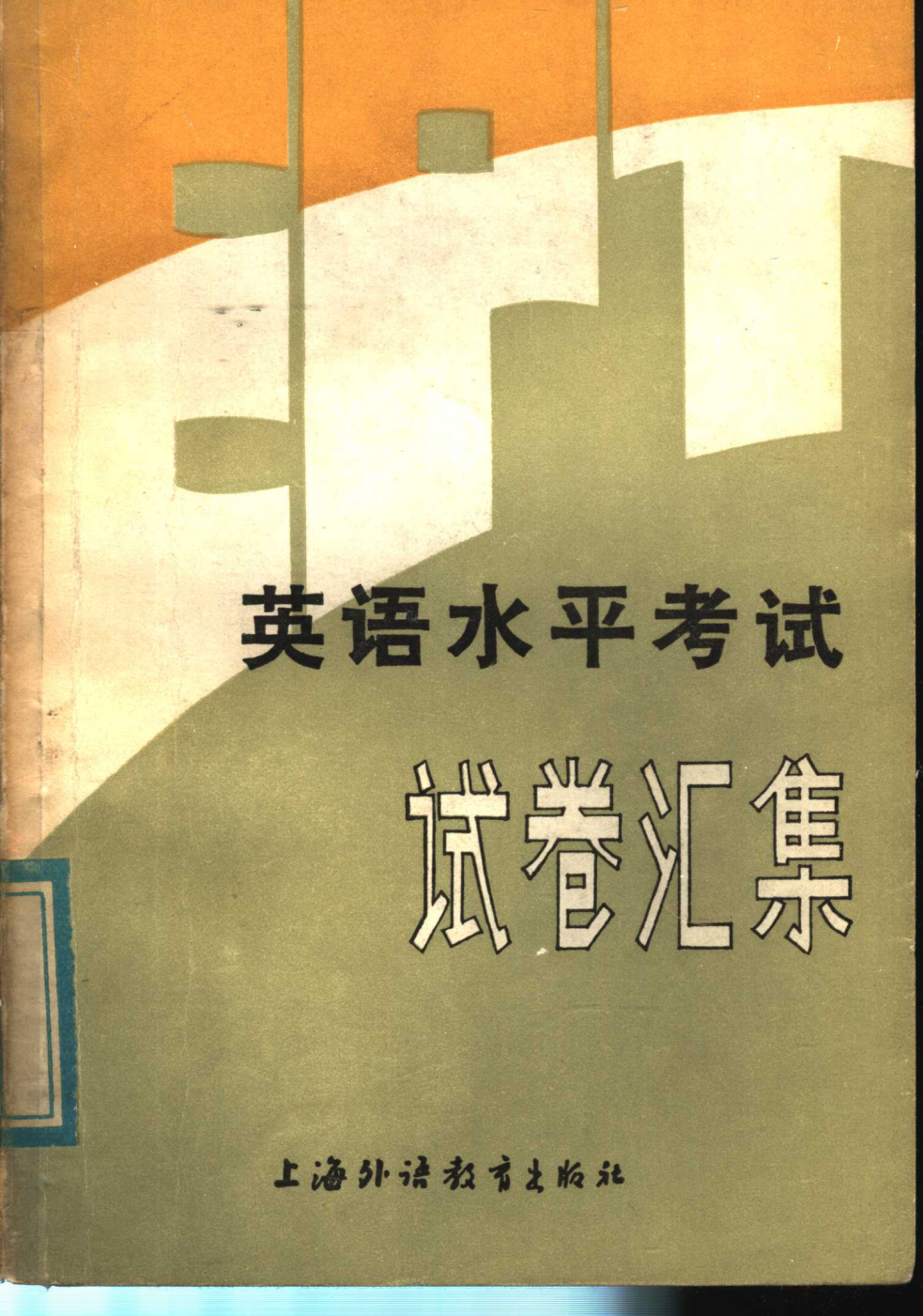




英语水平考试

试卷汇集

上海外语教育出版社

英语水平考试 (EPT)

