



T 统计分析 在语言研究中的应用

TONGJI FENXI ZAI YUYAN YANJIU ZHONG DE YINGYONG

徐 鹰 编著



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS



统计分析

在语言研究中的应用

TONGJI FENXI ZAI YUYAN YANJIU ZHONG DE YINGYONG

徐 鹰 编著



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

· 广州 ·

图书在版编目 (CIP) 数据

统计分析在语言研究中的应用/徐鹰编著. —广州: 华南理工大学出版社, 2018. 10
ISBN 978 - 7 - 5623 - 5823 - 7

I. ①统… II. ①徐… III. ①统计分析 - 应用 - 语言学 - 研究 IV. ①H0

中国版本图书馆 CIP 数据核字 (2018) 第 234881 号

统计分析在语言研究中的应用

徐 鹰 编著

出 版 人: 卢家明

出版发行: 华南理工大学出版社

(广州五山华南理工大学 17 号楼, 邮编 510640)

http: //www. scutpress. com. cn E-mail: scut13@scut. edu. cn

营销部电话: 020 - 87113487 87111048 (传真)

责任编辑: 朱彩翩

印 刷 者: 虎彩印艺股份有限公司

开 本: 787mm × 1092mm 1/16 印张: 15 字数: 384 千

版 次: 2018 年 10 月第 1 版 2018 年 10 月第 1 次印刷

印 数: 1 ~ 1 000 册

定 价: 45.00 元



版权所有 盗版必究 印装差错 负责调换

前 言

随着语言学研究范围的扩大和研究方法的丰富,统计方法已成为语言研究中的必备工具。在现代语言学研究的各个领域(如二语习得、语言测试、心理语言学、语料库语言学、话语分析等),研究者经常使用各种统计方法,以便得到客观、可信、准确的结论。

本书由笔者近年来为华南理工大学外国语学院研究生开设的“语言研究中的统计学”课程讲义整理而成。为了让研究生充分掌握科学研究所必要的统计方法,笔者对研究生进行了课程需求调查,结果发现他们一致希望该课程应该注重理论和实践相结合,不仅应用浅显的语言将统计方法的原理阐述清楚,而且更应对每一种统计方法配套贴近实际研究的案例。为此,笔者在国内公开发表的实证研究中精心收集各种研究方法的使用实例,并且要求学习者从研究设计的角度对统计方法的使用进行反思。事实证明,这种教学方法取得了良好的效果。在课程结束时,研究生们普遍反映对统计学不再焦虑,而且自身的研究素养得到了较大提升。

本书主要结合研究设计(尤其是定量研究)讨论了统计学在语言研究中的应用,不仅对各种统计方法的原理和使用条件进行了系统介绍和梳理,而且提供了操作实例,详细分析了采用某种统计方法的步骤,力求做到理论和实践相统一。本书共包括16章。第1章到第4章是研究设计和统计学基本概念,重点介绍了定量研究,并讨论了其他主要的研究方法以及统计分析的基本概念。第5章到第11章介绍了各种统计方法,结合实例详细讨论了各种统计方法的使用范围、操作步骤、结果报告和解释。第12章到第15章以非实验研究的问卷调查为例,重点讨论了如何对问卷质量进行统计分析。第16章是对前文各章的总结,对比了各种统计方法的使用条件。每章结构大致相同,包括基本原理、操作实例、小结、思考和练习、延伸阅读文献五部分。

本书对广大外语教师和外语专业研究生有一定的参考价值,尤其适于

对统计分析有兴趣但是缺乏基础的初学人士。本书以实验研究和问卷调查为重点，力求回答“用统计分析能研究什么”“如何用统计分析进行研究”“为什么要用这种统计方法”三个问题，有助于让学习者跳出以统计分析为“术”的固有观念，而将统计分析作为开展科学研究的必要之“道”。

