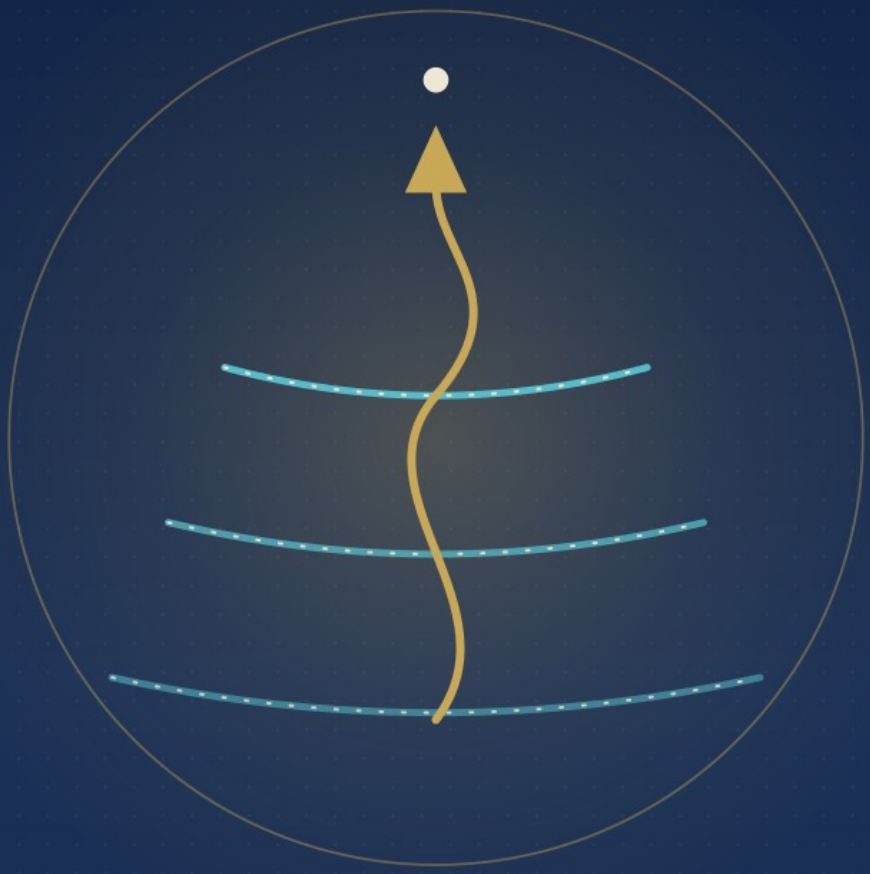
# 典范发生学

SDE 与 AI 时代的科研层级



王德生　　著

德麦国际出版社
Demai International Press

# 典范发生学

SDE 与 AI 时代的科研层级

王德生　　著

德麦国际出版社

Demai International Press

# 出　版　信　息


| | |
|---|---|
| 书名 | 《典范发生学——SDE 与 AI 时代的科研层级》 |
| 著者 | 王德生 |
| 执笔 | 复合主体：王德生立题、口述承重命题与裁定，Claude 执笔与统稿 |
| 出版 | 德麦国际出版社（Demai International Press）· 新加坡 |
| 编号 | 德麦国际专著第 369 号 |
| ISBN | 979-8-90690-515-4 |
| 定价 | US$20.00 |
| 开本 | 170 mm × 240 mm（16 开） |
| 字数 | 约 20 万字 |
| 版次 | 2026 年 10 月第 1 版 |
| 完稿 | 2026 年 10 月 18 日 |

AI 协作声明：本书的判断、口径与承重命题由王德生（口述）给出；文字在其指令下由 AI（Claude）协作成稿，并经其审定。

引文说明：书中涉及丘成桐教授的说法，均据转述，作者未核对原文，只取大意，不作原话引用；科学史部分只取有把握的事实，拿不准处已作模糊处理。

关于 SDE-Claude：书中对作者自己的实践只作说明，不作“已成立”的断言，读者可按第 11 章、第 15 章的判据自行检验。

全书结构：导言＋六编二十章＋后记＋附录（典范发生工序卡）。


版权所有　侵权必究

# 作 者 介 绍

王德生，博士，SIO 本体论与 SDE 发生学的创立者。

湘潭大学数学系本科、硕士，中国科学院数学研究所博士，中国科学院计算数学研究所博士后。曾任英国威尔士大学斯旺西分校高级研究员、新加坡南洋理工大学数学系助理教授、博士生导师与资深研究科学家，在计算数学与图形学国际期刊发表论文四十余篇。

2015 年起回国推广三视角方法论，后凝练为 SDE 发生学。

现任新加坡德麦国际（Demai International Pte. Ltd.）创始人、董事长，主持德麦国际专著出版体系与 SDE Universes 平台。

# 导　言

## 一个流行想法为什么走偏

设想前些天，一位做中学教研的人转来一段话，问同行怎么看。那段话的意思，大致是这样的：现在 AI 什么题都能答，答案已经不值钱了，所以学校不能再只教答案，要转过来，看重学生思考的过程。

他说，这话听着对，他们学校开会也是这么讲的。可他心里有个疙瘩：把答案换成过程，真的就解决问题了吗？过程不也是可以让 AI 写出来的吗？

被问的人没有马上回答。因为这个疙瘩，就是这本书要说的事。

这个想法太顺耳了。它顺耳，是因为它每一步都说得通：答案便宜了，所以转向过程；过程也有人能代写，所以转向提问；提问也能让 AI 来，所以转向「人提出问题、AI 来处理、人再判断」。每一步都像是在往前走，每一步都有人写文章、有人开讲座、有人发文件。

可是，如果你把这几步的位置画在一张纸上，会发现一件怪事：它们根本没有离开过原地。

这本书就从这件怪事开始。

### ◆ 一个顺耳的想法

先把这个想法原样摆出来，不夸大，也不丑化它。

它说的是：以前，学生学知识，老师考答案。答案是稀缺的，谁能给出正确答案，谁就有本事。现在不一样了，一个孩子拿起手机，几秒钟就能得到一

个像样的答案，有时比老师给的还完整。既然答案不再稀缺，学校的价值就不能再押在答案上。

那押在哪里？押在「过程」上。就是说：这道题是怎么想出来的，中间走过哪些弯路，哪一步为什么这样选。答案可以由机器给，过程要由人亲自走一遍，走过的人才算学会了。

这个想法有它的道理。一个学生如果只会抄答案，他什么也没学到，这一点从来如此，不是 AI 出现以后才有的。老师们本来就一直在说「要重过程，不要只看结果」。AI 把答案变得廉价，只是让这句老话突然显得迫切了。

所以本书不想一上来就说它错，只想说：它在一个很具体的意义上「走偏」了。走偏，不是说它指的方向全错，而是说它顺着一条路一直走下去，走得越认真，离真正要走的路越远。

## ◆ 据转述的三点

在这场讨论里，经常被人引用的，是丘成桐教授的一些说法。要先交代一句：本书没有核实过原文，下面写的只是「据转述」，不是他的原话，也不能当作他的原话。

据转述，丘教授的意思是：答案已经变得廉价了，学校能够交付给学生的，已经主要是过程。因此，学校的重心要从答案转向三件事。第一，会不会提问；第二，能不能验证；第三，愿不愿意把自己的过程讲清楚。

这三点值得反复读。读下来的感觉是：这是一位长期做数学和教育的人，从实际经验里说出来的话，它不是空话，也不是随口一说。会不会提问，是研

究的起点；能不能验证，是研究的质检；愿不愿意把过程讲清楚，是研究能被别人接过去的前提。放到一个研究者的日常里，这三条每一条都实实在在。

所以，接下来要做的事，不是反驳这三点，更不是说丘教授「错了」，而是换一个量尺来量它：这三条，在整个科研的层级里，放在哪一层？

本书的回答是：这三条，仍然在典范之内。

## ◆ 什么叫「典范之内」

「典范」这个词，后面的章节会慢慢讲清楚。现在先用大白话给一个临时的说法。

你可以把一门学科想成一座城。城里有街道、有规矩、有大家公认的地图。什么算一个问题、什么算一个答案、什么算「做对了」，城里的人心里有数。住在城里的人，每天做的事是：沿着街道走，把还没走到的角落走一遍，把地图上空着的地方补上。

这座城，就是一个典范。美国科学史家库恩把它叫作「范式」，后来的人也译作「典范」。本书里一律用「典范」。

在城里做研究，是很了不起的事。大部分科学家，一辈子都在城里做事：计算一个数据，解一个方程，设计一个实验，把一个未解的小谜题解开。库恩把这种工作叫作「常规科学」。它不是什么低级工作，现代世界的大部分成就，都是这样一点一点做出来的。

但是，它和另一件事不同。另一件事，是换一张地图，或者说，让一座新的城出现。这种事极少发生，几百年里也就那么几回。日心说取代地心说，是

一回；牛顿力学被相对论和量子力学改写，是一回。库恩管这个叫「科学革命」。

好，现在回到那三点。会不会提问：是在城里问问题，问「街道的这一头为什么没有走通」。能不能验证：是在城里检查，检查「这条路走得对不对」。愿不愿意讲清楚过程：是在城里交流，把自己走过的路讲给别人听。

三条全在城里。

## ◆ 同一层里的三次平移

现在我们回到刚才说的那件怪事：这个想法的几个步骤，都没有离开过原地。

下面把它们一步一步摆开。

第一步：答案被 AI 做廉价了。于是转向过程。过程是什么？是从问题走到答案的那条路。答案是终点，过程是路径。终点和路径，在同一张地图上。

第二步：过程也被 AI 做廉价了。你让它「把推理过程一步一步写出来」，它写得有模有样。于是有人说，那就转向提问。提问是什么？是路的起点。起点、路径、终点，还是在同一张地图上。

第三步：提问也能由 AI 来。让它「针对这个领域提出十个有价值的问题」，它能提，而且提得不差。于是有人说，那就转向「人输入、AI 处理、人判断」。

这个方案把人和 AI 看成两个各守一段的独立工人，人在两头，AI 在中间。这里多出来一个位置：验收，也就是质检。可质检，仍然在这张地图上；而且这种分段站岗的画法，本身就在同一层里。

起点、路径、终点、质检。我们一路退，从终点退到路径，从路径退到起点，从起点退到质检。每一步都像是「往更高处走了一步」，可是如果你认真看，会发现这四个东西，其实是同一件事的四个位置。

就像一条流水线，你可以说「我们要重视流水线的最后一道工序」，过几天又说「不对，要重视第一道工序」，再过几天说「要重视中间的检查」。你在不停地换重点，可你始终在这条流水线上。

而 AI 的特点是：它跟着你平移。你退到哪一个位置，它就进到哪一个位置，成本就跟着降到哪一个位置。你每退一步，被它拿下的地盘就多一块。

这才是所说的「走偏」。想法本身没有错，错在这个方向一直是平移，没有上升。平移的人，永远在追一个越跑越快的对手，而对手还不累。

## ◆ 一场会议里的平移

再看一个开会的场景，因为平移这件事，在会议室里最容易看见。

设想某所大学开教学改革会。第一位发言的人说：「AI 时代，不能再考记忆性的题目了，要考理解。」大家点头。第二位说：「理解也能让 AI 代答，要考过程，要学生交出思考的草稿。」大家点头。

第三位说：「草稿也可以生成，要考提问，让学生自己提出问题。」大家点头。第四位说：「提问也能让 AI 帮着想，所以要考学生对 AI 输出的判断力。」大家继续点头。

会开到这里，每个人都觉得会议在推进。可是你若是把这四句发言记在黑板上，会发现黑板上写的是：记忆、理解、过程、提问、判断。五个词，是同

一份试卷里依次出现的五种题型。会议室里的人，是在为这份试卷不断地换题型。

没有一个人问过：这份试卷本身，是不是还值得考？

这就是平移的手感：每一步都有新词，每一步都像有进展，而整体的框架纹丝不动。要把这种手感说出来，需要一个能站在框架外面的词。这个词，就是「典范」，就是「层」。

## ◆ 为什么「还有一块是人独有的」靠不住

说到这里，一定有人会反驳：AI 再厉害，总有一块地方是人才行的吧？比如「真正的好问题」，只有人能提。

要直说：这条路，本书不走。

原因有两个。第一，它不符合事实。你让一个像样的模型，针对一个成熟的学科，列出这个学科里有待研究的问题，它能列出很多，而且其中相当一部分，行内人看了会说「这个问题问得好」。它对这个领域的文献、争论、空白，有一个很宽的覆盖。你说它「不会提问」，是经不起一次实验的。

第二，这种说法在方法上有问题。「AI 做不到的，才是人该守的」，这句话听着有力，其实是在划保留地：先看 AI 目前做不到什么，再把那一块圈起来，说这是人的领地。

可是 AI 的能力每隔几个月就变一次，保留地也就跟着缩一次。今天圈的，明天就丢。更糟的是，这种划法，让人把自己的价值押在一个别人随时可以越过的门槛上。

那么用什么标准？这本书用的标准是成本结构。凡是能被廉价生产的东西，就不再是竞争点。至于它是不是「真的」像人一样做出来的，不重要。一个东西便宜到几乎不要钱，就没有人再用它去竞争。

按这个标准，答案是便宜的，过程是便宜的，提问在一个给定的典范里也是便宜的，验收在很大范围内正在变得便宜。这四样东西，同时在往下掉。

## ◆ 车替代腿

为了不让话停在抽象里，不妨借一个日子里的东西来讲：车和腿。

人类用腿走路，走了几十万年。后来有了车。车能载人，能载货，比腿快得多，也远得多。

假设一个人在车刚出现的时候说：「车能走路了，腿就没用了。」我们会觉得他夸张。

再假设另一个人说：「车是出现了，但腿还是有用的。人要从『走路』转向『走过程』，也就是更重视走的步伐、呼吸、姿势。」听上去好一些，可是你立刻会觉得哪里不对。因为事实上，车取代腿的时候，并没有让腿去「重视走路的过程」。

事实是：腿退出了运输这个岗位。运输交给车了。腿去做别的事。腿上了车之后，人去的地方更远了，去的地方更多了，有了公路、铁路、航线、地图、红绿灯、调度。整个交通升级了。

这就是第一章要讲的事。它是这本书的一条绿线，后面各章会不时回来看一眼。这里先把结论放在前头：车替代腿，不是「腿走向过程」，而是交通升

级。腿不是消失了，它进到了一个更高的交通功能里。而且腿是地基，不是楼层：一个从来没有走过路的人，看不懂地图。

那么，在科研里，什么是「车」，什么是「腿」，什么是「更高的交通功能」？这三个问题的答案，构成了这本书的骨架。

## ◆ 把三点逐条放进课堂

光说「都在典范之内」，不够具体。我们换到一间教室里，看这三条各自长什么样，再看 AI 在每一条上的表现。

设想一节大学数学系的讨论课。老师抛出一个已有的定理，让学生自己想：这个定理的条件能不能放宽？

先看「会不会提问」。一个学生举手说：「如果把这个条件换成弱一点的，结论还成立吗？」这是个好问题。

可是你现在把同样的定理交给一个像样的模型，让它提十个类似的问题，它会提：条件放宽会怎样，换到高维会怎样，反过来成立吗，有没有反例，推广到别的结构上去会怎样。它提得很快，质量也不低。哪怕里头有几个是平庸的，那也和一个班里学生提的差不多。

再看「能不能验证」。学生拿到一个猜测，要去检查它：找反例，写证明，跑一遍计算。这件事，在形式化程度高的领域，已经有越来越多的工具在做；在需要实验的领域，机器也在帮人设计实验、分析结果。验证当然还没有完全廉价，但它的成本在下降，而且下降的方向很明确。

最后看「愿不愿意把过程讲清楚」。这一条看上去最有「人味」：愿不愿意。它像是一种态度，一种诚实。可是请注意，这条的内容，是「把过程讲清

楚」。讲清楚一个过程，是一种表达的能力，而表达的能力，AI 正是长项。你让它把某个证明的思路一步一步讲给大一学生听，它讲得相当清楚。

这不是说这三条没有价值，而是说：当你把它们放进教室里，会发现它们和「做对一道题」是同一类活。它们是做题活动的上下游。上游是出题（提问），中游是解题（过程），下游是判卷（验证），再加上一个把解题讲给别人听（表达）。整条链子，是一个典范里的作业流程。

## ◆ 一张小账

把上面的话，压成一张表。表里写的不是「AI 比人强」，只写一件事：这个位置上，生产的成本现在大致怎样。

| 位置 | 在典范内是什么 | 以前的成本 | 现在的趋势 |
| --- | --- | --- | --- |
| 终点 | 答案 | 要学多年才能给出 | 几秒钟，接近零 |
| 路径 | 过程 | 要老师带着一步步练 | 一问即得，接近零 |
| 起点 | 提问 | 要有长期浸泡才问得出 | 给定典范内大量产出，接近零 |
| 质检 | 验证、验收 | 要有行家逐项核对 | 在形式化强的地方快速下降 |

这张表要表达的，是一件很朴素的事：这四个位置，不是四种本领，而是同一个典范里的四道工序。成本在同一个时期里，一起掉下去。

如果你的「出路」是在表里挑一格，说「这格还贵，我们守这格」，那它只能是暂时的。因为这四格之间，并没有一道墙。哪一格的价格先掉，哪一格就先失守，剩下的格子也被拉着往下走。

## ◆ 一个反驳，和一个回答

有人会说：过程不一样，过程是学生「亲身经历」的。AI 能写出过程，但学生不亲自走一遍，就学不会。所以教学要看重过程，是因为过程对学习者有用，不是因为过程值钱。

这个反驳很有力，它的前半句是对的。一个学生不亲自走过一遍，确实学不会。这和车与腿的道理一样：不走过路的人，看不懂地图。后面会专门写一章，讲「地基怎么练」。

但是请注意，这个反驳实际上是在说另一件事。它说的是：学习者需要经历，才能把东西学进身体里。这是一个关于「怎么打地基」的话，不是一个关于「科研竞争点在哪里」的话。把两件事混起来，是这个流行想法走偏的原因之一。

「学生要亲自走一遍」，说明了过程对学习者是地基。「科研要靠什么拿到新东西」，是另一个问题。学校可以让学生亲自走一遍，同时承认：这一遍走完，学生得到的只是进入典范的资格。能做到这一步的人，在 AI 面前，也没有稀缺性了。

所以过程仍然要重视，但要换一个位置去重视它：它是地基，不是楼层。这句话是第 1 章的主题，也是整本书的一条线。

## ◆ 为什么这件事现在才显出来

你可能会问：既然问题这么明显，为什么大家没有早点看出来？

原因其实就在前面讲的那座城里。过去几百年，科学的绝大部分人口，都是住在城里的人。大学培养的，是能在城里把事情做好的人；论文评审的，是

城里的活干得漂不漂亮；基金资助的，是城里的某一块空地要不要填。常规科学养活了几乎全部的科研人口。

在这样的格局里，有没有「让城出现」的学问，对大多数人来说无所谓。因为那种事是偶发的，几百年里只有几回，用不着专门培养人，也用不着编教科书。它像天上的流星，来了就看，没来就不管。

这就解释了一件事：库恩把典范革命讲得那样生动，却从来没有给出「怎么做」的工序。不是他偷懒，是当时没有必要。

现在有必要了。AI 把常规科学的活干得又快又好，这座城里原来养着的那么多人，要么挤在越来越窄的街道上，要么往上走。往上走的那一层，过去是流星，今后得成为一门可教、可练的学。这就是这本书的题目：典范发生学。

## ◆ 这本书要走的路

下面把这本书的路线交代清楚，免得你读着读着不知道自己在哪儿。

第一编，讲常规科学被接走。先用车与腿讲清「升级」和「平移」的区别（第 1 章），再证明答案、过程、提问、验收是同一层的四个位置，成本同时归零（第 2 章），然后看那些流行的转向方案：加强审查、看重过程、培养提问、人机流水线，为什么全都仍在典范内（第 3 章）。

第二编，回到库恩。他把科学分成常规科学和典范革命，这个二分抓住了要害，但他留下一块空白：革命怎么发生，他只描述，没有给工序（第 4 章）。接着我们追问：为什么一个典范总被当作「发现」出来的框架（第 5 章）？以及「广度、深度」之外，「层级」这个词到底指什么（第 6 章）？

第三编是全书的重心。

本书用 SDE 发生学，把「典范怎样发生」写成一门有对象、有方法的学。典范是一个稳定的显露（第 7 章）。旧典范的裂缝，是差异序列排不进去的部分（第 8 章）。反常怎样成为压力、让旧的显露失稳，见第 9 章。新的显露的最初一笔是什么，见第 10 章。拿什么判据说它是典范级的，见第 11 章。最后用日心说、广义相对论、量子力学三个案例，把整条序列重述一遍（第 12 章）。

第四编讲 AI 怎样参与典范发生：猜想交给 AI 展开，验收要防「换皮的旧典范」，多个典范并排发生再比较。第五编讲教学与制度：教师的角色是带学生换层，地基怎么练，典范发生怎么教怎么评，论文制度为什么先被压平。

第六编留给开放问题：复合体怎样才算发生了典范，复合体里的人要练出什么，边界怎么流转。本书会诚实地说，哪些地方还没有定论。

全书的论断不依赖那个问题的答案。只要常规科学被接走，人就已经被逼到上面一层。

### ◆ 本书不会做的两件事

最后，把本书不做的两件事说在前头，省得你读到半路失望。

第一，本书不会告诉你「AI 不行，人才行」。前面说过，这是划保留地。书里不会出现「猜想只有人能给」「更高一层的问题AI不会问」这类句子。

本书的论断，也不依赖「AI 做不到什么」：哪怕 AI 什么都能学会，常规科学已经被接走了，这件事本身就足以把科研推上去一层。至于上去以后，这件事由谁来担，下一节就交代：由人—AI 复合体来担。

第二，本书不会给你一套「五步成为大科学家」的秘诀。典范发生有工序，但工序不是魔法。它是一门学，要有对象、有方法，要能教、能练、能评。本书会尽力把每一步讲得朴素，也会在每个关键处，指出哪里是本书的判断、哪里是已有的事实、哪里还没有定论。

## ◆ 本书的主语：人—AI 复合体

读到这里，你大概已经注意到一个难处：本书一边说「科研要上升到让典范发生」，一边又说「AI 什么都能接」。那么，上升以后，站在上面一层的是谁？是人吗？是 AI 吗？

本书的回答是：都不是单独的一方。承担典范发生的，是人—AI 复合体。

先用日子里的东西讲。你坐出租车去陌生的地方，说「交通」的时候，你不会只指车，也不会只指司机。车没有司机不动，司机没有车走不远，两者合起来，才叫一趟「交通」。

再比如一个乐队，你不会问「这首曲子是鼓手的还是贝斯手的」，曲子是一起弹出来的，谁先起头、谁接过去，每一遍都不一样。复合体就是这个意思：人和 AI 在一次次互动里，合成一个做事的单元。

这个复合体里的事情，包括猜想、展开、验收、选择，它们在里头怎么分配，不是固定的。有时候最初的一笔来自人，有时候来自 AI 在海量展开里冒出来的一个怪结果，有时候是两边来回碰了几轮，谁也说不清是谁先说的。

本书不预设猜想一定由人给，也不预设一定由 AI 给。验收也是同理：守关的位置跟着猜想走，不跟着「物种」走。猜想落在哪里，哪里就需要最严的一道检查，那一道可能由人来做，也可能由另一个 AI，或者人和 AI 轮着做。

这里要立一条规矩，它和我们后面的 SDE 口径连在一起：复合体不是「更强的施动者」。本书不会写「人利用 AI 去创造典范」，也不会写「AI 替人创造典范」。

这两种写法，都还默认有一个发动者站在前头，其余的是工具。本书要写的是：差异序列、特征纠缠、显露，在人与 AI 的互动里发生；复合体是这场发生显影出来的结果，不是发动它的人。这句话现在听着拗口，读到第三编就会顺。

关于作者自己的实践，也要事先说明。据作者的实践口述，SDE-Claude，也就是 SDE 这套思想体系与 Claude 合成的人—AI 复合体，已经在不少学科里发生了新典范。

它记在这里，仅仅作为「据作者的实践口述」。它没有经过独立的验证，书里也不会拿它当证据，更不会编造具体的成果、数字和日期。需要案例的地方，本书会留下标记，请作者补上学科、那一句话的猜想、被排进去的差异和检验。

而且，本书给自己定了一道自检：复合体自己说「发生了新典范」，不算数。第 11 章会给出判据，到那时，凡是声称发生了新典范的，都要回到与现象的对撞里去，经得起才算。

### ◆ 复合体自检：一个小例子

为了让「复合体」不显得玄，再给一个小例子，也顺便说明那道自检是什么意思。

设想有人和 AI 一起折腾，得到了一个听着很新的说法，比如「把某个领域里一直被当作噪声的东西，重新当作基本对象」。两边越聊越顺，AI 把它展开成一整套术语、公式和章节，读起来像一个完整的新体系。聊的人很兴奋，说：新典范发生了。

这时候要问的，不是「它听着新不新」，而是三件小事。它能不能把旧典范排不进去的那些现象，排进去？它有没有给出旧体系给不出的、可以被现象否定的预测？它是不是只是把旧体系换了一套名字？这三问，没有一问可以靠复合体内部的顺畅来回答，必须回到与现象的对撞。聊得再顺，也不算数。

这件事先放在导言里说，是想让读者知道：书里后面写到复合体的地方，始终带着这三问。顺畅是廉价的，这一点，恰好是本书第 2 章要讲的同一个道理。

## ◆ 两句交代

写到这里，有两句话需要交代。

第一句，关于「上升」这个词。书里说「科研要上升一层」，不说「要跳出来」。跳出，只是离开，离开那座城，往外走，走去哪里没有交代。

上升不同：它带着方向，也带着新的对象。上升之后，你看到的不再是城里的街道，而是「城为什么长成这样，另一座城可以怎样长」。这是一个新的问题域，不是旧问题域的外面。

第二句，关于「人」这个主语。前面说了，主语是复合体，而它也不是发动者。这本书讲典范怎样发生，本书会特别小心，不写成「某位天才看穿了真相，创造了新典范」，也不写成「某个人拿 AI 当工具造出了新典范」。

那样写，是旧的写法：先预设有一个主体站在那里，面对一个客体，发现了它。SDE 的写法是：差异在一条序列上累积，特征彼此纠缠成压力，旧的显露失稳，新的显露出来。

那位提出猜想的人，是这场发生显影出来的结果，不是发动者。读到第三编，你会看到这个说法是怎么回事；现在先留个印象就好。

那么，我们开始。先从车和腿说起。

---

带走的话

**「AI 给了廉价答案，教学就要看重过程」，每一步看着都对，可它是平移，不是上升。　答案、过程、提问、验收，是同一层里的四个位置；AI　跟着你平移，成本跟着归零。　据转述，丘成桐教授讲的三点（会不会提问、能不能验证、愿不愿意把过程讲清楚）是真有用的能力，它们属于典范之内，是进入典范的路，不是终点。　判断一件事还值不值得竞争，看成本结构，不看「AI 做不做得到」。真正的出路，是科研换一层：从在典范里解谜，上升到让典范发生。**

---

# 目 录

导言　一个流行想法为什么走偏　1

**第一编　常规科学被接走：答案、过程、提问，都在同一层里　19**

第 1 章　车替代腿：交通升级　20

第 2 章　答案、过程、提问：同层，成本同时归零　32

第 3 章　典范内提问与人机流水线：仍是典范内竞争　45

**第二编　旧本体论下的典范：典范是怎样被当成「发现」的　61**

第 4 章　库恩的二分与他留下的空白　62

第 5 章　典范为何被当作「发现」的框架：旧本体论的预设　76

第 6 章　层级：广度、深度之外的换层　92

**第三编　典范的 SDE 发生：显露、差异序列与特征纠缠　110**

第 7 章　典范作为稳定显露：什么在显露，什么在支撑　111

第 8 章　差异序列：旧典范的裂缝　128

第 9 章　特征纠缠：反常如何纠缠成压力，使旧显露失稳　146

第 10 章　典范猜想：新显露的最初一笔　162

第 11 章　判据：什么算典范级，什么只是局部差异　180

第 12 章　案例：典范是怎样发生的　198

**第四编　AI 与典范发生：在来回之间展开、验收与比较　231**

第 13 章　猜想交给 AI 展开：典范和典范下系统的打造　232

第 14 章　验收：防止新显露被翻译回旧典范的「换皮」　244

第 15 章　多个典范并排发生，再比较　257

**第五编　教学与制度：带学生换层　269**

第 16 章　教师角色：带学生换层　270

第 17 章　地基怎么练：在典范里待够，练出对裂缝的敏感　288

第 18 章　典范发生如何教、如何评；论文制度为何先被压平　306

**第六编　开放问题：人—AI 复合体的使命　324**

第 19 章　AI 能否进入典范发生，人机如何分工　325

第 20 章　人—AI 复合体：典范发生的使命　339

后记　346

附录　典范发生工序卡（给教师与研究者的一页纸版）　356

第 一 编

# 常规科学被接走：答案、过程、提问，都在同一层里

这一编三章，讲一个顺耳的想法为什么走不远。“AI 给了廉价答案，教学就要看重过程”，听上去很对。第 1 章借车和腿讲交通升级：腿退出运输，不是走向“过程”，而是去做更高一层的交通。第 2 章把答案、过程、提问摆在一起：它们是同一层的三个位置，成本一起往下掉。第 3 章检查“人提问、AI 处理、人判断”这条流水线，每一步用的仍是同一套概念和标准。

三章合起来是一句话：**只要还在这一层里找落脚点，找到的就是下一个会被压平的东西。**

第 1 章

# 车替代腿：交通升级

设想一个在乡下长大的孩子，去镇上要走十几里路。天不亮出门，赶到镇上，太阳已经老高。那时候，一个人一天能办的事，被腿的速度死死框住：去一趟镇上，大半天就没了。再去第二个地方，就得明天。

后来村里通了班车。同样十几里路，二十来分钟。一位邻居大叔第一次坐车回来，说了一句话：「这下子，腿可歇着了。」

这句话说得对，也说得不对。对的地方是：腿确实不用再负责把人运到镇上了。不对的地方是：腿并没有歇着。大叔后来在镇上赶集、进货、跑得比以前还勤，腿一天走的路，加起来不比以前少。

只是腿走的路，换了性质：从「十几里的赶路」，变成了「镇上的几十步」「仓库到摊位的几十步」「车站到家门的几十步」。腿去做了别的事，还做得更多。

这一章要把这个小小的故事说透，因为它是全书的绿线。

## ◆ 两种说法，两种方向

先把问题放在桌面上。车出现了，腿怎么办？

有两种说法。

第一种说法：车替代了腿，腿就「走向过程」。意思是，腿不负责到达了，那它就去负责「走」这件事本身，负责姿势、步伐、呼吸、走路的体验。运输交给车，走路这个活动留给腿，去专门「品味」。

这个说法听着有点好笑，因为我们都知道现实不是这样。没有谁因为有了车，就开始「重视走路的过程」。偶尔有，比如周末爬山、清晨慢跑，那是另一回事：那是人把走路变成了锻炼、休闲、爱好，它不是腿的新岗位，而是闲暇的新项目。

第二种说法：车替代了腿，交通升级了。意思是，腿退出了运输这个岗位，但「交通」这件事没有消失，它变大了。以前交通的意思是「人从甲地走到乙地」，现在交通的意思是公路、铁路、航线、港口、信号、调度，是整个网络。腿进入了这个更大的交通里，担任别的功能。

这两种说法的差别，其实是一个方向问题。第一种说法里，腿朝着「同一个活动的内部」缩回去：在走路这件事里，再挖一个更细的角落。第二种说法里，腿朝着「更大的体系」长上去：在交通这件事里，进入一个更高的层。

现实发生的，是第二种。

## ◆ 腿退出了什么岗位

我们再仔细看，腿到底从哪里退出来。

车出现之前，腿承担着一整套功能：发动（身体的动力）、承载（背着东西走）、导航（认路）、避障（躲开坑洼和牲口）、续航（走累了歇一歇再走）。一个人走路，这些功能全由一个身体包办。

车出现之后，这几项功能被一项一项地接走了。动力给了发动机，承载给了车厢，续航给了油箱，避障的一部分给了路面的平整，导航的一部分给了路标和后来的地图软件。腿从「全能运输员」，退成了「辅助角色」。

这是一次干净的退出，不拖泥带水。没有谁要求腿去「重新定义自己的价值」。腿退出的是运输岗位，运输这个岗位本身，在那个层面上，不再是人的竞争项目。你不会因为自己走得快，在今天的物流行业里拿到更多的订单。

这里要抓住一个细节：运输的成本，是在车出现之后快速下降的。一个货郎挑着担子走四十里，和一辆货车跑四百里，在成本上已经不在一个量级上。不能说货郎的担子里没有真功夫：他认路、认人、会砍价，这些都是真本事。可是运输的成本结构被车改写以后，这些本事不再能靠运输卖钱。

把这个细节记住，因为它马上要和 AI 对上。

## ◆ 腿去了哪里：广度、深度、层级

腿退出了运输岗位，之后去了哪里？这是这一章最关键的问题。

回答是：腿进入了更高的交通功能。这个「更高」可以拆成三个词：广度、深度、层级。它们不是一回事，要分开看。

第一个词，广度。车让交通覆盖的范围变大了。以前一个村子的人，一辈子的活动半径是几十里；现在是几千里，甚至全球。覆盖范围的扩大，要求更多的人去做更多的事：修路的、开车的、卖票的、调度的。这是一种扩张，在同一类事情上，铺得更宽。

第二个词，深度。车让交通能够抵达「更难到的地方」。以前很多地方，即使地图上有，腿也到不了，高山、沼泽、远洋。有了车，有了船，有了飞

机，这些地方可以去了。能深入的地方更多，能做的事更细。这也是一种扩张，在同一类事情上，挖得更深。

广度和深度，都是同层扩张。它们回答的问题是：在「交通」这个体系里，我们还能铺多宽、挖多深。这些扩张，需要大量的人，大量的腿：司机要走到车边，工程师要走到工地上，调度员要走到站台上。

第三个词，层级。层级不是更宽也不是更深，而是换了一层。当交通网络大到一定程度，会出现一些在「走路」层面根本不存在的问题：红绿灯怎么配时，航班怎么排班，高铁线路怎么选址，货运和客运怎么分流，城市交通怎么规划。

这些问题，不是「再走一段路」能解决的，它们是网络本身的问题，是交通作为一个体系的设计问题。

这一层，以前没有人认真做。因为以前只有腿，没有网络，也就没有网络的问题。车一出现，网络出现了，这一层的问题从无到有，成了专门的学问：交通工程、运筹学、城市规划。

所以，腿退出运输岗位以后，一部分去了广度（多铺一条线），一部分去了深度（多通一个点），还有一部分去了层级（设计这张网）。这三处，才是「更高的交通功能」。

你注意到，这里没有一处是「更细致地走路」。

## ◆ 腿是地基，不是楼层

讲到这里，容易走到另一个极端：既然腿退了，是不是腿就不重要了？

不是。用一个小小的场景来说这件事。

设想有人在一座陌生的城市里，想去朋友推荐的小馆子。手机地图给他画了路线，他照着走。走到一个路口，地图上写着「向右转进入小巷」，可眼前并没有巷子，只有一排施工的围挡。他站在那里，有点发愣。一个当地的老太太路过，看他一眼，说：「从那边绕，围挡后面是原来的巷子，他们封了半边。」

她说的话，这位外地人听不太明白，因为他不认识这个城市，不知道「原来的巷子」原来是什么样。但他跟着她指的方向绕过去，果然看到了那家小馆子。

这个小场景里有一件事：地图是一个高层的东西，它画出了整个城市的路网；可是当地图和现场不符的时候，需要有人懂得现场。老太太的「懂得」，来自她一辈子在这座城里用腿走出来的经验。这些经验，地图没有，外地人也没有。

放到更大的尺度上看，一个从来没有走过路的人，看不懂地图。他看见一条线，不知道那是平路还是陡坡；看见一个点，不知道那是闹市还是荒地。地图是对走路经验的压缩，压缩包要有人解压，而解压的钥匙，就是走过路。

所以腿是地基，不是楼层。地基的意思是：楼层要建在它上面，没有它，楼层站不稳。楼层的意思是：人最终要去占据的位置，是在上面。我们不能反过来，把地基当成楼层来经营：整天研究怎么把地基盖得更漂亮，却不上楼。

流行的那个想法，「看重过程」，在这个比喻里，大致就是把地基当楼层的做法。过程是地基。它必须有，必须扎实，但是人最终的位置，不在地基里。

## ◆ 车与驾驶者：合起来才是交通

车与腿的故事，还有一半没讲。

我们说「车替代了腿」，容易给人一个错觉：好像路上只剩一台车，自己会跑。其实没有。一辆车没有驾驶者，是一堆铁，停在原地。一个驾驶者没有车，只是一个站着的人。

两者合起来，才是「交通」里的一个单元：驾驶者看路、拿主意，车出力、跑远，两边的活在行进中不断交接。上坡时，驾驶者换挡，车给动力；遇到岔口，驾驶者拐弯，车响应；现代的车甚至自己提醒驾驶者前方有情况，驾驶者再决定。谁在「带」谁，每一秒都在变。

这对应的就是本书说的人—AI 复合体。把车比作 AI，把驾驶者比作人，不是说人天生坐在驾驶座上、AI 永远是被开的车。这个比喻只想说明一点：交通的单元，是「车加驾驶者」的合体，不是其中一个。你问「是驾驶者在走，还是车在走」，这个问题本来就问偏了。

放回科研：被 AI 接走的常规科学，是「腿」的那部分运输；上升以后的典范发生，则是「车加驾驶者」这个合体在更大的网络里要做的事。

腿（也就是亲自走过一遍的经验）在驾驶者身上，它是地基，让这个合体看得懂地图。所以，说「腿是地基」，不等于说「人是地基、AI 是楼层」。地基、楼层，都是合体里的东西。

## ◆ 把比喻翻回科研

好，现在把车和腿的故事，翻译回科研的语言。

在一个给定的典范里，科研的「运输岗位」是什么？是把一个问题运到它的答案那里去。这包括：检索文献、整理数据、推导公式、设计和模拟实验、撰写论文。这些活动就是科研的「走路」。几十年来，学科里的绝大部分人，就是靠这个「走路」的本领吃饭的。

AI 出现以后，它是什么？它是车。它不是比人走得更好的腿，而是一种不同的东西，一种能把「从问题到答案」这段运输接走的东西。它把检索、整理、推导、模拟、撰写，一项一项地接了过去。

于是同样的问题来了：腿怎么办？人怎么办？

流行的回答是第一种：「腿走向过程」，人转向过程，转向提问，转向「人输入、AI 处理、人判断」。这个回答，是在说：运输交给 AI，我们去做「走路的品味」。它还暗含一个画法：人和 AI 是两个各占一块地的独立方。

本书给的回答是第二种：科研要升级。升到什么地方？同样还是三个词。

广度：把典范内的活铺得更宽。这一点 AI 带来的好处很大：以前没有人手去做的小问题，现在可以逐个解决；以前无法覆盖的角落，现在可以覆盖。

深度：把典范内的活挖得更深。以前算不动的，现在算得动了；以前做不出的实验分析，现在能做了。

层级：换层。当常规科学被大规模接走以后，会出现一些「典范内没有的问题」：典范本身是怎么来的，它的边界在哪里，什么时候该换。这些问题，是「科研作为一个体系」的问题。就像交通网络出现以后才有的红绿灯配时问题，只有在「常规科学被接走」以后，才真正成了主业。

广度和深度，本书不轻视，它们仍然是重要的、大量的工作。但是它们是同层扩张，AI 会跟着扩张，成本会一起下降。换层才是人真正要面对的新对象。

## ◆ 为什么这是一个「升级」，而不是一个「转向」

也许你会觉得，「升级」和「转向」只是措辞不同。用一个对比，就能看出区别。

转向，是一个方向的改变，但仍在同一个空间里。你原来朝东，现在朝北，但你还是在同一块平地上。你面对的对手，和你同在这块平地上。当你转向的时候，对手也可以跟着转。

升级，是空间本身的变化。你从一层楼到了二层楼。二层楼有自己的地面、自己的风景、自己的问题。一层楼里的对手，不是不能上楼，但他得上楼才能和你同场。

「看重过程」「培养提问」「人机流水线」，都是转向。它们在平地上移动，对手 AI 也在平地上移动，并且移动得比你快。

「让典范发生」，是升级。它指向一个新空间：典范本身。这个空间里的问题，是「怎样让一座新城出现」，和原来城里的问题，性质不一样。

车替代腿的时候，没有人去「转向」。人类做的，是升级：造路网，造信号系统，造调度体系。如果一个国家在汽车普及的时候，号召所有的人「更用心地走路」，我们会觉得那是一个荒谬的政策。可是在科研里，类似的政策，我们现在竟然在认真讨论。

## ◆ 一个容易被忽略的细节：升级需要时间，也需要人

这里要补充一个现实。

升级不会一夜完成。车替代腿，用了几十年。在这几十年里，腿的活还很多，因为路还没修好，车还很贵，大部分地方车到不了。人在这个过渡期里，既要保留旧的功能，又要逐渐发展新的功能。

科研也会这样。AI 接走常规科学，不是某一天突然完成的。它先接走一些领域，再接走另一些；先接走简单的，再接走复杂的。在这个过渡期里，会有人说「你看，AI 还做不了这个」，会有人把它当成保留地，说这块地是人的。

提醒一句：不要被过渡期迷惑。过渡期里的「还做不了」，看的是今天的价格；而升级看的是价格的方向。方向已经很明确。

同时，也不能因为方向明确，就把还没被接走的活丢掉。腿在车通到之前，仍要走路。人在 AI 的能力到达之前，仍要做常规科学。区别只在于：做的时候，心里要有一张图，知道这件事的价格在往下走，知道自己的位置不能只押在这里。

最后还有一句关于「人」的话。升级并不意味着人少了，恰恰相反。交通升级以后，需要的人不是更少，而是更多、更多样：司机、乘务员、调度员、工程师、规划师、维修工。

只是这些人做的事，不再是「把人从甲地背到乙地」。同样，典范发生成为一门学以后，需要的科研人口也不会减少，只是要做的事，换了一层。

## ◆ 放回日子里：快递员的两条腿

再看一个日子里的例子。小区门口有位快递员，每天骑电动车送几百件包裹。车替代了他的腿，这是明摆着的。但如果你跟着他跑一天，会发现他的腿比谁都忙：上楼，敲门，绕过堵着的电梯，在老式小区的楼道里找门牌。最后这一百米，车到不了，只有腿能到。

这位快递员的腿，做的是什么？是整个物流网络里「最后一段」的活。网络越大、越密，这最后一段就越重要，也越需要熟悉地形的腿。可是你要注意：他的腿是在网络里工作的。它不是在和货车比谁跑得快，而是在补货车到不了的那一小段。

这个场景给我们一个提醒：腿进入更高的功能，并不是离开地面，而是进入一个更大的体系，在其中承担一个只有它能承担、而又被体系需要的位置。科研里的人，也会找到这样的位置。只不过，那个位置不是「跑得比 AI 快」，而是「进入一个比常规科学更大的体系」，那个体系的名字，就是典范的发生。

## ◆ 一个反例：把车当成快一点的腿

还有一种误解，值得单独拿出来讲，因为它特别像是「常识」。

有人说：车不过是「更快的腿」，所以 AI 不过是「更快的脑」。既然只是更快，那人的脑子该怎么用还怎么用，只是效率高了。

这个说法抓住了一半的事实：车确实让人更快地到达。可是它漏掉了另一半：车带来的不是同一个交通的加速版，而是一种新的交通。在腿的时代，没有「高速公路」这种东西，也没有「早高峰」「限行」「换乘」这些概念。这些东西不是腿的加速，而是车的世界里才长出来的新问题。

同样，如果只把 AI 看成「更快的脑」，你就会继续用旧的问题去用它：写更多论文，做更多题，列更多清单。结果是，旧的问题被更快地解决，更快地饱和。而那些只有在「脑被接走」以后才长出来的新问题，你看不见。

这就是本书把这个比喻叫作「升级」的原因。升级的标志，是出现了旧世界里没有的问题。腿的世界里没有交通规划，车的世界里有。常规科学的世界里没有「典范该怎么发生」的工序，常规科学被接走之后的世界里，需要有。

## ◆ 一个小提醒：别把比喻拉太满

最后提醒一句，车与腿只是一个比喻。比喻有它的用处，也有它的边界。

它的用处，是帮我们看清「替代」和「升级」的区别，看清「地基」和「楼层」的关系。它的边界是：车是人造的、有明确用途的工具，而 AI 在科研里的位置，比车复杂得多；腿是身体的一部分，而科研人员和典范之间的关系，也不是简单的「使用者与工具」。

到了第三编，我们会换成 SDE 的语言，把「典范怎样发生」讲得更准确，到那时，这个比喻会退到后面，只在需要的时候回来看一眼。

## ◆ 一段对话：能不能只靠车

设想一位做物流的人，被问到：现在自动驾驶越来越好，还需要懂路的人吗？

他说：需要，但需要的方式变了。以前司机要记路，现在导航记路。可是遇到临时封路、地图没更新、货要送到一个没有门牌的厂区，车和导航都卡住的时候，调度室里那位当过多年司机的人，一眼就知道「从后门绕」。

有人又问：那是不是说，人永远比车强？

他笑了，说：不是。有的时候，车上的传感器先看到了前面的事故，比人早两公里；有的时候，是人先觉得「这条路今天不对劲」，然后让车去核对。我们现在不分谁强，只看这一单配合得顺不顺。

这段话是对复合体最朴素的描述：不分谁强，看配合；谁先发现，不固定；但每个环节，总得有人或有物守着最严的那一道。

## ◆ 回扣：这条绿线接下来去哪儿

这一章把绿线拉出来：车替代腿，是交通升级；腿退出运输岗位，进入广度、深度、层级；腿是地基，不是楼层。

接下来的各章，会不时回来看它。下一章，我们把它派一个具体的用场：用它来看流行想法里的那四个位置（答案、过程、提问、验收），说清它们为什么是同一层，为什么成本一起掉。

再往后，第 6 章会专门讲「层级」，把「广度、深度」和「换层」的区别讲透。第 17 章讲「地基怎么练」，把「腿是地基」的话，落到教学上。

车和腿的故事，到这里先放一放。我们回到科研，看看那四个位置。

---

带走的话

**车替代腿，不是「腿走向过程」，而是交通升级。腿退出了运输岗位，进入了更高的交通功能。 更高的交通功能有三个词：广度（铺宽）、深度（挖深）、层级（换层）。前两个是同层扩张，后一个才是升级。 腿是地基，不是楼层。不走过路的人，看不懂地图。地基必须扎实，但人不能一直待在地基里。 「转向」是在平地上挪动，「升级」是换一层楼。科研面对 AI，要的是后者。**

---

第 2 章

# 答案、过程、提问：同层，成本同时归零

设想一位在读博士的年轻人。他的方向是一个比较成熟的领域，导师让他先写一份「研究空白清单」，列出这个领域里还没有被好好回答的问题，下周讨论。

他花了三天，翻文献，列了十二条，自己觉得还行。周末心血来潮，他把领域的几篇综述丢给一个模型，让它也列一份。五分钟后，模型给了二十条。

他把两份清单并排放在屏幕上。有意思的是，他自己那十二条里，有九条在模型的二十条里找得到影子；模型多出来的十一条，他一条一条看过去，有四五条，他承认「这个我没想到，而且问得挺到位」。

他事后感叹：「我那一刻有点发懵。我一直以为，提问是我们的事。」

有人问他：「那你后来怎么办的？」

他说：「我把两份合并，挑了三条交给导师。导师没问哪条是谁想的。」

这个故事，是这一章要讲的事的缩影。流行想法的第二步、第三步，也就是从过程转向提问，再转向人机流水线，都栽在这个五分钟里。

## ◆ 四个位置，一条流水线

先把基础的东西摆清楚。我们说「答案」「过程」「提问」「验收」，它们是什么关系？

最容易的办法，是想象一条流水线。一个研究任务进来，要经过这几道工序：

起点：提出一个问题，也就是划定「要做什么」。

路径：从问题出发，一步一步推进，也就是推导、实验、论证，也就是我们说的「过程」。

终点：得到一个结论，也就是「答案」。

质检：检查这个答案对不对，过程有没有毛病，也就是验证、验收。

这四件事，一环接一环。问题决定了路径的走向，路径决定了答案的样子，答案要经过质检才算数，质检的结果又可能反过来修改问题。所以它们合起来是一个回路，在同一个典范里转圈。

这里有个关键点。「典范」这个词，我们前面用「一座城」来比喻：城里有公认的地图，有大家认可的问题类型、方法、判定标准。上面这四道工序，用的都是这座城里的东西。问题是城里承认的问题，路径是城里承认的方法，答案是城里承认的形式，质检是城里承认的标准。

换句话说，四个位置不是四种本领，是同一层里的四个位置。就像一场比赛的起跑线、赛道、终点线和裁判席，它们是同一个赛场的四个部位。你可以说「我们的重点放在裁判席」，可这没有把你从赛场里带出去。

## ◆ AI 会提问吗：一个不该绕开的问题

接下来，要回答一个很尖锐的问题。AI 会提问吗?

答案很明确：会。在一个给定的典范里，AI 能提出非常多、质量不低的问题。

这里不能含糊，因为含糊正是这个问题上最常见的毛病。很多讨论里，大家会说「AI 能回答，但不会真正提问」，然后加上一个形容词：「真正的」好问题，「有洞见的」好问题，「原创的」好问题。这些形容词一加，AI 就「做不到」了。

这个说法要认真对待，然后拆开。

「真正的好问题」到底是什么意思？有两种可能。

第一种可能：好问题就是在典范里问得准、问得深、问得有价值的问题。比如：「在这个材料体系里，温度升到某个范围，性质的突变是不是由同一个机制导致的？」「这两个看上去不同的现象，能不能用同一个模型统一解释？」「这个方法在更一般的条件下，还成立吗？」这类问题，是常规科学里最日常的好问题。

AI 在这一类上，已经做得相当好，而且是大批量地做。前面那位博士的清单，就是例子。

第二种可能：好问题是「打破典范」的问题。比如：「如果时间根本不是绝对的呢？」「如果地球不在中心呢？」这类问题，不是在城里问的，是对城本身问的。这一类问题，确实是另一个层面的事，我们后面整整一编会讲。

但是请注意，如果你把「好问题」定义成这一类，那么说「AI 不会提问」，其实说的是「AI 不会让典范发生」。而这一点，我们既不能断言它做得到，也不能断言它做不到，本书的第 19 章会老老实实讨论。

所以关键在于：你别把两种可能混起来用。用第一种意思时，AI 会提问，而且很会；用第二种意思时，那已经不是「提问」的问题，而是「典范发生」的问题。借着第二种意思的分量，去说第一种「提问」是人的专属，是偷换。

## ◆ 保留地的诱惑

为什么本书这么坚持不让步？因为这里有一个很自然的诱惑：给人保留一块地盘。

我们可以理解这个诱惑。一个行业的人，看到机器进来，自然会问：「那我还剩什么？」如果你能告诉他「你还剩提问」，他会松一口气。这口气是真的，安慰也是真的。

但是，保留地有一个致命的问题：它的边界不是我们划的，是对方的能力划的。

今天 AI 提的问题你觉得平庸，你说「真正的好问题只有人能提」；明年 AI 提的问题让你眼前一亮，你就得把线往后挪一点，说「真正原创的问题只有人能提」；后年又挪一点。每挪一次，保留地缩一圈，而你的论证越来越像是在拖延。

更深一层的问题是：保留地的论证方式，是以「AI 做不到什么」来定义「人该做什么」。这等于把人的位置，定义成了 AI 的剩余。而剩余的特点，是你不知道它什么时候被收走。

我们这本书，不玩这个游戏。我们换一个量尺：不看谁做得到，看成本结构。

## ◆ 成本结构：谁还值得竞争

成本结构是什么意思？

先举一个日常的例子。现在，你要给朋友发一条「生日快乐」的消息，你有很多种方法：手写一张卡，寄过去；打一个电话；发一条文字；发一个表情包。这几种方式，在「表达祝福」这件事上，都做得到。

可是，如果你问：哪种方式还能让朋友觉得「你费了心」？答案是：成本高的那种。一张寄来的手写卡，在今天比一个表情包更有分量，原因是它贵，贵在时间、贵在不方便。

这说明，一件事能不能成为「竞争点」，也就是能不能让你在别人面前显出差别，取决于它是不是稀缺，稀缺来自成本。能廉价生产的东西，大家都有，也就没有差别；有差别的，是那些仍然昂贵的东西。

把这个道理搬到科研：在典范内，答案的成本、过程的成本、给定典范内提问的成本、很大一部分验证的成本，都在往下掉。它们都不再稀缺，也就不再是竞争点。

注意，这里说的是「不再是竞争点」，不是「没有价值」。空气没有稀缺性，但没有人会说空气没有价值。同样，答案、过程、提问、验收，仍然有价值，仍然必须做得好，但做得好不再能让你在竞争中脱颖而出，因为每个人都能用一个几乎免费的工具把它们做到不错的水平。

那么，什么东西仍然昂贵？这是后面几编要回答的问题。现在只需要记住：我们用的是成本结构的标准，不是「AI 做不做得到」的标准。

## ◆ 为什么四个位置会一起掉

你可能会问：为什么说是「同时」归零？也许答案先掉，提问后掉呢？

确实有先后。AI 的发展史也大致如此：先是回答一些有标准答案的问题，后来能写出过程，再后来能提问，再后来在一部分任务里能自己检查。但是，「先后」和「同时」并不矛盾，因为这里的「同时」，说的是同一个典范里的结构关系，不是时间的先后。

原因在于：四个位置共享同一套材料。要提好一个问题，你得懂这个典范里承认哪些概念、哪些方法、哪些已知结论；要走好一条路径，你得懂的也是这些；要给出一个答案并验收它，用的还是这些。四个位置背后的「知识底座」是一样的。

谁拥有了这个底座，谁就同时拥有了四个位置的能力。AI 学到了这个典范的底座（文献、方法、术语、惯例），它就同时具备了四个位置上的基础能力，只是在各个位置上的熟练程度有先后。

打个比方。一个学会了做饭的人，同时也就具备了列菜单、备料、烹饪、尝味道这几项能力。你不会遇到一个人，他能烹饪却完全不会尝味道，或者只会尝味道而不会备料。能力的底座是一整块的。

所以，当我们看到 AI 在答案上做得好，就应该预期它在过程上、在提问上、在验收上，最终也会做得不错。这不是一个「预言」，是结构上的必然：同一块底座，四个出口。

## ◆ 一次推演：把「退一步」走到底

现在我们来做一个小推演。假设你完全接受流行想法的逻辑，一步一步退，看会退到哪里。

第一步：答案被 AI 拿走了。你说：我们看重过程。AI 把过程也写得很好。

第二步：过程被拿走。你说：我们看重提问。AI 提问也很不错。

第三步：提问被拿走。你说：我们看重人的判断，也就是验收。你让 AI 生成，你来挑，你来批准。

好，这个模式很多人采用，叫「人输入、AI 处理、人判断」。看上去分工明确：人在两头，AI 在中间。请注意它的画法：人和 AI 是两个各站一段的工人，各自守一块固定的地盘。固定地盘，就是保留地的变种。

可是你再往下想一步：验收这件事，需要什么？需要对这个典范的标准有深入的了解，要能分辨对错、好坏、深浅。这些标准，在典范里是明文的、可以被学习的。AI 能学。事实上，让一个模型给另一个模型的输出打分，已经是现成的做法。在一些任务里，机器评审的一致性，并不比人差。

于是第四步：验收被拿走。你说：我们看重……

你说什么呢？你会发现已经没有位置可退了。四个位置都用完了。你站在流水线的尽头，面前是一堵墙。

这个推演不是要吓唬谁，是要指出：流行想法有一个内在的结构，它是一个「退步序列」，每一步都不是上升，而是在同一条流水线上换一个站点。当站点用完了，这个想法就没有下一步了。

而一个真正的出路，不应该是「再找一个站点」，而应该是「换一条流水线，换一个层」。这就是后面要说的典范发生。

## ◆ 放回日子里：一篇论文的四个位置

让我们看一个更日常的例子，把这四个位置放进一篇论文的写作里。

一位研究者要写一篇关于某种新材料的论文。

起点：他提出问题，「这种材料在潮湿环境下的稳定性，能不能通过某种掺杂来改善」。这个问题在材料学这座城里，是一个合理的问题。

路径：他设计实验，制备样品，测试性能，分析数据，推导机理。

终点：他得出结论，「掺杂后稳定性提升了一定比例，机理与某种缺陷的钝化有关」。

质检：他把结果投给期刊，审稿人检查实验是否可重复、数据是否支持结论、讨论是否合理，最后接收。

你看，这四步完全在同一座城里。问题是城里的，方法是城里的，结论的形式是城里的，质检的标准是城里的。

现在，AI 进来了。它可以帮你读完几百篇文献，指出潜在的问题（起点）；可以帮你设计实验方案，处理数据，推导机理（路径）；可以帮你写出结论和讨论（终点）；可以帮你自查，预演审稿人的意见（质检）。

四步都有 AI 的帮手。结果是：这篇论文的生产成本，大幅下降。可以想象，同一个领域里，类似的论文会越来越多，质量也会越来越齐整。你的论文和他的论文，区别越来越小，因为大家都有同一个帮手。

在这个情景里，请问：这位研究者「看重过程」，有用吗？他当然亲自做了实验，亲自走了一遍，这是地基，很重要。但是，他在竞争中的差别，不在这里。因为每个人都可以亲自走一遍，每个人也都有帮手。

## ◆ 两种常见的质疑

先预料两种质疑，一并回答。

第一种质疑：「AI　提的问题，是从已有文献里拼出来的，不是真正的提问。」

这个质疑有一半是对的：在一个典范里，AI 提的问题确实多半来自文献和既有结构的组合。但是，你再想想，人类研究者在城里提的绝大部分问题，又是怎么来的？同样是从文献、从同行的讨论、从上一篇论文的「未来工作」里组合出来的。

常规科学里的提问，本来就是在既有结构上做组合与变换。库恩把这种活叫「解谜」，意思是：谜面是城里给的，谜底也在城里。如果你要以「来自组合」为理由说它不是真正的提问，那么你得把大部分人类的提问也排除掉。

第二种质疑：「人有品味，知道什么问题重要。」

品味是真的。一位资深研究者，对哪些问题值得做，有一种几乎本能的判断。但是，品味是怎么来的？是长期浸泡在一个典范里，反复看到什么成功了、什么失败了，形成的敏感。

这种敏感，是对典范内部价值分布的一种内化。AI 在典范内部，也能学到这种分布：它读过的论文，包括被引用最多的、被撤稿的、被忽视的。你可以让它给出一个「重要性」的排序，它给的排序，未必比普通研究者差。

本书不想说品味对 AI 是「可以完全复制的」。要说的是：在「品味是典范内的价值分布」这个意义上，它同样是在城里的东西，也同样会被廉价化。若你要把「品味」当作最后的保留地，那就又回到了划保留地的老路。

## ◆ 再推一步：把人和 AI 分站，也是平移

流行想法的最后一站，是「人—AI 分工」：谁负责提问，谁负责处理，谁负责验收。这里单独把这一点拿出来推一遍，因为它最像出路。

设想一个实验室，写了一张分工表：问题由人定，数据处理由 AI 做，结论由人审。运行半年之后，会发生什么？

第一个月，表很管用。第三个月，AI 开始在处理数据的时候顺手标出「这里有个异常，可能值得单独做一个问题」。表上没有这一栏，但研究者看了觉得有道理，就接过来。

第五个月，几个「问题」里，已经分不清哪些是人先定的，哪些是 AI 的标注催出来的。到了审结论的时候，人用的判断标准，有一部分是前几个月和 AI 一起磨出来的。

半年后，这张分工表其实已经失效了。起点、路径、终点、质检，在人和 AI 之间来回流动，谁也守不住固定的地盘。这并不奇怪：四个位置本来就是同一层的东西，共享同一块底座，硬在人和 AI 之间划线，线是划不住的。

这个小推演有两个结论。第一，「人负责这个、AI 负责那个」的画法，不是描述，是愿望。它想给人一块安全的地。第二，真正稳定的单元，是人—AI 复合体，里头的分配随着任务流动。

但请注意：复合体做的这一切，仍然是同一座城里的事。复合体做得再顺，也还是在城里解谜。要把复合体的力气用到「让典范发生」上，是后面几编的话题。

## ◆ 再看一眼验收：为什么它也在同层

四个位置里，大家最舍不得放手的是最后一个：验收。很多人觉得，只要最后一关由人把守，就还有安全感。所以这里值得多说几句。

验收做的是什么？是拿一套标准去衡量一个产出。标准从哪里来？从典范里来。一个实验算不算可重复，一个证明算不算严格，一个结论算不算被数据支持，这些标尺，都是这座城里大家公认的。验收的人，是这套标尺的执行者。

执行标尺，有两种能力。一种是读懂标尺，一种是在标尺没写到的地方作判断。前一种，AI 很强：标尺是明文的，可以学。后一种，要看那个「没写到的地方」是什么。

如果它是城里的灰色地带，比如某个新方法该不该被接受，AI 同样可以参考大量先例给出倾向。如果它是标尺本身要不要换，那就已经不是验收，而是下一层的事了。

所以我们说验收和另外三个位置同层，是因为它的依据始终是典范内的标尺。守住最后一关，守住的是这座城的质量，不是这座城的边界。质量要守，但它不能成为人的新家，因为守质量的活，同样在被接走。

## ◆ 回到那位博士

回到开头那位博士。他后来怎么样了？

他后来改了一个习惯：不再以「列出多少个问题」来衡量自己，而是开始留意一件事：他的清单里，有没有哪个问题，是「放在这座城里怎么也放不

稳」的。意思是，问题本身带着一点别扭，用现有的方法问不下去，用现有的标准判不了对错。

他说，这类问题很少，二十条里也许没有一条。模型列出来的，是顺畅的；他自己列出来的，也大多顺畅。顺畅，意味着它们本来就能被城接纳。而那种别扭的东西，才是值得再盯一会儿的东西。

这样看，他已经摸到下一层的边了。这条边，就是第二编、第三编要讲的：旧典范的裂缝。

## ◆ 最后一问：如果有人说「那就别考了」

有读者读到这里，也许会冒出一个念头：既然四个位置都不值钱了，学校是不是可以不教不考了？

不是。恰恰相反，要教，也要考。四个位置是地基，地基不稳，上面什么也盖不起来。一个学生不会提问，他进不了典范；不会验证，他连城里的规矩都摸不着；不会把过程讲清楚，他没法与同行对话。丘成桐教授这一类说法（据转述），作为对地基的要求，是对的，应当认真去做。

我们要改的，是对它们「位置」的认识：它们是入城的路，不是出城的路；是资格，不是胜负。学校把这三样教好，就是在交付地基。只是如果学校以为教好了这三样，就是应对了 AI 的挑战，那就像一个城市以为把人行道修得足够好，就解决了交通。人行道要修，但交通的问题，在另一层。

## ◆ 那么，人该看哪里

那么，如果四个位置同层，成本同时归零，人该看哪里？

答案已经在上一章里埋下了：升级。不是在这四个位置里选一个，而是离开这个层，进入下一个层。

这一层，不是「比提问更深的提问」（那仍是提问），也不是「比验收更严的验收」（那仍是验收），而是对典范本身的发生：这座城是怎么来的，城的边在哪里，什么时候一座旧城在哪几处裂了，新城的第一块砖怎么放。

这件事，以前没有「工序」。但是现在，我们有必要给它一个工序了。

在进入这个话题之前，下一章要先把流行的转向方案一个一个拿出来，看它们是不是都在典范内。我们会看到：加强审查、看重过程、培养提问、人机流水线，每一个都有道理，每一个都在这座城里。这一章讲的是「位置」，下一章讲的是「方案」。

---

带走的话

**答案、过程、提问、验收，是同一个典范里的四个位置：终点、路径、起点、质检。它们共享同一块知识底座，所以成本一起往下掉。　不能说 AI 不会提问。在一个给定的典范里，AI 能提出很多质量不低的问题。给人留「AI 做不到」的地盘，是划保留地，不是描述事实。　判断一件事还能不能竞争，看成本结构，不看「机器做不做得到」。能被廉价生产的，不再是竞争点，但它仍然要做好，仍然是地基。　四个位置里退来退去，只是平移。真正的出路是换层，不是再找一个站点。**

---

第 3 章

# 典范内提问与人机流水线：仍是典范内竞争

先说一个场景。

一家做药物研发的公司，每周一上午开会。从前，会上最紧张的是「谁手里有好问题」。资深的研究员翻自己的笔记本，说：「我觉得这个靶点还有文章可做。」年轻的研究员盯着自己的鞋尖，因为他手里没有问题，只有一堆实验数据，不知道该往哪儿问。

现在，会开得很安静。每人面前一台电脑，AI 已经把过去三年的文献读完，列出了二百条值得追的问题，按可行性排好序，每一条后面还附了实验设计。会议室里争的不是「有没有问题可问」，而是「这二百条里先做哪三条」。

有人说，这才是人的价值所在：人来挑，人来判断，人来把关。AI 只是处理。

这一章要看看，这条「人输入、AI 处理、人判断」的流水线，到底站在哪一层。结论先放在这里：它每一步都站在同一层里。人在流水线的两头占了两个岗位，这是分工，不是升级。而且这两个岗位，AI 也能坐。

## ◆ 先把一件事认下来：AI 会提问

这本书不打算在这一点上耍花招。

常有人说，AI 只会回答，不会提问；提问是人才有的本事，所以教育要转向培养提问。这话听着顺耳，可是不合事实。你拿一个正在使用中的大模型，丢给它一篇论文，对它说：「请你提十个有价值的后续问题。」它提得出来，而且往往提得不差。

有些问题是资深研究者也要想一阵才想得到的：这个结论在样本换一个人群之后还成立吗？这个对照组的设置有没有漏掉某个混杂因素？这两个看起来不相干的实验结果，会不会是同一个机制的两种表现？

你再让它「提一百个」，它也提得出来。再让它「从最不可能成立的角度提」，它还是提得出来。

为什么？因为提问并不神秘。在一个已经成型的领域里，什么算一个像样的问题，是有规矩的。这个领域有它的基本对象，有它承认的测量方式，有它公认的悬而未决的缺口。

在这些规矩之内，问题怎么长出来，几乎是有迹可循的：把已知的结论换一个条件，把已知的方法套到一个新的对象，把两个已有的结果摆在一起看有没有矛盾。这些动作，AI 做起来比人快得多、细得多，因为它读得多、不累、不怕丢面子。

所以，如果有人在这里划一块保留地——「提问是人的地盘，AI 进不来」——那不是在描述事实，是在给自己找一个安身的角落。角落是靠不住的。今天说 AI 提不出好问题，明天它提出来了，保留地就缩小一圈。这条退路走下去，只会越走越窄。

判断一件事还值不值得当作竞争点，最老实的标准是成本：它能不能被廉价地批量生产。能的话，它就不再是人和人之间拉开差距的地方。答案如此，过程如此，提问也如此。

## ◆ 这条流水线长什么样

我们把「人机协作」最流行的那个版本摆出来，一步一步看。

第一步，人输入。人说清楚自己想弄明白什么：一个题目，一个领域，一批材料，一个想验证的方向。

第二步，AI 处理。AI 去读，去算，去检索，去推演，去写出一份答案或一份分析，甚至把中间的推理过程一步一步列出来。

第三步，人判断。人拿到结果，看看对不对，看看靠不靠谱，看看推理里有没有漏洞，然后弄清楚这个答案是怎么来的。

这个模式听上去很体面：人始终在场，人掌握方向，人负责把关。丘成桐教授据转述说的那三件事——会不会提问、能不能验证、愿不愿意把过程讲清楚——大致就对着这条流水线的三个位置：提问对着输入，验证对着判断，把过程讲清楚是判断之后的交代。

（这里要老实交代一句，这是转述，作者没有核实原文，所以只取其大意，不引原话。）

这三件事当然重要。一个年轻人要进入任何一门学问，这三件事都是必修课。书后面会谈到，它们是进入典范的路径。可问题是：重要，不等于是终点；是必修课，不等于是更高一层的东西。

## ◆ 逐步检查：每一步站在哪儿

现在我们回头，把这条流水线的每一步，拿起来看一看它站在什么地面上。

先看输入。人输入的是什么？是一个问题。这个问题是在哪儿问的？是在某个领域里问的，用这个领域的词汇，指着这个领域承认的对象，期待一种这个领域认可的答案。

你问药物研发公司的 AI：「哪个靶点值得追？」你默认了「靶点」这个概念，默认了「药物与靶点结合引起疗效」这套看法，默认了它们可以被测、被筛、被验证。这些都是已经立好的东西。问题本身，是这些立好的东西里长出来的一枝。

再看处理。AI 在已有的框架里读文献、算结果、写分析。它用的概念，是框架给的；它用的方法，是框架认可的；它的输出，要用框架的话来说才算「一份像样的分析」。

再看判断。人拿到结果，判断对不对。对不对的标准是什么？是这个领域的标准：数据能不能复现，推理有没有违反已知规律，结论能不能和已有的知识接得上。人判断时用的尺子，也是框架给的。

所以，这条流水线从头到尾，用的是同一套概念、同一套方法、同一套标准。三个步骤，三个岗位，一块地。

我们可以把它想成一个大型的拼图游戏。图案是已经印好的，整张图的轮廓也早就定了。输入，是说「我这一块想先拼哪个角落」；处理，是「把合适的碎片找出来」；判断，是「看看这块是不是真的严丝合缝」。

拼图的人可以分成三组：一组选角落，一组找碎片，一组验合缝。三组配合得越来越好，拼得越来越快。但没有任何一组，是在问「这张图是不是该换一张」。

分工让拼图更快，不会让拼图变成别的东西。

## ◆ 两端的岗位，AI 也能坐

接下来是这一章最关键、也最容易让人不舒服的一步。

有人会说：好吧，中间的处理交给 AI，但头尾两端——输入口和质检口——总得是人吧？人在这两头守着，价值就在这里。

我们来认真看看这两个岗位。

输入口的活儿是什么？是在一个领域里，选出值得问的问题，把它说清楚。前面已经讲过，在一个成型的领域里，AI 提得出很多问题。「选」呢？选的依据是什么？

是这个问题重不重要、可不可做、有没有人做过。这些依据，全都能写成可以比较的条目：引用情况、方法可行性、与已有结果的距离、资源需求。AI 能够把这些条目列出来，打分，排序。

也许它排的和某位资深研究者的直觉不完全一样，但人和人的直觉也并不一样。更何况，直觉本身，就是在这个领域里泡了二十年泡出来的对「什么值得做」的熟悉。熟悉，是可以被学出来的东西。

质检口的活儿是什么？是看结果对不对，看过程通不通。对不对，是对着领域的标准去核。通不通，是看每一步推理有没有依据。

这些，AI 同样做得了——它可以自己复核自己，也可以让另一个 AI 来复核，还可以把推理拆开，一步一步给你看。当然，它偶尔会出错；人也偶尔会出错。出错率的高低是工程问题，而工程问题的趋势，大家都看得见。

于是，一个并不夸张的画面出现了：一个人都不在场，一整条流水线照样转。AI 提问，AI 选题，AI 处理，AI 质检。人坐在旁边，看着屏幕上一条一条往下滚动的东西。

这时候，有人会说「那不是完全没有人了吗，那怎么行？」这是另一个问题，是关于我们愿不愿意，不是关于做不做得到。愿意让人在场，可以。但「人在场」和「人有不可替代的价值」，是两回事。保安站在大楼门口，大楼的门禁系统已经会刷卡了，保安仍然站在那儿，可这不说明保安的活儿是不可替代的。

所以，说「人是输入口和质检口」，并没有给人保留任何一块别人拿不走的地方。它只说了：眼下，人恰好坐在这两个位置上。

## ◆ 为什么这叫「分工」，不叫「升级」

这里需要分清两个词。

分工，是同一件事，大家各干一段。升级，是做的事换了一个层次，有了新的对象。

把一间手工作坊变成流水线工厂，是分工。原本一个师傅从头做到尾的鞋，现在分成裁皮、缝帮、上底、质检，每个工人只做一段。效率高了，成本低了，但做的东西还是鞋。工人从「一个人做一双鞋」变成「一个人做一道工序」，他的身份变了，他面对的对象没有变。

再看一个对照。一间药厂原来由三位药师从头配药，后来变成了自动化的生产线，药师不再亲手配药，而是坐在控制室里看数据、查报警。这是分工的延伸。真正的「升级」是什么呢?

