

外国语言文学丛书

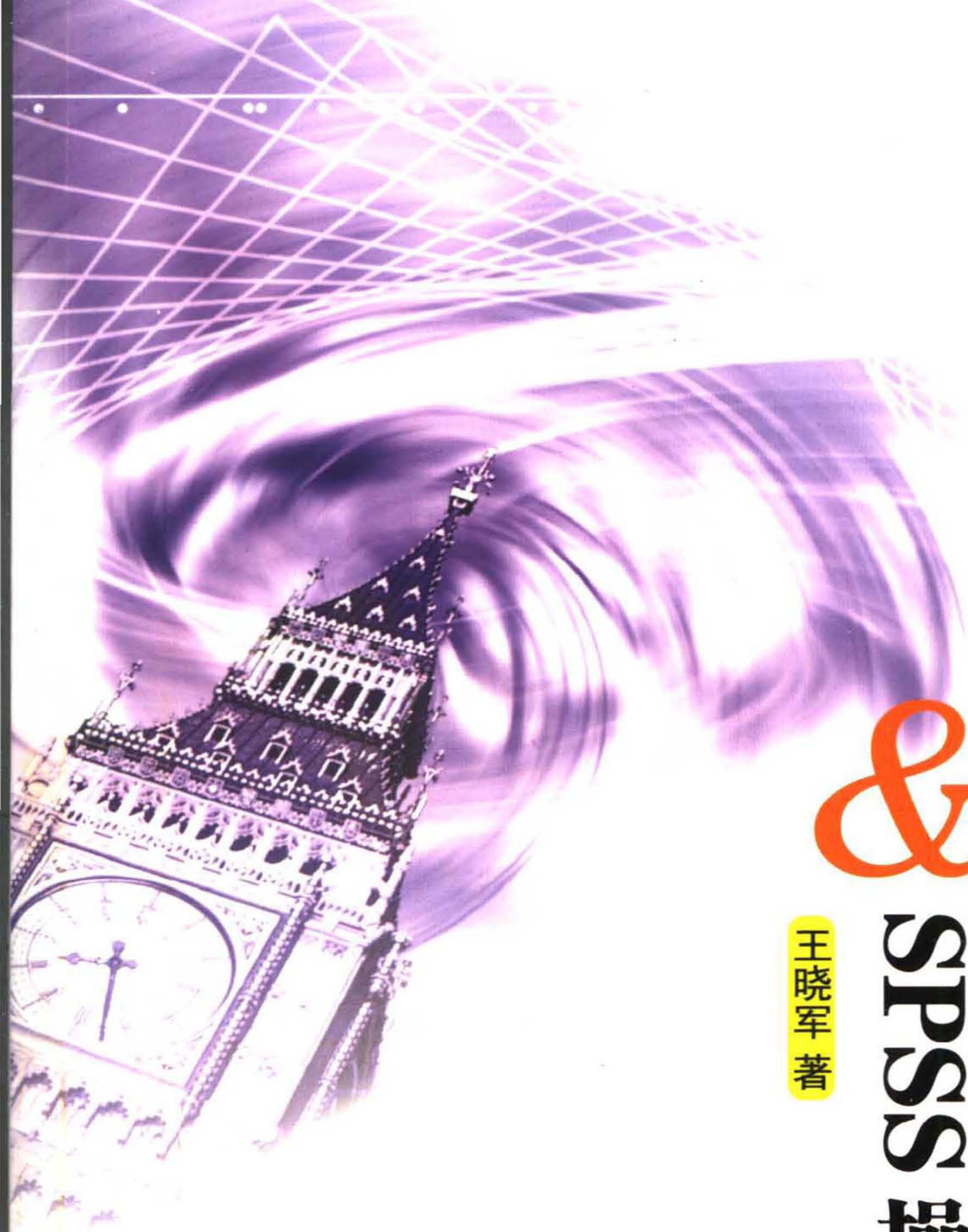
*A Communicative Perspective in Evaluation of the
Usefulness on An English Reading Test in Western
Region of China & Operation Method of SPSS*