由于作者学识有限，书中错漏之处在所难免，敬请广大读者批评指正。

编者

2017年12月

目 录

第1章 研究方法概览	1
1.1 定量研究介绍	1
1.1.1 定量研究在语言研究中的现状	2
1.1.2 什么是定量研究	2
1.1.3 定量研究的哲学基础	3
1.1.4 定量研究的适用范围	4
1.1.5 定量研究的主要类型	4
1.2 实验研究	4
1.2.1 实验研究的设计步骤	4
1.2.2 实验研究的原则	6
1.2.3 实验研究的优、缺点	7
1.3 准实验研究	8
1.4 小结	8
1.5 思考和练习	9
1.6 延伸阅读文献	9
第2章 变量	10
2.1 变量的定义	10
2.2 变量的类型	10
2.2.1 自变量	10
2.2.2 因变量	10
2.2.3 调节变量	11
2.2.4 控制变量	11
2.2.5 介入变量	11
2.3 各种变量之间的关系	11
2.4 变量的测量	12
2.4.1 称名变量的测量	12
2.4.2 顺序变量的测量	13
2.4.3 等距变量的测量	13
2.4.4 比率变量的测量	13
2.4.5 变量的转换	13
2.5 如何在 SPSS 软件中进行变量操作	14
2.5.1 变量的输入	18
2.5.2 建立新变量	22
2.5.3 变量的分组	24



2.6	小结	26
2.7	思考和练习	26
2.8	延伸阅读文献	26
第3章	数据	27
3.1	数据类型	27
3.1.1	称名数据	27
3.1.2	顺序数据	27
3.1.3	等距数据	27
3.1.4	比率数据	27
3.1.5	称名数据、顺序数据和等距数据之间的转化	27
3.1.6	范畴性数据	28
3.1.7	连续性数据	28
3.1.8	范畴性数据和连续性数据的区别	28
3.2	频数分布	29
3.2.1	频数分布表	29
3.2.2	精确上、下限	30
3.2.3	频数分布图	31
3.2.4	如何在 SPSS 中画直方图和柱状图	31
3.3	集中趋势的测量	35
3.3.1	平均数 (Mean)	35
3.3.2	中位数 (Median)	36
3.3.3	众数 (Mode)	36
3.3.4	选择集中趋势统计量的标准	36
3.3.5	如何在 SPSS 软件中计算平均数、中位数和众数	37
3.4	离散趋势的测量	37
3.4.1	全距 (Range)	37
3.4.2	方差 (Variance)	38
3.4.3	标准差 (Standard Deviation)	38
3.4.4	如何在 SPSS 软件中计算全距、方差、标准差	39
3.5	标准分	39
3.5.1	如何比较分数	40
3.5.2	标准分的内容	40
3.6	描述性统计和推断性统计	40
3.7	小结	41
3.8	思考和练习	41
第4章	概率与假设检验	42
4.1	概率的基本概念	42
4.1.1	先验概率	42

4.1.2 后验概率	42
4.2 正态分布	42
4.2.1 正态分布的特征	43
4.2.2 标准分、正态分布、百分位之间的关系	44
4.2.3 正态分布表	47
4.2.4 正态分布理论的应用	47
4.2.5 T 分数	47
4.2.6 正态分布的检验	47
4.3 从样本统计量估计总体参数	50
4.3.1 总体	50
4.3.2 样本	50
4.3.3 随机抽样	50
4.3.4 总体参数与样本统计量	51
4.4 区间估计	51
4.4.1 区间估计的概念	51
4.4.2 样本平均数的抽样分布	52
4.5 区间估计的应用	53
4.5.1 样本抽样呈正态分布的区间估计	54
4.5.2 从小样本对总体平均数进行区间估计	54
4.5.3 估计精确度的样本容量	55
4.6 假设检验	55
4.6.1 假设的类型	55
4.6.2 显著性水平	55
4.6.3 显著性检验中的两类错误	56
4.6.4 单尾和双尾检验	56
4.6.5 检验统计值和临界值	56
4.6.6 参数检验和非参数检验	56
4.7 小结	57
4.8 思考和练习	57
第5章 两个独立样本的比较(组间设计)	58
5.1 参数检验——独立样本 t 检验	58
5.1.1 独立样本 t 检验的基本原理	58
5.1.2 独立样本 t 检验的效应量	59
5.1.3 独立样本 t 检验的操作实例	59
5.2 非参数检验——Mann-Whitney U 检验	63
5.2.1 Mann-Whitney U 检验的基本原理	63
5.2.2 Mann-Whitney U 检验的效应量	63
5.2.3 Mann-Whitney U 检验的操作实例	63
5.3 小结	66



