

新英语水平测试指南

孟宪忠 主编

NEW
EPT
EPT
Guide



青岛海洋大学出版社

新英语水平测试指南

A New Guide To English Proficiency Tests

主编 孟宪忠

副主编 石世杰 张金廷 孟悦

青岛海洋大学出版社

5
(鲁)新登字 15 号

新英语水平测试指南

孟宪忠 主编

*

青岛海洋大学出版社出版发行

青岛市鱼山路 5 号

邮政编码: 266003

新华书店经销

山东日照市印刷厂印刷

*

1993 年 9 月第 1 版 1994 年 7 月第 2 次印刷

32 开(850×1168 毫米) 10.5 印张 330 千字

印数 4101—8100

ISBN 7-81026-568-7/H·53

定价: 6.60 元

GF54/05

前 言

《新英语水平测试指南》问世了。希望广大读者能够喜欢。

《新英语水平测试指南》是以国家教委组织制订的《高等学校英语专业基础阶段教学大纲》为基本依据,以国家教委组织编写的《英语水平测试》中的试题为样板,以先进的英语测试理论为指导,由山东省高校英语专业经验丰富的教师集体编写而成。

本书可供大学本、专科英语专业和非英语专业学生使用,也可供参加托福考试(TOEFL)、英语水平测试(EPT),研究生招生英语考试或其他考试的人员使用。它也是广大英语自学者和爱好者的良师益友。

《新英语水平测试指南》全书共分三大部分,即语言测试简介、试题设计与解题技巧,一套一级英语水平测试题及详解,一至四级英语水平测试题各两套。

书中的约一半试题,曾在四个学校作过预测,并根据有关统计数据,对试题作了调整,以确保试题的信度、效度和难度等。

如老师要使用《指南》的样题对学生进行考试,请自行设计答题纸(Answer Sheet)。

参加本书编写的除主编和副主编外,还有下列几位老师(按姓氏笔划为序):丁新善、王怀梅、王宏军、王盼妮、刘书琴、邱云鹏、李长磊、张首翔、张莉萍、国文、姚保慧等。在编写过程中,李长磊老师负责电脑打字,整理编排,做了大量的工作,刘云同志也曾帮助打字。

编写过程中,我们曾参阅了大量的国内外有关考试学和语言

测试的论著,试题集和其它资料,恕不一一注明出处,在此一并致谢。

因时间仓促,水平有限,错讹之处在所难免。诚望专家、学者和广大读者批评指正。

编 者

1993 年 7 月

于潍坊

CONTENTS

语言测试简介.....	(1)
试题设计与解题技巧	(14)
English Test Paper (Band One A)with Explanations	(21)
English Test Paper(Band One B).....	(60)
English Test Paper(Band One C).....	(90)
English Test Paper(Band Two A)	(116)
English Test Paper(Band Two B)	(142)
English Test Paper(Band Three A)	(172)
English Test Paper(Band Three B)	(203)
English Test Paper(Band Four A)	(234)
English Test Paper(Band Four B)	(262)
Key to English Test Paper(from Band One B to Band Four B)	(291)

语言测试简介

实践证明,要提高英语测试水平,具备一定的关于语言测试的基本知识是十分必要的。因此,希望使用本书的朋友们能认真阅读此文。

语言测试是一种教育测量。教育测量基本上是一种间接测量,它是将教育现象数量化的过程。它利用考试这种形式,通过被测量对象解答试卷,对教育现象进行测量,最后以等级或分数的形式表示出测量结果。这种教育测量通常测量的是,通过一段学习,被测量的人大脑中发生了些什么变化,哪些知识记住了,哪些不仅记住了,并且演化成了某种能力。显然,这种测量不能象用尺子测量长度,用称秤物体重量那样直接地进行,而是通过学生解答试卷的好坏,大致估计出被测量者的知识及能力状况。正是由于教育测量是一种间接测量,所以诸多与测量目的有关的因素,诸如:试题出得难或易,考场组织的好与坏,评分准确与否等,都会影响测量结果的准确性、可靠性。