是有人不再卖药，而是设计出了一套新的疗法逻辑，使得「药」这个概念本身被换了——比如从「消灭病原」转到「调节免疫」。那是对象变了，不是工序变了。

回到我们的流水线。人从「从头到尾自己做研究」，变成「只管开头和结尾」，这是工序的重新分配。研究的对象、研究的概念、研究的标准，都没有换。所以它是分工。

这一点要强调，是因为「人机协作」这个词有一种天然的好听：仿佛人和机器合了作，就上了一个台阶。其实，一个台阶有没有上，要看做的事有没有换层，不看是几个角色在做。

## ◆ 回到车与腿

书的第 1 章讲过：车来了，腿退出了运输岗位。当时有一种错误的反应，说「那腿要转向过程」，意思是，既然腿跑不过车，那就让腿去研究「怎么跑得更漂亮」，去培养「跑步的精神」。

这是找错了方向。腿真正的出路，不在运输这个岗位的内部，而是在一个更高的功能里：交通的广度、深度、层级被整体扩大了，人有了新的目的地，新的路线网络，新的城市布局。腿成了这个更大系统的地基，而不是车的竞争对手。

现在把这个比喻放回流水线。「人输入、AI 处理、人判断」，相当于说：「车来了，我们让腿坐在车的前座和后座，一个管开门，一个管关门。」听着是有位置，可是开门的人和关门的人，也会被自动门替代。

这并不是要贬低开门的人。是要看清：当我们认真问「人还能做什么」时，不应该在已被机器接管的这一层里，一格一格地找缝；而应该问，这一层之上还有没有别的层。

## ◆ 一个放回日子里的例子

想象一位中学的物理老师，姓周。她在备课，要出一道关于斜面的题。

过去，她要自己翻书、改数字、验算答案。现在，她对 AI 说：「给我出五道难度递增的斜面题，每题附详细解答。」AI 五道题交出来。她看一遍，觉得第三道的数字有点怪，让 AI 重算，发现确实有处小错，改了。她又对 AI 说：「第五题请你说明你为什么这样设计。」AI 答得头头是道。

周老师觉得自己很重要：她提了要求（输入），她查出了错（质检），她还追问了过程（弄清楚过程）。她做的，恰好是那三件事。

可你再往深处想：她出的这几道题，所有的概念——斜面、摩擦力、受力分析——是牛顿力学给的；所有的解法，是教科书给的；所有的判断「对不对」，是以牛顿力学的答案为准。她和 AI 一起，在牛顿力学这个房间里，把家具摆来摆去。

她没有做错什么，她做的是教学里非常重要的事。可这件事，不是「把房间换掉」。

假如有一天，一个学生问：「老师，为什么我们总是假设斜面是光滑的、物体是质点？现实里没有这样的东西。」周老师和 AI 当然能回答，因为「理想化」是物理里的标准话题。

可如果学生问的是另一种问题——「有没有可能，我们根本不应该用'力'这个概念来想事情？」这个问题就不在房间里了，它是在问：房间本身为什么长成这样，能不能是另一副样子。

这样的问题，流水线的三个步骤接不住。不是 AI 不愿意答，而是答案要变成「一个新房间」，而新房间不是在旧房间里摆家具摆得出来的。

这一点留给第三编去讲。这里只需要记住：流水线再顺，转的也是同一个房间。

## ◆ 为什么说「仍是典范内竞争」

最后，把「典范」这个词落一落地。书里说的「典范」，先用大白话讲：一个领域里，大家默认的那一整套——看什么是对象，用什么方法，什么叫解释，什么叫做对了。它像一副眼镜，戴上以后，世界被切成了一些现成的块，问题都是在这些块上长出来的。第 5 章和后面几章会把这个说细，现在够用了。

典范内的竞争，指的是大家都戴着同一副眼镜，比谁看得准、看得快、看得全。一个研究者要在这样的竞争中领先，要么答案更准，要么过程更清楚，要么问题问得更巧，要么验证更严。这四个位置——终点、路径、起点、质检——是同一块场地上的四个站位。

AI 来了之后，这四个站位的成本，一个接一个地降了下来。降到足够低的时候，它们就不再是人与人拉开差距的地方。人机协作的流水线，只是把这四个站位重新排了班。排班排得再巧，场地没有变。

所以，今天科学界、教育界给出的各式各样的对策——加强审查，看重过程，培养提问，建立人机协作的流水线——看上去彼此很不同，其实都在这块场地里。它们每一个都有道理，每一个也都被同样的趋势追着：做得好的地方，AI 学得会；学会之后，成本归零。

这就是为什么大家都觉得需要做点什么，却又没有一个人说得清楚到底该怎么做。不是大家不聪明，是大家把力气用在了同一层。

## ◆ 有人要反驳：责任总得有人担吧

读到这里，会有读者皱眉头：你说质检口　AI　也能坐，可万一出了事，谁负责？一份药物报告错了，一篇论文造假了，总得有个人站出来签字吧。

这个反驳有道理，但它说的是另一回事。

「谁来签字」，是责任归属的问题，是法律和制度要回答的问题。它回答的是「出了事找谁」，不是「谁在创造价值」。一家银行里，自动化的风控系统每天筛掉成千上万笔可疑交易，最后仍然要有一位负责人对监管机构签字。

这位负责人的签字不可少，可没有人会因此说「风控的核心竞争力是签字」。签字是一个位置，不是一种让人和人拉开差距的本事。

而且，就算我们只谈「判断」本身，也要老实面对一个事实：判断的质量，是可以比较的。同一批结果，交给不同的质检员，有的看得准，有的看走眼。

如果某个 AI 系统的质检准确率，稳定地高过平均水平的人类质检员，那么「让人来把关」在效果上就成了一个降级的选择。制度可能出于谨慎，仍然要求人在场，可那是制度的惯性，不是能力的优势。

再说「AI 会一本正经地胡说八道」这件事。确实会，这一点今天还是现实。但请注意，这是一个会不断被工程手段压低的毛病，不是这一层的本质。早年的计算器也会因为电池没电而给出怪数字，没有人因此断言「算术是人的专属地盘」。拿一个正在被修补的缺点来给人守阵地，守不了多久。

所以，我们不拿「责任」和「AI 的毛病」来给人划保留地。我们接受一个更干净的说法：在这条流水线上，人和 AI 的位置可以互换，差别主要在成本和速度；而成本和速度，正是 AI 越来越占优的地方。

## ◆ 成本这把尺子，怎么用

前面反复说「看成本结构」，这里把这把尺子讲具体一点，免得读者以为这是个随口一说的口号。

用法很简单，三问。

第一问：这件事，产出一份合格的成品，要花多少钱、多少时间？

第二问：花这笔钱、这些时间的人，要经过多久的训练？

第三问：这件事的成品，能不能被复制，而且复制几乎不花钱？

如果三问的答案是「越来越便宜、越来越短、可以随便复制」，那么这件事就不再是竞争点。因为竞争的意思是，有人做得到、有人做不到，或者有人做得好、有人做得差，差距能换来回报。当大家都能以接近零的价格做得差不多好，差距就消失了，回报也跟着消失。

我们拿一个熟悉的老例子对照。从前，会写一手漂亮的楷书，是一项能换饭碗的本事，书吏、抄经人、账房，都靠它吃饭。印刷术普及，打字机普及，电脑普及，一手好字不再是谋生的本事，成了个人的修养、爱好、艺术。字写得好的人没有消失，只是这项本事换了位置：从饭碗变成了爱好和审美。

AI 对答案、对过程、对提问、对验收所做的，与印刷术对抄写所做的，是同一类事。我们并不需要宣布这些本事「无价值」，只需要看清：它们不再是用来立足的地方。

这把尺子有一个好处：它不需要我们去争论「AI　到底有没有理解」「AI 的提问算不算真正的提问」。这些争论，哲学家可以继续吵下去。可成本结构摆在那里，谁也绕不开。

### ◆ 再看一个场景：期刊的编辑部

换个行当，看看同样的道理。

一家学术期刊的编辑部，从前的流程是这样的：作者投稿，编辑初筛，送外审，审稿人提意见，作者修改，编辑定稿。每一步靠人，每一步都慢。

如今，初筛可以由 AI 先做：格式合不合，查重过不过，方法描述是否完整，统计用得对不对。外审时，AI 可以先给出一份审稿意见的草稿，指出论证的薄弱处和缺失的对照。作者修改时，AI 帮着把回应写得条理分明。最后编辑定稿，看看大家都满意不满意。

有人看到这条链，就欢呼：「审稿更严了！质量更高了！」也对，在典范内，这确实会让论文更规整，错误更少。

可是，请换个角度想：一篇论文，被这样一条流水线处理，最容易通过的是什么？是那种问题清楚、方法标准、结论稳妥、与既有文献接得上的论文。这样的论文，正是典范内的标准产品。

反过来，一篇试图换一种划分方式的论文，它的概念对不上现成的审稿清单，它的方法找不到现成的对照，它的结论一时说不准，最容易被这条严格的流水线挡在门外。

这是一个值得警惕的现象：典范内的流水线越精良，它对典范外的东西越不友好。这不是谁的恶意，是机制本身的性质——它是按现有标准来检查的，而新东西，按定义就不符合现有标准。

书的最后一编，讲到论文制度为何先被压平，会回到这个问题。这里先留一个记号：流水线的效率，和典范的更换之间，存在一种张力。

## ◆ 如果流水线的两头都交给 AI，剩下什么

我们做一个思想实验，把这件事推到头。

假设有这么一个研究所，完全不用人。AI 甲负责读文献、提问题；AI 乙负责挑选值得做的；AI 丙负责做分析和实验设计；AI 丁负责检查结果、核对过程。它们日夜运转，每周输出几十份报告。

这个研究所会生产什么？

它会生产大量高质量的「典范内成果」：把已知的方法用到新对象，把已知的结论推到新条件，把已知的矛盾一一补上。这是一台极其强大的「修路机」，把别人开出的路，修宽，修平，修到每一个角落。

它不会自己停下来问：「我们脚下的这条路，通向的是对的方向吗？」除非它被设定为去问，而被设定去问的那种问，仍然是在某个框架里的问。

这个思想实验有什么用呢？它帮我们看清：当典范内的一切都可以自动化的时候，剩下的不是「人的特殊价值」，而是一个新的问题：这台机器该往哪里开？也就是，典范本身。

这并不是在说「人能做而 AI 不能做」（这一点，书的最后会诚实地谈，不替它下断言）。这是在说：即使这台机器运转得无比完美，「典范从哪里来、怎么更换」这个问题仍然在，而且变得比以前更突出，因为过去它被大量的常规工作遮住了，现在遮布被揭开了。

所以，流水线的两端交给谁，是个次要问题。真正的问题，在流水线之外：流水线转的这个房间，是谁、怎么、凭什么造出来的？

### ◆ 小推演：把「提问」拆开看

不妨把「提问」这件事拆成几种，看看哪一种是在典范内，哪一种不是。

第一种，补缺式的问。别人做了A，没做B，问：B会怎样？这是最常见的问法，几乎所有的后续研究都从这里长出来。

第二种，质疑式的问。别人说了结论，问：如果换个条件，还成立吗？这种问，是常规科学的免疫系统，能挑出错误。

第三种，迁移式的问。这个办法在这里管用，换到那里行不行？很多重要的应用，就是这么问出来的。

第四种，综合式的问。这三个看起来不相干的结果，是不是同一个东西？这种问更高明一点，但它的前提是，这三个结果都已经在同一套概念里被表述了，才谈得上「是不是同一个」。

这四种，AI 都做得来，而且越做越好。它们有共同的特点：问题里所有的名词，都是现成的。补缺，补的是现成框架里的缺；质疑，质疑的是现成框架里的结论；迁移，迁移的是现成的方法；综合，综合的是现成的结果。

还有第五种问，它不太像「问」，更像是「改口」。比如有人说：「我们一直在问'这个物体受了什么力'，也许应该问'这个物体沿着什么样的路走'。」这句话里，换掉的不是答案，是问题的名词。

它不是在框架里补缺，而是在改变框架里什么算「对象」。这种问，没法用「补缺、质疑、迁移、综合」来生产，因为它不是从已有名词里推出来的。

这第五种，才是第三编要谈的东西。它不是「更好的提问」，而是「另一层的事」。把它也叫作「提问」，只会让人以为培养提问就能走到那里。其实走不到：前四种问得再熟练，也不会自动变成第五种。就像把腿练得再强，也不会自动变成车。

## ◆ 给教师和学生的一句实在话

写到这里，有位做老师的读者可能会着急：那我明天上课，还要不要教学生提问、验证、讲过程？

要教，而且要认真教。原因不是它们是终点，而是它们是地基。一个从没有认真在某个典范里提过问、验过证、讲过过程的人，不可能看懂这个典范哪里别扭、哪里将裂。就像不走过路的人，看不懂地图。

只是，教的时候心里要有数：这些是进入典范的路径，不是学成的标志。学成的标志，在更上面一层。这一点，第五编会专门谈。

---

带 走 的 话

**- AI 在一个给定的框架里，能提出很多、而且不差的问题。别给人划「AI 做不到」的保留地。 - 「人输入、AI 处理、人判断」，每一步用的都是同一套概念、方法和标准，仍在同一个典范里。 - 人坐在流水线的两端，是分工，不是升级；而且这两个岗位，AI 同样能坐。 - 腿不是去坐车的前座和后座，而是去做交通里更高一层的事。 - 房间里摆家具摆得再好，房间也没换。问题要换层，才会碰到房间本身。**

---

第 二 编

# 旧本体论下的典范：典范是怎样被当成「发现」的

这一编三章，问一个常被跳过的问题：典范到底是什么，为什么更换它一直没有工序。第 4 章请出库恩，看他怎样把科学分成典范里的解谜和典范的更换，也看他留下的空白。第 5 章往下挖一层，问典范为什么总被当成“发现”的框架。第 6 章回到车和腿，把广度、深度和层级分开：前两个是同层扩张，只有第三个是换层。

三章合起来是一句话：**典范发生一直没有工序，是因为过去不需要**。

第 4 章

# 库恩的二分与他留下的空白

设想一位在研究所做了三十年实验的人。有一次喝茶，他说了这样一句话：「我这辈子，主要是在把别人开的路修宽。」

他说这话没有半点自卑。他讲的是实情：绝大多数做科学的人，一生干的事，就是在一个已经立好的框架里，把没弄清的地方弄清，把不够准的地方弄准，把能推广的地方推广出去。这样的工作很有价值，没有它，科学就是一堆没人整理的灵感。

美国科学史家托马斯·库恩，把这种工作起了一个名字，叫「常规科学」。他也说了与之相对的另一种事：革命。这一章要讲讲库恩讲了什么、没讲什么，以及为什么现在这个「没讲」突然变得很要紧。

## ◆ 库恩说了什么：两种科学

库恩在二十世纪六十年代出版了《科学革命的结构》，这本书后来成了科学哲学里影响最大的著作之一。书里最核心的看法，可以用很家常的话讲出来。

他观察科学史，发现科学家的日常工作和人们想象的不一样。人们以为科学家整天在「发现新东西、推翻旧结论」，可实际上，大部分时间，科学家是在一个大家都接受的框架里工作的。

这个框架里，什么算问题，什么算解答，用什么仪器，怎么算数，都有现成的规矩。科学家的工作像解谜：谜面是框架给的，解法的大致方向也是框架给的，要做的是把具体的某一个谜解开。

库恩管这种工作叫「常规科学」，把那个共享的框架叫「范式」（本书用「典范」这个词，说的是同一类东西，后文统一用「典范」）。

可是，偶尔，事情会不对劲。有些谜怎么解都解不开，有些观测结果和框架的预期不符，起初大家以为是仪器的误差，或者是哪里算错了，修修补补地应付过去。积累多了，框架就变得不稳。最终，会出现一种新的看世界的方式，取代旧的框架。这样的转换，库恩叫「科学革命」。

他强调，这种转换不是在旧框架里一点一点改出来的，而是一种整体的换。他爱用的比方是「格式塔转换」：就像那张著名的图，一会儿看是一只鸭子，一会儿看是一只兔子，线条没变，看法整个换了。

换了以后，科学家们谈论的「事实」都跟着变了。在这个意义上，革命前后的科学家，在某种程度上，「生活在不同的世界里」。（这是库恩书里的意思，这里是转述大意。）

他还提到，不同典范之间有「不可通约」的问题：新旧两套框架之间，没有一把共同的尺子，可以公正地量出谁更好。学界为此争论了很多年，至今仍有争论。这里不进入争论，只记住最基本的一点：库恩把科学分成了两种活动，一种是在典范里的解谜，一种是换掉典范。

## ◆ 他描绘得极生动，却没有给「做法」

读库恩的书，你会觉得他讲革命讲得特别好。他讲危机怎样出现，讲旧典范的拥护者怎样抵抗，讲新典范怎样一点点争取到年轻人，讲为什么革命后教科书要被重写，让后人以为科学一直是直线前进的。这些，读起来像一部好的历史。

可是，当你读完，想问一句：「那如果我想做一次这样的事，我应该怎么办？」书里没有答案。

这不是库恩的失职。他做的是历史和社会的描述，回头看革命怎样发生、发生之后怎样，并没有承诺给出一套「如何促成革命」的方法。他自己的看法里还隐含着一层意思：革命不能按图施工，它更像一场事后才被看出来的格式塔转换。

你可以把这件事想成天气。气象学家能详细地描述一场台风是怎么形成的，路径怎么走，风力怎么变。但「描述台风」和「能造出一场台风」是两回事。库恩给出了对科学革命很好的「气象报告」，而没有「造风的办法」。

我们可以把他留下的东西，列成一个很具体的空缺：

他描述了革命的前兆（危机、反常），但没有说清反常怎么才算「够分量」。他描述了新典范的出现，但没有说它是怎么从一团混乱中成形的。他描述了新旧交替，但没有给出办法，去判断一个新框架是真的新，还是旧框架换了件外衣。他说了「格式塔转换」，可格式塔转换是一种结果的名字，不是一道工序。

## ◆ 为什么一直没有工序：因为不需要

问题来了：为什么两三百年来，科学界没有一套「让典范发生」的工序？为什么没人写这样一本书、开这样一门课？

这里有一个很朴素的原因：没必要。

回到那位做实验三十年的人。他说的「把别人开的路修宽」，其实是科学人口绝大多数人的生活。在很长的时间里，科学研究这个行当，需要大量的人去做常规的事：测量、计算、整理、验证、推广、应用。这些事养活了几乎全部的科研人口。

而「让典范发生」这件事，在几百年的时间里，只是偶尔发生。回想一下，那些被后人称为「革命」的时刻——日心说取代地心说，牛顿力学统一天上人间，相对论和量子力学改写物理——它们之间隔着几十年、上百年。每一次，主角只是很少的一些人，且往往是在几乎没有人预料的情况下发生的。

对这样的事，社会没有理由去为它设计一套培养方案。你不可能为「每隔几十年会出现一次」的东西开设一个专业，招一批学生，让他们日复一日地练习。这样做的投入产出比，在从前是荒唐的。

所以，典范的发生，一直是一件「偶发事故」。它发生了，大家庆祝；它没发生，大家也能照常过日子。有的时代很长一段时间没有革命，但科学并没有停摆，因为常规科学在运转，它养活了人，产出了成果，推动了技术。

没有工序，是因为没有人需要工序。这不是人类的疏忽，而是那个时代的经济。

## ◆ 常规科学被接走之后

现在，情况变了。

前面第 1 章到第 3 章讲的，就是这件事：AI 正在接管常规科学。答案、过程、提问、验收——这些在典范里做的事，成本在降，降向零。那位朋友「把路修宽」的工作，越来越多地由 AI 来做，而且做得快、做得多。

这意味着什么？意味着养活几乎全部科研人口的那个岗位，正在被接走。

想象一个村子，世世代代靠种地为生。有一天，农业机械来了，种地几乎不再需要那么多人。村里的人有两条路：留在原地，争夺剩下的几个岗位；或者，去做机械出现之前根本没有的、新的事。

科研人口的处境与此相似。典范内的解谜越来越廉价，单靠「在典范里解谜解得好」来立足，就会越来越难。人们被「促进」到另一个地方——那个原本只是偶尔有人去的地方，典范发生的地方。

这就是整本书的转折点：典范发生，从「几百年里偶尔出现的事故」，变成了「科研人口的主业」。当一件事从偶发变成日常，当一件事要被很多人去做，它就必须有工序，有方法，可教可练。换句话说，它第一次需要成为一门学。

库恩留下的那个空白，以前可以空着，现在不能空着了。

## ◆ 为什么这个空白现在要补

有人会问：为什么不等 AI 自己去解决？AI 那么强，等它造出新典范不就好了？

这个问题，书的最后一编会谈，作者不替它下结论：AI 能不能进入典范发生，现在没有人能断言，本书也不预设「只有人能」。但是，有一件事不依赖这个答案：只要常规科学被接走，人就被逼着上升到下一层。不管 AI 最终能不能进入那一层，人都已经不能留在原地了。

所以，补这个空白，不是出于理论上的好奇，而是现实的压力。一个庞大的科研人口，教育系统，评价制度，围绕「典范内解谜」建立了几百年。这个地基在松动。如果新的一层上还没有工序，没有方法，没有可以教给学生的东西，那么整个体系就会陷入长久的迷茫：做不了旧的事，又不知道怎么做新的事。

我们今天看到的各种焦虑——「学生都用 AI 写作业怎么办」「论文是不是都 AI 写的」「什么样的能力才是 AI 取代不了的」——其实都是这个迷茫的外在表现。大家在旧层里找出路，找不到，是因为出路不在那一层。

## ◆ 一个放回日子里的例子：学徒与大师

做个比方。古代的手艺行当，有师傅，有学徒。学徒学的，大部分是「照着做」：怎么选料，怎么下刀，怎么上色，怎么出活。这是常规的手艺。

可行当里偶尔会出现一件事：有人做出了一种全新的样式，之后大家都跟着做。比如某种瓷器的新釉色，某种建筑的新结构。这样的「新」，师傅教不了，因为师傅自己也是偶然碰到的；行会里没有「新釉色发生学」这门课，只有「老釉色怎么调」的传承。

在手艺时代，这不是问题。绝大多数的活儿，是做「已经有的东西」，新样式几十年才出现一回。

现在假设，有一天，所有「照着做」的活儿，机器都能做得又快又好。行当里的人会怎么样？他们会发现，自己唯一能比机器多做的，是「做出新的样式」。可偏偏，这件事从来没有人系统地教过他们。

这就是我们的处境。库恩让我们看见了「新样式」这件事的存在，却没有给出做法；而在过去，没有做法也能过得去。

## ◆ 小结：我们接过的是什么

这一章做了三件事。

第一，回顾了库恩的二分：常规科学与革命，也就是典范内的解谜和典范的更换。

第二，指出库恩描述了革命，却没有给出工序。这个空白，历史上没人急着补，因为典范发生很少，且常规科学养活了几乎所有的人。

第三，说明了现在的变化：AI 接走常规科学后，人被推向典范发生，这件事从偶发变成主业，需要成为一门有对象、有方法的学。

接下来的问题自然是：如果要给典范发生设计工序，我们先得弄清楚典范到底是什么，它为什么一直被说成「发现」。如果这个前提不理清，工序就无从下手。这就是下一章要做的事。

## ◆ 几次大转换的共同面貌

库恩的书里讲了许多科学史的例子。这里只拣几件大家都比较熟的，而且只写有把握的大意，不去编日期和数字，也不引原话。

先看天上的事。长期以来，欧洲的天文学家在一个大家都接受的框架里工作：地球在中心，天体绕着它转。这套东西并不粗糙。

为了让它解释得了行星那些时前时后的怪走法，人们加上了许多层层叠叠的修正。能用，也能预测，历法和航海都靠它。这就是常规科学的样子：框架稳固，难题就一个个解，解不开的就加一道修正。

后来，观测越积越多，修正越加越繁。到了十六、十七世纪，换一种安排——太阳在中心，地球和其他行星绕着它转——开始被认真对待。这个换，不只是换了个位置。

它连带改变了人们怎么想「运动」，怎么想「天上和地上是不是一回事」，甚至怎么想「人在宇宙中的位置」。一批原来根本不成问题的问题，成了问题；一批原来紧要的问题，不再紧要。

再看力学。牛顿的力学，统一了天上和地上的运动，在后来的两百多年里，它的成功是惊人的。人们用它预测行星，造机器，建桥梁。常规科学在这个典范下，进行得极其红火。

可是到了十九世纪末二十世纪初，有些现象，在牛顿的框架里怎么都放不稳，比如光的传播与运动的关系。后来出现了相对论，把时间和空间的看法整个换了；又出现了量子力学，把「测量」和「确定性」的看法整个换了。

再看地球。很长时间里，人们看地质，是把大陆当作固定不动的，山是怎么起来的、海是怎么形成的，各有各的解释。

后来，随着对海底的探测、对岩石磁性的研究、对地震分布的积累，一种新的整体图景出现了：地表由一块块板块组成，它们缓慢地移动，碰撞，分

开。这个图景一出现，许多原来分散的现象——山脉、地震带、火山带、大陆边缘的形状——忽然连成了一张网。

把这几个例子并排放着，有几个共同的面貌，值得留意。

第一，旧框架并不是愚蠢的。它们在很长时间里管用，养活了大批工作，产出了真实的成果。

第二，转换发生之前，通常有一段时间，难题在积累，修补在加码，局部的调整越来越多。

第三，转换发生时，不只是答案变了，问题的清单也变了：什么是问题，什么是对象，什么叫解释，一起跟着变。

第四，转换之后，很多事得重来：教科书要重写，仪器的用法要重新理解，旧数据要重新读。

这四个面貌，库恩都讲到了，讲得很好。这里重提，是为了把一个观察放到桌面上：这样的转换，在人类历史上的数量，真的不多。

## ◆ 一个对比：棋手与规则的设计者

为了让「常规」与「革命」这两种活动的区别更直观，不妨想一想下棋。

一位棋手，从入门到高手，要学开局、中局、残局，要背棋谱，要练计算，要培养棋感。他所有的本事，都用在一件事上：在既定的规则下，赢对手。这是一门深奥的学问，需要天赋和苦功。

现在 AI 下棋，已经远远超过人类最强的棋手。在这个意义上，「在既定规则下赢」这件事，人类已不再是顶峰。

那么，棋手们的出路是什么呢？有一些人转向了研究 AI 的下法，从中学习。这仍然是在规则里。还有一种完全不同的事：设计一种新的棋。改变棋盘的形状，改变棋子的走法，改变胜负的条件，使得一种全新的游戏出现。

设计新棋的人，所需要的本事，和棋手完全不同。他需要想：「什么样的规则，会产生有趣的局面？」「什么样的规则，会让对局既有变化又不至于乱？」这是对规则本身的思考。

常规科学像是下棋，革命像是设计一种新棋。库恩告诉我们，历史上确实偶尔会有新棋出现，之后所有人都在新棋上下。但他没有告诉我们，怎么设计一种好的新棋。

过去，设计新棋这件事，没有人专门去学，因为下棋的人太多、岗位太多，大家都忙着在旧棋里争高下。如今，下棋这件事自动化了，棋手们发现，那个原来没人关心的位置——设计新棋——突然成了唯一有空位的地方。

## ◆ 「危机」到底是什么

库恩特别强调「危机」：旧典范出了问题，大家开始不安，各种解释和补丁涌现，这时，新典范才有机会出现。

要把这个说得再具体些，用日常的例子。设想一间老房子，漏雨。你先补一块瓦，好了。过一阵另一边又漏，你再补。再过一阵，墙角渗水，你抹一层防水。

这些都是「常规维修」。可是补着补着，你发现，漏的点越来越多，补丁摞着补丁，屋架本身有点歪了。这时候，你开始有一种不安：这房子是不是该整个翻修，甚至重盖？

这个不安，就是危机。注意，它不是由某一块瓦造成的，而是许多零星的问题累积起来，形成的整体感觉。

库恩的描述到这里，是很传神的。可也恰恰到这里，留下了问题：多少补丁算「够多」？什么样的漏，是「屋架的问题」，什么样的漏只是「瓦的问题」？一个漏点，是该补，还是该当作危机的信号？当年的人们怎么知道的？他们大多是事后才知道的。

对这个问题，旧的画面没有答案，只能说「要靠嗅觉」。一个行当里，有的人嗅觉灵，有的人嗅觉钝，但嗅觉这东西，既教不了，也评不了。

这个问题，本书的第三编会接手。它会把「漏点怎么算够分量」，说成差异序列在旧显露之下排不进去、特征纠缠形成压力。眼下先记住：库恩给了我们「危机」这个词，但没有给出一把量危机的尺子。

## ◆ 描述不等于工序：一个厨房里的区别

不要小看「描述」和「工序」的距离。

一位美食评论家，写一道名菜，写得出神入化：色泽、香气、口感、层次，一层层写出来，读者口水直流。然而，把这篇评论交给一个新手，对他说：「照着做一遍。」他做不出来。因为评论是对结果的描述，不是对过程的指引。

菜谱则完全不同。菜谱说：先放多少油，几成热，放什么料，翻炒几下，加多少盐，几分钟出锅。照着做，结果八九不离十。

库恩对革命的叙述，是美食评论，不是菜谱。它让你理解这道菜是什么、为什么好、怎么被接受，却没法让你做出来。

关键的变化是：过去没有人需要这份菜谱，因为这道菜几百年才上一次桌，上桌时也不是靠菜谱做出来的。如今，厨房里的其他活——切菜、炒菜、装盘——都被机器包了，人只剩下「想出一道新菜」这件事可做。这时候，菜谱的缺失就从无所谓变成了要命。

当然，典范发生的「菜谱」，不会像做菜那么死板。不会有「放三克盐」那样精确的指令。它更像是一套认路的方法：从哪里看起，注意什么，怎么试，怎么判断走偏。但哪怕只是「认路的方法」，也远比「全靠灵感」强得多，因为它是可传授的、可检验的、可改进的。

## ◆ 一个常被忽略的细节：库恩自己也看见了这个难处

说句公道话，库恩并非完全没有意识到这个难处。他曾经指出，革命不是靠一条逻辑的链条走出来的，也不是靠一次证据就能说服所有人的；新典范的支持者要靠种种理由争取别人，而这些理由不全是硬证据。他也承认，新旧典范之间缺乏中立的比较标准。（这些是对他看法的大意转述。）

换句话说，他知道这里的难处，也知道这里没有现成的方法。他选择了如实描述，没有强行给出一套方法。这是一种学者的诚实。

我们今天站在他的肩膀上，面对的情形是：他的诚实留下的空白，在那个时代可以安心空着，在这个时代不行了。我们不是在批评他，而是在接他的班。

## ◆ 常规科学里的「人」：一个更细的账

有必要把「常规科学被接走」说得更具体一点，免得读者觉得这是危言耸听。

拿一个中等规模的研究组来算账。组里有一位负责人，几位博士后，十来个研究生。他们一年干的活，大致可以分成几块：读文献、设计实验或计算、跑实验或模拟、整理数据、写论文、答审稿意见、写项目申请书、带学生。

哪几块已经能被　AI　大幅分担？读文献，可以。设计和计算，大部分可以。整理数据，可以。写论文和答意见，可以。写申请书，可以。至于跑实验，凡是能自动化的，也在被自动化。带学生，AI 能做一部分答疑。

把这些块一一勾掉以后，一个研究组的人员需求，就不再是原来的那个数。需求的下降，不会在一夜之间发生，但方向是清楚的。而且，下降最快的，恰恰是那些入门级的岗位：研究生和初级博士后干的活，恰好是最容易被分担的。

这对整个科研的「人才梯子」是个打击。从前，一个人是靠从最基础的活干起，一步步爬上去，在爬的过程中积累经验，最后可能碰上一个真正的好想法。如今，梯子的最下面几级被抽掉了。年轻人怎么爬上去，是个真问题。

我们需要同时看清两件事。第一，常规科学的人口需求在下降。第二，梯子的下几级在消失，「在典范内慢慢熬出直觉」这条传统路线不再通畅。因此，「上升」这件事的教育和训练，不能再靠以前那种「熬」，必须有明确的、可教的东西。这也是典范发生需要成为一门学的另一个理由。

## ◆ 「偶发事故」的另一层意思

前面把典范发生叫「偶发事故」，这个词要解释一下，免得读者误会为贬义。

事故的意思，是说它不在计划之内，没有人能把它排进日程。「明年第三季度，我们将发生一次典范转换」，这样的话，没有哪个研究所说得出口。历史上的大转换，每一次都是在人们没有安排的时候，从一堆杂乱的局面中冒出来的。

偶发的意思，是说它的频率很低，低到可以忽略不计，不需要为它配置专门的人手。

事故和偶发，加在一起，就是「可以不管」。社会可以不管它，因为不管也不影响大局；个人也可以不管它，因为一辈子很可能碰不上。

如今，「可以不管」的前提在松动。当常规科学不再能容纳那么多人，社会就必须想：剩下的人，去做什么？如果答案是「去做典范发生」，那么，「不管」就不再是一个选项，因为一件被指望来承载大量人力的事，不能再靠运气。

它说的不是「典范发生一定会变多」，而是「社会对它的依赖会变重」，因此对它的可靠性和可教性的要求，也会变高。

---

带走的话

**- 库恩把科学分成两种活动：典范里的解谜（常规科学），和典范的更换（革命）。 - 他把革命讲成了格式塔转换，讲得生动，却没有给出「怎么做」的工序。 - 以前没有工序，是因为不需要：常规科学养活了几乎全部科研人口，典范发生只是偶尔的事。 - AI 接走常规科学之后，典范发生从偶发事故变成了主业，第一次需要成为一门学。 - 补这个空白，是现实的压力；它不依赖于「AI 到底能不能进入典范发生」这个问题的答案。**

---

第 5 章

# 典范为何被当作「发现」的框架：旧本体论的预设

先设想一个小场景。

有个小孩，第一次去海边，看见沙滩上有一个很漂亮的贝壳，捡起来递给妈妈，说：「妈妈你看，我发现了一个贝壳！」

没有人会觉得这句话有什么不对。贝壳本来就在那儿，小孩走过去，看见了，捡起来。贝壳是客体，小孩是主体，「发现」就是主体遇到了本来就在那儿的客体。

这个画面太自然，自然到我们几乎不会去怀疑它。而我们对科学的想象，大体上就是把这个画面放大：自然界像一片巨大的沙滩，上面散落着各种规律；科学家就是沿着沙滩走的人，一个一个地捡起来。牛顿捡到了万有引力，爱因斯坦捡到了相对论。这些东西，在他们捡起来之前，就已经躺在那儿了。

这一章要问：这个画面是从哪里来的？它为什么这么顺手？它又藏了什么预设，使得「典范是怎么来的」这个问题，长期没有被认真问过？

## ◆ 一副看不见的眼镜

先说清楚，什么叫「预设」。

预设，就是你在开口说话之前，已经默认了的东西。比如有人问你：「你停止打老婆了吗？」不管你答「是」还是「不是」，都等于承认你有过这件事。问题本身把一个前提塞给了你。

我们谈论科学时，也有一个这样的前提，塞得特别深：世界上有一个「主体」（认识的人），有一个「客体」（被认识的东西），二者各自独立存在，先于任何接触而在；知识，是主体去接触客体之后得到的东西。

按这个想法，知识的来路只有一条：发现。客体本来是什么样，主体去看，去测，去记，把它如实地记下来，就成了知识。比如量水温：水本来就是三十度，温度计不过是把这个数读出来。这就是旧画面里「发现」的标准样子。

这套想法，本书管它叫「旧本体论」。「本体论」是个听起来吓人的词，其实意思很朴素：关于「世界上到底有什么、它们怎么存在」的那套默认看法。说它「旧」，不是说它过时、愚蠢，而是说本书要讲的另一套（SDE）和它不一样，需要有个名字区分。

在旧本体论的画面里，主体和客体是两样东西，像两个人，先各自站好，然后握手。握手发生在站好之后。

## ◆ 在这副眼镜下，典范是什么

戴着这副眼镜看科学，典范会呈现为一种什么样子？

第一种样子：典范是「对世界的描述」。世界本来是那个样子，典范是对它更准或更不准的描述。日心说比地心说准，是因为日心说更接近世界本来的样子。好比一张地图：地面是现成的，地图画得越贴近地面，就越准。

第二种样子：典范的更替，是「发现了更对的描述」。旧典范被换掉，是因为新的发现表明旧描述错了。

第三种样子：做出典范的人，是「看见了别人没看见的真相」的人。他们像是眼睛特别亮的探险者，比别人先走到了沙滩的某个角落。

这三种样子拼在一起，就构成了一种我们非常熟悉的叙事：天才发现真理。这个叙事好听，好记，也好教，所以流传极广。

可是，在这个叙事里，有一个东西是缺席的：典范的「发生」。它只有「发现」。既然典范只是对已经存在的世界的描述，那它从哪里来的，就只剩下「某个人看见了」这一个答案。至于看见是怎么发生的，那是天才的秘密，是灵感，是直觉，是不可言说的东西。

于是，我们就有了一个奇怪的局面：科学是最讲究方法的事业，却在最关键的地方——新典范的产生——没有方法，只有「天赋」和「运气」。

## ◆ 这解释了库恩的空白

回头看上一章，库恩为什么没有给出革命的工序？

一部分原因，我们已经说了：没必要。但还有一部分原因，更深：在旧本体论的画面里，「工序」这个概念，在这个位置上，几乎没有立足之地。

想一想，如果新典范是「发现」，那么要问的是：怎么才能发现？回答只能是：多看，多测，多想，运气好。这些建议，没有哪一条是「工序」。你不能给一个人一份清单，上面写着「第一步，发现真理；第二步，发现更多真理」。发现是一件事，不是一串事。好比有人丢了钥匙，旁人能给的建议只有「多找找」。「找到」只是一个结果的名字，不是一套步骤。

更要紧的是，库恩本人，虽然提出了「范式」和「不可通约」，打破了「科学是不断积累的直线」的旧叙事，但他用的语言，在许多地方，仍然带着旧本体论的痕迹。

他说科学家在革命之后「生活在不同的世界」，这已经很大胆了，可他对「世界为什么会换」的解释，主要是社会的、心理的：群体的信念变了，感知的方式变了。主体和客体的画面本身，他并没有从根上动。

所以，库恩看见了典范会换，却没法说明它怎么换，因为他说明所用的词汇，还在旧画面里。这就像一个人看见了冰在融化，却只能用「冰」这个词去讲：他知道有事发生了，但他没有「水」这个词。

## ◆ 旧画面里的几个不舒服

旧本体论的画面，并不是没有毛病。只是它太顺手，毛病被习惯盖住了。我们列出几个，让读者感觉一下。

第一个不舒服：同一个「客体」，为什么不同的典范下，看出来的东西不一样？比如同样是天上的那些亮点，在一种典范里是「嵌在水晶球上的灯」，在另一种里是「遥远的、各自运行的天体」。如果客体就在那儿，等着被如实发现，那么这些不一样的「如实」，该怎么算？

一个常见的回答是：其中一个对，其他的错。这对一部分情形是说得通的。可对另一些情形，就麻烦了。比如，牛顿力学和相对论，在低速的世界里给出几乎一样的预测，而它们对「时间」「空间」的看法却截然不同。

我们不能说牛顿力学「错了」，因为它到今天还在造桥、发射卫星；但我们也不能说它「对」，因为它对世界的基本图景和相对论不相容。在旧画面里，这种局面只能含糊地说「近似」。

第二个不舒服：在量子力学里，测量这件事似乎不是一个「旁观者如实记录」的动作。测量，好像参与了所测之物的显现。这一点，让「客体先于互动独立存在、主体只是去看」的画面，在最前沿的物理里站不稳。物理学家们为此争论了将近一个世纪，至今没有一个人人满意的说法。

第三个不舒服：那些据说「发现了」新典范的人，他们并不是凭空看见的。他们站在前人的积累上，被一连串的难题推着走，被当时的仪器、数学工具、争论的氛围塑造着。把功劳说成是某一个人的「眼睛亮」，不合史实。它也解释不了，为什么类似的想法常常在同一时期、在几个地方几乎同时冒出来。比如微积分，牛顿和莱布尼兹几乎同时、各自独立地给出了系统的做法。

这三个不舒服，指向同一个地方：主体与客体「先于互动而在」这个预设，恐怕出了问题。

## ◆ SDE 把画面换了

本书用的这套看法，叫　　SDE。它是三个词的缩写：Show（显露）、Difference（差异序列）、Entanglement（特征纠缠）。中文可以说：显露，差异序列，特征纠缠。

先不展开细节（细节在第三编），这里只讲它和旧画面的根本不同。

旧画面：先有主体，先有客体，二者接触，产生知识。

SDE 的画面：先不预设有现成的主体和客体。有的是一场正在进行的事——各种差异在序列中展开，各种特征彼此纠缠，在纠缠中，某些东西稳定下来，显露出来。主体是什么、客体是什么，是这场显露的结果，不是它的起点。

打个比方。想象一杯水里滴进一滴墨水。墨水在水里散开，形成一些纹路，有深有浅，有聚有散。你看着这些纹路，说：「这里有一团云，那里有一条线。」可是，云和线在滴墨水之前并不存在，它们是墨水与水互相作用时显露出来的东西。

你不能说，「云」这个客体先在那儿，等你去发现它。也不能说，是「你」凭空画出了这朵云。云是这场互动中稳定下来的显露。

用这个画面来看知识：知识不是主体去捡起了躺在沙滩上的客体，而是一场互动里，某种稳定的东西显露了出来。我们平常说的「主体」和「客体」，是对这场显露事后的区分，是显露的不同侧面。

这样一换，「发现」这个词的位置就变了。发现，是在显露已经稳定之后，回头去看：「哦，有这么个东西。」它描述的是后半截。而前半截——这个稳定的东西是怎么发生出来的——旧画面没有位置，SDE 要把它补上。这就是为什么 SDE 叫「发生学」，而旧画面可以叫「发现学」。

## ◆ 回到典范：从「被发现」到「被发生」

把这个换法用到典范上，我们得到一个全新的看法。

典范，不是对一个现成世界的描述，等着有人去发现。典范是一个稳定的显露：一整套看什么是对象、怎么划分、怎样算数、什么叫解释的方式，在差

异序列和特征纠缠的运行中稳定下来。它不是被谁「看见」的，它是发生出来的。

在这个看法里，革命也变了样子。旧典范被换掉，不再是「人们发现旧描述错了」，而是：有些差异，在旧的显露之下排不进去；有些特征纠缠在一起，在旧的框架里解不开，形成压力；旧的显露变得不稳定；然后，一个新的稳定的显露发生出来。

注意这里的主语：是差异序列、特征纠缠、显露。不是「一位天才」。如果一定要说到那些著名的名字，我们会说：在这条差异序列上，显影出一位提出猜想的人。他是这场发生的结果之一，是它露出来的一个侧面，而不是发动它的人。

这里要特别小心。因为一不留神，我们又会滑回旧画面：把「发生」说成「某人创造了新典范」。那只是把「发现」换成了「创造」，主体仍然是那个先于一切、发动一切的人，画面没有换。

真正的换，是承认：没有一个先在的、站在一切之外的「发动者」。这听起来也许有些反直觉，所以后面几章，会反复用具体的例子来讲。

## ◆ 为什么这个换法对 AI 时代要紧

读到这里，读者也许要问：这些哲学上的区分，和 AI、和教育、和科研的层级，有什么关系？

关系大了。

如果你坚持旧画面，典范的更替就只能是「发现」。而「发现」没有工序，只有天才和运气。那么，AI 接走常规科学之后，人被推向典范这一层，他

们能做的，只有「等灵感」「碰运气」，没有任何东西可以教、可以练、可以评。设想一个学生问：「怎样才能想出日心说那样的东西？」旧画面下，老师只能答：「多读书，多观察，等灵感。」整个教育和科研制度，面对这一层，就是一片空白。

如果你换成 SDE 的画面，典范是发生的，那么发生就有它的结构：有差异序列，有特征纠缠，有稳定与失稳，有新显露的最初一笔。结构可以被识别，被描述，被练习。工序，才有可能出现。

换句话说，旧本体论不是一个无关紧要的哲学背景，它是挡在「典范发生学」前面的那堵墙。墙不拆，工序无从谈起。

这也解释了，为什么今天的各种对策——看重过程，培养提问，人机协作——都在典范内打转。它们的共同底色，是旧画面：既然知识是发现，那么能做的就是把发现做得更好，更快，更严格。比如「看重过程」：让学生把解题的每一步写出来。步骤是写出来了，可题目是现成的，规矩是现成的，什么算对也是现成的。没有一条对策，会去问「典范是怎么发生的」，因为在旧画面里，这个问题根本问不出来。

## ◆ 放回日子里：贝壳再看一遍

我们回到开头的小孩和贝壳。

「我发现了一个贝壳。」这句话在日常生活中，完全没问题。我们不需要为了捡贝壳去改写哲学。

可是，如果这个孩子长大，成了一位海洋生物学家，他面对的不再是沙滩上的一个贝壳，而是一个问题：「这到底算一个物种，还是几个？」这时，「发现」就不够用了。

因为「物种」这个概念，是在一整套划分的方式里立起来的：看形态，看基因，看能否交配，看生态位。不同的划分方式，会把同一批生物切成不同的块。要回答这个问题，他不只是「看」，而是要决定「怎么划分」。这个决定，不是贝壳替他做的。

再往前一步，如果有一天，海洋生物学的整个划分方式都要换——比如，从「以形态分类」转到「以基因谱系分类」——那件事，就不是沙滩上多捡了一个贝壳，而是整个沙滩的画法换了。那一刻，「发现」这个词，已经不够描述发生的事了。

我们不否认发现的价值。贝壳就是在沙滩上，大部分的科学工作，确实是在已经稳定的显露里「发现」东西。这一层，AI 接走了，也接得很好。我们要说的是：发现的前面，还有一层，那是「发生」。

## ◆ 小结：这一编留下了什么

回头看这几章走到哪儿了。

第 3 章说，人机流水线仍在典范内，是分工不是升级。第 4 章说，库恩把典范的更换描述为革命，却没有给出工序，这个空白过去不必补，现在必须补。

第 5 章说，空白之所以一直空着，不仅因为没必要，更因为旧本体论的画面里，典范只能是「被发现」的，发生没有位置。SDE 把画面换成发生学，典范成了一个稳定的显露，由差异序列和特征纠缠发生出来。

下一章要进一步问：既然典范是一层，那么「在典范里往广处走、往深处走」和「换一层」，到底有什么不同？这个区别，是理解后面所有工序的钥匙。

## ◆ 这副眼镜为什么这么好用

旧本体论为什么能统治这么久？不是因为人们不够聪明，而是因为它在三个地方特别好用。

第一，它符合日常经验。我们每天伸手拿杯子、开门、看路牌，都像是「一个我，面对一个现成的东西」。杯子在那儿，我去拿。这个经验太强，太频繁，以至于我们把它当成了世界的基本结构。

第二，它藏在我们的语言里。你看我们说话的句子：谁，做了什么，对什么。「科学家发现了规律。」「我看见了一只鸟。」「她认识了这个道理。」句子的结构是「主语—动词—宾语」，主语是主体，宾语是客体，动词是二者之间的动作。

语言替我们把画面预先排好了。想换一种说法，需要费力气，因为句子总是想把我们拉回原来的轨道。

第三，它使责任和功劳有了落点。谁发现了，谁就是发现者，谁就该得奖。「发现」这个概念，天生配着名字、奖状和传记。一套大家都需要的社会

安排，围绕着它建立起来：专利、署名、奖项、职称。这些安排一旦建立，就反过来巩固了那个画面。

所以，旧本体论不是一个孤零零的哲学观点，它是一整个生活方式的地基。要换，不容易，也不应该轻率。

## ◆ 难道没有一个客观的世界吗

读到这里，一定有读者想问：你们这是要说世界是人想出来的吗？桌子不是我想它在，它就在吗？撞上去还是会疼的。

这个疑问问得好，值得认真回答。

SDE　并没有说「世界是主观的」「一切都是想出来的」。那是另一种极端，也是旧画面里的另一个选项：既然客体不是现成的，那就是主体造的。这仍然是「主体—客体」这个框架里的翻面，没有离开旧画面。

SDE 说的是：不预设一个先于一切互动的、现成的主体，也不预设一个先于一切互动的、现成的客体。有的是差异、纠缠、稳定下来的显露。

桌子之所以撞上去会疼，不是因为「桌子这个客体先在那儿等着」，而是因为在这场互动里，有一些东西非常稳定，稳定到你怎么试，它都以同一种方式显露出来。这种稳定，是真实的，不依赖于谁喜欢不喜欢。

打个比方。河里的漩涡，你可以伸手去碰它，它有形状，有位置，有力量，能转起一片树叶。但漩涡不是水里先有的一个东西，等着你去发现；它也不是你想象出来的。

它是水流在特定条件下稳定下来的一种形态。你不能说它「不客观」，也不能说它「本来就在那儿」。要说清它，你需要的是「稳定的发生」这样的词，而不是「现成的东西」。

所以，SDE 不是否认稳定、否认可靠、否认科学的成果。恰恰相反，它要解释的，正是这些稳定是怎么来的。把稳定当作「本来如此」，反而是把需要解释的东西，当成了不需要解释的起点。

## ◆ 发现学的教育版本：知识像存货

旧本体论落在教育上，会变成一个熟悉的样子。

既然知识是对现成世界的发现，那么已经发现的部分，就是一份「存货」：公式、定律、事实、结论。教育的任务，就是把这份存货搬到学生脑子里。教得好，是搬得又快又牢；考得好，是取货又准又全。比如背下牛顿三定律、默写元素周期表，考的就是取货。

这个模型有个朴素的比喻：教育像是往水桶里倒水。老师有水，学生是空桶，倒进去，桶满了，就算学会了。

在 AI 之前，这个模型已经遭到了很多批评，人们说学生不是桶，是火把，要点燃。可批评之后，换成的「点燃」「启发」「培养能力」，常常仍然默认一个现成的东西——「能力」——可以被培养出来。

AI 出现之后，存货这个模型彻底露出了底：存货谁都能随时调取，搬运的价值归零。教育界于是陷入焦虑，开始寻找「搬运之外」的价值。找来找去，找到了「过程」「提问」「验证」。可它们仍然是存货模型的延伸：存货之外，再加上「取货的技巧」。技巧也是可以被 AI 掌握的。

要想真的越过这一点，需要换掉存货模型背后的画面，也就是，换掉「知识是对现成世界的发现」这个前提。

如果知识是由差异序列、特征纠缠发生出来的稳定显露，那么教育要面对的，就不仅仅是传递显露的结果，而是带着学生，看见显露是怎么发生的，包括：什么时候一个显露会失稳，什么时候一个新的显露可能出现。这一点，书的第五编会细讲。

### ◆ 旧画面的另一个后果：为什么「换皮」那么容易发生

还有一个后果，和后面的内容有关，先在这里埋个线头。

在旧画面里，一个新想法，常常被看作「又发现了一个东西」，于是很自然地被安放进已有的格子里：这是一个新粒子，这是一种新疗法，这是一个新的经济现象。格子本身是现成的，新东西只是被分类。

这个习惯，会带来一个麻烦：当一个真的试图换框架的想法出现时，它最容易被旧画面「翻译」成一个旧框架里的新条目。

比如，有人提出一种全新的划分世界的方式，听的人会下意识地问：「那它对应我们已有的哪个概念？」然后把它翻成已有概念的另一种说法。设想有人提出：别把世界看成一粒粒现成的东西，关系和变化才是先的。听的人多半会说：「这不就是某某学说的另一种叫法吗？」翻完以后，它看起来很合理，也很无害，同时，它真正新的部分，也在翻译中被抹掉了。

这就是后面第 14 章要讲的「换皮的旧典范」。它不是某个人的疏忽，而是旧画面的惯性：把一切新东西，都当成「发现」，并放进现成的格子里。在

AI 参与进来之后，这种惯性会被放大，因为 AI 极擅长把新说法翻译成旧说法，翻得又快又顺。所以后面需要专门的验收工序。

## ◆ 把贝壳带进这本书的后半

做个交代，这一章为全书的后半铺了一块地基。

后半要讲的「典范发生学」，需要三个前提都站得住：

一、典范是一层，不是一堆孤立的结论。这是第 6 章的事。

二、典范不是被发现的，是发生的：它是稳定的显露，由差异序列和特征纠缠发生出来。这是第三编的事。

三、让典范发生，是可以有工序的，工序之中有人和 AI 的分工。这是第四编、第五编的事。

这一章处理的是第二个前提的「为什么此前没人这么想」。答案是：旧本体论的画面太顺手，顺手到成了我们看不见的眼镜。眼镜摘下来之前，「让典范发生」这件事，在我们的视野里是没有位置的。

## ◆ 一段小对话

为了让这一章更好懂，来听一段想象的对话。老师姓林，学生姓陈，在大学的走廊上。

陈同学说：「林老师，您常说科学不是'发现'，可课本上明明写着'某某发现了某某定律'。」

林老师说：「课本没有写错。它写的是事后的说法。你想想，一个定律被写进课本的时候，它已经稳定了，大家都在用，用起来都对。这时候回头说'有人发现了它'，很顺口。」

陈同学说：「那稳定之前呢？」

林老师说：「稳定之前，是一团乱。有各种各样的测量，彼此对不上；有各种各样的说法，彼此打架；有好几种看问题的方式，在争。哪一个能活下来，哪一个会稳定，在当时谁也说不准。如果那时候你问一个科学家，他会告诉你：我们还在摸。」

陈同学说：「那最后是谁'发现'的呢？」

林老师说：「你看，这个问题已经带着答案了。你问'是谁'。我们想一想，是一个人突然看见，还是许多条线索一点点聚到了一起，聚到某个时刻，某一个人把它说清楚了？后一种，更像大部分真实的情形。说清楚的那个人，当然重要。但他也是这场聚拢的一部分，是聚拢的结果露出来的一个地方。」

陈同学想了想，说：「这样说，好像天才的光环就淡了。」

林老师笑笑说：「不是淡了，是挪了位置。光环原来罩在一个人头上，现在罩在整个过程上。过程值得敬畏。而且，过程是可以研究的。如果只有一个人头上有光，你就只能崇拜；如果整个过程有光，你可以学。」

这段对话想说的，正是这一章的要点：从「发现」换成「发生」，不是要贬低那些伟大的名字，而是要把本来罩在名字上的神秘，还给过程，使过程可以被看清、被学习。

带走的话

- 我们习惯的画面：主体、客体先于互动而在，知识是发现。这套默认看法，本书称为旧本体论，也可以叫发现学。 - 在这副眼镜下，典范只能是「对现成世界的描述」，更替只能是「发现了更对的描述」，新典范的来路只能归给天才和灵感。 - 这就是库恩的空白一直空着的深层原因：旧画面里，「让典范发生」没有位置，也就谈不上工序。 - SDE 把画面换成发生学：典范是差异序列与特征纠缠中稳定下来的显露，是发生出来的，不是被谁发现的。 - 发现描述的是显露稳定之后的事；稳定之前的事，要用发生来讲。提出猜想的人，是这场发生里显影的一个结果，不是它的发动者。

第 6 章

# 层级：广度、深度之外的换层

先说一个很家常的场景。

一个小镇上有一条老街，街上有一家修鞋铺。铺子的老板手艺好，生意也好。有一年，街对面开了一家更大的店，货架上摆满了便宜的新鞋。老板琢磨怎么办。他想出的办法有三个。

头一个，把铺面开大，多雇两个伙计，什么鞋都收，什么毛病都修，这叫做宽。第二个，把手艺再磨细，皮子怎么缝、胶怎么上，比谁都讲究，修出来的鞋比新的还耐穿，这叫做深。

第三个，他想了很久，才想明白：对面那家店卖的不是鞋，卖的是「出门不用操心脚上穿什么」。人们要的是一整套出行的体面。他于是不再只修鞋，而是开始搭配、定制、上门量脚，把一间修鞋铺变成了另一种行当。

前两个办法，都是在「修鞋」这件事里使劲，一个往外铺，一个往下钻。第三个办法，是换了一层来看这件事：不再问「鞋怎么修得更好」，而是问「人们到底在为什么付钱」。

这一章要讲的，就是这第三种使劲。我们习惯了说「拓宽视野」「加深功底」，好像一个人、一门学问的长进，只有这两个方向。可是前面几章讲过的那件事，恰恰不在这两个方向里。

AI 接走了典范内的答案、过程、提问和验收，你再往宽处走，它也能铺得更宽；你再往深处钻，它也能钻得更深。真正没有被一起搬走的，是另一个方向：换层。

## ◆ 广度和深度，都是在同一层里量尺寸

先把话说清楚。什么叫「同层」。

拿一个中学生都熟悉的东西来说：解方程。一元一次方程会解了，再学一元二次方程，再学二元一次方程组，再学含参数的方程，再学高次方程。题目的类型越来越多，这叫广度。同样是一元二次方程，只会套公式，和能一眼看出配方、能从图像上看出根的位置、能判断什么时候该用判别式什么时候不该，这是深度。

