



北京语言大学对外汉语
教材研发中心规划项目

对外汉语教学
专业教材系列

语言测试概论

✻ 王佶旻 著



北京语言大学出版社
BEIJING LANGUAGE AND CULTURE
UNIVERSITY PRESS

对外汉语教学
专业教材系列

语言 测试 概论



北京语言大学出版社
BEIJING LANGUAGE AND CULTURE
UNIVERSITY PRESS

图书在版编目(CIP)数据

语言测试概论 / 王佑旻 著. —北京: 北京语言
大学出版社, 2010. 3

对外汉语教学专业教材系列 / 刘珣, 崔永华主编

ISBN 978-7-5619-2695-6

I. ①语… II. ①王… III. ①汉语—测试—对外汉语
教学—教学参考资料 IV. ①H195

中国版本图书馆 CIP 数据核字 (2010) 第 174153 号

书 名: 语言测试概论

责任印制: 汪学发

出版发行: 北京语言大学出版社

社 址: 北京市海淀区学院路 15 号 邮政编码: 100083

网 址: www.blcup.com

电 话: 发行部 82303650/3591/3651

编辑部 82300090

读者服务部 82303653/3908

网上订购电话 82303668

客户服务信箱 service@blcup.net

印 刷: 北京联兴盛业印刷股份有限公司

经 销: 全国新华书店

版 次: 2011 年 4 月第 1 版 2011 年 4 月第 1 次印刷

开 本: 700 毫米 × 1000 毫米 1/16 印张: 22.25

字 数: 383 千字

书 号: ISBN 978-7-5619-2695-6/H·10060

定 价: 46.00 元

凡有印装质量问题, 本社负责调换。电话: 82303590

序

今年元月初，刘珣教授写信给我，告以“对外汉语教学专业教材系列”已基本撰写修改完成，将陆续出版，嘱我为之写一总序。对外汉语教学的本科与研究生教育是我校对外汉语教育体系中非常重要的一个有机组成部分，也是我校对外汉语教育师资培养和研究生教育的重要方向。因此，我为这套教材的出版感到欢欣鼓舞，写点东西是责无旁贷的。

“对外汉语教学专业教材系列”是专门为对外汉语教学的硕士课程编写的系列教材，著作者都是北京语言大学的资深教师和青年才俊，其中有颇负盛名的资深教授，也有崭露头角的中青年学者。这支队伍是我国对外汉语教学领域的精英和领军人物，他们的学术成就可以代表我国对外汉语教学领域的最高水平。对外汉语教学是一门交叉学科，语言学、教育学、心理学、现代教育技术等学科是这个学科的基础。因此在这一书系的作者队伍中，有对外汉语教育的专家，如刘珣教授、崔永华教授、杨惠元教授；有教学理论和教材研究的专家，如张宁志副教授；有语言学专家，如张旺熹教授；有第二语言习得研究专家，如施家炜副教授；有心理学专家，如江新教授；有教育学、对外汉语教学研究专家，如姜丽萍副教授；有语言测试和语言教育专家，如郭树军副教授、杨翼副教授、王佑旻博士；有跨文化交际研究专家，如毕继万教授；有现代教育技术专家，如郑艳群教授。这些同志多年来在对外汉语教学与研究的第一线，为以对外汉语教学为专业的本科生、研究生开课，有着丰富的教学经验和研究经验。对外汉语教学涉及很多因素，教师、教材、课程、教育理念、教学模式、教学环境等等，我认为在所有的因素中人的因素是最重要的。

对外汉语教学也是一门实践性很强的学科，因此基础理论研究与应用研究同样重要，在研究生的培养方案中，两者不可偏废。这套教材包括《对外汉语教学设计》、《课堂教学理论与实践》、《对外汉语教学论》、《对外汉语学习论》、《第二语言习得概论》、《对外汉语教学评价》、《语言测试概论》、《对外汉语教材论》、《语言教学研究方法》、《跨文化交际与对外汉语教学》、《汉语多媒体教学课件设计》、《对外汉语教育学引论》（修订本）、《对外汉语教学论文选

