



# 汉语水平考试(HSK2.0) 反拨效应的多维度探析

周琳 肖媛 郑洁

**摘要:**通过混合研究方法对 HSK2.0 在国际中文教育中的反拨效应进行的多维度探析发现:HSK 对学习者的汉语学习产生了积极影响,影响强度较大,影响的长期性和阶段性共存,考试内容和形式使考生对语言综合能力的提升有所忽视;HSK 对教师教学产生了积极影响,影响强度低于学生,题目设置对教师 HSK 辅导产生了一定影响,但“以考促教”的理念在常规课程教学中落实不够理想;考试与目前国际中文教材、教学内容的契合度有待提高;考试辅导材料的“应试”的色彩较强且需要更新。应发挥 HSK 评价导向作用,激发学习者汉语学习的内生动力;优化考试资源体系,推动教、学、评相融合。

**关键词:**反拨效应;语言测试;汉语水平考试(HSK)

**DOI:** 10.13734/j.cnki.1000-5315.2023.0610

**收稿日期:**2022-09-20

**基金项目:**本文系北京市社会科学青年基金项目“基于动态系统理论的汉语二语学习者心理词汇发展研究”(17YYC018)和北京语言大学梧桐创新平台项目(20PT01)的研究成果。

**作者简介:**周琳,女,天津人,文学博士,对外经济贸易大学中国语言文学学院副教授,主要研究方向为第二语言习得、国际中文教育,E-mail: linzhou0328@qq.com;

肖媛,女,辽宁抚顺人,汉考国际教育科技(北京)有限公司研发人员;

郑洁,女,河北张家口人,对外经济贸易大学中国语言文学学院研究生。

## 一 引言

1984年,原北京语言学院(现北京语言大学)开始研制汉语水平考试(HSK)。1992年,HSK正式升级为国家级考试。至1997年,HSK形成了包含基础、初中等和高等在内的完整考试体系,我们将之称为HSK 1.0。2004年,为更好地满足海外不断增长的汉语学习新要求,原孔子学院总部/国家汉办组织专家,在充分调查海外汉语教学实际情况的基础上,借鉴国际语言测试研究新成果着手研发新考试,于2009年推出了新HSK(一至六级)并推行至今,我们称为HSK2.0。HSK2.0秉承“以考促学、以考促教”的理念,在国际中文教育事业发展中发挥了极为重要的作用。2021年3月24日,教育部、国家语言文字工作委员会发布了《国际中文教育中文水平等级标准》(以下简称“《等级标准》”),提出了“三等九级”的新框架<sup>①</sup>,为国际中文教育、测试与评估等提供了新的参考依据。因应新形势,HSK 3.0研发也提上了日程,它将依据《等级标准》的新框架,在HSK2.0六个级别的基础上向九个级别延伸并进行革新。在这样的背景下,对HSK 2.0的反拨效应进行深度分析,可为国际中文教育、HSK 3.0的革新、国际中文考试材料编写等提供一定参考。

## 二 研究综述

<sup>①</sup>中华人民共和国教育部、国家语言文字工作委员会《国际中文教育中文水平等级标准》(GF0025-2021),2021年3月24日发布,2021年7月1日实施,第1页。

对反拨效应(washback)的认识可追溯到20世纪五六十年代,但深入的研究则始于20世纪90年代。有学者将反拨效应定义为“考试使教师和学生去做在没有考试的情况下不一定会做的事”,提出了15个反拨效应假说并呼吁展开实证研究<sup>①</sup>。后来,在应用语言学领域,考试对教与学的影响被定义为反拨效应,研究成果包括理论探究与模型建构以及实证研究,如:反拨效应作用机理、影响反拨效应的因素、如何规避负面反拨效应并提高正面反拨效应等<sup>②</sup>,研究对象主要是TOEFL、IELTS等。

