

华文教学研究丛书

总主编 郭 熙  
副总主编 李嘉郁

# 华文测试与教学评估研究

主 编 黄霆玮 副主编 王 芳



商  
籍

华文教学研究丛书

总主编 郭 熙  
副总主编 李嘉郁

# 华文测试与教学评估研究

主 编 黄霆玮  
副主编 王 芳



商  
德  
中  
书  
馆

图书在版编目(CIP)数据

华文测试与教学评估研究/黄霆玮主编. —北京:  
商务印书馆, 2016

(华文教学研究丛书)

ISBN 978-7-100-11343-4

I. ①华… II. ①黄… III. ①汉语—对外汉语  
教学—教育评估—研究 IV. ①H195.3

中国版本图书馆 CIP 数据核字(2015)第 127291 号

所有权利保留。

未经许可,不得以任何方式使用。

华文测试与教学评估研究

主 编 黄霆玮

副主编 王 芳

---

商 务 印 书 馆 出 版

(北京王府井大街 36 号 邮政编码 100710)

商 务 印 书 馆 发 行

北 京 冠 中 印 刷 厂 印 刷

ISBN 978-7-100-11343-4

---

2016 年 9 月第 1 版

开本 880×1230 1/32

2016 年 9 月北京第 1 次印刷

印张 8 1/4

定价:30.00 元

总主编	郭 熙				
副总主编	李嘉郁				
主 编	黄霆玮				
副主编	王 芳				
编 者	黄霆玮	王 芳	原 鑫		
作 者	陈 宏	何子煌	胡林生	柯华葳	
	梁荣基	罗庆铭	宋如瑜	苏启祯	
	王汉卫	吴福焕	吴英成	谢泽文	
	张郁雯	赵守辉	周清海	祝新华	

# 前 言

郭 熙

一直以来,汉语教学被简单地两分:一是国内的语文教学,二是对外国或外族人的第二语言教学。近年来,华文教学被逐步地单列出来,无论是教学还是研究方面都取得了不小的进展。华文教学以华侨华人子女为对象,以语言文化传承教育为目标,旨在比较全面地培养这些“祖语生”使用母语文的能力。在世界各地非母语环境下开展母语教学,有各自的特殊情况,因此,华文工作者在勤力教学的同时,也围绕华文教学开展了大量的研究工作,发表了一大批研究成果。这些成果内容丰富,涵盖面广,它们源于不同的华语区,也发表在不同的华语区,表现出不同的风格和特点。它们有的重在理论,有的重在实践,针对性、目的性很强。

另一方面,随着华文教学事业的发展和队伍的扩大,对华文教学研究成果的借鉴和吸收也成了华文教育工作者的迫切需求。但是这些文章发表在不同时期和地点,查找起来比较困难。我们觉得有必要把相关的文献搜集起来,进行编辑整理,于是就有了这套《华文教学研究》丛书。

这套丛书共 10 册,分别为:《华文教学理论研究》《世界华文教育现状研究》《华文教学方法研究》《华文听说教学研究》《华文



读写教学研究》《中华文化教学研究》《华文教师培养与培训研究》《华文教材编写研究》《华语习得研究》《华文测试与教学评估研究》。

这次编选的总体思路是将相关论文的主要观点、研究方法、材料和结论集中起来,并尽可能地形成一定的框架,篇幅上则对原作尽量压缩,以便多选录一些。为了使读者对各册所涉领域的研究有一个总体的认识,我们请各册的主编分别进行了综述,希望这些综述能够反映该领域的实际。

选文尽可能考虑了不同观点的吸纳,但种种原因,未必能都照顾到,盼读者和作者谅解。

文献来自各个华语区,术语、行文风格和语言习惯都有差异。考虑到反映各地的实际面貌,除了一些技术性更动,语言方面大体保持原貌,必要时也加了一些注释。

这是我们第一次对世界各地华文教学研究进行梳理和总结,加之经验不足,时间紧迫,材料搜集困难,丛书一定有许多问题和不足。我们期待着读者的批评指正,更期待着更多的人关心华文教学研究。

