

第 5 章

数据管理

除了上帝，任何人都必须用数据来说话。

——爱德华兹·戴明(美国著名质量管理专家)

数据可以说伴随人们的一生。从你呱呱坠地的那一刻，就有一堆数据被记录下来——妊娠周数、出生时刻、体重、身长、体温等；在你成长的过程中，总是被大量数据所环绕——年龄、住址、学习成绩、工作经历、婚姻状况等；就算你离开人世之后，还是摆脱不了数据的纠缠——死亡时间、死亡原因、生前工作单位、生前声誉、生前贡献……如果没有这些数据，你无法“客观”地认识自己和评价别人。

人类的文明与进步，从某种意义上讲是通过数据收集、处理和总结而达成的。在史前时代，人类的祖先在没有发明记事的媒体工具时，其实已经开始使用数据了——从父辈和周围人的口耳相传中，知道了哪些环境可以居住，哪些动植物可以食用，哪些情况暗藏危险；有了文字之后，我们通过记录下来的历史数据来获取更多的经验教训——“秀才不出门，全知天下事”“以史为鉴，可以知兴替”；到了近代自然科学萌芽之后，数据的重要性逐渐提升到了一个前所未有的高度——不论是在哪个领域，科学家们很重要的一项工作就是做实验采集数据，因为科学发明需要通过这些数据来推导或证实。

那么，什么是数据呢？传统意义上的数据是指“有根据的数字”，例如我们常说的实验数据、统计数据就是以数字的形式表现出来的，这些其实只是狭义上的数据。随着技术的进步，数据的内涵得以扩大，可以指代许多“结构化的信息和情报”，例如我们经常提到的一个词——数据库，其实就是指符合一定格式的信息的汇总。数据库里的数据可以是某个机构所有成员的基本情况，包括姓名、年龄、通信方式、学历以及履历等(文字信息)，这些已经超出数字的范畴。进入信息时代之后，数据的含义更加宽泛，包括任意形式的信息，例如互联网上的全部内容、档案资料、设计图纸、病例、影像资料等。可以说数字、文本、音频、视频、图形等各种形式的记录组成了广义的数据。

5.1 决策的依据

我们经常痛苦于“彷徨”和“纠结”，这两者的区别在于：彷徨是因为无路可走，纠结是因为有太多路可走(图 5.1)。的确，生活中总是需要我们做出取舍，进行选择。而决策的失误往往直接导致失败，因为选错了方向走错了路，加倍努力也很难取得预想的效果。有句话说得好：“朝相反的方向奔跑，停下来就是前进。”

那么做决策的依据应该是什么呢？直觉，占卜，还是古圣先贤的哲言？我们常说“事实胜于雄辩”，这个“事实”往往指的就是数据，所以说数据才应该成为我们做决策的基本依据。下面就先从古今中外的军国大事的决策中看看数据的重要作用。



图 5.1 纠结的痛苦

5.1.1 多算胜少算

《孙子兵法》十三篇为后世兵家所推崇，被誉为“兵学圣典”，这部著作总结了春秋以前战争胜负的许多经验，是兵家必读之书，是世界三大兵书之一（另外两部是克劳塞维茨的《战争论》和宫本武藏的《五轮书》）。

孙武，生于公元前 545 年，是中国古代兵法的集大成者。他在《孙子兵法》中就提到了以数据作为决策依据的思想：“夫未战而庙算胜者，得算多也；未战而庙算不胜者，得算少也。多算胜少算，而况于无算乎！吾以此观之，胜负见矣。”而且，他还在书中对使用数据的具体方法也做了论述：“一曰度，二曰量，三曰数，四曰称，五曰胜。”其中，“度”是指国土的大小，“量”是指粮草资源的多少，“数”是指军队的数量，“称”是指双方实力的对比。孙武的意思是，战争的胜负可以通过这 4 个因素进行估计，而这 4 个因素本质上都是数据，作战双方都不断刺探对方的实力，试图获得准确的数据，同时也不断释放数据“烟幕”，以迷惑对方，掩盖自己的实力。

