

英语专业测试中的“盲点”剖析

丁黎明

(三明学院 外语系,福建 三明 365004)

摘 要:英语专业测试中存在许多“盲点”,这些“盲点”涉及测试目的、测试内容、命题技巧、成绩分析以及评估等各个环节,在很大程度上影响了测试效度和信度,进而给教学带来不少负面效应。依据语言测试理论对种种“盲点”的生成原因和负面影响进行剖析,有助于测试者认清语言测试的本质,克服目前存在的问题,以提高英语专业测试的质量。

关键词:英语专业;语言测试;盲点

中图分类号:H319.6 **文献标识码:**A **文章编号:**1672-447X(2008)06-0123-04

一、问题的提出

测试的功能被扩大化是个不争的事实。在大多高校,成就考试作为评价学生学业的重要手段,取代了终结性评估,因而在学籍管理中起着举足轻重的作用,英语专业考试也不例外。事实上,不论从测试理念到考试目的,从考题设计到命题形式,或是从考试实施到成绩分析,语言测试实践中的误区和盲点是普遍存在的。

所谓“盲点”通常是指容易被人们忽视或不易发现的某些问题。这里讨论的“盲点”即是测试者在英语专业测试中出现的误区或时常忽视的问题。通过对语言测试理论的研究和实践的反思,不难发现在英语测试的各个环节中存在着这样或那样的“盲点”。毫无疑问,这些隐性的问题在不同程度上影响了测试的效度和信度,进而对语言学习和教学产生了负面反拨效应。

二、“盲点”与剖析

英语专业测试中的盲点可能出于测试的每一个环节。这里选择测试目的、测试内容、试题数量与测试时间、测试提示语、测试分析及成绩评估六个方面进行剖析。

1. 测试目的不够明确

测试目的是语言测试要解决的首要问题,对于测试者这往往是最大的盲点所在。这一盲点的表征是:考试与教学大纲脱节。测试者往往对形式的关注胜于对内容的关注。他们

首先解决的是如何考,即决定采取什么方式、什么题型以及相关权重等问题,而不是明确应该考什么,因此在实践中留下了较大的模糊性。导致这种模糊性的原因至少有以下三点:其一,由于大学教育的特点更具包容性和开放性,不同的教师对同样的教材会做出不同的处理,教学内容和重点也不尽相同。其二,课程大纲对教学目标和内容的描述大多是指导性和概括性的,可操作性不强。其三,专业课程的成就考试基本上没有考试细则可依,大多由任课教师独立命题。事实上,成就考试与水平考试有着本质的不同。“成就考试的最大特点显示在它在教学大纲的关系上。成就考试可以说是基于教学大纲内容上的考试,即 syllabus-based。”^[1] 这种考试的目的应该“在于告诉学生他们学了些什么,强调的是调动积极性,而不是区分性”。^[2] 因此脱离教学大纲的成就考试具有盲目性,它难以调动学生学习的积极性,还可能误导学生的学习行为和方法,如过于注重题型练习和题海战术,甚至导致学生迷失学习方向。

2. 语言能力与知识失衡

“语言测试的主要用途在于推断受试者的语言能力”。^[3] 但种种迹象表明英语测试中重知识轻能力的倾向普遍存在。^[4] 如在测试内容方面,语言知识的比例往往超过语言的能力;就测试形式和方法而言,分离式试题多于综合性试题,客观测试题多于主观测试题,间接测试多于直接测试。显然结构主义类型的测试模式依然在语言测试中占据着主导地位。导致语言测试中知识与能力失衡的原因主要有两个方面:一是测试者对语言能力结构缺乏认识;二是对语言能力测试手段的贫乏。前者是认识上的误区,后者则是力不从心。从理论

收稿日期:2008-09-10

作者简介:丁黎明(1956-),福建古田人,三明学院外语系副主任,副教授,研究方向为英语教育、英语测试。

上讲,如何明确界定语言能力结构(language ability construct)、如何将结构效度(construct validity)落实到测试任务(test task)中以及这种结构是否与测试目的紧密相关,都是测试者务必深度思考的问题。就可行性(practicality)而言,由于人为的避重就轻心理客观存在,测试的理想往往不得不向测试的现实条件妥协。Hughes曾敏感地发现:“测试内容往往是取决于测试的容易度,而不是测试的重要性”。^[12]这可谓一语道破天机。不可否认,这种避重就轻以及加大精确度和客观性的策略有助于提高测试的信度,但同时应该警觉:获取这种信度是以牺牲效度为代价的。因为测试内容的重要性是决定内容效度的重心所在,重要内容的缺失必将导致测试效度的缺失。

