

# 思考，快与慢

## 序言

我想每位作者都会在脑海中勾勒读者因为读自己的书而受益的情形。我的这本书就像是办公室饮水机旁的闲谈一样，读者对它不会感到陌生。人们在饮水机旁或交流，是闲谈，很随意。我希望这本书能丰富人们的词汇，这样大家在谈论别人的判断与决策、公司的新政策或是同事的投资时，这些新词汇就能派上用场了。书中为何还会涉及闲谈的内容呢？因为利用闲谈来发现和分析别人犯的错误比分析自己的错误更容易，也更有意思。在人生最辉煌的时候，我们很难对自己的信念和需求产生怀疑，越是在最需要质疑自己的时候越难做到这一点，但我们可以从他人的真知灼见中受到启迪。很多人都会不由自主地推想朋友和同事会对自己的选择作何评价，而这些预先推断的中肯程度和内容都是十分重要的。对别有见地的闲谈有所期待是进行严厉的自我批评的强大动力，其作用比新年计划更大，更能提升我们在工作和生活中的决策能力。

一名内科医生要想成为优秀的诊断专家，就必须掌握多种疾病的大量特征，每种疾病都有其病理和症状，有其自身可能出现的发病前兆及病因、病情发展、结果和治疗或干预手段等。学医，当然也包括对医学用语的学习。要想对他人的判断和决策有更深入的了解，就要有更丰富的词汇，日常生活用语中的词汇还远远满足不了这一要求。我们很可能在闲谈中发现人们出错的方式各不相同，而一再出错就会令人印象深刻，而且我们可以预料到这种情况在特定环境中会再次发生。例如，一个潇洒自信的演讲者刚踏上舞台，观众往往就会对其大加赞赏，他可能并没有那么优秀，但这好评如潮的结果在他踏上舞台的一刹那便可预见。这种先入为主的情形是有据可依的，即光环效应，这一效应使我们更容易预测结果，更容易赏识和理解这位演讲者。

当别人问你正在想什么时，你一般都能回答上来。你觉得自己知道脑子里在想什么，通常就是一种意识自然而然导出另一种来，但这不是大脑工作的唯一方式，也不是其典型的工作方式。大多数印象和想法都是从意识经验中得来的，而人们是感知不到这一过程的。你无从知道自己是如何晓得面前的书桌上有盏灯，无从知道自己如何能通过电话听出爱人有些不耐烦，也无从知道自己如何毫无意识地成功规避了一场车祸。印象、直觉、决策，所有这些脑力活动都在无声地进行着。

本书讨论的很多内容都与直觉的成见相关。然而，对过失的关注并不意味着我们在诋毁

人类智慧，这就如同关注疾病并非否定健康一样。大多数人在通常情况下都是身体健康的，他们做出的决策、采取的行动往往也是恰当的。生活中，我们往往跟着印象和感觉走，凭直觉引导行事，而且我们觉得这种直觉和偏好很可靠，这种自信通常也是正当合理的。不过，也不尽然。我们经常在自己出现失误的时候还信心满满，此时，旁观者往往比我们自己更容易发现这些失误。

因此，我和大家一起在“饮水机旁交谈”，其目的就是：给读者提供更丰富精确的语言来讨论他人乃至自己在判断和决策上的失误，提升发现和理解这些失误的能力。至少有些例子表明，正确的判断能有效地干预错误，以降低错误的判断和决策经常造成的损失。

### 关于判断与决策的有趣研究

本书展现了我对判断和决策的理解，这种理解受到近几十年来心理学领域新发现的影响。然而，本书中心思想的形成还要追溯到1969年那个幸运的日子。当时我在耶路撒冷希伯来大学的心理学系教书，那天有个研讨会，我请我的一位同事阿莫斯·特沃斯基（Amos Tversky）在会上发言。他当时被视为决策研究领域的一颗新星，不过我觉得在其涉足的任何领域中，他都是耀眼的新星，因此我知道我们那天一定会交谈甚欢。很多认识阿莫斯的人都认为，在自己所见过的人中他是最聪明的。他才华横溢，十分健谈，魅力非凡。他有着超强的记忆力，记得很多有趣的笑话。他还拥有一种超常的能力，那就是利用记住的那些笑话阐明自己的观点。有阿莫斯在，你永远也不会感到沉闷。那时，他32岁，我35岁。

那天，阿莫斯给同学们讲了密歇根大学正在研究的一个项目，这个项目试图回答这样一个问题，即人是否是优秀的直觉型统计者。我们都知道人是优秀的直觉型文法家：4岁的孩子虽然对世界上存在语法这件事完全没有概念，但她在说话时会努力遵循语法规则。人们对统计的基本原则是否也有这种直觉感受呢？阿莫斯指出，研究得出的结论是附条件的肯定（肯定，但是附有一定条件）。我们在研讨会上进行了激烈的讨论，最终认为较为稳妥的结论应当是附条件的否定（否定，但是附有一定条件）。

阿莫斯和我很喜欢这种交流活动，我们认为直觉型统计者是个很有意思的话题，要是一起探索的话会很有意思。那个周五，我们在里蒙餐馆吃午餐，那里是波西米亚人和耶路撒冷的教授们最中意的去处。我们两人打算对一些经验丰富的研究人员的统计直觉进行一番研究。在此前那场研讨会上我们曾得出结论：人的直觉是有缺陷的。虽然这些年来一直在教书，在运用统计学原理，但我们也没能培养出一种直觉，无法利用这种直觉感知从小样本中观察到的统计结果的可靠性。我们的主观判断是存在成见的：我们特别容易相信在没有足够证据的基础上得出的研究结果，而且研究中对观察样本的收集也不足。我们两人此番研究的目的，

就是要看看其他研究人员是否也和我们一样有着同样的苦恼。

我们准备了一项调查，其中包括一些研究中出现的实际统计问题。阿莫斯收集了数学心理学协会与会的专家小组的回复，包括曾出版两本统计学教科书的几位作者的问卷。不出所料，我们发现那些专家同行也跟我们一样，总是夸大其词，他们认为一个实验的原创性结果可以被成功复制的概率很大，即使用一个小样本也可以做到这一点。关于实验要收集多少观察数据的问题，他们给一个假定的毕业生的建议也很糟糕。如此看来，即使是统计学家，也算不上是出色的直觉型统计者。

在撰写这些发现时，阿莫斯和我都觉得我们在一起工作是件很享受的事。阿莫斯总是很风趣，有他在的时候，我也变得幽默了，所以我们总会在轻松愉快的气氛中度过几个小时连续不间断的工作时光。工作中的乐趣使我们变得格外有耐心，人在放松惬意的情况下，更容易取得完美的结果。也许最重要的是，我们把批评的态度都抛在门外了吧。我和阿莫斯都是爱挑剔、好辩论的人，他甚至比我更甚，但在我们合作的这些年里，我们从没有不假思索地否定对方。事实上，我发现我们在合作时，阿莫斯总能更清楚地看出我模糊的观点中要表达的意思。我们两人中，阿莫斯的逻辑思考能力更强，他的意见总是有据可依，言之凿凿，令人信服。我则凭直觉走，深受心理学的影响，我的很多观点也都是从心理学中得来的。我们俩有很多相似之处，因此很容易理解对方；我们也有很多不同的地方，这些差异常常令对方吃惊。我们重新安排各自的日程，这样就有很多工作日可以在一起工作，工作之余，我们常会一起散步。此后的14年里，共同合作就成了我们生活的中心，对我们两人而言，那些年里所作的研究是我们一生中最精彩的篇章。

我们很快便形成了固定的工作模式，并且一直多年保持这一模式。我们的研究采用的是对话形式，对话中的问题是我们自拟的，那些凭直觉做出的答案也是经过我们两人共同检验过的。研究中的每个问题都是一个小实验，仅一天中就会做很多实验。我们并没有刻意寻求自己提出的那些统计问题的正确答案，只是想确认和分析直觉的回答——大脑中最先出现的、即使知道是错的我们也愿意拿来分析的回答。当时，我们认为其他人也会有我们两个人都有的直觉，事实也正是这样。如此说来，直觉对判断的影响便显而易见了。

我们曾经很高兴地发现，我们俩对几个认识的孩子的未来职业的预想竟如出一辙。我们确信那个3岁大却善辩的孩子将来会做律师，那个有点呆板的孩子可能成为教授，那个体谅他人、循循善诱的孩子可以做个心理咨询师。当然，这些预测都是荒谬的，不过却很有意思。我们都清楚一点，那就是我们对这些孩子的直觉，来自他们自身的特点与特定职业特点的相似度。这种有趣的做法使我们当时就在脑海中创立了一种理论，即预测角色的相似度。此后我们做了许多实验来验证和详细阐述这个理论，下面便是一例。

在你思考下文中的问题时，请记住史蒂夫是从一个有代表性的样本库中被随机挑选出来的：

邻居如此描述这个孩子：“史蒂夫非常腼腆，少言寡语，很乐于助人，却对他人或者这个现实世界没有兴趣。他谦恭有礼，做事井井有条，中规中矩，关注细节。”请问史蒂夫更可能从事哪种职业，图书管理员还是农民？

很显然，史蒂夫的个性和典型的图书管理员有着惊人的相似，但这些与职业密切相关的统计学因素却很少有人关注。你们是否注意到，在美国，农民与图书管理员的比例超过20：1。由于农民数量要多得多，所以那些“谦恭有礼，做事井井有条”的人也常常只能成为坐在拖拉机上的农民，而不可能是坐在图书馆咨询台后的管理员。但是，我们发现实验对象往往忽略这些相关的统计数据，而仅仅依赖于相似度来作出判断。于是，我们提出如下观点：人们把相似度当成一种简单的启发手段（简单地说就是经验法则）来作艰难的判断。对这种启发性手段的依赖必然会造成其预测带有成见（系统性失误）。

还有一次，阿莫斯和我想知道我们这所大学的教授们的离婚率是多少。我们注意到这个问题立即勾起了我们脑海中的记忆，我们俩不由想起自己知道或听说的那些离了婚的教授。于是我们就凭着脑海中这些事例对这个离婚率问题作出判断。我们把这种依靠记忆作出判断的方法称为可得性法则。在一项研究中，我们让调查对象回答一个简单的问题，这个问题与指定的一篇英语课文中的单词相关：

请思考字母 K。

请问字母 K 是更多地出现在单词的首字母位置上还是第三个字母位置上？

玩拼字游戏的人都知道，想起以某个字母开头的单词要比想起它在第三个字母位置上的单词容易得多。字母表中任何一个字母都适用于此法则。因此我们料到，尽管有些字母（比如 K、L、N、R、V）出现在第三个字母位置上的频率更高，但是被调查对象的回答肯定会夸大所有字母出现在单词首字母位置上的频率。这种情形再一次表明，对经验法则的依赖必然会导致人们判断时的成见。例如，我曾一度认为通奸在政客中较普遍，在医生或律师中则没那么多见，但我最近对此产生了怀疑。我甚至曾为原来的那个“事实”作出若干解释，包括权力的催情效果、出门在外受到的诱惑等。而我最终意识到，政客的过错只不过更容易被曝光而已，而医生和律师的过错却少有人报道。我的直觉印象可以完全归咎于记者对主题的选择和我对可得性法则的依赖。

阿莫斯和我用几年的时间来研究并记录在不同任务中体现的直觉思考所存在的成见，这些任务包括对事件的概率赋值、对未来进行预测、对假设进行评估，以及对频率进行预估等。在合作的第五年，我们撰文将这项研究的主要发现发表在《科学》杂志上，这份杂志的读者包括很多领域的学者。那篇文章（本书的最后附有全文）描述了直觉思考的简单快捷，列出了在启发法中表现出的大约20种成见，还包括启发法在判定中的作用。

科学史学家常指出，某一特定领域的学者在任何时候都愿意和他人分享关于本学科的观点。社会科学家也是如此。他们把一切问题都归结为人性，认为大多数关于人类特有行为的讨论都应以此为背景，这一观点几乎从未受到质疑。关于人性，20世纪70年代的社会科学家广泛接纳了两种观点。第一，人大体而言都是理性的，其想法通常也是合理的。第二，恐惧、喜爱和憎恨这样的情感能够为人们失去理智的大部分情形作出解释。我们这篇文章虽然没有直接讨论上述观点，却是对这两种观点的挑战。我们记录下正常人思考时出现的系统性失误，认为这些失误是由认知机制的构造造成的，并非由情感引起的思想腐化导致的。

这篇文章所受到的关注远远超出了我们的预期，而且它至今仍是社会科学著作中被引用次数最多的文章之一（2010年有300多篇学术文章参考了这篇文章）。其他学科的学者也觉得这篇文章很有用处，启发法和成见等概念被广泛应用于众多领域中，包括医学诊断、法律判决、情报分析、哲学、金融、统计学和军事战略等。

例如，学习政策的学生就曾注意到，可得性法则能解释为什么有些事人们记得很清楚，而有的却被遗忘了。人们是根据从记忆中提取信息的容易程度来估测事情的重要程度的，而这往往也与媒体报道的广泛程度有关。常被提到的话题就在脑中变得鲜活，而其他的则会慢慢被遗忘。也就是说，媒体选择报道的内容和人们脑中存在的信息不谋而合，所以专制政体对独立媒体施压的现象也不是偶然的了。因为重大事件和名人很容易引起公众的兴趣，媒体能借此煽动狂潮也就见怪不怪了。例如，在迈克尔·杰克逊死后的几周里，电视台几乎未报道别的事。相反，媒体对那些带有批评性的、不能引起公众兴趣、掀不起大波澜的事往往很少报道，比如说去年日趋下滑的教育标准，还有医疗资源的投资过剩等。（我在写这篇文章的时候发现自己选择的“很少报道”的例子都是受可得性限制的。我选为例子的话题经常被提到，那些同等重要却不常被提到的事我往往想不到。）

有一点我们当时并没有充分意识到，即“启发法和成见”这样的心理学概念在其他领域中也具有广泛的启发作用，这便是我们这项研究的一个附带成果：我们总是把为自己和被调查者设计的全部问题都写进文章里，这些问题可以为读者提供范例，使其认识到自己的想法是如何受认知性成见牵绊的。我希望你在读到“史蒂夫是个图书管理员”这样的问题时能有切

身的体验，这样能帮助你更好地体会到相似度在引导可能性上的力量，并且能体验到我们多么容易忽略相关事实的统计。

这些实证材料的使用可为不同领域的学者（主要是哲学家和经济学家）提供一次不寻常的机会，使他们关注自己在思考时可能出现的纰漏。看到自己的纰漏，这些学者才更有可能质疑当时普遍存在的那种武断想法，即人类很理性、很有逻辑性。方法的选择很重要：如果我们只报道传统实验的结果，这篇文章就不会那么令人关注，也不会令人如此难忘了。而且，那些持怀疑态度的读者会将自己的判断失误归咎于参与这些心理学研究的大学生，认为是这些学生一贯不负责任的做法使他们不愿相信实验结果。当然，我们摒弃传统的实验方法，采用事例展示的方式，并非只为影响那些哲学家和经济学家。我们采用这种方法，是因为将事例展示出来更有意思。我们很幸运，因为我们选择了正确的方法，其他各方面的选择也做对了。本书重复出现的一个主题就是，幸运在每个成功的事例中都扮演着重要角色。我们总能很容易地发现，这个事例中一个小小的改变就会将伟大的成就变得平淡无奇。我们展示的这些事例也不例外。

对本书的反馈也不全是正面的。我们对成见予以关注的做法受到非常多的批评，反对者认为我们过分否定了人类的思维能力。与对常规科学的反应一样，有些研究者对我们的观点加以改进，有些人则另外提出了一些貌似合理的见解，但有一个观点如今是得到普遍认可的，即我们的大脑容易受系统性误差的影响。我们关于判断（能力）的研究对社会科学产生的影响远远超出我们当时的预料。

研究完判断这一论题后，我们马上将目光转向面对不确定因素时的决策过程。我们的目标是创立一种心理学理论，研究人们在简单的赌博中如何作决定。例如：投硬币时如果是正面朝上你就能得到130美元，背面朝上就输掉100美元，你愿意打这个赌吗？这些简单的选择很久以来一直被用来检验各种与决策相关的问题，例如人们如何在确定的事物和不确定的结果之间进行权衡。我们的研究方法没有变：还是花很多天设计一些选择题，而后分析我们根据直觉进行的选择是否与通过逻辑判断作出的选择一致。在作判断时，我们会观察自己做出决策时出现的系统性成见，还会对一贯违背理性选择规律的直觉性选择进行观察。在《科学》杂志刊出那篇文章5年之后，我们又发表了《前景理论：风险决策分析》一文，据统计，该文中提出的决策理论比我们此前对判断的研究更具影响力，该理论也为行为经济学奠定了一定的基础。

在合作过程中，阿莫斯和我经常交流思想，两个人的智慧总要胜过一个人的想法，良好的关系也使我们的工作有趣且高效，这段时光是我人生中宝贵的财富。后来，我和阿莫斯离得远了，很难继续共同研究这一课题。我们在判断和决策制定方面的研究使我在2002年获得

了诺贝尔经济学奖。如果阿莫斯没有于1996年去世（时年59岁）的话，他应该和我一起去领这个奖。

本书并非为展示我和阿莫斯共同合作的早期研究，过去几年里很多作者已经出色地完成了这项工作。我的主要目标是，在认知心理学和社会心理学最新发展的基础上展示大脑的工作机制，在这些发展中有一些内容比较重要，其中一项就是我们认为瑕瑜互见的直觉思维。

阿莫斯和我没有对直觉下准确的定义，只简单说明了判断启发法“很有用，但有时也会导致严重的系统性误差”。我们的重点放在成见上，因为人们的大脑在高速运转时。研究成见非常有意思，而且成见为启发性判断提供了研究依据。我们没有自问在面对不确定因素时所有的直觉性判断是否都是通过我们研究的启发法而产生的。不过现在我们知道事实并非如此。专家们的直觉往往很准确，这种准确性与其说是启发法在发生作用，还不如说是长期实践的结果。我们现在可以描绘一幅更加美好、更加和谐的画面，在这幅画面里，熟练和启发法皆可成为直觉性判断和决策的选择依据。心理学家加里·克莱因（Gary Klein）曾讲过一个故事：一支消防队进入一座房屋，屋子里的厨房着火了。他们刚开始用水管浇厨房，指挥官喊道：“全部撤离！”其实他自己也不知道为什么要这样做。在消防员全部撤离的一刹那，厨房的地板轰然塌陷。事后指挥官才回想自己曾意识到这场火并不大，但他的耳朵特别烤得慌。这些他所谓的“对危险的第六感”闪进他的脑海，虽然不知道哪里不对劲儿，但他知道情况不妙。最后大家才知道这场火灾的火源根本不是厨房，而是消防员脚下的地下室。

我们都听过一些关于专家的直觉的故事：比如某位象棋大师路过街边棋局，无须驻足观看就知道“白方三步之内将杀”；又如某位医生只需一瞥便能作出全面的诊断。专家式直觉像谜一样吸引着我们，但它们不是谜。我们每个人每天都会多次表现出很强的直觉能力。大多数人在接电话时听到第一个词就能感知对方是否生气了；刚进门就能发现自己是大家谈论的对象；对细微的信号能迅速做出反应，断定旁边车道上的汽车司机正处于危险中。我们日常的直觉能力并不逊于一位经验丰富的消防员或者医生，只是无处施展罢了。

魔法不属于准确的直觉心理的范畴。也许对此有最精辟论述的人要数伟大的赫伯特·西蒙（Herbert Simon）了，他对多位象棋大师进行过研究，发现在练习数千个小时之后，这些大师看到的棋盘上的棋子和我们眼里的棋子是不一样的。西蒙写道：“这个棋局已经给了我们提示，根据这个提示我们可以搜寻到大脑存储的信息，而这些信息就能给出答案。直觉只不过是人们的认知而已。”看到这种说法，你也许就能感受到西蒙对神化专家直觉的做法不以为然了。

一个两岁的小孩看到一只狗时会说“小狗狗”，这没什么好大惊小怪的，因为我们已经对孩子认识事物并叫出很多事物名字这种“奇迹”习以为常了。西蒙认为灵性的直觉所创造的“奇迹”也是如此，不值得大惊小怪。如果专家在新情境中能察觉熟悉的因素，并且采取十分得体的做法来顺应这种情境，这便是正确的直觉。在喊出“小狗狗”的同时，孩子的大脑中就会出现正确的直觉性判断。

不幸的是，专业的直觉并非全部是从真实的经验中得来的。很多年前，我采访了一家大型金融公司的首席投资官，他告诉我他刚刚买了福特汽车公司上千万美元的股票。我问他是如何做出这样的决策的，他回答说刚参加了一个车展，感觉很好。他的解释是：“这还用问，他们生产的汽车太霸道了！”他非常明确地说，他相信自己的感觉，他对自己和自己的决定都很满意。我很吃惊地发现，他忽视了一个任何经济学家都会注意到的相关问题，福特公司的股票最近不是走低吗？而他反而在跟着自己的直觉行事：他喜欢汽车，喜欢福特公司，也喜欢持有其股票的感觉。按照我们的理解，买入股票要谨慎，这位仁兄恐怕根本不知道自己在做什么。

关于这位投资官投资福特公司股票的问题，阿莫斯和我所研究的特定启发法恐怕是解释不了了，但现在关于启发法还有一种更宽泛的概念，这种概念能对上述行为作出更好的解释。这一宽泛的概念是一个重大进步，即情感因素在我们对直觉判断和决策的理解上发挥了比以往更大的作用。这位投资官的决策在今天可能会被视为受启发法的影响，而判断和决策是直接受好恶这样的情感所左右的，没有什么思忖和推理可言。

当碰到像走哪步棋或决定是否投资股票这样的问题时，直觉思维机制就会充分发挥其作用。如果某个人有相关的专业知识，她就能更好地认清情况，头脑中形成的直觉性解决方案也很可能是正确的。象棋大师看到复杂的棋局时大脑会迅速反应：刹那间想到的那几步棋一定是好棋。要是问题很难，一时也想不到巧妙的主意，直觉就可能发挥作用：脑海里可能马上会有个答案，但这个答案却不一定是原题的答案。这位投资官所面对的问题（我是否要投资福特公司股票）就很难，但更简单且相关的问题（我喜欢福特汽车吗）却很快在他的脑海中形成了答案，并且让他就是否作出投资这个问题也作出了相应的选择。这就是直觉启发法的核心观点：当面对难题时，我们往往会对相对简单的问题进行回答，却忽略了自己已经置换了原始问题这个事实。

有时，我们无法自然地凭直觉找出问题的解决方案——不论是专业的解决方法还是启发式的答案。在这种情况下，我们往往想要找到一种更慢、更严谨、需要投入更多脑力的思考形式，这就是本书中提到的慢思考。快思考既包括直觉思维的不同形式，比如专家式的和启发式的，也包括感觉和记忆等所有无意识的大脑活动，这些活动会让你知道桌上有盏台灯或者



想起俄罗斯首都的名字来。

在过去的25年里，已经有很多心理学家对快思考和慢思考的区别进行了研究。我用两个因素来描述人的思维活动，即系统1和系统2，在后文中我将对分类原因进行更详细的阐述。系统1和系统2分别产生快思考和慢思考。我认为直觉和严谨思考的特点就像是大脑中两种性格的特征和性情。在近期的研究中，系统1的直觉性作用比我感觉到的还要大，它是做出的决策和判断的幕后主使。本书大部分内容是关于系统1的运作以及系统1和系统2间的相互影响的。

本书共分为五部分，第一部分讲述的是通过双系统判断与做出决策的基本原理。这部分内容详细说明了系统1的无意识运作和系统2的受控制运作的区别，并且说明了系统1的核心，即联想记忆是如何不断对世界上所发生的事作出连贯的解释的。关于直觉性思考的自主且无意识过程的复杂性和丰富程度，以及这些自主过程如何能解释判断的启发法等问题，我试图说出自己的见解，目的是要引入一套用于思考和表达思想的语言。

第二部分对判断启发法的研究作了更新，还探索了一个难题，即为什么很难具备统计型思维。我们思考时总是会把多种事情联系起来，会将一件事情比喻成另一件，会突然想起一件事来，但统计学要求同一时间把多件事情串联起来，而这一点系统1是做不到的。

本书第三部分描述了我们大脑有说不清楚的局限：我们对自己认为熟知的事物确信不疑，我们显然无法了解自己的无知程度，无法确切了解自己所生活的这个世界的不确定性。我们总是高估自己对世界的了解，却低估了事件中存在的偶然性。当我们回顾以往时，由于后见之明，对有些事会产生虚幻的确定感，因此我们变得过于自信。我对这个问题的看法受《黑天鹅》（The Black Swan）的作者纳西姆·塔勒布（Nassim Taleb）的影响。我希望我这“饮水机旁的闲谈”能明智地借鉴以往经验，同时抵制后见之明和虚幻的确定之感的诱惑。

第四部分的重点，是在决策制定的性质和经济因素为理性的前提下讨论经济的原则。1979年，阿莫斯和我发表了关于前景理论的决策模式，此部分在双系统下对前景理论的重要概念提出了新的看法。余下的几章讲的是人们从理性角度出发做出决策的几种方式。可悲的是，人们总是孤立地看待问题，表现出框架效应，即决策的制定往往因为对所回答问题不合逻辑的选择而受到影响。系统1的特征完全能解释这些观察结果，这对标准经济学所倾向的理性假设发起了很大的挑战。

第五部分是近期研究中关于两个自我的区别性描述，即经验自我和记忆自我，二者间没有共性。例如，我们可以让人们体验两种痛苦。其中一种比另一种要更痛苦，因为体验的时

间更长。系统1有一大特点，即记忆的自主形成是有其原则的，如此一来，较为痛苦的那段体验会留下更深刻的记忆。所以，此后当人们选择要回想哪段经历时，他们自然会受记忆自我的引导，将其自身（即经验自我）处于不必要的痛苦中。两种自我间的区别被用来测试人的幸福感，而我们发现使经验自我快乐的事不一定会让记忆自我满足。两种自我同时存在的个体要如何去追求幸福，这一问题引起了把居民的幸福看做政策目标的个人和社会的众多思考。

最后的章节是按倒叙来探索本书所述的三个区别的：经验自我和记忆自我的区别，古典经济学和和行为经济学（从心理学借鉴而来）的区别，以及自主的系统1和需费脑力的系统2的区别。书中还谈及了有价值的闲谈的好处，以及哪些内容有助于提升判断和自行决策的效能。

在最后，我附上了我和阿莫斯一起写的两篇文章，第一篇是我早期写的关于在面对不确定性因素时作出判断的评论。第二篇发表于1984年，总结了前景理论和我们关于框架效应的研究成果。文章中有被诺贝尔委员会引用的投稿，你可能会惊讶地发现这些投稿是多么简单。读这些能让你明白我们早期的知识有多少，也能让你知道我们这几十年来的进步有多大。

## 思考，快与慢

### 第一章 一张愤怒的脸和一道乘法题

#### 第一部分 系统1，系统2

想要观察你在自动模式下的大脑活动，请看下面这幅图。

审视这位女性的脸时，你正在很自然地将我们平时所说的观察和直觉思考结合在一起。你可以确定且迅速地判断这位年轻女性的头发是黑色的，你还知道她正在生气。此外，你还可以根据自己的观察对其行为进行推测。你感觉到这位女性正要说一些刻薄话，也许声音又大又刺耳。推测被观察对象即将做出的举动并不难，这种活动是无意识的，而且毫不费力。你并非有意评论她的情绪或是推测她可能要做的事，你对这张照片的反应跟你做过的事情没有丝毫关系，一切都是自然而然发生的。这就是快思考的一个例子。

现在，请看下面的问题：

你能立刻知道这是一道乘法题，也许你还会想到若有纸笔，就能算出答案。你还会对答案的大体范围有个模糊的直观认识，能很快知道12 609 和123不可能是答案。但如果不花点时间来计算的话，你就无法确定568不是正确答案。由于想不出一个准确的答案，你认为自己得想想是不是要做这道题。如果你还没有做这道题的话，就该试着做一做，哪怕完成其中的一部分也好。

按部就班的运算过程便是慢思考。首先，你会从记忆中重新提取读书时所学的乘法相关知识，然后加以运用。这个过程不容易，你得记住很多内容，你要知道自己算到哪一步了，知道下一步该怎样做，同时还要记住已得到的结果。这个计算过程是脑力工作，需要刻意、努力并且有序地进行——这也是慢思考的一个特征。这种计算不仅是大脑活动，身体也会参与其中，在计算时你的肌肉会紧张，血压会上升，心跳会加速。若在你解决这个问题时，有人在近处看你的眼睛，他会发现你的瞳孔也会扩大。结束计算时——得出正确答案（顺便提一下，答案是408）或是放弃计算都被视为结束计算——你的瞳孔便会恢复到正常大小。

### 行为与注意力

近几十年来，许多心理学家对人的两种思维模式一直保持着浓厚的兴趣，这两种思维模式是由一张愤怒女性的照片和一道乘法题所引发的，他们还指出了两种模式的许多特征。这里我且采用由心理学家基思·斯坦诺维奇（Keith Stanovich）和理查德·韦斯特（Richard West）率先提出的术语，用以说明大脑中的两套系统，即系统1和系统2。

系统1的运行是无意识且快速的，不怎么费脑力，没有感觉，完全处于自主控制状态。

系统2将注意力转移到需要费脑力的大脑活动上来，例如复杂的运算。系统2的运行通常与行为、选择和专注等主观体验相关联。

系统1和系统2的定义广泛应用于心理学领域，但我在本书中所作的阐释更加深入，读者可将此书视为有两个人物的心理剧。

我们在审视自己时，往往更容易采用系统2，认为自己头脑清醒，富有逻辑，抱有信仰，善作抉择，能够决定自己想要什么和该做些什么。尽管系统2的运行体现在行动中，但自主运行的系统1才是本书的重点。我将系统1描述成自主而初始的印象和感觉，这种印象和感觉是系统2中明确信念的主要来源，也是经过深思熟虑后作出抉择的主要依据。系统1的自主运

