

语言测试

理论与英语教学

研究

薛晓萍 郭宝申 徐辉 编著



YUYAN CESHI LILUN
YU YINGYU JIAOXUE YANJIU

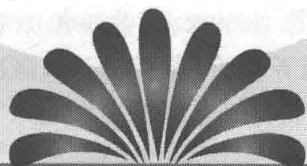
中国商业出版社

语言测试

理论与英语教学

研究

薛晓萍 郭宝申 徐 辉 编著



YUYAN CESHI LILUN
YU YINGYU JIAOXUE YANJIU

中国商业出版社

图书在版编目(CIP)数据

语言测试理论与英语教学研究 / 薛晓萍, 郭宝申,
徐辉编著. — 北京: 中国商业出版社, 2012. 6

ISBN 978-7-5044-7771-2

I. ①语… II. ①薛… ②郭… ③徐… III. ①英语—
测试—研究②英语—语言教学—教学研究 IV. ①H319

中国版本图书馆 CIP 数据核字(2012)第 127590 号

责任编辑:武文胜

中国商业出版社出版发行

010-63180647 www.c_cbook.com

(100053 北京广安门内报国寺1号)

新华书店总店北京发行所经销

北京市登峰印刷厂印刷

* * * * *

787 毫米×1092 毫米 16 开 15.125 印张 368 千字

2012 年 6 月第 1 版 2012 年 6 月第 1 次印刷

定价:36.00 元

* * * * *

(如有印装质量问题可更换)

前 言

21 世纪,经济全球化对人们的生活与生存方式的影响无处不在,国与国之间综合国力的较量日益突出。事实上,在综合国力竞争的背后,是教育、科技、人才的竞争,而外语教育在其中又体现出了特殊的价值。目前,外语教育几乎成为每个国家公民教育的一个重要部分,如果一个人不懂外语,那么他人生的发展将会遇到各种障碍。因此,在我国,学习一门外语,尤其是英语,已成为 21 世纪公民的基本素养之一。然而,当我国面对着一亿多个正在学习英语的人们时,进行语言测试似乎显得十分必要了。语言测试一方面可以使教师了解学生的学习目的与动机,采用怎样的学习方法与策略,影响他们学习的因素等;另一方面,可以使教师清楚使用什么样的教学方法和教材能使学生获得成功,如何对学生进行评价等。

《语言测试理论与英语教学研究》结合我国学生实际,将语言测试融入到英语教学中,做到了理论与实践结合。希望广大学者借此能够在语言测试和英语教学方面得到一定的提升。

本书共包括十一章。第一章到第三章分别对语言测试的概念、语言测试的产生与发展、语言测试的原则与设计进行了详细的叙述,为以下各章节的展开提供了理论框架与知识背景。在第一章,本书主要对语言测试的作用与目的、语言测试的种类与标准、语言测试与其他学科的关系、语言测试与英语教学进行了概述。第二章主要阐述了语言测试的历史演变和发展趋势,使我们对语言测试的来源和发展有一定的了解。第三章研究了语言测试的基本原则和设计,通过对这部分知识的了解,可以使我们更好地组织语言测试。第四章到第十章分别将不同的英语教学与语言测试相结合,包括语法、词汇、阅读、听力、口语、写作、翻译的教学与测试,分别介绍了它们的教学内容与目标、原则与方法,以及语言测试,使我们对各种语言测试有了更为深入的认识。第十一章重点对测试成绩的分析及解释进行了研究,首先阐述了考试分数的意义,然后列举了考试成绩数据分析的方法,最后探讨了考试成绩分析结果呈现及解释。

《语言测试理论与英语教学研究》一书在写作过程中参阅了大量的文献资料,借鉴了很多专家和学者的研究成果,在此对他们表示衷心的感谢。同时,本书观点新颖、语言朴实严谨、条理清晰,旨在为广大学生的英语学习提供理论与实践指导,同时也为英语教师的英语教学、其他志在从事英语教学的工作者、英语语言教学研究者及对英语教学感兴趣的人士提供些许可供借鉴的资源。