比较而言,中国大陆的反拨效应研究起步较晚,但对其理论模型建构及一些重要语言测试反拨效应的实证研究都有所涉及<sup>③</sup>,不过这些成果主要针对一些高风险英语测试,对有“汉语托福”之称的HSK(特别是HSK2.0)反拨效应的研究则相对匮乏。我们以“HSK”和“反拨效应”为主题词在CNKI数据库中检索然后对研究内容进行筛选,发现以HSK2.0为研究对象的期刊论文仅6篇。其中仅孔傅钰、张艳莉的研究是目前规模较大、较为系统的实证研究。前者针对HSK2.0对国内高校汉语教师教学的反拨效应进行了专题研究<sup>④</sup>;后者从学生、教师、教学管理者三方面入手,更为全面地考察了HSK2.0的反拨效应<sup>⑤</sup>。这两项研究还对影响反拨效应强度的因素进行了分析。从复杂动态系统理论(Complex Dynamic Systems Theory)的视角来看,反拨效应是极为复杂的现象,宜从多个维度进行探究<sup>⑥</sup>。对反拨效应的维度描述最全面且具影响力的是Watanabe提出的性质(Value)、强度(Intensity)、长度(Length)、特定性(Specificity)、意图(Intentionality)五个维度<sup>⑦</sup>。然而国内研究大多不区分反拨效应的维度或仅关注某一维度(如强度),鲜有研究从多个维度综合探究HSK2.0的反拨效应。鉴于此,本研究通过问卷调查和个人访谈相结合的方法对HSK2.0在国际中文教育中的反拨效应进行多角度探析。

### 三 研究设计

#### (一)研究问题

1.HSK2.0对汉语学习者的学习、国际中文教师的教学以及考试辅导所产生的反拨效应在不同维度有何种表现?

2.影响不同维度反拨效应的因素有哪些?

3.反拨效应对国际中文教学及HSK3.0的革新有何启示?

#### (二)反拨效应维度的界定

根据Watanabe提出的五个维度,对反拨效应的五个维度的界定如下:

1.性质:反拨效应是正面的还是负面的。

①J. Charles Alderson, Dianne Wall, "Does Washback Exist?," *Applied Linguistics* 14, no. 2 (June 1993): 115-129.

②J. Charles Alderson, Liz Hamp-Lyons, "TOEFL Preparation Courses: A Study of Washback," *Language Testing* 13, no. 3 (November 1996): 280-297; Liying Cheng, Yoshinori Watanabe, & Andy Curtis, *Washback in Language Testing: Research Contexts and Methods* (Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers, 2004), 3-210; Liying Cheng, *Changing Language Teaching through Language Testing: A Washback Study* (Cambridge: Cambridge University Press, 2005), 3-255; Dianne Wall, *The Impact of High-Stakes Examinations on Classroom Teaching: A Case Study Using Insight from Testing and Innovation Theory* (Cambridge: Cambridge University Press, 2005), 1-284; Anthony Green, *IELTS Washback in Context: Preparation for Academic Writing in Higher Education* (Cambridge: Cambridge University Press, 2007), 1-315; Chih-Min Shih, "A New Washback Model of Students' Learning," *The Canadian Modern Language Review* 64, no. 1 (September 2007): 135-162; Qin Xie, Stephen Andrews, "Do Test Design and Uses Influence Test Preparation? Testing a Model of Washback with Structural Equation Modeling," *Language Testing* 30, no. 1 (January 2013): 49-70.

③金艳《大学英语四、六级考试口语考试对教学的反拨作用》,《外语界》2000年第4期,第56-61页;黄大勇、杨炳均《语言测试反拨效应研究概述》,《外语教学与研究》2002年第4期,第288-293页;辜向东、彭莹莹《大学英语教师对CET及其反拨效应认识的历时研究》,《外语与外语教学》2010年第6期,第37-41、56页;亓鲁霞《语言测试反拨效应的近期研究与未来展望》,《现代外语》2012年第2期,第202-208页;徐倩《英语专业八级的反拨作用研究——对外语专家和英语学科负责人的一次调查》,《外语界》2012年第3期,第21-31页;刘戈《论语言测试中的反拨效应及研究方向》,《高教探索》2014年第4期,第64-68页;董曼霞《语言测试反拨效应研究亟需厘清的几个基本问题》,《外语教学理论与实践》2019年第3期,第50-57页。

④孔傅钰、张艳莉《汉语水平考试(HSK)对国际汉语教学的反拨效应研究》,《天津师范大学学报(社会科学版)》2021年第4期,第46-51页。

⑤张艳莉、孔傅钰《高风险语言测试的反拨效应:对考教学一体化的启示》,《外语电化教学》2021年第3期,第76-82、108、12页。

⑥王初明、亓鲁霞《从动态系统理论视角看语言测试的反拨效应》,《山东外语教学》2016年第04期,第35-42页。

⑦Yoshinori Watanabe, "Methodology in Washback Studies," in *Washback in Language Testing: Research Contexts and Methods*, 19-36.