作诱发了极其复杂的理念模式，但只有相对缓慢的系统2才能按部就班地构建想法。我还描述了继而发生的系统2的环境条件，在此条件下，系统1中随性的冲动及其诱发的联想都会受到抑制。读过此书后，你会觉得这两种系统各有千秋，各司其职。

依据复杂性进行粗略预估，试举以下系统1引发的自主行为的例子：

确定两件物品孰远孰近。

确定突然出现的声源。

将短语“面包和.....”补充完整。

看到恐怖画面后做出厌恶的表情。

察觉语气中的不友善。

回答 $2+2=?$

读大型广告牌上的字。

在空旷的道路上驾车行驶。

下象棋时看出一步好棋（前提是你是位象棋大师）。

理解简单的句子。

听到“畏首畏尾，追求完美”的说法时，知道其意指古板的员工。

上述所有思维活动都和那位愤怒的女性相关——这些活动都是自主发生且毫不费力的。系统1的诸多能力当中包括一些与生俱来的能力，这些能力与其他动物的本能一样。我们生来就能感知周围的世界，能够认识事物，可以集中注意力，会规避风险，会害怕蜘蛛，思维时快时慢。大脑的其他思维活动也因长期的训练而变得快速自主。系统1除了能将我们已有的知识（比如法国的首都是哪里）联系起来之外，还能使我们掌握一些技能，比如看出并理解一些社交场合的细微差别。有些技能只有专家才能掌握，比如象棋中出奇制胜的那几招；而其他的技能，普通人都能掌握。例如，要想看出刻板员工的个性有哪些相似之处，还需要我

们掌握大量的语言及文化知识，而大多数人都具备这些知识，这些知识就储存在我们的记忆中，不必刻意也无须努力便可随意存取。

以上列出的大脑活动是完全无意识的。你不必刻意学习便可领会一些母语中的简单句子；听到突然的一声响后会自发地确定声源；看到“2+2”就知道等于4；提到法国的首都时会不由自主地想到巴黎。其他一些行为可能会受到大脑的控制，比如咀嚼的动作，但大体上还是无意识的。系统1和系统2这两个系统都对注意力有控制作用。确定声源位置通常是在系统1控制下的无意识活动，随后系统2会立即被激发，产生有意识的注意力。拥挤的派对上，你也许会对大声而无礼的谈话置之不理，不过，即使你的头没转过去，你的注意力也已经转移过去了，哪怕只有一会儿。但是，要想将注意力从不想关注的对象上转移开来也容易，去关注另一个目标即可。

系统2的运作是高度多样化的，但所有这些运作方式都有一个共同特征：所有运作都需要集中注意力，如若注意力分散，运作也会随之中断。以下是一些例子：

赛跑时随时作好起跑准备。

关注马戏团里的小丑。

在一间嘈杂、拥挤的屋子里关注某个人的声音。

寻找某位白头发的妇女。

搜寻大脑记忆，判定声音是否表达惊喜。

保持比平常快的步行速度。

观察自己在社交场合的做法是否得体。

数出文章中某页字母 a 的出现次数。

告诉某人你的电话号码。

在狭小的空间里停车（除车库管理员外，大多数人停车的车位都很狭窄）。

比较两款洗衣机的总体功效。

填纳税申报表。

检验一个复杂的逻辑论证的有效性。

在上述各种场景中，你都必须集中注意力。若是没有准备好或者没有将注意力集中到正在做的事情上，你就会表现得差强人意，甚至是一塌糊涂。系统2具有某些改变系统1运作方式的能力，通过控制注意力和记忆力的一般自主运行功能的方法可以实现这些改变。例如，在繁忙的火车站等亲戚时，你若是刻意去找某位头发花白的妇女或是长着胡子的人，即使隔着一段距离你也很可能发现自己的亲戚；你也可以在记忆中搜寻哪些国家的首都是以字母N开头的，或者法国有哪些存在主义小说；而当你在伦敦希斯罗机场租车时，地勤人员有可能会提醒你“在我们国家是右驾左行”。

在上述所有事例中，没有哪件事情是水到渠成的，你会发现始终如一地保持某种状态需要付出持之以恒的努力，至少要一直耗费精力。

“注意力要集中”，这个耳熟能详的短语是很有道理的，如果你将原本应分配给某些活动的注意力分散开来，如果分散掉的注意力有限，还可以接受；但是如果你想透支你的注意力，将其过度分散到其他事情上，结果就会失败。这证明了有些费脑力的活动会相互影响，也说明了为什么同时进行几项活动很难，甚至是不可能的。就如同你不能一边向左转入路线复杂的道路，一边计算“ $17 \times 24$ ”这道乘法题的结果一样，两者无法兼做。当然了，你最好还是不要尝试。你可以同时做几件事情，但前提是这些事简单易懂，你可以在空旷的高速公路上一边开车一边和别人交谈而不出事故。有些家长可能会感到有些愧疚，因为他们在给孩子读故事时，脑子里还在想别的事。

每个人都能多多少少地意识到注意力是有限的，在社会生活中我们也会为此作出妥协。例如，当司机正在一条匝道上赶超一辆卡车时，车上的成年乘客会明智地停止与司机攀谈，因为他们清楚地知道，在这个时候让司机分心并不是什么好事，而且他们会认为司机此时会屏蔽掉别人的话语，仿佛暂时性失聪一般。

当人们太过专注于某件事时，就会屏蔽掉其他事情，即使是平时很感兴趣的事也不例外。在《看不见的大猩猩》（*The Invisible Gorilla*）一书中，克里斯托弗·查布里斯（Christopher Chabris）和丹尼尔·西蒙斯（Daniel Simons）两位作者为我们提供了一个最具戏剧性的证明。他们设计了一部两队传篮球的短片，其中一队穿的是白色球衣，另一队穿的是黑色球衣。观

看短片的人需要数出白衣球队的传球次数，忽略掉另一队传的球。这个任务比较困难，需要完全投入才行。短片播到一半时，一个套着大猩猩服装的女人出现了，她穿过球场，捶着胸，然后继续走动。这只“猩猩”出现了9秒钟。上万人看了这部短片，其中约有一半人并未注意到有什么异常。之所以这样，是因为这个计数任务——尤其是那个忽略黑衣球队的要求——造成了这种屏蔽。若没有那项任务，所有观看短片的人都会注意到那只“猩猩”。观看和定位是系统1的自动功能，但在执行时需要将一些注意力分配给相关的刺激物。该书的两位作者提到，在这项研究中，最值得注意的是人们在知道结果后的吃惊反应。那些没有看到“猩猩”的观众刚开始就确信场上没有“猩猩”——他们很难想象自己会错过这件吸引人眼球的事。这个关于“猩猩”的研究阐述了与我们大脑相关的两个重要事实：我们会忽视显而易见的事，也会忽视自己屏蔽了这些事的事实。

### 系统1遇到麻烦，系统2会出面解决

本书紧扣两个系统相互作用这一主题，其内容与结构简洁明了。在书中我将指出，当我们醒着时，系统1和系统2都处于活跃状态。系统1是自主运行，而系统2则通常处于不费力的放松状态，运行时只有部分能力参与。系统1不断为系统2提供印象、直觉、意向和感觉等信息。如果系统2接收了这些信息，则会将印象、直觉等转变为信念，将冲动转化为自主行为。通常情况下，一切都会顺利进行，系统2会稍微调整或是毫无保留地接受系统1的建议。因此，你一般会相信自己的最初印象，并依自己的想法行动。通常情况下，这样也挺好的。

当系统1的运行遇到阻碍时，便会向系统2寻求支持，请求系统2给出更为详细和明确的处理方式来解决当前问题。系统2在系统1无法提供问题答案时，就会被激活，这好比你碰到了“ $17 \times 24$ ”这样的乘法题，系统1无法给出答案，系统2便被激活来解决问题。当你遇到令人吃惊的事情时，同样会感到自己有意识的那部分注意力会瞬间激增。另外，在系统1所设定的世界里，电灯不会跳，猫不会像狗一样汪汪叫，“猩猩”也不会穿过篮球场。如若事物违反了系统1所设定的关于世界的模式，系统2同样会被激活。

“大猩猩”的实验表明，想要察觉到令人惊讶的刺激物，就要对其予以关注。然后，那种惊讶会激发并引导你的注意力：你会将目光集中在令你瞩目的对象身上，并在记忆中搜寻此事令人惊讶的原因。系统2还会起到持续监督你自身行为的作用——有了它，你在生气时也能保持应有的礼节；有了它，你在夜晚开车时也能保持警惕。当你就要犯错时，系统2就会受到刺激，加速运作。回想一下，在冒犯别人的话即将脱口而出时，想把话咽回去是多么困难。总的来说，你（或你的系统2）所想所做的大多数事情都是由系统1引起的，但当事情变得困难时，系统2便会接手难题，系统2出马，所有事情都会迎刃而解。

系统1和系统2的分工是非常高效的：代价最小，效果最好。通常情况下，这种分工很有效，因为系统1很善于完成自己的本职工作：它在熟悉情境中采取的模式是精确的，所作出的短期预测是准确的，遇到挑战时做出的第一反应也是迅速且基本恰当的。然而，系统1存在成见，在很多特定的情况下，这一系统易犯系统性错误。你会发现这个系统有时候会将原本较难的问题作简单化处理，对于逻辑学和统计学问题，它几乎一无所知。系统1还有一个更大的局限，即我们无法关闭它。如果看到屏幕上显示一个你认识的单词，你就能读出这个词——除非你的心思完全不在它上面。

自主反应和控制这种反应的意图之间存在冲突，这种冲突在生活中极为普遍。我们差不多都有过这样的经历：在餐厅里，自己的邻座是一对穿着怪异的夫妻，但我们会尽量不去盯着他们看。我们也清楚，如果看书时老是重读不知所云的内容，这就说明我们在强迫自己去读一本无聊的书。在冬季严寒的地区，许多司机都有着这样的记忆：当他们的车在冰上滑行失去控制时，他们必须放弃平时的做法，竭力去按照已反复演练过的指示进行操作：“如果车打滑，做什么都行，就是别踩刹车！”另外，每一个人都有强忍着不诅咒别人去死的经历。系统2的众多任务中就包括抑制系统1产生的这些冲动。换句话说，系统2负责人们的自我控制。

不是所有错觉都是视觉上的，还有思维方面的，我们将其称为认知错觉。我读研究生的时候，选修了一些与心理疗法的艺术及科学相关的课程。记得有一次上课时，老师跟我们分享了一些坐诊艺术。他告诉我们：“有时，你会碰到一两个这样的病人——他会像说故事一样，讲述自己以前遭遇的误诊，这些诊断五花八门，让人担心。他看过几个临床医生，但都没多大效果。这个病人还能清楚地描述医生是如何误解他的，但他很快就观察到，你和其他医生是不一样的，你能感同身受，充分理解他，并可以为他提供帮助。”此时，我的老师提高了音量，继续讲道：“千万别有接收这个病人的想法！将他赶走！他很有可能是位精神病患者，而且你也帮不了他。”

多年以后，我才知道那位老师当时是在提醒我们，要提防病态的假象。精神病研究领域的权威也证实了我们那位老师所给的建议是合理的。这与缪勒-莱耶错觉类似。没人教过我们如何体会患者的心情。所以我们的老师断言，我们对那位患者的同情心是不由自主的，这种同情心可能源自系统1。此外，也没有人教过我们不要总是相信自己对患者的感情。有人告诉我们，过多关注一个有数次治疗失败经历的病人是一种危险的信号——正如平行线两端的箭头一样，会让人产生错觉。这是一种认知错觉。我的老师教过我（系统2）如何识别这种错觉，也曾告诉我切莫相信这种感觉，更不要依照感觉行事。

提到认知错觉，最常被问及的问题就是能否避免这种错觉。上述各例传达的信息不容乐



观。因为系统1是自主运行的，我们无法随意使其停止，因此直观思维所导致的错误常常难以避免。我们不可能一直没有成见，因为系统2可能对系统1产生的错误毫无所知。即使对可能发生的错误有所察觉，也需要系统2进行强有力的调控和积极的运作才有可能避免。然而，作为一种生活方式，时刻保持警觉性并不是一件好事，想要这样做也并不实际。总是质疑自己的想法会使我们的生活非常枯燥无味，因为系统2在代替系统1进行日常抉择时总是耗时很长且非常低效。最好的解决办法就是妥协：学会区别常会出现重大错误的情境，在风险很高的时候，尽力避免这些错误。前文中曾提到过，发现别人的错误总比发现自己的错误更容易。

## 两个虚拟出来的角色

你可以将上述两个系统想象为大脑内部的两个人，他们有着各自不同的特征、能力和局限性。在书中我常会将这两个系统当做一些句子的主语，例如“系统2统计产品数量”。

在我所在的职业圈子中，使用这样的语言被视为一种过错，因为这种说法貌似通过一个人头脑中那两个小人的思维和行为去解释这个人的思维和行为。从语法上讲，这个关于系统2的句子与“管家监守自盗”的句型类似。我的同事指出，管家的行为实际上就解释了现金的丢失原因，由此他们也很怀疑关于系统2的那个句子同样解释了算错产品数量的原因。我的回答是，这个短小的主动句认为统计是系统2所为，而且这个句子是一种描述而不是一种解释。这样的描述只有在你已经对系统2有一定的了解后才有意义。看看下面这段让人纠结的话吧：“心算是一个需要作出努力的自主活动，人在心算时总会瞳孔放大，心跳加速，因此你绝对不能一边开车左转，一边心算得出结果。”

同样，“通常情况下，在高速路上行驶的任务要留给系统1来完成”这句话说明驾驶这辆车转弯是自主行为，不用费脑力，也意味着一个驾车新手能够一边与别人攀谈，一边轻松地在空旷的高速路上行驶。最后，“系统2防止了詹姆士因为受到侮辱而做出愚蠢的举动”这句话则表明如果其努力自控的行为受到干扰（比如他喝醉了），詹姆士就应该会有过分的举动。

尽管系统1和系统2是本书的主题，但我必须要澄清一个事实，即两个系统并不是真实存在的，它们只是我杜撰出来的角色。系统1和系统2不是标准意义上的实体，没有错综复杂的组成部分，也不是大脑中某个固定的部位。你可能会问：为什么要在如此严谨的一本书中引进两个名字并不漂亮的虚拟角色呢？答案很简单，这两个角色很重要，因为我们每个人都会有些奇特的想法无法解释。一个句子若是讲一个客体（系统2）做什么，则要比对这个客体及其特征进行描述更容易理解。换句话说，“系统2”比“心算”更适合当句子的主语。我们

的大脑，特别是系统1，似乎拥有一些特别的能力，能够构想和领会施动者的故事。这些施动者有其特有的特征、习惯和能力。你很快就会对那个监守自盗的男管家产生坏印象，认为他还会继续偷盗，而且你暂时无法忘掉这个人。这也是我对这两个系统语言所寄予的希望。

为什么将两个系统命名为系统1和系统2，而不是“自主系统”和“耗力系统”呢？原因很简单：说出“自主系统”比说出“系统1”所需的时间长，因此会占用更多大脑工作记忆（短时记忆）的空间。这一点很重要，因为任何事物占用了大脑的工作记忆，都会削弱你的思考能力。你可以将“系统1”和“系统2”当做昵称，就像鲍勃和乔一样，用这种拟人的方式去了解整本书中出现的各种角色。有了这两个虚拟的系统，我便能更从容地思考有关判断与决策的问题，而你也能更轻松地读懂我的文字。

示例—关于系统1和系统2的说法

“他有印象，只是其中一部分是幻象。”

“这纯粹是系统1的反应，她在意识到危险之前就果断采取了行动。”

“这是你系统1的想法，放慢速度，听听系统2的看法吧。”

## 思考，快与慢

### 第二章 电影的主角与配角

虽说不太可能，但假设本书真的被拍成电影的话，自视为主角的系统2一定只能做个配角。在这个故事中，系统2的典型特征表现在其各项活动都需要努力，但其自身却很懒惰，除了必需的努力外，她不愿多付出，哪怕是一点点。因此，虽然系统2认为是自己选择了人们的想法和行为，可实际上，这些选择都是在系统1的引导下完成的，系统1才是这个故事的真正主角。然而，一些至关重要的任务却只有系统2才能执行，因为这些任务需要付出努力和控制自我，由此方可抑制系统1产生的直觉和冲动。

瞳孔是人类思维活动的灵敏指示器

如果想让你的系统2 全力运转，你可以做做下面的练习。这个练习会让你在5秒钟之内

达到认知能力的极限。首先，编一串不同的4位数数字，并将这些数字写在一张索引卡上。然后，在桌上放一张空白的卡片。你即将要执行的任务叫做加1，以下是其具体做法：

敲打出稳定的节奏（最好是有一个节拍器，并将其设定为一秒一拍）。

移动空白卡纸，大声读出数字。然后等待两个节拍，说出一个新的数字（这个数字是将原来那个数字的每一位都加1得来的）。例如：卡片上的数字是5294，新的数字就应该是6305。另外，跟上节奏很重要。

很少有人在加1任务中能胜任超过4位数的数字，但如果你想挑战一下自己，可以尝试一下加3的任务。

如果想知道大脑在快速运转时身体正在干些什么的话，你可以这样做：在书桌上堆两摞书，将你的下巴放在其中一摞上，将一台摄像机放在另一摞上。打开摄像机，在你做加1或加3任务时，盯着摄像机的镜头看。然后，你可以通过摄像机真实的记录发现，你的瞳孔大小会随着你的努力程度而变化。

很久以前，我就开始练习加1任务了。在我职业生涯的早期，作为研究催眠实验的访问学者，我在密歇根大学待了一年。在寻找有意义的研究课题时，我在《科学美国人》(Scientific American)杂志中看到了心理学家埃克哈特·赫斯(Eckhard Hess)的一篇文章。该文章指出，瞳孔是人类心灵的窗户。最近，我又读了一遍这篇文章，备受启发。赫斯在文章的开头说道，他的妻子注意到当他在观赏美丽的风景图片时，瞳孔会扩大。文章结束处有两张吸引人的照片，照片是同一个漂亮女人，但其中一张照片中的她显得比另一张中的更加漂亮。造成这种不同的唯一原因是：在更漂亮的那张照片中，女人的瞳孔比较大，而另一张的瞳孔比较小。赫斯在文章中还提到了颠茄(belladonna)——一种使人瞳孔变大的物质，曾作美瞳之用。作者还提到，一些常去赶集的人常常戴着墨镜，因为这样就能隐藏自己对商品的兴趣了。

赫斯的一个发现让我特别感兴趣。他发现瞳孔就像是大脑运转情况的灵敏指示器——它们在人们进行乘法运算时会扩散，在人们解决更为困难的问题时扩散得更大。他的观察还表明，对脑力工作的回应与唤起情感是不同的。赫斯的这篇文章与催眠关系不大，但我认为，“大脑活动是可以看见的”这一想法是个值得研究的课题。杰克逊·比提(Jackson Beatty)是实验室里的一个研究生，他对这个课题同样很感兴趣。于是，我们一起展开了研究。

比提和我设计了一个类似于验光仪器的装置，受试者可将头倚在可固定住下巴和前额的支架上，然后一边盯着镜头，一边听事先录好的问题，并跟着节拍器的节拍回答这些问题。

每一个节拍都会触发红外闪光拍照。在每期实验结束时，我们都会很快把照片冲洗出来，并将它们投影到屏幕上，然后用尺子测量瞳孔大小。这种方法对年轻人和没有耐心的研究者而言都很适用：我们能很快知道实验的结果，而且这些结果总能说明一些问题。

比提和我很关注有节奏的任务，例如在加1任务中，我们能准确地了解受试者每时每刻的大脑活动。我们记录了跟着节拍器说出的一串数字，并指示受试者在保持节奏的情况下，逐一重复或是转换这些数字。我们很快发现，瞳孔的大小会逐秒发生变化，这也就反映了任务的难度在不断变化。瞳孔随时间变化的曲线图最后呈倒V字形。如果去做加1和加3任务，你会发现每听到一个新数字，任务难度就会加大一些，最后达到一个几乎令人难以接受的极限。那时，你会在节拍中或停顿时极快地说出转换后的数字，这就相当于“释放”了自己的短时记忆，然后，你才渐渐感到放松了一些。瞳孔大小的数据与受试者的体验非常吻合：数字位数越多，瞳孔扩散得越大；任务的难度与付出的努力相符合；瞳孔扩散到最大的时候也正是付出努力最多的时候。与立刻重复一个7位数相比，4位数的加1任务会使瞳孔扩散得更大。加3任务则更为困难，这项任务是我所观察到的要求最高的任务。仅仅在前5秒钟，瞳孔就扩散了50%，心跳每分钟增加了7拍。这是一个人能达到的最大工作极限——如果超过这个极限，人们就会自动放弃。当我们给实验受试者的数字超过他们所能承受的范围时，他们的瞳孔就会停止扩散或是收缩。

我们在宽敞的地下室套间里工作了几个月，套间里有相关闭路系统，可以将受试者的瞳孔投影在走廊的屏幕上；我们同时还可以听到实验室里的情况。投射出来的瞳孔直径大约是一英尺；观察受试者工作时的瞳孔变化是件非常有趣的事，引得那些来我们实验室参观的人纷纷驻足。我们预测受试者何时会放弃任务，自娱的同时也给参观者留下了深刻的印象。在心算一道乘法题时，受试者的瞳孔会在几秒之内变大并保持那样的大小，直到她算出答案或是放弃。我们在走廊里观察这些瞳孔时，时常会让受试者和参观者感到惊讶。我们会问受试者：“为什么你刚才停下来了呢？”实验室里的人经常会问：“你是怎么知道的？”我们回答：“因为我们看见了你心灵的窗户。”

我们在走廊里随意的观察有时和正式的实验一样能说明问题。在两个任务的间隙，我随意看了一下某位女性的瞳孔。她把头放在了装置上，所以当她与实验人员进行例行谈话时，我能够观察到她的瞳孔变化。我惊讶地发现，她的瞳孔并没有伴随谈话和倾听而发生明显的扩散或收缩。与我们研究的任务不同的是，平常的谈话明显只需要一点努力或是完全不费力——不会比记住两位或三位数需要的精力多。这是灵感迸发的时刻：我意识到我们选择研究的任务全都是需要付出特别多努力的。我的脑中闪现一个想法：我们大脑的生活步调（现在我爱用系统2的生活步调来代替）大多像是在悠闲地散步，有时候会变成慢跑，只有在极少数数的情况下，才会如短跑冲刺。执行加1和加3任务时，大脑就像是在短跑冲刺；而平时随意的

聊天，大脑就如同在漫步。

我们发现，如果人的大脑正处于冲刺的状态，就有可能（对次要信息）产生有效的屏蔽。前文提到的《看不见的大猩猩》一书的作者就是通过让观察者持续专注于数传球次数而对那只“大猩猩”视而不见的。我们通过加1任务提供了一个不那么夸张的例子。当受试者在执行加1任务时，我们会给他们看一串快速闪过的字母。我们要求受试者对加1的数字任务给予充分的重视，但是在这个任务即将结束时，他们也需要说出字母 K 是否在整个实验中出现过。这个实验的主要发现是，人们锁定和报告指定字母的能力在执行任务的10秒钟内发生了变化。如果字母 K 出现在加1任务的开始或结尾，几乎所有观察者都不会错过，但如果字母 K 出现在大脑活动最为频繁的中间时段，就算彼时他们正睁大眼直直地盯着这个字母，也会生生错过它。没能发现字母 K 的线形图与瞳孔大小变化所呈现的倒 V 形是一致的，这种一致性再次证明：瞳孔是衡量与思维活动形影不离的生理刺激的标尺，我们可以通过瞳孔了解大脑的运行状况。

就像是我家或公寓外安装的电表一样，瞳孔提供了一个关于你大脑使用率的参数。这个类比还可以有更深入的解释。你的用电量取决于你用电来做什么，是开灯还是烤面包。当你打开电灯或是烤面包机时，你就会获得所需要的电量。同样，我们也能决定自己要做什么，但做成这件事得花多少精力我们就说不准了。假设你见到一个4位数，比如9 462，然后被告知，你的性命就取决于是否能在10秒内记住这个数字。无论你多想活下去，付出的努力也不会比用同样数字执行加3任务时付出的多。

系统2和你家里的电表能力都有限，但它们对超负荷的负载反应不同。当用电超负荷时，断路器会跳闸，致使那条线路上的所有电器都断电。相反，如果大脑的使用超负荷，其处理则是有选择性且精确的：系统2会偏向最重要的活动，因此这个活动会得到其所需的注意力，其他“多出来的”注意力再慢慢被分配到其他任务中去。我们所做的猩猩实验要求受试者更加关注数字任务。我们确信他们按要求完成了任务，因为可视目标（指“大猩猩”）出现的时候并没有对主要任务造成影响。如果那个关键字母 K 是在大脑活动量最大的时刻闪现的，受试者往往会将其屏蔽掉。而当数字转换任务要求并不那么高时，受试者就更有可能觉察到这个字母。

注意力这种精细的分配是在大脑漫长的进化过程中形成的。快速判断最严重的困难或者快速锁定最佳时机并做出迅速反应能提高生存概率。当然，这种能力并不专属于人类。即使在现代人中，系统1也会承担起应对突发情况的任务，完成自我保护的最高使命。试想开车时，车意外地滑到了一大片油区，你会发现，在充分意识到这一点之前，你就已经采取了躲避危险的行为。

比提和我在一起工作的时间只有一年，但我们的合作对于各自今后的职业生涯都产生了很大的影响。他最终成为“认知瞳孔测量法”的权威，而我则写了《注意与努力》(Attention and Effort)一书。这本书在很大程度上是以我们此前的共同研究为基础而写的，与我后来在哈佛大学所作的后续研究也密不可分。通过类型多样的任务来测量瞳孔大小，我们知道了许多关于大脑工作的知识（现在我都把工作中的大脑视为系统2）。

当你对执行一个任务越来越熟练时，需要付出的努力程度就会降低。对大脑的各项研究证明，与行动相关的活动模式会随着熟练程度的加强而变化，一些大脑区域将不再参与其中。天才也是如此。通过观察瞳孔变化和大脑活动，我们发现高智商的人往往需要较少的努力便可解决同样的问题。普遍的“最省力法则”不仅适用于体力活儿，还适用于我们的认知行为。这个法则主张，如果达成同一个目标的方法有多种，人们往往会选择最简单的那一种。在经济行为中，付出就是成本，学习技能是为了追求利益和成本的平衡。因为懒惰是人类的本性。

我们研究的这些任务对瞳孔变化的影响差别很大。从基本水平来看，我们的受试者都是清醒的、有意识的，并时刻准备好投入到任务中去——也许觉醒水平和认知准备比平时还高、还充分。记住一位数或两位数或是学会将数字与词汇相联系（比如 3=门）会对基准线以上的记忆觉醒产生确切的效果。但是，这样做收效甚微，只有5%的瞳孔直径增大与加3任务有关。鉴别两个音调高低的任务也能有效地使瞳孔扩大。最近的研究还表明，抑制住自己读出干扰性单词的倾向同样会产生一定的效果，而在短时间内记住6位或7位数字则需付出更多精力。如你所体验到的那样，当被要求说出你的电话号码或是爱人的生日时，你需要作一番简单却重要的努力，因为你的回应是有逻辑的，你必须将整串数字记在脑中。而心算两位数的乘法题和加3任务则已经接近人们能够做到的极限。

为什么我们看不见那只大猩猩？

是什么原因使某些认知任务较其他的更加困难、更需付出努力呢？若注意力是种货币，那我们要买些什么样的产品呢？什么又是系统2能做而系统1不能做的呢？我们现在对这些问题给出假设性的回答。

想要同时记住不同的想法也需要耗费精力，其中有些想法需要按不同方案实施，另一些想法则需与一定的规则结合起来实施——在进超市前重新核实你的购物单，在餐馆吃饭时，在鱼和牛肉之间进行选择，或是根据小样本得来的信息归纳出一个令人惊喜的结果等，都属于此类实例。系统2是唯一一个可以按规则运行、能根据属性来对比物品、能深思熟虑作出选

择的系统。自动运行的系统1不具备这些能力。系统1能察觉简单的关系（比如“他们长得一模一样”，“儿子比父亲高得多”），还擅长整合关于一件事的所有信息，但不能快速处理多个独立的话题，也不能利用纯粹的统计学信息。如果一个人被描述成“本性怯懦，做事井井有条，循规蹈矩，关注细节”，系统1就会认为这个人像是个图书管理员，但系统2在结合了直觉以及图书管理员人数少的这个事实进行思考过后，却不会这样认为。只有系统2才能作这种判断。

系统2一个非常重要的才能是它能够处理“多重任务”，它可以提取记忆去执行抑制习惯性反应的指令。考虑以下的任务：数出这一页“的”字出现的次数。这个任务你以前从来没有做过，做起来很难得心应手，但是你的系统2却可以应付得来。着手这个练习并非易事，尽管在练习的过程中你会有所提高，但真正完成这项任务会很吃力。心理学家用“执行控制”来描述多重任务的执行和最终完成，神经系统科学家已经确认了大脑中负责执行功能的主要区域。当有冲突需要平息的时候，其中一部分区域也会活跃起来。另一部分是大脑前额叶，人类的这个区域要比其他灵长类的更为发达，它是与智力密切相关的重要脑区。

现在，假设你在看完这一页时，接到了另一个指示：数出下一页有多少个逗号。这项任务更加困难，因为你还要克制住不久前形成的倾向，即将注意力集中在“的”字上。近几十年来，认知心理学家们有很多重大发现，其中一项就是：从一个任务转换到另一个任务上需要付出努力，在时间紧迫的情况下尤其如此。完成加3任务和心算乘法之所以困难，也是因为两项任务都需要快速转换。要完成加3任务，你必须同时在工作记忆中储存好几个数字，并且每个数字都要与一个特定的运行过程相联系：得记住转换完的数字以便稍后说出来，一个数字正在转换中，还有一些数字正等着被转换。当前关于工作记忆的测试要求个人在两个高要求的任务间不停地转换，在记住其中一个结果的同时，还要执行另一个任务。能够很好完成这些测试的人大都能在一般智力测试中取得好成绩。然而，是否能够控制自己的注意力并不是一般智力的衡量标准。要想预测空中交通指挥员和以色列空军飞行员的表现，衡量他们控制注意力的能力比让他们作智力测试更为有效。