西部英语测试实证研究

宁夏人民出版社

& SPSS 操作

王晓军 著



*A Communicative Perspective in Evaluation of the
Usefulness on An English Reading Test in Western
Region of China & Operation Method of SPSS*

& 西部英语测试实证研究 SPSS 操作

王晓军 著

外国语言文学丛书

宁夏人民出版社

图书在版编目(CIP)数据

西部英语测试实证研究与 SPSS 操作 / 王晓军著. —银川:
宁夏人民出版社, 2006.12

ISBN 7-227-03235-3

I. 西… II. 王… III. 英语 — 语言能力 — 测试 — 研究
IV. H31

中国版本图书馆 CIP 数据核字(2006)第 152380 号

西部英语测试实证研究与 SPSS 操作

王晓军 著

选题策划 哈若蕙
责任编辑 朱晓灵
封面设计 赵倩
责任印制 来学军

宁夏人民出版社 出版发行

出版人 高伟
地址 银川市北京东路 139 号出版大厦 (750001)
网址 www.nxcbn.com
电子信箱 nxcbmail@126.com
邮购电话 0951-5044614
经销 全国新华书店
印刷装订 宁夏施尔福印刷有限公司
开本 880mm × 1230mm 1/32
印张 7.375
字数 200 千
印数 2150 册
版次 2006 年 12 月第 1 版
印次 2006 年 12 月第 1 次印刷
书号 ISBN 7-227-03235-3/H·31
定价 26.00 元

版权所有 翻印必究



ZUOZHE JIANJIE
作者简介

王晓军，女，1965年5月生，祖籍山西省霍州市。1986年7月毕业于宁夏大学外语系英语专业，获教育学士学位。2003年3月获上海外国语大学文学硕士学位。2005年7月赴英国南普敦学院培训。研究方向：应用语言学之测试学。现为宁夏大学外国语学院教授，外国语学院大学英语教学部主任，语言测试研究所主任，宁夏高等院校外语教学研究会理事。长期从事英语专业研究生语言测试学、非英语专业研究生及本科生的英语教学工作。参加教育部教改课题1项，参加并完成宁夏社科基金课题2项、宁夏教育厅教改课题1项，主持并完成宁夏大学社科基金课题2项，参加并完成宁夏大学社科基金课题3项，主持宁夏大学教改课题、外国语学院课题各1项。出版著作1部、参编教材1部，著作2部，发表学术论文30余篇。

:: 策划统筹 哈若蕙

:: 责任编辑 朱晓灵

:: 执行统筹 贺秀红

:: 装帧设计 赵 倩

NEIRONG JIANJIE

内容简介

本书在系统介绍语言测试最新理论与流程的基础上着重推出西部英语测试实证研究成果，是国内语言测试研究领域的一部新作，具有广泛的学术参考价值。

编委会

丛书主编 周玉忠

丛书编委 (以姓氏笔画为序)

兰宁鸽 安淑英 周玉忠 周 震

郭鸿雁 瞧秀梅



高校的中心工作是教书育人,培养德智体美全面发展的社会主义建设者和接班人。科研与教学结合是大学最显著的特点,也是大学区别于科研机构特征。科研与教学相辅相成,科研也是提高教学质量的条件和保障。高校犹如一架马车,教学与科研是车之双轮,二者相互依存,相互促进,缺一不可。

