

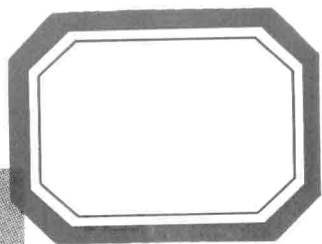
语言测试方法论

二十一世纪高等院校学术发展系列

补爱华 著



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS



言测试方法论

补爱华 著◎

21世纪高等院校学术发展系列



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

《语言测试方法论》是一部关于语言测试理论与实践的专著。本书通过对语言测试的原理、设计原则、标准和评估等基本理论的研究和探讨,旨在帮助从事语言教学工作的广大英语教师较为全面的了解和掌握语言测试知识,以便在教学实践中提高命题水平并保证考试质量,对广大外语教学工作者有较好的指导和借鉴作用。

图书在版编目(CIP)数据

语言测试方法论/补爱华著. —上海:上海交通大学出版社,2011

ISBN 978-7-313-07325-9

I. 语... II. 补廖 III. 语言—测试—研究
IV. H09

中国版本图书馆 CIP 数据核字(2011)第 090659 号

语言测试方法论

补爱华 著

上海交通大学出版社出版发行

(上海市番禺路 951 号 邮政编码 200030)

电话:64071208 出版人:韩建民

常熟文化印刷有限公司 印刷 全国新华书店经销

开本:787mm×960mm 1/16 印张:19.5 字数:367 千字

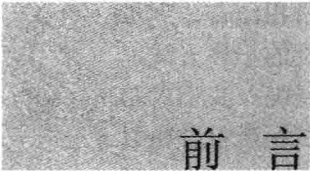
2011 年 7 月第 1 版 2011 年 7 月第 1 次印刷

印数:1~1030

ISBN 978-7-313-07325-9/H 定价:48.00 元

版权所有 侵权必究

告读者:如发现本书有质量问题请与印刷厂质量科联系
联系电话:0521-52219025



前言

Preface

外语教学离不开各种各样的语言测试，小到课堂练习、课外作业、单元测试，大到期中、期末考试以及目前最具规模的大学英语四、六级考试。语言测试已经成为英语教学的一部分。正确运用语言测试能有效地评估教学效果和教学质量，了解学生语言熟练程度，诊断和分析学生语言错误。熟悉和了解基本的英语测试理论和实践，不仅能够使广大英语教师理解各种考试，以便在教学中恰当地突出重点，减少盲目性，而且利于教师自身专业知识和技能的发展，以便在对学生的评价过程中合理使用英语测试方法。

语言测试是语言教师专业知识和技能的一个重要组成部分。英美发达国家把语言测试作为外语教师教育的一个主要课程，因为它从属于教师教育评价技能，在形成性评价或者终结性评价过程中必不可少。1990年，由美国的三个全国性教育机构联合颁布的《教师、学生教育评价能力的标准》，包含了七条规定。大致内容是教师必须具有选择适合于教学目的的评价方法的技能；发展适合于教学目的的评价方法的技能；实施、评阅和解释各种评价方法的技能以及恰当使用评价结果的技能等。虽然文件规定的是教师整体的教育教学评价技能标准，但是教育评价涵盖使用考试技术的内容。在美国，不具备教育教学评价能力的教师，必须接受专门培训才能继续从事教育工作。由此可见教师评价能力的重要性。

教师要有全面的教育教学评价知识，了解教育评价的多种方式。其中一个重要的评价方法就是考试。考试与教学是硬币的两个面，这一点教师们十分清楚。不过，使用包括考试在内的多种评价手段进行形成性评价和终结性评价还需要认识和培训。Weir (2001)指出：“我们都认识到在课堂上，测试没有起到形成性评价的作用。我们建议对此引起关注。”“重点应放在如何使考试更好地达到为学生进步提供有用反馈信息的目的。目前需要强调用于教学过程中的测量学生进步的、有效的运

用语言的考试形式。测试必须与语言学习的各个发展阶段相结合。”虽然大多数语言教师对考试是很熟悉的，但是作为教育评价的一种方式，怎样才能科学有效地使用考试来促进教学、促进学生发展呢？我们对与考试相关的命题理论和方法、试题类型及其考试重点、试题质量分析以及考试成绩解释等都应有所了解，这样才能掌握使用恰当的考试方式的恰当时机，取得促进学生发展的最佳效果。

