

汉语水平考试居家考试与纸笔考试和现场机考成绩比较研究

——以印尼泗水考点为例

Budi Kurniawan

Prodi Bahasa Mandarin, Universitas Kristen Petra

邮箱地址: budi.kurniawan@petra.ac.id

摘要

新冠肺炎疫情期间，汉考国际实行 HSK 居家网络考试。2020 年 5 月，印尼泗水 HSK 考点也开始举办居家考试。目前有关居家考试的研究还非常少，而考虑到居家考试的信度，需要研究居家考试的成绩，尤其是与纸笔考试成绩的比较。因为纸笔考试和现场机考得分也出现差异现象，基于三者是目前采用的考试形式，本研究对泗水考点居家考试和纸笔考试以及现场机考成绩进行比较。统计结果表明，HSK（一级）至（三级）纸笔考试和现场机考成绩基本高于居家考试。相反，HSK（四级）至（六级）居家考试成绩基本高于纸笔考试和现场机考。居家考试信度受几个因素的影响，即考试地点、考生群体、监控效力以及考试环境。

关键词：HSK；HSK 居家考试；HSK 成绩；泗水考点

Abstract

During the COVID-19 pandemic, Chinese Testing International implemented HSK Home Edition. In May 2020, the HSK test center in Surabaya, Indonesia also began to hold home edition tests. At present, there are very few studies on HSK Home Edition, and considering the reliability of HSK Home Edition, it is necessary to study the scores of HSK Home Edition, especially the comparison with paper-based test scores. Because the scores of the paper-based test and the on-site computer-based test are also different, and currently these three forms of HSK are all used, this study compares the scores of HSK Home Edition with paper-based and on-site computer-based HSK at the Surabaya test center. Statistics show that the scores of HSK (Level 1) to (Level 3) paper-based tests and on-site computer-based tests are basically higher than that of Home Edition. On the contrary, the scores of the HSK (Level 4) to (Level 6) Home Edition are basically higher than those of the paper-based test and the on-site computer-based test. HSK Home Edition reliability is affected by several factors, namely testing location, participant groups, proctoring effectiveness, and testing environment.

Keywords: HSK; HSK Home Edition; HSK score; Surabaya test center

一 前言

汉语水平考试 (HSK) 是考查汉语非第一语言考生汉语交际能力的国际标准化考试^[1]。作为汉语水平评估的国际标准, 参加 HSK 的考生人数越来越多, HSK 全球的覆盖率也越来越广。汉语考试服务网指出, 目前全球有 155 个国家 (地区) 设立 1208 个考点, 累计服务全球各类中文学习者 3000 多万人次^[2]。印尼 1998 年改革开放之后, 随着华文教育的复苏, 自 2001 年首次举办纸笔考汉考工作, 每年考生人数不断增加。近几年, 每年 3 月、8 月和 10 月都会举办大型汉语考试纸笔考工作, 每年考生近 2 万人^[3]。

印尼泗水 HSK 考点 (全名为“印尼国民教育部-中国汉语考试 (泗水)”, 以下简称泗水考点) 是印尼全国汉考委员会属下的分考点。泗水考点也从 2001 年开始, 每年都举办考试。除了传统纸笔考试形式以外, 从 2016 年 4 月起, 泗水考点也提供网络考试形式 (现场机考)。2016 年到 2020 年 1 月参加纸笔考试的 HSK 和 HSKK (口语考试) 总共 5908 人次, 参加机考总共 675 人次。

2020 年受新冠肺炎疫情的影响, 汉考国际开始实行 HSK 居家网络考试。汉考国际报道称, 截至 2020 年夏季, 54 个国家的 150 个考点已实施 HSK 居家考试^[4]。2020 年 5 月, 泗水考点也开始举办居家考试。

参加居家考试的考生要确保考试设备和环境符合要求。考生需准备一台电脑和一部手机, 电脑用来作答试题, 手机用来辅助监考。考试电脑需安装 Windows(7/10) 操作系统, 不支持在 Mac OS 操作系统下运行。屏幕分辨率须为 1366*768 或以上, 自带或外接独立摄像头, 像素 30 万以上, 房间光线充足, 摄像清晰。带摄像头的手机作为辅助监考设备, 下载安装所报名考点指定的监考软件 (目前泗水考点采用 Zoom)。

