

中小学教师测评素养的调查研究

——来自浙江和山东省的样本对比

何虹燕

青岛大学 外语学院

DOI:10.12238/jief.v5i4.6370

[摘要] 本研究调查了125位来自浙江省和山东省中小学英语教师的教育测评素养。通过SPSS软件分析发现:当前中小学英语教师的自我测评素养得分较高;自评的测评态度、测评知识和测评技能水平均呈现积极态度;两省中小学英语教师的测评态度、测评知识、测评技能和测评总体得分情况不存在显著性差异。最后通过相关性分析,得到了显著影响教师测评素养的8项因素。

[关键词] 测评素养; 英语教师; 对比研究

中图分类号: G633.41 **文献标识码:** A

Survey and Research on the Assessment of Literacy of English Teachers in Primary and Secondary Schools

—Comparison of Samples from Zhejiang and Shandong provinces

Hongyan He

School of Foreign Languages, Qingdao University

[Abstract] This study investigated the educational assessment literacy of 126 primary and secondary English teachers from Zhejiang and Shandong provinces. Through SPSS software analysis, it was found that the current self-evaluation literacy scores of English teachers in primary and secondary schools are relatively high; The evaluation attitude, knowledge, and skill level of self-evaluation show a positive attitude; There is no significant difference in the evaluation attitude, knowledge, skills, and overall scores of English teachers in primary and secondary schools in the two provinces. Finally, through correlation analysis, 8 factors that significantly affect teachers' assessment literacy were identified.

[Key words] assessment of literacy; English teachers; comparative study

引言

教学测评是教学进程中的一个重要环节,在促进学生学业成就发展中发挥着不可替代的作用(Stiggins, 1991; Taylor, 2009)。对于教师而言,如果测评缺位,则失去了对学生学习活动的有效监控和反馈的依据。因此,教师的测评素养成为了制约教学活动的关键要素,也成为了教师专业发展的目标之一(许悦婷, 2013; 许明, 朱文君, 2018)。近年来,我国颁布的不少文件都对教师测评素质的发展提出了要求。要扭转不科学的教育评价局面,就要求教师提高自身测评素养,能够运用测评的知识和方法正确评估学生的学习活动。英语教学被认为是我国最早引入标准化测试的学科领域。然而,伴随着教育评价领域的不断发展和完善,现实中英语教师的测评素养却仍旧存在着三方面问题:考试思维统领评价活动,五项技能不娴熟以及评价伦理基本失

范(郑东辉, 2022)。新时代的教师需要怎样的测评素养? 目前我国教师教育测评素养处于什么状态? 如何有效提升? 国内外相关学者对此展开了不懈探索。

本研究旨在探讨浙江和山东两省中小学英语教师测评素养的现状和差异,并试图发现影响两省中小学英语教师测评素养差异的原因。研究问题包括3项:两省的中小学英语教师测评素养的基本情况如何? 两省在整体测评素养、测评态度、测评知识和测评技能上存在怎样的差异? 哪些因素对中小学英语教师测评素养存在显著性影响? 在高信效度的测评素养测试量表基础上,本文提出了改编的测评素养问卷,并利用SPSS进行数据分析。

1 国内外英语教师测评素养研究现状

有关测评素养的研究最早可以追溯到美国在20世纪90年代

初发布的《教师教育测评能力标准》。2005-2019年间国际测评素养研究热点主题围绕研究现状、职前教师、兴趣项目研究、社会情感支持、教师教育和测评素养提升等多个方面。测评素养研究对于语言测评和教师发展的关注度较高,而对于课堂测评的关注度却稍显不足。对于英语教师的测评素养研究而言,除了一般教师所需具备的教学测评素养,还突出了英语学科的特点。

相比之下,国内在英语教师测评素养的研究目前仍处于萌芽期(肖杨田&潘鸣威,2020),仍然存在一些不足。如研究大多围绕着测评知识展开,而缺乏了专门针对教师测评方法、测评技能、测评伦理、测评意识等系统性的调查。此外,研究对象的地域性覆盖面不足,忽视了地域因素的不同带来的政策导向、测评理念和外部支持对教师测评素养的影响,其研究结果的代表性受到质疑。

2 研究设计

2.1 研究对象

本研究对来自浙江和山东两个省份的中小学英语教师进行了问卷调查。125名英语教师参与了本次问卷采集,其中63名来自浙江省,62名来自山东省,两省数据相对均衡。在所有调查对象中女性116名,男性9名,绝大多数都为女性教师,这基本符合中国英语教师的情况。数据经过初次检查,125份问卷都被认为是有效的,能够进行下一步的数据分析。

2.2 过程

为了方便数据收集,本研究使用了问卷星在线数据收集工具。本次问卷数据收集得到了很多一线的中小学英语教师帮助,其中包括两省英语教研名师的帮助,最后以滚雪球的方式收集到了数据。诚然,相对于两省众多的人口来说,本次收集到的样本数量并不够理想,但是已经能够基本满足本次研究。

