题，实验设计需要进行调整时，调整的原因和过程也应被如实记录。

关于实验记录本的核心——数据表，《研究手册》给出了一些如何组织设计的建议。即使在研究方案中关于实验方法的部分对于数据收集的过程有足够充分的描述，学生仍然要在实验记录本上花时间仔细想想如何用合适的表格把定性或定量的数据很好地呈现出来。因为实验设计中收集的数据类型不同，可能需要设计不同的数据表格，为后面的统计分析打下良好的基础。《研究手册》对此给出了很具体的指导建议，例如定性数据和推论表、定量数据表，以及其他变量影响数据表等。学生通过组织设计一个实验记录本的实践，可以深刻理解在科学研究过程中，如何有组织地收录、展示数据比选择用什么方法记录数据更重要。

磨刀不误砍柴工，当上述 3 个步骤学生都考虑周详，并且充分准备好之后，实验研究工作就可以开始了。这 3 个步骤的学习实践可以更好地培养学生严谨的科学态度，掌握专业的研究方法，以及提高把研究想法和基本设计转变成可实施的研究活动的能力。■

玩转数据而不要玩弄数据

文 _ 曾　筝 / 中国科协青少年科技中心　梁宗保 / 东南大学

记得在上大学时，一位给我们讲授心理学研究方法的老师曾经说过：“学习统计就像是往沙漠里面浇水”。时至今日，我仿佛依然能够感受到当时这句话在我心中引起的那种强烈共鸣。没错，我在统计学习上没少下工夫，但却收效甚微。花了很多时间去区分那些看上去长得很像的术语，试着努力地记住每一个公式的适用范围，几乎是怀着毕恭毕敬的心情去演算练习题……但是，在这门课程结束不久之后，我发现自己对各种统计方法的认识还是非常模糊。直到开始撰写毕业设计论文的时候，我还是对数据统计存在畏惧感。在完成论文的过程中，我对老师的话有了新的理解：如果沙漠里面存在一片绿洲，浇进去的水就会蒸发得慢一些，研究项目就是沙漠中的绿洲。

学生进行 STEM 项目研究的过程，是学习像科学家和工程师那样思考和行动的过程。只有通过基本的统计方法对数据进行整理和分析，才能有足够的证据解释和说明研究的问题，验证假设是否得到支持，以及围绕某一个课题进行深入细致的探讨。从某种程度上看，数据统计可以算是科学研究中用于交流的“通用语言”，因此，学习和运用数据统计方法，属于科学研究中的基本方法训练。最近，在我翻译《STEM 项目学生研究手册》（以下简称《研究手册》）的相关章节时，其中的数据分析和图表展示方法介绍，给我的“沙漠”再次浇了一点水。也许是将高中生作为阅读对象的缘故，这本《研究手册》用浅显易懂的句子解释了统计学中的专业术语，跳过了枯燥公式演绎推导环节，用实例告诉学生各种方法所适用的情况，有什么需要注意的问题。本书作者的撰文思路和语言风格，让我直观地看到数据统计和图表展示在 STEM 项目完成过程中的关键性作用。而且，如果运用得当的话，它们还会使学生的项目变得更加出色。

我有一位朋友在大学里当老师，他最近几年一直在指导学生的毕业论文。当我向他咨询学生在研究过程中数据统计方面出现的问题时，他很激动地表示“有话说”，并很快给我列出了他指导的学生中最常见的几个问题。当我再次仔细阅读《研究手册》的相关章节时，我发现，这些常常出现在大学生的研究中的数据统计问题，在这本针对高中生的项目指南中

就已经提到了，并且给出了有效的解决建议。

问题 1：学生对研究结果预测的主观愿望过于强烈，在进行统计分析时，他们更喜欢那些符合自己预期的数据，而对那些看上去不太“美观”的数据结果采取选择性地忽略，甚至非常草率地删除或者舍弃掉。

《研究手册》在实验记录和数据分析章节中反复强调：不要先入为主地看待你的研究和数据结果。书中还提到一些很实用的操作方法来避免这个问题。例如，永远不要在实验记录本中撕页或者删除文本；永远不要用修正液覆盖错误和修改的地方，而是在相应部分轻轻划一道删除线。当使用电脑进行统计计算时，即使发现数据运算结果对于你的假设没有显著作用，也不要轻易删除数据。

问题 2：在统计学课堂上，学生常把大量的时间花在对统计的原理和公式推导的学习上，对数据进行实际操作计算的经验相对较少。当我问学生关于平均数差异检验、方差分析、回归分析的原理，大多数人都能回答上来。然而，当我要求他们针对一些具体数据，根据自己的研究假设去进行方差分析或者回归分析时，用不了多久，他们就会跑回来问我：“老师，方差分析中的自变量可以有多个吗？回归分析中的自变量可以同时有多个吗？”

正如我在前面提到的那样，这本《研究手册》指导我们有效地避免这个问题，因为它在写法上更侧重于技能的实际运用。例如在计算一组数据的标准差时，有总体标准差和样本标准差两种方程，书中并没有从原理上详细剖析两个方程的区别，只是简单直接地告诉读者：“标准差是一个高估值，来弥补犯错误的可能性，STEM 研究中使用样本标准差计算方程更恰当”（第 103 页）。对于比较复杂的推断性统计问题，《研究手册》也是采用举重若轻的介绍方法。学生不是很难理解方差分析中自变量可以有多少个吗？该书作者给了一个中学生很容易读懂的例子：“在比较一个弹射器以 3 种不同弹力发射 1 颗 4 千克弹丸的平均距离时，可以用方差分析进行检验。”（第 134 页）