2.强度:反拨效应在教学某一或某些领域的影响强弱程度。

3.长度:考试影响所延续的时间,可分为短期和长期。

4.特定性:反拨效应可分为普遍效应和特定效应,前者指任何考试都会产生的反拨效应(如:考试会激发学生努力学习、会引起相应的备考行为),后者指某具体考试或考试某一方面所带来的影响(如:考试新题型增加对教和学所产生的影响)。

5.意图:指考试设计者或决策者对反拨效应的预期,可分为期望反拨效应和期望外反拨效应。前者指考试设计者或决策者所期望的反拨效应,后者指超出预期的反拨效应。

### (三)研究方法

本研究运用问卷调查进行定量分析,通过对教师和学生的深度访谈、问卷中的开放式问题以及考试研发人员内省的方法进行定性分析。采用多元测量方法,在定量统计的基础上,辅以定性分析来解释和验证量化结果,力求使多源数据和证据有效整合、彼此印证。

#### 1.问卷调查

##### (1)问卷设计

本研究参考以往研究对雅思考试反拨效应研究的学习者调查问卷<sup>①</sup>,结合与一线国际中文教师、汉语学习者、汉语考试研究人员的交流心得,围绕反拨效应的五个维度,针对学生和教师各设计了一套问卷初稿,部分内容采用李克特五级量表,1—5分别代表不同意/不符合、有些不同意/有些不符合、不确定、同意/符合、非常同意/非常符合。

我们先请调查对象之外的5名学生和3名汉语教师进行试测,根据试测者的反馈修改了部分项目,然后请一位语言测试专家与我们共同审读问卷,经商讨后认为问卷具有较好的内容效度。

学生问卷包括:基本信息、备考 HSK 情况、对 HSK 的认识和开放式问题。教师问卷包括:基本信息、对 HSK 的认识与感受、开放式问题。

##### (2)调查对象

学生问卷的调查对象为参加过 HSK2.0 的国内高校留学生和海外汉语学习者。本研究力求更为全面地考察 HSK2.0 的反拨效应,因此尽可能多地寻找不同国别的考生参与调查。被调查者来自韩国、越南、泰国、缅甸、日本、法国、俄罗斯、哈萨克斯坦、印尼、美国、马来西亚、埃及、巴基斯坦、阿塞拜疆等 22 个国家;年龄在 17—57 岁之间,平均 23 岁,女性占 51.86%,男性占 48.14%;62.69%为本科生、23.13%为研究生、5.97%为中学生、8.21%为其他学历。61.19%的学生参加 HSK 两次及以上。教师问卷的调查对象为国内高校国际中文教师、海外本土专/兼职汉语教师和外派汉语教师志愿者,年龄在 30 岁以下的占 45.05%、30—40 岁占 22.52%、40—50 岁占 26.13%、50 岁及以上占 6.31%,授课类型涵盖综合课、听说读写各项技能课、HSK 辅导课等。

##### (3)数据收集

使用问卷星正式发放问卷。共回收有效学生问卷 276 份(国内高校留学生 142 份、海外汉语学习者 134 份),有效教师问卷 157 份(国内院校教师 59 份、海外院校教师 98 份)。使用 SPSS 19.0 对量表数据进行分析显示,学生问卷的 Cronbach's  $\alpha$  系数为 0.814, KMO 值为 0.788, Bartlett's 球形检验显著( $p < 0.05$ );教师问卷 Cronbach's  $\alpha$  系数为 0.920, KMO 值为 0.886, Bartlett's 球形检验显著( $p < 0.05$ ),说明此次问卷调查具有较好的信度和结构效度。

#### 2.深度访谈

本研究采用半结构化访谈,在既有访谈大纲下对两名学生和三名教师的谈话进行引导与控制,同时也让受访者有一定的自由度,以便发掘更多信息。两名学生中一人曾参加 HSK 三级和五级,另一人曾参加 HSK 五级和六级。选择多次参加 HSK 的学生是为了考察 HSK 反拨效应的发展变化。三名教师中一人为国内高校专职汉语教师,教龄 10 年,参与过 HSK 教辅材料的编写工作;一位为公派志愿者教师,教龄 3 年,在海

<sup>①</sup>Anthony Green, *IELTS Washback in Context: Preparation for Academic Writing in Higher Education*, 323-349.

