



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

汉语教学评价

✿ 杨翼 著



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

清华大学出版社
Tsinghua University Press

英语数学评价

王德林 著

ISBN 7-302-10000-0

清华大学出版社

对外汉语教学
专业教材系列

汉语教学评价

※杨翼 著



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

图书在版编目 (CIP) 数据

汉语教学评价 / 杨翼著. - 北京: 北京语言大学出版社,
2008.11

ISBN 978 - 7 - 5619 - 2238 - 5

I. 汉… II. 杨… III. 对外汉语教学 - 教学评议
IV. H195

中国版本图书馆 CIP 数据核字 (2008) 第 179190 号

书 名: 汉语教学评价

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 82303647

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: 北京中科印刷有限公司

经 销: 全国新华书店

版 次: 2008 年 11 月第 1 版 2008 年 11 月第 1 次印刷

开 本: 710 毫米 × 1000 毫米 1/16 印张: 20.5

字 数: 357 千字 印数: 1—2000

书 号: ISBN 978 - 7 - 5619 - 2238 - 5/H·08227

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、王佶旻博士；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

前言

本书的名字叫做“汉语教学评价”，一谈到评价，我们不少教师就会想到上级领导或教务部门检查教师的教学，就如同一谈到测试，就会想到考学生一样。其实，这样理解是不全面的。那么，什么是“汉语教学评价”？评价与汉语作为第二语言的教学有什么关系？为什么要进行“汉语教学评价”？什么时候评价？我们的评价对象有哪些？如何进行“汉语教学评价”？这些问题是对外汉语教学工作感兴趣的，也是本书想着重阐述的一系列理论和实践问题。

在以往的汉语教学评价过程中，我们常常用测试来搜集信息，因此，不少人都以为评价就是“测试”，或者有时也喜欢称之为“考试”“测验”。然而，它们虽然跟评价关系很密切，但二者又有所不同。比如，评价是综合性活动，需要经历若干阶段和步骤，而测试却是单一性活动，作为一种测量工具，它只是评价的某一阶段中的某一步骤。评价的方式也是多种多样的，而测试只是其中的一种方式。因此，在论述汉语教学评价研究的一系列问题时，本书有时可能会涉及测试，但又从本质上区别于单纯的“语言测试”，重点在于全面探讨汉语教学评价的理论与实践、介绍该领域的研究现状与发展趋势等。

汉语作为第二语言的教学既指在中国进行的对外国人的汉语教学，也指世界各地的汉语教学，还指对中国国内少数民族的汉语教学。我们把这种教学简称为汉语教学。“汉语教学”这一概念不但可以简明扼要地涵盖上述三种教学情况，而且还便于与汉语作为第一语言的教学（即语文教学）相区别。因此，汉语教学评价是基于汉语作为第二语言教学的评价，它不但有教育评价方面的渊源，还跟语言教学理论、语言学习理论以及某些哲学文化思潮有着千丝万缕的联系，只有弄清楚它们之间的关系，才能很好地理解汉语教学评价的原理，很好地实施汉语教学评价实践的程序，因此，本书还将讨论它们之间的关系及其对教学评价的意义。

汉语教学评价又是一项理论与实践并重的技术，我们既需要领悟相关理论，又需要运用这些理论来指导我们的评价实践，解决我们汉语教学中遇到的一系列实际问题。比如，如何评价汉语学习者的学业成就，如何评价汉语教师的教学，如何评价汉语教材，等等。因此，本书还提供了一些具体的操作案例。

教学评价在第二语言教学研究领域占有重要的地位，近二十多年来国外在这方面的研究取得了较大成就，国内外语教学界也有一些成果，但汉语教学界的有关研究成果还很少。本书虽然进行了一些探索，但研究成果还不够成熟。另外，本人的研究水平有限，错谬之处在所难免，恳望专家和读者不吝指教。