因此,为了全面客观准确地反映被测量对象的情况,我们就应该尽量控制教育测量的误差,把教育测量搞得科学严谨。二十世纪四十年代在美国兴起的标准化考试,大大促进了上述目的的实现。

一、标准化测试

标准化考试(Standardized Tests)的实质就是考试的科学化,它是建立在教育测量理论上,以科学化的方式实施的一种评价教育效果的手段。它能认识并控制教育测量中的误差。从而保证测试目的的有效性 & 测试结果的可靠性。标准化考试不同于课堂考试。它规模大,正规化程度高,并具有统一的标准。它不仅能客观而准确地反映考生的实际水平,而且还可以作为标尺来校正

其他考试。这种考试立足于考生的个别差异,目的是通过比较,考查考生的学业总成就。

标准化考试是一项系统工程,是一套新的考试制度。一般说来,标准化考试是指考试全过程的标准化,从确定考试目的,编写考试大纲,直至阅卷评分及分数解释都要做到标准化。

标准化考试的先进性和优越性是过去的传统考试所无法比拟的。它是教育现代化的标志,是我国教育测试改革的方向。

标准化考试自本世纪四十年代在美国问世以来,得到迅速发展,现已在国际上广泛流行。我国于1985年开始对局部地区的个别学科进行试点并逐年扩展。目前我国出现的“EPT”(English Proficiency Test)英语水平考试,“MET”(Matriculation English Test)高等学校入学英语考试,全国高等院校英语专业四级和八级统考,全国高等院校公共外语的四级考试,全国师专英语专业统一测试等都属于标准化考试。

二、标准化考试的特点

标准化考试是严格按照教育测量理论的要求,用科学方法实施的一种大规模的正式考试。它具有下列特点:

1. 命题程序标准化。标准化考试的试题要由精通该学科知识并懂得教育测量学的专家共同编制。编制试题前不仅要制定考试大纲,明确考试的目的、内容范围、考试的形式、考试的题型及试卷设计、评分计分办法、考试时间等一系列问题,还要根据大纲制订命题双向细目表(Table of specification),然后据此命题。试题拟好后,要认真复查,进行筛选和修改并要经过预测,确定试题的难度、区分度、信度、效度等,获得相应的数据。标准化考试一般要建立题库。

2. 考试实施标准化。标准化考试要求实施过程规范化,严格控制实施条件与手续,并让考生对考试各方面的情况有充分的了解。

3. 阅卷评分标准化。标准化考试大量采用客观性试题。这种试题答案明确,因而评分客观,再加上用机器阅卷,所以可以较好地控制阅卷评分中的误差。标准化考试对试卷中少量的主观性试题也使用与传统考试不同的详细的评分准则进行评判。

4. 分数解释标准化。传统考试通常用原始分(卷面分)评价考生,这样做不科学,因原始分只说明考生在某次考试中答对的题数,它不能全面而精确地反映考生的成绩。而标准化考试对分数的解释是将原始分通过确定的参照标准转换成标准分,从而可以明确而科学地评价某考生的成绩。标准化考试通常采用百分等级分数 PR(Percentile Rank),标准分数(Z-score)及 T 分数表示。

三、常模

标准化考试不用原始分(Raw Score)评价考生。它通常根据某种参照标准,将其转换成标准分(Standard Score)。所谓常模就是一种参照标准。

常模(norm)是教育测量学中的一个术语,是指某一规定的考生总体在某次考试中的成绩的一般情况,单个考生的成绩好坏可参照常模,将其原始分转换成标准分来确定。标准分能清楚地表明该考生在考生总体中所处地位。

四、标准化考试的常用指标

标准化考试是以科学方式实施的严格控制的考试。它需要一些衡量考试成就的评价指标。这些指标不仅是考试质量分析的依据,也是编制试题的依据。这些指标主要是指信度、效度、区分度和难度。

1. 效度(Validity)