通过释放数据“烟幕”，以计诈敌，中国古代有不少著名的战例。战国时期，魏将庞涓率领十万大军进攻韩国，韩国不敌，向齐国求救。驰援韩国的齐军采用了军师孙臧的“减灶计”——开始的时候设 10 万个灶，其后设 5 万个灶，最后减到了 3 万个灶。庞涓见到齐军所留的灶迹不断减少，就判定齐军出现了大量掉队、减员的现象，因此撤下步兵，率领骑兵分队加速追击，结果在马陵中了孙臧的埋伏，兵败身亡。东汉时期，西北边陲的羌族入侵武都郡，名将虞诩率兵前往救援。因为一开始兵力不足，虞诩需要避免正面决战，等待援军集结，就在行军途中使用“增灶计”——让官兵每人各作两个锅灶，以后每日增加一倍。羌兵见此，认为郡兵已来接应，不敢纠缠，因此争取到了行军时间。到达郡府后，虞诩

集合全部军队,命令他们次日先从东门出城,再从北门入城,然后改换服装,往返多次。羌人不知城中有多少汉军,惊恐不安,最终被虞诩击败。

每战必胜的原因

林彪从红军带兵时起,身上就有个小本子,上面记载着每次战斗缴获和歼敌数量。这种几十年如一日收集数据的习惯是一般人是难以想象的。1948年辽沈战役开始后,林彪每天深夜都在东北野战军前线指挥所里听取军情汇报,由值班参谋读出下属各个纵队、师、团用电台报告的当日战况和缴获情况。林彪的要求很细,俘虏要分清军官和士兵,缴获的枪支要统计出机枪、长枪、短枪,击毁和缴获尚能使用的汽车也要分出大小和类别。经过一天紧张的战斗指挥工作后,作战室里清醒的估计只剩林彪和读电报的参谋了。

东北野战军攻克锦州后,与敌精锐廖耀湘军团二十余万在辽西相遇,一时间形成了混战。战局瞬息万变,谁胜谁负实难预料。一天深夜,值班参谋正在读着下面某师上报下属部队的战报,说他们的部队碰到了一个难度不大的遭遇战,歼敌部分,其余逃走。与其他之前所读的战报看上去并无明显异样,值班参谋就这样读着读着,林彪突然叫了一声“停!”他眼里闪光,问道:“刚才念的在胡家窝棚那个战斗的缴获,你们听到了吗?”在场的人睡意已深,似乎没有在意那些枯燥的数字,无人回答。林彪扫视一周,又接连问了三句:“为什么那里缴获的短枪与长枪的比例比其他战斗略高?为什么那里缴获和击毁的小车与大车的比例比其他战斗略高?为什么在那里俘虏和击毙的军官与士兵的比例比其他战斗略高?”其他人还没有来得及思索,林彪已经等不及了,他大步走向挂满军用地图的墙壁,指着地图上的那个点说:“我猜想,不,我断定!敌人的指挥所就在这里!”随后林彪口授命令,各部队坚决追击并打掉从胡家窝棚逃走的那部分敌人。廖耀湘万万没有想到自己精心隐蔽的司令部那么快就被发现和消灭了……

在中外战争史上,大规模、系统化运用数据的经典战役莫过于谢尔曼(William T. Sherman)将军在南北战争期间领导的“向大海进军”(March to the Sea)。1864年8月,谢尔曼率6万大军挺近南方的中心城市亚特兰大之后,采取了后世史学家认为整个南北战争中“最为大胆、最为关键的一次行动”:挥师东进,横穿佐治亚州,一路打到美国东海岸线。

兵马未动,粮草先行。在还没有飞机的时代,军队的行进路线安排必须充分考虑后勤补给。北宋年间的中国科学家沈括分析过,对一支10万人的军队而言,随军辎重就要占去1/3的兵力,最后真正能上阵打仗的士兵其实不足7万,如果一个士兵需要3个民夫供应,那就要征召30万民夫和额外的管理人员,但就是这样庞大的后勤规模,也只能支持行

每个民夫可以背6斗米,一个士兵可以自带5天的干粮。1个民夫供应1个士兵,两人同吃同行,可以维持18天,如果计回程的话,只能进军9天;2个民夫供应1个士兵的话,单程可以维持26天,若计回程,只能进军13天;3个民夫供应1个士兵,且每吃完1袋粮食就遣返1名民夫,单程最多进军31天,若计回程,只能进军16天。

军 31 天。沈括因此得出结论：凡行军作战，应该争取从当地获取粮草和补给，这是最为紧迫的事情，否则不仅耗费大，而且走不远，跑不快，作战能力极为有限！但是如何在当地获取补给，沈括却只字未提。