目前关于 HSK 居家考试的研究还很少, 可以说几乎没有。Chen (2022) 调查了欧洲举办居家考试的一些考点, 初步探讨居家考试考生的反应以及操作方面的一些困难^[4]。Chen (2022) 指出, 考虑到居家考试的信度, 需要进一步研究居家考试的成绩, 尤其是与纸笔考试成绩的比较。张晋军和符华均研究结果表明, HSK 网络考试 (现场机考) 平均成绩大部分高于纸笔考试平均成绩。因为纸笔考试和现场机考成绩也出现差异现象, 基于三者是目前采用的考试形式, 本研究将对居家考试和纸笔考试以及现场机考成绩进

行比较。通过于泗水考点三种考试形式的比较，希望能为居家考试的信度提供概述。

二 理论基础

目前的 HSK 是从 2009 年 11 月开始推出。HSK 共有六个级别，HSK（一级）、HSK（二级）为初级，HSK（三级）、HSK（四级）为中级，HSK（五级）、HSK（六级）为高级。除了 HSK（一级）和 HSK（二级）只评估听力和阅读，其他级别都评估考生的听力、阅读和书写能力。另外，用来评估口语能力，设有独立的 HSKK（汉语水平口语考试）。

每项语言技能（听、读、写）满分为 100。由于 HSK（一级）和 HSK（二级）仅评估听力和阅读能力，因此这两个级别的总分为 200。HSK（三级）到 HSK（六级）这四个级别的总分是 300。一级至二级总分 120 分为合格。三级至六级总分 180 分为合格，不过自从 2013 年 2 月起，五级和六级取消了合格线^[5]。

HSK 是一项结构完善、可信度高的考试^[6]。不过在纸笔考试和网络考试（机考）两种形式之间的成绩存在差异，机考成绩往往比纸笔考试成绩好^[7]。因此，有必要更好地确保机考和纸笔考的分数一致性^[6]。

印尼自从 2020 年新冠肺炎疫情爆发以来，与全球许多国家一样，教学活动搬到网上去。虽然网上课堂不是全新的东西，但已经习惯采用实体形式授课的教师和学生，还是要在很短的时间内去适应这种教学形式。在上网课的过程中，不仅在授课方面收到影响，更难的是如何进行考试^[8,9]。为了应对疫情影响，汉考国际推出 HSK 居家考试，让个人在家中在线参加考试。印尼许多考点从 2020 年 5 月开始组织居家考试，泗水考点也包括在内。

虽然 HSK 未来应将现代技术融入其中，但疫情前的 HSK 主要形式仍然是纸笔考试^[10]。居家考试的研发，可以说是一种突破，在进行居家网络考试的种种困难之下，居家考试的现代技术功能（人脸识别功能、雷同卷判定功能等）确实是现代技术优势的体现。然而，因为进行居家考试涉及到很多因素，不仅是考试系统本身，还有网络速度、考生本人家里的环境、远程监控的操作等，HSK 居家考试在实行过程中不免出现一些苦难与挑战^[4]。这些困难与挑战其实也存在与其他语言疫情期间的居家网络考试当中^[11-14]。

居家考试的答题方式与现场机考一样，因此也与机考一样在成绩方面需要与纸笔考试进行对比，确保可信度。目前张晋军和符华均对 HSK 四级至六级纸笔考试和机考的

对比研究结果表明^[7]，听力部分，现场机考得分高于纸笔考试。他们推测这可能是因为机考考生通过耳机听题，声音更清晰，让考生注意力更加集中。考试的阅读部分，成绩没什么差别。书写部分，HSK（四级）和 HSK（五级）机考成绩显著高于纸笔考试。他们推测的原因是键盘输入的因素，而至于 HSK（六级）差异不大是因为六级的书写题水平已经高到不受考试形式的影响。Chen（2022）认为，居家考试所面临的挑战有三个维度，即：信度、安全和隐私、公平性^[4]。这三个维度相互关联，公平性影响到安全和隐私，安全和隐私影响到信度。因此，将要探讨信度的话，也会涉及到居家考试的安全和隐私，以及公平性。换句话说，虽然 HSK 本身具有很高的可信度，但居家考试形式所涉及到的其他外在因素，可能影响到 HSK 居家考试信度。

