

The Psychology of Judgment and Decision Making

Scott Plous

决策与判断

[美] 斯科特·普劳斯 著 施俊琦 王星 译 彭凯平 审校



诺贝尔经济学奖超越常规地在1978年和2002年两次授予研究人类决策与判断的心理学家，他们是赫伯特·西蒙教授和丹尼尔·卡尼曼教授。本书系统地介绍了他们的发现，同时也阐述了决策与判断心理学中很多还没有被经济学家和其他社会及人文科学家所认识到的现象和研究结果。可以说，本书80%以上的内容超出了西蒙和卡尼曼工作的领域，因此可以想像，心理学家对经济学和其他相关学科的影响还有很多值得发掘的空间。难怪美联储主席格林斯潘曾经开玩笑地说“所谓的‘新经济’实际上就是心理学”。

——彭凯平

美国加州大学柏克利分校心理学教授

这本书充满了幽默感，对老的思想的新的解释，批判性思维的练习，从而把许多通常不相关的材料整合成了一本令人遐思发人深省的好书。可以说，从来还没有一个年轻作者的第一本著作能在一个重要的课题上对这么多的人产生如此大的影响。

——菲利普·津巴多

美国心理学会前任主席

如果能够有效地应用本书所描述的结果，你就会更好地避免决策偏差、误差和陷阱，而且你将会更好地理解他人所做的决策。

——斯科特·普劳斯

“本书是经典研究成果与该领域最新研究取向的完美结合……读者能亲身体验到这一基础研究在其日常生活当中的应用。”

——伊丽莎白·洛夫特斯

美国心理协会实验心理学分会前任主席

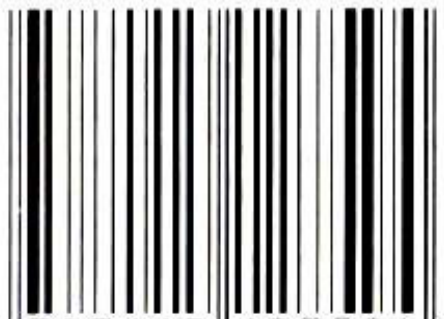
北京新曲线出版咨询有限公司

网址：www.ncc-pub.com

社会心理学精品译丛（第一批）

- | | | |
|-------------|---------|----|
| 《自我》 | 乔纳森·布朗 | 著 |
| 《决策与判断》 | 斯科特·普劳斯 | 著 |
| 《亲密关系》 | 沙伦·布雷姆 | 等著 |
| 《探索社会心理学》 | 戴维·迈尔斯 | 著 |
| 《态度改变与社会影响》 | 菲利普·津巴多 | 等著 |
| 《文化心理学读本》 | 钟年 彭凯平 | 编 |

ISBN 7-115-12759-X



9 787115 127594 >



ISBN 7-115-12759-X/F·569

定价：28.00元

人民邮电出版社网址：www.ptpress.com.cn

社会心理学精品译丛

The Psychology of Judgment and Decision Making

Scott Plous

决策与判断

[美] 斯科特·普劳斯

著

施俊琦 王星

译

彭凯平

审校



 人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目(CIP)数据

决策与判断/(美)斯科特·普劳斯(Scott Plous)著;施俊琦,王星译.

-北京:人民邮电出版社,2004.9

ISBN 7-115-12759-X/F·569

I.决… II.①斯… ②施… ③王… III.决策-判断 IV.C934

中国版本图书馆CIP数据核字(2004)第098797号

Scott Plous

The Psychology of Judgment and Decision Making

ISBN:0-07-050477-6

Copyright © 1993 by The McGraw-Hill Companies, Inc.

Original language published by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Simplified Chinese translation edition jointly published by McGraw-Hill Education (Asia) Co. and Posts & Telecom Press.

本书中文简体字翻译版由人民邮电出版社和美国麦格劳-希尔教育(亚洲)出版公司合作出版。

未经出版者书面许可,不得以任何形式复制或抄袭本书的任何部分。

本书封底贴有 McGraw-Hill 公司防伪标签,无标签者不得销售。

北京市版权局著作权合同登记号:01-2004-3331

版权所有,侵权必究。

决策与判断

◆ 著 [美] 斯科特·普劳斯
译 施俊琦 王星
审 校 彭凯平
策 划 刘 力 陆 瑜
责任编辑 陈浩莺
装帧设计 O. E. T.

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号A座
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
电话 (编辑部)010-64964059 (销售部)010-64983296
北京京科印刷有限公司印刷
新华书店经销

◆ 开本: 670×970 1/16
印张: 17
字数: 287千字 2004年9月第1版 2005年3月第2次印刷
著作权合同登记号 图字: 01-2004-3331

ISBN 7-115-12759-X/F·569

定价: 28.00元

本书如有印装质量问题,请与本社联系 电话:(010)64981059

作者简介



斯科特·普劳斯 (Scott Plous) 是韦斯利安大学的心理学教授。他以“最高荣誉”从明尼苏达大学毕业，并在斯坦福大学获博士学位。他还在斯坦福做了两年的政治心理学博士后研究，并在伊利诺伊大学做了两年的访问学者。

普劳斯获得过多种荣誉，包括国际和平与合作领域的麦克阿瑟奖，Gordon Allport Intergroup Relations 奖，IAAP 青年心理学家奖金，Slusser Peace Essay

奖。他所教授的科目是决策与判断、社会心理学、统计学，以及研究方法。他已经在各类期刊（《今日心理学》，《心理科学》，《咨询与临床心理学》，《冲突解决》，以及《应用社会心理学》）上发表了二十多篇论文。

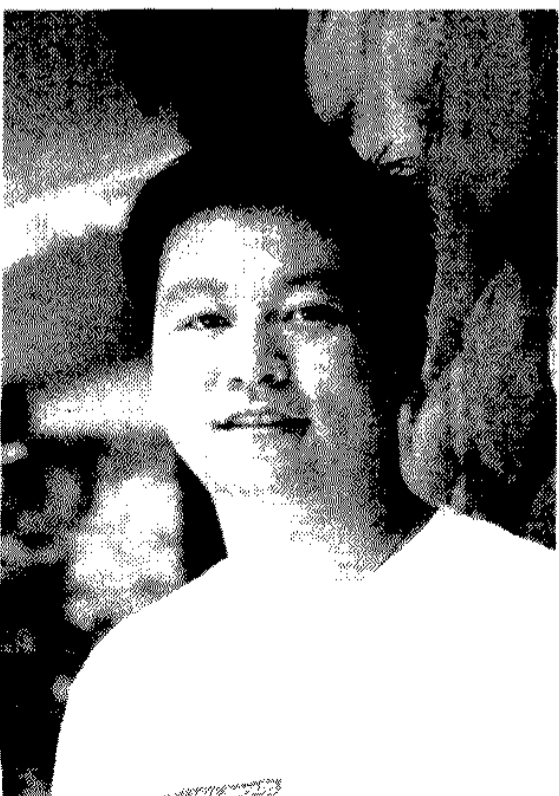
普劳斯现在的研究兴趣集中在有关动物和环境的道德问题上。1991年，他发表了针对动物权利保护者的大样本调查结果，现在，他正在编辑一个有关人类社会中动物所扮演角色的专题。除了这些学术工作外，普劳斯还是大量政治和商业项目的顾问。

主编简介



彭凯平 (Kaiping Peng), 美国加州大学柏克利分校 (U. C. Bekerley) 心理学教授, 文化与认知实验室主任。北京大学心理学系客座教授和北京大学光华管理学院国际博士导师, 国内若干知名高校兼职教授。主要研究领域为社会及文化心理学。是具有国际影响的华人心理学家。

译者简介



施俊琦, 浙江桐乡人, 北京大学心理学系讲师, 博士。主要研究领域: 消费行为学、行为经济学和行为金融学。

主编的话

社会心理学是在二次世界大战后兴起的一门社会科学学科，它研究的是人的心理和社会现象之间的关系，试图探讨人的思想、情感和行为是如何受到其他人的影响，这些影响包括实际的、想像中的和推测出来的人际作用。社会心理学家通常思考的问题有：我们如何认识他人（社会知觉），我们如何与他人打交道（社会互动）以及文化、社会、团体如何作用于我们（社会影响）等方面的内容。

众所周知，社会心理学研究向来有心理学的、社会学的和符号学的三种取向，其中心理学取向的社会心理学更强调实证的研究和对社会中个体心理的关注。本译丛以津巴多（Philip G. Zimbardo）主编的“麦格劳-希尔社会心理学系列丛书”为基础，从中遴选出精品（如《决策与判断》、《自我》、《亲密关系》、《态度改变与社会影响》、《探索社会心理学》），并在更大的范围内，补充一些在近年来有广泛影响的社会心理学新著（如汇聚了若干文化心理学经典范文的《文化心理学读本》等）。

十几年前，香港著名社会心理学家彭迈克（Michael Bond）就曾经说过：“心理学不幸是由西方人创建的，结果，西方的心理学研究了太多的变态心理和个性行为。如果心理学是由中国人创建的，那么它一定是一门强调社会心理学的基础学科。”确实，这门学科是我们中国人有可能做得比其他国家的学者更好的心理学领域，因为我们的文化几千年来就很强调人与他人、人与环境、人与社会的关系，而这些关系正好是社会心理学关注的焦点所在。可惜时至今日，中国的社会心理学并没有得到它所应有的关注。我们推出这套丛书的目的，一方面是为了让国内有志于学习、研究和应用社会心理学的各界人士较为系统地了解当代社会心理学的来龙去脉、重大发现以及最新前沿，而更重要的是，我们希望通过这套丛书，为推动中国社会心理学的发展以及提高中国社会心理学的国际影响贡献绵薄之力。

彭凯平

美国加州大学柏克利分校心理学教授

序 一

2002年11月，瑞典皇家科学院宣布2002年诺贝尔经济学奖将授予一位经济学家和一位心理学家。这位心理学家就是美国普林斯顿大学的丹尼尔·卡尼曼教授。在谈到卡尼曼教授的贡献时，诺贝尔奖评审委员会特别指出，卡尼曼教授成功地将人类决策和判断的心理学研究成果带到了经济科学的研究当中，为现代经济学理论和研究方法注入了新的见解，其贡献是卓著的，也是独特的。

人类的生活充满了判断和决策。从衣食的偏好到学校的选择，从个人职业的规划到国家政策的制定，人类的判断和决策是最常见的心理活动，也是具有决定作用的影响因素。这就不难理解，为什么有那么多优秀的心理学家，不管是来自社会心理学、认知心理学、发展心理学，还是来自文化心理学、教育心理学、工业心理学，都纷纷涉足这一领域；也就不难理解，为什么有那么多的社会和人文学科，不管是经济学、管理学、法学，还是政治学、历史学、哲学，无不从人类决策与判断的心理学研究中汲取知识和灵感。这也就是为什么诺贝尔经济学奖会超越常规地在1978年和2002年两次授予研究人类决策与判断的心理学家的原因。两次获奖的心理学家，一位是赫伯特·西蒙教授，另一位就是丹尼尔·卡尼曼教授。

本书系统地介绍了赫伯特·西蒙和丹尼尔·卡尼曼的发现，同时也阐述了决策与判断心理学中很多还没有被经济学家和其他社会及人文科学家所认识到的现象和研究发现。可以说，本书80%以上的内容超出了西蒙和卡尼曼工作的领域，因此可以想像，心理学家对经济学和其他相关学科的影响还有很多值得发掘的空间。难怪美联储主席格林斯潘曾经开玩笑地说“所谓的‘新经济’实际上就是心理学”。

值得一提的是，上述两位心理学家都对中国特别友好。早在1981年，西蒙教授就接受北京大学心理学系的邀请，成为第一个在中国讲授人类决策和判断的心理学家。那时，年仅19岁的我正是心理系的本科生，一下

子就被这一领域的严谨、生动和奥秘所深深吸引。西蒙的研究第一次证明决策与判断是人的思维活动，它不是建立在数学和逻辑基础之上的，而是建立在人的感情、理念和经验的基础上的。他把决策的原则定义成第一满意原则，也就是说，我们做出决策和判断的标准并不是建立在理性基础上的“最佳选择”，而是建立在人类心理上的“第一满意选择”。卡尼曼的工作就是在这一基本理论的指导下找到了那些影响我们非理性选择的因素。

我与卡尼曼教授的第一次个人接触是在1997年的夏天，那是在柏克利做教授的第一个星期。一天，有一位慈祥的老人敲开了我办公室的门，欢迎我到柏克利来工作，并告诉我这间办公室就是他当年工作过的地方，还说他最想访问的地方就是中国。这位老人正是卡尼曼教授，而他想访问中国的愿望也终于在7年以后的2004年得以实现。今年夏天，卡尼曼作为第28届国际心理学联合会的特邀嘉宾来到北京，在开幕式上做了关于人类决策与判断心理学的主题发言，发言的最后，他表达了希望看到更多中国心理学家从事决策与判断研究工作的愿望。我觉得本书的出版和发行是对这位慈祥老人的致敬，也是对他所表达愿望的呼应。

坐在加州大学柏克利分校卡尼曼教授曾经工作过的办公室里，审阅第一本系统介绍人类决策与判断的心理学教科书的中译稿，我似乎看到了卡尼曼教授欣慰的笑容。

彭凯平

美国加州大学柏克利分校心理学教授

序 二

如果你想知道心理学的热门课题，你只要读一读社会心理学。在过去的几年中，社会心理学已经成为心理学中了解人类思维、情感和行为的最重要的研究领域。这就是我们为什么越来越多地看见很多以社会心理学命名的研究领域。例如社会认知、社会发展、社会学习和社会人格等等。

社会心理学已经研究了许多复杂的社会问题。作为心理学领域最新的一批热心人，社会心理学者研究了几乎所有的个人和社会问题——从心理生理学到和平心理学，从学生的失败归因到艾滋病的预防教育。可以说，自从美国心理协会前任主席乔治·米勒呼吁心理学家“把心理学还给人民”，社会心理学家就已经是最早的响应者。

“麦格劳-希尔社会心理学系列丛书”是对许多杰出的研究者、理论家和工作者多年的辛勤劳动的赞颂。丛书的每一位作者都抱有严谨的科学态度，和尽可能把科学的信息传达给教师、研究人员、学生和普通民众的教育意愿。丛书的目的是尽可能地反映社会心理学的广泛领域，每一项选题都具有研究的广度和深度。使用这些书的教师可以将这套丛书作为基础课程的参考读物，或者是以这套丛书作为更深入地钻研某些课题的参考资料。

正如认知心理学家发展了决策研究的前沿课题，例如揭示了决策的“理性人”假设模型的局限性，社会心理学家已经把这项研究推广到了新的领域和方向。在这部新的创造性的作品当中，斯科特·普劳斯向我们展示了决策与判断的社会心理学分析如何帮助我们处理日常生活中的普通问题。普劳斯不仅把这些复杂的研究问题变成了通俗易懂的阅读材料，而且他还提出了许多新奇的洞见、有价值的新概念和有意义的结论。

我相信读者也会发现最令人高兴的是，这本书很巧妙地融合了高水平的学术探索和把复杂的学术思想巧妙地传输给包括本科生到商业和卫生界的专业人士直到国家级的民意领袖的技巧。这本书充满了幽默感，对老的思想的新的解释，批判性思维的练习，从而把许多通常不相关的材料整合

成了一本令人遐思发人深省的好书。可以说，从来还没有一个年轻作者的第一本著作能在一个重要的课题上对这么多的人产生如此大的影响。

菲利普·津巴多，丛书主编
斯坦福大学

前 言

今天，美国人可以在超市挑选 25 000 多种商品，他们也可以阅读将近 11 000 份杂志和期刊，他们还可以选择收看 50 多个电视台的任何一个频道 (Williams, 1990 年 2 月 14 日)。一句话，他们面临着一系列令人头疼的选择。

人们究竟是如何做出决策的？他们又是如何畅游在信息的海洋中而不被淹没？什么原因导致他们做出某种决策？

本书试图对这些问题做出初步的回答。本书的对象是希望了解决策与判断心理学的基础知识的非专业人士。它着重的是实验结果而不是心理学理论，是出人意料的结论而不是猜想，是对研究的描述而不是数学公式。一句话，这本书是想要大家高兴而且思考，同时也是为了传播和普及心理学的知识。

本书共分 6 部分，前两个部分主要介绍决策与判断的基本要素，包括知觉、记忆、情境和提问方式。第三和第四部分主要介绍决策的经典模式，并与最近有关判断偏差的新模式做对比。第五部分探讨由团体做出的判断和关于团体的判断。第六部分讨论决策与判断的一些常见的陷阱。每一章都设计成能够单独成立的专题，所以读者可以自由地跳跃不同的章节或随意地安排阅读顺序。

本书一个与众不同的地方就是第 1 章之前的读者调查，这个调查的问题是根据以后章节中涉及的研究问题改编和复制而成的。一旦你完成了这个调查，你就可以将你的答案与原来的研究中人们的反应做比较。有的时候你的回答可能与他们的反应相似，有的时候可能不一致。但是，无论何种情况你自己的决策与判断早已记录在案。由于你对读者调查的回答都将在本书中得到充分的讨论，所以，在你阅读任何其他章节之前完成这个调查是很重要的。

关于心理学的实验研究

对许多不熟悉心理学的读者来说，本书中使用的一些术语和实验方法可能看起来有些粗鲁和不仁道。比如说，实验的参加者通常被称为被试。还有，在有些实验程序中，研究的参与者故意被误导以至于他们不知道研究的真正目的。还有些研究结果将我们人类描绘成会产生严重的偏差和错误的决策者。

这些关注很重要，值得我们在此做出说明。

第一，我们使用术语“被试”的原因是它比其他词汇，例如人或个人更为准确，同时又不像参加者或自愿者拗口。被试这个词在心理学中是一个常用的概念，在本书中也使用得很普遍，但它并不意味着心理学家将实验的参与者看做非人类的物体。事实上，大多数心理学研究的真正被试并不是实验的参与者本身，而是参与者的行为。

第二，有些心理学研究利用误导往往是有其特殊的原因。在很多时候，被试都会读到很模糊的或者是误导的故事，这样的话他们的行为就不会受到实验真正目的的影响。比如说，一项关于群体动力学的研究中，一开始可能会告诉被试研究的目的是关于学习和记忆，这样的话他们就不会过分地关注他们群体内部的差异。在有些情况下，误导的使用是为了产生一般情况下不可能产生的特殊情形。例如，一项有关创造力的研究的实验可能会将被试随机分为两组，一组会得到他们解决问题的能力正面反馈，另一组得到负面反馈。如果被试事先知道他们会被随机分配这种情况，那么反馈的差异就失去了意义。

当然，我们有很多理由反对在研究中使用误导（Warwick, 1975年2月），但我们必须指出，美国心理协会（1989）早已经推出了一整套保护被试的原则。这些原则要求被试要得知足够的信息，要征得他们的同意。误导只能是研究的最后手段，被试也不应当受到伤害或感到明显的不舒适，被试也应该有随时退出实验的权利。被试的材料信息要高度保密，任何的误导在实验之后都应该加以澄清和解释。大多数大学都要求其研究人员严格遵守这些规定。

最后，相对于成功，心理学家更关注偏差和误差还存在其他的原因。其中一个最简单的原因就是这样的研究就是多一些（Kahneman, 1991）。许多学术期刊愿意发表与我们的常识相违背的研究结果，故此，所发表的很多研究都是有关决策失误而不是决策成功。从这个意义上讲，专业期刊

和报纸及杂志没什么两样，它们都偏好令人意外和令人着迷的消息。根据一项估计（Christensen - Szalanski & Beach, 1984），学术期刊引用关于错误推理的研究的数目大约是引用有关成功推理研究数目的6倍。由于本书是有关决策与判断研究的入门教材，于是它很自然地反映了学术界的偏好。

有关为什么我们如此关注人类的弱点还有一个重要原因。决策与判断的失败通常比成功更有启发性，即使成功是普通的原则。本书着重讨论偏差和误差，并不是要说明人类是不好的决策者，而是说明决策与判断的失败经常能够揭示决策的过程，就像汽车的失败往往能够揭示汽车是如何运转的。正如尼斯比特（Richard Nisbett）和罗斯（Lee Ross）所解释的：这种研究方式是“建立在类似于错觉或思维障碍研究的前提之上，这就是说知觉结构和过程的本质能够通过它们所产生的误差所揭示”（1980, p. xii）。

当然，对误差的强调还有容易应用的优势。一旦你能够发现引起某种偏差和误差的环境，你就可以避而远之或者是提前预备对策。为了帮助你找到对策，本书包括了很多来自医学、法律、商务、教育和核控（本人的专长）和其他领域的一些具体的例子。由于本书所描述的研究非常的普通适用，因此我决定向还不熟悉决策与判断研究的读者保证：如果能够有效地应用本书所描述的结果，你就会更好地避免决策偏差、误差和陷阱，而且你将会更好地理解他人所做的决策。

斯科特·普劳斯

目 录

读者调查	1
第一部分 知觉、记忆和情境	13
第 1 章 选择性知觉	14
第 2 章 认知不协调	20
第 3 章 记忆和事后聪明式偏差	28
第 4 章 情境依赖性	35
第二部分 问题如何影响答案	45
第 5 章 可塑性	46
第 6 章 问题的措辞和框架的影响	57
第三部分 决策模型	69
第 7 章 期望效用理论	70
第 8 章 理性决策的悖论	75
第 9 章 描述性决策模型	83
第四部分 直觉与偏差	95
第 10 章 代表性直觉	96
第 11 章 易得性直觉	106
第 12 章 概率和风险	115
第 13 章 锚定与调整	128
第 14 章 对随机性的知觉	135
第 15 章 相关、因果关系与控制	143
第 16 章 归因理论	153
第五部分 决策与判断中的社会性一面	167
第 17 章 社会影响	168
第 18 章 群体决策与判断	181
第六部分 常见陷阱	189
第 19 章 过度自信	190
第 20 章 自我实现的预言	202
第 21 章 行为陷阱	210
编后记：回顾	221
参考文献	229
推荐读物	258

读者调查

在读这本书之前，请务必完成这一问卷。这样，你才不会在做这个测验时因为受到本书内容的影响而产生判断上的偏差，也不会有“我早就知道这一切”效应存在（本书第3章将讨论）。如果你不想在书上直接做，你可以选择做在一张纸上，并把它当做书签。

1. 琳达是一位31岁的单身女性，坦率直言，而且非常聪明。她主修哲学，在念大学时对歧视、社会公平等问题非常感兴趣，而且积极参加了反核示威游行。你认为以下哪个选项最符合对琳达的描述：

- 琳达是一个银行出纳员
- 琳达是一个银行出纳员，而且是一个积极的女权运动参与者

2. 如果面临以下情况，你将会选择哪一种情形？

- 肯定（百分之百的概率）会输掉50美元
- 25%的概率输掉200美元，75%的概率什么也不会输掉

3. 约翰是一个嫉妒心强、顽固、挑剔、冲动，但勤勉、聪明的人。从总体上来讲，你认为约翰在多大程度上是一个情绪化的人？（划圈选择）

一点也不情绪化 1 2 3 4 5 6 7 极端情绪化

4. 吉姆是一个聪明、灵巧、勤勉、热心、坚定、实际而且谨慎的人。你认为他最有可能具有以下哪些特征？

在每组中选择一个特征：

- 大方——小气
- 不快乐——快乐
- 易怒的——脾气好
- 幽默——乏味

5. 以下这个问题只针对大学生：与和你同年龄同性别的大学生相比，以下事情发生在你身上的可能性有多大？请就每一事件选择你认为可能的概率。

2 读者调查

a. 出现酗酒问题

- 更可能出现这种情况的概率是 60% 以上
- 更可能出现这种情况的概率是 40%
- 更可能出现这种情况的概率是 20%
- 出现这种情况的概率与不出现这种情况的概率是相等的
- 更可能不出现这种情况的概率是 20%
- 更可能不出现这种情况的概率是 40%
- 更可能不出现这种情况的概率是 60% 以上

b. 将拥有自己的家庭

- 更可能出现这种情况的概率是 60% 以上
- 更可能出现这种情况的概率是 40%
- 更可能出现这种情况的概率是 20%
- 出现这种情况的概率与不出现这种情况的概率是相等的
- 更可能不出现这种情况的概率是 20%
- 更可能不出现这种情况的概率是 40%
- 更可能不出现这种情况的概率是 60% 以上

c. 毕业后的起薪超过每年 15 000 美元

- 更可能出现这种情况的概率是 60% 以上
- 更可能出现这种情况的概率是 40%
- 更可能出现这种情况的概率是 20%
- 出现这种情况的概率与不出现这种情况的概率是相等的
- 更可能不出现这种情况的概率是 20%
- 更可能不出现这种情况的概率是 40%
- 更可能不出现这种情况的概率是 60% 以上

d. 40 岁前会得心脏病

- 更可能出现这种情况的概率是 60% 以上
- 更可能出现这种情况的概率是 40%
- 更可能出现这种情况的概率是 20%
- 出现这种情况的概率与不出现这种情况的概率是相等的
- 更可能不出现这种情况的概率是 20%
- 更可能不出现这种情况的概率是 40%
- 更可能不出现这种情况的概率是 60% 以上

6. 作为一家航空公司的总裁，你已经将公司的 1 000 万美元投入在一个研究项

目中，目的在于研究一种不能被雷达探测到的飞机，也就是说，一种隐形飞机。当项目已经完成 90% 时，另一家公司已经开始销售隐形飞机。而且，那家公司的飞机在速度和成本方面似乎优于你们公司生产的飞机。问题是：你应该将剩下的 10% 的资金继续投入项目以完成这架隐形飞机的设计制造吗？

- 不会——再继续花钱在这个项目上是毫无意义的
- 会——既然已经花了 1 000 万美元在这一项目中，就应该继续把它做完

7. 在美国，以下哪个原因更可能导致死亡——被飞机上掉下的零件砸死还是被鲨鱼袭击？

- 被飞机上掉下的零件砸死
- 被鲨鱼袭击

8. 以下每对词组中的哪个词更可能是美国常见的死亡原因？

糖尿病——谋杀

龙卷风——闪电

车祸——胃癌

9. 考虑以下历史场景：距离超级大国 A 不远的一个国家在经历了某些政党制度的变革后，其政府决定扩大与超级大国 B 的贸易往来。为了阻止这种政府和贸易制度的改变，超级大国 A 向这个国家派出军队，并以军事力量帮助原来的政府复辟。

这个超级大国是哪个国家？ 前苏联 美国

你对你所选择答案的确定程度：

一点也不确定 1 2 3 4 5 6 7 非常确定

10. 再考虑下一种场景：20 世纪 60 年代，超级大国 A 向一个周边的小国发起了突然的军事侵略，企图推翻当时的执政政府。这场侵略失败了，而且当初的侵略者中，大部分被处以死刑或是监禁。

这个超级大国是哪个国家？ 前苏联 美国

你对你所选择答案的确定程度：

一点也不确定 1 2 3 4 5 6 7 非常确定

11. 在未来 10 年中，你认为最有可能发生的事件是：

美国和俄罗斯将爆发一场全面的核战争

4 读者调查

美国和俄罗斯将爆发一场全面的核战争，但一开始两国都不想动用自己的核武器，只是在卷入了一场局部战争之后，如伊拉克、利比亚、以色列或者巴基斯坦等国的战争，才被迫动用核武器。

12. 将一张纸对折为两半，再对折，再对折，直到对折 100 次以后，这张纸有多厚：

我认为最有把握的答案是，这张纸有____厚。

我有 90% 的把握，这张纸的厚度在____到____之间。

13. 如果算上 2 月 29 日，一年之中将会有 366 个生日。因此，如果一个小组中至少有两个人是同一天过生日，那么这个小组至少应该有 367 个人。如果小组中两个人同一天过生日的概率为 50%，那么这个小组需要____位成员。

14. 假设一项对 250 名神经症患者的研究发现，他们出现头晕症状和脑瘤的发病频率如下：

		脑 瘤	
		出 现	不出现
头 晕	出 现	160	40
	不出现	40	10

A. 在这组样本中，要计算头晕症状和脑瘤是否相关，需要上表中哪些空格中的数据。

左上角

左下角

右上角

右下角

B. 从表中的数据来看，头晕是否与脑瘤相关？

是 否 不确定

15. 已知一个城市中，八年级学生的平均 IQ 得分为 100。你从这些学龄儿童中随机挑选了 50 个人作为样本。第一个被试的 IQ 为 150，你认为你所挑选的样本平均 IQ 分值为多少？

答案：_____

16. 从总体上来讲, 你认为自己是一个性别歧视的人吗?

是 否 不确定

17. 如果把世界上所有人的血都装入到一个立方体容器里, 这个容器的宽度将是_____。

18. 你认为下列哪一个对罗夏墨迹测验的解释最能够预测男性的同性恋倾向? (如果你从未听说过罗夏墨迹测验, 可以跳过这一题。)

对以下阐述按照从 1 (最大可能预示同性恋) 到 6 (最小可能预示同性恋) 的顺序排序。

____ 性别模糊的体态

____ 同时具有男性和女性特征的体态

____ 臀部或者肛门

____ 生殖器官

____ 扭曲怪异的体态

____ 女性打扮

19. “记忆与大脑中的的存储器相关, 我们将一些东西存储进这个容器, 等到需要的时候再将它从这个容器中取出。有时候, 一些东西会从这个容器中漏掉, 这时候我们就会忘记一些东西。”你认为这一描述是否很恰当地描述了记忆的加工方式?

是 否 不确定

20. 一个人以 60 美元的价格买了一匹马, 又以 70 美元的价格售出。然后他以 80 美元的价格买进, 尔后又以 90 美元的价格出售。在买马卖马的过程中, 他一共赚了多少钱? 他在交易最后赢得了_____美元的利润。

21. (1) Absinthe (苦艾) 是一种:

烈酒 一种宝石

(2) 你对你的答案有多大把握?

0.5 0.55 0.6 0.65 0.7 0.75 0.8 0.85 0.9 0.95 1.00

22. 不用仔细计算, 在 5 秒钟之内给出这个式子的估计答案:

$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 =$ _____

6 读者调查

23. 假设你正在考虑是否给你的某项财产上保险以预防可能发生的损失，如火灾、失窃等。在计算过风险和保费以后，你发现在购买保险和不购买保险之间并不存在什么差异。

这时候，保险公司推出的一种名为概率保险 (probabilistic insurance) 的新保险计划引起了你的注意。在这一保险计划中，你开始只需支付一半的保费。而一旦发生损失，有 50% 的概率你可以付清另一半保费以获得全部赔偿；有 50% 的概率你只能拿回已付的保费并自己承担所有损失。

比如说，在一个月中的某个奇数日子发生了一次事故，你可以付清另一半保费从而获得全部赔偿；而如果这一事故发生在这个月的偶数日子里，你将会得到已付的保费，但必须自己承担所有损失。

请记住，获得全部赔偿的保费并不能够带来额外收益，只是刚好与成本相当。

在这种情况下，你会购买这种概率保险吗？

- 会 不会

24. 假设你在参加过的各种考试中都取得了很好的成绩，而别人参加了类似考试却并未获得好成绩。你对此会得出什么结论？（选择一个与你所想的最接近的答案）

- 解释 A：这些考试可能都比较简单
 解释 B：其他人的能力可能比较弱
 解释 C：我要么是善于考试，要么是对考试材料掌握得非常之好

在接下来的几页中，你可能需要回答以下问题。请仔细阅读这些问题，并继续做完这份读者调查：

- 蚂蚁吃了桌子上的甜果冻
- 蚂蚁在厨房
- 蚂蚁吃了甜果冻
- 厨房里的蚂蚁吃了桌子上的果冻
- 果冻放在桌子上
- 厨房里的蚂蚁吃了果冻

25. 如果你面临以下选择，你会选择哪一个？

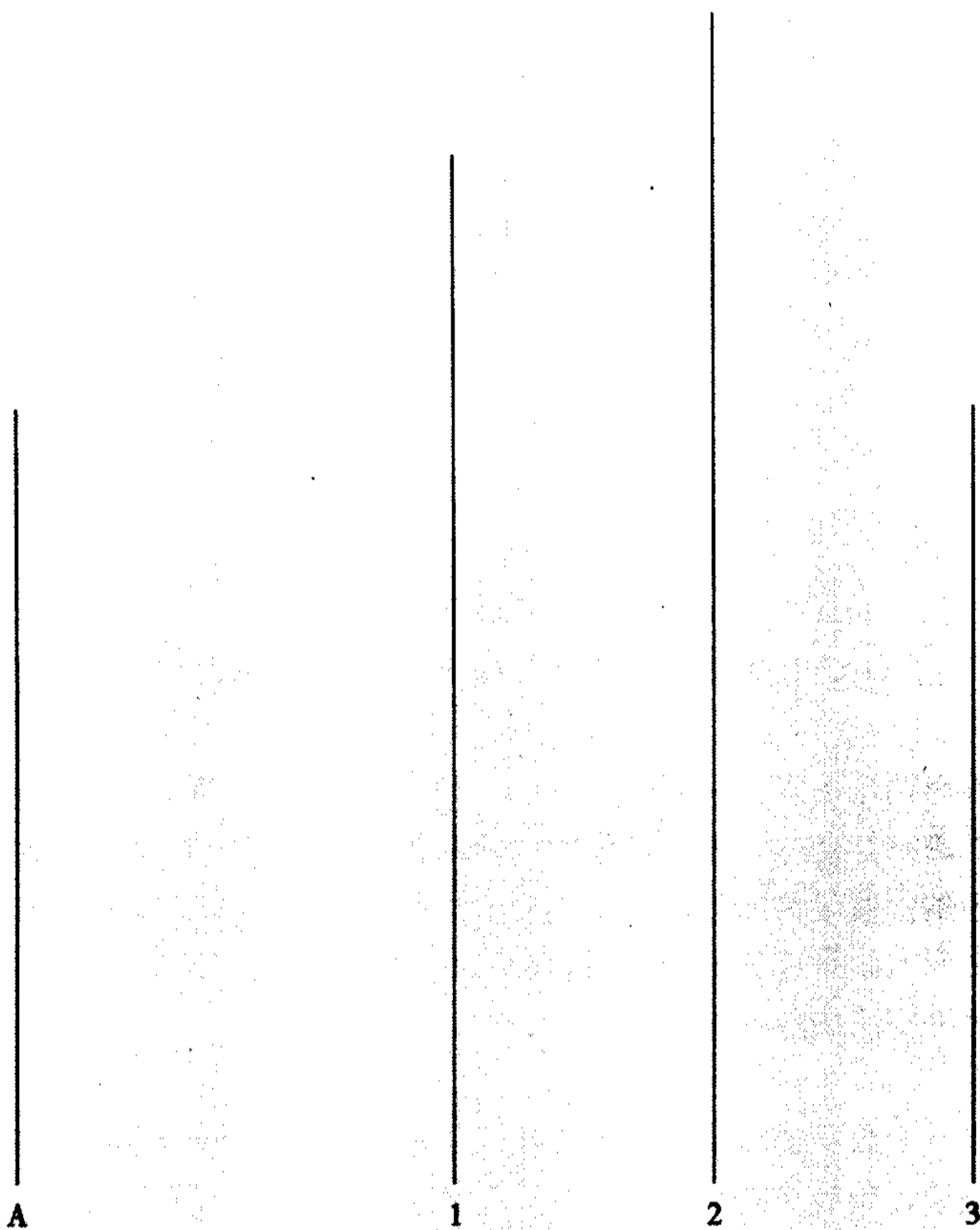
- 肯定会获得 240 美元
 25% 的概率获得 1 000 美元，75% 的概率什么也得不到

26. 如果你面临以下选择，你会选择哪一个？

- 肯定会输掉 750 美元
- 75% 的概率输掉 1 000 美元, 25% 的概率什么也不会输掉
27. 你认为现在国家面临的最重要的问题是什么?
28. (1) 你会选择下列何种赌博方式?
- 肯定会获得 100 万美元
- 10% 的概率获得 250 万美元, 89% 的概率获得 100 万美元, 1% 的概率什么也得不到
- (2) 你会选择下列何种赌博方式?
- 11% 的概率得到 100 万美元, 89% 的概率什么也得不到
- 10% 的概率获得 250 万美元, 90% 的概率什么也得不到
29. 假设有两个装满了扑克牌 (Poker chips) 的容器。第一个容器里有 70% 的扑克牌是红色的, 30% 是蓝色的。而第二个容器里, 70% 是蓝色, 30% 是红色的。假设你随机选中了其中一个容器, 并从中随机抽出 12 张扑克牌, 其中有 8 张是红色的, 4 张是蓝色的。那么你抽中的容器是那个装有更多红色扑克牌的容器的概率是多少?
30. 这是一个公平的掷硬币游戏。在这个游戏中, 如果出现字的那一面, 游戏便结束。游戏结束后, 你将获得 2^K 美元。 K 代表出现字之前掷硬币的总次数。也就是说, 如果第一次掷硬币时就出现字, 你将会得到 2 美元, 如果掷两次硬币才出现字, 你将获得 4 美元, 如果掷三次硬币才出现字, 你将获得 8 美元。总的来说, 也就是
- 直到出现字时掷硬币的总次数: 1 2 3 4 5 ... K
- 你将获得的美元: 2 4 8 16 32 ... 2^K
- 你愿意付多少钱来玩这个游戏?
31. 一个公平的掷硬币游戏。一共掷了三次, 每次硬币都是头的一面朝上。如果你要在下一次掷硬币上下注 100 美元, 你会赌哪一面朝上?
- 头 字 没有偏好

8 读者调查

32. 将下面的线 1、2、3 和线 A 做比较。哪根线和线 A 一样长？



- 线 A 和线 1 一样长。
- 线 A 和线 2 一样长。
- 线 A 和线 3 一样长。

33. 在下面的句子中，字母 f 出现了多少次？

These functional fuses have been developed after years of scientific investigation of electric phenomena, combined with the fruit of long experience on the part of the two investigators who have come forward with them for our meetings today.

字母 f 出现了____次。

34. 不要回头看前面的内容，请指出下列句子是否曾出现在你刚才所读到过的句子中。并用 5 点量表来衡量你对自己的答案的肯定程度，1 代表肯定程

度很低，5 代表肯定程度很高。

(1) 蚂蚁吃了桌子上的果冻

出现过 没有出现过

肯定程度: _____

(2) 厨房里的蚂蚁吃了桌子上的甜果冻

出现过 没有出现过

肯定程度: _____

(3) 蚂蚁吃了甜果冻

出现过 没有出现过

肯定程度: _____

35. 假设高中时期的测验分数在一定程度上与大学学业成绩 (GPA) 相关。给定下面的百分比分布，如果一个学生在高中测验中得分为 725，你预测他的 GPA 会是多少？

学生百分比分布	高中成绩测验	GPA
前 10%	> 750	> 3.7
前 20%	> 700	> 3.5
前 30%	> 650	> 3.2
前 40%	> 600	> 2.9
前 50%	> 500	> 2.5

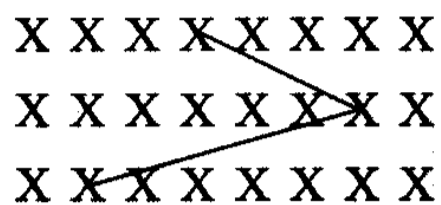
你预测这个学生的 GPA 是_____

36. 假如你投了某个候选人一票，这是否会改变你对他胜出本次竞选的概率的判断？

会 不会 不确定

37. 考虑以下两种结构 A 和 B:

结构A:



结构B:



一条路径将结构中第一排的 X 和最后一排的 X 连接起来，并通过这中间每一排的一个 X（注意每一排只有一个 X 被通过）。也就是说，结构 A 中的路径通过了 3 个 X（3 排 X，每排 1 个），结构 B 中的路径通过了 9 个 X（9 排 X，每排 1 个）。在上图中，每种结构我们都已经给出一条路径的例子。

(1) 哪种结构中，存在 3 条以上的路径？

- 结构 A 结构 B

(2) 结构 A 中大概有几条路径？

(3) 结构 B 中大概有几条路径？

38. 以下关于 X 和 O 的排列，你认为哪种排列更类似于随机排列（比如像掷硬币一样）？

- X O X X X O O O O X O X X O O O X X X O X
 X O X O X O O O X X O X O X O O X X X O X

39. 假设以下四张牌一面是字母，而另一面是数字。某个人告诉你：“如果某张牌的一面是一个元音字母，那么它的背面一定是一个偶数。”你需要翻开下面哪张牌来检验这个人说的话是否正确？

E

K

4

7

- E
- 4
- K
- 7

第一部分

知觉、记忆和情境

不依赖于情境的决策是不存在的。我们所做出的决策与判断都取决于我们看待和解释这个世界的方式。因此，本书的第一部分将讨论，我们的决策与判断如何受到选择性知觉、维持认知一致性的压力、记忆偏差布鲁纳以及情境变化的影响。

第 1 章

选择性知觉

“我们不是先看见再定义，而是先定义再看见。”

——沃尔特·李普曼（引自 Snyder & Uranowitz, 1978）

看看你眼前的东西。现在，再看看你的双手，再看看本书的封面。你所看到的东西在多大程度上是由你的预期所决定的？

如果你所看见的和大多数人一样，那就意味着你的知觉在很大程度上受到了你的预期的影响。即使对于一些近在眼前的东西的观察也很难不受已有观念的影响。你也许会认为，自己看待事情是完全没有偏差的。然而越来越多的证据表明，要避免知觉偏差几乎是不可能的。事实上，人们会选择性地去感知那些他们所期望和愿意看到的事物。

黑桃之所以被称为黑桃

杰米奥·布鲁纳和利奥·波斯特曼（Jerome Bruner & Leo Postman）于 1949 年发表的实验是关于选择性知觉的最早和最著名的实验之一。在这一实验中，他们利用视觉记忆测试镜向被试出示 5 张牌，每展示完 5 张牌为一个系列。视觉记忆测试镜向被试出示这 5 张牌时，会有很短的间隔时间，不同系列的间隔时间从十毫秒到一秒钟不等。他们向被试展示的牌与本书封面上的牌类似，请注意看一下这些牌是什么。

你是否注意到了这些牌独特的地方？如果只是不经心地扫一下封面，大多数人不会发现，其中一张牌的牌面是由 3 个黑色的心形组成。布鲁纳和波斯特曼的实验也发现，被试需要四次以上的展示才会意识到这张牌与普通牌的不同之处。而且他们发现，大多数人对于这种不协调的反应可以归为四种：支配、折衷、分裂和再认。

布鲁纳和波斯特曼认为，“知觉性否认”是引起支配反应的主要原因。比如，当看到这张有 3 个黑色心形的牌的时候，人们会非常肯定地认为，这张牌要么是正常的红心，要么是正常的黑桃。在第一种情况下，形状是占支配地位的，颜色被预先的期望所同化。而在第二种情况下，则是颜色占支配地位，形状被同化。在布鲁纳和波斯特曼的实验中，28 个被试中有 27 个人（即 96% 的

人)在一定程度上表现出这种支配反应。

另一个反应则是折衷。例如,实验中一些被试将一张有6个红色桃心的牌看成了有6个紫色桃心或6个紫色心形的牌。另外一些人则将一张有4个黑色心形的牌看成有4个带灰色桃心的牌。还有一些人将一张有6个红色梅花的牌解释为“被红色亮光照亮的梅花牌”(实验中的被试都是在视觉记忆测试镜上看这些牌的)。在布鲁纳和波斯特曼的实验中,有一半的人和11%的人分别对红色的牌和黑色的牌做出了这种折衷反应。

人们对这种不协调现象的第三种反应是分裂。如果反应是分裂的,人们将很难获得任何一种形式的知觉。分裂反应很少出现,然而一旦发生,其结果是非常戏剧化的。例如,一个被试声称:“我不知道这到底是怎么回事,甚至不确定这些是否是扑克牌。”同样,另一个被试说:“我辨认不出这些花色来,不管它们是什么,这甚至不像一张扑克牌。我现在不知道它是什么颜色,也不知道它究竟是黑桃还是红心。我现在甚至不确定黑桃到底是什么样的,我的天呀!”

最后一种反应是再认。但有时候即使被试意识到了有什么不对的地方,他们仍然说不清楚究竟是哪里出了问题。在完全弄明白这些牌的错误之前,已经有6个被试开始意识到这些牌上面的标志的位置有些奇怪。例如,一个被试在看到那张有6个红色桃心的牌时,他认为这些标志只是被倒转了。而另一个被试在看到那张有4个黑色心形的牌时则认为,这种不协调完全是由黑桃的放置引起的。

这些结果都表明,预期在很大程度上影响了人们的知觉。布鲁纳和波斯特曼认为:“我们的知觉结构在很大程度上是由我们的预期所决定的,这些预期建立在过去和情境的基础上。”当人们对某一特定情况有了足够的经验时,他们就会看到那些他们所预期看到的东西。

读者调查的第33题就提供了一个有关过去的经验将如何影响现在的知觉的例子。在这一问题中,读者需要计算下列句子中字母f出现的次数: *These functional fuses have been developed after years of scientific investigation of electric phenomena, combined with the fruit of long experience on the part of the two investigators who have come forward with them for our meetings today.*

大部分以英语为母语的人会低估字母f出现的次数(Block & Yunker, 1989)。正确的答案是11次(包括出现在单词of中的4次)。因为习惯于说英语的人会将单词of中的f发音为v,所以相对于那些没有经验的人来说,他们更难察觉到of中f的出现频率。事实上,正是过去的经验影响了他们的表现。

强有力的预期

假设你是一名男大学生,你参加了罗切斯大学酒精研究实验室的一项研

究。实验者会告诉你，你所要做的就是喝一杯伏特加和汤尼水（以奎宁调味的含矿物质饮料）的混合液体。二十分钟后，等酒精充分进入血液，他们会安排你与实验者的一名女助手交谈，她会想尽各种办法给你留下一个美好的印象。然后，实验者就会按照你的体重成比例地混合伏特加和汤尼水，并留下你一个人在房间里面享用这一饮品。

等你喝完这杯饮品后，一个女助手便会走进你的房间，在椅子上坐下来，望着你的眼睛。你开始与她交谈。你有多紧张？心跳会有多快？

特伦斯·威尔逊和戴维·艾布拉姆斯（G. Terrence Wilson & David Abrams）于1977年做了这一实验。他们发现，不管酒精是否真正被充分吸收了，相对于那些认为自己只是喝了一杯汤尼水的被试来说，那些认为自己喝了一杯伏特加和汤尼水的混合液体的被试的心跳速率增加得更少一些。也就是说，被试是否摄入了酒精对心跳速率其实并没有什么大的影响，真正影响他们心跳速率的是，他们是否相信自己摄入了酒精。由此证明，相对于血液中化学成分的实际变化来讲，心理预期有着更重要的影响。

戴维·麦克米伦等人（David McMillen, Stephen Smith, & Elisabeth Wells - Parker, 1989）进一步延伸了这一实验。他们采用了和威尔逊相似的实验方法，随机安排大学生饮用酒精饮料和非酒精饮料。而在实验之前，通过一些测定手法预先将被试分为两组：一组是高强度的“刺激寻找者”，即更偏爱冒险的一类人；另外一组则是低程度的“刺激寻找者”，也就是更加谨慎小心的一类人。喝完饮料的半个小时后，被试开始玩一个视频游戏。在游戏中，他们会沿着一条公路开车，并可以加速超过其他车辆。被试被告知，当他们玩这个视频游戏的时候，要想像成他们是在现实中开车。

麦克米伦和他的同事们发现，不管被试是否真的喝了酒精饮料，那些认为自己喝了酒精饮料的高程度“刺激寻找者”改变车道的次数和超车的次数要明显多于那些认为自己没有喝酒精饮料的高程度的“刺激寻找者”。与此形成对照的是，那些认为自己喝了酒精饮料的低程度的“刺激寻找者”则比那些认为自己没有喝酒精饮料的低程度的“刺激寻找者”更加谨慎小心。一些实验表明，在吸食大麻的人群中同样也存在这种很强的预期效应（Jones, 1971）。

在以上所介绍过的实验中，人们的知觉都受到了已有信念和预期的强烈影响。心理学家将这些影响称之为“认知”因素。实际上，知觉不仅仅受到人们的预期的影响，也受到了其愿望的影响。那些与希望、欲望以及情感等因素相关的影响被称为“动机”因素。本章余下部分将讨论，在选择性认知中，动机因素与认知因素是如何交互作用的。

什么时候挑起了争端

1951年11月23日，达特茅斯学院与普林斯顿大学的橄榄球队在普林斯顿

大学的帕尔默体育场进行了一场比赛。开球才一会儿，球场上的形势就表现得相当混乱。普林斯顿的一个明星球员，那个刚刚上过《时代》杂志封面的家伙，被打断鼻子后离开了球场。过了一会儿，达特茅斯的一个球员也带着一条伤腿离开。最后，普林斯顿赢得了比赛。但比赛结束前双方都得到了同样多的黄牌和红牌惩罚。

比赛结束后，双方都怒气冲冲、互相指责，并不断发表有关对方的苛刻评论。例如，比赛结束四天后，《普林斯顿日报》（*Daily Princetonian*，普林斯顿学生创办的一份报纸）的一名记者宣称：“观众们从来都没有看过如此糟糕透顶的表演——所谓的体育比赛。两个队都有错，但责任主要还是在达特茅斯队。普林斯顿队明显强于达特茅斯队，而且根本犯不着与达特茅斯队动粗。”就在同一天，《达特茅斯》（本科生报纸）指责普林斯顿的教练一直在向球员们灌输“以其人之道还治其人之身”的思想。接下来的一个星期里，这场关于发生了什么以及谁该负责的激烈争论在普林斯顿和达特茅斯学生中间愈演愈烈。

这场混战引起了达特茅斯一个叫艾伯特·哈斯托夫（Albert Hastorf）的社会心理学家和普林斯顿的哈德利·坎特里尔（Hadley Cantril）的研究员的注意。他们利用这场混战，于1954年做了一项有关选择性知觉的经典实验。

实验开始时，他们向163名达特茅斯学生和161名普林斯顿学生询问以下问题：“从你观看比赛现场或是比赛录像，又或者从你所读过的东西来看，你认为哪支队伍最先挑起了争端？”不出意料的是，哈斯托夫和坎特里尔发现，达特茅斯学生和普林斯顿学生的反应存在明显的差异（discrepancy）。达特茅斯学生中有53%的人认为双方都有过错，只有36%的人认为是由达特茅斯队挑起的这场争端。而普林斯顿的学生中有86%的人认为是达特茅斯队挑起的争端，只有11%的人认为双方都有过错。

哈斯托夫和坎特里尔开始思考，产生这种看法上的差异究竟是因为达特茅斯和普林斯顿的学生真的观看了不同的比赛（录像），还是因为他们虽然观看了同样的比赛却从不同的角度去解释所看到的事实。为了研究这一问题，他们在两所大学又找了一组学生，让他们观看这场比赛录像，并记录下他们所注意到的犯规行为。两所大学的学生都看的是同一个比赛录像，并用同一个评价系统来记录所观察到的犯规行为。

正如图1.1中所示，实验结果显示出强烈的选择性知觉效应。达特茅斯的学生观察到，两个队的犯规行为几乎一样多（达特茅斯队为4.3，普林斯顿队为4.4）。而普林斯顿的学生则观察到，达特茅斯队的犯规行为是普林斯顿队犯规行为的两倍多（达特茅斯队为9.8，普林斯顿队为4.2）。双方的知觉差异如此之大，以至于当普林斯顿送了一份比赛录像的拷贝给一些达特茅斯校友供他们集体观看时，一个曾经观看过比赛录像的达特茅斯校友发现，他在这部录像中竟然看不到达特茅斯队的犯规行为。他感到非常疑惑，甚至发了一份电报给普林斯顿要求其拿出完整的比赛录像。

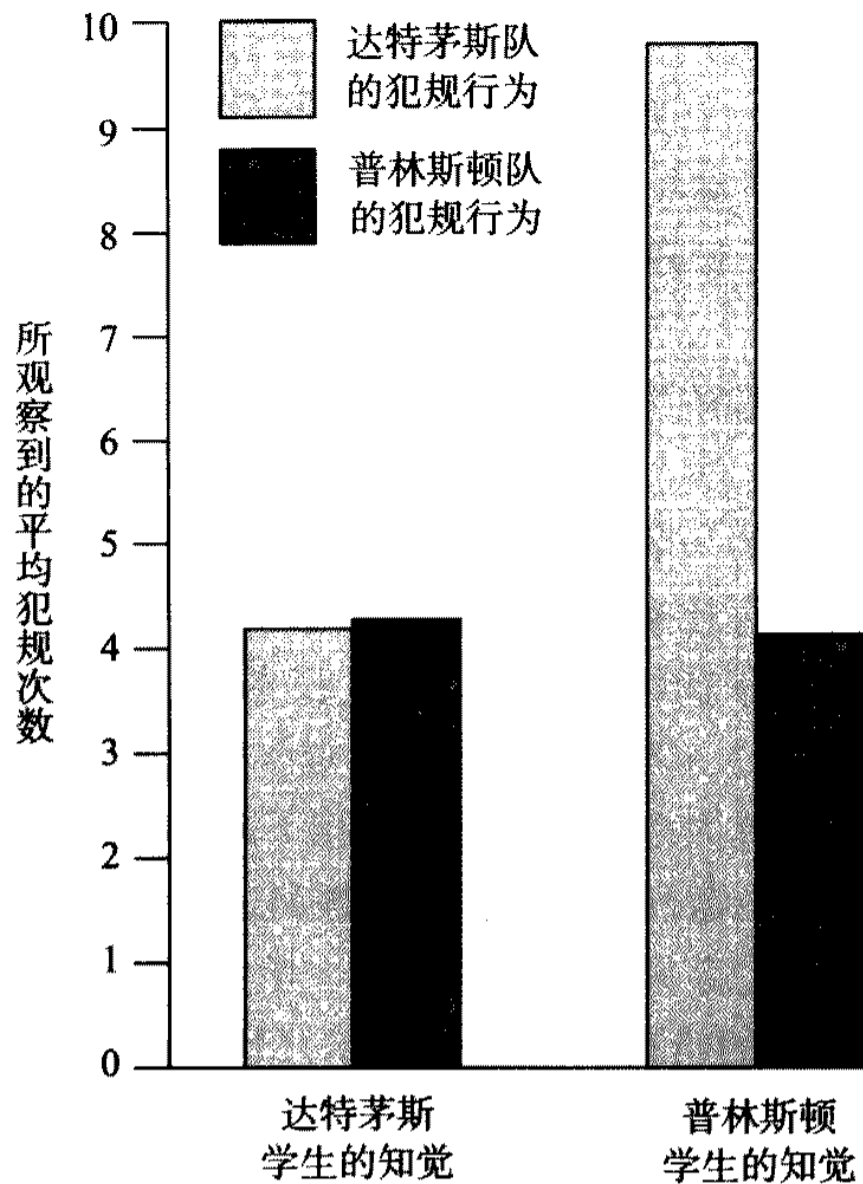


图 1.1 选择性知觉的一个例子 (引自 Hastorf & Cantril, 1954)

以这些实验结果为基础，哈斯托夫和坎特里尔（1954）得出这样一个结论：“看上去这个比赛好像是许多不同的比赛——人们常说，不同的人对于同一件事情有不同的态度。事实上，这种说法是不准确的，而且具有误导性。因为实际上对于不同的人来说，每件事情本身就是不同的，不管它是一场橄榄球比赛，还是一个总统候选人，又或是胡说八道。”1981年，约翰·洛伊和唐纳德·安德鲁（John Loy & Donald Andrews）小心地重复了哈斯托夫和坎特里尔的实验，也得出了相同的结论。

敌意媒体效应

多年以后，罗伯特·瓦伦，李·罗斯和马克·莱珀（Robert Vallone, Lee Ross & Mark Lepper）扩展了关于达特茅斯-普林斯顿争端的研究。他们在1985年提出，选择性知觉可能会导致每一党派的人都认为，大众媒体就某一事件的报道是偏向于其对立党派的。瓦伦、罗斯和莱珀称这一现象为“敌意媒体效应”。在1980年卡特对里根的总统选举中，他们第一次对此假说进行了研究。就在选举的前三天，他们对160位登记的投票人进行了采访，要求他们指出媒体对候选人的报道是否存有偏见，如果有偏见，那么这种偏见是有利的还是不利的。他们的研究发现，将近三分之一的回答者认为，媒体报道是存有偏见

的，而且他们认为在这些存有偏见的报道中，有 90% 是不利于其支持的候选人。

这些最初的发现引发了他们进一步研究的兴趣。瓦伦、罗斯和莱珀在 1985 年进行了第二次研究。他们邀请了 68 个支持以色列的大学生，27 个支持阿拉伯的大学生和 49 个持有一般看法或者说是中立的大学生，一起观看有关贝鲁特大屠杀的一组电视新闻报道片断（1982 年，一系列的阿以冲突导致了在黎巴嫩的 Sabra 和夏蒂拉对阿拉伯难民营平民的大屠杀）。这些新闻片断来自面向美国全国播出的十天内的六个不同的晚间和深夜新闻节目。

瓦伦、罗斯和莱珀发现，每一立场的人都认为新闻报道是偏向对方立场的。这支持了他们所提出的“敌意媒体效应”。支持阿拉伯的学生认为这些新闻片断普遍是偏向以色列的，而支持以色列的学生则认为这些新闻片断普遍是偏向阿拉伯的，中立学生的看法介于二者之间。另外，支持阿拉伯的学生认为，这些新闻节日本应该责备其他国家却责备了阿拉伯；支持以色列的学生则认为这些新闻节日本应该责备其他国家却责备了以色列。

和达特茅斯 - 普林斯顿事件的研究一样，瓦伦、罗斯和莱珀的研究也发现，双方的这种争论不仅仅是看法上的差异，也存在知觉上的差异。例如，就新闻报道中对以色列的评价来说，支持以色列的学生与支持阿拉伯学生所察觉到的对以色列有利和不利评价的次数是不一样的。绝大多数支持阿拉伯的学生认为，针对以色列的评价中，有 42% 是有利的，只有 26% 是不利的。而支持以色列的学生却认为，针对以色列的评价中，有 57% 是不利的，只有 16% 是有利的。另外，支持以色列的学生认为，大部分的中立者在观看这些新闻片断后会对以色列持有更负面的态度，而支持阿拉伯学生则不这么认为。

瓦伦、罗斯和莱珀的结论是，每一党派的人都会认为，媒体就某一颇有争议的报道是不公平的，且偏向他们所支持立场的对立面。他们认为，相似的知觉性偏差也存在于类似调解、仲裁或者其他情况中。在这些情况中，双方都十分忠于自己所处的立场。这种推断是有道理的。在第 2 章我们将会看到，当人们对某一特定原因或者某一特定过程做出承诺时，他们的知觉就会发生改变，从而与自己的这种承诺保持一致。

结 论

知觉的本质就是具有选择性。即便只是识别一张扑克牌——或是酒醉的感觉——也在很大程度上取决于认知和动机因素。因此，在做任何重大决策或判断之前，很值得停下来想一想并问自己一些关键的问题：我看待事物的方式是否受到了某种动机的驱使？我在看待和处理问题时是否夹杂了自身的预期？我是否与那些与我有不同预期和动机的人交换过意见？通过询问这些问题，决策者可以发现许多能够引起知觉偏差的认知和动机因素。

第 2 章

认知不协调

在开展有关选择性知觉的实验之后不久，费斯汀格（Leon Festinger）于 1957 年提出了“认知不协调”理论。自 20 世纪 50 年代以来，有关不协调理论的实验进行了许多次，其中许多是心理学领域非常有名和有趣的实验。下面内森·奥苏贝尔（Nathan Ausubel）的故事能够帮助我们理解认知不协调理论，了解这种认知不协调是如何影响人们的决策和判断的（1948；Deci, 1975, pp. 157 - 158）。

认知不协调理论的一个寓言

有一个犹太裁缝勇敢地在一个反犹太街区开了一家裁缝店。为了把他赶出这个街区，每天都有一群年轻人跑到他的店门口向他大吼：“犹太人！犹太人！”

失眠了几个晚上以后，裁缝想出了一个办法。那群人再来他店门口时，他宣布，任何称他为“犹太人”的人都将得到一毛钱。

得到激励后，第二天这群人更加高兴地跑来大叫：“犹太人！犹太人！”裁缝微笑着给了每个人五分硬币（他解释说今天只能付这么多）。这群年轻人还是很满意地离开了，毕竟五分钱也是钱。

接下来的几天，犹太裁缝就只给那群年轻人每人一分钱，并再次解释他付不起更多的钱。当然了，一分钱不再那么有激励作用，一些年轻人开始向裁缝抗议。

裁缝再次申明，他不可能付更多的钱，这些年轻人要么拿一分钱要么就离开。于是，这群年轻人决定离开，临走前还冲着裁缝大叫：“你只出一分钱，还想让我们叫你犹太人，真是疯了！”

为什么有这样的变化

为什么这群年轻人会“免费”骚扰裁缝，而当裁缝付给他们一分钱的时候他们却不愿意这样做了？认知不协调理论认为，这是因为人们往往想要减少或者避免心理上的不一致。当裁缝宣布他很乐意被称做犹太人并表示愿意为此付

钱时，他巧妙地改变了那帮年轻人称呼他为犹太人的动机，这一动机由单纯的反犹太主义变为了金钱激励。当裁缝宣布他不再付钱时，他便成功地诱发了一种心理上的不一致状态（或者说是“唤起不协调”），让这帮年轻人觉得好像是免费地奉承了裁缝一样。于是，当这群年轻人没有获得足够多的金钱时，他们开始对自己的行为和目标之间的不一致感到不满（其目标是要骚扰裁缝，而不是让他过得开心）。

无聊也可以变得有趣

相同的原理在费斯汀格和梅里尔·卡尔史密斯（Merrill Carlsmith）（1959）的实验中也得到了验证，这一实验是社会心理学领域最为著名的实验之一。60名男性斯坦福大学学生被随机分配到3种实验条件下。在“1美元”条件下，每个被试需要完成一项沉闷的实验任务，然后告诉那些在外面等待的人这是一项很有趣的实验，这样他们就可以得到1美元的奖励。在“20美元”条件下，每个被试需要做同样的事情，最后可以得到20美元的奖励。而控制组的被试则只是完成沉闷的实验任务而已，没有金钱奖励。

这项实验内容是什么呢？首先，在头半个小时之内，学生们用一只手把12个线轴放到一个盘子中去，然后倒空盘子，再把盘子装满，如此不断地重复30分钟。然后，在剩下的半个小时中，学生们用一只手转动小钉板上的48颗钉子，但一次只能转四分之一圈。实验工作人员对每个被试进行单独的观察，他们只是拿一个秒表坐在旁边，忙于在纸上做记录。

学生一完成任务，靠在椅背上的实验者就会对他说：

“我想向你解释一下这个实验，这样你就会知道为什么要你做这些事情了……实际上，有两个小组参加这个实验。在你所处的这一组中，我们会把被试带进屋子里，基本上不给他任何有关这个实验的介绍……而在另一组中，则会有一直受雇于我们的学生。我的工作就是把他带到隔壁的房间去，被试会在那里等候参加实验（就是你曾经等候过的那个房间）。我会向被试介绍，这个学生刚刚参加了我们的实验。然后这个为我们工作的学生在与下一组被试交谈的时候，会告诉他们，这是一个很令人愉快的实验，非常有趣，他自己非常喜欢这个实验，非常有趣……”

对于控制组中的被试，解释完这些以后，实验者会要求被试评价他们的实验任务是否令人愉快。而对于“1美元”条件和“20美元”条件下的被试，实验者会继续向他们解释道：

“这个平时为我们工作的学生今天不能为我们工作了——他刚刚打电话来说有其他要紧事情要做——所以我们想找一个可以受雇于我们，为我们做这件事情的人。你瞧，现在正有一个被试在那里等着参加另外一种条

件下的实验（看看表）……如果你愿意为我们做这件事，我们希望雇佣你做这份工作，而且以后如果发生类似的事情，你能够随时受雇于我们。对于你所做的事情（即现在为我们做这份工作并能够随时受雇于我们），我们会付给你1美元（或者20美元，视实验设定的条件而定）。你认为你会为我们工作吗？”

所有参加“1美元”和“20美元”实验的被试都同意被雇佣，并告诉等候在外的被试这是一个很令人愉快的实验。之后，实验者会要求他们对这项实验任务进行评价。费斯汀格和卡尔史密斯（1959）发现，在对这个实验的愉悦程度的评价中，“1美元”条件下的被试比其他两种条件下被试的评价要高得多。

费斯汀格和卡尔史密斯认为，那些只拿1美元却仍然需要对别人撒谎的被试经历了一种“认知不协调”。根据费斯汀格（1957）的理论，当人们同时有两种心理上不一致的想法（例如，在一定程度上相矛盾或是不相容的想法）时，他们就会处于认知不协调状态。在费斯汀格和卡尔史密斯的实验中，这种不协调的认知是：

1. 这项实验任务极端无聊。
2. 只是为了1美元，我（本来是一个诚实的人）就告诉其他人这是一项有趣且令人愉快的实验任务。

同时考虑这两种说法，可以看出，“1美元”条件下的被试为了一个并不怎么样的理由撒了谎（而在另一方面，“20美元”条件下的被试同意受雇则显然是为了一个他们认为非常不错的理由：20美元）。

费斯汀格在1957年提出，人们会尽一切可能减少认知不协调感。他认为，不协调是一种“消极驱力状态（一种令人厌恶的状况）”。他将这一认知不协调理论作为一种激励理论来进行阐述。根据这一理论，实验中的被试应该想要减少以上两种想法之间的不一致。

当然，对于第二种想法，被试无法改变多少。事实也的确如此，被试告诉了其他人这是一项令人愉快的任务，而且他们也确实为此得到了1美元（另外，他们肯定也不会改变他们对自己的看法，他们始终认为自己是诚实和正派的人）。而另一方面，关于这项工作是否沉闷的评价，被试却有较大的回旋余地。他们可以说，沉闷与否视当事人的感觉而定。

因此，费斯汀格和梅里尔·卡尔史密斯（1959）得出如下结论：“1美元”条件下的被试在告诉其他人这是一项很令人愉快也很有趣的实验时，产生了认知不协调感。为了降低这种不协调感，他们对实验的有趣性评价就相对要高些。而相反，“20美元”条件下的被试对这个实验的评价就更符合实情，也就是，非常之无趣。这一条件下的被试不需要刻意去降低不协调感，因为他们为自己的行为已经找到了一个很好的解释——他们得到了20美元作为“撒谎”的报酬。

自我知觉理论

故事并没有到此结束，对于费斯汀格和卡尔史密斯的实验，出现了另外一种解释。20世纪60年代中期，心理学家达里尔·贝姆（Daryl Bem）提出，认知不协调实验中的发现也可以用一种他称之为“自我知觉”的理论来解释。根据自我知觉理论，不协调实验中的发现与所谓的不协调无关；相反，它们与人们如何从所观察到的自身行为中推断自己的信念有关。

贝姆的自我知觉理论建立在以下两个假设之上：

1. 人们对自己的态度、情绪以及其他一些内在状态的认识和发现，部分是通过观察自己在各种情况下的行为来实现的。
2. 由于内部线索微弱、模糊和难以解释，人们的自我认识就与外部观察一样，需要推断得出。

从自我知觉的角度可以这样解释费斯汀格和卡尔史密斯的实验，“1美元”条件下的被试根据他们只得了1美元却对这个实验做出高度评价中推断出他们是很喜欢这一实验的（他人也会做出同样的推断）。另一方面，“20美元”条件下的被试则认为，他们的行为只是对较高的金钱激励做出了反应。自我知觉理论和认知不协调理论的差别在于，自我知觉理论是从人们如何推断行为原因的角度来解释实验结果，而认知不协调理论则是从人们天生所具有的减少内部冲突或不协调状态的动机的角度来解释实验结果。根据贝姆的理论，费斯汀格和卡尔史密斯实验中的被试即使没有经历一种不安状态，也仍然会给出同样的答案。

之后有很多实验对这两种理论进行了比较（Bem, 1972），但对于哪种理论能够更准确或者说更有用地解释这种“不协调现象”，仍然没有定论。很多年来，两种理论的研究者们都试图设计出一个能够得到最终结果的实验来验证他们所支持的理论，但是每一轮实验都只是引发了对方的新一轮实验。最后只能说，在不同的情境下，两种理论都是成立的（但是根据心理学的传统，对于两种理论都同样能够解释的实验，我们都以不协调理论的术语来指代）。

接下来的几个部分将进一步阐明，认知不协调对我们的很多决策与判断都具有重要的影响。许多引发这种不协调的情境必然是以下两种情况中的一种：决策前与决策后。对前者来说，不协调（或者说不协调产生的可能性）影响的是人们所做的决策。而对于后者，不协调（或者说不协调产生的可能性）是由已经做出的选择引起，而避免或者减少这种不协调则会影响到以后的行为。

决策前不协调的一个例子

一对父子开车外出，发生了车祸。父亲当场死亡，儿子的情况也十分危急。儿子很快被送到医院，准备实施手术抢救。当医生走进手术室见到这个病

人后，突然大叫起来：“我不能做这个手术，这是我的儿子！”

这种情况可能吗？很多人认为不可能。他们的推理是，如果这个病人的父亲已经在车祸中丧生，病人就不可能是医生的儿子。至少，在他们想到外科医生可能是病人的母亲之前，他们肯定会这样推理。

如果你一开始不知道这种可能性，而且如果你认为自己没有性别歧视，那么现在就有一个很好的机会来让你体会到不协调的感觉（请参见读者调查第16题来判断自我的性别歧视倾向）。而且，根据认知不协调理论，你会有动机去减少这种不协调，从而表现得比以前更加反对性别歧视。

1980年，吉姆·舍曼和拉里·戈尔金（Jim Sherman & Larry Gorkin）便是用这个女外科医生的故事来验证了这一假设。在这个名为“对待社会问题的态度和解决逻辑问题的能力之间的关系”的实验中，舍曼和戈尔金将大学生被试随机分配到实验的3个条件中。在“性别角色”条件下，学生有5分钟时间来回答这个有关女外科医生的题目；在“非性别角色”条件下，学生同样有5分钟时间来回答一个难度相当的有关点和线的问题；而在控制条件下，学生不需要回答任何问题。在“性别角色”和“非性别角色”条件中，实验工作人员会在5分钟后马上将答案告诉给被试（约有80%的被试在5分钟内未能解决问题）。

接下来，被试被告知这个实验已经结束，但实验者会给他们一本小册子，上面是另一本有关法律决策的实验研究（被试被事先告知，他们将参加“一组无关联的研究项目”）。他们告诉被试，另一项研究的主要调查者位于印第安纳州的南本德，在他们完成小册子上的题目后，答案将被装入信封通过邻近的邮局寄往南本德。随后，被试单独留下来完成这个有关法律决策的小册子。

实际上，这个有关法律决策的实验仍然是本次实验的一部分，也是为了收集被试有关性别歧视的信息而设的，只是设法不让他们察觉出它与先前进行的实验之间的关联。被试会读到一个反歧视行动的案例，在这个案例中，一名妇女宣称她因为性别原因而无法得到一所大学的教职。然后，被试需要指出，他们认为陪审团的裁决应该是什么，他们认为这所大学雇佣男性而非女性当教师的做法是否合理，以及他们对反歧视行动的总体评价。

舍曼和戈尔金（1980）发现，与控制条件下和“非性别角色”条件下的被试相比，“性别角色”条件下的那些没有能够回答出问题的被试更加倾向于认定这所大学存在性别歧视，而且这所大学雇佣男性而非女性当教师的行为是不合理的。另外，他们对反歧视行动政策表现出更多的支持。也就是说，在发现自己实际上有性别歧视倾向后，学生试图通过表现得更加“思想开放”来减少这种不协调（或者从自我知觉理论的角度来说，是为了表现出自己是一个非性别歧视者）。这种减少不协调的方法被称为“加固”（bolstering），并且已经被成功地运用于促进节约能源的运动之中。坎托拉、赛姆和坎贝尔（S. J. Kantola, G. J. Syme, & N. A. Campell, 1984）发现，以前的频繁用电者在被提

醒他们的频繁用电行为以及先前许下的节电承诺后，显著减少了其用电量。

有关决策前不协调的其他例子

正如安东尼·杜布（Anthony Doob）和他的同事1969年所做的实验所示，决策前不协调同样可以影响消费者行为。研究者将销售额相当的折扣店配对，一共12组，然后随机地指定每组内的一家商店以每瓶0.25美元的价格出售一种自有品牌的漱口水，而组内的另一家商店则以每瓶0.39美元的价格出售。9天后，以0.25美元价格出售的商店将价格提高至0.39美元（与组内的另一家商店售价相同）。同样的程序后来又运用到牙刷、铝箔、电灯泡和饼干上（从总体上来讲，这些物品的实验结果与漱口水的实验结果相同）。

杜布发现，一开始就以较高价格出售的商店最终售出了更多的漱口水。在12组商店中，有10组商店呈现了这样的结果，也就是说，一开始就以0.39美元价格出售漱口水的商店比一开始以0.25美元价格出售的商店售出了更多的商品。这一结果同样验证了认知不协调理论。

杜布和他的助手以消费者的“适应性水平”和减少不协调的需求来解释这个实验的结果。他们解释说：“当漱口水以0.25美元出售时，按这个价格购买商品或是注意到这个价格的消费者便倾向于认为该商品的实际价值就是0.25美元。他们对自己说，这就是一个0.25美元一瓶的漱口水。而当接下来的几个星期价格上涨到0.39美元时，这些顾客就会认为商品标价过高，而且他们不愿意再以比原来高出许多的价格来购买这一商品。”此外，不协调理论认为，当人们为一件事情付出越多，他们就越认为这件事有价值，而且越有压力为这件事情继续付出。这一原则不仅适用于购买商品，也适用于为了某一目标而做出的任何承诺和努力（另一个例子请参见阿伦森和米尔斯于1959年做的实验）。这一行为的最终结果与第21章将要提到的许多行为陷阱相类似。

决策后不协调的例子

决策后不协调是做出决策后而不是做出决策前产生的不协调感。20世纪60年代中期，罗伯特·诺克斯和詹姆斯·英克斯特（Robert Knox & James Inks-ster）对决策后不协调的问题进行了研究。他们在加拿大温哥华的赛马场采访了141名赌马的人，其中72人在过去的30秒中已经下了2美元的赌注，69人正准备在接下来的30秒内下2美元的赌注。诺克斯和英克斯特推测，刚刚已经作出某种行为承诺（下注2美元）的人，为了减少他们的决策后不协调，会比以前更加强烈地相信他们所挑选的马最终会胜出。

为了检验这一假设，诺克斯和英克斯特（1968）要求人们用7分制来评价他们所挑选马匹的获胜概率，1代表最小的获胜概率，7代表最大的获胜概率。

实验发现，那些马上就要下注的人对他们所挑选马匹的获胜概率的评价平均值为 3.48（也就是说，获胜概率与失败概率相当），而那些已经下注的人的评价平均值则为 4.81（也就是说，获胜的概率更高一些）。他们的假设得到了证实——当做出了 2 美元的承诺后，人们会更加相信他们的选择。

这一发现引出了一个有趣的问题：投票给某个候选人后你是否对这个候选人赢得选举更有信心了（请参见读者调查第 36 题）？1976 年，奥德·弗伦克尔（Oded Frenkel）和安东尼·杜布发表了有关这一问题的实验研究结果。

弗伦克尔和杜布采用了与诺克斯和英克斯特（1968）类似的基本实验程序。他们采访了刚刚投过票和马上要投票的选民。在一个实验中，他们调查的是加拿大某个省的选举，而在另一个实验中，他们调查的则是加拿大联邦选举。与诺克斯和英克斯特实验结果一样，弗伦克尔和杜布（1976）也发现：“在两次选举中，相对于投票前的选民，投票后的选民更倾向于相信他们所选的候选人就是最佳候选人，而且最有可能赢得选举。”

结 论

正如本章开头那个犹太裁缝的故事所表明的那样，认知不协调理论如果能够好好地加以运用，将会是一件强有力的武器。有关认知不协调的研究不仅丰富有趣，而且可以直接应用到许多情境中去。例如，根据杜布等人（1969）的实验结果，零售商店常常标明某种新引进的商品正处于促销期，以避免出现消费者适应性效应。同样，许多政治运动都试图赢得选民的承诺以制造决策后不协调效应（这种策略有时候被称为“登门槛法”）。在本书后面的章节，我们将讨论有关认知不协调的一些其他发现和应用。

费斯汀格的学生埃利奥特·阿伦森（Elliot Aronson）是研究认知不协调的著名教授之一，也是许多不协调实验的早期研究者（有兴趣的读者如果想了解更多有关认知不协调理论的内容，可以从了解阿伦森在 1969 年和 1972 做的实验开始）。因此，我们在这里引述阿伦森（1972）年有关认知不协调理论的一段话作为本章的结尾：

如果一个当代的马基雅维里*向一个当代的君主建议，他可能会根据有关决策后果的理论和数据提出如下建议：

* 尼科洛·马基雅维里（Niccolo Machiavelli, 1469 - 1527），文艺复兴时期意大利著名的政治思想家、政治活动家和历史学家。主张人性本恶，一旦统治者把人民当成善良者便会引来灾祸。其理论经过后人多方引申，其名字也就成为最高层次政治诈术的代名词。主要作品有《君主论》（1532），《李维史论》（1517），《战争的艺术》（1520），《佛罗伦萨史》。

1. 如果你希望一个人形成对某个事物的积极态度，那么就引导他，使他致力于得到这个事物。
2. 如果你希望一个人对某种错误行为的道德判断变得温和一些，那么就引诱他做出这种行为。相反，如果你希望一个人对某种错误行为的道德判断更加强硬，那么就引诱他，但不足以使他做出这一行为。

众所周知，态度的改变会引起行为的改变。但有关认知不协调理论的研究表明，态度的改变也可能是由于行为的改变所引起。根据认知不协调理论，为了保持一致性，人们常常不得不去改变他们的信念使之与行为一致。在第3章，我们将会看到，在许多情况下，人们会扭曲或者忘却他们原有的信念。

第3章

记忆和事后聪明式偏差

“今天是个不同寻常的日子。”

“我不明白，”艾丽丝说，“这实在太不可思议了。”

“这就是过去生活的结果，”女王和蔼地说，“一开始这总是让人感到不太适应。”

“为过去生活！”艾丽丝吃惊地重复道，“我从来没有听说过这样的事情。”

“但这有一个很大的好处，那就是一个人的记忆可以是双向的……只能回忆过去的记忆可不怎么样。”女王说道。

——路易斯·卡罗尔，《镜中世界》(*Through the Looking Glass*)

考虑下面这句话是否正确：“记忆可以被比做大脑中的储存器，我们将一些东西存进这个容器，等到需要的时候再从这个容器中取出。有时候，一些东西会从这个容器中漏掉，这时候我们就会忘记一些事情。”

你认为这句话是对还是错？（请参见读者调查中第19题你的回答）。在拉马尔（P. A. Lamal, 1979年10月）的一项实验中，大约有85%的大学生同意这种说法。但实际上，这句话存在严重的错误（此外，是否真的有东西从记忆中漏掉也是个问题）。

记忆并不是在我们过去经历发生的时候被我们存进记忆库中的拷贝。相反，记忆是在我们提取它的时候才建立起来的（Loftus, 1980; Myers, 1990）。在发生重建的这一瞬间，我们所用到的重建材料也就是那些填补缺失细节的逻辑推断、与原始记忆混合在一起的关联记忆（associated memory）以及其他一些相关信息。为了证明记忆的重建性，可以试试迈尔斯（Myers, 1990）所推荐的一个实验：闭上你的眼睛，回忆一个你经历过的愉快事情的场景。在你还没有在脑海中重现这一经历之前，请不要阅读下面的话。

你在这个场景中看见自己了吗？很多人都看见了。但如果你看见了自己，那么就表明你一定是重建了这个场景（当然，除非在原始经历中你就是在看着你自己）。

破碎的记忆

贝丝·洛夫特斯和约翰·帕尔默 (Beth Loftus & John Palmer) 在 1974 年所做的两个实验是有关记忆的重建性的最好证据之一。在第一个实验中, 45 名学生在一起观看有关一场车祸的 7 个不同的电影片断。这些片断从一些驾驶员教学电影中剪辑得到, 长度从 5 秒到 30 秒不等。

每放完一个电影片断, 学生需要回答一系列问题, 其中包括汽车的行驶速度有多快。1/5 的学生需要回答这个问题: “两车互相接触到对方的时候, 车子的行驶速度有多快?” 剩下的学生中各有五分之一的人需要回答类似的问题, 但“接触”变成了“撞到”、“碰撞”、“相撞”或者“撞碎”。

正如表 3.1 所示, 回答“撞碎”问题的学生对汽车速度的估计比回答“接触”问题的学生的估计平均高出了 9 英里。因此, 洛夫特斯和帕尔默得出结论, 问题的形式——即使只是有一个词不相同——也能够显著地影响到人们对一件事情进行记忆重组的过程。

第二个实验的结果更加引人注目。在这次实验中, 洛夫特斯和帕尔默让 150 名学生观看了一段长度为一分钟的影片, 这段影片描述了一起长度为 4 秒钟的多辆车相撞的事故。有 50 人被问到: “当汽车互相撞碎对方时, 车子的速度大约是多少?” 另有 50 人被问到: “当汽车互相撞到对方时, 车子的速度大约是多少?” 剩下的 50 人则不需要判断汽车的速度。一个星期后, 被试被重新召集, 这次他们不再看影片, 但需要回答一系列问题。洛夫特斯和帕尔默最感兴趣的问题就是, 这些学生的回忆中是否看到了这场车祸中撞碎的玻璃。

表 3.1 车子行驶的速度有多快, 当它们……?

动 词	平均估计速度
撞碎	40.8
相撞	39.3
碰撞	38.1
撞到	34.0
接触	31.8

注释: 这些是洛夫特斯和帕尔默所做研究中的平均速度估计值 (1974)。

表 3.2 你看见撞碎的玻璃了吗

回 答	实验条件		
	撞 碎	撞 到	参 照 组
是	16	7	6
否	34	43	44

注释: 引自洛夫特斯和帕尔默 (1974) 所做研究中的实验 2。共有三个实验条件, 每种条件下有 50 个被试。

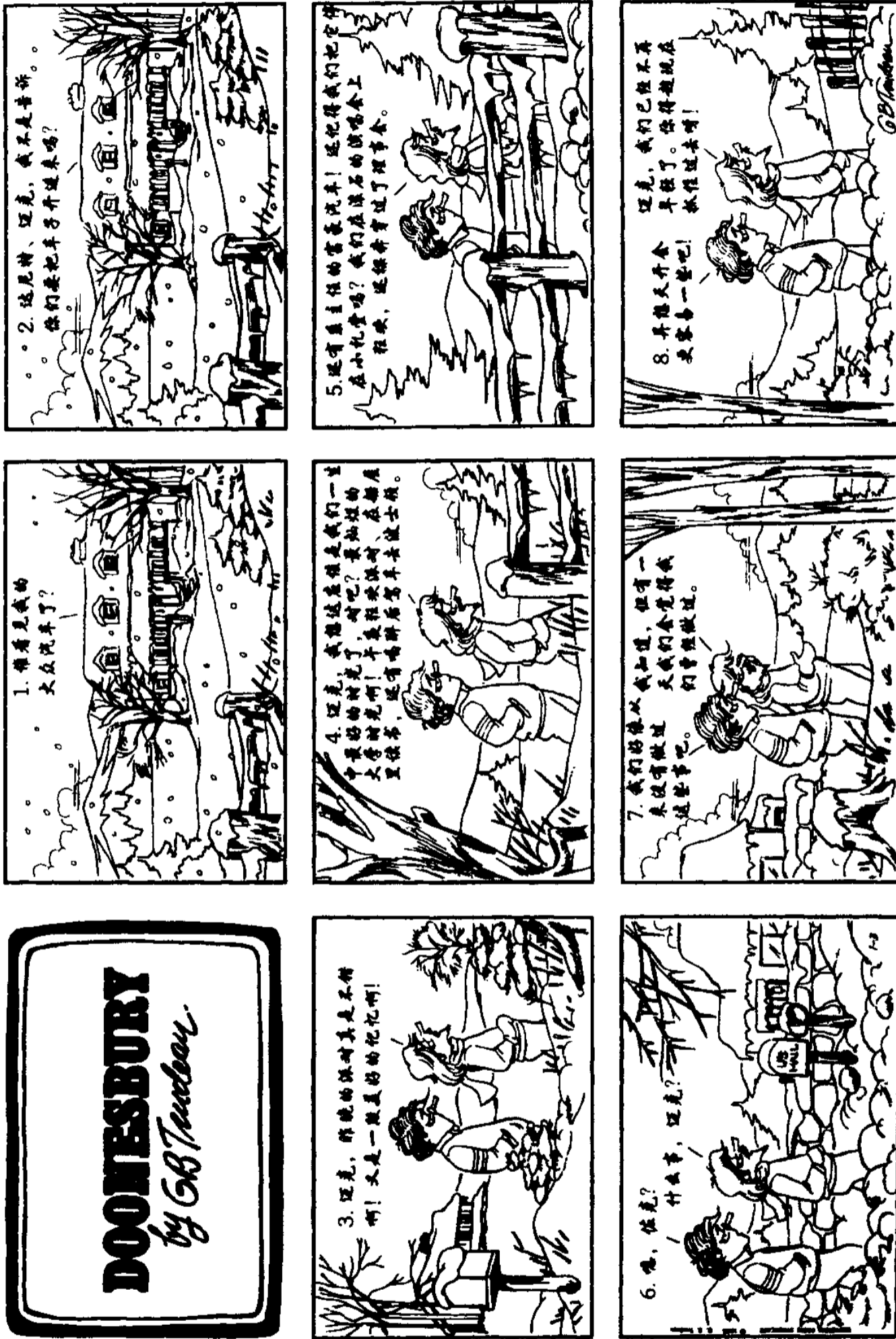


图3.1 重建的记忆(Doonisbury copyright 1992 G.B. Trudeau.Reprinted with permission of Universal Press Syndicates. All rights reserved.)

