

# 利用电脑设计与建立英语题库<sup>\*</sup>

赵燕怡

(外语教研室)

**摘要** 利用电脑设计与建立英语题库是实现标准化考试的有效手段之一。笔者大量收集了听力, 词汇, 改错, 阅读及写作等五个方面试题, 利用 IBM/PC 微机, 通过数据库管理系统 DBASE III, 进行编辑, 比较, 存储, 检索, 计算, 打印及设计试卷, 设计与建立了一个英语题库。同时笔者利用设计出的试题进行了两次测试试验, 从统计学的角度对分数分布, 试题效度及信度3个标准进行检查, 均达到指标要求。该题库可试行用于农业院校硕士生通过试。

**关键词** 电脑; 题库; 硕士研究生通过试; DBASE III

在现代化国家的教育制度和考试制度里, 标准化是不可缺少的。

建立题库 (item bank) 是实现考试标准化的一项重要基础工作, 特别对于英语的教学管理。建立题库是一个解决试题用后保密问题及保证测试标准的科学管理办法。从提高效率的角度看, 如果我们把两千份不同的试题保管在资料室, 要检索, 查阅与使用都非常麻烦, 而且保密性不强。如果把这些试题输入电脑, 只需两个很薄的软盘就可储存, 占用空间小, 保密性强, 易管理, 而且随时可使用。一旦题库一建立, 我们可省去很多时间与精力。例如试题不必每次考试前校对, 因为存入电脑的试题只需一次性校对便可储存备用以为将来使用。这样我们可以不断积累和更新试题, 利用电脑这一最新技术手段, 逐步做到考试专业化, 现代化和标准化。

由于电脑具有强大的储存信息能力和具有逻辑判断能力, 可以对两个数进行比较, 判断和选择, 并根据不同比较结果进行不同的处理。这样每一套题目之间就避免重复, 而且保证其测试的标准。例如笔者在词汇方面就输入了1 000道题, 如果按照每份试题含20道题, 就可拼出50份试卷。而这50份不同的试卷又可重新排列组合, 组成更多的题目。这样, 只要我们输入大量的资料, 即可建立各种类型的题库, 根据命题原则进行拼题。

目前农业院校硕士研究生每年都要进行通过试 (Qualifying Test) 测验, 但由于没有标准化试题, 很难比较学生修完一年后水平的高低。因此笔者根据国家教委对硕士研究生英语考核的要求, 内容, 题型, 格式以及一致性, 进行了英语试题库的设计与建立的尝试。首先大量收集了五个方面的题目<sup>[1-3]</sup>, 其中包括听力, 词汇, 改错, 阅读以及写作。然后按试题做好分类, 输入电脑, 归档, 以备随时检索调用。在计算机方面, 笔者选择 DBASE III 来建立题库。因为微机的数据库管理系统 DBASE III 是数据库技术中的杰出代表, 具有很强的数据处理功能, 即能对各类型的数据进行收集, 存储, 分类, 计算, 加工, 检索和传输, 数据库系统将输入的题目以记录为单位存储于数据库中。

• 校长基金资助课题  
1991-01-30收稿

使用电脑编制的试题,于1990年5月22日与1990年9月11日在89级的硕士生中进行了两次测试试验,时间均为2.5 h,共分100分,其中客观题占80%,主观题占20%。参加这两次测试的考生人数均是33人。测试结果如表1。现根据测试结果,从分数分布、试题效度、及信度等三大测试标准,对试题进行分析如下:

### 1 分数的分布 (Distribution)

采用下列公式分别计算出偏态值 ( $G_1$ ) 和峰值 ( $G_2$ )<sup>[4]</sup>:

$$G_1 = \frac{\sum (x-\bar{x})^3/N}{\sum (x-\bar{x})^2/N \sqrt{\sum (x-\bar{x})^2/N}} \quad (1)$$

$$G_2 = \frac{N \sum (x-\bar{x})^4}{[\sum (x-\bar{x})^2]^2} - 3 \quad (2)$$

根据表1的数据计算得

第一次测试:  $G_1 = 0.0420047$

$G_2 = -0.4776565$

第二次测试:  $G_1 = 0.0472809$

$G_2 = -0.192854$

两次测试的  $G_1$ ,  $G_2$  值都接近于零,说明这两次测试分数分布均属正态分布。

### 2 试题的效度 (Validity)

从统计学的角度分析,效度 (Validity) 指的是根据考试分数作出推论或预测的准确性程度。我把这两次测试成绩相比较,然后计算两者的相关系数。如果相关系数很高,也即效度高,说明我们所建立的试题库也是标准化的。