我国英语教育界对英语测试的了解和认识也只是在20世纪90年代才开始的。目前我国，英语测试课程是不少高等院校师范专业本科阶段的必修课程之一，也是英语语言教育硕士的必修课程之一。这是因为英语测试的基本理论和实践与学生评价技能已经是我国教师专业化的一个重要内容。长期以来，在我国教育第一线的绝大部分英语教师对语言测试的原理、设计原则、标准和评估都缺乏较为全面的了解，普遍存在着语言测试与语言教学相分离的现象。因此，充分了解并掌握语言测试的基本理论，全面掌握语言测试的知识，以便在实践中提高命题水平并保证考试质量已成为广大英语教师的当务之急。

本书的读者对象为从事英语教育工作的广大英语教师。本书对语言测试理论和实践进行了较为深入的探讨，对广大外语教学工作者有较好的指导和借鉴作用。

本书共分十章。第一章概述语言测试的历史及其发展、语言测试的意义以及与外语教学的关系。第二章主要介绍语言测试的类型。第三章探讨语言测试常用题型及特点。第四章探讨语言测试评估原则以及试卷命题流程。第五章探讨语言测试数据统计原理及其分析方法，其中重点介绍如何运用SPSS统计软件包进行语言测试数据统计与分析。第六章探讨听力测试设计的原则及其方法。第七章探讨口语测试设计原则及其方法。第八章探讨阅读测试设计原则及其方法。第九章探讨写作测试设计原则及其方法。第十章介绍和探讨目前国内外比较有影响力的大规模英语测试特点及其方法。

由于作者水平有限，书中难免存在差错与不足，恳请读者批评指正。

作者
2011年3月

目 录

Contents

第一章 语言测试与语言教学

1.1 语言测试	1
1.1.1 前科学语言测试	1
1.1.2 心理测量-结构主义语言测试	1
1.1.3 心理语言学-社会语言学测试	2
1.1.4 交际语言测试	2
1.1.5 语言测试的三种理论模式比较	5
1.2 语言测试的意义以及与外语教学的关系	5
1.3 计算机化测试	7
1.4 语言测试发展趋势	10

第二章 语言测试类型

2.1 分离式测试、综合性测试、交际性测试和语用测试	13
2.1.1 分离式测试	13
2.1.2 综合性测试	14
2.1.3 交际性测试	14
2.1.4 语用测试	14
2.2 水平测试、成绩测试、诊断性测试、潜能测试和编班测试	15
2.2.1 水平测试	15
2.2.2 成绩测试	15

2.2.3 诊断性测试	15
2.2.4 潜能测试	16
2.2.5 编班测试	16
2.3 直接测试和间接测试	16
2.3.1 直接测试	16
2.3.2 间接测试	16
2.4 主观测试和客观测试	17
2.4.1 主观测试	17
2.4.2 客观测试	17
2.5 常模参照测试和标准参照测试	17
2.5.1 常模参照测试	17
2.5.2 标准参照测试	18

第三章 语言测试常用测试题型

3.1 客观试题	19
3.1.1 多项选择	19
3.1.2 完型填空	27
3.1.3 听写	32
3.1.4 正误判断	33
3.1.5 简答	36
3.1.6 接续	39
3.1.7 配伍	41
3.1.8 排序	44
3.2 主观试题	46

第四章 语言测试评估原则

4.1 语言测试评估基本原则	49
----------------------	----

4.1.1 信度	49
4.1.2 效度	54
4.1.3 信度和效度的关系	56
4.1.4 难度	57
4.1.5 区分度	57
4.1.6 实用性	58
4.1.7 反拨作用	59
4.2 交际语言测试的评估原则	61
4.3 语言测试命题流程	64
4.3.1 设计阶段	64
4.3.2 操作阶段	65
4.3.3 考试实施阶段	65

第五章 语言测试数据统计与分析

5.1 语言测试常用数据统计与分析方法	67
5.1.1 集中趋势	67
5.1.2 离散趋势	68
5.1.3 正态分布与标准分	70
5.2 运用SPSS统计软件包进行语言测试数据统计与分析	73
5.2.1 数据准备	74
5.2.2 测试成绩统计分析的操作步骤	75
5.2.3 运用SPSS统计软件对语言测试成绩的实例分析	76
5.2.4 运用SPSS统计软件对试卷进行质量分析	83