评》(第二集)等,既有基础理论研究的内容,又有应用研究的内容。它总结了北京语言大学二十多年来从事“对外汉语”本科专业和“(对外汉语教学)课程与教学论”硕士专业的教学与研究的成果,在国内开设对外汉语教学方向或专业的学校中,这样成系统的课程和与之配套的教材还是不多见的。在当前加快汉语走向世界、大力发展对外汉语教学事业的新形势下,这套教材的出版对我国对外汉语教学学科的建设、特别是教师队伍的建设来说真如雪中送炭。我的导师陆俭明先生曾经说过:“对外汉语教学就是要研究怎么样在最短的时间里让学生掌握汉语。”这是我们最直接的目标。为了实现这一目标,我们要从汉语本身、教学方法、学习者习得过程和教学手段等方面着手进行深入细致的研究。这套教材也基本上体现了这样一种思想。

对外汉语教学作为国家和民族的事业得到了许多前辈学者的支持和鼓励,王力先生曾为这个专业题词曰“对外汉语教学是一门科学”,作为一门科学当然要有科学的范式和科学的态度。我想现在已经没有人认为“对外汉语是小儿科,只要会说中国话就能教”了。去年四月中旬,承蒙何九盈老师惠赠他老人家的新作《汉语三论》(语文出版社2007年3月出版),其中有一论叫做“全球化时代的汉语意识”,在这里他把对外汉语教学看做“一门崭新的学科”,并认为这是一门很有前途的学科。何九盈先生还在电话中特意强调:“对外汉语教学是一门独立的学科,这个学科要有国际视野,要有历史视角,要有跨学科意识,要有时代精神。”前辈学者的关怀和鼓励让我们信心倍增,也让我们感到肩上的责任重大。

一个学科的发展需要不懈的追求和持之以恒的努力。新中国的对外汉语教学已经有五十多年的历史了,曾经有许多先行者筚路蓝缕,孜孜以求。正是因为有了他们的探索和鼓吹,这个学科才有了今天的成就。我们还应该看到,这个学科还是一个比较稚嫩的学科,需要更多的人在这个园地里辛勤耕耘,需要更多地得到学界的关心和理解,需要在继承中发展、在发展中进步、在进步中成熟。

北京语言大学一直秉持“推广汉语,弘扬中华文化”的办学理念,在对外汉语教学的学科建设中,博采众长,融通诸家,不胶柱鼓瑟,不闭门造车,用科学的研究方法以聪明文思,用多样的文化活动以涵养性情。我希望我们培养的学生有理性的忧患意识,有扎实的专业基本功,有宽广的学术视野,有博大的人文情怀,有深沉的历史厚度,有高雅的审美情趣。

是为序。

崔希亮

2008年4月30日

目 录

第一章 语言测试的学科定位	1
第二章 语言测验的历史发展	8
第一节 前科学时期或传统时期	9
第二节 心理测量—结构主义时期或现代时期	17
第三节 心理语言学—社会语言学时期或后现代时期	20
第四节 国内外语言测验简介	22
第三章 语言测验的基本概念	37
第一节 语言测验的作用和目的	37
第二节 语言测验的种类	39
第三节 语言测验的质量标准	51
第四章 心理与教育测量学基础	54
第一节 测量及其相关概念	54
第二节 量表	58
第三节 测量的程序	61
第四节 统计学基础	63
第五章 语言能力	77
第一节 语言能力观点的发展演变	77
第二节 语言学理论对语言能力理论的影响	98
第三节 语言能力标准	104