外任教期间曾参与 HSK 辅导工作;一位海外本土专职汉语教师,教龄 23 年。

### 3. 考试研发人员内省

由 HSK 研发者结合学生和教师的问卷调查数据、开放式问题和访谈结果进行内省。

## 四 调查结果

### (一) 对学习者的反拨效应

我们根据学生版问卷中量表题的探索性因子分析结果,对题项所反映出的共同特征进行归纳,从以下方面分析 HSK2.0 对学习者的反拨效应。

#### 1. 学习者对 HSK 的态度和感受

调查显示,学习者对 HSK 重要性评价的平均分为 4.3,有 88.06%的人选择 4 分及以上。“备考 HSK 让我更喜欢学习中文”和“HSK 对我的影响是积极的”这两个项目的平均得分分别是 3.67 和 4.04。学习者对 HSK 考题的感受和考试与日常学习关系的看法详见表 1。

表 1 学习者对 HSK 的态度和感受调查结果

调查项目	评分					平均分
	1	2	3	4	5	
我认为 HSK 重要	2.24	2.24	7.46	44.03	44.03	4.3
我认为老师上课应多讲 HSK 相关内容	6.91	10.45	32.09	26.87	23.88	3.53
我认为有必要开设专门课程讲解 HSK	1.49	5.22	24.63	35.07	33.58	3.94
我认为上课使用的教材应该与 HSK 相关	5.97	10.45	24.63	33.58	25.37	3.62
我参加的 HSK 题量大	3.73	4.48	30.6	42.54	18.66	3.68
我参加的 HSK 难度大	3.73	18.66	27.61	38.81	11.19	3.35
准备 HSK 的过程让我更加喜欢学习中文	6.71	5.22	29.85	35.07	23.13	3.67
HSK 对我的影响是积极的	2.24	3.73	17.91	40.3	35.82	4.04

#### 2. HSK 对汉语学习的影响

60%以上的学生都会在考前花更多时间学习汉语,并根据 HSK 要求制定学习计划、调整方法、策略和学习内容等,详见表 2。

表 2 HSK 对学习者的汉语学习影响调查结果

调查项目	评分					平均分
	1	2	3	4	5	
我根据 HSK 考试内容制定学习计划	4.48	7.46	20.9	50.75	16.42	3.67
为了准备 HSK 我会花更多时间学习汉语	5.23	5.97	14.18	44.77	29.85	3.9
我根据 HSK 内容调整学习方法和策略	2.24	7.46	24.63	47.01	18.66	3.72
我会在考后分析每项成绩	9.70	8.96	20.9	42.54	17.91	3.52
我会根据 HSK 成绩调整学习计划	5.23	7.46	19.4	45.52	22.39	3.74
我会根据 HSK 考试成绩调整学习内容	2.99	11.19	22.39	44.03	19.4	3.66

#### 3. HSK 对学习者的反拨效应不同维度表现

(1)性质:调查显示,58.2%的学生对“备考 HSK 让我更喜欢学习中文”评分在 4 分及以上,76.12%的人

对“HSK 对我的影响是积极的”评分在 4 分及以上。大部分学生认为考试可以检验汉语水平,学习目标也更明确了。考试也促使学生采用不同策略学习。对一位参加过不同等级 HSK 的学生访谈显示,随着 HSK 级别的升高,其所采用的学习策略也逐渐多样。参加 HSK3 时学习策略较为单一,主要是背诵课文和造句;参加 HSK5 时,练习方法更加丰富,更倾向于分析句子语法、语篇的衔接等。可以说 HSK 对学习者的整体影响是积极的。