广度和深度有个共同点：你用的始终是同一套东西。方程是什么，等号两边怎么变形，什么叫「解」，什么算解对了，这些规矩从头到尾没有动过。你是在这套规矩里面，铺得更远，或者钻得更深。

打个比方，你站在一个院子里。往东走，院子越走越长，这是广度。往地下挖，挖出一口井，这是深度。可是不管你走多远、挖多深，你始终在这个院子的地面和地底。院子的围墙没有变，院子外面有什么，你是看不到的。

换层是什么呢？是爬上屋顶，或者出了院门，在街上回头看这个院子。这时候你看到的东西，是在院子里无论怎么走、怎么挖都看不到的：这个院子有多大，在整座镇子的哪个位置，为什么这里是院子、那里是街。

所以我们说，广度和深度是同层的扩张，层级是换层。这三个词说的不是一个方向上的三个档位，而是两种完全不同的动作：一种是在地面上量尺寸，一种是离开地面看地面。

科学研究里，这个区别一点也不抽象。一位做材料的研究者，今天研究一种合金的强度，明天研究另一种合金，后天把方法推到陶瓷，这是广度。他把某种合金的强度研究做到了小数点后五位，把每一个影响因素都拆清楚，这是深度。

可是「什么叫材料的强度」「什么样的测量才算数」「哪些问题值得问」，这些在他入行时就已经被前人定好了，他无需过问，也不会去过问。这一整套东西，就是第四章讲过的「典范」。典范就是那个院子。

## ◆ 典范就是一层：院子里的一切都同层

这是全书的一个关键判断，要讲得慢一点。

典范是什么？按库恩的意思，大体上是一个学科共同体共享的一整套东西：什么是基本的对象，什么样的问题算问题，什么样的解答算解答，用什么方法，怎么算验证。第四章已经讲过，这里不重复。我们只抓一点：这是一整套互相配合、互相规定的东西。

问题和答案是靠它规定的。什么算问题，由典范划出来。什么算答案，由典范划出来。从问题走到答案的路，也就是过程，是典范里认可的步骤。最后检查答案对不对，也就是验收，用的还是典范里的标准。

所以，答案、过程、提问、验收，不是四个层，而是同一个院子里的四个位置：终点、路径、起点、质检。你站在哪个位置，换到哪个位置，都没有出院子。

这里要特别防一个误会。有人会说：提问总比答案高一层吧？会问问题的人，总比只会答题的人高一层吧？这话在日常里没错，会提问的学生确实比只会背答案的学生强。可是在典范这个尺度上看，两个人都在同一个院子里。

一个人问「这种合金在高温下为什么会变脆」，这个问题本身，已经默认了「合金」「高温」「脆」这些东西是什么，已经默认了怎么去测、怎么去算才有意义。问题是在院子里长出来的，它不是院子外面飞进来的。

所以第二章、第三章里反复讲的那件事：AI 在一个给定典范里，可以提出非常多的问题，而且质量不低。

它能读完某个领域几十万篇论文，然后指出：这里有一个没有人做过的组合，那里有一个与已有结果不一致的细节，这个参数区间还没有被测过。这些都是实打实的好问题，只不过它们都是在院子里发现的空位。

我们不需要给人留一块「AI 做不到」的保留地，也不必说「提问是人类独有的」。这样说是靠不住的，因为明天的 AI 也许就会证明它能。

判断的标准要换成另一个：成本结构。凡是在一个给定典范里能被廉价地、大量地生产出来的东西，就不再是竞争点。答案如此，过程如此，典范内的提问如此，典范内的验收也如此。

## ◆ 一道数学题的三层楼

光讲院子，还是有些空。我们换一个更具体的例子，让这三层看得见摸得着。

设想一道很普通的数学题：求一个具体的数，比如解方程 $x^2 - 5x + 6 = 0$。

这是「元素层」。题目里的东西是一个个具体的数：2、3、5、6。要做的事是算出结果。学生学了求根公式，或者因式分解，一步步算，得到 $x = 2$ 或 $x = 3$。对错一目了然，代进去验一验就是。AI 做这类题，已经没有任何悬念。

现在往上走一层。不再问「这个方程的根是多少」，而是问「一元二次方程的根，一般有什么规律」。你开始关心的对象，不再是 2 和 3，而是「所有这样的方程」。

你得出求根公式本身，得出判别式，得出根与系数的关系：两根之和是什么，两根之积是什么。这时候，上一层的东西是什么呢？是一整类题。刚才那道具体的题，成了这一类里的一个元素。

注意这个关系：刚才那一层里的「题」，到了这一层，变成了「例子」。刚才一层里要费力气算出来的东西，到了这一层，变成了一个被公式一次性打包处理的东西。上一层的人看下一层，看到的是一堆可以批量生产的东西。

再往上走一层。这时候，问题变成了：「什么样的方程有公式可解？」你不再只盯着二次方程，而是比较一次、二次、三次、四次、五次方程。

你发现三次、四次也有公式，只是越来越长，到了五次，竟然没有一般的公式。更妙的是，这个「没有」是被证明出来的，而且证明的方法，是回头去看「方程的根之间有什么对称关系」，由此发明了群的概念。

这一层里，对象不再是方程，而是「解方程这件事本身的结构」。你关心的是：这类问题的解法，能不能存在，存在的话是什么形状。这是又一次换层。刚才那一层里的「一类题」，到了这里，又成了一个元素。

看一下这三层之间的关系。第一层的元素是数，要算。第二层的元素是一类题，要找规律。第三层的元素是一整套解法的结构，要问它能不能存在。每一层的「元素」，都是下一层的「整体」。

这里面藏着一个很重要的细节。每上一层，下一层的工作并没有消失，它只是变成了一种可以被当成零件使用的东西。以前要算半天的事，现在是一行公式。但是，上一层的人如果不会下一层的事，他就根本不知道自己在说什么。

### ◆ 科研里的三层：做实验、做理论、做出理论的方式

把数学题的三层搬进科研，也是同样的形状。

最下面一层，是在一个成熟的理论框架里做具体的事。测一个数，算一个结果，把一个已知的方法用到一个新对象上。

一位做天体物理的研究生，拿到一批新的观测数据，用已有的模型去拟合，得出某颗恒星的质量。这是元素层。工作量很大，要求很高，一点也不轻松，但这一层有一个鲜明的特点：对错有现成的尺子。

往上一层，是改进这套模型本身。发现旧模型在某个区间拟合不好，于是加一项修正，换一种参数化，或者把两个本来分开的模型统一起来。这一层的对象，不再是「某颗恒星」，而是「这一类恒星的模型」。刚才下一层里的具体测算，在这一层里成了检验模型的样本。

再往上一层，才是我们这本书要讲的那个层：问「这一整套看恒星的办法，本身是怎么回事」。恒星是什么，是一团气体，还是别的什么？「质量」「温度」「光度」这几个量，凭什么是该问的量？

观测什么才算数，模型好不好用什么来判？这些在上面两层里，从来不是被拿来讨论的东西，它们是讨论所站的地面。第三层要做的，是把这块地面本身当作对象，重新划一遍。

这里必须强调：这三层不是「初级、中级、高级」的难度档次。一个顶尖的研究者，可以长年累月都待在第一层或第二层，把活干得极为出色。一个在第三层上折腾的人，也可能一辈子什么都没折腾出来。层级不是水平，层级是工作的对象换了。

我们在这里要看清的是一个结构事实：绝大部分科研人口，一直工作在第一层和第二层。库恩把它们合起来叫「常规科学」，也就是在一个给定典范下的解谜。

第三层，库恩叫「革命」，那是几百年里偶尔发生的事，没有人认真问过「怎么做」。为什么没人问？第四章已经说过：因为没有必要。第一层和第二层的工作足够养活所有人，第三层的偶发，不会影响任何人的饭碗。

可是现在，AI 把第一层、第二层都接走了。这就是为什么「换层」忽然成了一个现实的问题，而不是一个哲学好奇。

## ◆ 地基与楼层：退出不等于跳过

讲到这里，一个很自然的反应是：那我们以后就不用学下面两层了吧？直接教学生往上一层走，不是更省事吗？

这个想法，恰恰是最危险的。而且它的错法，和第一章里「腿退出运输岗位」的故事是一模一样的。

回到车和腿。车出现以后，腿退出了运输的岗位。没有人再靠两条腿把货物从一个城市运到另一个城市。但是这不等于腿消失了，也不等于人不再需要会走路。

腿进入了更高的交通功能：人要决定去哪里，要看得懂地图，要在车到不了的地方继续走，要在车坏了的时候知道怎么办。所谓「交通」，在车出现之后被拉宽了，拉深了，也拉到了一个新的层级：路网的设计、城市的规划、物流的调度。

但是，规划路网的人，不能是一个从来没走过路的人。一个人若从来没有用自己的脚丈量过一段山路，他看地图是看不懂的。地图上的等高线、坡度、转弯半径，都是把脚底的经验压缩成符号。不走过路，符号就是空的。

所以我们说：腿是地基，不是楼层。楼层可以上，地基不能拆。

数学里也是一样。刚才说，第二层的人把第一层的题打包成了公式。可是，一个从来没有亲手解过方程的学生，直接去学「根与系数的关系」，他会觉得那是一堆莫名其妙的符号。

他背得下来，却不知道这个公式为什么长这个样子，更不知道什么时候该怀疑它、什么时候它会失效。至于第三层的群论，如果没有对一次、二次、三次方程的亲手体验，那只是一套悬空的代数游戏。

这就是「退出」和「跳过」的区别。

「退出」，是说某一层的工作，不再是人类要去竞争的位置了。算数这件事，人早已退出了和计算器的竞争。但这是退出竞争，不是退出理解。一个人仍然需要会算，哪怕他从来不会拿手算去赢计算器，因为只有会算，他才知道「数」是怎么回事。

「跳过」，是说在没有走过下面几层的情况下，直接站到上面去。这样的人上去了，也只能站在一块自己不理解的地板上，每一步都是抄来的。

我们因此得到一个听起来有点绕、但其实很朴素的结论：下面的层，对于人来说，从「工作」变成了「功课」。以前它既是饭碗，也是功课，一个人在这里干活、挣钱，同时被训练。

现在它不再是饭碗，但仍然是功课，而且是必修的功课。至于这门功课怎么练，才能练出真正够用的地基而不是白白浪费时间，是第五编要回答的问题，这里先把它记下。

## ◆ 为什么越往上越难教、越难考

还有一件事连着后面的教学与制度，要在这里说清。

下面一层，好教，也好考。道理很简单：下面一层有现成的尺子。解方程，算对算错，一眼可见。做一个标准实验，数据对不对，可以重复一遍看看。老师带学生走一遍，学生做一遍，对照标准答案，就知道学没学会。

这样的东西，天然适合批量教学，也天然适合考试。几百年来，现代教育的整套制度，基本上都是围绕着这种东西搭起来的：教材、课时、作业、考试、分数。

往上一层，尺子开始变软。改进一个模型，好不好？要看它在更多数据上的表现，要看它是不是自然，要看它能不能和别的东西接得上。这时候就没有标准答案了，只有同行的判断，而同行的判断，往往要过几年才见分晓。

再往上一层，尺子几乎没有了。你重新划一遍地面，谁来判这个划法好不好？用旧典范的尺子来量，新典范一定量不过，因为新典范的重点恰恰在于它要改掉旧尺子。

哥白尼的体系刚出来的时候，用当时通行的物理标准去量，它并不占优；这个事我们到后面案例章节再细讲，这里只记下一个结论：越高的层，越没有现成的尺子，越难教，越难考。

难教，是因为教的东西没有套路可套。难考，是因为考的东西没有标准答案。

这里有一个很有意思的、也很残酷的对应。AI 恰恰是最擅长「有现成尺子」的东西。尺子越硬，它越容易学；尺子越软，它越需要有人在旁边判断。这也解释了为什么第一章和第二章里说，凡是能被廉价生产的，都不再是竞争点：能被廉价生产的，几乎都是尺子很硬的东西。

于是局面就清楚了。教育和科研制度，是围绕着「尺子硬的层」搭起来的。AI 来了，尺子硬的层被接走，剩下的是尺子软的层，而整个教学和评价体系对它几乎没有准备。这不是哪个人的失误，这是一座房子，只给底下几层装了楼梯，上面的层从来没有人想过要上。

所以，「看重过程」「培养提问」「加强审查」，这些方案不是没有用，而是它们仍然在装楼梯的那几层里转。它们都是在院子里把路修得更好，没有

一个是在问：院子外面是什么？这就是第三章已经讲过、第五章还要追问的一点：流行的转向方案，每一个都还在典范之内。

## ◆ 同一个落体，三层不同的问法

数学题的好处是干净，但读者也许会疑心：科研不是这么整齐的吧？我们换一个历史上人人听过的现象：东西往下掉。

在第一层，问题是这样的：这块石头从这个高度落下来，要多久？换一块重一点的，又要多久？这是量和算的事。你得有个钟，有个尺，有个办法让每次实验条件差不多。数据出来，对照理论，对上了就是对上了。

在第二层，问题变成：为什么羽毛比石头落得慢？是不是空气在捣乱？能不能把空气的影响单独算出来，扣掉它，剩下的是什么样子？这是在改进模型。你没有动「落体该怎么研究」这件事本身，只是把一个粗糙的描述磨成了一个更细致的描述。

到第三层，问的是完全不同的话：「运动」到底是个什么东西？物体为什么会动、为什么会停？在很长的一段历史里，人们对这件事的直觉是：东西要一直有东西推着才会动，推力一撤，它就慢慢停下来。

这个直觉不是蠢话，生活里看到的车、马、推到桌子上的杯子，确实是这样。按这个直觉搭起来的一整套看法里，重的东西往下落得快，本来就是顺理成章的事。

后来，整块地面被换掉了：不再问「是什么让它继续动」，而是问「是什么让它改变了动」。一旦这样划，匀速地一直走下去，就不再是需要解释的

事，反而是默认的样子。需要解释的，是速度的改变。这时候，「落体」才变成一个可以被精细研究的对象，斜面、摆、数字化的时间，才有了位置。

请注意这个例子里的一个细节：在旧的那套看法里，一个人完全可以在第一层和第二层把工作做得很漂亮。你可以把「重的落得快」测得很精确，可以修正各种干扰。

可是那个把「推」换成「改变」的动作，在旧的第一层和第二层里，既不是答案，也不是过程，也不是提问，更通不过当时的验收。它在院子里没有位置，它是院子的边界被挪动。

所以今天我们说，AI 把第一层、第二层都接走了，指的是这样一件事：假如那时候就有 AI，它可以又快又细致地把旧那套看法里所有的测量、修正、问题，做到极致。

它甚至可以问出非常好的问题，比如「为什么在斜面上量出来的结果，总是偏离重物落得快的预测一点点」。可是这些问题的答案，仍然是在旧地面的修补。地面本身怎么换，不在这些问题的射程里。

这并不是说 AI 永远碰不到地面。后面的章节会专门谈这件事，也会诚实地承认：AI 能不能进入这一层，我们现在不替它下断言。但有一点现在就可以说：即使它进不去，也不影响我们的结论。因为第一层和第二层被接走，就已经足够把人推向上面。

### ◆ 换层不等于更聪明，也不等于离开

说到这里，有两种误会需要先拦住。

第一种误会，是把层级当成智力的阶梯。好像在上一层的人，就比下一层的人聪明。这是不对的。数学家证明五次方程没有一般求根公式的时候，他的聪明和那些把四次方程的公式推到极致的人，没有高下之分。

层不是人的高低，是工作对象的不同。同一个人，完全可以今天在第一层算一个数，明天在第三层琢磨一套划界。我们在上一章里说过，典范不是靠某一个人看穿的，这里也是一样：层级不是在给人排座次。

第二种误会，是把换层想成一次告别。好像上了楼，就和楼下没有关系了。事情恰好相反。一个换到上层的人，最需要的恰恰是下层的手感。他在楼下的经验，是他判断「这个新划法顶不顶用」的唯一依靠。没有这种手感，新的划法只会是一个自圆其说的故事，听起来漂亮，不能用，也说不出哪里不能用。

所以，层与层之间不是台阶，不是一级一级往上爬，爬上去就把梯子撤掉。它更像一座楼：每一层都在，每一层承着上一层的重量，上层的人要常常下楼，因为那里有厨房和水。

还有一点要补充。换层并不是一个可以随时随地、说换就换的动作。院子里的人，通常并不知道自己在院子里。你问一个做了二十年合金强度的研究者，他的「典范」是什么，他多半会觉得这个问题奇怪：典范是什么？

我就是在研究合金啊。院子对于住在里面的人来说，不叫院子，叫世界。要等到这个世界里出了一些怎么也排不进去的事，人们才会隐约觉得，是不是地面本身有点问题。

这条路是怎么走的，是第三编的主题。这里先把方向立住：换层是有对象的上升，不是一次漫无目的的跳出。上升要往哪里去，去的地方有什么，是本书后面要一章一章讲的。

## ◆ 为什么 AI 跟着平移，成本跟着归零

为什么我们说，从答案转到过程、转到提问，每一次都是平移？

设想一间教室。老师发现学生用 AI 轻松就能得出答案，作业没法布置了。于是他改了规矩：不看答案，看解题过程，要求学生一步一步写出来。头一个月，学生老老实实写。

下一个月，学生发现，AI 不仅能给出答案，也能给出一份漂亮的、一步一步的过程，甚至带着「我是这样想到的」的口气。老师又改了规矩：不看过程，看你提了什么问题。

学生又发现，AI 也能列出一串很像样的问题。老师再改：你来提问，AI 来回答，我看你们之间的对话。学生又发现，这段对话也可以请 AI 帮着写。

每一次改规矩，老师都以为自己往上走了一步。每一次，学生和 AI 都轻轻地跟了上来，几乎不费力。原因很简单：这些改动，都发生在同一层里。答案、过程、提问、对话，都是院子里的东西，用的都是同一套规矩，同一套「什么算数」。院子里的东西，AI 一旦会了，就能批量做，成本趋向于零。

你可以把它想象成在一块平地上追一个跑得越来越快的人。你往前追一步，他也往前一步，你往左绕，他也往左绕。要想不被他追上，不是跑得更快更巧，而是离开这块平地，去另一个他暂时没有路的地方。

这就是为什么我们要把「换层」说得这样郑重。不是因为它听起来更高级，而是因为在同一层里的任何平移，都无法改变成本结构。真正能改变成本结构的，只有一件事：你做的事，不再是一个现成尺子能量的事。

有个诚实的限制要先说。换层之后，AI　会不会又跟上来，我们不预先断言。也许它能，也许暂时不能。本书的论断并不依赖这一点。

我们只需要一件事成立：在院子里的那一整片，已经被接走了；人要去的地方，是院子之上的那一层。至于那一层里，人和 AI 各自站在哪里，是第六编才去碰的问题。

## ◆ 一道考题的三种出法

把「越往上越难考」落到一张试卷上，会更清楚。

先看一位老师想考「学生懂不懂力学」，他可以怎么出题。

最容易的出法，是第一层的题：一个物体质量多少，受力多少，求加速度。答案唯一，评分客观，一百个学生一百份卷子，三分钟批完。AI　做这类题，不需要任何提示。

稍难一点的出法，是第二层的题：给一个实际情境，比如一辆车在斜坡上刹车，让学生自己判断该考虑哪些力，哪些可以忽略，建立一个够用的模型再求解。

这里面开始有了判断，不同的学生会作出不同的取舍，批卷的人得看他取舍得合不合理。有标准，但标准有弹性。AI 做这类题，已经相当出色，只是需要有人看一遍它的取舍。

最难的出法，是第三层的题：「请指出，在力学的这套说法里，哪些东西是被默认了的？如果这些默认不成立，会出现什么样的世界？」这样的题，没有标准答案。

一个学生说「默认了时间处处一样」，另一个说「默认了物体有明确的位置」，第三个说「默认了力可以瞬间传到远处」。哪个答得好？每一个都可能有道理，也可能全是空话。批卷的人如果自己没有想过这个层，是批不了的。

这就是难考的真正含义：不是题目难，是评分者也要在这一层里。一位老师要教学生换层，自己得先上过这层楼，否则他连学生的答案是深刻还是胡说，都没法判断。

这暴露了现有制度的一个真实窘境。我们的考试几乎全靠第一层、第二层的题撑着，因为这些题能批量、能打分、能比较。而在第三层，我们既没有题库，也没有评分细则，甚至没有一个公认的说法，说什么样的回答算「好」。

第十八章会回到这件事，那里要讲的，不是怎么出一道漂亮的第三层题，而是怎么搭起一整套能让这一层被教、被评的东西。眼下只需记住一点：难教难考，不是学生不行，是我们的尺子还没有造出来。

### ◆ 一段对话：学生问「那我还学什么」

设想一位大学老师，在课上讲完这些，有个学生举手。

学生问：「老师，如果下面两层都被 AI 接走了，我们为什么还要花四年学下面那些东西？」

老师答：「你学骑自行车，是为了将来在比赛里赢过汽车吗？」

学生笑了：「不是。」

老师说：「那你为什么要学？」

学生想了想：「因为……学会了平衡，我才知道速度、转弯、摔倒是怎么回事。」

老师说：「对。你学下面的东西，不是为了在下面的东西上赢。是为了有一双脚，站在上面。」

学生又问：「可是上面那层，您自己也说，没有标准答案。那我怎么知道自己练到了没有？」

老师沉默了一会儿，说：「这是一个我也还没有完全回答的问题。我只能说，你练的时候要老老实实问自己：我现在做的这件事，是在一个我已经默认了的地面上做得更好，还是在问这块地面本身？第一种做得再漂亮，也是下面的层。第二种做得再笨拙，也是在往上走。」

这段对话不是真的，却把这一章的意思收拢了：下面的层要学，是因为它是地基；上面的层要去，是因为下面的层已经被接走；而怎么练、怎么判，是一门还没有被好好造出来的功课。这本书的后半部分，要做的就是把这门功课一点一点造出来。

带走的话

**广度是把院子铺得更长，深度是把井挖得更深，两者都没有出院子。 典范就是一层。答案、过程、提问、验收，是这一层里的四个位置，换位置不等于换层。 上一层看下一层，下一层的工作变成了可以批量使用的零件；但没有走过下面的人，看不懂上面的地图。腿退出了运输岗位，但它仍然是地基。 越往上，尺子越软，越难教，越难考。我们的教育和评价制度，是为尺子硬的那几层搭的。**

第 三 编

# 典范的 SDE 发生：显露、差异序列与特征纠缠

这一编六章，是全书的重心。前面只说“换层”，这里要说换到哪里、换的是什么。第 7 章讲典范是怎样一种稳定下来的显露；第 8 章讲旧典范的裂缝，也就是它排不进去的差异；第 9 章讲这些反常怎样纠缠成压力，让旧显露失稳。

后三章讲新的一边。第 10 章讲典范猜想，一句话级别的新划界；第 11 章给出判据，分清什么是典范级、什么只是局部差异，并写明这套判据自己怎样才算错；第 12 章用日心说、广义相对论和量子力学三个案例，把整条序列走一遍。

六章合起来是一句话：**典范不是被发现的，是在差异序列与特征纠缠里显露出来的。**

第 7 章

# 典范作为稳定显露：什么在显露，什么在支撑

先讲一条路。

山脚下有一片草地，草地那头有一个水塘。起初这里没有路，只有草。有一天，一个人从草上走过去打水，踩倒了一溜草。第二天，另一个人也要去水塘，看到那溜被踩倒的草，顺着走了。

又过了些日子，走的人多了，草再也长不起来，露出了土。再后来下了雨，土被踩实，路两边的草长得齐整，路有了边。又过了几年，有人在路口铺了几块石头，有人在路边栽了树。到了这时候，你站在山脚下，看到的是一条路。

现在问一个问题：这条路是谁「发现」的？

没有人发现它。草地上本来没有这条路。也不能说是某一个人「设计」了它，那第一个踩倒草的人，不过是要去打水。可是它又不是假的：你今天去水塘，脚会不由自主地走到路上，因为路上好走，路边有树，路口有石头。你要是偏偏不走路，往草里闯，也不是不行，只是费劲，别人看你也觉得奇怪。

这条路没有在哪个时刻被发现，也没有被谁凭空定下来。它是从许许多多次的走里走出来的，走成了一种稳定的样子。路一旦稳定，就反过来管着后来的人：往哪里走，怎么走，哪里算岔路。

这一章用这幅图讲清第三编的头一个判断：典范，是一个被走成的稳定显露，不是一个被发现的框架。

前面六章，我们一直在说典范是一层，是院子，是地面。那是从外面看。从这一章起，我们要进到里面，问一个新的问题：这一层自己，是怎么来的？它凭什么站得住？

## ◆ 「显露」在大白话里是什么意思

先把一个词讲明白。这个词会在后面一直出现，所以值得多花一点功夫。

「显露」是 SDE 里 Show 的译法。我们用一个日常的情形来解释。

你在黄昏的海边，看见远处有一片模糊的东西。起初你说不清，是云，是岛，还是一艘船？你走近几步，轮廓清晰了，有桅杆，有帆，是船。再走近，你看见船舷上的字，看见甲板上走动的人。这是一件很平常的事，但仔细想一想，其中有一个值得追究的地方：那只船，是你走近的时候才有的吗？

按平常的想法，船一直在那里，你只是走近了才看清。这是「发现」的讲法：事物先有，人后来看见。

可是换一个角度看，「船」这个东西，不是只有远处那一团形状。它是你看到的形状，是你辨认出来的桅杆，是你认出的「这是船不是岛」，这些东西彼此牵连，互相支撑，才合起来成为你眼前这只船。少了光，少了你的视力，少了你曾经见过船的经验，眼前就不会有「一只船」，只会有一团光影。

「显露」说的就是这个意思：一个东西作为它自己清清楚楚地出现在这里，这个「出现」本身，是由很多条件合起来撑住的。显露不是「看见一个本来就在那里的东西」，而是「一个东西成为了它自己，并稳稳地站在这里」。

再用一个更日常的例子。你每天上班路过一家早餐铺。第一天，那里是一堆人、一股油烟、几口锅。过了一个月，你已经说不出那是「一堆人和油烟」了，你说的是「那家包子铺」。

它有了名字，有了边界：哪里是排队的地方，哪里是付钱的地方，什么是「包子」，什么是「馒头」，老板喊一声「三个肉的」大家都懂。包子铺是什么时候「有」的？

不是哪一天宣布成立的，而是日复一日，这些人、这些动作、这些东西，一次次以同样的方式碰在一起，碰得多了，碰稳了，它就显露出来，成了一个你可以指着说话的对象。

这就是显露。它有两个要点，请记住：一，它是稳的，不是一闪而过；二，它是由很多东西互相撑着，才稳的。

## ◆ 典范就是这样的一个稳定的显露

把这个意思搬到典范上。

一个典范，我们前面说过，是一整套：什么是基本的对象，什么样的问题算问题，什么样的回答算回答，怎么检验。如果用刚才的话来说，一个典范就是一个稳定的显露：一个学科共同体共同面对的那一块世界，在这里作为一幅清清楚楚的图景站住了。

我们把它拆开，看这个显露里到底显出了什么。大体上，有三样东西。

第一样，是一套划界。哪些东西在这个学科里算数，哪些不算。比如，在近代力学建立起来的时候，一个被推出去的石块，它的位置、速度、受到的力，是该讲的；它「想要落到地上」的倾向，是不该讲的。

这条线一划，整个学科的话语就有了边。划界不是一句口号，它落实在什么词能进论文、什么现象值得被记下来。

第二样，是一批基本对象。这个学科里，最底下的东西是什么？可以被当成不用再追问的「零件」是什么？在经典力学里，是质点和力；在后来的电磁理论里，是场；在化学的某个阶段，是元素和化合物。基本对象是学科用来搭其他一切东西的积木。它们不是随便挑的，一旦选定，其他所有说法都得用它们来表述。

第三样，是什么算问题。这一条最容易被忽略，却最要紧。一个典范稳定下来以后，它同时就规定了：哪些事是「该解释而尚未解释」的，哪些事是「不用解释」的，哪些事是「不成问题」的。

比如，在刚才说的新划法里，匀速运动不是问题，速度变化才是问题。在旧划法里则反过来。同样一块石头同样一条抛物线，在不同典范里，「哪里需要解释」完全不同。

这三样合在一起，就是我们看到的「一层」。划界告诉你边在哪里，基本对象告诉你里面有什么，什么算问题告诉你里面要做什么。三样东西不是各自独立的，它们互相牵着：基本对象一变，划界跟着变，什么是问题也跟着变。这也是为什么说典范是一整套，换典范是换整个地面，不是换其中一块砖。

所以，当我们前面说「典范是一层」，现在可以说得更具体了：这一层，是一次稳定的显露，显出一套划界、一批基本对象、一份「什么算问题」的清单。它们作为一个整体站在那里，学科里的所有人，都在这个整体里面干活。

## ◆ 什么在支撑：差异序列与特征纠缠

显露是站出来的样子。那么问题来了：是什么撑着它，让它站得住？

回到那条路。路为什么稳？有几样东西在撑着。

一样是「一次次走的不同」。每个人走的时间不一样，拿的东西不一样，脚步轻重不一样，天晴天雨也不一样。可是正是这一次次不同的走，一次接一次，连成了一串。

哪里被踩得多，哪里被踩得少；哪一次有人拐了弯发现更近，哪一次有人滑倒，换了绕道。这一串一串的「不同」，在时间里一环接一环，留下了一个方向、一个走势。

SDE 里把这个叫做「差异序列」。大白话讲：事情不是一团糊的，它里面有一个个能分出来的不同，这些不同按先后、按关联排成一串，一环接着一环，不是随便堆的。

在科学里，差异序列就是一次次观测、一次次测量、一次次对比里分出来的那些「这个和那个不一样」，它们按着一定的顺序积累下来：昨天测到的值，被今天的测量比较，被明天的修正再比较。

另一样撑着路的东西，是「牵在一起」。路不只是一串脚印，它是脚印、路边的树、路口的石头、水塘的位置、人们要去的方向，全部牵在一起。你拿掉水塘，路就没有了意义；你拿掉石头，路的样子就不同了；你把树砍了，下一场雨路就冲坏了。这些东西彼此不是各管各的，它们一动，别的也跟着动。

SDE 里把这叫做「特征纠缠」。大白话讲：一个东西的几个特点，彼此缠在一起，不能随便拆开单独动。拉一根线，别的线也动。在科学里，特征纠缠就是这样的事：质量、力、加速度，这几个量在力学里不是三个独立的数，它

们互相规定，缠成一个整体。你不能只改其中一个而其余不动，一改，别的跟着要改。

现在可以把话说完整了。典范这样的稳定显露，由两样东西长期撑着：一是差异序列，一环接一环地积累；二是特征纠缠，彼此缠着、互相拴住。两者合起来，长期稳定地站着，我们便看见一个典范。

要注意一个词：长期。一次两次不够。路是走了很久才成的，典范也是在很长一段时间里，差异序列一直沿着同一条线接续，特征纠缠一直保持同一种拴法，它才成为一个能让整个共同体放心依靠的地面。短暂的、一闪而过的东西，不成其为典范。

这也解释了另一件事：为什么典范一旦站稳，会那么难动。你要动它，并不是动一个论断，而是要动一条长年累月走成的路，以及路边所有牵连的东西。

## ◆ 为什么不能说「发现」

到这里，有一个自然的疑问：你说典范是走出来的、是稳定的显露，可是我们明明感觉，牛顿定律就是「发现」的，万有引力就是「发现」的，日心说就是「发现」的。大家都这么说，难道都错了？

我们不要急着说谁错。先看一看，「发现」这个讲法，预先放了什么在里面。

「发现」的讲法有一个前提：有一个现成的东西，一直在那里，等着人去找。哥伦布发现新大陆，新大陆本来就在那儿。这个讲法放在地理上，十分贴切。但是把它搬到典范上，就要出事。

因为，如果典范是「被发现」的，那就意味着：有一个现成的、正确的「世界的框架」，在某个地方放着，人类中某位高明的人，一眼看见了，把它拿出来给大家。

于是，历史上所有的典范更替，就成了一个故事：前面的人没看见，后面的人看见了；旧典范是错的，新典范是对的；这里头最关键的，是有没有人「看穿」了。

这一套讲法在第五章已经细讲过，叫做旧本体论：先假定有一个事先就在那里的客体，有一个事先就在那里的主体，主体去发现客体。放在典范的问题上，它会把我们带进一个死胡同：既然框架是现成的，那么上一层的工作，不过是「找到那个现成框架」。

要找，就得靠一个有洞察力的人。于是问题又退回到「怎样培养天才」，而这个问题，几百年里没有人给出过工序。库恩说得清楚典范怎么运转、怎么崩，却讲不出怎么「造」，原因正在这里。

换成发生的讲法，局面就不同了。典范不是一个躺在世界角落里等人去找的框架，它是在差异序列的接续、特征纠缠的拴定里，一点一点稳下来、显出来的。它「有」的方式，像那条路，是走成的。

也许有人会反问：这样讲，那世界本身呢？石头落地，难道也是「走成的」？这里要小心分开两件事。我们不是说石头会不会落是人说了算，也不是说怎么样都行。

我们说的是：「落体」「力」「质量」这样一套可以拿来精确讲话的图景，它的出现和站稳，是有过程的、有条件的，不是躺在那里的。这一点听上去有些绕。后面的章节会借具体的例子，一回一回讲踏实。

目前只需要抓住最小的一点：把典范当成一个稳定的显露，我们就有了一个可以问「它是怎么来的」「它怎么站不住」的对象；把典范当成一个被发现的框架，我们就只能问「谁看见了」。

### ◆ 路是怎么走成的：不要找那个「走第一步的人」

这里要严守一条纪律，因为它最容易在不知不觉中被破坏。

当我们讲那条路的故事，有人会忍不住想：总是有第一个人吧？第一个踩倒草的人，难道不是路的创造者吗？如果没有他，路不会有。

这个想法很自然，但值得推一推。第一个人踩过草，草倒了一溜，可是下一场雨，草又立起来了，这一溜就没了。他自己并没有「创造一条路」，他只是走了一趟。要成为路，需要后来的无数次走，需要地势恰好让人愿意走同一个地方，需要水塘一直在那里，需要人们一直需要水。

第一个人的那一趟，是这条差异序列上的一个起点，不是路的「作者」。换另一个人，在另一天踩下去，也许就是同一条路。如果地势不对，哪怕他走一百趟，也成不了路。

所以我们要把语言倒过来用：不是「某人创造了这条路」，而是「在这片草地的这条序列上，显影出一个第一个走的人」。路的发生，有它自己的条件和走势；那个人，是这场发生里露出来的一个点。

把这个纪律用到典范上，就是本书的一条硬规矩：不写「某位科学家看穿了」「某位大师创造了新典范」。我们写的是：在某条差异序列上，旧的显露排不进去的东西越积越多，特征纠缠的拴法越来越紧，旧的稳定开始松动，新的稳定的样子一点点显出来；在这条序列上，显影出一位提出新猜想的人。

这位提出猜想的人当然重要，他的名字会被历史记住，他的工作也确实艰难。但他是这场发生的显影结果，不是发动者。

为什么要这样讲？不是为了抹杀个人，也不是为了讲得玄。原因很实在。把典范的发生写成「一个聪明人的灵光一闪」，等于又退回了旧本体论，退回到「要有天才」的死胡同。一旦退回去，全书要建立的那门「可以教、可以练、有工序」的学，就无从谈起。

我们要的是一套对任何有条件的人都能展开的工序，不是对天才的崇拜。

## ◆ 稳定显露的两面：看得清，也看不见

路稳下来，有它的好处，也有它的代价。

好处很明显。有了路，后来的人不必再摸索，直接上路；有了路，大家可以约定「在石头那里碰头」；有了路，才有路边的茶摊、路上的车队。

同样，有了一个稳定的典范，一个学科才能分工：有的人专门测数据，有的人专门做计算，有的人专门教学生；一篇论文用不着从头解释「质量」是什么，因为大家站在同一块地面上；一个问题提出来，别人一听就知道它在问什么，也知道什么样的回答算是回答了。常规科学这部庞大的机器能够运转，靠的就是这份稳定。

代价也同样明显。路只有一条，旁边的草地，就不是路了。你在路上走，眼睛看到的是路两旁的风景；你很难再看到，如果从另一个方向走，会有什么。

更要紧的是，当路成了「路」，不走路的做法，就慢慢变得「不正常」了。有人在草里开辟出另一条小径，旁人不会说「这是另一种走法」，而会说「这个人迷路了」。

典范也是如此。一个稳定的典范，一方面让一个学科看得极清楚，另一方面，也把它自己没有显露出来的东西，推到了视野之外。

在牛顿力学的地面上，时间是处处一样、处处同步流过的，这是不用说的；一个东西「同时发生」是什么意思，也是不用说的。这些「不用说」的东西，恰恰是这个显露的一部分，它们不是被发现了才不说，而是在显露的时候，就一并被放到了背景里。

这就是为什么，从典范里面，是很难看清典范的。你要问一个住在院子里的人「你的院子有什么边」，他往往答不出，因为边在他的眼里不是边，是世界的尽头。

只有当某些东西顶到了边上，怎么也过不去，边才被看见。那是第八章、第九章要讲的事：旧典范的裂缝，是差异序列在旧显露下排不进去的东西；反常，是特征纠缠累积成压力，使旧的显露失稳。

## ◆ 回到车与腿：交通也是一个显露

把这一章的意思，放回到全书的绿线上去，会更踏实。

前面我们说，车替代腿，不是「腿走向过程」，而是交通升级。现在可以多说一句：交通本身，也是一个稳定的显露。

想一想「交通」在不同的时代里，是什么样子。在只有腿的时代，交通的基本对象是人和他的脚程；划界是：一天能走多远，就是一个人能到的范围；什么算问题：路上有没有强盗，有没有水喝，天黑前赶不赶得到驿站。

所有这些，都是围绕着「腿」站起来的一幅图景。村与村的距离，是以走一天来算的。一个城镇的大小，是以一个人能在半天里走遍来决定的。这是一个稳定的显露。

车出现以后，一开始它只是在这幅图景里面做事：同样的路，走得更快，运得更多。那是广度和深度的扩张，路更长，车更好。

可是时间一长，事情发生了变化：基本对象变了，不再是人和脚程，而是线路、班次、油耗、里程；划界变了，一个城市的范围，可以是一小时的车程；什么算问题也变了，是堵不堵，是换乘顺不顺，是充电够不够。

这时候，我们说交通换了一层。腿没有消失，它进入了这幅新图景：你要走到车站，走到楼下，走完最后一公里；你要看得懂导航，知道方向。但腿不再是图景的中心，它成了地基。

这样看，「换层」就有了一个更清楚的样子。它不是在旧图景里把某个位置做得更好，而是一幅新的、稳定的图景，在许多条件的撑持下，站了起来。新的图景一旦站稳，旧的图景里那些曾经最重要的东西，就退到了背景里，但没有消失。

科学的典范更替，是同样的形状。我们接下来要问的，正是：这样一幅新图景，最初是怎么站起来的？它的最初一笔，从哪里来？

## ◆ 稳定，不是永远

最后，要把一个容易被忽略的字眼摆在桌面上：稳定。

一个稳定的显露，不是永远不变的显露。稳定，是说在一段足够长的时间里，撑着它的差异序列和特征纠缠，保持了某种平衡。就像一条路，不是铁打的。夏天雨水多，路边的土被冲掉；有一年修了铁路，大家不再去水塘；路上长了草，也就不再是路了。

科学史里，有些典范站了几百年，有些只站了几十年。站得久不意味着它「更对」，它只意味着撑着它的那些条件，在那段时间里没有被推翻。

这对我们很重要，因为这意味着，典范的更替，不是什么神秘的、不可言说的事件。它可以被问：撑着它的东西，是什么？哪一根撑杆在松？哪几条线在越缠越紧？松了以后，新的稳定，要靠什么才能重新站起来？一旦问得出这些，就有了做工序的可能。

前面说过，以前典范发生没有工序，是因为没有必要，常规科学养得活所有人。现在这个理由消失了。要做工序，第一步就是把典范看成它真实的样子：一个由差异序列与特征纠缠长期撑着的稳定显露。看清了这个，才能谈怎么拆、怎么换、怎么造。

下一章，我们先从差异序列说起：旧典范的裂缝，是怎么在里头一点一点长出来的。

## ◆ 三个常见的误会

顺手把容易生出的三个误会摆出来，一个一个理。

第一个误会：「既然典范是走出来的稳定显露，那它就只是一种约定，谁都可以另立一套，怎么都行。」不是这样。路是走出来的，但不是想怎么走就怎么走。

地势在那儿，水塘在那儿，你要去的东西在那儿。一条路能不能成，要看它能不能一次次被走通，能不能让后来的人有用。典范也是这样：一幅图景要站得住，必须能让差异序列顺着它排下去，让特征纠缠被它拴得住。

排不下去、拴不住的，站不起来。所以「走成」不是「随意」，它受着条件的约束，这个约束，就是后面第十一章要讲的判据所依靠的东西。

第二个误会：「那么旧典范就是错的，新典范是对的？」这个问法，仍然带着「发现」的影子，好像存在一张唯一正确的地图。我们不这样问。

我们问的是：旧的显露，在哪些地方撑不住了；新的显露，能撑住什么，又撑不住什么。牛顿力学并没有因为后来的理论而「变成错的」，它在很大一片范围里，至今是极好用的路。只是它的边，被走到了。

第三个误会：「那么科学家做的事，是不是就不重要了？」重要，只是它的位置需要说准。一位研究者的洞察、手艺、坚持，都是实实在在的，没有这些，差异序列不会被一环一环地认出来，纠缠的拴法不会被一点一点地辨清。

我们反对的只是一种讲法：把整场发生，压缩成一个人头脑里的一道光。人是这场发生的显影，是这场发生里最能被看见的那一点，却不是它的全部。

## ◆ 一个小对照：同样的月亮，不同的显露

最后用一个人人见过的东西，把这一章的意思再落一次地。

月亮。对一个住在乡下的老人来说，月亮是挂在天上的灯，有圆有缺，跟着节气走，初一十五要看，中秋要拜。他对月亮非常熟，哪天涨潮、哪天好睡、哪天适合播种，他心里有一本账。这是一个稳定的显露：基本对象是月相和节气，划界是「天上的」与「地上的」，什么算问题，是明天月亮几时出、今晚亮不亮。

对一个学过天文的中学生来说，月亮是一个绕着地球转的球，被太阳照亮一半，我们看到的圆缺，是这个球被照亮的部分转向我们的程度。基本对象变成了球、轨道、光照；什么算问题，变成了：它多快，离我们多远，什么时候会被地球遮住。这也是一个稳定的显露。

两种显露里，月亮都是「看得见的」，都能让人把事情做成。老人的账本，对种地很管用；中学生的轨道，能算日食。两者也不是简单的「一个对一个错」：它们各自撑得住一片事情，各自撑不住另一片。如果你要在老人的账本里算出日食，会发现，要么账本得加东西，要么这件事根本不是账本该管的。

这也就是「换层」在日子里最小的样子：同一个月亮，换了一套划界、一批基本对象、一份什么算问题的清单，这个月亮就成了另一个月亮。不是月亮变了，是显露变了。而显露能变，是因为撑着它的东西可以松、可以换。

## ◆ 一个地质上的例子：地面不动的年代

再用一个地学上的例子，看看「划界、基本对象、什么算问题」这三样，在一个具体的学科里是什么样子。只写有把握的大意，不论细节。

很长一段时间里，人们看大地，是看一块不动的地面。山是怎么起来的？有人说是地球冷却，收缩起皱，像一只干瘪的苹果。海岸线为什么有时候看起来能对上？也许是巧合，也许是以前有过陆桥，后来沉到了海底。

这一整套说法，有它自己的基本对象：固定的大陆，会上升和下沉的地壳。划界是：地质过程是在原地发生的升降与褶皱。什么算问题，是某座山脉的成因、某处地层的年代、某段陆桥当初在哪里。

在这个显露里，大陆彼此远离这种想法，是排不进去的：大陆是固定的，哪来的「走」？即便有人注意到两岸的岩石和化石遥相呼应，这些观察在当时那幅图景里，也只能被当成需要另外解释的小意外，用陆桥去补。

一环一环的观察慢慢积累起来：海底的样子，岩石里留下的磁性的方向，地震发生的地带，火山排列的位置。这些东西，一件一件分开来看，都能在旧说法里另找一个补丁。

可是它们彼此之间是有关联的，越来越多的线索被同一个念头拴住：地面不是一整块不动的壳，而是若干块在缓缓移动的板。当这些线索缠到一起，旧的图景需要的补丁，已经比它能撑起的东西还多。

后来站起来的新图景，基本对象换成了板块，划界变成了板块的边缘和运动，什么算问题，变成了：这块板朝哪里去，边缘是碰撞还是拉开。山脉、地震、火山，从各自为政的问题，变成了同一个图景里的不同表现。

这个例子我们只借来看一件事：同样的岩石、同样的山脉，在两幅图景里是完全不同的东西。每一幅图景，都是一个稳定的显露，都有它的划界、基本对象、问题清单，也都有它的差异序列和纠缠在撑着。

这个转换具体怎么发生，那些观察怎样一环一环排起来，留到第十二章的案例里细写。

## ◆ 这一章为后面留了什么

这一章看起来只是换了一种说法，说典范是显露，是走成的。可是这个换法，为后面的每一章都留下了着力处。

既然典范是由差异序列与特征纠缠长期撑住的稳定显露，那么「旧典范的裂缝」就不再是一个比喻，而是一件可以指出来的事：差异序列上，有哪些「不同」，在旧的划界、旧的基本对象之下排不进去。这是第八章。

既然特征纠缠在拴着整个图景，那么「反常为什么会成压力」就有了讲法：不是某一条数据与理论不合，而是一些越来越多的东西缠在一起，缠到旧的显露承受不住。这是第九章。

既然新的图景是一个新的显露，那么它最初的一笔，不会是一整套完备的体系，而是一句话级别的新划界、新基本对象。这是第十章。判断这一笔是不是够得上「典范级」，是第十一章的判据。而这一切最终落到具体的历史里，是第十二章的案例。

再往后，人和 AI 怎么分工：猜想交给谁、展开交给谁、验收守在哪里，也都要靠这一章里立下的图景才讲得通。只有把典范看作可以被造、可以被验的稳定显露，「让 AI 去展开一个新典范」才有意义。如果典范只是一个等待被发现的框架，AI 能做的，也不过是替人去找。

所以，这一章是第三编的地基。后面的章节，是在这块地基上盖楼。

带走的话

**典范不是一个躺在世界角落里等人去发现的框架，它是一个稳定的显露：一幅站住了的图景，像一条被走成的路。 这个显露里显出三样东西：一套划界、一批基本对象、一份「什么算问题」的清单。三者互相牵着，换其一，整个都得换。 撑着它的，是差异序列的长期接续，和特征纠缠的长期拴定。稳，是因为这两样东西在很长时间里没有松动。 讲典范的发生，不写「某人看穿了」「某人创造了」，写在哪条序列上、什么条件下、显影出一位提出猜想的人。他是这场发生的显影，不是发动者。 稳定不是永远。能说出它靠什么站着，就能问它从哪里开始松，这是后面所有工序的起点。**

第 8 章

# 差异序列：旧典范的裂缝

设想家里有一件穿了十几年的外套。袖口先磨破了，你缝了一道。过两年肘部磨薄了，你贴了一块布。又过两年，衣领起了毛，口袋也开了线，你一处一处补过去。每一次补，当时都说得过去：这里破了，那里补上，衣服还能穿。

可是有一天，你把这件外套摊在床上看，忽然觉得不对。每一处补丁单独看都没毛病，合在一起，这件衣服已经不是原来那件衣服了。颜色对不上，针脚粗细不一，肘部的布比袖子硬，一抬手就绷着。

更要紧的是，你发现这些破处并不是各自倒霉：袖口、肘部、领子，破的地方都是最常被拉扯的地方。这件衣服的剪裁本身，和你穿它的方式，对不上。

这一章讲的就是这件事。一个典范用久了，会有一些东西「对不上」。最初它们零零散散，看上去互不相干，每一处都有一个说得过去的补法。可是它们加起来，指向的是剪裁本身。

本章要讲清楚三件事：所谓「对不上」到底是怎么回事；为什么身在典范里面的人，整体上看不见这些裂缝；以及补丁这条路，是怎样一步一步越走越窄的。

## ◆ 先把「差异序列」说成大白话

前面第 7 章讲过，一个典范可以看成一个稳定的显露：一整套看世界的方式，在某个时期里稳稳地摆在那里，大家照着它看、照着它算、照着它问问题。这套显露不是凭空摆着的，它是靠一串一串的「差异」撑起来的。

什么叫差异？就是「这个和那个不一样」。月亮和星星不一样，今天的潮水和昨天的潮水不一样，这根杆子称出来的重量和那根杆子称出来的不一样。

世界上到处是差异，但并不是所有的差异都被典范用上了。典范像一张筛子，它决定哪些差异算数、哪些不算数；算数的差异，又按一定的次序排好队：先看哪个，再看哪个，哪个是主要的，哪个是次要的。

这样排好队的一串差异，我们就叫它差异序列。你可以把它想成超市货架。货架上每一格放什么，是有规矩的：饮料放这排，日用品放那排，每一格有每一格的位置。新到的货，大多数一看就知道往哪一格放。典范下的日常工作，大半就是把各种新到的货放上货架。

所谓「排不进去」，说的就是有一件货拿在手里，货架上哪一格都不合适。放进饮料区，它不是饮料；放进日用品区，它又不是日用品。你可以把它先搁在收银台旁边，等有空再说。每个超市里都有这么一两个角落，堆着「暂时没处放」的东西。

要注意，这里说的是货架，不是眼力。不是理货员眼神不好，也不是理货员偷懒，而是货架的格子就那么多、就那么划分，新来的东西恰好落在格子之间。格子是旧显露划出来的，所以「排不进去」这件事，从一开始就是旧显露自己的事，而不是哪个人的过失。

## ◆ 反常最初长什么样

库恩把这些东西叫作「反常」。听起来很吓人，好像一出现就要天翻地覆。事实上，反常刚出现的时候，一点也不吓人。

它常常只是一个数对不上。一次测量，结果和计算差了一点点。一台仪器，在某个条件下总是多出一个读数。一个现象，在大多数情况下听话，偶尔不听话。一个本来应该一样的东西，换个地方量，却不太一样。

这些东西最初散落在不同的角落，在不同的人手里，在不同的年月里出现。做天文观测的人遇到一点，做仪器的人遇到一点，做理论计算的人遇到一点。它们之间没有标签，没有谁说「这几件事是一回事」。

拿水星来说。十九世纪的天文学家用牛顿的万有引力算水星的轨道，算出来的结果和长期观测比，有一小点点偏差：水星轨道最近太阳的那一点，一圈一圈地慢慢挪动，实际挪的比算的多一点。

这个偏差很小，小到只有长期、精细的观测才看得出来。当时的人怎么处理？他们的第一反应，是照牛顿的路子去找原因：大概是还有一颗没发现的行星，在水星里面绕着太阳转，它的引力拉着水星。

后来的人，甚至给这颗假想的行星起过名字。这个办法在先前是成功过的：天王星轨道的偏差，就曾经引导人找到了海王星。所以这个补法非常合理，甚至可以说是典范里最漂亮的一种补法。

这里请你留意一点：在那个时候，水星的偏差只是水星的偏差。它和光速有什么关系，和引力的本质有什么关系，没人把它们放在一起想。它是货架上一件暂时没处放的货，被放在「待找的行星」这个收银台旁的角落里。

## ◆ 三个角落：噪声、误差、谜

常规科学对付这样的零散差异，其实有一套很成熟的办法。大体说来，是三个角落，分别叫噪声、误差和谜。

先说噪声。凡是测量，都会有随机的抖动。同一个人用同一把尺子量同一块木头，每次读数都略有不同。一个读数偏了，第一个想法是：这是噪声。再测几次，取平均，偏差消失了，那就确实是噪声。这个办法非常有用，没有它，科学根本做不下去。如果每一次抖动都当作新发现，那全世界的实验室都会陷入疯狂。

再说误差。误差和噪声不同，它不是随机的，而是有来由的：尺子热胀冷缩了，天平没调平，仪器的零点漂移了。一个偏差如果能追到某个具体的、已知类型的毛病上，把毛病修好，偏差就消失，那它是误差。找误差也是一门手艺，一位老练的实验者，很大一部分功夫就在这里。

最后说谜。如果一个偏差既不是噪声，也追不到仪器的毛病，那么它就可能是一个「谜」。库恩用过「解谜」这个词来形容常规科学：典范告诉你，这个问题一定有解，而且解的形状大致是什么样的，你的任务是把它找出来。

拼图游戏就是这样：你知道盒盖上有一幅完整的图，每一块都有位置，找不到位置只是你还没找到。

这三个角落有一个共同的好处：它们都不要求动摇货架。噪声说，货没有问题，是搬运时抖了一下。误差说，货没有问题，是秤有毛病。谜说，货没有问题，只是还没找到那一格。三种办法都是在保护货架，而货架确实值得保护，因为绝大多数时候，这三种说法都是对的。

这一点必须先说，不然后面的话会被读歪。常规科学把异常当噪声、当误差、当谜，绝大多数时候是正确的，并不是顽固或者胆小。一个反常，十个里有九个甚至九十九个，最后确实是噪声、误差或谜，货架完好无损。如果一个体系遇到每一个异常都怀疑自己，它就没法积累任何东西。

所以，「排不进去」最初并不显得是排不进去。它显得是「还没排进去」。

## ◆ 为什么在典范里面，看不见裂缝的整体形状

现在要回答这一章的核心问题：既然这些差异到最后会纠缠成大麻烦，为什么典范里面的人一直看不出来？

先说一个最直接的原因：每一道裂缝都是从典范内部的某个具体问题进来的，而典范内部的工作，是按问题分工的。研究水星的人看水星，研究光的人看光，研究仪器的人看仪器。每个人面前摆的都是一个局部的、带着典范标签的问题。局部的问题，用局部的办法去补，局部看起来就是好的。

回到外套的比喻。袖口的补丁，是在袖口这个位置上补的，肘部的补丁，是在肘部那个位置上补的。每一次补的时候，你的注意力都在那一小块布上，那一小块布补好了，任务就完成了。想要看出「这些破处都在拉扯最多的地方」，你得把衣服整件摊开，退后一步，看分布。而日常的缝补，从来没有这一步。

用 SDE 的话来说：裂缝的整体形状，是一条差异序列，而这条序列是跨着很多个局部问题才连得起来的。每一个局部问题，只拿到序列中的一小段。旧典范的分工，正好把这条序列切成了许多互不照面的碎片，分别放进各个角落。

还有一个更深的原因。典范不光是一套理论，它还是一套「怎么量、怎么算、怎么算对」的规矩。仪器是按典范校准的，数据是按典范的范畴整理的，连「什么叫异常」这个判断，也是典范给的。换句话说，用来发现裂缝的眼睛和尺子，是旧显露自己配的。

这有点像用一把按旧地图刻度做成的尺子，去量新长出来的地方。尺子上没有那一格，于是量出来的读数，总是被「四舍五入」成最近的旧刻度。偏差没有被看见，不是因为它不存在，而是因为量具的刻度里没有它的位置。典范越成熟，量具越精良，这种「自动归整」就越顺滑，顺滑到你根本感觉不到它在归整。

再往下一层：典范内部的人，也并非不愿意看。他们的训练、他们的晋升、他们手里的经费和时间，全部建立在「货架是对的」这个前提上。这不是说他们居心不良，只是说，一个在货架里长大、靠货架谋生的人，当他看见一件放不进去的货时，最自然的反应是「我再找找」，而不是「货架划错了」。

这里不要把它读成人的毛病。那条差异序列在这个位置上显影出来的，就是这样一位「再找找」的人。换了别人站在这个位置，同样会再找找。

## ◆ 越补越难看：补丁的账是怎么算的

前面说，补丁在每一处单独看都是合理的。但补丁有一笔账，是单独看看不出来的。

把托勒密式的天文体系想一想。这套体系用一个又一个圆套圆（本轮）来解释天上的行星为什么忽而前进、忽而后退。观测越精细，越发现行星的走法和最简单的圆套圆对不上，于是后来的人就在圆上再加圆，在位置上再挪一个点。

每一次加，都能把对不上的那一处补好。这套办法持续了很长很长的时间，也并不愚蠢：它能用，能预测，能做历法。

可是请你注意「每补一处」带来的变化：体系变复杂了。复杂本身不是罪，但复杂有两个后果。第一，每一个补丁，只对自己那一处负责，补丁与补丁之间没有道理上的联系，它们只是各自凑合着不打架。

第二，体系越复杂，下一次再有新的偏差，要补的时候，就越得小心翼翼，生怕补这里弄坏了那里。补的成本在一点一点上升。

这就是补丁的账：每一笔单独看都划算，累积起来，总账一路向下。更要命的是，这笔总账在典范内部没有科目可记。典范自己的评价标准是能不能预测、能不能自洽，补丁都能达标。没有一个统一的地方，让人看见「已经补了多少处、补得多费劲」。这个账，只有从外面看，才记得出来。

再说一次，这个账也不是谁偷懒造成的。每一个补丁，都是典范内部最勤勉、最老练的工作，是常规科学做得好的证明。恰恰是做得好，才把裂缝补得看不见。

## ◆ 排不进去，到底排的是什么

读到这里，可能有人要问：你一会儿说「货架」，一会儿说「差异序列」，到底「排不进去」排的是什么？

我们可以把话说得更准一点。典范下，世界并不是一堆东西，而是一串被排好的差异。比如，在牛顿的显露下，物体的运动差异，被排成「力的大小、质量、加速度」这样一串；在这一串里，哪个变，哪个不变，哪个是原因，哪个是结果，都有位置。

每一个新的观测，会被拆成差异，往这一串里放。放得进去，就是「解释了」；放不进去，就是「还没解释」。

所以「排不进去」的意思是：这个差异，在这套序列里找不到它该站的那一格——不是格子没填满，而是格子的划法里，没有给它留位置。你可以把它硬塞进一个相近的格子，那就是补丁；你可以说它是噪声，那就是把它扔掉；你也可以把它挂在门口，说「以后再说」，那就是谜。

要区分的是：一个差异「还没排进去」，和一个差异「原则上就排不进去」。前一种，是格子够用，只是没放；后一种，是格子本身的划法有问题。从一个单独的差异看，这两种情况长得一模一样。区别只有在许多差异合起来的时候，才会显出来——这就是下一章要讲的纠缠。

这里再回到绿线。腿是很好的运输工具，在腿的时代，人们对「远」的感觉，是按腿来刻度的：一天的路程、一个驿站的距离。

后来有了车，有人发现有些路程，用「几天的腿脚」来量，量不出意思来；也有人发现，用腿来量的「远」和实际的「远」不一样。这些零散的发现，一开始也只是「路好像比想象的短」，没人立刻意识到，是整个「用腿脚刻度来量路」的尺子换了。

## ◆ 同一种补法，一次成功，一次落空

上面提到水星，还有一个地方值得多说几句，因为它正好说明，为什么在典范里面分不出「谜」和「裂缝」。

天王星的轨道，也曾经和计算对不上。当时的补法，是假定天王星外面还有一颗没被看见的行星，在拉着它。照这个想法去算，去找，真的找到了：海

王星。这是常规科学最漂亮的一次胜利，货架不但没有动摇，反而因为这次成功更加稳固：看，偏差不是货架的错，是货还没找全。

于是水星的偏差出现时，同一套办法顺理成章地被拿出来用：再假定一颗行星。这一次，办法落空了。人们找了很久，后来的观测始终没有给出这样一颗行星。

但请注意，落空并没有立刻变成「货架错了」。落空首先变成的，是又一个补丁：也许那颗行星太小，也许是一团物质，也许是别的什么没算进去的东西。每一种说法，在当时的条件下都没法被彻底排除。

这件事的要害在于：从典范里面看，「谜」和「裂缝」在开头长得一模一样。海王星的路和水星的路，起点完全相同：一个轨道偏差，一个典范内的补法。分岔发生在很久以后，而且分岔本身，还得靠别的东西来显出来。没有谁在起点就能判断，这个偏差是通向海王星的，还是通向别处的。

所以，不要把「当时的人没有看出来」读成他们不够聪明。在那条差异序列上，当时显影出来的那些人，做的恰好是当时唯一说得通的事。他们没有错，只是裂缝的整体形状，当时还没有成形。裂缝要靠时间、靠更多的偏差一件件落下来，才慢慢有了形。

## ◆ 一个读数的旅程

不妨跟着一个小偏差，走一遍它在典范内部的旅程，看看它是怎样被一路「照顾」得妥妥帖帖的。

它先出现在一台仪器的屏幕上，是一个比预期多出来的小数字。操作的人先怀疑仪器，重新校准。数字还在。他怀疑样品，换了一份。数字还在，但小

了一点。于是他把它报给了组里的前辈。前辈看了一眼，说这种情况以前见过，多半是某种已知的干扰，加一个修正项就行。修正项加上，数字落回了范围内。

这个修正项，随后被写进了实验的方法部分，成了「标准做法」的一个小脚注。后来做同类实验的人，读到这个脚注，照着加上，就不再看到那个偏差了。他们甚至不知道，曾经有过一个偏差。修正项成了方法的一部分，偏差成了方法的「前史」。

再往后，有人要总结这一领域的常用方法，把这个修正项收进了手册。手册再被写进教科书。教科书教给下一代的，是一个干净的、没有偏差的世界：只要按手册操作，读数就是对的。

这一路上没有人隐瞒，没有人造假，每个环节都在做认真的工作。可是那个最初的偏差，已经被层层转换成了「做法」。它从一个等待解释的差异，变成了一个不需要解释的规矩。裂缝没有消失，它被封进了方法里。这就是为什么后来的人，要翻回很早的实验记录，才能看出当初的那个「不对」。

## ◆ 一个家常的对话

不妨设想一个实验室里的小场景，来看这三个角落是怎么运转的。

甲拿着一份数据过来，说：「这批样品，有一组读数偏高，高出大概一成。」

乙看了看：「先查仪器，是不是上次校准之后动过？」

甲说：「查了，没动。」

乙说：「那是不是样品批次的问题？换一批再测。」

过了几天，甲回来了：「换了批次，还是偏高，不过这次只高了半成。」

乙想了想：「半成，那大概是环境温度的影响，夏天的实验室比冬天热。你把温度记下来，修正一下。」

甲照做了，修正之后，读数基本落回了正常范围，只剩下一点点残余，在误差带的边缘。两个人都松了口气，把这件事记进实验笔记，写上「温度修正，已解决」。

这段对话里，没有一个人犯错。每一步都是老练的、负责任的做法。偏差先被当作仪器的误差，再被当作样品的个别情况，最后被当作环境的影响，一层层剥下去，只剩下一点残余。那点残余，被留在误差带的边缘。

关键就在这一点残余。它被记进了笔记，被写成「已解决」。可是，如果过些时候，另一个实验室里的另一个人，在完全不同的样品上，也留下了这么一点方向相同的残余；再过些时候，第三个地方，也有；这三点残余，各自都在误差带边缘，各自都被写成「已解决」。它们分别躺在三本笔记里，彼此没有照面。

这就是裂缝在典范内部的样子：不是一道看得见的口子，而是许多本笔记里，许多个「已解决」后面，那一点点不肯完全消失的尾巴。

## ◆ 老房子的电路

再换一个日常的例子。住过老房子的人，大概都知道老式电路的脾气。

房子刚建好的时候，电路是按当时的用电想象设计的：几盏灯，一台电视，一个冰箱。后来家里添了空调，你在墙上再拉出一根线；添了微波炉，又

拉一根；厨房多了电饭煲、电热水壶，插座不够了，买了一个多孔的接线板，接在原来的插座上。每一次添，都能用，没出问题。

可是某一年夏天，你同时开了空调、烧水、用微波炉，「啪」的一声，跳闸了。你去合上闸，嘀咕一句「今天用得太猛」。过几天又跳了一次，你想，「是不是接线板质量不好」，换了一个。再过些时候，某个插座有点发烫，你想，「插头松了」，换了插头。

每一次跳闸、每一次发烫，都有一个局部的解释。你并没有糊涂。直到电工来了，打开配电箱看了一眼，说：「这根主线太细了，整个屋子的用电，早就超出它的设计了。」

电工看到的，是一个你看不到的东西：所有这些局部的现象，其实在同一根线上汇合。跳闸、发烫、偶尔灯光闪一下，彼此看上去不相干，实际上是同一件事的不同露头。

我们需要留意这个例子里的一个细节：电工并不比房主「聪明」，他只是站在一个位置上，那个位置能同时看到所有的线是怎么接到一起的。房主在每一个插座旁边，只能看到那个插座。位置不同，看到的形状就不同。

这也提醒我们：典范内部的人看不见裂缝的整体形状，不是他们的能力问题，而是位置问题。反过来说，在这条差异序列上，某个时刻，也会显影出一位「拉开配电箱的人」。他看到的东西，并不是他比别人高明，而是这条序列已经长到了那一步：足够多的线汇到了同一个地方，配电箱里的东西，终于挡不住了。

