

# 英语语言测试五议

广东佛山科学技术学院 周 胜

**摘要:**英语语言测试有很多值得研究和讨论的问题,如测试内容规范,平衡样板,题目措辞,语法题与简单词汇,词汇题与功能词等。了解和注意这些问题将有助于实现语言测试的科学人文化。

**关键词:**测试内容规范;选样;词汇;语法;题目措辞

在语言测试实践活动中,有很多具体问题需要从理论层面弄清楚并在实践中切实加以关注,以求避免不可靠、无效或低效度的测试结果。但笔者只能在这有限的篇幅内根据问卷调查结果从实证的角度讨论应该注意的五个问题:测试内容规范,平衡样板,题目措辞,语法题与简单词汇,词汇题与功能词。

## 一、问卷调查

1999年底,笔者获准全国基础教育外语教学研究资助金项目《对中学外语教学测试误区的研究》(批准文号:99GJWYY)。为了认真完成该项目,笔者与项目成员一道设计了一份调查问卷,分为A卷和B卷,共160道题。受试者包括四川省各级各类中学100多所的近600名英语老师。发出问卷800份左右,有效卷为563份,其中参加A卷的受试者包括未分开统计的本、专科学历教师267人;参加B卷的教师包括187名专科学历教师和109名本科学历教师。要求受试者匿名填写,否则无效。据此,笔者认为此次调查结果较有代表性。

## 二、调查结果及讨论

### 1. 考试应该有内容规范

人们普遍认为考试应该有内容规范(test specifications)(Alderson, et al 2000; Bachman, 1990 & 1996; Harrison, 1983; Hughes, 1989; 高兰生、陈辉岳, 1996; 桂诗春, 1986; 李筱菊, 1997; 舒运祥, 1999; 刘龙根等, 1997; 邹申等, 2000等),因为内容规范为考试的官方文件,规定了考试的内容、方式和试题编写及组织管理人员应遵循的规则等事宜。Alderson等(2000:38)认为“考试内容规范应该包括:考试目的、考生情况、考试层次、考试结构、语言课程或课本描写、试题组成部分及比重、试卷数量、目标语掌握程度、题型、文本长度、所测语言技能及成分、考试任务、考试

方法、考试要求、评分标准、每一水平层次考生的典型行为描述、每一水平层次考生在真实生活中所能做的语言行为描述、样卷及考生考试行为样卷”。“毫无疑问,这样的一个文件对涉及一项考试的每一个人都十分重要:对考试设计者,这是保证考试质量的“蓝图”;对命题者,这是命题的依据;对教师,可以知道如何利用考试提供的反馈信息提高教学质量;对考生,了解了考试的题型、难度、时间长度等,有利于发挥自己的真实水平;对用户,了解了考试的性质和分数意义,有利于正确进行决策,防止考试的误用”(杨惠中 2000, 见 Alderson 2000:17)。

调查发现,94.39%的受试者对“考试是否应该有内容规范”之说持肯定态度,其中完全赞同的有76.00%、基本赞同的有10.90%、部分赞同的有7.49%,说明多数受试者知道考试内容规范的重要性,令人欣慰。

### 2. 学校考试应该是课堂教学的平衡选择

学校考试是“衡量一个人在某一特定课程或教学计划中所学的语言已达到何种程度的考试”(成绩考试, Richards等, 2000:4)或者“显示出学习者哪些方面技能或知识已经掌握而哪些方面尚未掌握的考试”(诊断性考试, Richards等, 2000:133),也就是说,学校考试需要满足内容效度的要求,即“考试任务在何种程度提供相关且具有代表性的任务领域选样,在成绩考试中尤为重要,是编写学校考试时需要考虑的首要问题”(Gronlund, 1981)。代表性指:(1)所测内容应该反映所学内容,而这里的“反映”指所测内容应该是最能代表考生语言能力的重难点,因为一份学校考试试题所能考查的内容远远少于一门课程或一个学期所学的内容;(2)考试应该以适当的比例反映有代表性的内容,因为不同的考点具有不同的权重。这就意味着,如果要想证明考试是否以适当的比例平衡选择了课堂教学内容,只需将考试任务与描述该任务领域的考试内容规范进行比较即可。调查发现,94.10%的专科学历受试者和93.52%的本科学历受试者同意学校考试完全应该(34.80%;44.00%),基本应该(47.00%;43.10%)或部分应该

(12.30%;6.42%)以适当的比例反映课堂教学内容。因此可以看出这一观点为大多数受试者,也可能为大多数教师所接受,值得肯定。本科学历受试者和专科学历受试者之间的相关系数为0.964,他们在这一问题上的看法呈高度正相关,即相互关系极为密切。

