



生命科学实验指南系列

Experimental Design for Biologists 生命科学实验 设计指南

〔美〕D.J. 格拉斯 著

丛羽生 等 译

实验项目

项目 A

项目 B

蛋白质的功能是什么？

E3泛素连接酶的功能是什么？

MuRF1
的功能
是什么？

MDM2

Skp2

哪里是MuRF1最合適的位置？

蛋白质的功能是什么？

MuRF1

泛素连接酶

Hect域 RING域

F-box域

磷酸化酶

酪氨酸

肌醇

受体

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸

酪氨酸



科学出版社

www.sciencep.com

生命科学实验指南系列

生命科学实验设计指南

Experimental Design for Biologists

〔美〕 D. J. 格拉斯 著

丛羽生 等 译

科学出版社

北 京

图字: 01-2007-4869 号

内 容 简 介

随着人类和其他物种基因组的破译,生物信息科学的发展,生命科学已成为 21 世纪最为活跃、发展最为迅速的先导性学科之一。国内各高校十分重视生命科学的发展,相继设立了本科生创新科研项目以及研究策略和实验设计方面的课程。研究生培养也以科学研究水平为标准。本书作者基于多年教学科研经验以及与著名专家学者交流的基础,对生物学研究策略和实验设计进行深入细致的解析。

本书可作为生命科学专业高年级本科生和研究生的教材和参考资料。

Experimental Design for Biologists, by David J. Glass.

Copyright © 2007. Translation rights arranged with the permission of Cold Spring Harbor Laboratory Press.

图书在版编目(CIP)数据

生命科学实验设计指南/(美)格拉斯(Glass, D. J.)著. 丛羽生等译. —北京:科学出版社, 2008

(生命科学实验指南系列)

ISBN 978-7-03-021341-9

I. 生… II. ①格…②丛… III. 生命科学-实验-指南 IV. Q1-0

中国版本图书馆 CIP 数据核字 (2008) 第 031866 号

责任编辑:夏 梁 席 慧/责任校对:包志虹

责任印制:钱玉芬/封面设计:耕者设计工作室

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2008 年 5 月 第 一 版 开本: B5 (720×1000)

2008 年 5 月第一次印刷 印张: 11 1/4

印数: 1—4 000 字数: 213 000

定价: 35.00 元

(如有印装质量问题, 我社负责调换〈环伟〉)

前 言

这本书是我在生物制药公司 Regeneron Pharmaceuticals 开设一门关于实验设计的短期课程时的产物。我要对在课程中给予我回应的所有人表示感谢。像很多想法一样，将这门课写成一本书的想法最初来自于一番酒后之言——当时我正在和科学家同事 Dan Finley, Fred Goldberg 和 Allan Weissman 讨论实验设计并没有在研究生院里被广泛教授给未来的科学家这一事实，并且批判理性主义和科学实践之间有着巨大的差距。

当然，如果没有出版商想出版的话，这本书还是不能够问世。因此，我非常感激冷泉港实验室出版社的 David Crotty。在 Ginger Peschke 和 Maria Smit 的帮助下，非常能干的 Siân Curtis 将手稿编辑出来。Jan Argentine 和她出版社的同事审查了稿件。特别感谢 Siân 和 Jan 自始至终的反馈和热情。也同样非常感谢 Rene Steuer 专业的产品介绍，Susan Schaefer 的排版，以及 Denise Weiss 专业的设计。

在编写本书的过程中，Kumar Dharmarajan 给了我很大的帮助。Kumar 是我实验室曾经的学生和实习生，并在本书完成大半的时候成为了哥伦比亚大学的一名医科学生。因此他扮演了“未来的读者”的角色，并为每一章都提出了宝贵的意见，找出表述不明的地方，并帮助改写。非常感谢他为此做出的贡献和献出的宝贵时间。我实验室的 Brian Clarke 也阅读了大量手稿，筛选出了许多需要明确的章节。

Woody Fu——我以前的学生——现在承担了许多工作，包括卡通制作。他为本书绘制的三幅漫画如此之好，以至于我们为无法进一步发挥他的才能而感到遗憾。

还要感谢诺华公司的全体成员对我这一工作的支持。诺华公司对于继续教育非常重视。发现本书激起了公司的热情是一件非常快乐的事情。特别的感谢还要献给统计学家 Leah Martell 和 Mathis Thoma，他们仔细审查了书中的数据以避免统计学上的误述。如果书中仍存在问题，那都是我一个人的错。听到读者的反馈我将非常高兴，你们提出的意见和问题将在再版时进行改正。我的电子邮箱是：expdesignbiology@gmail.com，请随时与我联系。