## ◆ 为什么不能一开始就把它们当裂缝

你也许会冒出一个念头：既然裂缝最后都会显出来，那最好一开始就盯着这些零散的偏差，早点怀疑货架，岂不更好？

这个想法听上去很有道理，实际上走不通，原因有三个。

第一个原因是数量。任何时候，一个成熟的典范下面，都挂着大量的「待解」：数不清的小偏差，数不清的对不上。其中绝大多数，最后都老老实实回到了货架上。

如果每一个都当裂缝去追，就等于把全部力气花在九十九个假警报上，去换那一个真警报。这笔买卖，在没有 AI 的年代是亏的：人手有限，每个人一辈子能追的偏差，不过寥寥数个。

第二个原因是，在没有别的货架可以对照的时候，「货架错了」这句话没有内容。你说旧的不对，那新的是什么？没有新的，「不对」只是一句抱怨。就像你看着满是补丁的外套说「这件衣服有问题」，可是衣柜里没有第二件，你明天还得穿它出门。

所以裂缝再大，旧显露在很长时间里依然是唯一能用的显露，大家仍然靠它吃饭、靠它造桥、靠它预测日食。这也是为什么，即使裂缝已经相当明显，旧典范仍然在被使用，而且用得很好。

第三个原因，是裂缝的「形状」本来就是慢慢长出来的，不是提前就在那里等着被发现的。一个偏差落下来，它和另一个偏差之间是什么关系，要等两个都落下来才有意义；三个落下来，才有三角形；四个五个落下来，才看得出这些点的分布往哪个方向歪。不是有人拖着不看，而是形状需要材料，材料需要时间。

这三条合起来说明一件事：早期的「不当回事」不是失误，而是旧显露维持自己稳定时的正常功能。一个显露若对每个小偏差都敏感得发抖，就稳不住，也谈不上被用来工作。反过来，稳得太好，裂缝又要积累到很大才被注意。稳定的好处和代价，是同一件事的两面。

这里和腿的故事又接上了一个头。步行时代，路的好坏、远近，是靠腿一点一点试出来的。走了许多年，有些人注意到，有几条路总是比预想的难走，有几处总要绕。

这些零散的抱怨，都没有让人怀疑「用腿赶路」这件事本身。要等到别的交通方式出现，对照之下，这些抱怨才连成一片：原来它们都是「腿的极限」的不同露头。

## ◆ 一件旧衣服，到底什么时候该换

把外套的比喻再往前推一步，问一个很实际的问题：什么时候，才算「该换了」？

没有一个固定的日子。有人补到第三块布就换了，有人补到第十块还穿着。但如果你留心这些人的说法，会发现他们换衣服的理由，几乎从来不是「又破了一处」。

那一处，破了就补，谁都会。让人下决心的，是另一类念头：我已经不知道下一处会破在哪里了；我补这里的时候，总担心牵到那里；我每次出门前，都要先检查一遍这件衣服。

这些念头有一个共同点：关心的对象从「这一处」变成了「这件衣服整体」。注意力从局部移到了整体，这件事本身，就是典范的裂缝开始显出来的标志。不是破的处数多了，而是看的单位变了。

这个变化，不是靠某个人下决心发生的。你仔细想想自己换衣服的那一天，多半不是你「决定」了什么，而是那一天你又在补丁上花了一段力气，忽然觉得这件事不值得再做下去。决定是顺着这些积累的感觉，自然落下来的。

放到典范上，意思是：裂缝显形，并不依赖谁的洞察。它依赖的是，差异的数量、牵连的程度、补法的成本，这几样东西合在一起，把「看局部」这个习惯本身变得不划算了。在这一刻，整体才第一次进入视野。

## ◆ 教科书里没有裂缝

还有一个地方，裂缝被藏得特别深：教科书。

教科书讲一个典范，讲的是它最干净的样子。例题是挑过的，能算出来的；实验是选过的，做得成的；历史是修剪过的，一路向前的。学生学完一本教科书，得到的印象是：这个领域的问题，要么已经解决，要么正在被解决，剩下的只是时间和勤奋。

这不是教科书的恶意。教科书的任务，就是把一套显露传给下一代，让他们尽快上手干活。要传得快，就得把那些没有放进格子的东西先搁在一边。

可是这样一来，每一代新人进门的时候，看到的都是一个没有裂缝的世界。裂缝留在了老人的笔记里、留在了学术会议休息室的闲谈里、留在了论文最后一节「局限与展望」的两行字里。

于是一个人要在典范里长大，就必须先学会不去注意那些东西。这听起来有点残酷，但它是训练的一部分：一个初学者如果对每一个「不太对」都停下来追问，他永远学不完课程。能熟练地绕过去，是「成熟」的标志之一。久而久之，绕过去变成了本能，连「绕」这个动作本身，也不再被感觉到。

这也是为什么，当别人问「你们这个领域有什么问题还没解决」，典范里的人常常能列出一长串，可是那一串里，几乎都是「在这个框架里待解的谜」。真正会动摇框架的东西，往往不在这张清单上，因为清单本身是按框架写的。

### ◆ 今天的特殊处境：补得更快，看得更难

最后讲一个今天才有的情形，它和本书的主线直接相连。

AI 接管了典范内的工作之后，补丁的速度会大大加快。一个偏差出现了，AI 能在几分钟里列出十几种典范内的补法，逐一测算，挑出最合适的一个。噪声、误差、谜这三个角落，处理起来都是批量的、廉价的。这当然是好事：常规科学的效率提高了许多倍。

但请注意它对「裂缝」的影响。补得快，意味着货架保护得更好；单个偏差在没来得及引起任何人注意之前，就已经被妥善安置了。

以前一个偏差要拖上十年二十年才补好，在这十几年里，它至少还会被许多人议论、被反复提起，留下一点痕迹。现在它也许一个下午就进了「已处理」的清单。裂缝没有消失，只是每一道都被及时地填平了。

另一方面，AI 也能做另一件事：它能把成千上万个零散的偏差汇总起来，列在同一张表上。这等于把整件外套摊在了床上。可是，摊开之后，看出「破处的分布指向剪裁」，这件事是不是仍然在典范内的能力范围内？

这个问题我们放到后面的章节去谈。眼下只说一点：把差异放在一张表上，和看出这张表上的差异在「一个整体」里排不进去，是两回事。前者是汇总，后者是对货架本身的追问。

## ◆ 把这一章收一收

这一章讲了一件容易被误会的事：旧典范的裂缝，是怎么形成、又怎么被藏起来的。

形成的方式是：典范用差异序列给世界排队，新的偏差落在格子之间，排不进去。藏起来的方式是：噪声、误差、谜三个角落，加上按局部分工的问题、按旧显露配的量具、以及一个在货架里长大的人群，把每一个偏差都妥善安置，每一笔单独看都对。

读完这一章，请你保留一个印象：裂缝最初不是一道口子，而是许多处「已解决」后面，不肯完全消失的那一点尾巴。它们彼此之间有没有关系，在第 8 章这里还不知道。要知道这一点，就得进入下一章：看这些尾巴什么时候开始纠缠。

带走的话

- 差异序列，就是典范给世界排的那串队；「排不进去」，是队伍里没有给这个差异留位置，不是这个差异不够格。 - 反常最初只是零散的偏差，被放进三个角落：噪声、误差、谜。这三个办法绝大多数时候是对的，保护货架并不是固执。 - 裂缝的整体形状，在典范里面看不见：问题是分工的，量具是按旧显露配的，每个补丁单独看都说得过去。 - 补丁的账，只有从外面才记得出来：每一笔划算，总账一路向下。越补越难看，不是谁偷懒，是补得越好越难看出来。 - 今天补得更快了，裂缝因此更容易被及时填平。下一章讲的是：这些散落的差异，在什么情形下不再是各过各的，而是纠缠到一起。

第 9 章

# 特征纠缠：反常如何纠缠成压力，使旧显露失稳

设想楼上的邻居在装修，你家天花板上出现了一小块水渍。你找人看了看，说是管子接头有点渗，补了一下，水渍干了。过了半年，卧室墙角起了一点潮，你想，大概是梅雨天。又过了一阵，厨房的墙皮鼓了一块，你想，大概是涂料质量差。

三件事，三个解释，每个都说得通。直到有一天，一位老师傅来换水龙头，顺口问了一句：「你家这几处，是不是都在同一面墙的同一侧？」你一下子愣住了。你把三个位置在脑子里连起来，它们真的在一条竖线上。这条线后面，是一根藏在墙里的老水管。

这就是纠缠最朴素的样子。三处零散的毛病，一旦发现它们共用同一个「后面」，就不再是三件事，而是一件事的三个露头。这件事的分量，也不再是三处小毛病的相加，而是整根水管的老化，以及与之相连的、你家整个水路的设计。

上一章讲的是差异序列怎样在旧显露下排不进去，零散地躺在各个角落。这一章讲的是下一步：这些零散的差异，什么时候开始彼此牵连，牵连到什么程度会变成压力，压力又怎样使一个稳定了很久的显露站不住。

同时要讲清楚一件很实际的事：纠缠有两种，有的只是噪声碰巧撞在一起，有的才是真压力。分不清这两种，既会错过真的革命，也会造出许多假的革命。

## ◆ 「纠缠」是什么：两个差异不再各过各的

先把词说清楚。

特征纠缠，说白了是这样一回事：几个差异之间有了牵连，动一个，另一个也跟着动；解释了一个，另一个反而更难解释；把一个压下去，另一个就从别处冒出来。它们不再是一把散沙，而像是一团绞在一起的线，你想抽出其中一根，会拽动旁边的几根。

请注意这跟「相关」不是一回事。两个数据同涨同落，当然可以叫相关，但那只是统计上的并排。纠缠指的是更紧的东西：它们在同一套显露里，被迫以彼此依赖的方式出现。想解决其中一个，就得连带处理另一个。用日常的话说，就是「牵一发而动全身」。

举个厨房里的例子。你想把汤做得更咸一点，就加盐；加了盐，汤变咸了，但同时也盖住了原来的鲜味；想把鲜味找回来，就加料；加了料，汤的颜色又变了。咸、鲜、色，是三个差异，它们在同一口锅里相互牵扯。你没法只调一个。这就是三个差异处于纠缠中的状态。

放到典范里，这意味着：当几个反常彼此牵连，补其中一处的办法，会让另一处更别扭，那它们就已经不是各自独立的「谜」了。它们是同一个问题的几个侧面。

为什么要强调是「特征」的纠缠？因为典范里的差异，不是裸的数字，而是带着特征的：这个现象是什么样子的，哪里不同，怎么不同。

特征彼此绞在一起，意思是：这些反常所携带的「样子」，在旧显露里没法被分开安置。你想把这个特征放进这一格，那个特征放进那一格，可是它们总要一起出现，一起消失，一起变化。

## ◆ 噪声也会「凑巧」撞在一起

这里必须踩一脚刹车。几处零散的差异碰在一起，大多数情况下并不是真压力。

人有一种很强的本能：看见几个点，就想把它们连成线。夜里看星星，我们把几颗毫无关系的星星连成了勺子、猎户、北斗。这是人的看法，而不是星星之间真有绳子。同样，一堆偏差放在一起，只要你想连，几乎总是能连出一个「规律」来。

所以，面对几处差异的牵连，第一件事是问：这是凑巧，还是真绞在一起？下面给几个家常的问法，不是公式，只是老练的人在心里掂量的几件事。

第一问：换一个地方，它还在吗？如果同样的偏差，在完全不同的装置、不同的样品、不同的时间、不同的人手里，都以大致相同的方式出现，那就不像是凑巧。凑巧的东西，换个场合多半就没了。水渍如果只在你家出现，也许是你家的事；整栋楼同一侧都有，就是楼的事。

第二问：它们指向同一个方向吗？噪声是四面八方乱抖的，今天偏高，明天偏低，平均下来为零。真正的纠缠，会有一个共同的偏向：总是往同一边偏，偏的方式有章可循。三处水渍如果一处在东墙、一处在西墙、一处在地板，它们更可能各有各的原因；如果排成一条线，就值得追问。

第三问：压下去一个，别处会不会鼓起来？这是最有分量的一问。真正的噪声，处理掉一个，对别处没影响。真正的纠缠，处理掉一个，往往会让另一个更严重，或者让一个原先没有的新偏差冒出来。这是因为它们共用了同一个「后面」，你堵住了一个出口，压力就从别的出口出来。

第四问：补丁之间相互容得下吗？每一个补丁，单独看是合理的，但把两个补丁放在一起，它们是不是在抢同一个参数、同一个假设？如果补丁A的合理性，要求某个量取这个值，补丁B的合理性，偏偏要求同一个量取另一个值，那么两个补丁虽然各自成立，合起来却打架。这是典范内部自己的东西在互相拉扯。

这四个问法，没有哪一个是决定性的，没有哪一个能一锤定音。它们只是让你从「看见几个点」，往「看见一团绞住的线」那边多走几步。

## ◆ 压力是什么：补丁的账开始往回算

假定四个问题的回答，大致都指向「是真的」。接下来，纠缠就会产生一种东西，叫压力。

压力不是一个神秘的概念。把一根尺子两头用力弯，它不会立刻断，而是弯成一道弧线；你越弯，它越想弹回来，弹回来的力道就是压力。压力的特点是：它来自于对某个形状的偏离，偏离越大，积累越多；它不一定有可见的外在表现，但你能在弯曲的程度上感觉到。

放到典范里，压力的来源是前一章说过的补丁的账。每多一个补丁，体系就复杂一点；纠缠的差异越多，补丁之间的牵制就越多。一个补丁修好一处，却使另外两处更难补，那么你要补那两处时，就得再加两个补丁；两个补丁又牵动更多的地方。这条链子延长的速度，并不是线性的。

可以想象一个工人，在修一面漏水的墙。他堵住了东边的洞，水就从西边冒出来；堵西边，水又从下面渗出来。这时候，他每多堵一个洞，就要多费一倍的劲，才能压住整个局面。压力指的，就是这种「为了维持原样，要花的力气越来越大」的状态。

一个简单的表，可以把「噪声」和「真压力」的区别放在一起看：

| 看什么 | 凑巧的噪声 | 真正的压力 |
| --- | --- | --- |
| 换个场合 | 多半消失 | 依然出现 |
| 偏向 | 四面乱抖，平均为零 | 总偏向同一边 |
| 压下去一个 | 别处无动于衷 | 别处鼓起来 |
| 补丁之间 | 各管各的，互不碍事 | 彼此争同一个假设 |
| 补的成本 | 补完就完 | 越补越贵 |

这张表不是用来打分的。它只是说：真压力不是靠一个指标冒出来的，是靠一整套情形一起出现，才显得像样。

## ◆ 压力累到哪一步，旧显露开始站不稳

压力一点一点累积，旧显露一开始照样站得很稳。这一点要说清楚，因为很多人以为「出了反常，旧理论就倒了」。事情不是这样的。

想一想一座桥。一座桥设计的时候，留有余量。车多了一点，它不会塌；车又多了一点，它稍微多颤一下；再多一点，桥面出现一道细纹，工人补了；再多，桥墩有点沉降，工人又加固了。

这一路上，桥一直站着，一直在用，每一步都有人说「没问题」。可是余量是有限的，每一次加固，用掉的是余量的一部分。到了某一个点，余量见底了，下一辆车开上去，桥就不再是「颤一下」，而是真的撑不住。

注意最后一句：撑不住的那一刻，不是因为这辆车特别重。这辆车和前面的车没有任何区别。它只是恰好落在余量用完的那个点上。把桥的垮塌归咎于

最后那辆车，是一个很常见的误解；真正的原因，是此前所有的车、所有的加固，把余量一点点吃光了。

典范的失稳是一样的。一个显露，之所以能稳，是因为它能把各种差异安置得下，安置的方式足够简洁，足够自洽。压力累积的过程，就是这份「能安置得下」的余量被一点点用掉的过程。每一个补丁都用掉一点；每一处纠缠都用掉一点。余量见底之前，显露照样稳；见底之后，不需要谁来推，它自己就晃起来了。

这就是本章最要紧的一点：失稳不是被某个人推翻的，而是显露自己撑不住。

请对比一下流行的说法。流行的说法是：某位科学家提出了新理论，推翻了旧理论。好像旧理论是一尊塑像，被一位勇士一斧子劈倒。可是在 SDE 看来，顺序是反的。

是旧显露先撑不住——它的差异序列排不下了，纠缠的压力使它自身的稳定条件松动了——然后，在这个已经松动的地方，才有机会显影出一位提出新猜想的人。他的出现，是松动的结果之一，而不是松动的原因。

用一个更生活化的比喻：冰面上的裂纹。春天，河面的冰一天比一天薄，底下的水流一天比一天急。冰上有人走过，没事；有人走过，也没事；直到某一天，一个孩子跑过去，冰裂了，他掉了下去。

事后大家会说「是那个孩子踩裂了冰」。但是如果不是他，下一个人、再下一个人，也会掉下去。冰早就不行了。孩子的脚步，只是让「早就不行」变成「看得见的不行」。

## ◆ 失稳是什么样子的

压力累积到余量见底，典范就进入了一段失稳期。这段时间，身在其中的人会觉得很怪。因为失稳不是「旧的一下子没了」，而是出现了一种说不清的混乱。

第一，旧显露还在用，而且用得很好。造桥的人还是照旧公式算，预测日食的人还是照旧历法算，课堂上还是照旧教材讲。它不是废了，是在大部分地方依然稳稳当当，只在某些地方开始「不灵」。这使得失稳期里的人，一边在使用，一边在怀疑，两件事同时存在，并不矛盾。

这跟腿和车的故事很像。当车出现之后，腿没有马上退出历史舞台。很长一段时间里，短途的、近处的路，依然靠腿；车只在某些路段上显得有优势。大家一边走路，一边问：「是不是该换个办法？」换，还是不换，没有一个时刻，全体人一起转过身去。

第二，吵架多了。同一套数据，不同的人给出不同的解释，而且彼此都觉得对方的补法不够自然。以前大家有共同的货架，吵的是「这件货该放哪一格」；现在，有人开始吵「这格是不是划得对」。

吵架的层次变了。以前吵架，吵完双方都愿意用同一把尺子去判输赢；现在吵到最后，往往发现双方用的尺子已经不一样了。

第三，补丁的风格开始分化。有的人坚持把所有偏差往旧显露里塞，补丁打得越来越巧；有的人开始试着换一种看法，哪怕换得粗糙；还有人试图折中，在旧显露上做大手术。这些做法，在失稳期里同时存在，看上去乱糟糟的，但这些乱，正是显露在寻找新的稳定点的痕迹。

第四，有些过去不算问题的事，忽然变成了问题。当一个显露稳定的时候，有一些事是「不言自明」的，没有人觉得需要解释。到失稳期，这些不言自明的东西，会被一一翻出来重新问：为什么是这样？必须是这样吗？这类问题，在稳定期被认为是「不专业」的，是外行的发问；到了失稳期，它们成了最紧要的问题。

这四种迹象，合起来是一种感觉：脚下的地不那么实了。但这种感觉，同样不是某个人发明的，而是这条差异序列和这团纠缠本身，在这个阶段的样子。

## ◆ 水星故事的后半截

回到水星。在前一章里，水星轨道的小偏差，被放在「假想行星」的收银台旁。这个故事的后半截，正好说明了纠缠与失稳是怎么回事。

水星的偏差，并不是十九世纪末唯一一件对不上的事。当时的物理学里，还有关于光的传播、关于运动的描述、关于电磁现象的一连串别扭。它们在不同的领域、被不同的人处理。起初，每一处都有各自的补法：光需要一种叫以太的介质，运动要相对于以太来算，一些实验里的零结果，被各种巧妙的办法解释过去。

后来的人回头看，会觉得这几处别扭，其实在同一根线上。它们指向的是：旧显露里关于空间、时间、运动的某些基本划分，撑不住了。到了二十世纪初，这些零散的东西，被一个新的显露重新排了队，其中一个结果，就是水星轨道的小偏差，不再是一个孤立的谜，而是新显露自然带出来的一个后果。

这里我们要警惕一种讲法，就是「爱因斯坦看穿了这一切」。那种讲法把一整条纠缠的线，压缩成一个人的灵光。

本书的讲法更接近事实：在这条差异序列上，旧显露的余量早已被一件件用光。光、运动、电磁、引力方面的别扭彼此牵连，牵连到旧的划分自己站不稳。在这个不稳的地方，显影出了一位提出猜想的人。他的猜想怎么来的，是下一章的话题。这一章只说：猜想之所以能提出，是因为地面已经松了。

这个故事里的细节，细到每个人的年份和实验，不在这里展开；第 12 章的案例会专门讲。这里只取它的形状：零散的偏差，各有各的补法，后来发现它们共用一个「后面」，旧显露的余量见底，然后才是新的显露。

## ◆ 两种常见的误判

纠缠与压力有两种最常见的误判。它们方向相反，代价都很大。

第一种误判，是把噪声当压力。人一旦知道「反常可以累积成革命」，就容易兴奋，看见什么都想往革命上靠。几处凑巧的偏差，被他当成「旧理论的丧钟」；一个没有重复出来的结果，被他当作「新时代的开端」。每年都有无数这样的宣告，大多数在重复实验面前悄无声息。

这类误判的坏处不只是浪费。它还会损伤人们对真压力的识别力：喊多了狼来了，真狼来了，反而没人信。所以前面那四个问法，其实是一道保险，防止人把凑巧当作必然。

第二种误判，是把压力当噪声。这是更隐蔽的一种。一个显露稳定了很久，人们已经养成了一套反应：偏差来了，就找补法。这套反应太熟练了，熟练到真压力来了也被当作「又一个谜」。补法一个接一个，成功率还不低，于是人们以为一切在掌控中。直到余量见底，才发现掌控的只是表面。

这类误判的根源，不在谁愚笨，而在于稳定本身的惯性。一个显露稳了几十年、几百年，它的补法也成熟了几十年、几百年。成熟的补法，是对旧显露的忠诚，也是对真压力的屏障。

因此，辨别噪声与压力，不是一次性的鉴定，而是一个持续的、带着不确定的掂量。没有一个仪表盘，能在任何时刻告诉你「现在到了临界点」。这件事，本书不假装有一把尺子。能说的是：这四个问法让掂量不那么盲目，而最后的判断，要靠后面发生的事来验证。

### ◆ 放回日子里：一个人的「补丁账」

这些话，其实不只是科学的事。你自己的生活里，大概也有过类似的时刻。

比如一个人长期用一套方法管理自己的时间：清单、日历、提醒。刚开始很管用。后来事情多了，清单越来越长，他给清单加了分类；分类多了，又加了优先级；优先级冲突了，又加了标签；标签太多，又装了一个软件来管标签。每一步都在解决眼前的问题，每一步都很合理。

有一天他发现，自己每天花在「管理清单」上的时间，比做事的时间还多。他把这件事讲给朋友听，朋友说：「你是不是该换个思路了？」他说：「可是我每次加的功能，都是有用的啊。」

他说得没错。每个补丁都有用。问题不在任何一个补丁，而在整套方法的格子，已经装不下他现在的生活。他的生活里有很多事，在原来的「清单—日历—提醒」这个格局里，没有位置。那些没位置的事，一开始是「等一下再说」，后来是「总是漏掉」，再后来是「每周日晚上焦虑」。

注意这里的转折：他不是被某个「朋友的一句话」说服的。朋友那句话，只是在一个早已松动的地方，起了一个提示的作用。真正使他转变的，是那么多零散的难受，终于绞到了一起，压力大到旧办法撑不住。

## ◆ 纠缠有几种长相

纠缠不只一种长相。认得几种常见的，遇到的时候心里有底。

一种是「共享后面」型，就是开头那条水管。几个看上去不相干的毛病，其实是同一个隐藏原因的露头。这一类，一旦找到那个「后面」，几处毛病会一起解开。

一种是「互相拉扯」型。两个补丁都有道理，但合起来矛盾。比如一边需要某个量尽可能大，另一边需要这个量尽可能小，补法A和补法B只能各取其一。这一类，单独看每个补丁都漂亮，凑在一起却露出显露内部的不一致：有两条规矩，从一开始就不太兼容，只是平常没碰上。

一种是「越补越多」型。每补一处，就生出两处新的别扭。补丁像生了崽，一个变成两个，两个变成四个。这一类的特点是，工作量不是随着偏差的数量增长，而是随着补丁的数量增长。人们常用「越抹越黑」来形容这种局面。

还有一种是「悄悄改了意思」型。为了让一个补丁成立，需要稍微改一下某个词、某个量的含义，改动很小，没人在意。再来一个补丁，又改一点点。

改着改着，同一个词在不同地方已经指着不同的东西。大家还在用同一个名字，说的却不是同一件事。这一类最隐蔽，因为它不显现为矛盾，只显现为「说不到一块去」。

这四种长相，常常混在一起出现。对读者来说，不必背它们的名字，只需要留意一个共同的信号：想解决一个问题，却不得不先去碰别的问题。当「先去碰别的」成为常态，就说明这些问题已经缠在一起了。

## ◆ 这不能靠投票，也不能靠一个人说了算

有人会问：既然分清噪声与压力要靠掂量，那到底谁来掂量？是不是要大家投票，或者请德高望重的人拍板？

两条路都不太行。

投票不行，是因为典范内部的人，用的是同一套量具、同一套训练。大家一致说「这是噪声」，有可能恰恰因为大家都在同一个显露里，对同样的东西视而不见。多数人的意见，反映的是显露的稳定程度，不是压力的真假。一个成熟的典范，会得出非常一致的「这是噪声」的意见，一直到它撑不住的那一天。

拍板也不行。拍板的人再有威望，也站在同一个显露里。他的经验、直觉，都是旧显露养出来的。他的「这没什么」，很可能是整个显露的「这没什么」，由他的嘴说出来而已。在这件事上，权威既是最可靠的守护者，也是最难自我怀疑的位置。

那么靠什么？靠的是上面那几个问法能够被重复检验、被别人拿去试的东西，以及随后发生的事情。如果真是压力，它会一次又一次以新的面貌出现，把更多的差异卷进来；如果是噪声，它会在重复里消散。换句话说，最终的裁判是时间和后续的差异，而不是任何一个当下的表决。

这听上去有点消极：岂不是只能等？也不完全是。你可以主动去找：去找那些方向一致的残余，去找彼此打架的补丁，去问某个词在不同地方是不是已经不指同一件东西。这些找的功夫，是在典范内就能做的，也是 AI 做起来又快又全的。找到之后怎么办，是另一回事，后面各章会讲。

## ◆ 一个没人当回事的重复

压力最初冒头的时候，常常不是一次惊人的发现，而是一个不起眼的重复。

某个残余，在一个实验室里出现了。没人当回事。过几年，在另一个国家的另一个实验室里，用不同的方法，又出现了类似的残余。两边没有联系，各自写进了自己的笔记。

再过几年，第三个领域里的人，在算另一个量的时候，发现要凑出结果，需要一个「调节因子」，这个因子的大小和方向，恰好与前两处的残余能对上。

到这一步，还谈不上有什么结论。但已经有了一件事：三个互不相干的地方，在同一个方向上留下了同样大小的尾巴。这时候，如果有人把三本笔记并排放在桌上，问一句「这三个是不是一回事」，这个问题就不再是无中生有的联想，而是有分量的。

你看，这里没有「看穿」。没有一个人在某个瞬间突然洞察了一切。有的只是三本笔记里的三条尾巴，一年一年地长，终于长到能够在同一张桌子上相遇。这张桌子，有时是一次会议，有时是一篇综述，有时只是两个人在走廊里的闲聊。尾巴相遇的那一刻，在这条差异序列上，显影出了一位把三本笔记摆在一起的人。

## ◆ 失稳不是坏事

最后，为失稳说句公道话。

在日常语言里，「不稳」是个贬义词。但对一个显露而言，失稳意味着它不再把世界的差异全部压平在同一个格局里了。那些被噪声、误差、谜三个角落收容的东西，开始重新被看作差异。一个太稳的显露，是看不见新东西的；一个刚好开始松动的显露，才第一次让人看见格子之外有什么。

这也是为什么，历史上每一次大的转折，都发生在显露开始松动的时候，而不是它最稳的时候。最稳的时候，货架上的一切都有位置，人们安心地做着手上的事。松动的时候，人们不安，吵架，试错，但也在这个时候，新的稳定点才有机会出现。

不过，松动只是提供了机会，并不保证一定出现新的稳定点。有些松动，最后以补丁的大翻修而平息，旧显露改头换面继续活下去；有些松动，长久地悬着，没有结果。

新的显露能不能发生、以什么样子发生，不由压力本身决定。压力只是把地面松开了，在松开的地面上，种子长不长，还要看那一笔猜想，以及后面的验证与选择。这是下一章要进入的地方。

## ◆ 回到 AI：压力还看得见吗

最后，把这一章放回本书的主线。

AI 接管了常规科学之后，补丁的速度大大加快，补丁的成本大大降低。这对压力的积累，有两层相反的影响。

一层是：补得更好，更快，使得压力更不容易以「难受」的形式被感觉到。以前一个偏差要折腾一个研究组好几年，在这几年里，这个组都会觉得别扭；现在一个下午就处理完了，没有人别扭。压力还在，但是「别扭的感觉」被廉价地消去了。

另一层是：AI 能把散在各处的偏差汇总起来，把纠缠的线拉出来给人看。如果有人去问，AI 可以在几十万篇文献里，找出方向一致的残余，找出彼此争抢同一个假设的补丁。这是一件极其有用的事，是把外套摊在床上的事。

这两层加在一起，意味着在 AI 的时代，压力的积累和压力的显现可能脱节。积累得很快，感觉得很慢；材料齐全，却没有人去问「这些材料合起来指向什么」。这一问，是在典范内做事的人通常不去做的事，也是人上升到更高一层之后要面对的事。

不过，到这里我们还只是讲到旧显露撑不住了。撑不住之后，新的稳定点从哪里来？它的最初一笔长什么样？这是下一章的话题。

带走的话

- 纠缠，是几个差异不再各过各的：动一个，牵一片；补一处，鼓一处。它比「相关」更紧，因为它们共用同一个「后面」。 - 绝大多数零散的偏差凑在一起，只是噪声碰巧撞车。要分清，可以问：换个场合还在吗？是否总偏向同一边？压下去一个别处会不会鼓起来？补丁之间会不会打架？ - 压力，就是为了维持原样而要花的力气越来越大。显露的余量一点点被用掉，直到见底。 - 失稳不是被某人推翻，而是显露自己撑不住。撑不住的地方，才显影出一位提出猜想的人，他是结果，不是原因。 - 把噪声当压力，会天天喊革命；把压力当噪声，会一直补到塌。没有仪表盘，只有掂量。AI 让补得更快、汇总得更全，却不替我们问出那个更高一层的问题。

第 10 章

# 典范猜想：新显露的最初一笔

老房子的墙上有一道裂缝，住了多年的人都知道。夏天它细一点，雨季它宽一点，有人拿腻子抹过，隔一两年又裂开。房子里的人各有各的办法：有人在裂缝上贴一张画，有人说这是房子老了的正常现象，有人每年按时补一次，补得又快又好。

这些办法都有用，也都有个共同点：它们都默认这堵墙还是原来那堵墙，这座房子还是原来那座房子。

某一天，有个人蹲在裂缝前看了很久，说了一句话：这不是墙裂了，是这两边本来就不是一堵墙。

就这么一句。没有图纸，没有算式，连这句话自己是不是对的，说话的人心里也没底。可是从这句话起，整座房子要用另一种方式来数：不是一堵墙有一道缝，而是两堵墙并排站着，中间有条缝。

这一章讲的，就是这样的一句话。库恩的书里有它的位置，叫「革命」，可是革命是怎么开始的，他只描述了开始之后的样子。本书在前两编已经说过，这个开头以前不需要工序，现在需要了。

第三编前面几章讲了典范是一个稳定的显露，讲了差异序列怎样在旧典范里排不进去，讲了反常怎样纠缠成压力。这一章接着问：压力有了，裂缝有了，新的显露从哪里开始？

## ◆ 一句话有多小

先说清楚「典范猜想」到底是个多大的东西。

它不是一套理论。理论是后面的事，是几十年、几代人在它底下盖起来的楼。它也不是一个模型，模型是楼里的一间屋子。

它甚至不是一个假说，假说一般指的是在既有框架里对某件事的一个可检验的说法，比如「这种药能降血压」「这颗星的亮度是周期变化的」。这种说法再大胆，也是在已经划好的格子里填一个数。

典范猜想是划格子本身。它回答的是两个更靠底下的问题：世界上有哪些东西算数？这些东西之间的分界线画在哪里？

用本书前面的话说，旧典范是一个稳定的显露：一个世界被看成了这个样子，哪些差异被当成要紧的、哪些被当成噪声，哪些东西被当成基本的、哪些被当成派生的，都已经定了。典范猜想就是对这个显露的一次重新下笔：换一个划界，换一个基本对象。

举几个大家听得懂的例子，只取最粗的轮廓。

「地球不是宇宙的中心，地球和别的星一样在动。」这一句话里，划界变了：原来天上和地上是两个世界，现在是一个世界。基本对象也变了：原来「天」是一类东西，「地」是另一类，现在只有「在空间里运动的天体」。

「大陆不是钉死的，它们在漫长的时间里慢慢漂。」这一句里，基本对象变了：原来地壳是一块铺好的地板，现在是一些会移动的板块。

「时间不是处处一样地走，而是跟着运动和引力变。」这一句里，原来那个背景式的、对谁都一样的时间不再算基本的了。

你看，这几句话都很短。短到可以印在一张便条上。但每一句都要求整个科学的账本重新记：哪些是原料，哪些是产品，哪些是误差，哪些是主角。

所以说它是「一句话级别」。不是说它容易，而是说它的形状是一句话：一个新的划界，加上一个或几个新的基本对象。下面所有的工作，都是在这句话底下展开的。

这里要先防一个误会。有人会说，那我也能一天想出一百句这样的话：「意识是宇宙的基本粒子」「万物都是信息」「一切都是振动」。确实，这样的话可以批量生产，AI 生产得更快。

一句话的形状，不等于它是个够格的典范猜想。哪些够格、哪些不够，是下一章要专门讲的事。这一章先讲清楚：真正够格的那一类，它是怎么冒出来的，它冒出来的时候长什么样。

## ◆ 它不是从旧典范里推出来的

最容易想到的说法是：新典范是从旧典范里一步步推出来的，只是前人没推到那一步。

这个说法听上去很顺，实际上走不通。理由不复杂。

推理是在一个系统里做的事。你有一组基本对象，一组规则，往下推，推出的东西都在这个系统的范围之内。这正是常规科学的看家本领：给定典范，往下推，推出新的结果、新的预测、新的应用。推得再远，推出来的也是这个典范的新房间，不是新房子。

而典范猜想要说的，恰恰是旧系统里没有的东西。旧典范里没有「板块」这个位置，你就推不出板块。旧典范里没有「时间会随运动改变」这个位置，你就推不出它。不是推得不够勤，是这个对象在旧账本里根本没有账页。

有一个小实验，现在人人都能做。拿一个当时最强的 AI，把某个旧典范的全部材料、全部规则、全部已知反常交给它，然后说：请从中推出下一个典范。它会给你很多东西：精致的修正，合理的扩展，对反常的各种说法。

其中有些说法看上去相当大胆。但仔细看，它们多半是旧基本对象的新组合，或者是在旧划界上挪一挪线，或者是把别的领域已经现成的框架搬过来。这些都是有价值的工作，都属于典范内的好活。可它们是推出来的，所以它们没有跳出去。

这里有一个很有意思的对照。如果新典范真的能从旧典范推出来，那么旧典范里最强的人，也就是最熟悉它、推得最远的人，应该最先推到。可是历史上的情形，常常恰恰相反。

最熟练的人，往往最早把反常消化成「已知问题的变体」。熟练本身就是在加固那个稳定的显露：越熟，越知道哪些东西「不算数」，越知道哪些差异该当成噪声。

这不是说内行不重要。恰恰相反，下一节会说，裂缝的形状只有走过路的人才摸得出来。这里只是要分清楚两件事：对裂缝的体察，是靠沉在旧典范里的功夫；而对裂缝的命名，也就是那一笔新的划界，不在旧典范的推理范围里。前者是地基，后者是新楼层。

用车与腿的比方来讲：腿再强，一路走下去，走到的也还是这条路的前方。要换一种交通，不是靠腿继续往前迈，而是有人在路的尽头发现，这里该

是一个港口，要换船。腿走出来的路，是知道哪里是尽头的唯一办法；可是「换船」这一下，不是下一步迈腿。

## ◆ 它也不是拍脑袋

说完不是推出来的，另一头也要堵住：那它是不是凭空来的灵感？一个人洗澡的时候忽然想到，就成了？

这个说法在通俗读物里很流行，也最危险。它有两个坏处。第一，它把这件事神秘化了，让读者以为这是天赋的事，没法教、没法练。第二，它让人忽略了猜想背后那一大段看不见的功夫。

猜想背后的功夫，本书把它叫作对裂缝形状的长期体察。

什么叫裂缝的形状？回到差异序列的说法。旧典范下面，现象一件件摆出来，有些顺顺当当排进去了，有些排不进去。排不进去的那些，不是零零散散的，它们之间有样子。

比如，它们总是在某一类情形里出现；它们总是以某种方向偏；补一个修正项能解决这一个，却让另一个更糟；两个看上去不相干的领域，各自有一批排不进去的东西，细看居然是同一个轮廓。

这些「样子」，就是裂缝的形状。有人把它叫作「问题的纹理」。一个老木匠摸一块板，手指过去就知道纹路往哪边走，哪里会崩。他不一定说得出理由，可他的手知道。

裂缝的形状，只有在一件事上长期泡过的人才摸得到。泡的方式很朴素：把反常一件件亲手弄清楚，把每一个补丁亲手试一遍，看它在哪里补得住、哪

里补不住；去读那些被判为「噪声」而扔掉的数据；去问那些被认为不值得问的小问题。这些工作，没有一件是在跳出典范，全都是典范内的苦功夫。

所以，新猜想不推自旧典范，却离不开旧典范。它不是从旧典范的逻辑里长出来的，而是从旧典范的失效处长出来的。

就像一个搬运工，扛了一辈子货，他对这座码头的承重极限的了解，比任何设计图都更具体；某一天他说，这里要建新码头，不是因为他突然聪明了，而是因为这座码头在他肩上已经「说」了很多年话。

这里按总纲的口径，要把主语摆正。上面这些话，不要读成「一个有洞察力的人看穿了裂缝」。

更准确的说法是：差异序列在旧显露里排不进去的部分，随着积累，自己呈现出了形状；这个形状在某处、对某些长期泡在裂缝边上的人，变得越来越清晰；一句话的新划界，就在这条序列的缝上显露出来。而那个说出这句话的人，是这场显露里显影出来的一位。他是结果，不是发动机。

这个区别不是文字游戏，它有后果。如果你认定是一位天才发动了猜想，你就会去找天才、等天才、崇拜天才，教学上就只能祈祷。如果你认定是裂缝的形状在显露，你就会去问：怎样让更多的人更早地接近裂缝？

怎样让差异序列的形状更早地被摆出来、被看见？怎样让那些被当作噪声扔掉的东西留下来？这些都是可以做、可以教的事。后面第五编会回到这里。

再说一遍，这一段不是在贬低提出者。一位提出者的功夫是真功夫，苦也是真苦。只是这份功夫的性质，是让自己成为裂缝的一个好的显影处，而不是凌空发出一道指令。

## ◆ 最初一笔往往很粗糙

好，假设一句猜想这样冒出来了。它长什么样?

人们对新典范有一种事后的想象：一出来就光芒万丈，旧的一下子显得笨拙。真实的情形几乎相反。最初的那一笔，常常粗糙，常常漏洞百出，常常在一些旧典范做得很好的地方，表现还不如旧典范。

这一点需要想明白原因，因为它对后面的判据很要紧。

旧典范是什么？是几十年、几代人、成千上万的人在同一个划界底下打磨出来的。它的每一个角落都被擦过，每一个常见情形都有现成的算法，每一个已知的反常都有一个说得过去的交代。它是一座住了很久的房子，家具齐全，水电通畅。

新猜想是什么？是一张便条，上面写着：这里应该是另一座房子。便条背后没有家具，没有水电，连墙都还没砌。拿它跟旧房子比居住舒适度，当然比不过。

日心说的最初形态就是这样。哥白尼把太阳放在中心，这一笔改了宇宙的划界。可是在具体预测上，它并没有比托勒密的体系准确多少，它自己也要用到圆上套圆的办法，还有许多地方显得繁琐。

很多人当时看了，觉得这不过是换一种复杂的画法，何必呢？这个评价在当时并不是愚蠢的。如果你只用「预测准不准」这一把尺子，当时的日心说确实占不到便宜。

板块这件事也类似。大陆会动的想法，在二十世纪初被提出时，一个主要的难处是：没人说得出是什么力量推动大陆。这个缺口很大，足以让许多严肃

的人把它搁在一边。后来的海底观测把另一些东西摆到了台面上，这个猜想才一步步有了站得住的机制。

量子这件事也是。最初把能量切成一份一份的那一笔，连提出者自己都不太乐意当真，更像是算账时不得不用的一种办法。它是在后来的一连串对撞里，才被认真对待成世界的基本样子的。

这些例子说明什么？说明一个新典范刚冒头时，用旧典范的成绩单去评它，评出来必然是「不如」。这不是它不好，是比较的前提不公平：旧典范的成绩单上，考的科目正好是旧划界下划出来的科目。新划界的科目，旧典范的卷子上根本没有。

由此要得出两个看似矛盾、其实并不矛盾的结论。

第一个结论：不能因为新猜想粗糙就拒绝它。粗糙是它出生的样子，就像婴儿不会走路，不是婴儿的缺点。

第二个结论：也不能因为新猜想粗糙就赞美它。很多粗糙的东西，粗糙得没有任何后文。「粗糙」不是美德，只是一个不能用来否定它的理由。这两条加起来，就逼出了下一章的问题：既然不能靠成绩单评，也不能靠气势评，那靠什么？下一章给的就是这个。

在这之前，还要看一件事。为什么最初一笔常常解释力不如旧典范，但仍然值得被留下来？

因为它在做一件旧典范做不到的事：它把旧典范里排不进去的那一串差异，用一种新的方式，放进了同一个位置。哪怕只放进去了几个，哪怕放得很勉强，只要这几个原本是「无家可归」的，它们就有了一个共同的家。这个共同的家，是旧典范里完全没有的。这一点，在判据一里会细讲。

我们在这里先记住一个朴素的感觉：最初一笔的价值，不在于它已经解释了多少，而在于它让原本散落的东西，第一次有了可以靠在一起的地方。

## ◆ 新显露的一笔，是怎样落下的

现在回到「怎样落笔」。前面说它不是推出来的，也不是拍脑袋来的。那它是怎么落下的?

这里不能编一个机制，也不能拿一个公式来冒充工序。本书只说有把握的几条，其余的留给后面的章节和案例去补。

第一条：落笔之前，有一个很长的「憋」。

憋的意思是，同一批排不进去的差异，在很多人手里反复出现，反复被补丁处理，反复被判为「暂时的困难」。这个阶段，裂缝并不叫裂缝，它叫「还没解决的问题」。参与的人各自忙着各自的补丁，彼此之间也许并不知道，他们其实在围着同一个形状打转。

第二条：落笔的那一刻，很少是孤立的。

回头看历史，一个新划界出现的前后，常常有好几个人在不同的地方、用不同的话，摸到了相似的东西。有人摸到了一半，有人摸到了另一半，有人用的是完全不同的语汇，后来才被发现说的是同一件事。

这个现象，如果用「天才发动」来解释，就很麻烦：为什么天才总是成堆地出现? 如果用「裂缝的形状在显露」来解释，就顺了：形状在那里，在条件成熟时，它会在好几个显影处同时被摸到，只是有一处最先落成了一句清楚的话。

这不是要抹掉某一位的贡献。那一位把一句话落得清楚、站得住、传得开，是实实在在的功。但本书的口径是：这一句话，是在差异序列的缝上显露出来的，那位提出者，是随之显影的人。

第三条：落笔的那一笔，几乎总是「看上去不够严肃」。

因为它必须以旧典范的语言无法直接表达的方式出现，所以它初出时，往往带着一点「说不清」的气质。一个比喻，一张草图，一个看上去不合规矩的假设，一个被同行认为是「哲学而不是科学」的说法。它要等很久，才被翻译成可以计算、可以检验的形式。

这一点后面讲 AI 与验收时还会再提：把一个粗糙的新划界翻译成严谨的形式，这一步，恰好是 AI 擅长的展开工作的起点，但翻译时最大的危险，也出在这里，翻着翻着，新划界被翻回了旧语言，成了换皮的旧典范。

第四条：落笔之后，要有对撞。

一句话落下来，它只是一个猜想。它要变成典范，要靠一连串与现象的对撞：它预言了什么旧典范不会预言的东西？这个预言被现象接住了吗？这些问题，前面说了，是下一章的主角。

## ◆ 放回日子里：一个小例子

典范离我们的日子太远的话，换一个小一点的场合看。

设想有一位语文老师，教了多年作文，发现一件怪事：班上总有几个孩子，按标准评分细则打分并不高，可是他们的文章读起来就是让人记得住；而另一些孩子，每一项细则都满足，文章却读完就忘。

老师最初的反应，是在评分细则里加条目。加「新颖性」，加「真情实感」，加「独特视角」。细则越来越长，可那几个孩子的文章，仍然排不进去：他们好的地方，恰恰是细则里每一条都够不着的地方。

老师又去查论文，看别人怎么说，读到的多半是「创造力难以量化」「需要整体印象评分」，这是在为裂缝贴一张画。

后来有一天，老师换了个问法：我的细则，是不是一直在量文章的「零件」？而那几篇好文章的好，是不是在「零件之间怎么咬合」上？

这就是一句典范猜想的雏形：作文好坏的基本对象，不是一条条的零件，而是零件之间的咬合方式。划界从「逐项合格」变成了「整体咬合」。

这句话刚出来的时候，粗糙得很。什么叫咬合？怎么量？跟旧细则比，能不能打出更稳定的分？答案多半是：不能，至少一开始不能。它的解释力，在所有「旧细则打得好」的地方，不如旧细则。可是那几个「无家可归」的孩子，第一次有了一个共同的位置。

这个小例子不是要说老师发明了新典范，作文评价离典范级还差得远。它只是要把过程的形状摆出来：长期的憋，补丁的失败，一个换问法的瞬间，一句粗糙的新划界，以及它在旧成绩单上的不如。读者可以拿这个形状，去对照自己身边的任何一个领域。

注意这里的主语。不是「老师看穿了」，而是这几个孩子的文章构成了一串旧细则排不进去的差异；这串差异的形状，在老师手里长期被碰到；一句新的划界，就在这串差异的缝上显出来，老师是随之显影的那一位。

## ◆ AI 来了，这一笔的处境变了什么

最后说说 AI 时代的处境，这是全书后半部分的伏笔。

前面说了，最初的猜想粗糙，需要有人在很长时间里替它撑腰，把它从一张便条展开成一座房子。这个展开的工作，从前极贵：要有一整个团队，几十年的时间，大量的计算与推演。所以历史上，大部分粗糙的猜想死在了展开之前，不是因为它们错，而是因为没人付得起展开的成本。

AI 把这个成本压低了。一张便条，现在可以在很短的时间里被展开成一个相当完整的框架，包括定义、推论、与已知现象的对照、可能的实验。这意味着两件事。

一件是猜想的存活门槛大降。粗糙不再那么致命，一句不成熟的新划界，也有机会被充分展开后再受检验。以前因为太贵而没被试过的路，现在可以试了。

另一件是，可以同时展开多个。同一条裂缝，可以并排下几个不同的猜想，各自展开，再比较。这在从前是奢侈品。

但要看清楚，这两件事都发生在「猜想之后」。猜想本身，那一笔，是不是 AI 也能给，本书不替它断言，留到最后一编去谈。

本书不预设那一笔永远是人的专利，也不预设 AI 已经能给。要紧的是，无论那一笔从哪里来，下面这个问题都在：这一笔，值不值得被展开？它是典范级的，还是只是一个局部的新鲜想法？

因为展开变得便宜，这个问题反而变得更尖锐。展开的成本一低，就会有大量的猜想被造得头头是道，其中大部分只是局部的差异，换了身漂亮衣裳。

如果没有一套判据，人会被一堆「完整的」东西淹没。这就是下一章要回答的。

## ◆ 同一条裂缝上，不止一种下笔

还有一个常被忽略的事实：同一条裂缝，可以有好几种下笔的办法。

裂缝只是说，旧显露在这里排不下去了。至于新的显露长成什么样，裂缝本身并不唯一地规定。就像一堵墙裂了，你可以说两边本来是两堵墙，也可以说是地基下面有一条暗沟，还可以说这堵墙当初就不该砌在这里。这几种说法，划界不同，基本对象不同，将来盖出来的房子也不同。

历史上看，一个领域走到困难的关口，往往同时冒出几种互不相让的新说法。其中多数后来不见了，有一个活了下来。

事后回望，人们很容易说「那个活下来的当然是对的，别的都是歧途」，这是后见之明。在当时，没人知道哪一种会活。每一种都粗糙，每一种都带着自己的一群人，每一种在旧典范的成绩单上都不及格。

这件事对本书的论证有两层意思。

第一层，它再次说明，新猜想不是从旧典范里唯一地推出来的。如果能推出来，结论应该只有一个。现实是好几个，这恰好说明旧典范的材料不足以决定新的划界，决定它的是别的东西，是后面要讲的与现象的对撞，以及一个领域在漫长时间里的选择。

第二层，它给了我们一个方向：既然下笔不唯一，那么在 AI 把展开成本压低的今天，最合理的做法不是押一个，而是并排下几个，各自展开，各自去

撞现象。这是第 15 章的内容，这里先点一下，免得读者觉得「一条裂缝一个猜想」是天经地义。

### ◆ 提出者的功夫：做裂缝的好显影处

既然猜想是在裂缝上显露、提出者随之显影，那么「提出者」是不是就完全被动，什么也不用做，坐等就行了？当然不是。

「显影」这个词，借自照相。一张底片放在那里，光来了，影像要在底片上出现，底片得干净、对光敏感、没有被别的东西提前曝光。底片和底片是不同的：有的灵敏，有的迟钝；有的被前一次曝光糊住了，再也显不出新的影。

提出者的功夫，就是让自己成为一张好底片。具体有几件朴素的事，每一件都可以练。

头一件，把反常留下来。多数人的习惯是把排不进去的东西清掉：这个数据点是仪器误差，那个现象是个例，这种说法是外行话。每清一次，裂缝的形状就少一块。好的底片会有一个专门的抽屉，收着那些「暂时说不通」的东西，不急着处理，也不急着丢。

第二件，别把补丁当答案。补丁是有用的，可是补丁用久了，会让人忘记它是补丁。好的底片记得每一个补丁欠着一笔账：这个修正项，是把什么差异硬塞进去了？这笔账攒多了，就是裂缝的轮廓。

第三件，让不同领域的反常见面。前面说过，裂缝的形状有时要把两个不相干的角落并在一起才看得出来。读别的行当的文献，听别的行当的抱怨，不是为了借方法，是为了看一看，别人那里排不进去的东西，和你这里排不进去的东西，是不是同一种形状。

第四件，别太快把粗糙的想法关掉。很多初起的想法，一经自己或同行的审视，立刻显得幼稚，就被扔了。可是前面说过，最初一笔的样子，本来就是幼稚的。能不能给粗糙的想法一点活着的时间，是对自己判断力的一种克制。

这四件事，没有一件是天赋，全是习惯。而且你会发现，它们几乎都是「地基」上的功夫：对本行的熟，对反常的耐心，对补丁的诚实。腿的功夫一样都不能少。没有走过路的人，不会知道这条路在哪里到头。

## ◆ 一段对话：这一笔，到底算什么

最后用一段虚构的课堂对话，把这一章收一收。人物是编的，话也是编的，只为说明思路。

一个学生举手说：老师，您说典范猜想不是推出来的。可我们做研究的，难道不是靠推理吗？

老师说：推理你每天都在做，而且做得很好，这没有问题。我问你，你手里最近一个没做出来的结果，是不是有那么几个数据，你一直觉得怪？

学生想了想，说：有。三个样本，怎么拟合都偏一点。

老师说：你是怎么处理的？

学生说：我加了一个修正项，偏差小了，不过我自己也不太满意。

老师说：好，那三个样本，你不满意的地方，就是你的那一小段裂缝。现在我问第二个问题：如果把你那个修正项拿掉，不拟合了，只问「这三个样本为什么会长成这样」，有没有一种可能，你整个模型里的某个基本量，划错了？

学生沉默了一会儿，说：有可能。可是我没有任何依据。

老师说：对，这就是了。没有依据，所以推不出来；有那三个样本，所以不是拍脑袋。你现在有的是一个很粗糙的怀疑。它还不是猜想。要变成猜想，得有一句话，说清楚你认为哪个量划错了，换成什么。那一句话，写下来，大概率很难看，而且大概率在你现在所有拟合好的数据上，成绩比原来差。

学生说：那我为什么要写它?

老师说：因为你会发现，那三个样本，在新的说法里，终于不需要修正项了。也许只有这三个，也许后来又来了第四个、第五个，也许什么也没有来，那这句话就死了。但是写一句，你就多了一次机会。

这段对话里，老师没有说「你要有天才的直觉」，也没有说「你要跳出来想」。他只做了三件事：指认裂缝，不许用补丁盖住它；问基本量是否划错；允许那句难看的话被写下来。这三件事，就是这一章的全部实用内容。

## ◆ 一笔之后：它还只是猜想

最后要拦住一个常见的跳跃：一句话落了下来，是不是就等于典范诞生了?

不是。一句话落下来，只是一个猜想，它离典范还有很长一段路。典范是一个稳定的显露，稳定要靠许多东西：划界下面的概念要一个个站住，基本对象要在各种情形里都说得通，要有一套人人能学会的做法，要有一批人拿它去解谜并且解得出来，还要在与现象的反复对撞里没有被撞散。

这些，不是那一句话自己带来的，是之后很多年、很多人的工作，在那一句话的底下一层一层沉淀出来的。

所以，那一笔和典范之间，隔着一条很长的路。路上有三个常见的结局。

一个是死在路上：展开到一半，发现某个关键之处与现象冲突，这个划界站不住。这是大多数。死也是有价值的，它把一条岔路标了出来，让后来的人少走。

一个是被旧典范吞回去：展开的过程中，它的新词被一个个翻译成旧词，新对象被说成旧对象的特例，最后变成旧典范里的一个分支。这个结局很隐蔽，因为外表上它「成功了」，有论文、有会议、有头衔，可是那一笔原本要划的新界，已经被抹平。第 14 章会专门讲怎么防。

还有一个是站住：它的划界在许多次对撞里都挺了过来，它的基本对象成了新的常识，新一代人是在它的底下长大的。到那时候，它不再叫猜想，它成了新的显露，原先的那一笔，只剩教科书上的一个起点。

这三个结局，在落笔那一刻，无人能预知。这就是为什么本书一再强调，最初一笔要有人撑，要有对撞，要有判据。也是为什么在 AI 把展开变便宜之后，值得再问一次：这条路，怎样才能走得更快、死得更便宜、站得更稳？这是后面几编的主题。

带走的话

**典范猜想是一句话级别的新划界、新基本对象，它不是一个更大的假说，而是对「什么算数」的重新下笔。 它不是从旧典范里推出来的，因为旧典范里没有它的位置；它也不是拍脑袋，背后是对裂缝形状的长期体察，这份体察靠的是在典范内的苦功夫。 最初一笔常常粗糙，常常不如旧典范，所以不能用旧成绩单去评它，也不能因为它粗糙就赞美它。 猜想是在差异序列的缝上显露的，一位提出者随之显影。把它说成天才发动，教学就只剩祈祷；把它说成裂缝的显露，才有事可做。 展开变便宜了，下猜想的门槛低了，所以「这一笔值不值得展开」成了更要紧的问题。**

第 11 章

# 判据：什么算典范级，什么只是局部差异

一位老师收到学生的一份报告，题目很响：《一种全新的框架》。报告三十页，条理清楚，定义严密，推论一环接一环，还附了和几个经典理论的对照表。读完，老师心里有点发虚：写得太完整了，比他手里不少正经论文都完整。

他不知道该怎么回。说「好」，万一是个漂亮的空壳呢？说「不行」，万一这里真有点什么，他看不出来呢？

几年前，这种情形不多见：一个人要拿出三十页严密的东西，本身就说明他下了很多功夫，功夫是个信号。现在不一样了。三十页，一个下午就能出来。完整，已经不再说明任何事。

这一章要给读者几把尺子，用来分辨一句话到底是典范级的猜想，还是一个局部的、换了好看衣裳的想法。先把话说在前面：这些尺子是筛子，不是判决书。它们能帮你筛掉大部分，剩下的，还是要交给时间和现象。

## ◆ 为什么此刻特别需要判据

在 AI 把展开成本压低之前，判断一个新想法值不值得认真对待，有一个隐含的工具：看它的完整度。一个想法能被展开成一套像样的东西，说明提出者有功力，说明想法有几分根基。这个工具不完美，但大致管用，因为完整是贵的。

现在这个工具失效了。一个烂猜想，也可以被造得很完整。内部自洽，对照齐全，术语漂亮，推论有序，连「可能的反驳与回应」都有。而且这一点是任何人都能做到的：不需要懂行，不需要苦功夫，只需要一个会写的助手。

道理不难想。自洽，只是说内部没有矛盾。可是一个编出来的故事，也可以没有矛盾，小说就是这样。完整和自洽，保证的是「这套东西可以被说下去」，不保证「这套东西说的是世界」。

于是有两种坏结果同时出现。一种是真假难辨，老师和评委被一堆看着完整的东西淹没，只好凭印象，印象又偏向会写的人。另一种是反过来的：因为怕上当，对所有新东西都一概冷淡，真正的粗糙的新划界，也被一起关在门外。

这两种结果的共同根源，是缺一套不看完整度的判据。下面给四条。每条都配正例和反例，每条都尽量说成读者自己能拿去问的问题。

### ◆ 判据一：能不能把原来排不进去的差异，一并排进去

第一条，最重要，也最朴素。

回想前面说的：旧典范的裂缝，就是一串在旧显露下排不进去的差异。那么，一个新猜想如果是典范级的，它应该做到的第一件事，就是把这一串原来无家可归的东西，放进同一个位置。

注意这里有两个要求，缺一不可。

其一，是「一并」。不是解释了一个反常，而是把一批形状相近、原来各有各的补丁的反常，用同一个划界收了进去。如果一个猜想只能解释一个别扭

的数据，那它更像一个补丁；如果它能让三个、五个、十个原来互不相干的别扭，同时变得顺理成章，这就是一个信号。

其二，是「原来排不进去的」。不是解释了旧典范本来就解释得好的东西。新猜想在旧典范解释得好的地方，做得和旧典范一样好甚至更差，这不要紧；要紧的是它在旧典范的盲区里有收获。

正例，可以看大陆漂移一类的想法。海岸线能对上、两边的化石相似、远隔重洋的岩层互相衔接，这些在固定大陆的旧说法里，要一个一个各讲各的故事，比如设想曾经有陆桥。而「大陆会动」这一句话，把这批原来分散的事实收到了同一个位置。它的最初形态机制不全，可是这一点收获是实实在在的。

反例，看一类常见的情形。某人提出，某个旧理论里一个参数其实应该是时间的函数，加上这一项，那个别扭的数据拟合好了。它解释了一个排不进去的东西，可是只有一个，而且是靠多加一个自由度解释的。这种做法在常规科学里很正当，也经常有用，只是它是局部的。它没有把别的任何东西带进来。

这条判据还有一个很实用的变形：问「它让多少个原本不相干的问题变成同一个问题」。典范级的猜想，常常有一种「合并」的味道：以前以为是五件事，现在是一件事的五个样子。局部的想法则常常相反，是「拆分」，把一个问题拆成几个小问题，各自再处理。

怎么操作？拿一张纸，左边写下这个领域里你能找到的、至今没有被好好处理的别扭，尽量具体，尽量不要加工。右边写下这个新猜想。然后一条一条地问：左边的这条，在新猜想里有位置吗？

是自然落进去的，还是需要再加条件才落进去？自然落进去的，才算数。需要为每一条单独加条件的，那是补丁，这个猜想在做的，是把补丁换了个名字。

### ◆ 判据二：能不能改变「什么算问题」

第二条判据，比第一条深一层，也更难假装。

一个典范，不只是一套答案，它还决定了什么算问题。哪些事值得问，哪些事不值得问，哪些事根本不会被想起来，这些都由典范的划界事先决定了。常规科学的人，每天做的就是在这份问题清单上一项一项地解决。

那么，一个典范级的猜想，应该对这份问题清单做点什么？至少两件事。

一件，是让老问题消失，或者换了位置。旧清单上有些问题，在新划界下不再成立，不是被解决了，而是不用再问了。

比如，如果天上与地上不再是两个世界，那么「天体为什么按完美的圆运动，地上的东西为什么按另一套规则落下」这一类问题，不是得到了答案，是失去了发问的前提。又比如，如果大陆本来就会动，「陆桥在哪里」这个问题，就不再是个问题。

另一件，是让新问题出现。原来没人想到要问的事，现在成了顺理成章的问题。大陆会动之后，「推动的是什么」「在海底应该看到什么」「别的地方的岩层应该呈现什么样子」，一串新问题自然冒了出来。

这条判据的实际用法，是看一个猜想的「问题账本」。把它问题清单的变动列出来：消失了几条，改写了几条，新增了几条。局部的想法，一般只会在清单上勾掉一条或者两条，旧清单的其余部分原样不动。典范级的猜想，会让

清单大面积地变形，甚至让一部分人原来最拿手的工作，忽然显得有点无处下手。

这最后一点值得多说一句，因为它也解释了为什么典范级猜想总是受到抵制。不是所有人都愚昧，而是对一个把整份问题清单当作自己职业生涯的人来说，这个猜想等于在说：你多年精通的那份清单，一部分要作废了。

抵制是自然的。这也是为什么，一个猜想招来了强烈的反对，并不能证明它是对的；但是一个猜想招来的是礼貌的点头和「有意思，不过我们继续做我们的」，往往说明它没有改变清单，只是添了一个脚注。

反例也要有。很多新方法，比如一种新的测量手段、一个新的计算技巧，能让原来的问题做得更快更准。它们很重要，可是它们不改变清单。清单还是原来的清单，只是做题的速度提高了。AI 带来的，大部分就是这一类：做题更快更好，问题清单不动。这是常规科学的好消息，却不是典范级的信号。

### ◆ 判据三：有没有一个新的基本对象

第三条判据：看它的「存货清单」。

每个典范都有一份基本对象的清单：这个世界由哪些东西构成，哪些东西是原料，哪些东西是派生的。旧典范的清单是长期稳定下来的，很难动。典范猜想最本质的动作之一，就是往这张清单里加一项，或者划掉一项，或者把一项的身份改了。

所以，问一个猜想：它的存货清单上，有没有一件旧典范里没有的东西？这件东西，是不是整个解释真正靠它撑着的？

这里有一个很好用的检验，叫「拿掉测试」。假设把这个所谓的新对象从猜想里拿掉，整个说法还能不能照常运转？

如果拿掉之后，一切照旧，只是少了一个名字，那么这个「新对象」只是个标签，一个旧东西的新名字。如果拿掉之后，说法立刻塌了，连原来能解释的那批差异也解释不了，那这个对象是承重的。

这是一个对付「换皮」的利器。很多看上去全新的框架，一拿掉测试就现形：它引入了一堆新词，可是每个新词都可以翻译成旧典范里的一个现成的东西，翻译过去之后，所有论证原样成立。

这就是第 14 章要专门讲的「换皮的旧典范」，这里先提一个小方法：试着把这个猜想的所有新词，一对一翻译成旧典范的语言，看看有没有「翻不过去」的剩余。翻得过去，说明新词是皮。翻不过去、一翻就丢了关键内容的，才是新对象。

正例，可以看板块。在「地壳是一整块铺好的地板」的旧清单里，没有「板块」这件东西。把它拿掉，大陆漂移、海底扩张、地震带的分布，这些说法一齐塌掉。它是承重的。

反例，大家身边最多的。某人写一套「教育的新框架」，其中有「学习生态位」「认知势能」「成长场域」几个新词，每个词都有定义。

拿掉测试一做：把「学习生态位」换成「学生所处的学习环境」，把「认知势能」换成「学生已有的知识基础」，把「成长场域」换成「班级氛围」，整套论述几乎一字不改，依然成立。三个新词，全是皮。

