
HSK 试卷架构对 1—3 级考生成绩的影响*

——以英语母语者为例

郑 英 英国南安普顿大学现代语言和语言学系

郑 玥 汉考国际教育科技(北京)有限公司

张家维 中国澳门科技大学国际学院

提 要 本文对 HSK 的考试结构和考生成绩进行了探索性研究。利用考生成绩分析与考生访谈相结合的研究方法,对 HSK1—3 级的试题质量、考生遇到的挑战及挑战背后的原因进行了一系列探究。本文首先介绍了研究背景及研究问题;其次,阐述了项目的研究方法,并对考生的成绩数据进行了分析,辅以访谈结果;最后,根据研究结果,对 HSK 考试题型的优化及国际中文教学方法的改进提出建议。

关键词 HSK; 难度; 信度; 区分度; 反拨作用

Abstract This article describes an exploratory study of the relationships between the HSK (Level 1-3) test design, test questions, and the performance of UK test takers in 2019. Using quantitative analysis of the 1,833 test taker scores and qualitative interviews with seven test takers, a series of mixed method inquiries were made on the quality of HSK test quality indices, including test difficulty, reliability, and test item discrimination. In addition, the challenges encountered by this group of test takers and the reasons behind these challenges were investigated. Based on the results, suggestions were made to improve the test design to promote more effective Chinese language teaching.

Key words HSK; difficulty; reliability; differentiation; backwash

* 本研究为汉考国际科研基金重点项目“HSK 考生特点、考试成绩与 HSK 考试效度/难度的关联研究”(项目编号:CTI2018A03)成果之一。

一、引言

目前,许多国家掀起了学习中文的热潮。一些国家和地区在中学和大学开设了中文课程,中文学习已形成一定规模(Fitria等,2019)。例如,马来西亚、澳大利亚等国的中文课堂如雨后春笋般涌现(Bardovi-harlig、Su,2018)。截至2016年,“一带一路”沿线国家来华留学生人数为207746人,与2015年相比,同期增长13.6%(安亚伦、段世飞,2018)。相应地,测查对象是母语非汉语人士的,考核考生在日常生活、工作和学习中的汉语水平的汉语水平考试(HSK)(以下简称“HSK”)发展迅速(赵琪凤,2016)。HSK于1990年问世,为满足市场需求,2009年HSK进行了提质升级。总体而言,HSK已形成一定的规模,具有广阔的市场前景(唐雪婷,2016)。本文探讨的对象为2009年改版后的HSK。

在关于HSK的研究中,部分研究关注试卷质量及HSK与国际中文教学的关系。以HSK1、3、5级试卷为例,研究者发现其区分度、信度良好,试卷难度较低(罗民等,2011)。有的学者从HSK阅读试题对国际中文教育的反拨作用入手,发现HSK5级考试阅读题目比重上升,且注重材料的真实性,重视对考生语言交际能力的考查。因此,在中文教学中,教师应注重培养学生在真实语境下的交际能力,而不是将提升成绩作为最终目的,从而发挥测试的积极反拨作用(格萨茹拉,2014)。

上述研究聚焦HSK试题质量及HSK与国际中文教学的关系,而本文以HSK1—3级考试为落脚点,对HSK考生成绩、试题结构进行定量分析;对以英语为母语的HSK考生进行访谈,作为定量分析结果的补充。希望分析结果能够促进HSK试题的优化以及国际中文教学方法的改进。

二、研究问题及研究方法

(一) 研究问题

根据上文阐述的HSK现状,本文希望通过考生成绩了解试卷架构,通过考生访谈了解他们对HSK的观点、态度,从而为HSK未来的改进提供参考。故提出以下两个研究问题。

(1) HSK1—3级试题的信度、难度、区分度等指标质量如何?

(2) HSK1—3级考生遇到的挑战以及导致上述挑战的原因是什么?