5.4 思考和练习	66
5.5 延伸阅读文献	67
第6章 两个相关样本的比较(重复设计或匹配设计)	68
6.1 参数检验——配对样本 t 检验	68
6.1.1 配对样本 t 检验的基本原理	68
6.1.2 配对样本 t 检验的效应量	69
6.1.3 配对样本 t 检验的操作实例	69
6.2 非参数检验——Wilcoxon 配对符号秩次检验	72
6.2.1 Wilcoxon 配对符号秩次检验的基本原理	72
6.2.2 Wilcoxon 配对符号秩次检验的效应量	72
6.2.3 Wilcoxon 配对符号秩次检验的操作实例	72
6.3 非参数检验——符号检验	75
6.3.1 符号检验的基本原理	75
6.3.2 符号检验的操作实例	75
6.4 小结	78
6.5 思考和练习	78
6.6 延伸阅读文献	79
第7章 多个独立样本的比较(组间设计)	80
7.1 参数检验——单因素方差分析	80
7.1.1 单因素方差分析的基本原理	80
7.1.2 单因素方差分析的效应量	83
7.1.3 单因素方差分析的操作实例	83
7.2 非参数检验——Kruskal-Wallis 检验	93
7.2.1 Kruskal-Wallis 检验的基本原理	93
7.2.2 Kruskal-Wallis 检验的操作实例	93
7.3 小结	95
7.4 思考和练习	96
7.5 延伸阅读文献	96
第8章 多个相关样本的比较(重复设计)	97
8.1 参数检验——单因素重复测量方差分析	97
8.1.1 重复测量方差分析的基本原理	97
8.1.2 重复测量方差分析的效应量	98
8.1.3 重复测量方差分析的操作实例	98
8.2 非参数检验——Friedman 检验	104
8.2.1 Friedman 检验的基本原理	104
8.2.2 Friedman 检验的效应量	104
8.2.3 Friedman 检验的操作实例	104
8.3 小结	107

8.4 思考和练习	107
8.5 延伸阅读文献	107
第9章 相关分析	108
9.1 相关分析的基本概念	108
9.1.1 相关分析的特征	108
9.1.2 相关分析的用途	110
9.1.3 相关系数的类型	110
9.2 皮尔逊积差相关系数	110
9.2.1 皮尔逊积差相关系数的计算	113
9.2.2 皮尔逊积差相关系数的操作实例	117
9.3 斯皮尔曼等级相关系数	118
9.3.1 斯皮尔曼等级相关系数的计算	119
9.3.2 斯皮尔曼等级相关系数的操作实例	119
9.4 点二列相关系数	121
9.4.1 点二列相关系数的计算	121
9.4.2 点二列相关系数的操作实例	121
9.5 Phi 相关系数	121
9.6 肯德尔和谐系数	124
9.6.1 肯德尔和谐系数的计算	124
9.6.2 肯德尔和谐系数的操作实例	125
9.7 小结	128
9.8 思考和练习	128
9.9 延伸阅读文献	129
第10章 线性回归分析	130
10.1 简单线性回归	130
10.1.1 简单线性回归的基本原理	130
10.1.2 简单线性回归的操作实例	132
10.2 多元线性回归	138
10.2.1 多元线性回归的基本原理	138
10.2.2 多元线性回归的操作实例	138
10.3 小结	147
10.4 思考和练习	147
10.5 延伸阅读文献	148
第11章 卡方检验	149
11.1 卡方拟合优度检验	149
11.1.1 卡方拟合优度检验的基本原理	149
11.1.2 卡方拟合优度检验的操作实例	150
11.2 卡方独立性检验	152