洛夫特斯和帕尔默发现，询问学生“当汽车互相撞碎对方时，车子的速度大约是多少？”不仅导致这组学生对速度的估计更高，而且导致这组学生在回忆中看到车祸中撞碎的玻璃的人数比例也更大。表 3.2 的结果显示了不同实验条件下统计意义上的差异。这一结果的有趣之处在于，这起车祸中并没有任何撞碎的玻璃——估计“撞碎”车辆速度的被试只是在重建关于这场车祸的记忆时才认为玻璃撞碎了。

甜蜜的回憶

正如洛夫特斯和帕尔默的实验所示，储存在脑海中的记忆并不是固定的。1971年，约翰·布兰斯福德和杰弗里·弗兰克斯（John Bransford & Jeffrey Franks）的研究进一步表明，每个记忆并不是与其他记忆分开来单独储存的。布兰斯福德和弗兰克斯（1971）向参加实验的大学生列出了有关一件事情的一组句子。举例来说，其中一组句子如下（见读者调查第6页）：

- 蚂蚁吃了桌子上的甜果冻
- 蚂蚁在厨房
- 蚂蚁吃了甜果冻
- 厨房里的蚂蚁吃了桌子上的果冻
- 果冻在桌子上
- 厨房里的蚂蚁吃掉了甜果冻

大约五分钟后，学生将看到另一组句子，然后指出其中哪些句子在第一组句子中出现过，并且他们需要用5点量表来表示他们对自己所选答案的确定程度。读者调查中的第34题即包含了第二组的句子，以及用来测评他们确定程度的量表。

事实上，第34题中只有一个句子在第一组句子中出现过：“蚂蚁吃了甜果冻。”而且，大多数布兰斯福德和弗兰克斯实验中的被试对这一句子确实在以前出现过持有中等程度的确定态度（确定程度为2到4）。

而更有趣的回答出现在第34题的（2）：“厨房里的蚂蚁吃了桌子上的甜果冻。”虽然这个句子并没有在第一组句子中出现过，但布兰斯福德和弗兰克斯实验中的学生却很肯定地认为他们曾经见到过这个句子。你的判断如何呢？

第34题的（2）有一个很明显的特征，就是它是全部几种关系的组合，而这些关系没有全部同时出现在第一组的任何一个单句中。原始的句子没有直接说厨房里的果冻是甜的，也没有直接说厨房里的蚂蚁吃掉了甜果冻。只有将第一组中的几个句子组合在一起，才能够得到第34题的（2）。

所以说，人们并不是简单地记住句子；他们构造并记住了一个总体的情境。一旦一部分信息与其他信息相结合，有时候就很难分清楚哪些信息是新

的，哪些信息是早已经存在于脑海中的了。

我早就知道一切

人们也很难估计，如果他们已经得知某种结果，那么他们受到这些信息的影响将会有多深。例如，如果他们已经知道了某个心理学实验的结果，他们会认为，这些结果本来就具有很高的可预测性——至少，如果在知道结果之前做判断的话，他们不会认为这些结果具有那么高的可预测性（这也是为什么本书要先进行读者调查的原因之一）。另外，即使要求人们在假设他们不知道实验结果的情况下做预测，相对于那些确实不知道结果的人来说，他们的回答也更接近实际结果一些。也就是说，如果让人们回过头来估计这些实验结果发生的概率，他们所预测的概率比那些在实验之前就进行预测的人要高一些。

这种倾向被称之为“事后聪明式偏差”（hindsight bias），或者“我早就知道一切”效应。事后聪明式偏差是这样一种倾向，它将已经发生的事情视为相对不可避免和显而易见的事情，却忽略了自己的判断实际上已经受到已知结果的影响。许多实验发现，事后聪明式偏差广泛存在于选举（Leary, 1982; Synodinos, 1981）、医疗诊断（Arkes, Wortmann, Saville, & Harkness, 1981）、怀孕测试（Pennington, Rutter, McKenna, & Morley, 1980）、购买决策（Walster, 1967）、比赛（Leary, 1981），以及其他一些领域中。这些实验运用了不同的实验技巧、指导语和被试人群，得到了同样的结果（有关事后聪明式偏差及相关效应的总结，参见 Campbell & Tesser, 1983; Christensen - Szalanski & William, 1991; Hawkins & Hastie, 1990; Verplanken & Pieters, 1988）。

巴鲁克·菲施霍夫和鲁思·贝思（Baruch Fischhoff & Ruth Beyth）于1975年发表了有关事后聪明式偏差的第一个研究。菲施霍夫和贝思（1975）的实验中所采用的主要事件是1972年尼克松对中国和前苏联的访问。实验的第一阶段，是在访问进行前，要求几组以色列学生对尼克松访问中国和前苏联的15种不同结果的概率做出估计。例如，对于中国之行，被试需要估计以下事件的概率：美国是否会在北京建立大使馆，尼克松总统是否至少会与毛泽东会晤一次，尼克松总统是否会宣布这是一次成功的访问，等等。同样，对于前苏联之行，被试则需要估计诸如美国与前苏联建立太空联合发展计划、被捕的前苏联犹太人是否有可能与尼克松会谈等事件的可能性。

实验的第二阶段是在访问进行后的两周至六个月内，让学生重新回忆他们以前所做过的预测。例如，回答过有关中国之行问题的学生被告知：“你应该还记得吧，在两周前，你曾经回答过一份有关尼克松总统访华的问卷，并对这次访问的一些可能结果做出了概率估计。人们的预测质量与其对预测的记忆能力之间的关系，令我们很感兴趣。因此，我们希望你能再做一次问卷。这份问卷与两周前你所做过的问卷一模一样，希望你能够回忆一下上次所做过的概

率预测。如果你记不住上次是怎么做的概率预测，那么就假设你现在是在尼克松总统访华之前，要对各种可能的结果做出概率预测。”

同时也要求学生指出，就他们所知，哪些可能的结果确实实现了。菲施霍夫和贝思希望知道，学生们所回忆出来的预测是否比他们当时实际所做的预测更加精确。

大体上讲，菲施霍夫和贝思（1975）的实验结果如下：对于那些他们认为确实已经实现了的结果，3/4的学生回忆出来的预测概率比他们当时实际所做的预测概率要高；而对于那些他们认为确实没有实现的结果，大部分学生回忆出来的预测概率比他们当时实际所做的预测概率要低。如果初次预测比回忆预测早几个月，那么事后聪明式偏差就会非常显著。当初次预测和回忆预测相隔3~6个月时，84%的学生显示出事后聪明式偏差——得知了尼克松访问的结果之后，他们发现这些本身难以预测的结果都具有了相当高的预测性。

减少事后聪明式偏差

1977年，保罗·斯洛维奇（Paul Slovic）和巴鲁克·菲施霍夫（Baruch Fischhoff）发表了他们的实验成果，这一实验研究了如何在已知实验结果的情况下（也就是一种“早就知道了结果”的感觉）减少事后聪明式偏差的问题。斯洛维奇和菲施霍夫发现，如果人们开始考虑有哪些原因可能导致实验出现不同结果，事后聪明式偏差就消失了。

在这一研究中，被试首先要阅读四篇有关生物学、心理学和气象学研究的简要介绍。“预见”条件下的被试被告知，这些研究马上就要开始，而“后见”条件下的被试则被告知，这些研究已经完成了。在读完每个研究的摘要后，被试需要估计第二次实验能够得到第一次实验的结果的概率（每一实验都可能出现两种结果）。也就是说，“后见”条件下的被试被告知，某种特定的结果已经得到验证，而“预见”条件下的被试则需要假设这种结果已经得到验证。

斯洛维奇和菲施霍夫（1977）发现，相对于“预见”条件下的被试，“后见”条件下的被试对于第二次实验是否能够重复第一次实验的结果给出了更高的概率估计。但是，如果要求“后见”条件下的被试考虑每种实验结果出现的原因，两组被试之间的概率估计差异就会大幅度地减小。这一研究表明，虽然事后聪明式偏差仍然存在，但程度却大大减轻了。

概括起来说，这个故事的寓意如下：如果你希望减少事后聪明式偏差，那么你应该仔细考虑，过去的事件在什么条件下可能会出现不同的结果。如果你只考虑为什么事件会出现这样的结果，很可能你就会高估这一结果出现的必然性以及类似结果在未来出现的可能性。实际上，菲施霍夫（1977）发现，如果只是单纯地告诫人们存在事后聪明式偏差并鼓励他们避免出现这种偏差，还不足以消除这种偏差。要避免受到事后聪明式偏差的影响，很重要的一点就是要

考虑另一种结果可能在什么情况下出现。

结 论

伊恩·亨特（Ian Hunter, 1964）在《记忆》一书中讲述了这样一个故事：有两个英国心理学家，他们秘密地记录下了剑桥心理学会的一次讨论。两周以后，这两位心理学家联系那些参与过讨论的人，请求他们写下他们所能够回忆起的任何关于这次讨论的内容。在将这些记录与原始记录相核对时，两位心理学家发现，回答者省略了讨论过程中90%以上的细节内容。另外，在这些能够记起的观点中，有半数观点都是错误的。回答者还回忆出一些原来并不存在的意见，并将一些随意的谈话说成是冗长的演说，或是将那些曾经是隐含的意思说成是明确的意见。

这个故事突出了做好准确记录的价值。就算是那些最富有经验的决策者也很可能存在记忆偏差。除了做好仔细的记录并保管好有关过去事件（如回忆、重要谈话、协议等）的记录之外，再没有其他什么更好的方法能够避免出现这些记忆偏差了。正如本章所引用的研究所示，记忆从本质上来讲具有重建性，并高度取决于情境因素。第4章将进一步探讨情境对决策和判断的影响。

第 4 章

情境依赖性

正如前 3 章所示，一种刺激所产生的作用在很大程度上取决于其所处的情境。也就是说，决策者并不是孤立地去感知和记忆某个事件，而是根据他们过去的经验和事件发生时的情境去理解和解释新信息。在不同情况下，同一个人对同一刺激（如某种人格特征）的认知可能会完全不同。情境依赖性可以解释许多知觉错觉产生的原因（见图 4.1）。在决策与判断领域，情境依赖性主要有四种表现方式：对比效应、初始效应、近因效应和晕轮效应。

对比效应

这是一个在你自己身上（或者在你朋友身上）就可以进行的简单实验。全部所需的工具只是三大碗水。在第一个碗中倒入热水，第二个碗中倒入温水，第三个碗中则是冰水。接下来，将一只手浸入热水中，而另一只手浸入冰水中，保持 30 秒。一旦你的手适应了水中的温度，迅速将浸入热水中的手浸入温水中，5 秒钟以后将浸入冰水中的手也浸入温水中。

和大多数人一样，你可能会有一种奇怪的感觉。最开始浸在热水中的手会让你觉得温水是凉的，而最开始浸在冰水中的手则会让你觉得温水是热的。实际上，如果你让一位朋友重复这一实验，事先不告诉他温水的温度，他也很可能难以辨别温水的温度。因为两只手同时经历了一种对比效应，但发生在每只手上的效应的方向却恰好是相反的！

许多早期心理学研究常常以这种知觉性判断作为主要研究对象，如温度识别、颜色识别和重量估计等。这些早期心理学实验研究的结果证实了对比效应的存在。例如，穆扎夫·谢里夫，丹尼尔·陶布和卡尔·霍夫兰（Muzaffer Sherif, Daniel Taub & Carl Hovland）在 1958 年就发表了一篇有关重量判断的颇有影响的文章。谢里夫、陶布和霍夫兰发现，如果被试最开始举起的是一件很重的物品，当他们稍后举一件较轻的物品时，他们对这件物品的重量判断要低于物品的实际重量。

斯坦利·科伦和乔尔·米勒（Stanley Coren & Joel Miller）在 1974 年所发表的研究是有关对比效应的最有趣的研究之一。科伦和米勒注意到，当一位 5.1 英尺高的体育播音员采访一群篮球运动员时，他显得非常矮，而当他采访一群

赛马手时却显得非常高。但是，当这位播音员站在一匹高大的赛马旁边或是一个露天体育场旁边时，他看起来却并没有变矮。

科伦和米勒从这一观察结果中推测，只有当用于对比的刺激物彼此相似时，对比效应才存在。为了验证这一假设，他们让12名志愿者观察图4.2中四个象限的图像。图中左上角象限的图像是著名的艾宾浩斯错觉的一个例子。艾宾浩斯错觉指的是，两个同样大小的圆圈，一个圆圈被比它本身更小的圆圈围绕，另一个圆圈则被比它本身更大的圆圈围绕，前者会比后者显得更大一些。

然而，科伦和米勒发现，当围绕物的形状越来越不像圆圈时，艾宾浩斯错觉会逐渐消失（正如实验中被试的判断一样）。当围绕物由圆圈变成六边形时，艾宾浩斯错觉变弱了；当围绕物变成三角形时，艾宾浩斯错觉就更弱；而最后围绕物变成不规则的多边形时，艾宾浩斯错觉降低至最小程度。通过比较图4.2中的四个象限，不难发现，即使是简单的大小判断也在很大程度上取决于情境的变化。

初始效应

所罗门·阿施（Solomon Asch）于1946年发表的研究是关于情境依赖性的一个经典研究。阿施以其在“从众性”（conformity）方面的研究最为著名（将在第17章进行讨论）。另外，他也是“印象形成”（impression formation）研究的先驱。在有关“印象形成”的大多数实验中，阿施都会让被试来描述他们对某个人的印象，这个人假想的，而且具有某些特定的特征。他在文章中提到了10个运用这种范式进行的实验，但现在我们只探讨其中的一个。

在这一实验中，阿施要求被试描述他们对一个人的印象（这个人的特征正如读者调查的第3题所述）。一半的被试所需要描述的对象是一个“嫉妒心强、顽固、挑剔、冲动，但勤勉、聪明”的人；另外一半的被试要描述的对象是一个具有相同特征的人，但这些特征的描述顺序改为“聪明、勤勉，但冲动、挑剔、顽固，而且嫉妒心强”。

阿施发现，在这些描述特征中，顺序靠前的特征比顺序靠后的特征对人们印象形成的影响更大一些。这种情况被称为“初始效应”。如果你在做读者调查中的第3题时，认为约翰是一个高度情绪化的人，那就表明你已经受到了初始效应的影响——有关情绪方面的特征，即嫉妒心强、顽固，如果被排在前面，那么相对于它们被排在后面的时候来说，对你印象形成的影响会更大一些。

如果“聪明”排在“嫉妒心强”和“顽固”前面，同样会产生初始效应吗？根据诺曼·安德森1965年的实验结果，很可能还是会产生这样的效应。安德森发现，初始效应并不意味着，只有最先出现在序列内才能导致初始效应的产生。事实上，初始效应是有关进入位置及其对判断的影响的一个总体描述。也就是说，第一印象固然是最重要的印象，第二和第三印象也同样显示出

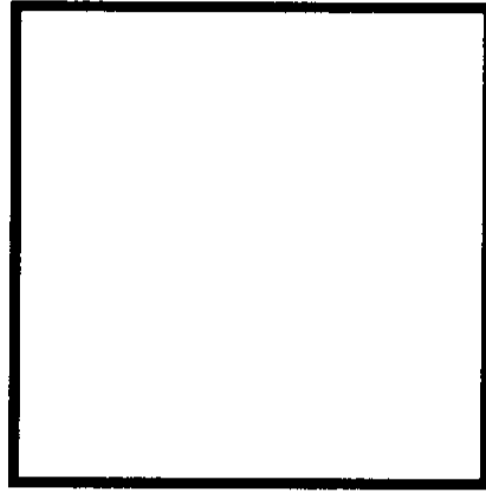


图 4.1 为了检验视知觉的情境效应，请将这一页对着光源看。多数人认为背面的正方形边框和两条平行线不再平行。(Adapted from Black and Yaker, 1989.)

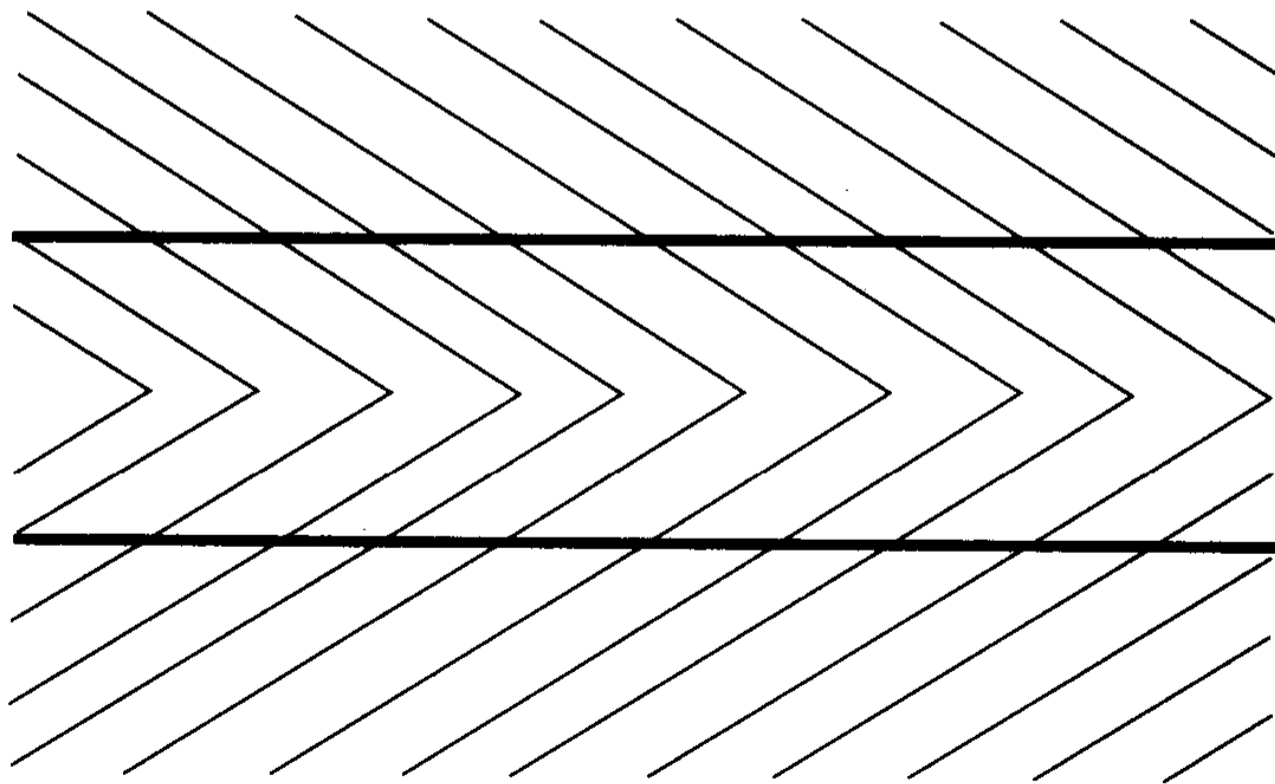
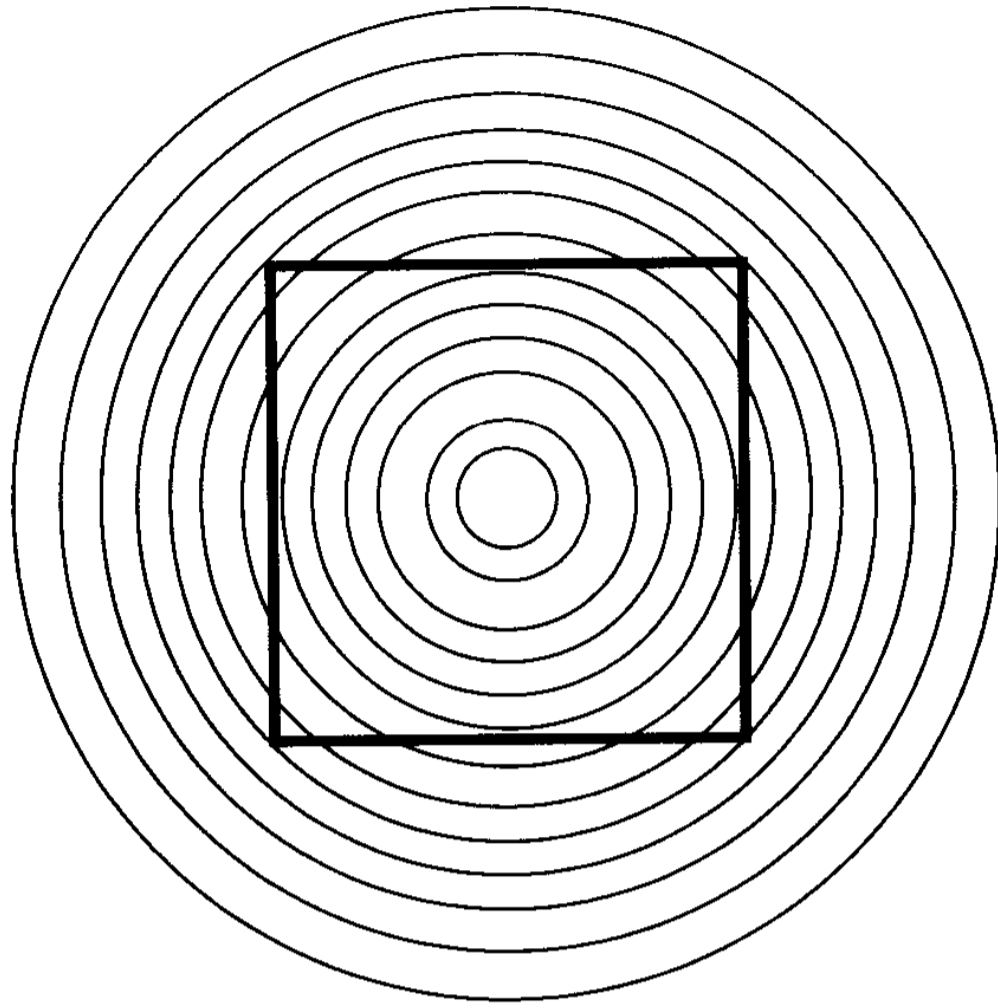


图 4.1 (续)

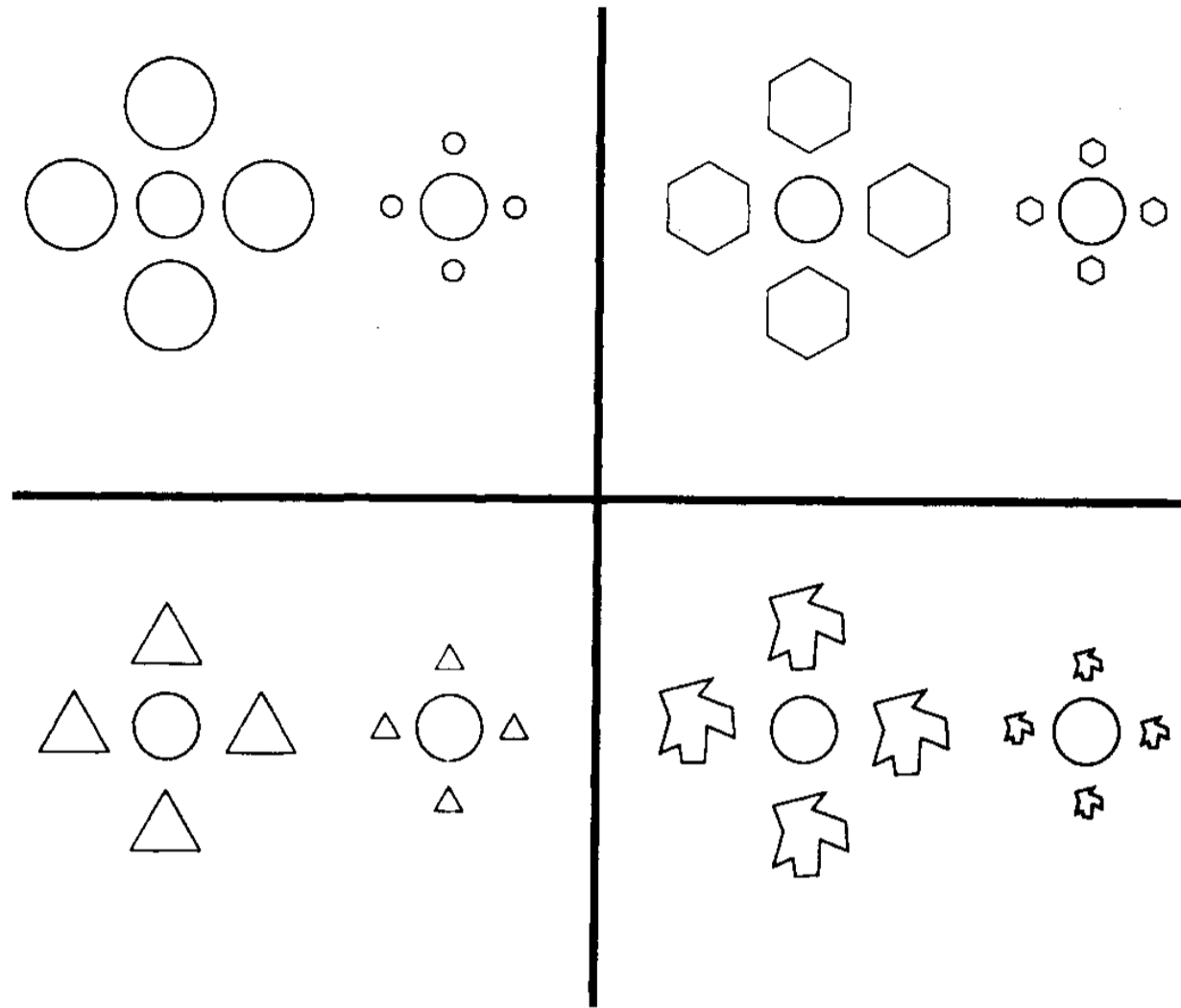


图 4.2 图中左上角象限是最为典型的艾宾浩斯错觉，科伦和米勒对此进行了一定的变换，以此阐明对比效应的作用方式。（引自 Coren & Miller, 1974）

很大的初始效应。

近因效应

初始效应不仅作用于人们的印象形成过程。在许多需要对序列信息做出评价的情况中，它都会发生作用。例如，当人们面临一个有争议问题的对立面信息时，初始效应使人们往往更容易受到先出现的信息的影响。

但事情也并不总是如此。在一些情况下，最后出现的信息会比最先出现的信息的影响更大。这种模式被称为“近因效应”。当人们能够更加清楚地记得最后出现的信息而不是最先出现的信息时，近因效应便起作用了。

一个有趣的问题是，这两种效应中谁的影响更强一些。比如，假设你参加了一场公开辩论，而且假设你可以选择最先发言或者最后发言。你会选择哪一个？如果你最先发言，你的发言可能会产生初始效应，而如果你最后发言，你的发言则可能产生一种近因效应。哪种选择才是最佳的？

诺曼·米勒和唐纳德·坎贝尔（Norman Miller & Donald Campbell）于 1959 年对这一问题进行了研究。他们对一次法庭审理的记录进行了编辑，该案件中的当事人宣称他所受的伤害是由一个有产品缺陷的蒸馏器所造成。案件的审理过程按下列方式重新编排：有利于原告的材料出现在一个部分，而有利于被告的材料则出现在另一部分。也就是说，有利陈述包括了原告方证人的证词，原

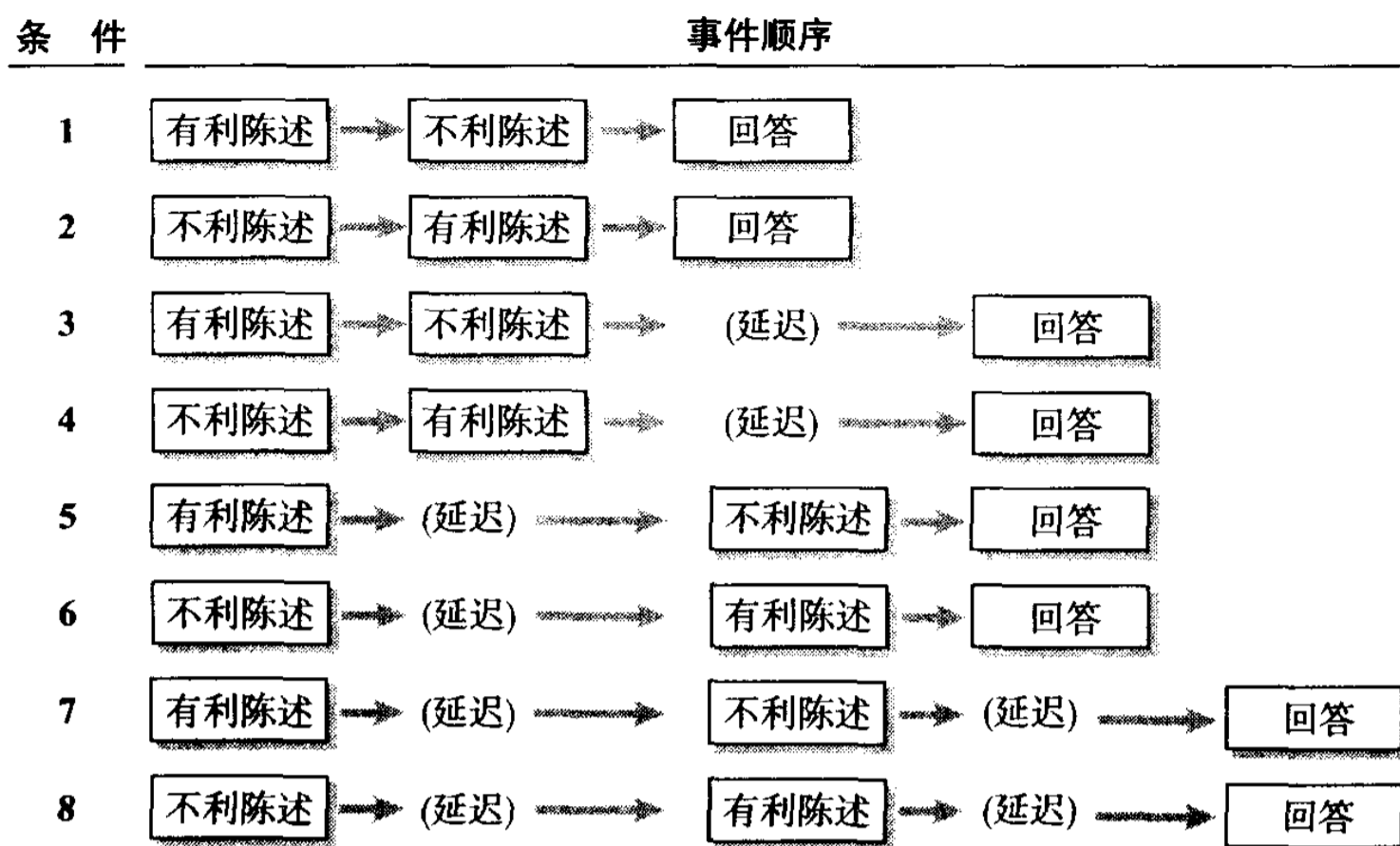


图 4.3 诺曼·米勒和唐纳德·坎贝尔（1959）比较了 8 种实验条件，考察了近因效应的作用形式。

告律师对被告方证人的诘问证词，以及原告律师的开案和结案陈词。不利陈述则包括了被告方证人的证词，被告律师对原告方证人的诘问证词，以及被告律师的开案和结案陈词。

米勒和坎贝尔对记录进行编辑后，他们记录了不同人阅读不同角色（律师、证人等等）陈述的过程。每一记录持续大约 45 分钟，陈述按 8 种不同的顺序编排（见图 4.3）。在一些实验条件下，被试需要在听完连贯的陈述（即有利 - 不利，或者不利 - 有利）后马上做出判断。而在另一些实验条件下，实验被分成几个阶段，中间间隔一个星期。

米勒和坎贝尔发现，在一些情况下会出现初始效应，而在另一些情况下则出现了近因效应。也就是说，在一些情况下，人们受到第一部分陈述的影响更多一些，而在另一些情况下，人们受到第二部分陈述的影响更多一些（不管这种陈述是有利的还是不利的）。如果被试听完连贯的陈述后，间隔一个星期再作出回答（如条件 3 和条件 4 所示），就会出现初始效应。而如果两部分陈述之间存在延迟（如条件 5 和条件 6 所示），就会出现近因效应。而条件 1、条件 2、条件 7 和条件 8 则没有出现明显的初始效应或者近因效应。

米勒和坎贝尔（1959）同时也发现，近因效应是由记忆的差别引起的。在一个多选测试中，条件 5 下的被试往往能够记得更多的有关不利陈述的事实，而条件 6 下的被试则往往能够记得更多的有关有利陈述的事实。

因此，再回过头来看我们最开始提出的问题：假设你可以在一次公开辩论中选择最先发言或最后发言。如果对方会立即回应你，而且在辩论结束后要延迟一会儿人们才会作出反应，你应该选择最先发言。比如说，如果你所辩论的

问题将在一周后进行投票表决，你应该选择最先发言。而如果双方各自的发言之间有一定的时间间隔，而且人们在双方发言结束后马上要作出反应，你应该有效地利用近因效应，选择最后发言。

斯蒂芬·霍克（Stephen Hoch, 1984）所做的一些有关人们决策方式的实验也得到了与以上所述基本相同的结果。霍克要求人们列出为什么某件事将来可能会发生的原因（“有利的原因”），以及为什么这件事将来可能不会发生的原因（“不利的原因”）。对一些事件，人们先列出有利的原因，而对另一些事件，人们则先列出不利的原因。霍克发现，如果人们连续列出有利原因和不利原因，就会出现很强的初始效应。但如果人们在列出有利原因和不利原因之间有3分钟的间隔，就会出现近因效应。这些发现与我们长久以来使用的销售技巧的实质是一致的。销售者常常会鼓励顾客列出购买原因（有利原因）和不购买的原因（不利原因）。如果在列出各自原因中间没有间隔，顾客就可能在无意间受到初始效应的影响。

晕轮效应

对比效应、初始效应和近因效应都说明，当情境不同或刺激出现的顺序不同时，刺激的影响是不同的。另一种情境依赖性效应被称为“晕轮效应”（halo effect），这是桑代克在1920年命名的。桑代克发现，在军队主管评价他们下属军官的智力、体能、领导能力和性格时，有关这些不同方面的评价通常是高度相关的。按照桑代克的说法，当飞行小队的指挥官对其下属的空军学员做出评价时，有关智力和体能的评价呈现出0.51的相关性，有关智力和领导能力的评价呈现出0.58的相关性，有关智力和性格的评价则呈现出0.64的相关性。^{*} 桑代克也发现，不同教师评价之间的正相关被用来确定薪酬和提升。例如，作为教师的一般优点与其外表、健康、守时、才智、诚实、真挚方面的评价有很强的相关性，而对教师声音的评价则与其在智力、“对社区事宜的兴趣”方面的评价呈现出很强的相关性。

在其最初的有关晕轮效应的文章中，桑代克（1922, pp. 28 - 29）得出如下结论：“即便是一个很有能力的工头、雇主、教师或者部门负责人，也很难把别人当成一个不同品质的集合体来看，很难对他的每一种品质作出独立的评价。”如今，我们知道桑代克的发现部分是由于技术原因导致的（如何设计评价量表），但他的总体理念却经受住了时间的考验。因为即便是运用复杂的测评技术，晕轮效应仍然会发生作用（Cooper, 1981; Feldman, 1986）。

* 相关系数的范围通常为+1.00和-1.00之间。正相关是指一个变量会随着另一个变量的增加而增加，负相关是指一个变量会随着另一个变量的增加而减少，0相关意味着变量间不存在任何关联。因而，在桑代克所举的例子中，智力与性格存在较高的相关。

自桑代克的研究开始，学术界进行了不少有关晕轮效应的实验。例如，有关“美丽晕轮效应”的实验显示，相对于那些长相欠佳或是长相一般的人来说，相貌有吸引力的人往往被认为更快乐，拥有更好的职业，更容易结婚，而且在性格特征方面也更加讨人喜欢（Dion, Berscheid, & Walster, 1972）。戴维·兰迪和哈罗德·西加尔（David Landy & Harold Sigall, 1974）也发现，相对于长相一般或者长相欠佳的作者来说，相貌有吸引力的作者的文章通常会得到更高的评价。

与对比效应的研究一样，晕轮效应的许多先驱性实验也是由阿施（1946）进行的。例如，在一个实验中，阿施让一半的被试描述他们对一个人的印象，这个人“聪明、伶俐、勤勉、热心、坚强、讲求实际而且谨慎”。同样，他让另一半的被试描述他们对另一个人的印象，这个人“聪明、伶俐、勤勉、冷漠、坚强、讲求实际而且谨慎”。也就是说，两组被试所听到的描述基本相同，只是这个人在第一组中被描述为热心，而在第二组中被描述为冷漠。

然后被试会看到几组描述人的特征的词组（这些词组大多为反义词构成），他们需要指出词组中的哪个词与他们所形成的印象最为一致。这些词组包括：慷慨大方/胸襟狭窄，不快乐/快乐，脾气暴躁/和善，以及幽默/不幽默。与早先桑代克的发现一致，阿施也发现，在描述中引入一个中心特征，如热心或者冷漠，能够影响被试对这个人的整体印象。比如说，在需要描述他们对“热心”的人的印象的被试中，有75%~95%的人认为这个人同样应该是一个慷慨大方、快乐、和善和幽默的人（你可以检查自己在读者调查中第4题的答案，与这些被试的回答相比较）。与之形成对照的是，在需要描述他们对“冷漠”的人的印象的被试中，则只有5%~35%的人认为这个人是个慷慨大方、快乐、和善和幽默的人。

这些结果不仅仅只是简单的纸笔测试的副产品。哈罗德·凯利（Harold Kelley, 1950）发表的研究中，利用真人代替特征描述对晕轮效应进行了研究。他同样发现了若干年前阿施实验所观察到的晕轮效应。实验中，一些学生受到引导，从而相信他们的老师是一个热心人。这些学生认为这名老师不仅考虑周到、脾气好、好交际，而且受欢迎、幽默、仁慈。而且，如果他们相信这是名热心的老师，这些学生便倾向于与老师交流更多。“热心”条件下的被试学生中有56%参与了课堂讨论，与之形成对照的是，“冷漠”条件下的被试学生中只有32%参与了课堂讨论。这些结果显示，晕轮效应确实会对人们的社会交往产生作用，并会影响随后的行为。

与对比效应、初始效应和近因效应一样，晕轮效应也显示出人们对某种刺激的反应方式具有情境依赖性。确实，不依赖于情境而起作用的刺激物是不存在的。情境因素对人们的反应方式有很深的影响——不管这种刺激物是几何图形、性格特征、法律论据，还是胡说八道。

结 论

情境效应处处可见，以至于人们有时候反而会忽略它们的存在。实际上，很难想像一个知觉独立于情境之外而存在的世界。这样的世界会是什么样子，与情境无关的判断有什么意义？

在我主持的一次研讨会上，一个学生指出，情境效应对人们社会关系的性质提出了意义深远的问题。比如，对比效应说明，在绝对条件下去了解一个人是不可能的。当你判断一个朋友的诚实程度时，这一判断是相对的，在一定程度上取决于你所认识的其他人的诚实程度。根据对比效应理论，当你得知你所认识的其他人欺骗了你的时候，你便会发现，这位朋友原来是那么的诚实——即使你朋友的行为并没有发生变化。你对你某个朋友的诚实程度的判断好像与其他人的诚实程度无关，但正如这一章节所描述过的研究所示，社会判断在一定程度上取决于情境的作用。

即便是快乐也具有情境依赖性。菲利普·布里克曼等人（Philip Brickman, Dan Coates & Ronnie Janoff - Bulman, 1978）在对伊利诺伊州彩票中奖者的研究中便发现了这一点。彩票中奖者的报告显示，与那些处于相同地理位置的非中奖者参照群体相比，他们在很多日常生活行为中感受到的乐趣更少一些，这些行为包括看电视、与朋友交谈、吃早餐、买衣服，等等。布里克曼等人解释说，这种差别出现的原因，部分是由日常行为与中彩票后的兴奋所形成的对比效应导致的。

那些职业的说客很早就认识到了对比效应的力量。比如说，地产经纪在劝说买家买下其想要出售的房屋之前，常常会让买家看一所破旧或者是标价过高的房子，以利用二者之间的对比效应。政党候选人经常宣扬他们过去的一两个成功事例，以利用晕轮效应向选民们展示自己所具备的各方面能力。广告商则费尽心机地为他们所促销的产品制造出吸引人的情境。

当然，情境效应的作用并不是无限的。不管其他人是否撒谎，一个惯于说谎的人看起来总不会是诚实的，而一个不合格的政客也只能有限度地利用晕轮效应。一个以同心圆为背景的正方形，它的边看起来可能会是弯曲的，但这并不能真地把一个正方形变成圆形。但即便如此，任何有关决策与判断的有意义的分析也都应该考虑情境效应的影响。另外，一些研究者认为，通过实验展示的情境效应实际上低估了日常生活中的情境效应（Hershey & Schoemaker, 1980）。

第二部分

问题如何影响答案

本书的第一部分探讨了刺激的情境影响人们的知觉方式。在这一部分中，我们扩展了这一命题，探究问题的形式对人们回答的影响。第5章和第6章主要解决以下两个问题：（1）人们的决策与判断在多大程度上取决于问题的措辞？（2）问题的措辞以哪些具体方式影响人们的回答？

可塑性

任何刺激的发生总离不开一定的情境，问题也一样。本章将讨论问题的语境和措辞对决策与判断的影响。

在一些情况下，在两种不同语境下提供同一个选项会得到不同的回答。例如，假设你面临如下的选择：

选择 A： 100% 的概率输掉 50 美元。

选择 B： 25% 的概率输掉 200 美元，75% 的概率什么也不输掉。

你会如何选择？（参见你在读者调查中第 2 题的答案。）在保罗·斯洛维奇等人（Paul Slovic, Baruch Fischhoff, & Sarah Lichtenstein, 1982）的实验中，80% 的人会选择 B。当面临损失时，很多人都属于“风险偏好”的，他们往往愿意冒险去承受更大的损失（在这个例子中为 200 美元），也不愿意承受具有相同期望值但却是确定的损失（25% 的概率输掉 200 美元与 100% 的输掉 50 美元具有相同的期望值，因为从长期来看，二者具有相同的期望收益）。

然而，稍微想一想就会发现其中的问题。如果人们在面临损失时更愿意承受风险，那么保险公司就没有什么生意可做了。保险行业的运营基础就在于，人们愿意承担一个确定的损失（保险费）去避免更大但不确定的损失。以保险业的术语来讨论确定的损失时，人们是否会表现得不同？如果这个确定的 50 美元的损失被说成是能够用来避免潜在的 200 美元损失的保险费，人们会如何在 A 和 B 中做出选择？

斯洛维奇，菲施霍夫和利希滕斯坦发现，在这种情况下，65% 的人更倾向于承受确定的 50 美元损失。这可能是因为保险费着重突出了潜在的更大损失，也可能是因为保险公司成功地塑造了一种社会规范，在这种规范之下，人们表现得更加审慎，因此更倾向于支付保费而不是承受更大的损失。不管是哪种原因，当以保险业的术语来重述这一问题时，同样的问题确实导致了不同的选择结果。如果确定的损失是以保险费的形式来表达，大多数的人会变得“风险规避”而不是“风险偏好”；他们倾向于承担确定的损失而不是冒损失更多的风险。

斯洛维奇，菲施霍夫和利希滕斯坦（1982）也发现，当面临付出 5 美元还是以千分之一的概率输掉 5 000 美元的选择时，人们会表现出相似的偏好逆转。

表 5.1 你会做什么样的选择

风险选择方案			赌博条件下， 偏好确定损失 的百分比	有保险时，偏好 确定损失的百分 比
概 率	可能的损失（美元）	确定的损失（美元）		
0.001	10 000	10	54	81
0.01	10 000	100	46	66
0.01	100 000	1 000	37	76
0.1	10 000	1 000	29	59

注释：从右边的两列可以看出，保险条件下的人数百分比要比赌博条件下高出 20% 到 40% 以上。数据引自赫尔希和休梅克（1980）的研究。

虽然在简单偏好的语境下，只有大约 2/5 的人会选择付出 5 美元，但如果将 5 美元说成是保险费时，就会有大约 2/3 的人选择确定的损失。约翰·赫尔希和保罗·休梅克（John Hershey & Paul Schoemaker, 1980）的实验也得到了同样的结果（见表 5.1）。

顺序效应

人们的答案也可能受到问题或者选项顺序的影响。这些影响通常很小，但在某些情况下，其影响却可能变得非常深远。比如说，如果两个问题都是关于同一主题，而且回答者需要表现得一致，那么回答者就会照着前一个问题的答案的思路来组织后一个问题的答案。

在霍华德·舒曼和斯坦利·普雷瑟（Howard Schuman & Stanley Presser）《态度调查中的问题与答案》（1981）一书中，他们以一项有关新闻自由的态度调查的结果来阐明了这一效应。舒曼和普雷瑟随机选取一组美国成年人作为实验样本，向他们询问了以下两个问题：

1. 你认为一个共产主义国家，比如前苏联，是否应该允许美国的新闻记者进入其境内采访，并将他们所得到的新闻发回美国？
2. 你认为美国是否应该允许其他共产主义国家的新闻记者入境采访，并将他们所得到的新闻发回他们所在的报社？

一半的被试会按照上面这种顺序来回答问题，而另一半的被试则以相反的顺序来回答问题。

舒曼和普雷瑟发现，如果被试被问到的第一个问题是有关美国的新闻记者，82% 的人认为美国记者应该被允许自由进入共产主义国家。与这一答案一致的是，75% 的人也认为美国应该允许共产主义国家的记者自由进入美国采访。另一方面，如果被试被问到的第一个问题是有关共产主义国家的新闻记

者，那么只有 55% 的人认为应该让共产主义国家的新闻记者自由进入美国进行新闻报道。在这种情况下，大概是为了保持一致，避免让自己看起来有双重标准，只有 64% 的人认为应该允许美国记者自由进入共产主义国家（相较于 82%，64% 的比例更接近于 55%）。所以说，相同问题的不同顺序会产生截然不同的答案。

选项的顺序不同也能够影响到人们的回答。选项顺序效应通常比较微弱，而且如果问题的答案只是简单的二选一（例如“同意”和“不同意”），或者答案选项非常多，那么选项顺序效应基本是不存在的。顺序效应中最常见的是近因效应，也就是说，当某一答案选项出现在备选答案中的最后时，其被选择的频率更高。

舒曼和普雷瑟（1981）发现，在有关离婚问题的两种不同表达方式中，出现了较强的近因效应。在一次全国性的民意调查中，一半的调查对象被问到：“你认为在这个国家里，离婚程序应该更容易、更难、还是维持现状？”

在对这一问题的回答中，23% 的调查对象认为，离婚程序应该变得更加容易一些，36% 认为应该更加难一些，另外 41% 的人则认为应该维持现状。

剩下的另一半实验对象也回答了相同的问题，但最后两个答案选项的顺序被调换：“你认为在这个国家里，离婚程序应该更容易、维持现状、还是更难？”

在这种情况下，26% 的实验对象认为，离婚程序应该变得更加容易一些，29% 倾向于维持现状，另外 46% 的人则认为应该更加难一些。在两种表达方式下，最受欢迎的答案总是最后被提到的选项。

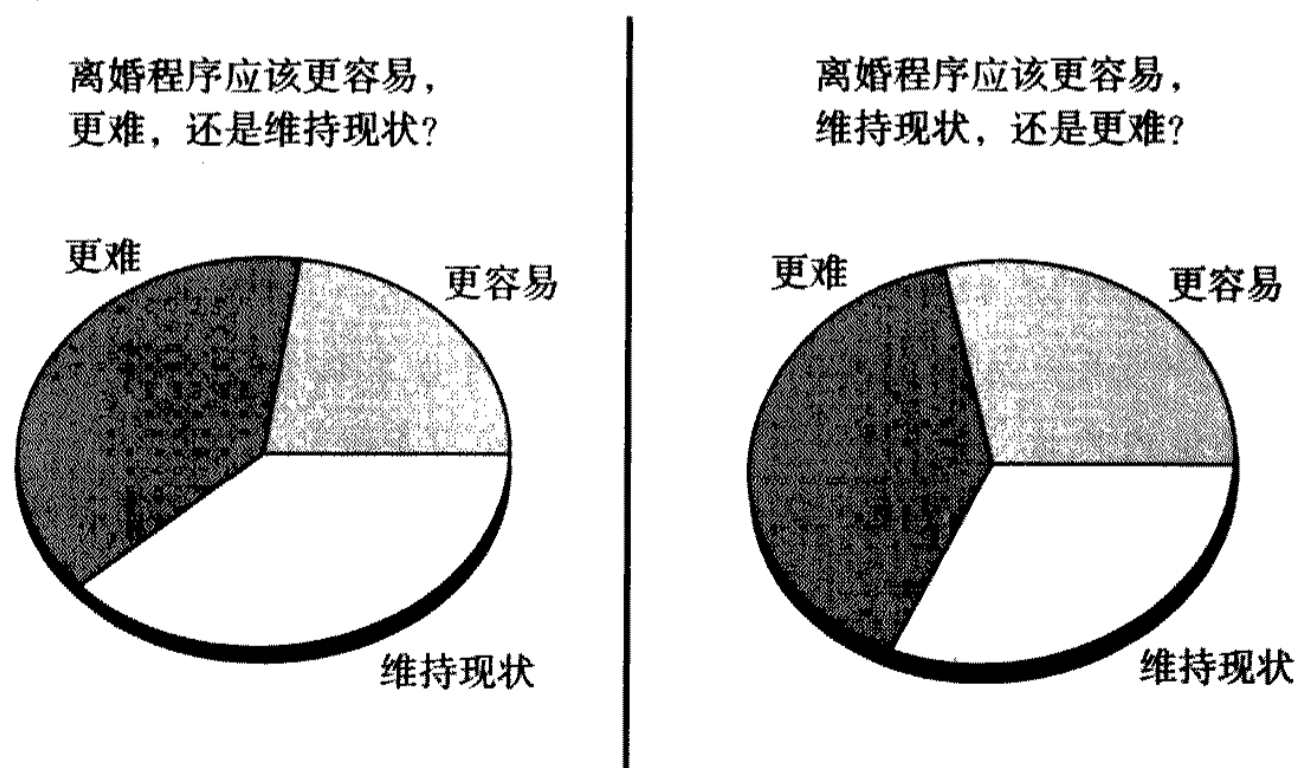


图 5.1 在有关离婚的两种版本的问题中，回答者都倾向于选择最后一个选项（引自 Howard Schuman 和 Stanley Presser 于 1981 年所做的一项全国调查的数据）。

虚假意见

虽然人们回答问题的方式受到语境和顺序的很大影响，答案的可塑性也是有限度的（研究者将答案的可延展性称为可塑性）。如果人们对一个问题已经非常熟悉，那么语境和顺序产生的边际变化一般小于30%（“边际”指的是调查对象给出每一个答案的百分比）。但是，如果人们对一个问题所知甚少，那么他们就会更容易受到语境和顺序变化的影响。另外，如果人们对一个问题一无所知，一部分人会表现出完全可塑性；在某些特别的询问方式下，一部分人会对这一问题发表意见，但实际上他对此并不真正了解。这种看法被称为“虚假意见”。*

尤金·哈特利（Eugene Hartley）在1946年发表的实验是有关虚假意见普遍程度最早的几个实验之一。在对几百名大学生的调查当中，哈特利发现，超过80%的人倾向于从“社会距离”（一种被广泛用于评价社会中人们彼此之间的感受的指数）的角度来评价 Danireans、Pireneans、Wallonians 和其他32个民族。这里的蹊跷在于，实际上并不存在所谓的 Danireans、Pireneans 或者 Wallonians——哈特利只是人为地设置了这些民族，以检验有多少学生会假装他们对此持有自己的意见。

第二年，*Tide* 杂志上刊登了一个类似的民意调查，就一个虚构的《Metallic Metals 法令》向被试询问如下问题：

以下哪种说法与你对《Metallic Metals 法令》的看法最为一致？

- a. 对全美国来说，这是一项很好的措施。
- b. 这是一项不错的措施，但应该留待各个州独立解决。
- c. 对其他国家来说或许不错，但不应该在这里实行。
- d. 没有任何意义。

虽然实际上并不存在《Metallic Metals 法令》，但却有70%的被试表达了他们对该法令的意见。其中41%的人认为应该将它留待各个州独立解决，15%的人认为对全美国来说这是一项很好的措施，11%的人认为不应该在这里实行，

* 正如汤姆·史密斯（Tom Smith, 1984）所说，意见和虚假意见（或者有时候被称为态度和虚假态度）实际上是一种连续状态的两个终端，但并没有清晰和量化的定义。然而，区别这二者仍然是非常有用的（正如冷和热都是一种连续状态的一部分，但却是截然不同的）。

对这一术语的另一补充：全书中，“意见”和“判断”通常是可以互换的，而“态度”则通常指带有评价性的判断（在好-坏，喜欢-不喜欢，同意-不同意等维度上面进行判断）。例如，对于罐子里有多少豆子，人们形成的是意见或是判断，而很少是一种态度。大多数的心理学家在使用“态度”这一术语时，指的都是带有评价性的判断。而在实际中很难区分态度、意见和判断。

另外3%的人则认为这项法令没有任何意义（Gill, 1947年3月14日）。

这篇文章同时还讨论了另一项调查的结果，在这项调查中，被试被问到：“你支持还是反对乱伦？”（在20世纪40年代，“乱伦”一词并不如今天这样为人们所熟知。）在发表意见的被试中，有2/3的人反对乱伦，而竟然有1/3的人支持乱伦！

政治事件中的虚假意见

在回顾了有关虚假意见的实验之后——其中包括舒曼和普雷瑟（1981）自己的一些实验，舒曼和普雷瑟得出的结论是，这一问题固然很重要，但并不像早期研究所认为的那么严重。在他们自己所做的调查中，只有1/3或1/4的被试对一个模糊的立法问题给出了虚假意见。其他一些研究者也得出了类似的结论（Bishop, Oldendick, Tuchfarber, & Bennet, 1980）。

被试中有25%~35%的人给出了虚假意见，这也许不算十分严重。但应该记住，在许多民主国家（包括美国），只要有30%的民众便可以选出一国的总统。政治论战的结果往往都是由一个差距很小的百分比来决定的。所以说，虚假意见对许多政治事件具有决定性的影响。一般说来，那些对某一问题知之甚少的人加上那些对此闻所未闻的人，往往能够形成一个绝对的多数。

在有关外交和军事政策的判断中，虚假意见实际上非常普遍。这是由于人们只有很低的政治参与程度和很少的政治意识，却面临一种不得不对此发表意见的社会压力。可以参考一下普劳斯（1989年3月）所给出的一些有关美国“政治文盲”的例子：

- 1988年的一项对美国成年人的盖洛普调查中，有近一半的人不知道南非是一个存在官方种族隔离制度的国家，有3/4的人不能说出官方已经承认拥有核武器的四个国家的名字。
- 1985年针对大学生的一项调查发现，45%的人不能说出曾遭受核武器袭击的两个城市的名字。
- 根据《纽约时报》所刊登的1979年的一项政府调查的结果，40%的高中高年级学生误以为以色列是一个阿拉伯国家。
- 1983年的一篇名为“萨尔瓦多不在路易斯安那”的《华盛顿邮报》文章指出，在一项全国性的调查中，有3/4的人不知道萨尔瓦多的地理位置（这篇文章的题目是从一名调查对象的答案得来）。
- 其他调查显示，68%的回答者不知道美国并没有有效对付弹道导弹的装备，81%的人错误地认为，只有当前苏联先动用其核武器时，美国才会动用核武器。

“政治文盲”现象的存在，使得解释有关外交和军事政策的公众意见变得

更加困难了。例如，如果以色列被误认为是阿拉伯国家，支持以色列有何意义？如果很多人认为美国能够抵御弹道导弹的进攻，那么支持战略防御行动还有什么意义？要使这类政治判断有意义，必须首先将真实意见和虚假意见区分开来。

过滤虚假意见

历史上，很多民意调查都没有对问题进行过滤。大多数这样的调查都没有采取措施排除那些实际上没有意见的人，回答选项中也没有包括诸如“没有意见”或者“不知道”这一类的选项。不过，越来越多的研究者都开始意识到“过滤器”的作用。设置过滤器正是为了排除掉那些实际上对某一给定问题并没有什么意见的人。

设置过滤器的方式有很多种。在一些民意测验中，被试首先会被问到他们是否听说过某件事情或是读到过有关某个问题的报道。如果回答是肯定的，他们才被要求发表有关这个问题的意见。如果不是，则回答其他问题。在另一些民意测验中，则首先询问被试是否就某一问题进行过深入的思考，或者更直接地询问他们是否对某一问题有自己的看法。还有一种过滤方法就是明确地设置“没有看法”或是“不知道”这样的选项。

一般来说，设置过滤器能够有效地排除那些虚假意见。但在某些情况下，却可能使调查结果产生偏差。例如，“不知道”往往与受教育程度或是对某一问题的关注程度呈负相关。如果过滤掉那些教育程度或是关注程度相对较低的人，那么调查结果就很难代表全体民众的看法。

为了评价过滤器的效果，舒曼和普雷瑟（1981）进行了一系列实验性民意调查来比较有过滤器条件下和无过滤器条件下的效果。根据其发现，他们得出以下结论：在大多数有过滤器的情况下，（1）至少有 1/5 的人会选择“不知道”；（2）做出某一特定回答的相对人数比例并没有受到显著影响（例如，回答“是”的人数与回答“不是”的人数的比例）；（3）不同问题答案之间的相关关系并没有受到很强的影响。

比如，1974 年的一次全美民意调查中，以下这样一组问题的答案就典型地反映了舒曼和普雷瑟的实验结果：

无过滤器的条件：“阿拉伯国家在试图与以色列达成真正的和平。你是否同意以上的说法？”

同意（17%）

不同意（60%）

不知道（主动提到）（23%）

有过滤器的条件：“阿拉伯国家在试图与以色列达成真正的和平。你对此

有看法吗？（如果有）你同意以上的说法还是不同意？”

同意（10%）

不同意（45%）

没有看法（45%）

虽然在有过滤器的条件下，回答“没有看法/不知道”比无过滤器的条件下多出了22%，但是回答“不同意”与回答“同意”的人数比例在两种条件下基本上是一样的（回答“不同意”与回答“同意”的人数比例大约为四比一）。

当然，人数比例接近并不意味着过滤器不重要。比如说，假设一个民意测验人希望民意测验的结果表明，绝大多数人并不同意这样的说法：“阿拉伯国家在试图与以色列达成真正的和平。”这种情况下，民意测验人就可能不使用过滤问题。而如果民意测验人希望民意测验的结果表明，绝大多数人同意这样的说法，那么使用过滤问题就是一个好方法。这种操纵手法已经产生了不少极具煽动性的新闻。

令人困惑的态度矛盾

选择和意见的可塑性与态度矛盾密切相关。可塑性通常指的是人们对同一个问题的不同版本的回答的差异程度，而矛盾则指的是两种相关态度之间的差异（态度-态度的矛盾）。詹姆斯·普罗思罗和查尔斯·格里格（James Prothro & Charles Grigg）于1960年发表的实验是有关这一矛盾的最有说服力的实验之一。

普罗思罗和格里格感兴趣的是，美国民众是否会支持一些广为流传的民主原则的具体运用。这些民主原则在当时确实已经为绝大多数美国人所接受。比如说，这些原则包括：

1. 公务员应该由大多数人投票选举产生。
2. 每个公民应该有平等的机会影响政府政策。
3. 少数派可以自由地批评多数派的决策。

普罗思罗和格里格概括出这些一般性原则后，列出了10个由此派生的问题，这些问题要么描述这些原则，要么与这些原则相矛盾。例如：

如果一个共产主义者想在这个城市里发表支持共产主义的演讲，应该允许他做这样的演讲。

或者：

在城市公投中，只有那些对公投问题有充分认识的人才 有资格进行投票。

然后，普罗思罗和格里格在密歇根州的安阿伯（地名）和佛罗里达州的塔拉哈西（地名）随机抽取了一组已登记的选民作为样本，询问他们对这 10 个问题的态度。

研究结果非常惊人。对于每个问题，调查对象都难以达成 90% 的一致。而且，他们对于这些民主原则具体运用的态度判断往往与这些民主原则本身相矛盾。例如，51% 的调查对象认为，只有那些对公投问题具有充分认识的人才 有资格进行投票。这无疑是反民主的。另外，79% 的人认为只有纳税人才有资格进行投票，也只有 44% 的人认为可以允许共产主义者公开宣传共产主义。普罗思罗和格里格（1960, p. 293）对这些具有讽刺意味的结果进行了总结：“假如我们承认美国是一个民主社会，我们就不能完全肯定地说，民众对民主原则的普通认同是一个民主社会存在的必要条件。”

这一研究只是态度 - 态度矛盾的一个极端的例证。一些后续的研究进一步证实了普罗思罗和 Grigg 的结论。人们对某一原则所持有的抽象态度往往与人们对这一原则的具体运用所持的态度无关。因为，在面临具体运用时，往往存在许多不可避免的复杂因素，如情境约束、其他与之相冲突的原则，等等。而且，正如本章下一部分将要说明的，关于态度 - 行为矛盾的研究显示，抽象态度与具体行为也只有很微弱的联系。

故地重游

1930 年，斯坦福一个名为理查德·拉彼埃尔（Richard LaPiere）的社会学家与一对中国夫妇结伴周游美国。两年间，拉彼埃尔和那对夫妇游遍了全国，总共到过 184 个餐馆，67 个旅馆、汽车旅馆和专门招待旅行者的家庭。虽然在当时存在着很严重的反华倾向，但在这些接待过他们的 251 个地方中，只有一次让拉彼埃尔感受到了种族歧视。事实上，拉彼埃尔认为，有 72 次他的同伴甚至受到了“超出一般规格”的优待。由此，拉彼埃尔认为，美国人对中国人并不存在偏见。

但在抽象概念的意义上，这种偏见却十分明显。旅游结束的 6 个月之后，拉彼埃尔向他到过的地方的业主送去一份调查问卷：“你会接待中国人吗？”在他坚持不懈的努力之下，回收了 128 份问卷，其中 81 份来自餐馆、咖啡厅，47 份来自旅馆、汽车旅馆和专门招待旅行者的家庭。在这些问卷中，有 118 份问卷表示他们不会接待中国客人，有 9 份表示要视情况而定，只有 1 个汽车旅馆的妇女很肯定地说她曾经在夏天接待了一对中国夫妇。拉彼埃尔抽取了相同区域内的一些他们没有到过的地方进行调查，也得到了完全一样的结果：118 份回答否定，9 份回答视情况而定，1 份肯定（如欲了解更多细节，参见 LaPiere, 1934）。这些发现说明，人们可能会有一些抽象概念，但这与他们的实际行为只有很微弱的联系，或是完全无关。

在拉彼埃尔的研究结果发表3年以后，斯蒂芬·科里（Stephen Corey, 1937）也发表了具有相同结论的实验。科里感兴趣的是对待作弊的态度和作弊的实际行为之间的关系。为了测量对待作弊的态度，他让67名大学生完成了一个关于作弊的态度的实验。该实验看起来好像是匿名的，但科里用了一种很隐秘的手段来识别这些被试。利用这种方法，他能够将这些真实的看法与作弊的具体行为所联系起来。

科里度量作弊行为的方法如下：让这些学生每周参加一次考试，连续5周。对每次考试进行秘密评分，然后将看起来没有评分的试卷返回给学生，让他们自行评分并将评分结果返回。学生报告的分数与他们实际分数的总差异就是作弊的实际行为度量。作弊的平均程度为：每40~45道题中有两题不符。

科里的发现表明，人们的态度和行为之间的相关关系几乎为零。这些学生对待作弊的态度似乎与他们自己的作弊倾向没有多大联系。真正与作弊有关的是他们在考试中的表现。用科里（1937, p. 278）的话来说：“学生是否会在考试中作弊，很大程度上取决于他是否为这次考试做了充分的准备，而不是他自己所宣称的对待作弊的态度。”

我们这个时代的寓言

1973年，约翰·达利和丹尼尔·巴斯顿（John Darley & Daniel Baston）发表了他们的实验结果，这是有史以来最为生动的有关态度-行为矛盾的说明。达利和巴斯顿感兴趣的是，哪些因素决定人们会对处于困境中的人施以援手。被试是神学院的学生，他们要从一幢大楼走到另一幢大楼去做一场演讲，演讲的主题可以是这些学生擅长的某项工作，也可以是一个撒玛利亚人寓言（圣经中帮助处于困境中的人的故事）。实验助手会告诉学生，这些演讲的时间应该控制在3~5分钟内，而且会有另外的实验助手为他们作记录。然后，在学生去往目的地的途中，他们会遇到一个需要帮助的人。达利和巴斯顿想知道，究竟帮助他人的行为是否与下列因素有关：（a）学生要做的演讲主题是否与帮助他人有关，（b）他们在途中的匆忙程度。

在非常匆忙的条件下，实验助手会突然对被试学生说：“噢，你已经迟到了。你本来应该几分钟以前就到那儿的，我们最好马上去。那边的助手已经等你一会儿了，所以你最好快点，最好一分钟就赶到。”在一般匆忙的条件下，实验助手则会告诉学生：“那边的助手已经准备好了，请赶快过去吧。”而在不太匆忙的条件下，实验助手则会说：“还有几分钟那边才能准备好，但你最好现在就走，即使要等，也不会等太久的。”

为了到达对面的大楼，每个学生都必须穿过一条胡同。在这条胡同中，达利和巴斯顿安排了一个衣着破烂、垂头丧气地坐在门口的人，他垂着头，闭着眼睛，面无表情。当被试经过的时候，这个人会咳嗽两声，低着头痛苦地呻

吟。如果学生停下来询问他有什么不舒服，或者提出要帮忙，这个人会含糊不清地说：“噢，谢谢（咳嗽）……不用了，我还好。（停顿一下）我早就有这个气喘的毛病了（咳嗽）……医生已经给我开了这些药，我刚才吃了一片……我只要坐下来休息一下就好……但还是很谢谢你的关心。”

如果学生坚持要将这个人送进大楼，这个人会接受学生所提供的任何帮助，并感谢他停下来帮助。学生离开后，这个人会马上按以下的五点量表对学生做出评价：

- 0 = 完全没有注意到这个人可能需要帮助
- 1 = 察觉到这个人可能需要帮助，但是并没有停下来帮助他
- 2 = 没有停下来，但间接地提供了帮助（比如，将这个人的情况告诉其他人）
- 3 = 停下来，询问这个人是否需要帮助
- 4 = 停下来询问后，坚持要将这个人送进大楼里去

达利和巴斯顿（1973）发现，相对于不是很匆忙的学生来说，匆忙中的学生停下来提供帮助的可能性要小一些。但是，演讲主题并没有明显地影响到学生是否会停下来提供帮助。事实上，有些要去讲有关撒玛利亚人寓言的学生为了不迟到，在经过胡同时甚至跨过了那个人！这些结果生动地说明，抽象概念——在这一案例中，也就是帮助他人脱离困境的重要性——与实际行为有极大的区别。

对矛盾的再研究

态度和行为是否总是这样的矛盾呢？1969年，心理学家艾伦·威克（Allan Wicker）发表了一篇研究回顾文章，文章认为答案是肯定的。威克的这个结论对态度研究来说是一个很大的打击。威克列举了46项研究作为其文献回顾的基础。在这些研究中，有关态度和相对应行为的实验是在不同条件下进行的。被试的来源十分广泛，包括了大学生、保险经纪人、工厂工人、妇产科护士，等等。同样，研究的主题也非常宽泛，包括了公共住房、足球、民权运动，等等。

在这46项研究的基础之上，威克（1969，p. 65）总结道：“态度与行动并不是那么密切相关。态度与表露在外的行为实际上是无关系的，或只有很微弱的联系。”根据他的研究，态度与行为的相关系数往往趋近于零，只有在很少的情况下这一系数会超过0.30。两年之后，威克（1971）做了更进一步的研究，他认为，即使抛弃“态度”这一概念也未尝不可。

可以想像，这些结论当然会令态度研究者们感到十分不安。没有多久，修正学派诞生了。修正学派的拥护者们认为，在某些条件满足的情况下，态度与行为是一致的。这些条件包括：（1）必须仔细选择态度和行为的测量方法，应

该尽可能有效和可靠；(2) 只要有可能，应该尽量以多种方式来研究态度和行为；(3) 为了避免出现干扰变量，研究态度和行为的时间应该被安排在一起；(4) 在行动、行动的对象、行动发生的情境以及行动发生的时间方面，态度和行为的测量手段应该相匹配。

在1977年发表的一篇文献回顾中，艾克·阿杰增和马丁·菲什拜因(Icek Ajzen & Martin Fishbein)说明了其中一些条件的重要性。阿杰增和菲什拜因从具体对象和行动的角度来定义态度和行为的符合程度，并将100多项研究中的态度-行为关系分为三类：高符合程度，中等符合程度和低符合程度。当态度和行为的符合程度较低时，阿杰增和菲什拜因发现，大多数情况下，态度和行为之间不存在显著相关。而另一方面，如果对态度和行为进行了仔细测量，且它们在具体对象和动作方面的符合程度较高时，态度-行为的相关系数则至少为0.40。也就是说，如果态度是关于某项特定的动作，且这一动作指向某个特定的对象，那么态度就能很好的预测行为。但是，如果态度指向的对象并不是行为所指向的对象，态度和行为相一致的程度就降低了。阿杰增和菲什拜因认为，拉彼埃尔(1934)之所以认为态度和行为只有很微弱的一致性，是因为其研究中态度的对象(普遍意义上的中国人)和行为的对象(一对特定的中国夫妇)相距甚远。

结 论

可以用一句俄罗斯名言来概括了本章所讨论过的大多数研究结果。正如那位著名的俄国沙皇亚历山大所说的(大意如下)：“跨越人生并不是像跨越一片田野那么简单。”应用到决策与判断的研究上来说，也就是：“衡量一种态度、看法或是偏好，并不是像问一个问题那么简单。”

态度、意见和选择往往具有惊人的可塑性。在许多情况下，问题的措辞对人们的回答有非常重大的影响。因此，我们对问题的结构和情境要特别小心。第6章将给出一些例子来说明，微妙的措辞变化是如何影响决策和判断的。

第6章

问题的措辞和框架的影响

1986年英国举行的一项盖洛普民意测验中的一个问题是：自己的国家拥有核武器是否会让他们感到“安全”时。40%的人对此持肯定态度，而50%的人则对此持否定态度（剩下的10%没有明确意见）。然而，当测验问卷中的“安全”被改为“更加安全”时，答案的百分比却倒转了过来：50%的人认为拥有核武器让他们感到更加安全，36%的人认为拥有核武器让他们感到更加不安全（Lelyveld, 1986年10月5日）。

这些结果说明了措辞微妙变化的重要性。有时候，即使只改变一两个词语——不管是问题本身变化还是答案选项变化——也会对人们回答问题产生深刻影响（Borrelli, Lockerbie, & Niemi, 1987）。考虑一下下面这个例子。

一项取巧的任务

这是在1969年。你是一名卓越的盖洛普民意测验组织的民意测验专家，同时也是一个为达到目的而不择手段的人。你的一项秘密任务就是要进行一次民意测验，并使这项测验的结果表明，美国公众希望美国加快从越南战争撤军。你怎样才能达到目的呢？

在这一年的6月，盖洛普告知问卷回答者：尼克松总统已经宣布，未来的3个月内将有25 000军队将从越南撤出。回答者需要回答“撤军应该加快还是应该减速”（“保持现状”并没有被直接包括在答案选项中，但如果回答者主动提到，采访者仍然需要记录这个答案）。几乎一半的回答者（42%）认为应该“加快”，16%认为应该“减速”，另外29%认为应该“保持现状”（Converse & Schuman, 1970年6月）。

你将这些结果向媒体公开，第二天全国报纸的头条都是：美国民众希望加快撤军速度。现在，你便可以轻松地坐在椅子上，偷偷享受自己的成功了。

但不久之后，另一家民意调查机构哈里斯（Harris）做了一个类似的调查，只是问题变成了：“你认为总统的撤军速度是过快，过慢，还是恰到好处？”在对这一问题的回答中，有半数的人（49%）认为撤军速度刚刚合适。只有29%的人倾向于加快撤军速度，而6%的人认为撤军速度过快（Converse & Schuman, 1970年6月）。

拿起报纸，现在你看到的头条又成了：美国民众希望维持现在的撤军速度；先前的民意测验是错误的。

“先前的民意测验是错误的？”你问自己。盖洛普进行的这次民意测验肯定是没有错误的。但如果先前的民意测验也没有错误，那么究竟哪个测验是正确的呢？

其实，与其探究究竟哪个民意测验是正确的——两种情况都可能成立——还不如仔细考虑一下这两个测验结果各自有什么样的含义。盖洛普使用的是一个没有明确包括中间选项的“迫选”问题。因此，即使人们只有非常非常小的偏好，也只能选择加快或减速撤军。没有中间选项的问题，常常能够有效地检测出人们的一般倾向。但如果要做到十分严谨，报道盖洛普测验结果的头条应该改为：如果只能在加快或减速撤军中选择一项，美国民众希望加快撤军速度。因为，盖洛普民意测验的结果并不代表真正不满意撤军速度的美国民众的比例。

与之形成对照的是，哈里斯调查为回答者提供了一个明确的中间选项。增加一个中间选项的效果与增加一个“没有看法/不知道”过滤器的效果差不多；通常，会有10%到40%的人转而选择中间答案，但是边际比例基本上会维持不变（在这个例子中，边际比例是认为“过快”和“过慢”的人之比）。但是设置这样一个中间选项也有缺点。这是一个“安全”答案，因此在某些情况下可能会产生虚假信息。

开放式问题

引入一个中间选项可能会鼓励人们选择中立。另外，引入一些其他选项也

表 6.1 有关重要性的问题

问 题	选择每个答案的百分比	
	开放式问题	封闭式问题
公立学校教育质量	1	32
污染问题	1	14
合法堕胎	0	8
能源短缺	0	6
所有其他回答	98	40

注释：这些数据来自舒曼和斯科特（1987）所做的调查，他们在调查中询问了8名被试一个开放性问题“你认为这个国家当今所面临的最重要的问题是什么？”，询问171名被试一个封闭式问题“你认为以下哪个问题是这个国家当今所面临的最重要的问题：能源短缺，公立学校的质量，堕胎合法化还是污染问题，如果其中没有你想要选择的答案，你可以填写其他答案。”

可能导致选择这项答案的人数增多。霍华德·舒曼和杰奎琳·斯科特 (Jacqueline Scott) 于 1987 年发表的研究便说明了这样一种现象。

在一次全国性调查中,舒曼和斯科特要求调查对象回答:“当前我们国家所面临的最重大的问题是什么?”正如表 6.1 的第一行所列,如果以“开放式问题”(调查对象给出自己的回答)的形式来询问调查对象,那么只有 2% 的人认为,当前最重大的问题是能源短缺、公立学校教育质量、合法堕胎或是污染问题(参见你在读者调查中第 27 题的回答)。然而,当舒曼和斯科特将这些回答频率不高的答案纳入到“封闭式问题”中时,情况就完全不同了(在封闭式问题中,调查对象需要从答案列表中选出答案)。正如表 6.1 的第二列所示,大部分的调查对象选择了一个不常见的答案作为国家现在所面临的最重大问题。将这些不常见的回答在答案选项中列出后,它们的被选频率就提高了 30 倍。

封闭式问题有时候可能会把那些平时很少想到的不常见答案纳入选项中,这样做确实会扭曲调查结果。但有时开放式问题同样也会对回答产生误导。比如说,当人们被问到“近代世界上发生过的最重大的事件是什么”时,只有不到 2% 的人提到了计算机的发明。但当计算机作为一个答案选项出现时,大约有 30% 的人都选择了这个答案 (Schuman & Scott, 1987)。在这个例子中,使用封闭式问题比使用开放式问题更能够体现人们真实的看法。

即便是回答一些非常普通的问题也容易受到答案选项的数量和类型的影响。比如说,一项研究就发现,被试对电视使用时间的回答会随着问题答案的种类而变化 (Schwarz, Hippler, Deutsch, & Strack, 1985)。在这一实验中,被试用了两种量度来表示他们每天看电视的时间。每种量度都以半小时为间距,分为六种水平。第一种以“最多半小时”开头,以“超过两个半小时”结尾。

表 6.2 什么才是事实

问 题	平均答案
你经常头痛吗,如果是,多久一次?	2.2 次/周
你偶尔会头痛吗,如果是,多久一次?	0.7 次/周
你试过多少种其他(头痛)药品? 1? 5? 10?	5.2 种产品
你试过多少种其他(头痛)药品? 1? 2? 3?	3.3 种产品
这部电影有多长时间?	130 分钟
这部电影有多短时间?	100 分钟
篮球运动员有多高?	79 英寸
篮球运动员有多矮?	69 英寸

注释:第一对问题摘自伊莉莎白·洛夫特斯的研究(1975),第四对问题摘自理查德·哈里斯的研究(1973)。所有差异都具有统计学上的显著性。

而第二种则以“最多两个半小时”开头，以“超过四个半小时”结尾。对于第一种问题，有84%的人回答说每天看两个半小时左右的电视。而对于第二种问题，则只有63%的人表示每天看两个半小时左右的电视。

正如诺伯特·施瓦茨（Norbert Schwarz）和他的同事（1985）所说：“量表不仅仅是被试用来描述他们行为的‘衡量工具’。更确切地说，被试在评价他们自身的行为时，往往利用了答案选项中所提供的行为范围作为参照框架。”框架化的答案选项——正如问题本身一样——从来都无法做到完全中立，它们通常传递了一种隐含的可以接受的答案（见表6.2，一些隐含了答案的问题）。正如下一部分将要阐述的那样，当一些答案比另外的答案更能具有社会称许性时，答案偏差也可能产生。

社会称许性

1982年，曾经负责为里根总统做民意测验的理查德·沃思林（Richard Wirthlin）在一次全国性调查中发现，58%的调查对象同意这一说法：“我们应该反对美国冻结其核武器，因为这对于减少那些业已存在的数以万计的核武器来说毫无益处，而且还会使前苏联处于核优势状态。”但就在几分钟后，又有56%的调查对象表示同意这样的说法：“我们应该赞成美国冻结其核武器，因为这样才能开展我们现在迫切需要实施的行动，那就是阻止世界上其他国家建立核武器储备，减少爆发核战争的可能性。”实际上，有27%的调查对象同时赞成两种说法。沃思林这样描述道：“这是我们所问过的问题中，最具有矛盾性的一个问题。”（Clymer, 1982年5月6日）

为什么会出现这种情况？当调查对象对冻结核武器缺乏一个坚定的看法时，他们就可能会试图给出最具有社会称许的答案。第一个问题将支持冻结核武器与让前苏联占据核优势等同，而第二个问题则将冻结核武器与世界和平相联系起来。几乎没有一个美国人会赞成让前苏联占据核优势，但也没有人会反对世界和平。

当人们对一个问题并没有很坚定的信念时，他们通常会对那些“最广为流传的句子”做出反应，这些句子使得人们的态度往社会所赞许的方向靠拢。例如，《美国新闻与世界报道》曾经公布了一项民意调查的结果。在这一调查中，58%的调查对象赞成对尼加拉瓜反抗者进行援助以“防止共产主义影响的扩散”，但只有24%的人赞成对“试图推翻当权政府的人们”进行援助（Budiansky, 1998年7月11日）。对于大多数的美国人来说，“预防共产主义”是一项被认可的行为，而“推翻当权政府”则不然。

其实早在1940年，哈德利·坎特里尔（Hadley Cantril）在两次有关美国是否应该参加第二次世界大战的民意测验中就得出了相同模式的结果。只有13%的调查对象同意“美国应该做出比现在更多的努力去帮助英国和法国”。但当

第二组调查对象被问到“你认为美国是否应该做出比现在更多的努力，去帮助英国和法国打败希特勒”时，这个数字就攀升到了22%。在当时，帮助“打败希特勒”比单纯地帮助英国和法国更具有社会称许性。

允许还是禁止

在另一组早期的民意测验中，鲁格（Rugg, 1941）向不同的调查对象分别询问了以下两个问题之一：

你认为美国是否应该允许公开发表反民主的演说？

你认为美国是否应该禁止公开发表反民主的演说？

虽然这两个问题看起来好像是问的都是同一件事情，但鲁格却发现它们导致了完全不同的回答。当人们被问到美国是否应该允许公开发表反民主的演说时，62%的人持否定看法。另一方面，当人们被问到美国是否应该禁止公开发表反民主的演说时，只有46%的人作出肯定回答（从逻辑上来讲，禁止与不允许发表这类演说是等同的）。在对此发表意见的调查对象中，回答“不允许”发表反民主演说的人竟然比回答“禁止”发表反民主演说的人多出20%！

舒曼和普雷瑟（1941）曾在20世纪70年代将鲁格的实验重复做了3次，每次都得到了与鲁格（1941）类似的结果。在跨越了30年的4次调查中，相对于“允许”类问题来说，“禁止”类问题使得支持演说自由的人数大增（有可能是因为“禁止”更容易让人联想到自由的丧失）。舒曼和普雷瑟的实验与鲁格的实验只有一个明显区别，那就是，20世纪70年代美国人更能够容忍自由演说一些。所以在对这两个问题的回答中，70年代支持自由演说的人都比40年代支持自由演说的人多出了将近30%。

允许-禁止的差异在其他话题的一些实验中也得到了验证。例如，希普尔和诺伯特·施瓦茨（Hans-J. Hipper & Norbert Schwarz, 1986）就发现，在偷窥、（青少年禁看的）X级片、高速公路上用盐来化雪等问题上，“不允许”和“禁止”就存在明显的差别。其他一些民意测验还发现，调查对象中只有29%的人赞成进行“反堕胎”的修宪，但却有50%的人赞成进行“保护未出生生命”的修宪（Budiansky, 1988年7月11日）。

框 架

正如以上研究结果所示，人们对损失（比如禁止）和获得（比如允许）的反应是有差别的。但直到心理学家阿莫斯·特韦尔斯基和丹尼尔·卡尼曼（Amos Tversky & Daniel Kahneman）普及了“框架”这一概念之后，这一差别的重要性才得到人们的广泛认识。特韦尔斯基和卡尼曼（1981）认为，决策框架是

“决策者所拥有的有关动作、结果以及某一特定选择可能引发的有关情况的一系列概念”。特韦尔斯基和卡尼曼（1981）认为，决策框架部分是由问题形式决定的，部分是由社会规范、习惯和决策者的性格特征决定的。在特韦尔斯基和卡尼曼的研究中，他们着重研究了不同问题形式所产生的不同效果。