DAVID J. GLASS

目 录

前言

1 确定实验项目	1
2 假说是科学研究课题的框架	5
3 科学地处理不切实际的假说	14
4 把问题作为科学研究的基准框架	18
5 是什么构成了实验问题的合理解释?	29
6 如何用实验结论来描绘现实	31
7 为实验建立模型	44
8 设计实验	49
9 将模型合理化	57
10 设计实验方案	62
11 实验重复	87
12 阴性对照的重要性	98
13 阳性对照的重要性	111
14 方法和试剂	126
15 研究对象对照	134
16 假定对照	145
17 实验者对照	153
18 有关生物教条主义的描述	159
19 简短概述	165
索引	167

1 确定实验项目

在医学院校中，学校要教授学生基础的医学技术：如何进行身体检查，如何解释症状，还有如何进行不同的诊断等。在法律学校中，最基础的专业知识是在实习的法院中获得的，学生学习如何撰写摘要，如何依照法律程序进行准备，律师如何在各种情况下保持自己的职业道德，还有如何以一名职业律师的思维方式进行思考。

在本书的写作过程中，作者发现一个可笑的事实：虽然良好的实验设计、规范的实验操作，以及严谨的表达能力是成为成功科学家的必要条件，但是在美国的大学中，这些即将获得博士学位的高才生却几乎不能进行正式的实验设计。一项对博士课程的调查显示：几乎在所有的大学，生物学、生物化学，以及相关的博士生所修课程都是在讲述目前已经被证明的科学知识，却没有强调如何应用正式的实验方法来解决未知的问题，如何通过已有的现象总结出新的理论和如何保持正确怀疑的世界观认知论^①。目前美国已经有少数人意识到这种教育缺陷，并通过统计学进行分析，但是如果进行大规模的教育改革，还需要更为科学的规范。这些统计数据可以在很大程度上促使人们发现实验设计的重要性。而如果学生具备较为开阔的实验思路，一定的实验设计能力，最为受益的将是博士生自己和他的导师。

对于一个见习科学工作者来说，一种正式的程序性教育能否可有可无呢（毕竟作为科学工作者，实验数据比哲学理论更有说服力）？是不是有什么现象值得我们关注从而进行实验呢？是不是千百年来建立起来的科学的哲学思想最终都要升华到现代的科学实验呢？这些问题的答案是否定的。本书的主旨在于提醒科学家：以上问题能够影响实验的选择和对结果的理解。因此，科学的实验方法学可以引导科学家发现更为有用的成果。

在通往科学成功的道路上，必然会遇到挫折，而思想哲学能够帮助科学家分析成功与失败。在正规教育中，许多科学家刻意避开科学与哲学“紧密结合”的理念，主张 100% 的自然科学，因为他们认为所有的自然现象都可以通过实验结果进行诠释。而作为实验数据，最重要的就是其自身的准确性，不能因为自己的思想而篡改实验数据。在实验中，我们还会遇到与预期结果不符的数据，而其实

^① 认知论是研究科学的科学，它通过各种参考资料来研究某种科学的局限性与有效性。科学的认知论是本书的基石，它对科学工作者具有特殊的重要意义。

这些数据同样是重要结果，科学家有可能通过这些数据的分析得到相关联的意外发现。因此，作为实验者来说，必须建立严谨的科学态度，在实验过程中必须慎用自身的假设，以免曲解实验结果。

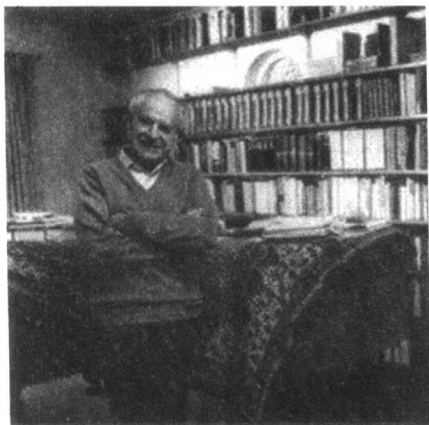


图1 卡尔·波普尔。

让我们回到建立良好的科学观念的问题上来。不到100年前，哲学家卡尔·波普尔（图1）首次提出了实验假说的概念。“理性批判”是波普尔的主要思想，他指出如果要证明一个理论的正确将非常困难，而相对地，如果要证明一个理论错误就要容易得多。因此他提倡实验的设计应该从假设出发。而实验假说更可以帮助学者从问题以及经验中归纳推理出新的理论。与此同时，波普尔以及其他哲学家（例如伊曼努尔·康德，大卫·休谟和伯特兰·罗素）认为个人经验总是有限的，因此如果通过

有限的个人经验来制定规则或者得出结论将导致错误。而且通过休谟的归纳推理，一个适应于过去的经验，在未来很可能是有瑕疵的。因此，与其说经过多元的分析与讨论去证明某一概念的正确性，不如从反面出发去证明另一概念的错误。在这种反证法中，只要能找到一个反例，就足以推翻原来的概念（这几句话概括了好几卷、数百页的波普尔的著作。这种思想将在本书的多个章节反复重申，而本书的要义就在于帮助学生快速转变为科学工作者，为他们指明做好实验的必要条件。而本书哲学方面的内容将仅限于较浅的层次，只会指出必需的哲学观点，以帮助科学家培养自己的兴趣以及为其思维方式的形成提供参考）。^①