(二) 研究对象和方法

本文主要采用混合研究方法。在定量分析中,以英语背景的中文学习者与分析对象,以2019年参加HSK1—3级考试的1833名英国考生的数据,包括性别、年龄及其在听力、阅读、写作各部分的分数和总分为样本,对考生成绩进行描述性分析,对试卷的难度、信度、区分度进行评估,对性别这一因素进行T检验,并对年龄这一因素进行T检验以及方差分析。

以18岁为分组界线,聚焦未成年考生与成人考生这两大群体。由于18岁以下与18岁及以上年龄组的样本年龄构成差距较小,因此,HSK1、2级的数据分为两个年龄段:未满18岁、18岁及以上,采用T检验进行量化分析。鉴于HSK3级样本年龄构成趋向多样化,且18岁以上的样本比例较高,为平衡小组样本数量,HSK3级数据分为三个年龄段:18岁及以下、19—30岁、30岁及以上,采用ANOVA分析手段对三组数据进行分析。

在定性分析中,采取半结构化访谈形式,探讨HSK1—3级考生在考试中遇到的挑战及其背后的原因。上述参与访谈的考生与定量分析中的考生语言背景相同,都以英语为母语。通过半结构化的

访谈,了解初学者在考试中遇到的挑战及挑战背后的原因。受访者共计 7 名考生,情况如下:

	澳大利亚	加拿大	美国	英国
1 级	1	-	1	-
2 级	-	-	1	1
3 级	-	1	1	1

三、研究结果分析

(一) 定量分析

1. HSK1—3 级考生成绩分析

首先对 HSK1—3 级考生的成绩进行描述性分析。HSK1、2、3 级样本容量分别为 584、620 和 629,研究包含对考生成绩的最小值、最大值、平均值、标准差的统计。鉴于 HSK1 级、2 级不涉及书写题,3 级才开始增设,因此本文将重点探讨听力和阅读两个模块。HSK1—3 级的听力题型分为四部分,分别是判断对错、图片匹配、短对话单项选择、长对话单项选择。HSK1—2 级的阅读题型分为四部分,包括判断对错、句子匹配、选词填空、图片匹配。3 级分为三部分,包括句子匹配、选词填空、单项选择。HSK1—3 级考试结构如表 1 所示。

表 1 HSK1—3 级考试结构

	听力		阅读		书写		合计总分
	题目数量	满分	题目数量	满分	题目数量	满分	
HSK1	20	100	20	100	-	-	200
HSK2	35	100	25	100	-	-	200
HSK3	40	100	30	100	10	100	300

就考生平均分而言,HSK1—3 级的阅读平均分普遍低于听力,但其标准差高于听力(见表 2),表明阅读这一题型的离散程度高于听力。HSK1、3 级阅读的标准差尤为突出,分别为 21.17 和 20.18,一定程度上说明这两级考生在阅读方面内部差异较

大,高分段和低分段考生人数较多,存在两极分化的趋势。在日后的教学中,教师可能需要对语言水平相对较低的学生给予关注,注意低分段学生的学习情况;也可考虑同伴互助,让高水平学生适当辅助水平较低的学生,提升低分段学生的语言水平。

表 2 HSK1—3 级考生成绩的分析

HSK 等级	数量	试卷结构	最低分	最高分	平均分	标准差
HSK1	584	听力	13	100	87.21	16.13
	584	阅读	14	100	83.31	21.17
	584	总分	27	200	170.54	34.85
HSK2	620	听力	49	100	90.96	9.73
	620	阅读	29	100	84.94	16.49
	620	总分	81	200	175.90	24.03
HSK3	629	听力	9	100	87.37	12.19
	629	阅读	0	100	81.64	20.18
	629	总分	9	200	169.00	30.16

2. HSK1—3 级试题信度及区分度分析

信度和区分度是评估考试质量的重要指标(Bavafa, 2010)。信度可评估测试结果是否具有可靠性、考试能否真实反映考生水平;区分度是鉴别试题能否区分出考生水平的关键指标。在 SPSS 中,使用 Cronbach α 系数反映信度,系数达到 0.5,表明试卷信度质量合格(施俊、杨勇,2017)。本研究中 HSK1—3 级试题的 Cronbach α 系数分别为 0.84、0.72、0.76,表明考试有较高的内部一致性,试卷具有较高的信度,结果如表 3 所示。

