

现代科研指北

于淼

2017-07-04

Contents

1	前言	5
2	科研在搞什么鬼	7
2.1	科研老鸭汤-科学哲学沿革	7
2.2	科研职业化-问题为导向	9
2.3	科研精细化-体面的博士	10
2.4	科研与科普	10
3	科研现状概览	11
3.1	国内版	11
3.2	国际版	11
4	思维工具篇	13
4.1	科学思维	13
4.2	模型思维	13
4.3	统计思维	13
4.4	估算法	13
5	实验	15
5.1	实验设计原则	15
5.2	定性实验	15
5.3	定量实验	15
5.4	思想实验	15
6	数据处理	17
6.1	多重比较	17
6.2	多重检验	17
6.3	回归	17
6.4	预测	17
6.5	仿真	17
7	文献阅读	19
7.1	文献管理	19
7.2	文本挖掘	19
7.3	荟萃分析	19
8	学术生活	21
8.1	学术出版	22
8.2	学术会议	22
8.3	审稿	22
8.4	学术合作	22
8.5	课题组管理	22
8.6	学术声誉	22

8.7 学术道德／伦理	22
9 专题一：基于互联网数据的科研	23
10 专题二：现代科研兵刃谱	25
11 专题三：现代科研小红花	27

Chapter 1

前言

才疏学浅，不知何为真，仅通少错之法，故不敢言南，仅指北。或曰：现代科研挖坑／跳坑指南

Chapter 2

科研在搞什么鬼

2.1 科研老鸭汤-科学哲学沿革

做科研一般都不讨论哲学，太多形而上的东西，说也说不清道也道不明还无法证伪。但懂一点科学哲学还是很有必要的，不然很容易研究着研究着就会觉得自己做的东西是垃圾，是谋生的工具，虽然从某个角度看也没错，但科学哲学无疑是应对这种心态最好的老鸭汤。

2.1.1 古希腊

哲学是爱智慧这个梗就不多说了，扯古希腊也显得俗套，反正有了古希腊人才有了理性跟逻辑的提法。古希腊前面的历史可理解成经验性知识的发展，知识多了就要有规律总结出来，逻辑和理性可看作用来生成规律的知识。其实哲学就是认识世界的知识，泰勒斯有一套，毕达哥拉斯有一套，赫拉克里特有一套.....大家能自圆其说就来一套，对不对另说，不服就辩论，赢了就是真理，输了就是谬误。如果说逻辑与理性出自于这些街头巷尾的辩论我一点也不会奇怪，因为两套理论对比，总得有两方都认可的法则才有结果，理性或许就是这种普遍性知识的产物。

不过辩论有诡辩这一说的，苏格拉底看不下去了就说你们这些人都觉得自己对，但有可能是不对的，反正我自知我无知（这句是我认可最有智慧含量的句子，还有一句：天下没有免费的午餐）。老苏不怎么关注解释万物，有点回归个人或社会的意思。到了柏拉图直接就理想国了，世界形物均为理型的影子。再到了亚里士多德就不怎么废话了，直接取消理型世界的存在，认为万物有因，这个因就是所有问题的因，寻找到最终因，真理就明了了。这货还不知道这个看法后来发展成第一推动问题，宗教界觉得只有全能的主有这能耐，就把亚里士多德的理论吸引到宗教哲学里去了。

同时，我们现在所提到的科学源于日本，可理解为分类的知识，而最早对人类知识体系分类的就是亚里士多德。他还很神奇的将自己的目的论揉到这个分类里去了，所以这个体系很完整，能解释的东西很多，所以后来几百年大家就都用了这个体系。其实这时候科学知识更适合分到亚里士多德所谓的自然哲学这个科目里，这个科目特指自然现象的规律及探索方法。这里需要注明的是数学更多是工具，数学化不一定就代表科学，另一个需要注意的是逻辑学，这货的三段论十分精彩，以至于要不是后来哥德尔横空出世，任谁也动不了根基。

2.1.2 中世纪

中世纪黑暗吗？如果看天气应该跟现在差不多，但这个黑暗的印象大致源于天主教对知识的垄断，而知识也反过来服务宗教，而宗教理性在一定程度上促进了我们对世界的认识，所以盲目对立宗教跟科学没必要，很多前期知识都是不少富有宗教热情的理性人士总结的。只不过知识很多种，科学在那年代连个独立的名字都没有，所以你看，牛顿写本书叫《自然哲学的数学原理》，跟现代意义上的科学没啥关系，所以你管他信仰什么呢。这个时候，科学知识跟形而上学还是分不开，很多知识有严谨的数学形式但你无法证实，很多天文学知识就这德行，你去看看托勒密体系，圆环套圆环的也能解释现象，那哥白尼的日心说为什么就流行了呢？因为是真理？因为结构简单？还是因为他用了别人看不懂的语言写出来的？或者说反对者死绝了新理论就流行了？总之，没有实证的理论的流行不会是你想的那么简单，但一般来说，人们都喜欢简单且解释面广的理论，宗教也这样，毕竟美的东西都是上帝赐予的。在科学这个提法之前，用一套知识来解释世界是各代学者所向往的，才不管验证什么的，理性重于事实。