时间制约是人们付出努力的另一个驱动因素。执行加3任务时，你的匆忙一方面是因为节拍器，另一方面是因为记忆负荷。你就好比是同时向空中抛出好几个球的马戏团演员，无法承担减速的后果。记忆减退的速率催促你的步调，迫使你在完全忘记这些信息前不断进行更新和演练。任何需要你同时记住许多想法的任务都是匆忙的。除非你运气较好，有很大的工作记忆容量，否则你就只能硬着头皮继续工作。慢思考最耗费脑力的思考形式就是那些催你思考的形式。

你肯定已注意到，在执行加3任务时，你的大脑会不同寻常地高速运作。即使你靠脑力

劳动谋生，在日常工作中也极少有类似加3或是类似马上记住6位数数字这样极具挑战性的任务。我们通常会分几个简单的步骤来执行任务，以避免大脑超负荷运行。这样的话，我们可以将中间结果储存在长期记忆中或是记在纸上，而不是简单地堆积在工作记忆中。我们不紧不慢地绕着远路向目标靠近，通过最省力法则来管理我们的思维活动。

#### 示例— 注意力和努力的问题

“我不会在开车的时候想破脑袋去解决这个问题。这是一个会使瞳孔扩散的任务，太费神了！”

“他在运用最省力法则，能不多想就不多想。”

“她没有忘记开会的事，只是会议开始时她完全在想别的事情，根本没有听到你在说什么。”

“我的脑海中最先出现的想法是来自系统1的直觉。我必须得从头到尾思索一遍，三思而后行。”

## 思考，快与慢

### 第三章 惰性思维与延迟满足的矛盾

我每年都要在伯克利待上几个月，在那里，我最大的乐趣就是每天在山间小路上散步4英里，领略旧金山湾的风景。通常我会记录散步所用的时间，也由此对自己在这个过程中所付出的努力有相当的了解。我发现自己大概用17分钟就可以走完1英里的路程。当然我也耗费了体力，以这一速度行走比我坐在靠椅上要消耗更多的热量，但行走中我并没有感受到精神压力，也没有内心矛盾，更无须催促自己前行。以这个速度散步，我还能边走路边思考。事实上，我觉得散步能唤醒身体的感应，使大脑思维更加敏锐。

系统2也有一个自然的速度。大脑没有专门处理某项任务时，你可以分些精力随意观察自己周围发生了什么。除非你非常小心谨慎，或者自我意识很强，否则观察周围环境或大脑的活动是不需要付出太多努力的。开车时你就能做出一些小决策，读报纸时也能汲取一些信息，和爱人或同事随意说说每天的开心事等，这些都不需要付出多少努力，也没有什么压力，跟散步没什么两样。



边散步边思考其实是一件很轻松、很惬意的事，但在某些极端情况下，这些活动似乎在争夺系统2有限的资源。只需一个简单的实验就可证实这个假设：在和朋友悠闲地散步时让他心算出“ $23 \times 78$ ”的结果，而且要立刻就算出来，这时他肯定会停下脚步来算。我的体会是，我可以在散步时思考，却无法利用短时记忆来完成这样一项复杂的心算任务。如果我必须在规定时间内构建一个复杂的理论，我希望无人打扰，而且坐着思考要比站着强。当然，不是所有的慢思考都必须集中精力、认真计算的。和阿莫斯悠闲散步就是我人生中的最佳思考时间。

加快散步速度会完全改变我的散步体验，因为加快速度会使我的连贯思考能力明显下降。只要一提速，我就要注意逐渐加快行走速度，要刻意保持更快的速度，将一连串想法加以总结的能力便相应下降了。我在山上行走能保持的最快速度是每14分钟走完1英里，不过这样一来，我根本什么事都想不了。沿着小路快速行走不仅要付出体力，还需要大脑的自我控制，以防止自己减速。自我控制和仔细思考很明显要抢夺努力的有限预算。

通常情况下，大多数人保持连贯的思维或时不时积极思考都需要自我控制力。尽管没有作过系统的研究，但我认为，不断转换任务和提高大脑运转速度从本质上说是不会让人感到快乐的，人们总是尽可能避开这种情况，这就说明了为什么最省力法则能成为法则。即使没有时间的限制，保持连贯的思维也需要此法则。有人曾观察并记录我写作的一个小时内查收电子邮件或打开冰箱的次数，这可以说明我想要逃离写作的欲望，也可以得出一个结论——我的自我控制力完全达不到工作的要求。

好在并不是所有认知工作都令人厌恶，有时并不需要意志力的支撑，人们也能花很长时间和大量精力进行一项工作。心理学家米哈里·契克森米哈（Mihaly Csikszentmihalyi）对这种无须作出努力的状态的研究比别人都多，他将这种状态命名为心流，而且此名称已成为一个心理学术语了。体验过心流的人将其描述为“一种将大脑注意力毫不费力地集中起来的状态，这种状态可以使人忘却时间的概念，忘掉自己，也忘掉自身问题”，他们对这种状态所带来的愉悦感的描述非常吸引人，契克森米哈称之为“最优体验”。很多活动都能带来心流体验，不论是画画还是摩托车比赛。我认识几个作者，出书就是他们的最优体验。对于一个作者而言，这样容易满足是件幸事。心流巧妙地区分了两种努力形式：对任务的关注和对注意力的严格控制。以每小时150英里的速度骑摩托车和在象棋大赛中角逐都需要付出努力，然而在心流状态下，集中注意力关注吸引人的事并不要求自我控制。因此，我们要将所有资源都用于手头上的任务才好。

又累又饿的保释官更可能否定保释申请

自我控制和认知努力是大脑工作的形式，这一观点已得到广泛认同。有几项心理研究表明，人若既有认知任务在手又同时受到诱惑的影响，就容易屈从于诱惑。如果有人要求你在一两分钟内记住一串7位数的数字，并且告诉你记住这些数字是你的首要任务，而当你将注意力集中在这些数字上时，却有个人端着甜点让你选一种：是选择让人既爱又恨的巧克力蛋糕呢，还是选择什锦水果沙拉。有证据显示，尽管大脑里装满了这些数字，你却更有可能选择诱人的巧克力蛋糕。系统2在忙碌时，系统1对行为的影响会更大。而且，系统1也更偏爱甜食。

当人们忙于认知活动时，更有可能作出自私的抉择，会用带有性别歧视的字眼，并在社交场合作出肤浅的评判。记住和重复这些数字会减轻系统2对行为的控制，当然，认知负担不是自我控制减弱的唯一因素。喝几杯酒，或者一夜没睡也会产生同样的结果。早起的人的自我控制力会在晚上受到影响，而夜猫子的自我控制能力则会在早晨受到影响。过多关注自己完成一项任务的结果，就会给其短时记忆增加毫无意义的思想负担，进而影响其整体表现。结论非常明显：自我控制需要集中注意力，需要付出努力。换种说法就是，控制思想和行为是系统2的任务之一。

心理学家罗伊·鲍迈斯特（Roy Baumeister）和他的同事们所做的一系列令人惊讶的实验最终表明，所有自主努力的不同形式——认知上的、情感上的或者身体上的——至少都能对集思广益有所帮助。他们的实验中要求受试者进行的是连续性任务而不是同时发生（不相关联）的任务。

鲍迈斯特的小组屡次发现，刻意掌控意志和进行自我控制很辛苦。如果你必须强迫自己去某件事，而此时这件事又面临一个新的挑战，你就会很不情愿或是根本无法进行自我控制。这种现象被命名为自我损耗（ego depletion）。在一次典型的展示活动中，我们要求受试者一边看一部能引起感情共鸣的电影，一边抑制自己的情绪反应。在随后的耐力测试中，他们表现得很糟糕。该耐力测试的内容是握住测力计——这个动作会让人越来越不舒服——看他们能保持多长时间。受试者在实验的初始阶段作出的抑制情感的努力，会削弱其忍受维持肌肉收缩而带来的痛苦的能力，自我损耗型的人因此会很快产生退出实验的冲动。在另一项实验中，受试者要首先经过自我意志的损耗，他们会吃不同的食物，包括小萝卜、芹菜等，同时还要抑制住吃巧克力和饼干的想法。后来，当这些人面对困难的认知任务时，会表现得比通常情况下更轻易放弃。

如今，我们知道关于自我控制的情形和任务很多，也很复杂。这些情况既包含思想斗争又要抑制自身倾向的需求，见下面的例子：

不去想北极熊

对震撼人心的电影抑制情感的共鸣

对矛盾的事作出选择

试着让他人眼前一亮

对同伴不好的行为委婉回应

和其他种族的人交流（这些人还有种族偏见）

自我损耗的前兆也多有不同：

改变日常饮食

疯狂购物，花很多钱

反应过度，有挑衅的意味

对有把握的任务花费较少的时间

在认知任务和逻辑决策的制定中表现得很糟糕

其证据很具有说服力：对系统2有高需求的活动同样需要自我控制，而发挥自我控制力既有损耗又很枯燥。与认知负担不同，自我损耗至少会令人丧失一部分动力。在一项任务中控制自我后，在另一项任务中就感受不到自己在努力，但只要你真的想做，就一定能做到。在几次实验中，如果受试者有强大的动力抑制自我损耗的影响，他是完全能够做到的。相反，如果你必须在执行某项任务的过程中以短时记忆记下6个数字，此时即使不断努力也是行不通的。自我损耗和认知投入并非同一种思维状态。

鲍迈斯特小组最惊人的发现，用他自己的话说就是，大脑能量这一概念不仅仅是个比喻。神经系统消耗的葡萄糖比身体其他部位消耗的都要多，而且需要付出脑力活动的成本显然要比葡萄糖高。在积极进行复杂的认知推理或者忙于要求自我控制的任务时，人的血糖就会下

降。这种情况和短跑运动员在短跑时肌肉中的葡萄糖储备量下降是相同的。这一概念的大致含义是，自我损耗的影响能通过注射葡萄糖得到缓解，而且鲍迈斯特和他的同事们也通过几个实验证实了这一假设。

他们有一项研究是让志愿者们看一个无声的短片，其内容是对一位女士的采访。这些志愿者的任务是解释她的肢体语言。在这些志愿者执行这项任务时，屏幕上有一串单词慢慢闪过，而他们事先已经被告知要忽略这些词，如果发现自己的注意力转移了，他们只能再次将注意力拉回到这位女士身上。研究认为，这种自我控制行为可以引起自我损耗。所有的志愿者在执行第二个任务前都喝了些柠檬汁，其中一半的人所喝的柠檬汁加了葡萄糖，而另外一半人的饮料中混有代糖。然后所有的志愿者开始执行第二项任务，在这项任务中，他们必须克服直觉才能得到正确的答案。通常在自我损耗的人群中，直觉性的错误常有发生。结果是喝了含有代糖饮料的人出现了损耗现象，而喝了含有葡萄糖饮料的人却没有。在大脑中储存一定量的糖，可使自己表现得不那么糟糕。能导致葡萄糖降低的任务是否同样会导致瞳孔放大和心率加快等情形发生呢？要想证明这一说法，恐怕还需要些时间和更深入的研究吧。

《美国科学院院报》最近报道了损耗对判决的影响这一问题的研究情况。在这项研究中，8位不知情的受试者全部是以色列的保释官。他们每天都要审阅保释申请，而且不是按这些保释申请的时间顺序审阅的，这些保释官在每份申请上所用的时间很少，平均只有6分钟。（弃权相当于拒绝保释，只有35%的申请能获准通过。这些保释官作出每个决定所用的时间都有精确的记录，而且他们一日三餐的餐歇时间也有记录，分别是早餐时间、午餐时间和午间休息时间。）这项研究的设计者对两次餐歇间所能获得批准的申请数量进行了预测，结果发现每次用餐过后，获得批准的申请数量都会增加，有约65%的申请得到了批准。在保释官下一次用餐前的约两个小时内，批准率就开始稳步下降，在用餐之前刚好达到零。如你预想，这种结果令人难以接受，但各位设计者已经认真审核并排除了很多其他的原因。对这个数据最合理的解释却带来了负面信息：又累又饿的保释官容易否定保释申请。疲劳和饥饿都有可能影响他们的决定。

### 脱口而出的错误答案

系统2的一大主要功能是监督和控制思想活动以及由系统1引导的各种行为，使得一些想法直接体现在行动上，或者抑制或改变其他想法。

例如，下面是一个相对简单的难题。别费力去分析它，凭直觉做做看：

球拍和球共花1.10美元。

球拍比球贵1美元。

问球多少钱？

你会马上想到一个数字，这个数字当然就是10，即10美分。这道简单的难题之所以与众不同，是因为它能引出一个直觉性的、吸引人的但却错误的答案。计算一下，你就会发现。如果球花费10美分的话，总共就要花1.20美元（球10美分，球拍1.10美元），而不是1.10美元。正确答案是5美分。我们可以假设那些最终得出正确答案的人也想到了这个答案，只是他们不知通过什么办法成功抵制住了直觉的诱惑，最终给出了正确的答案。

肖恩·弗雷德里克（Shane Frederick）和我对基于两个系统的判断理论进行了共同研究，他用这个球拍和球的问题来研究一个核心问题：系统2对系统1的各类方案监视得有多严密？他进行了这样的推理：那些说球为10美分的人让我们了解到这样一个重要的事实，那就是这个人没有认真验证这个答案是否正确，而且他的系统2倾向于直觉性的答案。其实他只要稍稍动脑想一下，就会否定这个直觉性的答案。此外，我们还知道给出直觉性答案的人忽视了一个明显的生活提示，他们应该想一想怎么会有答案这么明显的问题呢。没有验证答案，这是一个明显的失误，因为验证根本不费什么事，只是大脑工作几秒钟（这个问题的难度一般），或者肌肉动一动，抑或睁大眼睛而已，只要稍微动动脑子就可以避免这个令人窘迫的错误。认为答案是10美分的人显然不爱动脑筋，没有给出这个错误答案的人显然思维更活跃。

上万名大学生都回答了这个问题，其结果令人吃惊。哈佛大学、麻省理工学院和普林斯顿大学中50%以上的学生给出了这个直觉性的错误答案。在声誉稍差一点的大学里，则有80%以上的学生没有验证答案就脱口而出。这个球拍和球的问题是我们所作的第一个作了观测性的研究，这一问题将是本书反复出现的主题：很多人过于自信，过于相信自己的直觉。他们显然觉得认知努力没什么意思，会尽量避免费力思考。

现在我来为大家作一个逻辑论证——两个前提和一个结论。请你尽量快速地判定这个论证是否符合逻辑。想想结论是否与前提有关联？

所有的玫瑰都是花。

有些花会很快凋谢。

因此，有些玫瑰也会很快凋谢。

大部分大学生觉得这个推论是合理的。但事实上，这个论证是有问题的，因为玫瑰可能不会很快凋谢。就像球拍和球的问题，貌似正确的答案会马上在大脑中显现。摆脱这种现象需要作出努力，因为坚信自己的观点，认为“这是对的，是对的”，便使人很难去验证自己的答案是否符合逻辑，大多数人根本不愿费力去想明白这个问题。

这个实验使我们对在日常生活中做出的推论丧失了信心。这个实验使我们认为，当人们相信某个结论是正确的时候，他们很可能会相信支持这个结论的论证，哪怕这些论证不正确。如果系统1也参与到活动中来，人们总会先得出结论，然后才进行论证。

接下来，请思考以下问题，边读边作答：

密歇根州在一年中发生了多少起谋杀事件？

这个问题是肖恩·弗雷德里克设计的，仍旧是对系统2的挑战。回答这个问题的“诀窍”在于调查对象是否会记得底特律这个犯罪率相当高的城市就在密歇根州。美国的大学生都了解这个事实，都能准确无误地指出底特律是密歇根州最大的城市。然而，对事实的了解并不重要，重要的是在我们需要这个事实的时候却总是无法立即将其提取出来。记得底特律在密歇根州的人对该州谋杀犯罪率的估计要多于不知道这一点的人，但弗雷德里克那个实验的大多数受试者在看到密歇根州时，根本没有想到该州还有底特律这座城市。实际上，当人们被问及密歇根州和底特律市的谋杀犯罪率时，回答前一个问题的人总认为密歇根州的谋杀犯罪率低，而回答后一个问题的人则认为底特律市的谋杀犯罪率高。

受试者在回答上面的问题时没能想到底特律市，这既是系统1的问题，也是系统2的问题。提到密歇根州时能否想到底特律市，这在一定程度上要依赖记忆的自主功能，这一功能因人而异。有些人对密歇根州的一切耳熟能详：该州的居民比其他地方的居民更容易回想起关于该州的事实；熟悉地形的人能比棒球统计员回想的更多；聪明的人在大多数事情上的表述上比其他人表述得更好。聪明不仅是指推理的能力，也指在记忆中搜寻相关信息和在必要时调动注意力的能力。记忆功能是系统1的一种属性。但是，每个人都可以选择放慢速度，在记忆中积极搜寻所有可能相关的事实，就像他们在球拍和球的问题中可以放慢速度验证直觉性答案一样。不同的人仔细核对和搜索的程度也各不相同，这是系统2的一个特征。

球拍和球的问题、玫瑰的问题和密歇根州/底特律的问题都有共同点。在这些小实验中，答错问题显然在某种程度上说明了其原因是精神动力不足，而不是努力程度不够。能被好大学录取的学生当然能够在前两个问题上作推理判断，在密歇根州的问题上也能够进行思考，

能回想起该州的大城市及其犯罪问题。这些学生只要不想接受脑海中出现的似乎是正确的答案，就能够解决更难的问题。若他们满足于现成的答案而不想去思考则会很麻烦。用“懒惰”来形容这些年轻人的自我检测及其系统<sup>2</sup>似乎有些刻薄，却并非不公平。避免思维上懒惰的人可以被叫做“勤快人”。他们更机警，思维更活跃，不会满足于貌似正确的答案，对自己的直觉也常持怀疑态度。心理学家基思·斯坦诺维奇认为他们更理性。

今天得到1 000美元，1年后得到1万美元，你选哪个？

研究者曾经用各种方式来检验思考和自我控制之间的联系。有些研究者通过询问相关问题来论证其联系：如果分别根据自我控制能力和认知能力将人们分成不同等级，那么不同个体在这两个排序中是否会处于类似的等级呢？

在心理学发展史中有个很著名的实验，瓦特·米舍（Walter Mischel）和他的学生将一些4岁大的孩子置于残酷的两难处境中。这些孩子可以自行选择，是要一个随时就可以拿到的小奖励（一块奥利奥饼干），还是在充满考验的环境中苦等15分钟，然后得到更大的奖励（两块小甜饼）。每个孩子得单独待在一个房间里，面前有张桌子，桌子上有两件东西：一块饼干和一个铃铛。孩子可以随时摇铃通知研究人员并能领到一片饼干。根据描述，实验如下：“玩具、书、图画或任何其他会使孩子们分心的东西都不在房间里。研究人员离开房间15分钟后才会回来。不过，孩子要是摇了铃、吃掉饼干、站起身来或者表情痛苦，研究人员也会回来。”

研究人员通过一面单面镜观察这些孩子，他们等待期间的行为总会使观察者大笑。有些孩子成功地经受住了15分钟的考验，其成功原因是他们能把注意力从诱人的奖励上移开。10年或15年之后，那些忍住了诱惑和没忍住诱惑的孩子之间会出现很大差别。忍住了诱惑的孩子在认知任务——尤其是高效地重新分配注意力方面的控制力更强。当他们还年轻时，他们染上毒品的可能性更小。智力水平的巨大区别也随之出现：在4岁时表现出更强的自我控制能力的孩子在智力测验中得到了更高的分数。

俄勒冈大学的一个研究小组从几个方面研究了认知控制和智力之间的联系，包括通过提升对注意力的控制力来提升智力的做法。研究实验由5个阶段构成，每个阶段40分钟，他们让4~6岁的孩子去玩不同种类的电脑游戏，这些游戏是为检验注意力和控制力而专门设计的。在其中一项游戏中，孩子们要在尽量避开一块泥泞区域的同时，用一个操纵杆将一只卡通猫驱赶到一个长满草的地方。长草的地方会慢慢变少，而泥泞的地方会慢慢扩大，这就需要孩子们不断提高控制的精确度。测试者发现，注意力训练不仅提升了这些孩子的执行控制能力，而且他们在智力测验中的笔试成绩也提高了，并且这种提升的状态可以维持几个月不变。同

一个小组所作的其他研究确认了参与了注意力控制的具体基因，表明家长的教育方法也会影响这种能力，研究还证明了，儿童控制其注意力的能力和控制其情感的能力之间有着紧密的联系。

肖恩·弗莱德里克组织了一个认知反应测试，其中包括球拍和球的问题以及其他两个问题，之所以选这些问题是因为它们可以引发一个吸引人但却错误的直觉性答案（我们将在第5章中讨论这些问题）。他还对这项测试中得分很低的那些学生的特点进行了跟踪研究，发现他们更倾向于用脑海中最先出现的想法来回答问题，而不愿意费事去验证自己的直觉。这些人身上的系统2的监测功能往往较弱。那些遇到各种难题都跟着直觉走的人也易于接受系统1的其他暗示，特别是在冲动、不耐烦以及急功近利的时候，这些人就更易接受系统1的各种暗示了。例如，63%的直觉型受试者说他们宁愿这个月拿到3 400美元也不愿等到下个月拿3 800美元。而那些正确答出3道难题的人中只有37%持这样目光短浅的看法——希望能马上得到一笔钱，哪怕少些也可以。当被问到要想次日就收到自己订购的书愿意付多少钱时，在认知反应测试中得分低的人比得分高的人愿意多付一倍的价钱。弗莱德里克的发现表明，我们这场心理剧中的两个角色有着不同的“人格”。系统1是冲动、凭直觉的；而系统2则具备推理能力，它很谨慎，但对一些人而言，这个系统也是懒惰的。我们从不同人的不同特点中发现了相关性：有些人倾向于系统2，而有些人则更接近于系统1。这个简单的测试可以较好地反映出惰性思考的特点。

基思·斯坦诺维奇与其长期合作者理查德·韦斯特（Richard West）首先提出了系统1和系统2这两个术语（他们现在更喜欢将其称为第1类型过程和第2类型过程）。斯坦诺维奇和他的同事们用了几十年的时间来研究不同个体面对各种问题时的不同反应，这也是本书所关注的问题。他们用多种不同的方式问受试者同一个基本问题：为什么有些人比其他人更容易受判断成见的影响？斯坦诺维奇在《理性和反思性思维》（*Rationality and Reflective Mind*）一书中阐明了他的观点，对书中相关章节的主题进行了大胆而独到的论述。他对系统2的两个部分作了明确区分，这一区分十分明显，斯坦诺维奇称其为泾渭分明的两种“思维”。其中一种思维（他称其为算法）负责的是慢思考和要求很高的计算活动。有些人在这些脑力活动中比他人做得更好，他们在智力测试中超越他人，并且从一项任务转换到另一个任务上时，他们往往更快、更高效。不过，斯坦诺维奇认为，高智商并不能消除成见。要想消除成见，还需具备另一种能力，他称其为理性。斯坦诺维奇对理性之人的定义和我之前说到的“勤快人”有相似之处。他的核心观点是我们应当将理性和智力区分开来。在他看来，肤浅的或者惰性思考是一个反思缺陷，是一个理性错误。这是个引人注目且发人深省的想法。为了论证自己的想法，斯坦诺维奇和他的同事们发现，在某种情况下，球拍和球的问题以及此类问题比传统的智力测验（比如智商测试）更能反映出我们对认知错误的敏感度。关于智力和理性之间的区别是否会引出更多的新发现，时间最终会告知我们答案。



示例—自我控制的问题

“连续工作几小时，她也不会感到吃力，她处于一种‘心流’中。”

“在长达一天的会议之后，他的自我意识出现一定程度的损耗。因此他决定采用标准的操作规程，不再去想这个问题了。”

“他从来不去想自己的话是否有道理。他是特别习惯用懒惰的系统2呢，还是总是非常累？”

“不幸的是，她总是喜欢凭直觉随口就说，也许连表达感谢都词不达意吧，弱弱的系统2啊。”

## 思考，快与慢

### 第四章 联想的神奇力量

系统1的运行机制出人意料，在对其进行研究前，请先看以下两个词：

香蕉 呕吐

在刚才的一两秒钟里，你一定想到了很多，脑海中浮现出一些不愉快的图像和记忆，你的脸有些扭曲，露出厌恶的表情，而且你还可能不自觉地把这本书推得更远些。你的心率加快，手臂上的汗毛微微立起，而且你的汗腺开始分泌汗液。总之，你对那个令人反感的词的反应与对真实情况的反应相差无几。这些反应全是自发的，超出了你的控制范围。

你的大脑会不由自主地将香蕉和呕吐这两个词联系起来，暂时在二者间建立了因果联系，认为是香蕉引起了不适。这就是大脑的自然反应。结果，短期内你会对香蕉失去兴趣（不过别担心，这种感觉总会消失的）。你的记忆状态在其他方面也有所改变：你现在很容易就能识别出与“呕吐”相关的物品和概念，对它们的反应也很敏感，比如不舒服、臭味或者恶心等；看到和“香蕉”有关的词，比如黄色、水果，甚至连苹果和浆果都包括在内，也会有不良反应。

呕吐一般只在特定的情况下才会发生，比如宿醉过后和消化不良时。你肯定还能找出导致呕吐的其他相关词汇。而且，你的系统1注意到将这两个词相提并论并不常见，以前从没遇到过这种情况。你自己也有些惊讶。

这些复杂的反应快速地呈现在你的脑海中，而且都是自主发生的，无须费力。你左右不了它，也不能让它停下来。这是系统1的一个运行过程。你看到这些词后出现的一切反应都是循着一个名为“联想激活”的过程发生的：事物在你的大脑中唤起的想法激发出许多其他的想法，而且这些联想的行为在你的大脑中迅速扩展开来。连贯性是这种复杂的思维活动的重要特点，其中每个环节都是紧密相连、相互支持的。能引发记忆的词也会引发情感，还能引发面部表情变化和其他反应，比如常出现的紧张和回避倾向。面部表情和退缩行为强化了引起这两种反应的情感，这些情感反过来还会强化相应的概念。所有这些都是瞬间发生的，形成一种认知、情感和生理反应的自我强化模式，这种模式变化多样又能形成一个整体，被称为联想的连贯性。

在一秒钟左右的时间里，你就能自主且无意识地完成一件了不起的事。一旦发生了完全出乎意料的事——两个风马牛不相及的简单的词语被放在一起——你的系统1就将这两个词随意联系起来，试图弄清具体情况；还预估了可能会出现（从轻度到中等程度）的危险状况，帮你作好准备去面对这些极有可能会出现的的事件，以便为你将来的应对创造条件。同时，系统1还会对过去情形的骇人程度进行评估，为当前的事件发展创造条件。这样你就可以透彻了解过去，从容面对未来了。

在所发生的事情中有一点很奇怪，那就是你的系统1把两个词的简单联系看做真实的情况。你的身体对这一假想情形的反应要稍逊于对真实情况的反应，情感上的反应和生理上的反感表现可以从某个方面解释这个现象。正如近几年来许多认知科学家强调的那样，认知可以体现出来，你不只是用大脑思考，还用身体思考。

很久以来，人们一直都知道引起这些思维活动的机制是什么，这个机制就是联想。我们都能从自己的经历中了解到，有意识的思维活动中所有的观点都是井然有序的。17世纪和18世纪的英国哲学家曾经努力探求能解释这些续发事件的规律。在1748年出版的《人类理解研究》（*An Enquiry Concerning Human Understanding*）一书中，苏格兰哲学家大卫·休谟（David Hume）将联想的原则缩减为三个：相似性、时空相接以及因果关系。自休谟时代以后，我们对联想的定义发生了巨大变化，但他的三原则仍可作为定义所有联想的基础。

什么是观点？我愿意接受开放性的看法。它可能是具体的，也可能是抽象的，能通过很多方式表达出来：它可以是动词、名词、形容词或者只是握紧的拳头。心理学家认为，观点

是一张巨网上的节点，他们称其为联想记忆，这些节点和其他节点相联结。联结的类型多种多样：因果联系（病毒→感冒）；事物及其特性的联系（柠檬→绿色）；事物及其种类的联系（香蕉→水果）。有一方面我们超越了休谟，那就是我们不再将思维活动视为一连串顺序出现的有意识的观点。联想记忆的工作原理是什么？当前的研究中有这样一种见解：一瞬间会发生很多事。思维活动唤起一个看法不仅会引发另一个看法，它还会激发出很多其他看法，而这些看法还会让我们想到另外一些看法。此外，只有几个被激发出来的看法是有意识的思维活动；多数联想思维都是无声的，隐藏在有意识的自我之后的。如果说我们对大脑的运转认识有限，我们当然很难接受这种说法，因为这种说法和我们的体验背道而驰，但事实就是如此：你觉得自己很了解自己，但其实你错了。

启动效应：让人不知不觉微笑的铅笔

突破是科学研究的重点，而我们理解联想机制的第一个重大突破就是对测量方式的改进。几十年前，唯一研究联想的方法是向众人提问，比如“当你听到‘天’这个词时脑海中最先出现的词是什么”，研究者记录了不同回答出现的频率，比如“夜晚”、“晴天”，或者“日间长”等。在20世纪80年代，心理学家发现，看到一个词时，人的思维就会立刻产生变化，而且这种变化是可测量的，同时，很多相关词汇会被激发出来。如果你最近看到或者听到“喝\_\_”这个词，你可能会填上偏旁为“氵”的“汤”字而不会填成“场”。当然，如果你刚看到了“运动\_\_”这个词，你可能就会填上“场”字了。我们将此现象称为启动效应，也就是说“喝”这个概念在“汤”之前出现，“运动”先于“场”出现。