问题 3：大多数学生在原始数据的处理方面存在欠缺。开始正式的数据统计之前，研究者通常需要对原始数据进行处理，例如，缺失数据的处理，是否有特异值，是否存在录入错误，数据是否呈正态分布等。我发现大多数学生认为这个环节可有可无，当他们收集到实验的原始数据之后，就兴致勃勃地开始进行描述统计和推断统计。这时我就会问他们一个问题：你怎么判断原始数据是否满足正态分布？如果数据不是正态分布，做出来的方差分析会有什么后果？到这个时候，大多数学生才恍然大悟。但是，紧接着他们可能会说出一句更雷人的话：“老师，我们不知道怎么做正态分布检验。”

遇到这样的学生应该怎么办？《研究手册》的作者可能已经与这样的学生打过交道，因为书中“描述统计”一章详细地介绍了离群值的判断和处理方法，以避免“雷人”的问题再次出现。作者首先告诉学生出现离群值并不可怕，因为“专业的科学家在使用新的方法时也有可能出错，这很常见”（第 101 页）。“如果你确认离群值的出现是由于自己的错误、仪器故障，或者其他你知道的与离群值无关的原因导致的，那么，你可以在对数据进行计算和分析时把它剔除掉”（第 101 页）。但是，别忘了“在结果中还是要提到离群值，并且在论文的结论部分要进行分析和解释”（第 101 页）。这才是科学研究应有的严谨态度。作者给学生介绍了“经验法则”，通过四分位距的计算结果辅助判断某个数据是否属于离群值。至于判断总体数据分布是否为正态，作者并没有向高中生提出这么高的数学要求，她只是提到“当数据呈非正态分布时，集中趋势的值可能会有变化”（第 99 页），“可以从正态分布数据的平均值和标准差中推断出数据的更多信息”（第 104 页）。

问题 4：缺乏制作图表的技巧。学生在进行实验室数据或调查数据的统计分析时，很多人不会使用图表呈现自己的结果。例如，在描述两个变量是否存在线性关系时，大多数人只

知道用相关系数或回归系数表示，却不会使用散点图呈现。在描述一个变量是否随时间变化时，很多人不清楚可以用折线图或曲线图表示。一些学生想到了要用图表呈现数据统计结果，但是他们只会把统计软件中输出的图表直接拷贝到实验报告中，导致后期无法修改图表。一些学生不会自己在文字编辑软件中绘制图表，更谈不上制作美观的图表。

图形和表格是用一种可视化的方法展现原始数据、描述统计结果和推断统计结果。使用图表的目的不仅仅是美观，更重要的是突显那些从研究中得到的重要发现。该书按照定量和定性数据两类介绍了图表的使用和绘制技巧，每一种图表都辅以实例，让学生一目了然。作者还提到了图表的标题、标注和图例使用中需要注意的一些小细节，例如，“标注必须要清楚地表现出图表要展示的信息，数轴、表格的行列名称都要有明确的词语和单位描述”（第 128 页）。

问题 5：在计算完成后，不知道如何对数据结果进行解释。经常会有学生来问我：“老师，我的结果计算出来了，但是我不知道应该看哪个？有很多个 F 值，我应该看哪一个，是修正的还是未修正的。在修正的 F 值后，哪一种修正更好，F 值显著了，这两个变量是什么关系？……”

学生不知道怎么解释数据这个现象反映出来的问题本质是，他们没有搞清楚自己为什么要进行数据统计。其实，数据统计运算不是盲目的，而是要验证研究结果是否支持假设，从大的方面来说，是要为解决某一个问题而进行的。《研究手册》中给学生列出了 3 个问题，帮助他们审视自己的数据。这 3 个问题是：“数据真实地反映了什么，从数据中还能提出哪些问题？”“数据如何反映两个变量之间的关系？”“数据能支持假设吗？”在这 3 个问题之下，作者从大量 STEM 学生项目经验中总结出一系列关于“为什么”和“如何”的问题，帮助学生更加深入地思考自己的数据结果。最后还不忘提醒学生：“得到不支持假设的数据也是有价值的。无论如何，都不能改变假设来匹配数据，或者改变数据来迎合假设！”

从研究假设出发，对实验中收集到的数据进行统计和分析，这个过程需要研究者具备足够的耐心、细致和敏锐的观察力。关于数据分析这个事，还有一点你应该知道的是：无论多“大牌”的科学家和工程师，在某一篇研究论文中反映出来的数据分析工作量，永远比他们研究过程中实际运算的少得多。数据分析的过程和结果，在研究论文中的结果和讨论部分会用到，但是这两个部分对结果数据使用的要求有差异。结果部分只是对数据进行描述，而不必进行解释和分析；讨论部分，研究者则需要对结果数据的所有特征和趋势进行解释、分析和讨论。

“听到的会忘记，看到的能记住，只有真正做过才能理解”，《研究手册》向读者传达的正是这样一个基本理念。《研究手册》中结合中学生 STEM 项目实践案例，告诉读者对于数据做统计分析不是单纯的数学计算，而是项目研究完整过程的重要环节，这个环节运用到的很多技巧和细节，都是在研究过程中摸索出来的，也只有通过亲自实践才能真正地掌握。相对于其他统计学教程，这本《研究手册》的写法更加通俗易懂，贴近 STEM 学生项目的研究实际需要，我推荐用作中学生了解统计学的入门材料。■