试 卷 汇 集



上海外语教育出版社
Shanghai Foreign Language Education Press

英语水平考试 (EPT) 试卷汇集
英语水平考试 (EPT) 编写组编

上海外语教育出版社印刷出版

上海东体育会路166号

新华书店上海发行所发行

787×1092 毫米 1/32 6 印张 130 千字
1983 年 6 月第 1 版 1983 年 6 月第 1 次印刷
印数：1—130,000 册
统一书号：7218·089 定价：0.65 元

出版说明

英语水平考试(简称EPT)是在参考国外各种英语水平考试的基础上,结合我国实际情况和需要而设计的一种标准性的、客观性的测验。它既适用于测试报考国内外非英语专业的研究生和出国进修生的英语水平,也适用于测试现在各类学校修读中级或高级英语课程的学生,具有相当英语基础的在职职工、干部、教师、科技人员和广大自学者的英语水平。

英语水平考试每年在全国举行,考试方式和试题深浅程度大致相同。本书汇集了EPT考试大纲的征求意见稿、编号为SAM、M1、Y1的三套试题和答卷,并附有答案以及配有录音的听力材料。广大英语学习者可根据本书提供的试题范例进行自学,并按试题数目、计分标准和考试时间等项目测试自己的实际英语水平。

1983年6月

前 言

一种标准性的客观性的英语水平测验——EPT

桂 诗 春

近两年来，为了测量和评估我国学生、特别是准备派遣到说英语国家进修学习的访问学者和研究生的英语水平，在教育部高教一司的领导下成立了《英语水平测验》(The English Proficiency Test, 简称 EPT) 命题小组，试拟了一套 EPT 试卷，并在主要的出国预备部和一些院校里对不同程度的考生进行了测试。目前 EPT 虽处在试用阶段，但是它是测量出国留学人员英语水平的一种手段，受到各方面的关注。究竟 EPT 是一种什么性质的考试？它与别的类型的英语考试有何不同？EPT 有些什么特点？怎样才能正确使用它？怎样才能考好 EPT？EPT 有些什么问题和缺点？作为命题小组的成员之一，本文作者将试图回答上述问题；但是应该首先说明，文章的观点纯属作者本人的，并不代表命题小组。

EPT 是一种标准性的客观性的英语水平测验。下面将从三方面说明这句话：什么是语言水平测验？什么是标准性测验？什么是客观性测验？

一、什么是语言水平测验？

EPT 首先是一种英语水平测验。我们知道，测验根据

其目的的不同而区别为不同类型的测验。有的是成绩测验，其目的是检查和测量一个阶段内完成某一教学大纲或教材的进展情况；通过成绩测验，我们可以了解到考生学业的进步。有的是诊断性测验，其目的是了解考生某一方面的知识或技能的掌握情况，以便于更有针对性地进行教学。有的是学能测验，其目的是了解考生对某种知识和技能是否具有学习潜能。水平测验和上述三者都不同，其目的是了解考生对某种知识和技能是否具有所需的水平。四种测验中，头两种是我们在教学中经常采用的，也比较熟悉，例如我们在每个学年或学期都根据教学内容进行成绩测验，或是在某些教学环节里安排一些诊断性测验，以了解学生是否掌握某一特定的内容。第三种测验在我国还没有，在国外也是争议较多的一种考试，因为教育家们对什么是学能、有没有学能、学能有些什么组成部分、能否对它们进行测量等问题，都有不同的看法。这类测验在美国似略为普遍些，如美国考试服务公司为入大学和研究生院分别举办 SAT (Scholastic Aptitude Test) 和 GRE (Graduate Record Examinations) 考试。另外，P. Pimsleur 和 J. Carroll 还设计有语言学能测验。最后一种测验，特别是语言水平测验，则相当普遍，例如在美、英、澳、加都设有供测试留学生或移民的英语水平之用的测验。美国的 TOEFL 和英国最近设计的 ELTS (The English Language Testing Service) 都属于这一类。