第六章 听力测试

6.1 听力理解的实质	89
6.2 听力构件	90

6.3 听力测试的主要方法	95
6.4 听力测试原则	98
6.4.1 确保听力测试的真实性	98
6.4.2 精简多项选择题的选项	100
6.5 听力测试常用题型	101
6.5.1 听音辨音题	101
6.5.2 听音理解题	102
6.5.3 听写题	118
6.6 听力测试的准备与实施	125

第七章 口语测试

7.1 口语活动的特点	127
7.2 口语能力	127
7.3 口语测试原则	129
7.4 口语测试方式	129
7.4.1 直接型口试	129
7.4.2 半直接型口试	130
7.4.3 间接接型口试	130
7.5 口语测试评分方法	131
7.5.1 分析法	131
7.5.2 综合法	131
7.6 口语测试评分标准	132
7.7 口语测试常用题型	139
7.7.1 结构主义口试及典型题型	139
7.7.2 交际口试及典型题型	155
7.8 口语测试的准备与实施	163

第八章 阅读理解测试

8.1 阅读理解的实质	165
8.2 阅读理解测试的目的	166
8.3 阅读理解测试原则	168
8.3.1 注意把握和合理分配体现阅读能力三个层次	168
8.3.2 选择合适的阅读材料	168
8.3.3 合理把握干扰项的难易程度	169
8.4 阅读理解测试常用题型	173
8.4.1 多项选择题	173
8.4.2 简答题	178
8.4.3 信息转换题	179
8.4.4 配对题	180
8.4.5 正误判断题	183
8.4.6 翻译	185
8.4.7 填空题	186

第九章 写作测试

9.1 写作能力	189
9.2 写作测试原则	191
9.2.1 写作题目测试的写作技能应是命题者期望考生能够掌握的技能	191
9.2.2 明确写作测试任务	191
9.2.3 把握写作测试任务的难易度	192
9.2.4 考生根据题目所写的作文能真实反映考生的写作能力	192
9.2.5 命题的同时应该考虑评分的可信性	193
9.2.6 科学、客观地评价作文题的质量, 进一步提高写作测试效度	194
9.3 写作测试评分方法	194
9.3.1 整体评分法	194

9.3.2 分项评分法	197
9.4 写作测试作文题目的命题类型	200
9.4.1 文体学分类	200
9.4.2 考试实用分类	201

第十章 国内外大规模英语考试概述

10.1 全国大学英语考试	209
10.2 全国硕士研究生入学统一（英语）考试	216
10.3 高等学校英语应用能力考试	227
10.4 全国公共英语等级考试	234
10.5 高等学校英语专业四级、八级考试	243
10.5.1 英语专业四级考试(TEM 4)	243
10.5.2 英语专业八级考试(TEM 8)	248
10.6 全国国际商务英语认证考试	253
10.7 托福考试	256
10.8 雅思测试	269
10.9 美国研究生入学考试	282
参考书目	294

第一章 语言测试与语言教学

1.1 语言测试

语言测试经历了四个发展阶段：前科学时期、心理测量—结构主义时期、心理语言学—社会语言学时期、交际语言测试时期。对应后三个阶段三种语言测试模式，分别为心理测量—结构主义模式、心理语言学—社会语言学模式、交际语言测试模式。

1.1.1 前科学语言测试

语言测试是随着外语教学的发展而出现的，人们把20世纪40年代的语言测试统称为前科学语言测试。在这个时期，由于人们对语言缺乏科学的认识，认为语言是一门知识，主要包括语法知识、词汇知识和语音知识。外语教学也把语言当成一门知识传授。在教学方法上基本上是凭经验、遵循传统，主要采用语法翻译法(grammar-translation approach)，在此基础上产生的语言测试也就只考查学习者对这三方面知识的掌握情况，一般只有笔试，没有口语和听力测试。总的说来，该阶段的语言教学和测试完全以教师和命题人员的经验和主观判断来确定，基本谈不上有什么科学依据。