11.2.1	卡方独立性检验的基本原理	152
11.2.2	卡方独立性检验的操作实例	153
11.3	小结	157
11.4	思考和练习	157
11.5	延伸阅读文献	157
第12章	问卷调查	158
12.1	问卷的设计步骤	158
12.1.1	问卷题项的设计	160
12.1.2	问卷回答的设计	160
12.2	问卷调查的优缺点	161
12.3	小结	162
12.4	延伸阅读文献	162
第13章	问卷的项目分析	163
13.1	问卷的项目分析	163
13.1.1	问卷项目分析的步骤	163
13.1.2	问卷的项目分析的操作实例	163
13.2	小结	179
13.3	思考和练习	179
13.4	延伸阅读文献	182
第14章	问卷的因子分析	183
14.1	效度的概念	183
14.2	效度的类型	183
14.3	问卷构念效度分析的方法——因子分析	183
14.3.1	因子分析的作用	184
14.3.2	因子的概念	184
14.3.3	因子分析的概念	185
14.3.4	提取共同因子的方法	185
14.3.5	因子分析的条件	190
14.4	因子分析的操作实例	190
14.4.1	第一次因子分析	190
14.4.2	第二次因子分析	198
14.4.3	主成分分析法和主轴分析法的对比	198
14.5	小结	201
14.6	思考和练习	201
14.7	延伸阅读文献	201
第15章	问卷的信度分析	202
15.1	信度的概念	202
15.2	信度的类型	202

第 1 章 研究方法概览

1.1 定量研究介绍

语言研究方法主要包括定量研究、定性研究和混合研究（Creswell, 2009）。研究方法的选择不仅涉及研究者的世界观、研究步骤以及数据收集、分析和解释的方法，而且还由研究问题的性质决定。因此，不同研究者对同一个研究问题可能采用不同的研究方法，同一位研究者对不同的研究问题也可能采用不同的研究方法。本书将具体讨论定量研究，重点是实验研究和问卷调查。就研究设计而言，心理语言学和应用语言学倾向于采用实验研究，社会语言学倾向于采用问卷调查。定量研究的主要步骤见图 1-1。