测试中语言能力与知识的失衡往往导致了“高分低能”的结果,换言之测试的分数无助于说明受试者的语言能力。正如Palmer所言:“从语法和词汇测试中得到的分数未必能帮助我们推断受试者分析篇章结构的能力”。^[13]此外测试的反拨效应研究表明,语言能力由于测量的精确度难以把握而受到冷落,而测试中被冷落的东西往往成为教学中被忽视的东西。^[12-2]这便是一种具有危害的反拨效应。

3. 试题数量与时间配置不当

这个问题比较隐性,通常不会引起测试者足够的关注。题量多了,考生无法按时完成,教师自然是吃力不讨好。但更为常见的问题是题量不足。这种现象似乎在期中考试中更为普遍。显见的情形是考试时间过半,考生就开始交卷了。不少教师对此见怪不怪,似乎也不是原则问题。由于命题者即是阅卷者,大多教师都会将命题与评卷综合起来考虑。除了对试题质的方面的考虑,试题的数量与时间则是控制变量。减少试题必然省时省力,因而“偷工减料”便成了回避工作压力的捷径。而测试时间基本上按照常规而定,且执行中还可能酌情调整。实际上,试题数量和答题时间直接影响到测试的效度与信度。对于知识点而言(如语法和词汇),数量往往反映了覆盖面(coverage),测试的覆盖程度即是效度的一个重要指标;对于技能而言(如完成某项任务),量与时的变量都关系到答题(performance)的质量,此外还可能因受试者应试策略的不同而影响到测试的公平度。试题数量偏少往往会在增强答题的难度和偶然性的同时降低了测试的信度。因此在考试设计中,测试者务必用效度与信度的尺度来适度把握试题的数量与时间,尤其要避免题量过少和时间过多的问题。

4. 提示不明,节外生枝

与其他盲点相比,提示语的问题影响面相对较小,因此也更容易被人们忽视。试题提示语的作用在于告知受试者要做什么和如何做,必要时可提供范例加以说明。在英语专业成就考试中因提示不明而导致学生节外生枝的情形偶有发生。在通常情况下,此类问题多出现于某一特定的群体中。比如对于已经教过一年半载的学生,任课教师会认定学生熟悉并理解试题要求,因而忽略了提示语的清晰性。实际上这种认定毫无保障。尤其对于产生性运用试题,在缺乏范例的情况下,时常总有部分考生因误读或曲解提示语(一般均采用

英语表述)而影响了答题。比如要求将两个句子改写为主从复合句,结果变成了翻译或解释句子。虽然此类问题往往以个别形态出现,但它对考生的影响却是致命的,对测试信度的影响也是不容忽视的。此类错误本来是可以避免的,借用Hughes的话说,“命题者不该依赖于学生的心灵感应力来引发所希望的行为”。^[14]因此为了帮助受试者正确理解题意、熟悉测试的形式和答题要求,以避免答题失误,命题者理应提供明确的提示语和必要的范例,这无疑是提高测试信度的措施之一。

5. 试题分析表面化

对测试成绩缺乏深度分析是一个十分普遍的问题。通常情况下,教师在评卷、登分之后便上报、通报成绩,有时还会在评卷的同时让学生查卷并核实成绩。近年来随着评估的推广,教师手册中加上了分析报告。主要对分数的大致分布,如最高分、最低分、平均分等情况进行一般的统计和简要评述,大多属于表面文章,而对于更具意义的一些指标,如标准差(SD)、试题难易度(P)和区分度(D)等却很少涉及,更不用说测量误差和评分误差等内容了。其实关注成绩,本来无可厚非。无论对于学生还是教师,成绩总是一个敏感的指标。其原因不言自明,多少要归功于应试教育的大背景和分数驱动的管理机制。但问题是不可厚此薄彼,尤其对于成绩测试而言,对考分的关注显然不是目的所在。对于测试者和教学者,不仅要注意到学生的成绩,更重要的是通过深度分析发现测试背后所隐藏的信息。Hughes认为语言测试与教学的关系应该是一种伙伴关系(partnership)。^[15]不能指望测试只跟着教学走,应该要求它对好的教学发挥支撑作用,而在必要时对差的教学起到纠正作用。试题分析至少要取得两个方面的信息:其一,为学生提供诊断性意见。即让他们了解自己学习的现状,在看到进步的同时发现存在的问题,明确下一个阶段努力的方向。在可能的情况下,应该引导学生自我分析,了解个体差异,以便对症下药。其二,为教学提供指导性和纠正性意见。不仅要让教师了解学生的表现如何,更重要的是要反思此类问题,比如测试达到目的没有;就测试内容和题型,哪些有待改进;这种测试会产生哪些反拨效应;今后在教学方面应该注意哪些问题。可见加强测试分析有助于发挥积极的反拨效应,增进教学与考试的和谐发展,促进“教考合一”。^[16]