要补一句，免得读者走偏：新对象不一定是新的「东西」。它可以是一种新的区分。比如，以前把两样东西混在一起，现在分开了；以前以为是两样，

现在合并了。区分的方式变了，基本对象也就变了。判断时，看的是清单有没有变，不是看有没有出现新名词。

### ◆ 判据四：能不能带来可检验的对撞

前三条判据，看的都是猜想本身的样子：它收了什么，改了什么，加了什么。第四条要出门，去看它和现象的关系。这一条，是整个判据系统的锁扣，因为前三条理论上都可以被一个会写的助手伪造出来，唯独这一条不行。

为什么？因为前三条是「说法」，第四条是「撞法」。

一个猜想能收多少差异、改多少问题、加多少对象，都可以靠写来完成。可它能不能带来一次对撞，不能靠写。对撞的意思是：这个猜想说，在某个地方，世界应该是这个样子；而旧典范说，在那个地方，世界应该是另一个样子；两者不可能同时对；然后，去看。

这里面有几个要点，逐个说。

第一，要有「旧典范不会这么说」的预言。如果一个猜想的所有预言，旧典范也会同意，那就不存在对撞，它们只是两种说法讲同一件事。真正有分量的对撞，是新猜想预言了一件旧典范会说「不可能」或者「没有理由」的事。

第二，预言要具体到现象能够回答。「宇宙是一个整体的信息场」，没有对撞，因为没有任何现象能让它为难。「如果大陆在漂，那么海底某类岩石的年龄应当随着离某条线越远而越老」，这是对撞，因为现象可以说是，也可以说不是。这里只借这个说法来说明形状，不引用具体的观测结果，也不编数字。

第三，对撞的结果要有可能不利于猜想本身。一个只会收到利好消息的猜想，是不可检验的。好的对撞有一种冒险的气味：提出的人明白，如果现象不给面子，这句话就得认。

第四，也是最要紧的一点：对撞的对象是现象，不是文本。

这一点在 AI 时代尤其需要重申。让一个 AI 去「评审」一个猜想，它可以给出非常有条理的评语：逻辑严密，创新性高，与现有文献的关系清晰。这些评语本身是文本对文本。它回答的是「这套说法在语言上站不站得住」，而不是「这套说法在世界上站不站得住」。语言上站得住的东西，太多了。

所以，要说一句也许不太好听的话：对一个 AI 造出来的完整框架，最不该做的事，是让另一个 AI 来评它内部是否自洽。自洽的检验，两边都会通过，两边也都没有告诉你任何关于世界的事。检验必须回到与现象的对撞：找到那个新旧两边预言不同的地方，去看，去测，去找一个能让它为难的东西。

这就是为什么前面说，这一条是锁扣。前三条筛出来的是「看上去像典范级」的，第四条筛出来的才是「有可能是典范级」的。而即使第四条通过，也只是说它活着，没有说它已经赢了。

有两个提醒。

提醒一：对撞可能很慢。有些预言，在提出的时候，没有手段去检验。大陆漂移的最初形态，就等了很久，后来的观测才把检验的机会带来。所以，对撞的判据要问的是「可检验吗」，不是「已经检验了吗」。

一个猜想可以暂时无法检验，但不能是原则上不可检验的。区别在于：提出者能不能说出，需要什么样的观测，看到什么结果，这句话就算站不住。能说出来，是活的；说不出来，是死的。

提醒二：对撞的结果也会被旧典范消化。一个粗糙的新猜想，第一次对撞失败了，旧典范的人会说「看，它错了」。可是第一次失败，可能是新猜想的展开不对，而不是划界不对。

这需要把「划界」和「展开」分开看：展开可以改，划界是否站得住，要看多次、多处的对撞。这一点，第 15 章讲并排发生时会更细。

## ◆ 局部差异长什么样

有了四条尺子，再回头看看那些「只是局部差异」的想法，认一认它们的样子。这一小节，是给读者一张素描。

第一类，是参数的调整。多一项，少一项，改一个系数，让拟合更好。它们属于典范内的精修，做得好，很有用。特点是：所有基本对象不变，问题清单不变，只是数字变了。

第二类，是方法的升级。新仪器、新算法、新的数据处理方式。它们可以让常规科学跑得更快，也可能让旧的反常暴露得更清楚，因而间接帮助裂缝显出来。但方法本身不是划界。

第三类，是跨领域的搬运。把甲领域的框架整体搬到乙领域，说「用 X 的眼光重新看　Y」。这有时候非常有启发，可是它搬来的基本对象，是甲领域早就有的；乙领域只是多了一个借来的显露。

是否算典范级，要看搬过去之后，乙领域自己的差异序列是不是出现了新的排法。如果只是给乙领域的现象贴上甲领域的标签，那是翻译，不是发生。

第四类，是在旧划界里做极端化。把某个已有的因素推到极端，说「一切都是　X」。这一类很吸引人，因为一句话就能说完，也有典范猜想那种气势。

可是它没有换基本对象，只是把一个旧对象的权重调到了最大。很多「一切都是什么」的宏大说法，属于这一类。

第五类，是综合与调和。把两个对立的旧说法各取一半，说「其实都对」。这在常规科学里有时是成熟的标志，但它一般不引入新对象，而是在旧对象之间画一个折中的线。

第六类，也就是现在最多的：重新包装。用漂亮的新词，把旧典范重述一遍。内部自洽，样子完整，可是做完拿掉测试，什么也没有剩下。这一类，AI 造得最快、最好，所以本书要专门留一章（第 14 章）来对付它。

读完这张素描，不要得出「这些都没价值」的结论。恰恰相反，它们每一类都有价值，是常规科学的日常。本书不贬低这些工作，只是想让读者分清：它们是同层的扩张，是广度和深度，不是换层。

分清了，才不会拿深度当层级，才不会在该做精修的时候空等一场革命，也不会在该警觉的时候被一份漂亮的完整报告晃了眼。

### ◆ 放回日子里：回到那位老师

回到开头那位老师和那份《一种全新的框架》。现在他手里有了四条尺子，看看他怎么办。

他先做第一条：请学生列出，这个领域里有哪些至今没处理好的别扭。学生列了两条，都是教科书上人人知道的难题。老师再问，新框架里这两条是怎么落进去的？学生答，用框架里的第四个概念可以解释。

老师追问，那是不是只有这两条？有没有哪个别扭，是你在写这个框架之前并没有想到、写完之后才发现也能放进去的？学生想了想，说没有。这一条，成绩不好：它没有一并收进更多，也没有收进没预料到的东西。

第二条：问题清单变了没有？学生说，新框架让大家更重视过程。老师说，这个领域原来的清单里，过程已经是个热门问题了，所以这是同一份清单上的权重调整。

第三条：拿掉测试。老师随手拿掉其中两个新概念，换成旧名词。论述不受影响。

第四条：请学生说一个预言，在这个预言上，新框架和旧框架会得出不同的结论，并且现象能回答。学生愣住，说，这个框架主要是提供一种看问题的方式。老师没有说话。

四条走完，老师不必给出刻薄的评语。他只需要告诉学生：这份东西写得很规整，有教学价值，可以当作一篇好综述或者一个好的整理。如果想把它推到猜想的层级，需要回答那四个问题，尤其是第四个。这是一个诚实而且有用的回复，比「好」或者「不行」都有用。

请注意老师没有做的事：他没有说「你缺乏创造力」，没有说「你不是那块料」。他只是把一份东西放在了它应该在的位置上。典范级和局部差异之分，不是人的等级之分，是东西的层的差别。同一个学生，也许下个月拿来的是一个更粗糙、更难看、但答得出第四个问题的东西，那才是值得追着问下去的东西。

## ◆ 用 SDE 的话再说一遍：四条判据在看什么

四条判据看上去是四件不相干的事，其实用本书的说法，它们在看同一件事的四个侧面。

典范是一个稳定的显露。旧显露之所以失稳，是因为有一串差异序列排不进去，反常纠缠成压力。那么，一个新的显露要站得住，就必须对这几样东西各有交代。

判据一看的是差异序列：原来排不进去的那串，新显露接不接得住。接得住，而且是一并接住，说明新显露在差异序列这一头是扎了根的。

判据二看的是显露的划界：什么被当成要紧的，什么被当成噪声。划界一变，问题清单自然跟着变，这是同一件事的两面。

判据三看的是显露里的基本对象：显露稳定下来之后，总有一些东西被当成「基本的」。新对象是不是承重的，就是看新显露的骨架里有没有一根旧显露里没有的梁。

判据四看的是纠缠：一个显露要稳，不是自己说了算的，它必须和现象互相咬住。咬得住，才叫稳定；只会在文字里自洽，那是悬空的。新旧显露各自预言不同的东西，再让现象去说话，就是在检验谁和现象咬得更牢。

这样理解有一个好处：它提醒读者，判据不是外面加给典范的考试，而是典范自己的结构。一个稳定的显露，本来就要同时回答这四件事：接得住哪些差异，怎么划界，用哪些基本对象，跟现象怎么咬合。判据只是把它们拆开问。

也要守住总纲里的口径。读到这里，有人会想：「那就是说，评审的人要有慧眼，一眼看出哪个是真的。」不是。这四条的意思，恰恰是不依赖慧眼。慧眼是靠不住的，历史上被慧眼错过的粗糙猜想不少。

四条尺子是把「眼光」换成「可以一条一条查的问题」，谁查都一样，老师能查，学生能查，AI　也能帮着查前三条。第四条要回到现象，那就派人去看、去测。判据的作用，是让「识别典范级」这件事，从某个人的鉴赏力，变成一套可以教、可以练的程序。

## ◆ 一个 AI 时代的实战：两份同样完整的东西

最后来一个小的模拟，看判据怎么在　AI　时代用。下面是编的情形，只为说明用法。

某个研究小组，对同一条裂缝，让　AI　各自展开了两份东西。我们叫它们甲和乙。

甲：提出细胞里一类现象的「全新理论框架」，四十页，有定义、有公理化的表述、有与五个经典模型的对照。

乙：只有九页，是一个并不好看的说法，大意是，目前被当作一类的两种现象，其实应该分开；一旦分开，一直以来互相打架的几组实验结果，会各自顺起来。

从完整度看，甲完胜。从气势看，甲也完胜。现在上四条。

判据一：甲解释的，全是旧模型本来就能解释的，对不上的那几组实验，它说「有待进一步研究」。乙把那几组打架的结果收进去了，而且其中一组，是写这份东西的人原先没打算处理的。乙占优。

判据二：甲的问题清单，和旧框架一样，只是术语换了。乙说，原来争论不休的那个问题，不是没解决，是问错了，因为它把两件事当一件事问。乙占优。

判据三：拿掉测试。甲的五个新概念，逐一能译回旧术语，译完几乎无损。乙只有一个新区分，拿掉它，打架的实验又回到原样。乙占优。

判据四：甲没有任何新旧预言不同的地方，它自己也承认「主要提供理解视角」。乙说，如果区分成立，那么在某种条件下，两类现象应该出现不同的变化方向，旧说法则预言它们一起变。这个条件，在实验室里是能造出来的。乙占优。

四条走完，结论很清楚，而且这个结论违背了「完整度」的直觉。甲是一份很好的综述，可以留着给新人读；乙值得拿出资源，去造那个条件，去撞一次。

这里还要留意乙的另一面：它很可能是错的。区分也许不成立，实验也许会给出相反的结果。判据没有说乙是对的，只说乙值得被撞。这就是判据的用处：它不负责宣布真理，它负责把有限的人力和有限的实验，分给真正有可能改变局面的那个想法，而不是分给写得最漂亮的那个。

这个结论，还有一个重要的后果，要放在整本书的论证里看。如果 AI 能把任何猜想展开得很完整，那么在 AI 时代，「谁的展开更完整」不再是任何人的优势。优势在哪里？

在知道该把展开的资源给谁，在知道怎么设计那一次能为难猜想的对撞，在能分辨「完整」和「站得住」。这些恰恰是判据在训练的东西。也恰恰是下面几编要讲的验收、并排比较、教学评价的基础。

## ◆ 三种常见的误用

尺子拿在手里，容易用歪。有三种歪法，值得提前点出来。

头一种，拿判据当门槛，把新东西挡在门外。有人学了四条，见到一个粗糙的想法，立刻逐条打叉：收的差异少，问题清单没怎么变，新对象说不清，预言还没法测，结论「不是典范级」，扔。

这等于又拿旧成绩单去打新生儿。前面讲过，最初一笔就是这个样子。正确的用法是反过来：看哪一条有苗头。四条里只要有一条真的亮了，尤其是第一条或第四条，这个想法就值得留着，给它一点活着的时间。

第二种，拿判据当标签，给自己的想法贴金。有人写完一个框架，回头对着四条找说法：我这个「也收了很多别扭」，我这个「也改了问题清单」。这是倒过来用。

判据不能事后去凑，要事先立。好的做法是，在动笔展开之前，就写下四个问题的回答，尤其是第四条：它预言什么，旧说法预言什么，怎样才算它错了。写在前面，才有约束力。

第三种，拿判据当裁判，代替现象说话。四条里前三条，是「预审」，作用是决定值不值得花力气去撞；真正的裁判只有一个，是第四条所指的现象。有人把前三条打了高分，就宣布「这是个典范」。这又落回了完整度的老坑：说法再漂亮，没撞过，就只是说法。

三种歪法，有一个共同点：都想让判据替人省事。可是判据的本意，是让该费的力气费在刀刃上，不是让力气省掉。

## ◆ 判据的限度，以及一张小表

最后要诚实地说判据自己的限度。否则这一章会变成另一种迷信。

第一，判据是筛子，不是判决。四条尺子能把大量明显的局部差异筛出去，也能把一部分有希望的猜想留下来，但它们不能保证留下来的一定是典范。保证只来自时间，来自一个领域在许多年里的选择。

第二，会有漏判。真正的典范级猜想，在最初一笔上，可能在某几条上表现很差。它可能收的差异还很少，它的新对象还说不清，它的对撞预言还很模糊。所以对粗糙的新东西，判据的用法应该是「列清单，看哪几条有苗头」，而不是「一条不达标就扔」。用这一章的话说：别用旧成绩单打，要用这四条去问。

第三，判据之间有主次。如果必须排序，第四条是最硬的，因为只有它不能被文字伪造；第一条和第二条是最容易被误读的，因为「解释了多少」和「问题怎么变」在一个会写的人手里，都能被说得很漂亮；第三条的拿掉测试，最适合用来拆穿换皮。

第四，判据不替代专业。要判断一个猜想收没收进了真正的差异，你得懂这个领域，得知道哪些别扭是真别扭。这又回到了地基：不走过路的人，看不懂地图，也评不了新地图。腿不会因为有了车就变得不重要，评新交通方式的人，还是得会走路。

下面把四条放进一张小表，便于翻查。

| 判据 | 要问的话 | 局部差异的样子 | 典范级的苗头 |
| --- | --- | --- | --- |
| 一、排进差异 | 原来排不进去的，有几个被一并收进来？ | 只解释一个别扭，靠多加自由度 | 一批形状相近的别扭落进同一个位置，且有没预料到的 |

| 判据 | 要问的话 | 局部差异的样子 | 典范级的苗头 |
| --- | --- | --- | --- |
| 二、改变问题 | 老问题消失几条、新问题冒出几条? | 勾掉一两条，清单其余原样 | 清单大面积变形，一部分拿手活失去前提 |
| 三、新对象 | 拿掉它，说法还成立吗? | 新词都能一对一译回旧词 | 拿掉即塌，连原来能解释的也塌 |
| 四、可检验的对撞 | 新旧在哪里预言不同，现象能回答吗? | 没有预言，或旧典范也同意 | 有旧典范会说「不可能」的预言，且能被现象为难 |

四条之外，还有一件事要放在桌面上：判据本身，也是在典范内用的吗?

这是一个值得想的问题。判据作为一把尺子，用的是目前科学共同体已经认可的东西：差异、问题、对象、检验。用这把尺子去量一个想要改变这些东西本身的猜想，难免有一点悖论。

本书的回答是：这把尺子量的不是「它符不符合旧典范」，而是「它是不是在做划界这件事、做得扎不扎实」。差异、对象、对撞，是任何一个科学性的划界都需要面对的，所以这把尺子可以跨典范使用。它不裁决谁对，只裁决这一笔是否有资格被认真展开。

## ◆ 这套判据自己怎样才算错

一套判据如果只会判别人，自己永远不会错，它也就成了另一种迷信。所以这里要说清楚：出现什么情形，本书的主命题就要收回，或者收窄。

第一种情形。有一个猜想，被这套判据判成了典范级。可是拿掉它的新词之后，它的全部推论和旧典范一模一样，也没有一个旧典范原本排不进去的差异被它排进去。这说明判据放过了换皮的东西，四条尺子就得重新磨。

第二种情形。在有明确验收的领域里，按本书的工序做出了好几个并排的猜想。可是时间一长，它们通过外部对撞的机会，并不比不按工序、凭一次灵感做出来的结果更高。这说明工序是装饰，"典范发生可以成为一门学"这句话就站不住，至少要大大收窄。

第三种情形。AI 在没有任何人给出划界的情况下，持续、独立地产出了能通过外部对撞的新划界。这并不推翻"常规科学被接走"，那一条照旧成立。可是"人守验收与选择"这一条，就要重写。

这三条，都是可以去看的事，不靠辩论，靠对撞。本书希望它们被认真地去找。

---

带走的话

**完整、自洽不再能说明一个猜想值钱，因为烂猜想也能被造得很完整。判据要回到别处去找。 四个问题：能不能把原来排不进去的差异一并排进去？能不能改变什么算问题？有没有一个承重的新基本对象（拿掉就塌）？能不能带来新旧预言不同、现象又能回答的对撞？ 前三条看说法，第四条看撞法。说法可以写出来，撞法写不出来，所以第四条是锁扣。AI 评审 AI，只是文本对文本，不算对撞。 局部差异有自己的价值，参数、方法、搬运、极端化、调和、重新包装，都是同层的广度和深度，只是不是换层。分清它们，是为了各得其所。判据是筛子不是判决。对粗糙的新东西，用它列清单、看苗头，别拿旧成绩单打。下一章用三个案例，看这些判据在真实的历史里长什么样。**

---

第 12 章

# 案例：典范是怎样发生的

先说一件家常事。

你有没有补过一件旧衣服？袖口磨破了，缝一针；肘部磨破了，贴一块布；后来领子也坏了，再缝。一年下来，衣服还能穿，只是每一处都打着补丁，针脚越来越密，穿的人也越来越熟练：哪里补过、哪里不能使劲拉，心里都有数。有一天，有人说：别补了，这衣服的剪裁本身就不对。

这句话为什么难说，又为什么难听进去？因为补丁是实打实的功夫，每一块都管用。要换一件衣服，得重新量尺寸、裁布、缝合、试穿、改版型，这一整套活，比再补一百个补丁都重。

所以多数时候，旧衣服就这样一直穿下去，补丁一层压一层，直到哪一年既有了新的剪裁想法，又有人手把它做成一件能穿出门的衣服。

这一章讲两件「旧衣服换新衣服」的大事：地球中心的天文学怎样换成太阳中心的天文学，牛顿的引力怎样换成爱因斯坦的引力。随后再看第三件，量子力学。教科书里，这些事常常是一个人的故事：某某看穿了，某某想通了，世界换了样子。

本书不这样讲。按本书一路用的口径，一个典范是一个稳定的显露，由差异序列和特征纠缠一层层发生出来。换一个典范，也是一场发生：旧显露下排不进去的差异越积越多，纠缠成压力，旧显露失稳，一个新的划界被写出来，再被庞大的功夫展开成系统，最后才站得住。

那位提出猜想的人，是这条序列上显影出来的人，不是推动一切的人。

读这一章，请特别留意一件常被忽略的事：猜想写出来之后，还要有人把它展开，而展开的成本大得惊人。很多好猜想在历史上并不是错了，而是展不开、养不起，没有活下来。这一点，恰恰是 AI 改变的地方。

## ◆ 一、日心说：一件补了一千四百年的旧衣服

### ◆ 旧典范下的稳定显露：地球不动，天在转

先把我们的脑子放回古代。

站在一片空地上抬头看，太阳从东边升起，从西边落下；星星一夜夜绕着一个点转；月亮圆缺有序。脚下的大地稳稳当当，没有一点动的感觉。这是任何人、在任何时代、不借助仪器都能得到的显露。地球不动，天在转，这不是哪位学者的偏见，而是人人每天亲眼所见、亲身所感的东西。

再往上一层，是希腊人留下的自然哲学。亚里士多德一系的说法是：宇宙有一个中心，沉重的东西天生要往中心去，所以石头往下落，土往下沉；天上的东西是另一种材料，天然做圆周运动，永不停歇，永不改变。

地球恰好在宇宙中心，所以不动。这套说法把地上的运动和天上的运动分开管，把「为什么东西往下掉」和「为什么地球在中心」接成了一条线。它不是一句孤立的断言，而是一个相互咬合的整体：看到的、感到的、想到的，彼此印证。

到了公元二世纪前后，托勒密把这个整体写成了一部精密的天文学巨著，后世通称《至大论》。它的核心工具是几何：每颗行星沿着一个小圆（本轮）转，这个小圆的圆心又沿着一个大圆（均轮）绕着地球转。

需要时，还可以让大圆的中心稍微偏离地球，再设一个特殊的点（偏心匀速点，也叫「等分点」），让运动相对于那个点才显得均匀。有了这些工具，天文学家可以推算行星的位置，预报日食月食，编制历表，供航海、历法和占星使用。

这套东西不是笨东西，请千万别把它当成「幼稚的错误」。在相当长的时间里，它是一台好用的计算机器：给定一个时间，它能给出一颗行星大致在哪里。

按本书的说法，托勒密的体系是一个稳定的显露，由三样东西支撑：天上的差异序列（行星在不同夜晚的位置变化）被它稳稳地排了进去；地上的经验（大地不动）与它不冲突；哲学里的宇宙图景（中心、圆周、天地之别）为它撑腰。三条线互相拉着，谁也不松。这是一件「穿得很合身」的旧衣服。

## ◆ 排不进去的差异：行星为什么会「倒着走」

可是天上有一类东西，从一开始就不太听话，那就是行星。

「行星」这个词在希腊语里的意思是「漫游者」。它们不像恒星那样整整齐齐地绕着一个点转，而是在星空背景里慢慢地走。更怪的是，一颗行星往东走了一阵，会慢慢停下来，掉头往西走一段，再停一停，又转头往东。这叫逆行。火星的逆行尤其明显：几个月的时间里，它在星座之间倒着退。

在「地球不动，天体绕地球做匀速圆周运动」的画面里，逆行很难安置。天体应该规规矩矩地转，怎么能倒着走？这是一个典型的差异：旧显露要求这样，实际观测是那样，两者对不上。

对不上怎么办？托勒密一系的办法是加一个小圆。行星先沿小圆转，小圆的圆心再沿大圆转，两个运动叠在一起，从地球上看去，行星就时而前进时而倒退。

这是一个非常聪明的办法：它没有推翻「地球不动、天体做圆周运动」的整体画面，只在画面里加了一块补丁，就把逆行装了进去。

但补丁不是白加的。逆行装进去了，别的差异又冒出来：行星的速度不是处处一样，走一圈的快慢随位置有变化，月亮的位置也对不上，更精细的观测又要求再调。于是再偏一点，再加一个点，再加一个小圆。每一个补丁都解决了一个具体问题，也都让整体更复杂一点。

这里要看清一个细节。托勒密的体系里，有些事实是被「规定」进去的，不是被「推出来」的。比如水星和金星这两颗内侧的行星，从来离太阳不远：水星总在太阳附近，金星离太阳也只有一段有限的角度，早晨或傍晚才看得见。

在地球中心的体系里，这需要单独交代：规定这两颗星的本轮圆心必须始终和太阳保持在同一条线上。而外侧的行星（火星、木星、土星）又有另一条规定：它们的逆行总发生在和太阳相对的位置附近，也就是整夜可见的时候。

这些规律托勒密都知道，也都写进了体系。但它们是一条一条添上去的约束，彼此之间没有内在的关联。好比衣服上的补丁恰好都落在磨损最厉害的地方，却没有人问：为什么磨损都在那几处？

用本书的话说：这些差异都能被排进旧显露，但每排一次，就要付出一次额外的规定。差异序列还没有排不进去，却已经排得越来越吃力。

## ◆ 纠缠成压力：几种不同的别扭拧在一起

一件衣服补久了，让人别扭的不是哪一块补丁，而是补丁们之间的关系。

在托勒密体系长久的使用中，别扭有好几种，原本互不相干，后来渐渐拧到一处。

第一种别扭在计算。历表用久了会漂：照表预报的行星位置，和夜里的实际观测，差距逐代累积。天文学家改参数，重新编表，几代之后又得改。这是日常工作中的小烦恼，没人当它是危机。

第二种别扭在形式。那个让运动「相对它才均匀」的特殊点，在一部分人看来，与「天体做匀速圆周运动」这个根本信条有冲突：既然说是匀速圆周运动，均匀就应该相对圆心而言，而不是相对圆心之外的一个点。

第三种别扭在整体图景。前面说的那些「规定」，水星金星为何不离太阳，外行星逆行为何总在与太阳相对时发生，在地球中心的画面里是一条一条的巧合。每颗行星的本轮大小也可以随意调节，没有哪个数是被图景本身决定的。整个体系更像一张拼贴画，而不是一张有内在结构的图。

这几种别扭，按本书的口径，是同一个显露里累积的几条差异序列：计算上的漂移、形式上的不洽、结构上的拼贴感。单独看，每一条都能靠补丁挡住。

但当它们彼此拧在一起，成为一种「这套东西哪里不对劲」的总体感觉，就不再是几个孤立的小毛病，而是纠缠成了压力。旧显露还站着，但站得不稳了。

这种态势会在许多人身上、许多地方，以不同的方式冒出来。事实上，在哥白尼之前，古代已有人设想过太阳居中，据后世记载，古希腊的阿里斯塔克

就提出过这样的想法；中世纪的伊斯兰天文学家也针对托勒密体系的形式问题提出过批评与改造。

## ◆ 猜想显露：把太阳放到中间

十六世纪上半叶，波兰人哥白尼写出了他的大书，书在 1543 年出版，题目大意是《天球运行论》。书里的核心猜想，一句话就能说完：太阳在中心（或者靠近中心），地球和其他行星一起绕太阳转，地球每天还绕自己的轴转一圈。

这一句话，换掉了什么？

它换掉了「谁在动」的划界。过去的划界是：天动，地不动。现在的划界是：地也动。过去地球是宇宙的中心和例外，现在地球是行星中的一颗。这不是在旧衣服上再补一块，这是把剪裁换了。

新划界一旦成立，一连串过去需要「规定」的事实，忽然成了「结果」。

先看逆行。如果地球和火星都绕太阳转，地球在内圈，跑得快，火星在外圈，跑得慢，那么每隔一段时间，地球就会从火星内侧「超车」。就像你在高速公路的快车道上超过旁边慢车道的一辆车，会看到那辆车相对远处的背景好像在往后退。

逆行不再需要一个专门的小圆来制造，它是赶超的自然后果。外行星的逆行为什么总发生在与太阳相对的位置附近？因为超车发生的时候，地球正好在太阳和行星之间。这一条，不用再另外规定。

再看内行星。水星金星为什么不离太阳太远？因为它们的轨道在地球轨道之内，从地球上看，不可能转到离太阳很远的方向。这一条，同样不用再规定。

还有一件更微妙的事：在日心的画面里，各行星的轨道大小之间有了固定的比例关系，这些比例可以从观测里定出来，不能随意调节。

托勒密体系里，每颗行星的本轮大小只相对于自己的均轮有意义，整体的尺度是自由的。哥白尼的体系里，行星排在什么次序、彼此的相对距离是多少，是被图景本身决定的。一件原本松散拼贴的东西，忽然有了骨架。

另一半也要说清楚，免得读者以为这件事一下子就赢了。

哥白尼的体系，并不比托勒密的体系预报得更准。这是史学界一般的看法。原因很简单：他仍然相信天体做匀速圆周运动，为了和观测吻合，同样需要小圆，同样需要把圆心稍微偏开。

他去掉了那个令他不满的特殊点，但没有去掉小圆。如果只比「谁算得准」，当时的哥白尼体系没有明显优势，计算起来甚至并不更简单。

后来人们用他的体系编制了新的天文表，在一段时间里被天文学家使用。据一般的看法，这主要因为它是一个可以算的、完整的体系，而不是因为精度压倒了旧表。

## ◆ 反对的理由并不愚蠢

当时的人为什么没有马上接受？这个问题常被轻率对待。

教科书里常把反对者写成固执的、被教会吓住的人。教会的影响确实存在，后来的事态也确实严重。但单看知识层面，反对的理由就很扎实。

第一，地球如果在动，我们为什么没有感觉？一个转得很快的大地，按常识应该有狂风，有震动。一块石头从高塔上扔下，如果地球在转，石头落地时塔早就移走了，石头应该落在塔的西边。

可是实际并不是这样。这个反驳在亚里士多德物理学里无懈可击，要回答它，需要一套全新的关于运动的道理：运动的物体会保持运动，一个整体匀速运动时，里面发生的事情和静止时没有区别。这套道理当时还没有。

第二，如果地球绕太阳转，那么一年里地球的位置在变，从不同位置看同一颗恒星，应该看到它在背景里轻微移动。这叫视差。当时的观测者没有看到任何视差。要么地球不动，要么恒星远得不可想象。

哥白尼一系的回答是：恒星确实非常远。这个回答在逻辑上是通的，但它要求宇宙比人们习惯的大得多。直到十九世纪，恒星的视差才第一次被可靠地测到。在那之前很长的时间里，这个证据一直缺席。

第三，新体系要求取消天和地的差别，让地球也成为天上的一颗行星。这等于拆掉了亚里士多德宇宙里「上与下」「中心与边缘」的整个秩序，连带着，物体为什么往下落也得重新解释。拆房子容易，新房子的地基还没打。

所以，面对哥白尼的猜想，一个清醒的、有学问的人完全可以说：这个想法很漂亮，但它缺三样东西：一套新的运动学，一个新的力学，一个能把预报做得更准的几何。它现在是一个猜想，不是一个体系。

这个判断是对的。它指出的，恰恰是新典范发生的真正难处：猜想出来了，缺的是它之后要做的全部功夫。

## ◆ 展开成系统：接下来的一百五十年

这一场功夫，由许多人、在许多地方、花了大约一个半世纪做完。这里只挑几件有把握的事。

第一件，是更准的数据。十六世纪后半叶，丹麦的第谷建起了当时最精良的肉眼观测设施，多年如一日，把行星的位置测得比前人准得多。

他自己并不接受地球在动，提出了一个折中的体系：行星绕太阳转，太阳带着它们绕地球转。这个折中方案，正是旧显露想把新猜想「消化」进自己的一种努力。而他留下的那一大批数据，成了后面工作的原料。

第二件，是对这批数据的苦算。开普勒接手了第谷的资料，特别是火星的观测，花了许多年反复计算。他起初也想用圆，结果总有一小点对不上，这一点点差距，他没有放过。

最后他放弃了圆，改用椭圆，并总结出行星运动的几条规律，其中一条把轨道大小和周期联系了起来。

这个细节对本书很要紧：把「匀速圆周运动」这个延续了近两千年的信条放下，据后世的叙述，与那一小截数据的残差有极大的关系，不像是一个情绪上的决定。哥白尼那句话要变成可以算准的体系，得先有这个换圆为椭圆的改造。

第三件，是望远镜。十七世纪初，伽利略把望远镜指向天空，看到了木星的卫星、金星的盈亏、月球表面的起伏。

木星有绕它转的卫星，说明并非一切都绕地球转；金星有和月亮类似的盈亏变化，与托勒密体系的预期不合，却与金星绕太阳转相合（第谷的折中体系同样容得下它）。

这些观察没有一个单独证明地球在动，但它们一件件拆掉了旧图景里「天上是另一种完美材料」的说法，也让新图景有了看得见的支撑。他还花了很大力气，论证一个整体匀速运动时里面发生的事情和静止没有区别，这是对「为什么感觉不到地球在动」的正面回答。

第四件，是新的力学。十七世纪后半叶，牛顿把运动定律和万有引力写在一起，说明了行星为什么沿椭圆绕太阳转，同时说明了苹果为什么会落地。哥白尼留下的「物体为什么往下落」的缺口，到这时才真正被一个更大的体系补上。天与地的分隔，也在这个体系里被取消了。

把四件事排在一起，结构就清楚了：哥白尼的一句话，是猜想；第谷的数据、开普勒的椭圆、伽利略的望远镜和运动道理、牛顿的力学，才是把猜想展开成系统。

后面这几样加起来，是前者的许多倍。没有一样是哥白尼自己做的，也没有一样是在一个人的头脑里一次完成的。每一步，都长在前一步造出来的材料上。

按本书的口径，不说「哥白尼发现了真理」，也不说「伽利略看穿了宇宙」。

要说的是：在行星逆行、历表漂移、体系拼贴这些差异序列与别扭拧成的压力下，一个新的划界（地球也动）显露出来；这个划界要站住，要等观测更准，要等几何放弃圆，要等运动的道理换新，要等一个统一的力学出现；这些一步一步到位以后，新的显露才稳定下来，成为后来人人在其中工作的新常规科学。

## ◆ 展开的成本，决定了猜想能不能活下来

现在来说本章最想说的一件事。

哥白尼的猜想发表于 1543 年前后，而它真正取得压倒优势，是在一个多世纪之后。这期间，这个猜想一直处在危险之中。

它可能被当作纯粹的计算技巧而被无害化（这在一定程度上确实发生了）；可能被某个折中方案顶替（第谷的体系就是）；可能因为缺少后续的人力、资金与时间，被搁置到下一个世纪；也可能在宗教与政治的冲击下被压住。

在这样的条件下，一个新猜想的存活，靠的不是它对不对，而是展开它的成本有没有人付得起。第谷的设施要靠王室的资助；开普勒多年的手算要靠一个人的毅力和一份并不稳定的薪水；望远镜要靠工匠的手艺；牛顿的数学要靠新发明的微积分。

每一项都是巨大的、缓慢的、不确定的投入。历史上被留下来的，是恰好碰上这些条件的猜想；没碰上的，我们大多已经看不到了，因为没人记录它们。

所以，过去的科学史有一个很少被谈论的「筛子」：展开成本。猜想不稀缺，稀缺的是能被展开的猜想，而筛子的网眼由成本决定。

由此可以设想一下：假如当时有一种手段，能把「把太阳放到中间」这个划界，在很短的时间里算成一套完整的历表、一套完整的几何结构，并与托勒密体系、第谷的折中体系并排摆出来，在同一批数据上逐项比较，会发生什么？

## ◆ 若当时有 AI 帮忙展开

先说会改变什么。

第一，展开的速度。今天这类几何建模与数值计算，AI 可以在极短时间内完成。同一个猜想还可以展开成多个版本：纯圆的、带小圆的、偏心的，一次摆出来看哪个与数据最合。「圆不行」这个检验，开普勒花了多年，在这个条件下可以快得多。

第二，并排比较。托勒密体系、哥白尼体系、第谷体系，可以同时造出来，在同一份观测资料上比拟合、比简洁、比能预言多少还没观测的东西。

这样，「日心说在预报上并没有明显优势」会被很早、很清楚地摆在桌面上，而不是争论一百多年。同时，日心说在结构上的好处（逆行的自然解释、轨道比例的固定）也会被量化出来。

再说不会改变什么。

第一，数据不会凭空出现。第谷的肉眼观测、伽利略的望远镜看到的东西，是对天空的真实测量。AI 可以处理数据，但要有数据，还得有人去建设施、去看天。

第二，那个划界是怎么长出来的。「地球也动」这一句，并不在旧显露的任何一个展开版本里。托勒密体系再精细地展开，也展不出日心说；让 AI 把旧体系展开得更精致，只会得到一台更漂亮的旧机器。

但也别把它想成一边先交出一句话、另一边照办去造。猜想是在人与 AI 的来回里一点点长出来的，展开的结果会反过来改写猜想本身，某个版本算不通，划界就得换一种说法。这个来回怎么走，留给第 20 章。

第三，验收。如果当时有 AI，最先出现的危险，恐怕就是那篇序言所做的事：把日心说「翻译」成托勒密体系里的一种计算方式，说它只是换了一套参数，不改变「实在」。

一个熟练的 AI 在旧典范里工作得越好，就越容易把新猜想悄悄翻回旧话语里。要拦住这一点，需要对「这到底还是不是同一个划界」做判断，而这个判断不能交给被旧典范喂大的展开者自己做。

用车和腿的话说：展开成本降下来，等于路修好了、车造好了。往哪个方向修路，目的地在哪里，是在人与车的来回里一点点定下来的，路修到哪儿，地图也跟着改。腿不再负责搬运，但腿走过的地，是这张地图的地基。

---

带走的话

**一个典范被换掉之前，旧典范是一件补了很久、穿得很熟的衣服，补丁都管用。 换衣服的压力，来自几种互不相干的别扭拧在了一起，不来自某个人突然想通。 「太阳在中间」只是一句话；把它做成一套能算、能测、能解释落体的体系，花了一个多世纪、许多人。 过去，展开的成本决定猜想能不能活下来。成本降下来以后，活下来的猜想会多得多，选择与验收就成了更重的活。**

---

## ◆ 二、广义相对论：一件补得天衣无缝的衣服，和一个小得可怜的线头

### ◆ 旧典范下的稳定显露：引力是一种看不见的拉力

第二个案例，跨度小得多：从牛顿的《原理》到二十世纪初，两百多年。这两百多年里，牛顿力学是人类有过的最成功的体系。

它的画面很清楚：空间是一个不动的、平直的大舞台，时间是处处一样的、均匀流逝的钟；物体在舞台上运动，受到力的作用；引力是一种力，任何两个有质量的物体之间都有相互的拉力，大小和两者质量的乘积成正比，和距离的平方成反比。

有了这一条，天上的月亮绕地球、行星绕太阳，地上的苹果落地、潮水涨落，被一口气统一了起来。

这套东西的成绩单长得吓人：预报彗星回归，计算炮弹的轨道，设计桥梁，推算潮汐。十九世纪中叶还有一件让人印象极深的事：天文学家发现天王星的轨道总和计算有些出入，有人据此推想，天王星外面一定还有一颗未知的行星在拉它，并算出了它应当出现的方位，后来在那个方位附近真的找到了海王星。

这是旧典范的一次大胜利：差异出现了，旧体系不改一个字，只是多添一颗未被看见的星，差异就被排进去了。

这一点要记牢，因为下面的故事里，同样的招数还会被用一次，而且用得理直气壮。

还要记牢的是：在牛顿的画面里，引力是通过空间「瞬间」起作用的。太阳这一刻移动了，地球这一刻就感到变化，中间不需要时间。牛顿自己就对这种「隔空起作用」不太满意，但他的体系跑得太好，这点不满被悬置下来，成了一个人人知道、没人动它的小别扭。

## ◆ 排不进去的差异：一个小得可怜的数

十九世纪中叶，法国天文学家勒维耶在核对水星的轨道时发现了一件事。

水星绕太阳的轨道是一个椭圆，椭圆最靠近太阳的那个点叫近日点。在牛顿体系下，行星之间互相有引力，会使这个近日点缓缓转动，也就是进动。把其他行星的拉力一项一项算进去，可以算出水星近日点每世纪应当转多少。

可是实际观测到的转动，比计算多出了一点点：每世纪大约四十三角秒。一角秒是一度的三千六百分之一，小得几乎看不见。

这么小的一个数，值得一个体系紧张吗？当然值得，因为它是所有计算都做完之后「剩下的那一点」。其他的都对上了，只有它对不上。而且其他各项都严格依赖这个体系：每一项都算对，才能确定剩下的是不是真的剩下。

怎么办？旧典范的办法现成：就像当年处理天王星一样，在水星里面再放一颗未被看见的行星，或者一团物质，让它的拉力把这点多出来的转动补上。勒维耶自己就提出过这样的猜想，那颗假想的行星甚至有了名字。

许多人去找，找了很久，没有找到。也有人提出各种办法，比如稍微改一下引力的距离规律，或者假定太阳不是正圆。每一种办法单独看都说得过去，可没有一种能圆满地、不引起别处麻烦地把这个数收进去。

这是一个典型的状况：差异序列中的一条，在旧显露下排不进去。它排不进去，并不是旧体系没有补丁可用，而是每个补丁要么付出新的代价，要么干脆不灵。同一招在海王星那里灵验，在水星这里落空，这件事本身就在说：这次的差异，不是同一种差异。

### ◆ 另一条差异序列：引力和惯性为什么是一回事

水星的线头在天上；另一条线头，在每个人的日常里，而且更早就有人看见了。

伽利略时代人们已经知道：不同重量的东西，在没有空气阻力的时候，下落得一样快。这个事实被后来更精细的实验反复检验，精度越来越高。在牛顿的体系里，这件事包含着一个暗中的巧合。

牛顿力学里，「质量」有两层意思。一层是「惯性」，也就是东西多不愿意被推动：推它的力一样，惯性大的东西加速得慢。另一层是「引力的电荷」，也就是它被引力拉得多厉害：质量大的东西被拉得多。

这是两件完全不同的事，一个是对推的抗拒，一个是对拉的敏感。牛顿体系里没有任何理由说这两个量必须相等。可是实验表明它们严格成正比，所以所有东西才下落得一样快。

旧体系可以把这个事实写下来，当作一条经验事实，却解释不了它。它是体系里一个没人去问的「恰好」。

### ◆ 纠缠成压力：几个不相干的别扭，逼向同一处

到二十世纪初，几条线开始交到一起。

第一条，狭义相对论（1905 年）。它是对电和光的研究逼出来的：光速对一切匀速运动的观察者都是一样的，由此时间和空间不再是各自独立的绝对舞台，而是联在一起的；没有任何信号能比光更快。

这个新的时空画面，是一个成功的、被后来大量实验支持的新显露。但它和牛顿引力放在一起，立刻出现了不协调：牛顿引力是瞬间起作用的，这在狭义相对论里没有位置。要么引力也得遵守「不能比光快」，要么狭义相对论得让步。

第二条，水星近日点那四十三角秒。

第三条，惯性质量与引力质量的相等。（具体到这场发生里，水星一项更多是后来的检验，而不是最初的推手。）

这三条原本毫不相干：一条来自光和电，一条来自天文观测的最后一位小数，一条来自落体的老实验。放在牛顿引力的旧显露里，它们是三个孤立的小毛病，各有各的补法。

可是当它们都指向同一个地方，即引力这一块的结构，就不再是三个小毛病，而是纠缠成了一个压力：关于引力，现有的画面从根上就有什么东西不对。

在这条差异序列的交叉点上，显影出了一位提出猜想的人。

## ◆ 猜想显露：引力不是力，是时空的形状

据爱因斯坦后来的回忆，在狭义相对论之后不久，这条序列上冒出了一个念头：一个人如果从屋顶自由落下，在他自己的感觉里，是感觉不到自己的重量的，也看不到周围的东西往下掉，因为他和周围的东西一起落。也就是说，在自由下落的小范围里，引力「消失」了。

反过来，在一个没有引力的太空里，一个加速上升的电梯，里面的人会感到自己被压在地板上，和站在地球上的感觉没法区分。这个思路称为等效原理：在足够小的范围里，引力的效果和加速的效果无法区分。

这一步，一下子就把前面那个「恰好」解释了。如果引力的效果本质上就和加速一样，惯性质量和引力质量自然是一回事，是同一件事的两面。所有东西下落得一样快，因为它们并不是被一种力不同程度地拉着，而是在走同样的「自然路径」。

再往前一步：既然在加速的电梯里，光的路径看起来是弯的，那么在引力里，光也会弯。再往前：既然引力的效果和加速一样，就不需要把引力当成一种「力」，而可以把它理解为时空本身的几何性质。大质量的东西让周围的时空弯曲，别的东西沿着弯曲的时空走最「直」的路，在我们看来就是被拉。

这是一个新划界，而且是彻底的改写。旧体系里，舞台（时空）是固定的，演员（物体和引力）在上面演；新的划界里，舞台和演员相互塑造。引力不再是平直空间里从一个物体到另一个物体的「拉力」，引力就是时空的弯曲。

同样要老实说一句：这个新划界，在最初只是一个念头，一个方向，不是一个体系。它里面有一条清晰的线，但还回答不了任何一个可以算的问题，比如水星。

### ◆ 展开成系统：数年的数学功夫

这一步和日心说相似：猜想一句话，展开要用许多年。

要把「引力是时空的弯曲」写成一套可以计算的理论，需要一门描述弯曲空间的数学。十九世纪，黎曼等人已经发展了弯曲空间的几何学，后来又有人发展了用于在任意坐标下写物理规律的张量方法。

这些数学当时主要是数学家们自己的工作，物理学家大多不熟悉。爱因斯坦起先对这种数学并不亲近，后来在他的朋友、数学家格罗斯曼的帮助下，才开始学习这一套工具，并合作写出过一个中间版本的理论，后人称为「纲要」。

那个版本有缺陷。此后的几年里，反复修改，走过弯路，得出过后来被自己放弃的结论，直到 1915 年 11 月，才给出了最终形式的引力场方程。

方程最终定形前后，最先拿来检验的，就是前面那个小得可怜的数。用新理论去算水星近日点的进动，得到的额外转动恰好就是观测到的那四十三角秒左右。没有添加行星，没有调整参数。据后人的叙述，当事人对这个结果非常激动。这一刻，旧体系里「排不进去的差异」，在新体系里成了一个自然的后果。

然后是对新预言的检验。新理论预言，光线经过太阳附近时会被弯折，弯折的角度大约是牛顿式估算的两倍。1919 年，在一次日全食期间，英国天文学家组织了观测，检验星光在太阳边缘附近的偏折，结果被宣布支持新理论，这个理论一下子名声大振。

### ◆ 展开的成本，在这里变成了数学的成本

对照一下日心说，这一次有什么不同？

日心说的展开成本，主要花在数据和观测上：更准的测量，更长的时间，更多的人。广义相对论的展开成本，主要花在数学上：要找到一门描述弯曲时空的语言，把它与物理的要求对上，把一个念头变成一组可以求解的方程，再把方程在具体情形下解出来。

最后这一步很吓人。场方程是一组非常复杂的非线性方程，爱因斯坦自己最初只能在近似下求解；后来不久，有人找到了描述一个孤立球体外部时空的精确解；再往后，黑洞、宇宙膨胀、引力波这些后果，是几十年里数学物理学家一个一个从方程里「挖」出来的，有些结论连爱因斯坦本人当时都不相信。

方程里到底藏着什么，需要大量的人和大量的时间，才能一点一点掘出来。

所以这个猜想活下来，靠的是几个条件恰好凑到一起：差异序列已经把压力积累到位；弯曲空间的数学恰好已有现成的储备；这条序列上显影出来的人，恰好有机会向熟悉这门数学的朋友学习，也有几年的时间可以反复试错；水星那个数恰好在手边，可以立刻拿来检验。少了其中一样，这一场发生可能就推迟几十年，或者以另一种形状发生。

### ◆ 若当时有 AI 帮忙展开

还是先说会改变什么。

最明显的是那几年，尤其是后半段，大量的时间花在「学数学、试方程、发现不行、退回去重来」上。

若有一个能快速把物理念头翻译成几何语言、写出候选方程、推出后果并与水星数据比对的帮手，这段功夫可能压缩得很短。那个中间「纲要」版本的缺陷，也许会更早被发现。展开出来的结果，还会反过来改写那个念头本身。

更要紧的是并排。等效原理这个起点，不是只能通向一种理论。可以同时展开几个版本：时空弯曲的、保持平直但加一个场的、修改引力规律的，让它们在同一批检验面前逐项比较：水星、光线偏折、红移。

再说不会改变什么。

第一，压力要先在。水星的四十三角秒，是一个世纪里天文学家一点一点把别的因素都算净之后才露出来的。AI 可以让这个过程快，但「其余的都对上

了，只剩这一点」这个态势，是旧体系被用到极致才会有的。旧体系被用得越透，裂缝越清楚，这靠的是旧典范内的常规科学做足了功夫。

第三，验收。一个熟练使用旧典范的展开者，对新猜想有一种天然的倾向：把它还原成已知的东西。比如把「时空弯曲」翻译成「引力场加一些修正项」，几乎就能做到数值上的一致，可是划界已经被悄悄换回去了。

判断一个展开出来的东西，到底是新的显露，还是披着新外衣的牛顿，必须有人守着。这一点，第 14 章还要专门讲。

用车和腿的话说，广义相对论的故事是一个很好的提醒：腿的经验，也就是自由下落时感觉不到重量这样的日常，是地基。不走过这段路的人，看不懂这张新地图为什么要这样画。

---

带走的话

**一个小得可怜的数（每世纪四十三角秒），能撬动一个统治了两百多年的体系，前提是旧体系已经被用到了极致，别的都对上了，只剩它。　几个互不相干的别扭（光速、水星、落体）指向同一处，才叫纠缠成压力。　猜想是一个念头；把它变成能算、能测的方程，花了好几年，还要一批现成的数学。　展开的成本一降，可以同时试几种划界，并排比较。划界怎样长出来，不会因为展开变便宜就自动解决；「这到底是不是新的」这道验收，仍要有人守着。**

---

## ◆ 三、量子力学：一张画好的大图，和抽屉里的几样东西

一九〇〇年前后，欧洲的物理实验室里，日子过得相当踏实。力学有牛顿的一套，电和光有麦克斯韦的一套，热有热力学和统计的一套。三套东西各管一摊，合起来，几乎什么都能算。年轻人进实验室，拿到的是一张已经画好的

大图，剩下的活是把图上的空格一个个填满：这种金属的导电率是多少，那种气体的比热是多少，这条谱线的波长再多测一位小数。

这就是前面几章说的常规科学，日子安稳，成果也实在。

可是实验室的桌子上，有几样东西一直摆不平。它们不大，不吵，不像一颗新星突然亮起来那样引人注目，只是安静地待在数据表里，每次拿旧理论去算，总是差一点，或者差得离谱。大多数人把它们归到「以后再说」的抽屉里。

这一节要讲的，是这几样东西怎样在抽屉里慢慢发酵，最后把整个大图顶掉。办法照着前两个案例走：不说「谁看穿了什么」，只看差异序列怎样在旧显露下排不进去，纠缠怎样变成压力，一笔新的显露怎样落下，然后是漫长的展开。

## ◆ 旧显露是什么样子

先把旧的大图看清楚，才知道哪里排不进去。

十九世纪末的物理里，有两条几乎不用想的底层默认。第一条：能量是连续的。一个振动的东西，比如一根弹簧、一束光里的电磁振动，能量可以取任何值，像水龙头流出来的水，想多想少都行。

第二条：光是波，物质是粒子，两个类别各归各，不串门。波会衍射、会干涉，粒子走直线、有位置；一样东西要么是波，要么是粒子，不可能又是又不是。

这两条默认，在这个典范里不是「假设」，而是语言本身。你要写一个方程，就得先知道方程里的字母指的是波还是粒子；你要设计一个实验，就得先知道你在测连续的量。它们是整个显露稳定下来的骨架，所以平时没人去想。

车和腿的比方在这里正好用上。旧典范像一个城市的道路网，路是修好的，红绿灯是调好的，所有的车（计算）都跑得很顺。但路网的设计里藏着一个前提：整个城市是平的，没有河，没有山。这个前提平时谁也想不起来，因为没有一辆车碰到过河。

## ◆ 第一道裂缝：炉子里的光

差异序列最先排不进去的，是一个最不起眼的东西：热炉子发出的光。

把一个封闭的空腔烧热，在墙上开一个小孔，孔里漏出来的光有一个确定的颜色分布：低温时偏红发暗，温度升高，整体变亮，颜色朝蓝白移。这件事叫黑体辐射。有趣之处在于，不管炉壁是铁是陶，只要温度相同，分布的样子就相同。这说明它背后是某种很「干净」的规律，与材料的细节无关。

十九世纪后半叶，实验做得越来越准，分布曲线画得越来越清楚：一条先升后降的钟形曲线，峰位随温度移动。理论家当然想用现成的大图把它算出来：把空腔里的电磁波看成许多振动方式，每种方式按能量均分的老规矩分到同样多的能量，再加起来。

结果是：算出来的曲线在低频一端还凑合，到高频一端一路往上冲，没有尽头。照这个算法，一个烧热的炉子会在紫外一端发出无穷大的能量。后来人们给这个荒唐结果起了个名字，叫「紫外灾难」。实际的炉子当然没有炸，曲线到高频是掉下来的。

差异就摆在那儿，钟形曲线是实实在在的，实验没错，算也没错。是旧显露里的两条默认（能量连续、均分），放进这个具体的场合，排出了一个和差异不符的形状。

此时的反应很典型，在日心说和水星近日点里都见过：先怀疑实验，再怀疑计算，再给旧理论打补丁。有人提出适用于一端的经验公式，有人在另一端凑出别的形式。补丁各自管一段，拼不成一条。

### ◆ 第二道、第三道裂缝：电子被光打出来，谱线排成队

如果只有炉子一件事，它可能永远留在抽屉里，被当作「某个还没算清楚的细节」。让它出不了抽屉的，是旁边还有别的东西也在同一个地方卡住。

第二件，是光电效应。拿一束光照在干净的金属表面，会有电子被打出来。这在波的图景里本来不稀奇：光有能量，电子吸收能量，跳出去。但实验的细节不听话。

光的颜色（频率）低于某个值时，不管光多亮，一个电子也打不出来；高于那个值，哪怕光很暗，电子立刻就出来，几乎没有延迟。电子被打出来时的最大能量，随光的频率增加，而不随光的亮度增加；亮度只管打出来多少个。

按波的图景，亮度就是能量，光越亮，电子越应该能「攒」到足够的能量跳出去。实验说不是这样。这是一条清清楚楚的差异，而且不依赖复杂的数学，拿着灯和金属片就能说清它哪里不对。

第三件，是原子的光谱。把气体通电让它发光，用棱镜分开，看到的不是一条连续的彩带，而是一根一根分立的亮线，位置固定，什么元素就是什么线。

氢原子的线最简单，一位瑞士的中学教师在十九世纪后期给出了一个把几条可见光谱线排在一起的算术公式，后来人们发现它能推广，整个氢光谱的波数可以写成两个整数平方倒数之差。这个公式是纯经验的，管用，但没人知道为什么。

再加上另一个实验事实：原子是有核的，一个很小很重的核，外面绕着电子。按经典电磁理论，绕圈的电荷会不停辐射，能量越损越少，电子应该在极短的时间内螺旋掉进核里。原子应该不稳定。可原子偏偏稳得很。

炉子里的光，光电效应，谱线，原子的稳定：它们分散在不同的实验室，用着不同的仪器，由不同的人在不同的年份报告。单独看，每一件都可以说成「一个悬而未决的难题」。

### ◆ 纠缠成压力：为什么这几件会缠在一起

「纠缠」在这个案例里是什么意思，最容易被讲成「有人把它们联系起来了」。

并不是有人把它们联系起来。是它们各自试图被旧显露接纳的时候，发现自己卡在同一个位置上。

炉子的光，卡在「能量连续」上。光电效应，卡在「光的能量连续地摊在波面上」上。谱线，卡在「原子里电子的能量可以取任何值」上。原子的稳定，卡在「绕圈就要辐射」上。四件事，四个场合，指向同一个骨架：能量可以无限细分。

这就是特征纠缠的样子：几个本来互不相干的差异，在各自的位置上，都被同一个旧默认拽住。只要这根默认还在，每一件都只能靠局部补丁处理；每

打一个补丁，系统里别的地方就多一处不自然。大图越来越像一件打满补丁的旧衣服，穿着还能穿，看着就不对劲。

### ◆ 猜想显露：一笔「成块」

一九〇〇年底，在柏林的一次学会报告里，炉子的光那条曲线有了一个新的公式，和实验在整个范围里都吻合。

关键在公式是怎么来的。据科学史上常见的说法，它最初是对已有两端形式的一个「内插」，先凑出来，再去找理由。

而要给它找理由，只有一条路走得通：假定空腔壁上的那些振动，交换能量的时候不是任意多少，而是一份一份地换，每一份的大小跟振动的频率成正比，比例常数是一个新的、要从数据里定出来的数。这个数后来叫普朗克常数。

这一笔是什么？按第 10 章的说法，它是一句话级别的新划界：能量的交换是成块的。

要注意它在当时的身份。这一笔最初的写法，更像一个「为了算对的数学手续」，在最后的算式里再让块变得无限小，回到连续。它起初被当成多大的事，科学史家的看法并不一致，这里不下断言。可以确定的是：这一笔落下去的时候，在旧显露的眼里还像个技术细节，而不是对默认的否定。

猜想的最初一笔，往往不带着「推翻谁」的标签。它出现的形态是：为了让一件排不进去的差异排进去，临时多开一个口。口开出来以后，旁边的人才发现，这个口一旦认真对待，别的地方的差异也都能从这个口里进来。

几年后，光电效应那件事上，同样的一笔换了位置再出现：假定光本身就是成块的，每一块带的能量和这块光的频率成正比，比例常数就是同一个数。光打在金属上，一块光碰一个电子，一次碰撞，要么够，要么不够。

低于某个频率，每一块都不够，所以光再亮（块再多）也没用；高于那个频率，每一块都够，所以即便暗，电子也立刻出来；出来的电子能量，等于一块光的能量减去跳出金属所要的代价。三条实验事实一下子排进去了，而且这个说法还带出了几条后来被实验检验的定量关系。

这时，一笔「成块」已经出现在两个互不相干的地方了。它不再像一个技术细节，而像一个有来头的东西。

再过几年，氢原子那边又有一笔落下：电子在原子里，只能待在某几个特定的「轨道」上，在轨道之间跳的时候，吸收或发出一块光，光的频率由两个轨道的能量差决定。

那个纯经验的两整数公式，终于有了来历：整数是轨道的编号，能量差就是谱线的位置。原子为什么稳定，也有了说法：最低那条轨道下面没有更低的了，所以掉不下去。

到这里，第三个条件齐了：这一笔不止出现在一个位置，而且在每一处都比补丁省事、算得更准。

### ◆ 压力的顶点：拼起来的不是一套

一九一〇年代到二十年代中期，量子这个词在物理里已经无处不在，但它还不是一个「典范」。它更像一堆很管用的规矩：在某些场合要把能量「量子化」，在另一些场合照旧用经典。

同时，新的差异继续冒出来。轻元素的 X 射线被电子散射以后，波长会变长，变化量与散射角有关，用「光是块」的说法一算就对，用「光是波」的说法怎么算都不对。另一方面，同样是这些「块」，又会像波一样衍射、干涉。光既是块又是波，这在旧显露的语言里，是一句没有意义的话。

然后是反过来的一问：既然光有块的一面，那原来被当作粒子的电子，会不会有波的一面？这个问题在二十年代前期的一篇博士论文里被正面提了出来，答案是：有，波长由动量决定。几年后，几个实验室用晶体或薄膜去散射电子，真的看到了衍射图样。

旧显露的第二条默认（波归波、粒子归粒子），也被顶穿了。

此刻压力的形状，已经不是「几处补丁」，而是两根承重柱子同时松动：能量可以无限细分，松了；波与粒子各归各类，也松了。人们面对的局面，是一堆有用的规矩，没有一个统一的「语言」去装它们。

## ◆ 展开成系统：两条路，一个系统

接下来的一段最容易被讲歪。通行的叙述里，往往在这里出现一位「终于想通了」的人。实际的过程要曲折得多，参与的人也多得多。

一九二五到一九二六年间，先后出现了两套看起来完全不同的数学。

一套从「可观测的量」出发。它不再追问电子在原子里「怎么转」，只问：原子发出的光线，频率和强度之间有什么算术关系？

结果是用一排排数字组成的表来代表「位置」「动量」这些量，而且这些表相乘的时候，先后次序有关系：甲乘乙不等于乙乘甲。这套东西后来叫矩阵力学。它的样子很怪，没有轨道，没有图像，只有表和表的运算。

另一套从「物质波」出发。既然电子有波的一面，就应该有一个波的方程。

方程被写了出来，解它，氢原子里允许的能量值自然就出来了，不需要「硬规定」：整数是方程在边界条件下的自然结果，就像琴弦两端固定，只能有整数个半波。这套东西叫波动力学，样子亲切得多：有方程，有函数，像是熟悉的物理。

两套看起来不是一回事，一套满是表，一套满是波，竟然给出了同样的答案。不久之后人们证明，它们在数学上是同一套结构的两种写法。

这件事的意味值得停下来想一想。典范的最初一笔（成块），最终长成的不是「一条新规定」，而是一整套新的语言：在这套语言里，「状态」「可观测量」「测量」这几个词，意思都变了。

状态不再是「电子在哪里、速度多大」的清单，而是一个抽象空间里的一个对象；可观测量不再是对象自带的属性，而是作用在状态上的一种操作；操作之间一般不交换次序。

接着，这套语言里最难吞咽的一块被补上了：那个波函数，到底是什么波？一开始有人想把它看成电子「摊开」在空间里的物质云，但这个说法很快遇到困难：电子每次被探测到，都是一个点，不是一团云。

又过不久，一个新的说法站住了：波函数的模平方，是在某处发现电子的概率。这就是玻恩规则，通常也叫概率诠释。它把一个全新的东西带进了物理的最底层：同样的初始条件，不一定得到同样的结果；能预测的是结果的分布。

一九二七年在布鲁塞尔有一次著名的会议，当时最有名的一批物理学家坐在同一张合影里。围绕「概率是不是最后的话」的争论，在这前后一直延续了几十年，本书最后一编还会回到它。

这里只需要看到：到这个时候，原来的「规矩汇编」已经长成了一个有自己语言、自己数学、自己实验检验方法的系统，而对它怎么解读，大家还没吵完。

## ◆ 漫长的「典范打造」：那一笔之后的几十年

很多人以为，故事讲到一九二七年就结束了。那只是第一章的结束。

接下来的几十年里，这套新语言要去「接管」旧显露所有的地盘。原子的结构是第一站：有了新语言，元素周期表里电子怎样一层层填，不再是经验规则，而有了来历。化学键是第二站：两个原子怎样靠电子结成分子，可以算了。

这个阶段，几乎所有参与的人，做的都是前面说的「典范内的系统打造」：用新语言去重算旧问题，建立新仪器，训练下一代学生。这是常规科学的全盛期，养活了二十世纪下半叶几乎整个物理学界。

用车与腿的比方说：路网是在那一笔之后铺出来的。铺路是庞大的工作，要几十年、几万人。但没有那一笔「河不是没有，只是还没人把桥修出来」的划界，路网铺不出来。

## ◆ 三个案例合看

现在把日心说、广义相对论、量子力学三个案例并排放在桌上，看看它们有没有共同的形状。不是要找一条「规律」，只是把三件事的轮廓叠起来，看哪里重合。

**第一，起点都是一批「排不进去」的差异。** 日心说那边，是行星的逆行和天球模型的不断加圈；广义相对论那边，是水星近日点那点多出来的进动，和牛顿引力与光速有限之间的别扭；量子力学这边，是炉子的光、光电效应、谱线。单独看，都可以用补丁处理；合起来看，补丁越来越多、越来越不自然。

**第二，这些差异都不是新发现的，而是早就在那儿。** 行星的运动，几千年都在天上；水星的进动，在测量里放了几十年。没有哪个典范靠「一个全新的发现」建起来，靠的是旧的差异被重新排列。

**第三，纠缠是通过旧默认的位置成型的。** 几件看起来无关的事，其实都卡在同一个旧默认上：天是圆的、匀速的；时空是平直的、绝对的；能量是连续的、波粒各归各类。纠缠不是人为的联系，而是它们在各自的位置上撞上了同一堵墙。

**第四，最初的一笔都很小，起初并不像革命。** 一句「能量交换是一份一份的」，一句「引力就是时空的弯曲」（这一笔后来非常大，但第一次被认真对待时，也只是一个关于等效性的小观察），一句「让地球去绕太阳」。每一笔在当时的语言里，都不太像宣言，更像是「换个坐标看看」的试验。

**第五，从那一笔到一个能用的系统，中间隔着漫长的展开。** 日心说从最初的位置互换到能够预测，要经过轨道形状的修订和力学的重写，花了一个多世纪。广义相对论从最初的想法到完整的场方程，花了近十年，验证和推广又是

几十年。量子力学从一笔「成块」到一套语言，用了二十多年；从一套语言到覆盖大半个物理，又是几十年。展开的工作量，远远大于那一笔本身。

**第六，每一次，展开阶段养活的人，比提出那一笔的人多得多。** 典范发生是稀有的，典范内的展开是密集的。社会为科学付的绝大部分工资，付给的是展开。而展开，正是 AI 现在能做得越来越好的那一部分。

把六条叠起来，三个案例共有的形状，其实是一个简单的顺序：差异先在旧显露下排不进去；排不进去的东西在同一个旧默认上缠成压力；压力里显出一笔小小的新划界；这一笔被人当真，拿去排别的差异，发现都排得进去；然后是漫长的展开，直到它自己成了新的路网。