杨 翼

2008年5月于北京

目 录

第一章 汉语教学评价概论	1
第一节 汉语教学评价的基本概念	1
第二节 汉语教学评价的主要功能	11
第三节 汉语教学评价的主要类型	20
第二章 汉语教学评价的理论基础	29
第一节 教育评价理论与汉语教学评价	29
第二节 语言学及其教学理论与汉语教学评价	38
第三节 当代新型评价理念溯源及启示	48
第三章 汉语教学评价的源起和发展	56
第一节 汉语教学评价的源流与萌芽	56
第二节 汉语教学评价的现状与趋势	66
第三节 汉语教学评价体系的构建	84
第四章 汉语教学评价的一般过程	94
第一节 准备汉语教学评价	94
第二节 收集与解释教学评价信息	103
第三节 作出决策及反馈评价信息	123
第五章 汉语教学评价的形式与标准	141
第一节 汉语教学评价的指标体系	141
第二节 汉语教学评价的概括性问题	160
第三节 汉语教学评价标准的制定	162

第六章 汉语教学评价信息的收集	177
第一节 用测试法收集信息	177
第二节 用调查法收集信息	193
第三节 用观察法收集信息	207
第七章 汉语教学评价信息的质量	225
第一节 汉语教学评价信息的信度	225
第二节 汉语教学评价信息的效度	230
第三节 收集汉语教学评价信息的可行性	237
第八章 汉语学习者评价	242
第一节 汉语学业成就评价的特点	242
第二节 汉语学业成就评价的内容	250
第三节 汉语学业成就评价的方法	255
第九章 汉语教师评价	268
第一节 汉语教师评价的目的与类型	268
第二节 汉语教师评价的形式与内容	270
第三节 汉语教师评价的主要方法	280
第十章 汉语教材评价	287
第一节 汉语教材评价的基本问题	287
第二节 汉语教材评价的程序与方法	290
第三节 汉语教材评价的指标体系	302

汉语教学评价概论

汉语教学从本质上来说是一种第二语言教学，因此，汉语作为第二语言教学的评价研究的性质也就属于第二语言教学评价。本章的写作目的是让读者对汉语教学评价的全貌有个初步的了解。

我们在本章将讨论以下几个问题：什么是汉语作为第二语言教学的评价？为什么要进行汉语教学评价？汉语教学评价的主要功能是什么？汉语教学评价的主要类型有哪些？其中有些问题我们在后面的章节里还会进一步讨论。

第一节 汉语教学评价的基本概念

在讨论什么是汉语教学评价、为什么要进行汉语教学评价、怎样评价之前，我们需要先了解一些基本概念，弄清它们之间的关系，澄清一些模糊认识，以便使我们的讨论在明确的界定范围内进行。需要指出的是，在教育评价领域，对“评价”“评估”“测量”“考试”这样一些术语的定义也有分歧。我们将主要结合汉语教学评价的实际情况，同时参考某些有一定共识的说法，给出自己的定义。

一、评价与评估

在汉语中，“评价”是评定价值的意思。在生活中，人们常常在无意识地进行各种各样的评价。比如，在商店买东西，要评价货物；在餐厅吃饭，要评

价饭菜；在工作中，要评价工作；在人际交往中，要评价人物，等等。从教育评价的角度来说，广义的评价是用系统、科学的方法全面地收集和分析信息，从而对教育的价值作出判断的过程；狭义的评价是依据教育目标对学生经教育后所产生的行为变化作出价值判断，并为改善和优化教育途径提供依据的过程。（王孝玲，1999）

我国教育界常把“evaluation”译为“评价”，把 assessment 译为“评估”。Alan Davies 等（2002）的 *Dictionary of Language Testing* 一书在国际语言测试领域有着广泛的影响，他们在该书中，对语言教学中的“evaluation”和“assessment”给出了明确的定义：