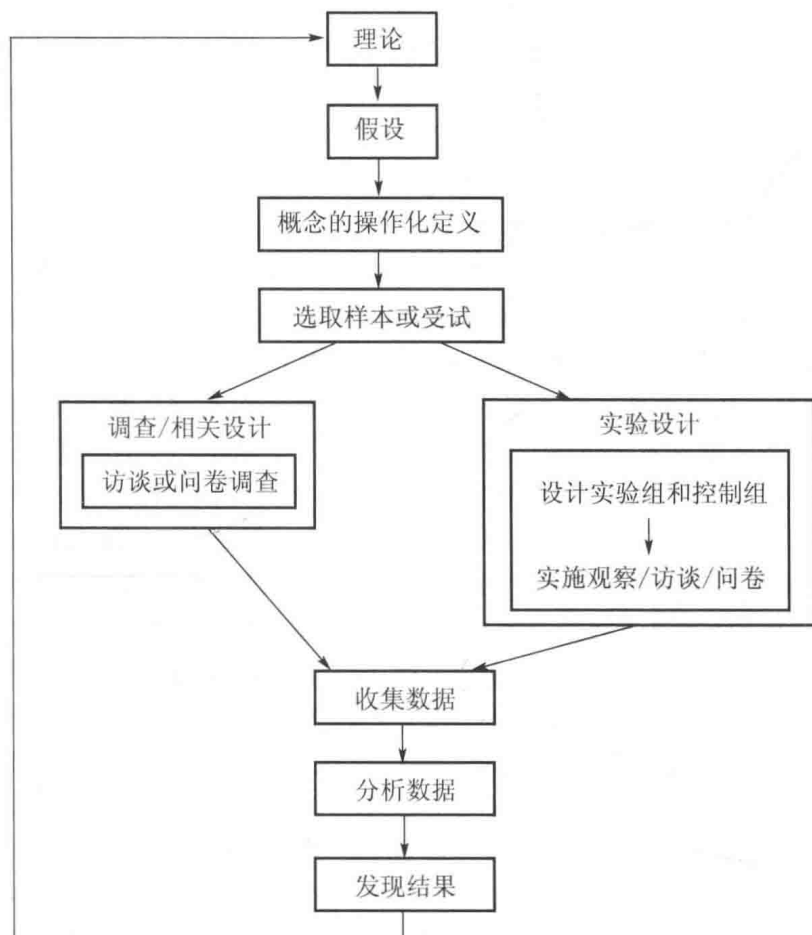


图 1-1 定量研究的主要步骤（Bryman & Cramer, 1999: 3）



1.1.1 定量研究在语言研究中的现状

目前,定量研究在语言研究中占据主导地位。在国际应用语言学界,Lazaraton (2005)对在1991—2001年期间发表在*TESOL Quarterly*、*Modern Language Journal*、*Language Learning*、*Studies in Second Language Acquisition*四大期刊上的论文进行了统计,发现在524篇基于数据的实证研究论文中,86%属于定量研究,而且在后三种期刊中定量研究的比例分别达到87%、95%、96%。

形成鲜明对比的是,我国的语言研究范式以思辨性为主,即使有一些研究采用了定量研究,但是局限于描述性统计分析,采用严格实验方法的研究不多,而且存在一些问题。吴旭东、张文忠(2002)对1997—2001年期间发表在我国七种主要外语教学研究学术期刊(《外语教学与研究》《外国语》《现代外语》《解放军外国语学院学报》《外语界》《外语教学》《外语与外语教学》)上的28篇实验报告进行了质量调查,发现:①半数以上的报告未能提供实验研究中的关键信息,如理论框架、研究假设、实验方案、实验步骤等;②半数以上的研究未能显示实验各阶段之间的内在关系,说明多数研究者缺乏充分借鉴相关研究成果的意识,而且对实验研究要求的精确和严谨还缺乏认识,尚未完全意识到实验研究各阶段必须环环相扣的重要性。

以上结果说明,定量研究是语言研究的主要范式,但是我国不少研究者并未充分掌握定量研究方法,尤其对于统计分析缺乏足够的认识。

1.1.2 什么是定量研究

Aliaga & Gunderson (2002)认为“Quantitative research is ‘Explaining phenomena by collecting numerical data that are analyzed using mathematically based methods (in particular statistics)’”,即定量研究是指通过收集能量化的数据并采用基于数学的方法(尤其是统计学)来解释各种现象的研究。该定义有三个重点。首先,定量研究和其他研究方法(定性研究、混合研究)一样,根本目的在于解释各种语言使用现象。其次,定量研究采用的数据必须能还原成数字,即能以数字的方式呈现。但是,几乎没有哪种语言使用现象本身就是以数字方式呈现的,因此需要借助研究工具将其转化为数字。例如,焦虑是一种心理状态,但Horwitz等(1986)通过开发外语课堂焦虑量表(Foreign Language Classroom Anxiety Scale)实现对外语课堂焦虑的测量,成功地将这种心理状态用数字呈现。总之,定量研究必须采用统计方法对数据进行分析。然而,定量研究中最重要部分并非统计分析,而是采用恰当的研究设计和正确的数据收集工具。与之相对的是,定性研究并不要求数据以数字的方式呈现,也不要求采用统计方法进行数据分析。

Creswell (2009)对定量研究的定义如下:“A means for testing objective theories by examining the relationship among variables. These variables, in turn, can be measured, typically on instruments, so that numbered data can be analyzed using statistical procedures. The final written report has a set structure consisting of introduction, literature and theory, methods, results, and discussion.”即定量研究通过分析变量之间的关系来验证某个理论。由于这些变量能够被各种工具所测量,所以采用统计方法能够分析这些量化数据。最后的研究报告通常采用固定的结构:引文、文献回顾和理论框架、研究方法、结果、讨论。Creswell的

定义对 Aliaga & Gunderson 的定义有三处明显的补充。首先,定量研究的目的在于对理论进行验证。其次,变量在定量研究中占有重要的地位。最后,定量研究报告必须遵循固定的体裁结构。

