

# EFL 写作水平测试影响因素的实证研究\*

邹琼

(湖南师范大学, 长沙 410081; 中南林业科技大学, 长沙 410004)

**提 要:**本研究通过对 12 名不同水平受试者两篇不同类型作文的测试, 并与他们的平时成绩和四级考试成绩相比较, 从卷面得分和语言运用两项指标, 分析探讨写作水平测试的“一致性”和“有效性”问题, 最后得出结论: EFL 写作水平受到作文题目、中/英文提示句和图表/图画的影响, EFL 写作水平测试要慎重命题, 尤其是四、六级考试。

**关键词:** EFL 写作水平测试; 作文题目; 中/英文提示句; 图表/图画

中图分类号: H319.3

文献标识码: A

文章编号: 1000-0100(2008)06-0139-4

## An Empirical Study of Influences on EFL Writing-level Tests

Zou Qiong

(Hunan Normal University, Changsha 410081, China;

Central South University of Forestry & Technology, Changsha 410004, China)

Experimenting on twelve different-level students with two different styles of compositions, and comparing the experiment results with students' terminal and CET-4 scores, this research, according to the compositions' total points and language use, explores the problems of uniformity and validity in EFL writing-level tests. It comes to the end that writing topic, Chinese/English hints, and chart/graph have great influences on EFL writing, so EFL writing-level tests should be prudently designed.

**Key words:** EFL writing-level test; writing topic; Chinese/English hint; chart/graph

### 1 引论

EFL 写作水平测试是一种对产出性语言运用能力的书面综合检测, 要提高测试的准确度, 必须考虑“一致性”和“有效性”原则。

#### 1.1 测试的“一致性”原则

##### 1.1.1 测试的内容与受试者的知识结构相一致

作文测试应与平时的写作训练区分开来。平时的写作训练主要是给学习者一个扩大词汇量和视野的机会, 教师可随意提供各种题型。然而, 作文测试要考虑的因素则比较多: 写作内容是否符合受试者的心理、语言和认知基础, 主题是否非常明确, 能否激发写作热情, 能否促使连贯性和可理解性输出。George Hillocks, Jr 概括了写作者写作时要使用的两种知识: 一种是实质知识 (substantive knowledge), 指事实、观点、事件、信仰等; 一种是形式知识 (formal knowledge), 指词汇、句法、修辞、篇章结构等 (George Hillocks, Jr 1986: 71-91)。形式知识用来表达

实质知识。英文写作应尽可能发挥写作者这两方面的知识, 表现他们在选择题材、遣词造句、运用语法和组织篇章等方面的能力 (杨玉晨 2005)。

##### 1.1.2 测试的难易度与受试者的认知熟悉度相一致

作文测试难易度的标准不容易确定。Brown *et al* 认为, 主要有两个维度 (dimension) 可以采用。一是信息内容的多少。要表达的信息内容越多, 说明任务越难。二是信息内容和任务类型内在关系的复杂程度。他把任务类型分为三种: 一是静态 (static) 任务 (三项任务中最容易的); 二是动态 (dynamic) 任务; 三是抽象 (abstract) 任务 (三项任务中最难的)。信息内容越多, 任务类型越多, 作文任务也越难 (陈慧媛 吴旭东 1998)。Foster 和 Skehan 也指出, 根据个人的、直接的信息设计的作文任务比那些根据个人不太熟悉的、遥远的信息而设计的作文任务要容易一些 (Foster & Skehan 1996: 299)。

认知熟悉度包括对话题的了解程度和对写作内容的

\* 本文系 2007 年度湖南省社科基金项目 (07YBA001) 和中南林业科技大学人才引进基金项目 (105-0014) 的阶段成果。

可预测程度。话题越不熟悉,写作内容的可预测程度越低,写作时对注意力的要求就越高,受试者认知负担就越重,写作任务也越复杂;反之,对话题越熟悉,认知负担越轻,写作任务也越简单。

### 1.2 测试的“有效性”原则

#### 1.2.1 从卷面总得分来分析

Heaton(1988)指出,测试的效度(validity)是指在多大程度上考了该考的内容,亦称测试的有效性。语言测试的效度种类繁多。内容效度(content validity)是指测试的内容要能体现出语言要素的运用和语言技能的发挥。共时效度(concurrent validity)是指两次时间相近的测试结果相比较,或与教师对学生的评估相比较,具有相同或相当的效果。例如,相距较短时间内测试两篇作文,受试者两篇作文的分数相当,说明这两篇作文具有共时效度,反之没有共时效度。区分度(discrimination)是指测试的自然结果按比例分布在不同的分数段(60-69,70-79,80-89,90-100),而不是人工预先设计好的按比例分布,它可以区分出不同水平学习者的差距。有效的作文测试可以反映出高低不同水平学习者的真实水平。过难或过易都会影响区分度:过难的作文使水平本来较高的学习者都无从下手,过易的作文会让水平低的受试者侥幸获得较好的成绩。