效度是指能测量到所要测量的东西的程度。即达到一定测量目的的有效程度。它要回答的问题有二,一是测验要测的是什么特

性,二是它对所要测量的特性测得有多么准确。效度是科学考试的必备条件。一个测验如果没有效度的证据,就不能对测验结果进行科学解释。效度资料可以从多种角度收集,由于侧重的问题不同,效度可分为许多种类。

2. 信度(Reliability)

信度是指重复测量同一事物的某种属性,所得结果的一致程度。它反映了测量结果的可靠性。一项有效的测量,如果能对同一对象测量多次,多次所测得的结果应是一致的、稳定的。一致程度越高,稳定性越大,这项测量就越可靠、可信。信度可用来解释个人在测试中所得分数的意义及用来比较不同测试的分数。信度容易受各种偶然因素的干扰。提高测试的信度,是提高测试质量的重要途径。人们常用下列方式提高测试的信度:①提高试题的区分能力,准确反映不同水平考生分数的差异;②适当增加试题的数量,扩大试题覆盖面,减小偶然得分、失分的机会;③尽量清除测试中的各种干扰因素。

效度和信度原都是统计学的概念,它们两者的关系是辩证统一的关系:信度是效度的前提,信度不高,效度也就无从谈起。反之,效度不高,信度也就失去了意义。

3. 难度(Item difficulty index)

难度是指测验和考试题目的难易程度。通常用通过率 P 来表示,其计算公式为 $P = \frac{R}{N}$ 。这里 R 代表正确回答试题的人数, N 为参加考试的总人数。难度是分析试题时所需计算的主要题目功能指标之一。难度也常用难度指数表示,其计算公式为,难度指数 = $\frac{H+L}{N} \times 100\%$ 。这里 H 为高分组得分之和, L 为低分组得分之和, N 为两组总人数满分之总和。高分组及低分组则是从被试总人数的两头各选 $1/4$ — $1/3$ 的分数。试题的难度指数控制在 0.5 左右为好,偏高或偏低都不利于区分考生。实践证明,一份好的试卷其试

题难易分布之比应是：最难最易的各占 5%，较难较易的各占 15%，难易适中的占 60%。这样的试卷有较高的信度和区分度。

4. 区分度(Item discrimination)

区分度是表示考试能区分考生能力大小和水平高低的指标。区分度好的考试，优秀生考高分，差生得低分。考试的区分度，也可说就是试题能够把不同水平考生按程度高低区分开来的能力。它是试题的质量指标之一，也是主要的题目功能指标，其本质是试题的效度。区分度的计算公式为 $D = \frac{n_H - n_L}{n}$ 。这里 n_H 为高分组正确回答试题人数， n_L 为低分组正确回答试题人数， n 为全部考生人数的 27%。区分度也常用区别指数表示，其公式为：区别指数 = $\frac{H-L}{N} \times 2$ 。这里 H 为高分组得分之和， L 为低分组得分之和， N 为两组总人数满分之和。

难度与区分度虽然有一定的联系，但它们是两个不同的概念。一般地说，调整试题难度是提高试题区分度的重要方法。难度适中的题目，往往有较高的区分度。

试题的区分度与试卷的信度也有密切的关系，试题的区分度高，试卷的信度也高。

五、标准化考试分数的解释

直接从考试中获得的原始分数，其本身没有实际意义，它没有可比性。标准化考试运用各种数学方法把原始分加以转换，而得到一种导出分数，导出分数是正确理解标准化考试所得结果的主要方法。常用导出分数有下列三种：

1. 百分等级分数 这是一种应用广泛的导出分数。它计算简便，意义明确，易于理解。该导出分数的计算公式为：

$$R_x = \frac{\text{考生人数} - \text{原始分所占等级}}{\text{考生人数}} \times 100$$

百分等级分数是由名次等级转换而来的,分数之间距离不相等,它对处于两端的原始分数反应迟钝,至使原始分数发生相当大的变化时,也不能引起百分等级分数的相应变化;而在原始分数集中的地方,甚至分数的微小变化,也可能引起百分等级分数的巨大变化。