### 3. 题目的措辞应该清楚、准确、简洁

根据 Grice 合作原则的要求,交谈双方应该遵守量的准则,即所说的话应该包括交谈目的所需要的信息以及所说的话不能包含超出需要的信息。将此用于考试题目的编写,就需要题目提供足够的信息以帮助考生理解,这也是语言教学与测试可理解性(comprehensible input)原则的要求,而且必须以清楚、准确、简洁的方式呈现信息,这样考生才能正确到位地理解信息。在考试实践中发现有两种情况值得注意。一种情况是,为了帮助考生理解,试题编写人员故意提供多余的信息,如额外的单词、短语、句子等。虽然本意是好的,但可能引出两个问题:(1)给考生增加额外的阅读负担,白白地浪费宝贵的考试时间;(2)使阅读而非考点本身成为该题的重心(桂诗春,1986:53;舒运祥,1999:120)。另一种情况是,为了“提升”试题的难度,有的试题编写人员故意模糊或非正规地呈现题目,从而引起误解和混乱。如:

(1) Last Sunday afternoon I met Alan in the street, he told me that he sold most of his belongings. He had scarcely \_\_\_\_\_ left in the house.

- A. everything                      B. anything  
C. nothing                          D. something

舒运祥(1999:114)解释到:“此题考试不定代词,题干的前半部分(第一句)与解题没有多大的关系,可以改成:

Having sold most of his belongings, he had scarcely \_\_\_\_\_ left in the house.”

(2) He went to school \_\_\_\_\_.

- A. because he wanted to learn more  
B. because he wanted to meet new friends  
C. because he wanted to get a better job  
D. because he wanted to please his parents

(Henning, 1982) “because he wanted to” 在选择项中重复四次,违背了实用性原则,因此建议将此放在题干中:

He went to school because he wanted to \_\_\_\_\_.

- A. learn more                      B. meet new friends  
C. get a better job                D. please his parents

(3) I would say it is not impossible for someone not as careful as you are to \_\_\_\_\_ this not very important point.

- A. overdo                            B. overlook  
C. overrate                          D. overstate

李筱菊(1997:271-272)解释到:“题干一连用了四个否定(还有选择项中的 over-也使词义否定,笔者注),似乎

故意要把所有受试者都搞糊涂。这当然没有必要。受试者糊涂了,他的真正能力也就体现不出来了。”这样的例子比比皆是。

调查发现,98.84%的专科学历受试者和 97.29%的本科学历受试者分别坚持题目的措辞完全应该(78.00%;78.00%)、基本应该(17.10%;14.70%)或部分应该(3.74%;4.59%)清楚、准确、简洁。相关系数为0.999。因此受试者赞同笔者的看法,而且看法一致。

### 4. 语法考题应该使用简单词汇

每一题目都有自己独特的考点,语法考题的考点就是语法。当呈现语法考题给考生时,我们最想了解的是考生是否掌握了该题所涉及到的语法项目(考点),而非其他。语法考题里的词汇不可以也不应该成为考点。难词用于引出语法考题时,至少牵涉到两个考点:语法和词汇。如果有考生没有做出答案,我们很难判定他是因为不知道语法内容或是不理解词汇意义而没能找对答案。所以,最好的办法是语法考题里的词汇应该简单,同理,词汇考题的语法结构也应该简单。调查结果在某种程度上肯定了这一观点。65.80%的专科学历受试者和 57.80%的本科学历受试者分别认为基本不可(31.00%;35.80%)或绝对不可(34.80%;22.00%)在语法考题使用难词。另外的 34.20%的专科学历受试者和 42.20%的本科学历受试者持相反的观点。两组受试者所做选择的相关系数为0.778,说明他们在这一问题上有着不尽一致的看法,因为专科学历受试者教授初中英语,学生所接触到的词汇简单,语言能力较差,而本科学历受试者大多教授高中英语,其词汇复杂些,学生的语言能力也强些,因此难词会影响初中学生理解语法考题,但却不一定影响高中学生找对答案。

### 5. 词汇考试应该考查词汇

词可以分为开系词(也叫实义词,包括名词、实义动词、形容词和副词)和闭系词(也叫功能词,包括介词、代词、限定词、连词、情态动词和基本助动词)。(Quirk *et al.* 1985:67-70, Richards *et al.* 2000:321, Bussmann, 2000:45 & 471)简单地讲,开系词组成词汇,而闭系词为结构词,属于通常意义上的语法研究范畴。这就意味着,如果是词汇考试,考题就应该来自那四类词,因为词汇考试旨在检查考生是否掌握了该词的意义,而非其他的。本次调查结果为:只有 11.80%的专科学历受试者和 16.51%的本科学历受试者完全同意(1.60%;9.17%)或基本同意(10.20%;7.34%)词汇考试就应该考词汇,而 88.20%的专科学历教师和 83.49%的本科学历教师分别坚持认为词汇考试完全可以(26.70%;12.84%)、基本可以(32.60%;39.45%)或部分可以(28.90%;31.20%)包含检查功能词的题目。相关系数为0.802,相互关系密切。笔者大胆猜测,受试者可能没有在意实义词和功能词,也许因为在教学实践中很少提及实义词和功能词或者因为在考试中没有专门 (下转第 24 页)