谢尔曼一方面从国家的人口普查部门获得了南方的人口、资源等方面的宏观数据，另一方面在亚特兰大搜寻了一切关于佐治亚州的地图、财税明细和各种统计表格。然后，他计划主动切断后方补给，以统计数据为“航标”，根据农场、牲畜、集市、车站等重要资源的分布，通过后勤参谋的精心计算，确定最佳的行军路线和在各地停留的时间。如图 5.2 所示，5 路大军沿着规划好的路线向东部沿海重镇萨凡纳突进，部队不仅在当地完成了补给，摧毁了敌方重要的基础设施，而且遭遇了最少的正面阻击。攻占萨凡纳之后，谢尔曼向联邦军总司令格兰特报告说，部队沿途消耗骡子 15 000 头，牛 10 000 余头，各种粮食都来自当地，和战争开始前的预测相差无几。经此一战，南方的战略资源被掠夺和破坏殆尽，几个月后南北战争结束。

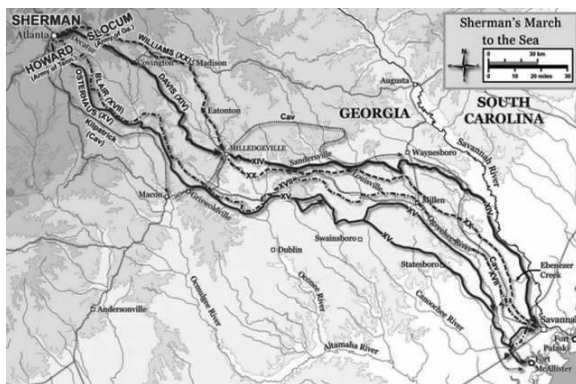


图 5.2 “向大海进军”的行军路线图

谢尔曼在战后给美国普查办公室主任约瑟夫·肯尼迪发去了感谢信：“此战证明，您给我提供的各种统计表格和数据价值巨大，没有它们，我不可能完成任务……”在回忆录中，他总结说：“历史上没有任何一次行军远征曾经建立在像这次一样完善和肯定的数据之上。”

在南北战争正式爆发的前一年，也就是 1860 年，这是美国的大选年。这一年，一位伟大的美国平民——亚伯拉罕·林肯 (Abraham Lincoln)——经历种种逆袭，最终入主白宫，登上了美国政治的中心舞台。而这一年也是美国第 8 次人口普查年，所以林肯总统可以直接从普查办公室主任肯尼迪那里获取大量的相关数据，来预测这次内战的最终胜负。

围绕着《孙子兵法》所论述的“度、量、数、称、胜”5 个角度，我们看一看当时美国南北

方的潜在军事力量对比：国土面积上，美国当时有 33 个州，其中 22 个州站在联邦一方（北方）；粮草资源上，虽然南方最大的优势就是其棉花出口，占据全国出口额的 60%，堪称国民经济的命脉，但是棉花毕竟不是粮食，北方的农业生产足以保证战时粮草的供应；军队数量上，全国 18~45 岁的青壮年劳力约有 69% 集中在自由州，即使中间的摇摆州全部倒戈，南方的力量也不过 31%；双方经济实力的对比上，联邦更是占有绝对优势，北方集中了全国 2/3 的铁路，90% 的工业常量。这些数据起到了“定心丸”的作用，林肯深信，如果打持久战，胜利最终将属于北方。

可以看出，无论是林肯的信心还是谢尔曼的传奇，都源于他们手头的数据是大量的、系统的、成片的，背后有专业人员给予支持和维护的。这种“有数可用”，得益于美国建国之后就开始的、长期的、周期性的努力和强大的制度保障。有无这种制度化的数据收集体系，才是近代战争中美国和其他国家在数据使用方面拉开差距的根本原因。

5.1.2 数据的文化

美国是个年轻的国家，其开国至今不过 200 多年，但数据在其政治活动和社会生活中的历史却几乎和其建国史相生相伴。美国的建国者一开始就把人口普查写进了宪法。他们认为，国家权力应该在人口之间平均分配，这个平均，必须用数据来说话。

当时的政论家、教育家诺亚·韦伯斯特(Noah Webster)继而指出，在所有的事实当中，用数据描述的事实是最准确、最锐利、最有说服力的。因此，描述一事实，增强客观性、减少主观性的最好方法，就是尽可能地使用数据。