1.1.3 定量研究的哲学基础

定量研究的哲学基础是“实在论”(realism)或“实证主义”(positivism),其主要理念是存在客观真理(The truth is out there)。研究者的工作就是要采用客观的研究方法揭示客观真理。实证主义是400多年来自然与社会科学领域的常用理论之一(Guba & Lincoln, 1994),认为现实是绝对的,外在于研究者,受恒定的自然法则所驱动。实证主义认识论的内核是二元论和客观主义(objectivism),强调知识和意义存在于意识之外;现实是稳定的、可观察的、可发现的。因此,研究者和研究对象之间有清晰的界限,二者是相对独立的实体,它们之间的相互影响应加以限制。研究者必须置身于研究之外,采用客观、准确的方法最大限度地保证研究的客观性。同时,人们相信通过开发可信、有效的测量工具,能够通过验证理论的方式来认识客观世界。

实在论和实证主义的本体论前提是存在一个客观世界,且该客观世界按照固定的因果法则运作。因此,实证主义者在认识论上强调发现或尽量接近这个客观世界,在方法论上强调研究手段的科学性、实验性、可操作性以及研究假设的可验证性。科学研究就是要对这些客观法则进行验证,既有可能证实,也有可能证伪。如果关于客观世界的任何陈述不能得到经验的证实或证伪,那么是没有意义的。

实证主义范式下的语言研究坚持自然科学的严谨性和精确性,形成了一系列标准化定量分析法与线性研究程序。定量研究采用演绎分析路径(deductive analysis),首先以概念模型为起点,即根据理论或前期实证研究结果提出关于某些变量关系及其条件的研究假设,尤其关注变量之间的因果关系;然后,对研究假设进行操作化定义、量化处理和实验验证。不难看出,定量研究从概念模型的建立,到经验数据的采集和分析,各步骤之间是一种线性关系,最终的研究目的在于对各种语言使用现象或因果关系进行客观描述和理性揭示。定量研究强调价值中立(value-free)(Denzin & Yvonna, 2005),即不允许研究者自身所可能产生的主观偏见,因此常采用特定的设计方案(如随机抽样、统计分析)消除研究者或被研究者的主观偏见,进而保证结论的客观性和普遍性。

实在论的上述观点受到了不少批评。人们最大的质疑在于,研究者本身就是观察的客观世界的一部分,无法真正做到和观察对象泾渭分明。大量实证研究说明研究目标和研究结果不仅或多或少为研究者信念所左右,而且也受到研究当时政治、社会观念的影响。

后实证主义(post-positivism)接受了上述批评,并对实证主义的部分观点进行了修正。后实证主义者认为人们不能直接观察客观世界,因为研究者本身就是客观世界的一部分,同时自然科学并不能为所有的人文社会科学提供一个标准的研究模板。尽管如此,客观实在仍然存在。虽然人们无法直接通过研究揭示客观真理,但是仍然可以通过科学研究接近客观真理。后实证主义者不再把找到客观真理作为研究的终极目标,而是代之以努力表征他们所能接近的客观现实。不同于实证主义者,后实证主义者不相信绝对真理,即研究结果永远不存在必然性。他们认为研究结果具有或然性,人们对研究结果的信心是一个程度问题,即在多大程度上能相信某项研究的结论。



1.1.4 定量研究的适用范围

定量研究常用来回答以下四种问题：

- (1) 研究问题需要用定量的方式回答。例如本学期有多少学生选修了统计学课程？
- (2) 各种变化只能采用定量研究进行分析。例如本学期学生的语言能力是否有进步？
- (3) 研究者不仅要对事物进行描述，更要进行解释。例如哪些文本特征能区分不同写作能力的学生？
- (4) 研究者要对某个研究假设进行验证。例如学生的期末考试成绩是否和学生焦虑水平存在关系？