这些效果可以通过下面这个赌钱的例子看出来（特韦尔斯基和卡尼曼的一个实验）。在决策 1 中，你必须从选项 A 和选项 B 中选择一个：

选项 A：肯定会获得 240 美元

选项 B：25% 的概率获得 1 000 美元，75% 的概率什么也得不到

你会做何选择？（可以参考你在读者调查中第 25 题的答案。）记录下你的选择后，再来看决策 2：

选项 C：肯定会输掉 750 美元

选项 D：75% 的概率输掉 1 000 美元，25% 的概率什么也不会输掉

特韦尔斯基和卡尼曼（1981）发现，在决策 1 中，人们更喜欢选择 A。84% 的被试选择了选项 A 而不是选项 B。正如第 5 章所提到的，当面临获利时，人们往往是规避风险的（“二鸟在林”对人们来说往往不如“一鸟在手”），但在决策 2 中面临损失时，人们则往往更喜欢赌一赌运气。特韦尔斯基和卡尼曼实验的被试中，有 87% 的人选择了选项 D。因为当面临损失时，人们更乐于冒险。这种风险偏好和风险规避同时存在的情况非常普遍，因此在特韦尔斯基和卡尼曼（1981）的实验中，有 73% 的被试同时选择了选项 A 和 D，只有 3% 的人同时选择了选项 B 和 C。

有趣的是，选择 B 和 C 实际上比选择 A 和 D 更加有利。只要将已选选项的结果加总（见图 6.1）就可以看出来。选择 A 和 D，不但可以肯定地获得 240 美元的同时，而且输 1 000 美元的概率只有 75%。也就是说，输掉 760 美元的概率是 75%，获得 240 美元的概率是 25%。而如果选择 B 和 C，则只有 25% 的概率获得 1 000 美元，75% 的概率一无所获，扣除掉一个确定的 750 美元的损失之后，得到的是 25% 的概率获得 250 美元和 75% 的概率输掉 750 美元。

也就是说，将选项组合后得到的选择方案是：

A 和 D：75% 的概率输掉 760 美元，25% 的概率获得 240 美元

B 和 C：75% 的概率输掉 750 美元，25% 的概率获得 250 美元

如果你选择了 A 和 D，那么相对于选择 B 和 C，你将获得更少的钱或是输掉更多的钱。不管是输钱还是赢钱，选择 B 和 C 都比选择 A 和 D 更为有利。

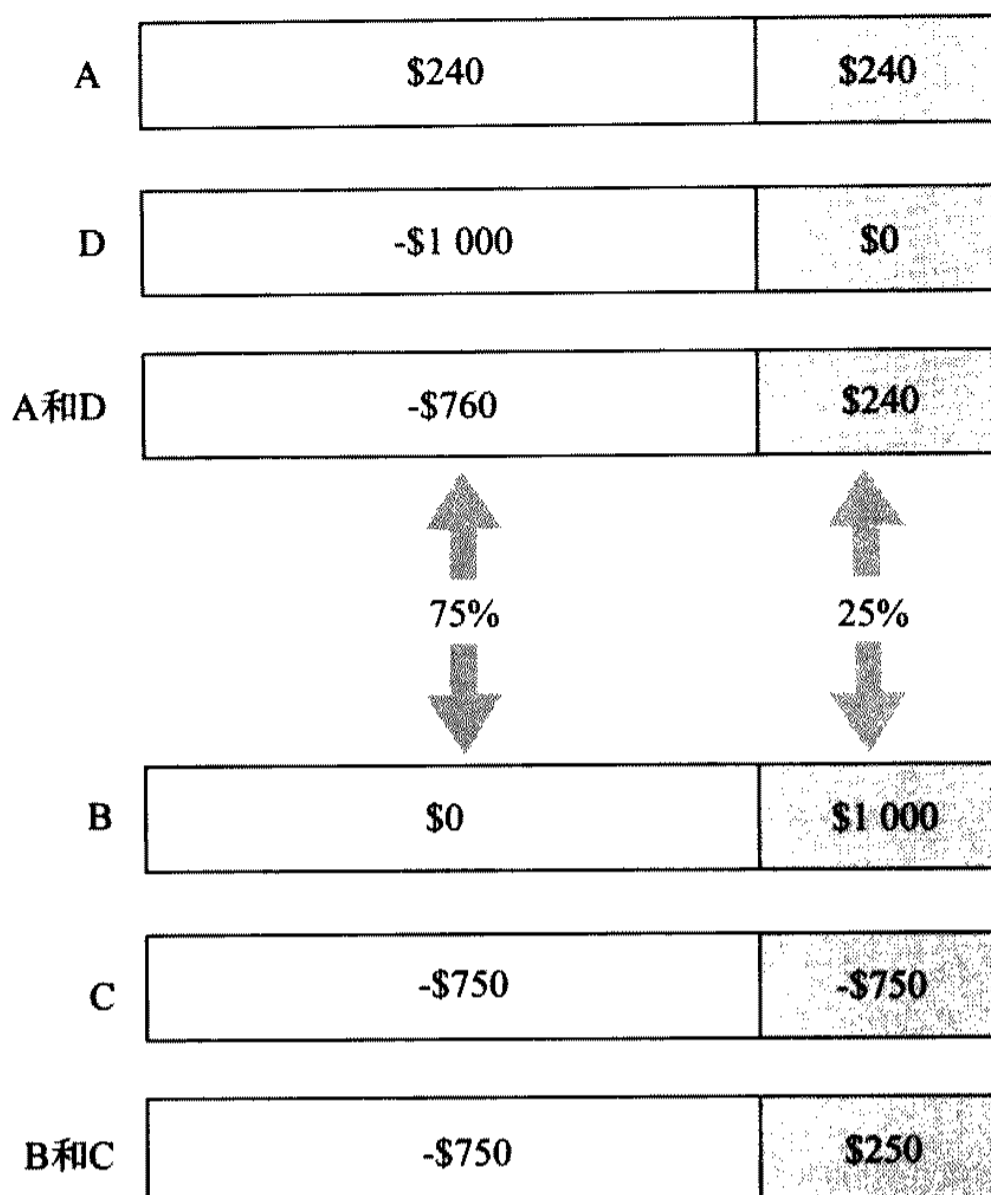


图 6.1 虽然选项 B 和 C 比 A 和 D 的结果更有利，特韦尔斯基和卡尼曼 (1981) 发现，与 B 和 C 相比，人们更多地选择了 A 和 D。

良药苦口难下咽

在医疗问题上，框架效应显得尤为重要。由特韦尔斯基和卡尼曼 (1981) 进行的“亚洲疾病”问题的研究是有关医疗方面框架效应的最早也是最有名的实验：

假设美国正在为亚洲即将爆发的一场非比寻常的疾病做准备，这场疾病可能会导致 600 多人丧生。如何与这场疾病作斗争，这里有两套方案。假设对每种方案结果的科学估计如下：

- 如果实施方案 A，能够挽救 200 人的性命。
- 如果实施方案 B，有 1/3 的概率挽回 600 人的性命，有 2/3 的概率无法挽救任何人。

你支持哪一个方案？

特韦尔斯基和卡尼曼发现，在这一框架下回答问题的人中，有 72% 的人属于风险规避型；他们选择能够肯定挽回 200 人生命的方案，而不是以如此多人的生命作赌注。

但如果改变问题的框架，特韦尔斯基和卡尼曼却得到了完全不同的答案。

他们给了另一组调查对象同样的问题，只是改变了对方案结果的描述：

- 如果实施方案 C，400 人会死亡。
- 如果实施方案 D，有三分之一的概率没有人死亡，有三分之二的概率 600 人都会死亡。

虽然这种框架下的方案 C 和 D 和前述框架下的 A 和 B 从数字上来说是等同的，但却有 78% 的回答者变得更愿意冒险了；他们更愿意赌一赌，而不愿意接受 400 条生命的丧失。

有关医疗决策的其他研究也得到了类似的结果（Wilson, Kaplan, & Schneiderman, 1987）。例如，《新英格兰医学杂志》（*New England Journal of Medicine*）上发表的一项实验显示，框架影响到了内科医生对肺癌治疗的决策（McNeil, Pauker, Sox, & Tversky, 1982）。这一实验的被试包括 424 名放射科专家，491 名学习过统计学和决策理论的研究生，以及 238 名患有各种慢性病但无需卧床的病人。他们向所有的 1153 名被试展示了两种肺癌治疗方法的信息概要（外科手术和放射性治疗）。读完这些概要后，他们要马上在两种治疗方法中做出选择。

一半被试所读到的信息概要是从病人的存活概率的角度来描述（比如说，68% 的概率能够多活一年）。而另一半的回答者所读到的信息概要，则是从死亡概率的角度来描述（比如说，一年之后死亡的概率为 32%）。由于肺癌手术的最大缺点就是在手术进行中或者手术结束后的短时间内死亡率较高，因此实验者推断，如果信息概要是从存活率的角度而不是从死亡率的角度来表述，那么选择外科手术的频率应该会高一些。

事实上，实验结果也的确如此。当以存活率表述时，选择手术而不是放射性治疗的人数百分比为 75%，而以死亡率表述时，则为 58%。这种选择结果不仅存在于病人中，也存在于内科医生和研究生中。因此说，即使是那些受过高等训练的专家所做出的生死攸关的决策，也会受到问题框架形式的深刻影响。

税收抵减的逻辑

在一篇引人深思的关于“经济推理”的文章中，哈佛经济学家托马斯·谢林（Thomas Schelling, 1981）向人们举了一个十分有趣的例子，在这个例子中，“重构”问题能够产生惊人的结论。谢林注意到，当时美国的所得税法规定，征收家庭所得税时，每个孩子可以抵减 1 000 美元的家庭收入（也就是说，在将家庭总收入转换为应纳税收入时，有一个孩子就可以少算 1 000 美元的收入）。这一抵减金额是固定的，它并不因家庭总收入的变化而变化。

但谢林对此提出一个问题，不允许那些富有家庭在每个孩子身上得到更大的税收抵减是否合理？毕竟，相对于贫穷家庭来说，富有家庭在每个孩子身上

的花费更多一些，而且一个富有家庭生育孩子的费用也比贫穷家庭的花费高出许多。

大多数的人都反对这一政策。人们认为，“完全没有道理进一步扩张富人的特权——更不能以牺牲公众的利益为代价。”

但谢林提出，如果税法以另外一种方式表达，情况也许不同。假设税法这样规定：一般来说典型的家庭都有2~3个孩子，那么对每个家庭征税时，加征一个“无子女附加税”以代替原来的税收抵减。也就是说，如果一个没有子女的家庭需要多缴税（而不是给有孩子的家庭予以税收抵减），那么一个无子女的贫穷家庭和一个无子女的富有家庭是否应该缴纳一样多的税收？

在这种情况下，人们就会说，富有家庭应该比贫穷家庭缴纳更多的税收。由于富有家庭在孩子身上的花费往往要比贫穷家庭更多，因此可以推断，相对于无子女的贫穷家庭来说，无子女的富有家庭能够负担更高的附加税。但问题在于，这一论述恰恰与刚才的有关富有家庭不能享受更多税收递减的论述相冲突。用谢林（1981，p. 55）的话来说：“所得税可以以无子女夫妇的家庭收入为基础，然后对有子女的家庭收入税予以调整；也可以是以有子女夫妇的家庭收入为基础，然后对无子女的家庭收入税予以调整。既然二者都能够得到相同的收入税，那么无论选择哪种方式应该不存在任何区别。”

但事实上却是有区别的。因为只要实施税收抵减，人们就会认为——并且好像有很好的理由——抵减额应该是固定的。

心理分账

决策者不仅会构架他们的选择——他们也会构架选择的后果。特韦尔斯基和卡尼曼（1981）将这种过程称为“心理分账”。心理分账（psychological accounting）是指，某一结果只是从动作直接结果的角度来描述（特韦尔斯基和卡尼曼称之为“简单账户”），还是需要综合考虑这一动作与其他行为的先前关系（特韦尔斯基和卡尼曼称之为“综合账户” inclusive account）。

下面这组问题（摘自特韦尔斯基和卡尼曼1981年所做的一个实验）就是心理分账的一个例证：

问题1：假设你想看一场戏剧，门票是10美元。当你走到戏院门口时，发现自己掉了10美元。你还会花10美元看戏吗？

在特韦尔斯基和卡尼曼调查的183名对象中，有88%的人表示他们仍然愿意花10美元看戏。大多数人并未将这10美元的损失与买票联系起来；相反，他们将这一损失划分到一个单独的账户上。

现在来看一看第二个问题：

问题2：假设你想看一场戏剧，并花了10美元买票。当你走到戏院门

口时，发现自己把门票丢了。座位并没有被标记出来，门票也不可能再找到。你会再花 10 美元买一张票吗？

在特韦尔斯基和卡尼曼调查的 200 名对象中，只有 46% 的人表示他们愿意再买一张票。虽然丢一张票的损失与丢 10 美元的损失从金钱上来说是等同的，但大多数的回答者显然将买第二张票的成本与买第一张票的成本加在了一起——而且认为这 20 美元是额外的花销。

特韦尔斯基和卡尼曼还做过另外一个有关心理分账的实验：

假设你准备买一件价值 125 美元的夹克和一个 15 美元的计算器。卖计算器的人告诉你，另一家分店里的计算器正在减价，只卖 10 美元。但是开车到那里需要 20 分钟，你会到那家分店去买吗？

对这一问题的回答中，有 68% 的被试表示愿意开 20 分钟的车以节约 5 美元。

但是如果计算器要 125 美元而夹克只要 15 美元呢？特韦尔斯基和卡尼曼就此问题询问了 88 名被试：

假设你准备买一件价值 15 美元的夹克和一个 125 美元的计算器。卖计算器的销售员告诉你，另一家分店里的计算器正在减价，只卖 120 美元。但是开车到那里需要 20 分钟，你会到那家分店去买吗？

在特韦尔斯基和卡尼曼的实验中，只有 29% 的被试表示他们愿意开车去。当计算器被标价为 125 美元时，多于 2/3 的被试表示他们不会为了节省 5 美元而开 20 分钟的车，虽然这一削价与先前问题中提到的削价没有什么区别。

这是为什么呢？特韦尔斯基和卡尼曼（1981）认为，被试是从简单账户的角度来结构化第一个问题的，因为 20 分钟的车程是直接和 5 美元的节省相比较的。而在第二个问题中，被试将整体购买价格综合进了账户中，因此 5 美元显得就不那么重要了，甚至可以忽略。乔治·夸特龙（George Quattrone）和特韦尔斯基（1988）进一步解释道，人们是根据“比率 - 差异原则”来做选择的，两个金额之间的固定（正）差额的影响随着它们之间比率的增加而增加。20 美元与 15 美元的比率为 1.33，这当然比 125 美元与 120 美元的比率 1.04 更有影响力。理查德·塞勒（Richard Thaler, 1985）在研究心理分账在消费选择中扮演的角色时也发现了类似的框架效应。

这些发现以及本章先前所提到的其他实验结果并不意味着一个聪明的销售员或是民意测验专家可以得到他们想要的任何结果。它们只是说明，问题的措辞和框架往往具有非常大的影响力，十分值得人们去关注。虽然这些结果看起来好像是不言自明的，但实际上——正如接下来的 3 章所表明的那样——人们花了很长时间来认识它们。

结 论

人们在得知某一公众民意调查的结果时很少会想到，这些问题是否经过了过滤，答案选项中是否包括了中间选项，问题是从收益还是从损失的角度来提问的，等等。大多数人往往会相信，被访者是按照他们自己的信念来回答问题的。但大量的研究表明，框架和问题的措辞能够显著地影响人们的回答。

对于那些有关决策与判断的调查和研究，一定要考虑人们的答案是否会随着以下因素而发生变化：

- 问题呈现的顺序
- 问题出现的情境
- 问题是开放式还是封闭式
- 问题是否经过了过滤
- 问题是否包含了某些时髦词句
- 答案选项的范围
- 答案选项出现的顺序
- 是否提供了中间选项
- 问题是从收益还是从损失的角度来提问的

如果你觉得这些因素的变化可能会改变人们的回答，那么在经过措辞改变测试之前，你的研究结果是不具有可信度的。正如保罗·斯洛维奇、戴尔·格里芬和特韦尔斯基（1990，p. 25）所说：“如果结果在经过了多种程序的测试之后仍然是一致的，那么我们就有理由相信这些判断；如果不是，就需要更进一步的分析。”由于判断往往容易受到问题措辞和框架的影响，最安全的做法就是以多种方法来测试和比较结果。

第三部分

决策模型

虽然决策和判断具有情境依赖性，但早期的决策模型常常假定，人们的态度和偏好是固定的，且不会随着引导方式的改变而改变。决策者被当成“理性人”来看待，他们追求效用最大化和自我利益，而且遵循理性行为的原则。这一部分的章节将对这些模型进行回顾，讨论它们各自的缺点以及其他一些决策理论。

第 7 章

期望效用理论

1713 年，一位名叫尼古拉斯·伯诺利（Nicolas Bernoulli）的瑞士籍教授提出了一个非常有意思的问题。翻译过来的意思大概是：伯诺利感兴趣的是，根据下面两个原则，人们愿意为玩这个游戏付多少钱：（1）掷硬币，直到硬币出现字，（2）如果第一次掷硬币就出现字，玩游戏的人可以得到 2 美元，如果第二次才出现字，玩游戏的人可以得到 4 美元，如果第三次才出现可以得 8 美元，如果第四次才出现可以得 16 美元，以此类推。为了参加这个游戏，你愿意付多少钱？（读者调查的第 30 题，你可以检查自己的答案。）大多数人都表示，只愿意花几美元来玩这个游戏。

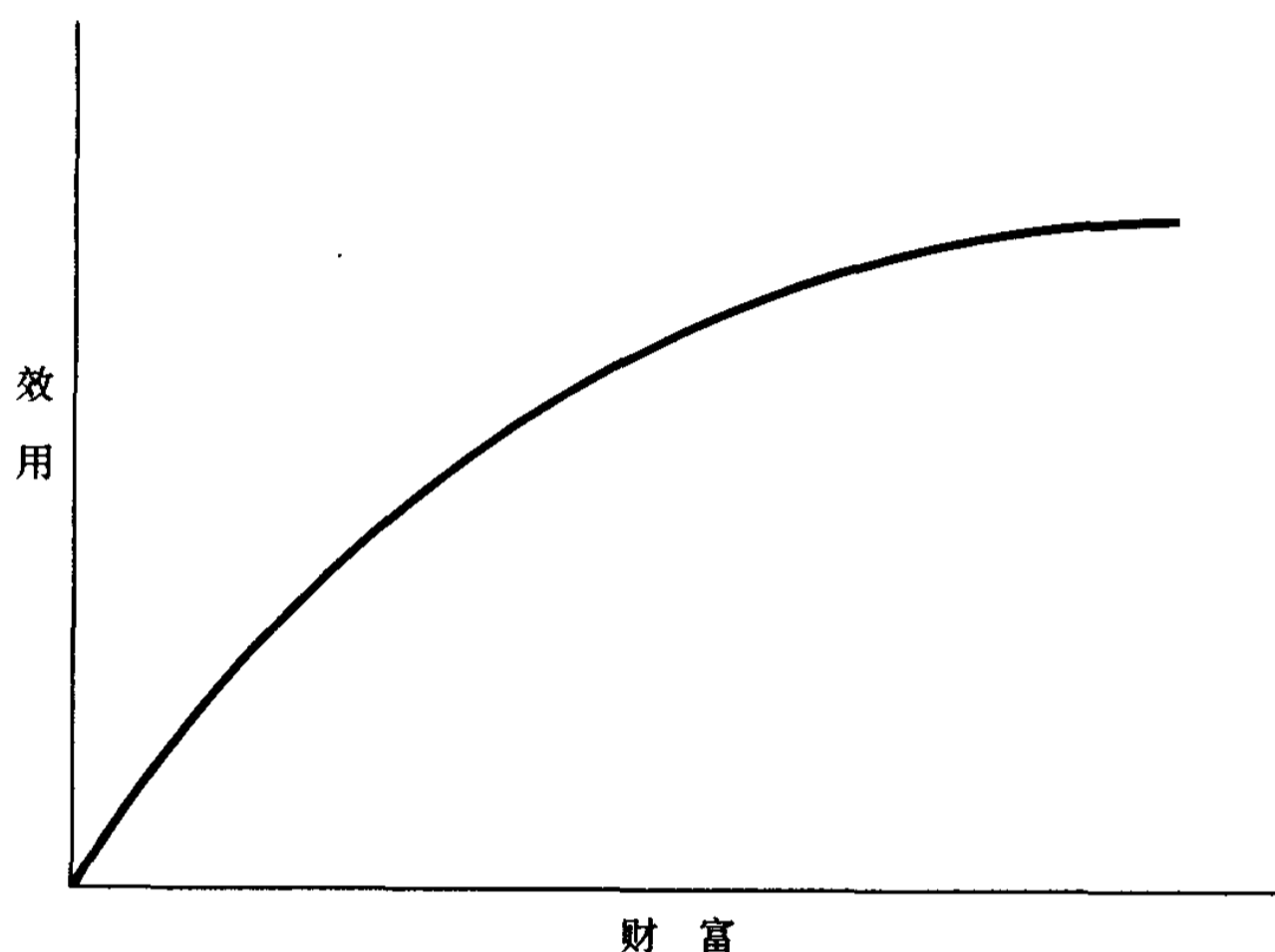
由于是伯诺利第一次提出了这个问题，因此它被称做“圣彼得堡悖论”。之所以称它为一个悖论，是因为这个游戏的期望值（如果游戏无限次地进行下去，其期望平均值）是无限的，但却几乎没有人肯付一大笔钱来玩这个游戏。为了说明平均回报是无限的，我们可以用每种结果的回报乘以其概率来计算伯诺利游戏的期望值。^{*} 第一次抛硬币就出现字（玩游戏的人可以得到 2 美元的回报）的概率为 $1/2$ ，第一次出现正面第二次出现字（玩游戏的人可以得到 4 美元的回报）的概率为 $1/4$ ，出现两次正面再出现字（玩游戏的人可以得到 8 美元的回报）的概率为 $1/8$ ，推而广之，期望值（EV）为（ K = 掷硬币的次数）：

$$\begin{aligned} \text{EV (游戏)} &= (1/2) (2 \text{ 美元}) + (1/4) (4 \text{ 美元}) + \\ &\quad (1/8) (8 \text{ 美元}) + \cdots + (1/2)^K (2 \text{ 美元})^K \\ &= 1 \text{ 美元} + 1 \text{ 美元} + 1 \text{ 美元} + \cdots + 1 \text{ 美元} \\ &= \text{无限多的钱} \end{aligned}$$

那么问题就在于，为什么人们不愿意付更多的钱来玩这个游戏，以得到无限大的回报。就在尼古拉斯·伯诺利提出这个问题的 25 年之后，他的堂弟，

* 本书的这一部分相对于其他部分来说，更加理论化和数学化一些。因此，有的读者可能觉得这一部分比前几章所讨论过的话题更具有挑战性。如果你对“期望值”一类的概念不太熟悉，不要放弃——大部分的观点即使没有什么数学知识也可以理解，以后的几章也很少再用到数学。

数学家丹尼尔·伯诺利对此进行了解释，并成为现代决策理论的起源。丹尼尔·伯诺利认为，金钱的价值或者说效用随着获取金额的增多（或者是已有财富的增多）而递减。丹尼尔·伯诺利书中的第24页写道：“同样是1 000 达克特，对于一个穷人的效用要比它对于一个富人的效用要大得多，虽然二者的金额完全相同。”伯诺利认为，金钱的价值可以这样来表示：



有了这样一个假定，即金钱效用随着财富增加而递减，伯诺利就能够解释，圣彼得堡游戏的期望值并非是有限的。

期望效用理论

并不是所有的学者都同意丹尼尔·伯诺利的理论确实能够解释圣彼得堡悖论（例如，Lopes, 1981；Weirich, 1984），但他的边际效用递减理论却为后来的选择行为理论奠定了基础。其中最为著名的便是“期望效用理论”。这是由约翰·冯·诺伊曼（John von Neumann）和奥斯卡·摩根斯坦（Oskar Morgenstern）于1947年提出的。冯·诺伊曼和摩根斯坦认为，期望效用理论是一种标准化行为理论。也就是说，经典的效用理论并不是要描述人们的实际行为，而是要解释，在满足一定的理性决策条件下人们将如何表现自己的行为。

这一理论的一个主要目的是为理性决策提供一套明确的基本假设或者说公理。当冯·诺伊曼和摩根斯坦定义这些公理后，决策研究者们就能够将期望效用理论计算出的数学预测结果与决策者的真实行为相比较。当研究者发现某一公理无法满足时，就可以对这一理论作修改并做出新的预测。这样，决策研究者们可以反复比照理论与实际，提出更多新的学说来。

理性决策的公理有哪些？期望效用理论的大多数公式都至少包含了以下六条原则中的一条：

- **有序性。**首先，决策者可以对任意两个备选方案进行比较。他们应该要么偏好其中一个，要么对两个都无所谓。
- **占优性。**理性的个体永远都不会采取一个被其他策略占优的策略（对于我们来说，采取策略等同于做决策）。如果一项策略与其他策略相比较，至少在某一方面比其他策略都好（意味着产生更大的效用），而且在其他方面与其他策略一样好，这项策略称之为弱式占优。如果一项策略与其他策略相比较，在所有方面都比其他策略好，这项策略称之为强式占优。举例来说，如果汽车 A 在质量、成本和外观方面来说都比汽车 B 更好，那么汽车 A 相对于 B 来说就是强式占优。但如果汽车 A 只是在质量方面比汽车 B 好，而在成本和外观方面与汽车 B 差不多，那么汽车 A 就属于弱式占优。根据期望效用理论，完全理性决策者绝不会选择一个占优策略，即使所选策略是一个弱式占优策略。
- **相消性。**如果两个有风险的备选方案所可能产生的结果中包含了某些完全相同且具有相同概率的结果，那么在对这两个方案进行选择时，就应该忽略那些结果的效用。也就是说，在进行选择时只需要比较那些不同的结果，而不是比较两种选择都具有的相同结果。相同因素应该相互抵消。
- **可传递性。**如果一个理性决策者在方案 A 和方案 B 中更偏好方案 A，在方案 B 和方案 C 中更偏好方案 B，那么这个人在方案 A 和方案 C 中肯定更偏好方案 A。
- **连续性。**对于任何一组结果，如果出现最好结果的概率非常大，决策者总是偏好在最好和最坏的结果中进行赌博而不是选择一个中间值。举例来说，如果完全财务损失的概率是 $1\ 000\ 000\ 000\ 000\ 000\ 000\ \dots\ 000$ 分之一，那么一个理性决策者肯定会偏好在 100 美元和完全财务损失中进行赌博，而不是选择一个确定的 10 美元的盈利。
- **恒定性。**恒定性原则表示，决策者不会受到备选方案的表现方式的影响。比如说，理性决策者不会在一个复杂的赌博（例如，两阶段的彩票，每一阶段的获奖概率为 50%，如果两阶段都获奖将得到 100 美元）和一个简单的赌博（例如，一次性彩票，有 25% 的概率赢得 100 美元）中有什么偏好。

冯·诺伊曼和摩根斯坦从数学上证明，如果决策者违背了这些原则，期望效用就无法达到最大化。举例来说，如果违背了可传递性原则，那么你对于方案 A、B 和 C 的偏好就具有不可传递性。你偏好方案 A 比方案 B 更多一些，偏好方案 B 比方案 C 更多一些，偏好方案 C 比方案 A 更多一些。这就意味着，我可以让你选择方案 C，也可以提议你放弃方案 C 并付一分钱而得到方案 B。因

为你偏好方案 B 比偏好方案 C 更多一些，你当然乐意接受我的提议，并付给我一分钱。

现在你选择了方案 B。同样的，我也可以提议你放弃方案 B 并再付一分钱而得到方案 A（相对于方案 B 来说，你更偏好方案 A 一些）。这样你就得到了方案 A。但是，由于你的偏好具有不可传递性，我仍然可以提议，你放弃方案 A 并再付一分钱而得到方案 C（相对于方案 A 来说，你更偏好方案 C 一些）。其结果就是，你还是回到了最初的状态，却损失了 3 分钱（或者是 3 美元、3 000 美元，或者其他）。也就是说，我可以继续利用偏好所具有的不可传递性，把它作为摇钱树，直到骗光你所有的钱。在以后几章中，我们将讨论违背理性行为中的可传递性原则和其他原则的一些例子。

理论的扩展

继冯·诺伊曼和摩根斯坦提出期望效用理论后，许多其他理论家对此进行了扩展，提出了多个变式。其中最著名的一个变式是由伦纳德·萨维奇（Leonard Savage）于 1954 年提出的“主观期望效用理论”。萨维奇的理论与冯·诺伊曼和摩根斯坦的理论相比的最大区别在于，萨维奇考虑到了主观的，个人的因素。1954 年以前，期望效用理论中使用的概率都是经典理论中的客观概率（例如，以相对频率为基础）。萨维奇将这一理论进行了推广，将人们对某个事件可能发生的主观概率也纳入进来。

如果客观概率不可能预先得知或是这种结果只会发生一次，这种推广就显得十分重要了。比如说，在主观期望效用理论的框架下，如果一些不可重复事件的发生可能性（如核战争）的概率无法通过计算相对频率来得到，就可以进行人为的估计。而经典效用理论则无法估计“核战争爆发的可能性”究竟有多大。

其他理论家也对经典效用理论进行了补充和完善。例如，邓肯·卢斯（Duncan Luce, 1959）和其他人一起建立了一个他们称之为“随机的”选择模型——这一模型认为偏好具有随机成分。在随机模型出现之前，效用理论家很难解释清楚为什么头一天偏好喝汤而第二天却偏好沙拉。卢斯解释这一问题的方法是，将对汤和沙拉的偏好作为一种随机概率而不是 100% 固定的选择。

彼得·菲什伯恩（Peter Fishburn, 1984）、尤德尔·卡梅卡（Udar Karmarkar, 1978）、约翰·佩恩（John Payne, 1973）和克莱德·库姆斯（Clyde Coombs, 1975）等人对期望效用理论进行了更进一步的扩展，并提出了其他一些替代理论。虽然期望效用理论常常被作为一个整体理论来进行探讨，但实际上期望效用理论指的是一个理论系，而不是某个特定的理论（虽然“期望效用理论”也常常作为冯·诺伊曼和摩根斯坦理论的一个简称）。

结 论

保罗·休梅克(1982, p. 529)对期望效用理论及其变量做了一次全面回顾,他认为:“可以毫不夸张地说,期望效用理论是自二战以来决策研究的主要范式。”确实,在决策理论中,这一理论所引起的研究和讨论比其他理论所引起的研究和讨论都要多。但是,正如第8章所示,经典期望效用理论却存在一些棘手的问题和悖论。这些问题使得部分决策研究者放弃了期望效用理论而去寻找更有用的替代理论。

第 8 章

理性决策的悖论

虽然期望效用理论的原则听起来好像是合理的，但在许多情况下决策者却会违背这些原则。例如，第 6 章的框架效应表明，决策者常常会违背恒定性原则（如果要进一步了解违背恒定性原则和占优性原则的例子，参见特韦尔斯基和卡尼曼，1986）。在这一章中，我们主要讨论违背相消性原则和可传递性原则的例子。

阿莱悖论

根据相消性原则，在两个方案中做出选择应该只取决于它们之间的差异——而不是两种方案所具有的共同点。两种方案所具备的相同因素不应该影响到理性人所做的选择。比如说，你要在两辆车之间进行选择，而它们具有同样的质量，那么质量这一因素就不应该影响到你的选择。

从表面看来，这一原则看似有理；如果两辆车有同样的质量，为什么要让质量的高低来影响你的选择呢？理性决策者应该是根据两种方案的不同方面来进行选择。但 1953 年，一位名叫莫里斯·阿莱斯（Maurice Allais）的法国经济学家发表了一篇文章，向相消性原则提出了挑战。在这篇文章中，阿莱斯（1953）简要概述了他的阿莱悖论——这一悖论显示了违背相互抵消原则的情况。让我们来看一看这一悖论是怎样起作用的。

假设我提供给你两种选择，A 和 B。如果你选择 A，你一定能够得到 100 万美元。但如果你选择 B，就有 10% 的概率得到 250 万美元，有 89% 的概率得到 100 万美元，1% 的概率什么也得不到。也就是说，你的选择如下：

选项 A：肯定会获得 100 万美元

选项 B：10% 的概率获得 250 万美元，89% 的概率获得 100 万美元，1% 的概率什么也得不到

你会做出什么样的选择？（参见读者调查中第 28 题（1）的答案。）实验结果是，即使选项 B 的期望值大于 100 万美元，大多数的人仍然会选择 A 以得到一个确定的数额。将选项 B 中的可能结果与其概率相乘后，可以计算出选项 B

的期望值 (EV) 实际上为 114 万美元, 比选项 A 的期望值更高:

$$\begin{aligned} EV(B) &= (0.10)(2\,500\,000 \text{ 美元}) + (0.89)(1\,000\,000 \text{ 美元}) \\ &\quad + (0.01)(0 \text{ 美元}) \\ &= 1\,140\,000 \text{ 美元} \end{aligned}$$

虽然如此, 大部分的人仍然愿意获得一个确定的 100 万美元的数额。

现在, 假设我提供给你另外一个选择。这一次, 选项 A 有 11% 的概率获得 100 万美元, 89% 的概率什么也得不到; 而选项 B 则有 10% 的概率获得 250 万美元, 90% 的概率什么也得不到。也就是说, 你的选择如下:

选项 A: 11% 的概率获得 100 万美元, 89% 的概率什么也得不到
选项 B: 10% 的概率获得 250 万美元, 90% 的概率什么也得不到

这次你会如何选择? (参见你在读者调查中第 28 题 (2) 的答案。) 大多数人会选择 B。他们通常会这样认为, 10% 与 11% 盈利概率的差别很小, 但 100 万美元和 250 万美元的差别却很大。而且, 选项 B 的期望值也最大。选项 B 的期望值是 250 万美元的 1/10, 也就是 25 万, 这是选项 A 的期望值 (100 万美元的 11% 是 11 万美元) 的两倍多。问题或者说是悖论在于, 在第一种情况下选择 A 的人在第二种情况下也应该选择 A——否则就违背了相互抵消的原则。

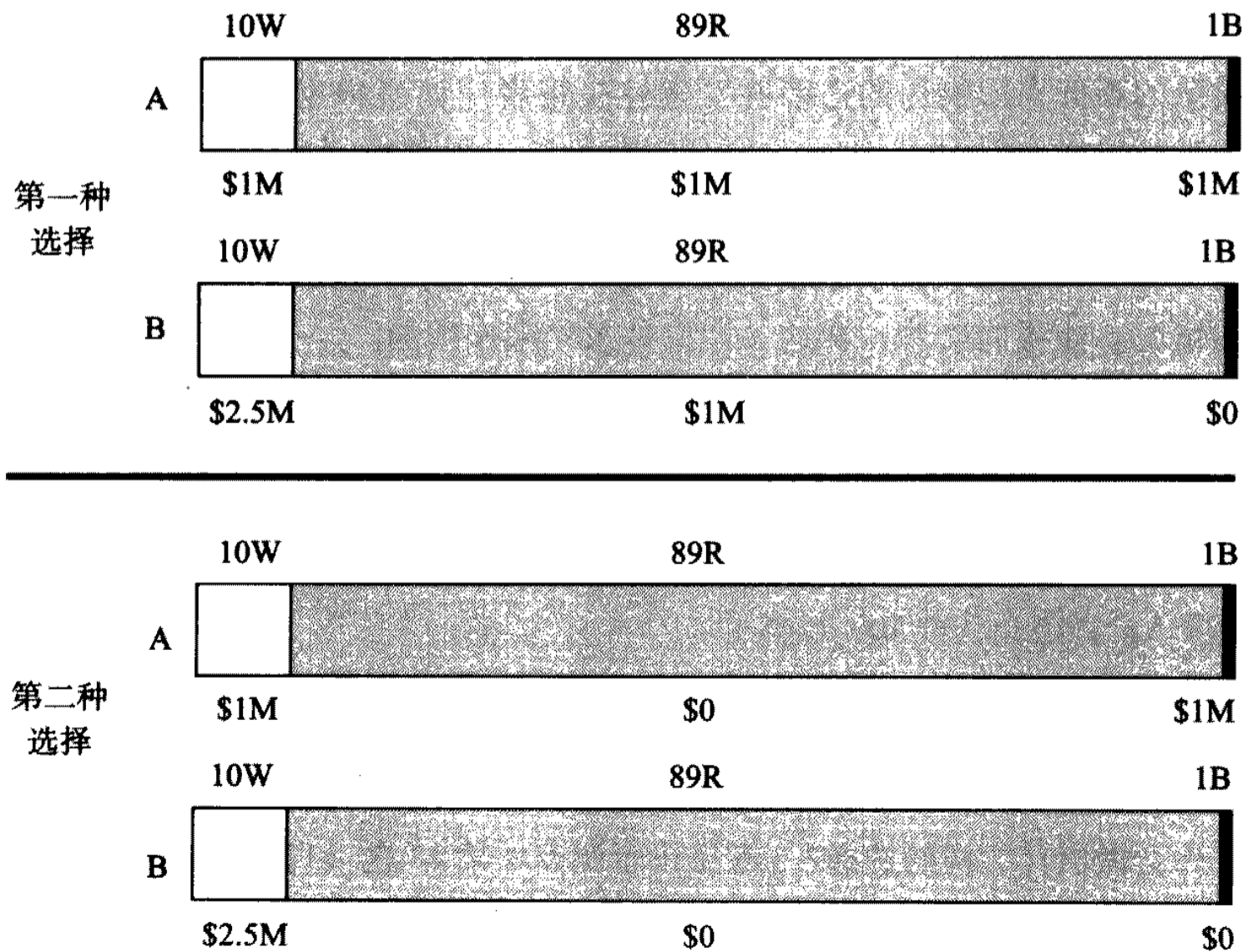


图 8.1 阿莱悖论

为了说明这一点，现在假设每个选择的结果是通过从罐子里的100个彩色小球中随机抽取来决定：89个红球（R），10个白球（W），1个蓝球（B）。在第一种情况下，选项A代表抽到红球、白球或是蓝球中的任意一个就可以获得100万美元（也就是说，不管抽到什么都可以得到100万美元）。选项B则代表抽到红球可以得到100万美元，抽到白球可以获得250万美元，抽到蓝球则什么也没有。按照同样的逻辑，在第二种情况下，选项A代表抽到白球和蓝球都可以获得100万美元，抽到红球则什么也得不到。而选项B则代表抽到白球可以得到250万美元，抽到红球或蓝球则什么也没有。

由此看来，你会发现两种情况实际上提供了完全相同的选择，只是在第一种情况下无论你选择哪一个选项，抽到红球你都可以获得100万美元，而在第二种情况下无论你选择哪一个选项，抽到红球都表示什么也没有。在两种情况下选项A都表示，抽到白球和蓝球可以获得100万美元；选项B则表示，抽到白球和蓝球分别可以获得250万美元和0美元。第一种情况下的选项A和第二种情况下的选项A相比较，除了89%的概率可以多获得100万美元以外完全相同；第二种情况下的选项B和第二种情况下的选项B相比较，除了89%的概率可以多获得100万美元以外也完全相同。

所以说，附加的等同结果——在第一种情况下红球代表100万美元，而在第二种情况下红球则代表一无所获——使得人们做出了不同选择。这种差异恰恰违背了相同因素相消性的原则。因为根据相消性原则，在对两种选项做出选择时，只应该以它们的差异为依据，而不以它们的共同点为依据。

埃尔斯伯格悖论

另外一个违背相消性原则的著名例子则由丹尼尔·埃尔斯伯格（Daniel Ellsberg, 1961）提出。埃尔斯伯格悖论是这样的：假设缸里有90个球，30个是红色的，剩下60个要么是黑色要么是黄色，其比例未知。从缸里抽出一个球，球的颜色将决定你的收益（见图8.2a的列表）。

你会在哪种颜色的球上下注——红色还是黑色？大多数人都选择了红球，以避免黑球和黄球混合比例的不确定性。但如果给出的收益列表如图8.2b所示，你又会会在哪个颜色上下注呢？在第二种情况下，大多数的人偏好在黑球或黄球上下注，而不是在红球或黄球上下注。这也是为了避免黑球和黄球混合比例的不确定性。也就是说，大多数的人在第一种情况下会选择选项1，而在第二种情况下则会选择选项2。

但根据相消性原则，人们在两种情况下应该选择同样的选项。这可以从图8.2看出，两种情况的结果只有一点不同，那就是在第一种情况下黄球表示什么也得不到，而在第二种情况下则表示100美元。由于在两种情况下选项1和选项2中的黄球都代表相同的数额（在第一种情况下代表0美元，在第二种情

图 8.2a 埃尔斯伯格悖论第一种情况下的结果

下注选项	30 个球		60 个球	
	红 色	黑 色	黑 色	黄 色
选项 1: 一个红球	100 美元	0 美元	0 美元	0 美元
选项 2: 一个黑球	0 美元	100 美元	100 美元	0 美元

图 8.2b 埃尔斯伯格悖论第二种情况下的结果。惟一的变化是黄球现在表示 100 美元而不是 0 美元。

下注选项	30 个球		60 个球	
	红 色	黑 色	黑 色	黄 色
选项 1: 一个红球或黄球	100 美元	0 美元	0 美元	100 美元
选项 2: 一个黑球或黄球	0 美元	100 美元	100 美元	100 美元

况下代表 100 美元), 因此两种情况下黄球的价值都不应该影响到对选项 1 和选项 2 的选择 (正如相同的质量不应该影响到对两辆车的选择一样)。但与期望效用理论相悖的是, 人们常常在两种情况下做出不同的选择。

不可传递性

另外一个理性决策原则是备选方案应该具有可传递性, 也就是说, 如果一个决策者在结果 A 和 B 中偏好 A, 在结果 B 和 C 中偏好 B, 那么在结果 A 和 C 中就应该偏好 A。第 7 章曾讨论过, 一个具有不可传递性偏好的决策者容易被别人当做“摇钱树”来使。图 8.3 给出了有关不可传递性的另外一个例子。

假设你必须在 3 个求职者中做出选择 (图 8.3 中的求职者 A、B 和 C), 你已经有了关于他们才智和经验方面的信息。再进一步假设你的决策原则如下: 如果两人的 IQ 差距超过 10 分, 选择更有才智的人; 如果两人的 IQ 相差不超过 10 分, 选择更有经验的人。

这听起来倒不失为一个合理的原则, 但如果你遵循这一原则就有可能面临困境。如果将求职者 A 与 B 相比较, 我们应该选择 B, 因为 A 和 B 的才智相差不超过 10 分, 而 B 却比 A 有经验得多。同样, 如果将求职者 B 与 C 比较, 我们应该选择 C, 因为 B 和 C 的才智相差不超过 10 分, 而 C 却比 B 有经验得多。但如果我们将求职者 C 与 A 相比较, 我们应该选择 A, 因为 A 的 IQ 比 C 的 IQ 高出了 10 分以上。也就是说, A 和 B 中应该选择 B, B 和 C 中应该选择 C, 而 A 和 C 中则应该选择 A。这种不可传递性出现的原因在于, 决策标准有两个维度——才智和经验——而这两个维度都是小幅递增且呈负相关的。

现实中的人们是否会违背可传递性原则呢? 1969 年阿莫斯·特韦尔斯基发表了他的研究, 在这项研究中有 1/3 的被试表现出了不可传递性。特韦尔斯基

图 8.3 以下的决策原则使得很难在 A、B、C 3 个求职者中做出选择，因为此时的偏好是不可传递的：如果两人的 IQ 相差超过 10 分，选择更有才智的人；如果两人的 IQ 相差不超过 10 分，选择更有经验的人。

		维 度	
		才智 (IQ)	经验 (年限)
求职者	A	120	1
	B	110	2
	C	100	3

向 18 名哈佛研究生展示了图 8.4 中所列出的 5 种赌博游戏。正如你所见，每个赌博游戏的期望值随着获胜概率递增，随着获胜金额而递减。随机挑选一对赌博游戏展示给学生，要求他们选出更喜欢的一种游戏。游戏的配对有 10 种可能 (A 和 B、A 和 C，等等)。学生做了 3 次选择后，特韦尔斯基从中挑出 8 名表现出不可传递性倾向的被试，让他们每周去一次他的实验室，进行一项为期 5 周的深入研究。

他发现，有 6 名学生表现出可信度较高的不可传递性。当两个选项具有十分相近的获胜概率时 (如赌博游戏 A 和 B)，被试会选择具有更高获胜金额的选项。与之相反的是，当两个选项的获胜概率相差悬殊时 (如赌博游戏 A 和 E)，被试会选择具有更高获胜概率的选项。也就是说，A 比 B 好，B 比 C 好，C 比 D 好，D 比 E 好，而 E 却比 A 好。特韦尔斯基 (1969) 采用了与上面所述的求职者选择相类似的问题，也发现了不可传递性的存在。

不可传递性并不只是实验上的一个新奇事物，它对决策者来说也具有非常重要的启示。举例来说，考虑一下决策研究者们所熟悉的“委员会问题”。在一个典型的委员会问题例子中，一个委员会里一般有 5 个成员：安、鲍伯、辛迪、丹和埃伦。他们的任务是聘用一名新教授，3 个候选人如图 8.5 所示。

假设你是委员会的主席，你知道每个人的偏好，并且你希望能够控制投票以使得阿尔·爱因斯坦被选中。你应该怎么做？

图 8.4 阿莫斯·特韦尔斯基 (1969) 采用以下的赌博游戏对偏好的不可传递性做了实验。将每个游戏的获胜概率和获胜金额相乘，即可得到每个游戏的期望值 (EV)

赌博游戏	获胜概率	获胜金额 (美元)	期望值 (美元)
A	7/24	5.00	1.46
B	8/24	4.75	1.58
C	9/24	4.50	1.69
D	10/24	4.25	1.77
E	11/24	4.00	1.83

图 8.5 在一个典型的委员会问题中，偏好排序如下。数字越低代表的偏好程度越高（例如，安在乔·施莫和简·多伊中更喜欢乔·施莫，而在简·多伊和阿尔·爱因斯坦之间更喜欢简·多伊）

候选人	委员会成员				
	安	鲍伯	辛迪	丹	埃伦
乔·施莫	1	1	2	3	3
简·多伊	2	3	3	1	1
阿尔·爱因斯坦	3	2	1	2	2

答案就是你应该避免在爱因斯坦和多伊之间直接投票，因为 5 个委员中有 3 个人都更加偏好多伊而不是爱因斯坦（安、丹和埃伦）。相反，你应该让他们就施莫和多伊之间选择谁进行投票，等到施莫赢了以后再安排在施莫和爱因斯坦之间进行投票。另一方面，如果你更希望雇佣多伊，你应该先就施莫和爱因斯坦之间进行投票，再在爱因斯坦和多伊之间进行投票。由于在成对比较中委员们的偏好具有不可传递性，因而安排投票进程的人能完全控制结果。

偏好逆转

违反不可传递性原则还不是最糟的，因为有的时候甚至连偏好都会随着情境的变化而发生逆转。萨拉·利希滕斯坦和保罗·斯洛维奇于 1971 年发表的实验是有关偏好逆转的最早的实验之一。利希滕斯坦和斯洛维奇推断，当需要从一对赌博游戏中做出选择时，其心理过程可能有别于单独在每种赌博游戏上出价的过程（也就是说，对一项赌博游戏单独设定一个金额）。具体来说，他们认为二选一的选择是由赌博的获胜概率决定的，而单独出价则主要是由输赢的金额决定的。

为了验证这一假设，他们进行了 3 个实验。在每一个实验中，他们首先向被试演示了几组赌博游戏。每组中的两个游戏具有十分相近的期望值，但一个游戏具有很高的获胜概率，而另一个则具有很高的获胜金额（见图 8.6）。在每组中做出选择后，被试还需要在单独考虑每种赌博的情况下出价。被试被告知，他们有一张玩赌博游戏的门票，他们需要说出愿意卖出这张票的最高价和最低价。

在第一次实验中，大学生们需要指出他们在每组赌博中更偏好哪个游戏，以及他们愿意以怎样的价格卖出其持有的门票。利希滕斯坦和斯洛维奇衡量偏好逆转的方法是，假定当两个游戏配对出现时，被试所选择的是获胜概率高的游戏，则计算高获胜金额游戏门票售价超过高获胜概率游戏门票的次数百分比。利希滕斯坦和斯洛维奇发现，73% 的被试总是表现出这种偏好逆转。第二

图 8.6 偏好逆转的实验材料 (引自 Lichtenstein & Slovic, 1971)

组	高获胜概率	期望值	高获胜金额	期望值
1	99% 的概率获得 4 美元 1% 的概率输掉 1 美元	3.95 美元	33% 的概率获得 16 美元 67% 的概率输掉 2 美元	3.94 美元
2	95% 的概率获得 2.5 美元 5% 的概率输掉 0.75 美元	2.34 美元	40% 的概率获得 8.5 美元 60% 的概率输掉 1.5 美元	2.5 美元
3	95% 的概率获得 3 美元 5% 的概率输掉 2 美元	2.75 美元	50% 的概率获得 6.5 美元 50% 的概率输掉 1 美元	2.75 美元
4	90% 的概率获得 2 美元 10% 的概率输掉 2 美元	1.6 美元	50% 的概率获得 5.25 美元 50% 的概率输掉 1.5 美元	1.88 美元
5	80% 的概率获得 2 美元 20% 的概率输掉 1 美元	1.4 美元	20% 的概率获得 9 美元 80% 的概率输掉 0.5 美元	1.4 美元
6	80% 的概率获得 4 美元 20% 的概率输掉 0.5 美元	3.1 美元	10% 的概率获得 40 美元 90% 的概率输掉 1 美元	3.1 美元

次实验基本上是第一实验的重复，只是在出价程序上略有不同。而第三次实验也发现，即使对每个被试单独进行了详尽仔细的指导，而且实实在在地进行了赌博游戏，人们也表现出了稳定的偏好逆转。

当然，利希滕斯坦和斯洛维奇发现的偏好逆转是在严格的实验情境之下产生的，它是否存在于实验情境之外的现实生活中依旧是一个问题。为了验证这一问题，利希滕斯坦和斯洛维奇在拉斯维加斯的赌场中重复了他们的实验。通过一台计算机和轮盘机，他们搜集了 44 名参与赌博人的数据（包括 7 名职业发牌人）。

实验结果十分令人震惊。在高获胜概率和高获胜金额的游戏选择了高获胜概率游戏的被试当中，81% 的人在高风险金额游戏上的出价更高。这一逆转比例甚至比第一次实验中发现的逆转比例还要大。所以说，偏好逆转似乎不仅仅存在于实验情境之中；对于那些有经验也有金钱激励的决策者来说，这一现象同样存在。

自这些早期的实验之后，一些研究重复和拓展了利希滕斯坦和斯洛维奇的发现 (Grether & Plott, 1979; Schkade & Johnson, 1989; Slovic, Griffin & Tversky, 1990; Slovic 和 Lichtenstein, 1983; Tversky, Slovic, & Kahneman, 1990)。偏好逆转很难改变，而且它们也不随着金钱激励而消失 (Tversky, Slovic, & Kahneman, 1990)。当人们需要在两种赌博中做出选择时，他们通常会特别关注获胜概率，而如果他们需要考虑到每种赌博的价值时，则往往看重其可能获得的金额。

违背期望效用理论真的就是不理性的吗

毫无疑问，人们常常违背期望效用理论的一些原则，但这种违背真的就意味着人们是不理性的吗？这些发现是否说明人们的决策是不理性的？

答案当然是否定的。这是因为，我们并不知道人们犯错误的成本与人们遵循这些理性原则（如相消性原则和可传递性原则）的成本相比（这一点将在以后做进一步讨论），哪个更大。正如利希滕斯坦和斯洛维奇所说：“被试为了将复杂任务简单化而采用的捷径实际上是非常有效的，因为他们减少了认知努力，而且结果与最佳策略相比相差无多。在使用这些捷径时决策者通常假设，这个世界（并不像我们进行的实验那样），并不想要利用他的这些捷径。”一项逻辑并不是那么严密的决策有时候却可能是理性的。因为如果从长期来看，这项决策相对于正常决策来说确实是提供了一条快而简单的捷径来实现效用最大化。

结 论

本章中所提到的研究看起来好像是针对理性原则的各项原则进行了单独的验证。但正如邓肯·卢斯（1990）所说，有时候很难区分出，一个实验究竟是在检验哪一项原则。比如说，偏好逆转通常可以用违背可传递性原则来解释，但近期的一些研究则表明，用违背恒定性原则来解释偏好逆转也许会更好（Bostic、Herrnstein, & Luce, 1990; Tversky、Sattath, & Slovic, 1988; Tversky、Slovic, & Kahneman, 1990）。但不管这些争论最后如何解决，有一点却是非常清楚的，那就是期望效用理论还不足以描述人们真实的决策行为。

自冯·诺伊曼和摩根斯坦研究之后，许多决策理论家试图将期望效用理论发展为一个描述性决策模型。但这些努力常常遭受失败。由于违反相消性原则、可传递性原则、恒定性原则和占优性原则情况的大量存在，因而许多以前是效用理论忠实支持者的理论家们都开始转向其他的决策模型。第9章将对其中一些模型进行讨论。

第9章

描述性决策模型

1977年杰伊·拉索（Jay Russo）发表了一项有关超市商品单位定价效应（单位价格指的是每盎司、每克或是其他每单位的商品明细价格）的实地研究。在这项研究中，拉索发现了一些有关人们购物方式的有趣现象。首先，当货架上的标签包含了单位价格信息时，购物者就会节约1%的单位采购成本。这一节省不是靠购买更低价格品牌的商品来实现，而是通过加大采购量来实现（1%好像不值多少钱，但请记住，在超市里的消费达数十亿美元）。第二，拉索发现，如果超市列出了不同品牌的价格比较，购物者的单位采购成本平均就会节约3%。在这种情况下，购物者的节省主要来自购买更便宜一些的品牌。

后一个发现确实有些令人吃惊，因为增加单价对比实际上没有提供任何新的信息，只是列出了已经标明的价格而已。根据期望效用理论的恒定性原则，决策不应该受到选项的表现方式的影响。但拉索（1977）发现，列出商品单价表会对消费者产生巨大影响。当不同品牌的单价同时出现在一张单子上时，购物者倾向于购买更便宜的品牌。超市可以通过将商品单价同时列出来影响消费者的选择。

满 意

期望效用理论做了一系列简化的假设，以便进行严格的数学分析。期望效用理论通常都会假设，决策者对过程中每一备选方案的结果及其概率拥有完全信息，而且决策者能够理解这些信息，能够间接或直接地推算出每一备选方案的有利和不利方面。最后，这一理论还假定决策者会在这些推算结果中做出比较，并选择能够实现效用最大化的方案。

但显而易见的是，决策者并不总是这样行事。有关备选方案的信息常常会缺失，或者本来就是不确定的，而知觉则往往具有高度的选择性，记忆又充满了偏差。有关不同方案的结果常常被错误地理解，而且正如拉索的研究所示，在没有协助情况下的决策者有时候并不去比较所有可能的结果。所以说，期望效用理论虽然是一个有用的标准化决策模型（在一定假设被满足的条件下，有关理性个体如何行为的模型），却并不是一个很好的描述性模型（实际中人们如何做决策的模型）。如果要描述人们实际的决策行为，就有必要考虑其他理

论模型。

诺贝尔奖获得者赫伯特·西蒙（Herbert Simon）于1956年提出的模型是最早替代期望效用理论的模型之一。西蒙认为，人们在做决策时，追求的是“满意”而并非最优。满意是指，选择一个最能够满足你需要的方案，即使这一选择并不是理想化或最优化的。比如说，在租房子的時候，人们倾向于寻找一套能够满足某些需求（价格、位置、空间、安全性，等等）的房子。他们并不会彻底搜寻所有可能的房子，然后选出能够产生最高效用的那一套。西蒙（1956, p. 129）写道：“不管有机体在学习和选择情境中的行为多么具有适应性，这种适应能力都远远无法达到经济理论中理想的‘最大化’状态。显然，机体的适应性往往只是能够达到‘满意’，而不是‘最优’。”

前景理论

自西蒙的论文之后，出现了许多替代期望效用理论的理论，但其中最广受承认的是“前景理论”（prospect theory）。这一理论是由丹尼尔·卡尼曼和阿莫斯·特韦尔斯基于1979年共同提出的。与期望效用理论相比，它在许多重要方面都存在差异。

首先，它用“价值”概念替代了“效用”概念。效用通常是从净财富的角度去定义，而价值则是从收益与损失（对某一参照点的偏离）方面去定义。而且，损失的价值函数与收益的价值函数也不同。如图9.1所示，损失的价值函数（横轴以下的曲线）是凸函数，而且更加陡峭一些。相比之下，收益的价值函数（横轴以上的曲线）则是凹函数，而且不是那么陡峭。这些差别引出了一些值得注意的结果。

由于损失的价值函数比收益的价值函数更为陡峭，所以损失比收益显得更加突出一些。例如，损失500美元的感觉比获得500美元的感觉更加强烈（见图9.1）。正如乔治·夸特龙和阿莫斯·特韦尔斯基（1988）所指出的，这种不对称性，或者说“规避损失”，与现任政客相对于其挑战者所具有的优势是一致的（也就是说，投票人常常认为，在政权交替中出现不利改变的潜在损失比出现有利改变的潜在收益要重要得多）。夸特龙和特韦尔斯基（1988）也注意到，这种规避损失倾向会使商讨和谈判变得更加复杂，因为“每一方都将自己的让步看做一种损失，而这种损失比让步带来的成果显得更加突出一些。比如说，在就导弹问题的谈判中，两个超级大国都认为解除导弹设施是安全上的一个重大损失，而实际上对方也会做相应的战备收缩而为自己带来安全利益。但双方在衡量这种损失和利益时，都不约而同地认为损失比利益更加重要一些”。

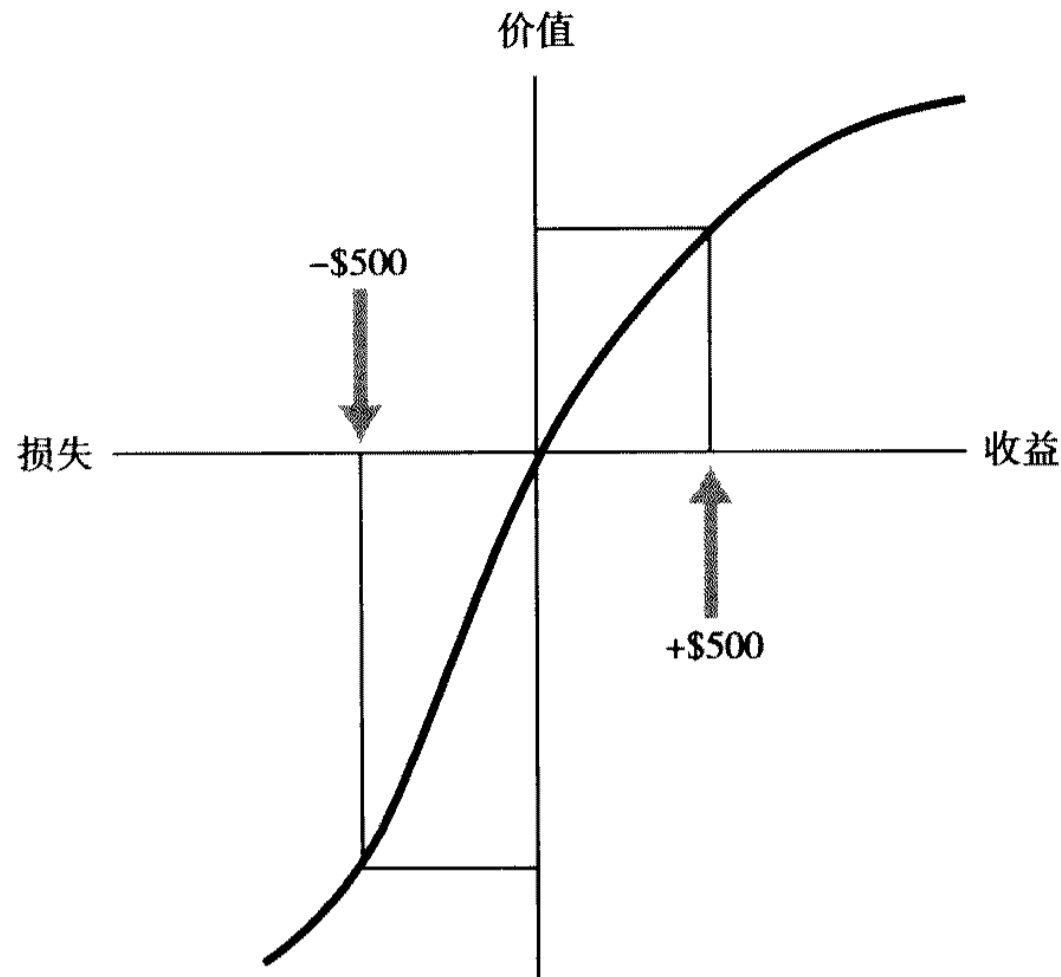


图 9.1 前景理论中的假定价值函数（数据摘自卡尼曼和特韦尔斯基 1979 年的研究）

规避损失所产生的另外一个结果就是“禀赋效应”(endowment effect)。禀赋效应指的是，当一件物品成为人们的禀赋时，它的价值便增加了 (Thaler, 1980)。比如说，当人们需要对自己所拥有的一件东西（如巧克力条、钢笔或是咖啡杯）定价时，这一价格通常比他们所愿意购买这件东西而出的买价更高 (Kahneman, Knetsch, & Thaler, 1990; Knetsch & Sinden, 1984)。根据理查德·塞勒的理论，这一效应之所以出现是因为损失（这一物品）的感觉，比获得同样东西的感觉更加强烈。这种不对称性常常被许多公司加以利用。它们常常向消费者提供试用产品，而对这些试用品的拥有权常常增加了产品的价值，使得消费者很不情愿归还这件产品。

不同于期望效用理论，前景理论认为偏好取决于问题的框架。如果相对于某一参照点，某项结果看起来是一种收益，那么其价值函数将是凹函数，决策者也倾向于规避风险。相反，如果相对于某一参照点，某项结果看起来是一种损失，那么其价值函数将是凸函数，决策者也变得更加偏好风险。以下一组问题是对这一点的一个很好的说明（摘自卡尼曼和特韦尔斯基 1979 的一个实验）：

问题 1：除了你现在所拥有的，再给你 1 000 美元。你现在必须在选项 A 和选项 B 中做出选择：

选项 A： 50% 的概率获得 1 000 美元

选项 B： 肯定会获得 500 美元

70 名被试回答了这一问题，其中 84% 的人选择了肯定会获得 500 美元。正如图 9.1 的上半部分所示，这一答案是十分合理的。因为在这一价值函数中，从 0 美元到 500 美元的价值增加，大于从 500 美元到 1 000 美元的价值增加。如果你认为第二个 500 美元的价值不如第一个 500 美元的价值，那么你不应该接受在第二个 500 美元上具有相等概率的赌注。

问题 2：除了你现在所拥有的，再给你 2 000 美元。你现在必须在选项 C 和选项 D 中做出选择：

选项 C：50% 的概率损失 1 000 美元

选项 D：肯定会损失 500 美元

在这种情况下，70% 的人选择了有风险的选项。正如图 9.1 的下半部分所示，当面临损失时人们通常会变得风险偏好。这是因为损失的价值从 0 美元到 500 美元要高于从 500 美元到 1 000 美元（所以肯定会损失掉 500 美元比 50% 的概率损失掉 1 000 美元更糟糕）。所以说，虽然这两个问题在数字上来说是一样的，却导致了不同的选择。由于价值函数呈“S”形，人们在面对收益时常常是规避风险的，而在面临损失时则常常偏好风险。而且由于价值总是依据一定的参照点来定义的，前景理论——而不像期望效用理论那样——认为，当参照点发生位移时，偏好也会受到影响（正如第 6 章所介绍的框架效应一样）。前景理论与期望效用理论的另一个不同之处在于它们处理结果概率的方式。经典效用理论假定，50% 的客观获胜概率对于决策者来说，就是 50% 的获胜概率。而前景理论则认为，偏好是“决策权重”的一个函数，而这些权重并不总与概率相对应。具体来说，前景理论假定，决策权重常常会强调小概率事件而忽视一般或者高概率事件。正如你在图 9.2 所见，对角线以上的权重函数代表小概率，而对角线以下的则表示高概率。

卡尼曼和特韦尔斯基（1979）利用了下面一组问题来说明人们具有强调小概率事件的倾向：

问题 1：在选项 A 和选项 B 中进行选择。

选项 A：千分之一的概率赢得 5 000 美元

选项 B：肯定会获得 5 美元

在对这一问题进行回答的 72 名被试中，有将近四分之三的人选择了有风险的选项。其实，每天都有很多人在做同样的选择——买彩票。但考虑第二个问题：

问题 2：在选项 C 和选项 D 中进行选择。

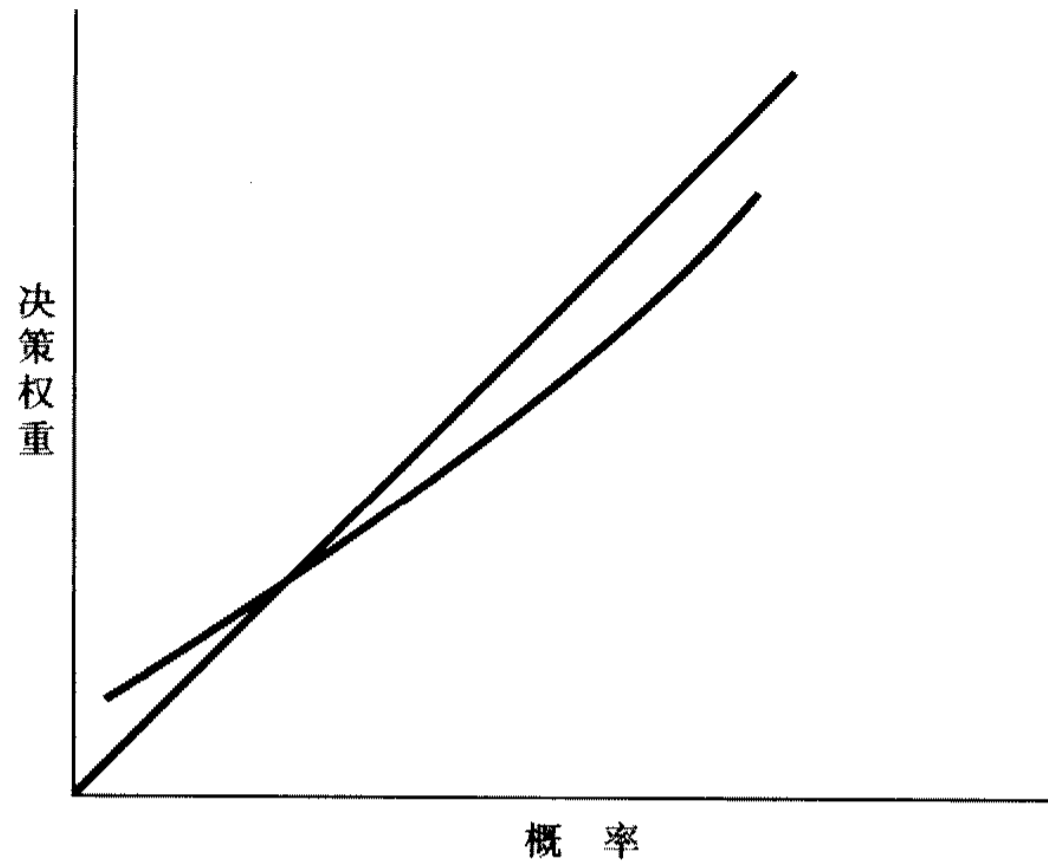


图 9.2 根据前景理论推测的决策权重（数据摘自卡尼曼和特韦尔斯基 1979 年的研究）

选项 C: 千分之一的概率损失 5 000 美元

选项 D: 肯定会损失 5 美元

在对这一问题进行回答的 72 名被试中，则有超过五分之四的人选择了确定的损失。卡尼曼和特韦尔斯基解释说，存在这种偏好是由于人们往往会过于看重大金额的损失——这种倾向对保险公司特别有利。

确定效应

期望效用理论与前景理论的另一个差别在于，前景理论隐含了“确定效应”。确定效应指的是，由同一个因素引起的结果概率减小，在结果最初就确定时所产生的影响要大于在结果最初只是可能时的影响。著名的莫里斯·阿莱斯（1953）期望效用理论反例是最早关于确定效应现象的描述。另外，经济学家理查德·泽克豪泽（Richard Zeckhauser）也用一个非常生动的例子说明了确定效应是如何起作用的。泽克豪泽发现，与从 4 颗子弹中拿走一颗的情况相比，大多数人愿意出更多的钱拿走俄式轮盘赌中惟一的一颗子弹。虽然拿走一颗子弹后，两种情况下被击中的概率都等量减少了，人们仍然会感觉到 0 颗子弹与 1 颗子弹的差异，比 3 颗子弹和 4 颗子弹的差异要重要得多。前景理论隐含了这一效应，而期望效用理论则没有。

为了详细了解前景理论是如何隐含确定效应的，可以参照卡尼曼和特韦尔斯基（1979）有关“概率保险”的一项调查结果。在这项调查中，卡尼曼和特韦尔斯基让大学生被试做这样一个假设：假设你正考虑给你的某项财产上保险

以预防可能发生的损失，如火灾、失窃等。在计算过风险和保费以后，购买保险和不购买保险对你来说是无所谓的。这时候保险公司新推出的一种名为概率保险的新保险计划引起了你的注意。在这项保险计划中，你一开始只需付一半的保费。然而一旦发生损失，你只有 50% 的概率可以获得全部赔偿（如果不要赔偿损失，则退还全部保费）。也就是说，在概率保险的情况下，你只需要付 50% 的保费，也只有 50% 的概率获得赔偿。这个问题看来好像已经为预防损失做了精心准备，但卡尼曼和特韦尔斯基同时也指出，还有许多其他预防性措施可以减少这些不利事件发生的概率（虽然不能完全消除这些概率），比如安装防盗自动警铃，换掉旧轮胎，戒烟，等等。

在回答这一问题时，80% 的学生表示他们不会购买概率保险（可以对比你在读者调查中第 23 题的答案）。对此，卡尼曼和特韦尔斯基解释道，将一项损失的概率（无论是多少，比如 p ）减少到原来的一半（ $p/2$ ）的价值不如将其概率从一半（ $p/2$ ）减至 0 的价值大。人们更愿意消除风险而不是减少风险，即使这项损失发生的概率在两种情况下都等量减少了。前景理论认为，由于决策中常常强调小概率，因而这些小概率事件的重要性总被夸大。因此，用前景理论可以很好地解释这一发现。相反，如果放在期望效用理论的框架下，则会认为概率保险比常规保险更有吸引力（有关这一说法的证明比较复杂，可以参见卡尼曼和特韦尔斯基书中的第 270 页，1979）。

虚假确定效应

除了确定效应，卡尼曼和特韦尔斯基（1981）还讨论了“虚假确定效应”（pseudo certainty）。拟确定效应与确定效应十分类似，只不过这种确定只是一种表象而不是真相。保罗·斯洛维奇，巴鲁克·菲什霍夫和萨拉·利钦斯坦（1982a）也在一篇文章中对虚假确定效应进行了详细的阐述。

斯洛维奇和他的同事向 211 名被试演示了以下两种疫苗接种项目之一。在概率防治条件下，被试需要回答他们是否愿意接受一种对某种疾病只有一半保护作用的疫苗，这种疾病能够传染 20% 的人口。也就是说，这种疫苗能够将受感染的概率从 20% 降低到 10%。只有 40% 的被试表示愿意接受这样一种疫苗。

在虚假确定条件下，被试被告知：有两种互相排斥且得病概率相等的疾病，每种疾病都可能会传染 10% 的人口。疫苗能够完全防治一种疾病的发生却对另一种疾病没有预防作用，因此染病的总体概率也是从 20% 降到了 10%。但在对这一问题的回答中，有 57% 的人表示他们愿意接种这一疫苗。对此，斯洛维奇、菲施霍夫和利钦斯坦（1982）解释说，虚假确定条件下的疫苗更受欢迎，是因为它们看起来好像是消除了某种风险而不仅仅是降低了某种风险。

营销学教授常常利用模拟技术来强调降价的知觉价值。比如说，干洗店如果提出一次洗三件衣服就可以免费洗一件，其效果要比降价 25% 更好。这是因

为，免费服务比折扣服务更有吸引力，即使免费服务并不真正意味着在总价格上有更多的优惠。

后悔理论

正如前景理论所指出的，决策者通常以一个参照点为基础来衡量其拥有的备选方案。这一参照点往往选择最具有普遍意义的一点，但在某些特定情况下，人们却是将某一决策质量的衡量建立在与其他不同决策的后果相比较的基础上。由于这种假想性结果的比较依赖于假设的事件，它有时也被称为“反事实推理”（counter-factual reasoning）（Dunning & Parpal, 1989）。

反事实推理构成了后悔理论（regret theory）的基础。这是一个有关选择的经济学理论，分别由戴维·贝尔（David Bell, 1982, 1985），格雷厄姆·卢姆斯和罗伯特·萨格登（Graham Looms & Robert Sugden, 1982, 1983, 1985）各自提出。卢姆斯和萨格登（1981, p. 820）认为：“后悔理论以两个基本假设为基础：首先，很多人都经历过后悔和欣喜的感觉；第二，在不确定的情况下做决策时，他们会预期这些感觉并将其作为决策时要考虑的一个因素。”比如说，如果人们面临这样一个选择：确定会得到1 000美元，还是通过掷硬币来决定是否获得2 000美元（出现人头）。他们很可能会选择确定的1 000美元，以避免未猜对时的后悔。

与前景理论对这种风险规避倾向的解释不同，后悔理论在解释这一选择时仍然采用了经典效用函数，只是加进了一个新的变量——后悔。加入这一变量后，后悔理论同前景理论一样，也能够解释许多悖论，包括阿莱悖论、埃尔斯伯格悖论、偏好逆转、对概率保险的回避等。卢姆斯和萨格登（1982）提出的后悔理论确实是前景理论之外的另一个选择。但这种后悔预期并不与前景理论相矛盾。而当涉及到死亡风险的决策时（如心脏手术），谈论负面结果的后悔却是毫无意义的。

多属性选择

在许多选择中，结果并不能根据某个单一的标准（如金钱或是患病的风险）来描述。与只有一种评价标准的情况不同，当需要在这些不同标准（如成本和质量）中做出选择时，通常并没有一个客观的最优化决策。因而人们只需要与他们的目标和价值观保持一致，而不是追求客观最优化（Einhorn & Hogarth, 1981）。所以，有关“多属性选择”的研究通常关注人们如何做出决策，而不是这些决策有多好。

在面临多属性选择时，人们会根据问题的类型而采用不同的策略。当决策者只需要在两种方案中选择其一时，他们常常会采用一种被称为“补偿性”的

策略 (Payne, 1982)。采用补偿性策略是指,以某一标准的高价值来弥补另一标准的低价值。比如说,汽车购买者会选择一辆外观时髦的车,即使它的质量相对另一辆来说要差一些。又比如,一个人可能会放弃一项长期稳定的工作而选择在哈佛任教,虽然暂时还没有终身教职,却能享有更高的学术声誉。

有几种方法能够解释这种选择策略 (cf. Hogarth, 1987), 其中一种被称为“线性模型”。在线性模型中,根据每一标准的重要性赋予其权重,然后将各个标准的值加权后得到总体价值指数。比如说,在选择研究生考生时,招生委员会通常会制定一个加权指数,其中包括了平均绩点、考试分数和推荐信。虽然人们通常不会严格地按照线性公式来做决策,但线性决策原则却能够产生与人们实际做出的决策十分相近的结果 (因此,它能够将人们的决策制定模型化,有时甚至能够取代人的决策)。

另一种补偿性策略的决策模型是“差异加法模型”。这一模型与线性模型有些类似,但有一点区别:在线性模型中,是将每一方案的各个标准的值加权,然后在这些方案中作出比较;但在差异加法模型中,则先比较每一标准上各个方案的差异,然后对这些差异赋予权重后进行加总。将注意力集中在差异上至少有两个优点——它不仅大大简化了两个备选方案,而且,作为一个决策制定模型,它更接近人们实际的决策行为。比如说,汽车购买者更愿意考虑两辆车各方面的差异,而不是 (像采用线性模型那样) 仔细考虑每辆车的各个方面后比较两辆车的加权价值。

还有一种补偿策略的决策方法被称为“理想点模型” (ideal point model)。理想点模型在运算上与线性模型类似,但两者的原理并不相同。在这一模型中,决策者心中有一个理想方案 (如一个理想的工作或是一辆理想的汽车)。然后他们将其拥有的备选方案在各个标准上与理想方案相比较,对这些差异进行加权。

非补偿性策略

人们在面临具有多个备选方案的复杂情况时通常采用非补偿性策略 (non-compensatory strategy)。与补偿性策略相反,这些策略不允许不同标准间进行互相协调。有四个著名的非补偿性策略方案:关联原则、析取原则、词典式策略和逐步淘汰制 (Hogarth, 1987)。

在关联原则 (conjunctive rule) 下,决策者可以排除那些在预定范围之外的备选方案。比如说,如果某个研究生申请者的 GRE 分数低于 1 000, 平均绩点低于 3.0, 或是申请信犯了 3 个以上的拼写错误,就会被排除在考虑范围之外。但关联原则只能达到满意,而不是最优。

如果决策者采用的是析取原则 (disjunctive rule), 那么只要申请者的 GRE 分数或是平均绩点足够高的话,决策者就可能会允许他们犯更多一些的拼写错

误。根据析取原则，衡量每一备选方案的标准是这一方案所具有的最好的属性，而不管它的其他属性有多差。举一个极端的例子，只要一个申请者的 GRE 分数够高，那么就算他的申请信是大猩猩替他写的也无所谓。

第三个非补偿性策略是词典式策略 (lexicographic strategy)。使用这一策略的决策者，首先甄选出最重要的衡量标准，然后选择在这一标准上最令人满意的方案。如果选出的方案不止一个，那么就甄选出第二重要的衡量标准，再在剩下的方案中进行选择。如此下去，直到最后只剩下一个方案。

第四个非补偿性策略是逐步淘汰制 (elimination-by-aspects, EBA)，由特韦尔斯基 (1972) 提出。从本质上来说，它是词典式策略的一种概率形式。根据 EBA 原则，每一衡量标准——或是方面——被选择的概率与其所具有的重要性相当。备选方案首先按照已选出的衡量标准进行比较，淘汰掉次要的方案后再选出进行比较的第二个方面，然后淘汰掉更多的方案，如此下去直到最后只剩下一个方案。特韦尔斯基 (1972) 是这样说明 EBA 原则的：“为了购买一辆新车，选出的第一个衡量标准可能是汽车的自动变速装置，这将排除那些不具有此项特征的所有汽车。假设在剩下的方案中，选出的另外一个评价标准是 3 000 美元的价格限制，那么超过这个价格限制的所有汽车都将被排除掉。这个过程一直持续下去，直到最后只剩下一辆汽车。”

优先标准

虽然有许多关于多属性选择策略的理论文章和数学论文，但有关这一命题的实验却相对较少。其中比较著名的是由斯洛维奇 (1975) 进行的一系列实验。

斯洛维奇感兴趣的是，人们如何在两个等价的方案中做出选择。他的假设是，如果两个方案具有相等的价值，那么人们倾向于选择在优先标准上具有更高价值的方案。斯洛维奇称之为“优先标准假设”。

为了检验这一假设，斯洛维奇先让被试找出两个等价的备选方案（比如说，对球队来说具有同等价值的两个棒球运动员）。他们会被问到与下面这个问题相类似的问题：

球员 1 打出了 26 个本垒打，击球平均得分数为 0.273；球员 2 打出了 20 个本垒打。如果说两名球员具有不相上下的能力和价值，那么球员 2 的击球平均得分数应该为多少？

被试找出等价的备选方案后（已经过了一段时间），斯洛维奇让他们评出他们认为更重要的衡量标准并在备选方案中做出选择。其实验结果有力地支持了“优先标准假设”。实验结果表明，被试确实倾向于选择在优先标准上具有更高价值的方案。

这表明，当面临等价的备选方案时，人们并不是进行随机的选择，而是先识别出对自己来说最重要的标准，然后选择那些在这一标准上具有更高价值的方案。

结 论

圣彼得堡悖论导致了期望效用理论的产生，而诸如阿莱悖论和埃尔斯伯格悖论之类的问题则引发了其他替代期望效用理论的理论研究。这些替代理论中最为广泛接受的便是前景理论。虽然许多决策研究分析家仍然采用期望效用理论作为一个标准化模型，但前景理论却是一个能够更加准确地描述人们实际决策行为的模型。它还能够被广泛应用到许多其他领域中。

比如，前景理论所隐含的损失回避倾向就意味着，相对于征收信用卡逾期费用来说，实行现金折扣政策能够使信用卡公司获利更多（Thaler, 1980）。逾期费用对人们来说是一种从自己腰包里流出去的损失，而现金折扣则是一种获利；因此，即使两种情况下的费用结构是等价的，交纳逾期费用的感觉仍然要比没有得到现金折扣的感觉更强烈一些。相同的策略也被许多商店所采用，它们给出一个“建议零售价”以改变消费者的参照点，使得它们的实际价格看起来好像要“省钱”一些（Thaler, 1985）。

损失规避同样被用于鼓励妇女们进行乳腺癌自我检查。贝丝·迈耶罗维茨和谢利·蔡金（Beth Meyerowitz & Shelly Chaiken, 1987）向女大学生出示了以下三本手册之一：（1）一本手册从防止乳腺癌产生的收益的角度来说明乳腺癌自我检查的重要性；（2）一本手册强调了不进行乳腺癌检查的可能损失；（3）一本手册两者都不提及。比如说，从收益角度来阐述的手册可能会告诉读者：“如果现在就进行乳腺癌自我检查，你就能够知道正常、健康的乳房应该是怎样的，这样你才可以注意到随着年龄变化而可能产生的很小但不正常的变化。”而从损失角度来阐述的手册则是：“如果你现在不做乳腺癌自我检查，你就不知道正常、健康的乳房应该是怎样的，而且很难注意到随着年龄变化而可能产生的很小但不正常的变化。”另一本两者都不提及的手册则省略掉了这些句子。

迈耶罗维茨和蔡金发现，4个月之后，阅读“损失手册”的女性中，有57%的人进行了乳腺癌自我检查，而阅读“收益手册”的女性和阅读两者都不提及的手册的女性中，分别只有38%和39%的人这样做了。对此，迈耶罗维茨和蔡金解释说，那些阅读“损失手册”的女性更愿意进行乳腺癌自我检查是因为，所预期到的健康预防损失比同样的健康预防收益要重要得多——即使两者在逻辑上是等同的。这些结果说明了损失规避倾向是如何被运用到提高社会健康状况和社会福利上来的。

前景理论相对于经典期望效用理论来说是一个很大的进步。确实，许多违背期望效用理论的现象都能够在前景理论中得到很好的解释。第四部分对各种

偏离理性原则的原因做了回顾。正如这些章节所示，决策者在判断和选择行为过程中很容易受到许多偏差的影响，但在许多情况下，这些偏差都是系统性的，而且是预先可以控制或预测的。

第四部分

直觉与偏差

当人们面临一个复杂的判断或决策问题时，他们通常会依据自己的直觉或者是一些常识来进行决策。大多数情况下，根据此类原则得到的答案往往会非常接近“最优”方案。但是在某些情况下，直觉就可能产生某些可预测的偏差和不一致。这部分将着重说明几种众所周知的直觉和偏差。

第 10 章

代表性直觉

人们是如何进行决策的？如何在不同的备选方案之间进行选择？如何对具体事件的价值或者可能性做出判断？这部分将关注以下两个相关的话题：决策者得出结论的整个过程，以及这样的过程可能导致的偏差。

阿莫斯·特韦尔斯基和丹尼尔·卡尼曼（1974）曾经建议决策者运用“直觉”或者是一般的常识来进行决策。利用直觉进行决策的优点是可以很少的时间和努力达到与理性决策相同的结果。例如，直觉可以很容易地估计某种结果出现的可能性，而不必用计算的方式（即把该结果过去的每一种发生可能性相加，然后再除以该结果发生的总数）。在绝大多数情况下，一个粗略的近似值就已经足够了（正如让人们满意的通常并不一定就是最优的方案）。

