同的事物组成——精神的或灵性的，以及物理的。柏拉图通过他于公元前387年建立的学院传播他的思想。这是西方文明的第一个高等学习机构。我称自己为学者(academic)，追溯到雅典英雄阿卡德摩斯(Akademos)，他的学校所在的橄榄园就以他的名字命名。

这些柏拉图式的观点后来被新约圣经吸收。它们构成了基督教灵魂教义的基础，灵魂将在时间的尽头复活，与上帝永远交流。这种对超然、不朽灵魂的信仰，认为它位于意识的核心，在思想史上反复出现，并被世界各地的许多信仰广泛认同。

许多读者可能不会对这种明显的二元论信仰有太多同情。然而，原教旨主义——对理性的、人文主义的、自由主义世界观的毫不妥协的拒绝，转而严格遵守关于身体和灵魂的教义和核心信仰——正在全世界兴起。基督教原教旨主义和极端伊斯兰变体都是如此。而且比以往任何时候都更多的年轻人愿意以他们特定的神的名义杀死他人和自己。这与尼采在狂乱中宣告“上帝死了！”“时的想法相去甚远。

当代处理心灵-身体问题的学术书籍在附带说明中——如果它们根本提到的话——摆脱了上帝和灵魂。作者以轻蔑的方式指出科学与这些过时思维模式的明显不兼容性。这与三四个世纪前的情况相去甚远，那时书籍和建筑都被奉献为“*ad majorem dei gloriam*”——为了上帝更大的荣耀！

笛卡尔，启蒙哲学家，假定太阳下的一切都由两种物质之一组成。你能触摸到的、具有空间延展性的物质是“*res extensa*”（广延物）；这包括动物和人的身体和大脑。你看不到的、没有长度和宽度的、激活人类大脑的物质是“*res cogitans*”（思维物），灵魂物质。

我们大脑的工作通常与当时最先进的技术进行比较。今天，它是庞大而复杂的互联网。昨天，它是数字计算机。上世纪，它是法国凡尔赛宫廷喷泉中神、半人马、海神、仙女和英雄的活动雕像。笛卡尔论证说，就像驱动这些简单机器的水一样，“动物精神”流过所有生物的动脉、脑腔和神经管道，使它们运动。与中世纪经院传统及其无休止的推测进行根本性决裂，笛卡尔为感知和行动寻求机械解释。在他对大脑和身体解剖的指导下，他论证大多数行为是由按大小、形状和运动区分的粒子的作用引起的。

但笛卡尔无法为智力、推理和语言设想机制。在十七世纪，没有人能设想算法（我们今天所说的详细的、逐步的指令的机械应用）如何使计算机下棋、识别面部和翻译网页。笛卡尔不得不求助于他神秘的、空灵的物质——“*res cogitans*”，以某种模糊的方式进行思考和推理。作为虔诚的天主教徒，他通过将“*res cogitans*”仅限于人类来维护人类与无灵魂动物之间的绝对区别。正如他非常明确地写道，狗被马车撞到时可能会凄惨地嚎叫，但它不感到疼痛。

如果说我在毕生探索心灵-身体关系中学到了什么，那就是：无论意识是什么——无论它如何与大脑相关——狗、鸟类以及众多其他物种都拥有意识。正如我在第3章中阐述并在最后一章中再次强调的那样，犬类的意识与我们的不同——首先，狗的内省能力要弱得多，而且不会说话——但毫无疑问的是，它们也体验着生活。

二元论的两位近期捍卫者，哲学家卡尔·波普尔和神经生理学家、诺贝尔奖得主约翰·埃克尔斯，在第7章中出现过。让我重复一下我在那里讨论他们关于自由意志观点时提出的一个观点。他们所倡导的二元论，即心灵强迫大脑执行其意愿，是不令人满意的，原因是25岁的波希米亚公主伊丽莎白三个世纪前就已经向笛卡尔指出的——非物质的灵魂通过什么方式指导物理大脑来实现其目标？如果灵魂是不可言喻的，它如何能操纵像突触这样的实际物质？

很容易看到因果关系从大脑流向心灵，但反向则很困难。任何心灵到大脑的交流都必须与自然法则相符，特别是与能量守恒原理相符。让大脑做事情，比如干预突触，需要灵魂必须执行的工作，而这必须被解释清楚。

两者之间相互作用的本质并非唯一的问题。灵魂如何记住任何事情？它有自己的记忆吗？如果有，在哪里？它遵循什么逻辑？当大脑死亡时，灵魂会发生什么？它会像幽灵一样漂浮在某种超空间中吗？在身体诞生之前，这个灵魂又在哪里？这些问题没有与我们对物理世界的认知相符的答案。

如果我们真诚地寻求对宇宙及其中一切事物的单一、理性且智力一致的观点，我们必须摒弃不朽灵魂的经典观念。这是一个深深嵌入我们文化中的观念；它渗透在我们的歌曲、小说、电影、伟大建筑、公共话语和神话中。科学把我们带到了童年的终点。成长对许多人来说是不安的，对少数人来说是难以忍受的，但我们必须学会看到世界的本来面目，而不是我们希望它成为的样子。一旦我们摆脱了魔幻思维，我们就有机会理解我们如何融入这个不断展开的宇宙。

我们这个时代的主导知识立场是物理主义——归根结底，一切都可以还原为物理学。无需诉诸除空间、时间、物质和能量之外的任何东西。物理主义——与唯物主义相差半音——因其形而上学的简约性而具有吸引力。它不做任何额外假设。

相反，这种简单性也可以被视为贫乏。本书论证物理主义本身过于贫乏，无法解释心灵的起源。在前一章中，我勾勒了一个增强物理主义的替代性解释。这是属性二元论的一种形式：整合信息理论假设有意识的现象体验与其底层物理载体是不同的。从信息角度来说，悲伤的体验是一个水晶，是一万亿维空间中一个极其复杂的形状，在质量上不同于产生悲伤的大脑状态。有意识的感觉源于整合信息；因果关系从大脑的底层物理流动而来，但不是以任何易于理解的方式。这是因为意识依赖于系统超越其各部分的总和。

把这个水晶想象成二十一世纪版本的灵魂。但是，*hélas*，它不是不朽的。一旦底层物理系统瓦解，水晶就会消失。它回到未成形的虚空，回到系统构成之前它所在的地方。

然而，在这种崩溃发生之前，构成这个水晶的因果关系可以上传到计算机上。这就是雷·库兹韦尔和其他技术专家希望能让他们永生的臭名昭著的奇点——书呆子的狂欢。一旦相关的整合信息被简化为电子模式，它就可以被复制或编辑、出售或盗版、与其他电子心灵捆绑，或被删除。

但是没有某种载体，某种机制，整合信息无法存在。简而言之：没有物质，就没有心灵。

## 宗教、理性和弗朗西斯·克里克

---

弗朗西斯·克里克体现了宗教与科学之间的历史敌意。他宽容的中产阶级英国成长环境中没有任何迹象表明这一点，但从与他的多次讨论中我知道，他感到有动力要将上帝从世界中驱逐，用基于自然力量的解释取代对生命和意识的超自然解释。他想要将上帝永远从理性和有教养的话语领域中放逐出去。

弗朗西斯在理解生命这一目标上取得了惊人的成功。尽管前生物世界中生命的起源仍然模糊不清，但其进化的概念架构都在那里。现在判断他在第二个目标上取得了多大进展还为时过早。

Francis对有组织宗教的反对在1961年变得传奇，当时他从剑桥大学丘吉尔学院辞职，抗议在学院场地内增建教堂的计划。Francis认为宗教在一个以科学、数学和工程为重点的现代学院中没有立足之地。温斯顿·丘吉尔爵士——学院就是以他的名字命名的——试图安抚他，指出建造教堂的资金是私人筹集的，没有人会被强迫参加那里的礼拜仪式。Francis回应说提议为学院筹建一个妓院：没有人会被强迫利用其服务，无论宗教信仰如何，都接受男性顾客。在他的信中还附上了十几尼的定金。可以理解，这结束了两人之间的进一步通信。

在我认识Francis的时候，他对任何宗教思想的强烈反对已经变得温和了。在他和Odile位于山顶的家中共进晚餐时，我们偶尔会讨论罗马天主教会及其对进化论、独身制等问题的立场。他知道我是在天主教环境中长大的，并且断断续续地继续参加弥撒。他从不深究我信仰的基础，因为他是个善良的人，想要免我摸索解释的尴尬——特别是因为我的信仰不会干扰我们在严格的经验主义框架内理解意识的追求。

值得注意的是，在他1994年的著作《惊人的假说》(The Astonishing Hypothesis)中，他概述了自己对心身问题的观点，并承认：“或者，某些更接近宗教观点的看法可能变得更加可信。”这个令人震惊的让步立即被以下说法中和了：“总有第三种可能性；事实支持一种看待心脑问题的新的、另类的方式，这种方式与今天许多神经科学家持有的相当粗糙的唯物主义观点以及宗教观点都有显著不同。”这不是政治正确的表达——远非如此。在我认识的所有人中，Francis最愿意接受新的、另类的，甚至激进的解释，只要它们与大部分既定事实一致、可验证，并开辟了新的思考和实验途径。