本书由吕梁学院薛晓萍、大连海洋大学郭宝申、沈阳理工大学徐辉编著完成,并由三人共同统稿。具体分工如下。

第二章,第三章,第六章:薛晓萍;

第七章,第八章,第九章第一节,第十章,第十一章:郭宝申;

第一章,第四章,第五章,第九章第二节、第三节:徐辉。

鉴于水平和时间有限,加之一些问题尚处于探索阶段,本书难免有不妥之处,敬请同行专家和读者批评指正。

作者

2012年5月

目 录

第一章 语言测试概述	1
第一节 语言测试的作用与目的	1
第二节 语言测试的种类与标准	3
第三节 语言测试与其他学科的关系	9
第四节 语言测试与英语教学	11
第二章 语言测试的产生与发展	14
第一节 语言测试的历史演变	14
第二节 语言测试的发展趋势	17
第三章 语言测试的原则与设计	22
第一节 语言测试的基本原则	22
第二节 语言测试的设计	39
第四章 英语语法教学与测试	60
第一节 语法教学的内容与目标	60
第二节 语法教学的原则与方法	63
第三节 英语语法测试	73
第五章 英语词汇教学与测试	82
第一节 词汇教学的内容与目标	82
第二节 词汇教学的原则与方法	85
第三节 英语词汇测试	91
第六章 英语阅读教学与测试	100
第一节 阅读教学的内容与目标	100
第二节 阅读教学的原则与方法	103
第三节 英语阅读测试	116
第七章 英语听力教学与测试	130
第一节 听力教学的内容与目标	130
第二节 听力教学的原则与方法	135
第三节 英语听力测试	144

第八章 英语口语教学与测试	154
第一节 口语教学的内容与目标	154
第二节 口语教学的原则与方法	160
第三节 英语口语测试	166
第九章 英语写作教学与测试	179
第一节 写作教学的内容与目标	179
第二节 写作教学的原则与方法	185
第三节 英语写作测试	198
第十章 英语翻译教学与测试	205
第一节 翻译教学的内容与目标	205
第二节 翻译教学的原则与方法	207
第三节 英语翻译测试	218
第十一章 测试成绩的分析及解释	225
第一节 考试分数的意义	225
第二节 考试成绩数据分析方法	226
第三节 考试成绩分析结果呈现及解释	231
参考文献	233

第一章 语言测试概述

语言测试发展到今天,已经成为一门独立的学科,其工具性价值逐渐凸显,影响领域愈来愈多,因而语言测试研究也逐渐成为显学。语言测试的作用与目的、语言测试的种类与标准以及其他有关语言测试的一般理论问题、语言测试与其他学科的关系、语言测试的过去、现在与将来、语言测试的总体设计等都成为研究的重点问题。另外,语言测试与英语教学的关系、语言测试在英语教学中的运用也受到了广泛的研究。本章正是在这样的背景下,对语言测试的基本知识进行阐述。

第一节 语言测试的作用与目的

马克思政治经济学一个基本的观点是,使用价值是价值的基础。对一门学科的研究,尤其是对语言测试这种具有实用性与工具性的学科的研究,首先需要明确的是,这门学科自身有多大的意义,即它的作用是什么,它能够实现什么样的目的?只有在此问题的基础上,得出满意的答案,下面的研究才是有意义的,也才是可行的。

我们此处使用语言测试的作用与目的,表面上看是两个方面,但实质上是一个问题。作用是从动态的角度探讨语言测试具有的用途,即其对人的有用性;目的是从静态的角度探讨经过动态的过程,语言测试所达到的一种状态,换句话说,语言测试目的是探求语言测试用途最终到达的一种状态。因而,用途与目的完全可以归结为一个问题,对其中一个问题的探讨,就是对另一个问题的诠释。

正因上述两方面的原因,在本书开篇第一节,我们旨在回答这样一个问题:为什么研究语言测试?同时,我们认为将此问题具体地分为语言测试的作用,或者更通俗地讲,语言测试的用途有哪些与语言测试的目的这两个方面是毫无意义的。因而,我们仅对语言测试的作用进行探讨。下面是国内外一些学者较有代表性的看法。

(1)厄尔(Penny Ur)教授在其专著《语言教学教程:实践与理论》(*A Course in Language Teaching: Practice and Theory*)中,将语言测试的作用概括为以下九点。