通常而言，直觉可以得到一个令人相对较为满意的答案。但是利用直觉进行判断的缺点是，在某些情况下，直觉判断可能导致一些系统性的偏差。在本章中讨论的直觉类型是“代表性直觉”（representativeness heuristic）。在某些情况下，使用代表性直觉可以导致一些可预测的偏差的产生。正如前文中所提到的，在这里更多的是关注偏差而不是成功的原因，主要是因为偏差能够更多地揭示一些隐含的过程。事实上，现在所有有关决策的理论都是基于决策偏差的研究而产生的。

A、B 和 C 的代表性直觉

根据特韦尔斯基和卡尼曼（1974, p. 1124）的理论，人们通常会根据“A在多大程度上能够代表 B，或者说是 A 在多大程度上与 B 相似”来判断事件发生的可能性。特韦尔斯基和卡尼曼将这样的一种原则称为“代表性直觉”。

什么是“A”和“B”？这将取决于你进行决策的情境。如果你在估计 A 来自 B 的可能性，那么 A 可能就是一个例子或者一个样本，而 B 则是一个种类或者样本总体。例如，A 可能是一个个体，而 B 则是一个群体，而决策的问题则可能是 A 成员属于 B 群体的可能性。另一方面，如果你试图判断 A 在多大程度上是 B 导致的，那么 A 可能是一个事件的结果，而 B 则是事件发生的过程或者原因。例如，B 可能是一个投掷硬币的过程，而 A 可能就是在一系列的投掷中有 6 次是人头，判断所关心的可能就是出现这种结果的可能性。由于代表性直

觉的定义是抽象的，理解起来有一定的难度，我们将举一些具体的例子来说明代表性直觉在特定情境下是如何起作用的，以及偏差是如何产生的。

读者调查中的第1题就是一个很好的例子。这个例子来自特韦尔斯基和卡尼曼（1982）的研究，内容如下：

琳达，31岁，单身，坦率直言，性格开朗。她所学的专业是哲学。当她还是一个学生的时候，她就非常关注歧视和社会公正问题，同时参加了反对核武器的活动。请从以下选项中选出可能性更高的选项：

- 琳达是一个银行出纳
- 琳达是一个银行出纳，同时是一个活跃的女权主义者

绝大多数人认为琳达是一个女权主义的银行出纳，而不仅是一个银行出纳。当特韦尔斯基和卡尼曼（1982）让86个人回答以上这个问题时，超过90%的人都是这样认为的。尽管有可能你也是这样认为的，但是这样的答案违反了概率的基本原则。两个独立事件（“银行出纳员”和“女权主义者”）同时发生的概率不可能高于单个事件发生的概率（例如，银行出纳员）。出于这样的原因，特韦尔斯基和卡尼曼（1983）将这种现象称为“结合谬论”（conjunction fallacy）（见 Leddo, Abelson & Gross, 1984; Morier & Borgida, 1984）。

你可以通过图10.1清楚地看到连接的原则。左边的圆圈代表所有的银行出纳员，右边的圆圈代表所有的女权主义者，中间重合的部分代表既是银行出纳员，又是女权主义者的个体。由于银行出纳员中有一部分并不是女权主义者，

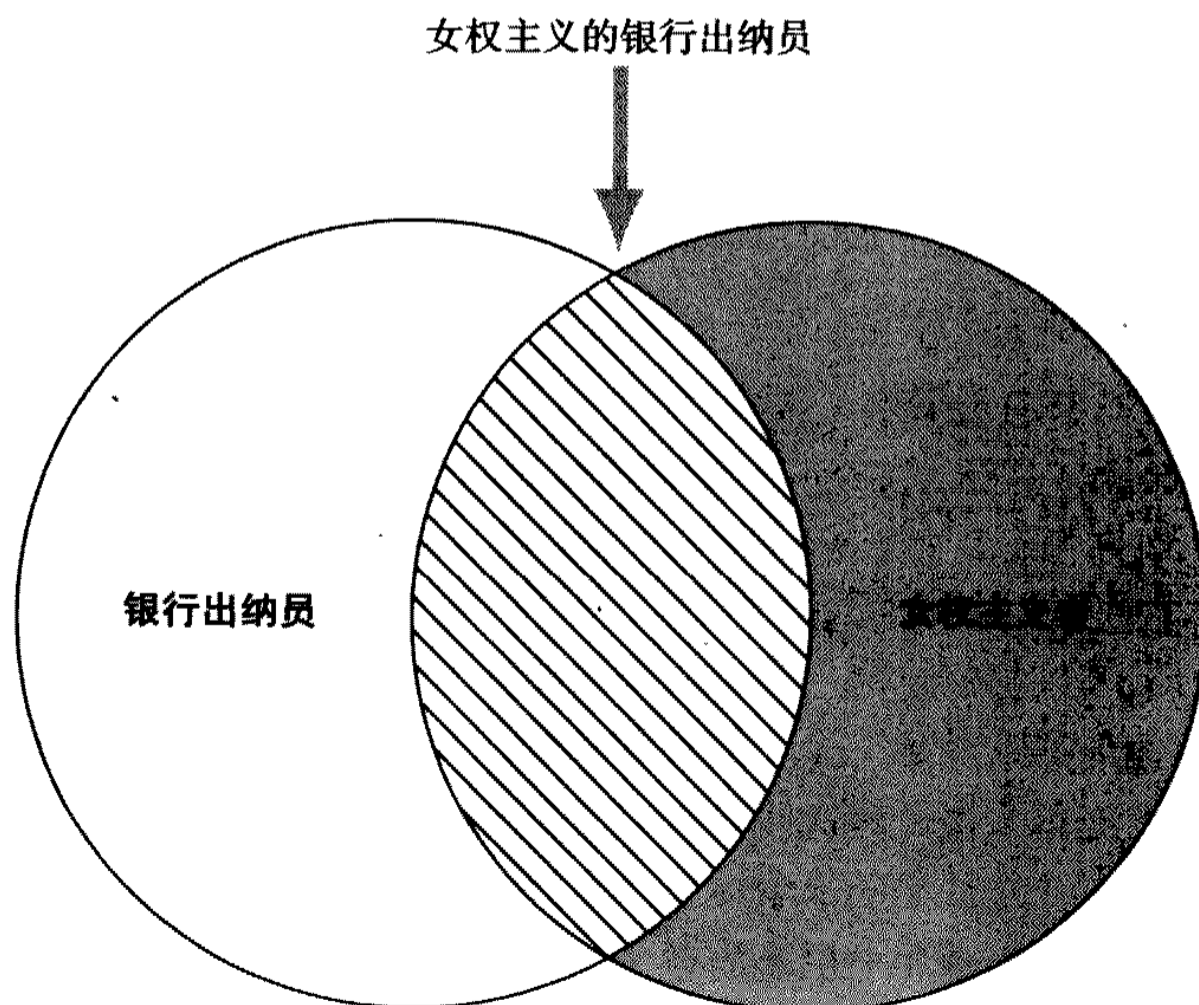


图 10.1 银行出纳和女权主义者的交集

因此成为银行出纳员（不论是否是女权主义者）的概率必定大于中间重合的部分个体。

为了确保人们不会将“银行出纳员”理解成为“在女权运动中不活跃的银行出纳员”，特韦尔斯基和卡尼曼（1982）又进行了一个补充实验，在这个实验中有两组不同的个体，其中一组个体进行选择的备选方案中只出现“银行出纳员”以及其他的选项，而另一组个体选择的备选方案中只出现“女权主义的银行出纳员”这一选项和其他的选项（这样就能确保了两个选项没有被直接进行比较）。即便是在这样的实验情境下，人们还是认为琳达是一个女权主义的银行出纳员的概率也大于认为琳达是一个银行出纳员的概率。

特韦尔斯基和卡尼曼（1982）在其他一些情境下也得到了类似的结果。“比尔”被认为“更可能是一个会计和爵士乐的演奏者，而不仅仅是一个爵士乐的演奏者”；同时，“一个温布尔登网球赛的选手更可能在输掉第一局以后而赢得整个比赛，其概率大于其仅仅输掉第一局的概率”；“美国的一位前总统更可能为未婚母亲提供财政援助和削减对地方政府的财政支持，其概率大于仅仅减少对地方政府的财政支持”。

根据上述结果，特韦尔斯基和卡尼曼（1982, p. 98）得出这样的结论：“随着情境中细节数量的增加，该情境发生的概率只会逐渐降低，但是它的代表性和由此带来的外显的可能性却会上升。我们相信，基于代表性的决策判断是人们喜欢选择毫无根据的细节化情境的主要原因。例如，‘被告离开犯罪现场’的陈述似乎比‘被告由于害怕被起诉谋杀而离开犯罪现场’的陈述更没有说服力。”

这样的结论得到了进一步的支持，读者调查中的第 11 题要求读者判断以下两个情境哪一个更可能发生：

在未来 10 年中，你认为最有可能发生的事件是：

- 美国和俄罗斯将爆发一场全面的核战争
- 美国和俄罗斯将爆发一场全面的核战争，但一开始两国都不想动用自己的核武器，只是在卷入了一场局部战争之后，如伊拉克、利比亚、以色列或者巴基斯坦等国的战争，才被迫动用核武器。

和银行出纳员问题一样，绝大多数人都认为更为具体的事件（由第三方拖累的全面战争的爆发）比一个一般性的事件（一场全面战争）发生的可能性更大。事实上，五角大楼针对这些具体的但完全不可能出现的情境，花费了几十年的时间制定了非常详尽的战争计划，生产了大量的武器来应付这样的情境的出现。根据特韦尔斯基和卡尼曼的理论，相对于一般的情境而言，表述非常具体的事件似乎更可能发生，因为这样的情境与人们对于具体事件的想像是一致的。

小数法则

代表性直觉的另一个结果被特韦尔斯基和卡尼曼（1971）称为“小数法则”。称做“小数法则”只是作者的玩笑话，主要是相对于统计学中的众所周知的“大数法则”（在统计学中，你从总体中抽取的样本容量越大，该样本的平均数与总体的平均数越是接近）而言。而人们所使用的“小数法则”则是认为从总体中抽取的随机样本相互之间是类似的，与总体之间的接近程度比实际的统计抽样理论所预测的要高得多。

例如，当你要求人们写下一个随机的投掷硬币的数列（不是真正去投掷硬币），人们试图将这个序列的每一个点都看上去像随机的〔特韦尔斯基和卡尼曼（1972）将这样的现象称为“局部代表性”〕。结果，在他们写下的答案中并没有很长的相同的序列，而相对于真实的随机序列而言，硬币的正面和背面之间的更替明显增加。在一个随机序列当中，如果只是看某一些局部的序列，它们看起来可能并不一定是随机的。为了证实这样的观点，你可以实际投掷100次硬币，并记录下正面和背面出现的次序，这样就可以大致模拟出一个真实的随机序列。

读者调查中的第15题就是对小数法则最好的说明。这个问题来自于特韦尔斯基和卡尼曼（1971）的一个研究，内容如下：

在一个城市中，8年级学生的IQ平均数是100。你从中抽取了50名学生来进行有关学业成就的研究。你抽取的第一个学生的IQ为150。你认为你抽取的这个50人的样本的IQ平均数为多少？

很多人对这个问题的回答是IQ的平均数将依然是100。但是事实上，正确的答案是这50个个体的IQ平均数应该为101。因为第一个孩子的IQ为150，剩余的49个孩子的期望IQ依然是100，因此这50个孩子的IQ总数为5050（ $150 + 4900$ ），然后除以50，因此这50个孩子的IQ平均数应该是101。

如果你的答案是100，而不是101，你可能假定在余下的49个学生中必定会出现一个IQ的低分将150的IQ高分“平衡”掉。但是这样的观点实际上就是假定偶然事件具有自我修正的功能，事实上，偶然事件并不具有自我修正的功能，出现一个高分也不一定出现相应的低分与之抵消；余下的样本只是对这个偶然事件进行“稀释”，使其平均数更加接近总体的平均数（在这个例子中总体的平均数是100）。特韦尔斯基和卡尼曼（1971）认为，人们倾向于认为偶然事件具有自我修正的功能，这样的偏差也是来自代表性直觉，因为人们总是希望随机抽取的样本能够很好地代表总体。

同样，特韦尔斯基和卡尼曼（1971）认为代表性直觉可以导致人们承认“赌徒谬论”（*gambler's fallacy*），这种观点认为，在一系列的坏运气之后必然会有好的结果出现（或者，用更为通俗的话讲，就是认为一系列结果相同的独立事件必然会跟随一个相反的结果）。读者调查中的第31题就是检验你是否相信

这样的赌徒谬论。题目的内容如下：

·假定你连续投掷了3次硬币（该硬币没有偏差），每一次的结果都是正面。如果你必须对下一次的投掷进行投注，金额是100美元，你会选择正面还是反面？

由于这个硬币是没有误差的，正确的答案是你对正面还是反面没有偏好，因为两者出现的机会是相等的。但是有一些人错误地认为在连续出现了3次正面之后，反面出现的概率更大。特韦尔斯基和卡尼曼认为出现这样的答案就是因为人们错误地认为一个随机序列必须具备局部代表性（即序列的每一个部分都必须看起来像是随机的）。

手热现象

对小数法则一个最形象的说明是由托马斯·吉洛维奇等人（Thomas Gilovich, Robert Vallone, & Amos Tversky, 1985）进行的研究。这次研究者关心的是篮球比赛中的人们对于“手热现象”的认知。一位手热的篮球运动员就是指其在投中一个或者几个球以后，再次投篮时命中的概率大于其投失一个球以后的再次尝试。

吉洛维奇等人的研究发现费城76人队的球迷——包括几个队员和教练——都认为存在这样的手热现象，尽管数据统计分析发现并不存在这样的现象。也就是说，这些人认为一位队员在命中几个球以后再次尝试投篮的命中率将提高，然而事实上投中下一个球的可能性与该队员总体的投篮命中率并没有显著的差异。吉洛维奇和他的助手们通过对波士顿凯尔特人队的罚篮情况进行分析以及对康奈尔大学篮球队的男女运动员进行实验室的模拟实验（更具体而言，是在体育馆中进行的实验）的结果都证实了这样的结论。

在很短的时间内，这些研究发现在美国的运动界引起了轰动。吉洛维奇等人怎么可以将手热的现象仅仅说成是一种错觉？任何一个打过或者观看过篮球比赛的人都知道运动员在有些时候手热，而有些时候手冷！篮球队有时候为了防守手热的球员甚至改变他们的防守战术。认为同一个篮球运动员进行的投篮情况在统计上是不相关的观点似乎很难被接受。

为了找出这一现象存在的原因，吉洛维奇等人（1985）进行了一个实验，在这个实验中被试要观看6个由“X”和“O”组成的不同序列（也许你可以将它们理解成篮球比赛中的命中和失败）。每一个系列中都包含有11个“X”和10个“O”，在不同的序列中两个字母交替的可能性分别为.40，.50，.60，.70，.80或者.90。例如，下面的序列“XOXOXOOOXXOXOXOOXXXOX”表示两个字母之间转换的概率为.70（因为在20个相邻字母之间，两字母进行转换的次数为14次）。

吉洛维奇等人发现被试更多地认为.70或者.80的序列是随机序列的代表,而不是正确地选择.50的序列。只有32%的被试认为.50的序列是一个随机的序列,而62%的被试将.50的序列定义为是“手热”序列。

如果你想看看自己在这项任务中的表现,请查看一下你在读者调查中第38题的答案。第一个序列(XOXXXO000XOXX000XXXOX)有一半的情况下字母之间进行转换(类似于一个随机的序列)。与之相对应的是第二个序列(XOXOX000XXOXOX00XXXO)代表.70的转换可能性序列——远远高于随机序列50%的数值。如果你认为在第一个序列中存在太多的连续性的相同字母,你必定会期望在序列中X和O有更多的转换(也就是人们为什么将一些随机的序列认知为“手热现象”)。本书的第14章将更加详细讨论个体对随机性的知觉。

忽视基线值

某些情况下,依赖于代表性直觉可能会使人们忽略“基线值”的信息(也就是一个事件发生的相对频率)。卡尼曼和特韦尔斯基在一系列的实验中考察了这一现象。例如,在一个实验中卡尼曼和特韦尔斯基(1973, p. 241)告诉被试:

一群心理学家对30名工程师和70名律师进行了访谈和人格测试,这些人在他们各自的领域内都相当成功。基于这样一些信息,心理学家对这30名工程师和70名律师进行了简单的描述。下面你将看到5个这样的描述,都是从100个描述中随机抽取的。对于每一个描述,请在一个0至100的标尺上选择你认为该描述有多大的可能性是对一个工程师的描述?

例如,下面一段话就是卡尼曼和特韦尔斯基用来代表一个工程师性格的简单描述:

杰克今年45岁。他已经结婚并有4个孩子。他通常比较保守,谨慎和雄心勃勃。他对政治和社会事件并没有多大的兴趣,他将他大部分的业余时间都用在了自己的爱好上面,例如家中的木工活、航海以及数字游戏。

使用同样的5个描述,卡尼曼和特韦尔斯基将第二组被试中工程师和律师的比例做了对调(即70名工程师和30名律师)。但是由于结果是相对应的,我们只是关注30名工程师组的结果。

每一位被试对这5个描述为工程师的可能性进行评价以后,研究者要求其估计一下从100个描述中随机抽取一个描述,该描述是工程师的描述的可能性。不出意料之外,被试大多认为这样的可能性接近30%。换一句话讲,在这样的情境中,人们使用了给定的基线信息。

在另一方面,当被试得到了一些有关的描述信息——即使有一些信息与工

程师或者律师的职业特点并没有任何的联系——他们将忽略这样的基线信息。例如，卡尼曼和特韦尔斯基（1973）特地选择了以下的描述，这样的描述与工程师或者律师的职业特点并没有关联：

迪克今年30岁。他已经结婚了，但还没有孩子。他拥有很强的能力和很高的工作积极性，他希望在自己的领域内取得很大的成功，同时他也很受其同事的欢迎。

相对于迪克的职业而言，这样的描述并不能提供有用的信息；因此在这种情况下，该描述是一个工程师的可能性应该为30%。但是卡尼曼和特韦尔斯基（1973）的实验结果发现，被试认为该描述为工程师的可能性的平均估计为50%。很显然，在这种情况下，被试忽略了情境的基线信息，而仅仅从一个中性描述的角度来进行判断。

还有大量的研究都发现在某些情境下人们可能会忽视基线信息（Bar - Hillel, 1980, 1990; Fischhoff & Bar - Hillel, 1984; Osberg & Shrauger, 1986）。例如，艾克·阿杰增（1977）的研究发现，当基线信息与人们对因果关系的认知相一致时，人们就会使用基线值。在一个实验中，阿杰增要求被试根据一些相关的因素（例如该学生每周的学习时间）或者一些非相关因素（例如该学生每周的收入）来预测该名学生的学习成绩。实验结果发现，人们在接受了相关的信息以后更加愿意使用基线值，而对于那些不相关的信息——即使是被告知这样的信息也可以对学生的学业成绩产生预测作用，他们还是很少使用。

非回归性的预测

人们在进行预测时往往缺乏对信息来源的诊断，结果可能造成“非回归性的”（nonregressive）预测。例如，在读者调查中的第35题（基于卡尼曼和特韦尔斯基在1973年进行的研究）的内容如下：

假定一个学生高中时的考试成绩与其在大学中的学业成绩（GPA）有中等程度的相关。基于如下图所示的百分等级（见图10.2），如果一个考试成绩为725分的学生，其GPA将是多少？

你的答案将是什么呢？绝大多数人的预测是其GPA将在3.5和3.7之间。如果高中时期的考试成绩能够很好地预测大学时期的学业成绩，那么这样的答案无可厚非。即如果两者之间能够一一对应，725的考试成绩应该对应的GPA为3.6。但是根据整个问题的背景，我们发现高中时期的考试成绩并不能很好地预测大学的GPA，因为两者之间只有中等程度的相关。由于存在着这样中等程度的相关，因此对该学生GPA最好的预测将在3.6和GPA的平均数2.5之间——这样才能体现“向平均数回归”。

图 10.2 高中考试成绩与大学学业成绩 (GPA) 之间的关系

学生的百分等级	考试成绩	GPA
前 10%	> 750	> 3.7
前 20%	> 700	> 3.5
前 30%	> 650	> 3.2
前 40%	> 600	> 2.9
前 50%	> 500	> 2.5

向平均数回归是统计中存在的一个现象，即较高或者较低的分数往往会伴随着一些更加接近平均数的分数，就像两位非常高的父母的孩子往往会更接近平均身高一些。因为 725 是一个很高的分数，如果你让该名考生再次考试，其成绩可能会更加接近平均分 500 分一些（根据同样的逻辑，它可以预测一个相应更低一些的 GPA）。你可以用这样的方法进行思维：如果你不知道一名学生任何方面的信息，你最佳的预测是他的 GPA 将是 2.5，如果高中的考试成绩和大学的学业成绩之间有完全的正相关，你的预测将是 3.6。因为高中的考试成绩与大学的学业成绩之间具有中等程度的相关，那么最好的预测就是居于 2.5 和 3.6 之间（比平均数高，但是没有 3.6 那么高）。

绝大多数心理学家都认为考试成绩是由两个独立的成分组成的：“真实分数”和“误差”。如果一个考试能够很好地测量一个人的能力，真实分数就是该学生在考试中应得的分数；同时误差因素包括所有与能力无关但是同样会影响考试成绩的因素（例如考试前一天晚上的睡眠时间、血糖水平、心境以及当时的照明情况等等）。在绝大多数的情况下，这些因素可以相抵消，但是在一些偶然的情况下，这些因素共同作用可以戏剧性地提高或者降低一个人的考试成绩。由于这样的波动与真实分数之间是没有关系的，因此下一次的考试成绩就更有可能向真实分数进行“回归”。

忽视这种回归现象可能会使决策产生重大失误。例如，卡尼曼和特韦尔斯基（1973）讨论了这样一个案例，一所飞行学校的指导员得出了这样的结论，对飞行员的出色表现进行表扬会导致其下一次飞行成绩的下降。这样的成绩下降是否意味着指导员应该停止对飞行员进行表扬？完全不是这样！如果只是考虑回归理论，即一个非常出色的成绩必将跟随着一个更加接近平均数的成绩。那么，糟糕的成绩得到惩罚以后也必将得到提高，而不用考虑这样的惩罚是否真正产生了作用。

理查德·尼斯比特和李·罗斯（Richard Nisbett & Lee Ross, 1980, pp. 163, 165）在他们有关人类推论的书中描述了人们对回归错误理解后产生的另一些后果：

人们对于一些简单的回归现象的错误理解（例如一件非常好或者非常

差的事件之后，必然会跟随着一些不那么好或者不那么差的事件，而不管其中是否存在随机因素）在我们的日常生活中是屡见不鲜的。这样的误解可能会使人们惊惶失措，以为产生了严重的“危机”（如犯罪率、疾病或者银行破产的突然上升，或者是销售额、降雨量或奥林匹克运动会的金牌数量突然减少），这样的一些事件对人们产生的影响看起来比实际情况要严重得多……仅仅在观察到一些简单的回归现象后，人们就会产生一些迷信：比如说非要做点什么去结束一连串的“坏运气”；或者什么都不敢做以免失去“好运气”。

乔治·迈尔奇（George Gmelch, 1978年秋）曾经是一名职业棒球运动员，后来成了一名社会科学研究者，他在他题为“棒球中的魔术”一文中列举了一些这样的迷信。根据迈尔奇的说法，纽约巨人队为了保持16场连胜的势头，不愿意洗他们的队服，生怕洗掉了他们的好运气。同样，1941年布鲁克林Dodgers队的队员Leo Durocher在三周半的时间内一直穿着同一双黑鞋，灰色的袜子，蓝色的外套，目的是为了保持连胜的势头。

向平均数回归同样也可以解释为什么非常成功的运动员或者运动队在登上了《体育画报》杂志的封面以后，其运动成绩马上就有所下降。通常来讲，能够登上该杂志封面的运动员都是取得了一系列非常优秀的的成绩之后，如果单单是从回归的角度考虑，其运动成绩有所下降也有其必然性。众所周知的“《体育画报》厄运”其实根本就不是厄运，而仅仅是向平均数回归而已。

人为的预测和数据的预测

人们忽视基线值和数据回归性的倾向性造成了许多令人吃惊和窘迫的现象。在社会科学中（Dawes, Faust & Meehl, 1989）有大约100个研究的结果表明“数据”预测（仅仅是依据给定变量和结果之间的实证关系来进行预测）的准确性等于或者高于“人为”预测（依据个体的判断而进行的预测）。换言之，与一般的常识相反，由决策者完成的决策的准确性往往偏低——即使该决策者完全掌握了数据的信息。

例如，有这样一项针对人为决策进行的研究（Lasky, Hover, Smith, Bostian, Duffendack & Nord, 1959），在预测患者是否需要重新进行心理治疗的决策中，将12名心理治疗专业人员的决策与患者的病历档案重量之间进行比较。患者的病历档案重量可以用来作为患者以往就医情况的一个大致估计。结果表明，专业人员的判断的准确性并没有显著高于病历档案重量的预测作用（两者的相关系数分别为.62和.61）。

很显然，专业的知识以及掌握的相关信息是其他因素所无法比拟的。但是使用人为决策的方法通常需要依赖于像代表性这样的直觉——容易受

到很多偏差的影响——因此人为预测的准确性很难超过仅仅依据数据关系而进行的预测。

结 论

有关代表性直觉的研究结果表明，有一些方法可以提高决策和判断的技巧，内容如下：

- ✓ **不要被很细节的情境所迷惑。**正是情境中的细节使得整个情境看起来更加具有代表性，但是同时也减少了其发生的可能性。一般而言，情境越是具体，其发生的可能性越低——即便这样的情境看起来能够非常好地代表最可能发生的结果。
- ✓ **只要有可能，无论什么时候都应该注意基线值。**当一个事件极少发生或者是非常普通的事件时，基线值就显得尤为重要。例如，由于录取的比例非常低，因此很多优秀的申请者都没有机会进入研究生院进一步深造（通常的理解就会认为这样的申请者并不具备所需要的学术能力）。相反，由于通过的比例非常高，许多不合格的驾驶员都顺利拿到了驾驶执照。当基线值是一个极端数值时，代表性往往成为发生可能性的误导因素。
- ✓ **记住偶然性并不具备自我修正的功能。**一系列的坏运气，就是一系列的坏运气。它并不意味着相应的好运气必然会来临，也不意味着事物总是一成不变的。如果一个随机的过程（就像投掷一枚没有偏差的硬币）存在一定的可能性产生一定的结果，那么过去的事件对将来发生的结果并不会产生影响。
- ✓ **不要错误地理解向平均数回归。**即使出现了一系列的坏运气，也不一定会有一系列的好运气与之相平衡（或者相反），但是一些极端的成绩往往会跟随着一些更接近平均数的成绩。向平均数回归是非常正常的，无论结果是否受到一些随机因素的影响。即使在某些时段，这些随机因素结合在一起可能产生一些非正常的结果，但是在接下去的情境中，成绩通常会回归正常。

将这些建议牢牢记在心中，就有可能避免许多由于依赖代表性直觉而产生的偏差，在下面的第11章里，我们将重点讨论另一个著名的直觉：“易得性直觉”，以及由此而产生的偏差。

第 11 章

易得性直觉

根据阿莫斯·特韦尔斯基和丹尼尔·卡尼曼（1974, p. 1127）的理论，易得性直觉也是人们经常使用的一种直觉。决策者“通常会依据一些容易想起来的事例来判断一种类别出现的频次或者事件发生的概率”。通常情况下，这样的直觉能够很好地发挥作用，在所有条件都平等的前提下，普通的事件要比不寻常的事件更容易被记住或者想像出来。决策者利用易得性直觉（availability heuristic）来估计事件发生的频率和概率往往能够将困难的决策内容简化。

但是，无论你使用何种直觉，在某些特定的情况下，这样的一般原则都可能失效并导致系统性偏差。有一些事件相对于其他事件而言更容易想到，并不是因为这样的事件更经常发生或者具有更高的发生概率，而只是因为这样的事件更容易被提取。可能是因为这样的事件是刚刚才发生的，也可能是这样的事件掺杂了很多情绪的因素。在本章中，我们将着重讨论以下几个问题：（1）在哪些情境下，易得性直觉可能会导致偏差？（2）当人们想像一事件发生以后，他们是否会认为这样的事件更加容易发生？（3）生动的信息与其他类型的信息是如何进行区分的？

歪曲的易得性

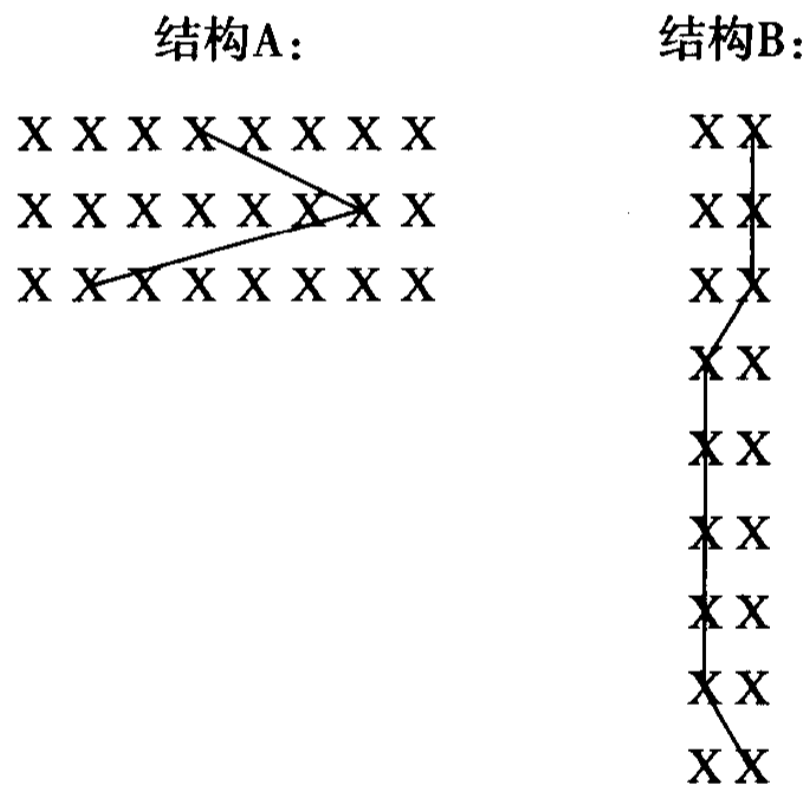
请你想像一下，在美国，下面两种情况哪一种更容易导致人死亡——被飞机上掉下来的零件砸死还是被鲨鱼咬死？绝大多数的人都认为被鲨鱼咬死的可能性更大（请查看一下你对读者调查第 7 题的回答）。相对于被掉下的飞机零件砸死的案例，被鲨鱼咬死的案例得到了更多的社会关注，这样的例子更容易被人们想像出来（从某种程度上讲，这也应该感谢像《大白鲨》这样的电影所造成的影响）。但事实上，在美国，死于掉下来的飞机零件的个案是被鲨鱼咬死的个案的 30 倍。在这个例子中，易得性直觉就是一个误导人们进行频率判断的因素。

读者调查中的第 8 题进行了一次比较，人们又发现一个惊人的答案（取自库姆斯和斯洛维奇 1979 年的研究）。例如，由于很少有媒体关注这样的问题，因此人们对于下列事实感到十分惊讶，那就是：每年死于糖尿病或者胃癌的人数是死于凶杀案或者交通事故的人数的两倍，而每年死于雷击的人数高于死于

龙卷风的人数。特韦尔斯基和卡尼曼认为，这样的统计数据与人们的直觉正好是相反的，人们的直觉会根据事件在大脑中唤起的难易程度来估计事件发生的可能性。由于交通事故、龙卷风或者谋杀几乎都是媒体的头条新闻，因此它比那些发生频率更高的事件——胃癌、雷电或者糖尿病——更“容易提取”。

当某些事件本身很难被想像时，易得性直觉也可能会导致偏差。例如，特韦尔斯基和卡尼曼（1973）问人们如下问题：在一般的英文词汇当中，以 K 作为开头字母的单词和以 K 作为第三个字母的单词相比，哪一个更多（不要考虑字母数少于 3 的单词）？在回答问题的 152 个被试当中，有 105 个被试认为由 K 作为首字母的单词要多。但是事实上，以 K 作为第三个字母的单词数是以 K 作为起首字母的单词数的两倍。因为人们更容易想到的是以 K 作为起首字母的单词，而不是以 K 作为第三个字母的单词，因此人们会错误估计了这两类单词的相对比例。

当结果在视觉上更加容易辨认的时候，易得性直觉同样可能会导致偏差。读者调查中的第 37 题正好可以说明这样一种现象：



请仔细观察下面 A、B 两个不同的结构。

这里的一条“通路”就是指将本结构中的每一层上的一个 X 连接起来。换一句话讲，在结构 A 中，一条通路就是连接三层中的 X（一共三层结构，每一层选取一个 X）；而在结构 B 中，一条通路就是连接九层中的 X（九层结构，每一层选取一个 X）。

(a) 在这两个结构中，哪一个包含的通路更多？

(b) 在结构 A 中，大约包含了多少条通路？在结构 B 中呢？

绝大多数的人都发现，在结构 A 中找到一条通路比在结构 B 中简单多了，因此他们都认为结构 A 中包含的通路较多。特韦尔斯基和卡尼曼（1973）针对这个问题进行了研究，发现 85% 的个体都认为结构 A 所包含的通路较多。而他

们对于通路数量的估计是：在结构 A 中包含的通路数大约为 40 条，而结构 B 中大约为 18 条。

事实上，这两种结构中所包含的通路数是相等的。在结构 A 中，第一层的 X 共有八种不同的选择，第二层的 X 也有八种选择；而第三层的 X 同样有八种不同的选择。因此在这个结构中通路数量为 $8 \times 8 \times 8$ (512)。在结构 B 中，一共包含的通路数量为 $2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 2$ ，其通路的总量也是 512。这样，这两种不同的结构所包含的通路数量是相等的，尽管结构 A 中的通路更加容易被识别（因为在结构 A 中每一层一共有 8 个不同的选择，而在结构 B 中，每一层只有两个不同的选择）。

一个想像的研究

1978 年，约翰·卡罗尔 (John Carroll) 发表了一项研究成果，他将易得性直觉与对一个事件的想像联系起来。卡罗尔的推理是如果一个容易被想像的事件能够被判断为更可能发生的事件，那么对一个事件的刻意想像可以增加其易得性，从而使它看起来更可能发生。他用两个实验来验证这样的假设。

第一个实验是在 1976 年美国总统大选的前一天进行的，参与者被要求想像他们在当天的晚上或者第二天一早从电视上看到了总统大选的结果。大约一半的参与者被要求进行如下的想像：

福特赢得了最后的大选，而卡特由于没有抓住几个关键的州而失败，福特最终赢得了中西部和西部的绝大多数州的选票，他一共获得了 316 个选举团的选票，而卡特赢得了 222 个。如果以州为单位来进行统计，福特最后赢得了 32 个州的选票，而卡特赢得了 18 个州和哥伦比亚特区的选票。

而余下的另一半参与者被要求进行如下的想像：

卡特利用其在南部和东部取得的不可逾越的优势而赢得了最后的大选，福特尽管横扫了西部但是也无济于事。卡特最后赢得了 342 个选举团的选票，而福特赢得了 196 个。卡特赢得了 28 个州和哥伦比亚特区的选票，而福特赢得了 22 个州。

这些情境中使用的是最新的民意调查的结果，参与者不仅要想像这样的情境是真实的，而且还要同时想像获胜者胜利之后的演讲，以及失败者失败后的情境。这样的想像力求尽可能真实。然后，当参与者完成这样的想像之后，让他们判断最后的大选的实际结果如何。

实验的结果表明想像卡特获胜的参与者更多地预期卡特能够获得最终的胜利，而想像福特获胜的参与者更多相信福特能够获得最终的胜利。根据卡罗尔的假设，对给定结果的想像可以使结果容易提取，同时也增加了对其发生概率

的估计。

在第二个研究中，卡罗尔（1978）要求匹兹堡大学的学生想像他们的橄榄球队在1977赛季取得了很好的战绩或者想像他们在该赛季糟糕的表现（匹兹堡大学的橄榄球队赢得了1976赛季的全国冠军，但是在1977年该队的教练和几位优秀的队员相继离开了球队）。尽管最后的结果并不能完全验证原来的假设，但还是显示对结果的想像可以增加对其结果发生概率的预期。例如，在35名想像球队取得好成绩的参与者中，63%的参与者认为他们在1977赛季可以取得很好的成绩；而在38名想像球队糟糕表现的参与者中，只有40%的人认为球队能够在1977赛季取得好成绩。结合这两个研究的结果，卡罗尔认为个体对结果的想像可以增加对其结果发生可能性的预期。之后，又有一些研究者的研究重复或者发展了这样的研究结果（Anderson, 1983; Gregory, Cialdini, & Carpenter, 1982）。

想像的局限性

如果一个事件的结果很难想像，将会出现什么情况呢？如果一个决策者很难想像出一个事件的结果，其知觉的结果的可能性是会上升还是下降？在1985年，吉姆·舍曼，罗伯特·恰尔蒂尼，唐娜·施瓦茨曼和金·雷诺兹（Jim Sherman, Robert Cialdini, Donna Schwartzman, & Kim Reynolds）就这个问题进行了研究。

舍曼和其合作者要求参与者阅读两种疾病的描述之一，这两种疾病在校园里发生的频率越来越高。这两种疾病在医学上都被称为是“Hyposcencia - B”，但是根据实验条件的不同，对这两种疾病的描述是完全不同的。在“容易想像”的实验条件下，参与者读到的疾病描述是非常具体的症状描述，例如肌肉疼痛，很低的能量水平以及经常性的头痛等等。而在“很难想像”的实验条件下，参与者阅读的是非常抽象的症状，例如模糊的方向迷失感，神经系统的功能失调以及肝脏发炎等等。

控制组的参与者就是简单地阅读了有关“Hyposcencia - B”的症状描述——无论想像是简单还是困难的——然后判断将来患上此病的可能性有多大。另一方面，实验组的被试要求他们“想像在3周的时间内他们感染上了这样的疾病，并且会经历到描述中的症状。”同时，他们被要求想像一下在3周的时间内，他们感染上这样的疾病以后的感受。

舍曼和其合作者发现，控制组的被试不会受到结果想像的难易程度的影响，但是实验组的结果是非常显著的。在实验组的被试中，相对于控制组的被试而言，“易于想像”的实验条件下，被试认为感染此病的可能性更大。舍曼等人（1985）认为对结果的想像并不能保证其出现的可能性更高；如果事件的结果是很难想像的，想像的努力就会降低人们对其发生可能性的预期。

否定

对事件结果的想像并不一定会提高人们对其发生可能性的预期，另一种情况是事件的结果是极端负性的。有些事件的结果是如此糟糕，以至于对事件结果的想像使他们否认这样的结果会在自己的身上发生（Rothbart, 1970）。

对于许多人而言，这样的极端结果事件的例子是核战争。在1989年，作者发表了一个研究，研究者询问了大约2000个人，让他们估计一下在未来的10年中发生核战争的可能性有多大。尽管在本书的第13章中将更加详细地阐述这个研究的过程和结果，但是其中的两个结果与易得性直觉是相关联的。首先，让参与者很形象地想像核战争爆发时的情景（例如，增加“结果的易得性”）并不能显著地影响个体对于发生核战争可能性的预期。第二，让参与者想像发生核战争的各种不同的途径（增加“路径的易得性”）同样不能显著影响他们的判断。第二个结果是令人惊讶的，因为有不少的研究已经发现路径易得性的重要性（Hoch, 1984; Levi & Pryor, 1987; Ross, Lepper, Strack, & Steinmetz, 1977; Sherman, Zehner, Johnson, & Hirt, 1983）。

人们对于这样一些情境——包括将自己最亲近的朋友和家人火化这样的情境——产生如此强烈的否定感，以至于认为核战争是不可能发生的。就是这样一种排斥感抵消了上升的易得性，使个体对于事件发生的可能性的预期保持不变。这样，如果一个事件的结果是非常恐怖的，就会使个体产生否定感，这样对其结果的想像并不一定能够提高个体对其结果发生可能性的预期。

生动性

与易得性相近的一个概念是生动性。生动性通常是指某事件多么具体和易于想像，尽管在另外一些情况下，它还会有其他的含义。有时候，指某事在情绪上的令人激动，或者是事件之间在时间或者空间上的接近性。许多研究的结果表明，决策者更加容易被生动的信息所影响，而不是平淡的、抽象的或者是统计的数据（Nesbett & Ross, 1980）。

例如，尤金·博吉达和理查德·尼斯比特（Eugene Borgida & Richard Nisbett, 1977）进行了一项研究，对比了一项大学课程评估的统计数据和一个更加生动的描述之间的有效性。参加本实验的被试，绝大多数都是来自密歇根大学心理学专业的学生，他们被安排到了三个不同的实验条件下：（1）基线条件，在这样一个实验条件下，他们读到的是一个在5点量表上的对课程的评分统计结果，这样一些统计结果都是来自于上一个学期选课同学对于该门课程的打分；（2）面对面的条件，在这样一个实验条件下，被试听取了1~4个不同学生委员会对课程的评估报告（这样的评估委员会也为不同的课程在5点量表上进行打分，评分的结果与上一个基线条件下的结果是相同的）；（3）没有评

估结果的控制条件，在这样的实验条件下，被试既没有听到也没有读到任何有关课程评估的信息。然后，在读完学生对课程的评估信息或者听完委员会对课程的评估报告之后（对于控制组而言，他们并没有得到任何的信息），学生被要求列出 27 门大学课程中哪一些是他们将来最有可能选择的课程。

因为基线条件下的被试读到的信息包含了几乎所有学生对该课程的评估结果，因此最“逻辑”的结果是在该实验条件下的被试最容易接受课程推荐中的所推荐的课程。但是正如表 11.1 所示，博吉达和尼斯比特发现结果正好相反，只是听取了少数学生的评估意见的学生比那些读过详细的综合统计分析结果的学生更容易接受所推荐的课程。事实上，基线条件的被试与控制条件的被试相比，在统计上他们并没有显著地多选被推荐的课程，而少选没有被推荐的课程。而只有面对面实验条件的被试与控制组的被试之间的差异是显著的。与控制组的被试相比，面对面条件的被试平均多选了 1.4 门推荐的课程，而少选了 0.9 门非推荐的课程。

这样的实验条件可以证明少量的对课程的生动评价可以超过综合性的统计数据。正如许多新车购买者都遇到过的，一个生动的关于车辆瑕疵的故事足以消磨掉你从《消费者报告》这样专业杂志的统计中建立起来的信心（Nisbett, Borgida, Crandall, & Reed, 1976）。同样，罪犯或者目击者对犯罪事实的生动描述也可以掩盖犯罪的统计数据 and 报告。因为这样生动的信息是“可得的”，与平淡的信息相比更容易被回忆起来，因此会使决策产生偏差。

在法律上的应用

生动信息的力量深得广告人、政客以及许多其他的“专业说服者”的青睐。而生动性能够起到决定性作用的一个领域是法庭。罗伯特·雷耶斯，比尔·汤普森和戈尔登·鲍尔（Robert Reyes, Bill Thompson, & Gordon Bower, 1980）通过一项研究来证实这样的观点，即信息的生动性可以影响陪审团最后的决策。实验是通过两个过程来完成的。

表 11.1 推荐生动性的力量

条 件	推荐的课程数量	非推荐的课程数量
面对面	4.73	0.50
没有评估结果（控制）	3.33	1.39
基线条件	4.11	0.94

注释：该表摘自博吉达和尼斯比特所做的研究（1977）。比起控制组学生来，面对面条件下的学生选择了更多的推荐课程和更少的非推荐课程，而基线条件下的学生与控制组没有显著差异。

第一步，被试读到了一个关于酒后驾驶的案例。被告在圣诞节晚会以后驾车回家，撞倒了一个停车的标志，最后与一辆垃圾车相撞。但是被告血液中的酒精含量还没有检测，需要在当时这样一些证据的基础上进行判断。辩方认为被告并没有酒后驾车。

在读完被告的一个简单的人格描述以后，被试将读到9条由辩方提供的无罪的证据，也会读到9条由控方提供的被告有罪的证据。这18条陈述每一条都是一个证据，将用一种生动或者平淡的方式提供给不同的被试。例如，控方的一条平淡的证据是这样描述的：

在他出门的路上，桑德斯（被告）沿着桌子蹒跚而行，将一个盘子打落在地上。

同样的证据，用生动的方式表述是这样的：

在他出门的路上，桑德斯（被告）沿着桌子蹒跚而行，将一个蘸酱的碟子碰到了地上，酱在粗毛毯的地毯上溅了一地。

同样，辩方平淡的证据是这样表述的：

在交叉询问中，垃圾车的驾驶员承认，在晚上，垃圾车是很难辨认的，因为该车的颜色是灰色的。

而同样的信息用生动的方式表述是这样的：

垃圾车的司机承认垃圾车是灰色的，他说：“因为它成天在垃圾堆里面打滚，你觉得我应该怎么做，难道把它涂成紫色吗？”

大约一半的被试获得的是辩方生动的证据陈述和控方平淡的证据陈述；而另一半的被试获得的是辩方平淡的证据陈述和控方生动的证据陈述。

当读完这18个陈述以后，被试被要求进行下面三个决策：（1）你认为在事故发生的时候，桑德斯喝得有多醉？（2）以你个人的观点，你认为桑德斯是有罪还是无罪的？（3）如果你是陪审团的成员，你必须依据“在毫无疑问的情况下才可以定罪”的原则，你会怎样进行裁决？（这三个判断最后将平均起来形成一个总体指数，判断被试认为桑德斯犯罪的程度。）第一阶段的实验结束，被试可以返回，但是被要求在48小时以后回来进行实验的第二阶段。

当被试来参加实验的第二阶段时，首先要求他们尽可能多地回忆两天以前读到的18条证据。他们要回答第一阶段结束时要求回答的同样的三个问题。告诉被试他们不用要求自己的回答与上一次相同，他们应该“就像第一次遇到这样的案例一样”做出自己的判断。

雷耶斯，汤普森和鲍尔的实验结果发现，在第一阶段，当被试读完18条证据之后，证据的生动性并不能影响他们的判断，但是在实验的第二阶段（48小时以后）却产生了显著的影响。在控方生动证据条件下的被试更多地认为被

告是有罪的。雷耶斯，汤普森和鲍尔（1980）对这一延迟现象的解释是生动的证据更容易被回忆起来。同样，在博吉达和尼斯比特（1977）的研究中发现，相对于平淡的信息而言，生动的信息最终可以产生更大的影响。这样的结果也可以解释为生动的信息更容易被提取。

不同的声音

为了证实以上的结论，有必要对有关“生动效应”的研究进行一次综述，以期了解在一般情况下，生动的信息相对于平淡的信息而言，是否能够产生更大的影响。谢利·泰勒和苏珊娜·汤普森（Shelley Taylor & Suzanne Thompson, 1982, p. 178）对有关“生动效应”的实验室研究进行了一次全面的总结，结果发现绝大多数的实验室研究都发现生动效应是混合的，或者在有些实验室研究中根本就不存在生动效应。他们的结论是：至少相对于实验室研究而言，即便生动效应确实存在，这样的效应也是相当微弱的。

这个结论当然可以调节一下以往的结论，认为生动信息能够产生很大影响的观点值得商榷。同时，我们也不可否认，在一定的情境下，生动效应确实是存在的。原因有三个。第一，泰勒和汤普森在下一般性结论的同时也提出了许多的特例。例如，他们发现使用案例的历史资料就比统计的或者抽象的信息更具说服力；在一定的情境下，使用录像的呈现方式就比书面或者口头的陈述更具说服力。第二，泰勒和汤普森也意识到，可以有很多种不同的方法来解释研究结果的缺损。在一个具体的实验条件下没有发现生动效应，完全可能是由于实验设计中存在一定的偏差，从而导致实验的结果产生偏差，在某些研究中，生动性的变量与其他几个变量都产生了混淆。最后，泰勒和汤普森也指出，在实验室条件下，可能促使被试关注日常生活中经常被他们忽视的一些信息，这样就导致生动效应没有办法通过实验室模拟的方法得以验证。因此，实验室的设计可能在很大程度上低估了生动效应在人们日常生活中的作用。

在最后的分析中，我们也可以发现，尽管生动效应在某些情境下是可以产生作用的，但是其适用的范围和效力都存在一定的局限性。同时，依据易得性直觉的一般原则，相对于平淡信息而言，对事件的生动描述可以提高人们对其发生概率和频率的判断。

结 论

在很多情况下，易得性直觉可以为我们提供很多关于发生频率和概率的准确估计。但是在某些情况下，它同样可以使人们的判断产生很大的偏差。例如，人们基于对胃癌这样的致死性疾病死亡率的认知来判断公共的健康水平。如果这样的疾病的死亡率被低估，人们就不太可能采取很多预防措施

(Kristainsen, 1983)。同样的，如果一些描述非常生动的，但是并不经常发生的致死原因被高估了，人们的视线也将随之转移，而忽略了普遍的危险。例如，有一些作家也指出，美国人高估了他们在国外旅游时遭到恐怖分子袭击的概率 (Paulos, 1986, November 24)。

解决这个问题的办法是公开比较这些以往被我们低估或者高估的危险。例如，美国癌症协会就可以推行一项公共信息运动，比较一下胃癌与其他公众认为更加危险的事故的死亡率，如谋杀和交通事故等。结果广告牌可能宣布：“今年死于胃癌的人数大于死于交通事故的人数。”毫无疑问，这样的比较可以使人们意识到胃癌的死亡率比他们想像的高得多（尽管这种宣传有可能使人们忽略交通事故的死亡率）。旅行社也可以使用同样的策略，告诉旅行者他们在国外遭遇交通事故的可能性远远大于受恐怖分子袭击的可能性。

谈到估计事件发生的概率和频率，易得性直觉是最重要的一个环节。但是大家需要牢记的是：易得性直觉也仅仅是影响人们对概率和频率判断的因素之一。在第 12 章中我们将讨论影响人们概率判断的其他因素，以及我们该采取什么样的措施来避免在决策中产生偏差。

第 12 章

概率和风险

概率和风险在我们的日常生活中无处不在：我们轿车的刹车，我们头顶的云彩，我们每天摄入的食物以及我们对他人的信任。但是风险通常很难进行量化，甚至概率事件中最基本的问题也可能成为一个很大的挑战。设想一下专栏作家玛丽莲·沃斯·萨万特（Marilyn vos Savant, 1990, 9 月, p. 13）对声名狼藉的“游戏秀问题”所做的讨论。

假设你正在参加一个游戏，你可以选择三扇不同的门：在其中一扇门的背后是一辆轿车，而另两个门后面是山羊。你选择了一号门，然后主持人说我们来看一下另两个门后面到底是什么，主持人打开了三号门，后面是山羊。然后他问你：“你是否会选择二号门？”这个时候你会改变主意，转而选择二号门吗？

在你继续阅读以前，请先做出您的选择。

许多人都认为在这个时候是否改变决策都没有什么差异。原因是他们认为在主持人打开其中的一扇门以后，剩余的两扇门背后有轿车的概率是相同的。然而，这并不是正确的答案。

正确的答案是你应该改变你的决策。对于这个问题，萨万特是这样来解释的：

当你从三扇门中挑选第一扇门后，你就有三分之一的机会得到大奖，而有三分之二的机会得到山羊。但是随后主持人上台来给你提供线索。如果轿车是在二号门，他就会选择三号门来给你看；而如果轿车是在三号门，他就会选择二号门来给你看。因此，如果奖品是在二号门或者三号门的后面，你只要改变你的决定，你就可以获得大奖，只要它是在二号门或三号门的后面！但是如果你不改变你的决定，那只有奖品在一号门后面的情况下，你才能够得奖。

矛盾信息的困惑

这里有另外一个难题：假设你是一位外科大夫，正在为一位妇女检查她是否得了乳腺癌。她的乳房上有一个肿块，但是依据你多年的临床经验，她患有

表 12.1 游戏秀问题的结果

轿车位置	你的选择	主持人打开	你改变决定	结果
一号门	一号门	二号或者三号	二号或者三号	你失败
一号门	二号门	三号	一号	你获胜
一号门	三号门	二号	一号	你获胜
二号门	一号门	三号	二号	你获胜
二号门	二号门	一号或者三号	一号或者三号	你失败
二号门	三号门	一号	二号	你获胜
三号门	一号门	二号	三号	你获胜
三号门	二号门	一号	三号	你获胜
三号门	三号门	一号或者二号	一号或者二号	你失败

注释：正如上表中显示，如果你采取改变决定的策略，在九种不同的情况下，有六种情况你将获得最后的胜利（adapted from Selvin, 1975, February）。因此采取改变决定的策略才是最佳的策略。

恶性肿瘤的概率只有 1/100。但是为了安全起见，你让她做了 X 光检查。这样的 X 光检查确诊恶性肿瘤的准确率为 80%，而检测良性肿瘤的准确率为 90%。

令你非常惊讶的是，检查报告显示该妇女乳房上的肿块是恶性肿瘤。

问题：你先前的观点认为该妇女患有恶性肿瘤的概率只有百分之一，而检查报告的结果准确性为 80% 或者 90%，现在你认为这位妇女患有恶性肿瘤的概率为多少？（在你继续阅读之前，请给出你的答案。）

根据戴维·埃迪（David Eddy, 1982）的研究结果，在 100 位外科大夫中有 95 位大夫认为这位妇女患癌症的可能性为 75% 以上。但是事实上，正确的答案应该是 7% 或者 8%——只有大家所认为的 1/10。很显然，大夫认为该妇女患癌症的概率大致与 X 光检查的准确性类似。决策研究者罗宾·迪尤斯（Robyn Dawes）将这样的错误称为“矛盾信息带来的困惑”。

为了让大家了解为什么正确的答案是 7% 或者 8%，我们首先需要了解统计学中一个不太为人了解的定律——贝叶斯定律。根据贝叶斯定律，估计该妇女患癌症可能性的方法应该是：

$$P(\text{癌症} | \text{阳性}) = \frac{P(\text{阳性} | \text{癌症}) P(\text{癌症})}{P(\text{阳性} | \text{癌症}) P(\text{癌症}) + P(\text{阳性} | \text{良性}) P(\text{良性})}$$

这里“P（癌症）”就是患者得癌症的概率，“P（癌症 | 阳性）”表示在检查结果为阳性的情况下，患者得癌症的概率。前者是一个简单概率，而后者是一个条件概率。

等式的左边就是埃迪要求外科大夫判断的概率，而等式的右边则可以通过

问题中给定的信息得到：

$P(\text{癌症}) = .01$ [最先估计患者得恶性肿瘤的概率为 1%]

$P(\text{良性}) = .99$ [患者没有患癌症的概率]

$P(\text{阳性} | \text{癌症}) = .80$ [癌症的检查结果下可能有 80% 的概率患者为阳性]

$P(\text{阳性} | \text{良性}) = .10$ [将良性误判成恶性的概率为 10%]

既然公式中的数据都是已知的，就可以将所得到的数据带入公式，计算出患者患有恶性肿瘤的条件概率为：

$$P(\text{癌症} | \text{阳性}) = \frac{(.80)(.01)}{(.80)(.01) + (.10)(.99)} = \frac{.008}{.107} = .075 \\ = 75\%$$

对于许多人而言，这样的结果是非常令人惊讶的，但是同样令人惊讶的是参加埃迪实验的外科大夫得到相关信息后的反应方式。埃迪是这样来描述的 (1982, p. 254)：

这些决策错误的大夫通常都假定 X 光检验呈阳性的患者得癌症的可能性就大致等于 X 光检查的准确性……第二个概率是在临床中得到的，他们对此都非常熟悉，但是第一个概率需要大夫进行临床的决策。即使不是绝大多数，也有许多大夫会把两者混淆起来。

尽管对于矛盾信息的困惑绝不仅仅局限于外科大夫，但是，在这个领域中，这个问题是非常重要的。在有关生死的决策中，很少有机会能够弥补。

如果在某种情境下贝叶斯定律并不能够起作用，你该怎么办呢？答案就是我们应该认真地审视统计学上被称为“先验概率”的概念。先验概率 (prior probability) 就是在新信息 (X 光的检查结果) 到来之前，对一个事件发生概率的最佳估计。在乳腺癌决策问题中，先验概率就是大夫起初对于患者得癌症的概率判断，为 1%。因为在这样的情境中先验概率是非常低的，而 X 光的检测准确率为 80% ~ 90%，因此在获得 X 光检查结果以后的概率判断应该不会比先验概率高出很多。如果实际的结果中先验概率是非常不可能发生 (或者非常可能发生)，然后“更新”源于不十分可靠的信息，这样的原则同样可以产生作用。根据像贝叶斯定律这样的标准理论，先验概率和修改后概率之间的绝对差异不会很大。

这样的事情绝对不会发生在我身上

对事件的概率判断同样会受到事件结果“效价” (就是事件的结果在多大程度上被认为是正性的还是负性的) 的影响。早在 20 世纪 50 年代，许多研究

都已经证明，在其他条件都相同的情况下，个体认为正性结果发生的概率比负性结果要高（Blascovien, Ginsburg, & Howe, 1975; Crandall, Solomon, & Kellaway, 1955; Irwin, 1953; Irwin & Metzger, 1966; Irwin & Snodgrass, 1966; Marks, 1951; Pruitt & Hoge, 1965）。当要求个体对事件的发生概率进行判断，或者判断两个结果哪一个更加可能发生，或者打赌一个结果的发生概率时，以上的现象都存在。即使是为了提高准确率而给予物质刺激的时候，这样的规律还是存在（尽管在这样的情况下，这种效应有可能降低）。

关于这种偏差的直接证据来自于戴维·罗森汉和塞缪尔·梅西克（David Rosenhan & Samuel Messick）在1966年进行的研究。在实验中，罗森汉和梅西克一共使用了150张带表情的图片（如图12.1），其中有一些图片是皱眉的脸孔，还有一些是微笑的脸孔。在其中的一个实验条件下，70%的图片是微笑的脸孔，而30%的图片是皱眉的脸孔；而在另一种实验条件下，30%的图片是微笑的脸孔，而70%的图片是皱眉的脸孔。被试的任务就是猜测当每一张图片翻过来时是微笑的脸孔的概率。

罗森汉和梅西克的实验结果发现，当图片中70%是微笑的脸孔的时候，被试的反应是比较准确的。在整个实验过程中，被试判断是笑脸的平均可能性为68.2%。但是在另一种实验条件下（70%是皱眉的脸孔），被试判断是皱眉的脸孔的平均可能性只有57.5%——明显少于70%的概率。在这样的情况下，个体对结果效价的估计影响了个体对结果发生概率的判断。

同样的发现并不仅仅局限于笑脸的实验中。在人们对生活中发生的正性和负性事件进行判断时也存在同样的现象。例如，在尼尔·温斯坦（Neil Weinstein, 1980）对Rutgers大学Cook学院学生进行的一项研究中，他要求学生回答下面的问题：“与Cook学院的其他学生相比——当然是与同性别的学生相比——你认为下列事件有多大的可能性会在你自己身上发生？”学生进行判断的

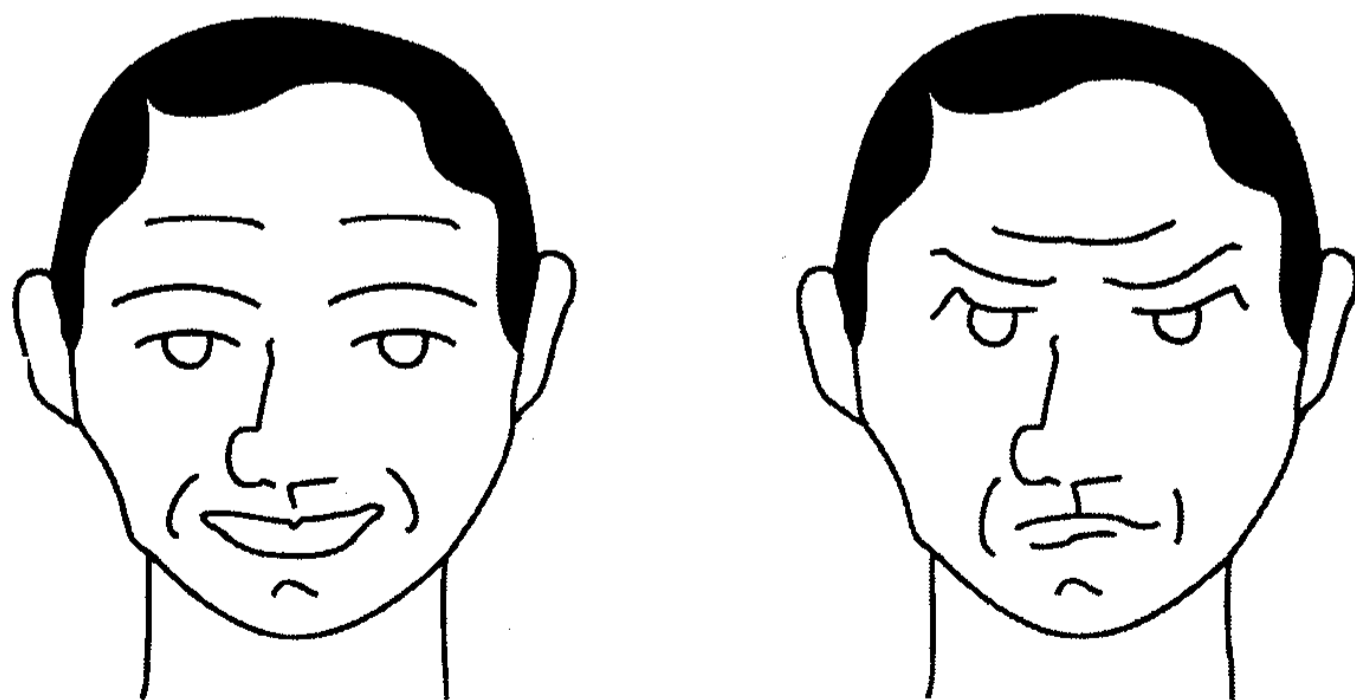


图 12.1 图为罗森汉和梅西克（1966）在他们所做的概率估计研究中所使用的刺激。

列表中有 18 件正性的事件和 24 件负性的事件，要求他们给出（相对于其他学生而言）他们认为以后会经历这样的事件的概率。

温斯坦发现，平均而言，学生认为自己经历正性事件的概率比其他人高 15%，而经历负性事件的概率比其他人低 20%。具体而言，他们认为自己毕业以后能够得到很高起薪的概率比其他人高 42%，拥有自己的家庭的概率比其他人高 44%，今后会染上酗酒的恶习的概率比其他人低 58%，在 40 岁以前得心脏病的概率比其他人低 38%。如果你是一名大学生，你可以做一下读者调查中的 5a~5d，然后比较一下你的答案如何。

复合事件

在概率理论中，单个的事件被认为是“简单”事件，同时多个事件被认为是“复合”（compound）事件。例如，一次开奖的彩票就是一个简单事件，而两次开奖（只有在两次开奖中都中奖，才能够获得最后的大奖）的彩票就是一个复合事件。如果一件复合事件是由 A 和 B 两个事件同时发生而组成的（例如在两次开奖的彩票中都中奖），这样的事件被称为“连续事件”；而如果复合事件是由事件 A 或者事件 B 其中的一件发生组成（例如两次开奖的彩票中只要有一次中奖即可），这样的事件被称为“非连续事件”。当需要估计一件复合事件的发生概率时，人们遇到的困难将会更加复杂。

首先，人们通常会高估连续事件的发生概率。约翰·科恩、切斯尼克和哈伦（John Cohen, E. I. Chesnick, & D. Haran, 1971）的研究结果显示，要求被试判断多个阶段的彩票中奖概率的时候就会出现这样的情况。当彩票分两个阶段进行，每一个阶段中奖的概率只有 50% 时，人们判断最后中奖的概率是 45%——几乎是正确答案的两倍（在这里你可以将彩票想像成一个没有偏差的硬币，综合考虑连续投掷两次的结果，你就可以发现连续两次出现正面的概率只有 25%）。同样，当每次中奖的概率只有 1/5 的时候，一个两阶段彩票的中奖概率只有 1/25，而被试的判断却是 30%。

在许多情况下，我们都可能高估事件发生的概率，但是与科恩、切斯尼克和哈伦的实验结果相比，这些高估只是小巫见大巫。他们的研究发现，当彩票涉及多个阶段，并有多个结果同时存在的时候，人们的估计更加不准确。例如，当要求人们判断一个 8 个阶段的彩票，每一个阶段中奖的概率是 1/8 时，人们认为最终中奖的概率是 1/20。这比实际的中奖概率高出了百万倍。

当人们需要估计一个复杂系统是否能够正常运作的时候，对连续事件发生概率的高估判断就显得尤为重要。例如，假设你能够建立一个导弹防御系统，该系统由 500 个独立的部件或者子系统构成（根据现在的技术，这样的一个导弹防御系统所需要的零部件的数量是该数量的几千倍），假设这个系统中的零部件在第一次使用时的可靠性是 99%，那么整个系统在第一次使用时的可靠性

如何？

结果会令很多人感到惊讶，整个系统的可靠性小于百分之一。

为什么可靠性会降到如此低的程度？在第 13 章中，我们将详细地论述——决策者在判断一个复合事件的发生概率的时候，更加倾向于“锚定”或者固着在其简单事件的发生概率上（在上面的例子中，每一个简单事件发生的概率为 99%）。一旦决策者将这样的简单概率固着在自己的头脑中，他们就很难根据大量的简单事件需要同时发生这一事实对概率的判断进行调整。因此，当复合事件是由大量的简单事件构成时，人们往往会高估其发生的概率。同样的原理也可以运用到非连续事件中，人们往往会低估非连续事件发生的概率。当非连续事件是由许多低概率的独立事件组成时，人们就会低估该复合事件发生的概率，这时的概率是简单事件中至少有一件发生即可。例如，人们通常会低估在连续几年使用避孕工具的情况下，至少发生一次怀孕现象的可能性（Shaklee & Fischhoff, 1990）。

马亚·巴-希勒尔（Maya Bar-Hillel, 1973）通过实验验证了人们在判断连续事件时容易出现高估，而在判断非连续事件时容易出现低估的现象。巴-希勒尔要求被试在两个不同的选项之间进行选择，一个选项是简单事件（例如，从装有 2 个彩球和 18 个白球的罐子中摸到彩球），而另一个选项是复合事件（从一个装有 10 个彩球和 10 个白球的罐子中，连续摸到 4 个彩球，每一次

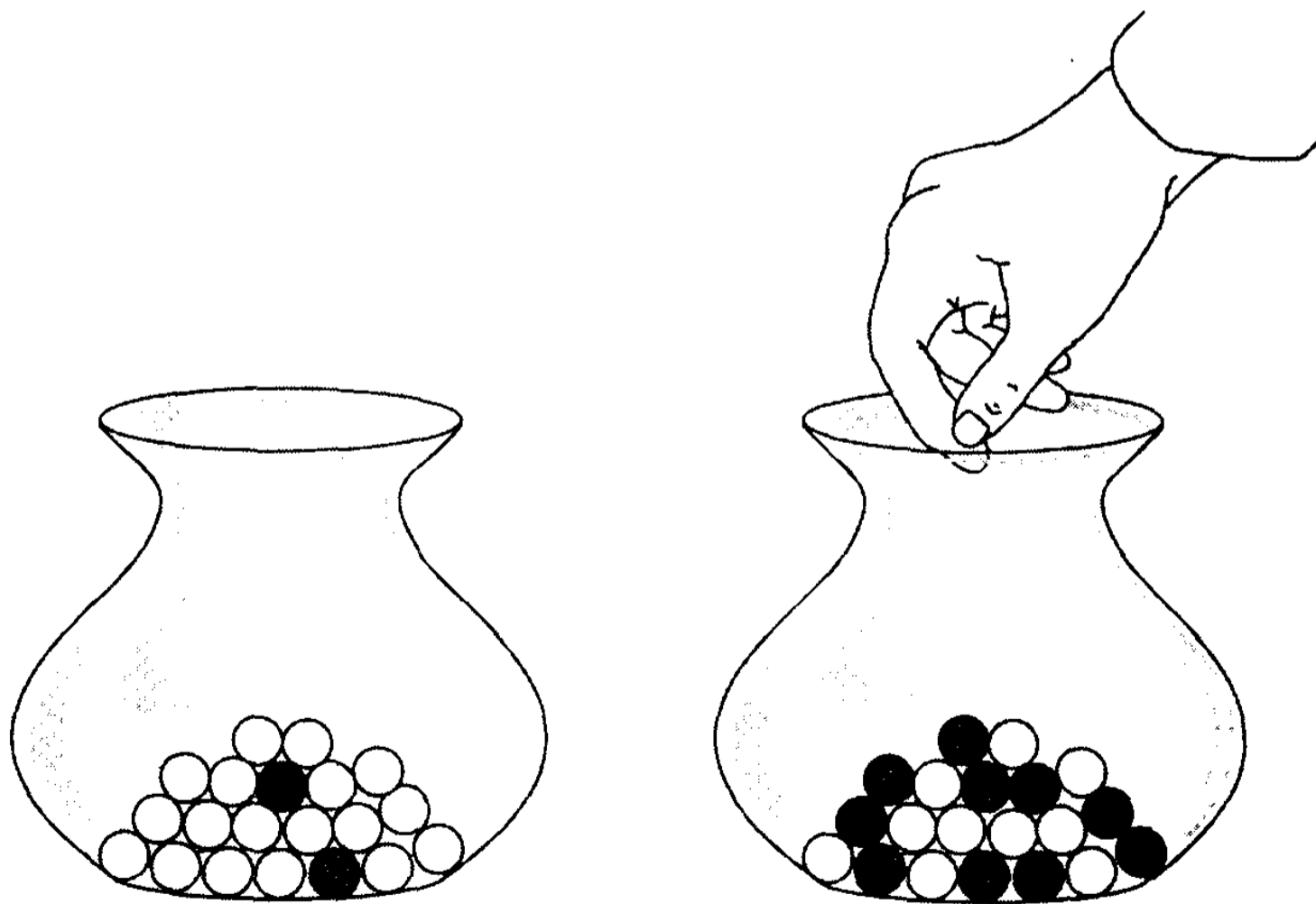


图 12.2 马亚·巴-希勒尔发现，当复合事件是相关联的时候，人们更喜欢复合性的赌博而不是简单赌博。我们在摸球时更愿意从右边罐子中连续摸四次，而不是从左边罐子中摸一次。与之相反，当复合事件不相关时，人们就更喜欢简单赌博而不是复合赌博（详见正文）。

摸完之后彩球复位，保持原有的状态)。被试将面对4~5对这样的选择(通常都是由一个简单选项和一个复合选项构成的)，并被告知可以根据自己的偏好自由选择，实验结束之后，还会根据他们的表现得到一定的报酬(以激励被试更加积极地参与实验)。

巴-希勒尔的研究结果发现，尽管所有选项的预期概率都是大致相同的，但是当复合事件是连续事件时，被试会更倾向于复合选项；而当复杂事件是非连续事件时，被试则更倾向于选择简单选项。例如，在上文中提到的两个选项中，15个被试中有12个选择复合选项，即连续拿出四个彩球(该选项发生的可能性是.06)，而不是简单选项(该选项发生的可能性是.10)。同时，当复合事件是非连续事件时(即只要其中有一次抓到彩球即获胜)，被试更加倾向于简单选项。在20个被试中有14个选择一次摸到一个彩球的选项(获胜的概率为.70)，而不是6次尝试的机会，只要有一次成功即可(在这种情况下，一次摸到彩球的概率为.20，而综合考虑6次的情况，最终获胜的概率为.74)。这样的结果显示，人们在判断连续事件发生的概率时倾向于高估，而判断非连续事件发生的概率时倾向于低估(Barclay & Beach, 1972; Wyer, 1976)。

保守主义

更加糟糕的是，一旦人们形成了对某一事件的概率判断，那么当呈现给他新信息的时候，他改变其原先判断的速度是比较慢的。这种不愿意改变先前概率估计的惰性被称为“保守主义”(conservatism)(Phillips & Edwards, 1966)。这样的保守主义都是相对于如贝叶斯法则这样的标准理论而言的，个体在接受到新的信息以后，改变概率估计的速度应该快得多(Hogarth, 1975)。

例如，在读者调查中的第29题中，沃德·爱德华兹(Ward Edwards, 1968)提供了这个案例：

假设有两个装满了数百万张扑克牌的罐子。在第一个罐子中，有70%的扑克牌是红色的，30%的扑克牌是蓝色的。而在第二个罐子中，有70%的扑克牌是蓝色的，30%的扑克牌是红色的。假设随机选择其中的一个罐子，从中拿出了12张扑克牌：其中有8张是红色的扑克牌，4张是蓝色的扑克牌。你认为你是从第一个罐子中拿出这12张牌的概率有多大？(请以百分数的形式给出你的答案)

如果你的答案和大多数人一样，你会认为概率在70%到80%之间。从本质上讲，你将原先的随机概率50%(两个罐子是随机选择的)调整到了70%到80%之间，主要是因为抽出的红色扑克牌大于蓝色。但是根据贝叶斯法则，正确的答案应该是97%(详细计算见Edwards, 1968)。比起数据可以允许的改变来，绝大多数人对概率估计作出的改变要保守得多。事实上，爱德华兹

(1968, p. 18) 认为: 对该数据的简单分析发现, 要花费两到五次观察才能达到一次观察就能引起的估计的改变。

对风险的知觉

在美国, 每六个死亡的人中就有一个是由于吸烟引起的 (Hilts, 1990, 9月26日)。平均而言, 一个重度男性烟民的寿命将减少 8.6 岁, 而一个重度女性烟民的寿命将减少 4.6 岁 (Cohen & Lee, 1979)。据估计, 平均抽 1.4 根烟增加的死亡概率等于居住在一个核电站 20 英里外长达 150 年的时间 (Wilson, 1979 年 2 月)。

这样的一些统计结果表明, 一些风险程度相同的事件可以通过不同的方式表达出来。对风险的知觉是高度主观性的, 人们所采取的预防措施往往取决于具体的风险呈现方式以及风险的种类 (Stone & Yates, 1991)。例如, 昌西·斯塔尔 (Chauncey Starr) 认为人们经常会“自愿”接受一些风险 (例如, 来自吸烟或者滑冰的风险), 而不愿意接受另外一些风险 (例如, 来自核电站的风险)。正如斯塔尔 (1969, p. 1235) 所指出的那样: “我们不愿意别人来指手画脚, 我们要自得其乐。” 尽管这样的情况并不总是发生 (Fischhoff, Lichtenstein, Slovic, Derby, & Keeney, 1981), 但是在很多情况下往往如此。

风险知觉是一个非常重要的课题——世界上的一些重要决策, 如全球的环境、国家间的安全以及疾病流行病学都越来越关注这样一个概念。但是当人们谈及风险这个概念时它的具体含义是什么呢? 风险是不是与特定行为的不良结果预期数量相等同? 是什么样的特征使得某一个事件显得更具风险性?