评价是为了作出决策而系统地收集信息。就一个语言教学项目进行评价是为了向该项目的涉及者（比如赞助人、管理者、教师或父母）提供信息，以便对项目的未来作出决策。这时，教学的有效性处于调查之中，语言测试被频繁地作为评价的组成部分来使用，有代表性的前测和后测或成绩测试有时将会伴随着一群控制对象来进行。

广义的评估是经常与测试一起使用的一个术语，但也更广泛地包含收集语言数据的意思，包括测试数据以及为评价目的利用面试、个案研究、调查问卷、观察技术这样一些工具。狭义的评估指不涉及测试的评估程序。评估的可能是一个学习者或一群学习者的语言能力，或者是实施语言教学的效果。更广泛地说，评估属于一种评价，它通过学习者评估和整个教学实施的评估（包括课程、方法、材料、资源、计划和教员）来实现。

从上述定义可以看出，“评估”类属于“评价”，它们在概念含义上非常接近。因此，在教育文献资料中二者常被人们交替使用，并无严格界限。

Alan Davies 等对“评价”的定义虽然很详细，也很清楚，但还不够简明扼要，而我们为了在后面开展更进一步的讨论，需要一个简明扼要的定义，以使大家能对“评价”这一概念产生清晰难忘的印象，这一点对大家理解本书的内容很关键，因此，我们把语言教学的“评价”的定义概括为：

评价是评价者为了作出决策而系统地收集信息，并对语言教学¹的任意元

1 这里的“教学”指教师的“教”和学生的“学”。

素的价值作出判断的过程。

我国教育界实际使用这两个术语时，在不同的范围和场合有不同的习惯用法。比如，在我国正式出版的学术著作中称为“评价”的占绝大多数；在我国教育主管部门颁布的文件中称为“评估”的占多数。可见，二者也没有严格的使用界限。

在汉语教学界，有不少教师习惯于使用“评估”这个术语，甚至称“教学评估”的人比称“教学评价”的人还要普遍。这是为什么呢？原因是上级领导或教务部门检查汉语教师的教学情况时，常常采用“评估”的说法，时间一长，人们都沿袭这种说法。几十年来，在汉语教学界，人们不仅普遍使用“评估”这个术语，而且还会把它跟上级领导或教务部门检查汉语教师的教学联系在一起，尽管这样理解是不准确的。

综合上述情况，结合本书主要想探讨的问题，我们倾向于使用“评价”这一术语，在本书中也将一直采用这一术语，而不在“评价”与“评估”之间随意变换，以保持全书术语的前后一致。在我们确定使用“评价”这一术语后，还需要进一步定义“汉语教学评价”，这个问题我们将在本节的第五个议题中讨论。

二、评价与测量

前面我们说过，评价是评价者为了作出决策而系统地收集信息，并对语言教学¹的任意元素的价值作出判断的过程。而测量则是依据一定规则，对事物的属性用数字进行量化描述的过程。（王孝玲，1999）

评价和测量既有联系又有区别。

测量所回答的是“程度”问题，它本质上是事实判断，而评价实质上是价值判断。但事实判断又是价值判断的基础，因为只有弄清了事实才能作出合理的价值判断。由此可见，测量是评价的工具。

测量强调数量化和客观性，一般要尽可能减少主观因素的影响。而评价则强调价值观和科学性，充分重视分析和判断。但测量的结果又是评价的重要依

1 这里的“教学”指教师的“教”和学生的“学”。

据，尽管它不是唯一依据。从这种意义上来说，评价又是测量的深化和发展。

测量是单一性活动。比如，某个留学生在中级汉语综合课第一学期期末考试中得了 75 分，全年级 60% 的人低于他的得分。对于测量者来说，他已经用考试分数描述了该考生当前的汉语水平，完成了他的测量任务。而评价则是综合性活动，它需要依据更多的信息资料、使用更多的方法来进行。比如，同样是这个留学生，对于评价者来说，只描述他当前的汉语水平还不够，还应该考查该生第一学期的期中成绩。如该生期中中级汉语综合课考试 85 分，全年级 78% 的人低于他的得分，而他的期末考试 65 分，全年级 30% 的人低于他的得分，这时评价者就需要找出该生成绩下降的原因，分析他的汉语技能缺陷，提出补救措施等。这个例子说明，评价者不仅要描述当前结果，还要考察其发展过程，诊断其问题所在，并提出补救措施。