- ①教师通过测试掌握学生的水平,进而确定下一步教学计划。
- ②学生通过测试了解自己的不足,进而明确他们需要努力的地方。
- ③评估教学效果。

- ④激励学生学习或复习特定的材料。
- ⑤让班级安静下来,集中精力听讲。
- ⑥掌握授课节点,使得整个课程结构更明晰。
- ⑦促使学生为在测试中获得更好或满意的成绩而努力。
- ⑧向学生布置任务,任务本身就是一种复习、练习甚至测试。
- ⑨使学生在在学习中获得进步,拥有成就感。

从内容上看,以上九点更像是对学校教学中测试作用的概括,可以大致分为以下两类。

从教师方面来看,测试可以使教师掌握学生所处的水平,并指导他们的教学;掌握授课节点,使得整个授课结构更明晰;通过测试激励学生学习。

从学生方面来看,学生可以通过测试把握自己的不足,通过测试中的好成绩获得满足感与成就感。

但是,这里需要指出的是,上述九点的概括不是在完全意义上的概括,而仅仅限于与语言教学有关的语言测试。

(2)海宁(Henning,2001)将语言测试的目的概括为以下六个方面。

- ①Diagnosis and Feedback.
- ②Screening and Selection.
- ③Placement.
- ④Program Evaluation.
- ⑤Providing Research Criteria.
- ⑥Assessment of Attitudes and Sociopsychological Differences.

(3)阿兰·戴维斯(1998)将语言测试的用途概括为六个方面。

①语言测试用于研究,即研究者通过运用语言测试的技巧获得有关深入开展讨论所需的数据。

②提供实验的信息。这是第一种用途的一部分,即把语言测试作为语言教学实验的标准。

③传递自身的信息。测试结构、内容和方法本身,承担着影响教学的责任。

④衡量学习者的进步。

⑤对学生进行选拔。

⑥关系到对课程、材料、方法的评估。

(4)我国学者刘润清(1991)教授在其专著《语言测试和它的方法》中则将语言测试的用途概括为以下三个方面。

①语言测试的第一个用途是用于语言教学。

②语言测试可以用来选拔人才。

③语言测试经常用于社会调查。

以上是从广义上对语言测试用途的一种概括。刘润清指出,他所要讲的语言测试并不是如此广义的语言测试,而只限于与语言教学有关的语言测试。

刘润清(2000)随后修订了自己就此问题的观点,完善了其内容,将语言测试的用途大致分为诊断、筛选、编班和研究四种。

从上述学者们的论断中,可以得出以下结论:第一,我们所说的语言测试的作用,或者说是

用途,有广义与狭义之分,这是源于语言测试本身具有广义与狭义之分的基础之上的。广义的语言测试的用途将与语言教学有关的语言测试仅作为研究的一部分;狭义的语言测试的用途将与语言教学有关的语言测试作为研究的全部内容。第二,学者们探讨的语言测试的用途,主要指的是与教学活动有关的用途。这里所谓的“与教学活动有关”指的不仅是教学活动,还包括选拔学生、编班活动以及与教学有关的科研活动。第三,虽然学者们的论述的具体内容不尽相同,用语差别也较大,但是本质上对语言测试用途的观点却大同小异,有共性可循。

一般来讲,语言测试的用途包括以下几点。

(1)诊断,即语言测试可以用于诊断。语言测试的诊断功能主要表现在以下两个方面。

①教师通过测试对学生的情况作出诊断,以便掌握学生的学习水平,明确学生的优势与不足。根据测试的结果,针对不同的学生,制定相应的教学措施与方法。

②学生通过对测试结果的分析与研究,明确自身的不足,进而改进自身的不足,克服自身的缺点,以取得更大的进步。

(2)筛选,即语言测试用于筛选。语言测试的另一个功能可以用于从一定数量的学生中选拔一部分达到预设标准的学生进入某一特定专业进行学习。比如,申请到澳大利亚留学,除了已经获得的学业成绩,有关人员的推荐信以及一些申请留学方面的行政程序之外,学生还要提交雅思成绩,甚至对成绩的最低值做一定的限制。这就是通过语言测试来进行筛选。

