

清华大学 英语水平测试： 理论与实践

清华大学外语系测试组 编

清华大学出版社

清华大学英语水平测试： 理论研究与实践

清华大学外语系测试组 编

清华大学出版社
北 京

内 容 简 介

本书从理论研究、实际运作、管理方法及政策制定等方面对清华大学英语水平测试进行了总结。它揭示了一个不断发展与完善的清华大学测试体系的建立过程。既有理论基础和支撑,又有实实在在和翔实的操作层面的阐述,还有科学有力的分析。此书是对清华大学外语测试实践经验的总结及相关研究成果的汇总,对于从事相关研究的人员、高校外语教师及有关管理人员均具有一定的参考价值,对于外语测试的发展具有一定的启发与借鉴作用。

读者对象: 广大从事英语教学、科研与管理的工作者。

版权所有,侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

清华大学英语水平测试: 理论与实践/清华大学外语系测试组编. —北京: 清华大学出版社, 2010. 7

ISBN 978-7-302-22698-7

I. ①清… II. ①清… III. ①英语—高等学校—水平考试—研究
IV. ①H310.42

中国版本图书馆 CIP 数据核字(2010)第 085802 号

责任编辑: 刘细珍

责任校对: 王荣静

责任印制: 杨 艳

出版发行: 清华大学出版社

地 址: 北京清华大学学研大厦 A 座

<http://www.tup.com.cn>

邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 北京鑫海金澳胶印有限公司

经 销: 全国新华书店

开 本: 148×210 印 张: 9.125 字 数: 250 千字

版 次: 2010 年 7 月第 1 版 印 次: 2010 年 7 月第 1 次印刷

印 数: 1~2000

定 价: 29.00 元

产品编号: 036608-01

清华大学外语系测试组编写人员

(按章节顺序排列)

策 划 张文霞

编 审 邢 如

编写人员

1. 张文霞、邢如
2. 邢如
3.
 - 3.1 庞红梅、杨莉、邢如、郭茜、张文霞、沈明波
 - 3.2 庞红梅、杨莉、邢如、郭茜、张文霞、沈明波
 - 3.3 张文霞、邢如
 - 3.4 庞红梅、杨莉
 - 3.5 邢如、郭茜
 - 3.6 张文霞、沈明波
 - 3.7 张文霞、沈明波
4.
 - 4.1 郭茜、邢如、张文霞
 - 4.2 郭茜、沈明波、张文霞、邢如
 - 4.3 邢如、郭茜
 - 4.4 邢如
 - 4.5 邢如
 - 4.6 刘梅华、李晓芳、张文霞
 - 4.7 沈明波
 - 4.8 张文霞、刘梅华
 - 4.9 王林晓、张文霞
 - 4.10 耿娟
5. 周咏

前 言

清华大学外语系测试组成立于1995年,负责清华大学英语水平I测试的有关各项事宜。历经十几年的理论与实践,测试组确立了清华大学英语水平测试体系。这是一套在国际最新的测试理论指导下并在实践中不断完善的测试体系,目的在于不断提高学生实际运用英语的能力。它既能检查学生掌握大学英语教学要求所规定的基本技能的程度,又能促使学生提高运用所掌握的技能完成实际工作的能力。

书中各章节从不同角度对清华大学英语水平测试中的不同领域进行了探讨与研究。编辑此书的目的在于回顾、总结多年的研究成果与实践经验,以促进对英语水平测试的认识与理解,为进一步提高英语测试的水平打下坚实的基础。同时也希望本书的出版能为广大从事英语教学与科研的工作者提供些许理论与实践方面的借鉴和参考。

在此特别感谢清华大学外语系测试组的创始人和策划人、时任外语系副主任官力教授,她为清华大学英语水平I测试奠基和发展付出了大量心血。感谢历年来测试组的成员,他们先后分别是孙郁根、贾明舫、邢如、郭茜、张文霞、吕燕彬、庞红梅、沈明波、杨莉、罗承丽、胡阳、刘梅华。感谢陈世敏、王欢、邓湘君、耿娟、申晓静等有关人员。特别感谢程慕胜教授多年来的支持与帮助。同时感谢清华大学