表 3 HSK1—3 级试题信度分析

HSK 等级	Cronbach α
HSK 1	0.84
HSK 2	0.72
HSK 3	0.76

同时,本文对 HSK1—3 级试题的听力、阅读、书写、试卷整体及每一题型的区分度进行了分析(见表 4)。区分度 D 值介于 -1 到 +1 之间,D 值与区分度呈正相关关系,D 值越趋近于 +1,试题区

分度越高。一般来说, D 值 > 0.3 的题目区分度较好, D 值 < 0.2 的题目区分度较差, D 值 < 0 的试题应被剔除(施俊、杨勇, 2017)。根据分析结果, 整体而言, HSK1—3 级的听力、阅读部分以及试

卷整体的区分度数值均大于 0.3, 试卷整体区分度数值分别为 1.00、0.86、0.34, 表明试题具有良好的鉴别力, 能够在一定程度上有效区分考生的不同水平。

表 4 HSK1—3 级考生成绩区分度分析 (D 值)

	听力					阅读						写作	试卷整体
	判断 对错	图片 匹配	短对话 单项 选择	长对话 单项 选择	听力部 分整体	判断 对错	句子 匹配	选词 填空	图片 匹配	单项 选择	阅读部 分整体	书写部 分整体	
HSK1	0.70	0.78	0.77	0.64	0.92	0.72	0.80	0.84	0.78	-	0.96	-	1.00
HSK2	0.67	0.76	0.61	0.59	0.71	0.80	0.74	0.77	0.58	-	0.83	-	0.86
HSK3	0.63	0.73	0.79	0.77	0.27	-	0.75	0.78	-	0.71	0.34	-	0.34

3. HSK1—3 级试题难度分析

难度是评判试卷质量的关键指标。难度用 P 值表示, P 值的区间为 0—1。试题或试卷的难度与 P 值成负相关关系, 即 P 值越大, 试题或试卷的难度越低(敖希琴等, 2016)。一般而言, P 值 < 0.3, 试题难度过高; P 值在 0.3—0.6 之间, 试题难度较

高; P 值在 0.6—0.8 之间, 试题难度适中; P 值 > 0.8, 试题难度偏低(施俊、杨勇, 2017)。总体而言, HSK1—3 级的 P 值分别为 0.84、0.87 和 0.84, 试卷难度略低。此外, 通过表 5 数据不难发现, 在 HSK1—3 级中, 阅读的难度整体略高于听力。

表 5 HSK1—3 级试题难度分析 (P 值)

	听力					阅读						写作	试卷整体
	判断 对错	图片匹 配	短对话 单项 选择	长对话 单项 选择	听力部 分整体	判断 对错	句子 匹配	选词 填空	图片 匹配	单项 选择	阅读部 分整体	写作部 分整体	
HSK1	0.84	0.83	0.88	0.93	0.86	0.88	0.71	0.75	0.89	-	0.82	-	0.84
HSK2	0.89	0.89	0.90	0.89	0.91	0.77	0.76	0.76	0.93	-	0.84	-	0.87
HSK3	0.77	0.87	0.81	0.81	0.87	-	0.79	0.68	-	0.82	0.81	-	0.84

4. 题型、年龄和性别对成绩的影响分析

就题型对成绩的影响结果而言, 阅读部分的“选词填空”和“句子匹配”相较其他题型略难(见

表 5), 其平均分也低于其他题型(见表 6)。但值得注意的是, 它们的标准差高于其他题型, 说明考生的高、低分比较分散。考生在此类题型中可能会

运用考试技巧,如猜答案、排除已选选项等,获得高分;或因对某一题答案存疑,导致多题犯错,存在“一对全对、一错错俩”的情况。相较而言,“图片匹配”和“判断对错”在 HSK1—2 级中平均分