1.1.2 心理测量—结构主义语言测试

心理测量—结构主义语言测试是建立在结构主义语言学和行为主义心理学理论基础之上，依据心理学领域的测量学为科学方法而形成。该阶段人们对语言第一次有了一个较客观、科学的分析和认识，认为语言是一套符号系统，是一种工具、手段，语言可分解为语言技能（主要包括听、说、读、写技能）和语言成分（主要包括语音、语法、词汇），学习语言就是要获得操作这套符号系统的技能，即训练对刺激作出反应的一套语言习惯。所学的和所考的就是操作这套符号的技能(skills)，即语言能力，也就是说，语言测试应该测试语言的应用，而不是有关语言的问题。该阶段的外语教学主要采用直接法、听说法和认知法。受该阶

段外语教学方法,尤其是听说法的影响,语言测试发生了巨大变革,出现了分离式语言测试(discrete test)。分离式测试通常把语言测试按听、说、读、写等语言技能分成若干部分,分别进行测试。

心理测量—结构主义语言测试在20世纪50、60年代在外语测试中占统治地位,直至80年代、90年代初仍为主流。心理测量—结构主义语言测试力求避开前科学语言测试主观性强和精确度差等弊端,采用统计学的方法来评价试题质量,解释、分析测试结果,并将定性和定量结合起来,努力实现客观和科学测试。其测试方法以“多项选择”、“句子改写”、“句子填空”等形式为主,该测试方法有如下优点:① 评分客观、公正,提高了阅卷信度。② 测试面广,可以对掌握某些语言特征的能力进行非常细致的考查,并兼顾口语和笔试。③ 适合大规模测试。但又有如下不足之处:① 出题难度大,如:不易把握多项选择中的干扰项标准。② 偏重判断能力,忽视语言表达能力,学生往往高分低能,对教学的负面反拨作用较大。总的说来,心理测量—结构主义语言测试注重语言知识的运用,强调听、说技能,尤其是听的技能。教学和考试用的目标语比前科学测试体系要真实。但由于该体系强调听、说、读、写等技能的分离处理,强调不同语言现象的分离处理,因此往往忽视语言情景,没有结合恰当的情景使语言具有更进一步的真实性。

1.1.3 心理语言学—社会语言学测试

20世纪70年代初期,语言测试界开始重视所谓整体综合法的研究,这一时期被称为心理语言学—社会语言学时期。综合法是针对采用以分离为主进行测试的方法而提出来的。离散题把语言加以分解,而综合测试正好相反。主张采用离散题的人认为,一次只可以测试一个语言点,而综合法则主张通过一次测试全面地评价学生的总体语言水平,学生必须综合地运用各种语言知识和技能。这一时期的语言学家认为语言不但是一个可以分解的体系,更是一种动态的、具有创造性的功能体系。社会语言学家提出了语言交际能力(communicative competence)的概念,认为使用语言不但要能够按照语法造出合格的句子,而且还必须具有在不同的语境中合理地使用这些句子的能力,这一点是孤立地测试语言点的单纯离散题做不到的。这一时期采用得比较多的是综合题,如完型填空(Cloze)、综合改错等。

1.1.4 交际语言测试

随着语言研究的深入,人们对语言的认识也进一步深化。60年代中期,著名

语言学家Chomsky在彻底否定行为主义语言理论的基础上创建了转换生成语法,提出了语言能力(linguistic competence)和语言行为(linguistic performance)两个不同概念。但他所提出的语言能力只是一个抽象的语言能力,由于人们交往时只有抽象的语言能力是不够的,还要考虑到使用语言时的语境,并且语言的使用过程是一个动态的交际过程,由此便产生了交际能力这一概念。Hymes (1972)指出:Chomsky (1965)的“能力”不涉及语言的使用,也不系统地考虑人们在特定环境中对语言的适当使用。他认为交际能力应包括四个社会文化因素:

(1) 可能性程度(degree of possibility)。某一语言系统中可能存在的东西,如语音、语法、句法、词汇、语义、语用等方面的知识。

(2) 可行性程度(degree of feasibility)。语言使用者个人心理方面的语言容量,如个人的记忆力和认知能力。

(3) 合适性程度(degree of appropriateness)。在交际中言语行为的表达是否得体,如对背景、目的、常规、参加者等因素的合理考虑。

(4) 出现程度(degree of performance)。言语行为能否发生。

Hymes认为,这四种因素相互作用构成一个整体。应该指出的是,尽管Hymes提出的交际能力试图使交际能力与一定的语境、对象、目的联系起来,但它仍和Chomsky(1965)所提出的能力一样,是一个绝对的、静态的和理想化的概念。在语言实践中没有任何人完全具有Hymes所说的交际能力,交际能力却应该是一个相对的、动态的、具体的、可比的概念。尽管如此,Hymes的交际能力理论仍对外语教学与测试产生了深远影响。