波普尔的理论目前已经渗透到了科研的每个理论，美国国家卫生研究院现在推荐，每一份基金的申请都从一个可行的假说出发。不过有趣的是，国家卫生研究院并不反对逆向的思维，即使与波普尔的理论相悖，期望直接证明一个假说的正确性也是被允许的^②。波普尔的理论作为单一的哲学思想是极其成功的，但这是否意味着假说是进行科学研究的唯一起点呢？答案是否定的。波普尔理论的实

① 如果你只需要单一的参考资料，那么试一下《科学发现中的逻辑》（*The logic of scientific discovery*）。作者：波普尔，2002年于纽约 Routledge。

② “许多由美国国家卫生研究院资助的研究项目是基于申请者所提出的重要假说，而不是先进的研究技术。设想以你的假说作为你课题的结构基础，并由此变成申请项目的基础。一般来讲申请是要提出问题，然后去证明这个假说的正确或者错误，而并不是指出查找问题的方法或是简单的数据罗列。”（下载于美国国家卫生研究院基金申请要求网页）[http://www.niaid.nih.gov/ncn/grants/plan/plan_c1.htm]。

际效果将在本章的后面进行具体讨论，并介绍与此接近的相关理论。

卡尔·波普尔并不是第一位科学哲学家，在科学实验中已经提出了一些候补的办法。大部分历史学家似乎都同意，现在的科学方法是建立在两种不同的思想上，一项由弗朗西斯·培根（图2）提出，另一项由雷内·迪卡尔（图3）提出。尽管这些内容超出了本书的范围，但读者应该了解，这两位伟大的科学家提出了现代科学的基石——归纳法和演绎法（这当然是对培根和迪卡尔思想理论的一次严重简化，但是在本书的以后章节中，很多方面都是从他们的论点出发的，而其中的一部分，纠正了波普尔理论中的一些错误）。



图2 弗朗西斯·培根。

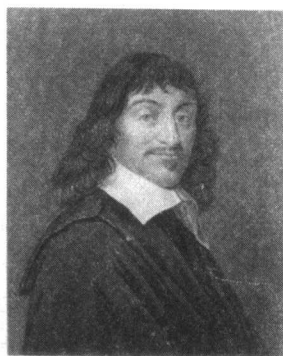


图3 雷内·迪卡尔。

由此我们可以看出，科学方法并不一定要牢牢地套在某一哲学思想背景下。我们必须把思想从实验室中解放出来，以科学的世界观观察试验的过程。我们能看到什么？什么是发掘未知知识的第一步？让我们把它叫做“实验项目”，就如韦恩的图表一样（图4）。实验项目包括所有的科学家未知的因子，它是科学家脑中的宇宙，说得更广阔一点，它是科学认知论。

当一个科学家研究某一项特定课题时，这个方面的内容就可以看作实验项目的一个亚集，在这方面，一个实验课题是指进行特殊的实验调查以研究某一特定的问题。很多科学家并没有意识到他们的课题被实地调查以外的东西限制住了，这些东西就好像一个框架。而这些框架就是哲学的思维构成，它用一样的思维去处理每一个课题，处理每一组实验数据（图5）。因为不同的思维框架就需要不同的实验方法，因而哲学思想的框架冲击了科学方法本身。在第2章中，我们将以波普尔的假说理论为基础，对这种思维框架进行进一步讨论。

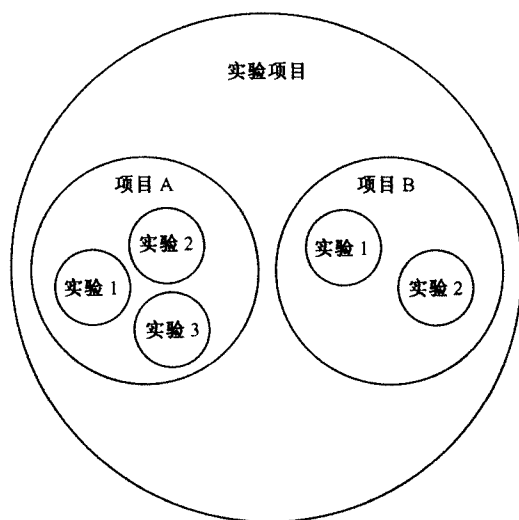


图 4 韦恩图表说明实验项目，它强调科学家必须对未知事物具有准确的预测。它可能涉及学习科学技术、训练、对论题的阅读理解能力并进行实际的实验。在科学家的“计划”中，实验是针对特殊的未知事物去回答一系列特殊的问题。实验既可以回答问题也可以验证假说的真伪（图中，项目 A 包括实验 1、实验 2 和实验 3，项目 B 包括实验 1 和实验 2）。