启动效应有多种表现形式。如果你的脑海中此时有“喝”这个概念（不论你是否注意到这一点），恰好此时有人低声跟你说了“汤”字或者在你前面远远的地方放上一碗汤，你肯定会比平时更快地意识到“汤”这个词。当然，你脑海中首先出现的不仅有汤这个概念，还有与食物相关的很多概念，包括猪肉、饿、脂肪、饮食和饼干等。如果最近用餐时几乎都是坐在摇晃的餐桌前吃饭，你的脑海中还会首先浮现摇晃这个概念。此外，启动联想的概念也能引起其他概念，尽管这种能力不算强。就像是池塘里的涟漪一样，概念的激活也是由联想概念这张大网上的一小部分向外逐渐展开来的。这个涟漪效应现在是心理学研究中最令人兴奋的探索之一。

关于记忆理解的另一项重大突破，是研究者发现启动效应不只限于概念和词汇。当然，单纯通过有意识的经验，你是无法了解到这一点的，但你一定要接受这个另类的想法，即你的行为和感情有时会受制于你自己甚至都没有意识到的事件。在一项实验中，心理学家约翰·巴奇（John Bargh）和他的同事们让纽约大学的数位学生从一个包含5个单词的词组中（例如“发现、他、它、黄色的、马上”）挑出4个单词来重组句子。其中一个小组的学生重组的

句子中有一半都含有与老年人相关的词汇，例如佛罗里达州、健忘的、秃顶的、灰白的或者满脸皱纹的。当他们完成这项任务时，又被叫到大厅另一头的办公室里去参加另一个实验。从大厅的一头走到另一头是这次实验的关键所在。研究者悄悄地测量了他们所用的时间。正如巴奇预料的那样，那些以老年为主题造句子的年轻人比其他人走得要慢得多。这个实验后来成了经典案例。

这个“佛罗里达效应”包括信息启动的两个阶段。第一，尽管没有人提过“老年”这个词，但上述那组词令人想到了年迈；第二，这些想法催生了一种行为，即缓慢行走，这个行为与老年人相关。所有这一切的发生都是无意识的。这些学生在回答问题时，谁也没有提及自己注意到这些词有什么共同点，而且他们都坚持认为自己在第一个实验中看到的那些词并未对实验后的行为产生什么影响。“老年”这个概念并非他们的自觉意识，但其行为却因此有了改变。这个由概念影响行为的启动效应被称为概念运动效应，值得我们关注。尽管你自己肯定没有意识到这个效应，但读了上述文字之后，“老年”这个概念肯定对你也产生了一定的影响。假设需要站起来接杯水喝，你起身离开椅子的动作就会比平常稍稍慢上那么一点，除非你恰巧不喜欢“老年”这一概念。有研究案例表明，如果对这一概念反感，人的动作就会比平时稍稍快那么一点。

在相反的情况下，概念运动效应也同样适用，德国一所大学曾做过的一项实验便证实了这一点。这个实验堪称巴奇和他的同事在纽约所做的那个早期实验的翻版。在实验中，研究人员要求参与实验的学生以每分钟30步的速度在房间里绕着圈走5分钟，这个速度是他们正常行走速度的1/3。这个简单的实验过后，学生们能更快地辨认出与“老年”相关的词汇，比如“健忘”、“年老”和“孤独”等。启动效应往往能产生连贯的反应：如果首先想到老年，你就会表现得像是上了年纪，而这种上了年纪的表现也会强化你关于“年老”的看法。

拿支铅笔放在齿间几秒钟，有橡皮的一端指向右边，笔尖指向左边。然后拿起笔来，咬住有橡皮的那一端，笔尖指向你的正前方。也许你并未意识到上述活动中有一个动作让你的眉毛皱了起来，而另一个动作却让你的脸上有了微笑。有人曾经让一些大学生给盖瑞·拉尔森（Gary Larson）的漫画《月亮背面》（The Far Side）评定幽默等级，同时要求他们在评级时咬一支笔。那些“微笑着的”学生（他们完全没有意识到自己在微笑）对这本漫画书幽默程度的评价比那些“皱着眉的”学生高得多。在另一项实验中，那些皱眉的人（眉头紧皱）在看到饥饿的儿童、争论不休的人以及事故现场等令人不舒服的图片时，往往会表现出更强烈的情感反应。

简单地说，常见的动作也会不知不觉地影响到我们的想法和感觉。在一项实验中，受试者按照要求用新耳机听一些信息。他们被告知这项实验的目的是检测音频设备的质量，研究

人员告诉他们要不断地摇晃脑袋，以检查设备是否有声音失真的问题，其中一半受试者要上下点头，而另一半则要左右摇头。他们听到的内容是电台的社论。那些点头（表示同意的动作）的受试者往往易于接受他们听到的信息，而那些摇头的受试者则易于否定那些信息。要强调的是，受试者并没有意识到上述问题，他们只是习惯性地将否定或接受的态度与其常用的身体语言联系起来而已。由此可见，老话讲“不管你怎么想的，都得心平气和”，这真是条很好的建议，只有真正做到心平气和，你才可能有回报。

你会自觉投到盒子里多少钱？

我们原以为自己作出的判断和选择是有意识且自主的，但那些关于启动效应的研究所带来的发现却颠覆了我们此前的认识。例如，我们中的大多数人都认为投票选举是经过深思熟虑的行为，它反映了我们对政策的认可程度与评价，不会受到不相关的事的影响。比方说，我们投票不应受到投票地点的影响，但事实上这个因素却实实在在影响了投票结果。一项关于2000年亚利桑那州选区投票模式的研究表明，当投票站设在某个学校时，加大教育投入力度的议案的支持率就会比投票站设在附近其他地方的要高。还有一项实验表明，当人们看到一些教室和学校储物柜的照片时，他们往往也倾向于支持关于教育的提案。这些图片对这些受试者的影响比其家长或其他选民与他们的意见分歧对其产生的影响还要大！对启动效应的研究从最初的实证研究——人们想到“年老”走路就会变慢——到今天已经取得了一定的发展，我们现在已经知道启动效应会影响到我们生活的方方面面。

让人联想到钱的事物往往令人不安。在一项实验中，受试者看到一个列有5个单词的单子，按照要求，他们得在5个词中选出4个组成以钱为主题的短语（比如“高、一份、薪水、桌子、工作”可组成“一份高薪工作”）。其他一些启动想象的实验则更挑战人的领悟力，包括背景中出现一个与钱没有任何关系的东西，比如一张桌子上放着一堆仿制钱币，或者一台电脑的屏保是水里漂着的一些美钞等。

起初脑海中有钱的概念的人比自己没有这方面联想时更独立。他们会持之以恒地解决一个非常难的问题，付出双倍的努力也在所不惜，实在迫不得已时才会向研究人员寻求帮助。这一做法清晰地表明其自力更生能力的提升。脑海中有钱的概念的人更自私：他们更不愿花时间去帮助另外那位假装对实验任务不大清楚的学生。当一位研究人员不小心将一捆铅笔掉到地板上时，脑子里想着钱（他们自己是无意识的）的那些受试者捡起的笔相对较少。在另一系列的一项实验中，受试者得知他们过会儿会和另一个人进行一次简短交谈，彼此熟悉一下，他们负责摆两把椅子，而研究人员则离开去找那个人。潜意识里有钱这一概念的受试者摆放椅子的距离（118厘米）会比没有这一概念的人摆放的距离（80厘米）更远。那些满脑子都是钱的受试者表现出更强烈的独处意愿。

上述所有发现有一个共同主题，即钱这一概念会滋生个人主义：不愿和他人在一起，不愿依赖他人，也不愿接受他人的请求。心理学家凯瑟琳·沃斯（Kathleen Vohs）完成了这项意义重大的研究，值得赞赏的是，她并未将自己的众多发现向公众和盘托出，而是留给读者更多的思考空间。凯瑟琳的研究意义深远，她的发现表明我们所处的文化环境中有很多能让人想起钱的事物，这些事物以我们意识不到的方式影响着我们的行为和态度，这些方式也许并不那么光彩。有些文化常常提醒人们尊重他人，另外一些文化则常让人们想起上帝，还有一些国家的人们会对着伟大领袖的照片顶礼膜拜。在一个专制国家中，到处挂着领袖的肖像不仅能向你传达“老大哥在看着你”的感觉，还会使你逐渐丧失自主的思想和独立的行为能力。

启动效应的研究证据显示，使人牢记“人终将一死”这一说法能让独裁的主张更得人心，因为在人们恐惧死亡的情况下，独裁的主张会让人心安。其他实验证实了弗洛伊德关于无意识关联中符号和比喻作用的见解。比方说，你可以看看下面这两个不完整的单词 W\_ \_H 和 S\_ \_P 是什么。如果有人近些天想起了自己某个难以启齿的举动，这些人往往会把这两个不完整的词填成 WASH 和 SOAP（“洗”和“香皂”），而很少会填成 WISH 和 SOUP（“希望”和“汤”）。此外，只是想到背后中伤某位同事就会使人更想去买香皂、消毒剂或清洁剂，而不是去买电池、果汁或者糖果。当人们感觉自己的心灵受到了玷污，往往也会引发他们清洗自己身体的想法，这种冲动被称为“麦克白效应”。

人们清洗的身体部位往往是那些令他们深感罪恶的部位。一项实验要求受试者通过电话或者电子邮件对一位假想中的人“说谎”。在随后对不同产品的需求测试中，那些通过电话说谎的人更想要漱口水，而不是香皂；而那些通过电子邮件说谎的人更想要的则是香皂，而不是漱口水。

当我向听众阐述关于启动效应的各项研究时，他们的反应通常是将信将疑的。这也不奇怪，因为系统2认为自己掌控一切，认为自己知道为什么要作出这样或那样的选择。你的脑海中也可能会突然出现一些问题：对情境进行微调怎么可能产生这么大的影响呢？这些实验是否表明我们完全任由情境摆布，随时要听从它的指示呢？当然不是。启动思维的影响力虽强，但并不见得很大。在100位投票者当中，只有几位开始时并不确定要选谁，如果投票点设在学校里而不是在教堂里，他们对与学校相关的问题会就作出不一样的选择，但也不排除有那么几个人会作出相反的选择。

然而，问题的关键是要接受相关研究的结果，而不是对此心存怀疑。这些结果不是捏造出来的，也不是统计上的偶然现象。你别无选择，只能接受这些研究的主要结论是正确的这

一事实。更重要的是，你必须承认这些结论对你自己来说也是正确的。如果你看到电脑屏保上有浮动的美钞，你帮那位笨手笨脚的陌生人捡起的铅笔数量可能比你没看到这些钱时更少。你不相信这些结论适用于你，因为这些结论与你的主观体验不相符，你的主观体验主要是由系统2决定的。启动效应来自系统1，而这个效应发生时，你根本就意识不到。

我用一个关于启动效应的完美展示案例来结束上面的阐述，这个案例是在英国一所大学一间办公室的茶水间进行的。多年来，这间办公室的职员一直都是自掏腰包买茶或咖啡，他们把每杯茶水和咖啡的建议价格写下来贴到墙上，上班时每次去接茶水或者咖啡时都会把相应的费用投到一个“诚实盒”里。某一天，有人在价格表的上方贴了张横条，上面既没有什么警告，也没作什么解释。在接下来的10周时间里，每周横条都贴有一张新的图片，图片上要么是一些花，要么是一双眼睛，好像盯着看着图片的人。没有人对这些新装饰发表过什么评论，但“诚实盒”里的钱却有了明显变化，如图4所示，这值得仔细研究一番。

实验开始的第一周（从图底端你可以看到相关信息），有一双睁大的眼睛在盯着来喝茶或咖啡的人，他们投进“诚实盒”的钱的平均值是70便士。第二周，图片上是鲜花，盒子中的钱的平均值少了15便士。这个趋势持续着，凡遇到有鲜花图片的那一周，盒子里的钱数就会减少。从平均水平来看，“眼睛周”时盒子中的钱是“鲜花周”时的3倍。显然，仅仅是一种象征性的监视符号便可促使人们改善自身的行为。正如我们预料的那样，产生这个影响的过程中没有任何意识的参与。现在你相信自己也难逃同样的模式了吧？

几年前，心理学家蒂莫西·威尔逊（Timothy Wilson）曾写过《我们是自己的陌生人》（Strangers to Ourselves）一书。现在你已经认识了自己身体中的那个陌生人，它也许在很多事情上都会为你做主，尽管你几乎从未察觉到它的存在。系统1带给你的各种印象经常会变成你的信念，而且是你作出选择和展开行动的动力源泉。它可以将当下的情形与新近发生的事情联系起来，再结合对近期的各种预期考虑，对发生在你身上或你身边的事情作出心照不宣的解释。系统1包含了对这个世界的认知模式，能立即估测哪些事情是正常的，哪些是出人意料的，它是你作出快速直觉性判断的依据，且这种判断十有八九是准确的，而你的所有判断活动几乎都是在这一系统的指引下毫无意识地完成的。然而，系统1也是你直觉中很多系统性错误的根源，这一点将会在后面几章中得到佐证。

#### 示例—启动效应

“这些人都穿着刻板的制服，看到他们时我们的大脑中是不会有有什么创造性想法的。”

“这个世界比你想象的要复杂得多，能否对它有个清晰的认识多半要看你的大脑的工作

方式。”

“他们的作用就是发现问题，而他们也的确发现了很多问题。”

“系统1编了一个故事，而系统2也相信了这个故事。我们每个人都有过这种体验。”

“我让自己微笑，这样做我也的确感觉好多了！”

## 思考，快与慢

### 第五章 你的直觉有可能只是错觉

在头脑清醒的时候——甚至在思绪并不那么清晰的时候——你的大脑一直在进行着多重运算，这些运算可以保留和更新一些关键问题的答案。例如：有什么新情况吗？存在危险吗？事情进展得顺利吗？我需要转移注意力吗？完成这个任务需要投入更多的精力吗？你可以将大脑想象成飞机驾驶座舱，里面有一套刻度盘，上面显示着这些重要变量的当前数值。系统1会自动对当前这些数值进行评估，因为这些数值的一个重要功能就是决定是否需要系统2提供额外帮助。

其中一个刻度盘测量的是认知放松度。认知放松度介于“放松”和“紧张”之间。放松是事情进展顺利的标志——没有障碍、没有新情况、没必要转移注意力或投入更多精力。紧张说明存在某种问题，且需要不断调动系统2参与其中。如若事情进展不顺利，你便处于认知紧张状态中。认知紧张同时还会受当时的努力程度和未得到满足的需求的影响。令人惊讶的是，一个简单的认知放松过程却与有着多种输入和输出活动的庞大网络相联结。图5对此作了解释。

#### 由记忆造成的错觉

错觉这个词会让人马上联想到视错觉，因为我们都很熟悉那些令人产生误解的图片。不过，并不是只有视觉才会产生错觉，记忆也容易产生错觉，并且这种现象更加普遍。

戴维·斯滕比尔（David Stenbill）、莫妮卡·比格特斯基（Monica Bigoutski）、莎娜·蒂拉纳（Shana Tirana），这些名字都是我自己编出来的。如果在接下来几分钟内再次看到这些名字，你很可能还记得自己在哪里见过它们。你心里清楚这些名字并不是那些为数不多的名人的名



字，而且有那么一段时间你会记住这一点。但假设几天后你看到一长串名单，其中包括一些知名度不高的名人和你从未听过的“新”名字，而你的任务是挑出名单中所有名人的名字，这时你很可能将戴维·斯滕比尔当成名人，尽管你并不清楚自己是否在与电影、体育运动或是政治有关的新闻中听说过他的名字。心理学家拉里·雅各比（Larry Jacoby）是第一位在实验室中论证了这种记忆错觉的学者，且发表了一篇名为“一夜成名”的文章。这种情况是怎样发生的呢？要想弄明白，先问问自己是如何知道一个人是否出名的。一些真正有名的名人，例如爱因斯坦、博诺·沃克斯、希拉里·克林顿等，你会在脑海中为他们建立一个信息丰富的记忆档案。但当你几天后再次见到戴维·斯滕比尔这个名字时，你并没有关于他的记忆档案，你有的只是一种熟悉感——你曾经在某个地方见过这个名字。

雅各比巧妙地阐述了这个问题：“熟悉感有着简单而又强烈的‘不可复返性’，这种‘不可复返性’似乎说明这种感觉是对过往经历的一种直接反应。”其一大特性便是错觉。正如雅各比和他的许多同事所说的那样，看到戴维·斯滕比尔这个名字之所以觉得很眼熟是因为你更容易辨认出这个名字。再次见到曾经见过的词时会更容易识别出来——无论是在你面前一晃而过或是在嘈杂的环境中让你指认，你都能更加轻松地识别出这些词，而且你读出这些词的速度要比读出其他词快（快百分之几秒）。简而言之，当你指认一个早已见过的词语时，会感到更加放松。正是这种放松使你产生了那种熟悉的感觉。

什么样的信息更容易让人信服？

“纽约是美国第一大城市”，“月亮绕着地球（公）转”，“鸡有4条腿”，在所有这些表述中，你在很短时间内便接收了大量信息，这些信息五花八门，各不相同。读过前面三个表述后，你很快就知道前两个是真实的，最后一个是错误的。但值得注意的是，判断“鸡有3条腿”是错误的，明显要比“鸡有4条腿”更容易。你的联想机制对后一种表述作出判断的速度相对缓慢，之所以这样是因为后一句话让你想到许多动物都有4条腿这一事实，你可能还会想到超市通常也会销售4只一包的袋装鸡腿。系统2负责信息转换，它也许会提出问题（关于纽约的那个问题是否太简单了），也可能要核对“（公）转”这个词的意思。

想想你最后一次参加驾驶考试的情形吧。你真的需要一张驾驶执照去开一辆载重3吨的车吗？或许你学习认真，连答案在哪一页的什么位置都记得，还知道答案的逻辑关系。但我搬到一个新的州后通过驾驶考试所用的方法跟这些丝毫不沾边儿。我只是快速地读了一遍交规，希望自己能交好运通过考试。我开车有很长一段时间了，还是知道一些交通法规的，但一旦遇到不知怎样回答的问题时，我只能靠认知放松来解决。如果某个答案看起来比较熟悉，我就猜测它可能就是正确答案。如果某个答案看起来比较生僻（或是非常极端），我便排除它。系统1让人产生熟悉感，系统2依靠系统1产生的这种熟悉感来作出正误判断。

我们还可从图5中得知，如果某个判断是基于认知放松或认知紧张作出的，那就一定会造成错觉。任何能使联想机制运行更轻松、更顺利的事物都会使我们心生偏见。想让人们相信谬误有个可靠的方法，那就是不断重复，因为人们很难对熟悉感和真相加以区别。权威机构和营销商都深谙这个事实。然而，心理学家发现，你不必完整地重复某件事情或某个想法，即使只说一部分，人们也可能相信你的话。经常看到“鸡的体温”这个短语的人会更容易接受“鸡的体温是144华氏度”（或其他随便什么温度都无妨）这一说法。只要熟悉其中一个短语，就会觉得对整个陈述都很熟悉，也会因此对陈述内容信以为真。如果你记不清楚某个陈述的来源，也无法将其与自己知道的事物联系起来，这时你就别无选择，只能跟着认知放松的感觉走了。

假设你需要写一则能让接收方相信的消息，当然，这则消息是真实的，但人们并不一定相信它是真的。此时，你完全可以利用认知放松来帮助自己，真相错觉的有关知识也可以提供一些具体的建议来帮助你实现这个目标。

这些建议总的原则是，任何缓解认知紧张的做法都会对你有所帮助，所以，你首先应该让字迹更清晰。请比较下面两个陈述句：

阿道夫·希特勒生于1892年。

阿道夫·希特勒生于1887年。

上述两种说法都是错误的（希特勒生于1889年），但实验显示，第一句话更容易使人信服。其次，你的信息应该印刷在质量较好的纸上，并且文字和背景间的反差要达到极致。如果你使用彩色字体的话，亮蓝或大红的文字会比绿、黄、灰蓝等色调更容易让人相信文字内容的真实性。

如果你很在意自己在别人眼里是否值得信赖、是否聪明睿智，那么说话时就言简意赅吧，能用简单句的时候就别用复杂句。在研究生中有一个广为流传的段子，说的是很多教授眼中那些最令人难忘的词汇。我在普林斯顿大学的同事丹尼·奥本海默（Danny Oppenheimer）反驳了这些学生的说法。在一篇名为“不实际地运用华丽辞藻的后果：使用长句毫无必要”的文章中，他指出用浮夸的语言来表达熟悉的概念是一种智商低下、可信度差的表现。

除了应该保持消息简洁外，还应使其易于记忆。如果可以的话，将你的想法以诗歌的形式表达出来，这样人们会更容易相信你的话。在一个多次被引用的实验中，受试者阅读了许

多他们并不熟悉的格言，例如：

危难时，敌人团结。

水滴石穿，绳锯木断。

亡羊补牢，为时不晚。

另一些学生读到的格言版本则是没有什么韵脚的：

危险团结敌人。

再轻的击打也能放倒一棵大树。

承认错误就相当于得救了一半。

由上可见，押韵的格言比没有韵脚的格言显得更加深刻。

求受试者通过两份经纪公司的报告去评估几家虚拟的土耳其公司的发展前景。每家公司都有两份相关报告，其中一份来自一家名字上口的经纪公司（比如 Artan），另一份则来自一个名字很拗口的经纪公司（比如 Taahhut）。两份报告有时会持不同意见。观测者能采取的最佳方法就是对两份报告进行综合评估，但他们并没有这样做。相比名字拗口的经纪公司的报告，他们更相信名字上口的经纪公司的报告。记住，系统2是非常懒惰的，它不愿付出努力。如果可能的话，消息的接收方会离所有需要付出精力的事物远远的，包括名字复杂的信息源。

所有这些都是很好的建议，但是我们不能因此而得意忘形。如果你的信息是明显荒谬的或是与接收方已知的事实相左，那么即使你把高质量的纸、鲜艳的颜色、押韵和简单的语言全都用上，也很难提高信息的可信度。做这些实验的心理学家并不认为其他人是愚蠢和容易骗的。他们认为，我们所有人的生活都受系统1所产生的印象指引，但我们通常不知道这些印象从何而来。你怎么知道某个陈述是否正确呢？如果信息富有逻辑性、与你所持的信念或偏好有联系，或是源自你信任和喜爱的信息源，你便会有认知放松之感。但问题在于，可能还有其他的原因导致你感到放松，包括字体的清晰度和优美的韵律，你无法轻易找到这些感觉的来源。这便是图5带给我们的信息：放松或紧张之感是由多种原因导致的，难以将它们理清。虽然这很难，但也不是不可能。人在备受鼓舞时，就能克服一些导致真理错觉的外在

因素。然而，在大多数情况下，系统2都会采纳系统1的建议，并继续运作。

股票代码上口，回报率也相对较高

联想关系的对称性是讨论联想连贯性的首要主题。正如前文所述，当人们将铅笔咬在嘴里或是把球放在眉心使自己“微笑”或“皱眉”时，也会感受到平时那种微笑或皱眉时所表达出来的情感。自我强化的相互作用同样也能在认知放松的研究中有所体现。一方面，当系统2参与到那些需要努力的事情中时，我们就会感到认知紧张。另一方面，无论是什么原因引起的认知紧张都有可能将系统2调动起来，改变人们处理问题的方式，使他们不再随意且凭直觉做事，而是转向更加专注的分析性思维。

前文所提到的球拍和球的问题，测试出人们在回答问题时有这样一个倾向，即不假思索地将脑海中出现的第一个想法当做答案。肖恩·弗雷德里克之所以选择了球拍和球的问题及其他两个问题，是因为它们可以唤起一个最直接但却是错误的直觉性答案。这个认知反应测试的另外两个问题是：

如果5台机器能在5分钟生产5个小零件，那么100台机器生产100个小零件需要多长时间？ 100分钟还是5分钟？

湖中有一片睡莲叶子，这片叶子以每天增长一倍的速度向外扩散。如果48天后莲叶就能覆盖整片湖面，那么其覆盖湖面一半的面积需要多长时间？

24天还是47天？

本页脚注中有两个问题的正确答案。研究人员召集了40名普林斯顿大学的学生来做这个认知反应测试，其中有一半问卷使用的是浅灰色的小字。拿到这部分问卷的学生可以辨认出上面是些什么字，但这样的字体会导致他们认知紧张。测试结果清楚地表明：当问卷上的字体清晰时，90%的学生至少会犯一个错误；但当字体勉强能够辨认时，他们的错误率却降到了35%。是的，你没有看错：字体模糊不清时，学生的表现更好。这是因为，无论认知紧张是由什么原因造成的，它都更有可能激发系统2来抑制系统1所给出的直觉性答案。

《大脑一放松，脸上现笑容》这篇文章描述了这样一个实验：让受试者快速浏览一些物体的图片，在播放其中一些图片时，先在整个物体出现之前用快得令人难以察觉的速度呈现其轮廓。研究发现，受试者在识别这些图片中的物体时会相对容易。实验人员对受试者面部肌肉的电脉冲进行测量，来记录肉眼难以观察到的细微而短暂的表情变化，并由此测出受试

者的情绪反应。不出所料，当图片上的物体更容易识别时，人们会微微一笑，眉头舒展，可见认知放松与良好的感觉相互关联似乎是系统1的一个特点。

正如所料，朗朗上口的单词会唤起人们的正面态度。若某公司有一个上口的名字，其股票发行第一周往往会取得比其他公司更好的成绩，虽然这样的效果会随着时间推移而逐渐消失。如果股票有一个读起来上口的代码（例如 KAR 或 LUNMOO），这只股票就会比有着 PXG 或 ROD 这样拗口代码的股票表现优异，并且这种优越性似乎还会持续一段时间。瑞士的一项研究发现，像 Emmi、Swissfirst、Comet 等有着顺口名字的股票会比 GeBerit、Ypsomed 这样名字拗口的股票带来更多的回报。

重复能引发放松状态和令人舒心的熟悉感。著名心理学家罗伯特·扎伊翁茨（Robert Zajonc）曾潜心关注重复某种刺激和这一刺激最终带来的轻微情感波动之间的关系，扎伊翁茨称之为曝光效应（mere exposure effect）。在密歇根大学和密歇根州立大学这两所大学的校报上所做的实验是我最满意的实验之一。在几个星期的时间里，一个类似广告的版块出现在校报的头版上，上面写着以下某个土耳其语单词（或像土耳其语那样发音的单词）：kadirga, saricik, biwonjni, nansoma, iktitaf。这些词语重复的频率各不相同，其中一个只出现过一次，另外一些分别出现过2次、5次、10次和25次。（在其中一所大学校报上出现得最多的单词在另一校报上出现的频率则最低。）没有人就这个问题作出解释，如有读者询问，得到的回答都是：“买下这个版块的人不愿透露身份。”

当这一系列神秘的广告结束后，研究人员在校园内散发调查问卷，询问学生对每一个土耳其语的印象，是有“好感”还是很“反感”。结果令人惊奇：相比只重复了一两次的词，受试者对那些重复次数较多的词更有好感。通过使用汉字、人脸和任意形状的多边形所做的实验也都再次证实了这一发现。

曝光效应并不依赖对熟悉程度有意识的体验，事实上，曝光效应完全脱离了意识：当重复的单词或图片迅速闪过，快到观察者根本没有意识到看见它们时，观察者同样还是更为喜欢重复率高的单词或图片。现在我们应该清楚的是，系统1能对系统2意识不到的那些事的印象做出回应。事实上，当人们完全没有意识到自己看见了刺激物时，其曝光效应会更强。

扎伊翁茨声称，这种只要不断重复接触就能增加喜欢程度的现象，是一个极其重要的生理现象，可推及所有动物身上。要想在一个危机层出不穷的世界生存，一个有机体对新奇刺激应该谨慎回应，随时保持警惕和逃离的状态。若一种动物对新奇的事物没有心存戒备，其生存概率就会很低。然而，如果了解到这个刺激物是无害的，最初的谨慎便会渐渐消失。扎伊翁茨称，曝光效应的产生是因为一个刺激的重复曝光并没有产生不好的影响，这样的刺激

最终会成为一个安全信号，而安全的就是好的。显然，这样的观点并不只适用于人类，同样适用于其他物种。为了证实这一点，扎伊翁茨的一个助手给处于孵化状态的鸡蛋放不同的音乐，随后，孵化出来的小鸡在听见孵化期间所听到的音乐时，发出的哀鸣要少得多。

扎伊翁茨为他的研究项目作了一个很有说服力的总结：

重复曝光的结果有益于机体适应其所处的有生命和无生命的环境。这一效应能使机体鉴别出安全的物品和栖息地，是最为原始的社会性依附的基础。因此，重复曝光构成了社会组织和社会整合的基础，而社会组织与社会整合又是心理稳定与社会稳定的基础。

积极情感和系统1的认知放松之间的联系由来已久。

创新是发生在能让人联想无限的环境中的

1960年前后，一个名为萨尔诺夫·梅德尼克（Sarnof Mednick）的年轻心理学家认为他已发现了创新的本质。他的观点虽然简单却很有力：创新与出众的记忆力有关（创新是极佳的联想记忆）。他设计了一个测试，名叫远隔联想测验（Remote Association Test, RAT），这个测验直到今天还常常被用来研究创造力问题。

实验带来了许多惊人的发现。第一个令人惊讶的地方是，参与实验的受试者猜测所给问题的答案的准确率要比他们随便选出答案的准确率高很多。这个结果很让人吃惊。很明显，认知放松感被联想机制中一个微弱的信号激发起来，这个信号在组内三个词语的关联得到还原之前就已经“知道”它们是相关的（有共同的联系）。认知放松在这种判断中起到的作用也在另一个德国小组所做的实验中得到证实：提高认知放松（启动、清晰的字体、事先看到词）的操作同样能增加人们将词视为相互联系的可能性。

另一个重大发现是，人的心情在这项直觉性任务中产生了重要影响。实验人员用电脑编写了一个“直觉参数”用于检测准确性。他们发现，通过让受试者想一些愉快的事，使他们在测试之前有好心情的话，猜测的准确率会提高一倍。一个更引人注目的结果是，心情不好的受试者完全无法准确地完成这个直觉性任务，他们的猜测还没有随便选择的准确率高。心情显然能够影响系统1的运行：当我们不舒服和不开心时，就会丧失自己的直觉。