(3)编班,即语言测试用于编班。这里主要适用于新生入学后分班的现实需要。但是,需要明确的是,这里的编班并不是真正地讲学生分为不同的班级,而是通过语言测试,掌握学生的知识水平,进而将知识水平相近的学生编入同一教学计划。这样做的目的在于提高教学效果与教学质量。近年来,通过语言测试,对新生进行编班十分盛行,不仅在初高中,而且在大学都十分盛行。以大学编班为例,一般通过英语测试,掌握不同学生的英语水平,然后分为三个级别:较差、合格、良好。这样教师在设计教学计划与教学材料时,英语水平较近的学生就会被编入同一教学计划。教师在进行教学时,因材施教,方便地进行有针对性授课。这样不仅能提高教学效率,也能提高学生学习效率。

(4)研究,即语言测试用于科研或调查。教师在教学之余,最主要的工作是科研。如为什么在相同的教学背景下,有的学生成绩优异,而有的学生成绩较差;影响学生英语成绩的因素到底有哪些?语言测试在这里的用途主要是指,通过语言测试,获得语言测试结果,进而对这一结果进行分析研究,将分析研究的数据或者结论作为科研的数据来源或者论据。从这种意义上讲,语言测试是科研的必要手段和工具。

(5)评估,即语言测试用于教学的评估。通过语言测试,可以对课程、材料以及教学方法进行反馈式评估。这是语言测试的一种特殊用途。从本质上说,这是对教学活动或者相关活动的检验,但反过来又有益于教学水平的提高。

第二节 语言测试的种类与标准

对某一事物进行分类是一个十分复杂的问题,究其原因主要有两个。第一,根据不同的标

准,可以分为不同的类型,而标准是无止境的。这就说明,分类问题,由于其标准的无止境,是一个永无止境的过程。第二,类别之间总是有或多或少的重叠部分,这一方面源于语言自身的特点,另一方面源于概念内涵与外延的杂乱交织。

因而,我们对语言测试进行分类,必须建立在以下两个方面的基础之上:第一,标准的选择必须有意义、有必要,无意义、无必要的分类在此不进行探讨;第二,分类限定在一定范围的大类之上,而对于范围较小的局部分类则不予涉及。

遵循上述两点的分类指导原则,根据不同的分类标准,我们将语言测试分为以下几个主要类型。

一、按不同的学习阶段来分

按学习阶段的不同,语言测试可以分为以下几种类型。

(一) 编班测试

编班测试(placement tests)。这里首先要明确一个问题,即翻译上的误解。英语中 placement 的含义更倾向于一种“定位”,具体到语言教学中,指的是确定学生的知识水平。因而,这一含义与汉语的“编班”有一定的区别,容易让人误解为“分班”,如果翻译为“定位”则可以避免这种误解,但语言测试研究者似乎已经习惯于这种说法,因而本书我们也沿用这一说法。编班测试更通俗的说法是“摸底考试”,一般发生在新学期开始,新生入学之初,学校组织一次摸底考试,目的不同于升学考试,不在于选拔或者达到资格,而是通过测试掌握学生的知识水平,考查学生的整体能力,因而这样的测试的设计必须全面,难易坡度明确,以拉大学生的成绩差别。通过分析学生成绩的特点,制定有针对性的教学计划,完善教学资料或教学技巧的不足。

(二) 随堂测试

随堂测试(classroom tests)。这里仍然需要纠正一个通常的误解,即将随堂测试理解为必须是一堂课结束后进行的小型测试。实际上,所谓随堂测试,强调的是学习阶段上的随堂测试,也就是说,它强调两个方面:其一,它强调时间,即发生在编班至期中、期中至期终两个阶段之间;其二,“随堂”,即测试必须发生在课堂之上,至于测试发生在本节课结束时,还是发生在下一次课开始时,亦或发生在几节课完成后的课堂测试,都是可以的。随堂测试具有以下几个特点。