三 研究方法

泗水考点从 2016 年 4 月开始举办现场机考，与纸笔考试轮流进行，直到 2020 年 1 月。疫情期间，也就是从 2020 年 5 月到 2022 年 6 月，只举行居家考试。由于参加现场机考口试的考生很少，除了现场机考口试以外，本研究采用 2016 年至 2022 年 6 月这期间内所有举办过的三种考试形式的得分为样本。具体信息见表 3.1。

表 3.1 泗水考点 2016 年至 2022 年 6 月得分的人次统计

	纸笔考试（人次）	现场机考（人次）	居家考试（人次）
HSK（一级）	818	19	141
HSK（二级）	1663	99	485
HSK（三级）	1038	122	485
HSK（四级）	1362	197	736
HSK（五级）	464	172	518
HSK（六级）	83	24	64
HSKK（初级）	197		79
HSKK（中级）	145		357
HSKK（高级）	13		171
共计	5783	633	3036

虽然 HSKK 网络考试与纸笔考试答题形式基本相同，但现场机考和居家考试的 HSKK 都不允许用纸笔手写提纲，而要通过系统打字输入提纲，因此本研究认为还有必要进行对比。另外，从 2021 年 9 月开始，印尼参加居家考试 HSK（三级）至（六级）须同时参加 HSKK。

笔者将对纸笔考试、现场机考和居家考试听力、阅读和书写，以及口试得分进行统计比较。由于参加纸笔考试，现场机考和居家考试三组考生之间不同，为了能知道差别的显著性，将采用方差不齐非配对 t 检验。

除了得分以外，笔者通过观察居家考试的筹备过程、主监考考前培训以及考试当天的情况所获得的数据，结合与统计学对比结果进行分析。

四 结果与讨论

统计了三种考试形式的得分之后，居家考试与纸笔考试 HSK 平均分对比结果见表 4.1，居家考试与现场机考 HSK 平均分对比结果见表 4.2。

表 4.1 居家考试与纸笔考试 HSK 得分对比

考试级别	听力		阅读		书写	
	居家	纸笔	居家	纸笔	居家	纸笔
HSK (一级)	88.90**	85.23	87.48**	82.38		
HSK (二级)	80.17	86.94**	68.03	77.87**		
HSK (三级)	83.99	84.96	71.76	73.62	66.61	73.64**
HSK (四级)	79.30**	75.43	77.62**	66.73	68.58*	66.53
HSK (五级)	74.15	72.73	59.87**	55.75	64.98**	60.73
HSK (六级)	73.09	70.48	67.78*	61.30	69.89*	64.66

*表示显著性达到 0.05 水平，**表示显著性达到 0.01 水平，下同。

表 4.2 居家考试与现场机考 HSK 得分对比

考试级别	听力		阅读		书写	
	居家	机考	居家	机考	居家	机考
HSK (一级)	88.90	95.79**	87.48	93.26		
HSK (二级)	80.17	84.94*	68.03	77.90**		
HSK (三级)	83.99	84.76	71.76	75.84	66.61	77.08**
HSK (四级)	79.30**	71.12	77.62**	65.46	68.58**	62.61
HSK (五级)	74.15**	68.03	59.87*	56.12	64.98*	62.52
HSK (六级)	73.09	72.38	67.78	64.38	69.89	70.29

表 4.3 居家考试与纸笔考试 HSKK 得分对比

考试级别	居家	纸笔
HSKK (初级)	72.85	72.82
HSKK (中级)	57.38	63.90**

HSKK (高级)