这些发现进一步证实，好心情、直觉、创造力、轻信以及对系统1不断增强的依赖性形成了一个关联群集。悲伤、警觉、怀疑、分析方法以及不断增强的努力程度等因素之间也是相互联系的。好心情使系统2放松对行为的控制：当人们心情好时，直觉和创造力会增强，

但也会放松警惕，易犯逻辑性错误，这种关联与曝光效应一样，在生理上也讲得通。好心情是事情进展顺利的信号，周围的环境是安全的，卸下防备并没有什么影响；坏心情则说明事情进展不那么顺利，有可能存在威胁，必须要保持警觉。认知放松与愉快的感觉互为因果。

除了认知放松和积极情感的关系之外，我们还可以从远隔联想测验中了解到更多信息。请简单考虑一下下面两组词：

睡觉 邮件 开关

盐 深 泡沫

在读第二组词时，你的脸上可能会露出一丝微笑，这是因为第二组词能引起一个共同的联想（答案就是“大海”）。你当然不可能知道自己的这一反应，但是只要对你的面部肌肉的电活动进行测量就能发现这一点。这种微笑反应出现在那些对于共同关联事物还并不了解的受试者的脸上。实验人员仅仅是给他们看了纵向排列的三个词语，然后指示他们看完后按下空格键。看到有共同联想意义的三个词而产生的认知放松似乎本身就能令人心情舒畅。

正如很多科学家所言，虽然我们证实了好心情、认知放松和对连贯性的直觉之间是相互关联的，但这种联系并不是必然的因果关系。虽然认知放松与微笑同时出现，但一定是好心情引发了连贯的直觉反应吗？答案是肯定的。证据源自一个巧妙的实验方法，这种方法现在越来越受欢迎。在这个实验中，一些受试者听到了实验人员讲的一个封面故事，也许正是这个故事给他们带来了好心情：他们都戴着耳机，听着音乐。实验人员告诉他们：“先前的研究显示，现在播放的音乐会对人的情绪反应产生影响。”这个故事完全使人丧失了对相关性的觉察。这一发现证明，在看到三个词（如果三个词相关则感到愉快，不相关则不愉快）后，短暂的情绪反应事实上是对相关性判断的基础。没有什么系统是系统1做不到的。情感变化现在是可以预料到的，正因为这些变化不足为奇，我们才不会随意将情绪变化归因于词语。

这项心理研究是迄今为止最为成功的一例，其成功之处在于将实验技巧与实验结果巧妙结合在一起，而这两者都是特色鲜明且极为惊人的。在过去几十年里，我们已经掌握了许多关于系统1自动运作的知识。我们现在掌握的很多知识听起来像是三四十年前的科幻小说。无论是模糊的字体会影响人们对真理的判断，还是提高认知能力，或是人们对每组三个词的认知放松感引起的情感回应会唤起他们头脑中的连贯印象，这些都不再难以想象。心理学研究已经取得了长足的发展。

示例—认知放松

“不要只是因为字体不清楚就否定他们的商务计划。”

“我们一定愿意相信这个观点，因为总有人这样想，但还是三思而后行吧。”

“熟悉了，就会喜欢，这就是一种曝光效应。”

“我今天心情很好，我的系统2也比平时懒惰，我得格外小心了。”

## 思考，快与慢

### 第六章 意料之外与情理之中

前面已经介绍了系统1和系统2的主要特点及功能，尤其是系统1，其介绍更为详细。我将我们的大脑比喻成一台功能非常强大的电脑，虽然按传统硬件标准来看，它的运行速度很慢，但通过不同想法联结而成的巨大网络上的相关联结，也能为我们展现这个世界的构造。而且，联想机制的不断激活也是自动完成的。我们（系统2）具备某种能力去控制记忆的搜寻活动，也能对其进行编辑，这样一来，在特定情况下我们才可以集中精力去追踪某个事件。接下来我们将对系统1的非凡之处及其不足之处作更详细的介绍。

#### 从第一次的惊喜到第二次的习以为常

系统1的主要功能是维护并更新你个人世界的模式，它呈现的都是常态下的思维模式。这个模式由许多联想和结果共同构成，这些联想由情境、事件、行为等概念引发，而结果则总是伴随某种规律出现，有可能是两者同时出现，也有可能出现时间只差那么一点点。随着这些联系的形成和加强，联想概念的模式逐渐展现出发生在你生活中的各种事件的结构，这一模式还将决定你对当下的理解和对未来的期望。

对惊喜的承受能力是你精神生活的一个重要方面，而且惊喜本身也是最敏感的指示，它可以表明我们如何理解这个世界，我们希望从这个世界中得到什么。惊喜主要有两种形式。有些期望是积极、有意识的——你知道你正在等待某件特别的事情发生。当时间迫近时，听到敲门声，你会期望是你的孩子放学回家了；打开门时，你期望听到自己熟悉的声音。如果非常期望的事情没有发生，你会很惊讶。不过，你不怎么期盼的事情也很多。你不会等待这些



事情发生，但如果真的发生了，你也不会惊讶。因为这些事在某种情境下是正常的，尽管可能不是你欣然盼望的。

一个插曲如果反复出现就可能不那么令人惊喜了。几年前，我和我的妻子在澳大利亚大堡礁的一座小岛上度假，岛上只有40间客房。我们去吃饭时，竟然遇见了一位老朋友——心理学家乔恩，我们当时感到很意外。朋友间相互问候，都觉得很巧。第二天乔恩就离开了度假村。大约两个星期后，我们去伦敦的一家剧院看电影。剧场关灯之后，一位来迟的人坐到了我旁边。幕间休息时，我发现我旁边坐着的竟是乔恩。我和妻子事后说起这两件事时，我们两人都意识到了两个事实：第一，这次碰面比上次更巧；第二，第二次碰到乔恩时我们的感受不如第一次那么惊讶。显然，第一次碰面在不知不觉中改变了乔恩在我们大脑中的印象，他现在是“只要我们外出旅行就会碰见的心理学家”。我们（系统2）知道这样想很可笑，但系统1却使大脑认为在陌生的地方碰到乔恩是件正常的事。如果我们在伦敦的那家剧院里碰到的是其他熟人，我们会比遇到乔恩感觉更惊讶。如果单纯从可能性来看的话，在剧院里碰到乔恩的概率比碰到我们那几百个熟人中的任何一个的概率都要小得多，然而现在看来，在那里碰到乔恩却似乎更正常。

在某些情况下，消极的期望很快会变成积极的，就像我们在又一次巧合中发现的那样。几年前，一个周日的傍晚，我们驱车从纽约赶往普林斯顿，很长一段时间以来，我们每周都会这样穿梭于两个城市之间。那天我们看见了不同寻常的景象：路旁有辆车着火了。我们第二周到达那段路时，又有一辆车着火了。我们发现与第一次相比，第二次遇到这种情况时明显表现得不那么惊讶了。这个地方现在就是“车着火的地方”。因为事件发生的情景相同，第二次事件引起了积极的期盼：每当走到这个路段时，我们都会想起着火的车，而且自己已经作好再看到一次的心理准备（当然，我们没有再看见过），几个月也好，几年也好，我们都有足够的心理准备。

心理学家戴尔·米勒（Dale Miller）和我曾经合写了一篇文章，在文中我们试图解释为什么我们觉得很多事情是正常或是不正常的。接下来我会引用那篇文章中我们用来描述“常态理论”的一个例子，不过我对这个例子稍稍作了改动：

在一家高级餐厅，一位观察者随意观察着邻桌的顾客。他发现第一位品汤的顾客往后退了一下，好像有些痛苦。接下来的一大堆事都因这个动作而发生了变化。一位服务生不小心碰到了一位喝汤的顾客时，这位顾客吓了一跳，而这根本没什么可惊讶的；另外一个顾客也喝了从同一个碗中盛来的汤，喝后他强忍着没有叫出来，对此我们也没有感到惊讶。这些事和其他种种事情本应是不正常的，但现在却显得正常了。因为这些事是对之前那些预期的肯定。这些事之所以看上去正常是因为它们重复了原来的情节，记忆中的这些情节相互联系，

对这几个事件作出了解释。

试想你就是餐厅中的那位观察者，第一次看到顾客对汤的反应时你感到惊讶，服务生在碰他时他吓了一跳，你也会感到惊讶。然而，第二次非正常事件会提取你对第一次事件的记忆，两件事联系起来似乎就合乎情理了。这两件事都符合一个模式，模式中的顾客是个特别容易紧张的人。另外，如果第一位顾客表现出痛苦状后，另一位顾客拒绝喝这种汤，这两件意外的事也会被联系起来，说明这碗汤肯定有问题。

“摩西上方舟时，每种动物各带了多少？”发觉这个问题有问题的人太少了，因而有人将此称为“摩西错觉”。摩西并没有带动物上方舟，带动物上方舟的是诺亚。就像那位往后退的喝汤顾客一样，摩西错觉可以通过常态理论来解释。动物登上方舟给人们营造了《圣经》的文化背景，而摩西也出现在这个背景中，虽然你对他并没有什么特别期待，但对他的名字你也不会感到惊讶。况且，摩西和诺亚这两个名字的英文元音和音节数都是相同的，这也容易使人将二人混淆。就像那些三个一组的词语会令人产生认知放松一样，你在不知不觉中就将“摩西”和“方舟”联系在了一起，很快地接受了这一提问。但如果句中的摩西换成乔治·W·布什的话，你就会认为这是一个很蹩脚的政治笑话。

若“关于水泥”这种说法与当前被激活观点的情境不搭，系统就会察觉到这一反常现象，就像你刚才的感受一样。你本来对“关于”之后是什么并无具体概念，但当“水泥”一词出现时，你就知道这个句子有点不对劲了。对大脑反应的许多研究表明，系统对违反常态的问题的察觉速度是惊人的，察觉过程也是微妙的。在最近的一次实验中，当人们听到“地球每年绕着麻烦转”这个句子时，听到“麻烦”这一莫名其妙的词之后不到0.2秒，对大脑活动进行检测就会发现一个特殊的图谱。更值得注意的是，当一个男声说“我确信我怀孕了，因为我每天早晨都想吐”，或者当有人操着上流社会的腔调在说“我后背上有一个很大的文身”时，同一个人对这些话的大脑反应速度是相同的。一定要立刻调动起来大量的知识，思考这些陈述与事实之间是否协调：一听这声音就知此人是上流社会的，而上流社会的人身上很少有很大的文身。

我们能相互交流，因为我们关于世界的知识和对文字的运用水平都相当。我要是提到桌子，没有其他特别的描述的话，你一定知道我指的是一张普通的桌子。你能确定这张桌子的表面大致是平的，而且桌腿远远少于25条。对于很多事物的常态我们都有认识，这些常态给我们提供了背景知识，使我们能够察觉到诸如怀孕的男性和文身的上流社会人士这样的反常现象。

要想了解标准在交际中的作用，请看“大老鼠爬过了小象的鼻子”这个句子。我敢肯定

你想象中的老鼠和小象的体型和我想象的差不了太多。常态下，这些动物的体型有典型的或平均的大小，而且，常态中还包括了这一类动物的变化范围及变异性。我们都不可能在脑中想象到比象还大的老鼠爬过比老鼠还小的小象的鼻子。但我们各自都能想象到比鞋还小的老鼠爬过比沙发还大的大象。系统1理解语言，也了解分类标准，它能辨明那些貌似正确的价值，也能廓清那些最典型事例可能出现的范围。

简的钱包是怎么丢的呢？

“弗雷德的父母来迟了，酒席承办商应该很快就到了，弗雷德很生气。”你知道为什么弗雷德会生气，也知道他生气不是因为酒席承办商还没来。在你的联想网络中，生气和不准时被看成是可能的因果关系，但生气和盼望酒席承办商来的想法却没有这种联系。在你读到这句话时，脑海中马上就形成了一个连贯的故事，你立即知道了弗雷德生气的原因。找到这种因果联系是理解一个故事的一部分，也是系统1的一种无意识行为。而系统2——也就是有意识的自我——在接收到这个主观理解后，也接受了这种因果关系。

纳西姆·塔勒布所著的《黑天鹅》一书中，有一个故事阐明了对因果关系的自主研究。他指出，萨达姆·侯赛因（Saddam Hussein）在伊拉克的藏身处被捕的那天，债券价格就开始上涨。早晨时，投资者显然想投资更安全的资产，而且彭博新闻社（Bloomberg News）打出了这样的头条：美国国债上涨，萨达姆被捕不会遏制住恐怖主义。半个小时后，债券下跌，标题修改为：美国国债下跌，萨达姆被捕刺激风险资产。显然，萨达姆的被捕是这一天的重大事件，因为自主搜索原因的方式影响了我们的思考，这件事就成了那天市场变动的原因了。这两个标题表面来看好像可以作为解释市场发生振荡的原因，但是对两个互相矛盾的结果作出解释的那条陈述其实什么都解释不了。事实上，所有的标题都要满足我们对逻辑连贯的需求：一件大事必然会带来一些后果，而这些后果也需要一些原因对其作出解释。我们对那天发生的事情所知有限，于是系统1便熟练地将这些知识片段组合成一个连贯的因果关系。

请读这个句子：

在纽约拥挤的大街上逛了一天，欣赏完美景后，简发觉自己的钱包丢了。

研究人员对读过这个句子的人（这些人同时也读过许多其他故事）进行了一次突击性的回想实验，发现这些读者认为“扒手”这个词比“景色”一词与该句子联系更紧密，尽管后者在句子中出现了而前者却没有。联想连贯性原则告诉我们实实在在发生的事情。丢钱包这件事可能有很多不同的原因：钱包从口袋里掉了出去，或者落在了餐厅等。但是，丢钱包、纽约以及拥挤的街道三种想法同时存在时，这些读者都把丢钱包的原因指向了扒手。在那个喝汤

的故事中，无论是第一位喝汤顾客在服务生碰他时做出的极端反应还是另一个顾客在喝汤时往后退，其结果都将对最初那次惊讶反应作出联想性关联的解释，因此整个故事看上去也是合理的。

比利时的贵族心理学家阿尔伯特·米乔特（Albert Michotte）在1945年出版的一本书（1963年被译成英文）中颠覆了几个世纪来关于因果关系的思考，使人们至少回到了休谟对想法关联的研究时代。过去人们普遍接受一个观点，即我们通过对不同事件之间的相互关联进行反复观察，推断其在自然状态下的因果关系。我们曾无数次看到一个运动中的物体碰触另一个物体，使之随即开始运动，且运动方向相同（也并不总是相同）。这和台球相撞的情形一样；同样，当你推一个花瓶时，也会发生同样的事。米乔特却持有不同观点，他说我们能像看见颜色那样直接“看到”因果现象。为了展示自己的观点，他在纸上画了一连串黑色方块，让这些方块看起来像在连贯地运动，一个连着另一个，而这些方块也立即动了起来。那些观察者知道二者其实并没有真正相连，但他们却有种强烈的“因果关系错觉”。如果第二个物体马上开始移动，他们就会认为这个动作是由第一个物体“引起的”。很多实验表明，6个月大的婴儿会将许多事件及其续发事件看做有因果关系，而续发事件一旦发生改变，他们就会觉得惊讶。我们显然从出生时就对因果关系有感觉，当然，这种因果关系并不依存于理性思维，它们是系统1的产物。

1944年，米乔特发表了他对物质性因果关系的实证研究。几乎就在同一时间，心理学家弗里茨·海德（Fritz Heider）和玛丽-安·西梅尔（Mary-Ann Simmel）运用一种和米乔特相似的方法展示了他们对“意向性”因果关系的看法。他们做了一部总长度只有1分40秒的电影，在这部微电影中，你能看到一个大三角形、一个小三角形和一个圆形。这些图形绕着一个貌似门开着的房子的图像转。观看者看到一个气势汹汹的大三角形正欺负一个小三角形，而那个圆形也受到了惊吓。圆形和小三角形联合起来共同对付大三角形的欺侮；他们还看到门周围（圆形和小三角形）的动作很多，最后小电影以一个爆发性的反击结束。毫无疑问，每个人对意图和情感都有很强的洞察力；只有患孤独症的人才体验不到这一点。当然，所有这一切全都在你的大脑中。你的大脑愿意甚至急切地想要辨别一些因素，确定这些因素的特点和特殊意图，并将其活动视为表达个人偏爱的举动。再次说明一下，我们生来就长于意向性归因：一岁以下的婴儿能辨别欺凌弱小的人和受害者；不管他们去抓什么东西，都知道要走最近的那条路。

按自由意愿行事的体验与物质性因果关系没有什么联系。虽然是你的手捏起了盐，但你并不认为这件事与一连串的物质性因果关系有什么联系。你感觉这个行为是由一个“无形”的你做出的决策而引起的，因为你想要在食物里加点盐。很多人发现将自己的心灵描述成行为的根源和原因是很正常的。2005年，心理学家保罗·布鲁姆（Paul Bloom）在《大西洋月

刊》(The Atlantic)上发表了有争议的观点,即我们生来就具有区分自然性和意向性的因果关系的能力,而这一能力也解释了宗教信仰的普遍性。他发现,“我们对物质世界和精神世界的感知完全不同,所以我们才可能对没有灵魂的身体和没有身体的灵魂进行想象”。我们注定能够感知得到的这两种因果关系模式使我们能够很自然地接受众多宗教中都存在的两个主要信念:无形的神是物质世界所有现象的最终原因,而不死的灵魂只是在我们活着时暂时控制我们的躯体,我们死后,灵魂也将离开躯体。在布鲁姆看来,因果关系的两个概念在进化过程中是分别形成的,同时也将宗教的起源追溯到系统1中去了。

因果性直觉的特点是本书一再出现的主题,因为人们总是很不恰当地将因果性思考用于需要统计论证的情景中。统计性思维总是根据事物的不同类别和总体性质得出个案的结论。可惜的是,系统1并不具备这种推理能力;而系统2通过学习可以进行统计性思考,但几乎没有人接受过必要的相关训练。

有了因果关系心理学做基础,我决心用两个比喻来描述心理过程,几乎不必考虑什么连贯性问题。我有时将系统1比喻成有着某些特性和偏好的媒介,有时又将其看做通过一个环环相扣的复杂模式来表现现实的联想工具。这个系统和工具是虚拟的;之所以用“媒介”和“工具”这两个喻体,是因为它们符合我们对原因的看法。海德的正三角形和圆形并不是真正的媒介,只是把它们比做媒介来思考问题更轻松,也更自然。这样做不用那么费力思考。如果按照所发生事情的特点和意图(即两个系统)来描述它,或者有时按照呆板的规律性(即联想工具)来描述它,我觉得你(和我一样)就会发现了解思维活动更加轻松。我并不想让你相信这两个系统是真实存在的,不想像海德那样,想让你相信那个大三角形是个欺负弱小的角色。

#### 示例——常态和原因问题

“若最后发现第二个申请人也是我的一个老朋友,我就不会像看到第一个来申请的朋友那么惊讶了。只要情节稍有重复,一种新体验也就不那么新鲜了。”

“当我们调查人们对这些产品的反馈时,一定要确定关注的不只是平均水平。我们应该看到所有的常态反应。”

“她接受不了自己只是运气不好这个解释,她需要一个有前因后果的解释,否则她会认为是有人在故意破坏她的工作。”

## 思考,快与慢

## 第七章 字母“B”与数字“13”

伟大的喜剧演员丹尼·凯（Danny Kaye）的一句话一直萦绕于我的脑际，从十几岁至今都不曾忘记。谈到一个自己很不喜欢的女人时，他说：“她最得意的姿态是忘乎所以，最喜欢做的事就是仓促下结论。”在我与阿莫斯·特维斯基开始讨论研究统计性直觉的合理性问题时，这句话就浮现在我的脑海里。现在，我觉得这句话正是对系统1的功能恰如其分的描述。如果结论可能是正确的，偶尔的错误所付出的代价也在可接受的范围内，而且这种仓促的做法可以节省很多时间和精力，那么这种仓促的结论就是高效的。如果对情况不熟悉，在风险又高并且没有时间去搜集更多信息的情况下过早下结论就很冒险。此时，很可能会出现直觉性错误，但这种错误也许可以通过系统2的有意干涉而得以避免。

是什么让你相信了那些荒谬之词？

心理学家丹尼尔·吉尔伯特（Daniel Gilbert）因著有《哈佛幸福课》（*Stumbling to Happiness*）一书而广为人知。一次，他在一篇名为“大脑怎样产生信任”的文章中，以17世纪的哲学家斯宾诺莎的理论为基础，提出了一个关于信任与不信任的理论。吉尔伯特提出，在理解一个陈述之前，一定会先试图相信它：如果这个陈述正确的话，你必须先了解它的观点究竟是什么意思。只有这样，你才能决定是否“怀疑”它。最初你产生相信某种观点的想法，是因为系统1的自主运作，这种运作包括构建这一情况下可能性最大的解释。吉尔伯特认为，即使是一个毫无意义的陈述也会唤起人们最初的信任。你试试他给出的例子：“白鱼吃糖果。”你有可能意识到一个关于鱼和糖果的模糊印象，这个印象的产生过程，就是联想记忆自动搜索“鱼”和“糖果”这两个概念之间各种联系的过程，这一过程会使这种很荒唐的说法看起来竟有些道理了。

吉尔伯特认为系统2的工作就是不信任/质疑，他用一个漂亮的实验来证明自己的观点。受试者看到一些很荒谬的说法，比如“一个叫丁卡的人是一团火焰”，他们要在几秒钟内说出这个句子的说法是“对”还是“错”。随后实验人员要检验一下受试者是否记得哪些句子是“对”的。这一实验还有一个条件，即受试者在执行任务的过程中还要按照要求记住一些数字。系统2的干扰产生了选择性效果：它使人们很难“不信任”那些错误的论断。在后续的记忆测试中，筋疲力尽的受试者最后竟认为许多错误的论断都是正确的。这个实验的寓意是深刻的：有系统2参与时，我们几乎会相信所有事情。因为系统1不仅好骗，还容易产生偏见，而尽管系统2掌管怀疑和不信任的大权，但是它有时很忙，不忙时也很懒惰，总会擅离职守。的确，已有证据显示，当人们劳累或是精力耗尽时，更容易受那些空洞却有说服力的信息影响，例如广告。

联想记忆的运算是导致“确认偏误”的原因之一。如果有人问你：“山姆友好吗？”你就会想到山姆的各种行为举止；而如果有人问你：“山姆是不是很不友好？”你就很难想起他的许多举动。有一项专门针对证据确认问题进行的名为“积极测试策略”的研究，这项研究主要考察系统2是如何验证假设的。一些自然科学家认为应通过驳斥假设以证实其是否成立，但其他人（通常是科学家）则主张通过寻找符合他们当前观点的数据来证实假设是否成立。系统1产生的确认偏误不加批判地接受了建议，夸大了极端的可能性以及不可能的事件。如果被问及海啸在今后30年内袭击加利福尼亚州的可能性有多大，你脑中所呈现的图像很有可能就是海啸，你会更容易高估出现灾难的可能性。这就如同看到同吉尔伯特提出的那些类似“白鱼吃糖果”这种无意义的陈述时产生的反应一样。

### 光环效应与群体的智慧

如果你赞同一个总统的政见，你可能也会喜爱他的声音及着装。喜爱（或讨厌）某个人就会喜爱（或讨厌）这个人的全部——包括你还没有观察到的方面——这种倾向就叫做光环效应。这个术语已在心理学领域使用了长达一个世纪，但仍然没能成为日常用语。这是一件憾事，因为光环效应这个说法很好地诠释了我们生活中普遍存在的一种偏见，这种偏见在我们塑造对人与环境的看法时起着很大的作用。系统1可以通过很多比现实更简单却更连贯的方式来表现这个世界，光环效应就是其中一种。

你在某派对上遇到了个名叫琼的女士，发现她既漂亮又善谈。现在，她的名字再次出现，并有可能是被叫去捐款。你知道琼有多慷慨吗？正确答案是：你事实上什么都不知道，因为没有理由可以让你认为善于社交的人在慈善方面会表现得慷慨。但你喜爱琼，当你想到琼时，那种喜爱的感觉会再次涌上心头。你自己慷慨，也喜欢慷慨的人。通过联想，你预先倾向于相信琼是慷慨的。现在，你认为琼是慷慨的，你可能会比以前更喜欢她，因为你又增加了一条令她讨你喜欢的特点。

在琼的这则故事里，我们并没有她慷慨程度的真正证据，而是凭借自己对她的情感回应作出猜测，用猜测弥补证据的缺失。在其他情况下，证据会逐渐出现，由第一印象产生的感觉会影响你对事物的解读。所罗门·阿希（Solomon Asch）的一个实验堪称心理学实验中不朽的经典。他对两个人进行了描述，并要求其他人对这两人的个性进行评论。你认为艾伦和本这两人怎么样？

艾伦：聪明—勤奋—冲动—爱挑剔—固执—忌妒心强

本：忌妒心强—固执—爱挑剔—冲动—勤奋—聪明

如果你像我们一样，你就会更喜欢艾伦一些。前几条列出的性格特征会改变后面出现的特征的含义。我们认为聪明人有理由固执，并且还会尊重他这一点。然而，一个忌妒心强又固执的人如果还很聪明的话，他身上就带有一些危险性。光环效应也可以化解歧义：如同“bank”这个单词一样，“固执”这个形容词也是有歧义的，但将它放在一定情境中歧义便会被化解。

对于这个研究主题，还有许多衍生出来的实验。在一项研究中，受试者要先考虑一下描述艾伦的前3个形容词，然后再考虑剩下那3个描述艾伦的形容词，但实验人员告诉他们，后3个词是用来描述另外一个人的。随后，当受试者根据这些描述假想出两个人时，实验人员问他们，这6个形容词是否有可能用来形容同一个人时，大多数受试者都认为不可能！

我们对一个人性格特征的观察顺序是随机的。然而，顺序的确很重要，因为光环效应注重第一印象，而后续信息在很大程度上都被消解掉了。在我刚做教授时，我评价学生论文的方式很传统。我会顺序一次取一本论文，一边读一边打分，然后计算出总成绩，之后再接着批改下一个学生的论文。最后，我注意到自己对每一本论文的评估都出奇地相似。我开始怀疑我的评分产生了光环效应，即第一次评分对接下来的所有评分都产生了一定影响。这样的机制十分简单：在我给某学生的第一份论文打了高分后，再碰到有模糊或是有歧义的陈述时都会对这个学生手下留情。这种做法看似合理。一个学生能将第一份论文写好的话，就不应在第二份论文中犯低级错误！但我评分的方法却存在一个严重的问题。如果某个学生写了两份论文，一份论点有力，另一份却经不起推敲，我则会因为批改的顺序不同而给出不同的分数。我曾经告诉学生两份论文的评分标准是一样的，但事实却并非如此：相比第二份论文来说，第一份论文对于总分的影响更大。所以这样的做法是不可取的。

我采取了一种新的做法。我阅读并批改了某个学生第一个问题的答案，然后接着改下一个学生的相关论述，而不是按顺序读完第一个学生的整本论文才评阅下一份。我确定将所有分数都写在了论文集的封底，以避免在批改第二份论文时出现偏见（即使是无意识的）。在改变方法后不久，出现了一个令人担忧的结果：我对评分的自信程度比原来更低了。原因在于我频繁地感到一种不适。当我对某个学生的第二份论文感到失望时，便在他的论文集封底记下一个低分，但却偶然发现我给他的第一份论文打了最高分。我也注意到了我会不经意地改变尚未写下来的分数以缩小两份论文之间的分数差，并且，我很难克制住自己不这样做。我对同一个学生的论文评分通常相差巨大。这样的前后不一使我感到不确定和沮丧。

我现在对于自己的评分行为感到失望和不自信，但我认为这是件好事，因为它表明现在



这个方法比原来的有进步。我先前所感到的一致性是伪造出来的，它使我有了认知放松之感，我的系统2也欣然接受了最后的分数。采用新方法之后，虽然第一个问题深深影响到我对之后学生的评估，但我允许自己这样做，因而学生某些问题答得好和不好的分数差别也就不是我有意为之的了。但我发现同一个学生在回答一个问题时表现优异，而对另一个问题的解答却很糟糕。当我改变方法时，这种令人感到不适的前后不一致就显露了出来：它不仅反映出随便拿一个问题去衡量学生水平的做法是不恰当的，还反映出我自己打出的分数同样不可靠。

我采取的避免光环效应的评卷方法遵循了一个普遍原则：消除错误的关联！为了了解这一原则的工作原理，我们设想一下：向大量观察者展示一些装有硬币的玻璃罐，让他们估计一下每一个罐里硬币的数量。詹姆斯·索罗维基（James Surowiecki）在他最为畅销的著作《群体的智慧》（The Wisdom of Crowds）一书中解释道，一个人单独完成这个任务的效果并不理想，但一群人共同作出判断时准确率就很高。有些人高估了硬币的数量，另一些人低估了它，但对所有判断进行平均估算得出的平均值就会趋近于准确值。这种机制很容易理解：每个人都观察着同一个玻璃罐，他们的判断都基于一个共同的基础。另一方面，每个人犯的错误都与其他人的错误无关，（在没有系统性偏见的情况下）这些错误的平均值趋近于零。然而，只有在每个人的观察相互独立、每个人所犯错误之间不相关联的情况下，降低错误率的奇迹才能出现。如果观察者持有相同偏见，就算将他们的判断汇总起来也难以降低错误率。允许不同观察者之间相互影响会减少样本量，进而影响小组估值的准确率。

想要从大量证据来源中获取最有用的信息，你应设法使这些来源相互独立。这也是警察办案时所遵循的规则。如果某个案件有多个目击证人，在录口供之前，这些证人是不能获准讨论案件的。这样做不仅是为了防止不怀好意的证人相互串通，还避免了没有偏见的证人相互影响。交流过各自目击过程的证人容易在证词中犯相似的错误，降低了他们所提供信息的总体价值。减少信息来源中的冗赘信息总是没错的。

企业高管需要花大量时间主持会议，独立判断原则（及解除错误关联）可以直接应用到这些工作中。一条简单的规则就能发挥作用：在开始讨论某个问题之前，先让与会的每一位成员各自写下简短的意见阐明自己的观点。这个过程很好地利用了小组里不同知识和见解的价值。而开放性讨论这一常规做法总会注重那些发言早而又强势的人的意见，使得其他人一味附和他们的观点。