美国人对数据依赖到何种地步，从 1787 年在费城召开的制宪会议中的一个片段——讨论“一个黑奴应该拥有多少权利和义务”——就可以看得出来：

既然人口的多少一定程度上决定了权力的大小，那南方拥有庞大的黑奴群体，是否也应该计入总数？一开始，大家都认为，奴隶本来就不拥有政治权力，因此不应该计入总数，但在后续的讨论中，人口的多少不仅成为分权的依据，还和纳税的义务挂上了钩，即人口多的州，国会占的席位多，也要缴纳更多的税收。南方则主张，黑奴既然不享受政治权利，也不应该承担义务，但北方又认为这样南方占了便宜。争论又起，最后的结论是，每个黑奴按 3/5 个白人（自由人）的标准纳入南方人口的总数，这个总数才是南方权力分配和纳税的依据。

每个黑奴等同于 3/5 个白人，这一规则被写进了宪法，成为种族不平等的历史明证。宪法颁布之后，曾引起很多追问，例如：为什么是 3/5，而不是 1/2 或者 2/3？当时主导辩论的汉密尔顿也说不清楚，他后来坦承：这是一个瑕疵，但当时必须找出一个数字，这个数字可能不完美，但比没有强。这就是美国人对数据的执着，哪怕是歧视，也要用数据来

衡量。

随着人口普查的作用从政治领域不断扩张,首先蔓延到了政策制定领域,然后是社会生活领域。人口普查也转化为向社会寻找“真正事实”的统计活动,通过收集充足的数据,国家可以掌握整个社会出生率、性别、年龄、婚姻状况、健康、职业、寿命等方方面面的情况。这就逐渐形成了一种数据文化:一方面提高识字率,减少文盲;另一方面要推广数学教育,减少“数盲”,以提高公民的思辨能力,使其学会独立思考。

美国的国父们都推崇数据文化,华盛顿、杰斐逊和富兰克林就是其中的突出代表。1788年,华盛顿曾经这样描述数学教育:“从某种程度上说,文明生活的方方面面都不可缺少数字的科学,对数学真理的追踪可以训练推理的方法和正确性,这是一项有益的活动,尤其适合理性的人类。”杰斐逊则建议,所有的小学除了教授阅读、写作外,还应该开设数学课。他认为:“就像身体的其他组织一样,大脑的功能也可以通过练习而改善、加强。因此,基于数学的推理和演绎,是人类了解深奥法则的有益准备。”到了1802年,数学已经正式成为哈佛大学入学考试内容。

在这样一批建国者的推动下,数学教育很快在这个新生国家普及,并影响到了美国的货币体系改革(见2.1.1节中“美国的货币体系改革”的论述)和测量单位的统一。这些工作对后世美利坚民族数据意识的形成、数据文化的建立,也产生了深远的影响。

19世纪30年代,英国哲学家托马斯·汉密尔顿(Thomas Hamilton)来到美国游历,并把他的亲身见闻写成了一本书《美国人及其作风》。他发现,美国人已经习惯于通过数据来做决策。例如,美国人会根据他人的财务状况对其进行分门别类,“我已经被清楚地告知,我的熟人当中谁有良好的名声和信誉以及他们每年的开支。”他最后在书中得出结论说:“我认为,在这群不断猜测、估算、预期和计算的美国人当中,算术就像是一种与生俱来的本能。”

类似的观察还有很多。1825年,费城的一名医生统计了7077名新生儿的体重,并制作了一张重量分布表,发放给新生儿的母亲,以方便她们对比掌握自己孩子的情况。他还监测了孕妇在280天孕期中每天增长的体重,并发放给孕妇作为其每天饮食标准以及体重增长的参考。

今天,现代化的医院一般都秉承了这种数据传统,从体检、诊断到治疗,几乎所有的医疗环节都以数据为支撑。例如,孩子一出生就要开始接受体检,身高、体重、头围是3个基本的检查指标,美国医院除了提供各项指标的大小,还会提供该项指标的百分位。

在美国做手术,术前病人或家属会被告知手术的风险,例如0.03%的死亡率、0.1%的感染率以及各种并发症的可能性。这些百分比的得出都建立在长期收集数据的基础上。2013年,美国外科医师协会(ACS)利用信息技术推陈出新,收集了2009—2012年全国393所医院、140多万病人的数据,在这个基础上开发了一个手术风险计算器(ACS/

