

新世纪英语教学丛书

英语测试教程

ENGLISH LANGUAGE TESTING

萧春麟 刘清华 著



上海交通大学出版社

内 容 提 要

本书分传统测试和交际测试两大篇。第一篇由导论(英语测试的理论和实践)、测试的分类与试题类型、英语测试的内容以及评估等章组成。第二篇是简介英语测试的发展到交际测试的过程,以及交际测试理论基础、特点和评估。

图书在版编目(CIP)数据

英语测试教程/萧春麟,刘清华著. —上海:上海交通大学出版社,2001

ISBN 7-313-02770-2

I. 英… II. ①萧… ②刘… III. 英语课-测验-师资培训-教材 IV. G633.413

中国版本图书馆 CIP 数据核字(2001)第 052458 号

英语测试教程

萧春麟 刘清华 主编

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码 200030)

电话: 64071208 出版人: 张天蔚

立信会计常熟市印刷联营厂印刷 全国新华书店经销

开本: 890mm×1240mm 1/32 印张: 9.75 字数: 324 千字

2001 年 8 月第 1 版 2001 年 8 月第 1 次印刷

印数: 1~3050

ISBN 7-313-02770-2/H·557 定价: 15.00 元

版权所有 侵权必究

前 言

在我国,英语卷的命题人绝大多数是从事英语教学的教师。他们虽然未学过语言测试之类的论著,但他们中不乏对测试感兴趣的人。由于各种有关测试的原版书较少,很少有借阅的机会,而翻译的版本又几乎没有。在命题过程中,他们只得遵循“拿来主义”,拼凑试卷,不顾各类测试试卷的测试目的、对象与各种测试的测试目标,那么,其效果必然较差,还不时出错。

本书分两大篇:传统测试和交际测试。第一篇由导论(英语测试的理论和实践)测试的分类与试题类型、英语测试的内容以及评估等章节组成。第二篇是简介英语测试的发展到交际测试的过程,以及交际测试理论基础、特点和评估。力求解决我国英语教师在命题操作过程中所遇到的问题。举例详尽,论述深入浅出,直截了当。各类学校的英语老师,包括小学英语老师,读来也不会有困难。在编写中,笔者参考了当代和近代一些语言学家(如 Alan Davies, Andrew Harrison, Bernard Spolsky, H. G. Widdowson, J. B. Heaton, L. F. Bachman, M. A. K. Halliday, D. Hymes, N. Chomsky, M. canale, D. P. Harris, G. N. Leech, R. Lado 等等)的有关著作和论文,并分析了我国各省市英语试卷约千份,介绍了他们的某些观点,采用了他们的一些材料,在此,谨向他们致谢。

由于语言学流派甚多,说法各异,各派对语言的结构与本质、测试的目的、功能与方法等看法不一,国内编写的此类书也为数寥寥。测试学受当今各种流派的语言学的影响,也孕育出各种不同的测试理论和方法。有一些方法已被几十年的实践证明是有成效的、影响广泛的,并且形成了相应的测试理论。正如语言学家 Bachman 在阐述当代两大测试理论流派时说,美国流派的主流,以 ETS 的 TOEFL 为典型,是“prothotypical psychometric/structural language test”,代表了“language

test development driven largely by measurement theory”;英国流派的主流，以 UCLES 的 CPU 或 IELTS 为代表，是一种测试“designed and developed largely on the basis of a consideration of language skills and abilities, with less emphasis on psychometric concerns”，代表的是“language test development driven primarily by applied linguistic theory”（1988: 2）。但是，世上万物在变，变则进，变则成，变则灵。无论是客观性还是主观性，分离式还是综合式，传统式还是交际式，还是任何什么性质的测试，都不可能永远垄断测试领域。此一时的某个方法可能很走红，彼一时就会受冷落；而从前遭弃用的，现今又招青睐。因此，本书不打算，也没有如此高的水平来判定孰是孰非。我们只想向读者介绍在测试中存在的种种方法，供广大读者、教师、专家参考、评论、选用。

本书原系上海中学英语教师职务培训课程使用教材，历经初审、复审至终审，曾经试用数轮，历时十余年，终于定稿。但因课程经费不足，再则印数有限，书稿未能及时问世，转而投向一贯支持学术研究的上海交通大学出版社。因笔者水平有限，书中的缺点和错误，恳请读者批评指正。