眼见为实的想法往往让我们仓促作出决定

我和阿莫斯早期合作时最美好的记忆中有一点很难忘，就是他总是乐此不疲地重复一个

搞笑情节。阿莫斯能惟妙惟肖地模仿他大学时代的一位哲学老师，他用带着浓重德国口音的希伯来语咆哮着说：“你必须时刻牢记‘Primat of the Is’。”我从未弄明白他的老师所说的那句短语是什么意思（我想阿莫斯也不明白吧），但是阿莫斯却还是会说那些笑话。每当我们遇到大脑对现有信息和未知信息的处理方法严重失衡的问题时，他总会想起那句莫名其妙的短语（我最后也总是这样）。

联想机制一个最基本的结构特点就是它只能回忆起已被激活的观点。无法从记忆中获取的信息（即使是无意识的）可能并不存在。系统1善于提取当前激活的想法来构建最可信的故事情节，但它不会（也不能）提取本系统中根本不存在的信息。

衡量系统1是否成功的方法是看它所创造的情境是否具有连贯性，而与故事所需数据的数量和质量关系不大。信息匮乏是常事，一旦出现这种情况，系统1则会仓促作出结论。请思考下面的说法：“明迪克会是一个出色的领导吗？她聪明又坚强……”你的脑海中一定会马上闪现出一个答案——“当然会”。你根据非常有限的信息选择了一个最佳答案，但是你却行动过早。试想一下，如果紧随其后的两个形容词是“腐败”和“严酷”，你该怎么办？

在对明迪克这位领导作出快速判断时，请注意有些事你并没有做，你没有自问：“在形成对某人是否具有领导才能的看法之前，我应该了解些什么？”系统1在第一个形容词出现后就开始自主运作了：聪明是好的，既聪明又坚强就更好了，系统1非常轻松地生成了这种想法。如果有新信息出现（例如明迪克思想腐败），这个故事就会被改写了，但是系统1并不会等待也不会出现主观上的不适。对于第一印象的偏好依然存在。

寻找连贯性的系统1和懒惰的系统2相结合，意味着系统2将会赞同许多直觉性的信念，而这些信念又准确地反映了系统1产生的印象。当然，系统2也能对证据采取系统而谨慎的处理方法，还能在作决定之前考虑出现的众多选项——设想你正在卖房子，你就会设法搜寻一些你不知道的信息。然而，即使是深思熟虑后的决定，系统1也能对其产生影响。系统1的信息输入从未停止。

在证据不足的情况下过早下结论对我们理解直觉性思考非常有帮助，本书也常提到这样的情况，我将会用一个冗长的缩写来代表这种情形：WYSIATI，意为“*What you see is all there is*”，即眼见即为事实。系统1基本上对于引起印象和直觉的信息的质量和数量都不敏感。

阿莫斯与他在斯坦福大学带的两名研究生作了一项与“眼见即为事实”直接相关的研究。受试者得到了一些证据并且知道这些证据是片面的，而研究者则是要观察他们拿到证据后的

反应。这些受试者接触到的都是法律案例，下面就是一例：

43岁的被告戴维·桑顿（David Thornton）是工会界代表。9月3日，他来到了“平价大药房”168号进行例行视察。他来到这家药店还不到10分钟，一名驻店经理就过来告诉他不能再站在店里和员工们说话，他只能利用员工休息时间在一间密室里见见他们。在工会与“平价大药房”的合约中，这一要求是得到允许的，但从未执行过。当桑顿拒绝这个要求时，经理告诉他，要么遵守约定，要么离开药店，否则他将被逮捕。此时桑顿暗示经理，在不妨碍生意的情况下，他一直以来都是在店里与员工交谈的，每次谈话时间也就是10分钟左右，从来也没有人反对过这一做法，他宁愿被抓也不愿改变例行的视察程序。于是，经理叫来了警察，警察以非法侵入的罪名逮捕了桑顿。在桑顿留了案底并被关押在拘留室一段时间后，所有的指控都被取消。现在，桑顿准备起诉“平价大药房”非法拘留。

所有受试者都阅读了这些背景材料，除此之外，每一组受试者还听了控辩双方代理律师所作的陈述。情况自然是工会方的律师将逮捕视为恐吓行为，药店方的律师则认为在药店进行谈话扰乱经营秩序，经理的行为是合理的。还有一些受试者就像陪审团成员一样听了控辩双方的陈述。双方律师并没有增加任何背景材料中未提及的有用信息。

所有受试者都充分了解了整个过程，那些只听到其中一方辩词的受试者能够很轻松地为一方写出辩词。然而，片面的证据陈述对判断有着重大影响。另外，只掌握一方证据的受试者比掌握了双方证据的受试者更有自信。这正说明人们根据已有信息勾勒出的故事的连贯性增强了他们的自信心。一个好故事最重要的是信息的前后一致性，而不是其完整性。的确，你常会发现：知道得很少反而可以把已知的所有事物都囊括进连贯的思维模式中。

眼见即为事实的理念有助于达成连贯性和认知放松的状态，从而使我们相信某个陈述是真实的。这一理念解释了我们能够快速思考的原因，解释了我们是如何弄清楚一个复杂领域中那些信息片段的含义的。很多时候，我们拼凑出的连贯情节与事实是无限接近的，完全可以用来支持理性活动。而我还会运用眼见即为事实原则对判断和选择中存在的很多偏见作出解释，以下便是其中的一部分：

过于自信：正如眼见即为事实原则指出的那样，无论是证据的数量还是质量都与主观自信关系不大。每个人对自身想法的自信程度主要取决于他们对亲眼所见的事情的讲述效果，即使他们几乎什么都没有看到也没有关系。我们经常考虑不到自己有可能尚未掌握对判断起决定性作用的那份证据，却总是认为眼见即为事实。此外，我们的联想系统更倾向于选择已被激活的连贯模式，抑制怀疑和歧义。

框架效应：同一信息的不同表达方式常常会激发人们不同的情感。“手术后一个月内的存活率是90%”的说法要比“手术后一个月的死亡率是10%”更令人安心。同样，说凉菜“90%不含脂肪”要比说“10%含有脂肪”更具吸引力。很明显，前述每组句子的深层含义都是相同的，只是表达方式不同而已，但人们通常能读出不同的含义，而且觉得自己的所见就是事实。

比率忽略：回忆一下史蒂夫，那个本性怯懦、做事井井有条，常被看做图书管理员的人。这个人物性格的描述是生动形象的，尽管你清楚地知道男性农民比男性图书管理员多，但在考虑这个问题时，你总会忽略这个事实。你觉得自己的所见即为事实。

#### 示例—过早下结论的问题

“她对这个人的管理技能一无所知。之所以对他印象很好，是因为曾经听他作过一次精彩的报告。”

“在讨论之前大家先独自考虑一下这个问题，这样可以避免观点的相互干扰，这样更利于集思广益。”

“他们看了一份优质的咨询报告后就作了一个重大决定，他们并没意识到自己掌握的信息其实很少。”

“他们并不想了解更多信息，因为那样可能会破坏整个故事情节。他们更愿意相信眼见即为事实。”

## 思考，快与慢

### 第八章 我们究竟是如何作出判断的？

你可以回答无数个问题，无论这些问题是别人问你的还是你自问的。同样，你能评价的事物特征也是无数的。你能数出这页中“的”字的出现次数，能比较自己家的窗子和马路对面那家的窗子哪个高，也能对你支持的参议员的政治前景作出评价，其前途无限光明还是前景堪忧，或是碌碌无为。这些问题由系统2来解决，系统2能调动注意力并通过搜寻记忆去寻找答案。系统2接受问题或提出问题：不管是提问还是回答，它都能引导注意力并搜寻记忆来找到答案。系统1以不同的方式运行，不断监视着大脑内外发生的一切，没有特定意图，也无须付出多少努力，只是对当时的情形作出全方位评估。这些“基本的评估”在直觉性判断中扮演了重要角色，因为人们常会拿它们来替代更难的问题——这也是启发法和偏见研究方法

的基本理念。系统1其他两个特点也支持用一种判断代替另一种判断的做法。其中一个特点就是系统1具备跨维度解读价值观的能力，你可以回答一个大多数人都觉得很简单的问题：“如果山姆的身高和智商一样，那么他究竟有多高？”此时思维快捷方式便开始运行了。系统2会集中注意力回答某个特定问题，或是对某种情况的特殊属性进行评估，集中的注意力又会自动运行其他的评价程序，包括一些基本判断。

看照片预测竞选胜出者，为何其准确率可高达70%？

随着人类进化不断完善，系统1可以对生物体生存必须解决的主要问题提供一个连续的评估，这些问题包括：事情进展得怎么样了？我们面临的是威胁还是机遇？一切都正常吗？我应该是前进还是退避呢？这些问题也许对于生存在城市中的人而言，不像对大草原上的羚羊那样紧急，但我们有不断进化的遗传神经机制，可以持续不断地对威胁水平进行评估。我们通常用好与坏来评价不同情形，要么说要避开这种情况要么说可以泰然处之，没有问题。人的好心情和认知放松与动物对安全和熟悉程度的判断是相当的。

如果想找一个“基本判断”的典型例子，想想只需一瞥就能区分朋友和敌人的能力就是了。这种能力能够提高人们在危险世界的生存概率，而这种专属能力也在不断增强。我在普林斯顿大学的同事亚历克斯·托多罗夫（Alex Todorov）曾经对与陌生人接触的安全性问题作出快速判断这一能力的生物学根源进行了探索。他认为我们生来就具有判断的能力，只需瞥一眼陌生人的脸，就能对这个人的两点重要事实作出判断：他有多强势（因此存在潜在的威胁性）；这个人有多可信（不管他的用意可能是友好的还是充满敌意的）。脸型为判断提供了许多暗示：方下巴就是强势的信号。面部表情（微笑或皱眉）是对陌生人意图的判断提示，方下巴加上瘪嘴唇也许就预示着有麻烦了。看脸形的精确性不是很高：圆下巴并不代表温顺，笑容（在某种程度上）也是可以伪装的。不过，即使对陌生人作出判断的能力不高，具备这种能力也是我们的生存优势。

这种古老的机制在现代社会得到重新利用：它对人们如何选举有些影响。托多罗夫向他的学生展示了一些人脸的图片，有时展示的时间只有0.1秒，他让这些学生按不同属性对这些面部图片进行评估，这些属性包括可爱程度和做事能力。结果所有学生对这些图片的评估结果非常一致。托多罗夫展示给学生的这些人脸图片并不是随意组合的，而是参加竞选的那些政治家的照片。大选结束后，托多罗夫将选举结果和普林斯顿大学学生所作出的能力评估进行了比较，这些学生当时并不了解这些候选人的任何政治背景，仅凭自己对这些照片的匆匆一瞥就作出了评估。事实证明，约有70%的参议员、国会议员和地方长官的竞选活动的胜

出者也正是那些在照片评估中获得较高评价的人。这一惊人结果在芬兰的全国大选中得到证实，同样的情况也发生在英国的地区选举中，澳大利亚、德国和墨西哥的众多选举中也发生过类似事件。令人惊奇的是（至少对我而言是这样的），在托多罗夫的研究中，能力评估远比可爱程度的评估对选举结果的预见能力强。

托多罗夫发现，人们总会结合力量和可信度两方面因素来评估一个人的能力。刚毅的下巴和自信的微笑便可告诉我们，这个人很有能力。没有证据显示这些面部特征确实能预示某些政治家可以当选，但关于人们对胜出和出局候选人的判断研究显示，我们往往在投票前就会对那些不具备我们认可的面部特征的候选人持否定态度。在他的研究中，失败者引起的（负面）情感回应更强烈，我将这个例子称做“判断启发法”案例，接下来的章节中会沿用这个说法。投票者尝试着对候选人将来的任职表现生成一种印象，他们又转而依靠系统2快速自主地作出一种更加简单的判断，这一系统只有在必要时才会作出这一判断。

许多政治学者也循着托多罗夫最初研究的路子继续深入研究这一问题，他们划定了一类投票者，这类投票者往往会不由自主地听从系统1 的指挥。这些投票者经常看电视，对政治却所知甚少，而那些政治学者在他们身上找到了自己一直在寻找的东西。不出所料，对于那些信息贫乏、爱看电视的投票者来说，面部特征表现出的能力对其投票的影响较大，其受影响程度约为那些信息丰富、看电视少的投票者的3倍。显然，系统1对投票选择的影响因人而异，下文中我们还会遇到一些体现个体差异性的例子。

当然，系统1理解语言，这种理解是建立在一些基本判断基础之上的，而这些判断通常又是在洞察事实和理解信息的基础上作出的。这些判断包括对相似度和代表性的判断，对因果关系的属性以及对联想和样本的可用性的判断。尽管判断的结果是用来满足任务要求的，但是没有具体任务时，这些判断活动照样也在进行着。

基本判断的内容很多，但并不是每个可能的属性都需要判断。例如，我们可以简单看看图7。

一眼看去，你便会对该图的很多特征有个初步印象。你知道左右两个长方体一样高，也很相似。然而，左边那个长方体的方块数和铺在平面上的方块数是不是一样，这可不是瞥一眼就能明了的事，而且你也想不出这堆方块能摞成多高的长方体。为了证实左面长方体的方块数目和中间的方块数目相同，你需要数一下这两堆方块，对比一下结果，这个活动只有系统2能完成。

平均长度与总长度是完全不同的数量概念

还有一个例子，请看这个问题：条线的平均长度是多少？

这个问题很简单，系统1无须任何提示就能回答。实验显示，人们在不到一秒钟的时间里完全可以精确地记下很多线段的平均长度。此外，观察者的认知系统即使正忙于记忆，这些判断的精确度也不会受到影响。认知系统未必知道如何用英寸或厘米为单位来描述平均值，但是如果让其判断另外一条线是否符合平均值，它们的判断也是非常精确的。对于一组线的长度标准生成一个印象无须系统2的参与，系统1会自主且毫不费力地完成这一任务，就像它记一组线的颜色和它们之间不相平行的事实一样轻松。我们也可以立刻对众多物品的数量生成一个印象，如果其数量只有4个或少于4个的话，印象会很精确；如果更多的话，就会变得模糊。

现在我们来讨论另外一个问题：中所有线的总长度是多少？这是一种全新体验，因为系统1无法为回答这个问题提供建议。为此题作答的唯一办法就是启动系统2，系统2会尽力估计平均值，评估或数出有几条线，用平均长度去乘条数得出结果。

仅凭一瞥系统1计算出一组线的总长度，其结果多半不对，这一点你很清楚。你认为自己绝不会这样做。事实上，这是该系统一个重要的局限性。因为系统1通过原型或一组典型事例来代表不同事物分类，它能解决好平均问题，但对总量问题就束手无策了。一个类别的规模及其所包含的实例数量，在我们判断总额变量时常常被忽略掉了。

在我们进行的众多实验中，有一项是根据那次损失惨重的埃克森-瓦尔德斯号（Exxon Valdez）原油泄漏事件的诉讼而设计的，我们询问受试者是否愿意掏钱买网来覆盖油池，因为这些油池常淹死迁徙的鸟类。受试者组成的不同小组分别表明了各组的意愿，他们愿意掏钱来拯救鸟的数量分别为2 000只、20 000只和200 000只。如果拯救鸟类是个经济善举的话，其价值大小就要看总数这一变量了，即拯救200 000只鸟应该比拯救2 000只鸟更有价值。事实上，3个组的平均捐款分别是80美元、78美元和88美元，与鸟的数量没有什么关系。3组受试者做出反应的对象为原型——一只无助的小鸟被淹死的可怕画面，鸟的羽毛浸泡在黏稠的原油中。实验人员屡次发现，在这样的情形下，受试者几乎完全忽略了数量的概念。

### 与强度等级匹配的描述

诸如你的幸福感、总统受欢迎的程度、金融骗子的合理惩罚和政治家的未来前景等问题有一个共同的重要特点：这些问题都涉及隐含的强度或数量概念，因而我们就可以使用“更”这个词对其进行描述：更幸福、更受欢迎、更严厉或（对政治家来说）更有力度。例如，

一个候选人的政治前景可能是“她在首轮竞选就会出局”这样的背运，也可能是“她有朝一日会成为美国总统”，身居高位。

接下来我们会了解到系统1的又一新能力。强度的等级在不同领域中都有“匹配”描述。如果罪行是颜色，杀人就应该是深红，颜色比偷窃更深。如果犯罪用音乐来表达，大屠杀就应该用强音，而停车不付钱则应该用弱音。当然，你对惩罚的强度也有类似的感觉。在传统的实验中，有些人用声音的大小来表达犯罪的严重性；其他人用声音大小来表达法律惩罚的严重性。如果你听到了两个声音，一个是表达犯罪的，一个是表达惩罚的，如果一个声音比另一个声音更响的话，你会有不公平之感。

请思考这个例子（后文中还会提到这个例子）：

朱莉4岁时就能阅读。

现在请将朱莉这个孩子的阅读能力与下面的强度等级进行匹配：

若某人的身高和朱莉的早慧程度一样，那他有多高？

你觉得6英尺怎么样？显然太少了。那7英尺呢？也许又太多了。你希望找到一个高度能匹配4岁孩子极强的能力。虽然很强，但并不超群。15个月大就能阅读才是超群的能力，这就跟一个人身高7.8英尺一样。

你的工作收入多高才能与朱莉的阅读能力相匹配呢？

什么罪行的严重程度可以与朱莉的早慧程度相匹配呢？

常春藤大学的毕业学分积点多高才能与朱莉的阅读水平相匹配呢？

上述问题并不是很难回答，对吧？此外，可以肯定的是与你同处一个文化领域的人作出的匹配与你的回答会很相近。我们发现，人们根据朱莉的阅读年龄这一信息预估她的学分积点时，他们通过一种范畴向另一范畴的转换来回答这个问题，并且选出了相应的学分积点值。我们也明白为什么这种利用匹配进行预测的模式从统计学角度来看是错误的，尽管对于系统1来说这很正常，但对于统计学家以外的大多数人来说，系统2也可以接受这种做法。

思维的发散性让我们作出直觉性判断



系统1任何时候都可以同时进行多种估算，其中有些估算是持续不间断的常规评估。只要眼睛是睁开的，你的大脑就会对视觉范围内呈现出的立体事物进行评估，这种评估是对这些物体的形状、空间位置和特性等因素的全方位评价。这一评估活动的运行或对违背期望的事物进行持续监督的行为都是无意识的。与这些常规评估不同，其他评估行为只有在需要时才会进行：你不会持续评估自己有多高兴或多富裕，即使热衷政治，你也不会一直不间断地评估总统的执政前景。偶尔的判断是主观自愿的，这种判断才是有意识的。

你不会不由自主地数出每个读到的词的音节数，但如果你选择这样做，就能数对。不过，想要使刻意计算的结果很精确并非易事：我们计算的结果往往比自己想要的或需要的要多。这种过量计算的过程就体现了“思维的发散性”。如同想用散弹猎枪瞄准一个点是不可能的一样（它射出的子弹是分散的），想要让系统1完全执行系统2的命令且不做多余的工作也很难，这一点与散弹枪很相似。我很久以前从书上看到的两个实验就表明了这一点。

其中一个实验让受试者听几对词，若他们听出这些词是押韵的，要马上按下一个键。下面两组词都是押韵的：

VOTE-NOTE

VOTE-GOAT

在你看来，区别很明显，因为你看到了这两组词，而受试者只能听到单词。“VOTE”和“GOAT”押韵，但它们的拼写不同，虽然受试者听到了这两个词，但他们也会受到拼写的影响。如果两个词的拼写不同，受试者听出它们是押韵的速度就会慢些。尽管要求是比较声音，但受试者同时也对两者的拼写进行了比较，而且与声音无关的不匹配因素妨碍他们迅速作出判断。刻意回答一个问题却引起了另一个问题，这一行为不仅没必要，而且对主要任务的完成也很不利。

在另一项实验中，受试者听了几个句子，如果句子是真实的，就要马上按下一个键，如果是假的，就按下另一个键。对以下这些句子的正确回应是什么呢？

有些路是蛇形的。

有些工作是蛇。

有些工作像监狱。

这三个句子从表述上来看都是错的。不过，你很可能已经注意到了第二个句子比另两个句子错得更明显，实验也证实了这一本质性不同。之所以存在这种不同，是因为第一句、第三句两个难句从比喻角度看是正确的。这次又是要进行一个预估活动却引起了另一个预估行为，而且，正确答案在冲突中更明显，但这个与回答并不相关的冲突却影响了系统的正常运行。在下一章中我们会发现，思维的发散性和强度匹配结合起来就可以解释为什么我们对很多自己不很了解的事情能够作出直觉性判断。

#### 示例—判断问题

“评价一个人是否有吸引力是一种基本判断，不管你是否想这样做，这种评价都是不由自主进行的，也会对你产生影响。”

“我们的大脑中有一些线路，这些线路可以从脸型来推断一个人统领大局的能力，即他看上去有些领导气质。”

“如果强度与罪行不匹配，惩罚则不可能公正。就像是你可以用光的亮度来与音量的大小匹配一样。”

“关于思维的发散性，有这样一个明确的例子：他被问及是否认为这家公司财力雄厚时，他想到的却是该公司令其钟情的产品。”

## 思考，快与慢

### 第九章 目标问题与启发性问题形影不离

关于你的思维活动，有一点值得注意，即你很少被问题难倒。的确是这样，偶尔你会碰到这样的问题： $17 \times 24 = ?$  你无法立即想到这个问题的答案，但这种让人目瞪口呆的时刻毕竟是少数。当大脑处于正常的状态时，你几乎对眼前出现的所有事物都会有直觉和想法。对一个人不是特别了解时，你就知道自己是不是喜欢他；你也不知道自己为什么会相信或者不相信一个陌生人；你没有作过调查分析，却能感觉到一家企业一定会成功。有时，对于一些无法完全看懂的问题，你也总能作出答案，而回答的依据是什么，连你自己也说不清道不明。

找个相对简单的问题来作答

我简单描述了如何从复杂的事情中提取直觉性的观点。对于有难度的问题，我们总是很难快速找到令人满意的答案，此时系统1就会找到一个相关问题来回答，这个问题比原来的问题更易作答。我把这种回答一个问题而绕开另一个问题的做法叫做“替代”。我还将采用以下术语：

“目标问题”就是你想要作出的评估。

“启发式问题”就是你绕开原来的问题去回答的那个更简单的问题。

“启发”这一术语是指协助寻找各种难题的恰当答案的简单过程，虽然找到的答案常常并不完美。这个词和希腊语 *eureka*（意为“找到了”）是同根词。

我和阿莫斯在共同工作的早期就想到了替代这一概念，它也是启发法和偏见研究法的核心内容。我们自问：人们在对可能性没有任何了解的情况下，是如何成功作出可能性判断的呢？我们得出的结论是，人们一定不知道通过什么方式将不可能完成的任务简单化了，由此，我们开始研究他们是如何做到这一点的。我们的答案是，当人们按照要求对可能性作出判断时，他们实际上是对其他的事情作了判断，并且认为自己已经完成了判断可能性的任务。遇到很难的“目标问题”时，如果脑海中马上出现了一些与之相关联且容易回答的“启发式问题”的答案，系统1通常便会采取这种“替代”的做法，采用替代问题的答案。

用一个问题替代原来的问题是一个解决难题的好策略，乔治·波利亚（George Pólya）在他的经典著作《怎样解题》（*How to Solve It*）中提到了替代问题：“如果你无法解决某个问题，就去解决另外一个简单点的问题好了——去找这个简单的问题吧。”波利亚的启发法是系统2有意实施战略性决策的过程。不过，我在本章讨论的启发法案例并不是精心挑选的，而是思维发散性造成的结果，是我们锁定问题答案控制能力不强的结果。

请看下面的“目标问题”。它们都很难，在给出合理的答案之前，你肯定会先处理些其他难题。幸福的含义是什么？未来6个月会有哪些可能的政治动向？其他金融犯罪的标准量刑是什么？候选人的竞争有多激烈？还有哪些环境因素或者其他原因需要考虑？想要真正解决这些问题根本不可能。但是你不要受制于完全理性的答案，经过认真推理之后，可以找到启发性的替代答案，尽管这种方法有时很好用，而有时却会导致严重的错误。

目标问题                  启发式问题

你愿意为拯救濒危物种捐多少钱？                  想到垂死的海豚时，我的情绪波动有多大？

你这些天对生活满意吗？                  我现在的心情如何？

今后6个月里总统受欢迎的程度如何？                  总统现在有多受欢迎？

欺骗老人的理财顾问应受到何种惩罚？                  想到金融骗子，我会有多生气？

这位竞选总统的女士能在政治道路上走多远？                  这位女士看上去像是个政治赢家吗？

思维的发散性可以使懒惰的系统2摆脱很多繁重的工作，快速找到难题的答案。“启发式问题”相对于“目标问题”更容易让人产生联想并作出回答。你对海豚和金融骗子的感觉，你现在的心情，你对自己看好的竞选人政治能力的印象，或者总统现在的立场等，都会立即呈现在你的脑海中。启发式问题给每个有难度的目标问题都准备了一个现成的答案。

上述说法中存在一定的不足：没有指出答案应该与原始问题存在关联。例如，我对垂死海豚的感觉一定要用钱来表示，这一点让人费解。而系统1的另一项能力，即强度匹配，可以解决这个问题。情感和捐款都是按强度分级的。我对海豚的感情可能很强烈，也可能很淡薄，而我捐款的数目会与我对其的感情强度相匹配，我脑海中出现的钱数就是匹配量。相似的强度匹配可能适用于所有的问题。例如，不同竞选者的政治表现可以从很糟糕到非常出众。同样，衡量政治成功等级也包括“她在首次竞选就会出局”的最差表现和“她将来会成为美国总统”的最佳表现。

思维发散性的自主过程和强度匹配可以使能映射到目标问题的简单问题有一个或多个答案。有些情况下会有替代答案，系统2会认可这一启发式答案。当然，系统2也可以拒绝这个直觉性答案，或者通过整合其他信息来改变它。但是，一个懒惰的系统2往往会遵循最省力法则，不经检验就认可某个启发式答案。你不会被难倒，也不必非常努力，甚至都没有注意到自己其实答非所问。此外，你也许都注意不到目标问题很难，因为你的头脑中会很迅速地闪出一个直觉性答案。

立体启发法：远处的物体看上去更高大

请看下面这幅图片并回答下面的问题。

如图所示，右侧的人是否比左侧的人高大？

我们很快就想到了那个最明显的答案：右侧的人更高大。但如果你用尺子量一下，你会发现实际上他们体型一样。你对他们相对体型的印象受到强大错觉的控制，这种错觉也正是替代过程的最佳解释。

图中人物经过的走廊被画成了透视图，显得平面越来越深。你的感知系统自动地将这幅图看做是立体的，而不是印在平面上的。在立体透视图，右侧的人比左侧的人离我们更远，显得更高大。大多数人都会对立体图感觉很强烈。只有视觉艺术家和有经验的摄影师才具备将这幅图看成是平面图的能力。对于我们这些人来说，此时就会有替代产生：对立体图像大小的印象左右了我们对平面图大小的判断。错觉的产生是由立体图的启发导致的。

此时出现的是一个真正的错觉，不是对问题的误解。你知道是图片中的3个人的体型造成了我们的错觉，因为它们是印在纸上的。如果有人让你估计这几个人的体型，我们通过实验就会知道你的答案会是用英寸来表达的，而不是英尺。你对问题没有疑问，但你会受到那个并没有被问道的问题的影响，比如：“这3个人有多高？”

上例中启发法的那个重要环节——立体大小替代平面大小——是自主发生的。图片中包含着一些暗示，这些暗示会让图片产生立体感。尽管这些暗示和要完成任务——判断纸上人物的体型大小——并无关联，你本该忽略这些因素，然而你却做不到。启发法的偏见是图中看上去更远的物体看上去更大。正如上例所示，基于替代的判断必然会如预想的那样有失偏颇。在这个例子中，基于替代的判断在感知系统中运作得比较隐蔽，根本就是不由自主发生的。

一些德国学生作过的一项调查堪称替代研究的最佳案例之一。这些年轻人完成的调查包括下面两个问题：

你最近觉得幸福吗？

你上个月有多少次约会？

实验人员对这两个问题的答案的关联度很感兴趣。那些回答自己有很多次约会的学生会比那些约会次数少的人更幸福吗？令人吃惊的是，答案是否定的，两个答案的关联度几乎为零。显然，学生在评价自己的幸福感时，首先想到的并不是约会。另一组学生也看到了这两

个问题，但次序正好相反：

你上个月有多少次约会？

你最近觉得幸福吗？

这一次的结果完全不同。在这种顺序下，约会次数和幸福感之间的相关度能达到心理测试的最高水平。其间发生了什么呢？

原因很明确，这就是替代作用的典型案例。约会显然不是这些学生生活的重心（第一次调查中，幸福和约会并不相关），但当实验者要求他们回想自己的浪漫生活时，他们确实有情感上的回应。有多次约会的学生想起了自己生活中快乐的事，而那些没有约会的学生想起的都是孤独和被拒绝的情节。因此看到第二个问题时，由（第一个）约会问题引起的那些情感就在大家的脑子里徘徊，影响他们对这个问题的回答。

这种任务并不常见，也不简单，需要仔细思考之后才能给出恰如其分的回答。但是，提到约会情况，这些学生就无须努力思考了，因为他们脑海中早已有了相关问题的答案。这个相关问题就是：他们对自己的恋爱生活满意程度如何。他们将被问道的问题被替换成了另一个自己心中已经有了答案的问题。

在这个实验中，我们同样可以采用与错觉研究相同的做法，可以问一问：这些学生感到困惑了吗？他们是否真的认为这两个问题——他们被问道的问题和他们回答的那个问题——是相同的？当然不是。学生不会一下子把浪漫时光和生活混为一谈，如果让他们说说对这两个词的理解，他们肯定会说浪漫时光和生活不是一回事，但其实他们需要回答的问题并不是两个概念是否相同，他们的问题是最近是否感到幸福，于是系统1就用已有的答案来作答了。