#### 1.2.2 从卷面语言运用来分析

近几年关于语言运用的研究大部分集中在不同口语任务对学习者的语言运用的影响方面。尽管口语与写作有不少差异,但两者在心理、语言和认知等方面仍有共同之处,都反映一种交际能力,即把内在的思想用清晰、得体的语言传递给信息接受者的能力。在衡量不同口语任务对受试者语言运用的不同影响时,Foster和 Skehan提出了口语任务的三条标准:流利度、正确度和复杂度(Foster & Skehan 1996: 324)。流利度是指作文中所用单词的总数量和丰富量,正确度是指作文中语法正确的句子数量与全文句子总数的比率,复杂度指作文中复合句的数量与全文句子总数的比率。我们也可以采用这三条标准来衡量受试者在写作中的语言运用,以此检验测试在多大程度上考了该考的内容。

## 2 研究方法

### 2.1 研究对象

本次研究对象是从湖南科技学院 2003级教育系教育技术学本科专业两个班遴选出来的 12名学生(7名男生,5名女生)。他们分别代表成绩为上、中、下水平层次的学习者。2005年 6月 18日他们参加了全国大学英语四级考试。测试分别在四级考试前一周的两个上午进行。

### 2.2 研究问题

本研究试图通过对两篇作文的比较和分析,回答以

下三个问题:

- (1) EFL写作水平是否受到作文题目的影响?
- (2) EFL写作水平是否受到中/英文提示句的影响?
- (3) EFL写作水平是否受到图表/图画的影响?

基于上述三个问题,笔者提出的假设是: EFL写作水平受到作文题目、中/英文提示句以及图表/图画的影响。本文旨在验证这个假设。

### 2.3 研究工具

本研究采用了两种研究工具:英语写作测试和定量分析软件。前者用于测量英语写作水平,后者用于对数据进行分析。

英语写作测试采用以下两个不同类型的作文题:

TV—a Good Thing or a Bad Thing(较易,以下称作文一)和 Changes in the Lives of Chinese Farmers(较难,以下称作文二)。这里的较易写和较难写是根据写作任务的信息内容多少和学生对话题的熟悉程度来确定的。其中作文一给出英文提示句;作文二给出中文提示句,并且作文二是图表作文,要求学生按照图表提供的数据,说明近 5年来中国农民生活的变化情况。两篇作文都是控制性作文(guided writing),要求在规定的 30分钟内写出不少于 120个单词的三段式文章。

使用的 4种分析工具分别是:Word Smith, Voca Profiler, SPSS(8.0版本)和 Microsoft Readability Index。Word Smith用于词汇检索和单词列表,统计英语作文中词汇的丰富度和流利度;Voca Profiler用于测量英语作文中的词汇密度;SPSS用于统计正确度、复杂度;Microsoft Readability Index用于计算作文的质量和水平。

### 2.4 数据收集和分析

两篇作文测试要求受试者不写姓名,只写给定的代码,代码顺序按他们二年级上学期英语期末考试卷面总分高低排列(见表 1)。两篇作文由两位不知测试目的的资深英语教师按四级考试作文评分标准客观、公正、单独评分,然后取平均值(见表 2,满分 15分, T1=教师 1, T2=教师 2)。为便于比较,表 1分别列出了受试者二年级上学期英语期末考试总分、作文单项分(其实是 96年 1月四级考试作文题 The Two-day Weekend,但改成英文提示句)和 2005年 6月 18日全国英语四级考试总分、作文单项分(折成百分制便于比较)。

表 1 受试者英语学期考试成绩和四级考试成绩

受试者		1	2	3	4	5	6	7	8	9	10	11	12
学 期	总分	85	83	81.5	79	69	66	65	63	54	53	50	54
	作文	14	12	13	11	9	9.5	9	7.5	7	6	5	5
四 级	总分	81	78	76	71.5	64	63	61.5	59	49.5	46	43	5
	作文	13	12	12.5	11	8.5	9.5	8	8	7.5	5	4	3

表2 受试者两篇作文测试得分

受试者	1	2	3	4	5	6	7	8	9	10	11	12	
作文一	T1	14	13	13	11	10	10	9	8	7	6	6	5
	T2	14	13	14	12	9	9	9	8	8	6	5	5
	平均	14	13	13.5	11.5	9.5	9.5	9	7.5	7.5	6	5.5	5
作文二	T1	10	9	9	8	10	11	9	10	8	8	7	6
	T2	9	9	10	9	9	10	10	10	9	8	8	7
	平均	9.5	9	9.5	8.5	9.5	10.5	9.5	10	8.5	8	7.5	6.5