20世纪80年代,Canale&Swain(1980)提出的交际理论丰富了Hymes的交际能力学说。在当时,该理论模式在应用语言学界影响广泛,为外语测试提供了比较科学的理论基础。他们认为交际能力包括四个方面的知识和技能:

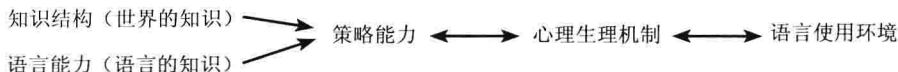
(1) 语法能力(grammatical competence)。指语言本身的知识(包括词汇、构词规则、语音、句法等方面),主要涉及正确理解和表达话语(utterance)的字面意义的知识,也就是句平面上遣词造句的能力。

(2) 社会语言能力(sociolinguistic competence)。指在不同的社会语言环境中适当理解和表达话语的能力。语境因素包括话题(topic)、交际双方的社会地位和交际目的。

(3) 语篇能力(discourse competence)。指把语法形式和意义融合在一起,用口头或书面形式连贯地表达同种类语篇的能力。语篇的完整统一通过语言形式衔接(cohesion)和语义的连贯(coherence)来实现。语篇能力也就是超句平面上组织语段的能力。

(4) 策略能力(strategic competence)。指为了加强交际效果或弥补由于缺乏交际能力等因素引起的交际中断所使用的策略,包括言语和非言语两方面。换句话说,策略能力就是在交际过程中的应变能力。

20世纪90年代,美国著名应用语言学家Bachman(1990)发展了Canale&Swain的交际能力模式,提出了新的交际能力理论模式。Bachman的交际语言能力学说是迄今为止我们所能见到的关于语言能力学说中最全面、最完整的学说。在20世纪70—80年代的测试界,最引人注目的问题是“怎样测试”和“测试什么”。Bachman的交际测试理论不仅涵盖了这两大问题,而且提出了测试的评价原则。Bachman借鉴和发展Halliday(1976)、Hymes(1972)等人的研究成果,提出自己的语言能力观:① 语言能力应包括语法规则如何使用语言达到特定交际目的的知识。② 语言使用是一个动态过程,语言能力的各成分之间互相作用。语言交际能力的构成及其成分的内在关系如下图所示:



(Bachman, 1990)

图中的知识结构指语言使用者的社会文化知识及关于现实世界的知识。语言使用环境包括交际双方、环境、话题和目的等。Bachman认为,语言交际能力包括三部分:语言能力、策略能力和心理生理机制,对这三个组成部分,每一个部分都可以细分。

Bachman的交际语言测试模式充分吸取了近代语言学及应用语言学领域的最新研究成果,对交际能力的认识更客观、全面、深刻,指出交际能力各构成因素之间不是一个简单的并列关系,是一个互动(interactive)关系,即语言的使用过程是一个动态的过程(dynamic process),各种知识、技能和心理过程交织在一起,互相影响,互相作用。Bachman交际语言模式的提出对语言测试产生了深刻影响,这就意味着在进行语言测试时,要把涉及交际能力的各种因素结成一个整体加以测试,从而出现了语言综合能力测试(integrative competence testing),这可以说是测试史上的一个里程碑。

交际语言测试是建立在交际教学法思想基础之上的现代测试方法,随着外语教学越来越强调交际能力及语言运用能力的培养,交际测试法也越来越倾向于将题型涉及到具体的现实生活中存在的或极其类似的活动或可能存在的任务,并

要求学习者借助所学的语言来完成这些任务,答题时,不仅要合乎语法,还要得体。测试类型也越来越倾向直接测试法,其重点是考查语言的实际运用,其结果更能反映学习者实际外语水平。与心理测量—结构主义语言测试时期的侧重语言技能的分离式测试方法不同的是,该测试侧重语言的交际能力。