2.3 问卷信息

基于已有教师测评素养的相关问卷,本研究参考了郑东辉(2014)教师评价素养问卷和林敦来(2011)的《中学英语教师评价素养问卷》。问卷分为三部分,第一部分是基本信息,第二部分是测评素养量表,主要涵盖了测评态度、测评知识和测评几个维度的量表,第三部分是测评经历。基本信息以单选题的形式展开,共10题。测评素养量表是自我测评形式,采用了李克特五级量表。其中,测评态度量表有7题,主要涵盖了测评重要性、测评公平和测评伦理。测评知识在内部一致性检验后对题项进行了调整,保留了9题,主要是测评的相关知识。测评技能从课堂教学和大规模测评两个角度,包括测评目标、工具、过程和结果,在调整题项后保留了17题。测评经历主要从选修课程、入职要求、课题研究、专题培训、自学经历、专家指导、编制试题、大规模测评阅卷共8个维度设置了8个单选题,此外还设置了1道开放题了解各地区对提升教师测评素养所作的努力。

根据SPSS软件结果分析,测评态度、测评知识、测评技能和量表整体的Cronbach's α 值均大于0.9,KMO值为0.92, $P<0.001$,说明量表具有很好的信效度。查看初始特征值,在主成分中显示3个维度,每个维度都匹配了量表在设计时的对应维度。

3 结果与讨论

3.1 测评素养总分与维度分

利用SPSS26进行了描述性统计分析,表3.1报告了本次125名中小学英语教师的总体测评素养得分和各维度得分。125名被试者的测评素养总分和三个维度的得分均值均高于4分,原始量表4分等于同意,说明本次调查中的中小学英语教师的测评素养在三个维度上均呈现积极态度。对四个得分进行排序的结果如下:测评态度>测评素养整体>测评技能>测评知识。同时,测评知识和测评技能的得分均值都低于测评态度,说明中小学教师普遍认为测评对于教学和促进学生学习很重要,但是在具体的测评知识和技能方面仍然相对缺乏。

表3.1 整体测评素养总分与维度分的描述性统计分析

	个数	最小值	最大值	平均值
测评态度	125	3.71	5.00	4.58
测评知识	125	2.56	5.00	4.16
测评技能	125	3.00	5.00	4.33
总体	125	3.24	5.00	4.34

本研究采用单样本T检验对两地教师测评素养总体得分进行了分析,测评态度、测评知识、测评技能和总体得分的P值均为0,小于0.05,说明本次调查中两地中小学英语教师在测评态度、测评知识和测评技能上的积极认同有显著性。

3.2 两省中小学英语教师数据对比

通过对两地的数据情况进行排序,结果如下:浙江来说,测评态度>测评技能>测评素养整体>测评知识;山东来说,测评态度>测评素养整体>测评技能>测评知识。本次调查中山东教师有56.5%属于研究生学历,而浙江85.7%的教师是本科学历,可见本次调查中的山东教师学历更高,但在测评知识、技能和整体得分上均不如浙江,说明学历因素并未对两地中小学英语教师的测评素养产生影响。

表3.2 浙江省和山东省教师测评素养数据对比

	浙江			山东		
	最小值	最大值	平均值	最小值	最大值	平均值
测评态度	3.71	5.00	4.5646	3.86	5.00	4.60
测评知识	3.00	5.00	4.2363	2.56	5.00	4.08
测评技能	3.18	5.00	4.3978	3.00	5.00	4.26
总体得分	3.36	5.00	4.3891	3.24	5.00	4.28

通过对样本进行了独立样本非参数检验。秩和检验结果显示四个维度的P值分别是0.848,0.191,0.158,0.266,均大于0.05,说明浙江省和山东省教师测评素养调查中测评态度、测评知识、测评技能和测评总体得分情况不存在显著性差异。

3.3 影响因素

本研究利用SPSS26.0对教师测评素养的影响因素进行相关性分析。影响因素分别包括自学经历、测评课题研究、职前课程选修、在职专题培训、政策制度、测评专家指导、命题经历、大规模测试阅卷经历8个方面。下表数据由斯皮尔曼相关性检验得到。

表3.3 两省独立样本非参数检验结果

dependent variables	Group	Percentiles	Z	P
测评态度	浙江	4. 7143 (4~5)	-0. 191	0. 848
	山东	4. 7143 (4. 25~5)		
测评知识	浙江	4. 2222 (3. 7778~5)	-1. 307	0. 191
	山东	4 (3. 6667~4. 6667)		
测评技能	浙江	4. 2353 (4~5)	-1. 411	0. 158
	山东	4. 1176 (3. 9412~4. 7794)		
总体得分	浙江	4. 303 (4~4. 9394)	-1. 113	0. 266
	山东	4. 2121 (3. 9621~4. 7273)		