外语系罗立胜教授、吕中舌教授等系领导的支持和许多教师的协助及清华大学出版社的大力支持。

由于本测试组在学术视野、时间精力等方面的局限性，本书中的纰漏在所难免，敬请专家、读者批评指正。

清华大学外语系测试组

2010年2月

于清华园

目 录

1	语言测试理论及清华大学英语水平测试简介	1
1.1	语言测试类型	1
1.2	衡量语言测试的评价标准	2
1.3	清华大学英语水平测试背景介绍	3
1.4	清华大学英语水平测试内容及特点	4
1.5	试题的不同发展阶段	4
1.6	阅卷方式	6
1.7	记分体制	6
1.8	管理方式	6
1.9	讨论及启示	6
2	题型确定与测试方法研究	8
2.1	改变公外传统听力测试的尝试	8
2.1.1	与传统听力测试的比较	8
2.1.2	出题细则	9
2.1.3	难点分析	10
2.1.4	值得思考的若干问题	11
2.1.5	如何评价考题	12
2.1.6	测试结果所反映的问题	12
2.1.7	测试改革的目的	13
2.2	阅读测试方法研究	14
2.2.1	对阅读理解的认识	15
2.2.2	阅读理解测试	15

2.2.3	阅读理解测试的选材	16
2.2.4	阅读理解测试的题型	16
2.2.5	阅读测试的效度与信度	17
2.2.6	清华大学英语水平 I 测试阅读理解测试实践 ...	18
3	清华大学英语水平 I 测试大纲	21
3.1	大纲前言	21
3.2	基本要求	22
3.2.1	词汇	22
3.2.2	语法	22
3.2.3	听的能力	22
3.2.4	说的能力	23
3.2.5	阅读能力	23
3.2.6	写的能力	23
3.2.7	译的能力	23
3.3	试卷构成与分数计算	23
3.3.1	试卷构成与时间分配	23
3.3.2	试卷分数的计算方法	24
3.4	听力部分	24
3.4.1	简介	24
3.4.2	考查技能	24
3.4.3	测试题材	25
3.4.4	题型说明	26
3.4.5	评分	29
3.5	阅读部分	31
3.5.1	简介	31
3.5.2	考查技能	31
3.5.3	测试题材	32
3.5.4	题型说明	32
3.5.5	评分	46

3.6	写作部分	47
3.6.1	简介	47
3.6.2	考查技能	48
3.6.3	测试题材	48
3.6.4	测试内容及形式	48
3.6.5	评分	49
3.7	口试部分	51
3.7.1	简介	51
3.7.2	考查技能	51
3.7.3	交际话题	51
3.7.4	试卷内容与结构	52
3.7.5	评分	54
4	英语水平 I 测试分析与研究	60
4.1	阅读新题型内部效度研究	60
4.1.1	内部效度研究	61
4.1.2	数据的整理与分析	62
4.1.3	分析与讨论	62
4.1.4	结语	68
4.2	评分规范化与口语信度研究	69
4.2.1	简介	69
4.2.2	理论回顾	70
4.2.3	研究方法	72
4.2.4	结果	74
4.2.5	讨论	75
4.3	英语阅读测试题中的完型填空	76
4.3.1	七种题型的测试与调查结果	76
4.3.2	正常卷和试验卷成绩	79
4.3.3	试验卷 A、B 卷平均分差异显著性检验	81
4.3.4	t 检验结论	82