2. 标准分数 它是从标准差(Standard deviation)为单位表示考试原始分数相对位置的一个量数。标准分数又称Z分数(Z-score)最适于科学地理解标准化考试结果的真实含义。不可比的原始分数转换为标准分数后,就可以进行直接比较。

标准分数的计算公式为: $Z = \frac{X - \bar{X}}{S}$ 这里X为考生原始分数, $\bar{X}$ 为考生的平均分数,S为标准差,其计算方法为: $S =$

$\sqrt{\frac{\sum (X - \bar{X})^2}{n}}$ 这里的n表示考生人数。标准分虽能确切表示考生的结果,但计算出来的分数必定一半为负值。其最大值为3,而最小值为-3,不便运算。为弥补这个缺陷,人们又提出T分数。

3. T分数 这是把标准分数转换成平均数为50,标准差为10的一个分数。其计算公式为: $T = 10Z + 50$,经过这样转换的原始分数,一半在50分以上,一半在50分以下,一般在20—80分之间波动。

百分等级分数、标准分数与 T 分数的换算关系如下:

丁分数	标准分数	百分等级分数
80	3	99.9
70	2	98
60	1	84
50	0	50
40	-1	16
30	-2	2
20	-3	1

六、标准化考试与客观性试题

标准化考试大都采用客观性试题,这是因为标准化考试要求从命题、答卷,到阅卷、评分等各环节都要严格控制,避免各种误差的出现。客观性试题,正确答案只有一个,不受评卷人主观因素的影响,又可用机器阅卷,不仅节省了阅卷时间,也可以极大地控制误差。客观试题使用方便,它不象主观性试题那样要求被试者自己组织答案,把大量时间花在书写上,因而节省了答题时间,从而使一次考试的试题量可以加大。试题量大,考查面广,这样就可减少偶然得失分的机会,提高了考试的可靠性,并且可以促使被试者严格按大纲规定,全面地复习,避免了传统考试中常见的那种划重点、猜题押宝的不良做法。

尽管客观性试题优点明显,但它的问题也不容忽视,这就是,它不能测量学生的综合能力,复用技能并为猜测得分提供了可能。为了弥补客观性试题的这一缺陷,目前,多数标准化考试添加了写作这样的主观性试题。外语考试中客观性试题的形式常用:选择题、是非(或正误)题、匹配题、改错题等。

七、标准化考试与水平测试

在测试的分类上,由于分类标准不同,因而就获得了多种类型

及名称。标准化考试与水平测试是既有联系又有区别的两个概念，不属同一范畴，不是一组同义词。

标准化考试 是根据测试的规模和正规化程度提出来的。它规模大、正规化程度高。而水平测试是根据测试的目的提出来的，它用来检查考生的水平是否达到某一标准，它不以某一个大纲或教材做依据来命题，因而不象成绩测试那样“学什么，考什么”，而是“需要什么，考什么”。

综上所述，由于出发点不同，一种考试可获得多样名称，它可是一种水平考试，但同时又是标准化考试。

八、当前国内外流行的几种标准化测试

(一)EPT

“英语水平考试”(The English Proficiency Test, 简称 EPT) 是国家教育委员会为我国学者设计的把英语作为一种外语的测试，专供选拔将去英语国家学习的出国留学学生及访问学者使用。它的主要目的是测量应试人员的实际英语水平。它的命题是按照《“英语水平考试”考试大纲》而制定的，而该大纲是根据中国学者在英语国家生活和学习的实际需要而设计的。因此，EPT 是一种水平考试。

EPT 还是一种标准化考试，它向考试使用者提供各种常模，每个考生的分数都能与其他考生的分数相比较。

EPT 试卷共有五个部分，包括有 120 个项目的多项选择题及一个书面表达部分，客观性试题与主观性试题紧密结合。EPT 考试历时 140 分钟，总成绩为 160 分。EPT 试卷的基本结构如下：