较高、标准差较低,结合表 5,表明这两种题型难度较低,适合初级考生。在 HSK3 级中,没有出现这两类题,而选词填空的比重提高,符合先易后难的考试设计原则,能够满足考试区分度的要求。

表 6 HSK1—3 级考生阅读部分题型的分析

	HSK1		HSK2		HSK3	
	平均分	标准差	平均分	标准差	平均分	标准差
判断对错	0.88	0.18	0.77	0.31	-	-
句子匹配	0.71	0.36	0.76	0.31	0.79	0.27
选词填空	0.75	0.34	0.76	0.23	0.68	0.32
图片匹配	0.89	0.24	0.93	0.20	-	-
单项选择	-	-	-	-	0.82	0.22

在听力部分(见表 7), HSK2 级整体来看,平均分较高标准差较低,相较于 HSK1、3 级而言,不易区分考生水平。从题型来看,“图片匹配”题在 HSK1、2、3 级中标准差较高,难度总体呈下降趋势(表 7),对语言能力越高的考生来说题型越发简单。此类题型以图片为辅助手段,意在考查考生通过图片辅助理解听力对话。当语言能力提高时,图片辅助作用变小。“长对话单项选择”在 HSK1—3 级中平均分体现下降趋势,难度上升(见表 7)。其标准差维持在 0.18—0.21 之间,表明这一题型能够较稳定地区分考生语言水平,符合语言学习过程的自然规律,较能反映考生的真实听力水平,能够体现考卷在不同层级难度递增的设计理念。

表 7 HSK1—3 级考生听力部分题型的分析

	HSK1		HSK2		HSK3	
	平均分	标准差	平均分	标准差	平均分	标准差
判断对错	0.84	0.21	0.89	0.13	0.77	0.19
图片匹配	0.83	0.29	0.89	0.19	0.87	0.41
短对话 单项选择	0.88	0.21	0.90	0.13	0.81	0.22
长对话 单项选择	0.93	0.18	0.89	0.20	0.81	0.21

此外,通过独立 T 检验(见表 8)可以看出,性别在考试成绩中没有显著性差异(sig.>0.05),这意味着不同性别的考生在 HSK1—3 级考试中的表现具有一致性,体现了试题在性别中的公平性。就年龄这一因素而言, HSK1、2 级分为两个年龄组,因 T 检验适用于两个样本组之间的比较,故采用 T 检验进行分析。由于 HSK3 级的年龄组为三组,因此采用了适合在三组之间两两比较的单因素方差分析(邹祎, 2016),因 Levene’s 方差不齐(P<0.05),故采用 Welch 方差分析和 Games-Howell 事后检验。

表 8 HSK1—3 级考生成绩的性别差异分析 (T 检验)

HSK 等级	试卷结构	组别 (平均数 ± 标准差)	组别 (平均数 ± 标准差)	t	sig.
		男性 (N=269)	女性 (N=315)	t	sig.
HSK1	听力	88.04±15.50	86.53±16.65	1.31	0.26
	阅读	85.10±19.62	81.77±22.33	1.92	0.06
	整体试卷	173.14±32.49	168.30±36.64	1.68	0.09
		男性 (N=339)	女性 (N=281)	t	sig.
HSK2	听力	91.08±9.54	90.81±9.96	0.35	0.73
	阅读	85.08±16.79	84.78±16.14	0.23	0.82
	整体试卷	176.16±24.10	175.59±23.95	0.30	0.77
		男性 (N=349)	女性 (N=280)	t	sig.
HSK3	听力	87.39±12.30	87.34±12.06	0.06	0.95
	阅读	80.45±21.12	83.12±18.89	-1.67	0.10
	整体试卷	167.84±31.06	170.45±29.00	-1.08	0.28
*sig.<0.05**sig.<0.01					