华盛顿的第一份工作是弗吉尼亚州的土地测量员,他深知数据对于认识客观世界的重要性,在第一次人口普查期间,他甚至亲力亲为,组织了美国的第一次农业调查;杰斐逊也曾做过土地测量员,除了是一位政治家,他还研究密码学、测量学和考古学;富兰克林则是一位政治家、外交家和科学家,年轻时曾沉迷于研究雷电,后来发明了避雷针。

百分位(percentile)是个体指标相对于全体的平均水平的参照。例如,体重38%的百分位,意味着这个个体的体重超过了38%的同类个体。

NSQIP Surgical Risk Calculator)。该计算器能针对病人的情况,计算 1557 种手术的风险及各种并发症的可能性,为医生和病人提供手术前的决策参考和准备。

反观我国,数据意识淡薄由来已久,甚至可以称之为国民性的一部分。从古至今的大量典籍里面,我们都能发现各种非常模糊、夸张的描述,比如各种正史文字中的“千余轻骑”“几十万大军”“向北百余里”“身高丈余”,仔细想想这些已经不是“差之毫厘”了,怎么都能算得上“谬以千里”。

涂子沛在《大数据:正在到来的数据革命》一书中提及,在国外留学期间,通过工作和生活中的对比,他感觉到了中国人缺乏“用数据说话”的素养。中国的语言表达方式中“重定性、轻定量”的特点非常明显,口语中常常使用“大概”“差不多”“少许”“若干”“一些”等等高度模糊的词语。比如中国菜的烹调方法,就会令美国教授抓耳挠腮、不知所措,其中关于“盐少许”“酒若干”“醋一勺”的提法,完全是跟着感觉走,让初学者无从下手。

中国近现代著名的思想家胡适就对一些人“凡事差不多、凡事只讲大致如此”的习惯和作风深感忧虑。1919年,他写下了著名的《差不多先生传》,活灵活现地描画了当时国人不肯认真、缺乏逻辑、甘于糊涂的庸碌形象:

差不多先生传

你知道中国最有名的人是谁?

提起此人,人人皆晓,处处闻名。他姓差,名不多,是各省各县各村人氏。你一定见过他,一定听过别人谈起他。差不多先生的名字天天挂在大家的口头,因为他是中国全国人的代表。

差不多先生的相貌和你和我都差不多。他有一双眼睛,但看的不很清楚;有两只耳朵,但听得不很分明;有鼻子和嘴,但他对于气味和口味都不很讲究。他的脑子也不小,但他的记性却不很精明,他的思想也不很细密。

他常说:“凡事只要差不多,就好了。何必太精明呢?”

他小的时候,他妈叫他去买红糖,他买了白糖回来。他妈骂他,他摇摇头说:“红糖白糖不是差不多吗?”

他在学堂的时候,先生问他:“直隶省的西边是哪一省?”他说是陕西。先生说:“错了。是山西,不是陕西。”他说:“陕西同山西,不是差不多吗?”

后来他在一个钱铺里做伙计;他也会写,也会算,只是总不会精细。十字常常写成千字,千字常常写成十字。掌柜的生气了,常常骂他。他只是笑嘻嘻地赔礼道:“千字比十字只多一小撇,不是差不多吗?”

有一天,他为了一件要紧的事,要搭火车到上海去。他从从容容地走到火车站,迟了两分钟,火车已开走了。他白瞪着眼,望着远远的火车上的煤烟,摇摇头道:“只好明天再走了,今天走同明天走,也还差不多。可是火车公司未免太认

涂子沛,江西吉安人,著名信息管理专家、科技作家,《大数据》《数据之巅》作者。毕业于卡内基·梅隆大学,系微软认证高级程序员,曾居美国硅谷。

胡适的这篇传记体寓言对过去中国人数据意识淡薄、不肯认真、拒绝精准的庸碌形象进行了讽刺。全文原载于1919年出版的《新生活》杂志第二期。

真了。八点三十分开,同八点三十二分开,不是差不多吗?”他一面说,一面慢慢地走回家,心里总不明白为什么火车不肯等他两分钟。

有一天,他忽然得了急病,赶快叫家人去请东街的汪医生。那家人急急忙忙地跑去,一时寻不着东街的汪大夫,却把西街牛医王大夫请来了。差不多先生病在床上,知道寻错了人;但病急了,身上痛苦,心里焦急,等不得了,心里想道:“好在王大夫同汪大夫也差不多,让他试试看罢。”于是这位牛医王大夫走近床前,用医牛的法子给差不多先生治病。不上一点钟,差不多先生就一命呜呼了。差不多先生差不多要死的时候,一口气断断续续地说道:“活人同死人也差……差……差不多,……凡事只要……差……差……不多……就……好了,……何……何……必……太……太认真呢?”他说完了这句话,方才绝气了。