定量研究通常不适用于以下情形：

- (1) 当研究者需要对一个问题进行深入探究时，应采用定性研究，具体可以采用民族志研究、访谈、个案分析。而定量研究常需要大样本，能有效提升研究的宽度，但对提高研究的深度作用有限。
- (2) 定量研究适合对理论或假设进行验证，但是无法构建理论或假设。理论和假设来自文献综述或探索性定性研究。
- (3) 如果研究问题较复杂，涉及的变量较多，那么定性研究更合适。因为定量研究只能处理有限数量的变量，而定性研究可以允许产生未预计到的变量。
- (4) 定量研究适合于建立变量之间的因果关系，而定性研究则适合对研究问题进行深入探讨。

1.1.5 定量研究的主要类型

定量研究主要包括实验研究和非实验研究。实验研究通常被认为是一种科学研究范式，而非实验研究通常等同于问卷调查，是人文社会科学常用的一种研究方法。本章将重点讨论实验研究，第12章将重点讨论问卷调查。

1.2 实验研究

社会科学中的实验研究遵循了自然科学的实验研究范式，以实验为研究的基础。Muijs (2004) 对实验的定义如下：“A test under controlled conditions that is made to demonstrate a known truth or examine the validity of a hypothesis”；指的是在受控条件下的某种测试，旨在验证某个已知的真理或研究某个假设的效度。这一定义的关键词是“控制”。是否对变量施加控制是区分实验研究和非实验研究的主要区别。“控制”的含义是指研究者力图对实验环境施加最大限度的控制，从而能专注于研究者打算研究的变量（即自变量）。研究者会操纵可能影响实验结果的各种变量，也可能会预测变量。相对而言，在非实验研究中，研究者无法对所有可能的外在影响施加限制。

1.2.1 实验研究的设计步骤

实验研究通常包括以下八个步骤：

(1) 确定研究目标

研究目标决定了研究者想要研究的问题以及如何进行研究,涉及研究的本体论和认识论。研究者必须清楚地陈述自己的研究目标。研究目标的设定必须符合实际,具有可研究性。例如,假设某位研究者要研究各种教学因素对学生语言能力的影响,那么教学方法、教材、教师等都会对学生成绩产生影响,因此需要选择某一个变量进行研究,不能胡子眉毛一把抓,如研究交际教学法是否会影响学生的语言能力。其次,研究者需要对研究的总体做到心中有数。总体是指研究结果所能外推到的人群。例如某项研究探讨交际教学法是否会对100位学生的语言能力产生影响,那么希望研究结果不仅适用于这些学生,而且能外推到采用这种教学方法的所有学生。

(2) 形成研究假设

研究假设是对某些事实的尝试性解释,并且能够被进一步的研究所检验。定量研究必须提出能够被检验的假设。研究者可以收集数据,然后采用统计手段进行分析,最终作出拒绝或暂时接受某种假设的决定。需要注意的是,如果假设得到验证,那么也是暂时性的,因为随着新数据的出现,假设可能会被拒绝。研究假设通常来自对研究目标的细致分解。在实验研究中,研究假设可分为零假设(null hypothesis)和备择假设(alternative hypothesis)。通常来说,研究者希望能拒绝零假设,而接受备择假设。例如,如果某研究的目标是交际教学法是否会影响学生的语言能力,那么零假设:交际教学法不会影响学生语言能力;备择假设:交际教学法会促进学生的语言能力。值得一提的是,在实际研究中,大部分研究者会对零假设进行检验,因为标准的统计方法通常是用来检验零假设的。

(3) 建立研究设计

标准的实验设计按“前测-处理-后测”设计,受试通常被随机分为两组(实验组和控制组),实验组接受处理(比如在考察交际教学法对学生语言能力的效果研究中,实验组的学生接受交际教学法),而控制组不接受处理(非实验组的学生不接受交际教学法)。两组都会在处理之前接受前测,在处理之后接受后测。前、后测通常采用相同的测量工具,如某个语言水平考试。当后测结束时,需进行统计分析考察处理是否产生了效果。具体的设计方案见表1-1。