除 HSK3 级的听力部分外, 如表 9 和表 10 所示, 年龄在 HSK1—3 级考试中具有显著性差异 (sig.<0.01), 成人考生的表现总体优于未成年考生。换言之, 在 HSK1—3 级考试中, 成人考生较易取得优异成绩, 未成年考生处于相对劣势的地位。潜在原因之一可能是成年人更具逻辑能力及分析能力, 相较于未成年考生更具解题优势。据 HSK3 级分析, 低年龄组到中年年龄组, 阅读平均得

分和整体试卷得分分别增加 10.02 (95% CI: 5.71—14.32) 和 12.21 (95% CI: 5.63—18.78), 差异均具有统计学意义 (P<0.001); 中年年龄组到高年龄组, 阅读平均得分减少 6.60 (95% CI: -12.03—-1.18), 差异具有统计学意义 (P=0.01)。教师可在日后教学中, 对不同年龄组学生的需求加以适当关注, 进行针对性教学。

表 9 HSK1—2 级考生成绩的年龄差异分析 (T 检验)

HSK 等级	试卷结构	组别 (Mean ± SD)	组别 (Mean ± SD)	t	Sig.
		≤ 18 (N=322)	>18 (N=262)	t	Sig.
HSK1	听力	84.85±17.19	91.39±13.13	-5.16	0.00**
	阅读	78.70±22.26	91.39±16.27	-7.90	0.00**
	整体试卷	163.55±36.58	182.78±27.68	-7.16	0.00**
		≤ 18 (N=339)	>18 (N=281)		
HSK2	听力	90.22±10.91	91.85±8.00	-2.14	0.03*
	阅读	81.66±17.75	88.90±13.85	-5.70	0.00**
	整体试卷	171.88±26.22	180.75±20.05	-4.77	0.00**

表 10 HSK3 级考生成绩的年龄差异分析 (ANOVA 单因素方差检验)

HSK 等级	试卷结构	组别 (Mean ± SD)			F	sig.
		≤ 18 (N=201)	19-30 (N=311)	>30 (N=117)		
HSK3	听力	86.08 ± 14.11	88.27 ± 10.72	87.17 ± 12.20	2.00	0.14
	阅读	76.05 ± 22.33	86.06 ± 16.40	79.46 ± 22.69	16.65	0.00**
	整体试卷	162.13 ± 34.02	174.34 ± 25.22	166.63 ± 32.68	10.77	0.00**
*sig.<0.05**sig.<0.01						

通过以上分析可以看出, 总体而言, HSK1—3 级具有较高的信度和区分度, 试卷难度较低。我们发现题型对考生成绩有一定的影响, 阅读部分, 选词填空、句子匹配是难点; 听力部分, 图片匹配可能不再适用于高水平考生, 长对话单项选择难度设计与考生水平相协调。在对年龄、性别差异进行研究后发现, 成人考生的表现优于未成年考生, 性别上无显著性差异。

(二) 定性分析: 考生访谈

虽然对试卷进行了定量分析, 但考生对于试卷难度的看法等在定量分析中难以体现, 故本研究还采用访谈的方式, 设计了如下问题: ①考生在作答 HSK 试卷时遇到的挑战是什么? ②个人认为导致上述挑战的原因是什么? 基于上述两个问题对受访者进行调查。就问题①而言, 在采访中, 两名 HSK1 级的考生认为, 在考试中听力是一大挑战, 其余五名 HSK2 级和 3 级的考生则认为阅读才是挑战。在访谈中, 一位参加过 HSK3 级的美国考生说道: “阅读对我来说非常难, 即使老师在课堂上对汉字进行了讲解, 学习之后, 阅读题仍然需要用很长的时间。”来自澳大利亚的参与 HSK1 级的考生则持有不同的观点: “在 HSK1 级考试中, 阅读容易, 听力很难。”在七名接受访谈的参与者中, 五名考生认为相较于听力, 阅读更具难度, 耗费更多的时间; 而两名 HSK1 级的考生则认为听力难于