正如现实情况一样, 对风险的知觉通常是非常复杂的。对于绝大多数人而言, 风险不同于对某一事件不良结果的预期, 它的含义更加丰富。

保罗·斯洛维奇和他的同事揭示了公众对风险的知觉中的三个基本维度。第一个维度被称为是“恐怖风险”, 其特点是“知觉到缺乏控制感, 恐惧, 灾难性, 不良后果以及利益与风险不相匹配” (Slovic, 1987, p. 283)。这样的维度与一般公众对风险的知觉是最为接近的, 该风险的最极端的例子是与核武器及核能相联系的风险。第二个维度我们称之为“未知的风险”, 这样的风险所包含的特点是“无法观测的, 无法了解的, 新颖的, 以及其所造成的伤害是延期的”。基因和化学技术所造成的风险就是这类风险中的极端例子。最后一个维度就是面对某一个特定风险时个体的数量。尽管还有许多其他的因素可以影响个体对风险的知觉——例如个体对概率估计的保守性、危险的真实性或可得性以及灾难发生的时间特征——但是, 保罗·斯洛维奇和他的同事的一系列实验都证明了这三个维度的存在。

研究人员发现的最有意思的研究结果之一就是, 普通个体与专家相比, 他们对风险的知觉方式是完全不同的 (Slovic, Fischhoff, & Lichtenstein, 1979,

4)。普通个体并不能很好地估计各种风险的大小，他们对风险的总体知觉总是从“灾难发生的可能性”以及“对下一代的威胁”角度来考虑，而不是根据其发生的可能性。另一方面，专家能够很好地估计风险的大小，他们对风险的知觉与其发生的概率是紧密相连的。

举例说明，一般的大众都会将核能视为是一种非常恐怖的风险（主要是基于其发生灾难的破坏力及其对下一代的影响），而专家则认为核能是一种相对较小的风险（主要是依据其发生的概率）。事实上，保罗·斯洛维奇和他的同事（1979年4月）让大学生评估30种不同的行为和技术的风险程度，学生们都将核能视为是最具风险的技术。当要求部分全国范围内广受尊重的风险评估家来评估相同的行为和技术时，他们将核能排在第20位——甚至排在骑自行车的后面。

事故可以使我们更安全吗

1980年6月3日，美国空间战略司令部（SAC）的官员正在进行日常的检查，查看是否有遭受前苏联导弹袭击的迹象。这样的事情通常来讲是不可能发生的，也没有迹象表明会发生类似的事件。

突然，有一台计算机发出了警报，前苏联已经从陆地和核潜艇上发射了携带核弹头的导弹。在几分钟之内，导弹将到达美国本土。

空间战略司令部立即做出了反应。在本土，上百架携带核武器的B-52战略轰炸机已经整装待命，随时准备出发；核潜艇的指挥官也整装待发；导弹发射的指挥官也已经将发射的钥匙插入发射位置。整个美国已经为一场核战争做好了准备。

然而，在第一个警报发出三分钟以后，大家发现这个信号是一个错误信号。美国的军队迅速解除了警报，并且开始了一系列的调查。几天以后又出现了同样的错误，这次，终于找到了事故的元凶，一枚价值0.46美元的计算机芯片失灵了，对于来袭的导弹数字信号，该芯片向电脑间歇性地输入了2，而不是0。

人们对于这个错误警报的反应是很有意思的。基本上讲，核武器的反对者认为这样的事件使他们感到更加不安全，而支持者则报告他们感到更加安全。例如，美国的发言人弗兰克·霍顿（Frank Horton）警告说，这样的错误信号“对国家安全是一个重大的威胁，可能会导致前苏联的过度反应，并引起双方的核战争”；而北美空间防御指挥部的长官詹姆斯·哈廷格（James V. Hartinger）将军在发生错误警报事件之后认为：“我现在对于我们的防御系统更加放心了，因为现实证明这样的错误我们是可以克服的”（U. S. Congress, 1981, May 19-20, pp. 131-133）。

同样的现象发生在三里岛事件以后（Mitchell, 1982; Rubin, 1981; Slovic,

Fischhoff, & Lichtenstein, 1982a; Weinberg, 1981)。艾伦·梅热 (Allan Mazur) (1981, pp. 220 - 221) 调查了 42 位科学家 (在三里岛事件之前, 这些科学家都对核能问题有过公开的表态), 他说: “没有一位科学家因三里岛事件而改变他们对核能的看法……绝大多数科学家对这次事件的解释与他们先前对于核能的观点一致。核能的反对者认为这是一场大灾难, 表明私营企业和政府管理人员并没有能力用一种安全的方式来管理反应堆。而核能的支持者认为在事件中并没有一人死亡, 辐射的扩散量也相对较小, 因此安全防御措施还是能够起作用的。” 双方都在寻找有利于自己观点的证据, 核能的反对者因为发生了这样的核泄露事件而感到非常的沮丧, 并以此为证据证明这样的灾难将来还有可能发生; 而核能的支持者认为这样的安全措施可以起到作用, 并相信将来可以避免这样的事故。

当然, 尽管人们在美国空间战略司令部和三里岛这样的事件之后, 会表现出这样的反应体现了人们决策的偏差, 但是这样的规则也可以在选择不同的曝光媒体和对公众人物形象的选择中产生作用。为了控制这些因素, 本书的作者进行了三个实验, 实验中, 持有不同政治观点的个体阅读了相同的非灾难性技术事故的描述。例如, 在其中的一个实验中, 24 名 Reserve Officers Training Corps (ROTC) 的学员 (核能的支持者) 与 21 位和平主义者 (核能的反对者) 阅读了下面四个相同的有关核武器技术事故的内容。这些事故——尽管是真实的, 但是都不为人知——包括:

- 美国的一个导弹模拟实验由于一个固定螺丝出现故障, 而导致一枚导弹发射升空。
- 一枚俄罗斯的导弹发生故障, 目标直指阿拉斯加。
- 一份假冒的美国海岸警卫队报告谎称总统遇刺, 于是副总统发动了第三次世界大战。
- 俄罗斯海军发出了错误的警报, 声称与美国进入战争状态, 要求舰艇的指挥官准备与美国军队进行战斗。

所有的被试在阅读完所有描述之后要回答一系列的问题, 其中有两个关键的问题。第一个问题是: “当一个人或技术故障发生以后, 在评估将来无意中发生核战争的可能性时, 哪一个因素显得更加重要——是事先准备的预防措施, 还是第一次发生的事故?” 在核能的支持者中大约有 2/3 的人认为, 最重要的因素是所采取的预防措施; 而在核能的反对者中, 除了一位被试以外, 其他所有人都认为最重要的影响因素是第一次发生的事故。正如研究者所预期的, 支持和反对核能的人对相同事件的关注点是不同的。

第二个问题被用来评估以上差异的结果。问题是: “一些人为的和技术的事故发生以后, 并没有引起两个超级大国之间的核战争。这样的事故是否使你更加有信心认为以后并不会发生核战争, 还是更加没有信心?” 答案也是非常

明确的，4/5 的被试都认为他们的信心发生了变化，但是改变的方向取决于他是核能的支持者还是反对者。核能的支持者中有 84% 的被试认为这样的非灾难性事故使他们更加有信心相信，核战争在将来是不可能发生的，但是对于核能的反对者来说，他们的信心反而下降了，见图 12.3。

在另外的两个实验中也发现了同样的现象（细节见 Plous, 1991）。对某项技术的支持者通常关注的是安全措施起到了应有的作用，这样的非灾难性事故使它进一步得到了确认；而反对者强调事故已经发生了，有一些重要的问题已经变得越来越糟糕。另外，对于将来类似事件发生的可能性问题，支持者和反对者所得出的结论也是相反的。当阅读了一个给定的事故描述之后，支持者认为将来发生类似事故的可能性降低了，而反对者认为将来发生类似事故的可能性提高了。

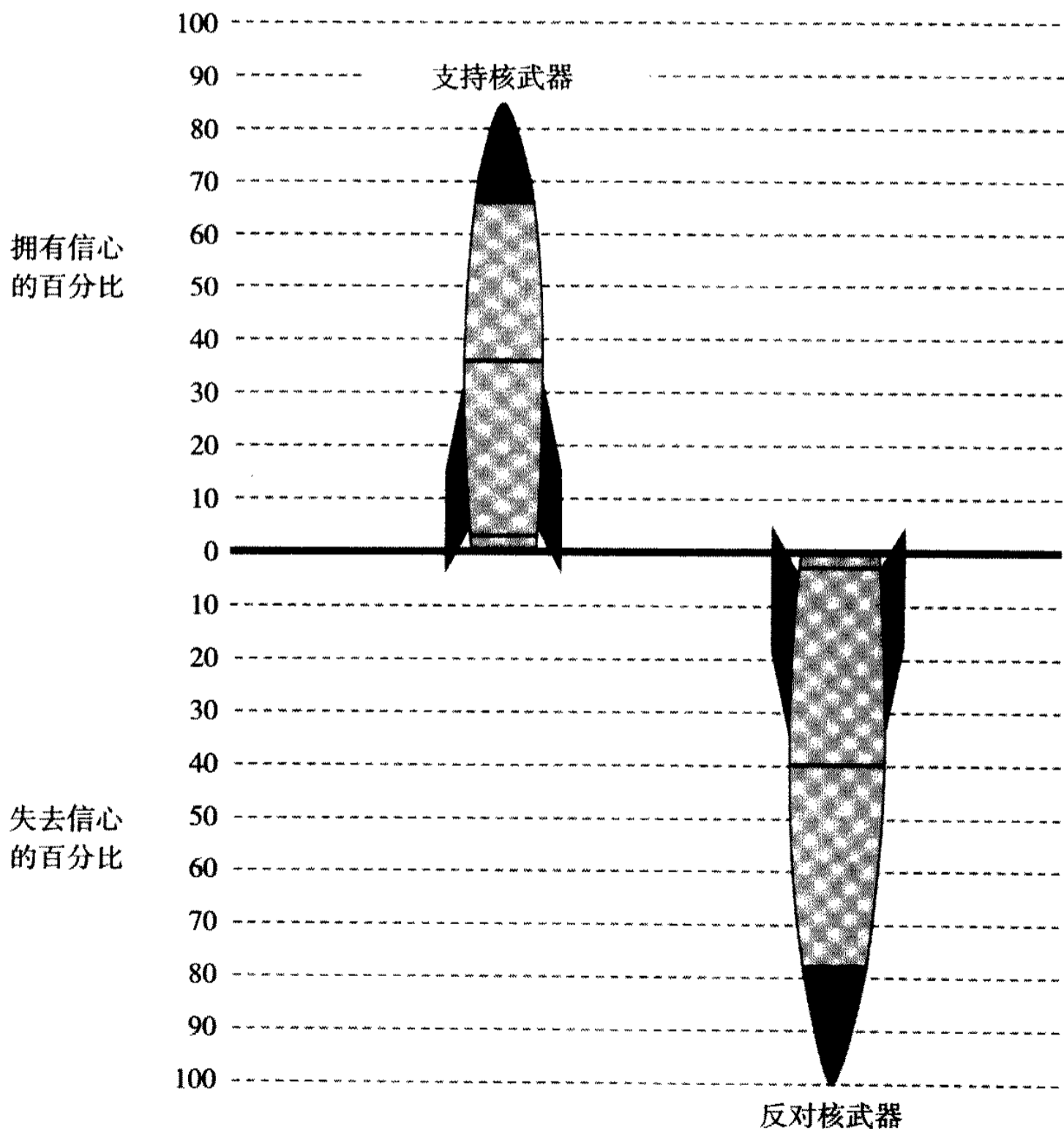


图 12.3 以上这些百分比基于对下列问题的回答：“几个严重的人为的和技术上的失误并没有导致两大集团之间核战争的爆发。这样的失误使您对将来核战争不会发生更加有信心？还是更加没有信心？或者两者都不是？”

无论这样的事故是由于人为的原因还是技术的原因造成的；是美国还是俄罗斯的事故；是与能源相关还是与战争武器相关；是预先知道的还是未知的，上面的研究结论都成立。态度的极化现象同样是很明显的，无论他们是支持者还是反对者。因此，在很多情境中，个体对于风险的知觉通常会受到其先前观点的很大影响。

建 议

决策者在对概率和风险进行判断的时候可以采取以下一些简单的措施来尽量避免偏差：

- ✓ **保持正确的记录。**记录过去发生的某些事件的发生频率，这样就有可能减少首因效应和近因效应的发生，减少可得性偏差以及其他一些由于信息呈现方式的不同而产生的偏差（例如，Hintzman, 1969; Ward & Jenkins, 1965）。
- ✓ **警惕如意算盘。**在许多情况下，人们往往会高估渴望事件发生的概率，同时也会低估自己不希望发生的事件的概率（尽管这样的规则存在着很明显的特例，例如在某些情境下，一些令人忧虑的结果的可得性可能会使人们的概率估计产生波动）。打消这种如意算盘的最好方法是找一个独立的第三方来进行评估。
- ✓ **将复合事件分解为简单事件。**如果一个复合事件是由一系列简单独立事件构成的（如这些事件的结果是互不相关的），一个有用的策略就是分别估计每一个简单独立事件发生的概率。如果这样的复合事件是连续性的，则将这些概率相乘。例如，在投掷硬币的时候连续投掷三次，都是正面的可能性（由三个简单事件组成的复合事件）将是 $.50 \times .50 \times .50$ （因此最后的概率是 $.125$ ）。相反，如果复合事件是非连续事件，将独立事件发生的概率相乘，然后用 1 减去这个概率即可。例如，连续三次投掷硬币，至少有一次出现正面的概率是 1 减去 $.125$ ，最后的概率应该为 $.875$ 。这样的方法只能使用在由统计意义上独立的简单事件构成的复合事件上（像投掷硬币），而如果简单事件的结果是相互关联的，那么这样的方法就不适用了。

结 论

正如前面三章的内容所示，人们对概率和风险的估计容易受到一系列偏差的影响。有一些偏差来自于代表性和可得性偏差这样的直觉，而另一些偏差则是动机因素的结果，还有一些偏差来自于个体事先的预期和信息呈现的方式。

尽管许多偏差是可以修正的，但这些偏差确实可以使我们在进行概率判断时犯错误。在第 13 章中我们将讨论另一种偏差的来源，我们称之为“锚定与调整”，它可以影响概率的评估和许多其他类型的判断。

锚定与调整

在你面前有一个幸运轮，幸运轮上面有一些数字，当你转动这个幸运轮以后，指针定在数字 65 上。下面你需要回答这样一个问题：非洲国家的数量在联合国国家总数中所占的百分比是大于 65% 还是小于 65%？你并不用仔细考虑这个问题，你的答案肯定是小于 65%。

然后问你，非洲国家的数量在整个联合国中占的实际百分比是多少？经过一番思考以后，你给出的答案是 45%。一位研究人员记录下你的答案，并感谢你的参与。然后你就离开了。

现在假定你是另外一个人，你之前并没有回答有关联合国的问题。你转动幸运轮以后得到的数字是 10，而不是 65。当这个幸运轮停止转动以后，一位研究人员问你：你认为非洲国家的数量在联合国国家总数中所占的百分比是大于 10% 还是小于 10%？你肯定认为应该是大于 10%。

然后问你，非洲国家的数量在联合国中所占的实际百分比是多少？

经过一番思考以后，你给出的答案是 25%。

事实上，上面的这些描述正是特韦尔斯基和卡尼曼在 1974 年进行的实验程序和得到的实验结果。被随机分到数字 65 的实验组的被试，给出的平均估计是 45%；而随机分到数字 10 的实验组被试给出的平均估计是 25%。特韦尔斯基和卡尼曼使用了“锚定与调整”的概念来解释这样的现象——也就是说，个体的判断是以一个初始值，或者说是“锚”为依据的，然后会进行并不充分的上下调整。

自从特韦尔斯基和卡尼曼进行了这个研究以后，锚定的现象在很多研究中都得到了证实。研究包括估计 5 岁以下孩子的母亲中职业女性所占的百分比；女性化学教授的百分比以及 NBC 播放的肥皂剧的比例等（Quattrone, Lawrence, Warren, Souza-Silva, Finkel, & Andrus, 1984）。在对一些非数字的情境进行判断时也会出现锚定的现象（Quattrone, 1982）。例如，乔治·夸特龙（George Quattrone）及其同事的研究（1984）发现，事先随机阅读的正性和负性的文章，可以影响个体对 SALT II 武器控制公约的态度。Daniel Cervone 和菲利普·皮克（Philip Peake）的研究（1986）发现，事先的“锚”可以影响个体对自己完成任务的估计——进而影响他们在这些任务中的坚持性。

非常令人惊讶的是，锚定效应并不会因为使用了金钱刺激（为了提高实验

的准确率) (Tversky & Kahneman, 1974; Wright & Anderson, 1989) 或者给予被试更加极端的锚定而消失 (Quattrone et al., 1984)。在一个实验中, 夸特龙和他的同事 (1984) 让被试在事先进行了一些判断之后进行数量的估计, 事先进行的判断包括询问甲壳虫乐队在流行音乐榜排名前十的唱片集的销量是否超过或低于 100 025 套, 大学教科书的平均价格是否超过或低于 7 128.53 美元, 以及旧金山的平均温度是否超过或低于 558 度。夸特龙的研究团队发现较高的锚定数值会导致个体较高的数值判断。根据这样的研究结果, 他们得出的结论是, 锚定效应是一个非常有效的现象, 效应的大小取决于锚和“锚前估计”(在呈现一个锚之前, 个体的平均判断) 之间差异的大小, 直到两者达到相近的水平。如果这是真实的, 那么对于谈判专家、广告人、政治家以及那些专门从事说服工作的专业人士而言, 一开始就采取一个极端的立场对于他们的成功是非常重要的。

思考一些无法想像的事例

遗憾的是, 以往进行的这些有关锚定的研究所针对的都是一些我们并不熟悉或者并不关心的话题: 例如在联合国中非洲国家所占的比例, NBC 播出的肥皂剧的比率以及女性化学教授的比例等等。相对而言, 很少有研究关注那些人们关心的热门话题, 或者是已经经过人们思考的问题, 或者人们已经相当熟悉的事件等等。属于这类问题的一个最为极端的例子也许就是, 人们对爆发核战争可能性的估计。与锚定研究中的其他问题相比, 核战争的爆发是很多人已经经过认真思考, 并且也是一个非常令人关注的话题。然后, 问题就是考察在给定一个较高或者较低的锚以后, 人们对于核战争爆发可能性的判断是否会受到影响。为了回答这个问题, 从 1985 年 1 月到 1987 年 5 月, 作者进行了一系列相关的研究, 一共采访了大约 2 000 名被试。

在低锚定的实验条件下, 要求被试首先回答美国与前苏联之间爆发核战争的可能性是否超过 1%; 而在高锚定的实验条件下, 要求被试回答两国之间爆发核战争的可能性是否超过 90% (在其中的一个实验中, 这个比率是 99%)。在被试回答这样的问题以后, 要求被试以百分比的形式来估计他们认为核战争爆发的可能性。第三组被试是属于没有锚定的控制组, 被试并没有回答第一个问题而直接判断核战争发生的可能性。

锚定刺激对被试核战争概率的估计造成很大的影响。被问及核战争的概率是否大于或小于 1% 的被试比起那些没有得到具体锚定的被试来会做出较低的估计; 相对而言, 被问及核战争的概率是否大于或小于 90% (或 99%) 的被试比起那些没有得到具体锚定的被试来会给出较高的估计。这样的差异并不受陈述方式——即核战争爆发或者核战争不爆发 (心理学上称为框架效应) 的影响, 同样, 被试对核战争的熟悉程度和关心程度也不会对结果产生影响。

这样的结果至少可以有两个含义。第一，问题中的一些数字参考值可能会产生意想不到的效果。例如，美国1984年进行的一项民意调查中问道：“你是否支持美国建立一套防御体系来抵御核导弹和轰炸机的攻击，如果这套体系能够击落90%的来自前苏联的核导弹和轰炸机？”（引自Graham & Kramer, 1986）。由于提及了90%这样一个数值（对于大多数人而言，这是一个很高的标准），很多人今后都会将整个战略防御体系的有效率锚定在90%上。第二，当支持者举出最好或者最坏的例子来表明他们的立场的时候，他们可能无法避免将个体的观点锚定在相反的方向上。例如，1986年《纽约时报》的一位专栏作家认为：“即使整个战略防御体系能够抵御99%的来自前苏联核武器的攻击也是远远不够的，因为逃脱防御体系的核武器所具备的强大攻击力还是可以给予致命的打击（Gliksman, 1986年2月13日）。”与他的目的恰恰相反，这位专栏作家给读者提供了一个锚定值，就是战略防御体系的有效率可以达到99%。

房地产到底有多真实

锚定效应并不仅仅存在于对核战争爆发的可能性以及战略防御系统有效性的估计上，例如，前一年的预算支出可能会影响下一年预算的制定，谈判中的起始位置可能会影响下一轮谈判。同样有证据表明，陪审团向法官陈述证据的顺序（从严重的到轻微的）——这也是现在正在使用的一种标准程序——可能会导致更严厉的判决（Greenberg, Williams, & O'Brein, 1986）。

格雷戈里·诺思克拉夫特和玛格丽特·尼尔（Gregory Northcraft & Margaret Neale, 1987）的研究显示，锚定效应还能够影响个体对房地产价格的判断。在这个研究中，一些房地产代理商有机会参观位于亚里桑那州图森（Tucson）的两个房地产项目——其中一个评估的价格是74 900美元，而另一个评估的价格是135 000美元。在他们参观的过程中，这些经纪可以得到一个10页纸的资料，其中包含了所有通常能够决定一个房产价格的因素（除了官方评估的价格）：标准的多重上市服务表（MLS），周边区域正在出售的房产的多重上市服务表以及周边刚刚售出的房产信息等等。所有代理商得到的房产信息绝大多数是相同的，其中只有一处存在差异：一部分代理商得到的评估价格比真实评估价格低11%~12%，一部分代理商得到的评估价格比真实评估价格低4%，一部分代理商得到的评估价格比真实评估价格高4%，其余的代理商得到的评估价格比真实评估价格高11%~12%。诺思克拉夫特和尼尔希望考察评估信息中的价格差异是否会影响房产代理商对价格的最终评估。

这些代理商有20分钟的时间参观整个房产，然后给出他们对房产价格的估计：（1）房产的评估价格，（2）房产广告中的销售价格，（3）购买该处房产的合理价格以及（4）销售商能够接受的最低销售价格。表13.1总结了房产代理商对价值135 000美元房产的最终评估价格（价值74 900美元的房产得到

表 13.1 房地产价格的锚定效应

资料中标明的价格	房地产代理商给出的价格平均数			
	评估价格	建议销售价格	合理价格	最低接受价格
119 900	114 204	117 745	111 454	111 136
129 900	126 772	127 836	123 209	122 254
139 900	125 041	128 530	124 653	121 884
149 900	128 754	130 981	127 318	123 818

注：该表来自于格雷戈里·诺思克拉夫特和玛格丽特·尼尔（1987）的研究。

了类似的结果)。正如你在表格中看到的，代理商都认为评估信息中给出的价格是偏高的（无论他们的评估信息中的价格如何），他们的四个估计值也证明了锚定效应的存在。但有趣的是，当被问及影响他们做决策的主要因素时，只有十分之一的代理商提到了资料中提供的价格信息。

这样的结果显得尤为重要，主要原因有以下几点：第一，证明在现实的生活情境中，锚定效应也是存在的。通过改变 10 页资料中的一条信息，诺思克拉夫特和尼尔就可以控制房地产代理商对地产价格的评估，波动幅度已经超过了 10 000 美元。第二，这样的结果表明专家并不能避免锚定效应。绝大多数的代理商都有多年经营房地产的经验，但是多年的经验并不能避免他们的决策受到锚定效应的影响。最后，只有很少的代理商将资料中的价格因素作为他们决策的重要因素。尽管有一些代理商可能不愿意承认自己的评估价格依赖于他人的评估价格，但可能许多的代理商确实没有意识到自己的评估受到了他人信息的影响。

锚定效应的其他例子

在读者调查中还有一些锚定效应的例子。例如 12a 的调查题中，让你估计一张纸折叠 100 次以后的厚度。只有很少人给出的答案大于几英尺或者几米，但是正确的答案远远大于此。假定一张纸的原始厚度是 0.1 毫米，折叠 100 次以后的厚度大约是 1.27×10^{23} 公里——是地球和太阳之间距离的 800 000 000 000 000 多倍！这样的答案通常会使人感到惊讶，因为他们的想像都来自于刚开始的几次折叠（一个非常低的锚定值），并不能很充分地估计平方的效应。正确的答案是这样得来的，折叠 100 次后纸张的层数一共是 2^{100} ，假定每一张纸的厚度是 0.1 毫米，则总厚度为 1.27×10^{29} 毫米，相当于 1.27×10^{23} 公里。

锚定效应的另一个例子来自于读者调查中的 17 题，问题是如果要将全世界人的血液都集中到一个立方体中，这个立方体的宽度将是多大？许多人在回答这个问题的时候，首先想到的是全世界的人口数量。这样的数据提供了

一个很高的锚定值，因此当他们估计立方体的宽度时，并不能充分地从高锚定值上调整下来。根据约翰·艾伦·保罗斯（John Allen Paulos, 1988）的估计，这个题的正确答案仅仅是 870 英尺。保罗斯当时假定的全球人口数量是 50 亿，平均每一个人血液的容量是 1 加仑。因为一立方英尺可以容纳 7.5 加仑，因此容纳全球所有人血液的容器需要 670 000 000 立方英尺。结果，立方体的宽度就是 870 英尺（ $870^3 = 670\,000\,000$ ）。

特韦尔斯基和卡尼曼（1974）进行了一个实验，同样说明了锚定效应。实验的内容正如读者调查 22 题中给出的乘法问题。当特韦尔斯基和卡尼曼要求一组学生估计 $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ 这个题的答案时，得到的平均数是 2 250；而当他们要求另外一组学生估计 $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$ 的答案时，得到的平均数只有 512。很显然，计算降序数列的学生首先锚定的答案是 8、7、6、5 的乘积，而计算升序数列的学生锚定的是 1、2、3、4 的乘积。但是事实上，这两组学生的答案都远远低于这个题目的正确答案。正确答案应该是 40 320。

最后，读者调查中的第 13 题是这样一个问题：

包括 2 月 29 日在内，一年中有 366 个可能的生日。因此在一个小组中，至少要 367 个人存在，才能确保至少两个人的生日是相同的。如果要确保有 50% 的可能性两个人的生日是相同的，至少需要多少人？

绝大多数人给出的答案是 183 人左右（一年中一半的天数）。而正确的答案只是需要 23 人。也就是说，一个随机的由 23 人组成的团队，其中至少有两个人同一天生日的可能性超过 50%。

我们考察一下这个答案的来由。首先我们需要考虑的是 2 个人不是同一天生日的概率（例如，其中一个人的生日是 3 月 11 日，而另一个人的生日不是 3 月 11 日）。这样的概率是 $365/366$ ，或者是 99.73%（为了简化答案，我们假定一个人的生日出现在一年中任何一天的概率是相等的）。

现在考虑 3 个人不是同一天生日的概率。这样的概率等于其中的两个人不是同一天生日的概率（如上文所述， $365/366$ ）乘以第三个人的生日与前两人不是同一天生日的概率（ $364/366$ ）。换一句话说，3 个人不是同一天生日的概率为 $(365/366) \times (364/366)$ ，或者 99.18%。

依据相同的逻辑，4 个人不是同一天生日的概率是 $(365/366) \times (364/366) \times (363/366) \times (362/366)$ ，或者是 98.37%。同样，23 个人不是同一天生日的概率为：

$$\frac{365 \times 364 \times 363 \times \cdots \times 344}{(366)^{23}} \text{ 或者 } 49\%$$

因此，我们可以看到 23 个人不是同一天生日的概率为 49%。也就是说，这个小组中至少有两个人同一天生日的概率超过 50%。

这样一个问题对于大多数人来说是很困难的，部分原因也是因为存在锚定效应。因为许多人对于这个问题的第一反应是一个很高的锚定值（例如183），然后他们发现很难将这样的值进行下调。这样的一种模式与巴-希勒尔（1973）在本书第12章中论述的研究结果比较一致。人们通常会低估一个非连续事件的概率——例如，23人团队中至少有两个人生日相同的概率。

在下面这个问题中，锚定效应同样存在：

为了使团队中至少有一个成员的生日为特定的某一天——例如，7月4日——的概率为50%，至少需要一个由多少人组成的团队？

在你阅读下面文字之前，请你猜一下这个问题的答案。

解决这个问题的方法与上一个问题是类似的。一个人的生日不是7月4日的概率为 $(365/366)$ ，同样，由于不同的生日相互之间是独立的，两个个体的生日不在7月4日的概率为 $(365/366) \times (365/366)$ ，或者是 $(365/366)^2$ 。因此，N个人的生日不在7月4日的概率为 $(365/366)^N$ 。通过这个公式的计算，当N等于254时，概率小于50%。也就是说，在一个254人的团队中，至少有一个人的生日在7月4日的概率大于50%。与前一个有关生日的问题类似，许多人认为这样一个问题的答案在183人左右，然后他们发现很难再从这样一个锚定值进行调整。

在讨论完这些问题以后，保罗斯（1988，p. 37）得出了以下的结论：“原则就是……一些看起来不可能发生的事情是有可能发生的，但是一个具体事件发生的可能性要小得多……一个自相矛盾的结论就是：想让一些不可能发生的事情不发生，是不太可能的。”只要一个23人的团队，就有50%的可能性使其中两个人的生日同在不确定的某一天；但是如果需要有50%的可能性使某人的生日在特定的某一天，需要的人数则是10倍以上。偶然事件是普通的，但是特定的偶然事件则不然。在第14章中我们将详细讨论几个著名的偶然事件。

结 论

锚定效应是普遍存在、不可动摇的。很多研究都指向了一个相同的目标：人们从一个锚定值进行的调整都是不充分的，无论是关于核战争爆发可能性的判断，还是房地产价格的判断，或者是对其他数值的判断。

人们很难抵御锚定效应的影响，部分原因在于对正确性的激励很少能够产生作用；部分原因在于锚定值通常不会引起人们的注意。首先需要注意的是，观察相关的数值中是否存在明显偏高或者偏低的数值。这些锚定值很可能使决策产生偏差。

在理想世界中，决策者可能会低估或者忽略这样的信息，但是在实际操作中是很难做到的（Quattrone et al., 1984）。因此，最有效的方法是针对原有的

极端锚定值，确定一个反方向的相等的锚定值。例如，在评估一个房地产的价格时，如果发现其价格是大大高估的，决策者这时可以想像一个相当低的销售价格作为另一个锚定值。这样的技术类似于我们在第6章中所推崇的发展多框架技术。

在这里，有一点需要进一步强调。由于极端的锚定值可能产生最大的锚定效应，而锚定效应通常又是不容易察觉的，因此我们要认识到，对最佳或者最差案例的讨论可能会引起不易察觉的锚定效应。例如，在考虑了理想状态下商业投资的可能性以后，就很难从现实的角度进行考虑。同样，在估计了对手的最大武器储备以后，就很难做出正确的武器储备评估。再一次强调，在做出最后的决策之前，考虑多个锚定值是非常有必要的。

对随机性的知觉

沃伦·韦弗 (Warren Weaver, 1982) 在他关于概率论和对运气的知觉一书中, 详细描述了他的邻居乔治·布赖森先生一次非同寻常的商务旅行。根据韦弗的描述, 布赖森先生坐火车在圣路易斯和纽约之间旅行。当他登上火车以后, 布赖森先生问列车员他能否在肯塔基的路易斯维尔停留一段时间, 这是一个他从来没有去过的城市。他得到了肯定的回答, 因此在路易斯维尔下了车。在火车站, 布赖森先生询问了当地住宿的情况, 并决定住在布朗酒店。他在登记处登了记, 房间号是 307。布赖森突然心血来潮, 问服务员是否有他的信件。服务员将标有“乔治·布赖森先生, 307 房间”的信件递给了他。这才发现这间房间的上一位住客也叫乔治·布赖森。

像这样的偶然事件发生的概率有多大呢? 许多人的自然反应就是这样的事件发生的概率一定小于百万分之一。毕竟, 并不是每一天在一个特定城市的特定旅馆的特定房间会连续入住两个都叫“乔治·布赖森”的人。

但是用这样一种思维方式来考虑偶然事件会令人误解。问题并不是这样一个特定事件发生的概率。问题是在某一个时间内, 某个城市的某家酒店的某个房间连续入住的两个客人拥有相同名字的概率有多大。正如我们在第 13 章中所讨论的有关生日的问题。一个非特定事件发生重合的可能性远远大于一个特定事件发生重合的可能性。

事实上, 像重名这样的偶然事件是经常发生的, 远高于人们对这类事件发生概率的想像 (见图 14.1)。下面又是两个其他的例子:

1990 年 10 月 5 日, *San Francisco Examiner* 报道了这样的新闻: 英特尔作为全球领先的计算机微处理器制造商, 正在起诉另一家芯片制造商侵犯了英特尔 386 微处理器的商标 (Sullivan, 1990 年 10 月 5 日)。英特尔了解到对方公司计划生产一种叫做“Am386”的芯片, 而且英特尔公司发现这个侵权行为的过程是非常巧合的。事情是这样的, 两个公司都有一个叫做迈克·韦布的员工, 这两个员工在同一时间住进了加利福尼亚州 Sunnyvale 的同一家酒店。然后, 当两个人都离开酒店以后, 酒店收到了一份给其中一个迈克·韦布的包裹。这个包裹——里面保存着有关 Am386 芯片资料的文件——被错误地传递给了英特尔的迈克·韦布, 他将这些证据转交给了英特尔的律师。

第二个巧合同样发生在 1990 年的夏天, 当时弗兰克·威廉·布马和他的妻

Mixup in Cars Proves Startling Coincidence

MILWAUKEE, April 2 (AP) — The police estimated the odds as “a million to one.”

Two men with the same surname, each owning a car of the same model and make and each having an identical key, wound up at a Sheboygan shopping center at the same time on April Fool's Day.

Richard Baker had taken his wife in their 1978 maroon Concord model of an American Motors Corporation automobile to shop at the Northgate Shopping Center in Sheboygan.

He came out before she did, put his groceries in the trunk and drove around the lot to pick her up.

The two were coming back from another errand when, “My wife said, ‘Whose sunglasses are these,’” Mr. Baker said.

“I noticed I had a heck of a time getting the seat up and it was way back when I got in,” he said. “My wife said: ‘Something's wrong. Look at all this stuff in here. This isn't our car.’”

A check of the license plates confirmed it: The Bakers had the wrong car.

Meanwhile, Thomas J. Baker, no relation to Richard Baker, had reported his car stolen.

Thomas Baker and the police were waiting for Richard Baker and his wife when they drove back into the shopping center parking lot.

The police tried the keys on the cars and both keys operated and unlocked both cars.

Sgt. William Peloquin said estimated the odds of such an incident as “a million to one.”

子特露迪将他们结婚五十周年纪念的相片寄往 *Grand Rapids Press* (密歇根当地的一家报纸)。大约在相同的时间,这家报纸同时收到了另一个弗兰克·威廉·布马和他的妻子内拉结婚五十五周年纪念的相片。令人感到惊讶的是,他们的结婚纪念日都在7月9日。考虑到这样的巧合本身就是一个有趣的故事,*Grand Rapids Press* 专门写了一篇文章详细讲述了两个弗兰克·布马其他的相同之处,例如他们都有一个叫做马西娅的姐姐 (Malone, 1990年9月9日)。之后这个故事又被美联社收入,最后又出现在 *National Enquirer* 杂志上 (Carden, 1990年9月25日)。

不可能事件

你是否能够想像一个巧合是这样的不可思议,以至于你开始怀疑这是否仅仅是一个偶然事件?下面一个由理查德·布洛杰特 (1983. 11. p. 17) 讲述的故事就是这样:

1914年,一位德国的母亲为她的男婴照了相,并将底片留在斯特拉斯堡的一家照相馆进行冲洗。在当时,胶卷的底片是单独购买的。这时第一次世界大战爆发了,她没有办法回到斯特拉斯堡去取回自己的相片。两年以后,这位妇女在将近200英里以外的法兰克福购买了胶卷的底片,目的是为了给她新近出生的女儿照相。当照片冲洗出来以后,她发现这些底片被两次曝光,她女儿的相片是在原来她儿子的底片上又再一次曝光。如此阴差阳错,她先前给儿子照的相片并没有被冲洗,而被误认为是没有使用的新底片,最终又重新回到了这位妇女的手中。

这个巧合是非常有名的,因为这样的故事是瑞士心理学家卡尔·荣格建立其“共时性”(synchronicity)理论的基础。根据荣格的理论,巧合发生的可能性远远大于人们基于概率的判断,而事实上有一种未知的力量在对整个宇宙的秩序进行调控。

从很多方面来看,荣格的理论类似于宗教对巧合的解释,似乎这一切都是上帝的杰作。一个流行甚广的例子出现在1950年的《生活》杂志 (*Life*) (Edeal, 1950. 3. 27)。当年的3月1日,在内布拉斯加州的比阿特丽斯,15位 West Side Baptist 教堂唱诗班的成员约定在下午的7点15分进行排练,但是由于各种各样的原因,这15位成员都迟到了。主持一家迟到是因为他的妻子在临出发的最后一刻熨坏了女儿的礼服;一些成员迟到是因为他们没有办法发动自己的汽车;而钢琴伴奏原本打算提前半个小时到达教堂,但是他吃过晚饭以后就睡着了。总体而言,大约有10种毫不相干的理由造成了这次集体迟到。

但是最后的事实证明,对于每一位迟到的人而言,他们都是非常幸运的。正如在《生活》杂志上报道的:“在7点25分,喊叫声充斥了比阿特丽斯的每

一个角落，West Side Baptist 教堂发生了爆炸。教堂的墙体向外倒塌，沉重的木屋顶被炸得四分五裂……消防队员说爆炸是由于天然气泄漏引起的。这些唱诗班的成员开始庆幸，正如其中的一位说道：‘这真是上帝的杰作’。”

当然，这样的“巧合”完全有可能真是上帝的安排，荣格的共时性理论也完全有可能是正确的。但是这样的解释并没有实证的证据加以支持或者反驳。即使是最不可能发生的事件也不能将无神论者转化为信教徒，因为无神论者总是可以运用成千上万概率组成的函数来解释不可能发生的事件（例如，Alvarez, 1965）。同样，绝大多数的信教徒也不会为无神论者的解释所动，因为他们的信念并不是仅仅建立在概率之上的。

那么，我们可以使用什么样的方法进行验证呢？回答下面两个相关的问题就是一个办法：（1）个体是否总是倾向于将有意义的模式看成一些随机事件的组合；（2）人们是否能够随机地行为？

对于第一个问题，我们在第 10 章中已经部分涉及了，当时我们着重讨论了篮球比赛中投篮命中率的错觉。下面一部分我们将对研究者进行的关于随机知觉的研究进行回顾。

运气和迷信

大约有 40% 的美国人都相信某些数字对于某些人而言是幸运数字（“Harper's Index”，1986. 10）。这种想法是否正确呢？

正如罗伯特·胡克（Robert Hooke）在《统计是如何教人撒谎的》一书中写道：

一个自称非常“运气”的个体总是有良好的感觉，因为至少直到现在他/她还是比较幸运的。如果这个运气能够延伸到将来，使得他/她比我们大多数人有机会赢得彩票，人们就会将这样的运气归结为迷信。而且，这种运气同样得到了那些自称不相信运气的人的称道。

关于迷信产生的最早的实验是由哈罗德·黑克和雷·海曼（Harold Hake & Ray Hyman, 1953）完成的。在 240 个练习中，黑克和海曼呈现给被试两种刺激中的一种——一组水平发光的氖灯泡或者是一组垂直发光的氖灯泡。在每一次实验之前，要求被试预测一下下一次刺激是水平的还是垂直的。黑克和海曼呈现给被试的是四种不同顺序中的一种，但是每一种刺激的顺序都是随机的——其中有一半的刺激是水平的，另一半的刺激是垂直的，且没有可以察觉到的模式（这里面存在一定问题，即刺激的序列是否是完全随机的，因此在这里使用“随机”一词主要是表示这样一个序列并不存在可以分辨的模式）。

黑克和海曼的研究结果发现，大约有一半的情况下，被试预测下一次的刺激是水平的，这里并不考虑预测前的刺激是否是水平的。从这方面讲，他们的

猜测类似于他们所观察到的随机序列。

但是被试的选择并不是完全随机的，他们的下一次预测要受到前一次预测正确与否的影响。如果前一次预测他们猜对了是水平的刺激，且那么在这一次预测中，他们有64%的可能性继续猜测是水平的刺激；如果前两次刺激都是水平的刺激，且均被猜中了，那么被试在这次预测时，有72%的可能性继续猜测是水平的刺激。换句话讲，尽管这样的序列是随机的，但是被试的猜测还是会迷信上一次猜测的正确与否。黑克和海曼（1953，p. 73）得出的结论是：“如果参加这次实验的被试就是普通的被试，（这就意味着人们）总是倾向于将一些模棱两可的事件序列知觉为更加有结构的事件。产生这种现象的主要原因是，他们总是用他们过去行为的经验来预测将来事件发生的规律，尽管事实上这些事件与个体过去的行为是完全没有关系的。”

约翰·赖特（John Wright, 1962）做了一个类似的实验。赖特呈现给被试的是一个仪表盘，上面是由16个各种形状按钮均匀组成的一个圈，圈的中央是第17个按钮。要求被试根据中间按钮的形状来预测下一个形状是什么。如果他们的预测是正确的，他们就会听到一声蜂鸣声并得到一分；如果他们的选择是不正确的，他们就什么都得不到，接着开始下面的实验。事实上，并不存在正确的答案，奖励的给出完全是随机的。在其中的一个实验条件下，被试获奖几率是20%；在另一个实验条件下，获奖几率是50%；而在最后一个条件下，获奖几率是80%。实验的结果与黑克和海曼发现的结果类似——对于这样一个随机的反馈，被试倾向于从中遵循一定的模式。他们通常会发展出一定的迷信行为，对某种形状的按钮有特别的偏好。这样的一种迷信行为在获奖几率最高（80%）的实验组中最为明显。

认识随机性

在黑克和海曼（1953）以及赖特（1962）的研究中，被试都要对许多的实验内容进行判断，而在每一次判断后都可以得到随机的反馈结果。在上面的两个研究中，被试都倾向于将随机的反馈结果认知为一种特殊的反应模式。但是这样的研究存在一个问题，就是实验任务本身的特性会引导个体从中找出一定的模式。如果心理学家是对知觉和学习这样的主题感兴趣，为什么他们要呈现给被试的是一个随机的反馈呢？至少应该给被试一个提前的警告，告诉他们反馈的结果可能是随机的，这样的做法可以给被试一个合理的机会使他们了解所获得的反馈是随机的。

1980年，克里斯托弗·彼得森（Christopher Peterson）发表了他的一个研究，实验的程序与黑克和海曼的实验大致相同，惟一的区别就是在有些实验条件下，被试不仅需要预测下一个刺激将是什么，还要判断实验所呈现的刺激是“完全随机的还是符合一定的规律”。其中一半的被试所得到的刺激是完全随机

的，而另一半的被试所得到的刺激则遵循一定的规律。彼得森的研究结果发现，如果明确地告诉被试刺激可能是完全随机的，那么被试通常能够将这样的刺激序列知觉为随机的或者接近随机。

当然，如果在前面的尝试中相同的选择被证明都是正确的，那么被试对随机性的知觉并不能完全消除被试的一些迷信。彼得森并没有发现，如果事先告诉被试某个序列可能是随机出现的，那么他们就不会出现迷信行为。他的研究只是发现，当被试知道随机是一种可能的时候，事实上许多人都可以进行正确的判断。但是正如我们在下一个部分中讲到的，对“随机性”含义的知觉并不总是正确的。

在随机中发现特定的模式

在1970年，荷兰研究者威廉·瓦根纳（Willem Wagenaar）发表了一篇有趣的对随机性知觉的研究报告。瓦根纳（1970a）给被试呈现的刺激是一系列的幻灯片，每一张幻灯片都是由7个系列的白色或者黑色的点组成的，而背景则是灰色的。在每一张幻灯片上，要求被试判断哪一个系列看起来最像一个随机的序列（就像前面投掷硬币的实验一样）。在每一张幻灯片中，其中一个系列有20%的可能产生重复（也就是黑色的点跟着一个黑色的点，或者白色的点跟着一个白色的点的概率），其中的一个系列有30%的可能性产生重复，以此类推，直到80%。在一个随机的序列中，黑点和白点转换的可能性是相等的，因此其重复的概率是50%（就像在投掷硬币的实验中，投掷正面以后再一次投掷正面的概率），所以，如果被试能够正确地分辨出7个系列中哪一个是随机的，他们就应该选择重复概率为50%的序列。

事实上，瓦根纳的研究结果发现，被试认为40%重复概率的序列最像一个随机的序列。也就是说，人们认为相同的点重复出现的概率在40%的情况下，这样的序列最像一个随机序列。^{*}与真正的随机序列相比，人们期望一个随机序列中不同颜色的点之间的转换更频繁。瓦根纳发现尤其是在对长度较长的序列进行判断的时候，这样的情况就更加明显——例如相同颜色的点重复出现6次或6次以上的序列。因此，对于一些人来说，他们倾向于将存在一定规则的序列看成是随机的，而将一些随机的序列看成存在一定的模式。正如我们前面

* 彼得·艾顿，安亨特和乔治·赖特（Peter Ayton, Anne Hunt, & George Wright, 1989）批评这一系列的研究误导决策者将随机性与“看起来像随机的”这样一个概念等同起来。一个足够长的随机序列中总是包含着一些看起来不像是随机序列的序列片断，在这样的情况下如果要求被试来选择随机序列，就可能使他们产生偏差，从而避免过多地选择重复。但是另外的一些研究发现，即使要求被试观察整个序列是否是随机产生的——而不是提及这个序列看起来如何，被试同样也会低估重复的概率。

所讲到的，托马斯·吉洛维奇（1985）等人在命中率的研究中发现了同样的结果。人们认为转换的概率在70%（相当于重复的概率为30%）的时候，这样的序列更加类似于一个随机的序列。

人们能够随机行为吗

人们对随机性进行判断的时候存在如此的困难，这就意味着如果要让他们自己来产生一个随机的序列，同样也会存在困难——事实也是如此。关于这个问题的讨论最早来自汉斯·赖克巴赫（Hans Reichenbach, 1949）的研究，他发现个体并不能自己制造一个随机的序列，即使是让他们竭尽全力来完成这样的工作。

这样一种能力的匮乏在保罗·巴肯（Paul Bakan, 1960）报告的一个实验中体现得更为明显。巴肯要求70名大学生来“制造一个随机序列”：假设以没有偏差的方式一共投掷硬币300次，正面（H）和反面（T）出现的随机序列将是怎样的。按照前面的论述，在300次的投掷中，将出现150次正面和反面之间的转换。但是巴肯的研究结果发现，接近90%的学生完成的随机序列中出现了过多的转换（平均为175次）。这些发现与瓦根纳（1970a）的研究结果类似——人们希望在一个随机序列中出现比实际情况更多的转换。

当要求个体的选择范围多于两种的时候，这种趋势就更为明显（Wagenaar, 1970b, 1972）。例如，当人们面对一组6个或者8个不同形状的按键时，要求他们组成一个随机序列，这时出现转换的次数要比只有两个选择时多得多。这样的实验结果对于我们日常生活的启示是，我们在日常生活中往往会遇到多于两种的选择（不仅仅局限于正面和反面，垂直和水平），任何一个结果的重复都可能使人们将它视为是一个非随机的序列，人们期望在随机序列中产生比随机序列应有的转换更多的转换。

学习如何随机地行为

在1986年，艾伦·纽林格尔（Allen Neuringer）的研究显示，通过一段时间的培训，可以使人们“随机”地行为。纽林格尔设计了一个计算机程序，这样的计算机程序能够给出5或10种关于随机性的统计测量指标，例如，行为的数量，在一个行为系列中某种选择所占的百分比等等。被试的任务就是在计算机的键盘上使用任意的速度来按两个不同的键，在他们按键的过程中，计算机会对他们的行为进行反馈。

纽林格尔观察到在实验的初期，被试总是产生一些非随机的序列，正如瓦根纳和其他研究人员发现的结果一样。但是，通过几千次的按键和得到相应的反馈以后，被试能够产生一个长度较长（6000次的按键）的随机序列，使用

纽林格尔在实验中用于评价随机性的统计指标，这样的序列完全可以达到随机序列的要求。这样的结果表明，人们可以使用一种类似随机的方式来行为，但必须是在他们经过足够的培训以后。尽管我们中间很少有人能够参加这样类似的培训——即使我们能够参加这样的培训，也很少有人愿意用这样一种随机的方式来行为——但是纽林格尔的研究还是具有重要的意义，因为它强调了人们对随机性的错觉并不是永恒不变的。经过足够的培训，我们可以随机地行为。

结 论

这些对随机性知觉研究的实际意义是什么呢？日常生活中最常见的结果并不会仅仅局限于硬币的正面和反面，或局限于垂直的或者水平的线条；公司中每天都在雇佣和解雇员工，运动场上到处是成功和失败，股价每天都在涨跌。在这样的一些结果中寻求一定的规律当然无可厚非，但是关于随机性知觉的研究结构表明决策者倾向于过分解释随机事件。

例如，尽管在证券市场中，股价的震荡类似于在华尔街上“随意漫步”（Fama, 1965; Malkiel, 1985），但还是有成千上万的人花费整天的时间对股价的走势进行预测。事实上，巴鲁克·菲什霍夫和保罗·斯洛维奇（1980）的研究发现，在得到了股价信息和趋势信息以后，65%的被试确信他们能够正确预测股市的走势，尽管这样的预测只有49%的正确率，几乎类似于投掷一枚硬币。

本章中的研究显示，我们很容易从一个随机序列中发现一定的模式。在目睹了三至四次相同的结果以后，绝大多数个体就可以得出结论，这个序列中必定存在一定的模式。当然，如果特定的结果在现实生活中并不是经常发生的，三到四次相同事件的发生确实具有一定的信息性（例如，很少进行人员调整的公司，在很短时间内发生了三次人员调整）。但是如果情境都是由独立事件组成，且结果相同，那么三到四次相同结果的发生并不是不寻常的，决策者应该尽量避免将相同结果的短期重复看做是有意义的。

相关、因果关系与控制

在过去的一个月中，你发现自己头晕的症状出现得更加频繁了。刚开始，你还以为是因为自己刚刚坠入爱河的缘故，后来，你觉得可能是因为上个月看了太多的心理学书籍。但是现在，你开始担心起来了——一位朋友告诉你，头晕可能是大脑长肿瘤的缘故。

所以，你决定了解一下大脑肿瘤的知识。在图书馆，你发现了一篇关于大脑肿瘤的研究论文，它对医院中 250 位神经科病人进行了如下的分类：

		大脑肿瘤	
		患者	无
头晕	有症状	160	40
	无症状	40	10

请仔细阅读上面图表中的数字，然后回答下面两个问题：（1）为了确定头晕和大脑肿瘤之间的关系，上表中的哪些数据是必需的？（2）根据上表中的数据，你认为头晕和大脑肿瘤之间存在联系吗？

后一个问题在心理学中被认为是“协相关测量”（covariation assessment）中存在的问题，或者是两个变量关系之间的测量（是否是“共同变化”的）。尽管这样的判断看起来还是比较简单（平均而言，几乎每一个人都知道多云就意味着多雨，高度越高重量越大的道理等等），但是事实上协同测量通常是比较困难的（Jenkins & Ward, 1965; Ward & Jenkins, 1965）。为了考察你是如何对这一问题进行判断的，请参见你对读者调查中第 14 题的回答（摘自 Nisbett 和 Ross, 1980, p. 91）。

通过上面的数据，许多人认为头晕和大脑肿瘤之间有关系，因为在同时患有头晕和大脑肿瘤的单元格中，病人的数量最多。有些人认为这两者之间存在正性的联系，因为同时出现头晕和患有大脑肿瘤的患者以及没有出现头晕症状同时没有大脑肿瘤的患者（左上角和右下角）数量之和远远大于出现头晕症状但是没有大脑肿瘤的患者以及没有出现头晕症状但是患有大脑肿瘤的患者数量之和（左下角和右上角）。然而，这样的决策原则通常并不能得出正确的答案；如果一个事件发生的概率（例如患有大脑肿瘤）取决于另一个事件（头晕）是

否发生，那么我们才能说这两个变量相关。

为了确定头晕和大脑肿瘤之间是否存在联系，上表中的4个数据都是必需的。我们可以比较头晕者中与不头晕者中患有大脑肿瘤的患者和没有患大脑肿瘤者的比例，前者为(160:40)，而后者为(40:10)。这样的比例也表明，不管头晕的症状是否出现，出现肿瘤的比例都是4:1。这一结果表明，神经科的病人患有大脑肿瘤的概率是非神经科病人的4倍，而与患者是否出现头晕的症状并没有关系。因此读者调查中第14题B的正确答案是变量之间并不存在联系。

绝大多数个体在协同测量中存在的困难引发了这样一个问题，健康专家是否也会错误地判断症状与疾病之间的关系。一位来自挪威的研究者简·斯梅兹伦德(Jan Smedslund, 1963)最早进行了相关的研究。在其中的一个实验中，斯梅兹伦德向护士们呈现了100张卡片，卡片上的内容是从病人的病例中摘取的。每一张卡片上包含两部分内容，病人是否出现某种特殊的症状(如果出现该症状则标记为+A，如果没有出现，则标记为-A)，以及病人是否患有某种疾病(如果患有该种疾病则标记为+F，如果没有则标记为-F)。在所有的卡片中，有37张卡片上的标记为+A和+F，33张卡片上的标记为-A和+F，17张卡片上的标记为+A和-F，而剩余13张卡片上的标记为-A和-F。他要求护士们根据这些卡片判断该症状与疾病之间是否存在联系，她们可以随意排列这些卡片，以便于进行正确的判断。根据斯梅兹伦德的实验结果，86%的护士错误地报告两者之间存在正性的联系，而只有7%的护士正确报告，两者之间并不存在正性的联系，还有7%的护士选择放弃，并没有回答这个问题。

当要求她们对自己给出的答案进行解释的时候，绝大多数认为两者之间存在正性联系的护士回答她们是观察了+A和+F出现的概率来进行判断的。根据这些结果，斯梅兹伦德(1963, p. 165)得出了这样的结论：“没有经过统计学训练的成人被试显然对相关的概念没有足够的认识……因此只要让他们进行统计方面的归因，他们总是倾向于仅仅利用++案例的发生频率来判断两者之间的关系。”尽管部分的研究结果也表明，个体在进行决策时会考虑更多的问题，而不仅仅是并存(如症状和疾病都存在的情境)的情境(Arkes & Harkness, 1983; Beyth - Marom, 1982; Shaklee & Mims, 1982; Shaklee & Tucker, 1980)。但是斯梅兹伦德的实验结果可以得出这样的结论：个体在判断两个事件的共存关系时存在很大的困难，因此他们往往依赖于两个事件同时发生的概率来做判断(Alloy & Tabachnik, 1984; Crocker, 1981)。

上帝能够满足祈祷者的请求吗

在理查德·尼斯比特和李罗斯(1980)对“协同测量”的讨论中，他们通过考虑上帝是否能够满足祈祷者的请求这样一个问题来说明个体对正性事件的

依赖。许多人都认为上帝能够满足祈祷者的请求，因为他们曾经向上帝请求，最后也得偿所愿了。这一判断的策略只是运用了一个 2×2 的列联表中的一项，这个列联表有两个维度，一个是“进行祈祷”和“不进行祈祷”，另一个维度是“得偿所愿”和“没有得偿所愿”。为了验证两者之间的联系，我们需要了解另外三个单元格中的信息：(1) 祈祷者的请求有多少没有得到应验；(2) 不管相信与否，没有乞求的事情有多少发生了；(3) 没有乞求的事情有多少没有发生（当然，这个是没有办法进行计算的）。

尽管这种方法表面上看起来是一种常识，但是个体往往很少关注那些没有发生的事情（Bourne & Guy, 1968; Nahinsky & Slaymaker, 1970）。正如尼斯比特和罗斯所指出的那样，这就是为什么福尔摩斯看起来在解决“银色马”（*The Silver Blaze*）疑案时是如此聪明的原因，他只不过考虑了“晚上发生在狗身上的怪事”。当巡视员报告“狗整晚都无所事事”的时候，福尔摩斯的回答是：“这是一件非常奇怪的事情”。尽管狗向入侵者咆哮是很正常的，但是没有咆哮并不是说明没有入侵者，也有可能入侵者是狗熟识的人，因此福尔摩斯缩小了怀疑对象的范围。

相关错觉

两个不相关的变量给人造成了它们相关的错误印象，这被认为是一种“相关错觉”（illusory correlation），首先对这一现象进行系统研究的研究者是洛伦·查普曼和琼·查普曼（Loren Chapman & Jean Chapman, 1967; 1969; 1971, 11月）。在第一个有关相关错觉的研究中，洛伦·查普曼（1967）在大屏幕上向被试呈现了很多不同的词对，如熏肉—老虎。屏幕左侧的单词总是熏肉、狮子、花或者船等等，而在屏幕的右侧总是出现蛋、老虎或者笔记本这样的词。尽管左边的词与右边每一个词配对出现的概率是一样，但是平均而言，被试报告当左边出现“熏肉”时，有47%的可能性右边出现“鸡蛋”，而当左边出现“狮子”时，右边出现“老虎”的可能性最大。人们认为这些在语义上存在关系的词之间存在着正性相关关系，但是实际上这样的相关关系并不存在。

在另一个实验中，查普曼（1967）对活跃在临床第一线的心理医生进行了调查。研究者写下了六类不同病人的简单描述，并要求医生判断这些病人在画人测验（一种非常流行的诊断测验）中的表现。总体而言，医生体现出了很强的一致性。例如，91%的医生都认为多疑的患者通常会画一个巨大的或者反常的眼睛，82%的人认为对自己的智力产生怀疑的患者会画一个巨大的或者反常的头部。但是事实上，并没有足够的证据表明画人测验中存在这样的模式。这种相关关系完全是一种错觉——是临床中刻板印象的结果。

查普曼（1967）同样向大学生呈现了相同的实验材料（这些学生从来都没有听说过画人测验），对病人的描述和患者所画的图形之间也是随机进行匹配

的。尽管实验材料之间进行了随机的匹配，但学生们却表现出了与心理医生同样的相关错觉（例如，绝大多数的学生都认为多疑的患者通常会画一个巨大的或者反常的眼睛）。在另一个实验中，查普曼（1967）允许被试在连续三天的时间内阅读相同的实验材料，或者给被试提供 20 美元以提高他们的正确率，希望考察他们在这样的情况下，是否还会存在相关错觉。实验结果是，在上述两种情况下，相关错觉依然非常强烈。另一些研究者证实，即使经过了很好的培训，相关错觉依然很难消除（Golding & Rorer, 1972）。

最后，在一个可能是查普曼所进行的系列研究中最著名的实验中（1969），他们要求 32 位心理医生使用罗夏墨迹测验来诊断男性同性恋者（在 20 世纪 60 年代，许多心理学家认为同性恋是心理疾病的一种——这种观点现在已经不再得到大家的认同）。这些心理医生观察了大量男性同性恋者的罗夏墨迹测验结果以后，认为相对于其他男性而言，男性同性恋者更加倾向于将墨迹解释为：（1）臀部或者是肛门，（2）生殖器，（3）女性打扮，（4）性别模糊的体态，以及（5）同时具有男性和女性特征的体态。事实上，这些反应与同性恋之间并没有必然的联系。另外，在 32 位心理医生中只有两位提到有两种墨迹的解释与男性同性恋之间存在共变的关系（将墨迹描述成怪兽的图形或者是半人半兽的图形）。根据这些结果，查普曼（1969）认为绝大多数的心理医生是基于对同性恋的刻板印象而形成了这样一种相关错觉。你可以查看一下你对读者调查中第 18 题的回答，以确认你是否也形成了类似的相关错觉。

尽管形成相关错觉的原因还不是十分明了，但是绝大多数的理论家将原因归结为易得性直觉和代表性直觉。易得性直觉的解释是这样的：相关错觉的产生是因为明显的或者突出的配对（例如熏肉—鸡蛋）更容易进入我们的记忆，因此会高估它们出现的频率（McAuliffe, 1980）。代表性直觉理论的解释是认为主要是因为一个事件可以成为另一个事件的典型代表（例如，某种墨迹解释可以被认为代表了男性同性恋）。这两个解释都得到了大量研究结果的支持（Mullen & Johnson, 1990），可以解释大量的相关错觉。

无法识别的相关

与观察到根本不存在的相关相反，我们可能会对一些现实存在的相关视而不见——这一现象可以被称为“无法识别的相关”（invisible correlation）。相关错觉产生的原因通常是由于期望两个变量之间存在关系（Hamilton & Rose, 1980），而无法识别的相关主要是由于个体缺少这样的期望。由于缺少对两个变量关系的期望，那么即使两者之间存在很强的联系也可能被忽略。

与相关错觉相比，无法识别的相关则显得更加普遍，而且可能导致复杂的结果。例如，直到最近，吸烟与肺癌之间的相关关系才得到确认，对大多数的美国人而言，肉类的消费量与结肠癌之间的相关关系至今还没有被人们所认

识。

关于人们在识别正性相关时存在困难的一个最为综合性的研究来自丹尼斯·詹宁斯等人 (Dennis Jennings, Teresa Amabile, & Lee Ross, 1982)。詹宁斯等人给大学生三组不同的数据系列, 第一组数据包含 10 对数据, 第二组数据包含 10 个不同人的身高和拐杖的长度数据, 而第三组数据包含 10 个人读一个字母表中的一个字母和唱一个音符 (在最后一个序列中, 这两个变量的价值在于字母在字母表中的位置以及音符的长度)。在每一组数据中, 成对数据的相关系数在 0 (完全没有线性相关) 和 1 (完全的正性相关) 之间。要求学生判断每组成对数据的相关系数是正性还是负性的, 它们相关的程度如何 (使用 0 到 100 的量表, 而不是 0 到 1)。

实验的结果显示, 直到数据的相关值达到 0.60 或者 0.70 (相对较强的相关) 的时候, 学生们才认为两组数据之间存在正相关。当变量之间的相关在 0.20 到 0.40 之间的时候 (研究表明人格特质与个体在具体情境中的行为之间的相关值也大致在这样的范围), 学生在 100 点量表上的评分只是在 4 到 8 之间——几乎等于零相关。只有当两个变量之间的相关达到 0.80 以上的时候, 学生在 100 点量表上的评分平均为 50, 而对于 1.00 的相关, 学生的平均评分也没有超过 85。换一句话讲, 对于许多学生而言, 他们没有办法识别中等强度的相关关系; 同时倾向于将很强的相关关系视为中等强度的相关。

正确判断

为什么人们在判断两个变量之间的相关关系的时候会存在如此多的困难呢? 让我们以协同测量中的一个典型的关系为例, 考察其中的复杂性。该问题就是创造性与智力之间的相关关系。珍妮弗·克罗克 (Jennifer Crocker, 1981) 在她关于协同测量的研究综述中, 总结了判断相关关系的六个步骤:

- 第一, 你需要了解哪些信息是相关的。例如, 你是否仅仅需要考察创造力与智力同时具备的情况, 或对于缺乏某一方面的例子是否也需要考虑在内? 这样的测量是否仅仅局限于成年人? 美国人? 还是著名的艺术家?
- 接着, 你需要在你考虑进行考察的人群中抽取一部分的样本 (如果可能的话进行随机抽取)。这是一个观测, 或者称为“数据收集”的阶段。
- 第三, 你必须对你观测到的现象进行分类和解释。哪些可以归结为创造力? 哪些可以归结为智力? 测量的结果是一个连续体还是可以分成不同的类别?
- 第四, 你要牢记你的分类标准, 并依此来估计符合事件和不符合事件的发生频率。
- 第五, 这样的估计必须用一种有意义的方式来进行整合。我们可以使用统

计公式来对频率估计进行整合，但是斯梅兹伦德（1963）的研究发现，人们通常还是倾向于使用非正式的策略。

- 最后，你需要根据整合的结果来判断是否存在相关的关系。

没有办法排除在以上的六个阶段都可能出现的偏差，因此人们在判断相关关系时存在各种各样的困难也就不足为奇了。例如，决策者在进行观测时很少使用随机样本，而是进行高度选择性的观测。正如我们在第3章中所讨论的，决策者的记忆也存在着各种各样的偏差，许多这样的偏差可以使相关关系的判断变得更加复杂。事实上，正是因为这些偏差的存在，个体现在的许多行为还是值得称道的。毕竟人们还是能够很好地观测到他们生活中许多重要的相关关系——如学习与考试成绩之间的相关，过度进食与体重增加之间的相关，种族主义言论与社会排斥之间的相关关系等等。

为什么我们在日常生活中可以成功地判断许多相关关系，而在实验室实验中的表现又如此糟糕呢？尼斯比特和罗斯（1980）给出了以下一些解释。首先，日常生活中的许多情境可以帮助我们进行相关关系的判断。例如，（1）许多是完全相关或者接近完全相关，例如将手掌置于火焰上，就会有发烫的感觉，又如摁下一个开关，灯就会亮等等；（2）有些相关关系是基于非常明显的、有区别的、动机性很强的刺激，例如暴风雨与下雨之间的关系；（3）刺激在时间或者空间上的距离非常近，例如轻抚你的小狗，小狗就会摇尾巴；（4）有些相关也很容易被探测到，如吃某些食物与消化不良之间的关系等。

第二，人们通常会基于某些事件来使用非统计性的接近原则，从而判断其中的相关关系。例如，他们会选择A事件的最具有代表性或者最极端的例子，来判断B事件是否发生。举例来说，在判断你朋友推荐的电影与你对该电影的欣赏程度之间的相关关系时，你可能就会想起你朋友以前推荐的最棒的一部电影。另外如果你希望考察不同城市的一般生活费用，你可能就会考察在每一个城市租一套两居室房子的租金。

第三点也是更为重要的一点是，人们不用重新去发现每一个相关关系。绝大多数的相关关系——无论是别人发现的，还是自己亲身经历的——都可能成为被认识到的相关关系中的一部分。例如，人们不必亲身去了解性交与怀孕之间的相关关系，很久以前人们都已经认识到了这种相关关系，同时也成为了被充分认识的相关关系的一种。

设想一下一个个体试图证实性交与怀孕之间的相关关系将会面临的困难。由于浪漫关系的许多方面都是与性行为有联系的，这些行为都可能成为怀孕的原因（掩盖了其性交之间的联系）。例如，亲吻、拉手、睡在一起或者感觉愉快可能被误认为怀孕的原因。许多身体变化也可能成为怀孕的原因，例如恶心、体重增加或者情绪不稳定等等。事实上，由于浪漫关系也可能在几个月后中止，因而停止性交也可能成为怀孕的原因之一。试想一下这样的信念可能引

起多大的灾难！*

因果关系

两个变量之间存在相关关系并不意味着其中的一个事件是另一个事件产生的原因——如果是这样，亲吻就可能导致人口的过度增长。尽管如此，人们还是常常将相关关系和因果关系等同起来。例如，图 15.1 中的广告就在暗示购买一个打字机就有可能提高一个人的学业成绩，但是对于这样一种相关关系（如果确实存在这种相关关系的话）还存在着很多不同的解释。购买了打字机而获得了较高的学业成绩的学生来自比较富有的家庭，而这样的家庭对孩子的写作更为关注；或者是购买打字机的学生比其他学生有更高的学习动机。可能有无数的原因可以解释购买打字机的行为与学生学业成绩之间的相关关系。

同样的武断推论来自最近广告业自身的一项推广活动。在《纽约时报》（1990，8月26日）的整版报道中，美国广告协会认为“做广告可以使你的投资回报率加倍”。这一论断的证据来自下面的几个相关关系：“比竞争对手投入更多广告的品牌的投资回报率是 32%，而比竞争对手投入较少广告的品牌的投资回报率仅为 7%。”即使这样的证据是真实的，也不能证明广告的投入与投资回报率之间的因果关系。因为可能是由于只有利润率较高的企业才能够负担高额广告费用，如果是这样的话，正是财务上的成功才导致了广告的投入，而不是相反的关系。

正如相关关系并不能证明其中的因果关系一样，因果关系也不能保证一定存在很强的相关关系。希勒尔·艾因霍恩和罗宾·霍格思（Hillel Einhorn & Robin Hogarth, 1986）将认为因果关系一定导致相关关系这种想法称为“Causation”。在他们针对这一现象的讨论中，他们让读者考虑下面的虚拟分类：

		怀 孕		总 计
		是	否	
性 交	是	20	80	100
	否	5	95	100
	总 计	25	175	200

非常奇怪的数据是竟然有 5 对怀孕的夫妇报告没有进行性交，可能的原因是他们的记忆出现了偏差，或者他们撒谎了等等。

* 这样的一种困惑并不是像看起来那样牵强。在早些时候，生了双胞胎的妇女通常被认为具有通奸的行为；而生下严重畸形婴儿的妇女则被认为进行了人兽性交（Ives, 1970）。

In an independent survey of high school and college instructors, fifty percent agreed:

STUDENTS WHO TYPE USUALLY RECEIVE BETTER GRADES.

For years, we at Smith-Corona have believed that typing can help students do better. But even we were impressed when we learned that half of the high school and college instructors agreed: typing does help students improve their spelling; typing does help students improve their punctuation; typing does help students organize their thought and that students who type usually receive better grades.

(Though the student's ability and the subject being taught are factors, fifty percent of the instructors agreed ... students who type usually receive better grades.)

If there is a student in your family, wouldn't this back-to-school time be the perfect time for you to purchase a typewriter. The one we suggest is the electric portable that is being used more than any other. The Smith-Corona® Cartridge Ribbon Typewriter.

This portable allows students to change ribbons by simply snapping in a cartridge. And in addition to our black fabric ribbon in a cartridge there are four different color cartridges students love to use for emphasizing thoughts. A correction cartridge that allows students to correct errors in less time than it takes you to read this sentence. And a carbon film cartridge that



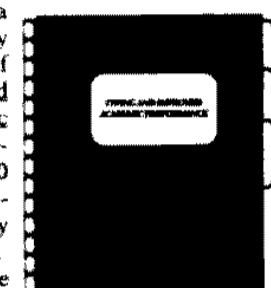
makes type from our portable look like type from more expensive executive typewriters. The carbon film cartridge is ideal for term papers and reports.


And we think you should also know that the Smith-Corona Cartridge Ribbon Typewriter is engineered to the same standards of durability, performance, and design that have made our typewriters Number One on campus ... before our cartridges ever came along.

SEND FOR A FREE SYNOPSIS OF THIS SURVEY.

Smith-Corona has prepared an easy to read synopsis of Typing and Improved Academic Performance, an independent survey of 400 high school and college instructors by Crossley Surveys, Inc.

For your free copy, see your Smith-Corona dealer or send a self-addressed envelope to Smith-Corona, P.O. Box 81, G.P.O., New York, N.Y. 10001.



SMITH-CORONA

SCM CORPORATION




图 15.1 相关关系 ≠ 因果关系 (Courtesy Smith - Corona)

就上表的数据而言，性交与怀孕之间只有中等程度的相关（只有 0.34）。如果基于这样的相关系数，一个不是很了解相关概念的个体很可能推测性交并不是怀孕的真正原因。相反，他可能将怀孕归结为其他相关系数更高的因素，这里存在的逻辑偏差就在于性交能够导致怀孕并不意味着性交要与怀孕有很高的相关。因果关系并不意味着两个变量之间一定存在着很高的相关。

正面，我赢了；反面，那是随机事件

人们进行的最基本的协同测量——对于人类的生存以及自身而言都是最关键和最重要的——是人们对于其行为与行为结果的判断。无论是以这样或者那样的形式，行为与后果之间的关系在心理学的理论中都占据了中心地位，无论是因果归因理论（Kelley, 1967, 1971），还是对自我效能的认知（Bandura, 1982, 1986），内外控理论（Lefcourt, 1982）或者其他的心理结构，关注的重

点都是这样一种协同变化。从这些文献中可以发现一个最有趣的现象，那就是人们通常相信他们对于概率事件能够拥有比实际情况更多的控制。这种信念通常被称为“控制错觉”（illusion of control），首先是由心理学家埃伦·兰格（Ellen Langer）提出来的。

在她最早进行的一个研究中，实验的名字就是“正面，我赢了；反面，那是随机事件”。兰格发现，在某些情境下，人们相信他们能够预测和控制硬币投掷的结果（Langer & Roth, 1975）。在预实验的“社会引导”下，被试被要求猜测实验者投掷30枚硬币的结果。事实上，被试预测的准确率是被事先设定的，每一个人都只能猜中其中的15次，但是其中大约1/3的被试开始的时候猜测的结果比较理想（头4次的猜测都是正确的），1/3的被试开始的时候猜测的结果不是很理想（头4次猜测的结果都是错误的），而另外的1/3被试的猜测结果就是随机水平。在猜测完30次结果以后，要求被试回答一系列有关他们猜测成绩的问题。

实验结果显示，刚开始猜测成绩较好的被试对他们猜测成绩的评估结果远远好于其他的被试。这类似于第4章中我们所讲到的“首因效应”。事实上，这些被试确实能够回忆起更多成功的预测，如果让他们再预测100次投掷，他们的预测也比其他人更加正确。兰格和她的合作者简·罗思（Jane Roth）也指出，大约40%的被试认为练习能够提高他们预测的成绩，25%的被试认为分心会影响他们预测的成绩——即使任务是预测一个偶然事件。这样的研究结果与赌博者投掷骰子的情形是类似的，他们如果想得到较小的数字，他们就会很轻柔地投掷；而如果想得到较大的数字，他们就会很重地投掷。在他们掷骰子的时候都是聚精会神的（Henslin, 1967）。兰格和罗思（1975, p954）得出了这样的结论：“起先，一个相对稳定的成功模式可以是个体产生对自己技能的归因，进一步导致个体确信他在将来也能够取得成功。”

兰格（1975）进行的另一个研究中，办公室的工作人员参加了一项博彩游戏，有些彩票上有他们比较熟悉的标志，而另一些则是他们不熟悉的标志。他们可以选择自己来挑选彩票还是由实验者进行分配。一旦被试得到了彩票，他们就有机会选择继续保留彩票，还是换成另外一种获胜机会更大的彩票。兰格的实验结果发现，自己选择彩票或者获得了熟悉标志的被试，相对于其他个体而言更加愿意保留原来的彩票。基于这一实验结果，兰格认为熟悉性和选择使被试产生了一种控制的错觉。卡米尔·沃特曼（Camille Wortman, 1975）也发现了类似的结果。

无法奏效的帮助

对控制的知觉尽管很大程度上是一种错觉，但它却在个体的健康问题上扮演了一个很重要的角色。例如，兰格和朱迪思·罗丁（Judith Rodin）（1976）

进行的一项现场研究表明，对于老年人的选择权和所担负责任的很小改变，就能够对他们的健康造成很大的影响。研究的被试是 91 位无家可归的老人，年龄从 65 岁到 90 岁不等，他们都居住在康涅狄格的一所养老院中。对于养老院的一层中居住的老人，管理者告诉他们可以自己控制自己的生活：如何分配他们的时间，如何摆放自己房间的家具，或者是改变养老院中的任何东西。同时，管理者给老人们一株植物，要求他们进行看护。而另一层的老人，管理人员通知他们养老院的工作人员将全权负责他们的生活，并尽可能地照顾他们的饮食起居，管理者同样给他们一株植物，但是由护士们负责进行看护。

三个星期以后，由工作人员全权负责的老人们中有 71% 的人认为他们疲惫不堪。相反，在允许他们自由支配自己生活的老人中，有 93% 的人认为自己的健康状况有所提高。基于他们自己的判断和护士的判断（他们并不知道整个实验的目的），这些老人更加高兴和活跃。他们比其他老人的思维更加活跃，花费更多的时间与其他老人或者外来的客人进行交流。兰格和罗丁（1976, p. 197）认为：“如此微弱的控制都能够产生效果说明，控制感对于这些老人、对于那些无法进行决策的人的重要性。”

这个实验存在的一个问题（和其他许多现场研究类似）就是无法了解实验的变化是否是产生实验结果的真正原因。是管理者关于控制的言论产生了作用？还是需要照看的植物产生了作用？或者就是住在这两层的老人本身就具有不同的特点？这里并没有确切的答案。但是既然这个实验已经完成了，其他还有许多实验同样也支持了控制感与健康之间的关系，特别是对于老年人而言（Rodin, 1986）。这些实验强调了相关判断在人们日常生活中的中心地位。

结 论

在人们的决策中最为常见的判断就是两个变量是否存在相关关系。在许多情况下，这些判断都是基于先前已经建立的关系；在另一些情况下，先前的联系并不足以使决策者就对全新的、矛盾的信息进行权衡，从而形成一个全新的判断。正如本章中的研究所显示的那样，这很容易使决策者深陷泥潭。

一些关于协同测量的研究为决策者提供了一些建议。首先，决策者不应该仅仅局限于正性的、已被确认的关系。在判断相关关系时，没有出现的事实往往是最为关键的。第二，在判断一个关系是否存在之前，决策者应该问自己，他们的判断是基于自己的观察还是期望。如果是前者，相关关系就有可能被低估；而如果是后者，相关关系就有可能被高估。这种差异与关于相关错觉和无法识别的相关的研究结果是一致的（Jennings, Amabile, & Ross, 1982）。最后，决策者必须很好地识别相关关系和因果关系，两个存在因果关系的事件之间可能只存在很弱的相关。第 16 章将非常详细地论述因果关系判断中的最重要的一环——对行为原因的判断。

归因理论

你遇到你的一个朋友，她一味称赞你的外貌。你去赴一个约会，但是结果很糟糕。你做了一个非常重要的报告，但是听众的反应比较负面。如果这些事情都发生在你身上，你将如何解释呢？

归因理论是关于人们如何进行“因果归因”的心理学理论，也就是对行为和行为的结果产生的原因进行解释。对这一理论的系统阐述来自 UCLA（加州大学洛杉矶分校）的心理学家哈罗德·凯利（Harold Kelley, 1967）的一篇论文。该论文的主要思想来自于弗里茨·海德（Fritz Heider, 1958）以及内德·琼斯和基思·戴维斯（Ned Jones & Keith Davis, 1965）的经典研究工作。正如期望效用理论那样，归因理论被认为是决策中的规范性理论（关于个体如何行为的理想化理论）；但是与期望效用理论不同的是，归因理论同样针对人们的日常行为提出了一个描述性模型。

在最原始的公式中，凯利将一个归因模型称为“变异框架分析”（根据统计学中的方差分析技术命名，或者称为 ANOVA）。根据这个模型，人们通常使用三种不同的方式来解释行为的原因：（1）行为人——情境中的个体可能是行为产生的原因；（2）环境——情境中的某些固有特征可能是导致行为的原因；（3）时间——特定时刻的某些因素引起了行为。凯利认为这三种归因方式主要依据三种不同的信息来源：

1. **共同反应**：在相同的情境中，其他人是否采取同样的行为反应？
2. **独特性**：其他的情境或者其他的刺激是否产生相同的行为？
3. **一致性**：同样的事情是不是每时每刻都会发生？

表 16.1 中包含了凯利归因理论所进行的某些预测。在这个表中，“高度的共同反应”表示绝大多数个体在相同的情境中的反应是类似的，“高度的独特性”表示只有独特的刺激或者情境才会引起该行为，而“高度的一致性”表示只要是相同的情境，那么每时每刻都会发生相同的事情。为了说明的需要，假设你是惟一的受测者（很低的共同反应），在许多测验上的得分都非常优秀（很低的独特性），在许多情况下都是如此（很高的一致性）。凯利会做出“个人归因”——用个人的能力来解释个人的行为（读者调查中第 24 题中 C 选项的解释）。事实上，这种对归因进行分类的方式是基于这样的假设：做出个人

归因的最准确的手段是对不同时间、不同地点的行为都进行测量。使用 ANOVA 这样一个模型就可以通过不同时间和情境中的测量来确定归因的方式，但这只是针对单个个体（见图 16.1）。

另一方面，如果针对不同时间和个体的测量只是在一个情境中进行，凯利则会进行“情境归因”。例如，一个测验（高度的独特性）每年都在进行，学生们总能取得很好的成绩（很高的共同反应倾向和很高的一致性），人们就会将得高分的原因解释为测验本身（情境），而不是参加测验的个体（见表 16.1 和图 16.2）。

根据凯利的理论，因果归因的基础是对共变关系（covariation）的判断（我们在第 15 章中已经进行了详细的讨论）。这一原则在凯利（1973, p. 108）的“共变原则”（covariation principle）中得到了充分的体现，这个原则认为“某效应可以归因为与其共同变化的一个因素上”。

凯利意识到这一原则过于简化，但是它为我们提供了一个判断人们进行因果归因的基本准则。由于时间归因理论是最早提出来的，因此，它引发了成百

表 16.1 归因理论进行的预测

归因预测	信息类型		
	共同反应	独特性	一致性
个人	低	低	高
刺激（情境）	高	高	高
环境（时间）	低	高	低

注：This table is adapted from an article by Bruce Orvis, John Cunningham, and Harold Kelley (1975).

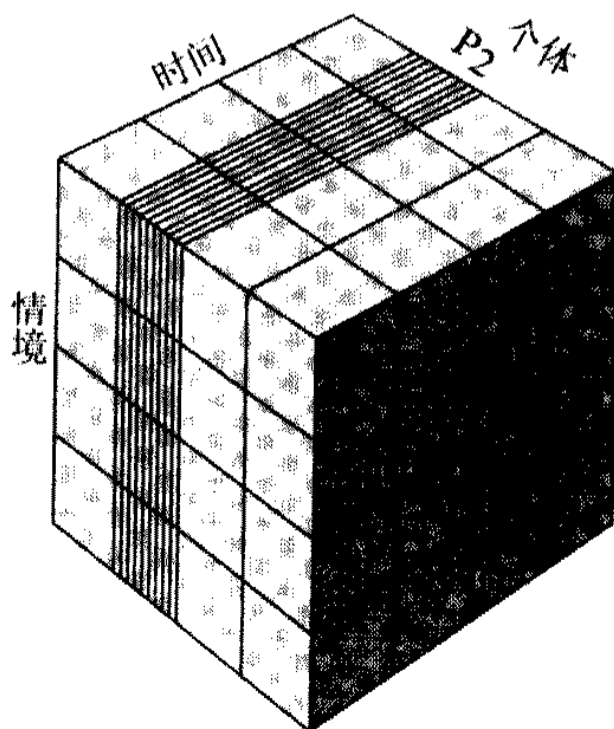


图 16.1 根据哈罗德·凯利的 ANOVA 模型，这样一种信息的模式会导致个人归因（基于个人因素进行解释）。（Reprinted from Kelley, 1973）

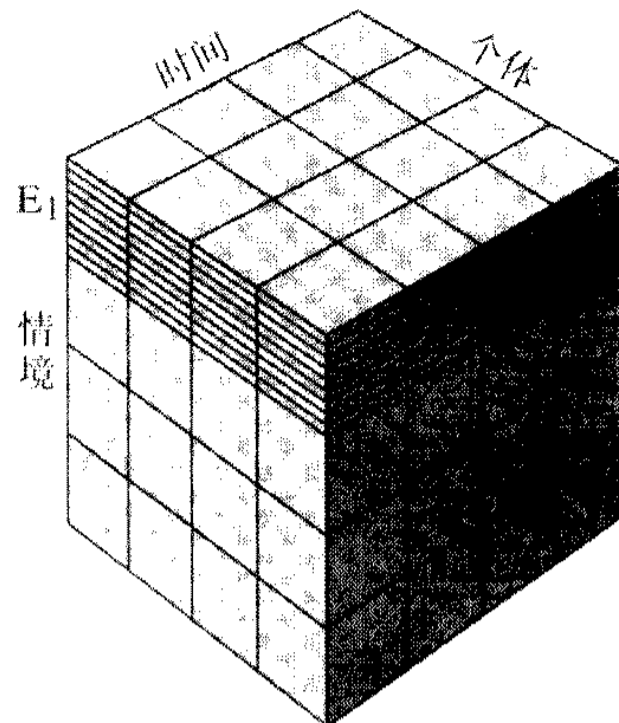


图 16.2 根据哈罗德·凯利的 ANOVA 模型，这种信息模式会导致情境归因（基于情境，或与刺激相关的因素进行解释）。(Reprinted from Kelley, 1973)

上千的研究来证实凯利的理论是否正确。事实上，在归因理论（Kelley & Michela, 1980）提出后的十年之内，大概有超过 900 个研究是针对归因进行的。这些研究对原始的理论进行了修正和扩展（Cheng & Novick, 1990; Försterling, 1989; Hewstone & Jaspars, 1987; Hilton & Slugoski, 1986; Orvis, Cunningham, & Kelley, 1975）。

归因理论能否解释个体和他人的行为？对于绝大多数行为而言，该理论是正确的；但是也存在着几个重要的例外。

缺乏共同反应

在某些情况下，个体不考虑共同反应（consensus）的信息，因此人们的判断会偏离归因理论的预测（Nisbett, Borgida, Crandall, & Reed, 1976）。对共同反应信息缺乏利用是忽视行为的基线水平的典型情况，因为在给定情境的前提下，共同反应就等同于行为的基线水平。例如，在著名的有关服从权威的研究中，斯坦利·米尔格拉姆（Stanley Milgram, 1963）发现，有 65% 的“达成一致”的被试都会完全服从实验者。但是，冈特·比尔布劳尔（Günter Bierbrauer, 1973）在观察者面前重复了米尔格拉姆的实验，基线水平同样被忽视了。观察者将被试的服从行为主要归结为个性因素，而不是实验中所有被试都面临的共同情境因素。

理查德·尼斯比特和尤金·博吉达（Richard Nisbett & Eugene Borgida, 1975）也认为个体在进行因果归因时会忽视这样的基线水平。尼斯比特和博吉达注意到，有大量的研究证明，一致性的信息和独特性的信息可以影响个体的因果归因，但是很少有研究证明共同反应信息在其中也扮演了重要的角色。在

验证凯利的 ANOVA 模型的最早实验中，莱斯利·麦克阿瑟 (Leslie McArthur, 1972) 的研究支持了这样的说法，他的研究发现，一致性信息和独特性信息对因果归因造成的影响是共同反应信息的几倍。

对这一观点的支持证据还来自尼斯比特和博吉达 (1975) 进行的两个实验。在这些实验中，大学生被试们得到了两个最近发表的研究信息——其中一个研究是由尼斯比特和斯坦利·沙克特 (Stanley Schachter, 1966) 完成的，而另一个研究是由达利和拉坦 (1968) 完成的。一部分学生得到了被试共同反应的信息，描述他们在实验中的表现如何；而另一部分学生则没有得到这样的信息。

之所以选取这两个实验是因为这两个实验的结果都包含了令人惊讶的共同反应信息。在尼斯比特和沙克特 (1966) 的研究中，34 名被试中有 32 名愿意接受电击，因为实验者告诉他们这只是一个测量“皮肤感应”的实验 (其中有将近一半的被试承受的电击太强烈，手臂都出现了痉挛)。更加令人吃惊的是，在达利和拉坦 (1968) 的研究中，15 名被试中有 11 名被试没有帮助一名非常明显的疾病发作的病人，直到这位病人出现明显的窒息症状 (其中有 6 名被试自始至终都没有提供帮助)。根据 ANOVA 模型，这样的共同反应的信息应该引导个体进行情境的归因，因为遇到相同的情境，绝大多数的被试的行为反应都是类似的。

尼斯比特和博吉达 (1975) 向被试询问了一系列的问题，其中有两个问题是最值得关注的。第一个问题是要求被试评价一个具体被试的行为——比尔，在电击实验中，他愿意承受的电击强度最大；或者是格雷格，在拯救病人的实验中，他完全没有提供帮助——在多大程度上是由于被试的个性因素决定的 (情境特征的对立面)。第二个问题是，假设学生参加相同的实验，他们会如何表现？

尼斯比特和博吉达的研究发现，共同反应的信息并不能“真正产生作用”。有些学生被告知实验中绝大多数被试的行为反应是类似的，但这样的信息并不能显著地使他们对比尔和格雷格的行为进行情境归因。在学生判断自己在实验中的表现时，共同反应的信息也没能产生作用。他人的行为与个体自己的行为无关。

尽管有一些研究者对尼斯比特和博吉达的研究结论提出了反驳的意见 (如, Ruble & Feldman, 1976; Wells & Harvey, 1977)，但是他们的研究结果还是得到了其他研究者的认同。当某些行为的倾向以一种统计的基线值的方式呈现给个体，人们往往会忽视或者低估共同反应信息的价值 (Fiske & Taylor, 1991; Kassin, 1979; Nisbett, Borgida, Crandall, & Reed, 1976)。尼斯比特和博吉达 (1975) 认为出现这种现象的主要原因是基线信息有时是比较抽象、模糊、令人难以捉摸的。根据这样的解释，相对于共同反应的信息而言，一致性信息和独特性信息看起来更加容易进行因果归因，因为这类信息通常是比较具体、生动和

突出的。正如下一节中所说的，突出（salient）因素——也就是能够引起个体注意的因素——相对于不突出的因素而言，更加容易引起因果归因。

突出因素

在许多方面，突出信息与信息的可得性与生动性相类似。相对而言，突出的、易得的、生动的信息能够产生更大的影响力。用比较通俗的话来讲：（1）如果一个事件经常发生，那么看到它的频率和可能性也就越大；（2）信息越是生动，就越容易被回忆起来；（3）事物越是突出，看起来就越像因果关系。人们对因果关系的知觉部分是由个体在环境中的注意指向决定的，而突出性可以影响个体的注意。

有关信息的突出性与因果归因之间联系的研究首先是由海德（Heider, 1958）提出来的，随后由谢利·泰勒和苏珊·菲斯克（Shelley Taylor & Susan Fiske, 1978）进行了总结。泰勒和菲斯克讨论了大量的证据，其中部分的证据来自他们自己完成的几个研究。例如，在泰勒和菲斯克（1975）完成的一个实验中，有6个观察者从三个不同的角度来观察两个个体的对话：其中两个观察者坐在一个发言者的后面，另两个观察者坐在另一个发言者的后面，而剩余的两个观察者坐在两个发言者的中间（见图16.3）。所有的观察者同时聆听发言者的发言，因此观察者之间的惟一差异就是他们观察发言者的角度（发言者的突出性）。然而，泰勒和菲斯克的研究结果发现，观察者往往倾向于认为在他们视野中的发言者掌握着整个对话的基调，确定了信息转换的类型以及使对方像自己一样进行反应（中间的观察者认为两个发言者具有相同的影响力）。

在另一个研究中，向被试呈现的是一张相关幻灯片以及六个男性就一个政治活动展开头脑风暴的录音（Taylor, Fiske, Close, Anderson, & Ruderman, 1977）。在其中一个实验条件下，录音中每一个人的声音都与一位白人男性的照片相匹配；在第二种实验条件下，三个成员的声音与白人的照片相匹配，而另三个成员的声音与黑人的照片相匹配；在第三种实验条件下，五个成员的声音与白人的照片相匹配，而另一个成员的声音与黑人的照片相匹配（事实上，在这三种实验条件下，使用的录音都是相同的）。研究的假设是突出的个体将被认为是更加具有影响力的个体。泰勒和他的同事发现，相对于另一种实验条件下的同一个黑人个体（显得不是那么突出），在第三种实验条件下，惟一的一位黑人发言者被认为说话更多，更具有影响力，同时给人留下了更加深刻的印象。也就是说，虽然是相同的个体在进行相同的行为，但是当小组中只存在这位黑人男性时，他就被认为更加健谈和更具有影响力。

在其他研究中，研究者使用了另外一些方法来控制被观测者的突出性——让一部分的被试穿着显眼的衬衫，而另一部分被试穿着灰色的衬衫；或者一部分人坐在摇椅上不停地摇晃，而另一部分人则静静地呆着；或者让一部分人坐

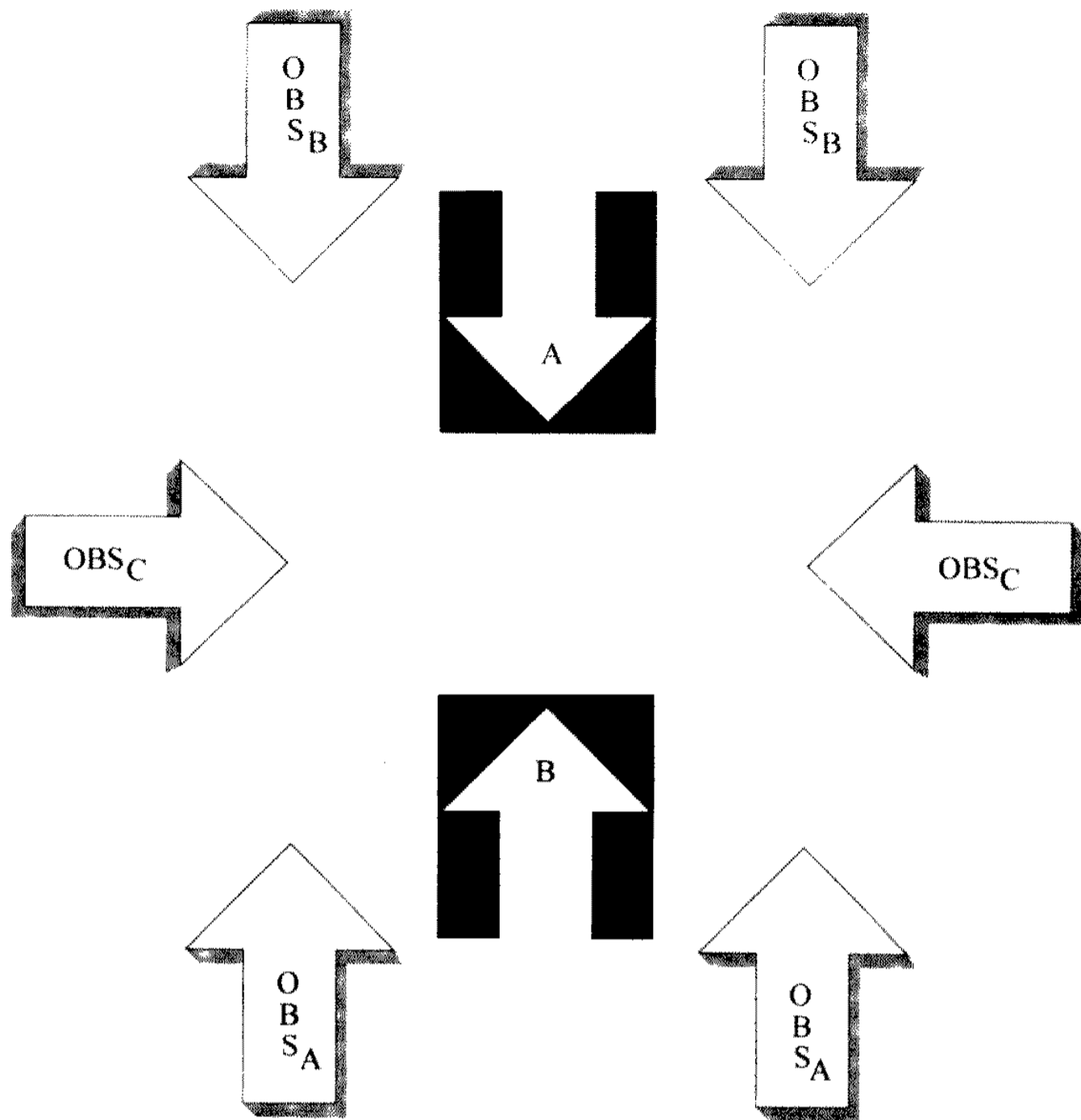


图 16.3 上图体现了泰勒与菲斯克（1975）进行的实验中对被试的座位安排情况。两个观察者面对发言者 A，两个观察者面对发言者 B，处于控制组的两个发言者处于发言者 A 和发言者 B 之间。泰勒与菲斯克的研究结果发现，观察者认为他们面对的发言者更加具有影响力。