### 2.5 信度与效度

为了保证实验成功,笔者遴选了成绩一向比较稳定的上、中、下不同层次的受试者,并充分调动使他们发挥出真实水平,比较他们作文一测试的得分和学期成绩以及四级成绩,排序基本一致,证明他们发挥出了实际水平。另外对评卷教师也严格要求和把关,表2显示两位教师给出的分数在统计上差异不大,说明评分具有一定的客观性。而且实验结果与先期的理论推测结果一致,说明实验具有一定的构筑效度(construct validity)。受试者实地作文的语料库和学期测试的信息来源也保证了实验的内容效度(content validity)。

至于实验的信度(reliability,亦称可靠性),笔者将受试者的编号按奇偶数分为两半,每一半中含有相同数目的上、中、下水平层次的受试者,然后将表1中的各数据和表2中作文一的各数据用SPSS提供的信度计算,得出不同水平层次的信度为.7590。这一信度是比较高的,也证明了实验具有一定的可靠性。

## 3 结果和讨论

本研究从卷面得分和语言运用两项指标来看测试的“一致性”问题,讨论作文题目、中英文提示句和图表图画对EFL写作水平的影响。

### 3.1 从得分情况来看

作文一表现出较高的区分度,受试者的得分从高到低比较合理地分布开来,而且分数的次序与平时成绩和四级考试成绩相符合。但是,各受试者自己的两篇作文分数相差很大。这两篇作文是在同一星期内写的,而且受试者又都很认真地完成,两篇作文的得分应该相差不多,应该具有共时效度。但实际情况是,分数反映出来的共时效度较差。其中受试者1、2、3、4是班上英语较好的学生,而他们作文二的得分与平时成绩中、下等的受试者差不多,其他受试者作文二分数的次序与他们正常成绩的次序也很不一致。这只能说明作文二的考试效度和区分度均较低。

笔者发现,在写作文一时受试者很快进入写作状态,而写作文二时,却花了较多时间解题。从提示句来看,作文一是用英语提示,受试者可以顺着英语思维尽快解码

语言信息,激活思维图式,诱发相关知识,预测写作内容,即刻进入写作状态。从话题的熟悉度来看,作文一是受试者相当熟悉的话题,其中的利与弊非常清楚,因此语言形式、词汇选择范围较广,便于用自己的词汇来组织语言、表达思想。

而作文二是图表作文,受试者首先要看懂图表,对图表给出的数据进行分析,逻辑地找出变化的原因,最后根据变化提出自己的观点。并且,中文提示句使受试者首先要用中文编码获得语言信息,激活相关图式,然后把中文编码转换成英语编码;或者首先把提示句的中文编码转换成英语编码,激活相关图式,然后编码语言信息进行表达。以上两种语言编码的转换都会使进入写作状态的速度变慢。在整个写作过程中,受试者的信息处理过程至少需要5个步骤:解码语言信息(看懂图表)、分析数据、转换语言信息(把中文提示句转换成英文)、编码语言信息、表达思想。而且由于图表和数据的限制,受试者无法回避一些力所不及的语言表达形式和词汇,抽象的内容又使他们必须凭自己的想象找出几条令人信服的理由,然后用英语表达。这一复杂过程极大地增加了受试者信息加工负担。

Skehan说过,较难的任务使受试者把大量的注意力资源耗费在信息处理上,结果只剩下较少的注意力放在语言形式上(Skehan 1998:7)。Ellis(1985)也认为在完成认知较复杂的任务时,学习者很可能把注意力放在了非语言的问题上(如分析数据等),而无法注意那些刚学到的尚无法运用自如的中介语形式(interlanguage forms)。本次测试结果正好与Skehan和Ellis的观点相吻合。

### 3.2 从语言运用情况来看

#### 3.2.1 流利度

单词总数上,作文一明显长于作文二(见表3);词汇丰富上,受试者在作文一中使用了他们平时很少用的三十多个词汇,如inexhaustible, incessantly, instinct, sub-consciousness, censor, materialize, compensation, redress, industrious, expertise, availability, plausible, bothersome, negligence, intake, overwhelming, refrain, beguile等,而作文二中只用了约十个,如permissive, reward, desirous, mutual, desperate, endeavor, nutritious, persistent, violate等。显然,作文一能较快地激活受试者的图式知识,使他们主动使用更多的表达式词汇和领会式词汇。这说明好的作文题目有助于实现流利表达,篇幅也会随之加长。因此,流利度与作文题目呈正相关。另外从表3可以看出,作文篇幅与作文质量没有相关关系,从表1可见作文质量与英语水平呈正相关。