第六章	测验的信度	113
第一节	信度的概念与真分数理论	113
第二节	估计信度的方法	117
第三节	影响信度的因素	124
第四节	测验的标准误	127
第七章	测验的效度	129
第一节	效度概念的发展演变	129
第二节	效度研究的种类与方法	135
第八章	测验分数	151
第一节	分数的转换	151
第二节	分数的合成	158
第三节	分数线的设定	161
第四节	测验常模	164
第五节	测验等值	170
第九章	测验的设计与开发	179
第一节	能力测验的开发	179
第二节	成绩测验的开发	196
第十章	命题技术	201
第一节	语言材料的选择与处理	201
第二节	试题编写原则及例解	204
第三节	不同类型测验的编写技巧	217
第十一章	口语测验	226
第一节	口语测验的历史发展	226
第二节	口语测验的题型与试题编写	230
第三节	国内外主要口语测验简介	235

第四节	口语测验的评分	244
第五节	口语测验表现的影响因素	253
第十二章	写作测验	258
第一节	写作能力及其认知模式	258
第二节	写作测验的开发与应用	267
第三节	国内外知名写作测验简介	271
第四节	写作测验的评分	278
第十三章	概化理论	281
第一节	概化理论的基本概念	281
第二节	概化理论的基本方法	285
第三节	概化理论的应用	288
第十四章	项目反应理论	295
第一节	项目反应理论的产生背景	295
第二节	项目反应理论的基本原理	300
第三节	项目反应理论在测验中的应用	310
第十五章	测验的反馈和研究作用	316
第一节	测验的反馈作用	316
第二节	测验的研究作用	323
参考文献		329
后记		344

第一章

语言测试的学科定位

语言测试作为一门学科，其地位近年来不断提高，一般认为，语言测试属于应用语言学，关于它的学科地位，Alan Davies 有这样一段论述：

“在过去的二十年里，语言测试经历了一个辩证的发展过程，从 1961 年 Robert Lado 写 *Language Testing: the Construction and Use of Foreign Language Tests* 时应用语言学的配角，转变为今天应用语言学中居主导地位的角色，原因在于，它使用明显有活力的方法去面对语言学习和语言应用中的主要问题。在过去的二十年中，语言测试一直在使自己从只具有实践意义向着既具有实践意义也具有理论意义的方向转变。唯其如此，它才既获得了学院派的尊敬，又开始具备了解释力。”

Davies 的话至少说明了两个问题：第一，作为一门学科，语言测试的地位越来越高；第二，语言测试是理论与实践相结合的学科。

为了了解语言测试究竟是怎样的一门学科，我们首先要讨论关于语言测试研究的几个基本问题。

一、语言测试研究的性质

在谈及语言测试研究时，我们首先要明确的是作为一种科学研究，它的基本性质是什么？科学研究可以分为两大类：经验科学和非经验科学。经验科学旨在探究、描述、说明和预言发生在我们所生活的世界上的事件。因此，经验科学的论断必须由我们经验中的事实来检验，而且仅当它们有经验证据的支持时，它们才是可以接受的。这类经验证据可以由许多不同的方式取得，如实

验、系统的观察、采访或调查、心理测试或临床检验、对文献和文物的考察等。而非经验科学，如逻辑和数学，是不需要依赖于经验证据的证明的。这也是两类科学的基本区别。

根据这个标准，语言测试应当属于经验科学。在应用层面上，语言测试的作用之一是测量人的语言能力、语言知识或语言技能。这些都是客观存在，也就是“发生在我们所生活的世界上的事件”，要对这些“事件”进行测量，其准确程度，最终是要由实践来检验的，也就是要获得“经验证据的支持”。同时在理论层面，我们的根本目的就是要对语言能力等“事件”做出描述、说明，甚至预测，而这一切，最终也需要由实践来检验，也仍然要获得经验证据的支持。