作 者

2001 年 8 月

目 录

第一篇 传统测试

第 1 章 导论	3
1.1 英语测试的历史与研究范围	3
1.2 英语测试与英语教学学科大纲及教材的关系	5
1.3 英语测试的目的	7
1.4 英语测试的命题权	8
1.5 英语测试的范围	8
1.6 试题与题样.....	11
思考题	13
第 2 章 英语测试的分类	15
2.1 英语客观性测试与主观性测试.....	15
2.1.1 英语客观性测试与主观性测试的区别与关系.....	15
2.1.2 英语客观性测试与主观性测试试题的优点和缺点.....	16
2.1.3 英语客观性测试的常见题型.....	17
2.2 英语结构测试与功能测试.....	18
2.3 英语分离式测试与综合式测试.....	20
2.4 英语直接测试与间接测试.....	21
2.5 英语学业测试,水平测试,学能测试,诊断测试,插、 分班测试.....	23
2.5.1 学业测试.....	23
2.5.2 水平测试.....	25

2.5.3 学能测试·····	26
2.5.4 英语诊断测试·····	27
2.5.5 英语分插班测试·····	27
2.6 常模参照测试和标准参照测试·····	27
思考题·····	29
第3章 英语测试中的多项选择与完型填空题 ·····	31
3.1 多项选择题·····	31
3.1.1 基本原则·····	32
3.1.2 题干(stem)、答案项(key/answer/correct option)、 干扰项(distracter)·····	34
3.1.3 答题·····	37
3.1.4 多项选择题的优缺点·····	37
3.2 完型填空题(cloze)·····	38
3.2.1 完型填空题的命题和阅卷评分·····	39
3.2.2 完型填空和多项选择兼容题·····	41
3.2.3 提示式完型填空(C-Test)与部分听写 (P-Dictation/Spot Dictation)·····	43
练习与思考题·····	44
第4章 英语测试的内容 ·····	46
4.1 语法结构测试·····	46
4.1.1 测试语法结构的常见题型·····	46
4.1.2 语法结构多项选择题的测试面·····	47
4.1.3 测试语法结构的多项选择题·····	52
4.1.4 测试语法结构的填充题·····	53
4.1.5 测试语法结构的转换题·····	55
4.1.6 测试语法结构的完型填空题·····	58
练习与思考题·····	58
4.2 词汇测试·····	62

4.2.1	词汇测试常见题型	63
4.2.2	用多项选择题型测试词汇的基本命题原则	70
	练习与思考题	73
4.3	阅读理解测试	74
4.3.1	测试阅读理解能力的多项选择题	76
4.3.2	阅读理解测试中多项选择题的一般命题原则	84
4.3.3	测试阅读理解能力的填充题(completion)和完型 填空题(cloze)	85
4.3.4	测试阅读理解的是非题(True or False)	86
4.3.5	测试阅读理解的跳读(skimming)与扫读(scanning) 能力	86
	练习与思考题	87
4.4	听力理解测试	89
4.4.1	辨音听力测试	90
4.4.2	辨重音、语调的听力测试	93
4.4.3	有视觉媒介的听力测试	94
4.4.4	句子与短对话的听力测试	97
4.4.5	演讲、讲话和长对话的听力测试	102
4.4.6	综合式听力测试	109
	练习与思考题	109
4.5	口语测试	111
4.5.1	朗读	112
4.5.2	转换	114
4.5.3	单向反应	114
4.5.4	双向反应	119
4.5.5	口试的评分	120
4.5.6	关于考务人员的事宜	129
	练习与思考题	129
4.6	写作测试	130
4.6.1	写作试题	131

4.6.2 翻译测试	138
4.6.3 阅卷评分	140
思考题	149
第5章 英语测试的实施与评估	150
5.1 预试与测试的实施	150
5.2 测试的评估	153
5.2.1 效度	153
5.2.2 信度	156
5.2.3 区分	158
5.2.4 英语测试评估中的常用指标及获取	159
思考题	168

第二篇 交际测试

第1章 交际测试的理论基础	173
1.1 Hymes 的交际语言能力	173
1.2 Canale 和 Swain 的交际能力理论模式	174
1.3 B. F. Bachman 的交际测试理论	175
1.3.1 交际语言能力(CLA)模式	176
1.3.2 测试方法层面(TMF)模式	181
1.3.3 测试的评价原则:一致性和实用性	183
第2章 交际测试的特点和本质	192
2.1 测试的目的	192
2.2 测试的内容	192
2.3 测试的方法	195
2.4 评分的方法	195
2.5 对测试的评价	197