在明亮的灯光下，而另一部分人坐在昏暗的灯光下——实验的结果都得到了类似的答案（McAuliffe & Post, 1977）。突出的个体相对而言更容易被进行因果归因，事实上，有一个研究甚至发现，当人们坐在一面大镜子前时，他们会倾向于认为自己更加具有影响力（Duval & Wicklund, 1973）。

基本归因误差

在海德（1958）具有里程碑意义的著作——《人际关系心理学》（*The Psychology of Interpersonal Relations*）一书中，他认为“行为决定一切”。这句话的意思就是在一个社会情境中，最突出的是人的行为（同样也就是暗指“行为”）。情境因素，例如房间的温度、具体的日期都是第二位的。个体处于大前方，而其他的事情都处于大后方。

在 1977 年，李·罗斯认为这种对于个体和个体行为的过分关注可以导致

观察者过度将行为归因为个性因素——如个体的能力、特质和动机——从而忽视了环境因素的影响。事实上，罗斯认为这样一种将行为过多地归因为个性因素的现象是如此的普遍，因此可以称做“基本归因误差”。之所以将这种现象称为“误差”，而不是简单地认为是一个归因偏差或者不同的观点，是因为这种个性的归因通常很明显是不正确的（Reeder, 1982）。例如，阿瑟·米勒和他的同事（1973）发现人们认为在米尔格拉姆（1963）研究中的被试相对而言是冷酷的、不适应环境的和具有攻击性的，尽管很明显，事实并非如此。

过度归因的另一个证据来自第一个发表的有关态度归因的研究。在这个由内德·琼斯和维克托·哈里斯（Ned Jones & Victor Harris, 1967）完成的实验中，研究者呈现给被试一篇文章，部分被试所读到的文章的作者明显是被迫持有某种观点；而另一篇文章的作者所持有的观点则是“自己选择”的。即使实验者明确地告诉被试作者是被迫采纳了这样的观点，被试还是倾向于将这样的结果归因为作者本身。这个研究发现非常明确，在其他许多实验中也得到了同样的结果（Jones, 1979; Snyder & Jones, 1974）。

基本归因误差甚至是持久稳固的，这一点在Pietromonaco和尼斯比特（1982）的研究“在基本归因误差中逆流而上”中得到了验证。在这个研究中，Pietromonaco和尼斯比特向大约一半的被试呈现了达利和巴特森（1973）所做研究的一个简短版本，研究指出，神学院的学生没有帮助一个需要帮助的人，因为他们要赶去作一个有关“行善”的演讲，而当时他们已经迟到了。在该研究中，帮助的行为取决于“着急”这样的情境因素，而不是宗教观念这样的个性因素（在没有时间压力的情况下，63%的学生提供了帮助，而在他们非常着急的情况下，只有10%的学生提供了帮助）。Pietromonaco和尼斯比特实验中剩下的一半被试并没有被告知达利和巴特森实验的具体结果。所有的被试都将面临两个与达利和巴特森研究类似的实验情境，让他们估计在每一个情境中，愿意停下来帮助需要帮助的人数比例。

Pietromonaco和尼斯比特（1982）发现，并不知道达利和巴特森实验结果的被试会犯基本归因误差。他们会根据宗教标准（事实上与帮助行为毫无关系的个性因素）来预测助人行为。如果实验的结果仅仅如此，这并不奇怪，因为实验被试并不知道达利和巴特森实验的具体结果，绝大多数的个体都会将助人行为归因为个性因素。但令人吃惊的结果是，对于那些已经了解达利和巴特森实验结果的被试，他们同样根据宗教标准来预测助人行为。事实上，与那些不知道达利和巴特森实验结果的被试相比，这些被试并没有显著地减少个性归因。

这样的结果又一次表明基线信息通常是被忽视的，而基本归因误差也是不容易改变的。即使个体得知在某一个情境中存在外在的压力，他们也会对相关情境中的行为进行个性因素的归因。也有一些研究者认为基本归因误差并不是那么“基本”（Harvey, Town, & Yarkin, 1981; Quattrone, 1982; Tetlock,

1985a), 但是毫无疑问的是, 观察者经常将行为过度归因为个性因素。

我的情境是你的个性

如果个体的行为主导了观察者, 那么什么东西可以主导行为人的思维呢? 这个问题的答案在于行为人关注自己所在的情境。只有在很少的情况下, 个体会认为自己的行为与他人的行为是类似的。

正是因为这种关注点的差异, 个体倾向于将自己的行为归因为情境因素, 而不是像观察者那样归因为个性因素。内德·琼斯和尼斯比特(1971, p. 2) 首先指出行为人和观察者之间的差异: “个体看待自己行为时主要关注行为当时所处的环境。而观察者则会关注行为人稳定的个性因素。因此我们认为在大多数的情况下, 个体倾向于对自己的行为做出情境归因, 而观察者倾向于对相同的行为作出稳定的个性归因。”

在有关这种行为人与观察者归因差异的研究中, 最为著名的一个实验由尼斯比特等人(Richard Nisbett, Craig Caputo, Patricia Legant, & Jeanne Marecek) 在1973年发表。尼斯比特和他的同事们完成了三个独立的实验, 每一个实验都发现, 个体在对自己的行为进行归因时与他人对此的归因有明显的差异。例如, 在其中的一个实验中, 尼斯比特和他的同事让男性的大学生对四个问题给出简短的解释: (1) 在过去的一年中, 他们为什么喜欢他们最经常约会的女孩儿? (2) 他们为什么选择现在的专业? (3) 在过去的一年中, 他最好的朋友为什么喜欢他最经常约会的女孩儿? (4) 他最好的朋友为什么选择现在的专业? 研究者对这些问题的回答进行了编码记分, 以确定他们的归因类型。例如, 有些回答是“她是一个非常轻松的人”, 或者“化学专业以后是一个很挣钱的专业”, 这样的回答被认为是情境因素归因; 而有些回答是“我需要一个能够轻松相处的人”, 或者“我希望挣很多钱”, 这样的回答被认为是个性归因。

尼斯比特和他的同事发现, 学生们在解释他们喜欢自己女朋友的原因时, 他们给出情境原因(与女孩的个性因素相关的原因)的数量是个性原因(与他们自己的需求、兴趣和特质相关的原因)的两倍。另一方面, 当他们解释他们最好的朋友喜欢自己女朋友的原因时, 他们给出的情境原因和个体原因的数量大致是相等的。同样, 在他们解释自己选择专业的原因时, 所给出的情境原因和个性原因的数量大致是相等的; 但是当他们解释他们最好的朋友选择专业的原因时, 他们给出的个性原因的数量是情境原因的4倍。

对上面约会女孩儿问题和选择专业问题的数据进行平均, 我们可以发现, 学生们在对他们自己行为进行解释时, 他们更多地倾向于进行情境因素的归因; 但是在对他人的行为进行解释时, 他们更加倾向于进行个性因素的归因。或者, 用另一种方法来诠释这个实验结果, 个体在对自己的行为和他人的行为进行归因时, 使用个性因素进行归因的数量大致相同, 但是他们在对自己进行

归因时使用的情境因素的数量是对他人进行归因时的两倍。尼斯比特和他的同事主要是根据行为人和观察者所获得信息在突出性上存在的差异来解释归因方式上的分歧。

1982年，戴维·沃森（David Watson）发表的一篇综述性的论文对这个问题进行了深入探讨。沃森希望了解行为人一观察者差异主要存在于个性归因中，还是情境归因中，或者两者中同时存在。他发现，在绝大多数研究中，行为人和观察者都倾向于更多地进行个性因素的归因，而不是进行情境归因。例如在尼斯比特等人（1973）的研究中发现的结果就是如此，尽管不同的个体在进行情境归因时存在差异。相对于他人的行为而言，人们更加愿意对自己的行为进行情境归因。

交换位置

如果行为人一观察者归因差异主要是来自两者的视角不同，那么如果交换两者的视角是否会出现同样的差异呢？这就是迈克尔·斯托姆斯（Michael Storms, 1973）进行实验的基本逻辑，在他们的实验中，行为人和观察者有机会相互交换位置。斯托姆斯录制了一段录像，其中有两个“行动被试”来完成一段初次见面的对话，而两位“观察被试”在镜头外进行观察。每一位观察者被分配来观察不同的行为人，同时有两架摄像机完成录像工作——其中一架摄像机拍摄的角度是从行动者1的角度观察行动者2，而另一架摄像机拍摄的角度是从行动者2的角度观察行动者1。一共有30组、120位被试参加了实验。

在完成了这个初次见面的交谈以后，被试被分配到三个不同的实验条件下：（1）相同观察角度组，在该组中，观察者所看录像的对象就是自己先前观察的相同的行为者，而行为者所看录像的对象是其谈话的同伴；（2）全新观察角度组：在该组中，观察者所看录像的对象是自己先前并没有观察的行为者（实际上采用是其观察的行为者的视角），而行为者观察的是其自己（实际上就是观察他们行为的观察者的视角）；（3）没有录像组：在这个实验条件下，行为者和观察者都没有看谈话的录像。

接着，斯托姆斯要求被试评价在多大程度上，行为者的行为是由其个人的特点或者情境的特点所决定的。其中行为者评价自己的行为，而观察者评价的对象就是在初次见面交谈时他们被要求观察的行为人。斯托姆斯发现，从相同角度观看录像的被试，或者没有观看录像的被试出现了行为人一观察者归因差异（见图16.4）。而从相反角度观看录像的被试却出现了相反的归因差异。根据这样的实验结果，斯托姆斯认为，视觉的观察角度对个体的因果归因会产生很大的影响。

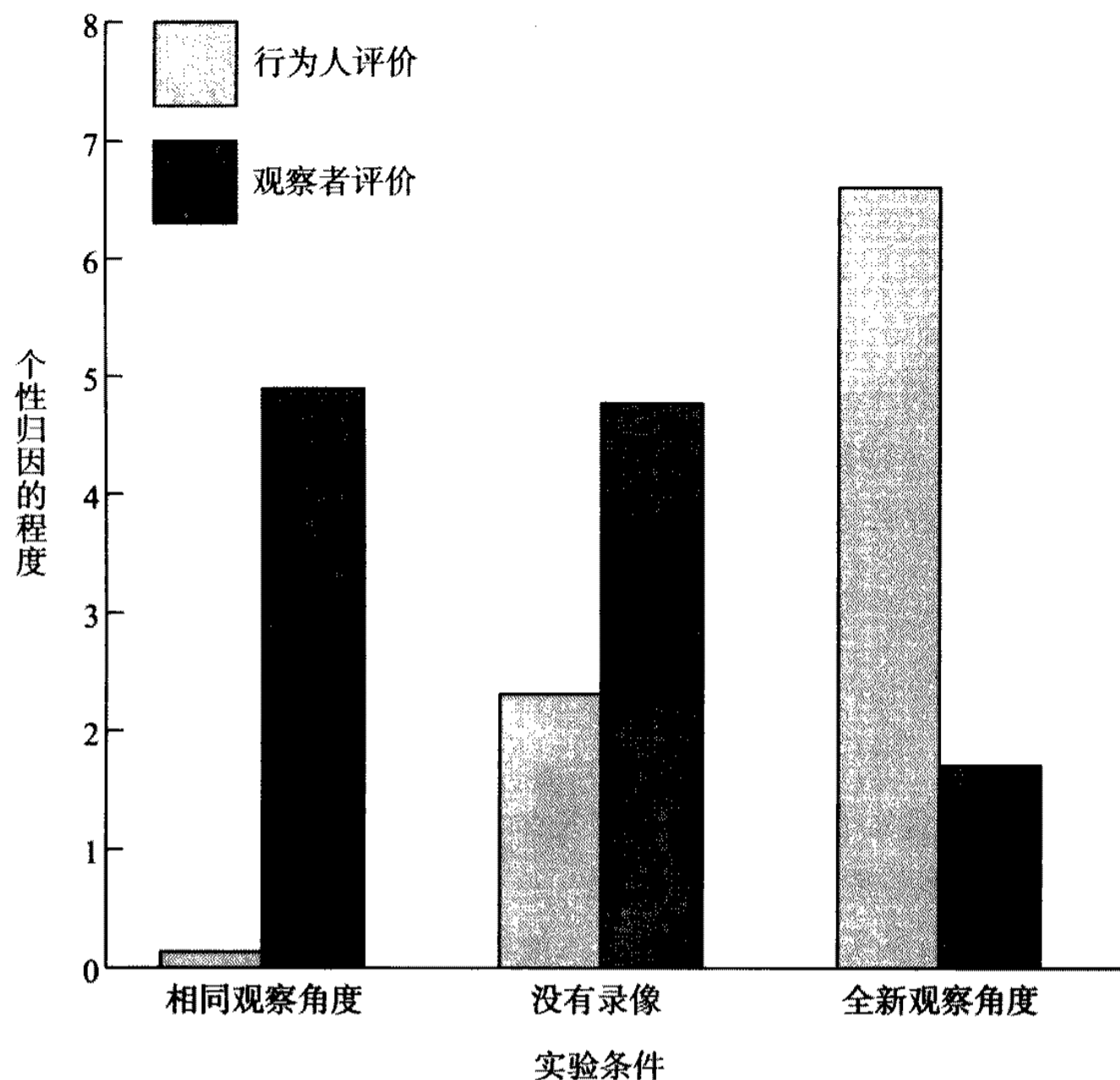


图 16.4 当行为者和观察者交换各自的视角以后，归因差异出现了翻转。（摘自 Michael Storms 1973 年的实验结果）

临床含意

绝大多数的研究都得出了一个相同的结果：当对行为的原因进行解释时，人们总是依赖于当时突出的因素。对于观察者而言，行为者就是突出的因素；而对于行为者而言，情境的需要是突出的。这样一个看似细小的差异却具有深远的含意。在许多社会判断中，将行为归因为情境因素还是个性因素，将起到关键性的作用，包括陪审团和假释官的决策，对受虐待妻子的知觉以及对学业失败的解释等等（Frieze, Bar-Tal, & Carroll, 1979）。

在心理治疗领域，因果归因显得尤为重要（Baston & Marz, 1979; Ross, Rodin, & Zimbardo, 1969; Shenkel, Snyder, Batson, & Clark, 1979; Snyder, 1977; Valins & Nisbett, 1971）。在临床治疗中对这样一个问题感兴趣，主要是因为治疗方案的制定很大程度上取决于将患者的问题归因于情境因素还是其个性因素（Batson, 1975）。如果情境因素是问题的症结，那么治疗师的目标就是努力改变其环境；如果个性因素是主要因素，努力的对象就是改变他自己。非常不幸的是，这样的判断很大程度上与治疗师的视角和接受的培训有关，而并非上文中分析的那么简单。本书的作者普劳斯和菲利普·津巴多曾经进行过一个研

究，结果显示归因偏差在很大程度上影响了治疗师的判断。

在该研究中，他们用信件调查的方式对著名的精神分析学家、资深的行为治疗师以及没有治疗经验的控制组的归因方式进行调查。实验者要求被试解释三个假设的问题，分别关于他们自己、他们最要好的同性朋友以及他们的患者。例如，其中的一个问题就是要求被试解释出现周期性失眠并伴有梦魇症状的原因：是情境性的还是个体性的原因，是生理的还是心理的原因。

实验的结果出现了很多有趣的现象。第一，精神分析学家倾向于将问题归因于患者的个性因素，而行为治疗师和控制组被试则倾向于将原因归结为环境的因素，或者是个体因素和环境因素共同起作用。第二，如果是自己的朋友或者患者遇到了假设的问题，精神分析学家倾向于将问题看成是心理的问题，但是如果是自己遇到相同的问题，精神分析学家则认为主要是躯体（生理）的问题。行为治疗师和控制组的被试则没有出现如此的归因。最后，与不拥有医学学位的治疗师相比，拥有医学学位的治疗师会更多地进行躯体因素的归因，而不是心理因素。根据这样的结果，研究者得出了如下结论：“就判断疾病的起因和推荐治疗方案而言，选择什么样的治疗师才是关键，与实际的问题并没有多大关系。”

其他的归因偏差

除了行为人一观察者归因差异、基本归因误差以及忽视共同反应的倾向以外，还存在许多其他的归因偏差影响个体的社会决策。例如，个体更加容易接受成功的努力，而不愿意承担失败的责任（Miller, 1976; Mullen & Riordan, 1988; Schlenker & Miller, 1977）。戴尔·米勒和迈克尔·罗斯（Dale Miller & Michael Ross, 1975）将这样一种现象称为“自我服务”归因偏差（self-serving bias in attribution），同时，研究者们认为，认识因素和动机因素的复杂结合（例如，错误的期望、对成功的渴望以及对自尊的保护等等）导致了这一偏差的出现。

与之相关的一种归因偏差被称为归因中的“自我中心”偏差（egocentric bias），它同样得到了大量研究的证实。这种自我中心归因偏差主要是指对于与他人共同完成的结果，个体认为自己承担了更多的责任。迈克尔·罗斯和菲奥里·西科里（Michael Ross & Fiore Sicoly, 1979）针对自我中心偏差进行了第一个实证的研究。在这个研究中，他们要求37对夫妇完成一份关于双方关系的调查问卷，问卷的内容是要求双方评估他们在进行20项活动中所承担的责任，这些活动包括做早饭、洗碗、沟通情感等等。研究结果发现，在其中的16项活动中都出现了自我中心的归因偏差（将夫妻双方对该项活动的责任相加，所得到的数值超过了100%）。

尽管有人认为自我中心归因偏差的产生仅仅是因为双方为了争取更多的正

性结果或行为。但是罗斯和西科里的研究发现，对于不良行为，同样存在自我中心归因偏差。例如，他们的研究发现，夫妻双方都认为他们的婚姻关系存在的冲突主要是由自己引起的。这样的结果表明自我中心归因偏差的产生并不仅仅是个体为了提升自我。在某些情况下，其产生的原因可能是个体自身的行为更加易得（更容易从记忆中提取）（Thompson & Kelley, 1981）。

与上面所提及的偏差相关的另一个偏差是谢利·泰勒和 Judy Koivumaki (1976) 提出的“正性偏差效应”（positivity effect）。正性偏差效应是指个体倾向于将正性的行为归因为个性的因素，而将负性的行为归因为情境因素。在泰勒和 Koivumaki 的研究中，他们向一些夫妇呈现了一系列正性和负性的行为列表，这些行为是由他们自己，他们的配偶、朋友或者一个熟人完成的。例如，正性的行为包括赞扬他人，与他人高兴地交谈或者玩得很开心等等。负性的行为包括与他人进行激烈的争论，粗鲁地对待他人或者忘记去做某件事等等。泰勒和 Koivumaki 要求被试判断这些行为在多大程度上是由个性因素或是情境因素造成的。他们的实验结果发现——无论是谁表现出这样的行为——对于正性的行为，他们总倾向于进行个性归因，而对于负性的行为，他们更倾向于进行情境归因。

当然，泰勒和 Koivumaki 并没有要求被试解释，如果行为人是他们不喜欢的人时，这些行为产生的原因；如果他们这样提问，他们就可能发现“负性效应”（negativity effect）（例如，Regan, Straus, & Fazio, 1974）。事实上，托马斯·佩蒂格鲁（Thomas Pettigrew, 1979）将这种负性归因偏差称为“最终归因误差”（ultimate attribution error）。佩蒂格鲁是一位广受尊重的研究种族歧视方面的专家，他警告说：“如果将种族的因素考虑进来，这样的归因方式就有可能使我们认为受贬低的个体的行为结果完全是由他们不变的、基本的个性特点决定的——这就是种族主义立论的根基”。

佩蒂格鲁的警告并不是空穴来风。1976年，伯特·邓肯（Birt Duncan）将这种归因偏差与种族歧视联系起来。在他的实验中，实验者要求学生观看一段录像，录像中的两个人由讨论发展到激烈的争论，在录像的最后，其中的一个人还推了对方一下。实验共分四个不同的实验条件：（1）推人者是黑人，而对方是白人；（2）推人者是白人，而对方是黑人；（3）双方都是白人；（4）双方都是黑人。实验的结果表明，被试不仅认为黑人推人者比白人推人者更加暴力，而且他们倾向于对黑人的推人行为进行个性归因，而对白人的推人行为进行情境归因。

另一种形式的归因偏差——这种偏差经常会强化偏见的判断——是个体认为他人的多样性并不如自己多。这样的偏差首先是由德国的研究人员丹尼尔·卡默（Daniele Kammer, 1982）提出来的，紧接着美国的研究人员特里·巴克斯特和路易斯·戈德堡（Terri Baxter & Lewis Goldberg, 1987）重复了这个实验。卡默呈现给被试 20 对双极的特质量表，例如安静—健谈，警觉—冒失等

等。研究者要求被试判断在不同的情境中，自己与自己最好的朋友在这些维度上是如何变化的。卡默的研究结果发现，个体认为自己的行为在不同的情境中更具多样性，而朋友的行为则不然。其他的研究结果还发现个体认为自己更多才多艺，而且与他人相比，自己的行为更具有不可预测性（Sande, Goethals, & Radloff, 1988），同时他们更加愿意将自己的行为归因于一些暂时的状态——心境、想法和感觉等（White & Younger, 1988）。因此，这种将对他人的种族的、民族的和性别的观点相对泛化的做法进一步加剧了刻板印象。

结 论

你如何解释发生在你自己身上的事情？你经常将它们归因于自己的个人能力、兴趣和特质，还是归因于外界的因素？这样的归因只是针对某一个特定的情境，还是普遍存在的情境？

本章的开头简短地描述了三个情境：（1）你遇到你的一个朋友，她一味称赞你的外貌；（2）你去参加一个约会，但是结果很糟糕；（3）你做一个非常重要的报告，但是听众的反应比较负面。这些情境都是取自由马丁·塞利格曼和他的同事（Peterson, Semmel, von Baeyer, Abramson, Metalsky, & Seligman, 1982）开发的归因风格问卷。塞利格曼和他的同事使用这样的问卷来判断抑郁的个体是否与他人拥有不同的归因风格。在较早的一篇文章中，林恩·艾布拉姆森等人（Lyn Abramson, Martin Seligman, & John Teasdale, 1978）假设抑郁的个体倾向于将负性的事件归因于某些稳定的、普遍的和内部的原因（例如，“像往常一样，约会很糟糕是因为我就是个倒霉的人”）。塞利格曼和他的同事完成的实验证实了这一假设（Seligman, Abramson, Semmel, & von Baeyer, 1979）。基于100多个研究的“元分析”（Sweeney, Anderson, & Bailey, 1986）统计结果也得出了相同的结论。归因的风格是直接与健康情绪幸福感联系在一起。

幸运的是，本章中所谈到的很多归因偏差都是可以避免的。例如，减少基本归因误差的一种方法就是更加关注共同反应的信息。如果在一个相同的情境中，绝大多数个体的行为都是类似的，这时做出个性因素的归因是不合适的。相反，观察者应该从情境因素中寻找对行为的解释。

另一种“消除误差”的方法是反问自己如果身处相同的环境，自己将如何行为。研究表明这种观点采择的方法可以扭转行为—观察者归因差异（Galper, 1976; Regan & Totten, 1975）。我们将在编后语中进一步讨论这项技术。

由于因果归因往往取决于事件发生当时最为突出的因素，因此寻找隐藏的因素也是相当重要的。约翰·普赖尔和米切尔·克里斯（John Pryor & Mitchell Kriss, 1977）的研究发现，即使是措辞方面细微的变化都可以影响信息的突出

性，从而改变因果归因。例如，“弗雷德喜欢这辆车”这样的论断就比“这辆车弗雷德很喜欢”这句话更可能引起个性因素的解释，因为在前一句中弗雷德更为突出。这种对问题措辞的敏感性与第5章和第6章中的研究结果是相吻合的，这就意味着因果归因和其他许多判断相似，是具有“可塑性的”。

第五部分

决策与判断中的社会性 一面

本书前面几部分关注的主要是单个决策者的行为。但在很多情况下，其他人会对决策者产生很大影响。这一部分就是关于社会因素如何影响人们的决策和判断（第17章），并且比较了群体行为和个体行为的异同（第18章）。

第 17 章

社会影响

像别人看我们一样看我们自己可能会证实我们对别人最坏的猜忌。

——富兰克林·P. 琼斯 (cited in Peter, 1977)

正如第 16 章所说的那样，人们在做归因的时候经常会轻视或忽略多数人的意见。那么这是否意味着决策者不在乎他人的行为和态度呢？正好相反，即使是最独立的决策者都会受到社会因素的强烈影响。

事实上，菲利普·泰特洛克 (Philip Tetlock, 1985b) 已经指出，社会因素在人们的决策和判断中起着关键作用。泰特洛克 (p. 325) 说：“过去用认知实验方法研究决策和判断，所关注的焦点是狭窄而有误导嫌疑的。这方面的研究需要得到扩展，把社会和组织的背景因素考虑进去。”泰特洛克建议研究者把决策者视为“政治家”，他必须对自己的“选民”（如他的朋友、家人和同事）负责，并且总是关注“如果我这样做别人会有什么反应”、“如果别人质疑我的观点，我应该怎样说服他们”这样的问题。外部评价的重要性——以及这些评价对人们行为的影响——正是实验社会心理学最早的研究成果之一。

社会易化

社会心理学家戈登·奥尔波特 (Gorden Allport) 曾经写道 (1954, p. 46)：“第一个实验研究的问题——实际上也是实验研究头三十年关注的惟一问题——就是：当有其他人在场的时候，个体的行为表现与独处的时候相比会有什么变化？”虽然对这个问题的研究早在 19 世纪下半叶就已经开始，但是直到 1965 年才有一个完整的答案。在那一年，罗伯特·扎伊翁茨 (Robert Zajonc) 发现，对于那些简单的、熟练的行为，人们在有旁观者在场时通常能做得更好；但是对于那些复杂的、还没有掌握好的技能，人们在有他人在场时则会表现得没那么好。扎伊翁茨认为这一被称为“社会易化” (social facilitation) 的效应，至少部分地来自于他人在场时个体产生的唤醒状态。后来的研究则显示，即使其他人并没有实际在场，只要个体预期自己的行为会受到其他人的评价，这种促进或损害的效应也会发生 (Henchy & Glass, 1968)。

尽管研究者在各种言语和数学任务中都发现了社会易化效应，但最直接的

证据来自于在一个大学撞球室里进行的研究 (Michaels, Blommel, Brocato, Linkous, & Rowe, 1982)。在这个研究中,一些隐藏的观察者在玩撞球的人分成技术高于平均水平和低于平均水平两组,然后记录这些球手在有观众和没有观众两种条件下的击球成功率。结果如图 17.1 所示,观众促进了那些技术高于平均水平的球手的表现,却损害了那些技术低于平均水平的球手的表现。查尔斯·邦德和琳达·泰特斯 (Charles Bond & Linda Titus, 1983) 在对共涉及 20 000 名被试的超过 200 个研究的元分析中也得到了 (尽管没有那么明显) 类似的结论。有观众在场会损害个体完成复杂任务的正确率,同时对完成简单任务的正确率会稍有改善。

社会性懈怠

社会性懈怠 (social loafing) 并不是他人在场影响个体表现的惟一方式。在社会性懈怠效应被提出约 30 年以后,沃尔瑟·莫德 (Walther Moede, 1927) 报告了一个实验,这个实验发现,当人们作为群体的一员做一件事情时,就不会像自己独立完成时那么努力。这是莫德的一个学生林格尔曼 (Ringelmann) 做的实验,他发现,与参与两人组、三人组或八人组拔河时相比,当个体独自一人与对手拔河时会拉得更卖力。平均而言,两人组的时候单个被试的努力程度只有单人组被试的 93%,三人组中单个被试的努力程度只有 85%,而八人组中单个被试的努力程度则只有单人组被试的 49%。艾伦·英厄姆等人 (Alan

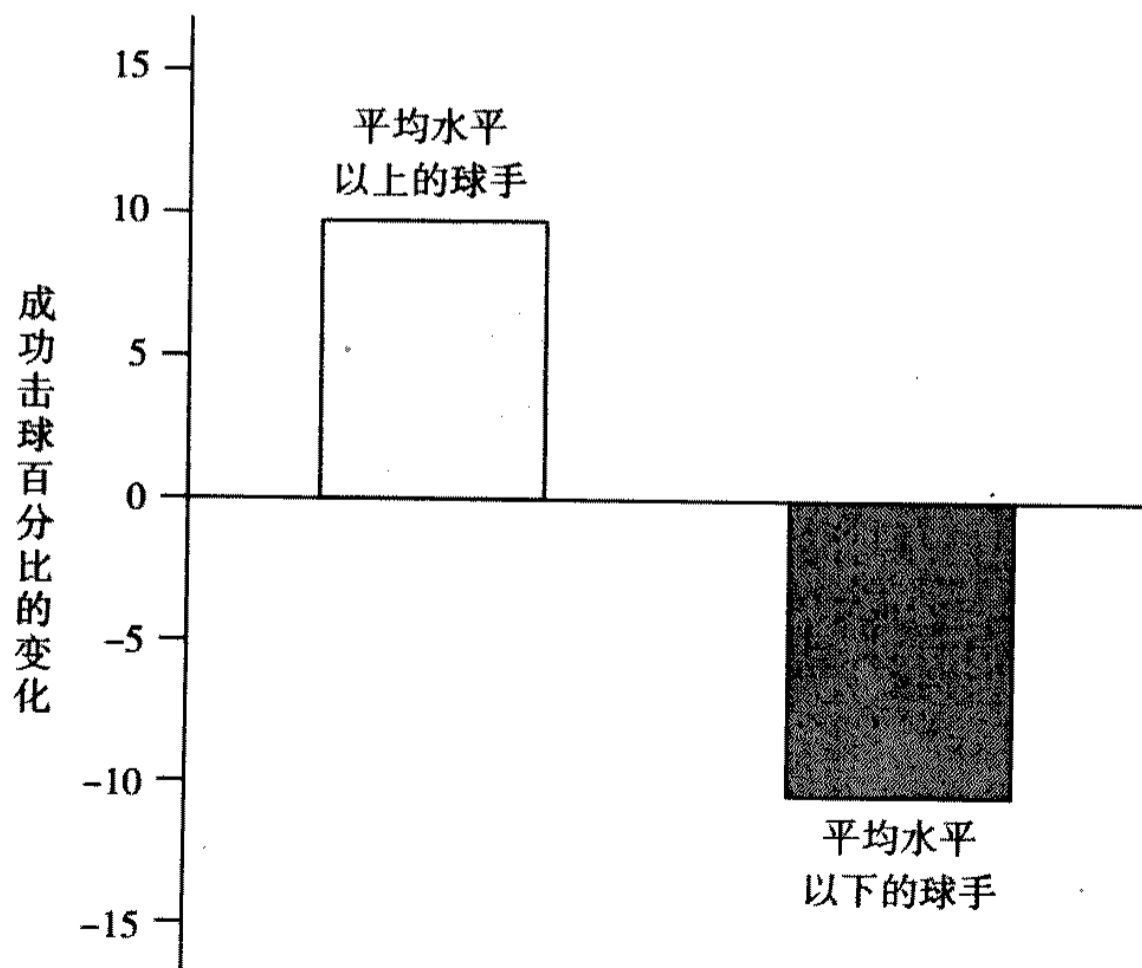


图 17.1 观众的在场会提高技术处于平均水平以上的球手的表现,却会损害平均水平以下的球手的表现 (摘自 Michaels 等研究者 1982 年的研究)。

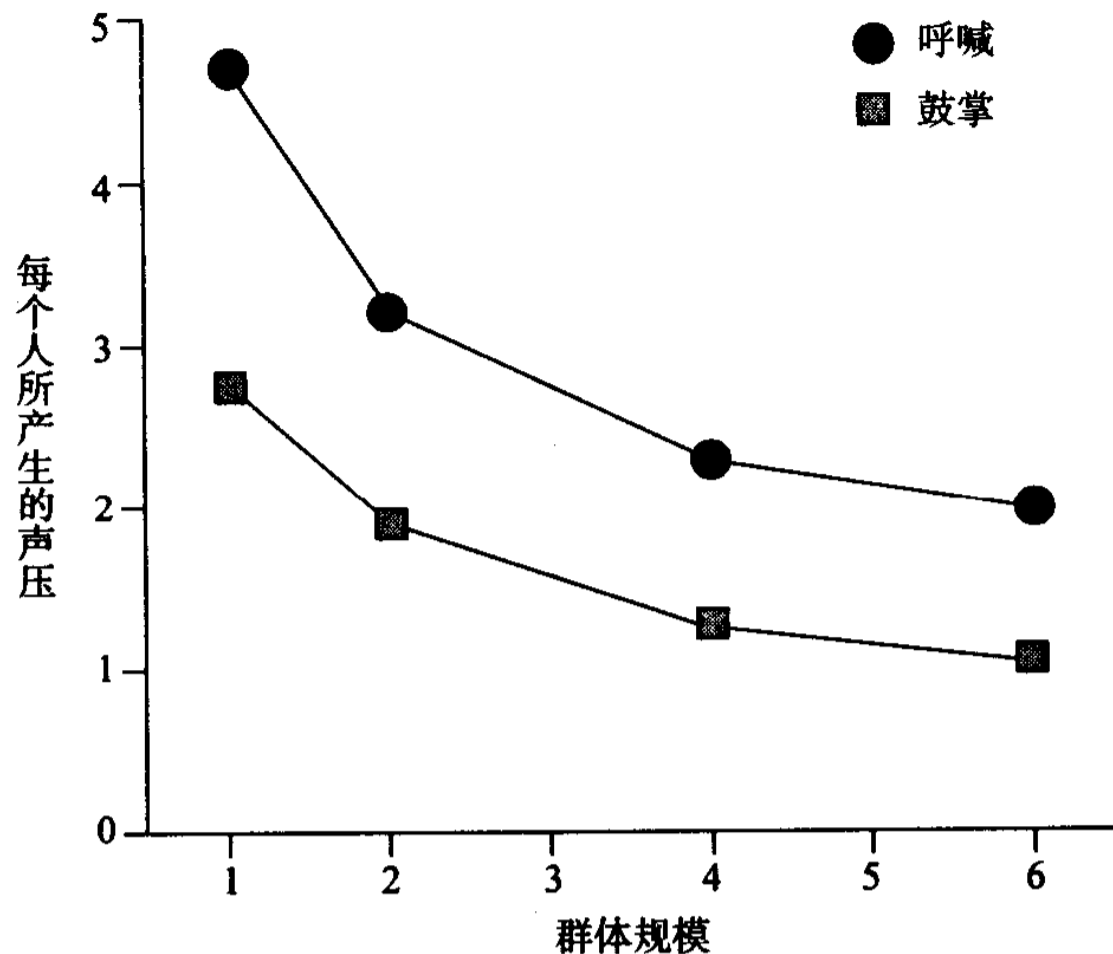


图 17.2 本图表现了社会性懈怠效应。平均而言，随着群体成员的增加，单个个体的贡献下降。从折线的走势来看，单个个体独自呼喊或鼓掌时制造的平均声响（每平方厘米的达因数）约是六人组中单个个体发出声响的 2 倍。（摘自 Latane, Williams 和 Harkins, 1979）

Ingham, George Levinger, James Graves, & Vaughn Peckham, 1974) 巧妙地用不同的方法重复了这个实验，研究者给被试蒙上双眼，使他们相信自己正与同组的其他成员一起与对手拔河，而实际上他们是单人应战。这种不同的实验方法却得到了与林格尔曼的实验相同的结果。比布·拉坦等人 (Bibb Latane, Kipling Williams, & Stephen Harkins, 1979) 将其称之为“社会性懈怠”效应，而且他们还发现，当要求人们尽可能大声地呼喊或鼓掌时也会出现这样的现象 (见图 17.2)。

是什么原因造成这种“社会性懈怠”的现象？目前还没有明确的答案。拉坦，威廉斯和哈金斯认为，社会性懈怠效应的出现是因为处于群体中时人们不会像独自行动时那样直接感觉到自己的努力和最终结果之间的关系。与这个差异相关的是，对最终结果所担负的责任会在群体成员之间分散，而独立行动时个体则要对结果负全部责任。责任分散会对决策和判断产生强有力的影响。下面 3 个事件就是这种影响力的体现：

事件 1：“放开那个女孩！”

这是 1964 年 3 月 13 日凌晨 3:20，纽约 Kew Gardens 火车站旁边的停车场。基蒂·吉诺维斯 (Kitty Genovese)，一位 28 岁的酒吧经理，刚刚下班回家。她把自己那辆红色的菲亚特汽车停好，关了车灯，锁上车门，然后向着自己 100

英尺以外的公寓走去。

突然她发现在停车场的另一边有一个男人。

基蒂转头就向附近林荫道的一个报警亭跑过去，但是还没有来得及跑到那里，就在一盏路灯下被那个男人捉住了。基蒂尖叫了一声，马路对面一幢10层高的公寓楼里亮起了不少灯，有些窗户打开了，基蒂大声喊：“噢，我的上帝，他刺中我了！救命啊！”

其中一间公寓里有个男人往下喊：“放开那个女孩！”

攻击者畏缩了，扔下基蒂跑了，基蒂在流血。很快公寓楼里的灯又熄灭了。基蒂尽力想回到自己的公寓，但是那个男人又出现了，又刺了她一刀。这一次基蒂发出了惨叫：“我要死了！我要死了！”公寓楼里的灯又亮了起来，邻居们打开了窗户，那个男人吓得钻进自己的车里飞奔而去。

过了一会儿，基蒂挣扎着站了起来。一辆开往肯尼迪机场的公共汽车开了过去。这时候是凌晨3:35。

最后基蒂总算挣扎着回到她的公寓门口，却在这儿昏过去了。那个男人又折了回来第三次刺中了她——这一次是致命的一击。

关于这个故事的令人心寒的结语是：当这桩谋杀刚刚发生的时候，《纽约时代》杂志对它的报道只有在第26页上的短短5句话。可是两个星期后，这一事件却出现在头版的一篇长篇报道中。为什么？因为警方的调查发现，不少于38名“可敬的、遵纪守法的市民”目睹了这次谋杀，可是却“没有一个人在袭击的过程中打电话报警”（Gansberg, 1964. 3. 27, p. 1）。警方在基蒂开始遭袭击后30分钟才接到这些证人的第一个报案电话。也就是说，作案者在半个小时的时间里多次来回袭击基蒂，基蒂也不断呼救，但这38名目击者却没有一个人去帮她或者从自己家打电话报警。

事件2：隧道里的故事

1990年7月14日，星期六，意大利佛罗伦萨附近的一条公路。马可·莫雷蒂正开着车带他6岁大的女儿范尼莎到海滨去。正当他驶入一条隧道的时候，马可突然心脏病发作。马可挣扎着把车停在一边，让范尼莎自己设法回家。随后，33岁的马可不幸身亡。

范尼莎挪到车外，走到川流不息的汽车中间求救。飞驰的汽车行驶速度太快，掠过的气流把范尼莎撞倒在地上。但是她没有停下来——她抓爬着、流着血、哭着——绝望地向过往车辆求救。在接下来的30分钟里，范尼莎走了超过一英里的路，期间有数百辆车飞驰过她身边，却没有一个人停下来帮助她。最后终于有一位驾驶员把车开过来帮助她了，很快警察也到了。

就像基蒂·吉诺维斯的故事那样，这个故事也成了全国报纸的头条。意大利人都在扪心自问，为什么连这样的事都能发生？他们都把它视为一种信号，

代表着意大利不再是那个曾经是那么温暖、充满同情心的地方了。例如，一位意大利社会学家下面的话就被广泛引用，他说，范尼莎事件“代表了一个我们倾向于视而不见的意大利——那个冷漠的、正处于冰河期的意大利，大部分时候每个人都只会想着自己以及和自己多多少少有密切关系的少数几个人。”（Haberman, 1990年7月19日，p. A11）

事件3：谁会捡起这些东西

这是20世纪70年代早期，美国哥伦布、西雅图或亚特兰大其中一座城市的一部电梯里。就在电梯门关上以后，站在里边的一个黑衣人“一不小心”掉了8~10根铅笔在地上（或者有些情况下是一些便士或镍币——是什么东西并不重要）。

作为比布·拉坦和詹姆斯·达布斯（1975）的一个现场研究的一部分，这样的实验模式被用于1497个不同的场合，有145人在合计4813名旁观者面前掉下不同的东西。拉坦和达布斯对下面的问题感兴趣：“谁会帮丢东西的人捡起这些东西？”他们的发现之一是随着电梯里人数的增加，旁观者帮助丢东西的人的可能性就会减小。

旁观者干预

贯穿于前面每个事件中的共同主线就是：在相对较大的群体中，给予帮助的责任会被分散。当人们面临着是否干预某件事的决策时，他人在场就会影响到他们如何决策。干预与责任分散的这一关系最初是在比布·拉坦和约翰·达利（1969, 1970）的一系列关于旁观者干预效应的精彩实验中被发现的。

在其中一个实验中，拉坦和达利邀请一些学生来参加一个关于“在城市的大学里生活会遇到的一些问题”的访问。学生到达以后首先在休息室里等候，或者独自一人，或者还有两位主试的助手装扮成的参与者（这两位助手在整个实验过程中都要保持被动），或者还有两位真正的参与者。然后，当被试坐在休息室里的時候，一股白烟开始从墙上的一个通风孔里冒出来。拉坦和达利想看看和独自一人时相比，当有他人在场时这些学生被试是否更不那么倾向于报告冒烟一事？

结果拉坦和达利发现，独自一人的情况下有3/4的学生被试报告了冒烟一事，其中一半人是在2分钟内报告的。相反，当与两位不采取任何行动的假扮的被试在一起时，只有1/10的学生被试报告了冒烟的事。其他人虽然咳嗽、揉眼睛甚至打开窗户——但就是不去报告！

至于三名真正的被试一起等待的情形，直接将他们与独自等待的被试的反应作比较并不恰当。如果独自等待时有75%的学生报告了冒烟事件，那么对于

三个独立的被试一起等待的情况，相应的比例就应该是98%（因为所有三个学生都不报告的概率是 $.25 \times .25 \times .25 = .02$ ，所以 $1.00 - .02 = .98$ ）。但实际的情况是，三人组时只有38%的被试报告了冒烟事件。

为了确保这些结果不是由一些无关的外部因素造成的（如独自一人与和他人在一起时对火灾的恐惧感不同），拉坦和达利决定在一个不同的背景下重复这个实验。在一个表面上名为“市场调查”的实验中，研究者让参加的学生独自等待，或者与一个朋友一起等待，或者与一个陌生人一起等待，或者与一个不采取行动的假扮的被试呆在一起。然后在他们正在等待的时候，研究者让他们相信隔壁屋里有一人跌到并且摔伤了。这是通过播放一段高保真录音带来实现的。录音带里首先传出一阵碰撞声，然后是一个女人的呼喊声：“噢，我的上帝，我的脚……它……它……动不了……噢，我的脚踝！我……不能……不能……挪开这些东西。”

听到她的喊声后，独自等待的学生中有70%的人会跑过去帮忙。这意味着如果两个人中的一个会独立于另一个采取行动，那么至少其中一个人去帮助那位女性的可能性为91%（也是因为每个人都不提供帮助的可能性是 $.30 \times .30 = .09$ ，这样 $1.00 - .09 = .91$ ）。相反，不熟悉的两个人坐在一起时，只有40%的被试主动提供帮助；而认识的两个人坐在一起时，则有70%的被试主动提出要帮忙。这表明，尽管朋友一起等待时比与陌生人一起等待时会有更大的可能性采取干预行为，但是与独自一人的情况相比，朋友之间还是会出现抑制效应。当比较被试独自一人和与假扮的参加者在一起时的行为时，来自他人的抑制效应也清楚地表现出来。在后一种情况下，只有7%的被试采取了干预行为。

拉坦和达利在其他几个研究中也发现了类似的结果。而且自他们最初研究这个问题以来，已经有很多其他研究者也证明了他人在场会抑制帮助行为。实际上，在拉坦和达利最初的那个研究完成后的十年间，研究者做的56个实验中有48个实验结果显示，有他人在场会使被试给予帮助的倾向降低（Latané & Nida, 1981）。对这56个研究的结果求平均，那么当人们独处时会有75%的可能性采取干预行为，而处于群体中时则只有53%的可能采取干预行为。

是否有某类人群能够避免旁观者导致的抑制效应的影响？从比布·拉坦和斯蒂夫·奈达（1981）所做的文献分析的结果来看，在美国社会中只有一类人能有这样的免疫力：9岁以下的儿童。那个阶段以后的人在决策时都会受到他人在场的强烈影响。

社会比较理论

关于社会易化、社会性懈怠和责任分散的研究都倾向于支持泰特洛克关于决策者是优秀政治家的观点。人们经常从他人身上直接获取行动的线索，而且他们非常关注他人对自己的看法。这些研究也显示出，人们通过把自己与他人

做比较来评价自己的观点和能力 (Suls & Miller, 1977)。描述人们如何进行此种决策判断的最精细的理论就是“社会比较理论”(social comparison theory)。

社会比较理论是费斯汀格在 1954 年提出的——他也是认知不协调理论的创立者。费斯汀格认为，人们具有评价自身能力水平和自身观点恰当性的需要。那么在缺乏客观的、非社会性的标准时，人们就会把自己和其他人做比较。费斯汀格尽可能做到详细和严格地对社会比较理论做界定。该理论包括 9 个假设、8 个推论和 8 个衍生观点，或者说一共有 25 个主要观点。

大致来看，这些命题最中心的观点是：

假设 I：人们具有评价自己的观点和能力的自然倾向性。

假设 II：在缺乏客观的、非社会性信息的时候，人们会通过与其他人的观点和能力做比较来评价自己的观点和能力。

推论 IIIA：在面临选择时，人们更愿意和那些观点和能力与自己接近的人做比较。

最后一个观点已经有丰富的研究成果，而且大部分研究都支持费斯汀格 (1954) 的看法。人们通常会与那些和自己相似的人做比较，而不是与那些和自己不同的人相比。例如，如果你是一名本科生，你很可能把自己的学术成就与其他本科生的学术成就做比较，而不是与研究生或者教授做比较。同样，助理教授一般会与其他助理教授做比较，而不会与资深教授做比较。

从那些与自己相似的人身上获取行动的线索

在哈维·霍恩斯坦、伊莱沙·菲什和迈克尔·霍姆斯 (Harvey Hornstein, Elisha Fisch, & Michael Holmes, 1968) 的一个有趣的现场实验里，社会比较的相似性作用得到了很好的体现。在这个研究里，在曼哈顿中心区有超过 100 名行人看到路上有一个写着地址却没有贴邮票的信封，信封里有一个遗失的钱包和一封写给钱包主人的信。钱包里有 2 美元现金，一张签给 E. M. H 公司 (这几名研究者名字的首字母组合!) 的支票，一张写着某人电话号码的纸片，以及钱包主人的身份证，上面有钱包主人的姓名、电话号码和地址，以及其他一些东西。钱包主人名叫迈克尔·欧文——这个名字是经过仔细挑选的，以避免含有种族或宗教的暗示——而且在所有情况下，钱包和里面的东西都是一样的。

有变化的是写给钱包主人的信。在某些实验条件下，写信者是与这些路人相似的人 (即一个能流利使用英语、像是本地人的写信者)，而在其他条件下，写信者则是一个与路人不同的人 (即一个几乎完全不懂英语的外国人)。另一个实验条件是不同的写信语气，在 1/3 的情况下，写信者用一种正性的语气写这封信，在另外 1/3 的情况下用一种中性的语气写信，在剩下 1/3 的情况下则

是用一种负性的语气写信。例如，一位操流利英语的人（“相似的他人”）用负性语气写的信是这样的：

亲爱的欧文先生：

我找到了您的钱包，如今物归原主。所有东西都在里面。我必须说的是，照看这个钱包和不得不把它还回来确实是一件很不方便的事。归还这个钱包的整个过程实在是麻烦，我觉得相当恼怒。我希望您会为我付出的努力而感到感激。

另一方面，一位不相似的他人用负性语气写的信是这样的：

亲爱的欧文先生：

我正在你们国家旅行，并不熟悉你们的行为方式。但是我发现了你的钱包，东西都在里面了。

对这个钱包负责并将它物归原主是一件很费劲的事情。其中的过程实在是繁琐，希望你能够对我有所感激。

霍恩斯坦、菲什和霍姆斯（1968）的假设是，当写信者是与自己相似的人时，人们所受到的信中语气的影响会大于写信者是不相似的人的时候（也就是说，他们预测当写信者是与被试相似的人时，正性语气的信会比负性语气的信导致更高的归还率，可是当写信者是不相似的人时，信的语气就不会产生影响）。正如图 17.3 所示，实际的结果符合研究者的预期。当写信者是本市人时，正性语气的信使人们归还钱包的次数是负性语气的信的 7 倍。相反，当写信者是一名外国人时，正性语气、负性语气和中性语气的信导致的归还率实际上是一样的。尽管这些结果本身不能说是结论性的，但它们的确支持了费斯汀格（1954）关于人们会从与自己相似的人那里获得行动线索的假设。

社会性痛觉丧失

1978 年，肯尼斯·克雷格和肯尼斯·普克钦（Kenneth Craig & Kenneth Prkachin）的研究展示了社会比较能在多大程度上影响人们的知觉。这些研究者发现，当人们把自己与一个有高的疼痛耐受性的人做社会比较时，这种比较实际上会导致人们体验到比其他情况时较少的疼痛感。

在这个研究中，研究者对被试的左前臂进行电击。最开始的时候电击轻微得几乎觉察不到，但随着实验的进行，电击强度一点一点地加大（每次增强 0.25 毫安）。每次电击后，被试在一个 0~100 的量表上标出自己觉得不舒服的程度。当他们把某次电击评价为 100 时，实验就会结束。

这个研究共有 2 个实验条件：忍耐模范条件和控制条件。在忍耐模范条件下，与每个被试配对的是一名假扮成被试的研究者的同事，他的疼痛评价总是

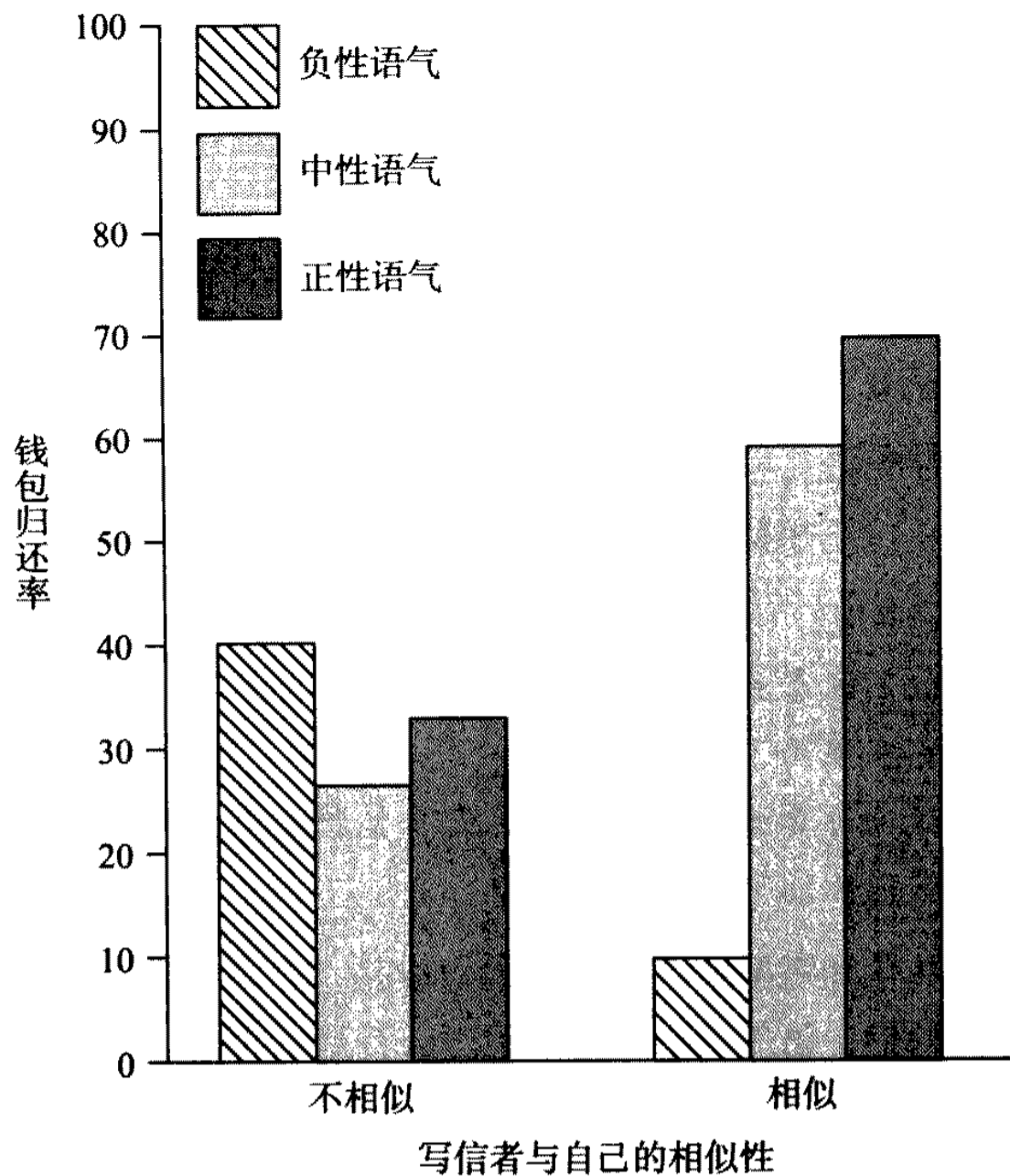


图 17.3 霍恩斯坦、菲什和霍姆斯 (1968) 进行的研究中, 各种实验条件下归还失落钱包的人数百分比。

比真正的被试低 25% (真正的被试总是先做评价)。在控制的条件下, 假扮的被试只是在一旁观察被试, 只有真正的被试会评价电击引起的疼痛 (这控制了他人在这场这一点)。

克雷格和普克钦发现, 当被试与忍耐力强的假扮被试一起配对时, 被试不仅会有较低的疼痛评价, 而且他们真的显得体验到较少的疼痛。在忍耐模范条件下, 虽然电击强度提高了, 但被试却比控制条件下的被试有较少的心率反应和较低的前臂皮肤电压 (这是对生理唤醒的两种测量指标)。因此, 社会比较既影响疼痛评价也影响躯体反应 (人体真正的生理反应)。

从众

请仔细看看“读者调查”第 32 题中 1、2、3 三条线。哪条线的长度与线 A 一样? 如果你和大多数人一样, 你会觉得第 3 条线和 A 一样长, 而且你会几乎毫无疑问地相信自己是对的。

可是如果你是在一个房间里和其他 7 个人一起判断, 而这 7 个人都一致声称第 1 条线和 A 一样长, 你又会如何反应呢? 你是会坚持自己的判断, 还是会

屈服于多数人一致的意见？

实际上这就是所罗门·阿施（1951；1955，11；1956）在他关于从众的经典实验中设计的情境。在他的多数实验里，7~9个男性大学生围着一张桌子就座，然后主试要求他们一起根据一张卡片上的一条“标准”线的长度，从另一种卡片上的三条“比较”线中找出哪一条和“标准”线等长。实验有18轮，每一轮的判断问题都非常简单。一组独立的判断者作为控制组，平均的判断正确率高达99%。被试不知道的是，和他们坐在一起的其他人实际上是主试的同伙。在多数实验中，每组只有一名真正的被试。

在实验的第一轮，标准线的长度是10英寸，三条比较线分别长 $8\frac{3}{4}$ 英寸、10英寸和8英寸。沿着桌子按顺时针的顺序，每名学生给出自己的判断，每个人都指出10英寸长的比较线与标准线一样长。到这里为止一切都是正常的。

在实验的第二轮，学生看到的标准线是2英寸长，而比较线分别长2英寸、1英寸和 $1\frac{1}{2}$ 英寸。这次学生们也是毫不费劲地找出了等长的线。

可是到了实验的第三轮，一些奇怪的事情发生了。这一次的标准线长3英寸，但是假扮的被试并没有选3英寸长的比较线，而是一致认为 $3\frac{3}{4}$ 英寸长的比较线与标准线一样长。那名真正的被试——他总是坐在接近最后发言的某个位置——不知道的是，阿施让假扮的被试在18轮实验中的12轮中都一致地给出错误答案。阿施想看看真正的被试是否会屈服于这种从众的压力。

阿施发现在这12轮实验中，有约 $\frac{1}{3}$ 轮实验中的被试会从众，即选择主试的同伙一致赞同的错误答案。 $\frac{3}{4}$ 的被试有不少于一次的从众，而 $\frac{1}{3}$ 的被试在12轮实验中的一半或更多轮中选择了从众。也就是说，大多数人做出了与自己的判断相反的选择，而且很多人在多数实验中都有从众的表现。

阿施还做了一些略做修改的实验。例如，在一组研究中，他改变了小组的人数（见图17.4）。当被试只与一名主试的同伙一起判断时，他们几乎完全不受同伙错误答案的影响。多数被试几乎在每一轮实验中都回答正确。当被试的判断与2名同伙的意见相反时，被试从众的概率是13%。当被试的判断与3名同伙的意见相反时，被试从众的概率是33%。阿施发现过了这个点以后，从众的概率就不会再上升，尽管后来其他研究者的研究结果显示，在持相反意见的多数被试人数超过3人以后，从众的概率还是会继续上升（Latané & Wolf, 1981）。在不考虑这个问题的情况下，一个非常明确的结果就是，当小组人数较少时，多数人的意见已经足以引起较高的从众水平。

在另一组实验中，阿施让另一名真正的、并且一般会给出正确判断的被试坐在第四位发言人的位置上，或者通过让一名会给出正确答案的实验者同伙坐在第四位发言人的位置上，从而打破了多数人一致意见的条件。阿施想看看当多数人的意见不再一致时会发生什么事情。

结果他发现这种做法对被试的反应有很大影响。从众的水平变成原来的四分之一——也就是说，当多数人的意见不一致时，被试从众的次数只是多数人

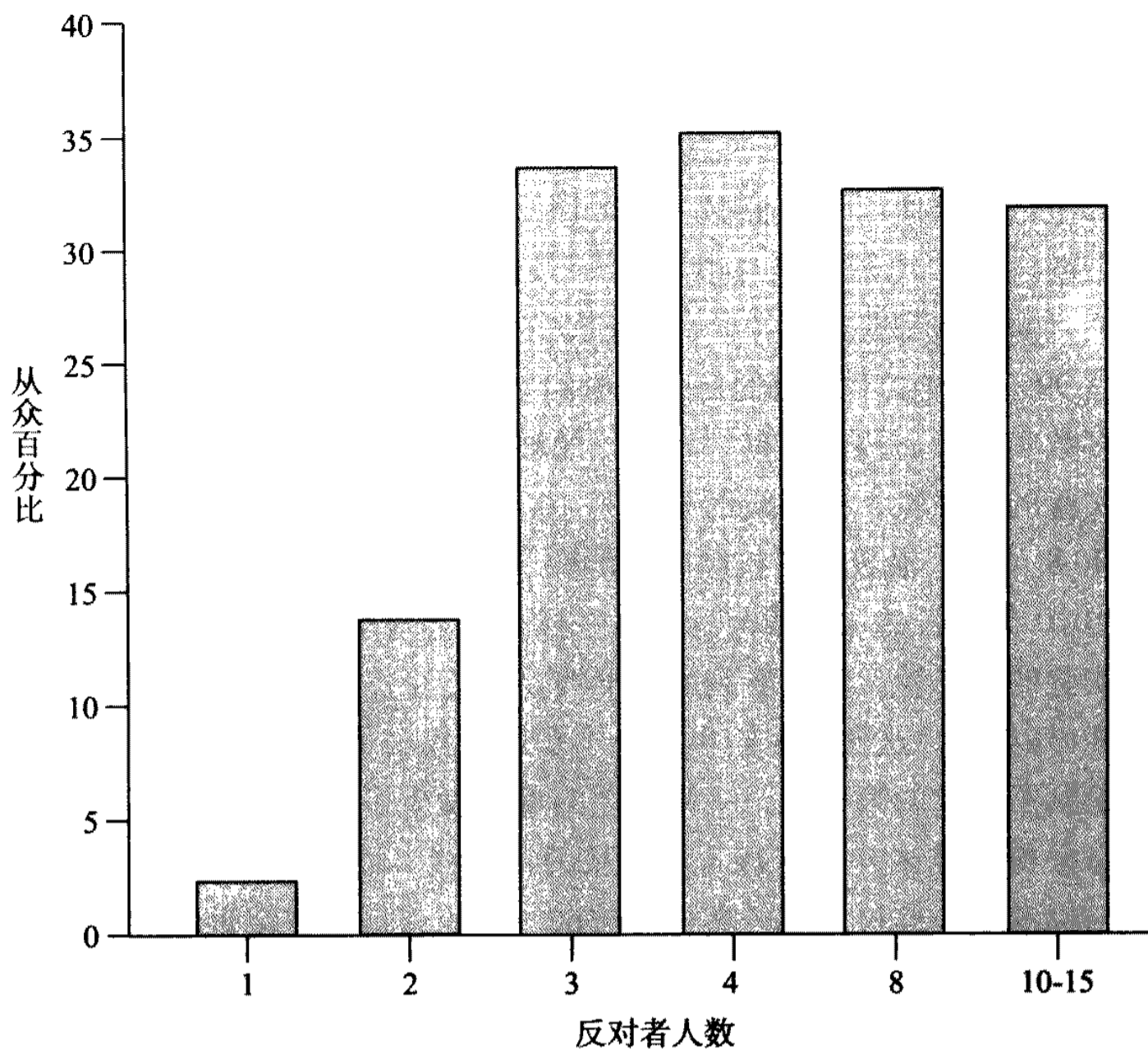


图 17.4 在所罗门·阿施关于从众的实验中，他发现三人组的多数意见与大 5 倍的组的多数意见在导致从众上的效果相似。(Asch, 1951)

意见一致条件下的约 1/4。实际上，阿施发现三个意见一致的多数派对被试判断的影响力，远远超过 8 个多数派加上一个异议者的条件。这些结果显示，当情境中存在强大的从众压力时，单个异议者会起到较大的作用。

少数派影响

1969 年，瑟奇·莫斯科维奇 (Serge Moscovici) 和他的两名同事把阿施的实验范式反转，用于研究少数人对多数人的影响 (Moscovici, Large, & Naffrechoux, 1969)。与阿施观察作为多数派的实验者同伙对一两名真正被试的影响不同，莫斯科维奇考察了少数几名实验者同伙对作为多数派的真正被试的影响。莫斯科维奇和他的同事发现，即使在少数派不是特别强大或者没有很高威望的情况下，意见一致的少数派也能产生显著的影响力。

在莫斯科维奇最初完成的一组实验中，每个小组包括 2 名实验者同伙和 4 名真正的被试。每个人都坐在一个屏幕前，主试告诉他们这个实验是关于颜色知觉的。然后屏幕上呈现了一系列的幻灯片。在每一轮实验中，小组成员要判断幻灯片的颜色和光强度。所有的幻灯片都是不同强度的蓝色，但是实验者同

伙在某些实验中却把它判断为绿色。莫斯科维奇和他的同事发现，被试在8%的实验中做出了和实验者同伙相同的判断，32%的被试把幻灯片颜色判断为绿色的次数不少于一次。相反，控制组的被试在792次判断中只有2次错误地判断为绿色。

因此，莫斯科维奇和他的同事证明了少数人能够对多数人产生显著的影响。但是存在一个前提——少数人的意见必须保持一致和稳定。如果两名实验者同伙随机地把幻灯片判断为蓝色或者绿色（其中1/3次的判断是蓝色），那么被试只在1%的实验中把幻灯片判断为绿色。只有在实验者同伙坚持同一个判断的情况下，少数人才会有显著影响（关于这方面研究的回顾与扩展，见Maass & Clark, 1984; Nemeth, 1986; Nemeth & Chiles, 1988）。

群体盲思

当群体具有凝聚力而且相对不受外界影响的时候，群体忠诚和从众压力会导致欧文·贾尼斯（Irving Janis, 1982）所说的“群体盲思”（groupthink）。贾尼斯对群体盲思的定义是：群体盲思“是指心理活动的效率、对现实的检验以及道德判断的一种退化，这种退化是来自于群体内压力”（p. 9）。群体盲思导致错误决策的最有名的例子就是美国的猪湾事件，在这一事件中，约翰·肯尼迪总统的内阁提出了一个入侵古巴的拙劣建议。更近期一点，詹姆斯·埃瑟和乔安妮·林多尔夫（James Esser & Joanne Lindoerfer, 1989）发现，1986年，美国在不安全的天气条件下发射挑战者号航天飞机的致命决策中就有群体盲思的因素在内。

根据贾尼斯的观点，群体盲思有8个普遍的特征：

- 群体中大多数或所有成员都抱有一种自己“无懈可击”的错觉，这种错觉导致过分乐观以及过度的冒险。
- 集体努力对警告采取忽视的态度或者对其进行合理化。
- 对群体固有的道德观加以毫不质疑的信任。
- 对提出建议者充满刻板印象，认为他们都太邪恶而不值得与之谈判，或者认为其太弱或太愚蠢而不会构成严重威胁。
- 任何不赞同多数人意见的群体成员都会感受到压力。
- 关于全体意见一致的共同错觉。
- 对与表面上一致的群体意见相背离的想法进行自我审查。
- 自封的群体“精神保卫者”，使群体远离任何可能破坏群体自满感的信息。

如果一个群体表现出这些迹象，那么它就很有可能受到了群体盲思的影响。关于如何才能避免群体盲思，贾尼斯推荐了几条预防措施。第一而且最重要的是，群体领导应该明确鼓励不同的意见和批评——包括对他们自身观点的

批评。第二，群体领导应该避免在一开始就表明自己的个人偏好。杰伊·拉索和保罗·休梅克（Jay Russo & Paul Schoemaker, 1989）观察到一些日本企业正是这样做的，他们开会的时候是由职位最低的群体成员首先发言，然后是职位次低的成员，依此类推。这样一来，群体成员就不会害怕自己的观点与比自己职位高的成员不一致。

第三条预防措施是与其他群体——或者与其他领导——一起考虑同一个问题（这样就可以比较不同的答案）。第四，群体成员应该与受信赖的同事定期对群体进行审议，并且向群体报告讨论的内容。第五，群体应该要求群体外的专家或者有资格的同事参加群体的会议，并且鼓励他们挑战群体的一致意见。事实上，贾尼斯指出，避免群体盲思最好的方法之一就是正式任命某位群体成员担当批评者的角色。每条措施都是使不同的意见合理化，都是在利用持不同意见的少数人所能起到的减少群体盲思的作用。

结 论

本章的主题是简单的：由于人类本质上具有社会性，因此他们的判断和决策很容易受到社会因素的影响。即使在独自决策的时候，他们也常常根据他人会做何评价的预期来决定自己的行为。因此，任何时候要对人们的决策和判断作出全面的解释，都必须把社会因素考虑进去。

在某些情况下，这些因素虽然能影响决策，但却没有影响决策背后的判断。例如，很多人会迫于从众压力和群体盲思而做出某个选择，但他们自己本身并不赞同这个选择。在另外一些情况下，判断本身也被歪曲了。例如阿施发现，在他的实验中有少数被试的知觉的确被歪曲了，而且他们并没有意识到多数人的意见影响了自己的判断。类似的，拉坦和达利也发现很多没有做出反应的旁观者，在责任分散的条件下也改变了自己对情境的知觉。正如拉坦和达利（1970, p. 65）所说的：“（这些）被试几乎无一例外地都说如果遇到‘真正’的紧急情况，他们一定会第一个去帮助受害者。”

本章所讨论的社会影响因素只是支配着我们日常生活的很多社会因素中的少数几个。这里只是举例说明，并不是要穷尽所有影响因素。而且，就像在讨论社会性懈怠效应时所显示的那样，社会影响因素既可能妨碍决策与判断，也可能改善决策与判断。第18章将通过比较群体表现和个体表现，继续就这个问题进行讨论。

群体决策与判断

如果一个群体成功地避免了群体盲思，那么是否就意味着它能比个体做出更好的决策和判断？群体也会和个体一样使用直觉并出现偏差吗？

相对来说，在个体和群体水平上对直觉和偏差进行比较的研究较少，所以没有综述对以上两者加以对比。然而，在已发表的研究中，多数都显示个体和群体在归因和判断时产生的偏差是类似的。本章的第一部分讨论了几种类似情况，接下来的第二部分集中阐述群体和个体的差异。

群体误差和偏差

如第 16 章所述，基本归因误差是过度地将个体的行为归因于人格或态度等内在特质上。例如，人们通常认为评论反映了作者本人的观点，即使在他们被告知作者是被迫采取这一见解时也是如此（Jones & Harris, 1967）。群体归因误差（group attribution error）是对一个群体做出无根据的内在归因的倾向。首先对群体归因误差这一概念进行阐述的是斯科特·艾利森和戴维·梅西克（Scott Allison & David Messick, 1985, p. 564），他们认为，“基本归因误差和群体归因误差有一个共同的基础，那就是我们倾向于忽略行为的外部决定因素，并且想当然地认为行为和一些内在态度之间存在关联。”他们在研究中发现，被试推断，当一位城市管理者制定出节水措施时，公众会比没有采取措施时更加支持节约用水，尽管几乎没有什么理由可以认为公众立场可被一位政府雇员的决定所影响。黛尼·麦凯（Diane Mackie）和斯科特·艾利森（1987）也报告了类似的发现。

“自我服务”偏差和“群体服务”偏差之间也存在相似的归因。前者指人们把自己的成功归因于内部因素，而把自己的失败归因于外部因素。后者指群体成员把群体的成功归因于群体的内部因素，而把群体的失败归因于群体的外部因素。唐纳德·泰勒和珍妮特·多里亚（Donald Taylor & Janet Doria, 1981）比较了校际运动员的自我服务偏差和群体服务偏差，发现群体服务偏差至少与自我服务偏差程度相当。运动员的群体服务偏差表现为，他们更多地将群体的成功归因于出色的群体合作，而较少将失败归因于糟糕的群体合作。

另一个群体水平的归因偏差是“外群体同质性偏差”（outgroup homogeneity

bias)。与人们总是认为自己比别人更具多样性 (more varied) 一样, 群体也认为他们自己的成员比其他群体的成员更具多样性 (Mullen & Hu, 1989)。例如, 在一个实验中, 普林斯顿大学 4 个学生俱乐部的成员为他们所在群体的成员和其他 3 个群体的成员打分, 评定他们的人格特征, 例如内向/外向、傲慢/谦逊、悠闲/紧张等等 (Jones, Wood, & Quattrone, 1981)。结果表明, 无论哪个群体的学生都把自己群体的成员评价得比其他群体成员更加多样化。

你可能会说: “他们当然会认为自己的群体更加多样化——因为在自己所在的俱乐部里, 他们认识更多的成员。”但研究表明, 外群体同质性偏差与学生认识的群体内和群体外成员数量无关。而且, 外群体同质性偏差也发生在那些互相之间有大量接触的群体之间, 比如女性和男性之间 (Park & Rothbart, 1982)。所以, 外群体同质性偏差并不仅仅是熟悉程度不同的结果。

外群体同质性偏差的一个重要结果是加深了刻板印象。由于外群体成员被认为是相对同质的, 所以他们的个体性往往被低估了。不幸的是, 这种倾向在国际关系中尤为普遍。例如, 前苏联领导人安德罗波夫上台时, 前美国国家安全顾问布热津斯基写道: “把这些人分为保守派或自由派、鹰派或鸽派、斯大林主义者或非斯大林主义者都是错误的, 关键在于他们都是野蛮而且粗暴的” (Schmemmann, 1985, March 3, p. 55)。不久以后所发生的事件就证明了布热津斯基犯了多么严重的错误。

很少有研究探讨像可得性和代表性直觉这样的直觉是否也在群体判断中起作用。而在琳达·阿尔戈特等人 (Linda Argote, Mark Seabright, & Linda Dyer, 1986) 的研究中, 个体被试和由五人组成的群体解决律师—工程师问题 (见第 10 章), 研究者发现, 群体比个体更加依赖代表性直觉。斯科特·廷代尔等人 (Scott Tindale, Susan Sheffey, & Joseph Filkins, 1990) 也发现, 当面对类似于读者调查中第 1 题这样的问题时, 四人群体比个体更容易犯连词错误 (使用代表性规则的另一个结果)。

所有上述结果意味着什么呢? 显然, 尽管我们还需要做更多的研究, 但这些初步的发现表明, 个体水平的直觉和偏差在群体决策和判断中仍然起作用。而且, 有时在群体中会出现比个体还大的偏差。以代表性直觉为例, 阿戈尔特等人 (Argote, Seabright, & Dyer, 1986, p. 74) 总结道: “当评定范畴内的全体成员时, 群体讨论似乎扩大了个体代表性判断的倾向。”

群体极化

群体讨论容易增强群体成员达成一致的倾向, 这种现象称为“群体极化” (group polarization) (Moscovici & Zavalloni, 1969)。首先指出这一现象的研究者是詹姆斯·斯托纳 (James Stoner, 1961), 他发现, 被试在参加群体讨论后, 似乎更愿意提倡和拥护冒险的行动了。斯托纳将这一变化称为“风险转移”

(risky shift)。在他的实验之后，很多研究对此进行了探索。毕竟，如果群体极化使得群体的冒险性比个体更高，那么这一发现对国家安全、商业和其他由群体做出重大决策的领域都具有深远的意义。

大概有 4/5 的风险转移实验是这样进行的：首先，被试要完成一份问卷，用于评估他们的冒险倾向。问卷通常选用由内森·科根和迈克尔·沃勒克 (Nathan Kogan & Michael Wallach, 1964) 编制的“两难选择”问题，问卷描述了 12 个模拟情境，情境中的主人公需要在冒险行为和保守行为之间做出选择。例如，在一道题目中，一位 45 岁的会计师要决定是冒心脏外科手术的风险还是让心脏的病情继续恶化（见图 18.1）。

在两难选择情境中，被试需要决定当成功概率达到多少时，他们将建议主人公选择冒险行为。这次测量称为前测。然后，大约 5 名被试组成一个群体，讨论这 12 个情境，并就每个情境的概率选择达成一致意见。这一阶段称为操作测验。操作测验的平均分与前测平均分的差值就是风险转移的测量值。

虽然这一范式有很多变体，但研究结果相当一致。群体讨论通常使人们提倡更冒险的行为。但当最初的倾向是保守的时，群体讨论有时就可能引起“保守转移” (cautious shift)。因此，很多研究者现在采用“选择转移” (choice shift) 这一更加概括的词汇来描述这个现象。

与选择转移研究一样，群体极化研究也延伸到了冒险以外的问题。例如，戴维·迈尔斯和乔治·毕晓普 (David Myers & George Bishop, 1970) 发现，对种族问题抱有偏见的学生互相讨论了这一话题之后，偏见变得更为强烈；而相对来说对此没有太多偏见的学生相互讨论以后偏见更少了。类似的，迈尔斯

B先生是一位45岁的会计师，最近医生发现他患上了一种严重的心脏病。这种疾病迫使B先生改变很多重要的生活习惯——减少工作压力、彻底改变饮食、放弃最喜欢的休闲嗜好。医生建议他尝试外科手术，如果手术成功，就可以完全治愈心脏病，但手术可能失败。

假定你为B先生提供建议。手术成功的概率值有如下几种。

请选择你能接受的最低的手术成功率，并在相应横线上打勾。

____ 无论手术成功的概率多大，B先生都不应接受手术。

____ 手术成功的概率是90%。

____ 手术成功的概率是70%。

____ 手术成功的概率是50%。

____ 手术成功的概率是30%。

____ 手术成功的概率是10%。

图 18.1 内森·科根和迈克尔·沃勒克 (1964, pp. 256 - 257) 编制的两难选择问卷中的一道题目。

(1975) 发现,“沙文主义者”和“女权主义者”在各自群体内讨论了女性问题以后,两者间的分歧增大了。这类研究基于“群体构成”(group composition) 范式,即群体由具有相同倾向的人构成(这一倾向在群体讨论后得到增强)。

另一类群体极化研究使用这样的范式:问题本身产生了一种倾向,群体讨论使这一倾向极端化。例如,戴维·迈尔斯和马丁·卡普兰(Martin Kaplan, 1976) 发现,当模拟陪审团仅得到较弱的犯罪证据时,群体讨论使判决更轻;而当陪审团得到确凿的证据时,群体讨论后判罪更重。可见,基于法律证据的初始倾向在经过群体讨论后极端化了(关于群体极化及其原因的综述见 Myers, 1982; Myers & Lamm, 1976)。

卖马问题

“有人以 60 美元的价格买了一匹马,并以 70 美元卖出。然后,他又花 80 美元钱买回了这匹马,再以 90 美元卖出。在这匹马的交易中,他共赚了多少钱?”

读者调查中的第 20 题最早出现在诺曼·梅尔和艾伦·索利姆(Norman Maier & Allen Solem, 1952) 的经典研究中。梅尔和索利姆发现,独自解答这个问题时,只有 45% 的大学生被试能够得出正确的答案。但当学生们组成 5 人或 6 人小组时,这一比率有所提高。在领导者不活跃的群体中(领导者仅是观察成员们的讨论),回答的正确率是 72%;而在鼓励型领导者的群体中(领导者鼓励所有成员表达想法),正确率是 84%。

为了了解一个有效的群体可能如何解决这一问题,让我们来听听发生在梅尔和索利姆实验室中的一次假想的群体讨论吧:

温迪 [群体领导]:我想这个问题可能有些迷惑性——也许不像表面上这么简单,而且我不信任心理学家。我们每个人都来讲讲想到的答案吧,再说说为什么这个答案是对的。

贝内特:我非常肯定答案是 10 美元。我解决这个问题的方法是假定这是股票交易。如果我以 60 美元一股的价格买进并以 70 美元的价格卖出,那我就赚了 10 美元。然后,如果我改变了主意,以 80 美元一股的价格买进了同一种股票,我就比卖出时的 70 美元多花了 10 美元,把先前赚的钱亏掉了。再以 90 美元的价格售出让我又赚了 10 美元,所以我得到这个答案。

吉尔:我的答案是 20 美元,因为这个人第一次卖马赚了 10 美元,第二次又赚了 10 美元。但听了贝内特的解释,我不敢肯定我的答案是对的。

温迪:不必担心你是不是对的——我们只是希望听到所有可能的答案。你呢,

史蒂文？

史蒂文：我认为正确的答案是30美元。他开始时花60美元买进，最后以90美元卖出，那他的利润不是30美元还会是多少呢？

埃米：如果他不需要借钱的话就确实赚了30美元，但他卖马得到70美元以后，要花80美元买回这匹马，就需要额外的10美元。这样剩下了20美元利润——30美元减去他额外借的10美元。

温迪：这么说你同意吉尔的答案？

埃米：是的。实际上，我想这个问题之所以让人困惑，原因就在于两笔交易用的是同一匹马。如果这道题里有两匹马，答案就很明显了。

贝内特：你的意思是？

埃米：是这样，假设他以60美元买进了A马，又以70美元卖出，他赚了多少钱？

贝内特：10美元。

埃米：现在假设他又买了第二匹马——B马——花了80美元，再以90美元的价格卖出，他赚了多少钱？

贝内特：又是10美元。

埃米：对，所以他最后一共赚了20美元。他的利润与交易的是一匹马还是两匹马无关——无论怎样他都赚了20美元。如果你喜欢用股票交易来想这个问题，那就假定交易的是两种股票而不是一种。不应该认为买进第二种股票花80美元意味着亏掉10美元——卖出比买进的价钱低才是亏了钱。

我们先听到这儿，现在请想像你就是这个群体的一名成员。埃米总结了她的发言以后，轮到你解释你的答案。你发言结束后，整个小组花几分钟来讨论这个问题。然后实验者询问每一名小组成员并记录下最终答案。你可以选择是保留做读者调查时的答案还是改变答案。你会怎样做呢？

大多数人在这样的情况下都会给出正确的答案——20美元。那些在一开始就得出正确答案的成员在后来的讨论中很少改变想法，而其他人通常在群体讨论后改正了答案。虽然梅尔和索利姆（1952）发现，无论群体领导者是否活跃和具有鼓励性，群体表现都胜过个体，但当群体领导者积极地鼓励所有成员与大家分享他们的见解时，正确率是最高的。

当讨论开始时群体里只有一名成员持有正确答案时，鼓励型领导者的作用是最明显的。在这样的情况下，群体中有76%的学生回答正确，而不活跃群体中答对的比率仅为36%。这一发现与上一章的两个结论一致：（1）避免群体盲思的最好方法是明确地鼓励成员发表不同见解（Janis, 1982），以及（2）在这一见解正确的情况下，少数人可显著地提升群体判断的正确率（Nemeth, 1986）。

梅尔和索利姆(1952)的研究表明,开放性的讨论可大幅度提高正确率。但在他们的实验中,如果群体中没有一个人的答案是正确的,那么群体讨论也无济于事。开始时,在参加实验的67个群体中,有63个群体内至少有一名成员持有正确答案,但大概有五分之一被试在群体讨论后仍然回答错误。埃德温·托马斯和克林顿·芬克(Edwin Thomas & Clinton Fink, 1961)用同样的卖马问题做实验也得到了相似的结果。在他们的研究中,群体讨论前,44个群体中有29个群体至少有一位成员持有正确答案,但只有15个群体在讨论后达到了全体正确。所以,群体讨论可以使正确率显著提高,但并不能保证所有群体成员都回答正确。

人多智慧多吗

群体判断在某种程度上比个体判断更为准确,尽管并非一贯如此。群体判断的准确性取决于很多因素,包括任务的性质和难度、群体成员的能力、成员间的互动方式等等。里德·黑斯蒂(Reid Hastie, 1986)发表了一篇综述,对很多影响群体判断的因素进行了总结,并比较了群体和个体在3种不同类型判断任务中的表现:(1)判断数量和大小,例如罐中豆粒的数量;(2)判断逻辑问题,例如卖马问题;(3)判断一般知识性问题,例如“苦艾是一种酒还是一种宝石?”