# 综 述

黄霆玮

测试是语言教学过程中的一个环节,有教学就有测试,在语言教学的过程中,随时都会遇到各种各样的测试。在教学过程开始前,人们常常通过测试来安置学生,给学生分班或者确定课程难度等;教学过程中,教师使用测试检验教学的效果,根据测试的结果及时查缺补漏;教学结束后,测试可以用来选拔人才、评价学习者的水平等。测试一直贯穿于语言教学过程的始终。

除了对语言教学工作者有用以外,语言测试还可以为研究者提供信息。语言测验是一种测量工具,是语言研究和语言教学研究的一个重要手段。<sup>①</sup> 每个测试都是基于某种理论的,人们根据各种不同的理论开发出不同的测试,对实际测试结果的分析能反映这种理论的可靠程度。

在华文教学领域,有关测试的研究起步较晚,但是华文教师和研究者们一直在自觉或非自觉地使用华文测试,他们通过测试了解学生的语言能力水平、评估教学情况。在华文测试实践的推动下,关于华文测试的理论探讨逐渐深入。近年来专门针对华侨华人的大型标准化测试已经出现,华文测试正逐步走向

---

<sup>①</sup> 郭树军《汉语作为第二语言的测试》,赵金铭主编《对外汉语教学概论》,商务印书馆,2010年。

规模化和科学化。

华文测试和 HSK 一样,都是汉语测试,区别首先在于针对的被试群体不同,华文测试的被试群体是海外华侨华人,HSK 的被试群体是把汉语作为第二语言学习的外国学生。汉语对于海外华侨华人是民族语,华文教育是海外华侨华人的母语或民族语教育。<sup>①</sup> 这一特性决定了华文测试在测试内容、形式以及难度上都有所不同。本书所收文章都是以华文测试为研究对象的,包括大型标准化测试,也包括教师自编的课堂测试,还有一些把测试用于教学评估、语言政策评估的文章。

### 一 华文测试在海外的的发展概况

虽然华文测试领域中还没有像 HSK 一样有广泛影响力的考试产品,但不论在国内还是在海外华侨华人聚居的地方,人们已经逐渐认识到华文测试的重要性,很多机构都在致力于华文测试的研发和开展。

陈玉兰(2009)介绍了印尼“华语文能力考试”的情况。<sup>②</sup> 这个考试的前身是印尼中文统一会考,中文统一会考是印尼“国家级外语统一会考”之一。2008 年印尼国家教育部提出了一系列会考改革方案,包括指定各个外语统筹处研制“外语能力考试”,统筹处可编写外语教材,成立语文能力考试机构,制定各个等级的最低标准教学大纲、考试大纲,并由语文能力考试机构颁发语

---

① 郭熙《论海外华文教学的性质和地位》,郭熙《华语研究录》,商务印书馆,2012 年。

② 陈玉兰《印尼基础华语文能力考试简介》,《第八届东南亚华文教学研讨会论文集》,菲律宾华教中心,2009 年。



文能力考试证书。改革后的“华语文能力考试”，共分为四个等级：基础、初级、中级和高级。考试的范围由印尼国家教育部校外教育司中文统筹处颁布的能力标准、基本标准划定。该改革的亮点之一是每个级别均设置了口语考试，进一步体现对目的语掌握能力的全技能（听、说、读、写）的要求，该考试已于2009年开始试考。

除印尼以外，菲律宾华教中心也设计了等级统一考试——华语等级测验（HDC），该考试的主要目的是检测华校学生学习和教师教学效果。HDC是一种成绩等级测试，测试以菲律宾华校的华语课本册数为蓝本开发，共分12个等级。HDC的每个题目都对应着教材的一个或几个知识点。测试结果可以为教学单位和教师提供诊断信息，例如学生在哪些方面已经达到教学要求，哪些方面还要加强训练。HDC已经建成了题库，每一份试卷都由题库电子文本自动生成，并且标有难度、区分度、效度等测试学指标。<sup>①</sup>有学者探讨了菲律宾华校在教学过程中使用的各种测试，提出要重视测试对学习效果的诊断作用以及对教学过程的评估作用。文章还明确华文测试的特殊性——文化教学的重要性，认为华语教学承载着“文化教学”使命，要努力保持并发扬学生的中华文化情结，培养并丰富他们的中华文化人格，促进中华文化与菲律宾文化的交流融合。<sup>②</sup>