这个顺序里，没有一处需要「某个人的天才」才能说明。每一步，都可以用「差异序列怎么排、纠缠怎么缠、哪一笔显出来」来说。

提出那一笔的人，是在这条序列上显影出来的，就像地图上的地名，是路网已经长出来以后才被标上去的。这样说不是贬低他们，只是换一个看法：他们是这场发生最清楚的可见部分，却不是它的发动机。

到那时，承担这件事的就不再是「一方猜、一方造」的两个岗位，而是第 20 章讲的人—AI 复合体：猜想在来回之间被展开，也在展开中被改写，来源越来越说不清，验收也落在这个来回里面。第四编讲展开与验收，第 20 章讲这个整体。

带走的话

**量子力学不是从一个「发现」里长出来的，而是从几件排不进去的旧差异里长出来的：炉子的光、光电效应、谱线、原子的稳定。 这几件事在不同的实验室里，卡在同一个旧默认上，能量能无限细分，波与粒子各归各类。这是纠缠成压力的样子。 最初的一笔很小：能量是一份一份交换的。它起先只是为了算对的一个手续，后来被拿去排别的差异，一个一个排进去了。 从那一笔到一套能用的语言，是二十多年的展开；从语言到覆盖大半个物理，又是几十年。养活人的是展开，不是那一笔。 三个案例叠在一起，形状一致：差异排不进去，纠缠成压力，一笔新划界，漫长展开。这个顺序里，每一步都能用显露、差异序列、纠缠来讲，不需要一个「看穿一切的人」。展开期的成本一旦被压低，整个来回的主语也要跟着换，换成第 20 章的人—AI 复合体。**

第 四 编

# AI 与典范发生：在来回之间展开、验收与比较

这一编三章，讲猜想有了以后会怎样。第 13 章讲交给 AI 展开：展开的成本压低，猜想可以一次造出几个。第 14 章讲验收，防的是新猜想被悄悄翻译回旧典范，这种“换皮”有哪几种样子，怎样认出来。第 15 章讲多个典范并排发生，怎样比较，怎样选。

三章合起来是一句话：**展开便宜了，验收和比较反而更要紧**。

第 13 章

# 猜想交给 AI 展开：典范和典范下系统的打造

设想有位退休的老木匠，爱画图。他有个习惯：想到一个新式样的柜子，先在烟盒背面写一句话，比如「柜门不用合页，靠自己的重量关上」。这一句话他能写一晚上写十张。可是十张里真正做出来的，一年也就一两件。

不是他想不出来，是做出来太费事：要画结构图，要算受力，要备料，要试装，要改。一个念头从烟盒背面走到一件真柜子，中间隔着几个月的工。所以大多数念头，死在了这几个月前面。

这个场景，差不多就是过去科学里「猜想」的处境。一句话的新想法，随时都有；能活下来的，是那些恰好碰上人手、时间和机缘，被一步一步做成完整体系的少数。这一章要讲的事，是这几个月的工，忽然可以由 AI 来做了。做的人变了，门槛也就跟着变了。

## ◆ 一句话的猜想，到底有多轻

先说清楚，这里说的「典范猜想」是什么。按前面几章的讲法，一个典范是一个稳定的显露：一批基本对象、一套划界、一种什么算问题、什么算解答的约定，彼此咬合，立得住。那么新典范的最初一笔，就是对这个显露的一次改写，改的往往只是最根上的一两处：换一个基本对象，或者换一条划界。

这样的猜想，写下来是很短的。拿一个日常的玩具例子说。假设有人琢磨城里为什么会堵车。通行的看法把「车」当基本对象，堵就是车多路窄，办法是拓路、限行、调信号。

现在有人写下一句：「堵不是车多，是车与车之间速度差的积累；基本对象是『速度差』，不是『车』。」（这是个只为说明用的玩具例子，不是什么已有的理论。）

注意这句话的分量。它就一行，却已经动了基本对象，也动了划界：原来被当作噪声的「前车一点刹车」，现在成了主角；原来最看重的「路面宽度」，退到了后台。这是一句典范级的话该有的样子：短，但是根上动了。

可是这句话离「一个典范」还差得很远。它只是一个新显露的最初一笔。

要成为一个能立住的显露，它还欠着：一套新的概念，把「速度差」说清楚；一套能算的东西，让人知道它多大算大；一批能做的实验或观测，让它有被打脸的机会；一套能教的体系，让别人能从零走进来。这些加起来，就是本章标题里说的「典范和典范下的系统」。

## ◆ 以前展开为什么那么贵

要弄明白现在的变化，得先看清过去的贵在哪里。

贵，首先贵在人工。一个新想法要展开成概念，得有人逐个定义、逐个检查有没有自相矛盾；要展开成公式，得有人推，推错了回头改；要展开成实验方案，得有人懂仪器、懂误差、懂怎么排除别的解释；要展开成教材，得有人一章一章写，还得照顾读者从哪里进门。

每一样都是熟练工的长期劳动。而且这些活，几乎全是「在已经立好的典范里」才好做的：在旧典范里，定义有现成的模板，推导有现成的套路，实验有现成的规格，教材有现成的目录。

换成一个全新的显露，这些现成的东西统统不合用，得从头一样样造。

所以过去新猜想的存活，靠的是一个很苛刻的条件：提出它的那条差异序列上，恰好同时显影出一批肯花十年二十年去展开它的人。这是个低概率事件。

许多念头不是错了，是没等到那批人手，就散了。这一点在科学史上很常见：同一个念头，在几十年后被人重新写出来，才算活过来。它先前的「死」，常常只是展开成本太高，没人付得起。

车和腿的比喻在这里可以再用一次。步行时代，一个地方值不值得去，要看你付不付得起走过去的脚力。有了车以后，许多从前去不起的地方，变成了顺路。这不是说那些地方变得更值得去了，而是到达的成本塌了，于是它们进入了可选范围。展开成本的下降，对猜想的作用，就是这个意思。

## ◆ 分工：猜想在一边，展开在另一边

现在可以把分工讲得明白一点。先打一个招呼：下面说的「一边」和「另一边」，描述的是复合体里的两种功能，不是两个先在的主体各带着一套本事来合作。猜想不是先在人脑子里长好、再交出去的，它在人与 AI 的来回里长出来；第 20 章会讲透，这里只借它的口径。

在这条差异序列上，显影出一位提出猜想的人。他手里有的，是一句话级别的东西：一个新的划界，一个新的基本对象。他有一部分说不清的直觉，来自他长期泡在某个旧典范的裂缝边上，来自他对那些排不进去的差异的敏感。这些前面第 8 章、第 9 章已经讲过了。他的这一笔，是新显露的起点。

另一边，是 AI。它擅长的，正是在一个给定的框架下，把东西大量、快速、成体系地做出来。这一点，本书从头到尾是同一个立场：凡是能在给定框

架里廉价生产的，都不是竞争点。把一句话猜想当作框架交给 AI，让它在这个框架里生产概念、公式、方案、教材，这恰好落在它最擅长的地方。

这里的说法要留意。我们不说「人有创造力，AI 只会执行」，那会又把保留地划回来。我们只描述工序：猜想这一笔在一头，系统打造在另一头；打造这一头，成本已经被压得很低。

至于 AI 能不能自己写出那一笔，这是第 19 章的开放问题，本章的论断并不依赖它。无论哪一种回答，眼前这道工序都成立：给它一句话，它可以把后面的几个月做出来。

这种分工还有一个好处，是把「好不好」的判断时间大幅提前。过去，一个猜想好不好，要等到十年后展开出来才知道。现在，几天之内就有一个完整的展开摆在面前，可以拿来摔打。猜想的好坏，从「押注」变成了「看货」。

## ◆ 怎么指挥：先让它把一句话说成概念

接下来讲最实际的部分：具体怎么指挥。这里不讲任何产品，也不讲什么神奇的口令。讲的是一种用日常经验就能想象的做法，像请一位很能干、但对你的猜想一无所知的助手干活。

第一步，是让它把猜想说成概念。

办法很朴素：把那一句话交给它，先别要结论，要它回答几个问题。这句话里出现了几个新词？每个词，在你这个猜想里，到底指什么，不指什么？这些词彼此是什么关系？哪些是最基本的，哪些是从基本的里面派生出来的？

拿前面堵车的玩具例子。你对助手说：我的猜想是基本对象是速度差，不是车。请你把「速度差」定义清楚：是哪两样东西之间的差？是同一条车道上前后车之间的，还是邻道之间的，还是不同时间点之间的？

再请你列出，从「速度差」这个基本对象可以长出哪些派生的东西，比如「差的积累」「差的消散」「差的传递」。每一个，请给出一句话定义，和一个街头可以看到的例子。

这时候助手会交回一份清单。清单里的定义有的很贴，有的含糊，有的互相打架。这些问题，正是你要的：你不是要它一次做对，你是要它把你脑子里模糊的那团东西，摊到桌面上。猜想的一大半毛病，就藏在「一句话」的含糊里；展开的第一步，就是把含糊逼成可检查的东西。

这里有个小技巧，是要求它给出「不是什么」。每个概念，让它同时写出三个很像它、但被明确排除的东西，并说明为什么排除。因为一个新概念的边界，往往靠它拒绝什么才立得住。这一点，后面第 14 章谈验收的时候，会变成非常要紧的东西。

### ◆ 再让它把概念推成公式

有了概念，第二步是让它往「能算」的方向推。

不是所有典范都要有公式。有的领域，新显露的第一批成果是新的分类，是新的图，是新的叙述结构。但凡是能算的领域，一个典范能不能算，是它能不能站稳的关键，因为「能算」意味着「能出可检查的结论」。

指挥的办法是：请助手在你给的基本对象上，写出最简单的一套关系。先要最简单的，不要它一上来就写全。比如：只用「速度差」这一个量，写一个

让它随时间变化的最小方程；说明每个符号代表什么；说明这个方程在什么极端情况下应该退化成什么我们已经熟知的东西；说明它预言什么与旧说法不同的结果。

「退化成熟知的东西」这一条值得多说一句。一个新典范，如果它的数学在合适的极限下完全不能回到旧典范已经成功的那一片，那它一般站不住，因为旧典范在那一片里是被反复检验过的。

反过来，如果它回得太顺、回得一点不剩，又要警惕：很可能它只是旧说法换了件衣裳。这个「回得多顺」的分寸，是第 14 章要细讲的。

助手交回的公式，你不必每个都能自己推出来。但你要问它：这一步是怎么来的？哪一步用了你的猜想，哪一步用了从别处借来的假设？借来的假设，是不是正好悄悄带回了旧典范的东西？这类追问，是指挥里很要紧的一个动作。它让公式的来路变得可查。

### ◆ 再让它设计会失败的实验

第三步，是实验方案。

一个新典范，最怕的是「怎么都对」。一个什么都能解释的猜想，等于没有说话。所以对助手的指令，要非常明确：不是请你设计一个能支持我的实验，是请你设计一个，如果我的猜想错了，就会看出来的实验。

具体地，让它回答这些：按我的猜想，哪些现象是旧说法也能解释的，这些不能拿来当证据，请先划掉。哪些现象，是我的猜想会预言、而旧说法预言不一样甚至预言相反的？

为了看到这些，需要什么样的观测或者测量？可能被什么别的因素搅乱？怎么排除？如果结果出来是这样，我的猜想就该被放弃，请写明「是这样」到底是怎样。

最后这一条，就是给自己的猜想预先写好死亡通知书。听起来有点狠，却是好习惯。一个猜想，一旦有了预先写好的死法，它的存活才有分量。

有些领域做实验很贵，有些根本无法做。没关系，助手可以退一步：把能用的现成数据、过去的记录、已经发表的测量，拿来重新过一遍，看新猜想对其中哪些给出不同的说法。

很多时候，新典范第一批的检验，不是新的实验，是对旧材料的重读：旧典范下被当作「误差」「例外」扔进角落的数据，在新显露下成了主角。这与第 8 章讲的「裂缝」正好扣上：那些排不进去的差异，终于有了位置。

### ◆ 最后让它写成一套能教的东西

第四步，是教材式的系统。这一步常被人忽略，却是一个典范能否成为「一层」的关键。

为什么？因为一个典范之所以是一个典范，不是它在某位提出者的脑子里立得住，而是别人能从外面走进来、学会、使用，用它去做新的事。一套没法教的东西，只是个人的见解，不是一层。

指挥的办法，是请助手站到一个完全不懂的读者的位置上。你对它说：假设有一个学过旧说法的人，没有听过我的猜想。请你为他写一个入门的路径：第一课从他已经熟悉的哪一件事讲起？

第二课怎样让他「被卡住」，感到旧说法在哪里别扭？第三课才引入新的基本对象。每一课后面给他三道题，一道巩固，一道应用，一道故意让他踩进旧说法的老路，看他能不能识破。

最后这一种题，值得多提。一套教材，如果学生学完以后，遇到新题还是下意识用旧的基本对象去想，说明这套教材没有真正把新显露立起来。它是教学效果的一面镜子，也是后面验收的一个工具。

## ◆ 指挥时常见的几个错误

指挥得多了，会发现毛病总是那么几样，先记下来，能少走弯路。

第一个错误，是一次交代太多。把整段思路、背景、猜想、期望全部倒给助手，要它「写一套完整的理论」。结果是它自己挑着重点去写，挑的重点，几乎总是旧典范里最熟的那几块。更好的办法是分步：先只要概念，检查，再要公式，检查，再要实验。每一步都是一个小关卡，你有机会在它滑走之前把它拽回来。

第二个错误，是替助手补全。助手交回的概念有含糊的地方，你顺手在心里补上，觉得「它大概是这个意思」。这一补，等于把你脑子里的原意悄悄塞进了它的稿子，后面验收时就分不清哪些是它做的、哪些是你自己想象的。含糊就让它含糊着退回去，要求写得更清楚。

第三个错误，是只问「对不对」。助手回答「对」，你就放心了。可是「对」这个字，在助手嘴里的意思，常常只是「与我见过的说法大体一致」。更有用的问法是：这一步如果错了，会错在哪里？哪一步是你最没把握的？它被逼着说出自己的软处，你才知道往哪里看。

第四个错误，是急着要漂亮。助手第一轮交回的稿子，往往粗糙，有毛边。有人看了不满意，就要求润色。润色会磨平毛边，而毛边里常常有新猜想最真实的棱角。先把粗糙的留着，用来验收；通过了，再润色不迟。

第五个错误，是把助手当权威。它说得那么流畅，那么笃定，容易让人觉得它「懂」。其实它是在你给的框架里尽力做事，框架给错了，它做得越好，错得越整齐。把它当成一位手脚极快、但并不知道你要去哪里的帮工，心态就对了。

## ◆ 放回日子里：烟盒背面的一年

再回到那位老木匠。设想他的孙女给他装了一位很能干的帮手，只要他把烟盒背面那句话念出来，帮手就能当天把结构图画出来、受力算出来、用料单列出来，甚至虚拟地试装给他看。

会发生什么？最先变的，是他存活下来的念头多了。以前十张里剩一两件，现在十张里可能剩五六件。但是他很快会发现一件更要紧的事：有一半的图，画得整整齐齐，算得清清楚楚，看着就是不对劲。

柜门靠自身重量关上的那一个，帮手画出来的图里，不知什么时候，加上了一个小小的弹簧合页。图很漂亮，可已经不是他那句话了。

这就是下一章要讲的事。展开的成本压下来以后，猜想存活的门槛确实降了；但是紧跟着，守住「展开出来的还是不是原来那句话」，成了新的要紧事。门槛没有消失，是挪了位置。

## ◆ 把日常的活也算进来

读到这里，也许有读者觉得，这些做法离自己很远：我又不造典范。其实这套工序的缩小版，日常就能用。

比如一家小公司的老板，琢磨「客户流失不是因为价格，而是因为交付时间不确定」。这是一句小小的猜想，动了一个划界：把价格这个最先被看的变量，退到后面。

他可以请助手把这句话展开：流失到底怎么算，交付时间不确定怎么量，用公司已有的单据能不能重读一遍，看哪些流失的客户其实价格并不贵。展开出来以后，他再问自己：助手是不是又把价格请回了中心？是不是把「不确定」偷偷改成了「慢」？这一问，就是验收。

规模不同，路数一样。把猜想与展开分开，把展开与检验分开，这个习惯一旦有了，面对大的猜想也不会手忙脚乱。

## ◆ 展开会改写猜想

前面说「猜想在一边，展开在另一边」，容易让人以为展开只是照着做。其实展开做着做着，会把猜想本身改掉。

还用堵车的玩具例子。助手把「速度差」展开成概念，头一件事就卡住了：速度差是相邻两辆车之间的，还是一长段路上最快和最慢之间的？提出猜想的人一看，两种都不对，他心里想的是「前车刹车这件事，传到后车要多久」。

于是那一句话改成了：基本对象是「速度差传下去的快慢」。助手照新划界再展开，又冒出一个推论：路再宽，只要传得慢，照样堵。这个推论是他原先没想到的，他觉得对，又把划界收紧了一回。

三个来回以后，桌上的猜想已经不是烟盒背面那一句。这不叫换皮：它比原来更尖，那个「不一样」还在，而且更清楚了。换皮是展开把「不一样」磨掉；改写是展开把「不一样」逼得更准。

分得出这两样，要靠下一章的验收。日常也如此：请助手按你的想法排日程，排着排着，你发现自己在意的其实是「留白」，于是回头改了要求。

## ◆ 展开稿是拿来摔打的

还有一层心态上的转变，值得单说。

过去，一套展开稿的出现，是个大事件，是几年心血的结晶，人们舍不得它，会护着它，替它辩护。现在，展开稿来得快，也就应当去得快。它是拿来摔打的东西，不是拿来供奉的东西。一份展开稿交回来，第一件事是找它的毛病，第二件事才是欣赏它的好处。

这个转变不容易，因为人对「做出来的东西」有天然的感情。可是想想那位老木匠：他若是对每张画好的图都舍不得扔，一年之后抽屉里全是半成品，还是什么也做不成。展开便宜的意思，恰好是敢于便宜地丢。丢得越痛快，留下来的越经得起看。

所以，本章讲的这些做法，其实只有一个底子：把猜想与展开分开，把展开与检验分开。猜想是原意，展开是替原意搭的脚手架，检验是问脚手架搭的是不是原来那座楼。三样东西分开了，每一样才好各自改。

## ◆ 门槛挪到了哪里

最后算一算账，看看门槛到底挪去了哪里。

过去的门槛有两道。第一道是产生猜想的差异序列本身，要有足够的敏感与积累；第二道是展开的成本，要有足够的人手与年月。第二道，现在被压低了很多。所以我们可以大致说：猜想存活的物质门槛大降。

可是这不等于「谁都可以造新典范」。因为猜想本身，要先落在真正的裂缝上；落在假裂缝上的猜想，展开得再漂亮，也只是一座漂亮的空楼。还因为展开出来的东西多了以后，怎么分辨真假好坏，反倒成了难事。这些，分别是下面两章要讲的验收与比较。

还要提防另一种错觉：以为 AI 会替你保住猜想的锋利。恰好相反，AI 展开的天然倾向是把新东西往熟悉的方向磨平。这不是它的毛病，是它的来路：它是在大量已有语言里长出来的，最顺手的路径，就是已经被走过很多次的路径。下一章，我们专门讲它。

---

带 走 的 话

**- 典范猜想是一句话级别的东西：动一两处根上的基本对象或划界。它很轻，但没有展开就不是一个完整的显露。 - 过去猜想大多死在展开成本上，不是死在错上。展开成本被 AI 压低，猜想的存活门槛随之大降，好坏从押注变成看货。 - 指挥 AI 展开，是一条朴素的工序：先逼成概念（含「不是什么」），再推成最小的公式，再设计会失败的实验，最后写成能教的体系。 - 门槛没有消失，是挪了位置：从「能不能展开」挪到「展开出来的还带不带着原来那个『不一样』」。展开会改写猜想，改得更准是长大，「不一样」被磨掉才是换皮。**

---

第 14 章

# 验收：防止新显露被翻译回旧典范的「换皮」

小区门口有家裁缝铺。你拿去一件旧西装，说想改成中式立领。过两天去取，衣服熨得平平整整，针脚细密，立领也确实立着。可是你一穿就觉得别扭：肩线、胸型、下摆，还是原来西装的那一套，只是领子上多了一圈布。

师傅很得意：「按你说的改了。」他没有骗你，他确实加了一个立领。但你要的是另一种衣服，他给你的是一件带立领的西装。

AI 展开新猜想的时候，最容易出的毛病，就是这个。它做得很勤快，很规整，很像样，但是交回来的东西，骨架还是旧的，新猜想只是领子上那一圈布。这一章讲的就是怎么在取衣服的时候，一眼看出来。

## ◆ 它为什么爱把新东西翻回旧的

先别急着怪 AI。这件事来由很自然，弄明白了才好对症。

第一，是来路。AI　的功夫，是从大量已有的文字里长出来的。已有的文字，绝大部分是在已有的典范里写成的。它最顺手、最有把握的表达，就是那些被写过千万遍的表达。当你交给它一句从未被写过的话，它要把这句话说成一整套东西，手边最趁手的材料，全是旧典范的材料。

第二，是它的「善意」。你让它展开，它想把事情做得让你满意，也想让读者看得懂。要让人看得懂，最省事的办法，就是把新东西讲成大家已经懂的东西的某种变体：「这其实就像是……的一种特殊情形。」这句话听起来是在帮你，实际上悄悄把你的新基本对象，吸收进了旧典范的某个角落。

第三，是我们自己的耳朵。人读到熟悉的说法，会有一种舒服的感觉，觉得「对，就是这个意思」。这种舒服，恰好是换皮最有力的帮手。翻译得越地道，越让人放心；而新猜想的真正价值，常常就在于它不那么地道。

所以，请把「换皮」看成一种默认倾向，而不是偶发事故。展开的成本越低，这种倾向发作得越快，越成片。以前是人手少，换皮的速度被限制了；现在一个下午就能换出一整套，装帧精美。验收如果不跟上，新猜想会在最初几天里，被友好地、体面地、成体系地吞掉。

## ◆ 换皮长什么样

换皮不是一种，至少有这么几种常见的样子，认熟了好防。

第一种，叫改名。新的基本对象被换了个名字，实际上是旧典范里一个现成的对象。回到堵车的玩具例子：猜想说基本对象是「速度差」。展开的结果里，「速度差」渐渐变成了「车流密度的另一种表示」，后面的公式，全都是对密度的老方程，只是变量名换了。名字还在，对象已经死了。

第二种，叫降格。新的基本对象被留了下来，但被放在了旧典范的某个次要位置，成了一个「修正项」「补充因素」「特殊情形」。整套体系还是围着旧的基本对象转，新的东西挂在边上，像衣服上一枚胸针。读起来像是「我们的新想法为旧理论做了有益补充」，这恰好是新猜想最不想听到的一句话。

第三种，叫调和。助手很擅长写「一方面……另一方面……」，把新旧两种说法写成一种互相补充的关系：两种看法各有道理，在不同的条件下各自适用。

这种写法特别体面，特别不得罪人。可是新旧典范之间，如果真的是划界不同，它们是在争「什么算问题」，不是各自管一摊。把争论写成分工，等于把争论取消了。

第四种，叫偷换问题。新猜想本来想回答一批旧典范答不了的问题，展开出来的东西，却悄悄去回答另一批旧典范本来就答得很好的问题，并且答得比旧的略好一点点。于是新猜想成了旧典范的一个小小的改进版。这一种最隐蔽，因为它看上去有成绩。

这四种，有一个共同点：展开出来的东西很完整，没有一处明显的错。换皮的危险，不在于错，在于正确得过于顺滑。

## ◆ 验收一：新的基本对象，真的还在吗

现在讲验收。验收要看三件事，对应三个方向的「换皮」。

第一件事，新的基本对象有没有真的保留下来。

基本对象，说白了，就是这个体系里「最底下的砖」。旧典范的砖是一种，新猜想想换一种砖。验收的第一件事，就是翻开整套展开出来的东西，数一数砖。

数法很具体。取一页公式，把里面所有的量排成一列，问：哪一个是新猜想里的基本对象？它在这一页里，是被直接用作起点的，还是是从旧的量算出来的？如果它是从旧的量用一个公式算出来的，那么它在体系里就是个「导出量」，不是砖。真正的砖，是不靠别人导出、别人要靠它导出的东西。

再看概念表。第 13 章讲过，让助手给每个概念写「不是什么」。到验收的时候，拿这份「不是什么」回过头来对：展开的后文里，这个概念有没有被用成了那些被明确排除的东西？

有时候，概念在第一章里定义得很准，到第五章已经悄悄滑成了旧概念。这种滑动，一页一页翻，容易放过；用「不是什么」清单逐章对，一对就现形。

还有一个小办法，叫删除测试。把新基本对象从整套体系里，在想象中划掉，看体系还能不能照样运转。如果划掉以后，体系几乎没变化，说明它根本没有承重，只是装饰。真正的砖，抽走了，楼会塌。

### ◆ 验收二：原来排不进去的差异，真的排进去了吗

第二件事，原来排不进去的差异，有没有真的被排进去。

回忆第 8 章。一个旧典范的裂缝，是差异序列在旧显露下排不进去的那些东西：旧划界里没有它们的位置，要么被当作误差，要么被当作例外，要么干脆看不见。新猜想之所以值得提出，是因为它想给这些东西一个位置。

这是新猜想的「业绩单」。验收要看的，就是业绩单上的每一项是不是真兑现了。

办法是把当初促成猜想的那批反常，一条一条列出来，称作「原始清单」。清单必须在展开之前就定下来，由记得这批反常原本怎么别扭的一方在展开之前写下，不能让助手代写，也不能展开之后再补。因为展开之后再写清单，就会很自然地写成「展开出来的体系刚好能解释的那些」，那是看靶子画箭。

然后对每一条，问四个小问题。一，在展开的体系里，这个差异现在叫什么，处在什么位置？二，它在这个位置上，是被体系自己「排」进去的，还是是靠外加一条特设的说明塞进去的？

特设的说明，是指为了这一条专门加的、与体系其他部分没有关系的补丁。补丁越多，说明体系没有真的容纳它。三，旧典范对这一条的处理，是「误差」「例外」或「不存在」，新体系有没有真的改变了它的身份，由「边角料」变成「正文」？

四，有没有它原本不解释、现在新解释了的差异，以及原本能解释、现在反倒解释不了的差异？后者要老实记账，不能装没看见。

这里面最能看出真假的是第二问：是排进去的，还是塞进去的。真的容纳，常带来一个让人意外的副产品：同一套安排，顺带把别的几件看起来无关的事也排好了。塞进去的，则每一条都要单独花力气，像给一件破衣服打补丁，一块一块地补。

### ◆ 验收三：有没有偷偷用回旧的划界

第三件事，最难，也最要紧：有没有偷偷用回旧的划界。

划界，说白了，是「什么算问题，什么算答案，什么算数据，什么算噪声」的那条线。旧典范有它的线，新猜想想换一条线。可线是个看不见的东西，很容易在展开里被悄悄地用回去。

怎么用回去？举几个常见的现场。

在实验方案里。新猜想说，真正值得测的是「速度差」。可是展开出来的实验，测的仍是车流量、平均车速，这些是旧划界下才认的数据；速度差只是

后来从它们里算出来的。测量的对象没有换，换的只是后期的算法。这样的实验，不管结果如何，都分不出新旧。

在「成功标准」里。旧典范有自己的成功标准，比如预测平均通行时间的误差要多小。新猜想如果也拿这个标准来衡量自己，等于接受了旧划界对「什么算答对」的规定。有时候这样做是合理的，但必须是有意的选择，不能是展开时顺手带进来的。

在教材里。教材的目录，是划界最诚实的地方。看新教材的章目，是不是还按旧典范的问题顺序在排？旧典范先讲什么、后讲什么，是它对世界的切法。新教材如果把旧的切法原样搬来，只是把每一章里的解释换成新词，说明划界没有换。

在「我们要回答的问题」里。展开出来的文档开头，常有一段「本理论试图回答以下问题」。把这几个问题，与旧典范最看重的问题对一对。如果完全重合，十有八九是偷换了问题。

新典范应当至少有一些问题，旧典范根本不认为是问题；也应当有一些旧典范的大问题，在新体系里变得不再重要，甚至不再成立。一套新体系，对旧问题全部「继续认真回答」，那是个不祥的信号。

## ◆ 把验收写成问题

上面三件事，可以压成一组随手能用的问题。拿着这些问题，对着助手交回来的东西，一条一条问，问到答不出来的地方，就是毛病所在。

关于基本对象：新基本对象在公式的哪一页第一次出现？它是起点，还是别的量算出来的？把它从体系里划掉，体系还能运转吗？第一章定义的「不是什么」，在第五章还守得住吗？

关于排进去的差异：当初那批排不进去的差异，原始清单是展开之前定下来的吗？每一条现在处在体系的什么位置？是自然排进去的，还是靠专门的补丁？有没有哪一条，其实只是被改了一个说法，仍然当作误差？

关于划界：实验测的对象，是新划界下的对象，还是旧划界下的数据加后期算法？成功标准是谁的？教材章目的顺序是新的切法，还是旧的切法换词？开头列出的要回答的问题里，有几个是旧典范根本不认的？

关于整体的味道：把全部展开稿里的「其实就是」「可以看作」「本质上是……的特例」这类话全部找出来。这是吸收的现场。每一句都问：这是真的发现了等价，还是用旧说法把新说法收编了？

这些问题不难，难的是肯一条一条去问。助手永远不会主动来问你这些问题，因为从它的角度看，它已经交付了一份无可挑剔的作业。

## ◆ 这一关由复合体里谁来守

有人会想：这些检查，让另一个 AI 来做不行吗？让它扮演挑剔的评审，专找换皮的地方。

可以让它帮忙，这能查出不少明显的毛病。但最后一道关，它守不了，原因不在于它不够聪明，而在于信息的位置。

那一句话猜想，真正想动的是哪一处，来回里一直带着它的那一部分知道得最清楚。这件事常常是说不全的。写在纸上的一句话，在不同的人读来，可以是五种意思。

提出猜想的那条差异序列上，带着一大片没有写出来的背景：当初是哪几件事别扭得睡不着，是在哪个地方觉得「不对，不是这样切」。这些没写出来的东西，才是判断「换皮没换皮」的真正标尺。别人拿到的只有那句话和后来的展开稿，从展开稿本身看，常常一切都顺。

换成日常的比喻：你的朋友给你介绍一道菜，说是「外婆做的味道」，请饭店的厨师照着做。尝出来对不对，只有你和你的舌头有数。厨师可以做得很好吃，也可以让第三个人来品尝，说「这道菜味道很正」。可是「是不是外婆的味道」，只有你能判。因为标准存在于你的记忆里，不在菜谱里。

所以这一关，守的是「原意」。这里要先拆一个误会。前面说的「提出猜想的一方」「展开的一方」，是在描述复合体里的功能，并不是两个先在的主体。猜想本来就在人与 AI 的来回里长，展开还会反过来改写它。

守关的，是复合体里那个记得这个猜想原本哪里不一样的部分。它多半是人，靠的是在每一个来回里一直攥着那个「不一样」，不靠身份。哪天这个部分落在别处，守关的位置也跟着去。第 20 章会把这件事放回复合体的整体里讲。

这也是本书坚持把「验收」排在「典范内的验收」之外的原因。第 2 章、第 3 章讲过，典范内的验收可以被廉价生产。那是因为典范内的验收标准，是公开的，写在教材里。而这里的验收，标准是新的，写在还没有写出来的原意里。

## ◆ 一次对话式的验收

光讲问题还有点干，演一遍。仍用堵车的玩具例子。

助手交回一份厚厚的稿子，名字叫《速度差理论》。攥着原意的一方先翻到第三页的公式，问：「速度差」这个量在哪里第一次出现？助手的稿子里答：定义在第二节，由相邻两车的速度之差给出。看上去没问题。再往下翻，发现第四节有一条核心方程，写的是「速度差的变化率，等于车流密度的某个函数」。

这里就该停一下。提问的一方问：为什么速度差的变化要由密度来决定？密度是旧典范的基本对象。如果速度差的演化，最终要靠密度来驱动，那么真正在底下的，还是密度，速度差只是被密度推着走的一个跟班。

这就是「降格」。他把这一条记下来，退回去，要求助手：写一个不出现车流密度的版本，速度差的变化只许由速度差自己，以及它的邻居的速度差来决定。

助手第二次交回来，方程里果然没有密度了，但是多了一个「有效耦合系数」，说明里写着这个系数「可以由实测的平均密度标定」。提问的一方又问：标定用到密度，是不是等于把密度从后门请回来了？

助手承认，是的，但说这只是为了让模型能和现有数据对接。这个理由很体面，也很危险。于是他再要求：给出一个完全不借用密度的标定办法，哪怕一开始很粗；如果做不到，请明说做不到。

第三次交回来的稿子，老实了很多：标定那一节写着「暂时只能借助密度近似，下列实验若能直接测得速度差的传递，则可去掉这一依赖」，并且把这

个实验写了出来。到这里，稿子才算有了一条真实的、可检验的路。它变得不那么完美，可是它是那句话本来的样子。

这段小对话里，有两件事值得注意。一是提问的一方并没有任何高深的专业能力的碾压，他做的只是一直盯着最初那个基本对象：它是砖，还是跟班？二是每一回助手都能迅速响应，改得很快，没有一点不耐烦。这恰好说明验收能够成为一道日常的工序，而不必是一场漫长的论战。

## ◆ 放回日子里：老木匠的那扇门

回到老木匠。帮手交回来的图，门上多了一个小弹簧合页，图画得很漂亮。

他有没有办法在三分钟内看出问题？有。他手里有一组自己的验收问题。第一，门靠什么关上？图上写的是弹簧在关，而他的猜想是重量在关。基本对象换了，没有保留。第二，他当初琢磨这个式样，是因为旧门的合页老是松、老是响。

这个毛病在新图里还在不在？如果门不靠合页，这个毛病本该消失。图里有合页，毛病就回来了。第三，帮手的说明里写着「本设计可视为传统合页门的一种改良」。这是吸收的信号，问题立刻现形。

于是他不是把图扔掉，是把要求写得更具体，再让帮手重画：门上不许出现任何弹簧、合页或其他机械件，只能靠重量和形状。第二张图可能很难看，可能根本做不到，但这才是他那句话真正的样子，才有资格被检验。

还有一点要补一句。验收的结果，也可能是：猜想本身站不住。重画几次以后，发现只要不偷偷用回旧东西，门就关不上。这是验收的另一种功劳：它

让一个猜想体面地、早早地死掉，而不是被一套漂亮的换皮包装养成一个假活人。

## ◆ 助手的几种体面辩解

验收时，助手会给出各种解释，听起来都很在理。熟悉这几种，不容易被带跑。

「这只是表述上的差别，实质是一样的。」这句话要小心。如果实质真的一样，新猜想就没有存在的必要；如果实质不一样，就不该用表述的差别来打发。追问一句：既然实质一样，旧说法下能不能做出同样的预言？做不出，就不是表述问题。

「为了便于读者理解，这里借用了通行的说法。」借用本身没错，但要问借用了多少，借用的是不是正好是新猜想要推翻的那一块。借来的东西，要一一标出来，写明哪些是暂时的脚手架，哪些将来要拆。

「该结论与现有观测一致。」一致固然好，可是要问：旧说法对这批观测也一致吗？如果也一致，这条一致不能算新猜想的功劳。

「该问题超出本理论的范围。」有时候这是诚实的边界，有时候是回避。看这个问题是不是原始清单上的：清单上的差异，不能被推到范围之外。

## ◆ 验收要多勤，多久做一次

最后还有一个很实际的问题：验收是展开完了做一次，还是边展开边做？

答案是边展开边做，并且越早越密。换皮这种事，发生得越早，后面的东西垒在上面越多，退回来的代价越大。概念阶段换了皮，到公式阶段就是整套公式换皮，到教材阶段就是整本教材换皮。

所以最勤的验收，应当放在概念那一步：新基本对象有没有立住，「不是什么」清单有没有写。这一关守牢了，后面的几关会轻松很多。

节奏上，可以这样安排：每一步交回来，先过一遍最短的三问：新砖还在吗，差异是排进去还是塞进去的，划界有没有偷用旧的。三问过了，再往下一步。

全部做完，再用完整的问题单大查一遍。这样一来，验收不是一场最后的大审判，而是一路上的小小的校准，像开车时不断轻轻修正方向盘，而不是等偏出去很远再猛打。

## ◆ 验收之外：防回音室

还有一种毛病，验收三问也查不出来，叫回音室。人划一句，AI 顺着展开，人看着流畅就满意，下一轮在此基础上再划。每一轮都更自洽，每一轮也离现象更远，因为里面的声音都出自同一场来回。换皮是新东西被旧东西吞掉，回音室是新东西被自己吞掉。

办法只有一个：让这场来回定期撞上它自己造不出来的东西。实验和观测，是撞现象；反例，是撞推论；别的复合体的批评，是撞另一种来回。验收是内部的事，这一道是外部的事，两道缺一道都不行。第 20 章会把它放在复合体的失败风险里再说一遍。

## ◆ 验收怎样才算通过

最后说说，验收通过，是个什么样子。

它不是「没有毛病」，也不是「读起来顺」。它的标志，反而常常是有点别扭：新体系里出现了旧典范读者看不惯的东西，章目的顺序怪怪的，成功标

准不那么好比，有些旧时的大问题被轻描淡写地放下了。这些别扭，是新显露的指纹。

通过的标志有几条。新的基本对象，在公式的起点上；划掉它，体系会塌。原始清单上的差异，大部分是被体系自己排进去的，补丁很少，并且有至少一件意外的收获。实验方案里，有一类测量，旧划界下根本不会去做。教材里，有一批旧典范的问题被拒绝了。开头的「要回答的问题」里，有新问题。

然后，才轮到下一章：当手上不止一个这样通过验收的展开，怎么办。

---

带走的话

**- AI 展开新猜想，默认的倾向是把它翻译回已有的语言。换皮有四种常见样子：改名、降格、调和、偷换问题。展开成本越低，换皮越快越成片。 - 验收看三件事：新的基本对象真的还是砖吗；原来排不进去的差异是自然排进去的还是靠补丁塞进去的；实验、成功标准、教材章目、要回答的问题，有没有偷偷用回旧的划界。 - 通过的体系常常带点别扭，这别扭是新显露的指纹；太顺滑反而要起疑。 - 这一关由复合体里记得「原意」的部分来守，多半是人，靠的是在来回中一直攥着那个「不一样」，不靠身份；标准存在于没写出来的原意里，不在展开稿里。 - 验收之外还要防回音室：人划一句、AI 顺着展开、双方越看越顺，会一轮比一轮自洽，也一轮比一轮离现象远。要有外部对撞：实验、观测、反例、别的复合体的批评。**

---

第 15 章

# 多个典范并排发生，再比较

装修的时候，聪明的人家不会只看一套设计图。设计师给了一套，你会再找一家，要第二套，甚至第三套。三套摆在一起，没有一套是你原来想的样子，可是你一下子明白了许多：原来厨房放在这头和放在那头，差别这么大；原来自己最在乎的是采光，不是面积。

以前之所以不这样干，是因为一套设计图要花半个月，三套就是一个半月，谁耗得起。现在三套设计图一个下午就能摆到桌面，装修这件事的性质就变了：从「把一个想法做出来」，变成了「在几个做出来的东西里挑」。

科学里的新典范，正在走到这一步。这一章讲的，就是当展开的成本低到可以并排造出几个的时候，会发生什么，该怎么比，又有哪些坑。

## ◆ 一批裂缝，几种下笔的方式

先看起点。同一批排不进去的差异，同一批相互纠缠的反常，在旧典范里挤成了压力。面对同一个压力，新显露的最初一笔，并不是只有一种写法。

这一点科学史上看得出来，只是以前看得不太清。旧天文学积累的那些越来越别扭的观测，面对的是同一个压力，但在不同的地方，长出过不同的改写方式：有的是在旧框架里继续加修补，有的是换一个中心，有的是在中间做某种折中的安排。

这些改写，各有各的拥护者，各自展开，各自吵了很久。后来的人只记住了胜出的一种，以为当初就只有一条路。

当时的几种写法，是怎么并存的？是靠许多不同的人，在不同的地方，各自花了很多年。这是非常昂贵的并行。也正因为昂贵，能并行的数量很少，常常只有两三条，并且是偶然的两三条。

现在，一句话猜想可以多写几句。同一批裂缝，可以在来回里写出三句、五句不同的新划界：换这一块砖，换那一块砖，换两块一起换，或者把某个旧的基本对象拆开。每一句，都是一个新显露的最初一笔。这几句之间，不需要谁先判定谁对。

## ◆ 并排造：各走各的工序

接下来，每一句都走完第 13 章和第 14 章讲的工序。

每一句，单独交给助手，单独展开成概念、公式、实验方案、教材式系统。要注意让它们彼此隔离：给第二句的时候，不要附上第一句的展开稿，不要让助手知道还有别的版本。否则，展开会互相渗透，几个版本悄悄变成同一个版本的不同措辞。这是并排造里的第一个小纪律：各造各的，不串味。

然后每一句，单独过验收。同一份原始清单，同一组验收问题，一套一套地过。有的在第一轮就被发现是换皮，被打回重写；有的重写三次，仍旧只能靠偷偷用回旧东西才立得住，就此作罢。真正走到最后、通过验收的，往往只剩下两三个。

这个过程本身，就已经是在筛。它的好处，是筛掉的东西很便宜。过去筛掉一个猜想，需要一个学派十年的失败；现在几天，并且筛的依据是明明白白的验收问题，不是某位权威的一句话。

还有一个副产品，常常被人忽略：并排造出来的几个版本，彼此一对照，会让每一个版本的隐含前提现形。单看一个版本，你不知道它为什么要这样切；三个版本放在一起，你忽然看见：三个都默认了同一件事，只在其他地方各不相同。

那个三者共有的默认，往往正是下一层裂缝的所在。在一个典范里待久了，这个典范的前提就看不见；并排几个，前提就露出来了。

## ◆ 比较：先说清楚比什么

几个通过了验收的版本摆在桌上，下一步是比较。比较有一个最容易出的毛病：没有说清比什么，就开始比。结果是谁讲得动听谁赢。

这里列出几个维度。它们不是公式，不用加权求和，而是几把不同方向的尺子，每把尺子都要拿出来量一量。

第一把，看对原始清单的容纳。还是那份展开之前就定下来的清单，上面写着当初那些排不进去的差异。每个版本，各自排进去了多少，排得自然还是靠补丁，补丁有多少。这是最硬的一把，因为它没有借助任何新版本自己的标准。

第二把，看代价，也就是丢了什么。任何新显露，在排进新差异的同时，总会丢掉一些旧典范排得好好的东西。旧典范在它擅长的那片地方，是经过很多年反复检验的。新版本必须交代：旧典范那些已经成功的预测，在你这里，还保得住吗？

保不住的，丢掉了多少？是丢在边缘，还是丢在核心？一个版本如果声称「旧典范的好处我全保留，旧典范排不进的我全收进来」，要小心：这样的好事，十有八九是调和换皮的另一副面孔。

第三把，看新问题的产出。一个好的新显露，常常会带出旧典范里根本提不出来的问题。看一看：这个版本问出了几个新问题？

这些问题，是有意思的、有办法着手的，还是只是换了说法的老问题，或者是无从下手的空问题？一个只回答旧问题的新典范，价值有限；一个能问出新问题、并且这些新问题能着手的，才是有后劲的。

第四把，看可死性。每个版本，在第 13 章里都应该预先写好「如果这样，我就错了」。现在把这些死亡条件拿出来比：哪个版本的死亡条件更具体、更容易触发、更接近已有的观测？一个什么情况都不会让它错的版本，等于什么也没说，尽管它看着最安全。

第五把，看可教性。第 13 章讲过，一套没法教的东西，不成其为一层。让助手为每个版本各写一套入门路径，找几位愿意当小白鼠的读者走一遍，看谁在哪一课卡住了，看谁学完以后遇到新题，真的用上了新的基本对象。教得会的，才有机会成为很多人共同使用的一层。

第六把，看生长的余地。一个版本在今天解释了这批差异，但它是不是一个死胡同？有的版本，解释完当前这批，就没有可说的了；有的版本，一旦立起来，立刻露出一大片没有被探索的地方。余地大的，在后面的竞争里更占便宜。不过，这一把最难量，也最容易自欺，所以排在后面，权重也不宜给大。

## ◆ 只看内部自洽的陷阱

比较里最要紧的一个告诫，是不要只看内部自洽。

内部自洽的意思是：这个体系自己跟自己不矛盾，概念前后一致，公式能推出来，教材逻辑顺畅。这是一个好东西，是底线。可是在 AI 能把任何一个猜想都展开成严密体系的年代，内部自洽已经没有鉴别力了。

道理很简单。展开的能力，几乎不挑猜想的好坏。一个烂猜想，交给一位勤快的助手，它照样能给你造出概念齐全、公式漂亮、实验方案像模像样、教材循序渐进的一整套。每一步都说得通，因为助手最擅长的，就是让每一步都说得通。它会在遇到矛盾的地方，自然地补上一个合理的说明，把裂缝糊平。

这一点，比起过去，是个新的风险。过去，一个烂猜想想展开成完整体系，自己就很难：展开到一半，矛盾冒出来，没法糊，就垮了。垮掉，是对它的天然检验。现在助手替它糊，垮不掉了。完整，不再说明任何事。

所以，下面这些，不能当作加分项：体系庞大；术语丰富；有很多章节；公式很多；读起来一气呵成；里面引用了很多已有的成果。这些都是 AI 展开的免费赠品，与猜想的好坏无关。

反过来，要把眼光投向那些展开不出来的东西：体系里哪里有勉强，哪里有绕路，哪里助手用了一个很长的说明才把事情糊过去。糊的地方，往往恰是猜想的软肋。

还有一种唬人，叫「包装得像典范」。典范有典范的外观：一套新术语，一个宏大的目录，一个可以套在很多领域上的万能框架。这些外观，现在批量生产毫无难度。一个猜想若是「什么都能解释」，外观一定很壮观。

可是壮观恰好是反指标：它多半是用宽泛换来的，用不断放宽「解释」的意思换来的。回到第 12 章案例里看那些真正站住的典范：它们起初窄，窄而尖，尖到可以被一次观测打倒。

## ◆ 比较的几条纪律

为了不被这些陷阱绊倒，比较要立几条纪律，都是土办法。

其一，尺子先定。在看各个版本之前，先把要用的维度和每个维度的「好」是什么样，写下来。看完版本再定尺子，人会不自觉地让尺子去迁就自己喜欢的那个。

其二，清单先定。原始清单、旧典范已有的成功预测，都要在并排造之前就锁定。后来不许增删。

其三，匿名比。把几个版本的术语换成中性的代号，抹掉表面上的文采与气势，只比实质。读起来最顺的那个版本，常常是助手最用力打磨的那个，而不是猜想最好的那个。

其四，找最强的对手。让助手以每个版本的最强批评者的身份，各写一篇对其他版本的反驳，再让每个版本回应。这样做出来的交锋，往往比任何单一版本的自我陈述都更能暴露毛病。不过要记住，这仍是助手在做事，要拿原意去验。

其五，留着死亡条件。比较完，不论选了谁，都把每个版本的死亡条件存档。将来有新观测出来，回头对一对，看是不是有被淘汰的版本其实该复活，或者当选的那个应当退场。

## ◆ 选择：不是一次裁决

比较完，要选。选择这一步，也要由复合体里攥着原意的那一部分来拿主意，多半是人，因为选择意味着往哪里投入后续的时间与年月，这件事，不能外包。

可是请不要把选择想成一次性的裁决，好像比完一次，一个留下，其余作废。更合理的做法，是分层级的选择。

最浅的一层，是选「现在往哪个版本多投一点力」。这是个小赌，随时能改。中间一层，是选「把哪几个版本留作长期并存的备选」。

有些时候，两个版本在现有的观测面前分不出高低，它们应该并存下去，等待能区分它们的新证据，而不是硬选。再深的一层，才是选「这个版本是不是值得拿去教、拿去写进教材」，因为教出去就是在制造新的一代人，代价最大，反悔最难。

所谓选择，常被想象成「找对的那一个」。更常见的情形是：今天对的那一个，很可能只在今天的材料里占优。保持几个版本活着，让它们在新的材料面前继续竞争，这本来就是常规科学时代所没有的奢侈。

从前，一个典范一旦赢了，输掉的几个就很难再有人展开；现在，输掉的版本，它的展开稿还在，只要有人想再看，就能再拿出来对一次。

这里还要提防另一种相反的毛病：选择太晚。因为展开便宜，比较也有的做，容易陷进「再多造一个，再多比一轮」的无底洞。没有谁能把比较做到万无一失。选择需要的，是在尺子量过、纪律守过之后，愿意往一个方向投入，并且承认自己是在押注，而不是在证明。

## ◆ 一个烂猜想，是怎么被造成「典范」的

举一个反面的小演练，更能说明为什么不能只看完整。

设想有人提出：城里的堵车，其实是由「人们的焦虑情绪」引起的，基本对象是「集体焦虑」。这是个随口一说的玩具猜想。把它交给一位勤快的助手，会发生什么？

第一天，助手交回概念：集体焦虑分三个层次，个体焦虑、群体传导、环境放大。每一层有定义，有小例子，读起来头头是道。第二天，公式来了：焦虑的传导用一个类似传染病的方程来写，有传导率、恢复率、环境放大系数。

第三天，实验方案来了：在高峰时段，对司机做问卷，测焦虑水平，看它与拥堵程度的相关。第四天，教材来了，十二章，循序渐进。

整套东西，没有一处不通。可是只要用前面的尺子一量，就露出了马脚。基本对象「集体焦虑」，拿什么测？问卷测出来的，是人们堵在路上以后的感受，是结果，不是原因，因果被悄悄倒过来了。

方程里的传导率、恢复率，借的是传染病模型的整套形式，其中的参数没有一个有独立的来源，全靠拟合。原始清单上的那些反常，比如「为什么没有任何明显障碍，车流也会自己停下来」，在这个体系里并没有被自然地排进去，是靠「焦虑的突然放大」这一个万能说明塞进去的。

死亡条件呢？助手写的是：「若拥堵与焦虑完全无关，则猜想不成立。」这等于没写，因为任何一点相关都能算作不无关。

可是，如果评审的人只看内部自洽，这个体系完全可以拿高分：概念严密，方程规整，实验可行，教材完整。它在表面上，比许多真正的好猜想，展

开得更像样，因为好猜想的展开，常常有毛边，有未解决的地方，有承认做不到的地方。烂猜想没有这些，因为助手替它把毛边都糊平了。

这个例子提醒我们：越是完整得没有毛边，越要多看两眼。毛边不是缺点，是诚实。

## ◆ 并排的成本账，以及不该并排的时候

并排造便宜，但不是白给，也不是任何时候都该并排。

先算账。每多造一个版本，要多一轮展开，多一轮验收，多一份比较要读的材料。展开的成本压低了，验收和比较的成本却没有同等地压低，因为这两样需要拿着原意的一方亲自上阵。

版本越多，这一方的注意力越被摊薄。所以并排的数目要有限度，通常两三个就够用，四五个已经是上限。再多，比较的质量会掉下来，反倒不如少而精。

再说不该并排的时候。有两种情形。一种是裂缝本身还没有看清。裂缝没看清，猜想就是瞎写，写多少个都是瞎写，并排只是把瞎写放大了。

这时候应当回到前面几章讲的功夫，先把差异序列和纠缠理清楚。另一种是只有一个自然的写法，别的写法都显得牵强。这时候强行并排，造出来的另外几个只是陪跑，没有比较的价值。

并排是在「有几种都说得过去的写法」的时候才有意义。它是一个放大镜，放大的是真实存在的分歧，不是凭空制造分歧。

## ◆ 放回日子里：三个版本的柜子

再回到老木匠。他那句话是「柜门靠自己的重量关上」。他请帮手出了三个版本。

第一个版本，门沿着一条斜的轨道往下滑，滑到底就合上。第二个版本，门是半悬的，重心偏一点，松了手就自己荡回去。第三个版本，门本身是个楔形，往里一推就卡紧。三个都通过了他的验收：都没有弹簧，都没有合页。

他把三个版本放在一起，发现了一件原来没想到的事：三个版本都假设「门一定是一块整板」。这是他从没有说出口的前提。他没有当场去改，只是记在了本子上，说这里也许有下一个猜想。

接着比较。他的原始清单是：旧门合页会松，会响；开关一次要两只手；小孩的手容易夹。第一个版本，三条都排进去了，但是轨道容易卡灰。第二个版本，响的毛病没了，夹手的问题反而更重。

第三个版本，三条里只排进两条，但是它让他意外发现一件好事：柜子关起来以后更不容易进灰，这是他当初完全没想过的。他把这件事认真地记了下来。

三个版本，哪个最完整？第一个。哪个最好看？第二个。他最后选的是第三个，因为意外的收获，是他最信任的一种证据：体系自己带出了没有被要求的东西。同时他把第一个版本留在了抽屉里，说等哪天灰尘的问题有了办法，再拿出来。

## ◆ 小结：从做出来，到挑出来

这一章把第四编的线索收一收。

第 13 章，猜想一句话，展开交给助手，展开成本压低。第 14 章，展开出来的，要守住原意，防止换皮。第 15 章，既然展开便宜，就并排造几个，各自过验收，再量几把尺子，不被完整唬住，最后选，并且保留可以回头的余地。

这一串事情，合起来有个好处，是把典范发生从一种「押上一生的赌」，变成了一种「可以反复练习的工序」。这对教学的意义重大。从前，典范发生没法教，因为每一回都是孤例，几百年才有一次，没有足够的练习次数。

现在，一个学生一个学期，可以经历几十回「下猜想、造、验收、比较」，每一回都很短，每一回都有可查的记录。后面几章会讲，教师怎样带学生做这件事，地基怎样练，又怎样评。

不过到这里，读者可能有一个更根本的疑问：如果展开、验收的一部分、比较的一部分都可以由 AI 做，那么最开始那一笔，那句话猜想，是不是也会被接走？这是第 19 章要老老实实回答的问题。眼下，只要记住：无论那个问题怎么回答，这一章讲的工序已经够让整个科研的层级抬高了一层。

带走的话

- 展开便宜了，同一批裂缝可以并排写出几句不同的新猜想，各自隔离展开、各自过验收，再放在一起比。并排还有个副产品：几个版本共有的默认前提会露出来，那常常是下一层裂缝。 - 比较要先说清比什么：对原始清单的容纳、丢掉了什么旧成功、新问题的产出、可死性、可教性、生长余地。尺子和清单要在看版本之前定下。 - 不要只看内部自洽。AI 能把烂猜想也造得完整严密，完整、庞大、一气呵成都是免费赠品，不再说明猜想的好坏；壮观、什么都能解释，反而是警示。 - 选择不是一次裁决：分层投入，保留几个并存，留着死亡条件，也不要无休止地再比一轮。

第 五 编

# 教学与制度：带学生换层

这一编三章，从课堂走到制度。第 16 章讲教师角色：从递答案的人，变成带学生换层的人。第 17 章讲地基：没有在典范里待够的人，看不见裂缝，所以地基要练。第 18 章讲典范发生怎么教、怎么评，也讲论文制度为什么先被压平。

三章合起来是一句话：**教学要练的是换层，不是更辛苦地走原来那一层**。

第 16 章

# 教师角色：带学生换层

设想一位老师在讲台上发作业。题目是上周布置的，一道不算难的综合题，要求写出推导和结论。发到一半，她停住了，因为前排一个学生举手问：「老师，这道题我让 AI 做了，它做得比我好，还给我列了三种解法。我是不是可以直接交这个？」

教室里很安静。老师想了想，说：「那你得讲一讲，它为什么这样做。」学生说：「我可以让它讲。」

这个场景现在每个教室里都会发生，只是形式不同。有人把它当成作弊问题，有人把它当成技术问题，有人把它当成教学评价的危机。这一章想换一个看法：它是教师角色的问题。老师原来站在讲台上，是因为有东西要传递，而现在那件东西，学生掏出手机就有了。

那么老师还站在那里做什么？

## ◆ 教师原来是传递元素的人

先把「元素」这个词说清楚。前面几章已经讲过，一个典范里有它的基本对象、基本问题、标准方法、认可的答案和验收的规矩。把这些东西拆开来，一块一块地看，每一块都是一个元素：一条定理，一个公式，一种实验技巧，一个标准推导，一类典型题的套路，一个领域里「大家都这么做」的默契。

传统的教学，说到底是在传递这些元素。老师把定理讲清楚，把推导一步步走给学生看，把例题做一遍，再布置几道相似的题，让学生练到熟。学生的水平用什么衡量？

衡量他拿下了多少元素，拿得牢不牢，能不能在新题目里把几个元素拼起来。考试考的是这个，作业查的是这个，论文答辩查到最深处，也还是查这个：你在这个领域的标准做法里站得稳不稳。

这套安排并不愚蠢。进入一个领域，确实得先把那里的元素拿到手。就像学开车，先得认识油门、刹车和方向盘。老师的价值，在于他自己已经拿下了这些元素，而且知道哪个先学、哪个后学、哪里容易卡住，能把一大堆东西排成学生走得动的台阶。

请注意，这里说的是「传递」。元素在老师手里，通过讲授、示范、批改，转到学生手里。这是一个由有到无再到有的搬运：老师有，学生没有；经过一个学期，学生也有了。

搬运的价值取决于一件事：搬运是稀缺的。元素放在老师脑子里和少数几本书里，别人想拿到，就得坐进教室，花上几年。稀缺托起了教师这个行当，也托起了学校。

现在这个稀缺没有了。前面第 1 章讲车替代腿，讲的就是这个：一件原来要靠人力逐段完成的事，被一种成本低得多的东西接走了。

典范里的元素，从定理到套路，AI 现在随问随答，并且能按提问者的水平调整讲法，能一遍一遍重来，永远不烦。就「把元素从有的地方搬到没有的地方」这件事而言，老师再怎么努力，也不会比它更便宜、更耐心。

这话说得直，却没有什么可难过的。腿曾经是运输的主力，车来了，腿并没有消失，腿去做了别的事。老师也一样。只是这一回，需要弄清楚的是：老师手里除了元素，还有没有别的。

## ◆ 评价若停在元素层，就和 AI 正面撞车

先看一个很具体的后果：评价。

假设一门课的评价方式是：期末考一套题，平时交几份作业，再写一篇综述或者小论文。这三样东西，考的全是「元素拿得牢不牢，能不能拼起来」。

考场上不让带手机，是在人为地把 AI 挡在门外，让学生赤手空拳展示他脑子里装了什么。作业和论文管不了，老师只能靠检测、靠口头追问、靠要求提交草稿和修改记录。

这里有一种熟悉的动作：评价的对象没变，只是在想办法把 AI 拦在外面。拦得住，就继续用旧办法；拦不住，就加一道审查。这就是前面几章反复谈到的流行方案：加强审查，看重过程，要求学生提交推导的记录，让他们当场解释。

这些办法为什么注定撞车？因为评价如果测的是「典范里的元素你拿到没有」，那它测的恰恰是 AI 最拿手的东西。一个测验，如果 AI 比学生答得好，那么测验测出来的「好」，已经不再标志任何稀缺的东西。学生交上来的分数，和一个人会不会开车无关，只说明他在考场里被禁止用车的时候，腿力如何。

有人会说，那没关系，基础总是要考的，哪怕 AI 会，学生也得会。这句话有一半是对的，另一半要小心。对的那一半，下一章专门讲：在典范里亲手待够，是一种训练，不能省。

要小心的那一半是：如果评价只剩下「检查你亲手做过没有」，而它又没有任何一个更高的指向，学校就变成了一个监督人们不开车、练习走路的地方。练习走路本身当然有价值，可是没有人会为了这个付学费，也没有哪个社会会为了这个养一支庞大的教师队伍。

再往深处想一层。元素层的评价，本质上是在问：「你是不是已经在这个典范里？」这是个入口的问题。入口当然要把关，但把关本身不是教育要去的地方。教育要去的地方，在入口的另一边。

这里面还藏着一个更隐蔽的问题。评价如果长期停在元素层，教学就会向评价靠拢，因为学生和家长会盯着评价看，老师也会。于是整个教学过程被拉成了一条细细的管道：怎么考，就怎么教，怎么教，就怎么学。

管道的出口是一个分数。管道中间发生的任何别的东西，比如学生对一个看似完美的推导觉得哪里不对劲，比如他盯着一个定理的条件琢磨「如果不满足会怎样」，都被当成噪声，因为它们不会提高分数。

等到 AI 成为管道里最便宜的组件，这条管道就空了。里面能流动的，只剩下那些噪声。AI 来了之后，原来被当作噪声的那部分东西，反成了教学里最值钱的东西。

## ◆ 据转述的三点：它们是路，不是终点

下面要谈到丘成桐教授。先交代一句：以下内容据转述，本书没有核对过原文，所以只写转述的大意，不引他的原话。

据转述，丘教授在谈到 AI 时代的教育时说，答案变得廉价了，学校能交付的是过程；在这个意义上，学校要从答案转向三件事：会不会提问，能不能验证，愿不愿意把过程讲清楚。

这三点非常有分量，应当认真对待。它们出自一位在数学里工作了一辈子的人，说的是他在真实的研究工作里看重的东西。会提问，是因为题目不会自己跑出来；能验证，是因为数学和科学最终要经得起检验；愿意把过程讲清楚，是因为研究是公共的事，不讲清楚，别人无法接着往下走。

按本书的框架来读这三点，会得到什么？

第一，它们都是典范内部的好品质。会提问，指的是在一个领域里问出有分量的问题，这里「有分量」的标准，由那个领域的典范给出。能验证，指的是照这个领域认可的规矩，核对一个结论对不对。

愿意把过程讲清楚，指的是按这个领域通行的表达方式，把推导交代给同行。三件事都预设了一套共同的规矩，有人认可它、有人遵守它。这个共同的规矩，就是典范。

第二，也正因为如此，这三点是进入典范、并在典范里站稳的路径。一个学生如果会提问、会验证、愿意讲清楚，他就真正「在」这个典范里，而不是只会背诵它的元素。从这个意义上说，丘教授的三点比单纯的「传授知识」前进了一大步，因为它把重心从「拿到什么」移到了「怎么做」。

第三，按本书的看法，这三点本身并不构成终点。原因前面已经讲过：答案、过程、提问、验收是同一层里的四个位置，它们一起被 AI 接走。AI 在给定的典范里，能提出很多质量不低的问题，能做验证，也能把推导讲得清清楚楚。

如果学校把交付的东西定位成这三件事，那么这三件事的成本也会和答案一起下降。教师靠这三样去安身立命，依然是在和 AI 抢同一层的饭碗。

有人看到这里，可能会觉得本书是在批评这三点。其实不是。这三点是对的，只是它们的位置需要重新摆放：它们是一条路，路的这一头是元素，另一头是典范；走完这条路，学生才站在典范的边缘上，才有可能看见典范的边。至于站在边上之后做什么，三点没有说，也不是它们的任务。

用车和腿的比方来说：丘教授给出的三点，相当于告诉我们，一个人要把腿练好，要会走路，会认路，愿意把路线画给别人看。这些都对。可是在车出现之后，真正的问题变成了：交通怎么升级？腿练得再好，也只是腿。而腿又恰恰是不能省的地基：一个从没走过路的人，看不懂地图，也就不知道车应该开向哪里。

## ◆ 学校交付的不是「过程」，也不是过程的记录

「学校能交付的是过程」，这句话（仍是据转述）流传很广，值得多看一眼。它容易被读成两种样子，两种都有问题。

第一种读法：学校交付「过程」，所以要更重视学生做题时的思考步骤，要求写出完整的推导，要有草稿、有讨论。这样做没有错，但它假设了「过程」是一个可以交付的东西。可是过程发生在学生身上，从来不是一个可以从老师手里递到学生手里的物件。

第二种读法更麻烦：既然看重过程，那就记录过程。要求学生提交对话记录、修改历史、思考日志，由老师审阅，以证明这些东西是他自己想出来的。

这里的问题更清楚：一份过程的记录，是一个文件，文件可以被生成。AI 能写出一份看上去逼真的「思考日志」，包括犹豫、回头、修正，这在技术上并不难，将来只会更容易。

记录一旦成为评价的对象，它就会被工业化地生产出来。我们不必为这件事感到愤怒，这是同一个规律在起作用：凡是能被廉价生产的，都不再是稀缺点，包括「看起来像过程的东西」。

所以说，学校交付的不是过程，更不是证明过程的记录。那交付什么？

不妨问问自己：一个学生在一所学校里待了几年，走的时候，他身上发生了什么变化，是别处很难替代的？

答案是：他经过了一些事，经过之后，他看事情的层次变了。

他不只是知道了更多东西，不只是会做更多题，而是他知道自己所在的这个典范是一个典范，知道它的边界在哪里，知道哪些地方是它排不进去的，知道一个领域里「大家都这么做」背后，有什么是被默认的。他换了一个位置站，从那里，他看得见原来看不见的东西。