(1) 随堂测试由于时间短,所以分量相对比较小。

(2) 随堂测试由于相对较随意,因而测试形式不固定,有时可以是试卷,有时是黑板板书,学生提交答案等。

(3) 由于课堂内容丰富,不同课堂之间教授材料不同,因而测试的题较多变,内容也多变。

(4) 随堂测试由于是对学生刚刚学过的知识进行测试,所以测试内容以课堂讲授内容为限,不宜过难。

(三) 期中测试

期中测试(mid-term tests)。这一类型的测试是与现实的学校学期设计有关的。一般学校会在学期到中间时期,会停课一周左右时间,进行一个正式的测试,测试的目的是对学生前半学期的学习情况进行系统诊断。这一测试也有利于弥补随堂测试的孤立、知识点单一等缺点。通过这样一个系统的测试,对已经学习的知识进行知识点之间系统的测试,学生在心理上也可以有一个系统的思考,以把握自己前半学期总的缺陷与不足。因而,设计的题目应当在这方面进行考量。

(四) 期末测试

期末测试(end-of-term tests)。这一测试与期中测试相类似,都是与现实的学校学期设计有关。但是,不同的是,期末测试考查的范围更广、考试时间更长、考查的分量更大。一般来讲,期末测试有以下三个目的。

(1) 促使学生巩固一学期已经学习的知识。

(2) 评价一学期的教学效果。

(3) 为下一学习的教学计划做准备。

由于以上目的,期末测试与前面的三种测试相比,通常更加受重视。在大部分学校,都将期末测试作为评定学生知识水平的标准,作为其他一些评比的依据,例如奖学金的评比,优秀学生中学习成绩方面的依据等。也正因为如此,学生对期末测试的重视较其他测试也较多。

二、按测试的目的或用途来分

按照测试用途或者目的的不同,测试可以分为以下四类。

(一) 成绩测试

成绩测试(attainment or achievement tests)是针对学生已经学习的知识进行考查,它一般依据一定的教学大纲,遵循一定的理论规律。上面讲到的随堂测试、期中测试以及期末测试都是成绩类型的测试。我国高考中的外语测试以及针对大学生的大学四、六级测试,严格的说,都是成绩测试。

(二) 水平测试

水平测试(proficiency tests)并不是针对学生已经学习的知识进行考查,与过去的教学内容与学习方式也没有直接的联系。它旨在考查学生的语言综合能力,例如美国的托福考试、(TOEFL)、英国的剑桥英语水平证书测试(University of Cambridge Certificate of Proficiency in English)、雅思考试(IELTS)。我国各种类型的英语职称测试、各种上岗测试等,都属于水平测试。

(三) 潜能(素质)测试

潜能(素质)测试(prognostic or aptitude tests)的目的在于把握学生学习某种语言的潜力与天赋,以明确学生应不应该学习某一语言,或者学习某种语言所要付出的努力。既然是对一种潜能的测试,那么它就不会基于某一大纲,或者事先给定某一考查范围。因而,区分潜能测试与水平测试的依据也在于这一点。

卡罗尔(J. Carroll)和萨本(S. Sapon, 1958, 1967)设计了 MLAT(Modern Language Aptitude Test),皮斯里尔(Pimsleur)(1964; 1966)设计了 PLAB(Pimsleur Language Aptitude Battery)。这是现有的著名的语言潜能测试模型。MLAT 与 PLAB 测试的内容不完全相同。具体来说,MLAT 主要考查以下四种能力。

- (1) 语音编码能力。
- (2) 语法敏感性。
- (3) 归纳式学习能力。
- (4) 机械记忆能力(后来被取消)。

而 PLAB 主要考查以下三种能力。

- (1) 言语职能。
- (2) 学习动力。
- (3) 听觉能力。

(四) 诊断测试

诊断测试(diagnostic tests)是针对成绩测试而言的,成绩测试关注的是学习成功的程度;而诊断测试,从某种意义上讲,更侧重于对失败程度的分析,通过这样的分析,以改进教学、调整教学计划、完善教学方法。因而,此类测试可以是单项,也可以是综合性的。前面讲到的编班测试就是诊断类型的测试。