他死后,大家都称赞差不多先生样样事情看得破,想得通;大家都说他一生不肯认真,不肯算账,不肯计较,真是一位有德行的人。于是大家给他取个死后的法号,叫他做圆通大师。

他的名誉越传越远,越久越大。无数无数的人都学他的榜样。于是人人都成了一个差不多先生——然而中国从此就成为一个懒人国了。

黄仁宇,美籍华人,密歇根大学历史学博士,以历史学家、明史专家、大历史观的倡导者而为世人所知。著有《万历十五年》《中国大历史》等畅销书。

瓷器是彻底的人造物,它和金属、玻璃(包括水晶)等材料不同,在自然界是找不到的,完全是人类活动的结果和文明的标志。

5.2 科学的方法

著名的华人历史学家黄仁宇曾经在《中国大历史》等著作中对古代中国文化的一些缺陷做了剖析。他认为:在中国传统的学问——理学和道学当中,一直都分不清伦理之“理”与物理之“理”的区别。这两个“理”混沌不分的结果,是中国人倾向于粗略的主观性,排斥精确的客观定量,从而养成了重形象、重概括、轻逻辑、轻数据的文化习惯。这种文化习惯,使中国人长期沉浸在含蓄、模糊的审美意识中,凡事只能在美术化的角度来印证,满足于基于相似的“模糊联想”,止步于用逻辑来分析,用数据来证明,最终将表象上的相似当作本质上的相同。

归根结底,传统中国人对数据的漠视缘于一种文化上的欠缺:随意、盲目、不求甚解、理性不足。从某种意义上讲,正是因为这种文化上的问题,近代科学最终在西方国家产生,中国近代的坎坷命运就此铸成。

5.2.1 站在前人的肩上

中国的瓷器是一个伟大的发明,它对世界的政治文化和人类的日常生活都产生了巨大的影响。尤其是在宋代和明代,中国瓷器在世界上每到一处,就会掀起一股奢侈品购置

的热潮,并改变当地人的生活方式、当地的文化,甚至改变当地的制造业。世界上还没有第二种商品能在几百年的时间里长期做到这一点。

葡萄牙国王曾经用 260 件中国瓷器装饰了桑托斯宫的天顶,这表明在当时欧洲最富有的皇室眼里,瓷器是美和财富的象征。大航海时代,西班牙人从美洲带走了 16 000t(约五亿两)白银,这些白银的 1/3 都用来购买了中国的货物,主要是瓷器和茶叶。这让中国赚足了欧洲人发现新大陆后 150 年的红利。在欧洲,还有后来的美国,中产家庭都有一个带玻璃门的瓷器柜(这种瓷器柜就叫 China),里面展示着各种瓷质的餐具。家里没有瓷器柜,会被认为没有品位。

我们也对历史上的中国名瓷耳熟能详,比如代表性的唐宋青瓷、元明青花瓷,还有宋代著名的五大名窑——汝、官、哥、钧、定(图 5.3)。这些都是人造的奇迹、祖先智慧的结晶,也是我们的骄傲。但你可能不知道,当今欧洲瓷器占据着世界高端瓷器市场 90% 的份额,其余份额由美国和日本瓜分,Made in China(中国制造)的瓷器只能出现在中低端市场。这究竟是什么因素导致的呢?而且欧洲人喜欢讲“中国人发明了瓷器,后来欧洲人再发明了它”,这又有什么鲜为人知的故事呢?