## 自然神论，或作为神圣建筑师的上帝

---

所有存在主义谜题中最大的一个原因是：为什么存在某些东西而不是什么都没有。当然，最自然的存在状态——在尽可能少假设的意义上——是空无。我不是指在物理学家手中被证明如此丰产的空间。我指的是任何东西的缺失：空间、时间、物质和能源。什么都没有，*rien, nada, nichts*。但我们在乎，这就是谜题。

年轻的Ludwig Wittgenstein在第一次世界大战的战壕中以及作为战俘时，在他的《逻辑哲学论》(Tractatus logico-philosophicus)中表达了对此的惊叹：“世界中事物如何存在并不神秘，神秘的是它的存在。”

宇宙学已经将这个问题追踪到创造本身的一点，那个无法想象的炽热大爆炸。它发生在137亿年前，在真正的深层时间中，完全超越了任何人类经验。正是在那里——尽管Stephen Hawking和其他人做出了最大努力——物理学遇到了形而上学。

是谁或什么设定了初始奇点的条件，当时一切都是被压缩成密度无限的单一点？它从哪里来？“无中不能生有”的原则难道不像适用于宇宙中的任何事物一样适用于整个宇宙吗？控制这个宇宙的定律从何而来？是谁或什么建立了量子力学和广义相对论？这些定律是必然的吗？宇宙能够遵循其他定律并仍然保持自治吗？一个不遵循量子力学的宇宙是可行的，甚至是可以想象的吗？

一个理性的解释是造物主(demiurge)，一个永远存在、现在存在、将来也存在的至高存在。在时间之外，这个实体创造了自然定律并意志大爆炸的存在。物理学然后产生了稳定的时空结构和我们的宇宙。在这最初的创造行为之后，神圣建筑师让宇宙自行发展，自由地通过机遇和必然性进化。最终，生物从原始泥浆中产生并建造神庙来赞美这个至高存在。这就是美国《独立宣言》所说的造物主或神圣天意。Thomas Jefferson和Benjamin Franklin是自然神论者，对这种自然主义上帝的信仰就这样被称呼。

科学对事物如何相互作用以及它们如何从一种形式变化为另一种形式给出了有效的描述。星系、汽车、台球和亚原子粒子以可以用数学捕捉的规律方式行动，因此可以预测，这简直令人惊叹。确实，一些物理学家和数学家——最著名的是Albert Einstein——正是因为这种“奇迹般”的事态而相信这样一个造物主。不难想象一个复杂到无法理解的宇宙。但自然神论者的上帝创造了一个不仅适宜生命存在，而且如此可预测的宇宙，其规律性可以被人类心智理解。

然而，我们徒然寻找这样一个超越自然力量的永恒存在的直接经验证据。上帝在我们的试管中不留残留物，在我们的气泡室中不留踪迹。上帝也不通过逻辑显示自己。正如天文学家-哲学家Immanuel Kant所论证的，所有证明上帝存在的证明都是有缺陷的。没有一连串无懈可击的论证从不可质疑的命题导致上帝必须存在的确定结论。（相反的情况也是如此——也无法证明没有上帝。）Wittgenstein再次简洁地表达：“上帝不在世界中显示自己。”

20世纪70年代初期，这场辩论出现了新的转折——即人择原理，这个观察发现宇宙对稳定、自我复制的生化系统极其友好。如果支配宇宙的物理常数和参数稍有不同，复杂分子以及生命就无法存在。从这个意义上说，“人择”是一个误称，因为该原理并非指人类生命；相反，它应该被称为“生物向性”或者“生物亲和性”原理。

以牛顿引力定律和库仑定律为例，后者支配着带电粒子相互吸引和排斥的方式。两个定律具有相同的形式，都表明力随任意两个粒子之间距离的平方递减。只有二次衰减项前的常数不同。有趣的是，两个相反电荷之间的吸引力必须恰好比它们相互间的引力吸引强10,000万亿万亿万倍，才能形成我们所知的生命。稍多或稍少一点我们都不会存在。另一个宇宙约束是，宇宙中所有正电荷粒子的总和必须等于所有负电荷粒子的总和；否则，电磁力将主导引力，恒星、星系和行星就无法形成。电子数量必须等于质子数量，误差不能超过万亿万亿万分之一。如果强核力

比实际情况稍强或稍弱，要么只存在氢元素，要么不存在比铁更重的元素。如果宇宙膨胀过快，质子和中子就无法结合成原子核。如果宇宙的初始膨胀稍微慢一点，构成早期宇宙的炽热混合物就会过热而无法形成原子核。简而言之，必须发生惊人数量的“巧合”才能产生一个稳定足够长时间且化学元素足够多样化的宇宙，以支持复杂的碳基生命形式。

一些人认为人择原理是同义反复的：如果宇宙对生命不友好，我们就不会存在来思考它的存在。这假设在时间或空间上存在无数个不适宜生命的平行宇宙，而我们恰好处在适宜生命的那一个中。问题是我們不了解这些其他宇宙——它们从未被观察到。也许我们生活在多元宇宙中，包含无限数量的非相互作用且不可观察的宇宙。可能如此。但假设无数个世界是一个非常强的假设，就像假设一个至高建筑师操控物理定律以促进生命形成的假说一样特设。

人择原理引发的激烈辩论没有平息的迹象。

剩下的既不是经验知识，也不是逻辑确定性，而是信念。一些人，如物理学家史蒂芬·霍金和伦纳德·姆洛迪诺夫，表达了他们的信念，认为一个尚未被证实的物理理论——M理论——将证明为什么宇宙必须以现在的方式存在。其他人认为这充其量是可疑的期票，并对不同的原理抱有信念。

## 有神论，或作为干预者的上帝

---

上帝拥有什么力量？至高存在能否影响其创造物中事件的进程？毕竟，人们祈祷时期望上帝会倾听他们——只要他们的意图纯洁、信念真诚——并会代表他们进行干预，治愈生病的孩子、稳定摇摆的婚姻或祝福新事业。如果上帝无法做到这些事情中的任何一件，那又何必呢？（我这里不关心祈祷可能带来的有益心理效果，如缓解焦虑。我追求更重要的目标。）

有神论是相信一位积极干预宇宙的活跃上帝。有神论与科学兼容吗？当宇宙之外的某种力量，不受自然法则约束，导致宇宙内部发生某事时，人们称之为奇迹。因此，问题需要重新表述为：奇迹与科学兼容吗？答案明确是否定的。

以耶稣的第一个公开行动（根据新约）为例——在迦拿的婚礼上将水变酒。这违背了一个基本原理——质能守恒。构成酒的芳香分子和醚分子必须来自某处。水分子可以转化为碳和构成酒的其他元素和分子，但这是一项需要巨大能量的核合成壮举。没有关于这种情况的报道。

每次测试这个守恒原理时，都发现它是可靠的，从极小到极大皆如此。因此，迦拿奇迹发生的可能性极小。

科学家遵循一个名为奥卡姆剃刀的演绎推理启发原理。以十四世纪英国修士和逻辑学家威廉·奥卡姆命名，它表明对于一个现象的两个同样好的解释，较简单的那个更可能是真的。更复杂的解释不如更简洁的解释可能。这不是逻辑原理，而是一个工作规则。

在重建异常事件、谋杀案或飞机坠毁时，调查人员永远无法绝对确定地确定发生了什么。但奥卡姆原理缩小了选择范围。奥卡姆剃刀四处挥砍，排除了没有明显动机、没有留下物理痕迹但辩护律师声称对谋杀负责的未知袭击者。奥卡姆剃刀来回削减，摧毁了秘密政府阴谋导致飞机坠毁的理论，因为这需要一连串不太可能的事件和许多人的积极参与。奥卡姆剃刀是一个宝贵的工具，从考虑中排除多余的实体。

至高存在将水变成酒的可能性如此离奇，可以用奥卡姆剃刀予以拒绝。更有可能的是，遵循物理定律的其他东西是原因。也许婚礼组织者在地下室发现了早已遗忘的酒瓶。或者客人带酒作为礼物。或者这个故事是为了巩固耶稣作为真正弥赛亚的声誉而编造的。记住夏洛克·福尔摩斯的建议：“当你排除了不可能的，剩下的无论多么不可能，都必须是真相。”

奇迹是不可能的。日常现实的结构编织得太紧密，不会被超自然力量撕裂。恐怕上帝是一个缺席的宇宙房东。如果我们想让事情在这里发生，我们最好自己处理。没有人会为我们做这件事。

## 启示和圣经能有所帮助吗？

---

传统上，关于超验的最重要知识来源是启示——对上帝的直接、第一手体验。扫罗在去大马士革的路上与活着的上帝相遇，使他从迫害耶稣早期信徒转变为使徒保罗，成为基督教最伟大的传教士。十七世纪的法国数学家、物理学家和哲学家布莱兹·帕斯卡尔同样以这种方式体验了上帝：在缝在他外套衬里的羊皮纸上发现了对他激烈体验的描述。来自所有宗教传统的圣人和神秘主义者的著作都包含与绝对的相遇和与宇宙合一的感觉。

如果我以这种方式体验过上帝，如果我看到了燃烧的灌木丛并感受到*Mysterium tremendum*的某种表现，我就不会写这些文字了。我不必求助于不充分的理性来弄清楚事情。

因为我只能依靠理性，我承认在考虑这种改变人生的经历的本体论（但不是心理学）有效性时持怀疑态度。作为丈夫、父亲、儿子、兄弟、朋友、恋人、同事、科学家、公民和历史的狂热读者，我仍然对受过高等教育和聪明人自欺欺人的能力感到惊讶。你和我都确信我们的动机是高尚的，我们比大多数人聪明，异性觉得我们有吸引力。

