

# 高考英语社会化:问题与出路<sup>\*</sup>

柯晓华<sup>1</sup> 曾用强<sup>2</sup> 罗海蛟<sup>1</sup>

**提要:** 党的十八届三中全会提出“深化教育领域综合改革”的总体要求,明确高考外语科目将实行社会化“一年多考”。这就迫切需要解决外语考试的公平公正性、导向作用和大规模考试评分的信度问题。本文分析了社会化考试的问题和实质,结合我国国情,提出两点改革思路:首先健全考试监管体制,出台中国教育考试标准、中国考试法,并成立考监局,推进管、考、用分离;其次构建社会化考试平台,从高考考题知识库、主观题自动评分和多次考试之间的分数换算等方面保证考试效度。

**关键词:** 社会化考试;一年多考;自动评分;分数等值;计算机化语言考试

**Abstract:** The Third Plenary Session of the Eighteenth CPC Central Committee raised the claim of deepening educational reform, establishing the social system of multiple tests of College Entrance Examination. However, three issues must be solved urgently as to the establishment of NMET: justice, guiding role and scoring reliability of large-scale examination. This paper analyzes the issue and innate character of social examination, and depicts two reform ideas based on China's national conditions: firstly, to improve the examination supervision system, to introduce novel educational examination standard and testing method, and to establish Test Supervision Bureau, so as to separate the power of administration, examination and test use; secondly, the social examination platform should be constructed, thus improving the assessment quality by using NMET knowledge base, automatic subjective/objective item scoring system, and score equivalent calculation between each examination.

**Key words:** social examination; multiple tests of College Entrance Examination; auto-scoring; score equivalent; computerized language testing

中国共产党十八届三中全会提出“深化教育领域综合改革”的总体要求,明确高考改革要探索外语等科目社会化一年多考。今年9月,《国务院关于深化考试招生制度改革的实施意见》强调外语科目将实行社会化考试和一年多考(国务院2014)。高考改革是牵涉到全社会的重大变革,要依据中国实情、遵循专业规范开展考试设计、实施和分数使用(桂诗春,杨惠中2014);有意识地让考试带来良好的导向作用,为教学、人才培养、社会负荷带来正能量,这是比考试本身更重要的目的。

以英语考试作为突破口,一个重要原因是英语的计算机化考试平台和主观题自动评分技术最为成熟,已被社会普遍接受,而且有些省市已经做过探索和试点。但是,高考英语要实现社会化一年多考,仍有若干问题迫切需要解决。例如,社会化考试是否需要政府的行政化资源?还是参考国外社会化考试的做法,完全由社会机构组织实施?目前,中国高考是行政化体制下考试,得以顺利实施是因为利用很多行政资源,比如命题老师封闭在一个地方出题、公安押送考卷、以政治任务来要求各高校改卷。本文从社会化考试的本质和特性提出高考英语社

\* 本文获广东省高校特色创新类项目(2014)“高考英语的计算机化考试方法及技术研究”的资助,特此感谢!

会化考试可能出现的问题,并依据我国实情,提出若干解决思路。

### 1. 高考英语社会化考试可能出现哪些问题?

考试是一种测量工具,以语言考试为例,是通过考察学生的行为,确定他们的语言能力,并以分数等形式反馈考试结果。研究考试就是探究其能否公平地、准确地开展测量工作,也就是考试的信效度问题,它们之间关系如图1所示。然而,在社会化背景下,考试被赋予多重社会属性,表现在社会资源对考试的作用、教育的不公平、考试的社会反拨效应等方面(桂诗春,杨惠中,2014)。可见,若高考英语变革为社会化一年多考,必然存在考试的信效度问题和考试的社会性问题。