图 5.3 中国名瓷,左图为汝窑杯盏,右图为元代青花瓷器——鬼谷下山

其实欧洲人制造瓷器的历史很富有戏剧性。由于和瑞典开战,萨克森公国的国王奥古斯都二世的财力几乎枯竭,于是他在 1706 年抓住了两个炼金术士来为自己炼制黄金,当然很快他就发现这件事是不可能的。由于在欧洲的瓷器售价堪比黄金,他就命令两个炼金术士研制瓷器,其中一个叫约翰·弗里德里希·伯特格尔的人因此而名垂青史。

从被奥古斯都二世软禁到阿尔布莱希茨堡(Albrechtsburg)到制造出欧洲的第一件瓷器,伯特格尔花了 4 年时间,做了 3 万次实验。他不仅记录了全部的实验过程和结果,而且把每一次实验之间的细小差异全都记录下来。与熟练掌握瓷器制造工艺却不明白其中的化学原理的亚洲工匠不同,这种科学实验和材料分析的方法让欧洲人对瓷器烧制的原理有了理性认识和定量的了解,他们可以通过细微调节瓷土中元素的配比和调整烧

1545 年,西班牙人在玻利维亚发现了银矿,第二年他们在墨西哥的萨卡特卡发现了更大的银矿,这个银矿至今仍然是世界三大银矿之一。

汝窑青瓷流传至今的真品已知的仅 67 件,“纵有家财万贯,不如汝瓷一片。”在香港苏富比于 2012 年举办的“中国瓷器及工艺品”拍卖中,北宋汝窑天青釉葵花洗经 34 口叫价,以 2 亿多港元成交。

伯特格尔的这些文件现保存于德国德累斯顿国家档案馆。为了保密,这些文件都是用密码书写的,只有他和助手看得懂。

珐琅彩瓷是一种将玻璃液化后烧制在瓷器表面的技术,不仅在瓷器表面营造出一种晶莹剔透的效果,也使得瓷器更加经久耐用。

1769年,韦奇伍德在自己开办的埃特鲁利亚陶瓷工厂里实行精细的劳动分工,把原来由一个人从头到尾完成的制陶流程分成几十道专门工序,分别由专人负责完成。这样一来,原来意义上的“制陶工”就不复存在了,分成了专门的挖泥工、运泥工、拌土工、制坯工等,他们必须按固定的工作节奏劳动,服从统一的劳动管理。

制过程来制造各种精致的瓷器。

伯特格尔的成功给萨克森公国带来了巨大的财富和荣誉,到了18世纪,德国麦森瓷器的售价已经是中国瓷器的两倍。今天麦森仍然是世界瓷都之一,并且在国际高端瓷器市场占有很大的份额。随后,奥地利和法国都在麦森瓷器的基础上不断研发新的工艺,比如西洋珐琅彩瓷器(图5.4,左)被欧洲人带到中国,康熙皇帝非常喜欢,下令在大内仿制,这实际上标志着中国在瓷器制造技术上已落后于欧洲了。



图5.4 欧洲制瓷工艺,左图为西洋珐琅彩瓷器,右图为韦奇伍德骨质瓷

18世纪中后期,“英国陶瓷之父”乔赛亚·韦奇伍德(Josiah Wedgwood)先是在工厂中搞出了一种叫作“流水线”生产管理方式,后来又把当时最先进的科技产品——蒸汽机引入瓷器制造。这些措施不仅极大地提高了瓷器的制造效率,而且不同批次的瓷器品质都能得到保障。他的后人在1812年还发明了骨质瓷器(图5.4,右),这种加入牛骨粉的制瓷工艺让瓷器更加结实,因此可以做得更薄,甚至薄到半透明的状态。正是从韦奇伍德的时代开始,瓷器首次在世界范围内供大于求。

从这段历史看来,欧洲人之所以在瓷器制造上超越中国,正是重视科学方法和数据记录的结果。欧洲人在研制瓷器过程中保留了全部的原始数据和实验报告,这样,前人每取得一点进步,后人都可以直接受益。例如,前面提到的伯特格尔把3万多次尝试的点点滴滴都保留了下来,同样,韦奇伍德在研制碧玉细炆器的时候,进行了5000多次实验,也把所有的细节都记录下来。相比之下,中国工匠更多的是具有对制瓷工艺的悟性,他们靠“师傅带徒弟”的方法将经验代代相传,而徒弟是否能超越师傅,则完全靠悟性。中间即使有一些发明和改进,却因为没有详细的过程记载,或许是出于保密故意略去,很多精湛的工艺都无法传世,比如宋代五大名窑的制作工艺大多失传了。这样后世常常不得不重复前人的失败,而无法直接“站在巨人的肩上”进行攀登,久而久之,造成了瓷器制造技术“起点很高,进步缓慢”的窘境。

这种对数据记录的不重视,不是中国瓷器制造业特有的问题,而是中国古代手工业普