阅读。

就访谈问题②而言, 在采访中, 两名 HSK1 级的考生认为之所以听力是一大挑战, 是因为汉语有四个声调, 相较于没有声调的母语而言, 难以辨别。来自澳大利亚的 HSK1 级考生说道: “汉语有四个声调, 区分二声和三声非常有难度, 字词的发音、意思很难记住。”其余五名参与者认为阅读是一大挑战, 如 HSK3 级的美国考生所说: “我认为汉字很难。汉字结构复杂, 我很难记住汉字的形状和它的意思, 记住汉字后, 也很难理解文章。听力比较简单, 虽然有时没能理解听力材料, 但可以根据听到的信息点与图片呈现的选项进行对应。”

根据访谈结果, 我们发现 HSK1 级考生认为听力是挑战, 而 2、3 级考生则认为阅读是挑战。听辨能力可能是导致 HSK1 级考生与 2、3 级考生观点不同的原因之一, 初级阶段考生听辨能力较弱, 目的语无法与母语对应时, 考生会受到一定的考验。总体而言, 定性探究与定量分析中的结论趋于一致, 阅读题目对考生而言更具挑战性。就导致挑战的原因而言, HSK2、3 级考生认为阅读较难的原因总体趋于一致, 即汉字较为复杂, 记忆汉字的形与义十分困难。两名 HSK1 级考生认为听力困难的原因也大致相同, 即汉语为声调语言, 辨音有难度, 语音与语义之间难以关联。

四、关于分析结果的讨论

(一) 汉字与考生母语文字的差异对阅读的影响

上文对考生成绩和访谈内容进行了定量和定性分析。通过以上分析,我们发现阅读是中文初学者的主要挑战。在访谈中,大多数受访者认为,相较于听力,阅读更难。在随后的采访中,受访者提及汉字结构复杂,并在访谈中反馈,认识汉字后,仍然难以理解文章内容。原因之一在于其母语文字与汉字相似度较低,对汉语学习产生一定的消极影响。持有上述观点的受访者均不属于汉字圈。他们的母语文字属于拼音文字,而汉字属于表意文字,汉字对他们来说很陌生。不同于拼音文字,汉字体系较为庞大(李大遂,2004),形体比较复杂,对考生的记忆而言,具有一定的挑战性(徐彩华等,2007)。本文的定量和定性分析,一定程度上反映出母语与目的语相似度较低会对第二语言的习得产生消极影响。

(二) 听辨能力与目的语特征对听力的影响

与访谈 HSK2、3 级考生的结果不同,两位 HSK1 级的考生认为听力是较大的挑战。造成这种情况的深层原因可从两个方面分析,一是目的语的特征,二是入门学习者的听辨能力。汉语是有声调的语言,而考生母语为非声调语言,且参加 HSK1 级时,考生的汉语听辨能力处于起步阶段,分辨声调并将语音与语义相关联比较困难。本文认为,中文学习者的听辨能力弱与汉语声调的特征共同作用是入门考生视听力为一大挑战的主要原因。

(三) HSK 题型对考生成绩的影响

本文发现 HSK 题型在一定程度上会影响考生的表现。在阅读部分中,句子匹配和选词填空属于一一对应的题型,同一选项只能选择一次,考生需

要花费一定精力理解文章内容。这在一定程度上降低了考生猜题的可能性。相反,在单项选择中考生根据阅读材料的碎片信息,利用猜测法或排除法,也有可能得出正确的答案。所以,在应对单项选择时,即使考生可能没有充分调动阅读能力,例如归纳、推理能力(王海军,2012),而是更多使用解题技巧作答,也取得了出色的成绩。

以 HSK3 级听力部分为例,1—10 题以图片形式呈现选项,与以文本形式呈现选项的方式相比,图片的形式相对直观,学生可以快速捕捉图片信息进行作答。因此,考生有时较难通过测试发现自身不足,从而对训练策略进行调整,考试在一定程度上也较难真正发挥积极的反拨作用。综上所述,题型在一定程度上影响考生的答题思路和解题策略,因此会对考生成绩产生一定影响,这与 Ihme (2017) 等人的研究具有一致性。