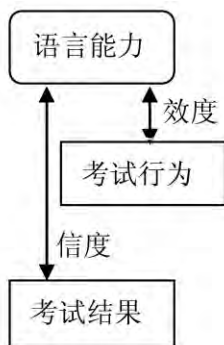


图1 信效度与考试

首先是社会化考试的信效度问题。(1)考试信度是指考试结果(即分数)的一致性、稳定性及可靠性。高考评分任务可以按照题型分为主观题和客观题两类,客观题由计算机评阅读题卡,信度十分高;传统的主观题评分任务由评分员在限定时间、地点内完成,由于评分员经过挑选和培训,他们所评的分数具有较高的信度。然而,当高考英语变为社会化一年多考,必定出现考卷数量的增加和考试地点的多样化,如何组织评分员是一个难题,如何保证所评分的信度是另一个难题。国外的托福、雅思和思科网络认证等社会化考试均采用计算机阅卷来保证信度。高考要实现计算机评阅主观题还要一段时间。另外,同一考生参加多次考试的分数要进行对比和换算,这也是考试信度问题。(2)考试效度是考试的准确性,它反映的是考试内容与考试大纲的吻合程度以及是否真实反映考生的能力,关键环节在于研制考题和实施考试等。过去,在行政主导的考试中,高考试题研制和考试实施都是由政府组织,以行政命令的方式让老师们定时定量完成出题、考务和阅卷任务,甚至动用警力维持秩序,保证考试顺利完成。然而,当高考转变为社会化考试,如何保证考试效度将成为一个重大问题。社会化组织是否一年多次请老师来研制考题?能否动用警力辅助考试?这些问题迫切需要解决。

其次是社会化考试的社会问题,这是指考试的社会性可能引发的问题,主要包括公平公正性、考试带来的负担、考试的导向等问题。(1)社会化考试首先要保证公平公正性。考试的实施和管理程序要提供公平的竞争环境,比如无年龄限制、具备完整的防舞弊程序、考试条件对城镇和农村的学生来说是公平的。社会化考试的组织方若缺乏监管,则试题研制和考试实施将可能引发道德问题、保密问题、安全问题等一系列问题。(2)考试负担有可能增长。高考改革的初衷是减负增效,通过考试社会化,使得学生有更多的学习机会和参加考试的机会;但是,由于考试的分数是高校选拔人才的主要依据,所以考生若第一次考试发挥得不好,必然要花费更多时间去反复备考。同时,由于学校的英语教师按照教学计划上课,不可能逐一辅导备考;学生一方面需要参加校外的辅导班,另一方面不得不参加多次考试以提高分数。因此,社会化一年多考未必能减轻英语学习负担,却有可能导致更严重的题海战术、应试教育。(3)反拨作用是考试效度不可分割的部分,高考是大规模高风险考试,对教学有着无可否认的导向作用。高考英语的正面反拨作用是:高中英语教学聚焦于语言运用能力,随之提高语言考试分数。然而,社会化一年多考有更多提高分数的机会,为了快速提高分数,英语教学可能专门操练语言知识和大搞题海战术,而这显然与考试改革的初衷背道而驰,也是社会化考试可能带来的社会问题。

## 2. 出路:考试监管的规范化和考试实施的专业化

就我国考试目前的情况而言,有效推进高考社会化改革,要从两方面入手:一是考试监管的规范化,二是考试实施的专业化。

考试监管的规范化是指从体制入手,推进管、考、用分离。首先,尽快出台中国教育考试标准和中国考试法,并成立考监局。教育考试标准由教育考试专家委员会制定,形成考试行业标准;考试法明晰考试过程中各方的权责;考监局要监督、帮助社会考试机构承担高考的复杂过程,包括试题研制、试卷安保、考试实施、试卷评阅等。可以借鉴美国和日本高校招考社会化的监管体制。

美国是当今世界上高等教育最为发达的国家之一,通过高度专业化的社会考试组织(教育考试服务中心和高校测试计划中心)为大学提供统一的招生考试服务,并接受市场的检验,通过竞争来不断提高考试的信度和效度,并依据大学学术要求不断调整考试内容(吴向明 2007)。大学作为考试服务机构的客户,可以选择是否使用入学考试成绩以及选择哪一种考试服务,选择考试成绩在录取参照中的比重等。著名的社会化考试包括:学术评估测验 SAT (Scholastic Assessment Test) 和高等学校检验 ACT (American College Test),前者由教育考试服务中心举办(该中心是全球最大的私人教育测验、评价与教育研究组织,它成功开发了很多标准化考试,包括托福、托业、GRE 等);后者由高等学校测试计划中心举办,是学术性向考试与学院成绩考试的综合考试。

日本与中国同属于儒家文化圈,其考试的选拔方式深受中国古代科举考试的影响,在日本学历主义一度盛行,以高校招生为例,也曾重视考试成绩,升学考试竞争激烈。但经过多次招考制度改革,目前形成了日本特色的“社会化”(王志武 2012):1990 年成立的“国家大学入学考试中心”是具有独立行政法人地位的专业考试机构,它既接受文部省的指导,又独立于文部省之外。国家大学考试中心负责提供专业考试服务,高校作为顾客,可以根据自身的特点和需求选择特定的考试科目及组合。考生也可以根据自身情况和学校的要求选择相应的科目组合参考。科目组合可以有效调节生源的市场,同时也是区分考生水平的有效手段,越是顶尖的高校,要求的科目越多越难,而一般的高校为了保证生源可能减少考试科目或降低难度。大学对考试中心考试成绩的使用上具有足够的自主权:可以直接使用其成绩作为入学唯一依据,也可以完全不使用该中心的考试成绩,或是选择考试中心的考试成绩与自主命题成绩按一定权重加权求和作为考生的最终入学成绩。