语言水平测验的一个最主要的特点是它并不根据任何一个教学大纲或任何一本教材命题。这是因为：

1. 参加语言水平测验的考生来源不一，根据单一的教学

大纲或教材命题则有利于一部分考生，而不利于另一部分考生。例如 TOEFL 在全世界一百三十多个国家里举行，各个国家和地区的考生使用不同的教学大纲和教材，根据任何一个大纲来命题都行不通。EPT 虽然只在我国使用，但是考生来源也不一样，有大学毕业多年的高校教师和科技人员，也有刚毕业不久的大学毕业生；对于在各个出国预备部接受英语训练的出国预备生我们使用的也是不同的教学大纲和教材，而且他们的起点和培训方式也不一样。

2. 即使有统一的教学大纲和教材，还要看它们所提出的要求和我们这个水平测验所规定的水平是否相一致。例如目前我国高等学校公共英语有统一的教学大纲，我们能否以这个教学大纲作为依据呢？也不行。因为这个大纲要求比较低，是针对大多数学生的目前的英语入学水平而订的，而英语水平测验是为了选拔少数优秀人材出国之用的。如果根据较低标准命题，势必造成分数普遍偏高，起不到拔尖的作用。

3. 更重要的是语言水平测验必须从实际需要出发订出一个较客观的较稳定的“水平”，使考生有一个看得见的目标。例如 EPT 有一个考试大纲（征求意见稿），它既是命题的依据，也是考生应试的指导。这个考试大纲是根据我国留学生到国外学习在英语方面的需要，并参考了同类英语水平测验而订的。EPT 强调听力和阅读能力。在听力测验中，又着重要求考生只听一遍就能听懂日常生活会话和一般性的讲话和学术演讲；在阅读能力测验中，又着重要求考生既能阅读范围较广的资料，又要有一定的速度。这些要求都是根据国外学习需要而提出来的。我们觉得，根据实际需要制订的英语

水平测验可以对我们的英语教学起到“反拨”作用。在一般的情况下，教学是指挥考试的；但是在某种条件下，考试也可以反过来指挥教学，推动教学。所谓条件，指的是考试必须有明确要求，较稳定的难易度和较固定的格式，并且定期公布考试的结果，等等。目前 EPT 正在朝这个方向努力，望能对出国留学学生的英语培训起一点“反拨”作用。我国多数公共英语的学生的水平还比较低，用 EPT 来要求他们，当然会脱离他们的实际。但是有这样一套试题，使他们看到努力的方向，仍然会有好处。何况标准性试题的分数是一个相对分数，不能把在校的公共英语学生考试成绩和出国留学学生的考试成绩来比较。这一点在下面还要谈到。

这里还想说明一下考试的要求和教学方法的问题。EPT 所要测量的是学生在英语掌握方面，理解与运用的综合能力，故听力和阅读能力的测验项目在整套试题中占的比重比较大，相反，语法部分所占的比重则较小。这当然不是说语法能力不重要。我们只是想在试题中反映这样的指导思想：语法能力的高低应当体现在对语言的综合理解和运用。这一点似和当前公共英语教学中一些同志所采取的逐词翻译、逐句分析的教学方法大相径庭。但是我们正是想通过 EPT 考试来对公共英语教学起一点“反拨”作用，使人们认识到语法分析能力不一定等于理解能力，公共外语教学不能以语法分析作为最终目的。

既然 EPT 的主要对象是出国留学学生，那么我们就要看 EPT 是否适合他们的情况。去年我国有四百多名考生参加了美国的 TOEFL 考试，其中有相当大的一部分是在北京、

上海等地出国预备部进行过一些考前准备。我们曾让他们参加过两次 EPT 考试, 并且事后比较过 EPT 和 TOEFL 的难易度:

	EPT 答对的百分比	TOEFL 答对的百分比
北京	68	68
上海	71	68
广州	63	63

广州外国语学院出国预备部还对 EPT 和 TOEFL 各个部分的难易度作过比较:

	EPT	TOEFL
听力	57.4%	48%
语法结构	81%	75%
词汇与阅读	67.8%	68.3%
Cloze	72%	—
写作	38.7%	—
总计	63%	63%