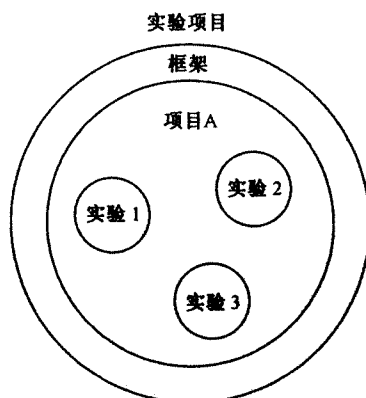


图 5 韦恩图表说明实验框架包含于实验项目之内。科学家的“计划”以假说或者问题为框架。实验的框架对于检测科学家计划成功与否以及把独立的实验结果统一起来具有重要意义。

2 假说是科学研究课题的框架

——批判性理性思维是关键吗？

在详细讨论将假说（与之对应的是批判性理性思维）作为实验的框架之前，必须意识到实践型科学家除了提出假说一般没有别的选择。正如第1章讨论的那样，某些资助机构（以及评论家，甚至某些科学期刊）要求对假说做出清晰明确的解释作为进行实验的理由。但是，在意识到批判性理性思维框架的缺陷以及深入了解与之对应的假说之后，即使仍以范式理论来指导工作，还是希望科学家能够运用适当的理性方式来解决问題。

假说与结论一样具有相同的语法结构

考虑以下假说：天空是红色的。首先，注意该假说的结构，它是一个声明性陈述。因为一般情况下提出假说是为了推翻它，因此假说必须采取能被推翻的声明性陈述形式。比如一个疑问句，就不具有可推翻性，也就不能够作为理性思维的结构框架。如果在实验中假说没有被推翻，那么明显可以得出结论：该假说是正确的。这实际上也正是卡尔·波普尔所指出的一种错误想法。他认为不可能证明某些结论是绝对正确的。实际情况很简单，就是在继续进行实验项目的时候，必须得出一个结论。在这个假说中，结论就是“天空是红色的”。

注意结论的结构与假说的结构是一致的。这就为科学家辨别未经证明的假说（需要通过验证推翻的）以及已验证的结论带来了更大的难度。

实验旨在验证而非推翻假说，因为假说倾向于 获得预期的“阳性”结果

某些实验方法已经逐渐成为标准。乍看上去，这些方法似乎很有用。举个例子就是运用已经建立起来的原理作为科学指导去否定一个假说。

如果我们回到“天空是红色的”这个假说，看起来似乎轻松地判断一下是否“红色”就可以验证这个说法正确与否。即使最没有经验的研究员也知道要引入一个阴性对照来判断红色是否存在。因此我们可以检测是否为红色，但是为了证明我们的假说，是否还需要引入其他颜色呢？显然没有必要；在假说里面，没有地方提到需要检测“红色”以及“非红色”以外的任何对象。

我们先来看下一个证明“天空是红色的”假说的实验设计：首先，科学家建立起一个系统能够精确地检测到红色。一个已知的红色物体作为阳性对照，比如一朵红玫瑰。另一个已知的非红色物体作为阴性对照，比如一块绿翡翠。我们用一个“红色检测仪”测量辐射光，将它校准，该检测仪应具有“红色”、“非红色”两个读出器。如果检测到红色，则会出现一个大大的“确定”标志。

准备好实验体系后，科学家将“红色检测仪”朝向天空，让其运行1小时。1小时后，什么现象都没有：“确定”的标志没有出现，因此记录的是一个阴性结果。为了确保“红色”出现的机会，实验的时间设计增加到了4小时。因为实验是从中午开始的，因此日落在实验设计时间内不会发生，得到的仍然是一个阴性结果。这样，科学家会认为可能是测量仪灵敏度不高，并且一天是有24小时的，以4小时为周期并不合理。因此，实验进行重复，以24小时为周期。经过一整天的测量之后，科学家观察测量仪，发现“确定”的标志出现了（在实验时间内黎明和日落时出现了红光）。由于得到了一个阳性的结果，这个实验可以看作是成功的。

科学家将该实验每天不间断地重复了一个月，这里需要指出的是，除了天气明显阴沉的一天外，每天都能检测到红光。科学家进行统计学分析^①，将检测到红色的次数（30次中出现了29次）与非红色的次数（30次中出现了0次，因为天空从来没有出现过翡翠绿的颜色信号）进行对比，发现对红色的检测是具有统计学意义的。因此可以得出天空是红色的结论。