第二个需要明确的是语言测试属于自然科学还是属于人文科学（或社会科学）。从学科发展看，语言测试是在语言学、心理学和测量学的共同影响下出现的。其中，测量学是自然科学；心理学有时归入自然科学，有时归入人文科学；语言学传统上被认为是人文科学，但 19 世纪以来，语言学发展迅猛，在研究方法上也越来越多地采用自然科学方法，虽然它还没有成为精密科学，但已经比其他人文科学具有了更多的自然科学的性质。语言测试在这样几个学科的交叉地带成长起来，它的自然科学成分就比典型的人文科学多。因而，我们可以把语言测试看成一门交叉学科，它融合了人文科学的眼光和自然科学的精神。

二、语言测验的对象

语言测验是一种测量工具，任何测量工具都有特定的测量对象。秤的适用对象是具有重量或质量的物质，尺子则适用于测量在空间上具有广延性质的对象，语言测验适用的对象是人的语言能力、语言知识或语言技能。语言测验有很多种，不同的语言测验有不同的对象。旨在测量人的一般语言能力的测验，我们称作语言能力测验或语言水平测验，它的测量对象是人的一般语言能力，这是一种潜在的能力，不可以直接测量，只能通过某种媒介间接地测量。教学中使用的成绩测验，测量的对象是学生在某一课程结束时所获得的语言知识或语言能力。能力倾向测验的对象是学习者学习某一语言时将会发挥出来的潜在能力。不论是哪种语言测验，它的对象都是人的某种特性，即语言能力。

三、语言测试研究的对象

语言测试的对象和语言测试研究的对象是两个不同的概念。语言测验是我们研制出来的工具，要使这个工具有效，我们既要研究工具的适用对象（即语言测试的对象），也要研究工具本身。比如说，要造一把尺子，我们要研究这把尺子测量的东西，也要研究用什么材料做这把尺子。语言测试也一样，我们既要研究语言测试的对象——语言能力、语言知识等，也要研究测试自身。所以，语言测试研究有两个大的方面，一个是语言方面，包括对语言理论、语言知识和语言行为的研究；另一个是测量学方面，包括对测量的可靠程度、精确程度的研究。

四、语言测试研究的基本问题和基本方法

语言测试从 20 世纪 60 年代起成为一门相对独立的学科。作为一个相对独立的学科，就要有特定的研究对象，要有学科的基本问题和基本研究方法。

第一个要弄明白的就是学科的基本问题。那么，什么是学科的基本问题呢？基本问题就是贯穿于一个学科的全部历史并且推动着学科发展的那些问题。比如说，存在和意识的关系问题是哲学的基本问题，资本问题是经济学的基本问题。那么，语言测试研究的基本问题是什么呢？语言测试研究的基本问题就是语言能力问题。对所有的语言测验来说，不论是开发还是应用，一个首要问题就是对语言能力有一个明确而清晰的定义。无论是什么样的测验，都不能回避语言能力问题，只是不同的测验对语言能力的不同方面或不同层次有所侧重而已。

除了语言能力问题这个贯穿始终的基本问题以外，另一个重要的问题就是测验的有效性问题，即测验在多大程度上测到了我们想要测的东西（语言能力）。这个问题涉及对测验工具即语言测验本身的研究，语言测验归根到底是一种教育测量工具，工具的有效性是研究者们应当始终关注并充分研究的课题。

第二个要弄清楚的是语言测试研究的基本方法。前面我们说过，语言测试在学科定位上是一种交叉学科，兼有人文科学和自然科学的性质，这样的性质

使得语言测试在研究方法上既要使用定性研究也要使用定量研究。定性研究是探索性研究的主要方法之一，调研者利用定性研究来定义问题或寻找处理问题的途径。在寻找处理问题的途径时，定性研究常常用于制定假设或是确定研究中应包括的变量。因此，掌握定性研究的基本方法对研究者来说是很必要的。定量研究是要寻求将数据定量表示的方法，并要采用统计分析的形式。一般考虑进行一项新的研究项目时，定量研究之前常常都要以适当的定性研究开路。有时候定性研究也用于解释由定量分析所得的结果。两种研究方法各有特点，我们以教育学研究为例来说明定性研究与定量研究的异同。