黑斯蒂总结出,对于数量判断,群体通常比个体稍微准确一些(黑斯蒂估计这一差别大约是 $1/8$ 个标准差)。珍妮特·斯尼泽克和贝姬·亨利(Janet Sniezek & Becky Henry, 1989, 1990)的研究显示,在某些情况下,群体判断的优势可能比黑斯蒂的估计数要大。他们运用了“标准化偏差”的测量方法,发现3人群体比个体的准确性高23%~32%——是黑斯蒂总结的研究中差别的2~3倍。

对于逻辑问题,黑斯蒂发现,群体通常比个体表现更好,但是群体中最优秀的成员独自解决问题的结果往往好于群体整体。同样,他总结出,在回答一般知识性问题时,群体通常比水平中等的成员表现更好,但群体中最优秀的成员往往超过群体的成绩或与之相当。通过总结这3种判断任务,黑斯蒂发现,群体比其水平居中的成员做出的判断更为准确,但群体中最好的成员经常胜过群体整体。

盖尔·希尔(Gayle Hill, 1982)在一篇题为《群体与个体表现: $N+1$ 是否大于1?》的综述中,也得出了类似的结论。在这篇文章里,希尔总结了50年来关于个体和群体表现的研究。她不仅讨论了二者在判断与决策上的表现,范围还涉及创造力和问题解决等方面。她的结论是,群体通常胜过其中一般的个体,但是群体经常不如其中最优秀的个体。希尔发现,对于容易的任务,增加群体的人数只能增加一个人能独立解决问题的可能性;对于困难的任

体合作的主要优势在于，群体成员能够汇聚所有人的资源并纠正他人的错误。

希尔还讨论了头脑风暴法。她比较了群体讨论中产生的想法和同样多的人独自进行头脑风暴得到想法的总和。希尔发现，独自产生想法后汇总要比群体会议更加有效。^{*}她总结道，群体讨论的优越性主要在于汇集的功能（也就是仅仅让更多的人来解决问题），而不是实质上的群体互动。从实际的角度来看，这意味着解决难题的最好方法是让几个人独立思考问题然后汇集他们的想法。

独裁的益处

虽然群体判断往往比个体判断准确，但这种准确性部分地取决于群体成员如何整合他们的答案（Davis, 1973）。斯尼泽克（1989）对此进行了探讨，并发表了相关的研究报告，斯尼泽克比较了5种群体决策技术：（1）“一致意见”（consensus），即面对面的讨论，直到形成所有成员都接受的意见；（2）“辩证”技术，即要求群体成员讨论可能导致判断偏差的因素；（3）“独裁者”技术（也称最佳成员技术），即面对面的讨论选择出一位成员，他的判断代表了整个群体；（4）“德尔菲”技术，即群体成员不直接见面，而是以某种顺序循环匿名提供答案，直到形成一致意见或稳定看法（这一技术的优势是避免了个别成员或对自己的判断过于自信的成员垄断讨论）；（5）“集体”技术，即禁止群体成员有任何形式的互动，只是把成员们的判断平均化，得到“群体”判断（在斯尼泽克的研究中，集体技术仅是缺乏社会互动的简单汇总，因而其结果可作为准确性的基线水平）。

每个群体由5名大学生组成，而且每个群体使用全部5种决策技术。从集体技术开始，然后以不同顺序采用其他决策技术。学生们的判断任务是估计一个校园商店下个月将售出多少商品，即服装、杂志、贺卡礼品和健康美容用品的营业额各是多少。斯尼泽克用与实际销售额的“绝对偏差”，来衡量判断准确性。

她发现，前4种技术（一致意见、辩证、独裁者、德尔菲）都比简单汇总（集体技术）的判断准确性高，其中独裁者技术的准确性最高，这一技术降低的绝对偏差是其他技术的3倍。有趣的是，在每个群体中，“独裁者”最后都修改了答案使其更接近集体平均数而使偏差值提高。也就是说，群体能够选出一个判断相当准确的独裁者，但这位独裁者总是变得更加民主，结果反而降低了最终判断的准确性。

当然，斯尼泽克的发现仅是基于来自特定群体（大学生）的特定人数群体

* 尽管希尔没有明确指出这一可能性，但社会性懈怠可能部分地导致了这一结果。伊莉莎白·韦尔登和吉纳·加格诺（Elizabeth Weldon & Gina Gargano, 1985, 1988）发现，当人们一起解决复杂的判断任务时，会出现社会性懈怠现象。

(5人)解决特定的判断问题(销量预期),将独裁者技术最优的结论推广到所有情境是不适合的。但斯尼泽克的实验表明,群体判断的准确性部分取决于群体采用的决策规则。她的研究结果也证明,在某些情况下,互动的群体胜过没有互动的群体。所以,至少在有些情境中,群体的优势不仅仅在于“更多的人”而在于“更多的人在一起工作”。

结 论

由于群体表现取决于很多不同的因素,所以难以进行高度概括(Tindale, 1989)。也是出于同一原因,这些不一致的和相反的研究结果难以综合。哪些因素引发了这些矛盾?是不同的任务,群体人数,还是决策规则?丰富的研究反而使结果的解释更加复杂。

另外的问题是,群体决策和判断没有受到像个体决策和判断那样的重视,尽管委员会、专家小组、董事会、陪审团和其他决策群体极其普遍。而且,个体和群体研究之间的界限常常因群体是由个体组成的而模糊不清。例如,研究群体归因错误的实质上是探讨个体如何对群体进行归因。类似的,选择转移也用群体决策和个体决策的差别加以测量,或对群体讨论前后的个体决策进行比较。只有前一个例子才是真正的群体决策。

尽管如此,群体决策和判断的研究暂有如下结论:

- 很多个体水平的直觉和偏差在群体中发挥了同样的效力。
- 群体讨论通常会加强已有倾向。
- 群体通常比一般的个体表现好一些,尤其是在一名指定的领导者鼓励全体成员发表见解的情况下。
- 群体中最好的成员通常比群体表现更好(有时可使用独裁者决策技术来利用这一点)。
- 几个人独自使用头脑风暴法比群体会议更为有效。

上述结论都得到了大量研究的支持,但由于群体表现受多种因素影响,在特定情境中应用这些一般性的结论仍需多加考虑。尽管决策者经常通过会竭尽所能,但合作并不一定是成功的保证。

第六部分

常见陷阱

本书的这一部分关注困扰决策者的三个常见问题。第19章探讨过度自信的问题，第20章讨论自我实现的预言，而第21章分析了一种被称做“行为陷阱”的特殊情境。同时每一章还提供了如何在实践中避免出现这些问题的技巧。

过度自信

这里发生泄漏的几率是一万年都难遇的。

——乌克兰能源与电气大臣 Vitali Skylarov 在切尔诺贝利核泄漏事件发生前 2 个月这样说道
(引自 Rylsky, 1986 年 2 月)

在人们的决策中，过度自信是一个最为普遍的问题，其所带来的潜在破坏性也是最大的。正如欧文·贾尼斯（1982）在他对群体盲思所进行的研究中提到的那样，美国人的过度自信使日本人在二战期间成功偷袭了珍珠港。同样，在美国的挑战者号航天飞机失事的灾难中，过度自信也扮演了重要的角色。在挑战者号航天飞机的第 25 次发射之前，美国宇航局（NASA）的官员对飞行风险的估计是十万分之一（Feynman, 1998, 2）。这样的风险估计大致等于航天飞机在三百年的时间内每天发射，也只可能产生一次事故。

约瑟夫·基德的例子

美国宇航局真的是对成功过分自信？还是仅仅需要看上去很自信？因为在这样的情境中，自信的程度是很难测量的，因此有关过度自信的最具有说服力的证据来自严格控制的实验中。

最早也是最著名的三个实验是由斯图尔特·奥斯坎普（Stuart Oskamp）在 1965 年完成的。奥斯坎普要求 8 位临床心理学家、18 位心理学专业的研究生以及 6 位心理学专业的本科生阅读一个有关“约瑟夫·基德”的个案研究。约瑟夫·基德是一位 29 岁的男子，曾经经历过“青春期失调综合症”。这个个案被分成四个部分。第一部分介绍了他是一个退伍的老兵，现在在一个花艺工作室做商业助理，第二部分介绍了基德 12 岁以前的童年时代，第三部分介绍了基德的中学和大学生涯，第四部分介绍了他服役的经历以及往后的生活。

参与实验的人在读完每一段故事之后都要回答一系列相同的问题。这些问题都来自个案，但是要求被试基于这样的信息对基德的人格形成一个临床判断。问题通常有五个迫选的备选答案，并且在每个项目之后，被试要估计他们的答案在多大程度上是正确的。这些信心的评分从 20%（在准确度的机遇水平

之外，毫无信心)到100% (完全确信)。

有些令人惊奇的是，心理学家、研究生和本科生之间的评分并没有显著差异，所以奥斯坎普将三个群体合并到一起来对结果进行分析。他发现信心随着被试阅读的信息量的增大而增长，但是准确度却并非如此。

在阅读完个案研究的第一部分之后，被试正确回答了26%的题目(稍高于机遇水平的期望值)，并且他们的平均信心得分为33%。这些数字显示出了相当接近的一致性。然而，当被试阅读了更多的信息时，信心与准确度之间的差距增大了(见图19.1)。被试阅读的材料越多，他们就会变得越自信——即使准确度并没有随着增加的信息而显著增长。在他们读完个案研究的第四部分之后，在奥斯坎普的被试中，有超过90%的人对于他们的回答过于自信。

该实验进行之后的若干年里，已经有一系列研究发现人们往往对他们的判断过于自信，特别是当他们难于做出准确判断的时候。比如说，利希滕斯坦和菲什霍夫(1977)进行了一系列实验并发现人们对自己的正确性有65%~70%的信心，但实际上他们只有50%的时候是正确的。

在第一个实验中，利希滕斯坦和菲什霍夫要求人们判断12个儿童的图画是来自欧洲还是亚洲，并且估计每一次判断正确的可能性。尽管仅仅有53%的判断正确(非常接近随机水平)，可是平均的信心评分却达到了68%。

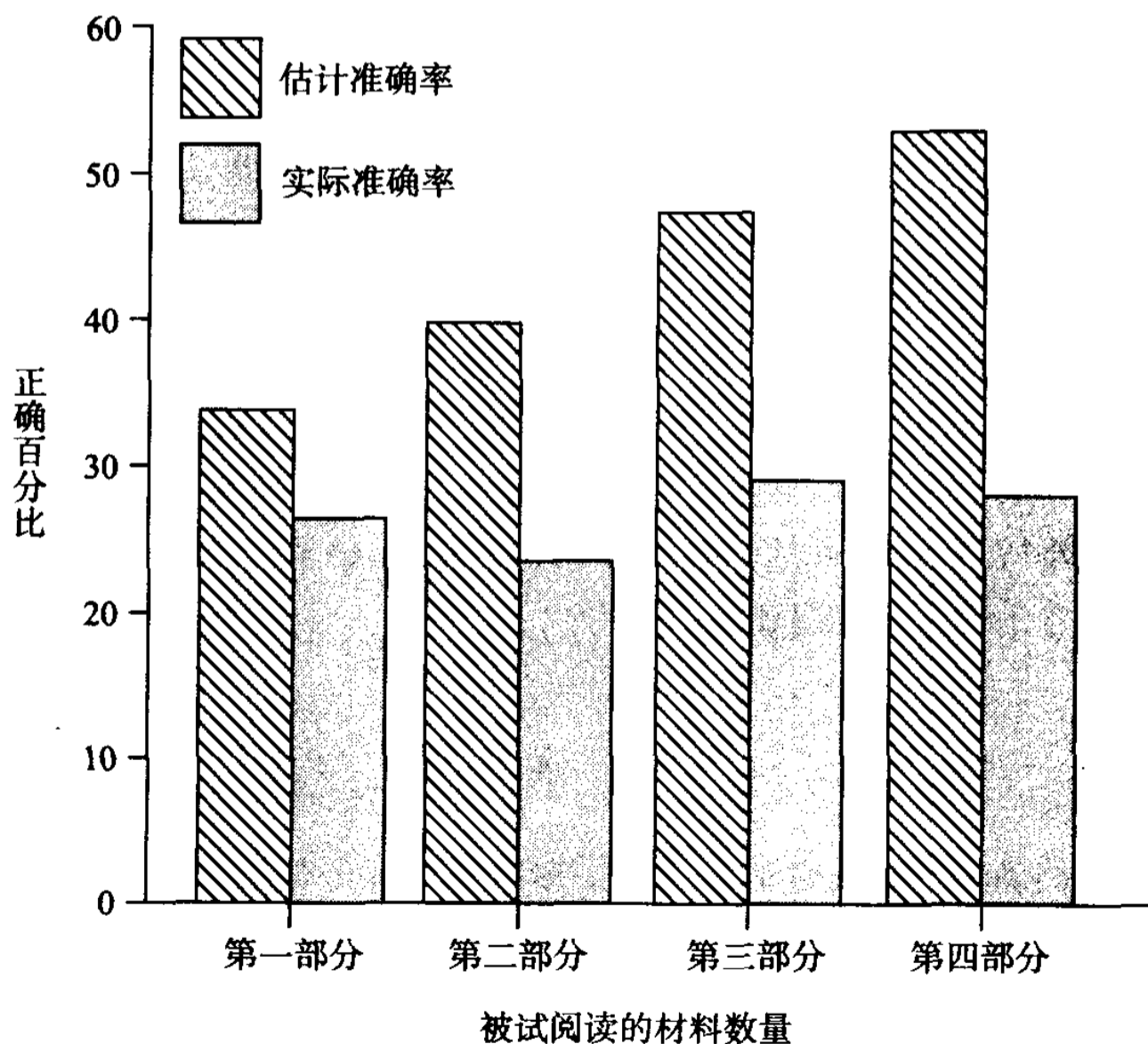


图19.1 斯图尔特·奥斯坎普(1965)发现当被试阅读了更多的信息时，他们估计的准确度(自信心)与实际的准确度之间的差异就越大。

在另一个实验中，利希滕斯坦和菲什霍夫给人们呈现了12只股票的市场报告，并请他们来预测在给定的时间段内这些股票将会看涨还是看跌。他们再一次发现，尽管这些预测只有47%是正确的（稍低于预期的机遇水平），但平均的信心评分达到了65%。

又完成了几个研究之后，对于在两可判断中准确度和信心之间的对应关系，利希滕斯坦和菲什霍夫得出了如下结论：

- 当准确度接近机遇水平时过度自信达到最大。
- 当准确度从50%增加到80%时，过度自信会随之减少；当准确度超过80%时，人们通常会变得不自信。换句话讲，在准确度达到80%左右时，准确度和信心之间的差距最小，当准确度偏离这一水平时差距会逐渐变大。
- 准确度与信心之间的分离与决策者的智商水平无关。

对于该研究的早期批评认为，产生这些结果的原因很大程度上是由于实验者总是询问人们一些晦涩或琐碎的问题，但近期的研究已经在一些更为普遍的判断上重复得到了利希滕斯坦和菲什霍夫的发现。举例来说，在一系列包含多于10 000个独立判断的实验中，李·罗斯和他的同事发现，当要求被试对他们自己的行为以及他人行为进行预测时，人们表现出了大约10%~15%的过度自信（Dunning, Griffin, Milojkovic, & Ross, 1990; Vallone, Griffin, Lin, & Ross, 1990）。

这并不是说人们总是过度自信的。戴维·罗尼斯和弗兰克·耶茨（David Ronis & Frank Yates, 1987）发现，过度自信部分取决于信心的评分如何得出以及决策者做出了何种类型的判断（一般知识性题目似乎会导致相对较高程度的过度自信）。同样有一些证据表明，桥牌高手、专业赌徒以及国家气象台的预报员——他们在做出判断后都会收到有规律的反馈——表现出较少或没有表现出过度自信（Keren, 1987; Lichtenstein, Fischhoff, & Phillips, 1982; Murphy & Brown, 1984; Murphy & Winkler, 1984）。不过，研究结果表明，就大部分人而言，过度自信现象是普遍存在的。

极度自信

如果人们确信答案正确又会怎样？在这种情况下他们正确的几率有多大？1977年，菲什霍夫、斯洛维奇和利希滕斯坦实施了一系列实验来考察这个问题。在第一个实验中，被试回答了大量一般知识性问题并且估计了答案正确的可能性。比如说，他们回答了苦艾（absinthe）是一种利口酒（liqueur）还是一种宝石，并且要在一个.50~1.00的量表上估计信心程度（这个问题作为读者调查中的第21题出现）。然后，实验者检验了那些被试们完全确信是正确回答

的准确度。

他们发现，当被试报告百分之百确信自己的答案时，他们的正确率往往只有70%~85%。你对自己在第21题上的回答有多大信心？正确答案为苦艾是一种利口酒，然而很多人将它与一种被称做紫水晶（amethyst）的宝石相混淆。

为了确保该结果不是由于对概率的误解所致，菲什霍夫、斯洛维奇和利希滕斯坦（1977）再次进行了一个实验，这一次，信心大小取决于回答正确的可能性与不正确的可能性之间的比率（比如2:1, 3:1等）。这样，被试可以采用3:1的正确可能性来表达自己的信心，而无需借助概率性的75%。

菲什霍夫、斯洛维奇和利希滕斯坦（1977）发现，当信心估计大约为3:1时，信心与准确度相当一致，可当信心从3:1提高到100:1时，准确度却没有明显提高。当人们将正确的比率设为100:1时，他们实际上只有73%的时间是正确的。即使当人们将正确的比率设为10 000:1~1 000 000:1时——表示完全确信——他们也仅有85%~90%的正确率（本应该相应给出6:1~9:1的信心评分）。*

最后，菲什霍夫、斯洛维奇和利希滕斯坦（1977）将实验重复了三次，将其作为一项附加检验以确保被试理解并且认真执行了任务。在一次重复中，比率和概率之间的关系在20分钟的发言中被加以详细解释。被试们得到一张说明各种比率估计与概率之间对应关系的图表，并且他们被告知使用比率评分来表达不确定性的微妙之处（特别强调了如何使用1:1和2:1之间的比率来表达不确定性）。然而，即便如此，被试还是对他们的答案表现出了毫无根据的信心。当实际比率大约为4:1的时候，他们至少给出了50:1的比率；当他们应该给出5:1的比率时，他们却给出了1 000:1的比率。

在另一次重复实验中，研究者询问被试是否会接受一项金钱上的赌博，这项赌博基于被试认为正确比率为50:1或更高的那些题目答案的准确度。在42名被试中，有39名愿意参与这项赌博——尽管他们的过度自信将导致总共140多美元的损失。在最后一次重复实验中，菲什霍夫、斯洛维奇和利希滕斯坦（1977）实际是让被试们进行了一次赌博。在该研究中，19名被试中的13名同意以他们答案的准确度来进行赌博，然而在他们认为正确比率为50:1或更高的那些问题中，12%的回答是错误的（并且如果实验者不加以免除，所有人都会损失1美元至11美元不等）。这些结果表明：（1）当人们真正确信自己是正确的时候他们仍然是过度自信的，（2）并不完全是由于漫不经心地执行任务或错误理解如何评价信心而导致过度自信。事实上，琼·西伯（Joan Sieber, 1974）发现，希望表现出色的动机越强烈，过度自信也会随之而增长。

* 尽管这些结果看起来可能与利希滕斯坦和菲什霍夫早期的结论（被试80%正确时过度自信达到最低）相互矛盾，但事实上并非如此。被试非常自信时他们的平均准确度只有70%~90%，这一事实并不意味着在70%~90%正确的时候他们总是非常自信的。

当过度自信变成一种严重的错误

当处在较之几美元更为利害攸关的时刻人们还会过度自信吗？尽管道德上的考虑明显限制了实验室中可以考察的范围，但至少有一条证据提醒我们，即使在人们的生命处在安危未定的时刻过度自信依然存在。这一证据来自对死刑的研究。

在一次对冤案的全面回顾中，雨果·贝多和迈克尔·拉德莱（Hugo Bedau & Michael Radelet, 1987）找到了350个有文件备案的实例，这些实例中无罪的被告都被判定在美国有重大的或潜在重大的罪行——即使被告被判断为“罪行在合理怀疑范围之外”。其中有五个案件，判决之前就已经及时发现了错误。但其他的被告就没有这么幸运了：67人被判处25年以下的监禁，139人被判处终身监禁（25年或更长），并且有139人被判处死刑。在贝多和拉德莱发表评论的时候，已经有23例死刑被执行。

校 准

“校准”是指信心与准确度的匹配程度。在给定的信心水平上，并且当所有判断中准确判断的比例与判断正确的期望概率相一致时，决策者就获得了最佳的校准。换句话说，认为正确可能性为90%的判断中有90%是正确的，认为正确可能性为80%的判断中有80%是正确的，依此类推。

当个体的判断被单独考虑时，就没有什么办法来校准了。一个在读者调查中对第21题回答了“.70”的决策者如何进行好的校准？仅有的一个评估校准的可靠方法，那就是在大量判断中将准确性和信心加以比较（Lichtenstein, Fischhoff, & Phillips, 1982）。

正如有许多种方法测量信心一样，我们有好几种用以评估校准的技术。一种方法就是简单地计算平均信心得分与全部准确判断所占比例之间的差异。举例来说，决策者可能在一组一般知识条目上具有平均80%的信心，但是他可能只有60%的回答是正确的。这样一位决策者就是过度自信了20个百分点。

尽管这种测量校准的方法非常方便，但它有时候却可能导致错误，比如，考虑一名整体准确率和平均信心得分都是80%的决策者。这个人是否达到了完美的校准呢？并非如此。这个人可能在一半题目的判断上抱有60%的信心而在其他题目上抱有100%的信心（平均起来达到了80%的信心），并且在两种信心水平上都达到了80%的准确度。这样的人在60%确定的时候信心不足，而在100%确定的时候又显得过于自信。

一个更为简洁的方法是在不同自信水平下检验准确性。当对准确度在不同的自信水平下进行单独计算时，可以创建一条“校准曲线”，其横轴代表信心，纵轴代表准确度。图19.2包含了两条校准曲线——一条表示天气预报员关于

降雨量的预测，另一条表示医生对于肺炎的诊断。正如你所看到的，天气预报员几乎完美地实现了校准；平均看来他们的预测非常接近实际的天气（与人们通常的观点恰恰相反！）。而内科医生的校准程度则很糟糕；他们的大多数预测都在对角线以下，说明他们过度自信。

还有其他的方法来评估校准，其中一些借助了复杂的数学工具。举例来说，最常用的方法之一是计算一个叫做“Brier 分数”（以统计学家 Glenn Brier 的名字命名）的数字。Brier 分数可以被分割为三个成分，其中之一就对应于校准。Brier 分数中的校准成分是每个类目下正确率和与该类目存在关联的概率之间的均方差的加权平均值（关于该技术中校准方面的详细介绍，见 Yates, 1990）。

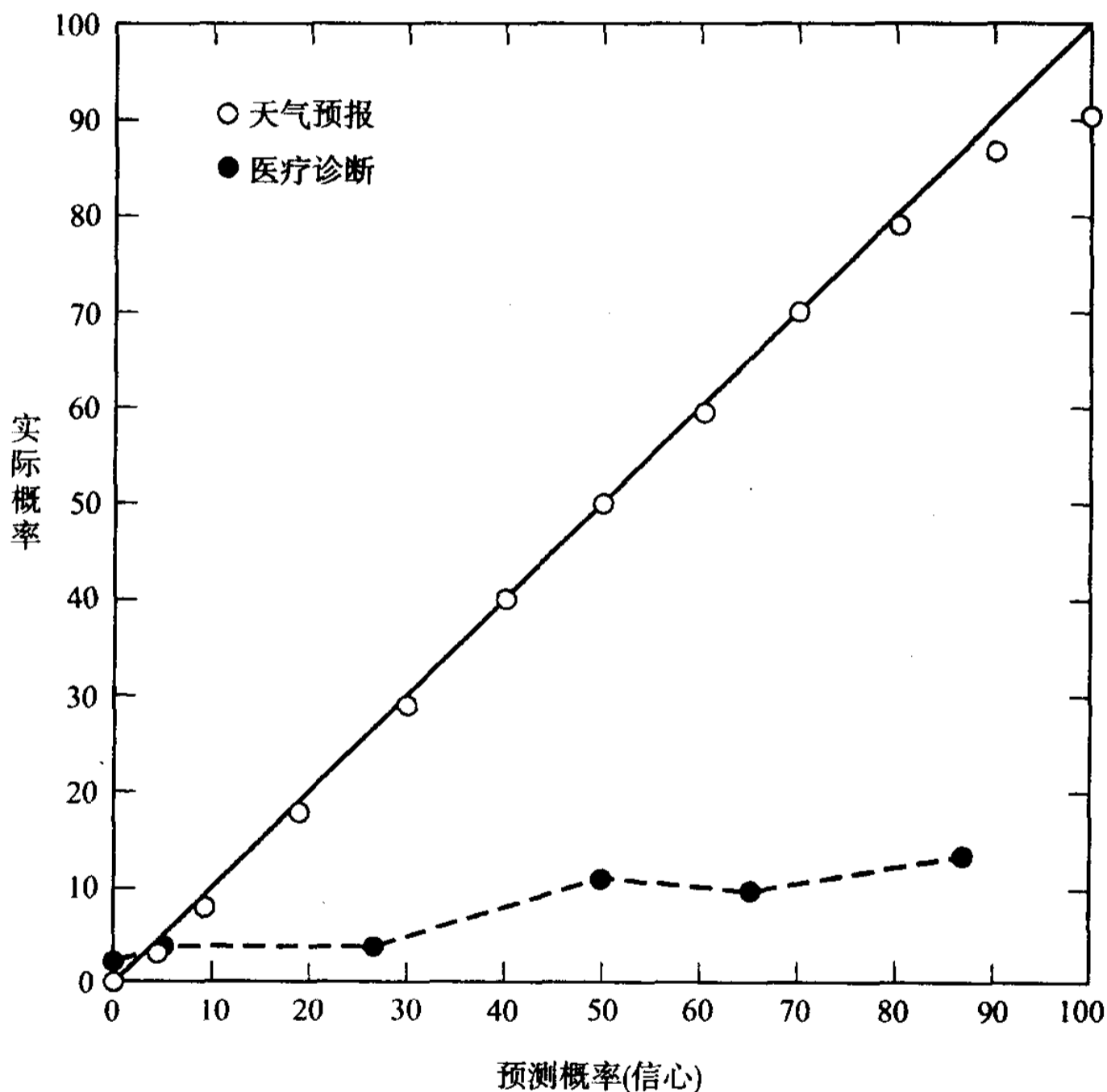


图 19.2 本图包含了天气预报员关于降雨量的预测矫正曲线（空心圆点），以及医生对于肺炎的诊断的矫正曲线（实心圆点）。天气预报员几乎完美地实现了校准，但是内科医生的校准表现出了很明显的过度自信现象（毫无根据地认为病人得了肺炎）。天气预报员的数据来自 Allan Murphy 和 Robert Winkler (1984) 的一份报告，内科医生的数据来自于 Jay Christensen-Szalanski 和 James Bushyhead (1981) 的一个研究。

一种最有意思的校准测量方法被称做“惊奇指数” (surprise index)。惊奇指数被用来对未知数量进行区间判断。举例来说,假定你对读者调查第 12 题的回答是在一英尺和一英里之间,并对此抱有 90% 的信心 (见读者调查第 12b 题来了解你真正的 90% 置信区间)。因为正确答案实际上应当大于一英里,所以这个答案被记做“一个惊奇”。惊奇指数其实就是置信区间之外的判断所占的百分比。

在一篇关于校准研究的主要论述中,利希滕斯坦、菲什霍夫和劳伦斯·菲利普斯 (1982) 进行了几项研究,这些研究要求被试给出 98% 的置信区间 (如,一些区间有 98% 的机会包含正确回答)。每个研究中,惊奇指数都超过了 2%。将所有信息可用的实验数据加以平均——总共接近 15 000 次判断——惊奇指数达到了 32%。换句话讲,当被试 98% 确信一个区间包含了正确答案时,他们做到了 68% 正确。过度自信再一次被证明是一种规律而非例外。

你过度自信吗? 爱德华·拉索和保罗·休梅克 (1989) 开发了一套自测题目来测量一般知识问题上的过度自信 (见图 19.3)。尽管一次全面的校准评估需要大量的判断,但该测验能够给你提供一个粗略的反馈,使你了解自己在某个信心水平上对一般知识问题的惊奇指数。拉索和休梅克对 1 000 多人实施了

过度自信的自我测量

对于以下的 10 个题目,请你给出一个较高的估计和一个较低的估计,使得正确答案落在你的区间中的概率在 90% 以上。其中对于你们的挑战是,你们所设定的区间既不能过窄 (过度自信),也不能过宽 (缺乏自信)。如果你只有 10% 的错误率——也就是只答错一题,你就获得了成功。

	90% 的信心区间	
	低	高
1. 马丁·路德·金去世时的年龄	_____	_____
2. 尼罗河的长度	_____	_____
3. OPEC (石油输出国组织) 成员国的数量	_____	_____
4. 基督教《旧约全书》的册数	_____	_____
5. 月球直径的英里数	_____	_____
6. 一架波音 747 飞机的重量 (磅)	_____	_____
7. 莫扎特是哪一年出生的	_____	_____
8. 一头亚洲象妊娠的时间 (天)	_____	_____
9. 伦敦到东京的飞行距离	_____	_____
10. (已知的) 最深的海洋深度 (英尺)	_____	_____
答案: (1) 39 岁; (2) 4 187 英里; (3) 13 个国家; (4) 39 册; (5) 2 160 英里; (6) 390 000 磅; (7) 1756 年; (8) 645 天; (9) 5 959 英里; (10) 36 198 英尺		

图 19.3 本测验将使你们了解到你们对于一般性知识是否存在过度自信的情况 (Reprinted with permission from Rosso & Schoemaker, 1989)。

该测验，发现只有不超过1%的人正确回答了9道或更多的题目。大多数人答错了4~7道题目（惊奇指数为40%~70%），反映出了过度自信的基本程度。

信心与准确度之间的相关

尽管会过度自信，但信心与准确度存在相关还是有可能的。举一个例子，假定一名决策者在70%自信的情况下具有50%的准确度，在80%自信的情况下具有60%的准确度，在90%自信的情况下具有70%的准确度。在这样一个例子中，信心与准确度之间存在着必然的联系，尽管这个决策者一直表现出20%的过度自信。

那么问题出现了，信心与准确度之间存在相关吗——如果不考虑决策者是否过度自信的话。如果当准确度提高时信心评分也增加的话，准确度就能够通过决策者的自信程度而得以预测。如果不是这样，信心就是一个准确度的错误指标。

很多研究都已经考察了这一问题，并且结果往往显示信心与准确度之间并不存在什么关系。为了加以说明，考虑下面两个关于军队历史的问题：

问题1：一个距离超级强国A并不太远的国家政府，在讨论过其政党体系的某些变化之后，开始拓展其与超级强国B之间的贸易。为了扭转该国政府及其贸易中的变化，超级强国A向该国派出了军队并且武装支持原来的政府。谁是超级强国A——美国还是前苏联？你对于自己答案的正确性有多大信心？

问题2：20世纪60年代，超级强国A对其边境的一个小国发动了一场突然袭击，目的是推翻当时掌权的政府。这次侵略行动以失败告终，并且多数侵略士兵被击毙或者囚禁。谁是超级强国A？并且再次回答，你对自己的答案有多确定？

这些问题在读者调查中作为第9题和第10题出现。如果在第一个问题中你猜是前苏联，并且在第二个问题中选择了美国，那么两道题目你都回答正确了。第一个问题描述的是1968年前苏联入侵捷克斯洛伐克，而第二个问题描述的是美国入侵古巴的猪湾。大多数人至少答错了其中一个问题，不管他们感到有多么自信。

在《今日心理学》（*Psychology Today*）1984年11月那一期里，菲利普·津巴多和普劳斯发表了包括这两个问题以及各种有关强权国家冲突在内的一份读者调查的结果。该调查包含了10个关于美国及前苏联的事件、声明或政策方面的描述，而在每个描述中，所有能够识别出美国和前苏联的标志都被删除了。读者的任务就是判断超级强国A是美国还是前苏联，并在一个9点量表上标出他们对每一个回答的信心有多大。

基于对 3 500 人的调查，我们能够得出两点结论。首先，答题者并不能够将美国和前苏联的军事行动区分开。即使他们仅仅通过抛硬币就可以平均正确回答 10 道题目中的 5 道，《今日心理学》的读者们（比一般公众更多地参与政治并接受过更好的教育）的总体平均值为 4.9 道题目正确。仅有 54% 的答题者正确地指出前苏联是入侵捷克斯洛伐克的超级强国 A，并且 25% 的答题者误认为是前苏联而非美国发动了猪湾侵略。这些发现说明美国人之所以谴责前苏联的行动和政策，很大程度上是因为他们是“苏维埃”，而不是因为他们与美国的行动和政策存在根本的区别。

我们发现的第二个结论是：人们的信心评分实际上与他们的准确度并无联系（每一名答题者信心与准确度的平均相关仅有 .08，接近零）。总体而言，答对了 9 或 10 道题的人并不比那些成绩较差的答题者具备更强的信心，并且高信心的答题者与低信心的答题者得分基本相同。

这并不意味着信心评分是随机得到的；高信心的答题者在一些方面还是表现出了与众不同之处。即使一般性样本中两性比例十分均衡，在全部的高信心答题者（比如那些在 9 点的信心量表上平均分达到 8 以上的人）中，却有 2/3 是男性，而且 80% 的年龄在 30 岁以上。高信心答题者增加国防开支的意愿是低信心答题者的两倍，同时高信心答题者对前苏联政府的不信任程度也是低信心答题者的两倍。这些答题者在调查中达到的平均成绩为 5.1 道题目正确——几乎就是机遇反应的期望值。因此，高信心的答题者难以将美国和前苏联的军事行动区分开来，但是他们对于这种错误知觉的区分非常自信并倡导增加国防开支。

正如前面所谈及的，许多其他的研究发现了在信心与准确度之间只有很小的相关或根本没有相关（Paese & Sniezek, 1991; Ryback, 1967; Sniezek & Henry, 1989, 1990; Sniezek, Paese, & Switzer, 1990）。这种一般性的模式在关于目击者证词的研究中表现得更为突出。总体而言，这些研究表明目击者对其证词所持的信心与证词的实际准确程度并无太大联系（Brown, Deffenbacher, & Sturgill, 1977; Clifford & Scott, 1978; Leippe, Wells, & Ostrom, 1978）。肯尼思·德芬巴赫（Kenneth Deffenbacher, 1980）曾经发表了一篇包含 43 个独立研究结果的综述文章，这些研究考察了在目击证人身上存在的准确度和信心之间的关系，他发现在 2/3 的“法庭相关”研究（如，在这些研究中，被试在观看一起事先策划好的模拟犯罪之前并未得到任何说明）中，信心与准确度之间存在并不显著的正性相关。这些研究发现使得《目击者证词》的作者伊莉莎白·洛夫特斯（Elizabeth Loftus, 1979, p. 101）警告说：“人们不能过于自信地为任何事情提供任何绝对的担保。”

在临床研究中也发现了类似的结果。在最初一项探讨该话题的实验中，路易斯·戈德堡（Lewis Goldberg, 1959）在临床诊断中评估了信心和准确度之间的相关性。戈德堡所感兴趣的是临床医生是否能够通过 Bender - Gestalt 测验（一

项被广泛用来诊断脑损伤的测验)来准确探测出器质性的脑损伤。他给4名有经验的临床心理学家、10名临床实习生、8名非心理学家(秘书)呈现了30个不同的测验结果。这些结果中的一半来自于脑损伤病人,另一半来自无器质性问题的精神科病人。被试要做出判断来指明每个病人是“器质性的”还是“非器质性的”,并在一个标有“断定的”、“相当确定的”、“这样认为”、“可能”或“瞎猜”的评价量表上指出他们的自信程度。

戈德堡发现了两个令人惊讶的结果。首先,全部三组判断者——有经验的临床医生,临床实习生以及非心理学家——都正确地对65%~70%的病人进行了分类。不存在基于临床经验的差别;秘书们表现得同那些具有4~10年临床经验的心理学家们一样好。其次,在个体诊断准确性和信心程度之间并没有显著的关系。判断者们在错误诊断的病例上所持有的信心通常与正确诊断的病例一样。后续研究也已经发现在癌症、肺炎的诊断过程中存在错误的校准(见图19.2),还包括其他一些严重的医学问题(Centor, Dalton, & Yates, 1984; Christensen - Szalanski & Bushyhead, 1981; Wallsten, 1981)。

如何才能减少过度自信

在两个考察如何改进校准的实验中,利希滕斯坦和菲什霍夫(1980)发现,在进行200次判断并得到集中的表现反馈之后,一开始过度自信的人能够更好地进行校准。同样,哈尔·阿克斯(Hal Arkes)和他的助手们发现在呈现五个富有迷惑性的难题之后,通过给被试提供反馈可以消除过度自信(Arkes, Christensen, Lai, & Blumer, 1987)。这些研究表明过度自信可以被消除,尽管它们的实用价值还比较有限。很少有人会因为希望能够更好地进行校准而接受特殊训练。

有用的是一项能够使决策者便于“携带”至各种决策中并加以利用的技术——重量轻、持久性强,并且易于在一系列环境下加以运用。的确,我们好像有一项这样的技术。最有效的改进校准的方法看起来非常简单:

停下来思考一下为什么你的判断可能是错误的。

阿舍·科莱特(Asher Koriat),利希滕斯坦和菲什霍夫(1980)最先证明了这项技术的价值。在他们的研究中,被试回答了两组二选一的一般知识性问题,第一组呈现控制条件的指导语,第二组呈现原因条件的指导语。在控制条件下,被试选择一个答案并估计回答正确的可能性(在.50和1.00之间)。在原因条件下,被试选择答案之前需要为每个备选答案列出支持以及反对的理由。

科莱特、利希滕斯坦,以及菲什霍夫发现在控制条件的指导语下,被试表现出了典型的过度自信,但是在列出正反两方面理由之后,他们表现出了非常

好的校准（与利希滕斯坦和菲什霍夫研究中那些被提供了集中反馈信息的被试大致相当）。在为每个备选答案列出支持以及反对的理由之后，被试变得不那么自信了，（主要是因为他们更多地使用 .50 而较少使用 1.00 了）并且准确度更高了（大概是因为他们对于自己的回答投入了更多的思考）。

在接下来的实验里，科莱特、利希滕斯坦和菲什霍夫发现，本质上并不是因为支持的理由导致了校准的改善；而是反对的理由使然。当被试列出理由支持他们所偏好的答案之后，过度自信并没有降低。而当被试思考他们偏好的答



图 19.4 从多个角度进行考虑的困难性。（Calvin and Hobbes Copyright 1990 Watterson. Dist. By Universal Press Syndicate. Reprinted with permission. All rights reserved.）

案如何可能是错误的时候，校准才会得到改善。尽管这些发现可能部分是由于“社会需要特征 (social demand characteristics)”的存在（如，被试感觉指导语暗示他们降低信心水平），但其他研究也证实了反对理由能够改善校准水平（如 Hoch, 1985）。

这些结果回顾了第3章讨论过的保罗·斯洛维奇和菲什霍夫（1977）的研究，在那个研究中，当被试思考为什么一些实验结果与他们所设想的有所不同的时候，后见之明的偏差 (hindsight biases) 减少了。从斯洛维奇和菲什霍夫的研究开始，有几个实验都已经解释了如何通过考虑其他结果或答案的可能性来减少各种判断上的偏差 (Griffin, Dunning, & Ross, 1990; Hoch, 1985; Lord, Lepper, & Preston, 1984)。

正如查尔斯·洛德，马克·莱珀和伊莉莎白·普雷斯頓 (Charles Lord, Mark Lepper, & Elizabeth Preston, 1984, p. 1239) 所指出的：“人们对于相反的可能性存在盲点，这种观察由来已久。在1620年，培根曾经写下‘肯定比否定能够激起更多的感动和兴奋是人类智力中特殊并且永恒的错误。’”在第20章，这个盲点——以及它所带来的某些后果——将会被详细探讨。

结 论

对过度自信进行研究是非常重要的。在大多数研究中，平均信心水平从没有超过准确度10~20个百分点以上。因此，只要决策者对自己判断的正确性不是非常确定，过度自信就不会成为一种灾难。与航天飞机的空间探索相似，最具破坏性的错误校准形式是不恰当的确信。

总之，本章的研究为我们提出了几条应对错误校准的策略：

- ✓ 首先，你可能会希望标记出那些需要特别考虑的判断。当判断难于做出或极端自信的时候过度自信达到最强。在这样的情况下，谨慎前行非常必要。
- ✓ 其次，你可能会希望“重新校准”你非常自信的判断以及其他人的判断。正如利希滕斯坦和菲什霍夫（1977）所观察到的，如果一个决策者抱有90%的信心而只有70%~75%的准确度，最好的办法可能就是将“90%的信心”看做“70%~75%的信心”。
- ✓ 同样，你可能会希望将具有“百分之百信心”的判断自动转变为较低程度的信心。百分之百的信心在预测人们如何行动的时候显得尤其没有根据 (Dunning, Griffin, Milojkovic, & Ross, 1990)。
- ✓ 最重要的是，如果你对某个答案感到极度自信，思考一下另外一个答案可能正确的原因。尽管你可能不会改变想法，但是你的判断却可能会得到更好的校准。

第 20 章

自我实现的预言

假如图 20.1 中的每张卡片都是一面有数字而另一面有字母，并且有人告诉你：“如果某张卡片的一面是一个元音字母，那么它的另一面就是一个偶数。”你需要翻开哪些卡片来判断这个人是不是在撒谎？（见你对读者调查第 39 题的回答。）

当彼得·沃森和菲尔·约翰森-莱尔德（Peter Wason & Phil Johnson-Laird, 1972）将这种类型的题目*呈现给 128 名大学生时，他们发现“E 和 4”是最为常见的回答（59 名学生这样选择），而“E”是其次最为常见的回答（42 名学生这样选择）。换句话说，大多数学生选择了那些翻过来就能够证实那条陈述的卡片。仅有 5 名学生给出了正确答案：“E 和 7”。

如果这个答案看起来难以理解，就试着这样考虑问题。要检验的规则是“如果是元音字母，那么就是偶数”，或者更为基本，“如果 X，那么 Y”。证明一个“如果……就”类型的陈述错误唯一的办法就是找到一个“是 X 而不是 Y”的例子（比如，元音和奇数）。因此，仅有的能够证伪的方法就是，找出那些具有元音字母和奇数的卡片（“E 和 7”）。而具有偶数或辅音字母的卡片并不重要。

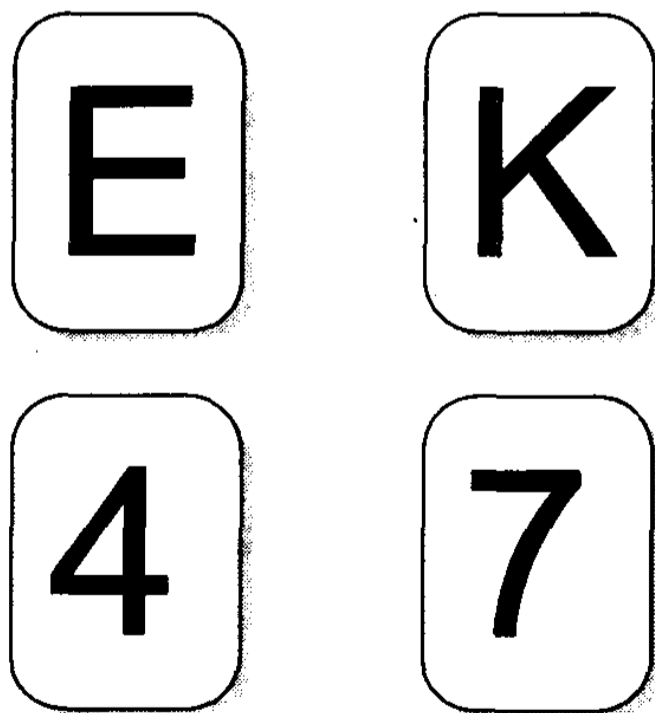


图 20.1

* 一些版本的问题使用了符号而不是 E、K、4 和 7，但是问题的逻辑结构都是一样的。

尽管看起来简单，这个问题对于大多数人来说却是极其困难的。罗宾·道斯（Robin Dawes, 1975）甚至发现五名“最受尊敬的”的研究数理的心理学家中有四位都不能解决这一问题。本章的焦点就在于为什么这样的问题如此困难，以及它能够给决策者带来什么启示。

再次猜猜看

“你将得到三个数字，它们符合我头脑中的一个简单规则。这一规则关注的是任意三个数字之间的关系，而与数字本身的绝对大小无关，比如它不是那种类似于‘所有数字都大于（或小于）50’的规则，等等。

你的目标是通过写出由三个数字组成的若干组数字组来发现这个规则……在你写下每组数字之后，我会告诉你这些数字是否符合这一规则……没有时间限制，但你应该通过尽量少的数字组来发现它。

记住你的目标不是简单地找出符合这一规则的数字，而是发现规则本身。当你非常自信地认为自己已经发现了这个规则时，不要过于匆忙，将它写下来……有什么问题吗？”

在检验假设的实验中，沃森（1960）将这些指导语——以及一组范例数字2, 4, 6——呈现给29名大学生。他头脑中的规则一直是“三个数字按照从小到大的顺序排列”，但是他发现只有六名被试第一次回答就正确说出了这条规则。作为粗略的解释，下面列出了一名典型被试的手稿：

被试4（一名19岁的女性）：8, 10, 12。

主试：那些数字符合规则。

被试：14, 16, 18。

主试：那些数字符合规则。

被试：20, 22, 24。

主试：那些数字符合规则。

被试：1, 3, 5。

主试：那些数字符合规则。

被试：规则就是以任意一个数字开头，然后每次加上2来得到下一个数字。

主试：那不是正确的规则。请继续……

与四卡片问题一样，沃森发现被试总是更多地去证实规则（如8, 10, 12）而不是证伪规则（如12, 10, 8）。这种倾向被称做“证实偏好”（confirmation bias）。尽管“证实偏好”已经成为了一个万能词组（Fischhoff & Beyth - Marom, 1983），它还是经常被用来表示决策者偏好与假设一致的信息，而不是那些不一致的信息。

在一项关于人们如何检验假设（预感、规则、理论等等）的深入分析中，乔希·克莱曼（Josh Klayman）和 Young - Won Ha（1987, pp. 220）认为证实偏好是“正性实验策略”（positive test strategy）的结果，这是一种有用的直觉，但是它“就像任何通用直觉一样……并不总是最优的，可能在某些情况下导致非常严重的困难。”在第15章已经讨论了一组困难——当决策者主要关注于正性的共生事件时，由共变评估所产生出来的问题（比如，Jan Smedslund 所做的护士研究，1963）。下一部分将给出另外一个例子。

自身永存的社会信念

1979年，马克·斯奈德和南茜·坎托（Mark Snyder & Nancy Cantor）发表了三个关于社会知觉中证实偏好的实验。第一个实验中，所有被试都得到了一位名叫简的女子的例子——她在一些情况下会表现得很外向而在其他情况下却表现得较为内向。举例来说，简在慢跑的时候会与陌生人自由交谈，但在超市里面她会感到害羞和胆小。阅读这个例子两天以后，被试会被要求评估两个命题：（1）简非常适合一个通常与外向性相联系的工作（房地产销售），或者（2）简非常适合一个与内向性相联系的工作（图书馆研究员）。一些被试被告知简正在申请这份工作，而另一些被告知她已经获得了这个工作。简单而言，我们将关注那些认为她正在申请这份工作的被试。

要判断简是否非常适合这项工作，以及她将会有多适合这个工作，这些被试需要列出与事例相关的所有依据。斯奈德和坎托发现，那些考虑简在申请房地产销售工作的被试往往列出她的外向性作为相关依据，而那些考虑简在申请图书馆研究员工作的被试往往列出她的内向性作为相关依据。换句话说，两种条件下的被试都把“证实”性证据看得比“证伪”性证据更有关联性。此外，这种证实偏好与人们对于简在所考虑工作上的适合性所做出的判断有关。被试越喜欢证实性的证据，他们就会认为简越适合——无论她申请什么职位。

第二个实验基本重复了第一个实验，但是有几处小的改动（如：将“非常适合”一词改为“适合性”来避免过于强调证实性）。而在第三个实验中，被试只需要列出他们想要了解一个工作候选者的哪些方面，从而来评估这个人有多适合房地产销售或图书馆研究员的工作。和前面一样，证实性的信息比证伪性的信息出现得更加频繁。比如，当考虑那些房地产销售工作的申请者时，被试提问“这名申请者待人友好吗？”比“这名候选者有多害羞？”的频率高得多。马克·斯奈德和威廉·斯旺（William Swann）（1978）所做的几个实验也获得了类似的结果。

这些结果带给我们的启示可以很好地拓展到职业领域之外。斯奈德和斯旺（1978, pp. 1211 - 1212）观察到：

【人们】可以为自己创造一个这样的世界，在这个世界中，假设会变成自我验证性的假设，并且信念会变成自身永存的信念……从这个观点出发，就不难理解为什么如此多关于他人的普遍信念（特别突出的是，明显错误的社会及文化刻板印象）会非常难以改变。即使一个人将要对这些信念产生足够的怀疑并主动地对它们进行验证，他仍然可能会“找到”那些他需要用来证实及坚持这些信念所需要的全部证据。这样，最终这个人将会持有一种安全的（但是毫无根据的）感觉，即这些信念一定是正确的，因为它们通过了十分恰当和精确的评估过程。

皮格马利翁效应

按照人们应该成为的样子去对待他们，并且你要帮助他们变成他们能够成为的样子。

——歌德（cited in Peter, 1977）

“自我实现的预言”（self-fulfilling prophecy）一词在1948年由罗伯特·默顿（Robert Merton）提出。用默顿的话说（1948, pp. 195-196）：“起初，自我实现的预言是指对情况的错误定义引发了一种新行为，这种新行为使得最初的错误概念变为了真实的。自我实现预言的外在效果是使得错误的盛行得以延续。因为预言者会引用真实的事件进程来证明他从一开始就是正确的……这就是社会逻辑的反常之处。”简言之，自我实现的预言是错误的概念，但它们是那种最终被证明是正确的错误概念。

在1968年，罗森塔尔和雅格布森发表了当前在自我实现预言方面最为著名的研究。在该研究中，小学教师得到了关于自己学生的一些诊断性信息，这些信息暗示，占总数20%的学生将会在未来的学业中获得巨大发展。在8个月以后的IQ测试中，这些学生的成绩确实比其他学生长得快。

这一发现值得关注是在于，这些高成就者是随机选出的。显然，教师给予了这些“巨大发展者”更多的赞扬和关注，这样导致的结果就是，这些学生确实比其他人发展得更好。罗森塔尔和雅格布森将这种现象称为“皮格马利翁效应”（来自萧伯纳的戏剧《皮格马利翁》，剧中希金斯教授通过教给一个卖花姑娘如何着装和说话，使她真正变成了一位“淑女”）。

从罗森塔尔和雅格布森的研究开始，已经有400多个实验对人际期望的自我实现性质进行了探讨，并有超过100个实验专门对教师期望的效果进行了检验（Brophy, 1983; Jussim, 1986; Rosenthal, 1987, December）。总之，这些研究表明，尽管效果经常是中等程度的，但教师期望对学生的表现具有重要影响（Brophy, 1983）。有趣的是，同样有“学生皮格马利翁”的证据存在。罗伯特·费尔德曼（Robert Feldman）及其同事发现，学生的期望——既对于他

们自己的表现也对于他们老师的表现的期望——能够影响师生关系和学生的表现，且影响程度并不亚于教师的期望（Feldman & Prohaska, 1979; Feldman & Theiss, 1982）。

在男人们的头脑里

尽管罗森塔尔和雅格布森（1968）探讨的是师生之间的互动，皮格马利翁效应却并不仅仅局限于教室之内。对于皮格马利翁效应最富戏剧性的描述之一就是马克·斯奈德、伊莉莎白·坦克（Elizabeth Tanke）和埃伦·伯奇德（Ellen Berscheid）（1977）所做的研究，该研究探讨了男人对女人吸引力的刻板印象是如何能够自我实现的。

实验的第一部分，在通过电话相互认识的过程中，随机配对的男性和女性被录下了十分钟的对话。然而，女性被试所不知道的是，实验者已经事先告知男性被试“实验中的其他人已经告诉我们，当头脑中形成了交谈者的形象时他们会感觉更舒服。”于是男人们用一次性成像的相机照了相，并得到了一张女性的快照（表面上是他们的搭档）。事实上，这些快照是从事先准备好的八张照片中随机选择的。这些照片中有四张是那些被评定为具有高吸引力的女性，有四张是被评定为没有吸引力的女性。这样就引导男性被试相信他们的搭档是外表吸引人的还是外表没有吸引力的。男性被试的快照被丢弃在了一旁，女性被试根本没有得到任何可以形成头脑图像的信息。

在相识谈话之后，每个男性被试完成了一份问卷，内容是关于他对与其谈话的女性的印象。问卷中包括了很多与一个有吸引力的人有关的刻板印象特征——善于交际、镇定、幽默等等。结果并不出人意料，与那些认为交谈同伴没有吸引力的男性相比，认为交谈同伴外表吸引人的男性在随后的评定中，认为同伴更加亲切、镇定、幽默以及善于社交。这是在第4章所讨论过的晕轮效应的一个例子。

更为重要的是几名独立评分者所做出的判断，他们倾听了谈话录音的剪辑。一组评分者在每段谈话中仅仅倾听男性的声音，而另一组评分者仅仅倾听女性的声音（每组评分者都既有男性又有女性，尽管在评分过程中并不存在显著的性别差异）。根据这些评分者做出的判断——他们既不知道实验假设也不知道被试实际的外表吸引力——与以为交谈同伴没有吸引力的男性相比，那些认为自己正在与一位有吸引力的女性交谈的男性被试听起来更加善于交际、温情、宽容、有趣、大胆、外向、幽默，并且有更强的社会适应性。并且从反应推测，与开始被知觉为不具吸引力的女性相比，那些开始被知觉为外表具有吸引力的女性在实际中听起来也更加富有吸引力（比如善于交际、镇定等等）——尽管被试是被随机分配到各种实验条件下的，而且男人的预想与女人实际的外表吸引力没有任何关系。

正如斯奈德、坦克和伯奇德（1977，pp. 661, 663）所提出的：“最初在男人们的头脑里形成的事实，现在已经在与他们互动的女人的行为中变成了真正的事实——一个即使是最幼稚的观察判断者也能加以辨别的行为事实，这些观察者仅仅评估了谈话录音中女性的贡献……知觉者最初的错误归因已经变得真实：刻板印象真正作为一种自我实现的预言而发挥了它的作用。”

自我实现的种族刻板印象

这类自我实现的预言——使刻板印象得以延续的一类——在种族歧视中扮演着关键角色。卡尔·沃德、马克·赞纳和乔尔·库珀（Carl Word, Mark Zanna, & Joel Cooper, 1974）发表了关于这个问题的一项研究。

沃德、赞纳和库珀在导言中解释说，人们往往会通过非语言线索表现出他们对别人的态度。比如说，当人们对某人持有正性态度的时候，他们就会让自己和那个人靠得相当近，表现出高度的目光接触、更为直接的肩部相向，以及更多的身体前倾。另一方面，人们往往避开那些名誉不佳的人。比如，他们会很快中止与这种人的会面，并且与他们保持更远的距离。沃德、赞纳和库珀在两个实验中考察了这种偏见——第一个实验中，白人被试面试白人和黑人实验助手，第二个实验中，白人实验助手面试白人被试。所有的被试和助手都是男性。

第一个实验中，被试会在实验室中遇到两个表面上也是被试的人（但实际上是主试的助手）。然后另外一个助手出现，再过一会儿主试走了进来。主试告诉四名“被试”他们将与另外四个团队在一场市场活动中进行较量。但主试解释说，他们的团队还需要一个人，并且团队中的一员将要对四个候选者进行面试以挑选出第五位成员。然后他们会进行一次事先做过手脚的抽签仪式来使得被试成为最后的面试官。

安排四次面试的原因在于隐瞒真正的研究目的（即，直接比较白人考官对待白人、黑人候选者有什么区别）。第一次面试（一直是位白人候选者）被简单地当做面试的热身，而最后一次面试从来没有发生过（告知被试申请者已经打电话取消约定）。令沃德、赞纳和库珀（1974）感兴趣的是第二次和第三次面试。一半被试会首先面试一名白人候选者，再面试一名黑人候选者；另一半被试会首先面试一名黑人候选者，再面试一名白人候选者（尽管后来发现顺序并不会导致差异出现）。在这些面试中，候选者——同样是实验者的助手——被训练过以保证其行为符合标准方式，并且他们不知道任何实验者设计该测试的目标。

沃德、赞纳和库珀（1974）测量了面试过程中的几个特征，包括：（1）面试时间长短，（2）面试官出现语言错误的次数（假定反映了不适感），以及（3）面试官和候选者之间的物理距离。第一个变量由主试测量。第二个变量由

独立的评分者进行计分，这些评分者边听面试的录音边对发言结巴、词语重复等诸如此类的内容做记录。第三个变量，物理距离，是通过一种特殊而富有创造性的方法加以测量的。当被试进入面试考场的时候，候选者已经坐下，而主试假装发现没有面试官的椅子了。然后他会要求被试去邻屋拿一把椅子，被试放置椅子的位置提供了其与候选者之间物理距离的测量指标。

沃德、赞纳和库珀（1974）发现被试面试白人候选者所花费的时间比面试黑人候选者多出了35%，在面试黑人时比面试白人时多犯了50%的语言错误，并且椅子放置的位置距离，黑人候选者比白人候选者要远7%。所有差异在统计上都是显著的。

但这还不是全部。在证明了这些差异之后，沃德、赞纳和库珀又进行了第二个实验来评估这种职业面试表现中的歧视效应。在第二个实验中，白人被试由一名白人助手面试，该助手采用第一个实验中对待黑人的方式或者对待白人的方式来对待这些白人被试。就是说，在前一种实验条件下，面试官坐得离候选者更远，犯更多的语言错误，并且更早地中止面试。

沃德、赞纳和库珀（1974）发现，那些被以第一个实验中对待黑人候选者的方式进行面试的被试在面试过程中表现得更加糟糕（由观看过面试录像的独立评判人进行评定），他们犯了50%以上的语言错误，并在后来认为面试官不够友好。因此，第一个实验中黑人被试所遭遇的对待方式使得白人被试在第二个实验中表现得很差——这为种族刻板印象如何能够自我实现提供了一个有力的证明。

结 论

尽管关于自我实现的刻板印象方面的研究主要局限在男性被试研究上（Christensen & Rosenthal, 1982），其他自我实现的预言和证实偏好已经在男性和女性被试身上得到了充分的证明（见三篇精彩的综述，Darley & Fazio, 1980；Miller & Turnbull, 1986；Snyder, 1984）。人们寻求证实性证据的倾向——不管是逻辑问题解决的任务、职业面试、在班级背景下，还是其他方面——是广泛存在且业已确定证实的。

此外，克利福德·迈纳特、迈克尔·多尔蒂和瑞安·特韦尼（Clifford Mynatt, Michael Doherty, & Ryan Tweney）所做的两个实验表明证实偏好可能难以消除。第一个实验中，迈纳特、多尔蒂和特韦尼（1977）创造了一种模拟研究环境，在这一环境中被试必须发现计算机屏幕上支配“粒子运动”的某个规律。被试随机得到三种指导语中的一种：（1）证实性指导语，他们被告知一名科学家的基本工作就是证实理论和假设；（2）证伪性指导语，他们被告知一名科学家的基本工作就是推翻理论和假设；（3）检验性指导语，他们仅被告知一名科学家的工作就是对理论和假设进行检验。

迈纳特和他的同事发现证伪性指导语实际上对证实偏好没有任何影响。如果不考虑被试得到了哪一种指导语的话，他们在全部实验的大约70%中会寻求证实性的证据。

在第二个实验中，迈纳特、多尔蒂和特韦尼（1978）扩展了证伪性的指导语。采用与前面相同的步骤，他们将被试随机分配到两种条件中的一种：（1）一个没有指导语的控制组，或者（2）一个强线索组，在这个群体中，被试会得到详尽的指导语，强调证伪和多假设验证的价值。然而，结果又是，这样的指导语效果并不明显或根本没有效果。

如何避免证实偏好以及自我实现的预言？尽管针对这一问题的研究相对较少，但有一种策略可能是关注动机因素（Snyder，待发表）。比如说，斯奈德、坎贝尔和普莱斯顿（1982）通过警告面试官他们面试的人可能会把某些问题看做思想封闭或偏见（如指向证实性刻板印象的问题）来消除证实偏好。一旦该研究中的面试官对这种可能性变得敏感，他们就会大致等同地采用证实性和证伪性证据进行衡量。

另外一个有价值的策略可能是采用鼓励证伪性回答的方式来组织问题。决策研究者杰伊·拉索和保罗·休梅克（1989）讲述了一个他们以前学生的故事——杰伊·弗里德曼采用这一策略取得了巨大成功。作为 Kidder, Peabody 和 Company 三家投资公司的顶尖分析师，弗里德曼将在制定财务建议之前收集证伪性证据。拉索和休梅克（1989, pp. xiv - xv）这样描述了弗里德曼的方法：

在收集公司的信息时他会故意问一些设计好的问题来“证伪”他认为真实的那些信息。比如，如果弗里德曼认为免洗尿布生意正在价格上变得没有竞争力，他会问执行官们一个相反的问题，如“一次性尿布的价格竞争是不是正在变得更有优势？”这类问题使他比其他竞争分析师更容易了解真实情况。

正如第19章所谈到的，通过思考自己的判断在哪些情况下可能是错误的，决策者能够降低过度自信并提高他们的决策质量（Koriat, Lichtenstein, & Fischhoff, 1980; Lord, Lepper, & Preston, 1984）。这样一种方法可能也会减少自我实现的预言和自我实现的刻板印象。但是在当前，这种可能性是不确定的，需要通过心理学的研究来证实——或证伪——其正确性。

行为陷阱

“感谢您打电话过来。目前我们的接线员正忙。请您在线等待，您的电话会按照顺序得到回应。”

一分钟过去了，两分钟过去了，你开始考虑是继续等待还是过会儿再拨。可能你被转到了无人接听的线上，好像电话在森林里空响，无人应答……

另一方面，挂断可能意味着重新开始。其他人可能会排到队伍的前面，你就丧失了原有的优先权。还是继续等吧，谁知道呢，可能你就是下一个了。

你又等了一会儿，三分钟，四分钟。谁打了那么久啊？你在想。

最后你拿定了主意。如果下个 60 秒内仍然没有接线员回应，你就挂断。30 秒过去了，40 秒过去了，50 秒了，但是还是没有回答。当最后期限到来的时候，你犹豫了一下，却仍抱有希望，但最后还是还是在沮丧中狠狠挂了电话。

这样的情形看起来是不是很熟悉？这个情况具备“行为陷阱”的所有特征。行为陷阱（behavioral trap）是指这样一种情境：个人或者群体从事一项很有前景的工作，最后却变得不尽人意并且难以脱身。这个定义类似于约翰·普拉特（John Platt, 1973）在有关社会陷阱的开创性工作中所提出的关于社会陷阱的定义。本章将着重思考普拉特、约翰·克罗斯和梅尔文·盖耶（John Cross & Melvin Guyer, 1980）所做的分析。由于陷阱可能是非社会性的也可能是社会性的，因此我们将使用更加宽泛的“行为陷阱”的概念，而不是传统意义上的“社会陷阱”。

陷阱的分类

1980 年，克罗斯和盖耶发表了关于陷阱和反陷阱的分类。用克罗斯和盖耶（1980, p. 18）的话说就是：“当我们避免可能有利的行为时反陷阱（不作为的过失，sins of omission）就发生了，而当我们进行可能有害的行为时则出现陷阱（作为的过失，sins of commission）。”如上所述，一个常见的陷阱就是等待接线员。常见的反陷阱则包括令人厌恶的清理工作（随着时间的推移情况将变得越来越乱）以及逾期的回信（这种情况下耽搁时间越久越令人尴尬）。

生活中存在着几种典型陷阱，每一种都有相应的反陷阱。根据克罗斯和盖耶的分类，我们把陷阱主要分为五类：

- 延期陷阱
- 无知陷阱
- 投入陷阱
- 恶化陷阱
- 集体陷阱

尽管这些陷阱常常结合在一起形成混合陷阱，但是每一种陷阱都基于不同的运作原理。因此，以下部分将分别讨论这些陷阱。

延期陷阱

如果你发现按时进食或者锻炼有些困难，那么你应该知道这是由于延期陷阱（time delay traps）的作用造成的。在延期陷阱中，暂时的满足与长期的后果相冲突。无知地开始食用一种好吃的甜食或者吸几根香烟可能在未来许多年以后导致肥胖或者肺癌。或者，在反陷阱中，对目前不愉快事件——对于某些人来说是有氧运动，对于另一些人可能是牙科检查——的逃避，最终可能导致心肌梗塞或者牙周炎。这类陷阱和反陷阱的关键在于短期的相对小的痛苦和愉快足以引起长期的、破坏性的甚至是致命性的后果。

任何短期后果与长期后果冲突时都可能成为一个延期陷阱。典型的冲突包括：喝酒的快感与第二天的宿醉，当前没有保护措施的性行为的快感与延期的感染艾滋病或者非意愿怀孕的可能性，可抛弃产品的方便性与长期的环境污染，受到信用卡影响的“先买后付”的选择与长期财务计划，以及体罚的短期效果与最终的糟糕结果。甚至伊甸园里的苹果也可以被看做是延期陷阱中的一种诱饵——诱惑的最终象征与它潜在的令人上当的后果。

无知陷阱

人们在延期陷阱中通常意识到了他们行为的长期后果。暴饮暴食者通常非常清楚体重上升的后果，吸烟者有时候甚至把香烟称为“癌症小棒”。关于体重上升或者罹患癌症的警告在延期陷阱中很难起到有效作用。

无知陷阱（ignorance traps）与此不同。在这些陷阱中，行为的负面后果并不被理解或者开始时并没有被预见到。例如，19世纪时吸烟者并没有意识到吸烟与肺癌之间的关系，如果当时人们知道这些信息的话，很多人可能不会去吸烟（当然，吸烟仍然具有延期陷阱的特征，上百万人尽管知道两者之间的联系却仍然深陷其中）。

当一段新生活开始时，无知陷阱通常较为常见。例如，大学生们有时会选择一个人并没有原先设想的那么有趣的专业，工人们有时发现他们在做一份与他

们期望大相径庭的工作，恋人们有时会发现现在的伴侣并没有原先那样吸引人了。这些陷阱是生活中不可避免的一部分，尽管有一些方法可以使被陷阱捕获的概率降到最低（减少或者避免不良后果的技巧会在本章稍后部分讨论到）。

一个特别悲剧性的无知陷阱是美国农业曾经对于杀虫剂的依赖。当合成的有机杀虫剂如 DDT 在 20 世纪 40 年代被引进时，它们看起来似乎是保护作物免受虫害的有效手段。美国农民很快就将它们作为控制害虫的手段。

然后，两件没有预料到的事情发生了：（1）鸟类和其他害虫捕食者开始相继死去，（2）害虫对那些使用过的化学药品产生出了抵抗能力。虫害开始增加了。人们发明了新型杀虫剂，然而再次出现了能够抵抗化学药物的害虫品种。经历了四亿年的进化之后，不经历战斗这些害虫是不会投降的。

数十年来，这场战役一直在美国的农场上进行着，但是每一轮新的“化学武器”只会引起更加严重的虫疫。美国农作物由于虫害而遭受的损失比例在 1950~1974 年间翻了一番（Robbins, 1987），根据加州大学的昆虫学家统计，在加州，25 种最严重的农业害虫中有 24 种是杀虫剂诱导产生或由杀虫剂导致破坏性增强（Luck, van den Bosch, & Garcia, 1977）。每年，美国使用超过 1 亿磅的杀虫剂，它们对野生动物、植被、水域及人类安全十分有害。

投入陷阱

克罗斯和盖耶（1980）在他们的分类中没有明确包含投入陷阱（investment traps），但是最近这类陷阱成为了大量研究的课题。当以前花费的时间、金钱或者其他资源让人们做出了他们本不会做出的选择时，投入陷阱就出现了。用决策理论的术语来说，这些陷阱导致了“沉没成本效应”（sunk cost effect）。

哈尔·阿克斯和凯瑟琳·布卢默（Hal Arkes & Catherine Blumer, 1985）在 10 个不同的小型实验中举例说明了沉没成本效应。在其中的一个实验中，要求一群被试解决以下的问题：

作为一个航空公司的董事长，你已经投资了 1 000 万美元来开发一个项目。目的是制造一种不会被传统雷达监控到的飞机，也就是隐形飞机。在该项目的 90% 已经完成时，另外一家公司已经开始为隐形飞机做市场宣传了。并且，很明显，与你们的飞机相比，他们的飞机速度更快、也更经济。问题是：你还会把剩余的 10% 的研究资金用于完成隐形飞机项目吗？

你可以通过查看“读者调查”的第 6 题去检查你的结果。阿克斯和布卢默发现 85% 的被试倾向于完成该项目，尽管完成后的飞机与市场上已有的飞机相比处于劣势。给予另外一组被试另一个版本的问题，其中并没有提到先前的投资，结果只有 17% 的人支持在该项目上投资。一千万美元的沉没成本造成了二者的差异。

在另外一个实验中，阿克斯和布卢默（1985）显示沉没成本可能具有持续的效应。在一个研究中，60位剧院的老顾客到售票口去买俄亥俄大学剧院的季票。这些人所不知道的是，他们随机买到了下面三种票中的一种：（1）普通的15美元的票，（2）打了2美元折扣的票，（3）打了7美元折扣的票。有幸买到打折票的被试被告知折扣是剧院部门的促销活动之一。

每种票有不同的颜色，于是阿克斯和布卢默（1985）能够收集每场演出的票根并且发现多少被试观看了每场演出。出于分析的目的，剧院的演出季被分成了两个部分，每个部分有6个月，期间共有5场演出。尽管阿克斯和布卢默在第二个演出季中并没有发现显著差别，但在前6个月中，他们发现那些买了全额票的顾客比那些买了折扣票的顾客（不管其折扣的幅度）观看演出的次数要多。所以，即使是微不足道的2美元投入也会持续影响行为长达6个月。

这个研究之所以非常重要，原因有以下两点。首先，它表明沉没成本效应不仅仅局限于纸笔测量；其次，它还显示了投入的不同可以对行为产生相对持久的影响。正如巴鲁克·菲施霍夫（Baruch Fischhoff）和他的同事（1981, p. 13）在他们名为《可承受的风险》（*Acceptable Risk*）一书中所述：“美国任何一个大型的水坝只要开工就不会半途而废的事实表明，一点点的水泥都能在一个关键问题中起作用。”

恶化陷阱

恶化陷阱（deterioration trap）与投入陷阱类似，除了行为的成本与收益随着时间而变化之外，这些陷阱——克罗斯和盖耶（1980）叫它们“可变强化陷阱”——出现在当原先高回报的行为逐渐变得不那么有收益或者变得更具有惩罚性时。

恶化陷阱的典型例子是海洛因成瘾（尽管海洛因成瘾也可以被看做是延期陷阱或者无知陷阱）。起先，海洛因使用者发现这种药物令人欣快。但是，随着时间的推移，他们产生了抗药性，因而需要更多的药物量才能达到同样的感觉。最后，海洛因使用者用该药物已是为避免不适症状的出现而不是体验愉悦感。起先的愉悦体验最终成了摆脱不掉的噩梦。

同样的过程也能出现在“杀虫剂成瘾”中。尽管杀虫剂的使用在一开始也可能是无知陷阱，但逐渐地，它转化成为恶化陷阱。据《生物科学》（*BioScience*）杂志的一篇报告，杀虫剂依赖性的工作原理如下：

起初的一段时间，由害虫所导致的粮食损失显著减少……但是最终，某种主要的、个别的或者杀虫剂引致的害虫产生了抗药性。这个问题是通过增加（多样化）杀虫剂和改变杀虫剂来解决的，但是替代品的有效时间变得更短暂，因此需要更加频繁地使用以达到与原来相同的控制力。此

时，种植者要想从这种策略中解脱出来，即便不是不可能，也会变得更困难了。随着他们继续使用杀虫剂，他们的问题也变得越来越严重（Luck, van den Bosch, & Garcia, 1977, p. 607）。

在那些没有看到这一过程如何发展的旁观者看来，恶化陷阱和反恶化陷阱常常会产生一种荒唐或自我毁灭性的行为。在斯金纳（1980, pp. 150 - 151）的回忆录《笔记》（*Notebooks*）中，他描述了下述行为的例子：

比尔的卡车是他惟一的收入来源，就像渔夫的船或者农民的奶牛与拉犁的马一样。岛上咸咸的海风、缺乏维护的道路，以及醉酒司机的虐待几乎要使它报废了。挡风玻璃上布满了呈辐射状的裂纹；挡泥板已经腐蚀成了薄片，弯折着，扭曲着；只有一块块的垫料还残存在座椅的弹簧中。

我曾经请比尔帮忙把我们的船运到山下去。卡车停在一个乡村商店前面的下坡路上。我上了车，坐到了残缺不全的座位上。比尔推了一下卡车，跳了进来，抓住挡把，借助仅有的一点速度挂上了挡。一阵激烈的震动后，马达开始咆哮。比尔……拼命地踩下油门，一只手始终在气门上。看到汽车发动了，他感到非常满意，于是快速地向那家商店倒车以便掉头。可是卡车熄火了，横在路上。我们三四个人一起推，其中包括两个从一辆小汽车里出来的年轻人，卡车堵住了他们的路……我们回到了那个下坡上，再一次将车发动起来，结果又熄火了。比尔不时地跳出来，打开发动机罩，用扳钳调整着什么。我们向错误的方向前进了大约160米，发动机咆哮着，并且在比尔加油的时候拒绝加速。最后他解释道，他的发动机启动装置被送去修理了，可能被游船送回来。如果他花两个小时时间去取回发动机，情况会怎样呢？而他却没去取。48小时之后他的车仍然停在下坡路上。没人愿意雇他的车了。

他为什么继续？从某种意义上讲他没有选择。他喝酒花光了他的收入……（但是他的）没有选择并不是这个故事的全部。他对那辆卡车狂热的关注是（对他所付出辛劳的回报的不断降低）的结果。比尔不愿从卡车那里一无所获。假如那是一匹马的话，他也许早就把它打死了，因为一匹年迈的老马也是它主人专注于在一项所费时间越来越长的工作上的强化物。比尔的卡车也正被他往死里打呢。

对于一个不知道比尔以前经历的旁观者而言，他的行为可能显得荒诞可笑。但是同样的动力在恶化的社会与感情关系中也同样起作用。当人际关系随时间逐渐变坏时，它们就构建了一个非常难以逃遁的反陷阱。

集体陷阱

与前面的陷阱不同的是，集体陷阱（collective traps）涉及多个人。在集体

陷阱中，对个人利益的追逐导致了对集体不利的后果。一个简单的例子是高峰期的交通堵塞。上百人倾向于在同一时间开车，但是假如每个人都按照自利的原则行事，那么大家都会遭殃。

集体陷阱——数学博弈论中“社会困境”（social dilemma）的近亲（Dawes, 1980）——受到的研究关注比其他所有陷阱加在一起还要多。最著名的集体陷阱是囚徒困境，在这种情境下，两个囚犯分别关在单独的牢房里并且被提供了如下的说明：

区律师：听着，比尔小子。如果你们中的任何一个都不认罪的话，我们有足够的证据让你和你的同伙入狱一年。我们真正想要的是，你们当中至少有一个人认罪。如果你认罪但是你的同伙不认罪，我们将判处你的同伙十年徒刑，而你将被无罪释放。但是，如果你不认罪但是他认罪了，你就会被关押十年。

野蛮比尔：如果我们都认罪呢——我们都会被判十年吗？

区律师：不。那样的话，我们会基于你们的诚实把每人的徒刑降为五年。

在一个标准的囚徒困境中，两个囚犯都面临着同样的选择——无论同伙选择什么，他们都最好选择坦白。假如他们的同伙拒绝承认，那么他们无罪释放；否则，他们起码不会被判十年徒刑。困境就在于如果每个囚犯都追求自己的利益而坦白，他们就会分别受到五年徒刑的惩罚，而这比两人都不认罪的惩罚要重（见图 21.1）。

		罪犯A	
		认罪	不认罪
罪犯B	认罪	5年 / 5年	0年 / 10年
	不认罪	10年 / 0年	1年 / 1年

图 21.1 囚徒困境问题。每一格对角线右上部分为罪犯 A 的陈述，对角线左下部分为罪犯 B 的陈述。

另一个著名的集体困境是生物学家加勒特·哈丁 (Garrett Hardin, 1968) 命名的“共有地的悲剧” (the tragedy of the commons)。在这个陷阱的经典版本中, 一个放牧部落使用公共牧场来放养牲畜。起初没有问题, 但渐渐地, 牛的数量达到了土地承载能力的极限。此刻, 增加一头牛对放牧人的效用有两方面——正面和负面的效用。正面效用为增加一头牛所带来的收益。这个收益完全归属于增加这头牛的放牧人。负面效用则是增加一头牛所导致的过度放牧。这个成本由该部落所有的放牧人共同承担, 并且对每个放牧人来说, 可以忽略不计。结果就产生了一个两难困境——每个人都从增加一头牛中获益, 但是对个人利益的追求导致了一个不尽如人意的结果。哈丁把共有地悲剧与另一些问题联系起来, 比如人口膨胀、污染、全球资源衰竭以及核武器的扩散等等。

共有地悲剧在很多方面类似于臭名昭著的“床垫问题”, 这是由托马斯·谢林 (Thomas Schelling, 1971) 最先描述的一个集体反陷阱。在床垫问题中, 成千上万辆汽车从科德角 (Cape Cod) 度完周末后, 由一条双车道高速公路返回, 就在此时, 一个床垫从货车的顶端不为人注意地掉了下来, 掉入了向北行驶的车道上。问题来了: 谁会停下来移开这个床垫呢?