在定量分析中,研究发现 HSK1—3 级试卷难度较低,部分听力考题中采用应试技巧即可获得一定分数,考试无法发挥正面反拨作用。鉴于报考 HSK1 级的考生处于入门阶段,下文就提升 HSK2、3 级听力题的难度进行探讨。2021 年发布的国家级语言文字规范标准——《国际中文教育中文水平等级标准》(GF0025-2021)(以下简称“《等级标准》”)将中文学习者水平分为三等九级。初等三级关于“听”的描述是,要求学生能够理解较长单句和简短复句为主的对话或一般性讲话,并且“能够通过语音、语调、语速的变化等辅助手段理解和获取主要信息”。依据《等级标准》,本文认为,可酌情减少以图片呈现选项的题目数量,适当增加以文本为选项的听力试题,以降低误解、混淆的可能性,测查考生的真实听力能力,如信息获取能力,而非让考生通过视觉辅助手段猜测出正确答案。单纯提高试题难度并非测试的初衷,考查相应能力才是试题

调整的旨归,上述题型的调整可在一定程度上给予考生有效的测试反馈,提高考生的听力能力,发挥语言测试的积极反拨作用。

五、结论与启示

在对试卷的定量分析中,我们发现,HSK1—3级试题信度与区分度表现良好,阅读略难于听力,试卷难度较低。就题型对考试成绩的影响而言,阅读部分,选词填空、句子匹配对考生而言是较难应对的题型;听力部分,图片匹配或不再适用于高水平考生,长对话单项选择难度设计与考生水平相协调。整体而言,考生成绩在年龄层面具有显著性差异,成人考生更具优势;在性别中无显著性差异。在定性研究中,访谈结果显示,阅读是入门级别学习者的一大挑战,汉字的复杂性、字形与语义的关联问题是造成挑战的潜在原因。

在国际中文教学中,教师需多加关注学生的考后反馈,改进教学方法,使教与考形成良性互动,发挥语言测试的积极反拨作用。HSK1—3级试题阅读中的选词填空、句子匹配是难点题,此类题型均需要考生拥有综合理解与运用字、词、句的能力,若只掌握零散汉字则无法有效理解文章。故教师在阅读课中应重视汉字的实际作用,通过汉字增强学生对词的记忆与掌握,培养学生的语块意

识。语块作为语言的“小团体”,可在运用过程中整体提取(Wray, 2002; 张博, 2020)。教师应有意识地总结语块,结合《等级标准》中的语言量化指标,与课文中的词汇、语法相呼应,作为课堂教学的补充材料,帮助学生逐步在头脑中构建“语块框架”,从而正确、有效地切分句子,准确理解句意。

本文对HSK1—3级考试的试题及考生遇到的挑战及其背后的原因进行了探究,并从试题设计及教学两方面对上述问题提出解决方案,以期为HSK试题的优化以及国际中文教学方法的改进建言献策。本研究存在一定的局限性,在定性研究的访谈环节中,样本数量较少,访谈结果的覆盖范围较为狭窄,不足之处将在日后的研究中得以完善。

作者简介

郑英,英国南安普顿大学现代语言和语言学系副教授,博士生导师,主要研究方向为语言测试中的量化验证及语言教师的职业发展。

郑玥,汉考国际教育科技(北京)有限公司研发专员,主要研究方向为第二语言教学、语言测试。

张家维,中国澳门科技大学国际学院助理教授,博士,主要研究方向为语言测试与评价、学习动机、教师职业发展与培训。

参考文献

- 安亚伦,段世飞,2018.“一带一路”倡议下的汉语国际教育:现状、问题及对策[J].湖南师范大学教育科学学报(6):45-52.
- 敖希琴,等,2016.基于SPSS的试卷质量分析与建模[J].安庆师范学院学报(自然科学版)(4):37-40.
- 格萨茹拉,2014.探析新HSK五级阅读测试对汉语阅读教学的反拨作用[J].内蒙古师范大学学报(教育科学版)(9):101-102,111.
- 李大遂,2004.对外汉字教学发展与研究概述[J].暨南大学华文学院学报(2):41-48.
- 罗民,等,2011.新汉语水平考试(HSK)质量报告[J].中国考试(10):3-7.
- 施俊,杨勇,2017.基于统计软件SPSS的试卷质量分析[J].电脑知识与技术(17):121-123,127.
- 唐雪婷,2016.新HSK研究述评[J].考试研究(4):90-96.