需要指出的是,在现实生活中,以上四类测试的界限并不是十分清晰。有些考试同时具有两种以上测试的性质,而主要是以一种测试的特征呈现出来。另外一个特征是,针对不同专业的学生,同一测试具有不同的测试类型交集。以我国每年全国硕士研究生入学考试中的英语测试为例:由于硕士入学必须要求具备一定的英语知识水平与能力,因而必须对学生是否具备这一能力进行考查,从此种程度上讲,它是一种成绩测试;但是,如果是英语专业的学生,那么这同一测试的测试类型可能就会有一定的变化。对于英语专业学生,硕士入学考试在测试成绩的基础上,还有另外一个目的,就是测试考生的学习语言的天赋与潜能,因而从这个意义上讲,它是一种潜能(素质)测试。

三、按不同的评分方式来分

根据评分方式的不同,语言测试可以分为主观性测试(subjective tests)和客观性测试(objective tests)。

(一) 主观性测试

海宁(2001)对主观性测试的定义为:

A subjective test is said to require scoring by opinion-ated judgment, hopefully based on insight and expertise. on the part of the scorer.

主观性测试的优点表现在以下两个方面。

(1)命题成本小。

(2)对考生的能力考查较准确,学生通过猜测性获得正确答案的成分小。

同时,主观性测试也有其缺点,体现在以下三个方面。

(1)很难确定正确作答的统一标准,评分较麻烦。

(2)受阅卷人员主观因素影响较大。

(3)考查的语言现象有局限性。

(二) 客观性测试

海宁(2001)对客观性测试的定义为:

An objective test is said to be one that may be scored by comparing examinee responses with an established set of acceptable response or scoring key.

客观性测试的优点体现在以下几个方面。

(1)答案固定,评分标准较统一。

(2)阅卷成本小,节约人力、物力和时间。

(3)受阅卷人员主观因素影响小。

(4)考查的覆盖面较大。

客观性测试的缺点主要表现在以下两个方面。

(1)存在不少猜测成分,对学生的语言能力的测试不准确。因此,客观性测试越来越多受到语言测试者的反对。但是,由于社会发展逐渐向高效率、快节奏发展。客观性测试虽然存在这样的缺点,在当今社会仍然十分盛行。

(2)客观性测试的另外一个缺点在于命题难度较大。以单项选择题为例,命题者必须对题目的设计有明确的认识,包括常规、干扰项与答案项的设计。既不能使得考生通过非知识性技巧进行排除以获得正确选项,亦不能设计难度过大,超出单选题分数承载的知识量信息。

可以看出,主观性测试与客观性测试的主要区别在于,测试结果或者分数评价的标准的不同。一般认为,主观性测试需要评卷人员对考生的作答作出自己主观的判断;而客观性测试不需要评卷人员掺入主观的因素。

一般来说,在语言测试中,名词解释、简答题、论述题、作文、翻译题以及口试等都属于主观性测试;而选择题,包括单项选择题、多项选择题以及其他变式选择题都属于客观性测试。

但是,这里仍然需要明确两个问题。

首先,这里所说的主观与客观,针对的是评卷人员,更具体地说,是针对评卷人员在阅卷过

程中,对考生作答的内容是否进行主观的判断,是否经过一个自己内心主观的评价,而不是指测试本身是主观还是客观。因为,从本质上讲,无论是主观性测试,还是客观性测试,测试本身内容的来源都是命题者的主观心理,即都是一种主观过程的结果。

其次,这里的主观与客观,并不是完全意义上的主观与客观,而只是一种相对的指示性言语。我们这里讲的客观性测试,也并不是说评卷人员在评阅客观性试题时,都不能存在任何的主观判断。事实上,评卷人员在评阅试卷过程中,总是受主观思维的主导,哪怕他批阅的对象是一些客观性试题;同样地,评卷人员在批阅主观性试题时,也并不是不涉及客观的标准。实际上,他总是受到评卷标准等一些客观标准的影响。

四、按应试者的测试行为与人们实际运用中的行为的相似性不同来分

根据应试者的测试行为与人们实际运用中的行为的相似性不同,语言测试可以分为直接测试(direct tests)与间接测试(indirect tests)。

(一)直接测试

休斯(Hughes,2000)认为,直接测试要求考试直接运用被测试的技能或能力。也就是说,测试者为应试者提供近似于真实的语言运用环境,要求应试者直接进行类似于真实的语言运用,以达到测试的目的。