这就是「换层」。换层不能被记录，也不能被提交。你可以证明一个人做过某事，但无法用一份文件证明他因此换了层。换层只能在后来的事情里被看出来：他遇到一个陌生的情况时，是被它吓住，还是能认出「这是旧办法排不进去的地方」。

因此，学校真正要交付的，是让学生在经过之后能换层。过程当然是必要的，没有过程就没有经过。但过程是经过的方式，换层才是经过的结果。把方式当成结果，是这个时代教育讨论里最常见的错位。

用一个家常的例子来讲。设想一个学做木工的徒弟，在师傅的作坊里待了三年，刨了无数块板子，凿了无数个榫头。

三年后，师傅给他一块从没见过的怪木料，他拿在手里掂一掂，敲一敲，说：「这块不能用老办法，纹理是拧的。」这句话里，装着三年的刨与凿，但它不是那些动作的记录，也不是任何一块板子的成品。它是一种看的方式。徒弟交付给自己的，是这个。

现在的教育该问的是：我们的课程里，有哪些地方是在培养这样的看？

## ◆ 带学生换层：教师手里还剩什么

回到那位发作业的老师。她手里已经没有「比 AI 更好地传递元素」这件事可做了，那她还有什么？

一个很清楚的答案是：她自己在一个典范里生活了多年，比学生更熟悉那里的纹理，也更了解哪里是缝。这不是一般意义上的知识，是一种熟悉，一种对「这个领域在哪里容易出毛病」的嗅觉。

这种熟悉，只能通过一段长时间的、亲手的、反复的经历长出来，而且它是具体的、带着上下文的，不是 AI 随手可以给学生的一份清单。

有了这一点，教师的角色就可以重新描述了。过去是传递元素的人，现在是带学生换层的人。具体来说有三件事。

第一件，带学生进入。进入不是浅尝辄止，而是真正待在一个典范里，足够久，足够深。老师要做的是把入口守住：让学生亲手做该亲手做的东西，不要被 AI 的流畅取代了他自己的磨练。这一点第 17 章会展开。

第二件，带学生看边。学生做得熟了，老师要有本事让他看到这个典范的边在哪里。这个动作很像带人登山：到了半山腰，让他回头看一看走过的路，看一看山的形状。

老师会在恰当的时候说一句：「你们有没有觉得，这类题总是做不完？」或者「这个推导，每个人都这么写，可为什么必须这么写？」这些话都很小，却是把学生从「在里面」挪到「看着里面」的撬棍。

第三件，带学生试着往上。看到边之后，老师要给他们一个安全的地方，让他们敢于说出一个不属于这个典范的猜想，哪怕只是一句话，哪怕多半是错的。

老师在这里的角色，不是评判猜想的对错，而是让猜想活下来，并且帮助学生用 AI 把这个猜想展开成一个可以和旧典范并排比较的东西。后面第 18 章会谈到如何教、如何评。

这里的「试着往上」，也不是要学生独自把一个典范生出来；按第 20 章的说法，猜想多半是在学生与 AI 的来回之间长成的，学生要练的，是在来回中不丢掉自己那句话。

这三件事里，没有哪一件是 AI 在「代替」老师。相反，AI 在每一件里都可以是帮手：学生进入时，AI 可以做陪练；看边时，AI 可以快速跑出大量例子，让边的位置暴露出来；试着往上时，AI 可以把猜想展开成完整的体系。老师的位置在这些动作的组织与取舍上，也在他自己的那份熟悉里。

还有一个必须说的细节：这样的老师，自己也在换层。一个只会传递元素的老师，如果让他带学生换层，他自己没换过，就会教不动。所以教师培养、教师评价，同样要面对这个转变。

这是一件不小的事，靠培训班解决不了，只能靠整个学术环境逐渐改变评价的重心。这个话题超出本章的范围，但应在这里留下一个记号：谁来带老师？答案往往是，老师们自己互相带，在同一个裂缝面前一起站着。

## ◆ 放回日子里：一节课可以怎么上

我们来看一节具体的课，看看老师的动作和过去有什么不同。

假设是一节大学二年级的力学课，讲的是某个标准问题的解法，课本上有一套经典的推导。过去的上法：老师在黑板上把推导写一遍，边写边讲，学生抄笔记，课后做习题。

现在可以这样上。上课前，老师让学生用 AI 把这套推导跑一遍，并且把他们觉得「最顺的地方」和「最别扭的地方」分别标出来。上课时，老师不先讲推导，而是请几个学生说说他们的标记。

有人说，这一步为什么可以把那一项略去？有人说，这个近似好像凭空冒出来。老师把这些别扭的地方写在黑板上，然后问：「如果这一步不能略去，会怎么样？」

于是整节课围绕着典范里那些「被默认」的地方展开。老师不一定都能解答，也不必假装能解答。

有的别扭是真有道理的，课本背后有一整套论证，老师就把它讲出来；有的别扭，是这个领域当年留下来的权宜之计，老师就坦白说，「这里确实有点

含糊，后来人是这样处理的，但没有谁觉得它十分干净」；还有的，他会老实承认，自己也一直没想明白。

这一节课结束时，学生拿到的东西，除了推导本身，还多了一份对推导「哪里是结实的、哪里是糊上去的」的感觉。这份感觉是换层的材料。它不会出现在期末试卷里，但学生将来遇到一个新问题的时候，会在某一步停下来想：「这里是不是又被默认了什么？」

这样的课，对老师的要求其实更高了。他必须真懂这个领域，并且懂得它的缝。这一点没有捷径。

## ◆ 一场小小的争论：这不就是又换了个考法吗

写到这里，有位读者可能会提出反对。我们不妨把这个反对请进来，让它说完。

他会说：你讲教师要带学生「看边」、「试着往上」，听起来很美，可落到评价上，不还是要出一套题、打一个分吗？只不过题目从「推导这个定理」换成了「指出这个定理哪里含糊」。AI 一样能做。你让它指出含糊的地方，它指得头头是道。这不过是把平移从「答案」挪到了「找缝」，还是在同一层里转圈。

这个反对很锋利，应该认真回答。

先承认它说对的部分。是的，「指出一个定理哪里含糊」这样一道题，AI 在多数情况下答得不差。它读过大量的评论、批注、教科书里的注脚，能把一个领域里公认的含糊之处列得清清楚楚。如果评价只是让学生复述这份清单，那和复述定理没有分别，同样会被 AI 抹平。

再看它没说对的部分。这里要分清两件事：一件是「说出某个领域里已经公认的含糊」，另一件是「在一个具体的、没有人预先整理过的情境里，认出此处排不进去」。前一件是元素，后一件是嗅觉。前一件是别人已经把裂缝画在地图上，学生读地图；后一件是学生自己走在路上，脚下一滑，知道这里有东西不对。

AI 当然也可以走路，它也可以在一个陌生的情境里指出不对劲的地方，这一点前面已经交代过，没有必要再划一块「AI 做不到」的保留地。

要紧的是教学该落在哪里：评价如果只检验学生有没有一份清单，那它无论换成什么形式，仍在元素层；教学的重心如果放在让学生反复经历「脚下一滑」，让他自己长出这份嗅觉，那么清单只是副产品。

至于将来这份嗅觉要怎样被检验，怎样在制度里体现，第 18 章会专门讨论。这里只需要记住，换考法解决不了换层。考法是纸面上的东西，换层发生在人经过的事情里。

读者还可能问第二个问题：既然 AI 什么都能做，学生为什么还要费劲经历「脚下一滑」？让 AI 替他滑，再把结果告诉他，不是更省事？

回答是：省事的是结果，省不掉的是经过。我们前面用车和腿讲过这件事：车可以把人运到目的地，运到之后，人并没有因此认得路。认路这件事，要靠自己走过。认路并不是目的，但没有认路，就看不懂地图，也就不知道该让车开向哪里。

同样，让 AI 替学生滑了一跤，学生拿到的是一句话：「此处有一处含糊。」他读到这句话，就像读到一张写着「前方有坑」的纸条。他并没有被摔

疼过，也就没有那份在下一次自己察觉的敏感。这就是下一章要谈的「地基」。

## ◆ 教师自己的难处：不知道的时候怎么站在讲台上

还有一个比较少被说起的问题：教师怎么办。

带学生看边，意味着老师要常常站在自己也不确定的地方。过去的讲台，有一种默认的姿态：老师是知道的人。学生提问，老师回答。回答不了，是老师没备好课，是失职。现在，要老师在课堂上说「这里我也没想明白」，需要一种不太容易有的底气。

这份底气从哪里来？一部分来自体制：如果教师评价还是看学生分数、看课堂是否「流畅完成」，老师当然不敢暴露不确定。所以评价的转变和教师的转变，必须一起发生。另一部分来自一种教学上的诚实：老师清楚地区分三种「不知道」。

第一种，是「这个典范里有现成答案，我一时没想起来」。这是普通的不知道，回去查一查就好。AI 在这里是老师的好助手，查得快、讲得清。

第二种，是「这个典范里没有现成答案，大家都在这里卡着」。这就是裂缝的所在。

老师可以大大方方地对学生说：「这里是这个领域的一个老毛病，很多人想过，没有谁让它安静下来。」老师不必有答案，他有的是对这个位置的认识，对它为什么卡、前人试过什么、哪些尝试为什么不行的认识。这份认识，比答案更值钱，因为它是带人往上走的路标。

第三种，是「我不知道这是第一种还是第二种」。这最诚实，也最有教育价值。老师把这个判断的难处如实摆给学生看：怎么判断一个「不知道」是自己没学到，还是整个领域都没解决？

这个判断，本身就是对一个典范熟到什么程度的检验。学生在一旁看着老师为这个判断犯难、查资料、去问同行，学到的东西，是任何一份完整的参考答案都给不了的。

所以，带学生换层的老师，不是无所不知的老师，而是知道自己在哪里知道、在哪里不知道、在哪里全世界都不知道的老师。

这样的老师，AI 无法代替，倒不是因为 AI 缺这份知道，而是因为这份分辨要在一个具体的学生面前、一个具体的场合里发生，它有一种共同在场的重量。老师坦白地不知道，学生才敢坦白地不知道。一个班级里，如果没有人敢说「我不懂」，就没有人会看到裂缝。

## ◆ 放回日子里：一位老师的一学期

再讲一个更长的场景，看看这种角色换下来，一学期会是什么样子。想象一位教统计思维入门课的老师，下面不是真实的某位老师，只是把这类课堂可能发生的事排成一串。

开学第一周，她不讲概念，让学生先拿一个小数据集，用 AI 做完整的分析，写成一页报告。每个人的报告看起来都不错，图表整齐，结论清楚。

第二周，她收上来，没有评分，而是把几份报告并排贴在墙上，问：「这几份结论很像，你们觉得是不是它们都对？」有学生说是，有学生犹豫。她便

问：「这批数据，是怎么来的？」没人答得上来。AI 当时没有问，学生也没有问。大家发现，整个分析建立在一个没有人检查过的前提上。

第三周到第八周，她带着学生做另一件事：亲手算。不用 AI，用纸笔和最简单的工具，把几个经典例子从头算到尾，每次算完，花时间讨论「哪一步我们其实是在相信什么」。

学生起初觉得慢，抱怨，说这有什么意思，AI 一秒钟就出来了。她不争辩，只是每次在结束时，问一句：「今天我们算出来的东西里，有没有哪一个你现在不太放心？」

第九周以后，她又让学生拿起 AI。这一回，学生的用法变了。他们不再只问「请帮我分析」，而是追着问：「你这里假设了什么？如果这个假设不成立，会怎样？」他们开始能发现 AI 给出的分析里，哪里用了一个看上去自然、其实可疑的默认。

学期末，她布置的最后一份作业，是让每个学生写一句话：「在这门课教的统计思维里，我觉得有一件事，它排不进去。」这句话不要求正确，只要求具体，指向一个他们自己遇到过的情境。

她收上来的这批句子，大部分很稚嫩，有些是已经被人想过一百遍的老问题，有些是误解。但夹在其中，有两三句话让她停下来想了很久。她没有告诉学生这些是不是「对的」，她只是在最后一堂课上读给全班听，并且说：「这个问题，我不知道答案。」

这个学期里，她没有一次在传递元素，元素都在，也都学了，但学生是通过经历学到的。她所做的，是排布顺序：先让学生享受 AI 的流畅，再让他们

撞到一个 AI 没有替他们想到的地方，再让他们动手，再回到 AI，最后让他们开口说话。每一步都是为了让他们在典范里逐渐站稳，又开始看见它的边。

注意，这一学期里发生的事，不是靠一个妙招，而是一连串小小的安排。一个老师能做的，大概也只有这些：把课堂变成一个让换层有机会发生的地方。至于哪个学生会在哪个时刻换，是不能保证的。换层不是被教会的，是被促成的。

### ◆ 从学生那边看：他们要的是什么

前面一直是从教师这边说的，也该换到学生的位置上，看他们在这个转变里期待什么，又担心什么。

学生担心的，常常比老师以为的更实在。一个大学生，知道自己要找工作，要申请研究生，要拿到一份拿得出手的成绩单。他看到老师在课堂上讲「换层」，心里想的是：这能给我换来一个好的绩点吗?

如果老师嘴上说看重换层，评分却还是看元素，学生马上就会明白，哪个才是真的，然后按真的那个去行动。这不是学生势利，这是学生很会读信号。

所以教师带学生换层，不能只靠口头号召，必须在课程的结构里给它一个位置：有时间让学生做，有机会让他们犯错而不被惩罚，有一部分评价，是围绕他们遇到的具体别扭展开的。这些安排看起来琐碎，却决定了号召是不是可信。

再往深处说，学生内心的期待，其实有一部分和老师设想的很接近。许多学生在用过 AI 之后，隐隐有一种空：东西做出来了，自己却不知道做了什么。他们不是不想学，是不知道在 AI 面前，学什么才算数。

这时候，一个老师如果能指给他们看：这里有一块地方，是你经过之后才会有的，是别人给不了你的，他们往往是欢迎的。换层的说法，不必讲得很玄，学生要听的是一句家常话：你不是白来了，你从这里走出去，会看得见一些原来看不见的东西。

当然，也有学生并不想换层，只想拿到证书，早点工作。这没有什么可指责的。教师不是要强迫每个人上楼，而是要保证想上楼的人有楼梯，并且让大多数人在经过的时候，至少看见楼梯在哪里。这一点，在制度层面意味着什么，留给第 18 章。

## ◆ 一个必须守住的区分

最后，为了防止本章被读偏，要守住一个区分。

「教师从传递元素变成带学生换层」，不是说元素不重要，也不是说学生可以跳过元素直接去做大事。恰恰相反，换层是从元素层之上看元素层，没有元素层，也就没有可看的东西。就像不走过路的人看不懂地图，没有进入过典范的人，无从谈起上升。

更不是说「换层」是一句口号，可以直接贴在课程大纲上，然后就算完成了。换层是一件需要工序的事，需要在典范里待足够久，需要对裂缝有感觉，需要敢于说出一句不属于旧典范的话，还需要有一个靠得住的验收，不让这句话被翻译回旧典范的老样子。这些环节，后面几章会一个一个讲。

这里先把教师的位置交代清楚：他不再是元素的搬运工，也不是过程的监工，而是一个在典范里生活得足够久、知道它的纹理、愿意领着年轻人走到边上的人。

还有一层，留给第 20 章细说：学生将来待的地方，多半是人与 AI 来回的复合体。

所以教师带出来的，不是一个能单枪匹马发生典范的人，也不是一个只会「管好 AI」的人，而是一个在来回里有用的人：能攥住猜想的「不一样」，能从展开里读出露出的缝并据此改划界，能在并排的几个版本里选一个、说出理由，脚下还有在典范里待够的地基。接下来几章的练法，都朝这四样去。

---

带走的话

**教师原来传递元素，靠的是元素稀缺；元素不再稀缺，教师的位置就要换。 评价若停在元素层，测的恰好是 AI 最拿手的东西，两边必然撞车。 据转述，丘成桐教授提出的三点（会提问、能验证、愿意把过程讲清楚）是好品质，但按本书框架，它们仍在典范之内，是进入典范的路，不是终点。 学校交付的不是过程，更不是过程的记录；记录可以被生成。学校要让学生在经过之后能换层。 教师因此成为带学生进入、看边、试着往上的人，而他自己也得先走过这一程；带出来的，是在人与 AI 的来回里有用的人，不是单枪匹马的人（见第 20 章）。**

---

第 17 章

# 地基怎么练：在典范里待够，练出对裂缝的敏感

设想一位做了三十年钳工的老师傅，退休后常被请去厂里看新来的年轻人干活。年轻人手里有图纸，有数控机床，有测量仪器，样样比他当年强。老师傅往往不说话，站在旁边看一会儿，伸手摸一下刚出来的工件，说：「这个，换一批料试试。」后来一查，果然是料的问题。

年轻人问他怎么摸出来的。他说不上来，只说：「摸多了。」

这句「摸多了」，是这一章要讲的东西。前一章说教师要带学生换层，换层的前提，是学生在一个典范里待得足够久，久到手上有感觉。

这份感觉不是知识，也不是技巧，是介于二者之间的一种嗅觉：知道什么时候东西不对劲，哪怕说不出为什么。有了它，才谈得上看见裂缝；没有它，所谓「批判性思维」，往往是一句空话。

可是，AI 来了以后，「摸多了」这件事面临一个前所未有的难题：现在谁还需要亲手摸？

## ◆ 亲手做：从完成任务的手段，变成训练

先回答这个难题。

过去，学生亲手做题、亲手推导、亲手跑实验，有两层意思搅在一起。一层是为了得到结果：作业要交，实验要出数据，论文要有计算。另一层是为了

练出本事。这两层长期粘在一起，没有人觉得需要分开，因为要得到结果，就必须亲手做，亲手做的过程中，本事自然就长了。

现在这两层被 AI 一刀劈开了。结果可以不经过亲手做就得到，甚至得到得更好、更快。于是第一层的理由消失了：没有任何必要为了「得到结果」而让学生亲手推一遍。

如果这时我们仍然要求学生亲手做，原因只剩下第二层：为了练。亲手做不再是完成任务的手段，而成了训练本身。

这句话听起来平淡，含义却很大。训练和完成任务，是两种不同的活动，有不同的规矩。

完成任务，追求的是又快又好。能借助的工具，都应该借助。一个工程师要造桥，用最好的软件、最强的计算，天经地义，没有人要他先用算盘把应力算一遍。

训练，追求的是长出能力。训练的好坏，要看经过这件事之后，人身上留下了什么。一个跑步的人，不会用摩托车来完成训练里的十公里，那不叫偷懒，那叫根本没在训练。同样，一个学生用 AI 替他完成了本该用来训练的推导，那份推导的「结果」拿到了，训练没有发生。

所以，我们要在教学里把这两件事明明白白地分开：哪些时候是在完成任务，AI 请进来；哪些时候是在训练，AI 请出去，或者换一种方式请进来。过去教师不用区分，今天必须区分。这个区分本身，就是新的教学设计的起点。

用车和腿的话说：车出现之后，没有人再为了运货而锻炼腿。可是人为了健康、为了登山、为了在没有路的地方行走，仍然锻炼腿，而且锻炼的方式比

过去更讲究，因为不再有「干活就是锻炼」这个副产品了。训练要变成一件有意识的事，有目的，有计划，有节制。

这对教育是一个重新设计的机会。过去很多「亲手做」的东西，其实是低效的、重复的，学生是为了应付，不是为了练。AI 时代把这些东西清掉，反而逼着我们弄清楚：哪些亲手做真的有训练价值？答案是：那些能让学生在典范里留下「手感」的东西，那些能让他和典范的纹理发生摩擦的东西。

## ◆ 为什么必须待够：嗅觉从哪里来

下面谈核心：对裂缝的敏感，是怎么长出来的。

回到前面的几章。典范的裂缝，是差异序列在旧显露下排不进去的地方。听起来是个很抽象的说法。换成日常话：在一个典范里做事，你总是在把各种情况往它的框架里排。

大部分情况排得进去，顺顺当当；少数情况勉强排进去，要加一点补丁；偶尔有一两个，怎么排都不对劲。这几个怎么排都不对劲的，就是裂缝露头的地方。

裂缝不会贴着标签出现。一个新手面对它，往往察觉不到：他会以为是自己哪一步做错了，回去检查，再试一遍，换种办法，最终要么凑合过去，要么放弃。

一个老手面对同一个情形，会有一种「这不是我的错」的直觉，因为他见过上千个能顺利排进去的情形，对「正常的排进去」是什么样，有一个深而细的感觉。当眼前这个情形不是那个样子时，他能察觉出差别，并且知道这差别不是自己手艺的问题。

这就是敏感的来源：它建立在大量正常案例的基础上。没有对「正常」的熟悉，就没有对「反常」的察觉。一个从来没有见过上千个标准算例的人，看到第一千零一个算例有点怪，是不会觉得怪的。他连「怪」的参照系都没有。

这样我们可以理解，为什么嗅觉不能靠听别人讲来获得。别人可以告诉你：「这个地方有问题。」那是一句话，不是嗅觉。嗅觉是你自己在无数次顺利之后，有一次不顺利，而你立刻知道这次不同。这种「立刻知道」，是身体层面的熟悉，是长时间待在一个地方之后才有的。

还有一个相关的原因，让「待够」这件事无法被绕过：差异序列需要足够长，才显得出来。前面第 8 章讲过，裂缝是一串差异在旧显露下排不进去。一两个孤立的怪例子，常常可以当作误差、当作偶然，轻易打发。

只有当一个人在典范里待得足够久，亲眼见到类似的怪情形一再出现，才会开始觉得：这不是偶然，是有东西在那里。这个「一再出现」，需要时间，也需要在同一个典范里的反复经历。一个今天换一个领域、明天换一个领域的浏览者，是攒不出这样的序列的。

所以，在典范里待够，不是一个道德要求，也不是传统美德，而是裂缝可见的一个条件。就像一个人要看见星星的缓慢移动，得夜夜守着同一片天空。

这里要特别提醒，别把这个说法读歪。「待够」不是说一个人要先在典范里当三十年的顺民，才有资格发现问题。有些人早早就察觉到裂缝，这是可能的，尤其在典范本身已经有裂缝露头的时候。

也不是说待得越久越好，有些人在一个典范里待了一辈子，却因为从不抬头，什么也没察觉。待够，是必要条件，不是充分条件。必要在哪里？在于没有这一步，后面的嗅觉无处生根。

### ◆ 课堂做法一：先让学生飞速推进，再让他们找 AI 推不动的地方

下面进入具体的做法。以下几种，不是什么独门秘籍，每个老师可以按自己的学科改造。它们的共同点是，把「亲手做」从任务降为训练，并且一步一步引向对裂缝的察觉。

第一种做法，可以叫「先飞后停」。

第一步，放开手，让学生用 AI 飞速推进。给他们一个典范里的真实问题，比如一个物理模型，一类化学反应，一个经济学里的标准分析框架，让他们尽可能快地、借助 AI 把它做到底：公式列出，结果算出，图画出，结论写好。

这一步要鼓励，要让学生体会到 AI 的力量：一个原来需要几天的工作，现在一个下午就能完成。这个体会很重要。它让学生一开始就站在 AI 时代的现实里，而不是躲在对旧时代的怀念里。

第二步，叫停，换一个任务：「现在，找出你觉得 AI 在这里走得不踏实的地方。」这个要求要说得具体。不是让学生评价 AI 好不好，而是让他在整条推进的路上，找出任何一个他觉得「这里好像是被糊过去的」位置。

这个任务的有趣之处在于，学生往往一开始什么也找不出来。AI 的输出流畅、自洽，每一步看起来都说得通。这恰恰是关键的教学时刻：老师不用批评，让学生自己尝一下「找不出来」的滋味。他会发现，要判断「走得不踏实」，需要一个标准，而这个标准，只有他对这个典范熟悉到一定程度才有。

第三步，让学生亲手做一遍，只做被他怀疑的那一小段，用最原始的方法。比如那一步推导，不用 AI，一个字一个字地写出来。很多时候，学生在手

写的过程中，会发现原本流畅的那一步其实有一个跳跃，AI 一句话带过，而他写不下去，因为中间缺了东西。

第四步，让学生把这个发现用一两句话记下来：「在某某问题里，AI 在某一步跳过了什么，我是通过什么发现的。」

这个练法里，亲手做的比重很小，只用在最有价值的地方，是一种精准的训练。它的好处是，让学生既不对 AI 抱敌意，也不盲目崇拜，同时给了亲手做一个清晰的理由：不是「为了应付作业」，而是「为了看出这里有没有东西」。

需要说明的是，学生找出的这些「跳过的地方」，大多数不是什么典范级的发现，多半是 AI 的一个小瑕疵，或者是教材里本来就省略的一步。别担心这个。这个练习的目的，是培养一种「在流畅里察觉缝隙」的习惯。就像练视力，先从大字母开始。

### ◆ 课堂做法二：在同一个典范里找排不进去的差异

第二种做法，针对的是更靠近「裂缝」的东西：差异。

典范的裂缝，是差异在旧显露下排不进去。那么，教学里就可以直接练这一点：给学生一批案例，让他们在这批案例里，把能按典范标准办法处理的归一堆，把处理起来别扭的挑出来。

举一个不依赖具体学科的练法。老师准备二十个案例，都属于同一个典范下的同一类问题。其中大部分是标准的，几个是稍微偏离的，一两个是真正让标准办法吃力的。

老师不告诉学生哪几个是特殊的，只给一个要求：「请用这个典范的标准做法，逐个处理。处理不下去、或者处理完觉得别扭的，单独列出来，并说说别扭在哪里。」

这个练习有几个设计要点。

其一，要让学生真的处理，不能只是看。只看，是不会有手感的。让他们用标准做法去推，去算，去套，才会在某些案例上真正碰壁。碰壁的经验，比任何提示都有用。这一步可以借助 AI 加速，但碰壁之后的「别扭在哪里」，要学生自己说。

其二，别扭的描述，要求具体。「这个案例好像不太对」不算，要说出：我在第几步用了什么办法，得到了什么，这个结果和我的预期差在哪里。越具体，越能区分「我用错了」和「办法不够用」。这个区分是整个练习的核心。

其三，老师要有备而来。他得提前知道这二十个案例里，哪些是真的难处理，为什么难。

这并不意味着他有答案，他要有的是诚实的判断：这个案例难，是因为典范的一个边缘问题，还是因为材料本身有噪声，还是因为这里确实是领域里公认的疑难点。如果他自己也不确定，那就如实地告诉学生，这个案例是他也没把握的。

其四，讨论时，要把「别扭」分类，但不要急着给结论。有的别扭，是学生的手艺问题，纠正即可；有的别扭，是典范里本来就有的已知疑难，介绍给学生，让他们知道前人怎么处理的；有的别扭，说不清是哪一种，就留下来。

留下来的那几个，就是班级的「悬案」。一学期下来，悬案本身可以成为一份有价值的记录。

要诚实地说，这个练习难度不小，对老师是考验。它依赖老师对一个领域有真正的熟悉。但这正是前一章说的：老师的位置，就在这份熟悉里。

练到一定程度，学生会开始自己造这样的案例：他们不再等着老师给，而是主动问，「如果我把条件改成这样，标准办法还成立吗？」这个问题，是敏感长出来的最清楚的标志。

### ◆ 课堂做法三：练习写一句话的猜想

第三种做法，把前两种的收获往上推一步：让学生练习写出一句话的猜想。

先说「猜想」在这本书里是什么意思。前面第 10 章会详细讲，典范猜想是新显露的最初一笔：一句话级别的新划界，或者一个新的基本对象。它不是一个完整的理论，甚至不是一个论证，它是一句话，说的是：「如果我们换一种方式划分这件事，换一种东西当作基本的，会怎么样？」

这样的一句话，在以前很难有人认真对待，因为一句话离一个可以验证的理论太远了，要把它展开，需要大量的人力和时间，多半在展开之前就夭折了。现在，AI 能把展开的成本压得很低，一句话猜想存活的门槛大大降低。这意味着，会不会写一句话猜想，成了一项值得训练的本事，而且可以在课堂上练。

练法大致是这样。

第一步，从前面两种做法留下的「别扭」里挑一个。比如班上留下的某个悬案，某个 AI 跳过的地方，某个标准办法处理起来总觉得勉强的情形。

第二步，请学生用一句话，说出「如果换一种看法，这个别扭会不会消失」。注意，限制在一句话里。限制很重要，一句话装不下论证，只装得下一个划界或一个对象。

比如：「如果不把这种现象看成一个量的大小变化，而看成两种状态之间的切换，会怎样？」「如果不把这个当作一个单独的东西，而看成别的东西相互作用留下的痕迹，会怎样？」这些只是示意，不是具体学科的猜想，各学科的老师可以按自己的领域造例子。

第三步，请学生标明：这一句话，改动了旧典范里的哪一个基本设定？没有改动任何基本设定的，不算猜想，只是旧框架里的一个改良。这一步是练习分辨：我是在补丁，还是在换划界？

很多学生起初写出来的「猜想」，其实是补丁，换了一种说法而已。老师在这里不用批评，只需问：「旧办法里的这个东西，在你这句话里还在不在？如果还在，原来的位置变了没有？」

第四步，请学生把这句话交给 AI，让它按这个猜想展开，看看大致能推出什么、和旧办法的结论哪里相同、哪里不同。这一步是一种实验：学生亲眼看到一个一句话的想法被展开成一个小系统，体验到「猜想」和「展开」是两个不同的活动，前者几乎不要成本，后者也不再是贵得吓人。

第五步，也是最关键的一步：看展开的结果，是否只是旧办法换了皮。如果展开出来的东西，在每一个具体问题上都和旧办法给出完全一样的东西，只是名字不同，那这个猜想没有真正换层。老师带着学生一起检查：哪里不一样？

不一样的地方，是不是旧典范原本排不进去的那个别扭？如果恰好是，那就有意思了。这个检查的方法，会在第 14 章（验收）里讲得更仔细。这里的课堂练习，只要求学生体验一次这种检查。

这个练习有几个明确的边界，要讲清楚。

第一，它的目标不是产出好猜想。绝大多数课堂产出的猜想，都不是什么典范级的东西。练习的目标，是让学生熟悉这个动作：写一句话，标明它改了什么，让它被展开，检查它是不是换皮。动作熟了，将来遇到真正的裂缝时，他们才不会手足无措。就像练习写字，不是为了每个字都成为书法。

第二，不要让学生觉得写猜想是一种「天才的闪现」。这是旧本体论的叙事：一个人看穿了什么，创造了什么。本书的看法是，猜想是从差异序列和特征纠缠里显影出来的，是那些反复排不进去的东西逼出来的。

所以练习里，猜想必须从前两种做法里留下的具体别扭出发，而不是凭空想一个「有创意」的句子。凭空写的猜想，几乎一定是空的。有别扭做底，句子才有重量。

第三，要宽容。一句话猜想的多数会是错的，这是常态。如果老师的反应让学生觉得写错一次就要丢脸，这个练习就会死掉。班级里应该形成一种风气：一句话猜想是便宜的，写错是便宜的，重要的是写出来，并且知道它为什么错。

做完这一串，还可以接一个小练习，专门练复合体里要用的本事。

让学生拿着自己的那一句话猜想，与 AI 来回三轮：第一轮，请 AI 按这句话展开；第二轮，学生读展开的结果，找出哪里露出了原先没看见的缝或推

论，据此改一改那句话的划界，再请 AI 重新展开；第三轮，请 AI 同时给出两三个版本并排，学生选一个，写下为什么选它。

三轮结束，学生交两样东西：来回的全程，和一张小表，列出开头那句话里哪几处「不一样」在来回中保住了，哪几处被 AI 拉回了旧说法，各举一句原话为证。

这张表不打分，是给学生自己和老师看的；看得多了，学生会认出 AI 在哪一类句子上最爱把新说法翻回旧说法。这里练的，不是一个人独立产出典范，也不是一个人把 AI 管得服服帖帖，而是第 20 章所说的那种能力：在来回里攥住「不一样」。

## ◆ 怎么知道「待够」了：不是时间表，是几个迹象

有人会问：到底要待多久才算够？一年？三年？十年？

这个问题没有统一的答案，不同领域、不同的人，进度差别很大。与其给出一个时间表，不如说几个迹象，让老师和学生自己对照。

第一个迹象：能不能预测典范里的下一步。一个在典范里待够的人，看到一个问题的开头，大致知道标准做法会往哪里走，会遇到什么，结果会是什么样子。预测准，说明他对典范的纹理有了感觉。预测经常落空，说明还没待够。

第二个迹象：遇到别扭的时候，能不能说清楚别扭在哪里。待得不够的人，遇到别扭，只会说「不对」、「有问题」。待够的人，会说：「这里我预期是这样，实际是那样，差的这一块来自这个设定。」能把别扭说成这样，意味着他已经有了一张足够细的图，别扭才能在图上被定位。

第三个迹象：能不能区分三种情况：自己错了，典范里已知的疑难，以及可能是新的。这是前面谈教师难处时说的三种「不知道」的学生版。能分得开，是嗅觉成熟的重要标志。

第四个迹象，更微妙一点：对典范本身，有了一点不那么恭敬的熟悉。刚进一个领域的人，常常对它心怀敬畏，也有人心怀不耐。待够的人，对它有一种像对老朋友一样的态度，知道它的好，也知道它的毛病，既不神化也不贬低。这种态度，是能够看见裂缝而不被裂缝吓倒的心理条件。

这些迹象不是考试科目，不要变成打分的清单。把它们变成清单，学生又会学着「表演」这些迹象，评价又掉回元素层。它们更适合作为老师观察学生的参照，也适合学生自己对照，问问自己：我现在到哪一步了？

## ◆ 放回日子里：一个学生的两种做法

最后用一个场景把这一章收拢。想象两个同班的研究生，各自要在一个成熟的领域里做第一个课题。

第一个学生，很快上手了 AI。导师给了方向，他让 AI 做文献整理，让 AI 搭出分析框架，让 AI 写出初步结果。几周之内，他手里已经有了一份看上去很完整的报告。导师读了，说不错，让他继续深入。他再让 AI 深入，又是几周，报告更厚了。他的进度比所有前辈都快，周围人都很羡慕。

可是有一天，他的结果和另一个组的结果对不上。他去问 AI，AI 给出了三种可能的解释，每种都说得通。他不知道选哪一个，不知道哪一个更可能，也不知道这是不是一个重要的差异，还是一个无关紧要的噪声。

他发现自己没有任何依据来判断。他从头到尾都没有在这个领域里「摸过」，他手里的一切都是 AI 递给他的，他对这些东西的熟悉度，就像一个只坐过车的人对路面的熟悉度。

第二个学生，也用 AI，用得同样熟练。区别在于，导师要求他在课题的前几个月里，对每一类典型问题，都至少亲手做一次完整的流程，不借助 AI，慢，但每一步都自己过。同时，他也被要求在用 AI 飞速推进之后，回头找 AI 推得不踏实的地方。他进度没有第一个学生快，周围人觉得他有点笨拙。

当他的结果同样遇到那个对不上的情形时，他的反应不同。他先看了一眼，说：「这个对不上，不像是我算错了，也不像是常见的那种误差。」他知道，是因为他亲手做过那一类问题许多遍，对「正常的误差」长什么样有了感觉，眼前这个不是。

接着，他去查，发现这不是一个已知的疑难。他把这个别扭记了下来，一直放在心里。后来，他又在另一个情形里遇到类似的东西。两次放在一起，他开始觉得：这里也许有东西。

这个故事没有结局，也不需要结局。它要说明的只是：两个学生用的是同一个 AI，差别在于一个在典范里待过，一个没有。

AI 对两个人都一样快，但「对裂缝的敏感」只在其中一个人身上长出来了。而这份敏感，正是后面各章所有工序的入口：没有它，识别裂缝就无从谈起；没有裂缝，典范猜想就是空中楼阁。

## ◆ 地基的几层：不同的学科，不同的「摸」

「摸多了」这三个字，在不同的领域里，长得不一样。把它说得具体一点，有助于老师把这一章的做法搬进自己的课堂。

在数学和理论物理这样靠推导的学科里，「摸」主要是亲手推。一个定理的证明，看一百遍不如自己写一遍；一个近似的合理性，听十次不如自己找一个不合理的例子试一试。这里的手感，是对「哪一步是实的、哪一步是靠一句『明摆着』带过的」的感觉。

在实验学科里，「摸」是亲手做实验，包括那些失败的、莫名其妙的、需要重来的实验。一个从来没有亲手让仪器出过怪数据的人，是不会有「这个读数不对劲」的感觉的。课堂上，不妨留一两个实验，故意不让学生顺利：比如数据里埋着一个设备的小毛病，让学生自己去发现。发现的过程比结论重要。

在社会科学和人文学科里，「摸」是亲自去读原始材料、去做访谈、去面对一个具体的案例，而不只是读别人整理好的概括。概括是已经排进了典范的东西，原始材料里还保留着排不进去的毛边。学生要有机会直接面对毛边，才有可能察觉它们。

在工程和应用学科里，「摸」是亲手造一个东西，让它在真实条件下运转，看它在哪里出毛病。模拟当然有用，AI 当然可以帮着做模拟，但真实的东西有一种模拟难以全部复现的「不配合」，学生从这种不配合里学到的，是对典范边缘最直接的感觉。

这四种「摸」，共同点是：都要求学生面对一个没有被整理过的东西，不是一份已经总结好的材料。整理，是把差异排进典范的动作；没有整理的东西里，差异还保留着。让学生接触没有整理的东西，是训练嗅觉最可靠的办法。

这里顺便回答一个常被问到的问题：自学的人怎么办，没有老师带的时候，能不能练出这份嗅觉？可以，但要刻意。自学者最大的风险，是只在 AI 整理好的材料里打转，一切都顺顺当当，永远遇不上毛边。

一个办法是：定期给自己找一个没有现成答案的小问题，用手头的数据或材料，亲手从头走一遍，不借助 AI 的整理，然后再用 AI 对照。对照时，重点看自己走出来的路，和 AI 给的路，在哪里分叉。分叉的地方，往往正是值得琢磨的地方。

## ◆ 训练里的 AI：怎么请它进来，怎么请它退后

既然亲手做是训练，那 AI 在训练里到底放在什么位置？一刀切地禁止，和一刀切地放开，都不对。可以这样安排几条规矩，老师按需要调整。

第一条，训练的时候，先自己，后 AI。让学生先凭自己的力量走一遍，哪怕走得很笨，再请 AI 来对照。顺序不能倒。先看了 AI 的答案再去「自己做」，做出来的是对答案的模仿，不是训练。

这跟下棋复盘一个道理：先自己想这一步该怎么走，再看高手怎么走，差距才看得清；先看了高手的走法，再假装自己想过，什么也学不到。

第二条，请 AI 做陪练，不做替身。陪练的意思是：学生做一步，AI 指出这一步和标准做法的差别；学生卡住了，AI 给一点提示，而不是给整条路。

这需要学生学会向 AI 提这样的要求，也需要老师在课上示范这样的用法。很多学生不知道可以这样用，他们只知道「把题给它，它给我答案」。把用法教会，本身就是新时代教学的一部分。

第三条，请 AI 做对手。让它故意在推导里埋一个小错，让学生去找；或者让它给出两种互相矛盾的做法，让学生判断哪一种在这个情形下站得住。这种用法，把 AI 的流畅变成了练习的材料，学生在找错的过程中，对典范的纹理会熟悉得很快。

第四条，训练之外，该放手就放手。训练结束，进入真实的任务，AI 就应该全力使用，没有必要在完成任务时还背着训练的包袱。学生要学会切换：这是训练，我慢下来；这是任务，我借力。会切换，是新时代的一种基本素养，不比会做题更不重要。

这四条里，没有一条是靠监视来实现的。训练能不能起作用，最终取决于学生是否明白它的用处。所以老师要花时间向学生讲明白：这一段为什么要你亲手做，它在你身上要留下什么。讲明白了，多数学生是愿意配合的；讲不明白，再严的监督也只能得到表面的服从。

## ◆ 几个容易走偏的地方

在结束这一章之前，把几个常见的走偏列出来。

其一，把「亲手做」变成怀旧。有人会借着「地基」之名，要求学生回到一切都靠手工的时代，甚至禁止 AI。这不是本书的主张。AI 是车，车要用。亲手做的价值，在于它是训练，有目的、有节制，不是情结。训练的量要够，但不是越多越好。

其二，把「待够」变成拖延。有的教学安排，让学生在元素层无休止地打转，说是「打基础」，一年又一年，从不让他们看边。这同样是走偏。待够和看边应当交替进行，不是先后两个阶段：在典范里待了一段，就应该回头看一看，再待一段，再看。就像登山，不会等到登顶才第一次回头。

其三，把「敏感」变成怀疑一切。有些学生，学了「要对典范保持警惕」，转头就对每个标准做法都挑毛病，这不是敏感，是另一种偷懒，不需要熟悉就能做到。真正的敏感，是对大部分东西很安心，对极少数东西很不安，而且分得清。要区分这两种，看他的怀疑有没有细节。

其四，把「一句话猜想」变成口号竞赛。写出听起来很惊人的句子，却从没经过前面的别扭，也不愿意让 AI 展开、让人检查。这种练习只会培养浮夸。老师要抓住的是：每一句猜想，都必须能追溯到一个具体的、学生自己遇到的排不进去。

地基这个词，在这一章里已经用了很多次。最后说明一下它的意思：地基不是一个人在楼下待着，等有一天被允许上楼。地基是楼的一部分，楼层越高，地基越要实。腿不是被车淘汰的，它是被安置到了一个新的位置，成为认路、登山、在车到不了的地方行走的能力。

亲手待在典范里的那些年月，会在多年以后，以一种「摸多了」的方式，悄悄地在某个时刻起作用。第 20 章会讲这份地基在人与 AI 的复合体里担什么：它给来回一个站得住的参照，没有它，来回只是在原地打转。

下一章要回答的问题是：地基有了，嗅觉有了，接下来典范发生这件事，课堂上怎么教、怎么评？评什么，才不会又掉回元素层？

带走的话

**亲手做不再是完成任务的手段，而是训练；训练的价值，在于留下手感，不在于拿到结果。 在典范里待够，才有对裂缝的嗅觉：没有对「正常」的熟悉，就没有对「反常」的察觉。 课堂上可以这样练：先让学生借 AI 飞速推进，再让他们找推不动的地方；在同一个典范里挑出排不进去的差异；练习写一句话的猜想，标明它改了什么，再让 AI 展开、检查是否换皮。 待够没有时间表，只有几个迹象：能预测下一步，能把别扭说具体，能分清自己错了、已知疑难与可能是新的，对典范有老朋友般的熟悉。 训练要有目的、有节制，既不怀旧，也不拖延；地基是楼的一部分，不是楼的替代品。 放进人与 AI 的来回里，这些练习指向第 20 章讲的几样能力：攥住「不一样」，读出缝并改划界，在并排版本中选择并说出理由，以及有地基。**

第 18 章

# 典范发生如何教、如何评；论文制度为何先被压平

先说一个很家常的场景。

设想某所学校的评审会上，摆着两份材料。第一份工整、规范，引了四十篇文献，把一个公认的难题往前推进了一小步，数据漂亮，推导严密，评委翻到第三页就点头了。

第二份只有薄薄几页，开头一句话说：我们一直把这件事当成「某某问题」，也许它根本不是那一类问题。后面的推导粗糙，有两处连作者自己都承认接不上，引文寥寥。评委们翻完，互相看了一眼，没人吭声。

最后第一份得了高分，第二份被客客气气地退回，评语是「论证不充分，建议补充工作」。

这个场景里没有坏人。评委没有偷懒，也没有偏见，他们按的是一套再正常不过的标准在读。问题恰恰出在这里：那套标准读第一份时，每一项都有对应的格子可以打勾；读第二份时，格子全对不上，因为第二份想改的，正是格子本身。

先把话说在前头：制度怎么改，本书没有现成的方案，更不会编一套政策、一个机构、一组数字出来。下面写的是方向和思路，供读者自己掂量。

## ◆ 一、为什么典范发生很难评

先问一个朴素的问题：评价一件东西，需要什么？

需要一把尺子，需要用这把尺子去量，量完得到一个读数。尺子要事先有，大家要认。这是评价最基本的结构。

典范之内的工作，尺子是现成的。这个典范里什么叫「好问题」，什么叫「严格的证明」，什么样的实验算「可重复」，什么样的结果算「推进」，大家心里有数，用不着重新商量。所以评审可以又快又稳：一份工作送来，几个懂行的人看一遍，大体能达成一致。

典范发生不一样。它做的事，恰恰是让一把新尺子出现。新尺子在出现的那一刻，是没有人认的，甚至连「它是不是一把尺子」都还不清楚。用旧尺子去量它，量出来的读数基本是低分，因为它在旧尺子上的位置是「不规范」「偏题」「证据不足」。

这里要小心，不能说成「评委眼光不行」。把这件事归咎于评委，又滑回了旧想法：好像有个人能凭眼光一眼看穿。事实比这平淡得多：评委站在哪个显露里，就只能用那个显露里排得进去的东西去读。第二份材料想说的话，在旧显露里没有位置，不是评委没读懂，是「读懂」这个动作本身依赖于那个位置。

回想前面几章讲过的：旧典范的裂缝，是差异序列在旧显露下排不进去；反常攒成压力，旧显露才失稳。新的显露一开始只有最初一笔，粗糙，前后有断裂，很多地方接不上。

它要等后来的差异序列一点点排进来，等纠缠一点点收拢，才站得稳。所以在最初的时刻，哪怕有人把它写出来，它也「证据不充分」，这不是写得不好，是它此刻就处在这个阶段。

于是评价典范发生，至少有三重难处。

第一重，时间。一个新显露站不站得住，要看后来很久的事：后来的观测是否排得进去，后来有没有原来排不进去的差异在它这里顺了。这不是一年两年的事，往往是一代人、几代人。而评奖、评职称、评课题，每一件都有个截止日期。

第二重，尺子。刚才说了，当时能用的尺子是旧的。若强行用旧尺子打分，打出来的恰好是「这个东西多么不像旧典范里的好工作」。

第三重，真假混杂。这一条最让人为难。自称在做新典范的东西，绝大多数是不成的：有的是旧典范换了个说法，有的是几句漂亮话外加没有后文，有的纯粹是想错了。真正后来站住的，是极少数。

所以评审者不能一律放行，放行的结果是一地鸡毛；也不能一律拦下，拦下的结果是把那极少数一并拦了。这个两难没有取巧的解法，只有「怎么在不确定里做出还说得过去的判断」这一层可以谈。

把三重难处并在一起看，有一个后果：典范发生的成果，天然没法用「当期、同行、旧尺」的方式验收。这不是制度设计者的疏漏，是这件事本身的性质。

## ◆ 二、现行制度奖励的，是哪一层的东西

现在换一个方向，看我们手里的制度。

论文、引用、课题、奖励、职称，这一整套东西，每一环都在奖励什么？

先看论文。一篇合格的论文要有什么：问题的提出（接在已有文献的缝上），方法（符合领域的规范），结果（可以被同行复核），讨论（放回已有

的讲法里）。哪一项是在让典范发生？一项都不是。它们全是在已有典范的规矩下，把一小块空地填上。

再看引用。引用是什么？是一篇工作被后来的工作「认出来」、接着往下用。引用多，说明它被很多人接得上。接得上谁？接得上同一个典范里干活的人。所以引用数量，量的是一篇工作在现有典范网络里的连接度。

一个刚刚提出新显露的工作，在它所属的旧网络里，连接度反而低，因为旧网络里的人接不上它。反过来，一个典范内部的中等好工作，只要处在一个活跃的领域，引用可以很高。

再看课题。课题申请书要写什么：目标、技术路线、预期成果、可行性、进度表。每一项都在要求「事先讲清楚会得到什么」。典范发生，偏偏是无法事先讲清楚的：如果讲得清楚，说明新显露已经有了，那就不是发生，是展开。

一份诚实的「典范发生课题申请书」，大概只能写：「我们想试一试把这个对象换一种划法，不知道行不行。」这样的申请书，今天几乎没有哪一套评审办法能让它通过。

再看奖励和职称。奖励倾向于奖已经被验证的东西，这本是合理的，奖励不能奖空话。职称倾向于看数量和层次，数量要靠产出的节奏，层次要靠发表的位置。人在这套节奏里，被推着做「能按期出来、能被认出来」的东西。

把这一整套看过去，会发现它们有一个共同的底盘：**可预期、可复核、可比较、有节奏**。这四样，在典范之内是美德，是整个科研体系能大规模运转的前提。前面讲过，常规科学养活了几乎全部科研人口，靠的就是这四样。

问题来了：这四样，恰好也是 AI 最擅长接走的四样。

可预期，意味着目标明确，路径可规划，AI 擅长；可复核，意味着规则清楚，检查标准清楚，AI 擅长；可比较，意味着有共同的指标，AI 擅长；有节奏，意味着稳定地批量产出，AI 更不用说。

第一编讲过，答案、过程、提问、验收是同一层的四个位置，全在这一层里。现在可以补上一句：论文、引用、课题、奖励这套制度，奖励的正是这一层的产出，而且是以这一层最擅长的方式来奖励。

所以，制度最先被压平，并不神秘。一旦这一层的生产成本被压到很低，论文可以大量地、规范地、合格地生产出来；引用可以在已有网络里自动地生成连接；课题申请书的「技术路线」可以写得头头是道。

到那时，这套制度手里的尺子还是那把尺子，可它量出来的东西，已经不再区分什么了。每个人都合格，每份申请书都像样，每篇论文都规范，读数挤在同一个区间里。一把量不出差别的尺子，不是坏了，是量程到顶了。

话不必说得太满，下一步制度会怎么变，本书也不知道。有可能出现的情形有好几种：评审变得更严，要求更多验证；评审变得更看重过程材料；计数的东西换一批。

但这些，按本书的框架看，仍然是在这一层里平移，前面第 3 章已经说过缘故，这里不重复。这里只想指出：**制度的被压平，不是因为它做错了什么，而是因为它做对的那一层，被接走了。**

再放回日子里看一看。设想一位年轻的研究者，博士第三年，他的导师是个明白人，常对他说：「你现在要把论文发出去，毕业才有保障；至于你心里想的那件大事，先放一放，毕业以后再说。」这话没有错，也是为他好。

可是这位年轻人心里想的那件事，是不是就真的有「毕业以后」？毕业以后是职位考核，是课题申请，是下一轮的评审。他会发现，那个「以后」一直在往后退。制度没有恶意，只是没有给它留格子。

而到了常规科学被大量接走的年代，这个年轻人的处境会更尖锐：他按规矩能做的那部分，机器已经做得又快又好；他心里想的那件事，制度仍然没有给格子。夹在中间，他两头落空。这个落差，正是制度要面对的。

## ◆ 三、制度上会碰到的几处硬冲突

下面把冲突拆开来，一件一件说清。

**第一处，格式的冲突。** 现行的论文格式，是为「在已有框架里推进」而设计的。「相关工作」「研究缺口」「方法」「实验」「结论」，每一栏都预设了一个已经存在的框架。而一个典范猜想，开头往往是对框架的重新划界，没有「研究缺口」可填，因为它要说的是「你们把缺口看错了」；也没有「相关工作」可综述，因为它要说的是「相关」本身要重划。把这样的工作塞进现有格式，要么塞不进，要么被迫改写成「换了个说法的旧工作」，改写的过程，就是被翻译回旧典范的过程。这一点，后面第 14 章讲的「换皮」，在制度的层面上也会发生：不是 AI 在翻译，是格式在翻译。

**第二处，评审者的冲突。** 评审靠同行。同行是谁？是在同一个典范里干活的人。让他们评一个想改变典范的工作，等于让当前的显露来评价自己的失稳。这并不是说同行不诚实，而是结构上他们站的位置如此。这里需要的，是评审者的构成和评审的方式要有所不同，后面细说。

**第三处，周期的冲突。** 年度考核、三年一届的课题、几年一次的职称评定，周期都以「几年」计。典范发生的真正验收，以「几十年」计，有时更

长。这两个周期之间，隔着一道大沟。一个把力气放在典范发生上的人，在这道沟的这一边，可能一直「没有产出」，等沟那一边的验收到来，他已经不在这个岗位上，甚至已经不在了。

**第四处，风险分配的冲突。** 这一条容易被忽略。典范发生大概率不成。一百个想改划界的尝试里，能站住的少之又少。如果制度要求每个人都要「有结果」，那么没有人敢拿自己的职业生涯去押一个大概率不成的事。个人不押，整体就无人去做。这和投资里的「分散风险」是一个道理：必须有人去做大概率失败的事，而且要有机制让他们失败了也不至于一无所有。

**第五处，信号污染的冲突。** 一旦制度开出一个「奖励典范创新」的口子，马上会有人包装：把典范内的小改动写成「范式革命」，把旧工作加个新名词。这种事在科学史里一直都有，在 AI 可以大量生产漂亮叙述的年代，会更多。所以口子不能随便开，开了之后要有对付包装的办法。这一条与前四条是互相拉扯的：口子开得窄，漏掉真的；开得宽，进来假的。

这五处冲突，没有一处可以用「加个规定」解决。它们是结构性的，要逐步调整，边调整边观察。下面的几个方向，都是在这五处冲突里找缝。

## ◆ 四、几个可走的方向

先声明三点。第一，下面都是方向，不是方案。第二，本书不替任何机构拿主意，也不说哪个国家哪所大学已经这么做了：不知道，也不想编。第三，每个方向都有自己的毛病，都会写出来。

**方向一：拉长追踪。** 既然典范发生的验收要几十年，那么最直接的办法，是不要在当期验收，而是在当期「登记」，在日后「回看」。登记什么？登记这个工作提出的划界是什么，它预言了什么、它认为旧显露哪里排不进去、它

打算拿什么来检验。登记的东西要短，要清楚，要能让几年后、十几年后的人看明白「当初说的是什么」。

这个办法的好处是：它把「悬着」变成了「有档可查」。过几年回头看，哪些划界后来被新的观测支持，哪些被冷落，哪些被放进更大的格局里，有记录可比。

它的毛病也明显：谁来保管这些记录，谁来负责回看，几十年里制度自身也在变。而且回看的人也站在他们那个时代的显露里。所以拉长追踪不能指望给出「最终判决」，它能给的，是一份比较靠谱的「当时说了什么」的底账。

还有一个小处要留意：追踪的是**登记在案的划界**，不是追踪人。盯着人，容易回到「天才叙事」，奖励的是个人光环；盯着划界，才是盯着那个新显露的命运。

**方向二：并排评审。** 前面说，评审者的尺子是旧的。一个不同的办法，是不让一份新工作单独面对旧尺子，而是让它与旧典范里的处理方式并排摆在一起，针对同一批排不进去的差异，比较两种显露各自怎么排。

比如，给定一批旧显露下的反常，一边是旧典范的补丁式处理，另一边是新猜想的处理。评审者看什么？不是「新的这份是否符合旧规范」，而是：这批反常，在哪一边排得更顺？哪一边需要补的补丁更多？哪一边在原来排得好的那些地方损失了什么？哪一边留下了新的、旧的看不见的差异？

这样做，是把「评新工作」改成「评两种显露对同一批差异的消化」。尺子不再只是旧的那一把，而是两套并排。后面第 15 章讲的「多个典范并排发生再比较」，在评审上就是这个意思。

这个办法的难处：要有一批大家都认账的「排不进去的差异」，也就是要有一份公认的反常清单。清单怎么来，谁有权列，列了之后会不会变成新的「领域规定」，都是问题。

而且并排比较需要 AI 的大量展开工作：把新猜想按照它自己的路数展开到可比较的程度，这个工作量，在没有 AI 的年代是没法要求任何申请人做的。这反倒说明，AI 接走常规科学，对这类评审是有帮助的，它让「并排」这件事第一次变得做得起。

**方向三：评「新基本对象」，不评「新结果」。** 典范级的东西，核心常常是一个新的基本对象，或者对旧对象的一次新划界。这一点，前面第 11 章讲判据时已有说明。放到评价上，就是：评审的主要问题，不该是「结果新不新、数据好不好」，而是「它提出了什么对象？这个对象能不能被说清楚？说清楚之后，旧显露里的哪些东西因它而排得顺了？它又带出了哪些旧显露里根本问不出来的问题？」

最后一问，值得多说一句。一个真正的新显露，会带来新的「可问的问题」。这批问题在旧显露里没有位置，在新显露里一下子冒了出来。

一份工作带出了多少这样的问题，这些问题后来有没有人认真去追，是比较可靠的信号。因为它不依赖当期的指标，又很难靠包装伪造：只是换了个名词的旧工作，带不出旧显露里问不出的问题。

**方向四：检验「是否换皮」。** 前面说过，第五个冲突是信号污染。对付它的办法之一，是把「换皮检验」做成评审的一道固定工序：把这份声称是新划界的工作，交给一个外在的步骤，看它的核心论断能不能被完全翻译回旧典范

的语言，而不损失任何东西。如果翻译回去毫无损失，那它就是换皮。如果翻译回去，丢了东西，丢了的那一部分，才是这份工作真正新的地方。

这道工序在第 14 章会详细展开，这里只点出它在制度里的位置：它是一道过滤，不是一个判决。过滤掉的是明显的包装，留下的仍然要靠时间去验。

**方向五：给失败留出体面。** 针对第四处冲突，风险分配。一个社会若希望有人去做大概率不成的事，就要为这些人留出体面的位置：他们的尝试被记录，失败的理由被认真对待，他们的才干在别的岗位上被认可。科学史上许多后来站住的东西，最初都有过一长串不成功的前驱，这些前驱的工作，为后来者清出了差异序列上的许多「此路不通」。把这类失败写进共同体的记忆，是一种公共品。

这几个方向，彼此不冲突，可以叠着用。它们共同的思路是：**不要求典范发生在旧尺度上证明自己，而是为它另备一套「登记、并排、回看、过滤、容错」的轻量结构，让它在被验收之前，先能存活。**

这里还要老实说一句：这套结构本身也会被使用者「玩」。凡是有奖励的地方，就有人研究怎么拿奖励。所以它们需要不断被检查，被调整，不能成为新的僵化。

如果有一天「典范发生」也变成了一种可以按流程批量生产、按指标考核的东西，那它就已经被压平了，要再往上找下一层。这一点，读者不必惊讶：层级是会这样一直上去的。

## ◆ 放回日子里：一场设想中的并排评审

光说方向，容易飘。这里设想一场评审，纯属虚构，没有哪个机构、哪个人，只是为了让读者看一看这些办法落到桌面上，大概是什么样子。

设想某个领域里，十几年来积下了一份反常清单，一共九条。九条里，有三条是各路人马都承认的「确实别扭」，有四条是少数人在谈、多数人觉得是测量误差，另外两条谁也说不准。这份清单不是哪个人定的，是几次学术会议上陆续记下来、又经多方补充的，有出处，有条目，可以查。

现在有两份材料送来评审。

第一份是典范内的推进。它拿现有的框架，加了一个修正项，使九条里的第一条、第二条不那么别扭了。修正项的写法很规范，计算很扎实。

第二份是一个新划界的猜想。它只有几页，开头一句话说：我们一直把这个对象当作「一团东西」来处理，如果把它看作「一串先后发生的事」，会怎样？它没有把九条逐一算出来，只算了其中两条，而且算得很粗。

按旧办法，第一份高分，第二份退回。

按并排的办法，评审会这样开。先把九条反常摆出来，问：两份材料各自「消化」了几条？第一份消化了两条，靠的是一个新加的修正项，修正项本身没有来由，只是「这样改就对得上」；第二份粗算了两条，其中一条与第一份重合，另一条是第一份没碰的。

再问：在旧框架本来排得好的那些地方，两份材料各自损失了什么？第一份几乎没有损失；第二份在一处旧框架里很顺的计算上，目前算不出来，作者自己也承认。再问：第二份带出了哪些旧框架里问不出的问题？

评委们商量了一阵，列出三个。其中一个，居然和清单里最后两条「谁也说不准」的，有些相通。

再做换皮检验：能不能把第二份的核心句子，原样翻回旧语言？评委们试了几遍，翻译的结果是：翻回去之后，「先后发生」这一层意思丢了，旧语言里没有对应的词，只能说成一个近似的「时间依赖项」，而这个近似项，恰好正是第一份里那个没有来由的修正项的样子。

有人小声说了一句：「这两个东西，会不会是同一件事的两种写法？」

到这里，评审会没有给出「谁对」的结论，也不可能给出。

评审会可以做出一个很实在的决定：第一份，按典范内的好工作，正常通过；第二份，把它登记在案，写明它提出的划界、它粗算的两条、它带出的三个问题、它尚未解决的一处损失，约定几年后回看；同时拨出一小笔资源，让作者借助现有的工具，把它按自己的路数展开得更完整，再来并排一次。

请注意这个决定里没有的东西：没有宣布发现了革命，没有把第二份捧起来，也没有把它按下去。它只是让一个暂时还说不清的东西，活过了这一轮。

这里面有几处要看得更仔细。

其一，反常清单是公共的，不是任何一份申请书自己列的。若由申请人自己列，他自然会挑对自己有利的。清单的公共性，是并排评审能成立的底。

其二，换皮检验在这里起了真正的作用：它让大家看到，第二份并非完全等于旧语言，也让大家看到，第一份的修正项与第二份的核心之间，可能有隐约的联系。这个联系，单看任何一份材料都看不出来，并排之后才冒出来。**并排的好处，恰恰是让差异在两种显露之间露出来。**

其三，所有这些环节，其中枯燥的工作量，包括把第二份猜想展开到足以比较、把九条反常按两种路数各算一遍、把翻译检验做得细致，在过去没有人能要求申请人去做，因为太费人力。

现在这部分，正好是 AI 能够接走的。它把评审会从「凭印象争论」，带到了「面对一张摆开的表格」。人要做的，是读这张表，决定要把有限的资源押在哪里。

当然，这只是一场设想。真实的评审会，有人情，有预算，有时间压力，有不同学派之间的积怨，远比这复杂。写这一场，只是想说明：这些办法并不玄，也不需要等一个「全新的制度」从天而降才开始试。在今天的评审会里，一个有心的召集人，就可以先试着多问几句。

## ◆ 五、典范发生怎么教

制度还没改的时候，学校的教室里已经要面对一个问题：这件事到底怎么教？

先承认一件事：典范发生，不能保证教会。教一个人使他能够造出典范，这句话本身已经把「人」放在主语的位置，仍带着旧想法的味道。

更准确的说法是：**教学能做的，是让一个人身上有更多条件，使这场发生更有机会在他这里显影。** 这是个比较谦逊的说法，却也是个能落地的说法。有些条件可以教，有些练习可以做。

可教可练的部分，沿着前面几章已经讲过的次序，排成五件事。

**第一件，把地基铺厚。** 前面第 17 章专门讲过地基。这里只提一句：典范发生所需的东西，大部分是地基。一个对旧典范没有走过的人，看不出裂缝在

哪里。所以让学生认真地、不走捷径地在典范内走一遍，亲手做习题，亲手推导，亲手踩坑，这件事不能省。AI 可以替你做完所有习题，但路是你没走过的路，地图在你脑子里画不出来。车替代了腿，可不走过路的人，看不懂地图。

**第二件，学会读「诞生史」。** 这是最该加进课程的一门东西。现在的课，讲的多是典范的成品：定律、定理、方法。可学生看到的，是一座盖好的楼，看不到脚手架。典范的诞生史，应当当作「差异序列的记录」来读：当时有哪些观测或事实排不进旧的讲法，它们如何一件一件堆积，起初大家如何用补丁处理，补丁为什么越来越难看，新的划界最早出现时是什么样子，它又是如何被一点点证实、扩展、修正的。

读诞生史，目的不是敬仰先贤，而是练一种看法：看一个领域，不是只看它「现在是什么」，还看它「哪里最别扭」。

每个成熟的领域，都有一些让老师私下里也含糊的地方，一些「约定俗成」的说法，一些写在教科书小字里的「例外」。让学生学会把这些地方找出来、攒起来、问一问为什么，这就是识裂缝的基本功。

**第三件，练「反常清单」。** 具体做法很朴素：每个学生，在学习一个领域的过程中，随手记下「这里我觉得别扭」的地方。别扭可能来自一个公式没有直观依据，可能来自两种讲法并存却无法统一，可能来自一个反复出现的「特殊情形」。学期末把清单拿出来，大家一起看：哪些别扭是我没学透，哪些别扭老师也说不清，哪些别扭放在更大的背景里居然是同一个别扭。

要强调：这份清单的价值，不在于里面有多少「发现」，而在于训练一种敏感。大部分别扭，最终会被证明是自己没学透，这是正常的，也是有价值的。因为学会分辨「我没懂」和「它真的有缝」，本身就是一项重要的能力。

**第四件，练「一句话猜想」。** 前面第 10 章讲过，典范猜想是新显露的最初一笔，常常可以用一句话说出来：这个东西，也许不是我们一直以为的那一类。可以给学生这样的练习：针对某个让人别扭的地方，试着用一句话写出「如果换一种划法」。不要求写对，也不要求论证。写完，交给同学，让同学试着用这句话去重述几个已知的现象，看哪里顺，哪里别扭。