你可能会反对以上陈述，认为那是一个精心编撰的攻击对象。但是难道这（至少部分）不是因为你^②知道天空通常是蓝色的（在白天）或黑色的（在夜间，你不考虑星星、月亮或其他星体发出的光的情况下），所以才意识到过多额外的对照可能会不利于证明天空是红色的？我们有必要对这些额外的对照进行讨论。由于有效的声明性陈述本质上是受限的，因此同时证明了假说的结构框架会使这些对照的存在意义变得模糊。记住，在大多数情况下，科学家都不知道真正的结果，因此真正指导测试的就是假说。虽然在“天空是红色的”这个例子中，科学家很明显应该测量整个光谱中的单色光才能得到更全面的结果，但是对于“NF- κ B^②通路的激活会引发炎症”的假说，我们就不会那么确定量度的对象了。没有关于NF- κ B通路的背景知识，即使和“天空是红色的”这个假说在情景设定

① 详尽的统计学讨论超过了本书的范畴。尽管最好的统计分析也无法使一个蹩脚的理论框架显得合理，值得注意的是，如果善加运用，统计学将成为你的朋友，避免你做出不合适的假设推断，让你在实验过程中及时改变各项参数。

② NF- κ B是一个蛋白质的名称。重要的是可能绝大多数本书的读者对于该蛋白质感到陌生：当需要解决的问题是未知的物体时，假说的作用就显得尤为重要，也更具有高度。这里，科学家就缺乏外部的知识去揭示“天空是红色的”假说中存在的问题。还是满足一下大家的好奇，NF- κ B行使转录因子的功能，即它能够诱导一系列的基因转录成为信使RNA。

上具有相同的问题，我们也无法判断这个假说的正确性。例如，在第二个假说中获取的数据可能存在着极大的变化，这取决于细胞的类型（在肝细胞中 NF- κ B 通路的激活会导致炎症，而在骨骼肌中却不会）。然而，这个假说正如所写的那样，不需要用实验表述组织特异性反应或者对于炎症的明确定义。因而，在第一个假说中，科学家将实验时间延长直到出现阳性结果为止，以此类推，在证明“NF- κ B 通路的激活会引发炎症”时，科学家会认为有必要寻找一个敏感的系统（即能够产生阳性结果的系统）来进行实验。同样，在第一个假说中，简单的事实就是天空偶尔会是红色的；在第二个假说中，则是 NF- κ B 通路似乎能够在多类细胞中诱导炎症的发生。这里需要指出的是对阳性结果的数据或阴性结果的实验设计所强调的是不同的。在第二个假说里面，因为 NF- κ B 应答机制是具有细胞以及环境特异性的，一个只能够获得阳性结果（发生炎症）的“筛选体系”会像看着天空只记录下了红色的结果那样提供不完整的实验结果。

关于 NF- κ B 假说，第二点需要指出的是：判断阳性结果的标准（即什么是炎症现象）可能会因为得到的阴性结果而改变。例如，假设炎症的标志物^①有 5 种，然而 NF- κ B 只能激活其中 2 种，那么阳性结果只在这两个标志物被激活的时候才存在。对于其他的标志物，这个实验设计是否合理仍然不为人知，并且科学家会觉得只报道阳性结果是合理的。在这种情况下，统计数据需要进行一定的修正，赋予其新的意义。几乎每一个合理的统计学分析都会让实验者进一步定义阳性结果的标准。如果假说会影响结果是阳性的还是阴性的，那么其他 3 个阴性标志物作为研究中的标志的可信度将会受到质疑。

我们必须在此强调一下最后一个观点。如果实验产生了一些阳性及阴性的数据，对于这些结果的不同处理过程实际上会促使科学家重新定义研究的课题。这是一个很严重的瑕疵。

假说针对数据类型建立起一个二元 （阳性及阴性）筛选系统

回到“天空是红色的”这个假说，很明显，科学家设计了“红色检测仪”作为数据筛选体系。该检测系统只能够记录到红色以及另外一种颜色，这种颜色只需要具有是非红色这个相关特性就可以。为了避免你认为在实验中以“绿色而非

^① 在生物学术语中，标志物是一个诊断标准。例如，右下腹部疼痛是阑尾炎的标志。高血清葡萄糖含量是糖尿病的标志。IGF1 蛋白含量的提高标志着生长激素的增加。

但是如果能够在实验前建立起成功的衡量标准（详见第 12 章）也是能够避免的。我们现在讨论一下为什么筛选过程在理性思维框架中极可能发生。

黑色”作为阴性对照已经足够了，还应认识到即使实验中将黑色作为阴性对照，观察的结果“天空是黑色的”也不会被理解成阳性结果。重点是科学家必须建立起一个特殊的阳性数据指标作为红色信号输出标志。在实验中，设置一个适当的筛选体系精确地记录阳性结果是很普遍的。即使整个夜晚都会检测到黑色，但是仪器仍然将黑色作为相对于红色的一个对照。因此，为了利于小概率的阳性结果的显示，大量表示阴性结果的情况（天空是黑色的）在无意中被舍弃了。