约会的案例并非特例，如果这些参与实验的学生先看到的问题是与父母或金钱的关系，而后马上看到那个关于幸福的问题，也会发生同样的情形。在两种情况下，满足感在特殊情况下影响了关于幸福感的表达。任何影响人心情的重要情感问题都会产生同样的效果。眼见即为事实。当人们评估自己的幸福感时，他们思维状态的作用就显得非常突出。

情感启发式：因为喜欢，所以认同

一旦加入情感因素，结论对论证的主导作用便会最大程度地凸显出来。心理学家保罗·斯洛维克（Paul Slovic）提出了“情感启发式”的概念，认为人们的好恶决定了他们的世界观。

你的政治倾向决定了你对各类论证的看法，即它们是令人信服的还是难以服众的。如果你对当前的医疗政策还满意，就会相信该政策能给自己带来很多好处，而且你相信在这个政策下，花同样的钱能做更多的事。如果你对其他国家采取鹰的态度，就很可能觉得其他国家相对弱小，更有可能对自己的国家俯首帖耳。而如果你的态度像只鸽子，就会觉得它们更强，不会轻易受到牵制。你对辐照食品、瘦肉、核能、文身或摩托车等事物的态度会左右你对这些事物的感受，它们是惠及生活还是充满风险。如果你对这些事物通通都不喜欢，很可能会觉得它们对你而言利少弊多，根本不会给你的生活带来什么好处。

结论的至高无上并不意味着你的思维完全停止运转了，也不意味着你可以完全忽略信息和合理解释得出自己的结论。在了解到自己并不喜欢的一项活动的风险其实远比自己想象的小时，你的看法乃至情绪就会发生变化（至少会有那么一点变化）。然而，这种较低风险的信息同样也会改变你对该活动益处的看法（你认为益处会更大），尽管你接收到的信息中没有任何与益处相关的信息。

在此我们又看到了系统2的另一个“特性”。到目前为止，我似乎已经将其描述成一个给系统1留有足够余地的默许监视器。我描述的系统2还具有主动搜寻记忆的功能、复杂计算功能、比较功能、规划功能和决策功能。在球拍和球的问题中以及许多其他的两个系统相互作用的例子中，系统2似乎总是处于最高决策地位，并有能力抵制系统1的建议，它能使事情放缓，开始进行逻辑分析。自我批评是系统2的功能之一。但在态度方面，系统2更像是系统1各种情感的赞许者而非批评者，也可以说是其各种情感的转让者而非实施者。它搜寻的信息和论据多半局限于与已有看法一致的信息，并不着意对其进行调查审核。积极且追求连贯的系统1为要求不高的系统2提供了各种解决方案。

#### 示例—替代和启发法

“还记得我们绞尽脑汁去回答的那个问题吗？我们是不是把它换成了一个简单点的问题呢？”

“别人问我们的问题是这位候选人是否会成功，但我们要回答的问题似乎是她是否能成功应对采访。咱们还是别顾左右而言他了。”

“他喜欢这个项目，因此他认为该项目投入少、回报高。这是情感启发式的一个典型案例。”

“我们将去年的表现作为依据来预测公司未来几年的表现。这个依据能作为有效参照

吗？我们还需要其他什么信息才能作出正确预测？”

下面列出了系统1的特点和活动。每个主动句都代替了一个陈述句，虽然表述更精确了，但却更难理解了，因此大脑便开始自主且快速地运行。我希望这个对大脑特性的描述能帮你形成一种对虚拟的系统1各种“特性”的直觉认识。与身边的很多人一样，你对系统1在不同情况下如何运作会有直觉的认识，而且其中大多数直觉都是正确的。

### 系统1的特点

生成印象、感觉和倾向；当系统2支持这些行为时，它们就会成为信仰、态度和意图。

自主且快速运行，只需付出较少努力，甚至不用付出努力，没有自主控制的感觉。

当发现（搜寻）特殊形式时，能接受系统2编控来调动注意力。

在接受了一定的训练后，能够做出熟练的回应，产生直觉。

为联想记忆激发出来的各种想法创造连贯形式。

将认知放松感和真理错觉、愉快的感觉以及放松的警惕感联系起来。

区分常态中令人惊奇之事。

推断原因和意图。

忽略歧义，按捺住心中的疑问。

夸大情感的一致性（光环效应）。

将注意力集中在当前的证据上，忽略不存在的证据（眼见即为事实）。

作一些基本估测。

通过常态和原型来表现集合，但不要将两者看成一个整体。



通过测量确定不同程度对应的匹配物（比如音量的大小）。

真正去计算，而不是空想（思维的发散性）。

有时用简单点的问题替代难题（启发法）。

对变化的感知比对形态的感知更敏锐（前景理论）。\*

对可能性作出过高估计。\*

对数量越来越不敏感（心理物理学）。\*

对损失的反应比获得更强烈（损失厌恶）。\*

严密设计决策问题，分别进行讨论。\*

\*带\*的内容将在本书第四部分作详细介绍。

## 思考，快与慢

### 第十章 大数法则与小数定律

#### 第二部分 启发法与偏见

一项研究对美国3 141个县的肾癌发病率进行了调查，调查显示该病的分布模式很值得注意。发病率最低的县差不多都位于中西部、南部和西部人口稀少的乡村，这些区域按照惯例由共和党管辖。对此，你有何看法？

刚刚过去的几秒钟里，你的大脑处于非常活跃的状态，这主要是因为系统2在运行。你谨慎地在记忆中搜寻着并作出假设。在这个过程中你也付出了一定的努力，你的瞳孔会扩张，心跳会适度加快。系统1也没有闲着，因为系统2的运行需要从联想记忆中获取事实和建议。你很可能会否认共和党的政策提供了肾癌防控方法这个想法，却会关注肾癌发病率低的县大多是乡村这个事实。这个例子是我从机智的统计学家霍华德·维纳（Howard Wainer）和哈里斯·泽维林（Harris Zwierling）那儿得到的，他们对这一案例的评论是：“人们很容易作出推

断，认为肾癌发病率低主要是由于乡村的生活方式很健康——没有空气污染和水污染，食品没有添加剂，保证新鲜。”这一点完全说得通。

现在，考虑一下肾癌发病率最高的县的情况吧。假设这些易发病的县差不多都位于中西部、南部和西部人口稀少的乡村，这些区域按照惯例由共和党管辖。霍华德·维纳和哈里斯·泽维林半开玩笑地评论道：“人们可以很容易作出推断，导致肾癌高发病率的直接原因是乡村生活的贫困——医疗条件差、高脂肪饮食、酗酒、嗜烟等。”当然这种说法肯定有问题，因为乡村生活方式不可能既是肾癌发病率高的原因又是其发病率低的原因。

问题的关键并不在于这些县处在乡村地区或是由共和党掌管，而在于乡村地区人口少。我们通过这个例子学到的不是流行病学知识，而是我们的大脑和统计数据之间的复杂关系。系统1非常擅长一种思维模式——自动且毫不费力地识别事物之间的因果联系，即使有时这种关系根本就不存在，它也会这样认定。当听到肾癌高发地区的情况时，你立刻会想当然地认为这些县与其他县不同是有原因的，一定有个理由可以解释这种不同。然而，正如我们所见，当系统1面对“纯统计学”的数据时是束手无策的，因为这些数据虽然可以改变结果出现的概率，却不能直接导致结果的发生。

根据定义，一个随机事件是不需要解释的，但一连串的随机事件就有规律可循。想象有一个装有大理石弹球的瓮，其中有一半的弹球是红色的，另一半弹球是白色的。然后，再想象有一个非常有耐心的人（或一个机器人）随意从瓮中取出4个大理石球，记录其中的红球数，再把球放回去，重复这样的做法数次。总结记录结果时，你会发现“2红2白”的结果出现的次数（几乎刚好）是“4个全红”或“4个全白”这种结果的6倍。这一倍数关系是个数学事实。你可以对这种从瓮中反复抽样的结果作出自信的预测，就像你能预测到用锤子砸鸡蛋的结果一样。尽管你无法预见蛋壳破碎的具体细节，但大概结果还是很确定的。两件事的不同之处在于：你想到锤子砸鸡蛋时感受到的那种明确的因果联系，在瓮中取样的设想中是找不到的。

相关的统计学事实与癌症那个例子也有联系。两个耐心的计数者轮流从瓮中取大理石球，杰克每次拿出4个球，吉尔拿出7个。他们都记录了每次拿到相同颜色弹球的次数——要么全白，要么全红。如果他们取球的做法持续的时间足够长，杰克拿到同颜色大理石的次数会是吉尔的8倍（两人的预期概率分别为12.5%和1.56%）。这个结果与锤子无关，也与因果联系无关，这仅仅是一个数学上的事实：一次拿4个弹球与一次拿7个相比，出现极端结果的概率更大。

现在，将美国人口想象成一个巨大的瓮中的弹球。有些球上标有 KC（即 Kidney Cancer

的简称)字样,表示肾癌。你抽取弹球样本,并依次按照所在县摆放,你会发现乡村地区的样本要比其他地区的少。如同杰克和吉尔所做的那个游戏一样,极端的结果(非常高或非常低的癌症发病率)容易出现在人口稀少的县,这个故事告诉我们的就是这些。

我们从一个令人费解的事实说起:肾癌的发病率在各县有所不同,且是有规律的,我用统计学理论对此作了解释:相比于大样本,极端的结果(高发病率和低发病率)更容易出现在小样本中。这样的解释不存在因果联系。某县的人口稀少既不会引发癌症,也不能避免癌症,只会使癌症的发病率比人口稠密地方的发病率更高(或更低)。这就是真相,没什么可解释的。在某个人口稀少的县,癌症发病率并非真的比正常情况更低或更高,只是这个县正好在某个特殊的年份赶上了抽样调查罢了。如果我们在第二年重复这样的分析,也能预测到在小样本中出现极端结果的一般模式,但在前一年癌症发病率高的县,这一年发病率并不一定高。如果是这样的话,则人口稠密或稀少的因素就无法对发病率作出解释了:这些表面因素就是科学家眼中所谓的假象,即观察结果完全依赖于调查方法的某一方面,在这个案例中,则依赖于样本大小。

我刚才说的例子也许会令你惊讶,但这并不是真相初次大白于天下。你早就知道应该更相信大样本,并且即使是对统计学一无所知的人也听说过大数法则。但是“知道”并非是非抉择问题,你可能会发现下列陈述放在自己身上很合适:

当你阅读这个关于流行病学的例子时,并没有立刻注意到“人口稀少”这一特点与此次调查有何关联。

对于采用4个样本还是7个样本所产生的不同结果,你至少会感到有一点惊讶。

即使是现在,想要确定下面两个陈述句所说的完全是一回事,你也要费些脑力:

(1) 大样本比小样本更精确。

(2) 小样本比大样本产生极端结果的概率大。

第一个表述清晰地陈述了一个事实,但直到感受到第二个表述传达给你的意思,你才意识到自己并没有真正理解第一个表述的意思。

上述内容概括起来就是:没错,你知道大样本的结果更精确,但你现在可能才意识到你并不清楚为什么它们更精确。不仅你一人如此,阿莫斯与我在一起进行的第一个研究表明,

即使是经验丰富的研究人员对样本效应也缺乏直觉，要么就是理解不到位。

小样本的出错风险可能高达50%

没有接受过统计学方面训练的人是出色的“直觉性统计学家”。我与阿莫斯在20世纪70年代早期的合作便始于对这个观点的讨论。他对我（在大学）的研究班及我本人讲过，密歇根大学的一些研究人员对直觉性统计抱有乐观态度。我个人对那个观点有种强烈的感觉：那段时间我发现自己并不是一个出色的直觉性统计学家，但是我也不相信别人会比我好多少。

对于一个研究型心理学家来说，样本变差没有什么特别的。它是个烦人且损失又大的麻烦事，会把每项实验都变成一场赌博。试想我希望证明6岁女孩的平均词汇量比同龄男孩的词汇量更丰富的假设。这个假设从整体来说是成立的，女孩的平均词汇量确实要比男孩的丰富一些。然而，尽管男孩与女孩差别很大，但你很可能会抽取到男女相差不太明显的样本，甚至会抽到一个男孩比女孩词汇测试成绩还要好的样本。如果你是那个研究者，这个结果对于你来说代价就太高了，因为它浪费了你的时间和精力，却无法证实一个实际正确的假设。使用一个足够大的样本是降低这种风险的唯一方法。选择小样本的研究者只能看自己是不是能选对合适的样本了。

想要对样本错误的风险作出评估，只需通过一个相当简单的步骤就可以实现。然而按照惯例来看，心理学家并不是通过计算来选定样本大小的。他们听从自己的判断，但这些判断往往是错的。在与阿莫斯发生意见分歧不久之前，我读过一篇文章，文章通过生动的观察结果展示了研究人员所犯的错误（他们现在仍在犯这种错误）。该文作者指出心理学家选择的样本通常都很小，致使他们有50%的风险不能够证实其正确的假设，而任何研究人员都不会在头脑清醒的情况下接受这种风险。对此有一个貌似正确的解释，即心理学家对于样本大小的决定反映了他们普遍存在的一个直觉性错误观念，即对于样本变差范围的错误看法。

这篇文章令我十分震惊，因为我在自己的研究中碰到了一些问题，却在这篇文章中找到了相关解释。与大多数研究型心理学家一样，我也墨守成规地选择了一些过小的样本，因此得到的实验结果毫无意义。现在，我知道了原因：那些奇怪的结果实际上就是我这种研究方法的典型产物。我的错误特别令人尴尬，因为我教过统计学，也知道该怎样计算样本的大小，以便将风险降至可以接受的程度。但是，我从未通过计算来确定样本大小。和我的同事一样，我被传统所禁锢，相信自己设计实验的直觉，也从未认真考虑过样本选择会带来的那些风险。阿莫斯来参加研讨会时，我已经意识到自己的直觉是错误的。在研讨会中，我们很快达成共识——密歇根的那些乐观派是错误的。

我与阿莫斯开始调查一个问题：只有我自己这么愚蠢还是我只是众多愚蠢的人之一，我们通过一项测试来证实这个问题，测试对象为一些数学家，想看看这些人是否也会犯类似的错误。我们设计了一份调查问卷，其中描述了真实的研究情境，包括一些成功实验的复制。问卷要求研究人员选择样品大小，对其决定可能带来的失败风险进行评估，并为那些正在设计自己实验的研究生提供建议。在“数学心理学协会”的一次会议上，阿莫斯收集了一组资深受试者（包括两本经济学著作的作者）的反应。结果很明显：我并不是唯一一个愚蠢的人。大多数受试者都会犯和我一样的错误。显然，即使是专家，在选择样品大小时也无法充分集中注意力。

我和阿莫斯将我们合写的第一篇文章命名为“对小数定律的盲信”。我们半开玩笑地解释道，“对于随意取样的直觉似乎符合小数定律，由此可以断言大数法则对于小数定律同样适用”。在文章中，我们还收录了一个措辞有力的建议，即研究人员认为他们“对于统计直觉应抱有一些怀疑，只要条件允许，都应采用计算方法来确定样本规模，而不是依靠直觉印象作决定”。

### 信任多于质疑的普遍性偏见

在一次面向300名老年人的电话民意调查中，有60%的人支持总统。

如果你只能用三个词来总结这句话，该怎么说呢？几乎可以肯定的是，你会说“老年人/支持/总统”。这些词概括了这句话的要点。这次民意调查被省略掉的细节——媒介为电话，样本为300人——本身意义不大，它们提供的背景信息并不怎么引人注意。即使样本数量变了，你的结论也不会发生变化。当然，一个完全荒谬的数字倒可能会引起你的注意。（例如一项对6名或6亿名老年选民的电话民意调查……）除非你是专业人员，否则不管样本是150还是3000，你都不会有什么不同的反应。这就是“人们对样本大小没有足够的敏感性”这一表述的意义。

这项民意调查包含了两方面的信息：新闻本身和新闻的来源。当然，你关注更多的是新闻本身，而不是其结果的可信度。但当可信度明显很低时，新闻所包含的信息也就不足为信了。如果得知“某党派小组操纵一项错误且带有偏见的民意调查，使结果显示老年人支持总统……”你当然会排斥这项调查的结果，不会相信这条新闻，这项由某党派进行的民意调查以及其错误结果不但没有令你信服，反而会成为另一条关于政治骗局的新闻。在这样清晰的案例中，你可以选择不相信其中的信息。但是你能把“我在《纽约时报》读到……”和“我在办公室闲聊中听到……”这两种说法完全区分开来吗？你的系统1能够区分出信息的可信度

吗？眼见即为事实的原则表明：不能。

如前所述，系统1并不善于质疑。它抑制了不明确的信息，不由自主地将信息处理得尽可能连贯。除非该信息被立刻否定，不然，它引发的联想就会扩散开，仿佛这条信息就是千真万确的。系统2能够提出质疑，因为它可以同时包含不相容的多种可能性。然而，保持这种质疑会比不知不觉相信其真实性更加困难。小数定律是普遍性偏见的一种表现，即对事物的信任多于质疑。类似这样的偏见在下面的章节中还会出现。

相信小样本能反映调查对象的整体情况，这一强烈偏见也是一个较大问题的一部分。这个问题就是，我们常夸大所见事物的相容性和连贯性。许多研究人员过于相信通过有限的几次观察得出的结果，这一现象与光环效应紧密相连。我们常常会觉得自己对某个人很熟悉也很了解，但事实上，我们对他却知之甚少。系统1在了解事实之前就根据零散的证据拼凑了一个饱满的形象。如果相信小数定律，急于下结论的机制就会运作起来。通常情况下，它会建构一个言之成理的说法使你相信自己的直觉判断。

对随机事件作出因果解释必然是错的

联想机制会搜寻原因。在统计规则方面，我们面对的困难是这些规则要求使用不同的方法处理问题。依据统计学观点，我们不应关注当前事件的成因，而应当关注其未来走向。这件事的发生并没有什么特殊原因，一切只是机缘而已。

因为偏好进行因果思考，我们在估测真实的随机事件的随机性时就会犯严重的错误。以在某家医院依次出生的4个婴儿的性别为例，男女出生次序明显是随机的。每个婴儿的出生是各自独立的。在前几个小时内出生的男婴女婴数量并不会影响到下一个出生婴儿的性别。现在，请考虑一下可能的序列：

男男男女女女

男男男男男男

男女男男女男

出现这些序列的可能性是一样的吗？人们的第一反应都是“肯定不一样啊”。但是，这样的反应是错误的。因为每个婴儿的出生都是独立的事，并且生男生女的概率也几乎相等，6个婴儿任何一种可能的性别顺序都与别的顺序概率相等。即使是现在，你仍然认为这个结论

是正确的，但它实际上是反直觉的，因为只有第三种顺序是随机的。如我们所料，“男女男男女女男”比其他两种顺序更有可能发生。我们追求模式，相信所处的是一个各方面都相互联系的世界。在这个世界里，规律（例如6个女婴的顺序）并不只是偶然发生的，它还是机械的因果联系或是人的意志的结果。我们并不期待在一个随机的过程中找到规律。但当探寻到一个可能的规则时，我们会抛开这个过程是真正随机的想法。随机过程会产生许多序列，以使人们相信这个过程完全是不随机的。如此你就可以看出来为什么假设的因果关系有进步发展的优势。它是我们从先辈那里继承的一般警觉性的一部分。我们会习惯性地搜寻环境变化的可能性。狮子可能随时都会出现在平原上，但注意到狮子出现频率的明显增长并采取行动则会安全许多，即使这种增长只是由于随机过程的波动而发生的。

对于随机性的广泛误解有时会带来重大影响。在我和阿莫斯合作的一篇代表性文章中，我们引用了统计学家威廉·费勒（William Feller）的阐述，他说，人们很容易在根本没有模式的情况下创建模式。“二战”期间，火箭弹在伦敦密集地轰炸。人们普遍相信爆炸不可能是随机的，因为地图显示，爆炸点在各地的分布有明显区别。一些人猜测没有被炸的地点住有德国的间谍。一份严谨的统计分析显示，爆炸点的分布是随机程序的一个典型代表，同样也是令人产生它并不是随机的这一强烈印象的典型代表。费勒评论道：“在没受过专业训练的人看来，这一连串轰炸行动就好像具有某种规律或趋势了。”

很快，我得到一次机会可以把我从费勒那儿学到的知识派上用场。1973年爆发的赎罪日战争中，我作出的唯一一项重大贡献就是建议以色列空军的高级官员停止一项调查。一开始，由于埃及地对空导弹表现出色，空战对于以色列来说很不利。以色列方面人员伤亡惨重，其人员分布也不均衡。有人告诉我说，有两支来自同一基地的空军中队，其中一支被击落了两架飞机，而另一支一架也没有被击落。为了弄明白那支不幸的空军中队到底做错了什么，相关人员对此展开了调查。我们没有理由认为其中一支空军中队比另一支更有效率，也并未发现他们在操作上有何不同。当然，飞行员的生活在很多方面会有所不同，据我回忆，其差异包括他们在任务之间回家的次数以及报告任务的执行情况等。我当时给出的建议是，司令部应该明白之所以出现不同结果仅仅只是因为他们运气不同而已，应该停止对飞行员的调查。我推断这次事件很可能是由于运气不佳，对不明显的原因进行随机调查必定是劳而无功的。与此同时，空军中队不断有人员损失，没有必要再给他们增加额外的负担，让他们觉得那些去世的伙伴做错了什么。

几年以后，阿莫斯和他的同学汤姆·季洛维奇（Tom Gilovich）、罗伯特·瓦隆（Robert Vallone）对篮球随机性的错误直觉所作的研究引起了轰动。运动员有时投篮顺手的“事实”普遍被运动员、教练和球迷们所接受。这样的推断是顺理成章的：如果一个运动员连续进了三四个球，你就会不由自主作出判断：这个运动员正处于“投篮顺手”的状态，得分率暂时增

加。两队队员都持这种判断——队员也更爱将球传给打得顺手的人，对方球队则会用两位防守球员防卫这位进攻球员。然而，对上千个投篮动作的分析结果却十分令人失望：在职业篮球比赛中，无论球是从球场上投出还是从罚球线投出的，根本没有“投篮顺手”这回事。当然，一些球员会比其他球员投篮更准，但进球与投篮未中都只是随机的。“投篮顺手”完全只是旁人所见，而且他们太快作出评判了，以至于感知不到随机事件中的顺序和因果关系。“投篮顺手”是一个影响深远的认知错觉。

公众对于这项研究的反应也是这项研究的一部分。这个发现令人惊讶，很快就受到了媒体关注，而大家普遍的反应都是不相信。当著名的波士顿凯尔特人队教练瑞德·奥尔巴赫(Red Auerbach)听说了季洛维奇及其发现(研究)时，他回应道：“这人是谁呀？他做了个实验是吧？不过我对他的观点不敢苟同。”在随机性中发现规律的想法往往不可逆转，肯定比某个人作了一项研究更有说服力。

对规律的错觉在方方面面影响着篮球赛场。你要用多少年的观察才能肯定一个投资顾问是有真才实学的？一个执行总裁实现多少次成功的兼并，董事会才能确定他对这项工作有着非凡的才能？简单说来，如果你听从自己的直觉，就常常会因为把随机事件看做是有规律的事件而犯错。我们都非常愿意相信生活中大多数事情并不是随机的。

在本章开头，我引用了一个美国癌症发病率的例子。这个例子本来是有意写给统计学老师看的，我是从前文中提到的两位统计学家霍华德·维纳和哈里斯·泽维林所写的一篇有趣的文章中看到这个例子的。他们写这篇文章得到了盖茨基金会17亿美元的赞助，用以调查那些最成功的院校有哪些特点。许多研究人员在那些名声显赫的院校中作调查，希望发现这些学校的与众不同之处，从而寻求成功教学的秘密。这项研究的结论之一是，这些著名院校规模普遍较小。例如在宾夕法尼亚州对1662所院校的调查中，排名前50的院校里有6所规模都较小，是(普通院校)的3倍多。这个数据使得盖茨基金会积极投入大量资金建立小规模院校，有时会采取将大的院校拆分成小的院校的方法。另外，其他著名的机构中至少有一半也采取了同样的做法，例如爱林伯格基金会和皮尤慈善机构也采取了同样的做法，美国教育部还启动了“小型学习社区计划”。

你可能会觉得上述做法很有道理。我们很容易从因果关系角度去解释小规模院校为什么可以提供优质的教学。我们认为，比起大规模院校，小规模院校可以给予学生更多的关注及鼓励，因此能培养出成就卓越的学者。但不幸的是，这样的因果分析是无意义的，因为得到的结论都是错误的。如果那些向盖茨基金会提交报告的统计学家们调查过最差学校的特点，他们会发现那些较差的学校也比水平一般的学校的规模小一些。事实上，规模小的学校办学水平并不一定更高，他们只不过更懂得变通而已。维纳和泽维林说，如果真有什么区别的



话，那就是大规模院校可以提供多种多样的课程，所以容易收到好的效果，特别是学生能拿到高分。

多亏了几年来认知心理学的发展，我们现在才能清楚地知道阿莫斯和我所瞥见的不过是冰山一角：小数定律包含在大脑工作的两个重要部分中。

？夸大对小样本的信任只是众多错觉中的一种——比起信息的可靠度，我们会更加注重信息本身的内容，其结果就是我们会将周围的世界变得比数据所能证明的更加简单和统一。在想象的世界中过早下结论比在现实中更有把握。

统计学家的很多观察研究都可归结到因果关系的解释上，但他们却不承认是这样的。许多事实其实只是巧合，包括事件的采样。对偶发事件作出因果关系的解释必然是错误的。

#### 示例—小数定律

“没错，自从这个新执行总裁接手后，电影制片厂已经制作了3部优秀电影。但是现在说他老练还为时过早。”

“统计学家能够估算出这个交易新手表现出来的强劲势头是否是偶然的，所以在咨询统计学家之前，我不会相信他就是个天才。”

“观察的样本太小，所以不能作出任何推断。我们不要遵从小数定律。”

“在我们有足够大的样本之前，我计划对实验结果暂时保密。否则，我们会有过早下结论的风险。”

## 思考，快与慢

### 第十一章 锚定效应在生活中随处可见

阿莫斯和我曾临时赶制过一个幸运轮盘，上面刻有0到100的标记，但我们对它进行了改装，使指针只能停在10或65的位置上。我们从俄勒冈大学招募了一些学生做这项实验。我们两人中有一个会站在一个小组前面，转动这个幸运轮盘，并让小组成员记下转盘停下时指向的数字，当然了，这些数字只可能是10或65。之后，我们问了他们两个问题：

你刚才写下的关于非洲国家占联合国（所有成员国）的百分比的数字大还是小？

你认为联合国中非洲国家所占的比例最有可能是多少？

幸运轮盘的转动根本不可能为任何事情提供有用信息，即使没有经过改装的轮盘也不可能，实验的受试者应该忽略它的影响，但他们没有做到这一点。那些看到10和65的人的平均估值分别为25%和45%。

我们研究的现象在日常生活中很普遍也很重要，因此你应该记住它的名字：锚定效应。人们在对某一未知量的特殊价值进行评估之前，总会事先对这个量进行一番考量，此时锚定效应就会发生。这一效应是实验心理学中最可靠也最稳健的结果，即估测结果和人们思考的结果很相近，就好比沉入海底的锚一样。如果有人问你甘地死时年龄是否大于114岁，你在估测他的死亡年龄时会比锚定问题是35岁（死亡）时更高。你在考量买房要花多少钱时，也会受到要价的影响。同样的房子，如果市场价格高，它就显得比市场价格低时更有价值，即使你决心抵制价格的影响也没有用。此类事例不一而足，锚定效应的事例不胜枚举。一旦你要考虑某个数字是否会成为一个估测问题的可能答案，这个数字就会产生锚定效应。

我们不是最先观测锚定效应的人，但我们的实验是第一个揭示其荒谬性的实证研究：人们的判断明显受到没有任何信息价值的数字的影响。若想解释轮盘的锚定效应是合理的，这似乎不太可能。阿莫斯和我在《科学》杂志上发表的论文中提到了我们的实验，我们在这份杂志上发表的所有研究成果中，这是最著名的一个。

但唯一的问题是：阿莫斯和我对锚定效应体系中的心理学现象的认识并不完全一致。他支持一种解释，我则偏向另外一种，而且我们从来没有找到解决这个分歧的方法。几十年后，无数研究者通过努力终于解决了这个问题。现在问题明确了：阿莫斯和我都是正确的，是两个不同的机制共同作用产生了锚定效应，即一个系统对应一个机制。锚定有一种形式，即它是在进行刻意调整时发生的，这也是系统2的一种运行模式。还有一种是由启发效应产生的锚定，是系统1的一种自主显示模式。

对锚定值的调整常常是不足的

阿莫斯喜欢将调整——锚定的启发式想法作为估测不确定值的策略：从锚定的数字开始，然后估测它是过高还是过低，接着让大脑从锚定数值上“转移”并逐渐调整你的估值。调整通常会过早结束，因为当人们不再确定他们是否应该继续移动时就会停止移动。在我和阿莫斯

意见出现分歧的几十年后，也就是阿莫斯去世的几年后，两位心理学家分别提出了有关这个（调整—锚定）过程令人信服的证据，这两位心理学家在其职业生涯早期曾和阿莫斯一起共事过，他们就是埃尔德·沙菲尔（Eldar Shafir）和汤姆·季洛维奇，参加研究的还包括他们的学生——阿莫斯的“徒孙”们！

为了验证这个想法，请拿出一张纸，从纸的底端由下而上画一条2.5英寸长的线，徒手画，不用尺子。现在请拿出另一张纸，从纸的顶端开始由上而下画一条线，在距顶端2.5英寸处停笔。比较这两条线。你很可能会估测第一条2.5英寸长的线比第二条短，因为你并不确切地知道这样的一条线究竟该是什么样的。这种估测有很多不确定的因素。你若从纸的下端开始画起，就会停在下面的不确定区域；如果你从纸的上端开始画起，就会停在上面的不确定区域。罗宾·勒伯夫（Robyn LeBoeuf）和沙菲尔在日常经历中发现很多关于这种机制的例子。为什么你在开车下了高速公路驶入城市街道时还会开那么快——如果是边聊天边开车，速度尤其容易快。调整不足就是对这一现象的最好解释。孩子在房间里喜欢把音乐开得很大声，父母却很恼火，两代人关系紧张，其原因也是调整不足。勒伯夫和沙菲尔注意到“家长要求孩子应将音乐调到‘适当’的音量，但听着特别大声音乐的孩子即使想这样做也难，还会认为家长忽视了他们真心作出的让步，这是因为他们无法从充分的高锚定值中调整过来”。这里提到的司机和孩子都有意地（将数值）向下调整，但二者都没有（将这个值）调整到位。