直接性测试侧重于考试形式以及考试内容与现实环境的近似性。比如在口试中,测试者通过近似现实的对话性测试,以考查应试者的语言口语能力;也正因为如此,直接性测试具有较明确的测试目的。

(二)间接测试

休斯(2000)认为,间接测试与人们的交际语言运用行为无关,是用非直接的方法考查应试者的语言能力。

间接性测试并不侧重考试形式与内容与现实的近似性。比如多项选择题、正误判断题。测试者根据应试者的测试成绩,推断应试者的语言能能力。

正是因为这一推断的存在,我们也可以说,直接性测试与间接性测试的分类标准也可以是应试者语言能力是直接被考查还是通过推断获得。如果是直接被考查,则是直接性测试,如果是间接地被推断,则是间接性测试。

五、按测试结果的解释方法和使用目的的不同来分

根据测试结果的解释方法和使用目的的不同,语言测试可以分为常模参照性测试(criterion-referenced tests)与标准参照性测试(norm-referenced tests)。

(一)常模参照性测试

所谓常模参照性测试是将同一次测试的结果,根据某一特定的等级序列进行排序,并据此

确定通过测试的合格分数线,高于此分的为通过考试,低于此分的为未通过测试。在此种测试中,应试者所得分数本身没有太大的意义,不能说明其具有的语言能力的高低。测试者以一组的形式出现,他们的分数确定他们在同组中的地位。学过语言学的人都知道,索绪尔在探讨词语的含义时,强调意义存在于与其他事物的比较之中,这里也是,某一特定应试者的分数的意义,存在于与同组其他应试者分数的比较之中。通常,在一个标准化的测试中,我们可以为其建立常模,即应试者在历次测试中的成绩,用标准分与平均分进行表示。然后,通过标准差的概念,确定一个特定的应试者的分数与平均分的距离,以确定其在整体分数中的位置。测试者一般根据现实的需要,设定一个通过测试的比例,比如10%。最典型的常模参考性考试就是托福考试。

(二) 准则参照性测试

准则参照性测试与常模参照性测试不同。具体来说,它测试的是应试者的特定知识或能力。它以事先设定的标准为依据。测试的目的在于检验应试者是否达到这一预先设定的标准,而对应试者彼此的分数则不问。更通俗地讲就是,应试者之间不具有竞争性,而只需满足标准的要求就可以。最典型的准则性参照标准测试是雅思考试。

除上述分类之外,语言测试还存在其他的分类方式。例如,根据测试内容方式的不同,可以将语言测试分为分离式测试(discrete-point tests)与综合性测试(integrative tests);根据测试的知识与能力的不同,可以将测试分为速度测试(speed or speeded tests)与强度测试(power tests)。对于这些分类方法,这里我们不再一一赘述。

第三节 语言测试与其他学科的关系

语言测试首先是一门综合性科学,其次是一门独立学科。

这种论断表面看是自相矛盾的,但实则具有合理性,是对语言测试学科定位的科学概括。其理由如下:

语言测试是一门跨学科的学问,它包括语言学(包括应用语言学、心理语言学、社会语言学等)、数学、教育学、教育测量学、学习论以及计算机科学等多门学科。杨惠中(1999)曾对此做过精辟的论断,他指出,语言测试从语言学、语言教学法和学习论取得科学内容,从心理测量学获得科学手段,语言测试是一门跨学科的综合科学。

那么为什么又说语言测试是一门独立学科呢?我们判断一门学科是否已经发展为一门独立的学科的标准是,或者说一般是,该学科是否已经具备自己的研究领域与研究方法。语言测试虽然是建立在众多学科基础之上的,但是随着其不断的发展,已经具备了自己的研究领域,已经拥有其独特的研究方法,因而,已经脱胎于众学科,发展为一门独立的学科。

虽然如此,语言测试与其他众学科的关系仍然是十分密切的,这里我们就对语言测试与应用语言学、心理语言学、社会语言学、学习论、数学以及计算机科学的关系进行概括性陈述。语言测试与其他学科的关系如图1-1所示。