4.3.5	正常卷和试验卷的 Z 检验	82
4.3.6	A 卷与 B 卷的对比	83
4.3.7	结语	85
4.4	英语能力分析	86
4.4.1	对清华大学本科生英语因子分的分析	86
4.4.2	试卷构成与测试成绩	87
4.4.3	因子分析结果	89
4.4.4	因子分计算结果	92
4.5	英语水平 I 测试试卷质量评估	95
4.5.1	试题区分度与难易度测试结果	95
4.5.2	因子分析结果	97
4.5.3	信度分析结果	104
4.5.4	分级不同的学生总分与单项分的对比	104
4.5.5	结语	105
4.6	清华大学英语水平 I 测试之学生评价	106
4.6.1	引言	106
4.6.2	文献综述	107
4.6.3	研究方法	109
4.6.4	结果和讨论	111
4.6.5	结论与启示	120
4.7	Lexical Errors in Chinese College Students' English Writing	124
4.7.1	Introduction	124
4.7.2	Development of Error Studies	124
4.7.3	Research Questions	126
4.7.4	Context	126
4.7.5	Results and Discussion	127
4.7.6	Conclusions and Implications	134
4.8	Investigating Cognitive and Metacognitive Strategy Use During an English Proficiency Test	134
4.8.1	Introduction	134

4.8.2	Research Methodology	139
4.8.3	Results and Discussion	142
4.8.4	Conclusions and Suggestions for Future Research	152
4.9	Internal Validity Research on Writing Subtests of TEPT I	154
4.9.1	Introduction	154
4.9.2	Literature Review	156
4.9.3	Research Design and Methodology	158
4.9.4	Data Analysis and Discussion	160
4.9.5	Conclusion and Implication	166
4.10	Analytical Study of Test-Taking Strategies of Chinese College Students in Reading Comprehension Test	168
4.10.1	Introduction	168
4.10.2	Literature Review	169
4.10.3	Research Design and Methodology	170
4.10.4	Result, Discussion and Conclusion	172
4.10.5	Limitation and Implication	176
5	英语水平 I 测试数据管理	178
5.1	英语水平 I 测试数据处理的流程	179
5.2	影响数据管理准确性的因素	185
5.3	水平 I 测试数据管理的发展方向	186
	参考文献	188
	附录 1 清华大学英语水平 I 测试试卷样卷	195
	笔试部分	195
	口试部分	212

样卷答案及解析·····	216
样卷听力部分文字材料·····	225
附录 2 英语水平 I 测试考务管理 ·····	234
考前事务·····	234
考场管理·····	235
日常管理·····	239
附录 3 外语系测试组与英语水平测试简介 ·····	244
测试组的工作·····	244
水平测试简介·····	244
附录 4 英语教学课程设置、教学管理及水平 I 测试介绍 ·····	250
附录 5 外语学分要求 ·····	252
多数院系对外语学分的要求·····	252
少数院系对外语学分的要求·····	252
附录 6 相关文件 ·····	254
教育部认证书·····	254
清华大学英语水平 I 证书·····	255
试题册·····	257
答题纸·····	262
笔试答题卡·····	270
口语评分卡·····	275
相关出版物·····	276

1 语言测试理论及清华大学英语水平测试简介

1.1 语言测试类型

语言测试是一门跨学科的综合性科学,它涉及语言教学、语言学、心理测量学及统计学等多种学科。语言测试的分类方法有多种。它可以按教学阶段分类,按用途分类,也可以按所测试的语言项目范围分类。

语言测试按照教学阶段可分为编班测试(placement test)、随堂测试(classroom test)、期中测试(midterm test)及期末测试(final test)。例如,我国新生的入学英语分级测试属于编班测试。此类试题覆盖面广,难度等级多,目的是把学生的分数档次拉开,以利于分级别组织课堂教学。随堂测试则覆盖面窄,难度小,目的是复习刚刚学过的知识,以帮助学生掌握学习重点。期中测试则应具有一定的综合性和系统性。期末测试范围则应更广,主要目的是促进学生巩固所学知识及评价一学期的教学效果。

语言测试按照用途又可分为水平测试(proficiency test)、成绩测试(achievement test)、潜能测试/aptitude test)及诊断测试(diagnostic test)。水平测试主要考查学生现有的语言能力,其特点是不考虑以前是否学过或学了多少,也不考虑是如何学的,美国的托福考试就属于此类测试。成绩测试则考查学生对所学知识的掌握程度,它需要参考教学大纲。上述提到的随堂测试、期中测试及期末测试都属于成绩测试。潜能测试用来衡量学生学习某种语言的潜能和天赋。诊断测试用来发现教与学两个方面的问题,其目的是为了改进教学。