(2)强度:学习者对 HSK 重要性评价的平均分为 4.3,有 88.06% 的人选择 4 分及以上(见表 1)。在参加 HSK 的原因(多选题)中,“学校要求”和“获得更多就业机会”分别占 57.47% 和 29.85%。61.19% 的被调查者多次参加 HSK,这些人中选择“工作/学习需要”和“没取得理想成绩”的分别占 41.46% 和 28.05%。因为 HSK 成绩已成为外国学生来华留学、申请奖学金的必备条件,并被越来越多的跨国企业作为员工招聘和晋升的重要依据,也是外国人来华工作、申请永居权的语言证明。毋庸置疑,对学习者来说 HSK 是一种高风险语言测试以至于本次调查中近 60% 的学生赞同或非常赞同他们的课程、教材应与 HSK 相关(见表 1)。可见,HSK 对学生的汉语学习的影响强度较大。

(3)长度:HSK 对学生的影响具有长期性和阶段性共存的特点。调查显示,近 70% 的学生会在考前延长学习汉语的时间并根据考试内容制定学习计划和选择学习策略,表 2 前三个关于考前准备的调查项目中,评分在 4 分及以上的占比分别为 67.17.36%,74.62% 和 65.67%;后三个关于考后行为的调查项目中,评分在 4 分及以上的占比分别是 60.45%,67.97% 和 63.43%。这说明,HSK 考试的影响贯穿考前和考后,考试结束后很多考生会根据自己的成绩对后续的汉语学习做出调整。不过,从访谈对象处获知,也有很多考生在考前 3—5 周延长学习时间,而考试后会立刻放松学习。

(4)特定性:同其他高风险考试一样,HSK 促进了学生努力学习汉语、引起了相应的备考行为。同时,HSK 2.0 的题目和考试体系设置也带来了一些特殊的影响。从学生备考的内容(多选题)来看,85.07% 的被调查者选择阅读练习,78.36% 选择写作练习,75.37% 选择听力练习、47.76% 选择口语练习、3.73% 选择了其他。被调查学生中参加过 HSK 初、中、高级考试的人数分别为 7.46%,23.88% 和 13.43%,比例较低。由于 HSK 2.0 分笔试(1—6 级)和口试(初、中、高级)两部分,笔试和口试相互独立,笔试成绩对学生影响较大,口试成绩则非必须,这使得很多考生重视笔试而忽视口试。

(5)意图:将“以考促学”作为理念的 HSK 为激发学习者学习汉语的热情、减少畏难情绪而将 HSK2.0 设计为阶梯式分级考试,以降低难度,且未对口语测试做硬性规定。问卷和访谈显示,现有考试体系造成考生对口语能力的忽视,还有相当一部分学生表示虽已通过 HSK 六级并获得了不错的成绩,但在真实交际中仍然存在障碍,感到考试成绩与实际水平有些不匹配。这些问题可能都是考试开发者未曾预料到的。

## (二)对教师教学的反拨效应

### 1. 教师对 HSK 的态度和感受

此类题目中,选择“赞同”(4 分)和“非常赞同”(5 分)的比重均超过 70%,平均在 4 分左右。绝大多数教师都认为 HSK 对学生的汉语学习和教师的教学很重要并产生了积极影响(见表 3)。

表 3 教师对 HSK 的态度和感受调查结果

调查项目	评分					平均分
	1	2	3	4	5	
HSK 对学生汉语学习有重要影响	2.7	3.6	11.71	46.85	35.14	4.08
HSK 对教师的汉语教学有重要影响	0	7.21	18.92	48.65	25.23	3.92
HSK 起到了检验学生汉语学习效果的作用	1.8	1.8	11.71	54.95	29.73	4.09
HSK 对教学的整体影响是积极的	1.8	2.7	22.52	46.85	26.13	3.93

### 2. HSK 对教师教学的影响

虽然教师普遍认为 HSK 对汉语教学十分重要,但 HSK 对教师教学影响的各项评分的均值都在 3 分左右,根据 HSK 调整教学内容、进度、顺序和方法的调查结果中,选择“符合”(4 分)和“非常符合”(5 分)的总比重介于 23.43%—56.76%之间(见表 4)。其中,仅“会将所授课程与 HSK 内容结合”、“会为学生安排 HSK 模拟测试”和“会根据 HSK 的考试情况进行反思”三个项目评分在 4 及以上的比重超过半数,分别是 56.76%,50.45%和 56.76%。