总之，测量限于对被评价对象（如学生）的定量描述，结果总是用数字来体现（如，玛丽答对了 30 道阅读题中的 25 道）。它既不包括定性描述，也不含有对所得结果的价值判断。而评价则不同，它可以包括对学生的定量描述（测量）和定性描述（非测量）两个方面，此外，评价还总是包含对结果的价值判断。因此，评价是一个更全面的、涵盖面更广的术语。

三、评价与测试、考试

汉语中的“测验”和“测试”在英语中都与“test”对应，而“test”在译为汉语时，有人喜欢译为“测验”，有人喜欢译为“测试”。一般来说，从事语言测试研究工作的专业人员，在学术著作中使用“测试”这一术语更为普遍，比如“语言测试研究”“英语测试理论与实践”等。另外，从事科技工作的人使用“测试”这一术语也十分常见，比如“测试仪器的性能”“测试软件的报告”等。学校的教师和学生使用“测验”这一术语更为普遍，比如“我们今天有个小测验”“下星期有语法测验”等。因为测验和测试在概念内涵上非常接近，几乎没有使用界限，所以常常有人交替使用二者。在教育评价领域，它们都是测量的工具，能引起人的代表性行为，以便对人的行为特性或心理特性进行测量与评价。因此，我们把它们视为同一概念，并这样来定义汉语教学中的“测试”：

汉语测试(或测验)是汉语教学评价中收集信息的一种形式,通常是由一组要求在固定时间内完成的汉语题目组成,并在相同的条件下对所有应试的学生施测。

评价与测试和考试之间的关系如何呢?

在汉语教学评价过程中,我们常常使用测试这个工具来收集信息。测试对汉语教学评价的信息收集有很大帮助,但是测试的作用相对来说也是有限的,因为它只能告诉我们汉语学习者成绩的某些方面,而不能告诉我们许多其他因素,而这些其他因素在汉语教学评价中却常常需要考虑。因此,在汉语教学评价中,也需要一些通过其他途径来收集信息的方法。例如,通过在常规汉语课堂教学中观察学生的汉语表现来收集信息,通过学生汉语讨论会中的个体发言或登录学生的汉语博客来收集信息等。所有这一切都能获得学生汉语学习效果和教师汉语教学效果的重要信息。甚至来自学校的学生档案记录、学生父母提供的信息等,也都是信息的重要来源。由于评价所要求的信息的多样性和广泛性,这些信息资源同样可以如同测试那样为汉语学习者作出某个方面的鉴定。

从汉语教学评价的狭义范围来说,考试(examination)是汉语教学评价中收集信息的一种较为严格的特定形式,它通常是由一组要求在固定时间内完成的汉语题目组成,并在相同的条件下对所有应试的学生施测。从这种意义上来说,考试也是测试。但是,从观念上来说,学生和教师认为“考试”在形式上更正规,在程序上更严格,在规模和题量上更大;而测试较为随意、松散,题量、规模较小。其原因是,考试往往是根据汉语教学的内容和目标,严格选择有代表性的内容与题目,并以较大的内容覆盖面,对学生的汉语技能和汉语知识进行测量与评价;而测试则不一定那么严格,测试的内容覆盖面可大可小。比如我们习惯这样说:从7月1号到10号是全校的期末考试时间。这时我们会使用“考试”这个术语。同时,我们又习惯说:这周我们有个单元小测验。这时我们会使用“测验(即测试)”这个术语。这说明人们对“考试”和“测试”的理解和看法都带有浓厚的约定俗成的色彩。因此,我们也可以更准确地说,严格意义上的特定的测试就是考试。