宁夏大学外国语学院自 2002 年 5 月成立以来,学院非常重视教学与科研的并行发展,制定了一系列科研奖励和管理办法,用制度管人、管事,充分调动了全院教师的积极性,科研热情空前高涨,科研成果不断涌现,极大地促进了教学工作。自 2003 年以来,全院教师每年发表论文在 100 篇以上,几乎达到每人年均 1 篇。在核心刊物上发表的论文趋逐年增多,学院科研工作实现了“量变”到“质变”的跨越式发展,从根本上改变了学院的科研局面,为学院学科与专业建设、学位点建设奠定了坚实的基础。英语语言文学从院级重点学科一跃发展成为校级重点培育学科。为了进一步建设好该学科并使其争取成为自治区级重点学科,学院及时提出了科研工作上水平、上台阶的工作思路,使学院科研工作由数量的“扩张型”变为质量的“提高型”。提出今后课题的申报、专著的出版等方面要有所突破。令人欣喜的是,2006 年学院在国家社科基金的申报上已有了历史性的突破:宁夏大学全校申报到 12 项,外国语学院占 3 项,位居第一,真是可歌可贺!为了使教师在学术著作出版方面也有所突破,学院决定用 3~5 年的时间出版一套“外国语言文学研究丛书”,共计 10 本,每年出 2~3 本。这一

计划得到了宁夏人民出版社的大力支持,哈若蕙副社长给丛书的出版设计提出非常好的意见,她领导的综合编辑室具体负责丛书的编辑出版。这就是我们这套丛书应运而生的背景。为出好这套丛书,学院成立了专门的编委会,由学院相关领导及学术骨干组成,由学院学术带头人、院长周玉忠教授任编委会主任,以确保丛书的选题、审阅、编辑和出版。我们相信,这套丛书的出版意义非同小可,会对学院今后的本科生、硕士生教学起到极大的促进作用,对学科与专业的建设与发展产生重大的影响和积极的推动作用。

最后,我们感谢宁夏人民出版社副社长哈若蕙女士以及综合室的各位编辑!感谢丛书编委会的各位委员,也感谢学院教师的参与和学院的大力支持。

外国语言文学研究丛书编委会

二〇〇六年十月

王晓军同志现任我院英语教授,大学英语教学部主任。在宁夏大学与上海外国语大学合办的硕士学位班中师从著名语言测试专家邹申教授研究语言测试学,获得语言学硕士学位。此后一直从事语言测试学的教学与研究工作,对我校大学英语教学的测试分析、总结做出了积极的贡献。晓军同志是位勤奋好学、富于进取的人。作为军嫂,她家务繁重;作为大英教师,她教学任务繁忙。但她各方兼顾,一切都处理得井然有序,令人佩服。教学之余,她潜心科研,成果不断涌现,堪称大学英语教学部的科研典范之一。我知道,这一两年她在构思一部有关西部语言测试方面的书,但未曾料到她能这样快就完成。所以,当她拿着书稿请我为其作序时,我既高兴又惊讶!随即欣然答应写上几句。

对于广大外语教师和外语工作者,语言测试是一个既熟悉又陌生的课题。说熟悉,是指经验的感受。说陌生,是指科学的分析。

谁没有经历过几百次测试?谁不曾领教过各种题型?似乎其间没有什么深刻的道理。但说到科学的分析,恐怕懂得运用测试理论知识去研究分析测试结果,解决教学科研中的问题的是很有限的。

《西部英语测试实证研究与 SPSS 操作》就是这样一部科学分析的专著。它在系统地综述了语言测试理论的基础上,对西部英语测试的特色进行了多角度的分析论证,肯定了成绩,指出了不足。填补了本地区语言测试研究领域的空白。本书数据资料翔实,论述条理清晰,理论功底扎实,实践经验丰富,研究方法科学,既

侧重理论分析的准确性,又注重测试操作的实用性,是国内语言测试研究领域的一部新作,显示了作者独具的才华和观察力。此外,本部专著对语言测试涉及到的 SPSS 应用操作实现过程、结果分析以及图形制作等专题进行了分门别类的详细介绍。因此,本书的出版无论对于语言测试工作者、外语教师还是研究人员都有较高的参考价值,同时,对推动本地区外语测试科学化有着积极的现实意义。