表 1.1 定性研究与定量研究的特点与差异

	定 性 研 究	定 量 研 究
目的	了解和解释现象 对问题的复杂性作出解释	描述和预测现象 对问题的一致性进行描述
行为	微观行为 了解对象的经历和事件 用词汇和言语进行描述	宏观行为 设计指标进行测量 相关性分析
方法	从理论假设开始 有目的地选择个案 开放式访谈，参与式观察	以验证理论假设结束 随机抽样 问卷调查
工具	研究者本人	测验和测量量表或问卷 计算机统计分析
结果	内部效度检验 通过辨识进行概括 描述性写作，可包含个人偏好	效度和信度检验 推断总体进行概括 客观性统计，无个人偏好

在科学研究中我们要把两种方法有机地结合起来。那么，在研究中应该怎样正确处理二者的关系呢？我们认为，定性研究为定量研究指引方向，定量研究为定性研究提供保障。没有定性研究，定量研究就没有内容；没有定量研究，定性研究就不能落实。定性和定量，都是现代科学研究中不可缺少的研究方法。其实，定性也好，定量也好，怎么研究都要基于逻辑。定性研究可以是经验的，也可能是非经验的，而定量研究，无一例外，都是经验的。在研究中处理好定性与定量的关系是十分重要的。我们知道，定性研究的主观随意性可

以很大，定量研究恰恰能够对主观随意性有所限制，我们的结论于是就更客观一些。科学讲究要遵循可重复性原则，可重复性是判别真伪科学的操作方式（王德胜，1996），而定量研究正是保证可重复性的一种手段。所以，不应该把这二者对立起来，而应当有机地结合起来。

五、语言测试研究涉及的领域

前面说过，语言测试是多学科交叉的产物，因此语言测试的研究就不能孤立地进行，它必然要涉及以下相关学科。

（一）语言学理论

语言测试的对象是人的语言能力，测试使用的材料主要是语言材料，从某种角度讲，人的语言能力是通过对语言材料的运用表现出来的，因此，语言理论和语言能力理论对语言测试来说，是不可或缺的。从语言测试的发展历史上也可以看出这一点。在 20 世纪五六十年代，占统治地位的是美国结构主义语言学，结构主义把语言看成是自足的结构，而这个结构又可以被分析成更小的组成成分，在这种语言学背景下，分立式测验应运而生。70 年代，乔姆斯基理论蒸蒸日上，于是 Oller 和 Spolsky 等人使用测验去寻找一种“一般语言能力”。80 年代以来，社会语言学蓬勃发展，交际性测试就成了语言测试领域的热点。到目前为止，语言测试总是在一定的语言理论的影响甚至是指导下发展的。

（二）语言能力理论

语言能力理论是关于人的语言能力的获得和形成的理论，它的发展与语言理论的发展并不是同步的。我们知道，结构主义语言学的理论体系中并不涉及人的语言能力。因此在结构主义时期，语言测试必须借助心理学中行为主义的学习理论来解释语言能力的形成，否则它就无法实现。乔姆斯基理论使这种状况有了改变，从某种意义上讲，普遍语法理论是把关于语言结构的理论和关于语言能力的理论结合了起来，它既是语言学理论，又是语言能力理论。普遍语法理论和交际理论本身已经包含能力理论，以它们为基础的语言测试就无须借助另外的能力理论了。