语言测试按所测的语言项目范围又被分为独立式测试(discrete point test)、综合式测试(integrative test)、交际式测试(communicative test)及语用测试(pragmatic test)。独立式测试考查语言学习者对构成语言的许多成分的了解和使用,考查对单项知识的掌握,如单个句子的选择填空题。综合性测试则同时注重考查语言的多方面知识与技能,如综合填空、阅读理解、翻译题等。交际性测试不仅注重语法正确,而且注重符合社会规范。语用测试注重考查语言在实际生活中的应用。然而,综合测试与语用测试的区别并不那么明显,而语用测试与交际性测试也很难区分。

1.2 衡量语言测试的评价标准

衡量语言测试的评价标准主要包括效度(validity)、信度(reliability)及实用性(practicality)。效度关注测试是否达到了目的。其中的表面效度(face validity)要求试题看上去是合适的;内容效度(content validity)要求测试应体现要考查的语言要素及技能;结构效度(construct validity)要求测试要有自己的理论依据,它也是解释测试分数的依据。信度则关注分数的可靠性和稳定性。信度体现在试题的覆盖面及代表性上。此外,评分因素也影响着测试的信度。实用性是指试题要便于使用。

此外,对试题区分度、试题难易度、整卷信度等的定量分析方法也随着计算机软硬件的不断方展而成熟与完善起来。而利用计算机技术进行因子分析(factor analysis)也可以使我们对语言能力有更深入的理解与认识。

清华大学英语水平测试顾名思义属于按照用途分类的水平测试,其所信奉的理论依据是语言测试考查的是语言的多方面知识和技能,语言测试不仅应考查学生的语言知识,而且应考查学生运用语言完成各种任务的能力及得体性。即采用了以结构主义语言学和行为主义心理学为依据,将语言分解为听、说、读、写四种技能的方法,也采用了以语言测试的心理语言学和社会语言学为理论依据的总体综合法及新兴

的交际法语言测试,尽量采用真实的语言材料进行测试,力图做到博采众长。

1.3 清华大学英语水平测试背景介绍

推进大学英语教学稳步发展需要树立全面的英语测试评价观,这一点毋庸置疑。大学英语四、六级测试作为一项全国统一的标准化测试而存在,但开发和采用校内英语测试体系反映了大学英语教学评价体系的多元化和个性化趋势。清华大学实施校内英语测评体系可以避免大学英语评价“一刀切”的现象。在坚持语言评价原则的前提下,各高校制定自己的英语测试标准可以兼顾到考生群体的具体需求,而且从操作层面上来说更简便易行。

测试应当是个性化的。建立院校自己的英语教学测试才能准确地衡量学生的英语水平,正确评估教师的执教水平,真正起到促进教学的作用。

清华大学英语水平测试是基于1996年教育部高等教育司八所重点院校外语教学测试改革精神面向大学本科非英语专业的学生所建立的测试体系,并于1999年12月通过大学外语教学指导委员会及有关专家的验证,得到了教育部高等教育司的正式认可。为此,清华大学外语系于1995年成立了测试组,专门负责清华大学英语水平测试的有关各项事宜。测试组成员为外语系具有一定教学及科研经验,且工作认真负责的一线英语教师,负责出题、阅卷、管理、评审、研究等各个环节的工作。历经十几年的理论研究与实践,测试组确立了较为完善的清华大学英语水平I测试体系。测试分为口试和笔试两部分,口试部分占15%。此外,题型设计方面加大主观题比例,问答、写作、填空等主观题占70%。

测试每年在清华大学举行两次,每次约3000人参加。2004年根据教育部高等教育司颁发的《大学英语课程教学要求(试行)》,对水平I测试进行了较大调整,并于2005年10月开始实行新的测试形式。