没有人能免于自欺和自我欺骗。我们都有复杂的、潜意识的防御机制，让我们保持对我们珍视的信念，尽管有相反的事实。9·11事件、伊拉克惨败和雷曼兄弟破产生动地证明了“精英”像其他任何人一样遭受这些常识失败。加州理工学院的同事、物理学家理查德·费曼说：“第一个原则是你不能欺骗自己，而你是最容易被欺骗的人。”我们病态地倾向于根据我们想要相信的来解释任何事件，这正是双盲实验协议在科学和医学中如此重要的原因。它们根除了实验者的隐藏偏见，否则这些偏见会污染结果。

鉴于这些关于人性的不舒服事实，我怀疑强烈的宗教体验，尽管无疑是真实感受到的，是否揭示了关于上帝实际存在的任何信息。人们面临的风险太高，无法保持客观。我不否认这样的经历，但我对它们的解释保持警惕。我准备好了湿毯子和理性的冷水。

在转向圣经——另一个宗教思想和教导的传统来源时，我同样持怀疑态度。数千年前生活的人们的经历和思想与我们对宇宙和我们在其中的位置的理解相关这一想法，在我看来是古怪的。《圣经》的书籍是在宇宙的真实年龄和范围甚至远未被想象之前写的，在人类和动物之间的进化联系被理解之前，在大脑被确定为心灵的所在地之前（旧约和新约都没有提到大脑一次）。

此外，观察到不同的社会和文化有截然不同的基础文本和传统，这些差异一些追随者愿意为之杀戮和死亡，这并没有增加我对这些接受的“真理”的信心。（如果我们在遥远的星球上发现文明，什么奇怪的神、仪式和信仰会加入这个万神殿？他们也能获得救赎吗？耶稣也为他们而死吗？）鉴于这个宗教市场，我应该在什么基础上选择一个而不是其他的？多年来，我像绝大多数人一样，相信我的父母所相信的。但那不是一个真正明智的选择。

《旧约》和《新约》、《古兰经》以及其他宗教文本都富有诗意、启发性，对人类持久的需求和欲望具有深刻的洞察力。它们提供了数千年来指导信徒的伦理基础。圣典一次又一次地提醒我们，每个个体都是更大整体的一部分，是信徒大家庭的一部分，也是创造物的一部分。当代文化、政治和社会生活围绕着贪婪和消费这个金牛犊转动。战争和冲突、股市崩盘、环境恶化以及水资源和石油短缺提醒我们，忽视这些基本真理会带来危险。但随着我们对宇宙了解的加深，这些神圣文本与当代世界的相关性在减弱。

让我举一个个人例子，说明本书讨论的科学见解如何在我的生活中产生了具体的影响。正如我之前解释的，许多——可能所有——动物物种都有主观的、现象性的感受。它们体验快乐和痛苦，会高兴和悲伤。鉴于这一认知，我们如何能够为了食用而在恶劣的工业条件下饲养动物，让它们远离自然栖息地？我们如何能够为了肉食而繁殖有感知能力的生物？我们如何能够将幼小的小牛锁在狭窄、封闭的箱子里，让它们无法转身或躺下，剥夺它们短暂生命

中的任何社交接触，只是为了我们能够食用它们白嫩的肉？这在今天尤其野蛮，因为营养丰富、味道好、便宜——而且更健康——的肉类替代品很容易获得。然而，对我来说，很难将这种理性的洞察付诸行动——肉的味道在我们的烹饪和味觉中根深蒂固。然后，在2004年，苏珊·布莱克莫尔(Susan Blackmore)，一位勇敢的英国心理学家，有着彩虹色的头发，为了她的一本书采访了我。我刚刚结束了一段关于小鼠意识的论述，呼吁不要像许多研究人员那样轻率地杀死小鼠，这时苏珊突然问我是否吃肉。我们默默地看着彼此一会儿，直到我叹了口气来掩饰被揭露为伪君子的尴尬。这件事真的困扰了我。

一年后，当我心爱的诺西(Nosy)去世时，我被触动去行动。在我与之生活过的六只狗中，我最喜欢这只极其聪明、好玩、强烈好奇的黑色德国牧羊犬。当她去世时，我心神不定；直到今天我仍然梦到她。那天晚上，当她在我怀中垂死时，我问自己，我怎么能为她哭泣却愉快地吃羊肉和猪肉？它们的智力和大脑与狗并没有太大不同。从那个夜晚开始，我停止了食用哺乳动物和鸟类，尽管有些不一致，我仍然食用鱼类。

十诫中没有一条教导我们避免食用有感知能力的生物的肉。没有一条指导我们照顾地球。十诫在做临终决定或处理生殖克隆方面没有帮助。我们需要一套新的诫命，适合我们的时代，正如现代动物权利运动的创始人之一、哲学家和伦理学家彼得·辛格(Peter Singer)所大力倡导的。

## 我在阿卡迪亚

---

我的成长经历让我对绝对的事物怀有渴望，并认识到神圣可以在所有事物中找到——狗的嚎叫、星空的景象、对周期表的思考，或者在大风攀登中冰冷双手的疼痛。

有时，我在这样的沉思中遇到了黑暗的一面。作为一个躺在床上的青少年，我会努力去理解永恒。时间永远持续下去是什么感觉？永远死去意味着什么？不只是死一个世纪，或一千年，不只是很长时间或真的很长时间，而是永远永远。概念艺术家罗曼·奥帕尔卡(Roman Opalka)试图探测无穷，捕捉它——一个数字接一个数字的稳定、不可理解的进展，从一到无穷，不断继续。在他生命的最后四十五年中，绘制这无尽的数字流是奥帕尔卡处理我们面前延展的无穷这一令人眩晕和恐怖概念的方式。

然而，我从未担心过自己的死亡。像许多追求极限的年轻人——无论是攀岩、摩托车赛车、金融还是战争——我不会思考结局。它真的不会发生在我身上。甚至我女儿伊丽莎白(Elisabeth)的死亡也没有让我从这种幸福的自满中惊醒。

直到我四十出头，我才真正意识到死亡也会来到我身上。我在第6章的开头页面中讲述了这个故事。一天晚上，我的潜意识反抗了。我醒来，抽象的知识变成了令人肠子打结的确定性：我真的要死了！

在接下来的几个月里，我思考着我个人湮灭的意义，面对着内心存在主义的遗忘和无意义的深渊。最终，通过某种无意识的重新校准过程，我回到了我的基本态度，即一切都应该如此。我无法用其他方式来描述：没有山顶的转变或深刻洞察的闪现，而是一种弥漫在我生活中的情感。我每天早晨醒来，发现自己身处一个充满神秘和美丽的世界。我对这一切的奇迹深感感激。

在这里，我是一个高度组织化的质量和能量模式，七十亿人中的一个，在任何客观的世界记录中都微不足道。不久之后，我将不复存在。对宇宙而言，我是什么？几乎什么都不是。然而，死亡的确定性使我的生命更有意义。我对生活的喜悦，对孩子们的爱，对狗的喜爱，跑步和攀登，书籍和音乐，钴蓝色的天空，都是有意义的因为我终将结束。这就是应该的样子。我不知道之后会发生什么，“如果”之后”在通常意义上存在的话，但无论是什么，我从内心深处知道一切都是最好的安排。

这种情感与我整体阳光乐观的性格相关，这主要由遗传因素决定，并在我成长的良好环境中得到放大。我无法为此邀功。

Edith是那个坚强、沉着、负责任的女人，她让我在近三十年里保持脚踏实地。她让我能够作为教授和科学家充分发展。她将自己的事业搁置多年，将我们的孩子培养成今天健康、聪明、足智多谋、负责任和美丽的成年人。这意味着我可以给他们读书唱歌哄他们入睡，到国外旅行，与他们一起徒步、露营、漂流，帮助他们做作业和学校项目，并享受作为父亲的所有其他乐趣，而不会对我的职业生涯造成任何重大牺牲。

我们还享受着数量不定的大型、毛茸茸、活泼的狗狗们的陪伴——Trixie、Nosy、Bella和Falko。除了孩子，它们是生活中最美好的事物。

与少数几位同事一起，我创办并指导了两个科学夏季学校：一个是在马萨诸塞州伍兹霍尔海洋生物实验室举办的计算神经科学学校，位于大西洋海岸；另一个是在科罗拉多州泰卢赖德举办的神经形态工程学校（工程师可以从神经生物学中学到什么），位于落基山脉。两个学校都仍然很受欢迎。每年夏天，我们一家在这些美丽的地方度过紧张的四周。这是我生命中最快乐的时光。

当我的儿子和女儿离家上大学时，那些平静快乐的日子结束了。我比其他任何事情都更想念他们。为了填补这个巨大的空虚，让我的精力得到疏导，我开始在内华达山脉和约塞米蒂山谷攀岩，在当地山区进行长距离越野跑，在死亡谷跑马拉松，等等——任何有挑战性的活动来对抗我日益增长的不安。我患上了严重的空巢综合征。

然后Francis离开了我的生活。当他的肿瘤科医生打电话确认他的结肠癌恶化复发时，我和他一起在他家的书房里。他望向远方一两分钟，然后回到我们的阅读中。这个诊断在午餐时与Odile讨论了，但那天就到此为止了。当然，那一夜我没有读到他内心的黑暗想法。但我确实记得早先的一次谈话，他向我坦白说，自己不太遥远的终点让他悲伤，但他决心不要通过无果的反思和沉思，也不要通过追求高风险的实验性治疗来浪费剩余的时间。在这里我看到他践行了这个决心。多么强大的心理控制！多么镇定！