在语言测试发展的初期，测试学家只是被动地应用现成的语言理论。结构主义说语言是由音位、词法和句法等不同的结构单位组成的，测试学家就按这种划分来开发测验，不再深究语言能力的本质。随着语言测试的发展和成熟，测试学家不仅追求测试的精确可靠，而且也追求测试的理论解释力，正如 Bachman 所说：“虽然很多外语或第二语言的水平测试仍然沿用 Lado 和 Carroll 的语言技能加语言成分的框架，但许多测试学家对语言能力已有更深刻的理解。” Oller 提出的“语用期望语法”，Bachman 等人提出的“交际性语言水平”和“语言交际能力”等概念，都是测试学家对语言能力的探索。在对语言能力的探索中，近年来大家有一个共同的想法，即企图寻找一个关于语言能力精确的有经验基础的定义，在这样一个定义的基础上，我们可以开发出一个“通用量表”，这个通用量表不受语境、水平、语种等条件的限制。Bachman 和 Clark 还对这种量表的优点作了如下描述：

这样的量表具有明显的优点，根据它开发的测验可作为定义和测量语言能力的标准，而这将不依赖特定的语言、语境和语域。建立在这个量表上的测试分数于是可以做跨语言和跨语境的比较。

可见，这样一个愿望，就不仅仅是满足于测到语言能力，而是向揭示语言能力的本质和性质进军了。虽然目前还没有严格意义上的通用量表，但欧共体研究和制定的“欧洲语言能力共同框架（Common European Framework, CEF）”已经开始在这一领域迈出了坚实的步伐。

（三）语言习得和语言教学理论

语言测试、语言习得和语言教学同属于应用语言学学科，它们有很多相同的研究特点。这些相同的特点已经引起了学者们的关注，1998 年 Bachman 和 Cohen 专门编纂专辑来讨论第二语言习得与测试研究的接口问题。他们指出，语言习得和语言测试研究所关心的问题是共通的，即如何描述和解释学习者语言行为的差异情况。学习者在习得过程中表现出来的语言行为情况各异，在测试任务中所表现出来的测试行为也各不相同，如何寻求导致这种情况的原因是习得研究者和测试研究者的共同任务。

实际上，在语言习得、语言教学研究和语言测试研究之间，有一种本质上互补与互动的关系，语言测试既为语言习得和语言教学研究服务，它自己也从语言习得和语言教学的研究中受益。语言测试可以提供关于学习和教学效果的

信息，语言教师经常用测试来诊断学生的学习情况，语言测试也常被用于评价课程设置、教学大纲、教学体系和教学方法。另一方面，语言习得研究和语言教学的实践的成果，也可用于设计和开发语言测验提供有用的材料。例如，语言习得顺序的研究对确定语言能力水平就很有参考价值。所以说，语言测试不是在真空中发展的，我们对语言习得和教学过程的深入理解，可以促进语言测试的发展，而语言测试对语言习得和教学的研究来说，也是一个既有实践意义又有理论洞察力的工具。

相关学科的相互借鉴是学科繁荣的基础，语言测试研究需要借鉴语言习得和教学成果，语言测试的研究者也需要具备这些方面的理论和知识。

（四）测量学理论

我们说过，语言测试是一种教育或心理测量，因而语言测试研究必然涉及对测量学的研究。语言测试所涉及的统计、测量理论大致有相关分析、方差分析、概化理论、项目反应理论、假设检验，等等，这些理论在测试中都有直接的应用。对于开发和研究语言测验来说，首先是能够正确地理解、使用 and 解释这些理论，因为这些理论是测试的工具基础，其基本作用是保证我们的测试可靠。如果我们误用了测量学的概念和理论，测试的可靠性和有效性就没有了保障。在语言测试的研究和实践过程中，我们也有可能改进测量理论，这当然是非常高的要求，但如果语言测试学家能够在工作中对测量理论作出一两项改进，那将是更大的贡献。这也就是我们常说的，测试研究不应仅仅停留在应用研究的层面上，而应该为理论作出贡献。

以上我们探讨了语言测试研究的学科性质和学科地位，研究的基本问题、基本方法和相关研究领域等重要问题。在学习语言测试之初就明确这些问题，对于我们今后的学习和研究是很有帮助的。