2.1.3 1500 年以后

时间点不太好找，但历史的发展是伴随知识的增长的。大航海时代为人类的知识提供了一个海量来源，文艺复兴带来了人性的解放，宗教改革让生活走出了政教合一，总之，经验开始比逻辑更为人接受。最开始是欧洲大陆的理性说与英国的经验论的争执，争论核心在知识的构建是从理性出发还是从经验出发，这两种观点打架几百年，到了 19 世纪末大家都不争了。因为实证主义一统江湖认为从经验中提取逻辑，然后再证实就 OK 了。这时候科学哲学才独立出来，而观察式的经验也开始让位于实验式的事实，人们不满足于被动接受知识，开始主动去寻找真相。

2.1.4 逻辑实证主义

当人自己把握了主动权，原有的常识知识就要被逻辑重新检验，而无法检验的就划到形而上学这一类里留给做宗教神学的人去讨论。换言之，从柏拉图开始的将现实世界与理想世界的区分被打破了，原来的哲学家都醉心于构建理想世界而不关心现实生活，而逻辑实证主义则要求通过生活的事实来寻找真相。换句话说，经验事实及逻辑推理被结合用在真理的探索上了。而经验事实的崛起则伴随着归纳法的崛起，事实成为知识的唯一来源，科学开始渗入并改造哲学方法论，这一转变真正让科学有了真理探寻的光环，一举扫清神秘主义与宗教束缚，直到今天还在深刻的影响着每一个科研工作者。

2.1.5 否证主义

但不久大家发现不对头，因为归纳法不如演绎法严格，得到的结论有局限性，不够严谨。这时候波普就说了，演绎法靠谱！大家都提假说，然后验证它，出现反例就把假说否了，不能否证就不科学，这就是证伪。一时间大家都接受了，神马佛洛伊德，历史唯物主义都因为自洽但不能证伪给端出科学圈了。

不久又有人感觉不对了，一方面演绎法很难产生新知识，另一方面貌似假说是无穷无尽了。证实比较费事，证伪容易但很多理论就垮了。为了调和这个矛盾，否证主义给出的答案是演绎法虽不能产生新知识，但假说的产生不是无缘无故的，而知识的进步应该通过大胆猜想的确证与谨慎猜想的否证来完成，一个推翻的理论必然联系着新理论的提出，这时不断发展的，而科学的任务就是处理进步问题而非回答真理问题。形而上学也并完全被排斥了，因为假说的提出有时就是没有事实证据的。进一步讲，波普尔将世界分成世界 1，也就是物理世界，世界 2，也就是精神世界，然后又分了个世界 3，也就是客观知识世界。这种三分法其实是将柏拉图的理型世界进化了，同时也留下了世界 2 的个人空间。每个世界都在进化，这就是科学发展的轨迹。一口吃不成胖子，我们就去试错吧！猜想与批判这一否证主义的核心思想也是当下科研中比较闪光与巧妙的实验设计动机来源。

2.1.6 历史主义

前面那些理论的提出者大都数理化出身，推理证明构建系统很在行，但没案例不成啊，得解释得了现象啊。其中一些人翻了翻了史书，发现很多发现不是通过证伪得到认可的，也不是建立在大量归纳的基础上，而是具有“历史性”。也就是逻辑不怎么灵光，然后他们就说咱以史为鉴吧！拉卡托斯就搞出了个硬核软核的理论，大意说一个理论是有生命力的，硬核部分无须质疑，有保护带，一时半会死不了。需要缝缝补补的是外围软核，什么时候硬核也不行了，就退出历史舞台了。这个解释保全了科学理论体系，也就是堵了民科的路，要知道民科最喜欢证伪，一个错误就否了整体，现在拉卡托斯说不成，得慢慢来，有历史的。

不久又有人感觉不对了，我怎么知道现在的硬核到底对不对？拉卡托斯这时就呵呵了，交给历史评价吧！库恩在这个背景下提出了范式，他本身有较强的历史功底，手头案例多，所以有了科学共同体这个说法。大意就是一个时代的真理主流说了算，这伙人挂了而接任的更多采取了另一种解释现象更多的理论，那这个理论就上位了，就革命完成了。前面那个时期比较压抑就叫前科学，后面上位了就是常规科学。之后又有新现象解释不了了就有了危机，这时候新理论又出现了，再搞一次革命就 OK 了。范式是来区别前科学与常规科学的，范式通常是一套当前时代科学共同体所使用的理论体系，而这个理论体系要比之前的更能解释更多的问题也更严格。这理论比拉卡托斯那一套通俗易懂，那年代搞政治的一看有革命二字纷纷表示深有体会，大力推广之，所以范式着实火了好一段时间。