几个月后，化疗的恶心副作用未能阻止癌症扩散，他在隔壁房间放下电话，从我身边拖着脚步走向浴室。当他回来继续电话交谈时，他在经过时干巴巴地说道：“现在我可以真正告诉他们，他们的想法让我呕吐了。”（有人试图说服Francis同意制作他的摇头娃娃。）面对不可避免的结局，Francis给了我一张他的真人大小黑白照片，就像我认识的他那样。坐在藤椅上，他讽刺地凝视着观者，眼中闪烁着光芒。签名写着“给Christof——Francis——看着你”，它在我的办公室里守护着我。

2004年夏天，Francis在去医院的路上给我打电话，告诉我他对我们最后一份关于屏状核功能的手稿的修改会延迟。屏状核是皮层下方的片状结构。然而他继续工作，在诊所里向秘书口述修正。两天后他去世了。Odile讲述了他在临终之际如何产生幻觉，与我争论快速放电的屏状核神经元及其与意识的联系，直到最后一刻还是个科学家。他是我的导师，我亲密的知识伙伴，也是我的英雄，因为他面对衰老和死亡时毫不畏缩的方式。他的离世在我的生活中留下了一个巨大的空洞。

我父亲在第三个千年的开始几周去世了，让我失去了可以寻求指导和支持的长者导师。矛盾的是，《意识的探索》一书的成功出版加剧了我的不安。我多年来为这个目标努力工作，这个目标有时显得非常遥远。现在比赛结束了，我感到无精打采，缺乏明确的使命。我的生活需要安纳普尔纳峰这样的挑战。

由于这些累积的背离，我与妻子渐行渐远，最终离开了她。说出来很容易，但这短短几个字涵盖了一段长时间内无法用纸笔描述的痛苦、困扰、疼痛和愤怒。（观看英格玛·伯格曼的阴暗杰作《婚姻场景》，你就能理解我的意思。）我经历了一场灼痛的危机，亲身体验到无意识塑造情感和行为的力量，这种力量逃脱了意识洞察。一旦这些力量被释放，我就无法控制它们。或者说我不愿意控制它们。但丁将“让理性屈服于欲望”的罪人打入地狱第二层是有原因的。这无疑是我人生的阴暗低谷。然而某种东西在驱使我继续前进。

斯宾诺莎创造了一个美妙的表达，*sub specie aeternitatis*，字面意思是“在永恒的形式下”。这是一个遥远的视角。从银河系中央黑洞的远上方俯视银河系。你看到一个由数千亿颗恒星组成的旋转圆盘，其中许多被微小的黑暗伴星环绕。这些行星中的一些孕育着生命。在其中一颗行星上，半智能的、暴力的、群居的灵长类疯狂地结合又分离。它们赋予这种狂热的蚁穴活动以巨大的宇宙重要性。与壮丽的银河系轮盘完成一次旋转所需的时间相比，这些配对只持续一眨眼的功夫、萤火虫的一闪、箭矢的一飞。

在深时间的这种天体之光下，我的痛苦开始变得微不足道。我的磨难并非毫无意义——我不是虚无主义者——但它们不应该也不会压垮我的生活。失去了我的中心太阳，我现在是一颗孤独的行星，在星际间的寂静空间中漫游。我正在慢慢重新获得一些内心平静，那种心灵的宁静，伊壁鸠鲁派称之为*ataraxia*(不动心)，我曾经长时间拥有的平静。

为了正视我的行为，我研究了科学对自主行为和自由意志的认知，这解释了第7章的起源。我从阅读中得到的是，我没有感觉中的那么自由。无数事件和倾向影响着我。然而我不能躲在生物冲动或匿名社会力量后面。我必须表现得好像“我”完全负责，否则这个词以及善恶行为的概念将失去所有意义。

在我危机的某个夜晚，我一边看奇幻动作电影《高地人》一边喝完了一瓶Barolo红酒，感到需要某种象征性的姿态。午夜时分，我决定跑到威尔逊山顶，那里海拔比帕萨迪纳高出五千多英尺。在戴着头灯跌跌撞撞地走了一个小时并感到恶心后，我意识到自己很愚蠢，于是转身回去——但在此之前，我向黑暗中大喊了诗歌《不可征服》的结尾句：“我是我命运的主人，我是我灵魂的船长。”这或许是我对自由意志问题立场的过度热情表达：无论好坏，我都是我生活的主要演员。

现在你已经到达了这个自传性章节的尾声，我可以承认此时已经显而易见的事情。我被驱使写这本书有三重原因——描述我寻找意识物质根源的探索，直面我个人的失败，并为我寻找宇宙统一观及我在其中作用的探索画下句号，这种观点既要体现偶然性也要体现必然性。

## 亮明我的立场

---

这是我故事的结尾。我乐观地认为科学即将完全理解心身问题。套用《哥林多前书》的话说：“现在我们是对着镜子观看，模糊不清，但到那时，我们就要面对面了。”

我确实相信某种深刻而基本的组织原理创造了宇宙并为某个我无法理解的目的而使其运动。我从小称这个实体为上帝。它更接近斯宾诺莎的上帝，而不是米开朗基罗绘画中的上帝。神秘主义者Angelus Silesius是笛卡尔的同时代人，他捕捉到了自因第一推动者的悖论本质：“*Gott ist ein lauter Nichts, ihn röhrt kein Nun noch Hier*”（上帝是一个透明的虚无，没有现在和此地能触及他）。

先驱一代恒星必须以壮观的超新星方式死亡，才能为太空播撒创造第二幕所必需的重元素——在围绕一颗年轻恒星以恰当距离运行的岩石行星上，自我复制化学物质袋的崛起。自然选择的竞争压力引发了创造的第三幕——被赋予感知力、具有主观状态的生物的兴起。随着它们神经系统的复杂性增长到惊人的程度，这些生物中的一些进化出了思考自身以及沉思周围壮丽美好而又可怕残酷世界的能力。

在时间的广阔循环中，有感知生命的兴起是不可避免的。德日进(Teilhard de Chardin)的观点是正确的，宇宙中的岛屿——如果不是整个宇宙——正在向更大的复杂性和自我认知演进。我并不是说地球必须孕育生命，或者双足大脑的灵长类必须在非洲草原上行走。但我确实相信物理定律绝大部分上有利于意识的出现。宇宙是一项正在进行的工作。这样的信念会招致许多生物学家和哲学家的严厉批评，但来自宇宙学、生物学和历史学的证据是令人信服的。

宗教传统鼓励我们向时间长河中的同路人伸出援手。与大多数世俗意识形态相比，宗教更加强调人与人之间的共同纽带：爱邻如己。宗教情感通过音乐、文学、建筑和视觉艺术表达，展现了人类最美好的品质。然而，从整体上看，它们对于理解我们存在之谜只能提供有限的帮助。唯一确定的答案来自科学。从智力和伦理角度来看，我最感兴趣的是佛教的某些流派。但这是另一本书的话题。

失去宗教信仰让我感到悲伤，就像永远离开童年温暖的家，那里充满温馨的光芒和美好的回忆。当我走进高耸的大教堂或聆听巴赫的《圣马太受难曲》时，仍然会产生敬畏之情。我也无法摆脱盛大弥撒的情感魅力、壮丽和仪式感。但失去信仰是成长、成熟以及看清世界真相的不可避免的一部分。

我被抛入宇宙中，这是一个光荣、奇怪、可怕且常常孤独的地方。我努力通过它嘈杂的表现形式——人类、狗、树木、山脉和星辰——来辨别永恒的天体音乐。

说到底，我怀着深刻而持久的惊奇感。跨越两千多年的历史，住在犹太沙漠小社区中的《死海古卷》无名抄写员很好地表达了这种感情。让他的赞美诗结束我的书：

我走在无边的高地上

知道有希望存在

你用尘土塑造的生命

将与永恒的事物为伴。

## 注释

---

2004年，我出版了《意识的探求》。这本书总结了弗朗西斯·克里克和我采用的方法，描述了对意识至关重要的神经生物学回路和心理过程。它包含数百个脚注，分布在四百页单倍行距的篇幅中，以及近千个学术文献参考。《忏悔录》则更为轻松。如果你想了解本书中提到的技术和实验，请参考《探求》、Laureys和Tononi编辑的卷册(2009)、众包且及时更新的维基百科，或以下简要注释。我还为《科学美国人·心智》杂志定期撰写专栏，名为《意识再现》，介绍当代意识研究。

# 第一章

---

Ridley(2006)撰写的弗朗西斯·克里克简短传记很好地描绘了弗朗西斯的性格。Olby(2009)的巨著更加权威，详细阐述了弗朗西斯的科学贡献。Olby倒数第二章涵盖了弗朗西斯与我的合作。

关于“困难问题”一词的起源和含义，请参考Chalmers(1996)。

## 第二章

---

Koch和Segev(2000)总结了单个神经元的生物物理学。

Mann和Paulsen(2010)描述了由数万个神经元产生的局部场电位对这些神经细胞放电的影响。Anastassiou和Perin的实验(Anastassiou et al., 2011)直接显示了神经元的放电如何通过微弱的外部电场同步。