思考题：

1. 语言测试是怎样的一门学科？
2. 定量研究和定性研究在语言测试研究中分别发挥什么作用？
3. 语言测试研究的基本问题是什么？

第二章

语言测验的历史发展

中国是最早使用考试的国家，考试在我国至少有两千年的历史。西周时期就有用考试选拔人才的事，所以《礼记·射义》说：“是故古者天子，以射选诸侯、卿、大夫、士。射者，男子之事也。”西周时期的国学建立了定期考试的制度，隔一年考一次，《礼记·学记》说，“一年视离经辨志，三年视敬业乐群，五年视博席亲师，七年视论学取友，是谓小成”，“九年视知类通达，强立而不返，是谓大成”。

汉代取士有“对策”、“射策”之制。“对策”是以政事、经义等设问，把问题写在简策上，让应试者对答。“射策”由主试者出试题，写在简策上，分甲乙科，列置案上，应试者随意取答，主试者按题目难易和所答内容定优劣。刘勰《文心雕龙·议对》：“又对策者，应诏而陈政也；射策者，探事而献说也。言中理准，譬射侯中的。二名虽殊，即议之别体也……对策者，以第一登庸；射策者，以甲科入仕。”隋炀帝大业二年（公元 606 年）设进士科，开始用科举选拔官员。从此，科举制度在中国一直延续了 1300 多年，直到清光绪三十一年（1905 年）才废止。

国外和语言测验有关的最早的记载，大概就是《圣经·旧约全书·士师记》里说的一件事：基利德人（Gileadite）在约旦河边巡逻，遇到有人接近，就让他说“shibboleth”这个词，以此鉴别他是否是刚被打败的依弗列姆人（Ephraimite），因为依弗列姆人发不出“sh”这个音。根据这个测验，42000 人当场被杀死。从技术上讲，这是一个只有一个题目的口语语音测验，只是在测验中完全没有教育因素，而是出于急功近利的政治目的。在我们看来，这算不上是什么语言测验（Spolsky, 1995）。

传统的考试制度——科举考试在 20 世纪初消亡，也正是在这个时候，现代考试和考试理论在西方出现。在以后的近一百年的时间里，考试的理论和实践在西方有了长足的进展，以致当代中国的考试不得不学习国外的经验，我们的考试研究，也不得不较多地参考西方的文献。我们现在所说的考试研究，主要是指 20 世纪发展起来的客观性考试。

关于现代语言测验理论和实践的分期，不同的人有不同的分法。我们认同以色列著名应用语言学家伯纳德·斯波斯基（Bernard Spolsky）的观点，把整个语言测验和测验理论的发展分为三个时期：前科学时期或传统（traditional）时期、心理测量—结构主义（psychometric-structuralist）时期或现代（modern）时期、心理语言学—社会语言学（psycholinguistic-sociolinguistic）时期或后现代（postmodern）时期。

第一节 前科学时期或传统时期

前科学或传统时期是从考试制度的确立开始一直到 20 世纪 50 年代。

一、考试制度的确立

前科学时期是客观性考试的萌芽时期。西方的考试制度是传教士从中国带去的。第一个详细介绍中国考试制度的西方学者是 16 世纪末意大利传教士利玛窦（Matteo Ricci, 1552-1610）。通过利玛窦的介绍，西方人知道了考试这件事。孙中山先生曾经说：“考试制度在英国实行最早，美国实行考试不过是二三十年，现在各国的考试制度，差不多都是学英国的。穷流溯源，英国的考试制度原来是从我们中国学去的。所以中国的考试制度就是世界中最古最好的制度。”

为了更好地了解考试制度在西方特别是在欧洲逐步确立的过程，我们来看一下这项制度在法国、德国和英国的发展历程。

在法国，先是教会学校建立了考试制度，然后，整个考试体系从天主教学校推广到普通学校。从 1747 年开始，巴黎的公立中学设立了会考，为的是奖