50.98

61.31*

结果显示,除了 HSK (二级)和 HSK (三级)以外,居家考试 HSK 得分都高于纸笔考试 HSK 得分。而与现场机考对比,除了 HSK (一级)、HSK (二级)、HSK (三级)以外,居家考试 HSK 得分都高于现场机考(六级的书写部分虽然机考得分高于居家考试,但显著性不明显)。这结果表示,与纸笔考试和现场机考对比,HSK 居家考试一级至三级得分一般更低,而四级至六级得分更高。笔者推测,得分的差别不是与纸笔和电脑两种不同形式有关,而更是与场地有关。一级至三级考生一起在考场答题的表现一般更佳,而四级至六级考生在家里考试的表现更好。居家考试的地点确实是影响公平性维度的一个因素。就是因为考生分布于不同地点参加考试,而这关系到网络和信号的质量,所以有可能影响考生答题。这么说来,参加四级至六级的考生家里网络更稳定,或者因为级别高了,他们就更注意网络,事先做好准备。

除了考试地点,有可能影响得分的差别是考生群体不同。参加一级至三级与参加四级至六级的考生群体不同,参加居家考试和参加现场考试的考生群体也不同。谈到群体,所涉及的因素就颇为复杂。年龄、学习时间、家庭、生活环境、学习环境、教材等各种因素都组成每位考生的特性。居家考试因为考生在家里参加考试,考生就有可能来自各种城市和地区,与现场考试显然不同。由于本研究数据有限,关于考生群体的特征还待进一步研究。回顾 Teng (2017) 的建议,HSK 作为标准化考试需要进一步考虑到各个地区中文教育不均等,以及学习者的多样性^[10]。居家考试由于不受空间约束(只是不能跨国考试),就更为突出这一点。

谈到信度,至于四级至六级居家考试得分高于其他两种形式,当然也不排除作弊行为。虽然居家考试采用手机作为辅助监考设备,但所看到的范围有限,因此还无法进行全程全范围监控。

张晋军和符华均推测,机考考生通过耳机听题,使听力得分更高^[7],而本研究出现纸笔考试听力得分高于机考和居家考试的现象(一级至三级考试),再说居家考试的考生不一定戴耳机,但在四级至六级得分当中,居家考试得分更高。笔者推测,考生注意力更受考试环境的影响,不仅是戴不戴上耳机的问题。在家里考试,考生单独在房间里答题,注意力更集中。这也可以用来解释为什么四级至六级居家考试阅读和书写得分都高于纸笔考试和现场机考。相反,一级至三级居家考试得分低于纸笔考试和机考,也可能因为注意力分散的影响,因为考生向监考老师提问一般都是开麦说话,说话的声音会

影响到其他考生。

HSKK 的对比结果显示, 参加 HSK (三级) 至 HSK (六级) 考生须同时参加 HSKK 显然对 HSKK 得分有很大的影响, 尤其是 HSKK (中级) 和 HSKK (高级)。两者平均得分都低于纸笔考试。不过以上提到的考生家里网络问题, 也可能影响 HSKK 居家考试得分。另外, 根据笔者的观察, HSKK 登录后的试录界面只显示几个按钮, 没有详细的操作说明, 让考生往往不清楚该怎样进行试录, 耽误他们进入试卷下载界面。

上面提到的网络或设备问题, 其实无论是现场机考还是居家考试都有可能发生。考试系统已经具备了补时功能。不过现场机考一般都有技术人员在, 能尽快解决网络或设备问题, 而在家里考试就要考生自己解决。考生解决不了就会影响到得分。

五 结论

通过居家考试与纸笔考试和现场机考的得分对比, 本研究发现 HSK 居家考试信度受几个因素的影响。第一是考试地点。因为考生各自在家里参加考试, 网络质量均不相同, 一有问题就可能影响得分。第二是考生群体。HSK (四级) 至 (六级) 的居家考试得分几乎都高于纸笔考试和现场机考, 本研究推测是考生群体不同于其他级别, 也不同于纸笔考试和现场机考。第三是监控效力。第四是考试环境, 房间里单独参加考试可能让考生注意力更集中, 或相反, 考生说话提问的声音会让其他考生注意力分散。HSKK 则因为新规定让参加 HSK (三级) 至 (六级) 必须参加 HSKK, 让泗水考点 HSKK 得分降低。

本研究以泗水考点为例, 有些级别的样本数量不是很多。因此还需要进一步研究更多的考点, 以获取更多的数据, 让研究结果更加全面可靠。另外也建议进行考点与考点之间的比较, 让我们对不同的考生群体有更好的了解。

参考文献

考试介绍--汉语考试服务网 [EB/OL]. [2018-03-08].