现实生活中，可以通过“咖啡因会导致癌症”这个假说确定出一个能够得出阳性结果的筛选体系。首先，应该进行流行病学研究，确定喝咖啡、茶和可乐饮品（所有这些都是含咖啡因的）的人是否比没有摄入咖啡因的人有更大的概率患上某种特定的癌症。假设调查中发现咖啡因引发的癌症一共有 100 种。即使统计学差异仅在一种癌症中存在，比如胰腺癌，那么这种癌症会被认为是阳性结果。而从其余 99 种癌症获取的数据将会被标记为阴性结果。这样科学家可以得出这样的结论：咖啡因导致癌症至少在胰腺癌中是成立的。或者，更确切地说，摄入咖啡因会增加个人患胰腺癌的概率，但是在其他癌症中没有什么影响。顺便提及一下，这是另一个运用完整的统计学知识来解释结果的例子。一个造诣高深的统计学家会知道由这些数据得出的任何结论都是不合适的，因为实验前期只是进行了调查研究，并且没有确定是适合每一种癌症的。不幸的是，在经过这些调查后经常会得出不合适的结论。举例来说，“咖啡因会导致癌症”这个假说的框架在“得到一个具有统计学差异的相关性”和“没有得到该相关性”之间建立起了一个二元阳性/阴性区分系统。在检测的 100 种癌症中仅有一种会得到阳性的结果，这种癌症被凸显出来了，因为标题是咖啡因会增加患胰腺癌的概率，而非咖啡因不能增加患其余 99 种癌症的概率。

咖啡因是不是真的可以增加患胰腺癌的概率，而不会增加其他癌症发生的概率呢？这当然是可能的，也可以从上述研究中得出结论。但是，科学家必须进行重复实验加以确定后才能够得出这样一个结论。如果第二次研究仍然在理性思维框架指导下进行，那么假说就应该改成“咖啡因会导致胰腺癌”。在第一个实验里，发现对 100 种癌症的大量调查研究中，胰腺癌的数据具有统计学意义。虽然要避免详细的阐述，但是少量的统计分析在这里还是很有用的：如果第一个研究中获取了合适的统计数据^①结果显示概率（ p 值）为 0.05（0.05 是常用的表示具有显著性差异的阈值），那么也有 5% 的随机可能情况证明该差异不是显著的。因此，在对 100 种癌症的研究过程中，随机出现有一种癌症具有阳性的结果也是不足为奇的；实际上，在 100 种癌症中，应该会随机出现 5 个阳性结果。但是如

① 可能运用的是 ANOVA 分析法得出的数据统计法。ANOVA 是用来分析多个群体中的计量型数据，以便比较变异的意义和分析其来源。

果第二个实验仅集中在分析一种癌症上，那么该种癌症随机出现阳性结果的概率也只有 5%。在两个独立的实验中，一种癌症随机出现阳性结果的概率应该是 0.05×0.05 ，或者说 0.25%，这意味着仅重复一次实验，就可以极大地增加实验结果的说服力（对于实验的重复将会在第 8 章进行详尽阐述）。这里需要强调的是这个阳性筛选体系至少将范围宽泛的假说集中到了 100 种癌症中的一种身上，而不是得出更可能的“咖啡因不会增加癌症的概率”这个结论。相信咖啡因会增加患胰腺癌的概率，那么肯定会得出这样的结论：在引起胰腺癌的通路中存在着某些对咖啡因特异敏感的物质。这比发现一种特殊的致癌物质会导致另一种癌症（如石棉会导致间皮癌）更特别，因此，否定阳性结果是不合适的。这就需要对这些结果进行适当的处理，任何阳性筛选体系的结果必须经过更深入的研究，在更准确、更直接的假说经过确认以后才能有所定论。在进一步的确定性实验中，假说必须试图推翻那些起初会产生阳性结果的特例。如果运用理性思维进行严格的推理，那么理性思维将会非常成功。

验证“咖啡因会导致癌症”这个假说有多种途径。在第二轮的实验设计中，科学家不会采取流行病学研究了，而是进行药理学实验，建立一个系统，用咖啡因含量逐渐增加的食物喂养大鼠——加入咖啡因量等同于那些极其依赖咖啡的人群的摄入量，进行一年的研究，对比没有摄入咖啡因、同窝生的大鼠，那些大鼠患上肿瘤的概率并没有增加。

实验又一次陷入了进退两难的境地：是应该肯定实验的阴性结果还是应该将实验继续进行下去？科学家很可能将实验继续进行下去，因为他们认为长年暴露在恶劣的外界环境中才会导致癌症，达到预期的效果^①，必须增加药物的剂量。因此，将大鼠摄入咖啡因的剂量增加到正常咖啡因日摄入量的 100 倍，每天如此，坚持一年。除了变得十分急躁之外，这些摄入咖啡因的大鼠确实比没有处理的阴性对照具有更高的肿瘤发生率，而这个概率实际上很有统计学意义。科学家可以得出结论：咖啡因会导致癌症，并且高剂量的咖啡因对于模拟人类数十年地摄入咖啡的状态是十分必要的。