1.1.5 语言测试的三种理论模式比较

在前面的论述中我们总结了语言测试理论发展的四个阶段和三种测试模式,这里我们对这些理论模式从所依据的语言理论、测试方式、效度、信度、命题难度以及对教学的反拨作用等方面加以比较,具体如下:

心理测量—结构主义模式	心理语言学—社会语言学模式	交际语言测试模式
分解语言成分	不可分语言能力	分解交际功能
分离式试题	(一般)综合性试题	综合性试题
间接测试能力	间接测试能力	直接测试能力
常模参照	常模参照	标准参照
效度低	效度低	效度高
客观评分	(一般)客观评分	主观评分
评分信度高	评分信度高	评分信度低
重视语言知识的准确性	重视语言知识与技能的准确性	重视语用的得体性
阅卷速度快	阅卷速度适中	阅卷速度较慢
命题难度大	命题容易	命题简便
需抽样分析	无需抽样分析	需抽样分析
便于大规模考试	便于中小规模考试	不便于大规模考试
测试听、读、写	测试听、读、写	测试听说读写综合能力
对教学有不良反拨作用	对教学有不良反拨作用	积极促进教学

(戴曼纯, 1999: 32-42)

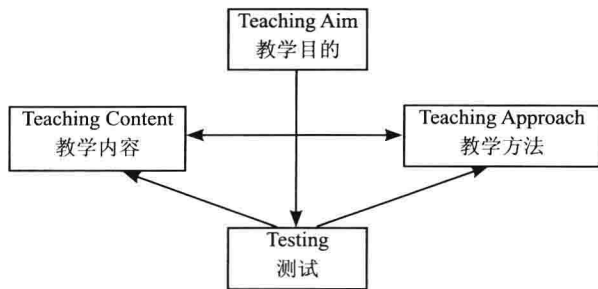
总而言之,语言测试经历了从注重语言知识到注重交际能力的转变,由科学前的经验型转向了当今现代语言测试的科学型,是一个质的飞跃。语言测试将随着相关学科的发展而日益完善。

1.2 语言测试的意义以及与外语教学的关系

纵观语言教学法的历史,语言教学主要涉及三个方面:“学什么,怎么教和怎么学(Davies, 1990)。”“学什么”是有关语言本身的研究,也就是人们对语言的认识。“怎么教”涉及语言教学法的研究。而“怎么学”是研究学习者的语言

习得过程。从语言测试理论发展的过程可以看出,语言教学和语言测试经历了由重视语言知识转向语言技能,再转向语言实际运用能力的过程。一种教学方法的流行往往会有—种相应的测试模式伴随产生,外语测试与外语教学法的发展进程即从语法翻译法、听说法到交际法的过程是一脉相承的。不同历史时期的语言教学法直接受语言观和语言学习观的指导和支配,而语言测试的内容和形式又直接反映了不同历史阶段的语言观和语言学习观。语言测试对教学的影响是广泛存在的,这种影响有正面的也有负面的。人们对语言测试的认识,较多地谈及其负面影响。如考试是教学的指挥棒,教学围绕考试转,忽视了对学生实际运用语言知识能力的培养,学生为考试而学等等。然而,科学合理的语言测试,其积极影响甚于其负面的影响。教学过程中,语言测试的基本运用是给教学评估提供重要的信息。通过对这些信息的分析和研究,授课教师深入了解学生的整体情况,从而能更科学、更有效地进行将来的教学工作。语言测试在教学方面至少能起3个作用:① 信息反馈作用。为教师和学生提供准确的反馈信息,帮助教师了解自己的教学效果;帮助学生了解自己的学习状况。② 诊断和教学把关作用。帮助教师和学生发现教学与学习过程中的长处与不足,进一步提高教学质量与学习效果。③ 评估与导向作用。准确、客观、公正地检测和评价学生的语言水平和能力,提高学生学习的动力和积极性。

语言学家Hughes(1989)把教学与测试的关系客观而形象地描述为“伙伴关系”(partnership),下面的图示反映了教学与测试的相互关系。



许多语言学家都很重视测试在语言学习中的重要性,Alderson(2000)认为语言测试在语言教学中处于中心地位,不仅如此,它在很大程度上对它所从属的应用语言学这一学科也有着不小的贡献。语言测试的必要性具体体现在以下四个方面(Alderson, 2000):

- ① 把应用语言学的理论运用到实践中。
- ② 为教学计划和大纲提供建立依据和标准。