通常, 答案是谁也不会。远离该床垫的、处于忙碌交通中的人们不知道问题在哪, 所以不会来移开。正在绕过床垫的人们已经等了如此之久, 以至于他们只想着如何绕过它。在漫长的等待之后, 他们最不情愿做的就是花几分钟时间把床垫从车道上移开了。而已经绕过床垫的人们不再有动力去挪开它。

床垫问题与紧急情况下的集体反陷阱很相似 (责任分散了, 旁观者干涉的速度就慢了)。这也可能部分解释了政治“冷漠”为何在美国如此普遍。不幸的是, 正如道格拉斯·霍夫施塔特 (Douglas Hofstadter, 1985, p. 757) 贴切的评论: “个体水平的漠不关心会导致社会大众丧失理智。”

你愿意为一美元付多少钱

心理学研究中最著名的行为陷阱之一是美元拍卖的游戏——这个游戏中结合了集体陷阱、投入陷阱、无知陷阱的特征。这个游戏由马丁·舒比克 (Martin Shubik, 1971) 发明, 在该游戏中, 1 美元被卖给出价最高的人。普拉特 (1973) 认为该拍卖游戏具有四个简单的规则:

1. 拍卖进行时出价者之间不得有任何交流。
2. 出价由 5 美分开始, 每次只能加 5 美分。
3. 出价不能超过 50 美元 (为避免竞价者狂热的热情)。
4. 出价最高的前两名都必须付出他们所出的价格, 即使这 1 美元只能给出价最高的那个人 (说到底, 拍卖商得弥补他的损失)。

尽管游戏听起来非常简单, 但有两个“无回报的点”值得注意。其一是当

两个出价最高的人所出总额超过1美元时，那么拍卖商就能稳获利润（例如，一个人出50美分而另一个人出55美分）。此时，拍卖在单个竞价者眼中是有吸引力的（1美元只要出55美分就能到手），但是个人利益的追求已经导致了竞价者整体的损失。

第二个要当心的地方是当第一个人出价超过1美元时。为了明白为何人们愿意出大于1美元的钱来获取1美元，我们考虑一个人出了95美分，而此时恰好有另一人出了1美元时此人的困境。如果是你，你怎么做？如果你在该点放弃，你就必定损失95美分。另一方面，如果你出价1美元零5美分，你就赢得了1美元，损失仅为5美分。问题是，你的竞争对手也面临着同样的状况。

于是结果是，拍卖常常会出现几美元的竞价。

拍卖游戏广受关注的一个原因是它很类似于核武器竞赛以及其他的国际冲突（Costanza, 1984）。在1980年，艾伦·特格（Allen Teger）出版了《投资过多，无法退出》（*Too Much Invested to Quit*）一书，整本书都是关于一美元拍卖游戏的，并且很多结论都可以直接应用到军事冲突上。据特格称，被试们起先是被自己的利益所驱使，但是渐渐地他们的动力变了。随着竞价的进行，被试们开始关注赢得竞争、保住面子、使损失最小化，并且惩罚竞争对手，因为对手使他们陷入窘境（通常，只有两个竞价者在最后还保持活跃）。特格发现，当开价达到1美元时，双方都觉得他们是被对方逼迫继续下去的，并且很多被试认为对方继续下去是疯狂的——他们并没有认识到同样的力量作用于双方的参与者。这个“镜像”酷似核武器竞赛。

深陷泥潭

一旦竞价者在一美元拍卖游戏中被套——“深陷泥潭”，巴里·斯塔（Barry Staw, 1976）这样形容——他们通常会继续打击对方直到对方最终放弃为止。乔尔·布罗克纳和杰弗里·鲁宾（Joel Brockner & Jeffrey Rubin, 1985, p. 5）把这种机制称做“诱捕”（entrapment），被定义为“一个决策过程，此过程中人们持续提高对已经被证明是失误的先前选择行为的忠诚度，目的是使他们以前的投入显得合情合理。”

对诱捕的最先研究之一是由斯塔（1976）所做的。斯塔呈现给商学院学生们一个假想但是非常详细的情境。这是一家开始亏损的高科技公司，他让学生设想自己是公司的财务副总裁。根据该情境，公司的董事们决定额外投资1000万美元的研发资金给两个最大部门中的一个——消费产品部或者工业产品部。在研究的第一部分，半数的学生被问及哪个部门应该得到额外的资金（并且要为自己的决定高度负责），而另一半学生则被告知公司中另一位财务官员已经决定哪个部门将获得这笔资金（且不需要为自己的决定负责）。然后大约半数的学生被告知在接下来的5年里得到资金的部门的确比未得到资金的部门

表现得更好（即，决策的确带来了积极的结果），而半数的学生则被告知相反的结果（即，决策带来了消极的结果）。

在该实验的第二部分，学生们得知，公司的管理者对公司价值的重新评估将决定另外 2 000 万美元的研发资金如何分配，并且学生们可以他们认为合适的任意方式将这笔钱在消费和工业部门之间进行分配。斯塔（1976）发现了“诱捕”现象——对一个失误行为持续增加投入。如图 21.2 所示，开始选择不成功并且要担负个人责任的学生平均分配了大约 1 300 万美元到以前选择的部门——比其他学生多出约 400 万美元。当责任重大时，失误导致了更多的投入，而非更少。

斯塔（1976）的实验激发了很多后续的研究，并且在他的研究之后，出现了几种对诱捕行为的理论分析（其中最优的两个是 Brockner & Rubin, 1985, 及 Staw & Ross, 1987）。尽管对诱捕的研究起步不久，但已经有实验证据表明：（1）在被动地维持现状的情境中（例如自动的再投资计划）比在主动选择继续与否的情况下更加有诱捕倾向（Brockner, Shaw, & Rubin, 1979）；（2）诱捕在竞争的社会条件下比非竞争的社会条件下来得激烈，这一点至少对男性而言是成立的（Rubin, Brockner, Small-Weil, & Nathanson, 1980）；（3）诱捕不仅在个

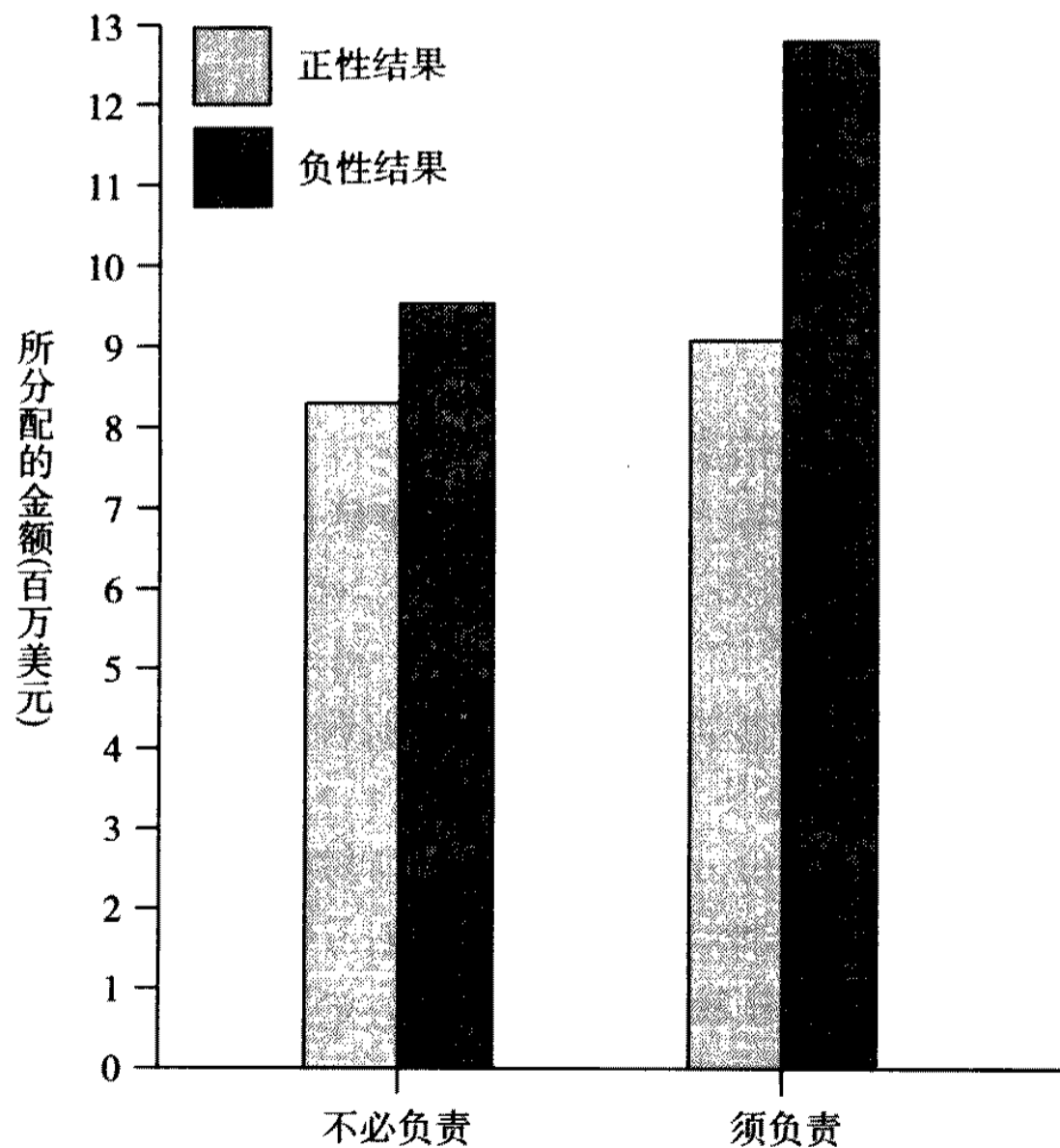


图 21.2 图为诱捕的一个例子。巴里·斯塔（1976）发现，相对于不必对先前的投资负责的学生以及先前投资获益的学生而言，起初投资失败的学生在后来的投资中会投入更多的资金在先前的失败项目上面。

体中，而且在群体中也会出现（Bazerman, Giuliano, & Appelman, 1984），尽管这可能仅对女性成立（Brockner & Rubin, 1985）。

在恋爱关系中也有诱捕现象存在。卡里尔·拉斯布尔特（Caryl Rusbult, 1980）发现大学生在角色扮演的实验中更加忠实于恋爱伴侣——并且较少可能与他人约会——如果他们的关系持续了一年而不是一个月。因此，在其他条件相同的情况下，学生在该关系中投入的时间是与他们未来的忠诚度直接相关的。

胜利解脱

尽管陷阱难以逃脱，但它们很少能够永远持续。最终，等待的人们挂断了电话。公司官员停止把钱投入到不良项目中。令人不悦的恋爱关系解除了。通常，问题不在于行为陷阱会永远抓住受害者不放，而是回头看时，人们希望他们更早地从陷阱中逃离出来。*

幸运的是，有一些方法可供减弱或避免诱捕（请回顾：Brockner & Rubin, 1985；Cross & Guyer, 1980；Staw & Ross, 1987）。一种由巴里·斯塔和杰里·罗斯（Barry Staw & Jerry Ross, 1987）提出的技术是在做出一个承诺之前“把结束的成本明确化”——就是说，在投入一项长期风险之前要清楚地考虑中止行为的代价。实验证据表明，事先将投入成本清晰化可以使诱捕减弱或者消除（Brockner, Rubin, & Lang, 1981；Nathanson et al., 1982）。

在布罗克纳和鲁宾（1985, p. 203）关于诱捕的专著中，他们建议决策者在任何可能的时候事先设定极限，并且用下述方法利用所设定的极限：

并不是说达到极限所设定的数量就马上主动退出，决策者应该利用他们所设定的极限点，作为一个重新衡量继续或终止行为的决策时间点，而无论他们事先已经投入了多少。也就是说，如果个人决定继续投资而超过事先设定的极限的话，应该以对未来的（而不是过去的）成本—效益分析为基础。

在商业情境下，斯塔和罗斯（1987年3~4月）建议应该先回答这样一个问题：“如果今天我是首次从事这个工作，发现这个项目正在进行中，我会支持它还是放弃它？”这个问题也可以很容易地运用到其他非商业情境下（比如，“假如今天我是第一次遇到这个人，我会被他吸引吗？”）

另一个技巧是让不同的人进行最初的和后续的决策（Bazerman, Giuliano,

* 有些集体陷阱，比如说人口膨胀、饥饿、环境恶化以及军事冲突等，可能是这条法则的例外。这些问题可能有永久性的烦人特征，并且有待时间检验人类是否有能力解决它们。

& Appelman, 1984; Staw & Ross, 1987)。例如，商业贷款可以由一位银行官员发放，然后由另一位人员来审查是否可以继续签约。这样做的好处在于，后来的决定是由不必为先前错误负责的人做出的（因此基本没有提高投入的理由）。但是，这种方法也存在缺点，即决策的不连贯性以及“组织记忆”（institutional memory）的潜在损失。

结 论

行为陷阱是生活中普遍存在的一部分，如果不加以关注，可能会导致严重后果。斯塔（1981）提出，很多有害的个人决策或者公共政策都是由连续的、不断升级的投入所导致（比如在越战中的种种事例）。普拉特（1973，p. 651）则更加极端，他说：“陷阱代表了当今几乎所有难以驾驭的，大规模的城市、国家以及国际间出现的问题。”

但是，陷阱并不总是不好的。正如布罗克纳和鲁宾（1985）所说，有些情况下，人们刻意让自己进入陷阱中。比如恢复中的嗜酒者，以前的吸烟者、暴饮暴食者常常鼓起勇气在陷阱中等待着，目的是可以使他们“陷在”健康的生活方式里。

当陷阱是合意的时候，决策者应该：

- ✓ 回避有关诱捕成本的信息。
- ✓ 尽量不要做出限定或评估继续这样做将花费的成本。
- ✓ 做出一个公开的、坚持下去的承诺。
- ✓ 与那些有相同目标的人们竞赛。

与本书中讨论的直觉以及偏好相同，行为陷阱并不总是好的或者坏的，心理学研究的目的也并不在于评判这些问题的好坏。而且，对诱捕行为的研究目的——以及一般性的决策研究——是更加局限的。它的目的是，让我们进一步了解决策过程是如何操作的，并且，在这样做的过程中，让我们的决策质量得到提高。

编后记：回顾

在读完了长达 21 章令人费解的问题和不合常理的结果后，现在我们需要回顾并反思一下以上种种的含义。决策者们是否根本就毫无理性可言？那些用于纠正决策偏差的策略是否有什么潜在的共通之处？这本书里面各种结论的局限在哪里？而研究决策的学者们对于上述理论又是如何应用的？以下的回顾部分将一一作出解答。

偏差不一直都是非理性的

在第 21 章中提到过，心理学研究的目的不在于判定偏差是好是坏。本书中所讨论的大部分研究结果同时具有积极和消极的一面。举例来说，有大量证据表明一个膨胀的自我形象有助于身心健康（Taylor & Brown, 1998）。尽管本书的章节中反映出“华而不实的非理性”在目前的决策研究中占主导地位（Lopes, 1991），但仍需强调的是，决策和判断中的偏差并不一定是有害或者非理性的。洛拉·洛佩斯（Lola Lopes, 1982）指出，在大多数时候，理性的含义不仅止于正确性。理性的策略可能就是决策者所期望获得的结果。

在一些情境中，决策者可能首要考虑的是如何避免灾难性事件的发生，其次才是效用最大化。在另一些情境中，他们可能选择让朋友和家人而非他们自己获利。而在其他情境中，他们会和那些理想化的“理性角色”一样，依照期望效用理论去追求效用最大化。决策者们可能追求任意数量的目标，那么仅仅用一个标准化的标尺来给理性下定义就不准确。

此外，即使在某些期望效用框架中决策者充当了“理性角色”，我们也没有理由假定最优的选择总是能够将效用最大化。洛佩斯已经在基本归因误差（见第 16 章）中指出了这一点。根据洛佩斯（1982, p. 633）的观点，基本归因误差“是一种倾向，即我们总是会把他人的行为归因于稳定的人格变量，而不是那些短暂且经常是模糊的情境因素。但有可能这根本不是一种误差，无论是在物理环境中还是在社会环境中，我们都必须尽可能好地预期和控制他人行为对自身的影响。因此，如果某些影响是可预期的，哪怕只是通过一些个体的某几个行为而微弱地表现出来，是否能把它找出来对于我们就显得很重要。”因此，这种倾向于将行为归因于个性的方式，也许会得到更为理性的归因结果。

究其原因，让我们来考虑一下图 A.1 中虚构且高度简化的损益分析图示。如果洛佩斯关于觉察个性因素重要性的假说是正确的，那么正确辨认个性因素将得到较高的收益（+3），而遗漏了个性因素将会得到较高的惩罚（-3）。另一方面，情境因素被精确地觉察（+1）或被错失（-1）所带来的损益不大。

为了讨论的方便，我们假设情境因素的发生几率约为 60%，进一步假设决策者不知道应该做出何种归因。那么根据图 A.1 中的情况，怎样才能做出一个“理性”的归因呢？

在此种情况下，个性因素的归因得到最大的平均报偿。习惯将行为归因于个性因素的决策者将获得平均 +.60 的效用（有 60% 的可能当时的报偿为 -1，但有 40% 的可能当时的报偿为 +3）；然而，习惯将行为归因于情境因素的决策者将获得平均 -.60 的效用（有 60% 的可能当时的报偿为 +1，但有 40% 的可能当时的报偿为 -3）。即使情境归因的正确率高于个性归因，效用最大化的归因方式还是将行为归因于个性因素（即基本归因误差）。

当然，图 A.1 中的报偿仅仅是假设中的。我们无法估计一生中无数次归因中成功与失败的效用，也无法评定情境因素和个性因素出现的相关频次。计算机模拟的直觉决策表明，我们为某些偏差所付出的代价常常不算大。举例而言，沃伦·索恩盖特（Warren Thorngate, 1980）发现，忽视了概率信息的直觉几乎和基于概率的决策效果一样好。唐·克莱门茨（Don Kleinmuntz, 1985）发现，当重复的判断任务提供了反馈时，忽略了基本比率信息的直觉表现也相

		行为的实际原因	
		个性	情境
归因	个性	+3	-1
	情境	-3	+1

图 A.1 图中的数字代表了四种不同的归因结果的想像值。根据这一模型，基本归因误差（右上角的单元格）所造成的损失并没有遗漏个性归因所造成的损失（左下角的单元格）大，而正确进行个性归因（左上角的单元格）所带来的奖励比正确进行情境归因（右下角的单元格）所带来的奖励要大。在这些情况下，决策者对行为过分进行个性归因是情有可原的。

当好。由此得出结论，在无法正常取得概率及基本比率信息的情况下，直觉有可能带来比标准化决策更有吸引力的选择。

以上的结论并非告诉决策者们发生偏差优于没有偏差。很多事例都可以证明决策和判断中的偏差可能导致重大问题（例如医疗中的误诊）。但是，上述发现试图将偏差置于决策和判断中的适当角度，并且证明，偏差不总意味着非理性。

一个普遍的思路

尽管理性比正确意义更大，但在绝大多数的情境下，决策者们仍然努力做出最为精准的判断。因此，决策研究者们一直致力于找出有效的“去偏差”技巧（*debiasing*）（即能够减少判断中偏差和误差的技巧）。在某些场合，只需要对潜在问题进行简单了解就可以避免上述问题。例如，熟悉和了解诱捕原理的人们不大可能被诱捕（Nathanson et al., 1982）。但多数情况下，仅凭警告是不足以消除偏差和误差的。

本书已就各种减少偏差技巧进行了讨论——消除过度自信，事后之明的偏差，框架效应，集一思维，归因偏差等等。很自然地，我们想知道这些去偏差技巧是否具有某些共性。

冒着过分简单化的风险，但没有刻意将去偏差策略缩减为共通主题，我们发现确实存在一种共性。

多数行之有效的去偏差技巧都考虑到了看待问题时使用不同的视角。

举例来讲，认真考虑发生谬误的原因通常能够降低过度自信的程度。事后之明的偏差能够通过事前思考不同结果产生的原因而得到降低甚至消除。当同样的问题被放在不同的框架中考虑时，框架效应自然会减轻。如果群体中存在一个能够和大家唱反调的角色，集一思维就能够避免。参与者—旁观者在归因中的差别能够通过交换二者视角的方法来消除。在所有这些情境中，“考虑对立面”都能够为判断带来更高的准确度。

与上面的主题相一致的是，还有证据表明移情能够消除某些社会判断中的偏差。迈克尔·斯托姆斯（Michael Storms, 1973）通过研究认为参与者—旁观者的归因差别能够通过转变二者的视角而消除（见第16章）。两年之后，丹尼斯·里甘和朱迪思·托腾（Dennis Regan & Judith Totten, 1975）发现对移情的指示也有相同效果。在这个研究中被试的指导语有两种，一种是指示被试只简单观察一场二人对话中的一人（命名为玛格丽特的某人），另一种是指示被试对她进行移情（想像玛格丽特在对话中的感受）。

里甘和托腾（1975）发现，简单观察被试对玛格丽特行为的归因多倾向于个性因素而非情境因素，而对玛格丽特发生移情的被试对她行为的归因则倾向

于情境因素而非个性因素。换句话说，对移情的指示产生了与之前斯托姆斯得到的结论相同的效果：移情消除了参与者—旁观者的归因差别。

一年后，鲁思·加尔珀（Ruth Galper）发表了类似研究，她也获得了相同的结论。加尔珀（1976，p. 333）对此归纳如下：

未被指示移情的参与者表现出典型的“观察者偏差”，认为人格因素的重要性高于情境因素。另一方面，被指示“移情”的被试表现出反转性的归因，他们认为人格因素的重要性低于情境性因素。而斯托姆斯（1973）证明了一个字面上的视角改变能够系统地影响参与者和旁观者的归因偏差，本研究证明了一个虚拟的“视角改变”能够引发旁观者不经意的“类似参与者”的归因。

这些研究都指出在最终判断前，参考其他相异观点的重要性（Anderson, 1982; Lord, Lepper, & Preston, 1984）。尽管没有一种去偏差技巧是万能的，但在看待问题时使用不同的视角通常能够提高决策和判断的质量和水平。

对于不好的消息……

抛开对去偏差技巧和其他主题的决策研究的种种有用之处不谈，本书的很多结论仍然存在许多局限。某些局限是由本书主题覆盖范围有限造成的，而另外一些则来源于决策研究本身所存在的问题。

正如前言中说过的那样，本书并未对决策和判断的研究情况作出详尽描述。而且书中的很多主题是有意选取的一些令人震惊的、娱乐的和能够激起大家兴趣的话题。很多相当重要的研究领域仅以最简略的方式做了说明，并且从实际出发，一些非常好的研究根本没有提及。以下举两个例子，来说明这些局限是怎样影响研究结论的普遍性的。

首先，第17章简略提过“承担责任和义务”对于决策的强大影响力。对自身行为有责任感的决策者不会像没有责任感的人那样过度自信（Tetlock & Kim, 1987），表现出较低的社会性懈怠（Weldon & Gargano, 1983），以及更高的复杂问题分析能力（Hagafors & Brehmer, 1983; Tetlock, 1983）。正如菲利普·泰特洛克等人（Philip Tetlock, Linda Skitka, & Richard Boettger, 1989, p. 640）所指出的：“一个人的思维模式部分取决于他为什么思考。”任何关于决策的整体讨论都必须考虑到上述关系，且被试对自己行为的负责任程度应当成为为研究结论评分的一个指标。

第二，决策和判断在相当大程度上取决于诸如决策者可使用的时间和决策者的情绪这类具有环境特征的因素。当时间紧迫时，决策者选用简单化的策略，只注意到少量线索，并作出相对而言低风险的选择（Ben Zur & Breznitz, 1981; Rothstein, 1986; Wright, 1974）。而情绪好的决策者可能更富有创造力

(Isen, Daubman, & Nowicki, 1987), 觉得负面事件的发生概率相对较低或不可能发生 (Johnson & Tversky, 1983; Salovey & Birnbaum, 1989; Wright & Bower, 1992), 并会冒其他人不乐意接受的低水平风险 (Isen & Geva, 1987; Isen & Patrick, 1983)。这些结论再一次说明, 我们从决策研究中得到的结论应被妥善地评价, 以防止过度概括化的发生。

本书中很多结论的局限还来源于决策研究本身的特性。首先, 纵观第4章到第6章的内容, 我们看到很多关于决策和判断的研究都是与内容相关联的 (Schum, 1990)。理查德·格里格和詹姆斯·考克斯 (1982) 通过随机安排大学生遇到两类问题的实验, 有力地证明了这类局限。实验条件为 (1) 第20章中讨论过的标准4卡片问题, (2) 一个4卡片问题的修改版本: 用短语“DRINKING A BEER” (喝啤酒), “DRINKING A COKE” (喝可乐), “22 YEARS OF AGE” (22周岁) 和 “16 YEARS OF AGE” (16周岁) 来代替“E”、“K”、“4”和“7”四个字母 (规则为如果一个人在喝啤酒, 那么这个人一定超过19周岁)。当使用喝啤酒的情境时, 大约3/4的学生能正确解决问题, 而当以字母和数字出题时, 所有学生都失败了。措辞的改变能够影响共变判断 (covariation judgment) (Crocker, 1982) 和联合谬误 (conjunction fallacy) (Fiedler, 1988)。

使用大学生充当被试的实验室研究方法也是目前决策和判断研究中的一大局限。尽管这些研究颇有价值, 但实验室设计的各种任务无法获取与自然环境下相同的决策过程。举例而言, 埃贝·埃布森 (Ebbe Ebbesen) 和 Vladimir Konečni (1975) 发现, 地方法官及高法法官在真实情境和虚拟情境下所做出的保释判决差异巨大。第19章中曾经指出, 专家级桥牌选手和专业气象预报员比一般人能够做出准确得多的判断 (Keren, 1987; Murphy & Brown, 1984; Murphy & Winkler, 1984)。另一方面, 许多研究显示, 专家们与普通大学生相比, 在实验中表现出了相同的偏差 (Dubé-Rioux & Russo, 1988), 或仅仅是偏差的程度略低 (Christensen-Szalanski, Beck, Christensen-Szalanski, & Koepsell, 1983; Smith & Kida, 1991)。很明显, 我们需要更多地针对专家决策者进行研究。

决策研究的另一个问题是缺乏跨文化取向。绝大多数的研究都局限于北美洲和欧洲, 极少研究世界其他地区的人是如何做决策的, 由此导致的结果是, 我们很难判定这些结论具有多大范围的适用性 (Fletcher & Ward, 1988)。以下的一些例子描述了在取得肯定结论过程中存在的困难。

- 唐纳德·泰勒和韦施纳·贾格 (Donald Taylor & Vaishna Jaggi, 1974) 发现了信仰印度教的印度人对偏差的归因方式与美国人相似, 但其他学者的研究发现美国人和印度人在归因风格上存在差异。与美国人倾向于出现基本归因误差相比, 印度人习惯将行为归因于情境因素而非个性因素

(Miller, 1984)。

- 詹姆斯·斯托纳 (James Stoner, 1968) 认为已经证明了美国、英国和以色列被试的风险转移模式, 但其他学者在美国、英国和亚洲被试中发现了他们在概率判断—存在的差异 (Wright et al., 1978; Yates et al., 1989)。
- 塔哈·阿米尔 (Taha Amir, 1984) 发现科威特人与美国人在阿施范例上的测试结果比较类似, 而 Noriyuki Matsuda (1985) 发现日本人在这种测试上的得分水平要高于美国人。因而, 上一年还认为是具有普遍意义的结论, 在下一年就变成了与文化有关的结论。
- 还有研究显示日本被试在归因中没有表现出自私的归因偏差, 这与美国、印度、南非和南斯拉夫的研究结果有差异 (Chandler, Shama, Wolf, & Planchard, 1981)。

上述结论仅仅指出了跨文化决策研究的复杂性中很小的一部分。在建立一个庞大的研究体系前, 向任何非西方文化国家提供决策研究的成果及建议都需要非常谨慎。

决策研究者易犯的错误

决策和判断的研究通常对另一个问题非常头痛——虽然很少提及, 但有深远的内涵。

决策和判断的研究人员与实验中的被试具有相似的出现偏差和误差的可能。

这个无法回避的问题的结果就是, 所有关于决策和判断的研究——毕竟这也是一系列决策和判断结果——都容易出现偏差和误差。

举例而言, 正如被试可能对他们的回答过分自信, 决策的研究者也可能对他们的理论过度自信。同理, 研究者易受认知不协调和投资陷阱的影响, 并且有理由假设正是前期工作的沉没成本诱使研究者进入陷阱。正如布罗克纳和鲁宾 (1985, p. 260) 在他们关于圈套和陷阱的著作的结束语中提到的: “优秀的陷阱研究者——和成为我们实验研究对象的决策者一样——必须能够在更大的框架下评估自己的行为。”

巴鲁克·菲施霍夫 (1991) 进一步观察到, 若研究者想要并期待着观察到某种现象, 就有可能过高评价其频率 (据推断, 这是由于与理论匹配的数据比不匹配的数据更容易被研究者注意到)。结果就是出现易得性偏差, 匹配数据出现的频次及 (或) 可能性就容易被过高评价, 由此研究者可能会产生一个虚构的——至少是夸大的——理论与数据之间的相关。这种可能性与选择性知觉、偏差的同化作用以及在不可信证据下对错误信念的盲目坚持等现象都相符 (例如, Lord, Ross, & Lepper, 1979; Ross, Lepper, & Hubbard, 1975)。

也许研究中最严重的问题在于，研究者们倾向于某些固定化的偏见和自我实现的预言（Jenkins, 1981; Rosenthal, 1976）。罗伯特·罗森塔尔和克米特·福德（Kermit Fode）（1963）作出这个题目下最早也是最戏剧化的研究。罗森塔尔和福德告诉12个实验者，他们将以一些“迷宫—智慧型”老鼠或“迷宫—蠢笨型”老鼠为实验对象进行实验，而实际上所谓的标签都是随机分配的。每个实验者分配到5只老鼠进行实验，要求实验者记录这些老鼠在50次走迷宫任务中的实验结果。

对50次实验结果取平均，罗森塔尔和福德发现那些“迷宫—智慧型”老鼠比“迷宫—蠢笨型”老鼠正确走完迷宫的数量高出50%。因此，即使对结果的预期是随机的，实验者们报告的结果实际上只是他们心中的预期。

总结一下，决策和判断研究常遭受自相矛盾的困扰：当其结论是正确的（即存在偏差和误差），其结论在某些程度上也受到偏差和误差的影响。与此同时，当其结论不正确（即不存在偏差和误差），其结论仍然可能有误。两种结果都无法摆脱偏差和误差的出现。这种自相矛盾的困境与奥地利逻辑学家 Kurt Gödel 之不完全法则（Incompleteness Theorem）的某些方面类似，他认为任何纯使用算术并规范化的逻辑系统，都存在内在矛盾。

当然，与其他研究领域相比，在决策研究中，偏差和误差并非起着更大的破坏作用。关键在于，不管多么不情愿，研究者在考虑普通人易犯的错误时必须把自己考虑在内。本书在结尾处强调以上内容的目的在于进一步明确这一方向，即使只以斯金纳（1980）描述过的方式：



“终于到了可以做决策的时候了！”

我习惯于通过描述一个解释人类行为的演讲者（包括其他演讲者的行为）来表现出一个行为主义学家对待他自己的态度，然后离开舞台。接着，他从舞台侧面探出头来说：“我也喜欢这样！”

参 考 文 献

- Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.
- Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35*, 303-314.
- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*, 888-918.
- Allais, P. M. (1953). The behavior of rational man in risk situations—A critique of the axioms and postulates of the American School. *Econometrica, 21*, 503-546.
- Allison, S. T., & Messick, D. M. (1985). The group attribution error. *Journal of Experimental Social Psychology, 21*, 563-579.
- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91*, 112-149.
- Allport, G. W. (1954). The historical background of modern social psychology. In G. Lindzey (Ed.), *The handbook of social psychology* (Vol. 1). Reading, MA: Addison-Wesley.
- Alvarez, L. W. (1965). A pseudo experience in parapsychology. *Science, 148*, 1541.
- American Psychological Association. (1989). Ethical principles of psychologists. *American Psychologist, 45*, 390-395.
- Amir, T. (1984). The Asch conformity effect: A study in Kuwait. *Social Behavior and Personality, 12*, 187-190.
- Anderson, C. A. (1982). Inoculation and counterexplanation: Debiasing techniques in the perseverance of social theories. *Social Cognition, 1*, 126-139.
- Anderson, C. A. (1983). Imagination and expectation: The effect of imagining behavioral scripts on personal intentions. *Journal of Personality and Social Psychology, 45*, 293-305.
- Anderson, N. H. (1965). Primacy effects in personality impression formation using a generalized order effect paradigm. *Journal of Personality and Social Psychology, 2*, 1-9.
- Argote, L., Seabright, M. A., & Dyer, L. (1986). Individual versus group use of base-rate and individuating information. *Organizational Behavior and Human Decision Processes, 38*, 65-75.

- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35, 124-140.
- Arkes, H. R., Christensen, C., Lai, C., & Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, 39, 133-144.
- Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General*, 112, 117-135.
- Arkes, H. R., Wortmann, R. L., Saville, P. D., & Harkness, A. R. (1981). Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology*, 66, 252-254.
- Aronson, E. (1969). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 1-34). New York: Academic Press.
- Aronson, E. (1972). *The social animal*. San Francisco: W. H. Freeman and Company.
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, 59, 177-181.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258-290.
- Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (Ed.), *Groups, leadership and men*. Pittsburgh: Carnegie Press.
- Asch, S. E. (1955, November). Opinions and social pressure. *Scientific American*, pp. 31-35.
- Asch, S. E. (1956). Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70 (9, Whole No. 416).
- Ausubel, N. (Ed.). (1948). *A treasury of Jewish folklore: Stories, traditions, legends, humor, wisdom and folk songs of the Jewish people*. New York: Crown Publishers.
- Ayton, P., Hunt, A. J., & Wright, G. (1989). Psychological conceptions of randomness. *Journal of Behavioral Decision Making*, 2, 221-238.
- Bakan, P. (1960). Response tendencies in attempts to generate random binary series. *American Journal of Psychology*, 73, 127-131.
- Bandura, A. (1982). Self-efficacy: Mechanism in human agency. *American Psychologist*, 37, 122-147.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Barclay, S., & Beach, L. R. (1972). Combinatorial properties of personal probabilities. *Organizational Behavior and Human Performance*, 8, 176-183.
- Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, 9, 396-406.
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44, 211-233.

- Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn*. Chicago: University of Chicago Press.
- Batson, C. D. (1975). Attribution as a mediator of bias in helping. *Journal of Personality and Social Psychology*, 32, 455-466.
- Batson, C. D., & Marz, B. (1979). Dispositional bias in trained therapists' diagnoses: Does it exist? *Journal of Applied Social Psychology*, 9, 476-489.
- Baxter, T. L., & Goldberg, L. R. (1987). Perceived behavioral consistency underlying trait attributions to oneself and another: An extension of the actor-observer effect. *Personality and Social Psychology Bulletin*, 13, 437-447.
- Bazerman, M. H., Giuliano, T., & Appelman, A. (1984). Escalation of commitment in individual and group decision making. *Organizational Behavior and Human Performance*, 33, 141-152.
- Bedau, H. A., & Radelet, M. L. (1987). Miscarriages of justice in potentially capital cases. *Stanford Law Review*, 40, 21-179.
- Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, 30, 961-981.
- Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, 33, 1-27.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). New York: Academic Press.
- Ben Zur, H., & Breznitz, S. J. (1981). The effects of time pressure on risky choice behavior. *Acta Psychologica*, 47, 89-104.
- Bernoulli, D. (1738/1954). Exposition of a new theory on the measurement of risk, translated by Louise Sommer. *Econometrica*, 22, 22-36.
- Beyth-Marom, M. (1982). Perception of correlation reexamined. *Memory & Cognition*, 10, 511-519.
- Bierbrauer, G. (1973). *Effect of set, perspective, and temporal factors in attribution*. Unpublished doctoral dissertation, Stanford University.
- Bishop, G. F., Oldendick, R. W., Tuchfarber, A. J., & Bennett, S. E. (1980). Pseudo-opinions on public affairs. *Public Opinion Quarterly*, 44, 198-209.
- Blascovich, J., Ginsburg, G. P., & Howe, R. C. (1975). Blackjack and the risky shift, II: Monetary stakes. *Journal of Experimental Social Psychology*, 11, 224-232.
- Block, J. R., & Yuker, H. E. (1989). *Can you believe your eyes?* New York: Gardner Press.
- Blodgett, R. (1983, November). Against all odds. *Games*, pp. 14-18.
- Bond, C. F., Jr., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, 94, 265-292.
- Borgida, E., & Nisbett, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology*, 7, 258-271.
- Borrelli, S., Lockerbie, B., & Niemi, R. G. (1987). Why the Democrat-Republi-

- can partisanship gap varies from poll to poll. *Public Opinion Quarterly*, 51, 115-119.
- Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect on the preference-reversal phenomenon of using choice indifference. *Journal of Economic Behavior and Organization*, 13, 193-212.
- Bourne, L. E., Jr., & Guy, D. E. (1968). Learning conceptual rules: II. The role of positive and negative instances. *Journal of Experimental Psychology*, 77, 488-494.
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.
- Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, 36, 917-927.
- Brockner, J., & Rubin, J. Z. (1985). *Entrapment in escalating conflicts: A social psychological analysis*. New York: Springer-Verlag.
- Brockner, J., Rubin, J. Z., & Lang, E. (1981). Face-saving and entrapment. *Journal of Experimental Social Psychology*, 17, 68-79.
- Brockner, J., Shaw, M. C., & Rubin, J. Z. (1979). Factors affecting withdrawal from an escalating conflict: Quitting before it's too late. *Journal of Experimental Social Psychology*, 15, 492-503.
- Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, 75, 631-661.
- Brown, E., Deffenbacher, K., & Sturgill, W. (1977). Memory for faces and the circumstances of encounter. *Journal of Applied Psychology*, 62, 311-318.
- Bruner, J. S., & Postman, L. J. (1949). On the perception of incongruity: A paradigm. *Journal of Personality*, 18, 206-223.
- Budiansky, S. (1988, July 11). The numbers racket: How polls and statistics lie. *U.S. News & World Report*, pp. 44-47.
- Campbell, J. D., & Tesser, A. (1983). Motivational interpretations of hindsight bias: An individual difference analysis. *Journal of Personality*, 51, 605-620.
- Cantril, H. (1940). Experiments in the wording of questions. *Public Opinion Quarterly*, 4, 330-332.
- Carden, G. (1990, September 25). Strangers, but they share unbelievable mirror-image lives. *National Enquirer*, p. 19.
- Carroll, J. S. (1978). The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. *Journal of Experimental Social Psychology*, 14, 88-96.
- Centor, R. M., Dalton, H. P., & Yates, J. F. (1984). *Are physicians' probability estimates better or worse than regression model estimates?* Paper presented at the Sixth Annual Meeting of the Society for Medical Decision Making, Bethesda, MD.
- Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments. *Journal of Personality and Social Psychology*, 50, 492-501.
- Chandler, T. A., Shama, D. D., Wolf, F. M., & Planchard, S. K. (1981). Multi-

- attributional causality: A five cross-national samples study. *Journal of Cross-Cultural Psychology*, 12, 207-221.
- Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Behavior*, 6, 151-155.
- Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 72, 193-204.
- Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.
- Chapman, L. J., & Chapman, J. P. (1971, November). Test results are what you think they are. *Psychology Today*, pp. 18-22, 106-110.
- Cheng, P. W., & Novick, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology*, 58, 545-567.
- Christensen, D., & Rosenthal, R. (1982). Gender and nonverbal decoding skill as determinants of interpersonal expectancy effects. *Journal of Personality and Social Psychology*, 42, 75-87.
- Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. *American Psychologist*, 39, 75-78.
- Christensen-Szalanski, J. J. J., Beck, D. E., Christensen-Szalanski, C. M., & Koepsell, T. D. (1983). Effects of expertise and experience on risk judgments. *Journal of Applied Psychology*, 68, 278-284.
- Christensen-Szalanski, J. J. J., & Bushyhead, J. B. (1981). Physicians' use of probabilistic information in a real clinical setting. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 928-935.
- Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 48, 147-168.
- Clifford, B. R., & Scott, J. (1978). Individual and situational factors in eyewitness testimony. *Journal of Applied Psychology*, 63, 352-359.
- Clymer, A. (1982, May 6). Nuclear issue yields run on pollsters. *New York Times*, p. B14.
- Cohen, B. L., & Lee, I. (1979). A catalog of risks. *Health Physics*, 36, 707-722.
- Cohen, J., Chesnick, E. I., & Haran, D. (1971). Evaluation of compound probabilities in sequential choice. *Nature*, 232, 414-416.
- Combs, B., & Slovic, P. (1979). Newspaper coverage of causes of death. *Journalism Quarterly*, 56, 837-843, 849.
- Converse, P. E., & Schuman, H. (1970, June). "Silent majorities" and the Vietnam war. *Scientific American*, pp. 17-25.
- Coombs, C. H. (1975). Portfolio theory and the measurement of risk. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes*. New York: Academic Press.
- Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin*, 90, 218-244.

- Coren, S., & Miller, J. (1974). Size contrast as a function of figural similarity. *Perception and Psychophysics*, 16, 355-357.
- Corey, S. M. (1937). Professed attitudes and actual behavior. *Journal of Educational Psychology*, 28, 271-280.
- Costanza, R. (1984). Review essay: The nuclear arms race and the theory of social traps. *Journal of Peace Research*, 21, 79-86.
- Craig, K. D., & Prkachin, K. M. (1978). Social modeling influences on sensory decision theory and psychophysiological indexes of pain. *Journal of Personality and Social Psychology*, 36, 805-815.
- Crandall, V. J., Solomon, D., & Kellaway, R. (1955). Expectancy statements and decision times as functions of objective probabilities and reinforcement values. *Journal of Personality*, 24, 192-203.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90, 272-292.
- Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin*, 8, 214-220.
- Cross, J. G., & Guyer, M. J. (1980). *Social traps*. Ann Arbor: University of Michigan Press.
- Darley, J. M., & Batson, C. D. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27, 100-108.
- Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist*, 35, 867-881.
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, 8, 377-383.
- Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. *Psychological Review*, 80, 97-125.
- Dawes, R. M. (1975). The mind, the model, and the task. In F. Restle, R. M. Shiffrin, N. J. Castellan, H. R. Lindman, & D. B. Pisoni (Eds.), *Cognitive theory* (Vol. 1). Hillsdale, NJ: Erlbaum.
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31, 169-193.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.
- "Death Odds." (1990, September 24). *Newsweek*, p. 10.
- Deci, E. L. (1975). *Intrinsic motivation*. New York: Plenum Press.
- Deffenbacher, K. A. (1980). Eyewitness accuracy and confidence. *Law and Human Behavior*, 4, 243-260.
- Diener, D., & Thompson, W. B. (1985). Recognizing randomness. *American Journal of Psychology*, 98, 433-447.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285-290.
- Doob, A. N., Carlsmith, J. M., Freedman, J. L., Landauer, T. K., & Tom, S., Jr.

- (1969). Effect of initial selling price on subsequent sales. *Journal of Personality and Social Psychology*, 11, 345-350.
- Dubé-Rioux, L., & Russo, J. E. (1988). An availability bias in professional judgment. *Journal of Behavioral Decision Making*, 1, 223-237.
- Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping blacks. *Journal of Personality and Social Psychology*, 43, 590-598.
- Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology*, 58, 568-581.
- Dunning, D., & Parpal, M. (1989). Mental addition versus subtraction in counterfactual reasoning: On assessing the impact of personal actions and life events. *Journal of Personality and Social Psychology*, 57, 5-15.
- Duval, S., & Wicklund, S. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, 9, 17-31.
- Ebbesen, E. B., & Konečni, V. J. (1975). Decision making and information integration in the courts: The setting of bail. *Journal of Personality and Social Psychology*, 32, 805-821.
- Edeal, G. H. (1950, March 27). Why the choir was late. *Life*, p. 19.
- Eddy, D. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: John Wiley and Sons.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, 32, 53-88.
- Einhorn, H. J., & Hogarth, R. J. (1986). Backward and forward thinking in decision making. *Selected Paper No. 62*. Chicago: Graduate School of Business, University of Chicago.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643-669.
- Esser, J. K., & Lindoerfer, J. S. (1989). Groupthink and the space shuttle Challenger accident: Toward a quantitative case analysis. *Journal of Behavioral Decision Making*, 2, 167-177.
- Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal*, 21, 55-60.
- Feldman, J. M. (1986). A note on the statistical correction of halo error. *Journal of Applied Psychology*, 71, 173-176.
- Feldman, R. S., & Prohaska, T. (1979). The student as Pygmalion: Effect of student expectation on the teacher. *Journal of Educational Psychology*, 71, 485-493.
- Feldman, R. S., & Theiss, A. J. (1982). The teacher and the student as Pyg-

- malions: Joint effects of teacher and student expectations. *Journal of Educational Psychology*, 74, 217-223.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology*, 58, 203-210.
- Feynman, R. P. (1988, February). An outsider's inside view of the Challenger inquiry. *Physics Today*, 26-37.
- Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle linguistic factors. *Psychological Research*, 50, 123-129.
- Fischhoff, B. (1977). Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 349-358.
- Fischhoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist*, 46, 835-847.
- Fischhoff, B., & Bar-Hillel, M. (1984). Diagnosticity and the base-rate effect. *Memory & Cognition*, 12, 402-410.
- Fischhoff, B., & Beyth, R. (1975). "I knew it would happen": Remembered probabilities of once-future things. *Organizational Behavior and Human Performance*, 13, 1-16.
- Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review*, 90, 239-260.
- Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., & Keeney, R. L. (1981). *Acceptable risk*. Cambridge, England: Cambridge University Press.
- Fischhoff, B., & Slovic, P. (1980). A little learning . . . : Confidence in multicue judgment tasks. In R. Nickerson (Ed.), *Attention and performance, VIII*. Hillsdale, NJ: Erlbaum.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 552-564.
- Fishburn, P. C. (1984). SSB utility theory and decision making under uncertainty. *Mathematical Social Sciences*, 8, 253-285.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Fletcher, G. J. O., & Ward, W. (1988). Attribution theory and processes: A cross-cultural perspective. In M. H. Bond (Ed.), *The cross-cultural challenge to social psychology*. Newbury Park, CA: Sage Publications.
- Försterling, F. (1989). Models of covariation and attribution: How do they relate to the analogy of analysis of variance? *Journal of Personality and Social Psychology*, 57, 615-625.
- Frenkel, O. J., & Doob, A. N. (1976). Post-decision dissonance at the polling booth. *Canadian Journal of Behavioural Science*, 8, 347-350.
- Frieze, I. H., Bar-Tal, D., & Carroll, J. S. (Eds.). (1979). *New approaches to social problems*. San Francisco: Jossey-Bass.

- Galper, R. E. (1976). Turning observers into actors: Differential causal attributions as a function of "empathy." *Journal of Research in Personality, 10*, 328-335.
- Gansberg, M. (1964, March 27). 37 who saw murder didn't call the police. *New York Times*, pp. 1, 38.
- Gill, S. (1947, March 14). How do you stand on sin? *Tide*, p. 72.
- Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Journal of Personality and Social Psychology, 17*, 295-314.
- Gliksman, A. (1986, February 13). Behind Moscow's fear of "Star Wars." *New York Times*, p. 27.
- Gmelch, G. (1978, August). Baseball magic. *Human Nature*, pp. 32-39.
- Goldberg, L. R. (1959). The effectiveness of clinicians' judgments: The diagnosis of organic brain damage from the Bender-Gestalt test. *Journal of Consulting Psychology, 23*, 25-33.
- Golding, S. L., & Rorer, L. G. (1972). Illusory correlation and subjective judgment. *Journal of Personality and Social Psychology, 80*, 249-260.
- Graham, T. W., & Kramer, B. M. (1986). The polls: ABM and Star Wars: Attitudes toward nuclear defense, 1945-1985. *Public Opinion Quarterly, 50*, 125-134.
- Greenberg, J., Williams, K. D., & O'Brien, M. K. (1986). Considering the harshest verdict first: Biasing effects on mock juror verdicts. *Personality and Social Psychology Bulletin, 12*, 41-50.
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. B. (1982). Self-relevant scenarios as mediators of likelihood and compliance: Does imagining make it so? *Journal of Personality and Social Psychology, 43*, 89-99.
- Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review, 69*, 623-638.
- Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology, 59*, 1128-1139.
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology, 73*, 407-420.
- Gross, E. J. (1964). The effect of question sequence on measures of buying interest. *Journal of Advertising Research, 4*, 40-41.
- Haberman, C. (1990, July 19). A child's nightmare in Italy is now the nation's shame. *New York Times*, pp. A1, A11.
- Hagafors, R., & Brehmer, B. (1983). Does having to justify one's judgments change the nature of the judgment process? *Organizational Behavior and Human Performance, 31*, 223-232.
- Hake, H. W., & Hyman, R. (1953). Perception of the statistical structure of a random series of binary symbols. *Journal of Experimental Psychology, 45*, 64-74.
- Hamilton, D. L., & Rose, T. L. (1980). Illusory correlation and the maintenance of stereotypic beliefs. *Journal of Personality and Social Psychology, 39*, 832-845.

- Hardin, G. (1968). The tragedy of the commons. *Science*, 162, 1243–1248.
- “Harper’s Index.” (1986, October). *Harper’s*, p. 11.
- Harris, R. J. (1973). Answering questions containing marked and unmarked adjectives and adverbs. *Journal of Experimental Psychology*, 97, 399–401.
- Hartley, E. (1946). *Problems in prejudice*. New York: King’s Crown Press.
- Harvey, J. H., Town, J. P., & Yarkin, K. L. (1981). How fundamental is the “fundamental attribution error”? *Journal of Personality and Social Psychology*, 40, 346–349.
- Hastie, R. (1986). Review essay: Experimental evidence on group accuracy. In B. Grofman and G. Owen (Eds.), *Information pooling and group decision making: Proceedings of the Second University of California, Irvine, Conference on Political Economy*. Greenwich, CT: Jai Press.
- Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology*, 49, 129–134.
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, 107, 311–327.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Henchy, T., & Glass, D. C. (1968). Evaluation apprehension and the social facilitation of dominant and subordinate responses. *Journal of Personality and Social Psychology*, 10, 446–454.
- Henslin, J. M. (1967). Craps and magic. *American Journal of Sociology*, 73, 316–330.
- Hershey, J. C., & Schoemaker, P. J. H. (1980). Risk taking and problem context in the domain of losses: An expected utility analysis. *Journal of Risk and Insurance*, 47, 111–132.
- Hewstone, M. R. C., & Jaspars, J. M. F. (1987). Covariation and causal attribution: A logical model for the intuitive analysis of variance. *Journal of Personality and Social Psychology*, 53, 663–672.
- Hill, G. W. (1982). Group versus individual performance: Are $N + 1$ heads better than one? *Psychological Bulletin*, 91, 517–539.
- Hilton, D. J., & Slugoski, B. R. (1986). Knowledge-based causal attribution: The abnormal conditions focus model. *Psychological Review*, 93, 75–88.
- Hilts, P. J. (1990, September 26). Major gains are cited for smokers who quit. *New York Times*, p. B4.
- Hintzman, D. L. (1969). Apparent frequency as a function of frequency and the spacing of repetitions. *Journal of Experimental Psychology*, 80, 139–145.
- Hippler, H., & Schwarz, N. (1986). Not forbidding isn’t allowing: The cognitive basis of the forbid–allow asymmetry. *Public Opinion Quarterly*, 50, 87–96.
- Hoch, S. J. (1984). Availability and interference in predictive judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 649–662.
- Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 719–731.

- Hofstadter, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern*. New York: Basic Books.
- Hogarth, R. M. (1975). Cognitive processes and the assessment of subjective probability distributions. *Journal of the American Statistical Association*, 70, 271-289.
- Hogarth, R. (1987). *Judgement and choice* (2nd ed.). New York: John Wiley and Sons.
- Hooke, R. (1983). *How to tell the liars from the statisticians*. New York: Marcel Dekker.
- Hornstein, H. A., Fisch, E., & Holmes, M. (1968). Influence of a model's feeling about his behavior and his relevance as a comparison other on observers' helping behavior. *Journal of Personality and Social Psychology*, 10, 222-226.
- Hunter, I. M. L. (1964). *Memory*. Middlesex, England: Penguin Books.
- Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology*, 10, 371-384.
- Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, 21, 329-335.
- Irwin, F. W., & Metzger, J. (1966). Effects of probabilistic independent outcomes upon predictions. *Psychonomic Science*, 5, 79-80.
- Irwin, F. W., & Snodgrass, J. G. (1966). Effects of independent and dependent outcome values upon bets. *Journal of Experimental Psychology*, 71, 282-285.
- Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, 52, 1122-1131.
- Isen, A. M., & Geva, N. (1987). The influence of positive affect on acceptable level of risk: The person with a large canoe has a large worry. *Organizational Behavior and Human Decision Processes*, 39, 145-154.
- Isen, A. M., & Patrick, R. (1983). The effect of positive feelings on risk taking: When the chips are down. *Organizational Behavior and Human Performance*, 31, 194-202.
- Ives, G. (1970). *A history of penal methods: Criminals, witches, lunatics*. Montclair, NJ: Patterson Smith.
- Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes* (2nd ed.). Boston: Houghton Mifflin.
- Jenkins, H. M., & Ward, W. C. (1965). The judgment of contingency between responses and outcomes. *Psychological Monographs*, 79 (Whole No. 594).
- Jenkins, J. J. (1981). Can we have a fruitful cognitive psychology? In J. H. Flowers (Ed.), *Nebraska symposium on motivation, 1980*. Lincoln: University of Nebraska Press.
- Jennings, D. L., Amabile, T. M., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.

- Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology*, 45, 20–31.
- Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, 34, 107–117.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York: Academic Press.
- Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, 3, 1–24.
- Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.
- Jones, E. E., Wood, G. C., & Quattrone, G. A. (1981). Perceived variability of personal characteristics in in-groups and out-groups: The role of knowledge and evaluation. *Personality and Social Psychology Bulletin*, 7, 523–528.
- Jones, R. T. (1971). Tetrahydrocannabinol and the marijuana-induced social “high,” or the effects of the mind on marijuana. *Annals of the New York Academy of Sciences*, 191, 155–165.
- Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, 93, 429–445.
- Kahneman, D. (1991). Judgment and decision making: A personal view. *Psychological Science*, 2, 142–145.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98, 1325–1348.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430–454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- Kammer, D. (1982). Differences in trait ascriptions to self and friend: Unconfounding intensity from variability. *Psychological Reports*, 51, 99–102.
- Kantola, S. J., Syme, G. J., & Campbell, N. A. (1984). Cognitive dissonance and energy conservation. *Journal of Applied Psychology*, 69, 416–421.
- Karmarkar, U. (1978). Subjectively weighted utility: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance*, 21, 61–72.
- Kassin, S. M. (1979). Consensus information, prediction, and causal attribution: A review of the literature and issues. *Journal of Personality and Social Psychology*, 37, 1966–1981.
- Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18, 431–439.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine

- (Ed.), *Nebraska symposium on motivation, 1967*. Lincoln: University of Nebraska Press.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*, 107-128.
- Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology, 31*, 457-501.
- Keren, G. (1987). Facing uncertainty in the game of bridge: A calibration study. *Organizational Behavior and Human Decision Processes, 39*, 98-114.
- Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211-228.
- Kleinmuntz, D. N. (1985). Cognitive heuristics and feedback in a dynamic decision environment. *Management Science, 31*, 680-702.
- Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics, 99*, 507-521.
- Knox, R. E., & Inkster, J. A. (1968). Postdecision dissonance at post time. *Journal of Personality and Social Psychology, 8*, 319-323.
- Kogan, N., & Wallach, M. A. (1964). *Risk taking: A study in cognition and personality*. New York: Holt, Rinehart and Winston.
- Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 107-118.
- Kristiansen, C. M. (1983). Newspaper coverage of diseases and actual mortality statistics. *European Journal of Social Psychology, 13*, 193-194.
- Lamal, P. A. (1979, October). College student common beliefs about psychology. *Teaching of Psychology, pp.* 155-158.
- Landy, D., & Sigall, H. (1974). Beauty is talent: Task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology, 29*, 299-304.
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311-328.
- Langer, E. J., & Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology, 34*, 191-198.
- Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology, 32*, 951-955.
- LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces, 13*, 230-237.
- Lasky, J. J., Hover, G. L., Smith, P. A., Bostian, D. W., Duffendack, S. C., & Nord, C. L. (1959). Post-hospital adjustment as predicted by psychiatric patients and by their staff. *Journal of Consulting Psychology, 23*, 213-218.
- Latané, B., & Dabbs, J. M., Jr. (1975). Sex, group size and helping in three cities. *Sociometry, 38*, 180-194.
- Latané, B., & Darley, J. M. (1969). Bystander "apathy." *American Scientist, 57*, 244-268.

- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* Englewood Cliffs, NJ: Prentice-Hall.
- Latané, B., & Nida, S. (1981). Ten years of research on group size and helping. *Psychological Bulletin*, 89, 308–324.
- Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37, 822–832.
- Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, 88, 438–453.
- Leary, M. R. (1981). The distorted nature of hindsight. *Journal of Social Psychology*, 115, 25–29.
- Leary, M. R. (1982). Hindsight distortion and the 1980 presidential election. *Personality and Social Psychology Bulletin*, 8, 257–263.
- Leddo, J., Abelson, R. P., & Gross, P. H. (1984). Conjunctive explanations: When two reasons are better than one. *Journal of Personality and Social Psychology*, 47, 933–943.
- Lefcourt, H. M. (1982). *Locus of control: Current trends in theory and research*. Hillsdale, NJ: Erlbaum.
- Leippe, M. R., Wells, G. L., & Ostrom, T. M. (1978). Crime seriousness as a determinant of accuracy in eyewitness identification. *Journal of Applied Psychology*, 63, 345–351.
- Lelyveld, J. (1986, October 5). Britain heads for nuclear war at polls. *New York Times*, p. E2.
- Levi, A. S., & Pryor, J. B. (1987). Use of the availability heuristic in probability estimates of future events: The effects of imagining outcomes versus imagining reasons. *Organizational Behavior and Human Decision Processes*, 40, 219–234.
- Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, 20, 159–183.
- Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Performance*, 26, 149–171.
- Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89, 46–55.
- Lichtenstein, S., & Slovic, P. (1973). Response-induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology*, 101, 16–20.
- Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, 7, 560–572.
- Loftus, E. F. (1979) *Eyewitness testimony*. Cambridge, MA: Harvard University Press.

- Loftus, E. (1980). *Memory: Surprising new insights into how we remember and why we forget*. New York: Ardsley House.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585–589.
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, 92, 805–824.
- Loomes, G., & Sugden, R. (1983). A rationale for preference reversal. *American Economic Review*, 73, 428–432.
- Loomes, G., & Sugden, R. (1987). Some implications of a more general form of regret theory. *Journal of Economic Theory*, 41, 270–287.
- Lopes, L. L. (1981). Decision making in the short run. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 377–385.
- Lopes, L. L. (1982). Doing the impossible: A note on induction and the experience of randomness. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 626–636.
- Lopes, L. L. (1991). The rhetoric of irrationality. *Theory & Psychology*, 1, 65–82.
- Lopes, L. L., & Oden, G. C. (1987). Distinguishing between random and non-random events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 392–400.
- Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47, 1231–1243.
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2098–2109.
- Loy, J. W., & Andrews, D. S. (1981). They also saw a game: A replication of a case study. *Replications in Social Psychology*, 1, 45–49.
- Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- Luce, R. D. (1990). Rational versus plausible accounting equivalences in preference judgments. *Psychological Science*, 1, 225–234.
- Luck, R. F., van den Bosch, R., & Garcia, R. (1977). Chemical insect control—A troubled pest management strategy. *BioScience*, 606–611.
- Maass, A., & Clark, R. D., III. (1984). Hidden impact of minorities: Fifteen years of minority research. *Psychological Bulletin*, 95, 428–450.
- Mackie, D. M., & Allison, S. T. (1987). Group attribution errors and the illusion of group attitude change. *Journal of Experimental Social Psychology*, 23, 460–480.
- Maier, N. R. F., & Solem, A. R. (1952). The contribution of a discussion leader to the quality of group thinking: The effective use of minority opinions. *Human Relations*, 5, 277–288.
- Malkiel, B. G. (1985). *A random walk down Wall Street* (4th ed.). New York: W. W. Norton & Co.
- Malone, B. (1990, September 9). Double trouble: Lives of two Frank William

- Boumas bedeviled by strange coincidences. *Grand Rapids Press*, pp. B1-B2.
- Marks, R. W. (1951). The effect of probability, desirability, and "privilege" on the stated expectations of children. *Journal of Personality*, 19, 332-351.
- Matsuda, N. (1985). Strong, quasi-, and weak conformity among Japanese in the modified Asch procedure. *Journal of Cross Cultural Psychology*, 16, 83-97.
- Mazur, A. (1981). Three Mile Island and the scientific community. In T. H. Moss & D. L. Sills (Eds.), *The Three Mile Island nuclear accident: Lessons and implications. Annals of the New York Academy of Sciences* (Vol. 365, pp. 216-221). New York: The New York Academy of Sciences.
- McArthur, L. A. (1972). The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 22, 171-193.
- McArthur, L. Z. (1980). Illusory causation and illusory correlation: Two epistemological accounts. *Personality and Social Psychology Bulletin*, 6, 507-519.
- McArthur, L. Z., & Post, D. L. (1977). Figural emphasis and person perception. *Journal of Experimental Social Psychology*, 13, 520-535.
- McMillen, D., Smith, S., & Wells-Parker, E. (1989). The effects of alcohol, expectancy, and sensation seeking on driving risk taking. *Addictive Behaviors*, 14, 477-483.
- McNeil, B. J., Pauker, S. G., Sox, H. C., Jr., & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *New England Journal of Medicine*, 306, 1259-1262.
- Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193-210.
- Meyerowitz, B. E., & Chaiken, S. (1987). The effect of message framing on breast self-examination attitudes, intentions, and behavior. *Journal of Personality and Social Psychology*, 52, 500-510.
- Michaels, J. W., Blommel, J. M., Brocato, R. M., Linkous, R. A., & Rowe, J. S. (1982). Social facilitation and inhibition in a natural setting. *Replications in Social Psychology*, 2, 21-24.
- Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378.
- Miller, A. G., Gillen, B., Schenker, C., & Radlove, S. (1973). Perception of obedience to authority. *Proceedings of the 81st Annual Convention of the American Psychological Association*, 8, 127-128.
- Miller, D. T. (1976). Ego involvement and attribution for success and failure. *Journal of Personality and Social Psychology*, 34, 901-906.
- Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*, 82, 213-225.
- Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. *Annual Review of Psychology*, 37, 233-256.
- Miller, J. G. (1984). Culture and the development of everyday social explanation. *Journal of Personality and Social Psychology*, 46, 961-978.
- Miller, N., & Campbell, D. T. (1959). Recency and primacy in persuasion as a

- function of the timing of speeches and measurements. *Journal of Abnormal and Social Psychology*, 59, 1-9.
- Mitchell, R. C. (1982). Public response to a major failure of a controversial technology. In D. L. Sills, C. P. Wolf, & V. B. Shelanski (Eds.), *Accident at Three Mile Island: The human dimensions* (pp. 21-38). Boulder, CO: Westview Press.
- Moede, W. (1927). Die Richtlinien der Leistungs-Psychologie. *Industrielle Psychotechnik*, 4, 193-207.
- Morier, D. M., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin*, 10, 243-252.
- Moscovici, S., Lage, E., & Naffrechoux, M. (1969). Influence of a consistent minority on the responses of a majority in a color perception task. *Sociometry*, 32, 365-380.
- Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology*, 12, 125-135.
- Mullen, B., & Hu, L. (1989). Perceptions of ingroup and outgroup variability: A meta-analytic integration. *Basic and Applied Social Psychology*, 10, 233-252.
- Mullen, B., & Johnson, C. (1990). Distinctiveness-based illusory correlations and stereotyping: A meta-analytic integration. *British Journal of Social Psychology*, 29, 11-28.
- Mullen, B., & Riordan, C. A. (1988). Self-serving attributions for performance in naturalistic settings: A meta-analytic review. *Journal of Applied Social Psychology*, 18, 3-22.
- Murphy, A. H., & Brown, B. G. (1984). A comparative evaluation of objective and subjective weather forecasts in the United States. *Journal of Forecasting*, 3, 369-393.
- Murphy, A. H., & Winkler, R. L. (1984). Probability forecasting in meteorology. *Journal of the American Statistical Association*, 79, 489-500.
- Myers, D. G. (1975). Discussion-induced attitude polarization. *Human Relations*, 28, 699-714.
- Myers, D. G. (1982). Polarizing effects of social interaction. In H. Brandstatter, J. H. Davis, & G. Stocker-Kreichgauer (Eds.), *Group decision making*. London: Academic Press.
- Myers, D. G. (1990). *Social psychology* (2nd ed.). McGraw-Hill: New York.
- Myers, D. G., & Bishop, G. D. (1970). Discussion effects on racial attitudes. *Science*, 169, 778-779.
- Myers, D. G., & Kaplan, M. F. (1976). Group-induced polarization in simulated juries. *Personality and Social Psychology Bulletin*, 2, 63-66.
- Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, 83, 602-627.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, 29, 85-95.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of con-

- firmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, 30, 395-406.
- Nahinsky, I. D., & Slaymaker, F. L. (1970). Use of negative instances in conjunctive identification. *Journal of Experimental Psychology*, 84, 64-84.
- Nathanson, S., Brockner, J., Brenner, D., Samuelson, C., Countryman, M., Lloyd, M., & Rubin, J. Z. (1982). Toward the reduction of entrapment. *Journal of Applied Social Psychology*, 12, 193-208.
- Nemeth, C. (1986). Differential contributions of majority and minority influence. *Psychological Review*, 93, 1-10.
- Nemeth, C., & Chiles, C. (1988). Modelling courage: The role of dissent in fostering independence. *European Journal of Social Psychology*, 18, 275-280.
- Neuringer, A. (1986). Can people behave "randomly"?: The role of feedback. *Journal of Experimental Psychology: General*, 115, 62-75.
- Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology*, 32, 932-943.
- Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. (1976). Popular induction: Information is not always informative. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (Vol. 2, pp. 227-236). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Nisbett, R. E., Caputo, C., Legant, P., & Marecek, J. (1973). Behavior as seen by the actor and as seen by the observer. *Journal of Personality and Social Psychology*, 27, 154-165.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., & Schachter, S. (1966). Cognitive manipulation of pain. *Journal of Experimental Social Psychology*, 2, 227-236.
- Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, 39, 84-97.
- Orvis, B. R., Cunningham, J. D., & Kelley, H. H. (1975). A closer examination of causal inference: The roles of consensus, distinctiveness, and consistency information. *Journal of Personality and Social Psychology*, 32, 605-616.
- Osberg, T. M., & Shrauger, J. S. (1986). Self-prediction: Exploring the parameters of accuracy. *Journal of Personality and Social Psychology*, 51, 1044-1057.
- Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- Paese, P. W., & Sniezek, J. A. (1991). Influences on the appropriateness of confidence in judgment: Practice, effort, information, and decision-making. *Organizational Behavior and Human Decision Processes*, 48, 100-130.
- Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, 42, 1051-1068.

- Paulos, J. A. (1986, November 24). Orders of magnitude. *Newsweek*, pp. 12-13.
- Paulos, J. A. (1988). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Vintage Books.
- Payne, J. W. (1973). Alternative approaches to decision making under risk: Moments versus risk dimensions. *Psychological Bulletin*, 80, 439-453.
- Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin*, 92, 382-402.
- Pennington, D. C., Rutter, D. R., McKenna, K., & Morley, I. E. (1980). Estimating the outcome of a pregnancy test: Women's judgments in foresight and hindsight. *British Journal of Social and Clinical Psychology*, 19, 317-324.
- Peter, L. J. (1977). *Peter's quotations: Ideas for our time*. New York: Bantam Books.
- Peterson, C. (1980). Recognition of noncontingency. *Journal of Personality and Social Psychology*, 38, 727-734.
- Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. Y., Metalsky, G. I., & Seligman, M. E. P. (1982). The Attributional Style Questionnaire. *Cognitive Therapy and Research*, 6, 287-300.
- Pettigrew, T. F. (1979). The ultimate attribution error: Extending Allport's cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5, 461-476.
- Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology*, 72, 346-354.
- Pietromonaco, P. R., & Nisbett, R. E. (1982). Swimming upstream against the fundamental attribution error: Subjects' weak generalizations from the Darley and Batson study. *Social Behavior and Personality*, 10, 1-4.
- Platt, J. (1973). Social traps. *American Psychologist*, 28, 641-651.
- Plous, S. (1989). Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology*, 19, 67-91.
- Plous, S. (1989, March). Political illiteracy: A threat to international security. *Swords and Ploughshares*, pp. 9-10.
- Plous, S. (1991). Biases in the assimilation of technological breakdowns: Do accidents make us safer? *Journal of Applied Social Psychology*, 21, 1058-1082.
- Plous, S., & Zimbardo, P. G. (1984, November). The looking glass war. *Psychology Today*, pp. 48-59.
- Plous, S., & Zimbardo, P. G. (1986). Attributional biases among clinicians: A comparison of psychoanalysts and behavior therapists. *Journal of Consulting and Clinical Psychology*, 54, 568-570.
- Pool, R. (1988). The Allais Paradox. *Science*, 242, 512.
- Prothro, J. W., & Grigg, C. M. (1960). Fundamental principles of democracy: Bases of agreement and disagreement. *Journal of Politics*, 22, 276-294.

- Pruitt, D. G., & Hoge, R. D. (1965). Strength of the relationship between the value of an event and its subjective probability as a function of method of measurement. *Journal of Personality and Social Psychology*, 69, 483-489.
- Pryor, J. B., & Kriss, N. (1977). The cognitive dynamics of salience in the attribution process. *Journal of Personality and Social Psychology*, 35, 49-55.
- Quattrone, G. A. (1982). Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, 42, 593-607.
- Quattrone, G. A., Lawrence, C. P., Warren, D. L., Souza-Silva, K., Finkel, S. E., & Andrus, D. E. (1984). *Explorations in anchoring: The effects of prior range, anchor extremity, and suggestive hints*. Unpublished manuscript, Stanford University, Stanford.
- Quattrone, G. A., & Tversky, A. (1988). Contrasting rational and psychological analyses of political choice. *American Political Science Review*, 82, 719-736.
- Reeder, G. D. (1982). Let's give the fundamental attribution error another chance. *Journal of Personality and Social Psychology*, 43, 341-344.
- Regan, D. T., Straus, E., & Fazio, R. (1974). Liking and the attribution process. *Journal of Experimental Social Psychology*, 10, 385-397.
- Regan, D. T., & Totten, J. (1975). Empathy and attribution: Turning observers into actors. *Journal of Personality and Social Psychology*, 32, 850-856.
- Reichenbach, H. (1949). *The theory of probability*. Berkeley: University of California Press.
- Reyes, R. M., Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availabilities of arguments. *Journal of Personality and Social Psychology*, 39, 2-12.
- Robbins, J. (1987). *Diet for a new America*. Walpole, NH: Stillpoint Publishing.
- Rodin, J. (1986). Aging and health: Effects of the sense of control. *Science*, 233, 1271-1276.
- Ronis, D. L., & Yates, J. F. (1987). Components of probability judgment accuracy: Individual consistency and effects of subject matter and assessment method. *Organizational Behavior and Human Decision Processes*, 40, 193-218.
- Rosenhan, D. L., & Messick, S. (1966). Affect and expectation. *Journal of Personality and Social Psychology*, 3, 38-44.
- Rosenthal, R. (1976). *Experimenter effects in behavioral research* (enlarged ed.). New York: Irvington.
- Rosenthal, R. (1987, December). Pygmalion effects: Existence, magnitude, and social importance. *Educational Researcher*, pp. 37-41.
- Rosenthal, R., & Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, 8, 183-189.

- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10). New York: Academic Press.
- Ross, L., Lepper, M., & Hubbard, M. (1975). Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32, 880–892.
- Ross, L., Lepper, M., Strack, F., & Steinmetz, J. (1977). Social explanation and social expectation: Effects of real and hypothetical explanations on subjective likelihood. *Journal of Personality and Social Psychology*, 37, 817–829.
- Ross, L., Rodin, J., & Zimbardo, P. G. (1969). Toward an attribution therapy: The reduction of fear through induced cognitive-emotional misattribution. *Journal of Personality and Social Psychology*, 12, 279–288.
- Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology*, 37, 322–336.
- Rothbart, M. (1970). Assessing the likelihood of a threatening event: English Canadians' evaluation of the Quebec separatist movement. *Journal of Personality and Social Psychology*, 15, 109–117.
- Rothstein, H. G. (1986). The effects of time pressure on judgment in multiple cue probability learning. *Organizational Behavior and Human Decision Processes*, 37, 83–92.
- Rubin, D. M. (1981). What the president's commission learned about the media. In T. H. Moss & D. L. Sills (Eds.), *The Three Mile Island nuclear accident: Lessons and implications*. *Annals of the New York Academy of Sciences* (Vol. 365, pp. 95–106). New York: The New York Academy of Sciences.
- Rubin, J. Z., Brockner, J., Small-Weil, S., & Nathanson, S. (1980). Factors affecting entry into psychological traps. *Journal of Conflict Resolution*, 24, 405–426.
- Ruble, D. N., & Feldman, N. S. (1976). Order of consensus, distinctiveness, and consistency information and causal attributions. *Journal of Personality and Social Psychology*, 34, 930–937.
- Rugg, D. (1941). Experiments in wording questions: II. *Public Opinion Quarterly*, 5, 91–92.
- Rusbult, C. E. (1980). Commitment and satisfaction in romantic associations: A test of the investment model. *Journal of Experimental Social Psychology*, 16, 172–186.
- Russo, J. E. (1977). The value of unit price information. *Journal of Marketing Research*, 14, 193–201.
- Russo, J. E., & Schoemaker, P. J. H. (1989). *Decision traps: Ten barriers to brilliant decision making and how to overcome them*. New York: Simon & Schuster.
- Ryback, D. (1967). Confidence and accuracy as a function of experience in

- judgment-making in the absence of systematic feedback. *Perceptual and Motor Skills*, 24, 331–334.
- Rylsky, M. (1986, February). A town born of the atom. *Soviet Life*, pp. 8–15.
- Salovey, P., & Birnbaum, D. (1989). Influence of mood on health-relevant cognitions. *Journal of Personality and Social Psychology*, 57, 539–551.
- Sande, G. N., Goethals, G. R., & Radloff, C. E. (1988). Perceiving one's own traits and others': The multifaceted self. *Journal of Personality and Social Psychology*, 54, 13–20.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- Schelling, T. (1971). The ecology of micromotives. *Public Interest*, 25, 61–98.
- Schelling, T. C. (1981). Economic reasoning and the ethics of policy. *Public Interest*, 63, 37–61.
- Schkade, D. A., & Johnson, E. J. (1989). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes*, 44, 203–231.
- Schlenker, B. R., & Miller, R. S. (1977). Egocentrism in groups: Self-serving biases or logical information processing? *Journal of Personality and Social Psychology*, 35, 755–764.
- Schmemmann, S. (1985, March 3). The emergence of Gorbachev. *New York Times Magazine*, pp. 40, 44–46, 55–57.
- Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 20, 529–563.
- Schum, D. (1990). Discussion. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn*. Chicago: University of Chicago Press.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Orlando, FL: Academic Press.
- Schuman, H., & Scott, J. (1987). Problems in the use of survey questions to measure public opinion. *Science*, 236, 957–959.
- Schwarz, N., Hippler, H., Deutsch, B., & Strack, S. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, 49, 388–395.
- Seligman, M. E. P., Abramson, L. Y., Semmel, A., & von Baeyer, C. (1979). Depressive attributional style. *Journal of Abnormal Psychology*, 88, 242–247.
- Selvin, S. (1975, February). A problem in probability. *American Statistician*, p. 67.
- Shaklee, H., & Fischhoff, B. (1990). The psychology of contraceptive surprises: Cumulative risk and contraceptive effectiveness. *Journal of Applied Social Psychology*, 20, 385–403.
- Shaklee, H., & Mims, M. (1982). Sources of error in judging event covariations: Effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 208–224.
- Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition*, 8, 459–467.

- Shenkel, R. J., Snyder, C. R., Batson, C. D., & Clark, G. M. (1979). Effects of prior diagnostic information on clinicians' causal attributions of a client's problems. *Journal of Consulting and Clinical Psychology, 47*, 404-406.
- Sherif, M., Taub, D., & Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *Journal of Experimental Psychology, 55*, 150-155.
- Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality and Social Psychology Bulletin, 11*, 118-127.
- Sherman, S. J., & Gorkin, L. (1980). Attitude bolstering when behavior is inconsistent with central attitudes. *Journal of Experimental Social Psychology, 16*, 388-403.
- Sherman, S. J., Zehner, K. S., Johnson, J., & Hirt, E. R. (1983). Social explanation: The role of timing, set and recall on subjective likelihood estimates. *Journal of Personality and Social Psychology, 44*, 1127-1143.
- Shubik, M. (1971). The Dollar Auction game: A paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution, 15*, 109-111.
- Sieber, J. E. (1974). Effects of decision importance on ability to generate warranted subjective uncertainty. *Journal of Personality and Social Psychology, 30*, 688-694.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129-138.
- Skinner, B. F. (1980). *Notebooks*. Englewood Cliffs, NJ: Prentice-Hall.
- Slovic, P. (1975). Choice between equally valued alternatives. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 280-287.
- Slovic, P. (1987). Perception of risk. *Science, 236*, 280-285.
- Slovic, P., & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 544-551.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1979, April). Rating the risks. *Environment*, pp. 14-20, 36-39.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982a). Facts versus fears: Understanding perceived risk. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982b). Response mode, framing, and information-processing effects in risk assessment. In R. M. Hogarth (Ed.), *New directions for methodology of social and behavioral science: Question framing and response consistency* (No. 11). San Francisco: Jossey-Bass.
- Slovic, P., Griffin, D., & Tversky, A. (1990). Compatibility effects in judgment and choice. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J. Einhorn*. Chicago: University of Chicago Press.

- Slovic, P., & Lichtenstein, S. (1983). Preference reversals: A broader perspective. *American Economic Review*, 73, 596-605.
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-173.
- Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin*, 109, 472-489.
- Smith, T. W. (1984). Nonattitudes: A review and evaluation. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena* (Vol. 2). New York: Russell Sage Foundation.
- Sniezek, J. A. (1989). An examination of group process in judgmental forecasting. *International Journal of Forecasting*, 5, 171-178.
- Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes*, 43, 1-28.
- Sniezek, J. A., & Henry, R. A. (1990). Revision, weighting, and commitment in consensus group judgment. *Organizational Behavior and Human Decision Processes*, 45, 66-84.
- Sniezek, J. A., Paese, P. W., & Switzer, F. S., III. (1990). The effect of choosing on confidence and choice. *Organizational Behavior and Human Decision Processes*, 46, 264-282.
- Snyder, C. R. (1977). "A patient by any other name" revisited: Maladjustment or attributional locus of problem? *Journal of Consulting and Clinical Psychology*, 45, 101-103.
- Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 247-305). New York: Academic Press.
- Snyder, M. (in press). Motivational foundations of behavioral confirmation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25). New York: Academic Press.
- Snyder, M., Campbell, B. H., & Preston, E. (1982). Testing hypotheses about human nature: Assessing the accuracy of social stereotypes. *Social Cognition*, 1, 256-272.
- Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Experimental Social Psychology*, 15, 330-342.
- Snyder, M., & Jones, E. E. (1974). Attitude attribution when behavior is constrained. *Journal of Experimental Social Psychology*, 10, 585-600.
- Snyder, M., & Swann, W. B., Jr. (1978). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36, 1202-1212.
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology*, 35, 656-666.
- Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, 36, 941-950.

- Starr, C. (1969). Social benefit versus technological risk. *Science*, 165, 1232-1238.
- Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 16, 27-44.
- Staw, B. M. (1981). The escalation of commitment to a course of action. *Academy of Management Review*, 6, 577-587.
- Staw, B. M., & Ross, J. (1987). Behavior in escalation situations: Antecedents, prototypes, and solutions. *Research in Organizational Behavior*, 9, 39-78.
- Staw, B. M., & Ross, J. (1987, March-April). Knowing when to pull the plug. *Harvard Business Review*, pp. 68-74.
- Stone, E. R., & Yates, J. F. (1991). *Communications about low-probability risks: Effects of alternative displays*. Unpublished manuscript, University of Michigan, Ann Arbor.
- Stoner, J. A. F. (1961). *A comparison of individual and group decisions involving risk*. Unpublished master's thesis, Massachusetts Institute of Technology.
- Stoner, J. A. F. (1968). Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 4, 442-459.
- Storms, M. D. (1973). Videotape and the attribution process: Reversing actors' and observers' points of view. *Journal of Personality and Social Psychology*, 27, 165-175.
- Sullivan, K. (1990, October 5). Chip rivals in comedy of errors. *San Francisco Examiner*, pp. B1, B4.
- Suls, J. M., & Miller, R. L. (Eds.). (1977). *Social comparison processes: Theoretical and empirical perspectives*. Washington, DC: Hemisphere Publishing Corporation.
- Sweeney, P. D., Anderson, K., & Bailey, S. (1986). Attributional style in depression: A meta-analytic review. *Journal of Personality and Social Psychology*, 50, 974-991.
- Synodinos, N. E. (1986). Hindsight distortion: "I knew-it-all along and I was sure about it." *Journal of Applied Social Psychology*, 16, 107-117.
- Taylor, D. M., & Doria, J. R. (1981). Self-serving and group-serving bias in attribution. *Journal of Social Psychology*, 113, 201-211.
- Taylor, D. M., & Jaggi, V. (1974). Ethnocentrism and causal attribution in a South Indian context. *Journal of Cross Cultural Psychology*, 5, 162-171.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social-psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.
- Taylor, S. E., & Fiske, S. T. (1975). Point of view and perceptions of causality. *Journal of Personality and Social Psychology*, 32, 439-445.
- Taylor, S. E., & Fiske, S. (1978). Salience, attention, and attribution: Top of the head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 249-288). New York: Academic Press.
- Taylor, S. E., Fiske, S. T., Close, M., Anderson, C., & Ruderman, A. (1977).

- Solo status as a psychological variable: The power of being distinctive.* Unpublished manuscript, Harvard University, Cambridge, MA.
- Taylor, S. E., & Koivumaki, J. H. (1976). The perception of self and others: Acquaintanceship, affect, and actor-observer differences. *Journal of Personality and Social Psychology*, 33, 403-408.
- Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive "vividness" effect. *Psychological Review*, 89, 155-181.
- Teger, A. I. (1980). *Too much invested to quit*. New York: Pergamon Press.
- Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, 45, 74-83.
- Tetlock, P. E. (1985a). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly*, 48, 227-236.
- Tetlock, P. E. (1985b). Accountability: The neglected social context of judgment and choice. *Research in Organizational Behavior*, 7, 297-332.
- Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology*, 52, 700-709.
- Tetlock, P. E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology*, 57, 632-640.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39-60.
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4, 199-214.
- Thomas, E. J., & Fink, C. F. (1961). Models of group problem solving. *Journal of Abnormal and Social Psychology*, 63, 53-63.
- Thompson, S. C., & Kelley, H. H. (1981). Judgments of responsibility for activities in close relationships. *Journal of Personality and Social Psychology*, 41, 469-477.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology*, 4, 25-29.
- Thorngate, W. (1980). Efficient decision heuristics. *Behavioral Science*, 25, 219-225.
- Tindale, R. S. (1989). Group vs individual information processing: The effects of outcome feedback on decision making. *Organizational Behavior and Human Decision Processes*, 44, 454-473.
- Tindale, R. S., Sheffey, S., & Filkins, J. (1990). *Conjunction errors by individuals and groups*. Paper presented at the annual meeting of the Society for Judgment and Decision Making, New Orleans, LA.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31-48.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79, 281-299.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76, 105-110.

- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1130.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453–458.
- Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90, 293–315.
- Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59, S251–S278.
- Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371–384.
- Tversky, A., Slovic, P., & Kahneman, D. (1990). The causes of preference reversal. *American Economic Review*, 80, 204–217.
- U.S. Congress: House, Subcommittee of the Committee on Government Operations. (1981, May 19–20). *Failures of the North American Aerospace Defense Command's (NORAD) Attack Warning System*. 97th Congress, First Session, Washington, DC: U.S. Government Printing Office.
- Valins, S., & Nisbett, R. E. (1971). Attribution processes in the development and treatment of emotional disorders. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior*. Morristown, NJ: General Learning Press.
- Vallone, R. P., Griffin, D. W., Lin, S., & Ross, L. (1990). Overconfident prediction of future actions and outcomes by self and others. *Journal of Personality and Social Psychology*, 58, 582–592.
- Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, 49, 577–585.
- Verplanken, B., & Pieters, R. G. M. (1988). Individual differences in reverse hindsight bias: I never thought something like Chernobyl would happen, did I? *Journal of Behavioral Decision Making*, 1, 131–147.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
- vos Savant, M. (1990, September 9). Ask Marilyn. *Parade*, p. 13.
- Wagenaar, W. A. (1970a). Appreciation of conditional probabilities in binary sequences. *Acta Psychologica*, 34, 348–356.
- Wagenaar, W. A. (1970b). Subjective randomness and the capacity to generate information. *Acta Psychologica*, 33, 233–242.
- Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin*, 77, 65–72.

- Wallsten, T. S. (1981). Physician and medical student bias in evaluating diagnostic information. *Medical Decision Making*, 1, 145-164.
- Walster, E. (1967). Second-guessing important events. *Human Relations*, 20, 239-249.
- Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, 19, 231-241.
- Warwick, D. P. (1975, February). Social scientists ought to stop lying. *Psychology Today*, pp. 38, 40, 105, 106.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content*. Cambridge, MA: Harvard University Press.
- Watson, D. (1982). The actor and the observer: How are their perceptions of causality divergent? *Psychological Bulletin*, 92, 682-700.
- Weaver, W. (1982). *Lady luck: The theory of probability*. New York: Dover Publications.
- Weinberg, A. M. (1981). Three Mile Island in perspective: Keynote address. In T. H. Moss & D. L. Sills (Eds.), *The Three Mile Island nuclear accident: Lessons and implications*. *Annals of the New York Academy of Sciences* (Vol. 365, pp. 1-12). New York: The New York Academy of Sciences.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39, 806-820.
- Weirich, P. (1984). The St. Petersburg gamble and risk. *Theory and Decision*, 17, 193-202.
- Weldon, E., & Gargano, G. M. (1985). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. *Organizational Behavior and Human Decision Processes*, 36, 348-361.
- Weldon, E., & Gargano, G. M. (1988). Cognitive loafing: The effects of accountability and shared responsibility on cognitive effort. *Personality and Social Psychology Bulletin*, 14, 159-171.
- Wells, G. L., & Harvey, J. H. (1977). Do people use consensus information in making causal attributions? *Journal of Personality and Social Psychology*, 35, 279-293.
- White, P. A., & Younger, D. P. (1988). Differences in the ascription of transient internal states to self and other. *Journal of Experimental Social Psychology*, 24, 292-309.
- Wicker, A. W. (1969). Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues*, 25, 41-78.
- Wicker, A. W. (1971). An examination of the "other variables" explanation of attitude-behavior inconsistency. *Journal of Personality and Social Psychology*, 19, 18-30.
- Williams, L. (1990, February 14). Decisions, decisions, decisions: Enough! *New York Times*, pp. B1, B5.
- Wilson, D. K., Kaplan, R. M., & Schneiderman, L. J. (1987). Framing of deci-

- sions and selections of alternatives in health care. *Social Behaviour*, 2, 51-59.
- Wilson, G. T., & Abrams, D. (1977). Effects of alcohol on social anxiety and physiological arousal: Cognitive versus pharmacological processes. *Cognitive Research and Therapy*, 1, 195-210.
- Wilson, R. (1979, February). Analyzing the daily risks of life. *Technology Review*, pp. 41-46.
- Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, 10, 109-120.
- Wortman, C. B. (1975). Some determinants of perceived control. *Journal of Personality and Social Psychology*, 31, 282-294.
- Wright, G. N., Phillips, L. D., Whalley, P. C., Choo, G. T., Ng, K., & Tan, I. (1978). Cultural differences in probabilistic thinking. *Journal of Cross Cultural Psychology*, 9, 285-299.
- Wright, J. C. (1962). Consistency and complexity of response sequences as a function of schedules of noncontingent reward. *Journal of Experimental Psychology*, 63, 601-609.
- Wright, P. (1974). The harassed decision maker: Time pressures, distractions, and the use of evidence. *Journal of Applied Psychology*, 59, 555-561.
- Wright, W. F., & Anderson, U. (1989). Effects of situation familiarity and financial incentives on use of the anchoring and adjustment heuristic for probability assessment. *Organizational Behavior and Human Decision Processes*, 44, 68-82.
- Wright, W. F., & Bower, G. H. (1992). Mood effects on subjective probability assessment. *Organizational Behavior and Human Decision Processes*, 52, 276-291.
- Wyer, R. S., Jr. (1976). An investigation of the relations among probability estimates. *Organizational Behavior and Human Performance*, 15, 1-18.
- Yates, J. F. (1990). *Judgment and decision making*. Englewood Cliffs, NJ: Prentice Hall.
- Yates, J. F., Zhu, Y., Ronis, D. L., Wang, D., Shinotsuka, H., & Toda, M. (1989). Probability judgment accuracy: China, Japan, and the United States. *Organizational Behavior and Human Decision Processes*, 43, 145-171.
- Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 269-274.

推荐阅读物

如果你想要阅读有关决策与判断领域的更多材料，那么以下这些书是我强烈推荐的。

强调管理和商业背景下的决策与判断的书籍有：

- Russo, J. E. , & Schoemaker, P. J. H. (1989). *Decision traps: Ten barriers to brilliant decision making and how to overcome them*. New York: Simon & Schuster.
- Bazerman, M. H. (1990). *Judgment in managerial decision making* (2nd ed.). New York: John Wiley & Sons.
- Hogarth, R. (1987). *Judgment and choice* (2nd ed.). New York: John Wiley & Sons.

如果你想了解不强调商业应用的非技术性的概述，可以读一读：

- Dawes, R. M. (1988). *Rational choice in an uncertain world*. San Diego: Harcourt Brace Jovanovich.
- Nisbett, R. E. , & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice - Hall.
- Gilovich, T. (1991). *How we know what aint so*. New York: Free Press.

如果想要获取更深入的材料，请阅读剑桥大学出版社出版的三本杰出的文集，以下就是这三本书，其中的第一本被认为是当代的经典：

- Kahneman, D. , Slovic, P. , & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Arkes, H. R. , & Hammond, K. R. (Eds.). (1986). *Judgment and decision making: An interdisciplinary reader*. Cambridge, England: Cambridge University Press.
- Bell, D. E. , & Raiffa, H. , & Tversky, A. (Eds.). (1988). *Decision making: Descriptive, normative, and prescriptive interactions*. Cambridge, Wngland: Cambridge University Press.

芝加哥大学也出版了一本非常好的演讲录，以纪念决策研究者希勒尔·艾