2.6 小结	198
第3章 交际测试的实践与展望	200
3.1 国外较著名的交际测试实践	200
3.2 我国的现状	203
3.3 交际测试的展望	204
附录 I 英、美主要考试机构	205
附录 II 高考英语全国试卷(节选)	218
附录 III 专业英语 8 级考试(TEM8)	230
附录 IV 美国的 SAT 考试(节选)	243
附录 V 美国的 TOEFL (节选)	253
附录 VI 英国的 CPE 考试(节选)	261
附录 VII TEEP 试卷(节选)	275
附录 VIII CET 口试(节选)	288
参考书目	296

第一篇 传统测试

第 1 章 导 论

1.1 英语测试的历史与研究范围

英语测试属语言学,是一门专门研究和评估学习者掌握、运用英语语言知识和能力的科学。英语测试是测试的一种。测试的狭义界定就是考试。考试的历史源远流长,据考证,考试始于我国公元前 201 年的汉朝。考的是儒家学说,目的是公开选拔“公务员”以取代世袭制,反腐倡廉。阅卷采取“背靠背”。筛选出的人选,最后由皇帝亲自“殿试”决定名次(B. Spolsky 1995:10)。直到 16 世纪末才由法国耶稣会会士和传教士,从中国介绍到法国,略加修改用于此后的法国选拔“公务员”考试。类似的考试于 17 世纪在欧洲兴起。如普鲁士,大革命后的法国,联合王国的英格兰和印度等。到 18 世纪、19 世纪,西欧各国和美国用于教学和选拔公务员的种种考试方兴未艾。从 19 世纪末至第一次世界大战,一种量化考试结果的“智力测试”在英国、美国被广泛使用。这种智力测试就是语言客观性测试的本源。客观性测试源于这么一个信念:智力测试可用于测量智力,也能用于测量认知力;客观性测试比以前的考试较公平合理。

首先使用客观性语言测试的是 20 世纪 20 年代初的美国。至 20 世纪 20 年代末,客观性测试大大地推动了当时的语言教学研究。20 世纪 30 年代心理测试产业的发展极大地激励了客观性测试技术的运用,但是仅用于当时的外语教学,而且局限于测书面语。至 1954 年,美国人有效地把这些有限的技术用于 TOEFL 考试。此后,由美国教育服务处(ETS)将 TOEFL 发展成现今世界上最成功的英语(作为外语)测试之一。

美国的“TOEFL 考试规划办公室”于 1963 年 7 月 1 日挂牌。当时,由教育行政部门的代表 ETS,经过两年多时间的协调、磋商和准备,创立以 David Harris 和 Leslie Palmer 为正、副主任,Lois McArdle 为行政助理,Elizabeth Shanahan 为秘书的执行机构。考试的名字是由“全国委员会(The National Council)”主席 Boynton 提议的。Boynton 要 Palmer 用“Test of English as a Foreign Language(TOEFL)”作为考试的名称。

虽然英国人运用这些技术较晚,但“剑桥大学地区考试特别委员会

(Cambridge University Local Examinations Syndicate)”也作了相应的尝试。

事实上,随着语言学研究的深入发展和不同时期产生了不同流派的语法大纲、教学法,不可避免,语言学家必须寻找与之相适应的语言测试。请看下面的表格。

“不同时期的测试法与语法大纲、语言学、教学法”对比表

时 期	理论大纲	教学法	代表人物	基本原则	测试法
二战前	无理论大纲	演绎法	K. Plotz	准确性。翻译和写作。教师为中心	教什么,考什么;怎么教,怎么考。评分主要是等第制
二战至60年代末	描写性语法,结构/生成语法,比较语法;心理语言学,社会语言学;心理测量,行为主义心理学等	归纳法;结构法;结构-情景法;听说法;视听说法;默示法等	Jespersen; Haislund; Chomsky; Bloomfield; Leech; Quirk 等	先听说,读写跟上。句型操练,情景对话等。正确、流畅性。教师为中心	听说读写分离式。阅卷评分客观性。评分为分数制。评估强调量化。以 R. Lado 的 <i>English Testing</i> 为代表
60年代末至今	交际/功能意念语法,语用学	交际式。开发学习者的潜能。在真实场景中实践,在讨论中学习。	Hymes; D. Wilkins; Halliday; H. Widdowson; Canale & Swain; F. Bachman 等	语言是交际工具。表达思想、办成事(task)。学习者为中心	要求考生完成 test tasks。强调真实性、可信性、有效性、沟通性、可行性、影响力诸方面的最佳平衡。评分采用等第,分数,评语等综合形式。代表人物: F. Bachman 等