我们还在三地对约 190 名参加过 EPT 考试的考生征询过书面意见。87% 的人认为 EPT 能反映自己的英语能力; 9% 认为不能, 其中 4% 认为题目偏浅; 4% 没有表示意见。对考试各部分的难易度看法如下:

	易	较易	适中	较难	难	没有表示意见
听 力	6%	25%	46%	20%	2%	1%
语法结构	19%	38%	33%	4%	2%	4%
词汇与阅读	6%	29%	54%	7%	2%	2%
Cloze	10%	39%	40%	8%	1%	2%

写	作	11%	29%	44%	12%	2%	2%
平	均	10.4%	32%	43.4%	10.2%	1.8%	2.2%

总的来说, EPT 和 TOEFL 的难易度是大致相等的, 但 EPT 的语法和听力部分较浅, 对写作有要求, 考生的意见和他们的考试结果相符, 认为题目适中的有 43.4% 人, 容易和较易的有 42.4% 人; 认为难和较难的只有 12% 人。

二、什么是标准性测验

标准性测验 (standardized tests) 是一种测量教学的标尺, 它不但可以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。我们是按照标准性测验的要求来建立 EPT 的, 但是还有不少问题有待解决。

标准性测验异于其他测验的一个显著特点是它要建立和提供标准, 这可以分几个方面来讨论:

1. 标准性测验必须能够适用于不同类型和不同水平的考生, 并且按照他们的实际情况提供一个测量的标准。因此标准性测验的分数是没有及格和不及格之分的, 但是它提供资料来说明某一考分在同类的全部考分中所占的百分比位置 (即“百分位”, percentile)。这样, 考生得到一个分数后, 他自己也好, 使用这个考试的单位也好, 都知道这个分数在同类考分中占什么位置, 代表什么水平。这就是所谓相对分数。表面看来, 相对分数似乎没有什么客观标准。其实不然, 标准性测验根据大纲统一命题, 通过试测校正试题, 定期反复举行, 统计成绩, 可以取得较稳定的标准。例如, TOEFL 经过多年的观察, 其分数的分布是正常的, 均数总是在 500 分左

右,而且 500 分的百分位是 50, 即取得 500 分约比 50% 的分数要高。经过两年多的观察,EPT 的分数也有渐趋稳定之势,其均数大概比 100 分略高一点,但是这个分数的百分位约为 43。下表是根据去年和今年在中南地区 494 名准备出国留学的考生参加三次 EPT 后的分数计算出来的百分位表,可供参考:

EPT 分数百分位表

分数	频数	累加频数	百分位
156—160			
151—155			
146—150	1	494	99.9
141—145	1	493	99.7
136—140	1	492	99.4
131—135	16	491	98
126—130	29	475	93
121—125	41	446	86
116—120	64	405	75
111—115	47	341	64
106—110	54	294	54
101—105	52	240	43
96—100	37	188	34
91—95	28	151	28
86—90	25	123	22

81—85	22	98	18
76—80	24	76	13
71—75	17	52	9
66—70	11	35	6
61—65	9	24	4
56—60	6	15	2
51—55	7	9	1
50以下	2	2	.2

应该注意的是，这个百分位表只适用于出国留学学生这种类型，对别的类型的考生，我们应该在他们多次试用试题后，才能根据其成绩推算其百分位表作为他们一类考生的标准。听说有的理工院校正在试用 EPT 的第一试卷来测量公共英语学生的英语水平。对 EPT 能否用来测验目前公共英语学生的水平，是大可以讨论和研究的。但是有的同志反对使用这套题目，是因为“题目太难，及格率太低”。题目过难是要考虑的，难了就降低了题目的区别性；但是标准性测验不存在所谓及格率的问题。我们不能要求它象一般的成绩测验那样，考生正常情况下都会取得较高的成绩；因为如果这样，考试的适应性就很差，水平略高的考生又会觉得题目太浅。因此，标准性测验的评分只能是相对的，只能在同类考生中作分数的比较，而不能把不同类型的考分来作比较。因此，试考过 EPT 的一些院校可否考虑做这样的研究：一是根据总平均分来推算试题的难易度；二是推算出公共英语学生(最好按年级)的 EPT 成绩百分位表；三是求算学生的