经济学也不遵守守恒定律：一家公司今天可能价值数十亿美元，第二天却只值数百万美元，尽管实地情况没有任何变化。同样的人在同样的建筑中使用同样的基础设施工作。钱去哪里了？答案是市场对公司未来的信念，即预期，突然蒸发了，随之而去的是市场的估值。与能量不同，金钱可以被创造和毁灭。

## 第三章

---

尽管有些过时，弗朗西斯对意识及其生物学基础的简洁介绍仍然是杰作(Crick, 1995)。

Craig Venter合成生物有机体的工作在Gibson等人(2010)中有描述。

Tyndall的引文来自他1868年在英国科学促进会数学和物理学分会上的主席演讲，题为“科学唯物主义”(Tyndall, 1901)。

Tannhäuser Gate的引文当然来自《银翼杀手》的最后一幕。这部由Ridley Scott执导的电影是史上最伟大的科幻电影。它大致基于几行后提到的《仿生人会梦见电子羊吗？》。这部由Philip K. Dick于1968年写成的杰出小说预见了“恐怖谷”效应，即机器人或计算机动画制作的几乎但不完全完美的人类复制品会引起厌恶感的心理观察。

Huxley的引文来自他1884年向十六年前聆听Tyndall演讲的同一个英国协会发表的一次杰出演讲。Huxley质疑笛卡尔认为动物只是机器或自动装置、缺乏意识感觉的信念。他基于生物连续性原因假设某些动物与人类共享意识的某些方面，但对意识的功能感到困惑。

动物意识的最佳入门书是Dawkins(1998)的简短著作《只通过我们的眼睛？》。或者参看百科全书式的Griffin(2001)。Edelman和Seth(2009)专注于鸟类和头足类动物的意识。

现代简明的神经解剖学入门是Swanson(2012)。

Krakauer的引文来自他1990年关于登山的优秀随笔集。

在磁共振扫描仪中观看《好人、坏人与丑人》的研究报告见Hasson等人(2004)。

与自我相关区域的失活在Goldberg, Harel和Malach(2006)中有论述。

## 第四章

---

我们对意识神经相关物的思考在不断发展（Crick and Koch, 1990, 1995, 1998, 2003）。哲学家David Chalmers在Chalmers (2000)中雄辩地总结了意识相关物概念背后的形而上学和概念假设（另见Block, 1996的文章）。Tononi and Koch (2008)提供了相关实验研究的最新进展。

Rauschecker和同事们（2011）在神经外科环境中通过电刺激视觉皮层表面来引发视觉运动知觉。

Macknik等人（2008）是最早指出舞台魔术为心理学和神经科学提供丰富洞察来源的研究者之一。

用于在几分之一分钟或更长时间内隐藏物体的连续闪光抑制(CFS)技术由Tsuchiya and Koch (2005)开发。

使用CFS来探索无意识心智的一个更新且优雅的例子是Mudrik等人（2011）。掩蔽方法的概要可以在Kim and Blake (2005)中找到。Jiang等人（2006）要求志愿者观看看不见的男性和女性裸体图片，Haynes and Rees (2005)对观看看不见光栅的志愿者大脑进行了fMRI研究。视觉词形区域及其与阅读的关系由McCandliss, Cohen, and Dehaene (2003)描述。

Logothetis (2008)很好地总结了fMRI在解读潜在的不同神经元反应方面的前景和局限性。他在双眼竞争神经基础方面的经典研究在Logothetis (1998)和Leopold and Logothetis (1999)中得到了综述。

一个日德联合研究团队（Watanabe et al. 2011）在人类初级视觉皮层中分离了视觉注意力和意识——通过刺激可见性来评估。受试者是否看到他们正在观看的内容对V1中的血流动力学信号几乎没有影响，而注意力则强烈调节该信号。

我在这里讨论的关于初级视觉、听觉和体感觉皮层中神经活动与意识视觉分离的大部分研究在《Quest》第6章中有参考。

Boly等人（2011）为严重残疾患者维持意识需要从皮层前部到后部的功能连接提供了支持。另见Crick and Koch (1995)中的图1。

哲学家Ned Block在注意力与意识关系的辩论中具有巨大影响力（Block, 2007）。van Boxtel, Tsuchiya, and Koch (2010)综述了将选择性视觉注意与视觉意识区分开来的无数实验。

## 第5章

---

Gallant等人 (2000) 描述了患者A.R.。

我最喜欢的神经学家Oliver Sacks (2011)的最新著作包含了对面盲症和其他神经缺陷的生动描述。Sacks是人类状况的热心见证者。通过研究人们如何应对疾病，他展示了他们和我们如何能够获得关于生活的智慧。

皮层电生理学先驱之一Semir Zeki在Zeki (2001)中创造了*essential node*一词。

失忆症患者H.M.的科学遗产在Squire (2009)中得到了总结。

Quijan Quiroga等人 (2005, 2009) 在人类内侧颞叶中发现了概念细胞(concept cells)，这些细胞对名人或熟悉人物的图片、文本和声音做出反应。这些细胞与所谓的祖母神经元密切相关 (Quijan Quiroga et al. 2008)。Cerf等人 (2010) 利用计算机反馈使患者能够通过思想控制这些神经元。

Owen等人 (2006) 和Monti等人 (2010) 使用磁共振扫描仪在少数植物状态患者中检测到意识时引起了国际关注。

Parvizi and Damasio (2001)探索了四十多个脑干核团与意识之间的联系。

Laureys (2005)描绘了我们关于死亡、大脑和意识之间联系的惊人动态观念。Schiff (2010)是一位专门研究大脑创伤后意识恢复的神经学家。

## 第6章

---

关于无意识的大量无稽之谈仍在继续传播。然而，在严格控制条件下对非意识进行的扎实实证研究已经复苏。Hassin, Uleman, and Bargh (2006) 提供了一些后弗洛伊德时代最佳研究的专著，而 Berlin (2011) 专注于无意识神经生物学的已知和未知。

Jeannerod 的研究在他 1997 年的著作中有所描述。

Varki (2009) 的信件将否认死亡机制视为可能的进化驱动因素。

证明你的眼睛能够分辨比你能看到的更精细细节的实验由 Bridgeman 等人 (1979) 和 Goodale 等人 (1986) 进行。Logan and Crump (2009) 表明打字员的手知道打字员的大脑不知道的事情。意识感知和无意识视觉运动动作的双视觉流理论在 Goodale and Milner (2004) 中得到了回顾。

麻省理工学院历史学家 John Dower (2010) 详尽分析了珍珠港事件和 9·11 事件之间的结构相似性和差异性，以及机构决策的其他病态。

启动实验取自 Bargh, Chen, and Burrows (1996)。Johansson 等人 (2005) 要求男性和女性选择两名女性中哪一个看起来更有吸引力，然后在大多数志愿者没有注意到的情况下交换了她们的图像。

## 第7章

---

对自由意志物理学的反思受到我阅读Sompolinksy (2005)的影响。

Sussman and Wisdom (1988)证明冥王星的轨道是混沌的。

果蝇表现出真正随机性的程度由Maye等人 (2007) 研究。

Turner (1999) 将早期宇宙中的量子涨落与今日天空中星系的分布联系起来。物理学家Jordan (1938) 提出了仍然流行的——在某些圈子中——量子放大理论，将基本粒子物理学与自由意志联系起来。Koch和Hepp (2011) 讨论了量子力学与大脑的可能相关性。Collini等人 (2010) 提供了室温下光合蛋白中电子相干性的有力证据。

关于在意志行为感觉之前大脑活动的最初报告由Libet及其同事 (1983) 发表。Soon及其同事 (2008) 对原始实验的脑成像变体引起了公众的关注。

关于自由意志神经心理学的文献越来越多 (Haggard, 2008)。Murphy、Ellis和O’ Connor (2009) 编辑了一本书，试图解决传统自由意志观念（植根于神学和日常经验）与现代心理学和生物学观念之间的一些紧张关系。

关于弓形虫如何搭便车进入老鼠大脑，迫使其改变行为使其更容易被猫吃掉的病态故事，由Vyas等人 (2007) 讲述。Lafferty (2006) 分析了这种寄生虫的脑部感染对人类文化可能产生的影响。

Wegner (2003) 出色地描述了正常生活和病理条件下自主行为的心理学。

两项神经外科研究通过电刺激大脑来启动”自主”行为 (Fried等人, 1991; Desmurget等人, 2009)。

## 第八章

---

Baars (2002) 的书描述了他的全局工作空间意识模型。Dehaene和Changeux (2011) 回顾了支持全局工作空间模型神经实现的成像和生理数据。

Chalmers关于信息论和意识的观点在他1996年著作的附录中有概述。

了解Tononi理论和思想最简单的入门是他2008年的宣言。我热烈推荐Giulio本人在其*Phi* (2012) 中对相关事实和理论进行的高度文学性和小说化处理，其中Francis Crick、Alan Turing和Galileo Galilei进行了一次穿越巴洛克的发现之旅。关于实际的数学计算，请参考Balduzzi和Tononi (2008, 2009)。Barrett和Seth (2011) 推导出计算整合信息的启发式方法。