再次，这两个结论（“咖啡因会导致癌症”以及“100 倍剂量的咖啡因增加准确地模拟了长期正常水平的摄入量”）很有可能被证明是正确的。但是这些确切的结论是从这些数据中获得的吗？还是实验假说的结构框架促使科学家不断增加剂量直到出现阳性结果为止？

也有可能正常剂量的咖啡因不会致癌，但是很高剂量的咖啡因（如正常剂量 100 倍的剂量）会触发癌变过程。关于这一点，我们会在第 16 章详细讨论建立精确模型的问题。有时，要在生理水平上设置一个实验模拟一个基因产物是不可

① 注意这里的措辞：“预期的效果”。科学家有必要更期望某一种结果吗？

能的；相反，极大地增加被检测药物的剂量产生有用的数据却是可能的。但是却不能因为预定的需要去证明一个未经证实的假说而进行这些实验。实验必须控制在正常环境浓度下模拟导致癌症的过程，如一个特定的机制会引发其他的结果。换句话说，如果我们得出咖啡因会导致胰腺癌，那么假说的结构框架就不需要更深入，例如，寻找其他更可能导致胰腺癌的因素进行研究。相反，应该从讨论中去掉这样的研究（为什么要多此一举，在假说被证明时去寻找其他的因子呢？）。

下面再举一个现实生活中的例子来说明建立的筛选系统不适合证明假说。让我们用“X 酶的激活会导致癌症”的假说来讨论，这是基于先前的一些发现的。假设科学家从结肠癌细胞中克隆了一个具有活性的 X 酶。为了验证这个假说，科学家构建了一个遗传载体，编码持续表达活性 X 酶的突变体。作为阴性对照，科学家构建了另一个遗传载体，编码正常的不具有活性的 X 酶。编码突变 X 酶和正常 X 酶的载体被注入细胞。转入持续表达活性突变体的细胞确实表现出癌细胞的表型，并且在注射到小鼠体内后会致瘤。而转入正常 X 酶载体的细胞没有发生癌变，也不会使小鼠成瘤。该实验重复了 10 次，仍然得到相同的结果。看起来似乎结果很好，我必须承认如果在我的实验室中得到这些数据，至少起初我会十分兴奋。但是我们还是需要更谨慎一点。

如果我告诉你实验所用细胞正常情况下不表达 X 酶呢？如果我进一步告诉你通常会表达 X 酶的细胞对于转入持续表达活性的 X 酶具有不同的反应呢？第二类细胞不会发生癌变，而是分化成了神经元。现在我们的假说“X 酶的激活会导致癌症”还能站得住脚么？我们开始怀疑假说的正确性。但是先别急，如果我们进行更多的实验，发现结肠细胞正常情况下不表达 X 酶，但是在特定的条件下不适当地表达该酶，而此时 X 酶确实会诱导结肠细胞发生癌变呢？我们现在可能会提出一个新的假说：X 酶会导致结肠癌。我们会觉得这个较全面的假说更合适么？毕竟，这个基因最初是从结肠癌细胞中克隆出来的。但是如果说在调查的 100 种结肠癌细胞中并没有发现 X 酶的活性。那么，即使突变的 X 酶能够导致结肠癌的发生，它也不能被认为是疾病中常见的调节因素。最后，有一个新的发现，在脑瘤中，当细胞表达 X 酶时，更可能发生的现象是：肿瘤被治愈并且分化成正常的神经元，这种转化是通过用另一种蛋白质处理癌细胞激活 X 酶实现的。因此我们从科学家致力于研究“X 酶的激活会导致癌症”的假说，转变到了发现 X 酶的激活可以阻断癌症的发生。那么，明确地阐述假说结构框架及实验过程在实验课题中到底是有利还是有弊？假说确实促进了检测酶是否会导致肿瘤发生的实验。但是，假说没有要求对癌细胞本身会不会表达基因进行分析。并且假说也没有要求科学家确定该基因通常情况下在哪些细胞中是活跃的，或者它的活性怎样影响这些细胞。最后，假说也没有要求研究正常发生的肿瘤中这些酶是否被激活。这个假说提出的唯一的疑问就是这个基因被激活的时候会不会导致肿瘤的发生。在第 4 章中，我们将会讨论，