EPT 考分和学生的成绩考试之间的相关系数。这样做的目的是看 EPT 能否如实地反映学生的英语水平。

2. 标准性测验所提出的标准必须是客观的、能经受时间的考验。一些标准性测验在最初建立的时候总要另外一些公认的标准性测验作为它的标尺。通常的办法是让考生同时参加新建立的测验和公认的测验两种考试, 找寻它们的相关系数。相关系数高就说明两种考试都能反映考生的水平, 例如去年我们就让出国预备部的考生差不多同时参加了 EPT 和 TOEFL 的考试。根据广州外国语学院出国预备部考生成绩统计, 证明 EPT 和 TOEFL 具有较高的相关系数:

EPT 与 TOEFL 的相关系数 人数: 62

EPT(样题)		EPT(莫一)		TOEFL(样题)		TOEFL(12月份)	
(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
(1)	—	.86	.86	.86	.85		
(2)	.86	—	.86	.86	.86		
(3)	.86	.86	—	.87			
(4)	.85	.86	.87	—			

当然找寻相关系数仅是一种间接的办法, 更直接的办法是对接受 EPT 考试而又到了国外学习的学生作纵向研究, 了解他们应付学习和生活的英语能力, 从而检查 EPT 的有效性。

3. 标准性测验必须保持其标准的均衡性和继承性。标准性测验经常举行, 甚至成为制度, 如 TOEFL 每年举行 12 次以上, EPT 也有两、三次。每次题目不同, 但却要求反

映出同样的水平。换句话说，如果考生的英语水平没有发生变化，他参加任何一次考试的成绩都应该是一样或基本一样的。这是标准性测验设计中难度最大、技术性最强的一个问题。首先，它要求每次题目的难易度应该大致一样，在 EPT 命题中，我们希望保持题目内容的广泛性（广泛性可减少偏颇），并且使题目的难易有较大的跨度，以测量出水平不同的考生。但是这都是命题人员的主观判断，实际的效果往往是出人意料的。因此，要保持考题的均衡性，还必须采用一系列的措施，如让考题先在同类考生中作试测，对考题中的每一个项目作项目分析（item analysis），找出其可靠性指数和区别性指数，对不理想的项目进行修改或更换，然后拼出一套较好的试题，在更大规模的考生中使用。每次考试之后还要拿试题和以前的试题平衡，并对分数进行调整。因此，一套标准性测验所评出的分数，往往不是考生原来所得的分数，而是一种转换了的分数，甚至是标准分。和别的标准性测验相比，EPT 在这方面的差距最大，因此我们还不能说 EPT 是一套十分完整的标准性测验。

三、什么是客观性测验？

测验按其评分方式的不同而区别为主观性测验和客观性测验两种。主观性测验包括口试、写作、翻译、回答问题等类型的测验项目，其优点是容易考出考生的综合运用语言的能力，有效性较高；缺点是考试范围较窄，而且评分标准较难掌握，可靠性较低。客观性测验包括正误题、多项选择题等类型的测验项目，其优点是考试范围较广，评分简单准确，

考试容易组织，可靠性较高；缺点是不能测量出考生的复用技能(如说、写的能力)，有效性较低，而且考试的副作用较大。

EPT 基本上是一种客观性测验。除第二试卷的写作部分外，其他的测验项目均为多项选择题。其主要考虑是：

1. 一种标准性测验的建立必须以可靠性为基础，并进而考虑其有效性和区别性。如果评分做不到客观可靠，则标准无法稳定，对不同时期参加考试的考生的水平就难以作出准确的测定。在 EPT 试用期间，我们注意使题目保持较高的可靠性，以下为中南地区 494 名考生参加最近三次 EPT 的结果：