我们所用求相关系数的方法为 Spearman 公式: (桂诗春 1986)

$$\rho = 1 - \frac{6 \sum (x-y)^2}{N(N^2-1)} \quad (3)$$

$N$  = 总人数     $X$  = X 卷分数     $Y$  = Y 卷分数

根据实验数据表明,当用积矩率法计算时,相关系数为:

$$\begin{aligned} r &= 1.44 \times \rho \\ \text{代入公式 (4)} \quad r &= 0.72 \end{aligned} \quad (4)$$

查对相关系数表,表明这两次考试分数相互关系密切。

### 3 试题的信度 (Reliability)

信度就是试题的可靠性。从试题的效度的相关系数以及试题种类数可通过公式计算其信度。一般来说,影响信度的因素有题目的多少(即试卷的长度)以及题目的区分度等等,如果信度系数超过0.90,就说明试卷的试题数已足够。将相关系数及试题种类代入公式<sup>[4]</sup>:

$$\alpha = N\bar{P} / [1 + \bar{P}(N-1)]$$

其中:  $\bar{P}$  = 相关系数平均数 = 0.7

$N$  = 试题种类数 = 5

求得第一次测试:  $\alpha = 0.92$

求得第二次测试:  $\alpha=0.93$ 

学生两次考试成绩的比较与统计

学生 编号	考试成绩		(X-Y)	(X-Y) <sup>2</sup>
	第一次	第二次		
1	55	64	-9	81
2	53	61	-8	64
3	54	61	0	0
4	51	55	-4	16
5	59	65	-6	36
6	53	57	-4	16
7	58	60	2	4
8	50	68	-18	324
9	51	60	-9	81
10	44	53	-9	81
11	52	52	0	0
12	48	59	-11	121
13	58	66	-8	64
14	51	66	-15	225
15	55	64	-9	81
16	65	73	-8	64
17	65	68	-3	9
18	65	62	3	9
19	61	73	-12	144
20	71	81	-10	100
21	67	72	-5	25
22	62	76	-14	196
23	53	55	-2	4
24	62	60	+2	4
25	55	57	-2	4
26	60	70	-10	100
27	48	49	-1	1
28	48	71	-23	529
29	48	54	-6	36
30	32	44	-12	144
31	52	63	-11	121
32	34	40	-6	36
33	60	65	-5	25
平均数 $\bar{x}=53.6$		$\bar{y}=61.7$	$\Sigma=2775$	
标准差 $S_x=8.6$		$S_y=9.1$		

两次测试得出的信度系数均超过0.90并且两次系数接近,说明该试卷信度高。

上述分析表明,用测试的三大标准进行检验,试题均达到了指标要求。这说明所试行建立的题库具有较高的科学性与可行性且该题库随机性强,库存量大。每次测试使用不同的试卷而又保证参照点稳定不变,这就提供了一个统一的,客观的,稳定的标准,由于采用DBASE III 编程序,使命题过程标准化,科学化。出题组卷速度快,试卷标准化,美观,清晰。该题库可试行用于农业院校硕士生通过试。

致谢 本课题承蒙李侯才教授,罗富和教授的指导,及刘平章教授、刘才兴教师、徐汉虹博士研究生的帮助。谨此致谢。

#### 参 考 文 献

- 1 邓平等.最新托福试题详解.北京:光明日报出版社,1990.1~289
- 2 朱月珍,陈厚勤.英语水平考试指南.南宁:广西人民出版社,1983.708~716
- 3 周世英,吴建国,王晓风.出国人员英语水平测试模拟试卷汇集.上海翻译出版公司,1985.37~430
- 4 桂诗春.标准化考试——理论,原则与方法.广州:广东高等教育出版社,1986.127~147

#### DESIGNING AND ESTABLISHING AN ENGLISH ITEM BANK WITH THE COMPUTER

Zhao yanyi

(Department of English)

**Abstract** To design and establish an English item bank with the computer is one of the most effective approaches to realizing standardized tests. With an IBM/PC microcomputer, an English item bank was established using a DATA BASE III (DBASE III) managing system and based on extensive collection of language material in five aspects namely, listening, vocabulary, identifying errors, reading and writing. By classifying and storing them in different categories, editing, finding, calculating, comparing, designing and printing could be done for a standardized test. Two experimental tests were carried out on postgraduates with the computer-designed test papers. Examining the test results with three statistical measurements, distribution, validity and reliability, it was concluded that the tests had reached those standards. As a result, this item bank can be recommended to work in qualifying tests for postgraduates in agricultural universities.

**Key words** Computer item bank; Qualifying test for postgraduates; DBASE III