6. 将期末考试作为终结性评估

这是一个致命的盲点。所谓终结性评估其特征就是一次性检验,令人遗憾的是这种检验通常依赖于终结性测试。在不少高校,期末考试即作为课程终结性测试,将考试成绩(通常以分数的形式)作为评价学生学业的主要依据纳入学籍管理。虽然部分学校在此基础上加入了一定比例的半期考和平时成绩,但分数的决定性作用和主导地位并没有改变。这种“以分取人”的评价机制在客观上夸大了考试的功能,强化了应试教育,甚至导致了学生学习方法和思维方式的偏差,进而影响到语言综合能力的发展,引发重知识、轻技能和“高分低能”的现象,甚至还可能引发弄虚作假等不良行为。以测试取代评估是以偏概全。测试理论研究早已证明:测试并不

等同评估。作为评估的一个部分,测试的功能和作用是很有限的。^[12]而作为一种测量手段,并非所有的测试都具有评估性。把终结性测试作为教学评估的唯一或主要依据显然是不够客观和全面的。因此专家呼吁要“改变现有的考评形式,以确保将学生能力的提高落到实处”。^[13]

三、应对策略

英语专业测试中存在诸多盲点,究其原因大多在于教师,但管理层面亦有不可推卸的责任。客观地说,要从根本上排除这些盲点几乎是不可能的,但是只要采取相应措施,是完全可以有可能缩小测试中的盲区、减少它们给语言教学造成的负面影响的。

1.提高语言测试理论素养

在英语专业测试中,大多教师扮演的是多元化的角色。他们既是教学者,也是命题者,同时又是阅卷者和评估者。因此教师的素质将直接影响到测试的效度与信度。然而英语师资队伍的建设与教学的需求并不相适应。据专家调查,教师短缺和进修资源匮乏乃是目前英语教师队伍建设的最突出两个问题。^[14]近年就一些地方高校的大学英语教师素质的调查还表明:教师普遍欠缺语言测试基础理论知识。比如对测试效度、信度等基本问题大多是一知半解,因而在测试的各环节中出现种种盲点也不足为奇。^[15]要走出应试教育的种种误区,转变传统的测试观念,必要前提是理论武装。测试理论素养的提升必定增强测试者对各种盲点的辨别能力和透视力。比如对于效度与信度的关系、测试内容和方法与反拨效应的关系、如何提高测试的效度与信度以及如何挖掘正面反拨效应等问题的反思有助于减少测试实践中的盲目性。毫无疑问,加强教师语言测试基本理论素养对于保障语言测试的质量具有战略性意义。自我进修与在职培训相结合是实现这一愿望的重要手段。这既需要个人克服困难实现自我进修与发展,又需要职能部门创造条件尽可能提供脱产进修或短期培训的机会。

2.加强测试手段的改进与创新

在提高理论水平的同时,专业教师还必须与时俱进,掌握语言测试的新方法和新技巧。就实践层面而言,以下三个方向的努力是很有必要的。其一,要掌握基本原理和方法,不断改进测试手段。测试者必须在把握各类题型优劣的基础上扬长避短,深入了解外语测试改革的新动向和新方法,通过提高测试的真实性和互动性来加强对语言能力的测试。其二,应加强计算机素养(computer literacy)。由于计算机科学和信息技术的日新月异,许多教师多媒体和网络技术的应用能力已跟不上教学发展的步伐,因而不能很好地适应外语教学改革和发展的需要。^[16]因此加强多媒体和网络技术的学习,是网络时代对教师角色多元化的必然要求。这不仅有利于发挥计算机在测试中的优势,提高测试、阅卷、分析和反馈的速度和准确性,而且还可以拓宽客观性试题的范围,鼓励学生开展自我评价。^[17]其三,还需加强统计方法在语言测试

分析中的运用。在提高语言测试艺术性的同时,提高测试分析的科学性。比如通过差异分布和回归系数等检验手段纠正主观性测试的评分误差。^[18]总之,加强测试手段的改进与创新有助于理论与实践的结合,有利于拉近和缩短测试与教学的距离。