- 王海军, 2012. 多项选择、填空与问答式阅读题型的效标关联效度对比研究 [J]. 西安外国语大学学报 (4): 80-83.
- 徐彩华, 等, 2007. 留学生汉字形误识别能力发展的实验研究 [J]. 语言教学与研究 (4): 63-70.
- 张博, 2020. “语素法”“语块法”的要义及应用 [J]. 语言教学与研究 (4): 12-24.
- 赵琪凤, 2016. 汉语水平考试的历史回顾及研究述评 [J]. 中国考试 (9): 47-53.
- 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[S]. 北京: 北京语言大学出版社, 2021.
- 邹伟, 2016. SPSS 软件单因素方差分析的应用 [J]. 价值工程 (34): 219-222.
- BARDOVI-HARLIG K, SU Y, 2018. The acquisition of conventional expressions as a pragmalinguistic resource in Chinese as a foreign language[J]. The modern language journal,102(4): 732-757.
- BAVAFA, 2010. The intentional targeting test: a necessary alternative to the disparate treatment and disparate impact analysis in property rentals discrimination[J]. Columbia journal of law and social problems, 43(4): 491-515.
- FITRIA A, et al., 2019. A needs analysis of Chinese language teaching materials in vocational high schools offering business and management majors[J]. Polyglot jurnal ilmiah 15: 1-16.
- IHME J M, et al., 2017. Assessment of computer and information literacy in ICILS 2013: do different item types measure the same construct?[J]. European educational research journal,16(6): 716-732.
- WRAY A, 2002. Formulaic language and the lexicon[M]. Cambridge: Cambridge University Press.

(责任编辑 刘智贤 赵青)

~~~~~

(上接 49 页)

- CHRISTENSEN C R, 1991. The discussion teacher in action[G]//CHRISTENSEN C R, et al. Education for judgment: the artistry of discussion leadership[M]. Boston, MA: Harvard Business School Press: 153-172.
- COTTON K, 1988. Classroom questioning[EB/OL].[2021-08-30]. <https://educationnorthwest.org/sites/default/files/ClassroomQuestioning.pdf>.
- CROOKES G, Chaudron C, 2001. Guidelines for language classroom instruction[C]//CELCE-MURCIA M. Teaching English as a second or foreign language. 3rd edition. Boston, MA: Heinle & Heinle: 29-42.
- GALL M D, 1970. The use of questions in teaching[J]. Review of educational research, 40(5): 707-721.
- GILL T G, 2011. Informing with the case method: a guide to case method research, writing and facilitation[M]. Santa Rosa, CA: Informing Science Press.
- LONG M, SATO C, 1983. Classroom foreigner talk discourse: forms and functions of teachers' questions[G]//SELIGER H, LONG M. Classroom-oriented research in second language acquisition. Rowley, MA: Newbury House: 268-286.
- MCKEACHIE W J, 2006. Teaching tips: strategies, research, and theory for college and university teachers[M]. 12th edition. Boston, MA: Houghton Mifflin Company.
- SANDERS N M, 1966. Classroom questions: what kinds?[M]. New York: Harper & Row.
- SOLOMON D, et al., 1964. Teacher behavior and student learning[J]. Journal of educational psychology, 55(1): 23-30.
- STROTHER D B, 1989. Developing thinking skills through questioning[J]. The phi delta kappan, 71(4): 324-327.
- WILEN W W, 1991. Questioning skills for teachers[M]. Washington, DC: National Education Association.

(责任编辑 刘智贤 赵青)