周玉忠

2006 年 7 月



随着国内外语言学研究的发展与深入,语言测试的研究也日益细化,研究成果层出不穷。但是,一直以来,由于历史与地理的原因,西部,特别是宁夏,在语言测试研究领域成果殊少,令人担忧。笔者有感于开展这方面的研究工作的迫切性和必要性,在2002年宁夏大学外国语学院语言测试研究所成立之后,即开始收集资料,着手研究,力求为这方面的工作尽一些绵薄之力。现在,数年的夙愿终于实现,笔者的心情是可以想见的。

本书分为上、下两篇,上篇为语言测试理论与SPSS操作。系统介绍了语言测试领域的最新理论研究与测试流程以及统计描述、信度分析、相关分析、回归分析在SPSS中的实现过程与结果分析。下篇为西部英语测试实证研究,由笔者的上海外国语大学文学硕士学位论文组成。

本书的写作出版受到宁夏大学外国语学院的关怀和资金支持,宁夏人民出版社哈若蕙副社长认真审阅了全书并提出了可行的修改意见,外国语学院院长周玉忠教授对书稿的审定提出了宝贵的意见,上海外国语大学邹申教授对本书的下篇部分进行了悉心的指导和匡正。这里,谨向以上单位和个人的支持和帮助表示诚挚的谢意。此外,本书的撰写受到母亲和丈夫的倾力支持,笔者一并表示感谢。

由于笔者水平所限,书中难免有不当之处,敬请读者批评指正。

王晓军

2006年7月于银川学林园

目录

序 / 001

前言 / 001

上篇 语言测试理论与 SPSS 操作

第一章 语言测试的性质与历史 / 003

1.1 语言测试的性质 / 003

1.2 语言测试的历史 / 003

第二章 语言测试的理论框架 / 007

2.1 信度 / 007

2.1.1 考试信度 / 009

2.1.2 评卷信度 / 010

2.1.3 如何提高信度 / 012

2.2 效度 / 015

2.2.1 表面效度 / 015

2.2.2 内容效度 / 016

2.2.3 结构效度 / 016

2.2.4 尺度关联效度 / 017

2.2.5 信度与效度的关系 / 019

2.3 真实性 / 020

2.4 交互性 / 022

2.5 后效作用 / 023

2.6 可操作性 / 024

CONTENTS

第三章 语言测试的用途及类别 / 026

3.1 语言测试的用途 / 026

3.1.1 语言测试用于诊断及反馈 / 026

3.1.2 语言测试用于筛选或选拔 / 026

3.1.3 语言测试用于编班 / 026

3.1.4 语言测试用于科研或调查 / 027

3.2 语言测试的类别 / 027

3.2.1 按学习阶段分类 / 027

3.2.2 按考试目的分类 / 028

3.2.3 按考试方式分类 / 030

3.2.4 按考试形式分类 / 030

3.2.5 按对考试分数的解释分类 / 031

3.2.6 按试卷的评阅方式分类 / 032

第四章 语言测试的总体设计 / 034

4.1 制定考试规范 / 034

4.1.1 考试规范与考试大纲的区别 / 034

4.1.2 制定考试规范的意义 / 034

4.1.3 考试规范与考试效度的关系 / 034

4.1.4 考试规范与信度的关系 / 035

4.1.5 考试规范与总体设计的关系 / 035

4.1.6 考试规范与考试评估的关系 / 035

4.2 考试规范内容的制定 / 036

4.2.1 确定考试目的 / 039

4.2.2 确定考试能力范畴 / 040

目录

- 4.2.3 确定考试对象 / 041
- 4.2.4 确定考试内容 / 042
- 4.2.5 确定考试类别 / 043
- 4.2.6 确定考试题型 / 044
- 4.2.7 确定试卷结构与分值 / 045
- 4.2.8 制定考试细目表 / 048
- 4.2.9 命题及编辑试卷 / 048
- 4.2.10 制定评分标准 / 049
- 4.2.11 保证考试实施 / 052
- 4.3 选择答题类评估方法 / 053
 - 4.3.1 多项选择题 / 053
 - 4.3.2 是非题 / 055
 - 4.3.3 匹配题 057
- 4.4 非选择答题类评估方法 / 058
 - 4.4.1 填充题 / 058
 - 4.4.2 转换题 / 062
 - 4.4.3 简答题 / 065
 - 4.4.4 论述题 / 067