表 4 HSK 对教师教学影响调查结果

调查项目	评分					平均分
	1	2	3	4	5	
您经常向学生介绍 HSK 的相关信息	7.21	15.32	27.93	33.33	16.22	3.36
您会将所授课程与 HSK 内容结合	7.21	14.41	21.62	39.64	17.12	3.45
您会布置与 HSK 相关的作业	14.41	17.12	25.23	30.63	12.61	3.1
您安排的测试与 HSK 题型相似	8.11	15.32	28.83	28.83	18.92	3.35
您会为学生安排 HSK 模拟考试	14.41	14.41	20.72	30.63	19.82	3.27
您对学生的 HSK 考试情况做过调查	17.12	16.22	18.02	32.43	16.22	3.14
您会根据学生 HSK 考试内容调整教学内容	9.91	16.22	38.74	24.32	10.81	3.1
您会根据学生 HSK 考试内容调整教学顺序	17.12	19.82	31.53	27.03	4.5	2.82
您会根据学生 HSK 考试内容调整教学进度	11.71	19.82	45.05	17.12	6.31	2.86
您会根据学生 HSK 考试内容调整教学方法	13.51	14.41	29.73	33.33	9.01	3.1
您会根据学生的 HSK 考试情况进行反思	10.81	9.01	23.42	41.44	15.32	3.41

我们还对国内和海外任教的教师在上述项目上的打分进行了独立样本 t 检验,未发现显著差异。对三位教师的访谈显示,由于完成教学计划是首要任务,教师不可能为学生的考试而改变教学计划,不过会适当放缓教学进度配合学生的备考需求,在这样的情况下,他们认为学生参加 HSK 辅导十分必要,这与问卷调查的结果一致(见后)。国内专职汉语教师提到,部分学生比较看重 HSK 成绩,而能否通过较高等级的 HSK 存在一定偶然性,通过高等级考试不一定代表拥有高水平,考试成绩和实际水平之间的偏差有时会让教师不能客观认知自己的汉语水平,还有一些学生认为通过 HSK 考试比日常的汉语学习更重要,将更多精力放在了应试上。被访外派教师志愿者教师认为 HSK 对教学影响不大,因为班上大部分学生是社会人士,不需要参加 HSK 考试,另一些学生会自己准备 HSK 考试,其所就职的学校也并未将 HSK 教学列为教师的工作内容。教龄 23 年的海外专职汉语教师表示,如果教材和 HSK 考试是依据同一个等级标准编写的,那么可以认为 HSK 与汉语教学间接相关,但具体到他自己的教学则并未受到 HSK 的影响。

### 3. HSK 对教师的反拨效应不同维度表现

(1)性质:从教师对“HSK 对教学的整体影响是积极的”的平均打分来看(3.93),HSK 对教师的反拨效应是积极的。绝大多数教师认为 HSK 起到了检验汉语学习效果的作用(4.09)。也有部分教师在开放式问题中表示,HSK 有些侧重考察语言知识,对实际语言运用能力的考察力度不够,使得学生要求教师增加语言知识讲解减少语言技能实练。

(2)强度:HSK 对教师影响的强度更多体现在主观认知上,客观上对教学的影响强度并不大,至少不如对学生产生的强度大。这是因为教学单位一般不会将学生 HSK 考试情况作为教师考核的内容,对多数教师来说,HSK 不属于高风险考试。虽然教师普遍认为“HSK 对教师的汉语教学有重要影响”(3.92),但 HSK 对教师教学的实质影响并不大,各项平均分在 3 分左右(见表 4),相当于“不确定”。得分最高的项目

是“您会将所授课程与 HSK 内容结合”(3.45),得分最低的项目是“您会根据学生 HSK 考试内容调整教学顺序”(2.82)。在选择“不会因 HSK 而调整教学内容”的教师中,63.93%认为是“教学计划有严格要求”造成的。