在新移民较多的澳大利亚和美国，也有汉语考试是专门针

---

① 郝爱廷《菲律宾华语测试的初期阶段》，《第八届东南亚华文教学研讨会论文集》，菲律宾华教中心，2009年。

② 苏山《浅谈华语教学考试——明确华语教学目标、规范华语教学考试》，《第八届东南亚华文教学研讨会论文集》，菲律宾华教中心，2009年。

对海外华侨华人的。早在 20 世纪 70 年代中期,澳大利亚维多利亚州(首府墨尔本)教育部就把中文纳入了大学入学考试体系,是西方国家以中文成绩作为大学入学标准的创始者。目前维州的中文会考分为三种,有母语组、非母语组和第二语言高级组。第二语言高级组即为在海外出生并居住至 12 岁的华裔学生。维州会考是一种选拔性测试,已有较长时间的历史,测试内容包括校内和校外两部分互相补充,考查内容完备,题型设计也很新颖。口试部分采用角色扮演和面试等形式,阅读、听力采用读后说、听后说等,会考也很重视中国传统文化,写作及口试部分都有“有关中国语言文化的特定题目”。<sup>①</sup>

2006 年,美国大学理事会在全美设置大学先修(简称 AP)中文课程,从 2007 年开始举办 AP 中文考试。AP 中文课程程度相当于大学二年级第二学期(或是 250 学时)的中文课程,一般认为考试题目很有挑战性,但由于参加考试的学生大多为海外华侨华人,所以绝大部分考生成绩相当高。AP 中文课程强调中国文化与语言相结合,测试的内容也多与文化有关,部分题目和中国的节日及文化常识相关。<sup>②</sup>

众多专门针对海外华侨华人的测试说明大家已经意识到华文测试的重要性,意识到一个统一的能力标准对华文教学的重要性。那么如何编制一个测试?怎么评价一个测试是好是坏?测试能对教学起到哪些作用呢?这些都是测试学科研究和关注的问题。

---

① 吴坚立《澳大利亚维多利亚州最新中文会考介绍》,《第八届世界华语语文教学研讨会会前论文集》(三),中国台湾世界华语文教育学会,2006 年。

② 陈雅芬《美国大学先修(AP)中文考试趋势分析:比较其他 AP 外语考试结果》,《第九届世界华语语文教学研讨会会前论文集》(三),中国台湾世界华语文教育学会,2009 年。

## 二 华文测试中的理论问题

编制一个新测试首先要考虑这个测试的用途,即为什么开发这个测试?为谁设计这个测试?然后再根据测试目的选择题型、进行试测、调整题目难度等。研究华文测试首先要思考的问题是开发华文测试的必要性,即为什么要开发华文测试?华文测试有存在的基础吗?

王汉卫(2009)总结了研制华文测试的三个基础:海外有相当数量的华人人口、海外华人讲的华语不同于一般意义上的第二语言、汉字和汉文化对海外华人有深入骨髓的影响。<sup>①</sup>人们常说“有海水的地方就有华侨华人”,千百年来海外华人讲华语、写汉字,较好地保存了民族语言文化特征。华语是海外华人社会的一种“非典型第二语言”,华人、华裔学习汉语,是学习本民族的语言,这不仅是学习使用一种交际工具,也是传承民族文化特征的自觉行为。华人、华裔在学习汉语的过程中还有一些特点,例如很多人会某种汉语方言而不会华语华文、听说和读写发展不均衡等。这些特点决定了华文测试与 HSK 等汉语作为第二语言的测试有诸多不同之处,我们总结归纳如下:

首先是测试对象不同。华文测试是专为海外华侨华人设计的。华文测试的被试群体是海外华侨华人,和被试群体为汉语作为二语学习者的测试不同。对象的不同决定了测试的内容、形式不同。对华文测试特点的思考最早来自于 HSK 的相关研究中。陈宏(1995)在研究 HSK(初中等)的信度问题时,发现不

<sup>①</sup> 王汉卫《论“华语测试”的三个基石》,《暨南大学华文学院学报》2009年第1期。

断增长的“有中国背景”的考生团体(这里指以新加坡为例的海外华人)并没有影响信度系数。<sup>①</sup> 信度是反映一个测试稳定性的指标,信度越高越好,说明测试越稳定,越能反映被试的真实水平。这是否说明 HSK(初中等)也适合这些“有中国背景”的考生团体? 经过严谨详细的论证,他认为“有中国背景”的考生参加 HSK(初中等)属于“常模误用”,也就是说,HSK(初中等)不适合这些考生。

其次是测试内容不同。对外国留学生的汉语测试一般会弱化中国文化,选用的语料都是有关日常生活的或科普类等不涉及传统价值观的文章。而对于华裔来说,保持民族文化、接受传统教育则是华文教学的重要内容。有研究者把“中华文化知识竞赛”和 HSK、HDC 并提,认为“中华文化知识”等反映华族民族特征的内容也应成为华人、华裔学习华语的一部分。同时,他也提到“华校学生不适宜参加 HSK 考试”的问题(郝爱廷,2009)。华文测试内容的特殊还表现在汉字上。对外国留学生的汉语测试对书写汉字要求不高,而华文测试则应更重视汉字在华文教学中的地位和作用。

华文测试的第三个特点表现在语言规范问题上。对外国留学生的测试以普通话为标准,语音、语法和词汇等都以普通话为标准。而华文测试以“华语”为标准,“华语”是以普通话为标准的华人共同语,<sup>②</sup>包括中国大陆、新加坡、中国香港、中国台湾等地的华语,各地华语有共性,也有差异。华文测试对轻声、儿

<sup>①</sup> 陈宏《关于考生团体异质程度对 HSK(初中等)信度的影响》,张凯编《汉语水平考试研究论文选》,现代出版社,1995年。

<sup>②</sup> 郭熙《论“华语”》,《暨南大学华文学院学报》2004年第2期。

化等语言现象应放低要求。周清海(2001)谈到了“语言区域变体”的问题,认为华文测试应注重华语的共性,减少地域特征,同时也要保留各自的语文特点。<sup>①</sup>不管是北京普通话、中国台湾所说“国语”、中国香港或者新加坡华语里的特殊成分,都不应该作为测试的项目,测试应以华语的共同核心为依据。同时,各地区不能避免地要接受自己的常用词汇、语音特点以及语法句式,这些都应包含在测试范围内。

华文测试和汉语作为第二语言的测验也有许多共通之处,华文测试可以为华文教学提供许多信息。梁荣基(1994)从认知角度探讨了选择题、是非题、听写等几种常用题型的作用,即如何从中分析提取对教学有用的信息,这项研究将测试测查的语文能力和认知层次相联系起来,对华文教师编写题目很有启发。任何测试最终都能得到一个分数,明白分数的意义并能合理地解释分数,应是教师的基本素质之一。谢泽文(2003)讨论了如何分析和解释测验分数,还介绍了原始分、标准分、Z分数和T分数等常见的分数计算方法。<sup>②</sup>这些文章介绍了基础的测量学知识,对华文教师和其他测试使用者正确理解分数、合理使用测试很有帮助。