其次,中国的高考英语社会化考试必须推进管、考、用分离。(1) 考试管理者不直接参与考试实施:政府部门只负责投入和依法监管社会考试机构的科学办考,不干涉具体考试事务。由政府出资成立社会考试机构,考试由其进行开发和实施。(2) 社会考试机构具有独立的考试开发、组织、评估权力和资质:每个考试都有详细的手册来解释其分数含义、提供有关其效度和信度的数据;考监局应为其提供必要的智库资源和安保政策,但不宜干涉其试题研制、考试实施、评分和分数使用;而社会考试机构应该成立考试委员会,以认真负责的态度组织和实施考试;借鉴美国、日本等的社会化考试的经验教训,应有专门机构来提供性质相同考试的“对比报告”(Comparability Reports)供考试使用者参考,专门机构运用语言测试的信效度理论,持续研究高考社会化的试题设计、题库和试卷生产、信效度校验、社会导向作用、分数使用等过程。(3) 考试使用者不直接参与考试开发:考试使用者(如高等学校)在监督下决定怎样使用考试的结果、有权决定接受哪一个或几个考试成绩,如果觉得有需要进一步了解考生情况,可

选择加设考试项目(如面试)。

考试实施的专业化是指必须由专业人士保证高考的信效度、搭建高考英语的计算机化考试平台。高考英语的社会化一年多考意味着考试规模的增大、考卷设计和评分的任务加重。为此,高考英语的社会化一年多考必须采用计算机化考试模式,提高考试的信效度,研究思路如图2所示。

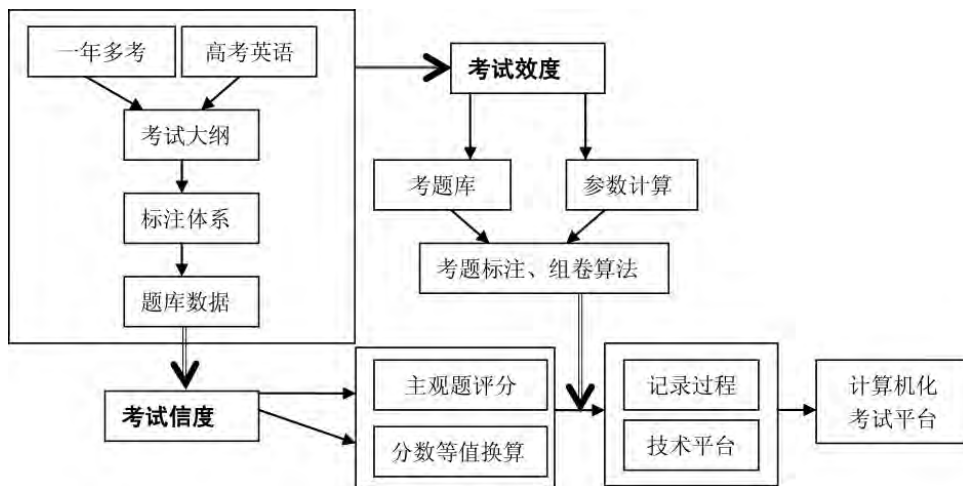


图2 高考英语改革的计算机考试研究思路

总的来说,高考英语改革“一年多考”的研究思路可分三步走:

(1)在厘清“一年多考”的基本模式、知识结构、考察目标等基础上,建立带项目参数的考题知识库、编制标注体系。对高考英语计算机化考试有参考价值的英语知识库包括 WordNet、FrameNet、VerbNet 和 PropBank 等。其中,FrameNet 最具应用价值,并被借鉴构造了很多英语知识库。FrameNet 是基于 Fillmore 框架语义学、借助词典释义和语料建成的知识库。“框架”是解释词义和组织词汇知识的基本手段。一个框架包含了若干“框架元素”,用以表征词汇之间的共现特征和相互关系,语义结构就蕴涵在这种关系中。框架语义描述基于大规模文本语料中抽取出来的例句,进行了手工标注的语义注释。因此,这个工程的核心是以机器可读的形式由人对语义知识进行编码。目前,FrameNet1.3 已构建 825 个语义框架,包括 1 万个词汇,其中 6100 多个被完全标注,另外还有 13.5 万句经过语义标注的例句库。FrameNet 在自然语言处理界得到了广泛的应用,尤其是它的例句库,可直接用来训练语义标注模型,也可用于机器翻译、问答系统和信息检索等高精度语言工程项目,同样也很适合英语语义分析,实现英语作文自动评分。