请看下面的问题：

乔治·华盛顿是何时当选总统的？

珠穆朗玛峰峰顶上的沸点是多少？

在思考这些问题时，最先发生的事就是你的脑海中出现了锚定数字，你知道这样是不对的，而且还知道正确答案的大致方向。你马上就知道乔治·华盛顿是1776年后成为总统的，你也知道珠穆朗玛峰峰顶上的沸水温度比100摄氏度低。你不得不去寻找从这个锚定值上移开的论据，将其调整到合适的数值。就像前文提到的那个线段的例子那样，在那个不确定区域的边缘，你不确定是否应该继续进行时就会停止。

尼克·艾普雷（Nick Epley）和汤姆·季洛维奇发现，调整就是刻意去寻找离开锚定数字的理由：按照要求一听到锚定值就摇头的人仿佛对此有些排斥，他们会离锚定值更远。而点头则会增强他们对锚定值的概念。艾普雷和季洛维奇也证实说调整是一项需要付出努力的活动。人们在自己的大脑资源耗尽时调整较少（离锚定较近），因为他们的记忆中存储着一些

数字或是因为他们有些醉了。调整不足是软弱或懒惰的系统2的一种失误。

我们现在知道阿莫斯至少对一些锚定效应的例子的理解是正确的，这些例子中包含了系统2为离开锚定值进行的具体调整。

暗示就是一种锚定效应

当阿莫斯和我就锚定效应进行争论时，我同意（他说的）有时会出现调整现象，但其实心里很不服气。调整是一个审慎且有意识的行为，但锚定效应的大多数例子中并没有相应的主观经验。请看下面两个问题：

甘地去世时比144岁大还是小？

甘地去世时多少岁？

你是否会把144岁向下调整来得出自己的评估？也许不会，但这个大得离奇的数字仍然会影响你的估测。我的直觉告诉我，锚定就是一种暗示。如果有人只是提起某件事，而他的话却促使我们去看、去听或是去感受这件事，此时的情形就可以用锚定来形容。例如，“你的左腿现在是否微微麻木了”这个问题常会使相当多的人回答说他们的左腿确实感到有些异样。

阿莫斯对直觉的研究比我更谨慎，他一针见血地指出，关注暗示对我们理解锚定效应没有什么帮助，因为我们不知道如何解释暗示作用。我不得不同意他是对的，但我不同意调整不足是锚定效应的唯一原因。我们为理解锚定效应做过很多实验，但是都以失败告终，最终我们放弃了进一步研究这一课题的想法。

现在，那个打败我们的难题得到解决了，因为暗示的概念已经明确了：暗示即一种启动效应，它会有选择地找出相应的证据。你完全不会相信甘地活到了144岁，但你的联想机制却会对一位逝去的老人产生印象。系统1理解句子的方式就是尽量相信其内容的真实性，它对相应想法的选择性激活会产生一系列的系统性误差，这些误差会使我们更容易受骗，更加坚定地相信自己的想法。我们现在明白为什么阿莫斯和我从前没有意识到锚定效应有两种类型：研究手法和理论观念，我们研究这个问题时很需要这两种类型但它们却没有出现。后来，其他人极大地发展了这两种类型。在很多情况下，我们都需要一个类似暗示的过程发生作用：系统1试图建立一个将锚定数字视为真实数值的世界。这是我在本书第一部分描述的联想一致性的一种表现。

德国心理学家托马斯·穆斯魏勒（Tomas Mussweiler）和弗里茨·斯特拉克（Fritz Strack）对锚定效应中联想一致性的作用所作的实证研究最令人信服。在一项实验中，他们问了一个关于温度的锚定问题，“德国每年的平均温度是高于20摄氏度还是低于20摄氏度”，或者“德国每年的平均温度是高于还是低于5摄氏度”？

研究人员让所有受试者快速扫过一些单词，然后让他们去识别这些词。研究人员发现，受试者看到20摄氏度后更容易识别和夏天相关的词（比如“太阳”和“沙滩”），看到5摄氏度后则能更加轻松地识别出关于冬天的词汇（比如“冰冻”和“滑雪”）。对相应记忆的选择性激发解释了锚定效应：大小不同的数字能激发起记忆中不同的观念体系，而这些带有偏见的观念则成为（受试者）估测年度平均温度的依据，据此作出的估测值也因此带有一定偏见。在另一个基于同一目的的简单实验中，研究人员让受试者估测德国汽车的平均价格。高锚定值的受试者会选择性地让奢华品牌（比如奔驰、奥迪）的汽车最先出现在脑海中，而低锚定值则会使人最先想到一些销售量大的汽车品牌（比如大众）。此前我们就知道任何启发都会唤起与之相应的信息。暗示和锚定效应都通过相同的系统1的自主运行得到了解释。尽管当时并不知道如何证实这一观点，但我对锚定效应和暗示之间存在联系这一直觉最终被证实是正确的。

作为这幢房子的主人，你能接受的最低售价是多少？

很多心理学现象可以通过实验得到证实，但事实上这些现象中几乎没有哪种是可以测量的。锚定的影响是一个特例，锚定效应可以测量，测量结果证实这一效应影响超大。有人曾经问过那些参观“旧金山探索馆”的游客下面两个问题：

最高的那棵红杉树是高于1 200英尺还是低于1 200英尺？

你认为那棵最高的红杉树有多高？

这个实验中的“高锚定值”是1 200英尺。而另外一组受试者看到的第一个问题则用了一个180英尺的“低锚定值”。两个锚定值相差1 020英尺。

不出所料，（关于那棵最高的红杉树有多高）两个组给出了完全不同的平均评估：844英尺和282英尺。二者的差距有562英尺。锚定指数就是两个不同答案的比率 $562/1\ 020$ ：55%。对于那些在评估过程中过度运用锚定的人而言，锚定效应的测量结果将是100%，对于那些能够忽略锚定的人而言，锚定效应的测量结果则是零。这项实验中观察到的55%的锚定指数

非常典型，在其他许多问题中也观察到了相似的锚定指数。

锚定效应在实验室中司空见惯，在现实世界中其影响同样毫不逊色。几年前的一个实验中，实验人员让一些房地产经纪人对一幢待售房子的价值进行评估。这些经纪人亲自去看了这所房子，还仔细研究了一本小册子，里面包括这幢房子的全面信息和售价。其中一半经纪人评估的售价比标价高很多，另一半评估的售价则比标价低很多。每位经纪人都给出了自己认为合理的售价，同时（假定自己是这幢房子的房主）她们还给出了自己能够接受的最低售价。实验人员随后问这些经纪人哪些因素影响了她们的决策。出乎意料的是，售价并非影响因素之一。这些经纪人十分得意，因为她们没有被这个因素所左右。她们坚称标价没有影响到她们对这一问题的回应，但实际上她们错了：锚定指数是41%。事实上，专业人士和对房地产一无所知的商学院学生一样，都受到了锚定效应的影响。后者的锚定指数是48%。两者的唯一区别是，学生承认他们受到了锚定效应的影响，而专业人士则否认这一点。

在人们对钱的问题所作的决定中同样体现出了强大的锚定效应，人们选择对一项事业投入多少时就会受锚定效应的影响。为了证实这一影响，我们告诉探索馆研究中的受试者，太平洋中的油管引起了环境污染，问他们是否愿意每年都捐钱“使5万只近太平洋海岸的海鸟不致受小面积海上溢油的影响，直到找到防止泄漏的方法或要求油罐所有者支付运行费用为止”。这个问题需要强度相匹配：实际上，我们是在要求调查对象预估应该捐献多少钱才能与自己海鸟困境的感受程度匹配。对其中一些游客，实验人员并没有直截了当地问她们愿意捐多少钱，而是先问了一个锚定问题，比如“你是否愿意花5美元来……”

若没有提到锚定问题，这些对环境都很敏感的探索馆参观者会说他们愿意拿出的钱平均为64美元。当锚定金额只有5美元时，平均捐款则是20美元。当锚定金额达到400美元时，人们的捐款平均数就达到了143美元。

高锚定金额和低锚定金额之间相差123美元。锚定指数高于30%，表明增长100美元的初始要价就能带来平均值为30美元的回报。

许多估测和捐款意愿的研究都曾体现出相似甚至更大的锚定效应。例如，有人曾问过那些在法国马赛重度污染地区居住的居民，如果能住在一个污染程度较低的地方，他们能接受生活开销提高多少。锚定指数在该研究中超过了50%。在网上购物过程中也很容易观察到锚定效应，网上相同的产品经常标出不同的“立购”价。“估测”在艺术品拍卖行业中是影响第一次竞拍的锚定价格。

锚定效应在某些情形下看起来也是合理的。毕竟，那些被问到难题的人肯定会去抓住这

根救命稻草的，况且这个锚定值也是根貌似合理的稻草。如果你对加利福尼亚的树所知无几，却又被问到红杉树是否高于1 200英尺，此时你可能就会认为这个数字与真实数字相差不远。因为是那些知道这种树真实高度的人想出的这个问题，所以这个锚定值也许是个有价值的提示。但是，锚定效应研究有一个重要发现，即锚定值显然是任意的，它也许和可能有信息价值的锚定值一样有效。当我们用轮盘来估测联合国中非洲国家所占的比例时，锚定指数是44%，还算是处在看似正确且能作为提示的锚定效应的影响范围内。有些实验已观察到大小相似的锚定效应，在这些实验中，受试者社保号的最后几个数字被用做锚定值（比如为估测他们城市中医生的数量）。结果很明确：锚定值没有影响，因为人们认为这些数字没有什么信息价值。

随机锚定的影响以一些令人不安的方式出现在生活中。一些任职经历平均为15年的德国法官先是读了一份案例，讲到一个妇女在商店顺手牵羊被捉住的案例，然后他们开始掷一副骰子，这副骰子被人提前做过手脚。因此，每次掷骰子的结果不是3就是9。骰子一停，实验人员就问这些法官是否会将那位妇女送进监狱，且其服刑的时间应该比骰子上的数大还是小。最后，实验人员问这些法官，他们给这个行窃的妇女判定的服刑期具体是多少。平均来看，那些掷了9的法官说他们会关她8个月，而掷了3的法官说他们会关她5个月，锚定指数是50%。

锚定效应何时适用，何时不适用？

到现在为止，你已经相信锚定效应无处不在，有时是由于启发效应，有时则是因为调整不足。产生锚定效应的心理机制使我们比自己预想的更容易受影响。当然，有些人愿意且能够利用我们的轻信。

例如，锚定效应解释了为什么限量购买是一种有效的营销策略。几年前，在艾奥瓦州的苏城的超市里，购物者遇到了坎贝尔汤罐头在作促销的情形，降价10%。有那么几天货架上写着“每人限购12罐”，而在其他几天里则写着“不限量”。购物者在限购时平均会购买7罐，是不限购时购买量的2倍。锚定效应不是唯一的解释，配给也显示货物很快就下架了，购物者应该对货物储存量感到有些紧张。但我们也知道12罐的可购买量会成为一种锚定，即使这一数字是通过轮盘产生的也不例外。

同样的策略在商讨购房价上也适用。在许多其他的活动中也是这样，在协商中率先出击的一方往往会占有优势，例如，当价钱是买方和卖方唯一要协商的事时就会出现锚定效应。第一次在集市上讨价还价也是这样，先发锚定有着重大的影响。我在教学生谈判时，给他们的建议是如果你认为是对方作出了无礼的提议，你就不应该提出同样无礼的提议，因为两者

之间有距离的话会使此后的商谈难以进行。你应该大吵大闹，夺门而出，或者威胁对方说自己也会这样做，要让对方明白以这个数字为基准的话，谈判将难以继续。

心理学家亚当·加林斯基（Adam Galinsky）和托马斯·穆斯魏勒提出了更好的方法来抵制商谈中锚定效应的影响。他们告诉谈判者，在商谈中要集中注意力搜寻大脑记忆来抵制锚定效应。激活系统2的做法会很奏效。例如，在第二个提议人将其注意力集中在对方能接受的最低值或对方无法接受的费用上时，锚定效应就会削弱或消除。大体来讲，有意地“为对方着想”的策略也许是抵制锚定效应的好方法，因为它否定了能产生这些效应的带有偏见的想法。

最后，试试看你能不能弄清楚锚定效应对公共政策问题的影响，即人身伤害案件的损害程度的裁定。这类案件的判决有时是很严厉的。类似医院和化工企业等单位常常是这类诉讼案件的被告，这些单位曾经游说各方为此类判决设置了一个上限。读本章之前你可能想过，给这些判决设置上限绝对是有利于潜在被告的，但现在你就不会这么确信了，如果上限是100万美元，其结果会怎样？这条规则会消除所有的严厉判决，但锚定也会阻止法官将许多较轻罪行的量刑判得更轻。这对违法者和大公司都有好处，比给小公司带来的好处多。

随机锚定效应还会使我们更加了解系统1和系统2之间的关系。人们一直利用判断与选择的案例来研究锚定效应，而判断与选择最终总是由系统2完成的。但是，系统2对从记忆中提取的数据进行加工，并由系统1进行自主的、无意识的运行，因此很容易受锚定效应的影响，而这种影响会使某些信息更容易让人回想起来。此外，系统2对这种影响一无所知，也无法控制。看到随机或荒谬的锚定值（比如甘地死时144岁）的那些受试者会自信地说，这个明显无用的信息并没有对他们的估测行为产生影响。事实上，他们错了。

在讨论小数定律时我们发现，如果一则消息没有马上被视为谎言，那么不管其可靠性如何，它都会对联想系统产生同样的影响。这个消息的重点是故事，随便根据什么信息编造的都无所谓，即使这则消息的信息量很少，质量很差劲也无所谓，因为眼见即为事实。当你读到一个拯救受伤登山客的人的英勇故事时，这个故事对你的联想记忆产生的影响和一篇新闻报道或电影简介大体差不多。锚定效应是由这个联想激发引起的。这个故事是否真实、是否可信一点都不重要。随机锚定的强大影响是锚定效应的极端例子，因为随机锚定显然没有提供什么信息。

我在前文中讨论了启发效应纷繁复杂的表现类型，在启发效应下，你的思想和行为也许会被完全不曾留意的刺激所影响，甚至会被你完全没有意识到的刺激所影响。启发实验的主要寓意是我们的思想和行为会受当时的环境影响，且这种影响比我们了解或想象的要大。很



多人发现启发效应的结果令人难以置信，因为它们和主观经验相去甚远。另外很多人则发现其结果令人不安，因为它们威胁着我们对中介和自主性的主观感受。如果不相关的大脑屏保在你意识不到的情况下影响你对陌生人的帮助，那你到底有多自由呢？锚定效应也以相似的方式威胁着你。你总能意识到锚定，甚至会对它格外关注，但你不知道它是如何引导和限制你的思考的，因为你不能想象如果锚定改变（或不存在）你会如何思考。但是，你应该假设任何一个公开谈判时的数字都对你有锚定效应，如果概率大，你应该抵制（你的系统2）该效应。

### 示例—锚定

“我们想要收购的公司给我们看了他们的商业计划，其中包括他们希望得到的收益。我们不应该让那个数字影响到我们的思路。将其放置一边。”

“计划是为最佳情况设计的方案。当我们预计实际结果时，要避开计划的锚定效应。想想计划出现失误的各种方式也不失为执行计划的一个方式。”

“我们商谈的目标是让他们锚定在这个数字上。”

“我们要清楚一点，如果那就是他们的提案，那么商谈就此结束吧，我们不想那样开展工作。”

“被告律师提出一个微不足道的证明，证明中提到了一个荒谬的小损失，这些律师就是想让法官们拿这个损失做锚定。”

## 思考，快与慢

### 第十二章 科学地利用可得性启发法

1971~1972年，阿莫斯和我在俄勒冈州的尤金度过了我们最为高产的研究时期。我们在俄勒冈研究院做客，在我们研究的领域——判断、决策制定和直觉性预测——中该研究院后来诞生了很多未来之星。主要负责接待我们的是保罗·斯洛维克，他曾是阿莫斯在安阿伯市时的同学，也是他一辈子的朋友。保罗当时即将成为风险心理学领域中的领军人物，他独领风骚几十年，也获得了很多荣誉。保罗和他的妻子洛兹带我们感受尤金的生活，很快我们也开始效仿尤金人，常去跑步、烧烤，带孩子去看篮球赛等。我们非常努力地工作，做了很多关于启发法的实验，也写了很多文章。晚上的时候，我就写那本《注意与努力》（Attention and Effort）。那段时间，我非常忙。

我们的项目中有一项是对“可得性启发法”进行研究。我们问自己，人们在想要估计某类事件的出现频率时到底是怎么做的，这些事件包括“人在60岁之后的离婚概率”或“（是否是）危险的植物”等。我们认为这种自问就是启发法。这些问题的答案很简单：从记忆中搜寻这类问题的实例，如果搜寻过程既轻松又顺畅，这些事的发生概率就会被判断为很大。我们将可得性启发法定义为通过“实例呈现在脑中的轻松程度”来判断概率的过程。我们的系统阐述似乎已经给出了这一方法的明确定义，但可得性这一概念此后仍在不断得以精炼。我们在研究可得性的时候还没有阐发两个系统的方法，而且我们并没有费尽心思去确定启发法是可以解决问题的主观策略，还是一个自主运行的无意识行为。现在我们知道启发法其实涉及两个系统。

我们之前思考的问题是，到底需要在大脑中搜寻多少实例才算是轻松回忆，以获得某一印象。我们现在知道了答案：一个也不用。例如，请考虑下面两组字母可以组成多少个单词。

XUZONLCJM

TAPCERHOB

你几乎不必去想什么实例，一打眼就知道其中一组字母比另一组组成单词的可能性更大，可能会多10个或者更多。同样，想要清楚地了解不同国家去年出现在新闻中的相关频率（比利时、中国、法国、刚果、尼加拉瓜、罗马尼亚等）你也无须回想具体的新闻报道。

意识到自己的偏见有利于团队关系融洽

与其他判断启发法一样，可得性启发法就是用一个问题替代另一个问题：你希望估测某一范畴的大小或某一事件的（发生）频率，但你却会提到自己想到相关实例的轻松程度。问题的替代必然会产生系统性错误。你会发现启发法是如何通过一个简单的过程导致偏见的，不直接说出（事件发生的）频率，而是列举那些使你轻松想起相关实例的因素，其中的每个因素都会成为偏见的潜在来源。试举几例：

你可以很轻松地回想起引起自己注意的突出事件。好莱坞明星的离婚事件和政客的性丑闻事件格外引人注目，想到这些实例并不难。因此，你很容易夸大好莱坞离婚事件和政客性

丑闻事件的频率。

一个大事件会暂时提高此类事件的可得性。飞机失事事件会有媒体来报道，这也会暂时改变你对飞行安全的看法，接着你又看到路旁有辆汽车着火了，于是这些事故会暂时盘踞在你的脑海中，你会觉得这个世界此时充满更多难以预料的事。

亲身经历、生动的图片和鲜活的例子比发生在别人身上的事、单纯的文字或是统计数据更容易让人回想起来。一个与你相关的判决错误会逐渐削弱你对司法体系的信任度，其影响程度比你在报纸上读到类似事件的影响更深。

你可以尽可能地抵制如此之多的潜在的可得性偏见，但那样做会令你身心俱疲。你必须通过自问一些问题努力重新审视自己的印象和直觉，比如“我们是否会因为小区内最近发生了几起偷盗事件就认为青少年盗窃是个严重问题”或者“我认识的人去年没有得感冒的，我是不是就没必要打免疫针了呢”。时刻对偏见保持警惕是件累人的事——但由此便可避免一个代价高昂的错误，因此付出努力也是值得的。

有个很著名的可得性实验表明，意识到自己的偏见可以使夫妻和睦相处，而且很可能在其他的合作计划中与他人的关系也很融洽。在一项广为人知的研究中，研究人员问夫妻双方的问题是：你为保持此地整洁作了多大贡献？用百分比来表示。此外，夫妻俩还回答了如“倒垃圾”、“发起社交互动”等类似问题。那么这两位自我估测的贡献率合计能达到100%吗，是更多或者更少？不出所料，他们自我估测的贡献率合计超过了100%。一个简单的可得性偏见就可以对此作出解释：夫妻二人记自己的努力和贡献比记对方的清楚得多，而且可得性的不同导致了对频率判断的不同。偏见不一定是自私的：这对夫妻还过多地将两人的争执归因于自己，尽管这一比例比两人在那些积极正面的事件中的自评比例小得多，但也是难能可贵了。同样的偏见对常见的观察同样适用，很多合作团队成员感觉他们做的事超出了自己的分内工作，还感到其他人并不感激自己作出的贡献。

我通常对人们控制偏见的潜能不是很乐观，但这次例外。成功去除偏见的案例还是存在的，即我们可以很轻松地识别出功劳分配问题是何时出现的，尤其当几个人同时感到他们的努力没有得到足够的认同时更是如此。自己周围的那些人通常也会付出超出100%的努力工作，只要你看这一点有时就足以缓和这种（心理失衡）情形。任何情况下，每个人都该牢记这一点。你做的事情偶尔会超出自己的分内事，但你应该知道，当你有可能有这种感觉的时候，你的团队里的每个成员也都可能有同感。

可得性偏见会影响我们对自己或他人的看法

20世纪90年代早期，人们对可得性启发法的理解有了重大进展。那时，由诺伯特·施瓦茨（Norbert Schwarz）领导的一组德国心理学家提出了一个有趣的问题：人们对某件事发生频率的印象是如何受到列举实例的具体数目这一要求的影响的？设想你自己就是那个实验的受试者：

首先，列出6个你果断行事的例子。

接下来评判一下你有多果断。

假如有人要求你列12件自己表现得果断的事（大多数人都会觉得很难列出这么多件事）。你对自己果断程度的判定会有所不同吗？

施瓦茨和他的同事观察到列举事件的任务可能会通过两种不同方式加强对特点的判断：

能回想起的事例数量。

事件在脑中呈现的轻松程度。

列举12个例子的要求使得两个决定因素相互排斥。一方面，你刚想起了几件自己做得很果断的事；另一方面，你想起前三四件果断的事可能很轻松，可说出12件就要挖空心思了，回忆起来也没有那么顺畅。哪个更重要呢，提取的数量还是提取的轻松感和顺畅性？

两个因素间的较量谁是赢家一目了然：那些列举了12件事的人认为和只列举了6件事的人相比，自己不够果断。而且，列举出自己表现不够果断的12件事的那些受试者最终却认为自己非常果断！如果无法轻松地想起懦弱的事例，你可能就会说自己一点也不懦弱。自我评估是由事件呈现在脑海中的轻松度来衡量的。轻松地想起某件事的体验比想起事情的数量更重要。

同一个小组的另外一些心理学家对顺畅性的作用作了一个更为直接的实证研究，实验中所有的受试者都按要求列出了6件果断（或不果断）行事的例子，同时还保持着特定的面部表情。“微笑者”要收缩颧肌，露出浅浅的微笑；“皱眉者”要皱着眉头。如你所知，皱眉通常伴有认知紧张，且其影响是对称的：执行任务时被要求皱眉的人付出的努力更多，体验到的认知紧张也更强烈。研究者预计皱眉者在提取果断行为时会有难度，因此会估计自己缺少果

断性。结果也的确如此。

心理学家喜欢产生悖论的实验，他们怀着极大的兴趣运用了施瓦茨的发现。例如，人们：

在回忆起自己多次骑自行车的经历后，依然相信他们自己不经常骑自行车。

在被要求用更多的论据支撑自己的选择时自信心下降。

在列出多个避免某种问题的方法后，却对避免此问题显得更加不自信。

在列出某辆车的优点后，对该车变得没那么感兴趣了。

加州大学洛杉矶分校的一位教授发现了一个利用可得性偏见的巧妙方法。他让不同组的学生列出改进课程的方法，要求不同组列出不同数量的方法。正如他所预料，列出较多改进方法的学生对这一课程的评价也更高。

也许这个有悖论的实验中最有趣的发现就是悖论不一定会被发现：人们有时依照提取到的内容而非提取的轻松程度来作出判定。你真正了解行为模式的证据就是你知道如何让其发生逆转。施瓦茨和他的同事们接受了这一挑战，去探索在何种情况下会发生这种逆转。

受试者想到那些行事果断的例子的轻松程度在任务实施过程中是不断变化的。前几件事情提取时很容易，但很快就变得越来越难。当然，受试者的思考顺畅性也会慢慢下降，但在列举6~12件事中顺畅性的下降速度显然要超出受试者的预料。结果表明受试者作出了推理：如果我在回想能体现自己果断行事的例子时遇到了超乎想象的困难，那说明我根本就不是个果断行事的人。请注意，这个推理建立在受试者未曾料到的情况之上——顺畅性比预期的更糟糕。此例中受试者运用的可得性启发法被叫做“无法解释的不可得性”启发法才更合适吧。

施瓦茨和他的同事们论证道，他们为受试者解释其在回想相关经历的流畅性（发生变化的问题），由此可以干扰启发法。他们告诉受试者在回想事件时他们会听到背景音乐，而且音乐会对他们完成记忆任务产生影响。实验人员告诉一些受试者音乐能帮助他们回想起相关例子，却告知其他受试者听音乐其回忆会不那么顺畅。不出所料，那些对顺畅性有所认识的受试者没有将音乐看成是一种启发法；而那些得知音乐会使回忆更困难的受试者无论完成提取6件还是12件的任务，对自己果断程度的估测没什么两样。其他的主要实验也得到了同样的结论：若通过展示曲线或直线的文本框，或屏幕的背景颜色，又或其他与实验预期不相关

的因素对体验到的顺畅性作出虚假解释的话，判断便不再受提取轻松程度的影响。

诚如我所述，用可得性来判断的过程包括了一个复杂的推理链。受试者会体验到，他们在提取事件时，顺畅性会逐渐降低。他们显然对顺畅性的下降率事先有过预计，但那些预计是不准确的：想起新事例的困难增速远远超出了他们的预计。正是这种超出预期的低顺畅性使那些被要求列举12个事例的人将自己描述成优柔寡断之人。了解个中缘由，低顺畅性也就不会再影响判断了。这个过程看似由一套复杂的推理组成。自主的系统1能胜任这项任务吗？

回答时其实根本就不需要复杂的推理。在系统1的基本特征中，其中一点就是这一系统具有设定预期的能力，当现实与预期相悖时它就会感到惊讶。该系统还会提取造成惊讶情绪的可能原因——通常是在近期所经历的各种惊讶体验中找到一个可能的原因。此外，系统2在运行中会重塑系统1的预期，因此一件本该令人惊讶的事就变得正常了。假设有人事先告诉过你那个住在隔壁的3岁大的小男孩经常戴着一顶礼帽坐在小推车里，那么当你真的看到他戴着礼帽时就不会像事先并不知道（这件事）那么惊讶了。在施瓦茨的实验中，受试者事先知道背景音乐可能是影响他们回想具体事例的一个原因，因此他们对提取12件事的难度也就不再那么吃惊了，这种难度也就不大可能影响他们对自己果断行事程度的判断。

施瓦茨及其同事发现，判断涉及自身情况的人往往更有可能关注他们从记忆中提取的事件数量，对顺畅度则不大关注。他们请两组同学参加对心脏健康风险的研究，其中一半学生有心脏病家族史，他们应该比其他没有这种家族病史的人更注重这个研究。所有的学生都要回想自己日常生活中可能会影响心脏健康的3种或8种行为（按照要求，有些学生需要回想的是危险行为，其他学生需要回想的是保护性行为）。没有心脏病家族史的学生对这项任务较随意，遵循的是可得性启发法。觉得想起8件危险行为很难的学生觉得他们相对安全，而那些努力回想安全做法的学生却感觉自己处于危险中。有心脏病家族史的学生则表现出相反的模式——他们在想起很多安全做法时觉得很安全，在想起很多危险行为时感到很危险。他们还更有可能觉得自己将来的行为会受到自己对危险做法评估的影响。

由此可见，事件在脑海中呈现的轻松程度体现出系统1的启发作用，然而当系统2越来越多地参与其中时，受试者关注的就不再是提取记忆的轻松度，而是回忆起来的若干事例的内容了。各种各样的证据都指向统一结论，即那些跟着系统1走的人更容易受可得性偏见的影响，比那些警惕性更高的人受影响的程度更大。在下面这些情况中，人们都在“跟着感觉走”，提取轻松度对他们的影响要大于其回想事例内容带给他们的影响：

当他们同时忙于另一件需要付出努力的任务时。

因他们刚刚想起生命中的一个快乐片段而心情大好时。

如果他们在抑郁量表中得分很低的话。

尽管对这项任务所给话题的了解达不到专家级水准，但他们也算是对此领域了解颇多的新手了。

当他们跟着感觉走却拿了高分时。

如果他们（或别人令他们感到）很强大时。

我认为最后一个发现尤其有趣。几位作者引用一句名言来介绍自己的文章：“我没有满世界做民意调查来告诉自己怎样做才对，知道自己的感受就够了。”他们进一步表明，对直觉的依赖只是个人品行特征的一部分。他们只想提醒人们，能力可以提升我们对自己直觉的信任。

#### 示例—可得性

“因为上个月发生了两架飞机相撞事件，她现在更愿意坐火车。这真是愚蠢，风险其实并没有真正降低，这就是可得性偏见。”

“他低估了室内污染的风险，因为媒体对此报道极少。这是可得性的影响。他应该看些统计数据。”

“她最近看的间谍电影太多了，因此她看什么都觉得有阴谋。”

“这位执行总裁连续多次成功，因此失败不会轻易在她的脑海中出现。可得性偏见使得她过于自信。”