结构框架的改变会不会更利于有效实验的设计^①。

以上情景显示根据假说的本质，可以建立起一个阳性/阴性筛选系统，但可能会使非阳性数据被无意中舍弃，也可能使阳性数据处理不当。

科学家只有在证实了假说的时候受到奖励， 而在推翻假说的时候则不会

以上情景证实了这样的情况：在获得阳性的结果前，实验设计都会几经波折。不要打击个人的整体性或者是指责他们目的错误，因为阴性的结果会使科学家面对实验设计是否正确的棘手问题，对结果的评判标准也可能会改变。换句话说，科学家可能产生这样的动机：持续改变实验的组成条件直到出现阳性结果为止。如果能够觉察到这种动机，就必须警惕。记住，掌握好统计学是很有用的，很多统计学框架需要科学家在实验研究前决定阳性结果的评判标准。

另外一些因素可能会鼓励人们去追逐阳性数据。有人认为人类的动机以及心理因素在实验中可以轻易地剖析出来，要控制这种科学中的人为因素，还是必须将几个明显的动机因素提出来。事实上，科学家证实一个假说会比推翻一个假说获得更多的奖励（即使这个假说并不期望得到确定）。高水平的（即使是低水平的）期刊通常不会发表阴性结果，到底是为什么呢？谁愿意读到一个观点，但是最终却被推翻了呢？更重要的是，资助机构凭什么愿意奖励提出的那些十分出色而最终却被证明是错误的假说呢？一个大学的系主任凭什么会继续任用一個不能成功找到支持假说的证据的研究员呢？

可能科研新手在意识到假说结构框架的深远意义时会十分惊讶。一旦实验的判断标准确定下来，那么它也成为了评判科学家的标准，阳性/阴性二元性也会被延伸到对科学家的评估中。而其他替代性结构框架将会更不容易受到这个因素的影响。

科学家在证明一个假说的正确性时会投入 情感的因素，替代性思维结构框架在 头脑中会使科学家更具有免疫力

关于心理因素对于研究员的影响已经进行了充分的探讨，但是却很少分析到

^① 顺带提及一下，X 酶确实是存在的。酪氨酸激酶 TrkA 受体在一种结肠癌细胞中通过突变被激活，对该受体进行了克隆。虽然并不明显，但是 TrkA 是结肠癌中一种普遍的调节因素。当在成纤维细胞中过表达时，TrkA 会使细胞发生转化而具有癌细胞表型。但是在神经组织起源的 PC12 细胞中被激活的时候，TrkA 会诱导细胞分化成神经元。因此，虽然故事有点复杂，但是它证明了一个观点：一个简单的假说可能会更不利于揭示这样一幅“画面”，因为它建立起来的是一个没有用的积极或者消极的思维框架。

实验工作者的实验设计以及数据解释^①。但是一旦意识到潜在的情感影响，则它是很容易消除的。

从根本上说，如果一个科学家构建了一个假说，他会为之自豪。但是，由于假说与结论具有相同的语法结构，科学家对自身的观点投入情感会使已证明及未经证明的观点之间的分界点变得更加模糊，或者说是鼓励科学家去证实假说。

同样，如果实验室里一个主要的研究员建立起一种假说，指导一个年轻的研究生去研究这个假说，至少有些担心这个研究生会将假说奉为真理而产生偏见。如果结果是客观的，那么这还不会成为一个难题，这是指数据收集员不用对数据进行解释，但是如果数据具有主观性成分，那么就要小心使观察者不要受到假说的影响。“盲”观察将在第12章中进行阐述。对于现在来说，要注意到，尽管研究员的偏见以及心理因素对结构框架的影响是可以控制的，但是除非意识到了这个问题，否则在实验设计中适当的控制方法还是很少被提及的。

其他的思维结构框架相对来讲 可能不会受假说问题的影响

作为一个实验员，你可能会完全地、强烈地反对最后几个观点：假说会不适当地促成收集支持其观点的数据的动机。但是如果存在着一个替代的结构框架使产生的结果更公正呢？如果有办法设计实验使所有的结果注定是阳性的，从而产生的都是有用的数据呢？即使你相信理性思维十分有效，难道那不是更受理性思维方法的推崇吗？

接下来的章节建议运用一种替代的思维框架，并且指出会导致假说无效的设计，从而凸显替代性思维框架的作用。

总结及告诫

本章中，对理性思维进行了一定的批判，但是这种哲学方法即使在已知的科学探索最鼎盛的时期也是占据统治地位的。因此，很明显可以看出假说确实使很多有用的实验能够进行下去。但是有经验的科学家会意识到不适当地运用假说会产生无用的解释，因此会预感到理性思维结构框架中存在的潜在问题。下一章会提供一种替代理性思维的方法——仍然可以提出假说，但是却不会使实验框架受

^① 但是，波普尔（《科学发现的逻辑学》，2002。纽约，Routledge）实际上驳回了心理学与一个观点的逻辑检测无关的想法。他反而强调在一个心理学能起作用的场景下，一个想法是浮现在脑海的心理过程，而不是有效的可以引起捍卫自身观点的心理学。大家可以在这冗长的话中读到讽刺的意味。