测试项目	测试目的	试题量	时间(分钟)
I. 听力	能否听懂正常语速的句子、对话及一般性题材的讲话	A 10 B 10 C 10	25
II. 语法结构与词汇	对标准英语语法结构的掌握程度,对词义及其用法的掌握,词汇量是否能适应阅读原著	A 10 B 10 C 20	30
III. 阅读	阅读能力	30	35
IV. 完形填空	理解和综合运用语言的能力	20	20
V. 书面表达	是否具有初步的书面表达能力	1(或 2)	30

EPT 考试每年在我国举行两次,通常在六月份和十二月份。目前我国共有北京、上海、广州等 9 个考点负责此种考试的实施工作。

(二)TOEFL

“托福”(TOEFL)是英语作为外语的考试(Test of English As a Foreign Language)的简称,它是一种考核非英语国家学生英语水平的考试。该考试具有相当的科学性和有效性。现在美国、加拿大等国家的许多大专院校,要求申请入校学习的非英语国家学生必须提供“托福”成绩,并视它为唯一正式的外国学生英语能力的证明,只有达到学校所要求的成绩的入学申请者,才能取得入学及申请奖学金的资格。

TOEFL 考试基本采用“多项选择法”,试题分为三大部分,共计 150 个题目,约需 110 分钟,(满分)总成绩约为 670 分左右。TOEFL 试卷的结构如下:

测试题目	测试目的	试题量	时间(分钟)
I. 听力理解	听懂、理解美国英语的能力	50 { A 20 B 15 C 15	35
II. 语法结构与书面表达	在标准英语书面表达中掌握重点语法的能力	40 { A 15 B 25	25
III. 阅读理解	理解不同类型阅读材料及理解和掌握词汇使用的能力	60 { A 30 B 30	45

TOEFL 考试从 1986 年开始,在特定的几次中,增加了 30 分钟的作文,作文一般为两种类型的文章:一种是比较两种相反的观点,支持其中一个并加以论证;一种是描述并解释图表。其目的是测试考生的写作能力。

TOEFL 由美国教育服务处(ETS—Educational Testing Service)主办,该服务处每年提供 6 次国际常规和 6 次非常规的 TOEFL 考试。我国为便利出国学习人员参加 TOEFL 考试,从 1981 年开始,经国家教育委员会批准,建立了中国国外考试协调处(CIECB——China International Examinations Coordination Bureau),负责承办考试业务,并在北京 17 所院校及全国六大政区的 28 所院校设立了 45 个考试中心。每年在我国举行四次常规 TOEFL 考试,分别安排在每年的一月、五月、八月和十月(五月及十月份的 TOEFL 考试中有作文)。

(三)GRE

GRE(Graduate Record Examination)是为前往美国或加拿大

等国各大院校研究生院继续深造或申请读研究生的考生提供的研究生入学考试。

GRE 考试分二种：一是一般能力测验(General Test)，一是专业测验(Subject Test)。考试的基本结构如下：

	测试项目	测试目的	试题量	时间(分钟)
一般能力测验	1. 语言能力	语言知识,词汇水平及逻辑思维和阅读理解能力	38×2	一般能力测验共有七个部分,每一部分考试时间为 30 分钟 30×7
	2. 定量能力	基本数学的潜在能力及对数理方面问题的理解及反应能力	30×2	
	3. 分析能力	分析及逻辑推理能力	25×2	
专业测验	在某一学科领域内已获得的知识 and 能力			170

GRE 考试由美国教育考试服务处(ETS)主办,每年提供五次常规考试。我国的 GRE 考试由中国国外考试协调处承办,每年二月、四月和十月,在我国的二十四个考试中心举行。GRE 一般能力测验的总分为 2400 分,专业测验的满分总分为 990。

(四)ELTS

ELTS(The English Language Testing Service)考试由英国文化委员会与剑桥大学地方考试管理委员会联合举办。它是英国许多大学和研究机构衡量外国留学生英语水平能否适应在英国学习某一专业或从事某项训练的主要依据。

ELTS 考试分两大部分：一为普通英语,测试运用英语的一般能力；一为专业英语,测试结合某项专业的学习技巧、写作、口头