1.4 清华大学英语水平测试内容及特点

清华大学英语水平测试分为笔试和口试两部分,笔试考查学生的听力、阅读、写作和翻译能力,口试考查学生的口头表达能力。笔试总时间为 150 分钟,口试时间为 15 分钟左右。听力测试是通过尽可能真实的听力材料,考查学生的语音基础和辨音能力,同时要求考生“能听懂英语讲课及简短会话、谈话和讲座,抓住中心大意、要点和有关细节,领会作者的观点和态度。”阅读部分要求学生阅读理解语言难度中等、题材较为广泛的文章,掌握中心大意以及说明中心大意的主要事实和有关细节,并能进行一定的分析、推理和判断,领会作者的观点和态度。同时,还要求学生能将阅读理解文章中的难句和关键句翻译成汉语,理解正确,译文达意。写作测试的目的是对考生英文写作所体现的产出能力或对语言知识的运用能力进行可信而有效的评估。写作测试内容根据不同写作任务的要求分两部分:第一部分主要考查学生一般个人信函(包括短信、便条、备忘录、电子邮件等)交往,询问或提供、解释一般客观信息,组织、描述和比较所提供的数据的写作能力;第二部分考查表达与本人生活有关的需求、愿望、好恶等,陈述问题并提出解决方案,表述、分析、论证或反驳和讨论某些观点的写作能力。口试是根据大学英语教学大纲对学生口语技能的要求设计的,要求学生能进行简单的日常会话,能就测试内容作简短回答和复述,能就熟悉的话题作简短发言,表达思想基本清楚。

清华大学英语水平测试主要有 5 个特点:1)考核内容全面,包括了听、说、读、写、译五项语言技能;2)主观题型的设计和评分规范合理且解释性强;3)强调国际学术交流中的英语知识和能力;4)侧重综合应用能力的考查;5)测试管理系统科学规范。

1.5 试题的不同发展阶段

第一阶段:1997 年至 1999 年

测试组成立初期所设计的试题包括听力理解、词汇与结构、阅读理

解及写作四大部分,每部分满分均为100分,并且其写作部分与授课内容相结合。1997年1月11日清华大学外语系对完成四级课程学习的公外学生进行了一次基础阶段英语测试。测试时间为150分钟。此次测试引起了巨大的震动,特别是听力部分。其中,题量的增加、时间的延长、分值的加大、主观题型的出现及语速的加快都是引起学生不适的原因。

此次对测试试题所做的重大调整就是加大了主观题的比例,提高到平均60%,特别增加了不限字数的问答题。

为了缓和测试给学生带来的压力,从1998年起测试调整为笔试部分的写作(15分)、听力理解(20分)、词汇与结构(30分)、阅读理解(20分)及口试(15分)。

第二阶段:2000年至2004年春季

从2000年起测试分为三部分,试卷一为听力理解(20分),试卷二为阅读理解(30分)及词汇与语法(20分),试卷三为作文(15分)及口语测试(15分)。写作部分不再与授课内容相结合,完成了向水平测试的过渡。

此次对测试试题所做的重大调整就是再次加大了主观题的比例,提高到70%,并继续沿用了不限字数的问答题。

第三阶段:2004年9月至今

2004年9月测试组制定并出版了全新的英语水平I测试大纲及样题。与以前的考题相比,新的考题做了如下变化:听力部分增加了针对如讲座、报告或演讲所出的填充题、概述题等;阅读部分增加了针对一篇1000字左右的较长文章的多种题型,取消了语法与词汇部分;写作部分由完成一篇议论文写作增加为完成两项写作任务,即一篇应用文与一篇议论文;口试部分增加了两考生针对用中文给出的提示信息进行对话的测试方式。此次改进使得考核内容更加全面,主观题的设计和评分更加合理,更加侧重综合应用能力的考查。

此次测试改革最重大的变化之一就是进一步加大了主观题所占的比例,由原来的70%提高到90%。另一个重大变化是对主观题的回答字数限制为三个以内,以利于制定与掌握统一的评分标准,增强分数的