人 数	均 数	均方差	可靠性指数	测量误差
494	103	19.2	.91	6

应该说，这和采用客观性测验方法有密切的关系。EPT 也有一个部分(写作)是属于主观性测验的。尽管我们在评分方面采取了种种措施，但是评分结果却仍然有较大的误差，下面是 EPT 各个部分的相关系数：

EPT 各个部分的相关系数 人数：

	语法结构 (1)	词汇与阅读 (2)	cloze (3)	听力 (4)	写作 (5)	总分 (6)
(1)	—	.70	.62	.55	.08	.77
(2)	.70	—	.66	.67	.25	.92
(3)	.62	.66	—	.64	.21	.81
(4)	.55	.67	.64	—	.17	.80
(5)	.08	.25	.21	.17	—	.39
(6)	.77	.92	.81	.80	.39	—

从表中可以看出属客观性测验的项目和总分都有较高的相关系数，说明它们是反映考生的总体语言能力的，另外各个部分之间的相关系数亦不低，说明这些单项分离的语言能力也是互有联系的。但是，属主观性测验的写作这一部分，它和总分以及和其他项目的相关系数都不高，和语法结构的相关系数最低(意味着不少考生的语法结构部分考得好，但写作的分数却不高，或是相反)。其原因可能有两个：一是写作的命题范围窄，考生容易猜题；二是评分不准确。因此在目前的情况下，主观性测验项目过多则势必影响试题的可靠性。有的同志建议在试题中增加翻译题，我们之所以没有这样做是出于下面的考虑：一是翻译是一种较专门的技能，不一定要包括在语言水平测验里。如果把翻译作为了解考生的阅读理解能力的一种手段，则有效性会降低，因为翻译还包括汉语的能力。另外翻译题的范围较窄，也不容易通过它来了解阅读理解能力。阅读理解能力中还有一个阅读速度的问题，这是翻译题考不出来的。二是翻译题的自由程度更大，评分更难掌握。

有的同志认为多项选择题鼓励猜测，分数是猜出来的，靠不住。我们承认多项选择题是有缺点的，特别是副作用不良，应予补救或消除。在试题中包括一些主观性测验项目是一种经常采用的措施，尽管它会给评分带来不少问题。但是也应该说，客观性测验是能够考出考生水平的。因为这种试题的项目比较多，若是水平不高，不可能把许多题刚好都猜对。特别是多项选择题的几个选择，除正确答案外，都有干扰的作用。没有一定的把握是选不出正确答案的。有的同志

认为答错了并不扣分，不见得公允。其实，扣不扣分主要是一个技术上的问题。实行扣分的结果无非是降低每个考生的总分，如果用相对分数来算分，情况是一样的。不过我们在评估考生成绩时必须心中有数，由于 EPT 的多项选择题是有四个选择的，因此其猜测的或然率是百分之二十五。EPT 满分为 160 分，如果一个考生只有 40 分左右，这就意味着他完全不懂，其分数应该相当于零分。为了避免考生考试时认定四个选择中的某一选择作为每个项目的正确答案，而取得过高或过低的分数，在编写多项选择题时，我们必须让正确答案平均分布在四个选择之间。例如一份试题有 100 个项目，每个项目有四个选择 A, B, C, D，那么 A, B, C, D 应各有 25 个是正确答案。这是就整份试题而言的，每一部分的正确答案倒不一定要平均分布，因为有的部分有 10 个或 35 个项目，不可能平均分布，另外每一部分都这样做，会给学生更多猜测的机会。

2. EPT 首先是用来测量出国留学生的英语水平的，文、理、工、农、医的考生都有。题目的范围应该比较广泛，才能适用于不同类型的考生。在这方面，客观性测验项目数量大，比主观性测验的复盖面要广得多。以 EPT 为例，在 160 分钟内考了听、读、写的能力和语法、词汇的知识，便于从不同角度去了解考生的语言水平。阅读是测试的重点，从内容上有自然科学、社会科学、人文科学等方面的阅读材料；从体裁上有议论文、叙事文、描写文、说明文和各种应用文；从要求上不仅考虑理解的程度，还要考虑阅读的速度。因为考生都有专业，选材如果是专业性太强的材料，又会有利于