<http://www.chinesetest.cn/gosign.do?id=1&lid=0>. 关于我们--汉语考试服务网 [EB/OL]. [2022-07-31]. <http://www.chinesetest.cn/goaboutus.do>.

卢鹏. 印尼文教部统筹处和雅协联办汉考居家线上考试 - 千岛日报[N/OL]. 千岛日报, 2020-12-17[2022-07-31]. <https://www.qiandaoribao.com/2020/12/17/印尼文教部统筹处和雅协联办汉考居家线上考试/>.

- CHEN J Y W. Implementation of the Hanyu Shuiping Kaoshi and its Home Edition During the Covid-19 Pandemic: A Survey of European Test Centres[J/OL]. *International Journal of Chinese Education*, 2022, 11(2): 1-12. DOI:10.1177/2212585X221100877.
- 张泉慧, 张晋军. 新HSK为什么要取消合格线[J]. *测量与评价*, 2014(9): 14-17.
- PENG Y, YAN W, CHENG L. Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test[J/OL]. *Language Testing*, 2021, 38(2): 326-337. DOI:10.1177/0265532220957298.
- 张晋军, 符华均. 新HSK纸笔考试与网络考试比较研究[J]. *中国考试*, 2015, 12(11): 54-57.
- SELVARAJ A, RADHIN V, KA N等. Effect of pandemic based online education on teaching and learning system[J/OL]. *International Journal of Educational Development*, 2021, 85(May): 102444. <https://doi.org/10.1016/j.ijedudev.2021.102444>. DOI:10.1016/j.ijedudev.2021.102444.
- GENIS-GRUBER A, WEISZ G. Challenges of Online Exam Systems in the COVID-19 Pandemic Era: E-Assessment of Perception, Motivation, and Performance[M/OL]//ISAIAS P, ISSA T, KOMMERS P. *Measurement Methodologies to Assess the Effectiveness of Global Online Learning*. Hershey: IGI Global, 2022: 108-136. DOI:10.4018/978-1-7998-8661-7.ch005.
- TENG Y. Hanyu Shuiping Kaoshi (HSK): Past, Present, and Future[M/OL]//ZHANG D, LIN C H. *Chinese as a Second Language Assessment*. Singapore: Springer, 2017: 3-19. DOI:10.1007/978-981-10-4089-4_1.
- MUHAMMAD A A, OCKEY G J. Upholding Language Assessment Quality during the COVID-19 Pandemic: Some Final Thoughts and Questions[J/OL]. *Language Assessment Quarterly*, 2021, 18(1): 51-55. <https://doi.org/10.1080/15434303.2020.1867555>. DOI:10.1080/15434303.2020.1867555.
- OCKEY G J. An Overview of COVID-19's Impact on English Language University Admissions and Placement Tests[J/OL]. *Language Assessment Quarterly*, 2021, 18(1): 1-5. <https://doi.org/10.1080/15434303.2020.1866576>.

DOI:10.1080/15434303.2020.1866576.

ISBELL D R, KREMMEL B. Test Review: Current options in at-home language proficiency tests for making high-stakes decisions[J/OL]. *Language Testing*, 2020, 37(4): 600-619.

DOI:10.1177/0265532220943483.

PAPAGEORGIU S, MANNA V F. Maintaining access to a large-scale test of academic language proficiency during the pandemic: The launch of TOEFL iBT Home Edition[J/OL]. *Language Assessment Quarterly*, 2021, 18(1): 36-41.

<https://doi.org/10.1080/15434303.2020.1864376>.

DOI:10.1080/15434303.2020.1864376.