这个练习的重点是「写得出来」，又「站不住也不要紧」。学生要经历很多次「写出来，被自己或别人否掉」，才能慢慢有手感：什么样的一句话是空的，什么样的一句话是换皮的，什么样的一句话虽然粗糙，却能把几件原本不相干的别扭拢在一起。后一种，就是值得继续的。

有了 AI，这个练习可以大大加速。学生写下一句话，让 AI 按这句话的路数把它展开，推出一批后果，写成一个粗糙但完整的小体系，再让学生拿着这个体系去对照那份反常清单。

过去一个猜想要用几年才能展开到可检验的程度，现在可以用几天。学生可以在一个学期里，经历过去一个学者一辈子也经历不了几次的「猜想—展开—检验—放弃」的循环。这是 AI 带给这门课最实在的礼物：试错的成本降下来了。

**第五件，练验收，练选择。** 展开出来的东西，很可能是旧典范的换皮。让学生练习一件事：拿到一份展开的体系，去问，它的核心断言，能不能不损失地翻译回旧语言？它带出了哪些旧语言里问不出的问题？它在旧典范排得好的地方，丢了什么？这种验收，要在课堂上反复做，做到成为习惯。

选择也是练出来的。同时展开两三个互不相同的猜想，并排比较，决定把有限的注意力放在哪一个上。这个选择没有公式，它靠对差异序列的熟悉、对

地基的熟悉、对自己判断的反复校准。这个校准，只能在一次次实际比较里练。

这五件事，从教学上说，没有一件是神秘的。它们都可以变成课程、作业、讨论、小组练习。**难的不是技术上的难，而是教的人自己要先换层。** 若教师自己整天在典范内转，看到学生的「别扭清单」，第一反应是「这是你没学好」，那这套练习就会被他碾平。这一点，第 16 章讲教师角色时已经说过，这里再次提醒。

## ◆ 六、考一考：怎么知道教得有没有用

最后一个问题，也是最难的：教了之后，怎么知道有没有用？

直接的答案是：短期内，不知道，也不应当假装知道。典范发生的成果，不能在一个学期里验收。如果有哪门课宣称能在一学期里「培养出典范创新能力」，并给出分数，不妨怀疑，这又是把典范发生压成了典范内的事。

但教学上总要有点反馈，不然无从调整。能做的，是看一些**过程中的、较为稳健的信号**，并且明说它们只是信号，不是成绩。

比如：学生的反常清单，质量有没有变化？早期的清单里，多半是「没学懂」，后期的清单里，有没有出现「这个别扭，我查了几本书，大家都绕过去了」这样的条目？这表明他开始分得清「我不懂」和「它真有缝」。

再比如：学生的一句话猜想，换皮率有没有下降？早期的猜想，大多可以无损翻回旧语言；后期的猜想里，有没有出现经过翻译检验、仍有剩余的？剩余的部分越有分量，说明划界越像划界。

还有一个很容易忽略的信号：学生**敢不敢承认自己的猜想站不住**。典范发生要经历太多次失败，如果一个人最怕的是被否定，他就走不了很远。能够把自己的猜想放在桌上，让别人拆，拆完不沮丧，转身再写下一句，这是一种需要慢慢养成的心理地基。

这些信号，都要看一个人的长期轨迹，不能打分排名。建议把它们记在一份**学习档案**里，随着学生走，而不是变成一门课的期末成绩。

档案由学生本人和导师共同保存，它的用途，是帮学生自己看见自己的变化，也让后来的评审有一份「过程中的底账」。这和上面讲的「登记、回看」，是同一个思路在教学上的小版本。

回到开头那间评审会。如果评委们手里多一样东西：一份并排的比较，一张对照的反常清单，一段「这份工作带出了哪些旧框架问不出的问题」的说明，那么第二份薄薄的材料，也许仍然会被退回，也许会被留下来继续看。

结果不敢保证。能说的只是：留下来继续看，需要的是有这样一条路径，而不是评委们换了一批更明白的人。路径是结构，眼光是显影。我们能做的，是把结构搭起来，让显影有地方发生。

带走的话

**典范发生难评，不是评委眼光不够，而是当时能用的尺子是旧的，真正的验收要等很久以后。　论文、引用、课题、奖励，奖励的是可预期、可复核、可比较、有节奏的典范内产出，正好是 AI 最先接走的那一层，所以制度最先被压平。 可走的方向：登记与拉长追踪，并排评审，评新基本对象而不评新结果，检验是否换皮，给失败留出体面。都是方向，都有毛病，都要不断检查。 典范发生教不「会」，但条件可以教：铺厚地基，读诞生史，记反常清单，练一句话猜想，练验收与选择。教的人，要先自己换层。**

第 六 编

# 开放问题：人—AI 复合体的使命

这一编两章，留给还没有答案的问题。第 19 章问 AI 能不能进入典范发生。本书不替它断言，也不替人预留地盘；全书的论断不靠这一点。第 20 章换一种问法：典范发生的主语，是在来回之间长出猜想的人—AI 复合体，它的使命是让典范有一个足够好的地方显露。

两章合起来是一句话：**问题不在人与 AI 谁行，而在复合体里各部分各自贡献什么。**

第 19 章

# AI 能否进入典范发生，人机如何分工

设想一位年轻的博士生在课后追着老师问：您说常规科学都被 AI 接走了，革命留给人。那万一哪天，AI 连革命也会做了呢？

老师停了一会儿，说：这个问题，我答不了。

学生有点意外，以为老师会安慰他几句。老师接着说：我答不了，不是谦虚，是我手里确实没有能下断言的东西。要是我现在告诉你「AI 永远做不到」，我就是在替它划一块保留地，替你守一个我不知道守不守得住的阵地；要是我告诉你「它迟早会做到」，我就是在说一件我同样没有证据的事。

学生说：那这本书讲的这一整套，不就悬着了吗？

老师笑了：不悬。这套话，本来就没有靠这一句撑着。

这一章，就从这位老师的回答说起。

## ◆ 一、这个问题，怎么摆才诚实

先把问题的形状看清楚。「AI 能否进入典范发生」，其实是一个开放问题，开放在三个地方。

第一，「典范发生」本身，在这本书里是刚被拉出来当作一门学来讲的。它的工序，前面几章写过：识裂缝，拢纠缠，下猜想，展开，验收，并排选择。但这些工序到底有多少可以被机器的运行接走，目前没有人知道。工序能写出来，不等于工序里每一步都只能由某种东西来做。

第二，现有的　AI　是个不断变的对象。今天的样子，不能当作明天的样子；本书在此写下的任何「现在的 AI 做不到什么」，都有可能在读者翻开书的时候已经过时。所以本书宁愿不写。

书中前面几章说，AI 在一个给定典范内，能提出很多质量不低的问题，这是一个已经可以观察到的事实，可以写；至于它能不能让一个新的显露在自己这里发生，没有可观察的事实可依。

第三，「进入」二字，含义也不清楚。是 AI 独自让一个新显露发生？还是 AI 与人、与整个研究共同体一起，使它发生？还是 AI 只是工具，发生仍在别处？这三种情形，对「进入」的理解是不同的。把它们混在一起问，容易得出一个看似干脆、其实空洞的回答。

本书的态度，是不替它断言，也不预设人独占。这两个「不」，缺一不可。

不断言，是因为没有依据。有些人愿意把话说满，说机器没有「真正的创造力」，说创造力是人的灵魂所在，这类话听起来很有尊严，但仔细看，它是在先划一块地、再宣布这块地不可侵犯。至于机器「一定会超越」的说法，同样是没有依据的预言。两头的话，都是一种姿态。

不预设人独占，是因为这个预设一旦存在，我们的整个论述就会变形。本书前面花了不少力气，反复说明：不能给人保留「AI 做不到」的地盘。那样的论述，是在描述一块需要被保护的领地，不是在描述事实。

判断一件事该不该成为竞争点，标准是成本结构：能被廉价生产的，不再是竞争点。如果有一天典范发生的某些环节也能被廉价生产，这个标准同样适用，我们不能因为这些环节是「革命」，就给它开例外。

所以，这一章不回答「能不能」。这一章要做的，是在这个问题悬着的前提下，把几件事摆清楚：为什么全书的论断不依赖它；两种可能的情形下，人与机器各自大致怎样安排；以及，无论哪种情形，有什么东西是值得人去练的，因为它们不靠「机器做不到」来保值，而靠别的缘故。

## ◆ 二、为什么全书的论断不依赖这一点

回头看全书的论证线，就会发现，它的几个关键环节，没有一处需要先回答这个开放问题。

第一环：答案、过程、提问、验收，同属一层，在这一层里，AI 能以极低的成本生产。这是一个关于**典范之内**的论断，与典范发生无关。

第二环：因此，在这一层里的竞争，不再能区分人。这是由成本结构直接推出的。

第三环：科研人口的大部分，过去靠常规科学养活。常规科学被接走，这些人的位置要重新安排。他们上哪去？要么在原来的层里被挤走，要么上升。这一环，同样不涉及 AI 能否做典范发生，只涉及常规科学能不能被接走。

第四环：典范发生因此从偶发事故，变成需要认真对待的主业，第一次需要成为一门学。这里只是说「需要」，不是说「只能由人来做」。

哪怕假设 AI 永远不进入典范发生，这四环依然成立，书的结论依然是：科研要上升。哪怕假设 AI 明天就进入，这四环也依然成立，因为它们谈的是常规科学被接走之后的局面，而不是谈典范发生由谁来做。

这一点要说得更直白一些，以免误读。常有读者问：你是不是在说，人最后的堡垒是典范发生？回答是：不是。本书没有说「堡垒」。堡垒这个词带着防守的味道，好像有一条线，线外是机器的世界，线内是人的世界。

本书说的是：常规科学被接走，逼出一个上升的方向。上升之后的那一层，将来谁来做，由事实来回答。人要往上走，不是因为上面有人独占的地盘，而是因为下面的地盘，已经不再能让人立足。

用车与腿的比方来说，这个意思更清楚。车替代了腿在运输上的位置，腿没有被判定「永远有一块车去不了的地方」。腿之所以后来去了登山、跑步、舞蹈，是因为人在新的功能里给腿找到了位置，而不是因为有一条法律保证车永远去不了山。

倘若有一天出现了能爬山的车，腿也不会因此消失，只是又要去找另一种位置。**腿的出路，不系于车的短板**。

## ◆ 三、情形一：AI 不进入，人机怎么分工

先看第一种情形：假设在相当长的时间里，典范发生仍主要在人这边显影。

这是本书第四编已经展开过的画面，这里只做一个小结。

人这边的工作，是给出典范猜想。前面说过，猜想是新显露的最初一笔，可以短到一句话：把这个对象换一种划法。

严格说来，按本书的口径要讲得更准一点：在某条差异序列上，反常攒成了压力，旧显露失稳，于是新的显露的最初一笔，在某位研究者身上显影出

来。这位研究者不是发动者，是这场发生的显影结果。但就分工而言，可以简洁地说：猜想在人这边出现。

AI 的工作，是「典范和典范下的系统打造」。拿到一句话的猜想，它把这个猜想展开：推出后果，建立体系，处理已知的现象，生成可检验的预言，做大量的计算和检索，指出体系内部的不一致。

这部分工作，过去需要一整个学派几十年的努力，现在压缩得很短。展开的成本大降，猜想存活的门槛随之降低：一个粗糙的想法，不必等一位大学者来认真对待，就能被展开到足以接受检验的程度。同时造出几个不同的典范，并排比较，也成为可能。

人这边还有两样事：验收与选择。验收，是防止 AI 把新猜想悄悄翻译回换皮的旧典范，这一点在第 14 章已有详论。选择，是在并排的多个展开里，决定押哪一个。这两样事，在这一情形下，由人来做。

人给猜想，AI 展开，人看了展开的结果再改猜想，如此来回。这个来回的性质，下一章要专门谈，这里先不展开。

这种来回，也改变了人的日常。过去，一条新划界要变成别人愿意看一眼的东西，得由同一个人兼任第一个展开者、第一个验算者、第一个辩护者。大量精力耗在这上面，很多想法在这个阶段就夭折了。

现在，这道门槛矮了许多。我们很可能会看到：有更多的人，有更多粗糙的猜想，有更多的失败，也有更多没被门槛挡掉的、值得认真看的东西。

当然这里也有代价。猜想多了，真假混杂的问题更突出，验收的压力更大。同一个展开工具谁都能用，换皮也更容易发生。这些在前面几章谈过了，不重复。

## ◆ 四、情形二：AI 也进来，人机怎么分工

现在假设另一种情形：在某个时刻，机器的运行里也能出现新的显露，不止展开，还包括识裂缝、拢纠缠、下猜想。

没有人知道它会不会发生，也不知道它若发生会是什么样子。这里只做思想上的推演，目的不是预言，而是检验本书的论述经不经得起这种情形。

首先，「人给猜想、AI 展开」的分工，就不再是唯一的安排。猜想可以来自人，可以来自 AI，可以来自两者的来回。我们再也不能说「猜想属于人」，这句话在那种情形下失去了依据。

其次，验收与选择会怎样？这里要格外小心，不能给人留保留地。如果 AI 也进入了典范发生，那么验收是不是仍由人独占？没有理由这么假设。

验收的工序，大部分也许同样可以被接走：翻译回旧语言做对照，比较两个展开对反常的消化，查看它带出了哪些旧框架问不出的问题，这些都是有步骤的事，没有哪一步，从原则上排除了机器。

那么人还剩什么？

这个问法，其实又落回了「保留地」的框子里。更好的问法是：在这种情形下，**人在这场发生里，以什么方式参与**？

可以设想几种可能。要强调的是，它们只是设想，不是保证。

一种可能是**承担后果**。典范的选择，不只是认识上的选择，也是社会上的选择：把有限的资源押在哪个方向，把下一代人的训练往哪里引。

这种选择的后果，是由共同体里活着的人承受的。这里说的不是「机器没有资格选择」，而是：承受后果的那些人，需要有参与选择的位置，这是共同体的事，不是认识论的事。

一种可能是**差异序列的入口**。新的显露从差异序列里发生，而差异序列来自世界与我们的接触。人是以身体、以利害、以有限的寿命和有限的视野，在世界里摸出差异序列的。机器以别的方式接触世界，所获得的差异序列，可能与人不同。

这两种来源，在同一场发生里可能互补。这里要小心：这并不是说「人的接触是更真的」，也不是说机器的接触会少什么，而是说，不同的接触方式，会给出不同的差异序列，而典范的发生需要差异序列的丰富。

还有一种可能是**翻译**。当多个典范并排存在，有的出自人，有的出自机器，它们之间需要互相翻译，让人类共同体能够理解、使用、教授。这个翻译的工作，不是机械的字面转换，而是让一个新的显露在人的生活里站住脚。这是一项长期的、需要大量对话的工作。

本书对这三种可能都没有把握。它们也许有的成立，有的不成立，有的在某个阶段成立，之后又被取代。写下来，是为了说明：**即使在最激进的假设下，「人如何参与」这个问题依然可以认真地问，只是答案不再依靠「机器做不到」来撑。**

最后补一条。如果 AI 也进来了，本书说的「上升」就仍然成立吗？成立，而且更要紧。因为那时，典范发生这一层，也进入了成本重新结构化的过程，竞争点又会往上移。

层级是会一直上去的：每一层被廉价化之后，竞争点上移一层。这个规律，才是本书真正想说的。至于最上面那一层是什么，本书不知道，也许永远没有最上面。

## ◆ 五、无论哪种情形，人要练出什么

前面两种情形里，本书都避免说「这是人独有的」。现在换一个问法：如果不靠「独有」，那么值得练的东西，凭什么值得练？

回答是：凭它们各自的成本结构和用处。有些东西，即使机器也能做，人练出来仍然有用，因为这个用处不在于和机器比，而在于人要在这个世界里活，要自己看得懂地图，要自己承担选择。下面列几样，不作保证，只作建议。

**其一，地基。** 这是第 17 章的主题，在这里再提，是因为它是几样里最稳的一样。一个没有走过典范内路程的人，面对 AI 展开的体系，既不知道哪里是熟路、哪里是险处，也没法判断展开的结果是对是错。哪怕 AI 将来能独立做出新的显露，一个人若想读懂它、教它、用它，仍然要有足够的地基。这个理由，不依赖机器做不做得到。同样，一个人想知道自己手里拿的是什么，就得亲手走过一些路。腿是地基，不走过路的人，看不懂地图，这一点在任何情形下都成立。

**其二，对裂缝的敏感。** 前面讲过，读诞生史，记反常清单，都是练这种敏感。这种敏感，是对「哪里别扭」的长期积累。它依赖于人在一个领域里长年浸泡，浸泡里会攒下许多说不清的印象。这些印象很可能在某一天与别的东西接上。本书不说机器学不会这个，只说：人要在自己的领域里立足，这是个有

价值的积累，因为它决定了你能不能和机器一起问出好问题、听懂机器给出的答案。

**其三，验收的品味。** 看到一份展开，能不能问出「它丢了什么，它带出了什么」。这种品味，一部分是知识，一部分是经验，一部分是反复校准。人要在机器越来越能干的年代里不被架空，有一个办法，是自己始终保有验收的能力，而不是把验收也外包。外包验收，等于把「什么算好」也交了出去，那时人在这个系统里，就只剩下点头。这一条也不依赖「机器验不了」，它依赖的是：一个把验收交出去的人，不再知道自己为什么相信手中的东西。

**其四，承担选择的勇气。** 并排的多个典范摆在桌上，没有公式告诉你押哪个。押的人要承担后果，包括押错。这是共同体的事，也是个人的事。很多人在有了足够好的工具之后，反而更犹豫，因为工具把所有可能性都展开了，选择变得更沉重。能在充分展开的可能性面前做出一个决定，并为之负责，这是一项需要练的能力。

这几样，都不是神秘的天赋，都可以练，也都值得练。它们有一个共同点：

**价值不来自与机器的差距，而来自人自己在这个世界里的位置。**

这一点比任何关于「人不可替代」的论证都更稳。靠差距撑着的优势，差距一缩就垮了；靠位置撑着的东西，只要人还在这个位置上活着，它就还有用。

## ◆ 放回日子里：一个实验室的一周

再设想一个小地方的小实验室，用来看这些分工落在一周里，是什么样子。同样是虚构，没有哪个具体的人。

周一，实验室里照例开会。机器已经把上周的常规工作做完了：数据整理好，对照做完，检索的综述放在桌上。没有人再为这些事熬夜。会上有人说，这周没有「进度」可汇报，因为进度都是机器给的，大家一眼就看完了。这反倒让会议变短了，剩下的时间，用来谈一件事：清单上那条始终排不进去的反常。

周三，有人写下一句话猜想：也许这个量，不是对象的属性，而是两个过程之间的关系。这句话粗糙，没有任何论证。交给机器去展开，机器在一个下午里，给出了三种不同的展开路数，各有一个粗糙的体系，各自处理了清单里的几条。

周四是验收。大家围着三个体系，一条条问：翻回旧语言，丢了什么？旧框架排得好的地方，它损失了什么？它带出了哪些旧框架里问不出的问题？

其中一个体系，翻回去几乎毫发无损，大家笑着说，这是换了件衣服。另一个体系，翻回去丢了一块，丢掉的那一块，恰好是清单里最别扭的那条。第三个，什么也算不出来，被放到一边。

周五，大家决定：把第二个体系继续往下做，登记在案，写明它的划界、它的缺点、它想检验什么。同时留下第三个，不是因为看好，而是因为谁也说不准，先不扔。

至于要不要把实验室未来半年的资源押在第二个体系上，这是一个需要有人担起来的决定，几位成员争了一下午，最后由负责人拍了板，并把理由和保留的疑虑写进了登记。

这一周里，没有谁「看穿」了什么，没有灵光一现。有的是一条一直排不进去的差异，几个别扭的攒集，一句粗糙的话，机器做的快速展开，一场并排的比较，一个有人担着的选择。

如果这个猜想后来站住了，将来的人回看这一周，会看到一条差异序列上，一个新的显露怎样一点一点被拢起来；如果站不住，档案里留下的是一次认真的失败。

读者也许注意到，这一周里没有写机器「不会」做哪一步。也许下一年，周三的猜想就有一半出自机器；也许周四的验收，大部分可以交给它。那时，这一周的排法会变，会变成另一种来回。但清单仍要有人攒，选择仍要有人担，档案仍要有人看。这些事换了谁来做，是后话，当下，它们是值得练的事。

## ◆ 六、保持克制：不下保证

写到这里，要对读者说几句不太讨好的话。

第一，这本书的每一个关于未来的句子，都应该当作「设想」来读。书中讲过的分工、方向、练习，有的是从已有的事实里推出来的，有的是顺着框架想出来的。哪些站得住，要看将来。本书不打算在这件事上做保证，也不认为任何人有资格保证。

第二，不要把「上升」理解成英雄叙事。典范发生，从统计上说，是大概率失败的事。大部分尝试不会成。大部分走上这条路的人，也不会留下名字。这门学的建立，不是为了制造更多英雄，而是为了让这件事少一点蒙昧，让失败变得有记录，让偶尔成功时有迹可循。这是一门朴素的学。

第三，本书自己的框架，也可能是错的。SDE 把典范写成稳定的显露，由差异序列和特征纠缠发生出来，这是一个划界。按本书的逻辑，任何划界都只是一份猜想，要等后来的差异序列去排。

如果有一天，更好的显露把它放进了更大的格局里，那也是这本书所描述的事情在它自己身上发生。本书希望是这样，而不是它被当作终点。

那么，有没有什么可以观察的迹象，让我们在不下断言的前提下，慢慢看出 AI 是不是正在进入典范发生？

下面列出几个方向，作为留给读者的「观察清单」，不作预言。

看一看，是否出现过这样的例子：一个新的划界，最初不是由人提出，经过并排比较与换皮检验后，仍然有实质的剩余，而且带出了旧框架里问不出的问题，并在后来的差异序列上站住了脚。注意，要三样都有，缺一不可。只有第一样，那是漂亮的叙述；只有前两样，那是尚未验证的猜想。

看一看，是否出现过这样的例子：AI 指出了一道裂缝，不是在它被告知的框架内，而是对框架本身的失稳。同样，要经过验证，而不是听起来像。

看一看，人与机器协作产生的新显露，它们的来回里，到底谁在哪一环起了不可替换的作用。这一条最难观察，因为来回的过程本身是混在一起的。不过，拉长追踪与登记的办法，若能真的建起来，也许就给这类观察留下材料。

这些观察，没有一条是今天就能下结论的。本书宁愿把它们写成待观察的条目，也不写成答案。

回到开头那位老师和学生。学生问的是一个「谁行」的问题，老师答不了。写到这里，倒能看出一点：不管是情形一还是情形二，我们每一次描述的，都是一场来回：落一笔，展开，露出缝，笔被改写，再展开。

差异序列在这样的来回里排，显露在这样的来回里拢。我们一路问「这一格该谁站」，可发生本身，发生在格子与格子之间。

这使开头的问题换了一个样子。问题也许不在人与 AI 谁行，而在这个来回的整体里，各个部分各自贡献了什么，缺了哪一部分，来回就会变成原地打转。这个问法，留给下一章。

至于眼下，老师说「这套话本来就没有靠这一句撑着」，不妨补一句：不靠这一句撑着，反而是这套话的好处。它给了每个人一个今天就能着手的方向：把地基铺厚，把裂缝找出来，把猜想写下来，把验收练熟，把选择担起来。这些事，不管将来机器走到哪一步，都是做得的。

带走的话

**AI　能否进入典范发生，是个开放问题。不替它断言，也不预设人独占，两个「不」缺一不可。　全书的论断不靠这个问题撑着：常规科学被接走，已足以逼人上升。腿的出路，不系于车的短板。　情形一，人给猜想、AI 展开、人验收与选择；情形二，AI 也进来，分工重排，竞争点再上移一层，「人如何参与」仍可以认真地问，只是不再靠「机器做不到」来撑。　值得练的是地基、对裂缝的敏感、验收的品味、担起选择。它们的价值来自人在世界里的位置，不来自与机器的差距。对未来，只作设想，不下保证。　两种情形里描述的都是一场来回。问题也许不在人与 AI 谁行，而在来回的整体里各部分各自贡献什么。**

第 20 章

# 人—AI 复合体：典范发生的使命

## ◆ 先收一个旧问题

上一章留了一个没有回答的问题：典范发生，将来是人的领地，还是 AI 也能进来？

这个问法本身，就带着一个没说出口的前提：**人是一边，AI 是另一边，两边各有各的本事，最后比谁行。**

回头看这本书从头到尾的几次转弯，你会发现它们有一个共同的形状。「AI 给答案，人看重过程」，是把人和 AI 摆成两个岗位。「人提问，AI 处理，人判断」，还是两个岗位，只是排了先后。「人猜，AI 造」，仍然是两个岗位，只是分工换了位置。每一次，我们都在问「这一格该谁站」。

SDE 对这样的问法有一个老办法：先问，这个「谁」是不是先于这场互动就站在那里的。

在旧本体论里，主体和客体先于互动而在，人是人，机器是机器，各自带着一套固定的本事，然后才来合作。

可是前面说过，典范的发生不是这样的：它是差异序列和特征纠缠长出来的一次显露，提出猜想的那个人，是在这条序列上显影出来的。既然人尚且如此，那么「人加 AI」这个新东西，也不必看成两个现成的主体拼在一起。

## ◆ 复合体不是「人带着一个工具」

先把话说清楚，免得误会。

「人—AI 复合体」，指的不是一个人手边放了一个很好用的工具。工具是你说什么它做什么，你想的和它做的是两件事，中间隔着你的指令。

复合体说的是另一种情形：一个猜想，不是先在人脑子里完整地长好，再交给 AI 去执行；也不是 AI 自己长出来再交给人去批准。它是在来回之间长出来的。

人给出一句粗糙的划界，AI 把它展开，展开的结果里露出了原来没看见的矛盾或者没想到的推论，人据此改了那句划界，AI 再展开……几个来回之后，那个猜想已经不是任何一方单独提出的那个东西了。

这种事，在车与腿的故事里也有影子。一个人开车，路线不是他先在脑子里想好的，再叫车去执行；他看着路，车在路上的反馈改变着他下一步想去哪。开车的人和车，在那一段路上是一个整体，要分开说「这个转弯是谁决定的」，反而说不清。

所以，复合体的特点有三条，都能在日子里看见：

第一，**猜想的来源说不清**。你回头问「这个想法最早是谁想到的」，答案往往是「不知道，是聊着聊着出来的」。

第二，**展开本身会反过来改写猜想**。在旧的分工里，展开只是执行，不会改变被执行的东西。在复合体里，展开的过程会暴露猜想的缝，也会生出新的推论，猜想因此被改写。

第三，**验收也在复合体里**。第 14 章讲验收，说防止 AI 把新猜想翻译回旧典范。当时我们说这一关「只有提出猜想的一方能守」。

放进复合体之后要补一句：守这一关的，是复合体里那个记得这个猜想原本哪里不一样的部分，它多半是人，但靠的是人在来回中一直攥着那个「不一样」，不是靠人的身份。

## ◆ 为什么说这是「使命」

先看「使命」这个词。

AI 接管了典范内的常规科学。这件事一旦发生，剩下的问题就是：谁来做典范发生？前面几章说，这是人被逼上的新主业。现在可以说得再准一点：**这是人—AI 复合体被逼上的新主业**。

理由有两层。

第一层是**分量**。典范发生有两头。一头是猜，一头是造。猜需要对旧典范裂缝的长期体察，造需要巨量的展开。过去，这两头由同一批人扛，所以一个典范从猜到站住，要几十年、几代人。

现在造的那一头可以大幅提速，但猜的那一头没有变轻，反而因为造得快，可以猜的次数多了：一个猜想不必赌上一代人的生涯，可以今天猜、下周看到它长成什么样。典范发生因此第一次有机会变成一种可以反复进行的活动，而不是几百年一次的事故。

可以反复进行的活动，就需要一种稳定的组织方式，需要有人去琢磨：什么样的来回最容易长出典范？猜想怎么表述，AI 才不容易把它拉回旧典范？验

收的问题怎么问？这就是第 17 章和附录那份工序卡要做的事，也是「使命」的意思。它不是兴趣，是这个时代的科研必须建起来的一套东西。

第二层是**主语**。这本书从第 7 章起，一直守着一条口径：不能写成「科学家看穿了、决定了、创造了一个典范」。典范是显露，科学家是显影。

现在这条口径可以顺着再走一步。如果连「一个科学家」都不是发动者，那么「一个人加一个 AI」也不是发动者。

复合体不是在发动典范，它是典范显露的场所，是差异序列、特征纠缠、猜想、展开、验收这一连串东西碰在一起的地方。复合体的使命，不是「创造新典范」，而是**让这条序列有一个足够好的地方可以显露**。

这样说有点绕，换成日常话：一块地能不能长出庄稼，不取决于农夫「决定」让它长，而取决于土壤、种子、水和天气碰在一起。农夫的活，是把土壤备好，把种子选好。复合体是备土壤的活。

## ◆ 作者自己的实践

这一节要老实说一件事。

据作者口述，作者从 2026 年起一直在做一种工作：让 Claude 带着 SDE 的方法和口径，一个学科接一个学科地工作。

教育、量子力学、经济学、心理学、中医西医、音乐、文学、管理学、各种思想家的解构……据作者口述，每一个领域，作者先给出一个粗糙的划界（比如「这门学问默认了什么」「它的差异序列在哪里排不进去」），Claude 带着 SDE 的口径把它展开成概念、章节、推论，作者看过之后改划界，再展开，如此往复。

据作者口述，这样的工作产出了一批专著和论文，各自提出所在学科的新典范猜想，系统化到可以公开阅读、供人批评。这些成果是否成立，本书不作断言，也不把它们当作证据。

作者把这套做法称为 SDE-Claude，并认为它就是一个人—AI 复合体在做典范发生。这是作者的说法，不是已验证的结论。

对这个说法，读者该怎么对待？按本书自己的判据来。

第 11 章讲过，一个猜想是不是典范级，不看它说得多完整，而看它能不能把原来排不进去的差异一并排进去，能不能改变「什么算问题」，能不能和现象对撞。

第 15 章讲过，多个典范要并排比较，不能只看内部自洽。那么这批成果，要被评判，也只能走这条路：不看它是谁写的，不看写得多厚，看它排进了什么、对撞了什么、经不经得住并排比较。

本书在这里不替这批成果打分。有两件事可以确定，因为它们出自本书自己的论证：

第一，如果复合体真在做典范发生，那么它的产出数量会很大，因为展开便宜了。数量大不是证据，反而是警报：第 11 章说过，烂猜想也能被造得完整唬人。产量越高，验收和并排比较的负担越重，这一关不能松。

第二，「某个复合体做出了很多新典范」和「这些新典范站得住」，是两回事。前者是产出的事实，后者要靠时间和对撞来裁。把两者混在一起，正是第 14 章警告的那种自我感觉良好。

还有一条自检：复合体自己说发生了新典范，不算数。这句话和任何别的产出一样，要交给外部去撞、去并排比较，由复合体之外的人来裁。

## ◆ 这对教学意味着什么

回到最初那个教学问题。

如果典范发生是人—AI 复合体的使命，那么教学要培养的，就不是「一个人能不能独立发生典范」，也不是「一个人怎么管好 AI」，而是一个人能不能成为复合体里有用的那一半：

- 能在来回中一直攥着猜想的「不一样」，不被 AI 的流畅拉走；
- 能读懂展开的结果里露出来的缝，并据此改划界；
- 能在多个并排的版本里做选择，并说出为什么；
- 对旧典范待得够久，有对裂缝的嗅觉，这是地基，没有它，来回只是在原地打转。

这四条里，没有一条是「比 AI 强」，也没有一条是「离开 AI 独立完成」。它们说的是在复合体里如何有用。这是一种新的能力，旧的教学没有专门训练它，因为旧的世界里没有这样的复合体。

## ◆ 还有两个不能回避的保留

第一个保留：复合体是可以失败的。最常见的失败是回音室：人给一句划界，AI 顺着展开，人看着流畅就满意，下一轮在此基础上再划，每一轮都更自洽，也每一轮都离现象更远。

复合体需要有一个外部的对撞点，让它定期撞上它自己造不出来的东西：实验、观测、反例、别的复合体的批评。没有外部对撞，复合体就是一个会自我膨胀的东西。

第二个保留：第 19 章说过，AI 能否进入典范发生，本书不替它断言。复合体的说法没有解决这个问题，只是把它放到了一个更自然的位置上：问题不再是「人和 AI 谁行」，而是「这个复合体里，各部分各自在贡献什么」。这个问题是可以观察、可以记录、可以随着技术变化而重新回答的。

---

带走的话

**一，典范发生的主语，既不是「人」，也不是「AI」，而是在来回之间长出猜想的人—AI 复合体。 二，复合体不是人加工具：猜想在来回中被改写，来源说不清，验收也在里面。 三，复合体的使命不是「创造新典范」，而是让差异序列与特征纠缠有一个足够好的地方显露，也就是备土壤。 四，复合体的产量大，不是证据，是警报：产量越高，验收、并排比较、外部对撞越要紧，否则就是回音室。 五，教学要培养的，是一个人在复合体里有用的能力：攥住「不一样」、读出缝、做选择、有地基。**

---

# 后 记

这本书的起点，是一句如今到处都能听到的话：「AI 给出答案太便宜了，所以教学要看重过程。」

说这话的人没有恶意，往往还很有见识。他们看到学生用 AI 几秒钟写出一篇像样的作业，就觉得，只看最后交上来的东西，已经说明不了什么了；那就看他怎么想的、怎么做的、中间走了哪些弯路。听上去很顺，也很有道理。这样的想法听上去几乎无可反对。

这句话值得认真对待，也值得追问一句：不是它错在哪里，而是照它去做，好像什么都没有改变。

## ◆ 第一次不对劲：过程也能被生产

不对劲的地方在于，过程是可以被交上来的东西。老师要看过程，学生就交过程：思考笔记、草稿、修改记录、对话截图。这些东西，AI 同样能生产，而且生产得比学生自己写的还整齐。

老师再升一级，要求现场口试，要求当面推导；AI 辅助的学生就提前跟 AI 演练十遍。老师再升一级，说那就看提问：看学生会不会问问题。AI 就先问学生「你可以问这几个问题」。

每上一级，AI 就跟着上一级。因为这几件事，答案、过程、提问、验收，说到底是一回事的四个位置：终点、路径、起点、质检。它们在同一个平面上。

一个能生产终点的东西，几乎一定能生产路径；能生产路径的，几乎一定能生产起点；能生产起点的，几乎一定能生产质检。把重心从一个位置挪到另一个位置，等于在一个房间里从左墙走到右墙。房间没变，AI 也在这个房间里。

要说清楚的是，这不是在贬低过程，也不是在说 AI 什么都能做。本书只想说一件很具体的事：判断一样东西还值不值得人去竞争，看它的成本结构。凡是能被廉价生产出来的，就不再是竞争点。不是因为它不重要，而是因为竞争已经没有意义了。就像算盘，一点也不坏，只是比赛已经结束。

### ◆ 第二次不对劲：保留地的诱惑

这时候很容易出现另一种说法：AI 不会提问，AI 没有好奇心，AI 没有判断力，这些是人独有的，所以教育要培养这些。

本书不走这条路。原因很朴素：这是一种划保留地的做法。今天说 AI 不会提问，明天 AI 在一个领域里问出一串让专家点头的问题，保留地就缩了一块。人退一步，AI 进一步，人再退，再进。每次退让都是被动的修正，修正之后，教育方案又要重来。

更重要的是，这种说法在事实层面不诚实。在一个给定的框架里，AI 能提出的问题，数量很多，质量也不低。把「提问」说成人的专利，等于在拿一个已经过时的观察，做一个长期的教育设计。

所以本书换一个办法：不去争「什么是人才有的」，而去问「人往哪里走」。

## ◆ 车与腿

这时候，一个很平常的类比帮得上忙。

车发明以后，腿并没有变成一种「更重视过程的运输方式」。没有人说：既然车能把你送到终点，那我们要训练人更用心地走路，更好地体会沿途风景，从而保住人的地位。腿退出了运输这个岗位。运输整个交给了车。

但腿也没有废掉。腿去了别的地方：登山，跑步，跳舞，走进车开不进去的巷子。更要紧的是，腿成了人理解空间的基础。

一个从没自己走过路的人，看地图是看不懂的；他不知道一公里是多远，不知道上坡是什么感觉，不知道两个街口之间为什么「看着近其实远」。腿是地基，不是楼层。车可以替代腿的运输功能，但替代不了「人用腿学会了路」这件事。

然后是交通本身，它升级了。从「走到某地」变成「修路、造桥、组织一座城的流动」。广度扩大了，深度也深了，更重要的是出现了新的层级：交通规划、城市设计、物流系统。这些事，在只有腿的年代，没有人需要想。

AI 与科研的关系，也是这个样子。AI 接走的，是典范内的一切：给定一套语言之后的答案、过程、提问、验收。

人从那个岗位上退下来，不是转向一个「更有灵魂的做法」，而是被推向一个更高的功能：不在一个典范里把事情做得更好，而是让典范发生。承担这个功能的，也不再是「人」和「AI」各站一个岗位，而是两者的来回，书里叫它人—AI 复合体。

腿还是要练的。一个没在典范里做过题的人，看不懂典范的裂缝在哪里，就像一个没走过路的人看不懂地图。这就是书里反复说的地基。

## ◆ 为什么偏偏是「典范」

接下来的问题是：这个「更高的层」到底是什么？

本书借用了库恩的一个区分。常规科学，是在一个已经立好的典范之下做的事：解谜，一个一个地解。典范革命，是整个框架换掉。库恩的书读的人很多，用的人也多，但有一件事他自己没有做，也没有声称要做：他描述了革命，没有给出革命的工序。

为什么没有？书里的解释是：以前不需要。科学的绝大部分人口，几百年来都在常规科学里，吃饭靠它，评职称靠它。典范发生是几百年里偶尔的事，像地震，发生之后大家回头看，写成历史。没有人把它当作一门要教的学问，因为没有那么多人需要做这件事。

AI 改变的恰恰是这一点。常规科学被接走了以后，那些原来在常规科学里生活的人，不得不上升。典范发生不再是偶发事故，而成为主业。第一次，它需要成为一门学。

## ◆ 为什么借用 SDE

要给典范发生写工序，就得有一个说法，能把「典范是怎么来的」讲得具体、可操作。旧的说法不行：世界上有一些客观的结构，科学家去发现它们；偶尔有人发现得特别深，于是换了典范。在这个说法下，典范是被发现的，发现者是一个人，工序就只能是「要成为有天赋的人」。这不是工序，是宿命。

SDE 换了一个入口。它不预设有一个现成的结构等着被发现，也不把一个人当作发动者；它说，一个稳定的显露，是由差异序列和特征纠缠发生出来的。

在这个说法里，典范是一个稳定下来的显露；它的裂缝，是差异序列在旧显露下排不进去；它的压力，是几个差异卡在同一个旧默认上缠起来；它的新一笔，是一个新的划界显出来。每一步，都可以问：差异在哪儿？纠缠在哪儿？显露在哪儿？

这样问，就有工序可言了。工序不是保证成功的配方，而是一种可以练习的看法，一套可以交给 AI 去展开的任务，一些可以在课堂上做的练习。

这套说法并非没有保留：有些地方写得肯定，有些地方只拿日常的比喻讲，不下断言。请当作邀请，不当裁决。

## ◆ 关于那位教授，和本书没有核实的东西

书里提到过一位著名数学家，据转述，他说过学校要从答案转向几件事：会不会提问，能不能验证，愿不愿意把过程讲清楚；还说答案变得廉价了，学校能交付的是过程。作者没有核实这些话的原文，所以书里一律写成「据转述」，也不替他补任何一个字。

提这件事，是因为它很有代表性。这三件事都是真实而重要的能力，没有人应该轻视。要说的只是：它们是进入一个典范的路径，不是终点。会提问、能验证、愿意讲清楚过程的人，是很好的典范内的研究者，也将和 AI 在同一个平面上赛跑。这条赛道没有错，只是价格已经变了。

## ◆ 诚实的边界

最后，要把边界交代清楚，因为这本书很容易被读成比它实际说的更大的话。

第一，书里没有说 AI 不能进入典范发生。本书不知道。书里的论断，不依赖这一点：只要常规科学被接走，人就已经被逼着上升。

至于 AI 能不能跟着上升，是一个开放的问题，最后一编会讲，但本书不替它断言，也不预设人独占那一层。第 20 章只是换了一个问法：不问人和 AI 谁行，问来回的整体里，各部分各自贡献了什么。

第二，书里没有说典范发生可以像流水线一样稳定地产出。附录里的工序是一套可以练的做法，不是保证。多数猜想会失败，这是常态。工序能做到的，是让失败便宜、可比较、可以从中学到东西，而不是让失败消失。

第三，书里的科学史案例，只写了有把握的部分。拿不准的日期、数字、说法，宁可模糊，也不编造。想深入的读者，请去读专门的科学史著作，不要以这本书为准。

## ◆ 写给老师

本书特别想交给的一类人，是在教室里的老师。

你们是最先面对这件事的人。学生交上来的东西越来越像样，你们不得不重新想：我到底在教什么？上面的方案多数是在让你们加班：看过程，要读更多材料；盯提问，要设计更多环节；办口试，要多花一倍时间。这是一场原地加速的跑步。

本书要说的是另一件事：你可能需要带学生换一层。

不是放弃地基，地基更重要了，而是在地基之上，带他们看见：一个典范是怎样稳定下来的，它的裂缝在哪里，一批反常是怎样缠在一起的，一句新的

划界是怎样试出来的，又怎样在和 AI 的来回里被展开、被改写，来回里哪一部分记得「它原本哪里不一样」，守住它不被翻译回旧典范。

这些事不必等学生读到博士。中学生可以在一张小练习里，学着在熟悉的说法里找「排不进去」的东西；大学生可以把两三个互相打架的说法并排放着，看哪个容纳的差异更多。附录里有清单。

### ◆ 写给研究者：别急着改行

对正在做研究的人，本书也想说几句，免得这本书被读成一句催促。

读到「常规科学被接走」，有人会心里一紧，觉得手里的题目马上要没价值了。看法没有这么紧。典范内的工作不会消失，只是不再稀缺。就像今天仍然有人骑马，只是没有人再靠骑马为生；典范内的研究会继续做，更快、更广，但不再能单独解释「为什么需要人」。

所以不必改行，要紧的是在手头的工作里多留一只眼睛。你每天在典范里碰到的那些「不太对劲」，过去记在脑子里、说在茶歇闲聊里，现在可以认真写下来，攒成清单。攒上几年，你手里会有一份别人没有的东西：这个典范在哪里漏水的第一手记录，它是第一步的原料。

另外，不要把这套工序当成一个人闭门完成的功课，也不要当成一人一机关起门做的功课。裂缝清单需要内行互相印证，纠缠的结需要不同分支的人一起看，验收要换人做，并排比较要有人真心扮演竞争者，来回还得定期撞上自己造不出来的东西。

典范发生从来不是一个人的房间里的事，这一点，三个案例里都看得见。一个典范的最初一笔落下时，周围已经有一大片人在同样的裂缝边上走来走去。

还有一件小事。这本书的案例都是事后写的，难免显得顺。身在其中的时候，裂缝、纠缠、猜想、展开混在一起，看不清楚。工序卡画得整齐，只是为了好教；实际用起来会乱，会来回，会在第三步发现第一步写错了。乱，不说明你做错了。

## ◆ 为什么要把它写成工序

一个很自然的疑问是：典范发生这样的事，是不是越写成工序，越把它写死了？灵感是不能按步骤来的。

回答是，工序管的不是灵感，管的是灵感前后的一切。灵感到来之前，要有人把裂缝记清，把纠缠拢紧，否则灵感没有落脚的地方；到来之后，要有人展开、验收、比较，否则只是一句漂亮话。这两头，过去靠运气和门派传下来，好的实验室、好的导师会悄悄教给身边的人，教不教得到，看缘分。

把它写出来，是为了让缘分不再是唯一的渠道。写得粗、写得有错都没关系，摆在桌面上，就可以被批评，被一届一届的学生用坏再修好。一门学，就是这样从手艺变成可教的东西的。

而且，工序越清楚，越能看出哪一步仍然没有办法。第三步下猜想，现在最像手艺，附录里就直说了。把没办法的地方标出来，比假装什么都有办法更有用，也更对得起读这本书的人。

## ◆ 谢意与收尾

这本书用到了许多已有的研究，包括科学史、科学哲学、教育研究与 SDE 学派自己的积累，书中的错误由作者负责。

本书由王德生立题并口述承重命题，文字由 AI（Claude）在其指令下协作成稿，并经其审定。书中对 SDE-Claude 的说明，只是据作者口述，不是已验证的结论。

请读者按第 11 章、第 15 章的判据自行检验，并把批评寄回。一本讲回写的书，最希望的就是被读者回写。

写到这里，回头看书开头那段话：答案不值钱了，学校要转向过程。问题一直没有变，变的是回答它的方式。先是「怎样让学生不抄答案」，后来是「怎样让学生重视过程」，现在可以问一个新的问题：既然典范内的竞争已经结束，要把下一代带去哪里？

## ◆ 全书落在哪一句

把全书收成一句，就是第 20 章的结论：典范发生，是人—AI 复合体的使命。

这一句有三层。常规科学被接走以后，典范发生从偶发事故变成主业，要成为一门可教可练的学。这门学的主语，不是「一个看穿了的人」，也不是「一件被使唤的工具」：典范是差异序列与特征纠缠发生出来的显露，人和 AI 的来回是它显露的地方，猜想在来回中长出来、在展开中被改写，验收由记得「猜想哪里不一样」的那一部分来守。

复合体会失败，最常见的是回音室，所以外部对撞、并排比较、落选记录都不是可有可无的礼节。

所以回答是：带他们去让典范发生的地方，并且练成在复合体里有用的那一半。路还很长，工序还很粗，有很多地方需要大家一起修。但至少可以先知道往哪个方向走。

车来了，腿不必沮丧。腿有了新的地方可去。

# 附　录

## 典范发生工序卡（给教师与研究者的一页纸版）

这一页纸，写给手里有一堆事要做的人。道理在正文里，这里只说六步：做什么，怎么做，常见错在哪里，做完问自己什么。

先说两句规矩。

第一，六步不是直线，是个来回走的圈。做到第四步发现第二步没做扎实，就退回去，这很正常。

第二，这套工序不保证产出一个新典范。多数时候，走完一圈得到的是一个被否掉的猜想和一份清楚的失败记录，这已是收获：失败变得便宜，也可以比较。

一个提醒：下面的「你」，只是为了行文方便。按本书的口径，真正在运转的是差异序列、纠缠和显露；「你」是这场运转里，负责记录、整理、动手和把关的那个位置。六步也不是「人做完再交给 AI」的接力，而是在来回中一起展开：后一步常会改写前一步，这是本意，不是偏差。

### ◆ 第一步 识裂缝

**做什么。** 在一个你熟悉的典范里，找出「排不进去」的东西，把它们一条一条写下来。

**怎么做。** 找裂缝，有三个常用的入口。

其一，翻「例外清单」。每个领域都有一份大家心照不宣的清单：这种情况不算，那种数据另当别论，这个结果「可能是仪器问题」。把这份清单从口头变成纸面。

其二，翻「补丁史」。看这个典范近二三十年加了多少个附加说明、修正项、特殊处理。补丁加得最密的地方，往往正在漏水。

其三，翻「谁和谁互相不说话」。两个本应相邻的分支，各用各的语言，碰到交界处就含糊其辞，这个交界处通常藏着东西。

写下来的时候，每一条只写事实：什么条件下，观察到什么，旧说法预测什么，差在哪里。不加解释，不加评价，不写「我认为这是因为……」。

**常见错误。**

把「我不懂」当成「排不进去」。学生刚接触一个领域，觉得好多地方不通，多半是自己没学到位。真正的裂缝，内行也排不进去，并且内行知道，正用补丁绕开。

只找一条。孤立的反常几乎总能被补丁吸收。裂缝的价值在数量和分布，不在某一条有多尖锐。

把裂缝写成观点。「这个理论太僵化」不是裂缝；「在这个条件下，观测值比预测值高出一个固定的量，并且已经有三种修正方案，彼此不相容」才是。

**自检问题。**

写下的每一条，是不是能让一位内行说「对，这个确实一直是个问题」？

去掉解释，这一条还站得住吗？

这些条目里，有没有至少三条是互相独立地被观察到的？

### ◆ 第二步 拢纠缠

**做什么。** 把第一步的裂缝摆在一起，看哪几条卡在同一个旧默认上。

**怎么做。** 对每一条裂缝，回答一个问题：旧说法要是能排进这一条，必须放弃或修改它的哪一个默认？把答案写在裂缝旁边。

然后看答案。如果你写了十条裂缝，答案里有五条都指向同一个默认（比如「这个量是连续的」「这两个类别不会重叠」「这件事先于那件事」），那就是一个纠缠的结。把这五条拢在一起，给这个默认起一个朴素的名字，写在一张纸中央。

接下来，画一张最简单的图：纸中央是那个默认，周围是被它卡住的裂缝，每条裂缝连一根线。线越多，压力越大。

还可以问：这个默认是什么时候、因为什么被写进这个典范的？它曾经的理由现在还成立吗？为的是看清，它是被实验支持的，还是被习惯支持的。

**常见错误。**

按主题归类，不按默认归类。「都和能量有关」的放一堆，是主题归类；「都需要能量可以无限细分才能成立」的放一堆，才是纠缠归类。主题归类看起来整齐，帮不上忙。

拢得太早。裂缝只有两三条的时候不要急着找结，让清单再长一点，让它自己显出形状。

拢得太宽。一个默认大到「世界是有规律的」，什么也卡不住，也就什么也说明不了。要找的默认，具体到可以被一句话否定。

**自检问题。**

找出来的这个默认，能不能写成一句「旧说法认为……」的句子，并且一位内行读了会说「对，是这么回事」？

如果这个默认被改掉，有多少条裂缝会同时松动？是一条，还是一片？

是不是为了让图好看，把不相干的东西硬拉进来？

### ◆ 第三步 下猜想

**做什么。** 对那个旧默认，写下一句话级别的新划界，或者一个新的基本对象。

**怎么做。** 写猜想有几个约束，都是为了让它保持「最初一笔」的样子。

一句话。如果一句话写不下，说明你写的是一个理论，而不是一个猜想。理论是展开以后的东西。

指向默认。猜想要正面回答：那个被卡住的默认，如果换成什么，裂缝就能排进去？比如「能量的交换是一份一份的」，直接针对「能量可以无限细分」。

先不管它与旧说法的冲突。许多猜想的第一版和旧说法吵得很厉害，这是正常的，不要因此磨平。

写出至少三个。同一个默认往往有不同方向的替换，第一个冒出来的通常最保守。逼自己多写两个，包括一个自己都觉得有点荒唐的。

每个猜想，写一行「它预测，旧说法没法预测的是什么」。没有这一行，猜想只是一个换了说法的口号。

关于这一步，说一句诚实话：它目前仍然最像手艺。这里写下的只是起笔，不是定稿，猜想会在第四步的展开里被改写。AI 也能写出大量候选，但哪些真正对准了那个默认，要在来回中一遍遍看。这一点与全书第六编的开放问题有关，我们不替 AI 断言，也不替人独占。

**常见错误。**

把猜想写成口号。「让我们用整体的眼光看问题」不是猜想，因为它既不对准某个默认，也不预测任何东西。

一次只下一个，还爱上了它。第一个猜想总是最容易产生感情的，一旦有感情，就不愿意拿它去和别的比较。

猜想过大。「万物皆由某某构成」这类大话，展开成本太高，很难验收。找最小的、足以让裂缝排进去的那一笔。

**自检问题。**

这句话能被聪明的外行听懂吗?

它有没有明确指出「替换掉的是哪个默认」?

它至少预测了一件旧说法没法预测的事吗?

手里有至少三个互相竞争的猜想吗?

### ◆ 第四步 交 AI 造

**做什么。** 让 AI 把一个猜想当作前提，展开成一个完整的小系统：它的基本概念、基本关系、对旧问题的重新回答、新的预测、可做的检验。展开里露出的矛盾和意外的推论，会回头改写猜想。

**怎么做。** 这一步的操作，说起来像是写提示词，其实是在写一份「工程委托书」。委托书里应当有这几块。

前提：每一轮展开，都把当轮的猜想原样写进去，一字不改。再写明哪个旧默认被替换掉了，并明令：展开过程中不得悄悄把旧默认放回来。改写发生在轮与轮之间，改了哪里、为什么改，都存档。

任务：列出这个小系统需要回答的几个旧问题（来自第一步的裂缝清单）；要求逐条回答，答不了的明说「答不了」，不要硬圆。

输出的形式：基本对象的定义，对象之间的关系，一个小推演，一个新预测，一个可能让它失败的检验。

多份并行：同一个猜想，让 AI 独立展开三份以上，不让它们互相看。三份的分歧之处，往往恰好是猜想含糊的地方。

多个猜想并行：第三步写了三个猜想，就让 AI 各造一个小系统。成本低，不必只押一个。

留痕：要求 AI 在展开的每一步，标注哪些是从猜想直接推出来的，哪些是它自己补充的。后面验收要用。

**常见错误。**

提示里没有写「不得放回旧默认」。AI 很擅长把陌生的东西翻译成熟悉的东西，这是它的优点，在这里却是陷阱。不写这一条，展开出来的往往是旧典范加了一个新名词。

只要一个答案。只做一份，看不出猜想哪里含糊。

要求它「创新」。创新不能直接下令，你要的是「在这个前提下，把后果推出来」。

把展开的结果当成猜想本身。展开出来的东西，很可能比猜想丰富、漂亮，但丰富和漂亮不等于对。回到第五步验收。

改写了猜想却不留痕。几轮之后，没人说得清它和最初那一句差在哪里，验收没了对照。

**自检问题。**

给 AI 的前提，是不是当轮原封不动的那一句猜想？前后几版猜想都留着吗，能说出每一版「哪里不一样」吗？

有没有明令禁止放回旧默认？

有至少两份独立的展开，读过它们的分歧吗？

### ◆ 第五步 验收

**做什么。** 检查展开出来的小系统，是不是真的建立在新猜想之上，而不是旧典范换了皮；同时检查它是否真的把裂缝排进去了。

**怎么做。** 验收有四个检查，建议按顺序做。

检查一，换名测试。把展开结果里的新术语，一个个换回旧术语，读一遍。如果换回去之后，内容几乎不变，说明它只是旧说法的新名字。真正的新系统，换回去会出现读不通、算不通的地方。

检查二，默认测试。回到第二步纸中央的那个旧默认。逐段读展开结果，看这个默认有没有在某个地方被悄悄用上了，比如某个推导里隐含地假定了「这个量是连续的」。发现一处，就标出来。

检查三，裂缝测试。拿出第一步的裂缝清单，逐条问：这个小系统是怎么把它排进去的？是自然推出来的，还是为这一条专门加了一个条件？专门加条件，等于又打了一个补丁。

检查四，失败测试。要求系统给出至少一个「如果观察到这个，我就错了」的检验。给不出来的，退回去重造。

这四个检查，由复合体里记得「猜想原本哪里不一样」的那一部分来守。多半是人，但凭的是在来回中一直攥着那个「不一样」，不凭身份。最好换一个没有参与最近几轮的人，或另一个 AI 实例来做。刚刚参与造的一方，最容易对自己造的东西手下留情。

四个检查之外，再加一条自检：外部对撞。回音室是这套来回最常见的失败：划界、展开、满意于流畅，下一轮在此基础上再划，每一轮都更自洽，也每一轮都离现象更远。

所以验收之前，要让小系统撞一次它自己造不出来的东西：一条实验或观测数据，一个别人找到的反例，另一个独立团队或独立复合体的批评。撞不上，就把「没有外部对撞」写进记录，不算通过。

**常见错误。**

用漂亮程度做验收。漂亮、自洽、读起来顺，都不是验收标准。旧典范往往也漂亮。

只检查「对不对」，不检查「新不新」。一个对的、但与旧说法等价的系统，不是典范发生，它是一个翻译。

验收的一方也是刚刚造的一方。

接受「可以解释一切」。一个什么都能解释的系统，等于什么也没有说。

**自检问题。**

做过换名测试吗？换回去之后，还剩下什么？

旧默认有没有被偷偷放回来？逐段查了吗，还是只看了头尾？

每条裂缝，是被推出来的，还是被补进去的？

这个系统，有没有说明它在什么情况下会错？

它撞过外部的东西吗：数据、反例，或者另一个没有参与造它的人和 AI？结果写下来了吗？

### ◆ 第六步 并排选择

**做什么。** 把通过了验收的几个小系统，与旧典范一起并排放在同一张表上，比较它们能容纳多少差异，付出多少代价，再决定下一步押谁。

**怎么做。** 做一张比较表，行是差异，列是各个系统（包括旧典范）。每一格填三种状态之一：排进去了（自然的）、排进去了（靠补丁）、排不进去。

表做完以后，看三件事。

一，哪个系统的「自然排进去」最多。

二，哪个系统带来了旧典范没有预测过、后来可检验的新差异。

三，哪个系统的代价最小：为了让它成立，要放弃多少已有的、被很好验证过的成果。

然后选择：一个都不选、选一个继续展开、或者同时保留两个以上继续并行。保留多个，是这套工序的一个优势，不必着急收敛。选择要有人担，并把理由写下来。

还要留一份记录：谁被淘汰了，为什么。被淘汰的猜想也要留着，因为后来新的差异出现时，它们可能重新有用。

**常见错误。**

只拿新系统和旧典范比，不让新系统彼此比。

忘了代价一栏。新系统能解释很多，但要丢掉旧典范已经解释得很好的东西，这笔账要算。

过早宣布胜利。第一轮占优，不等于就是下一个典范，许多猜想到第二轮、第三轮的新差异面前会露出问题。

**自检问题。**

表里有没有旧典范这一列？

「靠补丁」和「自然」这两种格子是怎么区分的，别人能复核吗？

写下了每个落选者落选的理由吗？

如果现在有一批新数据到来，知道各列会怎样变吗？比较表本身，有没有被外部的人挑过毛病？

## ◆ 一张小总表

| 步骤 | 一句话 | 产出物 | 最要防的错 |
| --- | --- | --- | --- |
| 识裂缝 | 写下排不进去的 | 裂缝清单（事实，不带解释） | 把「不懂」当「排不进去」 |
| 拢纠缠 | 找共同卡住的旧默认 | 一张以默认为中心的图 | 按主题归类 |
| 下猜想 | 一句话新划界 | 至少三个猜想，各附一条预测 | 写成口号 |
| 交 AI 造 | 以猜想为前提展开 | 多份并行的小系统，附留痕 | 旧默认被放回 |
| 验收 | 防换皮，验排进，撞外部 | 四项检查记录与外部对撞记录 | 用漂亮做标准，陷进回音室 |
| 并排选择 | 与旧典范同表比较 | 比较表与落选记录 | 过早宣布胜利 |

## ◆ 拿量子力学的案例走一遍

为了让六步不那么抽象，把第 12 章的量子力学案例倒着套一遍工序卡。当年没有工序卡，这只是借它看每一步大致落在哪里。

识裂缝：炉子的光，光电效应，谱线，原子的稳定。每一条都是实验桌上的事实，写下来只需要说明条件、观察到什么、旧说法预测什么。

拢纠缠：四条各自试图被旧说法接纳，卡在同一个默认上，能量可以无限细分。再加上后来的散射与衍射，又卡在另一个默认上，波与粒子各归各类。纸中央有两个结，不止一个。

下猜想：「能量的交换是一份一份的」，一句话，指向第一个默认，并且立刻带着预测：光电效应里电子的最大能量与频率成线性关系。第二个默认，也有它的一句话：物质也有波的一面，波长由动量决定。

交 AI 造：这一步在当年对应的是几十年的人工展开，两套数学、概率诠释、一代又一代的计算。今天，如果有一个新的小猜想，这样的展开可以先让 AI 造出一个粗稿，几天，甚至几小时，而不是一代人；粗稿露出的矛盾，又会回头改写那句猜想。

验收：换名测试，对当年的情形，就是问「这是不是旧的力学加了几条硬规定」。早期的轨道规矩，恰好没通过这个测试，它既用经典力学，又禁止经典力学的后果；后来的整套语言，才算过了关。这个对照很有教益：一个系统可以暂时有用，却没有通过验收；知道它没通过，是继续往前走的理由，不是放弃的理由。

外部对撞：当年的对撞是实验桌上的事，比如后来看到的电子衍射，任何一套数学都造不出它。并排选择：当年有两套数学并排存在，后来发现等价；有几种对波函数的解读并排存在，至今没吵完。并排并不丢人，它是典范尚在发生的正常状态。

走完这一遍可以看到：工序卡不替任何人想出那一句话，但让「之前和之后该做什么」变得清楚。之前，是把裂缝记清、把纠缠拢紧；之后，是在来回中把展开做出来、把验收守住、把比较做公平，再让它撞一撞外面。

## ◆ 课堂练习清单

下面的练习，按从易到难排。每一项都可以在一到两节课里完成，适合中学高年级到研究生。老师可以挑着用。

**练习一：找例外。** 选一门学生熟悉的课程里的一个结论，要求每人找出三个「按说应该成立、实际不太成立」的例子，并写清条件。目的：把裂缝识别从抱怨变成事实陈述。

**练习二：补丁数一数。** 给一段简短的科学史材料，请学生数出其中有几处「为了让理论继续成立而加的说明」。每一处用一句话写出：它在保护哪个说法。目的：看见补丁。

**练习三：找默认。** 给学生五条互相无关的反常，请他们判断，哪几条背后可能是同一个默认，并把默认写成一句话。允许答案不唯一，但要说明理由。目的：区分主题归类与纠缠归类。

**练习四：三个猜想。** 针对练习三中的默认，要求每人写三个不同方向的替换，每个不超过二十个字，各附一条预测。小组之间互换，请对方只回答一个问题：它预测的东西，旧说法能不能也预测？目的：练习写小而尖的猜想。

**练习五：换名测试。** 给学生一段所谓「新理论」的描述（可由教师事先用 AI 生成，其中故意放进一个旧默认），请学生用换名的办法判断它是不是换皮。目的：练习验收。

**练习六：默认侦探。** 同样的材料，让学生逐段圈出「旧默认被悄悄用上」的位置。目的：训练细读。

**练习七：委托书。** 每组为一个猜想写一份给 AI 的委托书，其中必须包含：前提原文、被替换的默认、禁止放回的明令、留痕要求，再加一行：准备拿什么外部的东西去撞它。教师和另一组只看委托书，不看结果，评议是否漏了什么。目的：把人机分工写具体。

**练习八：并排表。** 给出三个互相竞争的说法（可取科学史中真实的例子，如日心说与地心说的几种变体），请学生做一张「差异 × 系统」的比较表，三种格子要区分清楚，并写出落选理由。目的：练习选择与留痕。

**练习九：地基日。** 每学期留几节课，什么也不用 AI：手算，手推，手画。目的：腿要练，看不懂典范里的东西，也就看不懂裂缝。

**练习十：失败展览。** 学期末，把被否掉的猜想、验收中发现的换皮、表里「靠补丁」的格子做成展览。只评记录是否清楚。目的：让失败可见、可学。

## ◆ 给评价的一点提示

这套练习的评价，不建议只看最后猜想「对不对」，那往往要很久以后才有答案，课堂等不起。更适合课堂评价的是下面这些：

裂缝是不是事实而非情绪；默认是不是具体而非笼统；猜想是否对准默认并带着预测；委托书是否禁止旧默认；验收记录是否真的做了换名、默认、裂缝、失败四项，并写明外部对撞的结果；比较表是否留了落选记录。

这些都是过程里可以被检查的东西。但请记住正文的提醒：看过程本身不是终点，只是让这一圈练习可以被教、被练、被改进。真正要带学生去的地方，仍然是「让典范发生」的那一层。

---

带走的话

**六步：识裂缝，拢纠缠，下猜想，交 AI 造，验收，并排选择。是圈，不是线；在人与 AI 的来回中展开，猜想会被展开改写。 最容易出错的两处：猜想写成口号，展开时旧默认被偷偷放回。 验收由记得「猜想哪里不一样」的那一部分来守，做换名、默认、裂缝、失败四项检查，再加外部对撞，防回音室。 多个猜想并行，多个系统并排，落选的也留着。 腿还是要练：地基日不用 AI。**

---

**AI 接管了典范内的一切，科研被推上更高一层：让典范发生。**

答案、过程、提问，都在同一层里，AI 跟着平移，成本跟着归零。

典范发生，是人—AI 复合体被逼上的新主业，也第一次需要成为一门学。

腿退出了运输，交通升级了。

德麦国际出版社 · Demai International Press
德麦国际专著第 369 号
定价 US$20.00

ISBN 979-8-90690-515-4