#### 3.2.2 正确度

作文二的正确度(0.60)比作文一(0.71)低(见表4)。在写作文二时,正如上述分析,受试者认知负担较

重,很难把注意力资源集中在语言形式上,正确度必然降低。另外由于时间的限制,写完后没有时间检查,也是正确度降低的原因之一。而在写作文一时,受试者不须要把注意力放在诸如分析数据和转换语言编码等问题上,可以腾出较多时间注意语言形式,因此能最大限度地发挥他们的语言水平(Ellis 1985)。可以说,正确度与作文任务的难易度呈正相关。

表3 受试者两篇作文的单词数量  
(作文一单词数不包括给出提示句的单词)

受试者	1	2	3	4	5	6	7	8	9	10	11	12	平均数
作文一	225	255	230	240	201	211	221	223	217	186	172	195	214
作文二	176	193	225	219	129	182	124	172	131	156	150	178	170

### 3.23 复杂度

作文一的复合句(0.39)普遍比作文二(0.28)多一些(见表4)。原来笔者认为,两道作文题在认识过程中都具有一定的挑战性,都能引导受试者写出比较复杂的句子,但结果却出乎意料之外,各受试者复合句总数都没有超过句子总数的一半。为什么会导致这种现象?从表4我们发现,高分受试者的复合句基本上接近句子总数量的一半,而低分者却没有写出多少复合句。这一结果说明作文的复杂度与作文题目没有相关性,而与受试者的英语水平呈正相关:高水平者能写出较多的复杂句,低水平者多用简单句。

表4 受试者两篇作文的正确句、复合句数量  
(作文一统计的数量不包括给出的提示句)

受试者	1	2	3	4	5	6	7	8	9	10	11	12	总数	
作文一	句子总数	15	22	17	15	15	13	13	17	13	10	9	11	170
	正确句	15	18	15	10	9	9	9	7	7	7	7	5	120
	复合句	7	10	8	7	6	6	5	7	3	2	3	2	66
作文二	句子总数	13	12	15	10	12	10	13	9	7	8	7	8	124
	正确句	7	7	8	6	8	8	7	6	5	4	4	4	74
	复合句	5	4	3	4	4	4	4	2	1	2	1	1	35

## 4 结论与启示

### 4.1 结论

根据上述实验研究,笔者得出如下结论。

(1)作文题目设计直接影响受试者EFL写作水平的正常发挥:熟悉的话题和可预测的内容使他们有话可说,并更多地使用形式知识和实质知识。英文提示句比中文提示句更容易促使受试者用英语思维,激活相关知识,较

快进入写作状态。图表、图画作文分散了受试者的注意力,使他们不得不集中注意那些非语言形式而减弱了对中介语形式的注意。

(2)作文流利度与作文题目呈正相关,与作文质量没有相关性;作文质量与英语水平呈正相关。作文正确度与作文任务的难易度呈正相关:难易度直接影响作文的流利度、正确度和复杂度。作文复杂度与作文题目没有相关性,而与受试者的英语水平呈正相关。

### 4.2 对EFL写作测试命题的启示

根据本次实验研究,笔者提出如下建议:

(1)英语教师应慎重设计作文测试题,认真考虑题型,从多方面考察受试者已掌握的实质知识和形式知识情况,同时还应考虑测试效度和区分度,使测试真正反映受试者EFL写作水平。

(2)设计图表作文时,要考虑写作任务的类型不能太多,并给出较简单的英文提示句,这样有利于受试者尽快解码语言信息,进入写作状态。

(3)四级考试通常是提纲式作文(95%),但提示句多用中文(95%),如果改用英文,测试效度和区分度应该更高。

## 参考文献

- 陈慧媛,吴旭东. 任务难度与任务条件对EFL写作的影响[J]. 现代外语, 1998(2).
- 杨玉晨. 英文写作中的汉语思维模式和逻辑推理——中国学生英文习作案例分析[J]. 外语学刊, 2005(6).
- Ellis, R. *Understanding Second Language Acquisition* [M]. Oxford: Oxford University Press, 1985.
- Foster, P., P. Skehan. The Influence of Planning on Performance in Task-based Learning [J]. *Studies in Second Language Acquisition*, 1996(18).
- George Hillocks, Jr. The Writer's Knowledge: Theory, Research, and Implications for Practice [A]. In Anthony, R. Petrosky, David Batholmae (eds). *The Teaching of Writing* [C]. Chicago: University of Chicago Press, 1986.
- Heaton, J. B. *Writing English Language Tests* [M]. Beijing: Foreign Language Teaching and Research Press, 2000.
- Skehan, P. A. *Cognitive Approach to Language Learning* [M]. Oxford: Oxford University Press, 1998.