请参考我的*Quest*第17章了解裂脑患者的背景。Lem (1987) 对裂脑半球生活进行了小说化处理。

在人脑的860亿神经元中，令人惊讶的是其中690亿在小脑，160亿在皮质 (Herculano-Houzel, 2009)。也就是说，大约每五个脑细胞中有四个是小脑颗粒神经元，具有典型的四个短粗树突。剩余的10亿神经元分布在丘脑、基底神经节、中脑、脑干和其他地方。然而值得注意的是，先天性小脑缺失（罕见现象）或因中风或其他创伤失去部分小脑的患者，认知缺陷却很少。主要缺陷是运动失调、言语不清和步态不稳 (Lemon和Edgley, 2010)。

进化*animats*——生活在计算机内部的生物简化模型——经过数万代后证明，它们对模拟环境的适应性越强，其整合信息就越高 (Edlund及其同事, 2011)。

Koch和Tononi (2008) 根据意识的整合信息理论讨论了机器意识的前景。我们在Koch和Tononi (2011) 中提出了一个基于图像的测试，用于确定计算机有意识地看到图片意味着什么。

## 第九章

---

使用TMS-EEG测试睡眠中意识心灵崩溃的组合方法由Massimini等人 (2005) 描述。将此方法扩展到持续植物状态和最小意识患者的研究见Rosanova等人 (2012)。

神经解剖学家仍然不知道最大的哺乳动物蓝鲸，拥有17磅（8000克）的大脑，是否比人类拥有更多的神经元。更大的大脑并不一定意味着更多的神经元；不过，如果鲸类和大象拥有比人类更多的脑细胞，那将是相当震惊的。关于大脑大小和神经元数量的全面讨论，请参见Herculano-Houzel (2009)。

神经元细胞类型的简短入门是Masland (2004)。

光遗传学的应用正在爆炸式增长，数百个实验室在他们选择的位置和时间操纵基因可识别的细胞群体。考虑到开创性使用channelrhodopsin-2驱动神经活动的原始论文是在Boyden及其同事 (2005) 发表的，这是令人瞩目的。在我看来，将神经回路与小鼠行为因果关系联系起来的三个最优雅的光遗传学实验是Adamantidis等人 (2007；这是我描述的orexin研究)、Gradinaru等人 (2009) 和Lin等人 (2011)。当前技术的纲要可在Gradinaru等人 (2010) 中找到。

小鼠的Allen Brain Atlas在Lein等人 (2007) 中有全面描述，可在线查找。

Dehaene等人 (2001) 测量了志愿者观看可见和不可见单词时的fMRI反应。

## 第十章

---

我令人尴尬的采访可在Blackmore (2006) 中找到。

关于科学与宗教关系的文献浩如烟海。我发现著名自由派神学家Hans Küng 2008年的书很有帮助。

挂在我办公室里的Francis肖像由Mariana Cook为她的*Faces of Science* (2005) 收藏拍摄。

Francis字面上一直工作在claustrum论文上直到生命的最后一息。该论文作为Crick和Koch (2005) 发表。

哲学家Peter Singer 1994年关于传统伦理学在解决现代生死挑战方面不足的书非常有见地。

尽管结尾略显夸张，我书中的任何内容都不应被理解为过于严肃地对待我或我的生活。在Ty将Ramón y Cajal的皮层微回路纹在我左臂上的纹身店里，墙上贴着一句格言，可以作为另一种结尾：

生命的旅程不是要安全抵达坟墓，保持身体完好无损，而是要侧身滑入，彻底精疲力尽，大喊：“天啊...多么精彩的旅程！”

## 参考文献

---

- Adamantidis, A. R., Zhang, F., Aravanis, A. M., Deisseroth, K., & de Lecea, L. (2007). Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature*, *450*, 420 – 424.
- Anastassiou, C. A., Perin, R., Markram, H., & Koch, C. (2011). Ephaptic coupling of cortical neurons. *Nature Neuroscience* *14*, 217 – 223.
- Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*, 47 – 52.
- Balduzzi, D., & Tononi, G. (2008). Integrated information in discrete dynamical systems: Motivation and theoretical framework. *PLoS Computational Biology*, *4*, e1000091.
- Balduzzi, D., & Tononi, G. (2009). Qualia: The geometry of integrated information. *PLoS Computational Biology*, *5*, e1000462.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71*, 230 – 244.
- Barrett, A. B., & Seth, A. K. (2011). Practical measures of integrated information for time-series data. *PLoS Computational Biology*, *7*, e1001052.
- Berlin, H. A. (2011). The neural basis of the dynamic unconscious. *Neuro-psychoanalysis*, *13*, 5 – 31.
- Blackmore, S. (2006). *Conversations on Consciousness: What the Best Minds Think about the Brain, Free Will, and What It Means to Be Human*. New York: Oxford University Press.
- Block, N. (1996). How can we find the neural correlate of consciousness? *Trends in Neurosciences*, *19*, 456 – 459.
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*, 481 – 499, discussion 499 – 548.
- Boly, M., & Associates. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science*, *332*, 858 – 862.
- Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, *8*, 1263 – 1268.
- Bridgeman, B., Lewis, S., Heit, G., & Nagle, M. (1979). Relation between cognitive and motor-oriented systems of visual position perception. *Journal of Experimental Psychology. Human Perception and Performance*, *5*, 692 – 700.
- Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R., Koch, C., et al. (2010). On-line, voluntary control of human temporal lobe neurons. *Nature*, *467*, 1104 – 1108.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Chalmers, D. J. (2000). What is a neural correlate of consciousness? In T. Metzinger (Ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (pp. 17 – 40). Cambridge, MA: MIT Press.

Cook, M. (2005). *Faces of Science: Portraits*. New York: W.W. Norton.

Collini, E., Wong, C. Y., Wilk, K. E., Curmi, P. M. G., Brumer, P., & Schoes, G. D. (2010). Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. *Nature*, 463, 644 – 647.

Crick, F. C. (1995). *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Scribner.

Crick, F. C., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, 2, 263 – 275.

Crick, F. C., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, 375, 121 – 123.

Crick, F. C., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8, 97 – 107.

Crick, F. C., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience*, 6, 119 – 126.

Crick, F. C., & Koch, C. (2005). What is the function of the claustrum? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360, 1271 – 1279.

Dawkins, M. S. (1998). *Through Our Eyes Only? The Search for Animal Consciousness*. New York: Oxford University Press.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70, 200 – 227.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4, 752 – 758.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulations in humans. *Science*, 324, 811 – 813.

Dower, J. W. (2010). *Cultures of War: Pearl Harbor, Hiroshima, 9 – 11, Iraq*. New York: W.W. Norton.

Edelman, D. B., & Seth, A. K. (2009). Animal consciousness: A synthetic approach. *Trends in Neurosciences*, 32, 476 – 484.

Edlund, J.A., Chaumont, N., Hintze, A., Koch, C., Tononi, G., and Adami, C. (2011). Integrated information increases with fitness in the simulated evolution of autonomous agents. *PLoS Computational Biology*, 7(10): e1002236.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., et al. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience*, 11, 3656 – 3666.

Gallant, J. L., Shoup, R. E., & Mazer, J. A. (2000). A human extrastriate area functionally homologous to macaque V4. *Neuron*, 27, 227 – 235.

Gibson, D. G., & Associates. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science*, 329, 52 – 56.

Goldberg, I. I., Harel, M., & Malach, R. (2006). 当大脑失去自我：感觉运动处理过程中的前额叶失活。 *Neuron*, 50, 329 – 339。

- Goodale, M. A., & Milner, A. D. (2004). 视而不见：意识与无意识视觉的探索。英国牛津：牛津大学出版社。
- Goodale, M. A., Périsson, D., & Prablanc, C. (1986). 视觉引导伸手动作中的大幅调整不依赖于手的视觉或目标位移的感知。*Nature*, 320, 748 – 750。
- Gradinaru, V., Mogri, M., Thompson, K. R., Henderson, J. M., & Deisseroth, K. (2009). 帕金森神经回路的光学解构。*Science*, 324, 354 – 359。
- Gradinaru, V., Zhang, F., Ramakrishnan, C., Mattis, J., Prakash, R., Diester, I., et al. (2010). 用于多样化和扩展optogenetics的分子和细胞方法。*Cell*, 141, 154 – 165。
- Griffin, D. R. (2001). 动物心智：超越认知到意识。伊利诺伊州芝加哥：芝加哥大学出版社。
- Haggard, P. (2008). 人类意志：走向意志的神经科学。*Nature Reviews Neuroscience*, 9, 934 – 946。
- Hassin, R. R., Uleman, J. S., & Bargh, J. A. (Eds.). (2006). 新的无意识。英国牛津：牛津大学出版社。
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). 自然视觉期间皮质活动的跨被试同步化。*Science*, 303, 1634 – 1640。
- Haynes, J. D., & Rees, G. (2005). 从人类初级视觉皮质的活动中预测不可见刺激的方向。*Nature Neuroscience*, 8, 686 – 691。
- Herculano-Houzel, S. (2009). 数字中的人脑：一个线性放大的灵长类大脑。*Frontiers in Human Neuroscience*, 3, 1 – 11。
- Jeannerod, M. (1997). 动作的认知神经科学。英国牛津：Blackwell出版社。
- Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). 不可见图像的性别和性取向依赖性空间注意效应。*Proceedings of the National Academy of Sciences of the United States of America*, 103, 17048 – 17052。
- Johansson, P., Hall, L., Sikström, S., & Olsson, A. (2005). 在简单决策任务中未能检测到意图与结果之间的不匹配。*Science*, 310, 116 – 119。
- Jordan, P. (1938). 有机体的Verstärkertheorie（增强理论）的现状。*Naturwissenschaften*, 33, 537 – 545。
- Kim, C. Y., & Blake, R. (2005). 心理物理学魔术：让可见的变为不可见。*Trends in Cognitive Sciences*, 9, 381 – 388。
- Koch, C. (2004). 意识的探索：神经生物学方法。科罗拉多州Englewood：Roberts & Company出版社。
- Koch, C., & Hepp, K. (2011). 量子力学与高级大脑功能的关系：来自量子计算和神经生物学的经验。收录于R. Y. Chiao, M. L. Cohen, A. J. Leggett, W. D. Phillips, & C. L. Harper, Jr. (Eds.)所著的奇妙之光：物理学、宇宙学和意识的新光（第584 – 600页）。纽约：剑桥大学出版社。
- Koch, C., & Segev, I. (2000). 单个神经元及其在信息处理中的作用。*Nature Neuroscience*, 3, 1171 – 1177。
- Koch, C., & Tononi, G. (2008). 机器能有意识吗？*IEEE Spectrum*, 45, 54 – 59。
- Koch, C., & Tononi, G. (2011). 意识测试。*Scientific American*, 304 (6月), 44 – 47。