(3)长度:从调查来看,HSK 对于专门开设 HSK 课程的教师来说影响时间较长,对其他教师来说很少有长期影响。

(4)特定性:HSK 题目设置对教师的 HSK 辅导内容产生了一定影响。本次调查的教师有 35.05%教授 HSK 辅导课,分析他们对“您认为 HSK 对您的教学还产生了哪些影响”这一开放问题的回答发现,由于 HSK1-6 级题型多为选择题,教师在授课时也大量练习选择题而较少使用开放性问题。

(5)意图:从本次调查来看,HSK“以考促教”的理念在常规课程教学中并未得到很好的落实。此外,HSK1-6 级均包含阅读题,体现了考试对阅读能力的重视,教师也会花大量时间进行阅读训练,但部分教师认为,HSK 的阅读部分让学生惯于选择答案,对阅读理解能力的提升作用有限。此外,HSK6 级写作测试所采用的读后缩写题型使得考生更侧重训练记忆原文的能力而忽略了独立式写作训练。

### (三)对教材和考试辅导的反拨效应

学生问卷中“教材应该与 HSK 相关”这一项目的平均分为 3.62。“有必要开设专门课程讲解 HSK”的平均分为 3.94,68.65%选择 4 及以上。被调查学生中有 45.52%参加过 HSK 辅导班,这些人对“辅导班对 HSK 考试有帮助”的平均打分为 4.36,选择 4 分及以上的占 86.88%。可以看出,教材与 HSK 的相关与否在学生看来并不特别重要,但多数学生认为应开设 HSK 辅导课,参加过辅导课的学生大多认为对考试有帮助。有些学生表示,目前的考试辅导材料比较陈旧,而且 HSK 相关的辅导材料十分缺乏,这也在一定程度上影响了参加 HSK 的意愿。

教师问卷中“所使用的教材与 HSK 相关程度高”的平均分是 3.41,多数教师认为上课所用教材与 HSK 考试契合度一般。“学生需要上 HSK 辅导课”的平均分为 3.39。有教师表示,HSK 是水平测验,上课时应着重提高学生的语言综合运用能力,而不应过多关注考题和应试技巧,所以辅导班不是必需的。可以看出教师对 HSK 辅导课的认同度低于学生。通过访谈有编写 HSK 辅导教材经验的教师了解到,HSK2.0 希望把汉语教学从“知识传授”向“能力训练”引导,因此与 HSK 1.0 相比更重视对交际能力和实际语言应用能力的考查。然而有 HSK 辅导经验的教师则表示,其所使用的辅导教材倾向于考试相关语言点和题型的练习,“应试”的色彩仍然很强。

## 五 启示与建议

### (一)发挥 HSK 评价导向作用,激发学习者汉语学习内生动力

在本研究调查范围内,学习者普遍认为 HSK 是非常重要的考试,对自己的汉语学习产生了积极影响。不过学生多因学业和就业的需求参加 HSK,从动机来源上看,大部分考生动机属于被动性、诱发性的外部动机(extrinsic motivation),而非主动性、自发性的内部动机(intrinsic motivation)。一些因升学或工作压力而参加 HSK 的考生,如未取得理想成绩可能直接放弃学习汉语,单纯依靠外部动机难以使学生保持学习汉语的长久动力。而且外部动机过强的学生更为关注应试而忽视了汉语综合运用能力的提升,也与 HSK“以考促学”的理念有所偏离。应该认识到激发学习者汉语学习的内生动力是使其葆有汉语学习热情的源泉。如何更有效地通过 HSK 促进汉语学习,激发学习者的学习动机是需要思考的问题。HSK 2.0 在学习者心中是以结果性评价为主的高风险语言测试,而考试本身对学习者的汉语水平的诊断与反馈作用相对缺乏,这就导致学习者在备考前后花费较多精力,着重练习考试题型,提升应试技巧,而并未将 HSK 备考与日常汉语学习真正结合起来。2020 年 10 月,中共中央、国务院印发《深化新时代教育评价改革总体方案》,指出要“充分发挥教育评价的指挥棒作用”,“改进结果评价,强化过程评价,探索增值评价,健全综合评价”<sup>①</sup>,为各级各类评价改革提供了指南。HSK 改革正当其时,通过加强对评价结果的解释与运用,强化对学习过程的及时诊断