库恩的范式革命是格式塔式的转换，历史上一共也没发生几次，真正有益的是他对范式定义时要求要有自称科学的学科要有自己的理论体系与假设且对现实世界产生作用，这个理论自身并不要求科学家的态度是客观的，但范式自身要是客观的。这时候，大家都不愿搭理真理性这茬了，因为都清楚对错问题是历史性的。同时范式也把形而上学彻底请回到科学体系中了并认为对科学的发展是有益的，要知道波普尔虽然不拒斥形而上学但本质还是批判形而上学的。所以历史主义的强调使得真理相对化。

2.1.7 无政府主义

事实上你沿着这个思路走下去发现貌似科学发展跟三国演义差不多，不在于对不对而在于认可的多不多，有没有跟你闹革命的。当然因为实证主义的余威，理性与逻辑在科学研究中是绕不开的。这时候来了个更霸气的费耶阿本德，一拍桌子，科学跟别的东西没啥区别，不能特殊对待。后来流传到世上的就是那句 anything goes，很多人认为这货终结了科学哲学的发展。从 20 世纪初到六七十年代这个学科就完蛋了，这就是科学哲学的学科危机。

2.1.8 实用主义

逻辑确实打不过历史，原来那些搞科学哲学研究的还没死就没饭碗了，生存是硬道理。他们发挥了科学共同体的作用，把费耶阿本德斥为异类、后现代。但他说的话又绕不过去，这时候蒯因跳出来讲科学哲学还要发展，不能 anything goes，科学不科学总得有个标准。美国人想来想去想到了有用两个字，然后大家纷纷鼓掌。理性，历史都打不过生存这个命题。有用是硬道理，有用解释一切，然后就没有然后了。

2.1.9 其他

除此之外，由于逻辑讲求语义明确而严格，但要是日常交流用一堆符号估计谁也受不了，所以科学哲学也在语义学方面继续发展。英国人的经验论也促进了新实验主义与主观贝叶斯学派的发展，慢慢地科学哲学也开始接受一些非实在论的观点，而科学实在论是穿插在上述命题中的。

科学哲学从实证主义发展到今天，被各种新命题与发现折腾的够呛，从里面提一个片段就可以看到很多，科学是什么？它跟哲学啥关系？又对哲学发展有什么样的影响？总之，我们没有停下探索真理的脚步，答案在哪里也毫无头绪，只要不满足于现状，知识就存在进步的可能，同时须知人生苦短，自知无知是很重要的。

2.2 科研职业化-问题为导向

喝完老鸭汤我们还是回归现实吧，现代科研隶属于现代政治经济系统，满足社会的需求是其存在的基础，至于是否满足个人兴趣爱好与远大理想，可认为是副产品。当然，这是从社会层面说，具体到个人千差万别。

首先，我们要了解现代社会运行的基本模式，其中陌生人大尺度分工协作是现代社会最突出的特色。社会，简单说就是一群人而不是一个人生存的行为与知识模式集合。相比宗族或家庭为单位的原始聚居，古代与近代社会的发展不断突破着人们行为与知识范围的地理与血缘限制。

在原始聚居条件下，人们终生活动范围有限，语言隔阂等也限制了信息交流，好的生存模式很难传递到下一代或更远的地方，短暂的寿命基本都用在维持生存繁衍上了。当然，对原始部落的研究发现生活在其中的人并不比焦虑的现代人的快乐感受更少，但生活的自由度其实很有限（从另一方面讲，如果完全意识不到当今生活自由度可以改变其实也是一种内在幸福感，拥有更大自由度的人并不能完全体会到）。这种狩猎采集的原始聚居其实并不太需要共同的社会行为规则，但后来人们驯化了农作物与牲畜（其实很难讲谁驯化了谁，作物与牲畜也可能通过驯化更好的传播了基因），进而从流动走向了定居。