在语言测试领域,关于语言能力问题的研究是一项基础研究。我们测量语言能力,但目前对测量对象的认识目前还处于非常浅显的阶段,对于什么是人的语言能力,还没有一个定论。

---

<sup>①</sup> 周清海《语文测试里的语文问题》,《新加坡华文教学论文二集》,新加坡华文研究会,2001年。

<sup>②</sup> 梁荣基《从认知层次看语文测试的形式和作用》,《新加坡华文教学论文集》,北京语言学院出版社,1994年;谢泽文《测验分数的处理和解释》,《教学与测试》,新加坡华文教师总会,2003年。

在中文测试领域中,语言能力问题也一直备受关注。周质平(1984)讨论了口语能力的构成;张莉萍(2003)则从语义、句式、语气、话题、应对方式等向度研究听力能力;韩彤宇(2003)探讨了主观测验的评分标准对语言能力定义的影响;周清海(1998)分析了影响双语教育的因素,例如人的智力、性别,不同语言的特点等对华语学习的影响;梁荣基(2001)以新加坡特选中学的学生为研究对象,分析了他们在听、说、读、写方面的能力偏差,总体来说就是说比写强,词汇贫乏是他们书面表达的最大障碍。<sup>①</sup> 这些研究已经包括听、说、读、写等各种能力,也注意到了人的自然属性和社会属性对语言能力的影 响,这些研究成果对于语言教学、开发测试和制定语言政策有一定的参考价值。但有的研究还未区分华侨华人与其他汉语作为二语的学习者,讨论的是一种宽泛的华语语言能力,对专门的华人、华裔的语言能力研究还不够透彻深入。

### 三 华文测试的实践

根据所测量的语言技能分类,语言测试可以分成听、说、读、写等几种。这几种测试中,学界对阅读测试的研究最为集中和深入。这种现象和阅读测试自身的特点有关,阅读测试的试卷

---

<sup>①</sup> 周质平《中文口语能力测定之拟议》,《第一届世界华语语文教学研讨会论文集教材教法组》,中国台湾世界华文教育协进会,1984年;张莉萍《华语听力测验内容探讨》,韩彤宇《从不同的评分标准系统看对学生语言能力评估的影响》,《第七届世界华语语文教学研讨会论文集》(第七册),中国台湾世界华语文教育学会,2003年;周清海《文化、智力、性别与双语能力》,《华文教学应走的路向》,新加坡南洋理工大学中华语言文化中心,1998年;梁荣基《学生的双语背景和书写能力》,《新加坡华文教学论文二集》,新加坡华文研究会,2001年。



是书写材料,被试的回答也是书写材料。和听、说、写等其他测试相比,举行阅读测试比较简便易行,而且进行后期的结果分析也比较容易。

在阅读测试的研究中,有对阅读能力的研究,有对阅读题出题技巧和策略的研究,还有对阅读材料文本难度的研究,对新的阅读测验形式的研究。对阅读能力的讨论可以帮助华文教师等出题的人对阅读测验的整体设计有更清晰的认识,谢译文(2003)把阅读理解能力分为七个小类,三个层次,分别是:翻译、释意和推衍。考查阅读能力也应分出层次,一篇好的阅读理解应有各种层次的题目,较低层次的题目如通过速记和查找就回答的问题,也有较高层次的问题,如必须通过推衍才能预知的文中没提及的若干信息,作者没有明确表示的态度或心情等。谢文还简要介绍了如何计算题目区分度(discrimination)的简单方法。难能可贵的是,谢文还提到了要注意华文测试的特点,不能照搬英文测试的考查方法。祝新华、高慕莲(2006)采用“阅读测试层次——题型系统”作为研究框架,这种框架把阅读题目分成复述、解释、重整、伸展、评鉴、创意六个层次,分析了新加坡等地阅读试题中各种认知层次题目的分布情况。祝新华(2009)用新加坡小学离校考试中阅读理解试题的实际例子,详细分析了如何编写一个好阅读理解题。<sup>①</sup> 以上三篇文章资料翔实、客观

---

<sup>①</sup> 谢译文《阅读理解能力的测量》,《教学与测试》,新加坡华文教师总会,2003年;祝新华、高慕莲《中学现代文阅读试题存在的问题及改进刍议》,《新加坡华文教学论文四集》,新加坡华文研究会,2006年;祝新华《从老师的编题表现讨论华文科篇章阅读理解试题的编制策略》,《新加坡华文教学论文六集》,新加坡华文研究会,2009年。