- Krakauer, J. (1990). 艾格尔之梦。纽约：Lyons & Burford出版社。
- Küng, H. (2008). 万物的开端：科学与宗教。英国剑桥：Wm. B. Eerdmans出版社。
- Lafferty, K. D. (2006). 常见的大脑寄生虫弓形虫能影响人类文化吗？*Proceedings. Biological Sciences / The Royal Society*, 273, 2749 – 2755。
- Laureys, S. (2005). 死亡、无意识和大脑。*Nature Reviews. Neuroscience*, 6, 899 – 909。
- Laureys, S., & Tononi, G. (Eds.). (2009). 意识的神经学。纽约：Elsevier出版社。
- Lein, E. S., & Associates. (2007). 成年小鼠大脑基因表达的全基因组图谱。*Nature*, 445, 168 – 176。
- Lem, S. (1987). 地球和平。圣地亚哥：Harcourt出版社。
- Lemon, R. N., & Edgley, S. A. (2010). 没有小脑的生活。*Brain*, 133, 652 – 654。
- Leopold, D. A., & Logothetis, N. K. (1999). 多稳态现象：感知中的视角变化。*Trends in Cognitive Sciences*, 3, 254 – 264。
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). 有意识行动意图的时间与大脑活动开始时间（准备电位）的关系。自由意志行为的无意识启动。*Brain*, 106, 623 – 642。
- Lin, D., Boyle, M. P., Dollar, P., Lee, H., Lein, E. S., Perona, P., et al. (2011). 小鼠下丘脑攻击位点的功能识别。*Nature*, 470, 221 – 226。
- Logan, G. D., & Crump, M. J. C. (2009). 左手不知道右手在做什么：在熟练打字中对手的注意的破坏性影响。*Psychological Science*, 20, 1296 – 1300。
- Logothetis, N. K. (1998). 单个单元和有意识视觉。*Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 353, 1801 – 1818。
- Logothetis, N. K. (2008). 我们能用fMRI做什么和不能做什么。*Nature*, 453, 869 – 878。
- Macknik, S. L., King, M., Randi, J., Robbins, A., Teller, J. T., & Martinez-Conde, S. (2008). 舞台魔术中的注意和意识：将戏法转化为研究。*Nature Reviews. Neuroscience*, 9, 871 – 879。
- Mann, E. O., & Paulsen, O. (2010). 局部场电位振荡作为皮质独白。*Neuron*, 67, 3 – 5。
- Masland, R. H. (2004) 神经元细胞类型。*Current Biology*, 14(13), R497 – 500。
- Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). 睡眠期间皮质有效连通性的破坏。*Science*, 309, 2228 – 2232。
- Maye, A., Hsieh, C.-H., Sugihara, G., & Brembs, B. (2007). 自发行为中的秩序。*PLoS ONE*, 2, e443。
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus.*Trends in Cognitive Sciences*, 7, 293 – 299.

- Monti, M. M., & Associates. (2010). Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine*, 362, 579 – 589.
- Mudrik, L., Breska, A., Lamy, D., and Deouell, L.Y. (2011). Integration without awareness: Expanding the limits of unconscious processing. *Psychological Sciences*, 22, 764 – 770.
- Murphy, N., Ellis, G. F., & O’ Connor, T. (Eds.). (2009). *Downward Causation and the Neurobiology of Free Will*. Berlin: Springer.
- Olby, R. (2009). *Francis Crick: Hunter of Life’s Secrets*. New York: Cold Spring Harbor Press.
- Owen, A. M., & Associates. (2006). Detecting awareness in the vegetative state. *Science*, 313, 1402.
- Parvizi, J., & Damasio, A. R. (2001). Consciousness and the brainstem. *Cognition*, 79, 135 – 160.
- Quian Quiroga, R., Kraskov, A., Koch, C., & Fried, I. (2009). Explicit encoding of multimodal percepts by single neurons in the human brain. *Current Biology*, 19, 1 – 6.
- Quian Quiroga, R., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not “Grandmother-cell” coding in the medial temporal lobe. *Trends in Cognitive Science*, 12, 87 – 89.
- Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435, 1102 – 1107.
- Rauschecker, A. M., Dastjerdi, M., Weiner, K. S., Witthoft, N., Chen, J., Selimbeyoglu, A., & Parvizi, J. (2011). Illusions of visual motion elicited by electrical stimulation of human MT complex. *PLoS ONE* 6(7), e21798.
- Ridley, M. (2006). *Francis Crick: Discoverer of the Genetic Code*. New York: HarperCollins.
- Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A.G., Bruno, M.-A., Mariotti, M., Boveroux, P., Tononi, G., Laureys, S., & Massimini, M. (2012) Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain*, in press.
- Sacks, O. (2011). *The Mind’s Eye*. New York: Knopf.
- Schiff, N. D. (2010). Recovery of consciousness after brain injury. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences*, 4th ed. (pp. 1123 – 1136). Cambridge, MA: MIT Press.
- Singer, P. (1994). *Rethinking Life and Death: The Collapse of our Traditional Ethics*. New York: St. Martin’s Griffin.
- Sompolinsky, H. (2005). A scientific perspective on human choice. In Y. Berger & D. Shatz (Eds.), *Judaism, Science, and Moral Responsibility* (pp. 13 – 44). Lanham, MD: Rowman & Littlefield.
- Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11, 543 – 545.
- Squire, L. R. (2009). The legacy of patient H.M. for neuroscience. *Neuron*, 61, 6 – 9.

- Sussman, G. J., & Wisdom, J. (1988). Numerical evidence that the motion of Pluto is chaotic. *Science*, 241, 433 – 437.
- Swanson, L. W. (2012). *Brain Architecture: Understanding the Basic Plan*, 2nd edition. New York: Oxford University Press.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215, 216 – 242.
- Tononi, G. (2012). *PHI: A Voyage from the Brain to the Soul*. New York: Pantheon Books.
- Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: An update. *Annals of the New York Academy of Sciences*, 1124, 239 – 261.
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8, 1096 – 1101.
- Turner, M. S. (1999). Large-scale structure from quantum fluctuations in the early universe. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 357(1750), 7 – 20.
- Tyndall, J. (1901). *Fragments of Science* (Vol. 2). New York: P.F. Collier and Son.
- van Boxtel, J. A., Tsuchiya, N., & Koch, C. (2010). Consciousness and attention: On sufficiency and necessity. *Frontiers in Consciousness Research*, 1, 1 – 13.
- Varki, A. (2009). Human uniqueness and the denial of death. *Nature*, 460, 684.
- Vyas, A., Kim, S.-K., Giacomini, N., Boothroyd, J. C., & Sapolsky, R. M. (2007). Behavioral changes induced by *Toxoplasma* infection of rodents are highly specific to aversion of cat odors. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 6442 – 6447.
- Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothetis, N. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science*, 334, 829 – 831.
- Wegner, D. M. (2003). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience*, 24, 57 – 86.