表1-1 标准的实验设计

受试	前测	处理	后测
实验组	×	×	×
控制组	×		×

标准的实验设计可以有以下几种变化。

首先,实验组可以不止一个,即可以对不同情况的处理进行研究。同时,也可以设置多于一个的控制组。其次,严格的前、后测设计有时并不会在实际中采用,典型的情况是经常不会安排前测。但是,如果不安排前测,将无法确定实验组和控制组在后测上的差异是来自处理还是来自两组受试本身的差异。再次,在研究设计中要注意“安慰剂效应”(placebo effect)。在医学研究中,人们发现有些病人的康复不是因为治疗本身的效果,而是因为病人对治疗的信念。因此,常采用的措施是对控制组提供安慰剂(通常为普通糖丸)。如果接受安慰剂的病人得到康复,那么说明他们的信念产生了作用。换句话说,如果不提供安慰剂,将无从得知处理的效果是否真正产生作用。如何控制“安慰剂效应”比较复杂,因为安慰剂本身就是一种处理,所以一般可以安排两个控制组:一个安慰剂控制组和一个真正的控制组(即不接受任何处理)。最后,在研究设计中还要注意“霍桑效应”(Hawthorne effect)。该效应是指那些意识到自己正在被别人观察的个人具有改变自己



行为的倾向。例如,如果实验组学生知道他们正接受某种有助于提高他们语言能力的教学方法(如交际教学法),那么他们的学习动机会增强,从而产生更好的学习效果。

(4) 选择研究工具

一旦确定了实验设计,那么需要选择或开发合适的前、后测工具。“工欲善其事,必先利其器。”研究工具的质量对研究的成败至关重要。首先,研究工具必须有效度,即测量了想要测的东西。其次,研究工具必须有信度,即研究工具能保证研究结果是稳定的、可靠的。这里暂时不重点讨论工具的信、效度问题。在实际研究中,建议尽量采用前人研究中已经验证过的工具,最好不要自己开发。如果必须自己开发,则一定要对其信、效度进行检验。

(5) 选择检验研究假设的合适处理量

在实验设计中,必须确定检验研究假设的合适处理量。这一点类似于医学研究上的药剂量。合适的药剂量可以治愈疾病,太少的量则可能不起作用,太多的量则可能导致医疗事故。在语言研究中也必须注意处理的强度,如在探讨交际教学法对学生语言能力的效果研究中,教学强度不足(如一周一节课)可能产生不了明显的效果,教学强度太高(如一周十节课)则可能让学生产生疲惫、厌倦心态,并且占用了其他课程的学习时间,反而减弱了教学效果。因此,有必要确定合适的处理量,常用的手段是在先导研究(pilot study)中设置不同处理量的实验组。

(6) 分组受试

在实验设计中,为了让外界因素的影响尽量最小化,即保证实验组和控制组在实验之前的差异足够小,必须保证受试的分组没有偏差。因此,通常采用随机分组,即受试是随机安排到实验组或控制组。为了确定随机分组是否能控制偏差,可以在分组后对受试的各种变量(比如说年龄、性别、语言水平)进行无差异分析。

(7) 谨慎进行实验

在进行实验的过程中,必须严格控制外在因素的影响。在交际教学法对学生语言能力的效果研究中,实验组和控制组必须严格分开,如至少保证他们在不同教室学习。其次,必须保证处理不会产生实验者偏差。在上例中,实验组和控制组的授课教师非常关键,如果实验组的教师对处理(交际教学法)信心满满,而控制组的教师对非处理(一般教学法)没有信心,那么实验效果必然大打折扣。

(8) 分析数据

实验结束后,可以对数据进行分析,通常采用的统计方法是 t 检验或方差分析,具体见第5章和第7章。

1.2.2 实验研究的原则

实验研究的原则主要包括:

- (1) 要有适当的统计检验力,变量之间有实质性的差异。
- (2) 参加研究的人数和资源最少。
- (3) 避免犯第一类错误,即错误地拒绝了零假设。
- (4) 无关干扰变量应得到控制。
- (5) 研究设计应能有效地计算出实验效果和误差效果值,这两个值能证明实验处理