定居后的社会出现了更细致的分工，例如一个村落需要祭祀、防卫、生产、医疗等部门维持生存结构，这种分工有着自己的生命力，一旦产生会让整体受益，同时也会让这种结构加强。同样的，这种分工模式并不惟一，但如果两个定居的社会共同体产生利益矛盾，最后剩下的总是一种更有利群体生存的模式，这个模式下的规则并无道德可言，或者说这就是社会道德的起源。这同时也是一个路径依赖的过程，总会带有一些副产品，很多时候我们就是通过副产品来回溯过去。如同对进化过程的研究一致，使用幸存者就是最好的或最合理的逻辑是不恰当的，我们需要通过回溯来发现一些制度历史上的合理性与偶然性，逻辑自洽并不代表历史真相，这点对科研认识也是很重要的。然而，这个阶段的社会政治经济体制依然很大程度被自然条件所控制，多数规则要么偏向农业社会，要么偏向海洋经济，人类的视野逐渐开阔，但基于血缘与地域的多样化依然可以保留，直到更追求效率的技术与体制规则进一步交互作用，孕育出近代工业社会。

近代工业社会将分工与效率推向了极致，影响的范围从多个国家推广到了全球。伴随而来的就是一套基本抛弃血缘关系与多样性的陌生人交流法则，地理限制被信息技术与交通技术打破，所有国家都会遵循同样的工业标准，语言也尽可能一致，法律也会去遵循共通的法律。科学研究在这个过程中起了很重要的作用，而工业化也不断向科研提出需求，此时科学研究从精英们的兴趣爱好变成了巨大的财富来源，每一次技术革新都服务了社会，而几乎所有的社会经济体都会拿出资金支持科研。务实一点的国家或企业会对工程学优先发展，而对自然科学的支持则颇有情怀意味，毕竟一旦经济下滑，最先拿不到钱的都是基础科研等见效慢的学科。这种社会整体的功利主义自产生之时就展示了巨大的生命力，甚至不断影响了社会中个体的决策行为。

时至今日，现代社会基本延续了工业社会对分工与效率的追求，但维持文化多样性与个体-社会相互关系的思考不断涌现。现代社会塑造了个体认知，个体认知却反过来反思现代社会的诸多问题例如民族主义的崛起、环境保护、气候变化、社会隔离与歧视、机会公平、人口老龄化、战争暴力、谣言传播、经济危机、金融危机、人工智能等。这些问题的根源有相当比例是社会政治经济体制的构建过程出现了漏洞，而今天的科技发展把一些问题放大了，或者说这个系统需要打补丁了。毋庸置疑，科研对于现实问题的解决是一个靠谱的选择，其他选择例如宗教、回归原始生活更多的是一种消极的保守策略，选择那些方法并不会真的解决问题。这样如果给科研立块大牌坊，我想最好的题词就是从方法论层面解决社会问题。换言之，科研总是面向问题解决问题的一个社会分工，是一个职业，既不神圣也不低俗，从事这个职业的人总在用科学方法论解决实际问题，有时候也是揭示问题或为问题找一个解释。这个需求是根源，也就是说如果你科研自认为做的不错但跟现实脱节，那么即使留在象牙塔，也会面临自我认同与社会认同不协调的困境，需要你有额外的资源平衡。

放到经济视角下，这个职业也是有温饱小康问题的，也是一个利益集团，需要人代表到国会或人大去抢财政的分配，还要跟不同学科去抢所有的科研分配，充满了复杂的博弈过程，原来是陌生人之间，以后可能会发展到人跟机器或规则之间。这个职业有光环，但退却光环都是一个个为生计奔波劳碌的现代人，当然，不同人有着不同的生计标准。

2.3 科研精细化-体面的博士

精细化促进了分工效率限制了个人视野螺丝钉体面的博士

2.4 科研与科普

Chapter 3

科研现状概览

3.1 国内版

3.2 国际版

Chapter 4

思维工具篇

4.1 科学思维

4.2 模型思维

4.3 统计思维

4.4 估算法

4.4.1 费米估计

Chapter 5

实验

5.1 实验设计原则

5.2 定性实验

5.3 定量实验

5.4 思想实验

Chapter 6

数据处理

6.1 多重比较

6.2 多重检验

6.3 回归

6.4 预测

6.5 仿真

Chapter 7

文献阅读

7.1 文献管理

7.1.1 Zotero

7.1.2 Mendeley

7.1.3 EndNote

7.2 文本挖掘

7.2.1 关键词

7.2.2 作者

7.2.3 时空分布

7.2.4 影响力

7.3 荟萃分析

Chapter 8

学术生活

8.1 学术出版

8.1.1 期刊论文

8.1.2 会议摘要

8.1.3 专著

8.1.4 专利

8.1.5 软件

8.2 学术会议

8.2.1 口头报告

8.2.2 海报报告

8.2.3 听报告

8.3 审稿

8.4 学术合作

8.4.1 数据共享

8.4.2 社交网络

8.5 课题组管理

8.6 学术声誉

8.7 学术道德／伦理

Chapter 9

专题一：基于互联网数据的科研

Chapter 10

专题二：现代科研兵器谱

Chapter 11

专题三：现代科研小红花

Bibliography