# 索引

---

- 色盲症, 60
- 动作电位, 16, 46, 48, 51, 53, 65 – 66, 100, 103 – 104, 106, 113, 129, 142 – 146
- 能动性, 106 – 111
- 失认症, 81
- 运动盲症, 62
- Allen, Paul, 145
- Allen脑科学研究所, 145 – 147
- 杏仁核, 34, 48, 62, 73, 94, 109
- Anderson, David, 142, 176
- Angelus Silesius, 165
- 动物权利, 140 – 141
- 人择原理, 7, 155 – 156
- 人工智能, 25, 131
- 注意, 选择性, 37, 45, 48, 55 – 57, 80, 113, 168 – 169, 176
- Baars, Bernie, 122, 170, 173
- 基底神经节, 34, 49, 73, 80, 106, 170
- 偏见, 无意识, 83 – 88, 159
- 双眼竞争, 52 – 54, 168
- 生物物理学, 16 – 19, 100 – 101, 103, 167
- 银翼杀手, 27, 168
- 盲点, 50
- Block, Ned, 5, 168, 173
- Bogen, Joseph, 68 – 69

Braitenberg, Valentino, 15

蝴蝶效应, 97

秀丽隐杆线虫, 36, 118, 128

演算法. 参见 Professor Calculus

Capgras妄想症, 61

小脑, 42 – 43, 80, 89, 129, 170, 176

大脑皮层, 34 – 35, 38, 42 – 43, 47 – 49, 52 – 55, 62, 68, 70 – 74, 82, 89 – 90, 110, 113, 121 – 122, 127, 139, 141 – 142, 145, 168, 170

Changeux, Jean-Pierre, 5, 122, 170, 174

河道视紫质-2 (ChR2), **143 – 144**, 146, 171

选择盲点, 87

屏状核, 34, 73, 113, 163, 171, 174

联盟。参见 神经元联盟

一致性, 101 – 103, 170

昏迷, 24, 33, 70 – 71, 139

相容论, 93 – 95, 111

复杂化, 7, 133

计算机意识。参见 意识, 计算机

混沌, 97 – 98, 101, 111

复杂性, 7, 36, 63, 117, 119, 133 – 134, 141, 147, 165

细胞类型, 16, 43, 54, 141 – 142, 146 – 147, 171, 176

加州理工学院 (Caltech), 6, 17, 20, 32, 66, 69, 142, 147, 159

Chalmers, David, 3, 5, 26, 124, 167, 168, 170, 173 – 174

计算机编程, 15

概念神经元, **65 – 67**, 100, 169

意识

动物意识, 34 – 36, 53 – 54, 116, 145 – 147

计算机意识, 2, 131, 170

伴随现象的, 31, 129

不同于注意力, 55 – 57, 168 – 169, 176 – 177

定义, 32 – 34, 42

梦境意识, 33, 44, 50 – 52, 70, 138

反馈回路, 43, 48, 53 – 54, 89, 114, 146

功能, 29 – 31

-测量仪, 7, 72, 131, 138 – 140

自我意识, 36 – 39

视觉意识, 48 – 53, 71, 89, 125, 146

连续闪烁抑制 (CFS), 45 – 46, 48, 52, 55, 168, 177

胼胝体, 68 – 70, 126, 142

皮层。参见 大脑皮层

皮层-丘脑复合体, 34, 51, 70, 73, 76, 89 – 90, 111, 124, 140, 146

Crick, Francis, 5, 19 – 21, 30 – 31, 42 – 44, 48 – 49, 53, 68, 78, 90, 113 – 114, 121 – 124, 137, 146, 153 – 154, 162 – 163, 167 – 171, 174

Dante, Alighieri, 8 – 9, 11, 164

决策制定, 30, 43, 49, 85 – 88, 94, 101, 105 – 107, 111, 169

深度睡眠, 24, 33, 52, 70, 126, 132, 138 – 139

Dehaene, Stanislas, 5, 122, 147, 168, 170 – 171, 174, 176

尼伯龙根的指环, 15, 94 – 95

Descartes, René, 23 – 24, 68, 92, 102, 114, 123, 134, 150 – 151, 165, 168

决定论, 93, 95 – 96, 98, 100 – 102, 110

确定性混沌。参见 混沌

DNA, 5, 24, 89, 117, 145

狗, 35, 72, 90, 115 – 116, 120, 140, 151, 160, 162

Dostoyevsky, Fyodor, 27

双重面向理论, 124

二元论, 23, 149 – 152

Eccles, John, 103 – 104, 151

涌现, 7, 116 – 119

情绪, 38 – 39, 44, 61, 77, 115

使能因子, 73

纠缠, 102 – 103, 105, 114

肠神经系统, 28 – 29

Epicurus, 164

关键节点, 62

脑电图 (EEG), 18, 40, 54, 64, 71, 126, **138 – 139**, 170

进化, 2, 4, 7, 11, 19, 31, 36, 38, 101, 115, 129, 133 – 134, 145, 153, 169, 174

面盲症, **60 – 61**

信仰, 11 – 12, 149, 153, 156, 159, 166

反馈回路, 43, 48, 53 – 54, 89, 114, 146

第一人称视角, 24

Fried, Itzhak, 64 – 67, 110, 170, 174, 176

Freud, Sigmund, 6, 76 – 77

功能主义, 120 – 121

功能性磁共振成像 (fMRI), **47**, 72, 142, 147, 168

全局工作空间理论, 122, 170

祖母神经元, 169

困难问题, 3, 5, 26, 137, 167

Huxley, Thomas 31, 168

下丘脑, 73, 144

非决定论, 101 – 102

下颞叶皮层, 53 – 54, 82

信息论, 115, 123 – 124, 170

抑制性中间神经元, 142, 144, 146

整合信息理论 (IIT), 6, 60, **124 – 131**, 146, 152 – 153, 170, 177

丘脑板内核, 73

Jennifer Aniston神经元。参见 概念神经元

Laureys, Steven, 5, 139, 167, 169, 175

Leibniz, Gottfried, 26, 119, 131, 133

损伤, 43

自由意志论的自由意志, 92, 102, 104, 151

Libet, Benjamin, 105 – 107, 170, 176

Logothetis, Nikos, 53 – 54, 168, 176 – 177

Lucretius, 102

猕猴。参见 猴子

磁共振成像 (MRI), 6, 45, **47**, 60, 168

掩蔽, 46, 52, 55, 147, 168

麻省理工学院 (MIT), 17, 19

黑客帝国, 44, 91 – 92, 111

内侧颞叶, **64 – 67**, 169

心-身二元论。参见 二元论

MindScope项目, 145

最小意识状态 (MCS), 71 – 72, 139, 170

镜像测试, 36 – 37

猴子, 32, 35, 37, 49, 51, 53, 57, 82 – 83, 102, 141, 145

单子, 26, 119, 125 – 126

一元论, 15

Monod, Jacques, 4, 5

小鼠, 33, 35, 57, 141, 144 – 147, 160, 171

支原体, 24

新皮层, 34, 52, 122, 142, 145

神经元联盟, 16, 46, 54, 63, 66, 71, 73 – 74, 89, 90, 117, 146 – 147

意识的神经元（神经）关联, 42 – 44, 49, 51, 53, 67, 89 – 90, 114, 125, 128, 168, 174, 177

Newton, Isaac, 95 – 98, 155

Nietzsche, Friedrich, 4, 15, 75, 77, 150

非快速眼动睡眠。参见深度睡眠

光遗传学, 26, 143 – 144, 171

泛心论, 131 – 134

持续性植物状态 (PVS), 71 – 72, 170

药物遗传学, 144

Poggio, Tomaso, 15

Popper, Karl, 103 – 104, 151

前额皮层, 43 – 44, 49, 54, 73, 80, 82, 89, 94, 121 – 122, 140

初级视觉皮层, 47 – 55, 60, 82, 89, 146, 168, 174 – 175

启动, 85, 169

作为教授, 14, 17 – 18

教授卡尔库卢斯, 14

属性二元论, 152

面容失认症。参见 面盲症

锥体神经元, 44, 49, 68, 89, 106, 113, 122, 129, 142

Qualia, 27 – 29, 30 – 32, 43, 54, 111, 121, 128, 130 – 131

量子力学, 98 – 104, 110 – 111, 114, 154, 169, 170, 177

快速眼动 (REM) 睡眠, 52, 138

准备电位, 105 – 107, 176

Rees, Geraint, 5, 51, 168, 175

宗教, 3 – 4, 12, 86, 116, 153, 159, 165, 171

视网膜, 46, 48 – 51, 53, 82, 89 – 90, 142, 144

RNA, 89, 134

浪漫还原论者, 8

扫视, 50, 79

Sacks, Oliver, 20, 61, 169

选择性视觉注意。参见 注意力, 选择性

自我意识。参见 意识, 自我

夏洛克 · 福尔摩斯, 41, 68, 128, 158

银斑马, 128

Singer, Peter, 160, 171, 177

Singer, Wolf, 5

灵魂, 3 – 4, 23, 93 – 94, 103, 111, 116, 149 – 152

稀疏表征, 66

Sperry, Roger, 69 – 70

峰值。参见 动作电位

裂脑患者, 69 – 70, 126, 170

注意力聚光灯。参见 注意力, 选择性

奇异环, 38

同步化。参见 同步性

同步性, 18, 48, 53, 73, 125, 129, 167

Teilhard de Chardin, Pierre, 7, 133 – 134

丘脑, 34, 49, 70, 73, 142, 170

整合信息理论。参见 整合信息理论

热运动, 98, 100

第三人称视角, 26

平铺, 142

丁丁, 14, 68

Tononi, Giulio, 5, 6, 124 – 126, 128, 130 – 131, 134, 138 – 140, 167 – 168, 170, 174 – 177

弓形虫病(Toxoplasmosis), 108 – 109

经颅磁刺激(Transcranial magnetic stimulation, TMS), 138 – 140, 170

无意识。见 无意识加工(Unconscious processing)

无意识加工(Unconscious processing), 30, 47 – 48, 61, 72, 76 – 90, 111 – 112, 119, 129 – 130, 161, 163, 168 – 169, 175

V1。见 初级视觉皮层(Primary visual cortex)

素食主义(Vegetarianism), 160

植物状态(Vegetative state, VS), 34, 52, 71 – 72, 139 – 140, 169 – 170

瓦格纳, 理查德(Wagner, Richard), 14 – 15, 94, 127

沃森, 詹姆斯(Watson, James), 5, 20

王尔德, 奥斯卡(Wilde, Oscar), 2

僵尸代理或行为(Zombie agents or actions), 30 – 31, 33, 78 – 81, 89 – 90, 108, 122, 130, 146