<sup>①</sup>中共中央、国务院印发《深化新时代教育评价改革总体方案》,中华人民共和国教育部官方网站,2020 年 10 月 13 日发布,2022 年 12 月 25 日访问,http://www.moe.gov.cn/jyb\_xxgk/moe\_1777/moe\_1778/202010/t20201013\_494381.html。

和动态反馈,可以更有针对性地指导学习者查漏补缺,提升学习者汉语学习成效,这样不仅可将“以考促学”理念落到实处,也有利于将外部动机转化为内部动机,培养终身学习的汉语学习者。

要想更好地发挥 HSK 评价导向作用,首先需要增强 HSK 的即时诊断与反馈功能,应该能够让学习者通过考试切实体会到发展和进步、薄弱与不足,学生和教师可以根据反馈结果,有针对性地制定下一步学习计划。二是优化试题结构,强化汉语综合运用能力。HSK 2.0采用的是分立式语言测验,将语言能力按技能分为三个分测验,分别考察听力、阅读、书写能力,试题以大量的选择题为主,这都体现出了结构主义语言学的痕迹,又进一步促使教师在授课时大量练习选择题而较少使用开放性问题,这或许可以解释虽然考生 HSK 取得了不错的成绩,但在真实交际中仍然存在障碍的情况。HSK 应进一步聚焦对汉语综合运用能力的考察,以综合式语言测试代替分立式语言测试,在同一语言任务中考察听、说、读、写、译等能力,综合考察和训练学习者的汉语理解和产出能力。

目前 HSK 3.0正蓄势待发,HSK 3.0的研发应该是一次汉语考试系统升舱的过程,这不仅仅是扁平化地延伸考试级别、增加考试难度,更需要进行系统地设计与改造,保障 HSK 2.0和 HSK 3.0的有效衔接,增强科学性、规范性和有效性。虽然这些年 HSK 的发展取得了一定进展,发展出了纸笔、网考和居家考试等多种形式,但应对数字时代,特别是经历了疫情,为了更好地适应时代的要求,主办方还应在 HSK 等传统旗舰项目的基础上,推出更加灵活、多样的评价工具,如自适应考试、模块化考试、游戏化考试、借助元宇宙概念的真实沉浸式考试等。这些新型考试可以实现对学习者的语言表现的快速诊断,并能够同时满足视、听、说多模态试题呈现方式,在语言诊断针对性、评价场景真实性、趣味性、便捷考生、测试安全等方面具备独特优势,也可以激发学习者内动力,促进考试与汉语学习相融合,这应该是 HSK 3.0的重点任务之一。

## (二)聚焦标准引领,优化考试资源体系,推动教学评相融合

有什么样的指挥棒,就有什么样的办学导向,有什么样的标准,就有什么样的评价工具。在本研究调查范围内,绝大多数教师都认为 HSK 对学生的汉语学习和教师的教学很重要并产生了积极影响。尽管如此,由于学生的 HSK 成绩不与教师的教学效果评价挂钩,完成教学计划规定的授课内容是教师的首要任务,所以多数教师不会因为学生要参加 HSK 而对教学内容、进度、顺序和方法等进行大的改动。调查显示约有九成以上的教师都认为其所使用的教材与 HSK 的契合程度并不高,这种情况与 HSK 2.0的“考试原则”中所述“考试设计与目前国际汉语教学现状、使用教材紧密结合”的初衷存在一定偏差。这一方面说明 HSK 对教师教学的反拨作用有限,教评分离的情况普遍存在;另一方面则反映出 HSK 与现行教材存在一定程度的脱节。良好的教、学、评互动关系是评价作为桥梁,连接和调节教与学,教和学又可以进一步优化评价,通过教、学、评的不断良性互动,学习者在三驾马车的驱动下朝着学习目标不断前进。而目前的情况显然量不同衡、车不同轨,教、学、评无法达到有效对话。基于统一的标准开展基于标准的教育活动是解决这一问题的重要途径,这就包括基于标准研发教材、学习材料和评价工具等。权威统一的标准能够更有效地衔接教、学、评,增强教学评的连贯性和一致性,因此,《等级标准》应该成为教学与考试资源研发的重要依据。

[责任编辑:唐 普]