总的来看,汉语测试的概念比汉语考试的概念内涵更宽泛,汉语考试是汉语测试的一种特定形式,但它们又都是汉语教学评价的一种工具。在某些情况下,汉语测试、汉语考试活动本身就是某种汉语教学评价活动的一个组成部分。

四、评价、测量、测试和考试之间的关系

前面我们分别介绍了评价、测量、测试、考试这些基本概念，但是它们在概念含义上纵横交错，可能会使读者产生扑朔迷离之感。我们在这里借鉴 Bachman & Palmer (1997:23) 中的一张图，并略加改造来形象、直观地描述一下它们之间的关系，以帮助读者更好地理解这些概念和含义。下图共有 4 个圆圈，分别表示评价、测量、测试和考试。

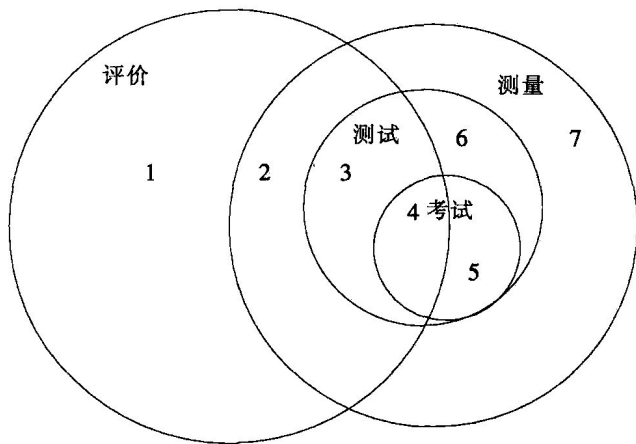


图 1-1 评价、测量、测试和考试之间的关系

从上图可以看出，圆圈覆盖的区域有 7 个，我们可以分别标上数字 1、2、3、4、5、6、7。根据 Bachman & Palmer 的观点，这 7 个区域可以代表 7 种不同的情况（高兰生，2002）：

1. 无测量、测试和考试的评价。

这种评价是定性地描述学生的行为表现，比如，用于诊断学习中的问题。特点是不通过测量，也没有测试和考试，不用数量陈述。

2. 通过无测试和考试的测量进行的评价。

比如教师通过平时的课堂观察，按优劣程度给学生排序，并给出分数或分出等级，以便作出评价。这时即使用了分数或等级来陈述评价结果，但这些分

数和等级也不是通过测试和考试得到的。

3. 通过有测试的测量进行的评价。

4. 通过有考试的测量进行的评价。

5. 通过考试进行的测量。

6. 有测试无考试的测量。

这种测量的目的不是为了评价，而是为了进行科学研究。例如，进行汉语作为第二语言的习得研究。

7. 无测试无考试的测量。

这种测量没有测试，也没有考试，不是为了评价。例如，为了研究的需要，根据母语的类别给每个学生分配一个号码，再计算出不同母语类别的人数分布。

对上述7种不同情况的描述提示我们：

(1) 并不是所有的评价都会涉及测量、测试和考试。

(2) 并不是所有的测量都是测试和考试。

(3) 也不是所有的测试和考试都是评价。

(4) 虽然测试、考试、测量都是评价的工具，但评价既可以借助定量描述型的测量工具来进行，也可以借助定性描述型的非测量工具来进行。

(5) 测试包含考试，但它们又都是评价或测量的一种工具。

(6) 汉语教学评价多数在3区和4区进行，有时候在1区进行，少数在2区进行。

五、汉语教学评价

我们在前面所进行的一系列定义和阐释都是为了更好地回答一个关键性问题——什么是汉语教学评价？对此我们必须清楚地给出定义，因为本书关于汉语教学评价的进一步探讨，都是在这个定义下进行的。同时，明确汉语教学评价的对象，充分认识它跟汉语教学的紧密联系，对我们理解“什么是汉语教学评价”这个问题也有很大的帮助。