3.完善考试与评价机制

改变“以分取人”的终结性评估机制是大多高校亟待解决的重要课题之一。《英语专业教学大纲修订说明》明确提出:“各校在教学过程中,应根据教学情况和需要以及各门课程的特点和要求,在各个阶段采取多种形式对学生进行考核和检测。”^[19]同时鼓励“采取灵活多样的方式对学生的学习情况进行评估。”对此,学校管理层和职能部分理应承担更多的职责。他们在强调规范化管理的同时往往忽视了决策的科学性,比如现实中某些对测试及成绩的管理条例恰恰强化了分数的功能和作用,从而背离了以人为本的科学发展观。因此要借鉴多维度的测试与评价手段,首先要求管理层和测试者解放思想、转变观念。唯有在这个前提下,才可能打破常规以建构起更加完善的考评体系。目前国内外测试与评估理论的研究表明:形成性评估(formative assessment)可以弥补终结性评估的缺陷,从而实现优势互补。^[20]形成性评估具有重在过程,以学生为主体和以人为本的特点。这种评估方法着眼于学生的各种真实的课堂表现和课外学习活动,关注点从考试成绩转向了学习行为,将主要通过对学习行为的观察、信息收集、建立档案、分析反馈等综合手段,让学生通过自评、互评和评教等方式参与到教学的评价过程中,发现问题、进而改进学习过程以达到提高学习效果的目的。这提高了评估的客观性和透明度,为学生提供了更大的学习空间和自主性,也呼唤着学生更强的自觉性。它不仅可以摆脱应试教育的阴影、促进教育模式的革新,而且可以有效地促进教学相长,有利于学生语言综合能力的发展。无疑是一种更加开发的、更具活力的、更加客观全面的评估方式。^[21]

四、结 语

事实上,语言测试中的盲点远不止以上这些。对语言能力测试的探索是永无止境的,测试虽然不是万能的,但在现有条件下摒弃测试也是不够现实的。剖析英语专业测试中盲点的意义在于进一步认识语言测试的目的和本质特征,加强对测试效度与信度意义的认识,以助于专业教师发现自身测试实践中存在的问题和误区,减少盲目性,并逐步克服及排除各种盲点。通过教师测试理论素养的提高和测试手段的改进,通过管理层观念的转变和评估机制的完善,英语专业测试的效度与信度会得到有效的提高,测试的反拨效应也会得到优化,语言测试与教学的伙伴关系将更好地朝着“教考合一”的和谐境界健康发展。

参考文献:

[1]邹申.简明英语测试教程(第二版)[M].北京:高等教育出版

- 社,2005.
- [2]桂诗春.语言测试:新技术与新理论[J].外语教学与研究,1989,(3).
- [3]Bachman,L. F. & A. S. Palmer. *Language Testing in Practice*[M].Oxford, UK.: OUP,1996.
- [4]韩宝成,戴曼纯,杨莉芳.从一项调查看大学英语考试存在的问题[J].外语与外语教学,2004,(2).
- [5]Hughes, A. *Testing for Language Teachers*[M].Cambridge, UK.: CUP,1989.
- [6]李筱菊.语言测试科学与艺术[M].长沙:湖南教育出版社,1997.
- [7]Bachman,L. F. *Fundamental Considerations in Language Testing*[M].Oxford, UK.: OUP,1990.
- [8]戴培栋,张雪梅.对我国英语专业本科教学的反思[J].外语界,2007,(4).
- [9]岑建军.我国高校外语教学现状[J].外语教学与研究,1999,(1).
- [10]骆爱凤,叶张煌.高校在职外语教师职业发展情况调查与研究[J].外语界,2007,(4).
- [11]张宜,王新,郭威,于森.大学英语教师素质调查报告[J].外语与外语教学,2003,(10).
- [12]袁小陆.关于大学 E-Teacher 的角色思考[J].外语教学,2006,(3).
- [13]曾用强.论英语口语考试的评分误差[J].考试研究,2007,(3).
- [14]高等学校外语专业教学指导委员会.高等学校英语专业英语教学大纲[M].上海:外语教育出版社,2000.
- [15]Black, P., & William, D. Inside the black box: Raising standards through classroom assessment[J].Phi Delta Kappan, 1998,80(2).
- [16]Nicol, D.J. & Macfarlane-Dick, D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice[J].Studies in Higher Education, 2006,31(2).
- [17]丁黎明.论构建英语专业形成性评估体系的必要性[J].高教发展与评估,2008,(6).

责任编辑:曲晓红

Analysis of Blind Spots in English Language Assessment for the English Major

Ding Liming

(Department of Foreign Languages, Sanming University, Sanming365004, China)

Abstract: In English language assessment for the English major, there exist various blind spots, covering testing purpose, content, test-writing, scoring, test analysis and so on, which have affected the validity and reliability of tests and teaching and learning as well. Backed up by testing theories, this article provides some helpful analysis of the blind spots for language testers, and proposes three practical suggestions to improve the test quality.

Key words: the English major; language assessment; blind spots