第五章 考试成绩反馈 / 068

- 5.1 分数的频数分布 / 068
- 5.2 分数频数的图形显示 / 069
- 5.3 分数的集中量 / 072
 - 5.3.1 平均数 / 072
 - 5.3.2 众数 / 073

CONTENTS

- 5.3.3 中位数 / 073
- 5.3.4 平均数、众数、中位数的关系 / 073
- 5.4 分数的差异量 / 073
 - 5.4.1 全距 / 074
 - 5.4.2 标准差 / 074
 - 5.4.3 方差 / 075
 - 5.4.4 差异系数 / 076
- 5.5 正态分布 / 077
- 5.6 偏度、峰度 / 078
- 5.7 四分位数、十分位数、百分位数 / 079
- 5.8 分数的转换 / 080
 - 5.8.1 原始分数 / 080
 - 5.8.2 转换分数 / 080
 - 5.8.3 Z 分数和 T 分数 / 081
- 5.9 题目难易度 / 082
- 5.10 题目区分度 / 084

- 第六章 统计描述与 SPSS 应用 / 088**
 - 6.1 平均数、中位数、众数、全距、标准差、方差、百分位数、频数、峰度、偏度 / 088
 - 6.1.1 SPSS 中的实现过程 / 088
 - 6.1.2 结果与讨论 / 091
 - 6.2 四分位数 / 093
 - 6.2.1 SPSS 中的实现过程 / 093
 - 6.2.2 结果与讨论 / 094

目录

6.3 十分位数 / 094
6.3.1 SPSS 中的实现过程 / 094
6.3.2 结果与讨论 / 096
6.4 百位分数 / 096
6.4.1 SPSS 中的实现过程 / 096
6.4.2 结果与讨论 / 097
6.5 标准化 Z 分数的线性转换 / 097
6.5.1 SPSS 中的实现过程 / 097
6.5.2 结果与讨论 / 099
6.6 标准化 T 分数的线性转换 / 099
6.6.1 SPSS 中的实现过程 / 099
6.6.2 结果与讨论 / 101
第七章 信度分析与 SPSS 应用 / 102
7.1 同质性信度 / 102
7.1.1 SPSS 中的实现过程 / 102
7.1.2 结果与讨论 / 105
7.2 分半信度 / 108
7.2.1 SPSS 中的实现过程 / 109
7.2.2 结果与讨论 / 110
7.3 再测信度 / 111
7.3.1 SPSS 中的实现过程 / 111
7.3.2 结果与讨论 / 112

CONTENTS

第八章 相关分析、回归分析与 SPSS 应用 / 113

8.1 相关分析 / 113

8.1.1 定义 / 113

8.1.2 SPSS 中的实现过程 / 114

8.2 回归分析 / 118

8.2.1 定义 / 118

8.2.2 一元线性回归分析 / 119

8.2.3 多元线性回归分析 / 131

下篇 西部英语测试实证研究

——从交际法角度评估西部地区阅读测试的实用性

Introduction / 141

Chapter One Theoretical Concerns in Language Assessment / 147

1.1 Validity / 148

1.1.1 Face Validity / 148

1.1.2 Content Validity / 149

1.1.3 Construct Validity / 150

1.1.4 Criterion-related Validity / 150

1.2 Reliability / 151

1.2.1 Test Reliability & Scorer Reliability / 152

1.2.2 Enhance the Reliability / 153