可靠,总结得出的编题建议非常实用。前两篇文章都谈到了在阅读测试中应注意不同层次题目的分布,这确实是出阅读题目时应注意的一个问题。

在教学中,教师要为学生选择恰当的教材和课外读物,编写教科书的人要选择合适的篇目进入课本,审查课本的人要鉴定课本是否适合学生水平,这些都与阅读文本的难易度测量有关。阅读材料的难易度又称为文本的可读性,研究可读性对于编写教材、编制阅读试题等都有重要意义。对阅读材料的难度研究是一个热点,相关的研究较多。谢泽文(1994)和苏启楨(1994)介绍了常见的集中测量阅读文本难易度的方法,并比较了这些方法的优缺点。阅读文本难易度的研究在国外已经有七八十年的历史,罗庆铭(2007)对国内外的相关研究做了一个总结性梳理,后采用实证研究的方法得出结论:篇章的生词量与它的难度等级存在正相关,也就是说,生词量会影响篇章难度。该文使用公式法计算华文篇章的难易度,在客观评价文本难易度上迈出了可贵的一步,文章在理论和研究方法上都值得借鉴。<sup>①</sup>

同样是阅读测试,不同题型的测试也有不同特点。梁荣基(1994)、吴英成(2010)、胡林生(2001)等分别讨论了多项选择题、完形填空和短句测试等常见的几种题型应用于实践中时应注意的问题,分析了这几种题型的适用范围等,对于测试的开发

---

<sup>①</sup> 谢泽文《阅读难度的测量》,《新加坡华文教学论文集》,北京语言学院出版社,1994年;苏启楨《华文阅读难度的测定》,《新加坡华文教学论文集》,北京语言学院出版社,1994年;罗庆铭《新加坡中学华文会考阅读篇章的难易度分析》,陈之权、张连航主编《理论、实践与反思:新加坡华文教学论文十三篇》,新加坡国立教育学院中文系,2007年。

者和华文教师有一定的参考价值。<sup>①</sup>

王汉卫(2008)根据华语的特点,提出新的一种阅读测试方式,这种方式以语素为基本单位。这种测试方式,不仅消除了使用选择题可以猜测的负面影响,而且对汉字的要求较高,符合华语测试的特性。<sup>②</sup>这种对新题型的探索对于听力、写作等测试也有一定的启发作用。

除了阅读理解以外,也有部分关于听力、说话的测试研究,但和阅读理解研究相比,不论在广度和深度上都还不够。何子煌(2001)探讨了听力测试的改革,文中综述了听力测试研究的现状,还借鉴了托福(TOFEL)等其他第二语言听力测试,并在此基础上提出了听力测试改革的建议,是听力测试领域中值得一读的好文章。<sup>③</sup>

以上我们讨论的是以教学内容为蓝本的测试,也称为成绩测试,这种测试可以达到检验教学效果的目的,是一种教学后的测试。下面我们讨论教学前测试,即用于分班的安置测试。中国台湾世界华语文教育学会于2002年开始着手编制专门针对华侨的侨生华语文能力测试,用于把在中国台湾就读大学的华侨学生分入不同的华文课程中,同时兼有评估华侨学生华文能力的作用。

① 梁荣基《利用多项选择题来测定学生语文能力的利弊》,《新加坡华文教学论文集》,北京语言学院出版社,1994年;吴英成《克漏字测验》,《汉语国际传播:新加坡视角》,商务印书馆,2010年;胡林生《短句测试的多元推理方式》,《新加坡华文教学论文集二集》,新加坡华文研究会,2001年。

② 王汉卫《华语阅读测试论》,暨南大学博士学位论文,2008年。

③ 何子煌《探讨华文听力测验的改革》,《新加坡华文教学论文集二集》,新加坡华文研究会,2001年。