英语测试,作为一种手段,始于教学。自有英语教学,作为其一部分的测试也就存在了。但是,它以一门学科出现,还是近几十年的事。英美等本族语言学家自20世纪初开始研究语言测试。经过几代语言学家、心理语言学家几十年的努力实践和潜心研究,至今已形成不同的理论体系和可供使用的实践方法。语言测试测试语言的什么?测试语言的“用法(usage)”,还是测试语言的“使用(use)”?当代语言学家 H. G. Widdowson 是这样区

分的,“用法(usage)”指考生(学生)掌握语法知识的程度(the extent to which the language user demonstrates his knowledge of linguistic rules);而“使用(use)”指考生(学生)为了进行有效交际而使用语法知识的能力(to which the language user demonstrates his ability to use his knowledge of linguistic rules for effective communication)(1978:3)。那么,在下文提及的众多测试类型中,哪些测试用来测量用法?哪些测量使用?什么测试是较可靠、可信的?一般来说,用客观性测试测量用法较有效,用主观性测试测量使用较有效。但是,这不是绝对的。较有效的测试是从测试的目标出发,只要能有效地测定考生的语言水平,用什么测试类型的试题都可以。至今,测试就其理论和实践问题而言,各持己见,众说纷纭。然而,作为从事语言教学的教师,和把英语测试用作专门用途的专业人员来说,终于有了可供选择使用的余地。

我国是在七八十年代才正式运用英语测试学。那么,究竟什么是英语测试?什么是客观性测试?什么是主观性测试?什么是交际测试?什么是标准化?什么是水平测试、学业测试、诊断测试、学能测试?测试与教学的关系如何?测试怎样分类?英语测试的内容是什么?各种测试的目的、对象、形式如何?怎样进行测试的总体评估以及深层分析?等等,这些是本科研究的范围和内容。本书将就上述问题逐个探讨。

1.2 英语测试与英语教学学科大纲及教材的关系

戴维斯(A. Davies)于1968年曾描述过测试与教学之间的关系:好的测试就像一名忠心耿耿追随其主子教学的仆人。这里,主从关系比喻生动。但中国当今英语教学的实际情况恰好相反。另一种说法以语言学家阿瑟·休斯(Arthur Hughes)为代表,他说,这层关系应该是伙伴关系。无论“主仆”说,还是“伙伴”说,有一点是共同的,即测试与教学紧密相关,是一个整体的两个不可分割的部分。测试是课堂教学的自然延伸,测试既可检验教学的质量,又可为提高教学质量提供有用的信息。很明显,测试对教学有反冲(backwash)作用。科学性较强的测试对教学产生较有益的反冲作用。缺乏科学性的测试则会对教学产生不良反冲。如近年流行的间接测试法(见第2章第4节),使学生的听、说、写能力呈下降趋势。

以英语测试中的学业测试(achievement/attainment test)为例。学业测试是检测教学是否有效贯彻英语学科大纲的手段。譬如,中学英语教师使

用的测试多属学业测试。学科大纲对学业测试的范围、深度和广度应有限制性和指导性。适度的学业测试会使教学更好地贯彻学科大纲,并产生有益的反冲。

学业测试内容的依据是教材。当然,教材良莠不齐不属本教程讨论范围。假设教材已体现了学科大纲的原则和要求,那么,英语学业测试内容的依据应该是教材。任何偏离教材的做法都应列入超纲行为。另一方面,好的学业测试可发现学生在使用教材中的困难或问题,也能发现教师在教学中的优缺点,为改进教学和教材提供有用的反馈信息。

测试对教学具有无可非议的导向作用。请看下例(引自某市某区初三统题),语音辨别:

- A. houses
- B. news
- C. newspaper
- D. horses
- E. newsreel

暂不说试题的 directions 含糊其词,现在就其几个选择项略作分析。D 项为大纲中明确规定的要求,该项本身可数名词以不发音的 e 结尾,未加 -(e)s 读作 /iz/; A 项名词 house 读作 /haus/, 作为特殊情况词末加 -(e)s, 那么 houses 应读作 /haʊziz/, 这恐怕已不在大纲之列。B 项与 C 项单词虽同出一源,但两个 -s 发音不同, B 项的 -s 发音是 /z/, C 项是 /s/; 而 E 项的 s 又读作 /z/。在实际使用时,即使是本族语使用者,在说话时,也极少,或者说不可能注意这些 -s 的不同发音。至于这些差别在语音学上的意义不属本教程讨论的范围。