(2)从考试效度出发,明确考试的构念,即考试到底在考察什么、考试的目的所在。传统的纸笔考试只产出一个维度的信息,就是成绩,无法记录考生的答题过程。计算机化考试并非简单地把纸张考试搬到屏幕上,它能记录和分析考生思维过程、不同语境下的语言运用、考生作答题目的每个行为、决策变化和时间间隔等信息(曾用强 2010),运用多维度数据描述考生的语言能力,产出考生答题过程记录,提高考试的效度(柯晓华,曾用强 2012)。这样一来,不同的考生即使回答了同样的测试项目,通过分析考试过程的数据记录,他们的语言能力依然具有可比性,使考试的效度得到保证。

(3)从考试信度出发,针对主观题自动评分和分数转换开展研究,提出具体算法、技术、实

验结果、测试平台,形成计算机化考试平台。计算机化考试一方面运用主观题自动评分软件快速准确地完成评卷任务,研制高准确率的汉语和英语主观题自动评分软件。以知识库为基础,提取训练集,进行语义标注、关键信息识别、多种评分模型的选取和自适应调整,结合文本形式特征,构建评分模型;以英语作文自动评分为例,分别从作文的语言形式、思想内容评判作文质量,辅以人工评分,提高高考英语作文评分信度。另一方面借助项目反应理论和项目参数(比如难度和区分度等)实现同一考生的多次考试之间的分数等值换算。构建带项目参数的高考考题知识库,根据每道题的难度系数,辅以区分度和猜测系数,建立题目特征函数,对同一考生的多次考试成绩进行换算,客观反映多次考试的分数可比性,实现多次考试之间的难度对等计算(即分数等值换算)。

### 3. 结论

本文在语言考试的信效度理论和社会属性基础上,分析高考英语改革、社会化一年多考可能出现的问题,借鉴国外社会化考试的模式和特点,尝试提出我国高考英语社会化考试的两点思路:考试监管的规范化、考试实施的专业化。高考英语改革不仅要适应社会进步,还要促进我国的国家外语能力、完善外语教育体系。与此同时,考试研究者应该以考生和育人的需求为导向,探索适合我国国情的外语考试模式,实现高等教育资源充分合理配置、中学英语教育效率最大化目标的机制。

### 参 考 文 献

- [1] 国务院. 国发[2014]35号. 国务院关于深化考试招生制度改革的实施意见[Z/OL]. [http://www.gov.cn/jzhengce/content/2014-09/04/content\\_9065.htm](http://www.gov.cn/jzhengce/content/2014-09/04/content_9065.htm) 2014-09-19.
- [2] 桂诗春, 杨惠中. 从社会视角看中国考试的过去和现在[M]. 科学出版社, 2014(即将出版).
- [3] 柯晓华, 曾用强. 基于阅读行为变量的阅读策略对比研究[J]. 外语测试与教学, 2012, (4): 42-50.
- [4] 王志武. 社会考试社会化发展的模式与机制研究[J]. 中国考试, 2012, (8): 39-45.
- [5] 吴向明. 美国高等学校招生制度的特点及启示[J]. 大学研究与评价, 2007, (4): 395-400.
- [6] 曾用强. 对计算机化考试的几点思考[J]. 外语电化教学, 2010(1): 131-135.

作者联系方式: 1. 广东外语外贸大学思科信息学院, 广东广州 510420

2. 广东外语艺术职业学院, 广东广州 510640

## “全国高校外语学刊研究会 2014 年年会”在天津召开

2014 年 9 月 26-27 日,“全国高校外语学刊研究会 2014 年年会”在天津外国语大学召开。会议伊始,天津外国语大学校长修刚教授代表东道主向与会的 20 多家高校外语学刊近 50 位代表致欢迎辞。研究会会长庄智象教授在随后的发言中指出,各高校外语学刊应以外语类学科国家标准的出台为契机,调整办刊定位,坚持学术品位,办出多元特色,同时确保和提升刊物的学术与编校质量,促进学术成果交流和先进文化传播,为外语人才培养、学校发展、学科建设做出学刊应有的贡献。庄会长还要求各学刊重视编辑团队的培养,尤其要注重年轻编辑的培训,以保证学术期刊健康、可持续发展。在会议进行过程中,与会代表就规范办刊、提升办刊质量、办出学刊特色、防止编辑过程中可能遇到的学术剽窃和抄袭及“一稿多投”等问题以及业内的一些新动向进行了热烈交流和讨论。