应该应用“教育要三个面向”的理论,先权衡我国现代化的需要与可能,再面对世界审视自己的需要与可能。然后展望未来以发展自己的需要与可能。总的说,就是本着“实践是检验真理之唯一标准”的理念,广泛借鉴中外有关理论、经验与方法。

#### 4. 宏观调控,和谐竞争

这是引用我国经济建设的成功经验于多种英语教材的编写与选用。

(1)宏观调控指教育部通过外语教育政策、课标和英语教材的审订与管理办法,以保证教材的多元化以及编写经验和资源的共享,从而促进多种教材在质与量两方面都能有序地发展。

鉴于多种教材的市场竞争须得有一定的宏观调控,以及众多编写单位编写基教英语教材的经验还有不足,加之教育部基教司官员较少,笔者认为可以在人教社成立基教英语教材编写中心,作为教育部这方面的办事机构。

此外,为了发挥多种教材的功能,最好健全英语教材选用机制。为了保证教师在选用上的发言权,教育行政官

员最好与选用脱钩;同时规范出版社的推销活动。因为英语教材决非一般商品,选用是否得当,会影响一代或几代人的发展,所以对其推销活动也需要制订与外语教育政策相吻合的规范。

#### 注释与参考

①本文只是粗线条地提出了笔者个人的思考,既不成熟,也未加申论。如能引发共鸣,当再就我国外语教育政策和英语教改方面的问题,分题加以论述。

②据传,各省市基教各学科教科书的供应书目,须由教育厅正厅长选定。

③除部订《课标》以外,上海等地均自订课标。

④这种衔接有隐性与显性两面。显性化衔接表现为内容深浅的升降,教学形式的呼应等等。隐性化的衔接主要表现为学习风格的民族性特点;比如,我国师生因受大一统文化的潜在影响,较易接受演绎式教学。

⑤参考张正东编著.外语教育学.第22节,科学出版社,1999年。

(上接第19页)考查实义词或功能词的单一考试。无论原因是什么,笔者主张,词汇考试还是只能检查考生对实义词的掌握,而语法考试只可包括语法项目。

### 三、结束语

本文讨论了五个问题,前三个问题的调查结果与假设的吻合程度达到了90.00%以上,第四个问题的吻合程度达到了60.00%以上,而第五个问题的吻合程度很低,只有14.16%。从本专科学历受试者对同一问题看法的相关系数看,除第一个问题外,在第二、三、四、五问题上呈高度正相关,说明本专科学历受试者对本文所讨论的多数问题有着相同或基本相同的看法。

#### 参考文献

- [1] Alderson, J., Charles et al. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press, 1995.
- [2] Bachman, Lyle F. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1990.
- [3] Bachman, Lyle F. et al. *Language Testing in Practice*. Oxford: Oxford University Press, 1996.
- [4] Bussmann, Hadumod. *Routledge Dictionary of Language and Linguistics*. Beijing: Foreign Language Teaching and Research Press, 2000.
- [5] Cronlund, Norman E. *Measurement and Evaluation in Teaching* (fifth edition). New York: Macmillan Publishing Company, 1981.
- [6] Harrison, Andrew. *A Language Testing Handbook*. London: Macmillan Press, 1983.
- [7] Henning, Grant. Twenty Common Testing Errors EFL Teachers must Avoid, Forum, 1982. In *English Language Testing*, lectured by Charles Read, 1982-1983.
- [8] Hughes, Arthur. *Testing for Language Teachers*. Cambridge: Cambridge University Press, 1989.
- [9] Quirk, Randolph et al. *A Comprehensive Grammar of the English Language*. England: Langman World Publishing Corp, 1985.
- [10] Richards, Jack C. et al. *Longman Dictionary of Language Teaching & Applied Linguistics*. Beijing: Foreign Language Teaching and Research Press, 2000.
- [11] Watson, Owen. *Longman Modern English Dictionary*. London: Longman Group Limited, 1976.
- [12] 李筱菊. 语言考试科学与艺术. 长沙: 湖南教育出版社, 1997.
- [13] 刘龙根, 孙怀庆. 外语考试学导论. 吉林: 吉林大学出版社, 1997.
- [14] 高兰生, 陈辉岳. 英语考试论. 南宁: 广西教育出版社, 1996.
- [15] 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.
- [16] 邹申. 简明英语考试教程. 北京: 高等教育出版社, 2000.
- [17] 周胜. 英语语言测试研究. 成都: 四川大学出版社, 2003.