有关名词复数的测试试题超纲、超教材的例子屡见不鲜。如:

chief	chiefs
cliff	cliffs
safe	safes
Germany	Germanys
woman cook	woman cooks
man doctor	man doctors
man-eater	man-eaters
stand-by	stand-bys
looker-on	lookers-on

仅此数例,足以说明一些教师、学生钻偏题、怪题,死记硬背,忽视语言

运用的倾向。殊不知,每次级考、高考、中考、会考甚至小学毕业考试的试题,都会影响下一轮毕业生与非毕业生的教学。试题题型的任何细微的变化,或“新”语言点、语法点的出现,会不可避免地导致此后对它们成千上万次口头的、笔头的操练。这种反冲的利弊得失,很值得我们深思。

课程改革与教材改革试图摆脱这种应试教学模式。但是在目前,我们还不能废除考试,因为考试毕竟具有不可取代的公正性,通过考试公开选拔才公平、合理。因此,重要的是用好考试这根杠杆,当务之急是探索、寻求具有适合中国国情和中国英语教学的英语测试的理论、原则和方法。

1.3 英语测试的目的

英语测试为英语教学服务(多见于学业测试、诊断测试),也为考核学员掌握英语的水平并加以运用的能力(多见于水平测试、学能测试)服务。

测试服务于教学的主要作用为:

- (1) 巩固已学知识技能,鼓励学生进取。
- (2) 评估学生运用英语能力。

我们就以随堂测试(classroom test)为例,其目的有:

- (1) 及时发现全班或一部分学生的学习难点。
- (2) 使教师随时调整教学内容与方式,提高教学效率。
- (3) 科学地而不是用偏题、怪题测定学生运用语言的能力。
- (4) 给学生表现自己语言能力的机会,激励学生的学习积极性等。

如果我们能及时获得学生的测试情况,学生就能了解自己前一阶段英语学习上的问题,成功的测试可以起到巩固已学知识的作用。好的测试势必成为教学的有效手段。

与随堂测试不同,公开测试(public test)则以选拔为其测量目标,其目的约有五个方面:

- (1) 测定考生已学到的语言知识。
- (2) 测定考生已达到的运用语言的水平。
- (3) 测定考生还需学些什么。
- (4) 诊断考生英语学习的长处与弱点。
- (5) 预测考生继续学习的潜能。

如果我们能获得考生以上五个方面较为可靠的信息,就有可能较全面、确切地为上一级院校或有关用人单位及机构提供考生运用英语能力

的实际情况,考生也能够较客观地根据自己的实际水平作出明智的选择。

1.4 英语测试的命题权

在英美等国家,人们普遍认为测试是少数几个语言测试机构与专家的事。显然,这些机构、专家掌握着语言测试的命题权。英国有 ELTS (English Language Testing Service), RSA (Royal Society of Arts Examinations Board), Cambridge, Oxford 等官方的或代表官方的测试机构。美国也有近似于官方的测试机构,较著名的是 TOEFL 的命题机构 ETS (Education Testing Service); Michigan Test 的命题权是 English Language Institute, University of Michigan; CELT (Comprehensive English Language Test) 的命题者是语言学家 David P. Harris 和 Leslie A. Palmaer 等。其实,测试未必真的只是少数几个语言测试机构与专家的事,语言学家 J. B. Heaton 曾说,有经验的教师已经是半个测试命题专家,只是他们还要学习测试学的原则与技巧,以及怎样运用这些原则与技巧。

在我国,绝大多数掌握着各级、各类、各层次英语测试命题权的人是英语教师。他们中的大多数人,正如 Heaton 指出的那样,是有经验的教师,但只能算半个合格的测试命题人。因为他们未曾有机会去接触英语测试的理论、原理、原则、方法、技巧,更谈不上怎样去运用、实践。因此,命题时常常犹豫不定,无所适从;或者干脆照搬照抄,不考虑试题的内容是否能与考生的实际情况相适应。

笔者曾对我国恢复高考以来,尤其是八九十年代的全国各省、市,主要是上海市各区县学校的近千份试卷作过抽样分析,发现仅由于缺乏测试知识(不含试题纯语言或打印等错误)而引起的平均出错率达5%以上,有的试卷高达10%以上,符合语言测试学原理、原则的试卷可谓凤毛麟角。本教程旨在同广大英语教师和命题人共同学习、讨论、探索国内外测试理论和实践,从而逐步使我们的英语试卷更可靠、更可信。

1.5 英语测试的范围

英语测试的范围受英语教学大纲的制约,什么样的教学大纲形成什么样的测试范围。为了阐述清晰起见,有必要粗略回顾一下英语语法,英语教学法与英语教学大纲的历史。