表3.4 斯皮尔曼相关性检验

	测评态度	测评知识	测评技能	整体测评素养
职前测评课程	0	-. 282**	-. 233**	-. 235**
	1	0. 001	0. 009	0. 008
政策制度	-0. 166	-. 314**	-. 300**	-. 313**
	0. 065	0	0. 001	0
在职专题培训	-0. 086	-. 271**	-. 267**	-. 235**
	0. 343	0. 002	0. 003	0. 008
自学经历	-. 218*	-. 421**	-. 354**	-. 389**
	0. 015	0	0	0
测评专家指导	-. 210*	-. 388**	-. 375**	-. 373**
	0. 019	0	0	0
测评课题研究	-. 193*	-. 360**	-. 388**	-. 361**
	0. 031	0	0	0
命题经历	0. 093	-0. 031	-0. 077	-0. 054
	0. 303	0. 735	0. 393	0. 552
大规模测试阅卷经历	-0. 087	-0. 171	-. 295**	-. 240**
	0. 336	0. 057	0. 001	0. 007

相关性分析结果显示, 职前测评课程、政策制度、在职专题培训、自学经历、测评专家指导和测评课题研究的P值均小于0. 01, 说明以上几个因素与测评知识、测评技能和整体测评素养的相关性显著。自学经历、测评专家指导和测评课题研究对测评态度的P值小于0. 05, 说明上述三项因素对教师测评态度具有较为显著的影响。此外大规模测试阅卷经历对教师测评技能和整体测评素养的P值小于0. 01, 说明这项因素能够显著影响教师测评技能和整体测评素养。而命题经历则与教师测评态度、知识、技能和整体测评素养不存在显著性影响。令人感到惊讶的是, 职前测评课程等7项与教师测评素养的相关系数均为负数, 这一发现引起了本次研究的极大关注。

4 总结

本研究对125位来自浙江和山东省的中小学英语教师进行了调查。通过SPSS26. 0的描述性统计分析, 并且从整体测评素养、自我测评态度、测评知识和测评技能分别进行了讨论。在影响因素探究中, 本研究发现职前测评课程等7项因素与测评知

识、测评技能和整体测评素养中的一项或几项的相关性显著, 而命题经历则不存在显著性影响。特别地, 本次相关性研究中职前测评课程经历等7项因素与测评素养得分之间的关系为负相关。在调查教师相关测评经历时, 68. 9%的教师修过测评相关的课程, 70. 6%的教师确认参加过测评相关的专题培训, 然而对比测评素养的得分情况, 测评知识在整体、浙江和山东样本中, 都处于最低值。因此有理由怀疑, 英语教师在职前或者在职期间的培训可能是无效的, 缺乏英语教育测评的针对性。此外, 77%的教师认同没有专门的测评专家指导日常教学评价实践, 84%的教师没有参加过专门的测评项目或课题研究, 以至于缺乏测评实践的专业性指导, 而可能将测评认知简单地定义为大型考试中的阅卷工作。因此, 在自评时的测评技能维度中打分偏高。从自学测评因素来看, 如果缺乏了系统的训练和指导, 教师容易盲目学习且无法将测评知识有效运用到实践中。并且也无法确定教师在自学时选读的测评相关书本的质量高低。在入职政策要求方面, 选择“是”的教师人数与“否”的人数差距不大, 但在影响测评素养的系数值占了30%, 这说明我国在将语言测评素养作为入职必要条件一项并未产生显著性效果。

由于研究时间有限, 本研究仍然存在以下问题有待改进。首先是自评问卷主观性过强, 若要获得客观性评价, 就需要后续研究从客观操作出发进行调查。其次, 本研究的样本数据仍旧是小样本, 对比人口规模庞大的山东省和浙江省来说, 在说服力上仍有限制。因此, 后续调查需要不断扩大样本基数, 从而提高问卷结果的可信度和代表性。

【参考文献】

- [1]Stiggins, R. J. (1991). Assessment literacy. Phi Delta Kappan,72(7):534-539.
- [2]Taylor,L.(2009).Developing assessment literacy. Annual Review of Applied Linguistics,29:21-36.
- [3]林敦来,高森.教师评估素养:理论与实践[J].外语教学理论与实践,2011,(4):29-37.
- [4]许明,朱文君.教师评价素养研究:近十年国内核心期刊的统计分析[J].教育观察(上旬),2018,7(7):8-11.
- [5]许悦婷.外语教师课堂评估素质研究述评[J].外语测试与教学,2013,(4):42-50.
- [6]肖杨田,潘鸣威.教师测评素养国际研究热点与趋势分析(2005-2019)——基于SSCI期刊论文的可视化分析[J].外语研究,2020,37(5):43-51,74.
- [7]郑东辉.教师评价素养发展研究[D].上海:华东师范大学,2009.
- [8]迟春苗.小学教师评价素养发展研究[J].课外语文(下),2018,(1):96-97.