

教师评价素养的内涵、结构组成 及测量研究综述

张志红

(北京师范大学 中国基础教育质量监测协同创新中心, 北京 100875)

【摘要】教师评价素养作为教师的核心专业能力,是学生学习的重要保障。关于教师评价素养的内涵界定和结构组成已经形成三种视角,包括单维度、多维度、社会文化性。从测量上看,主要有情境题、量表题及量规等方法;已有的测量工具比较落后,质量不好;测量维度的选择可以灵活多样,但对教师评价素养的多维度测量是未来发展需要解决的课题。基于已有研究经验及启示,需要立足我国文化背景,对我国教师评价素养进行本土化的认识、发展及测量。

【关键词】教师评价素养;学生学习;教师评价素养测量

【中图分类号】G451.6 **【文献标识码】**A **【文章编号】**1005-3476(2023)06-0060-10

DOI:10.16215/j.cnki.cn44-1371/g4.2023.06.003

评价是支持与改善学生学习的最强干预变量^[1],将评价素养视为教师的核心专业能力已得到广泛认同。斯蒂金斯(Stiggins)曾指出教师每日的专业活动有1/3到一半的时间用于学生评价^[2],波帕姆(Popham)甚至指出缺少评价素养等于专业自杀^[3]。因此,教师需要具备评价素养,发挥评价的最大潜能,发展高质量的评价来促进学生学习。当前,我国开启了以核心素养为导向的课程改革,教师作为评价者的角色越发重要。我国颁布的《深化新时代教育评价改革总体方案》提出的“四个评价”为学生评价定下了总的基调,也对教师评价素养的发展提出了新的要求和挑战。在新时代背景下,亟待加强对教师评价素养的发展与研究,以发挥评价促进学生学习的功能。

因此,有必要基于已有的研究,获取视角上的拓展和参考。对教师评价素养的兴趣可追溯到关于教师资格认证所要求的教育测量课程、

对测验的使用和选择、对学生评价技巧的掌握等内容的调查^[4]。然而,这些研究主要是对测量素养的调查。美国基于对高质量教育的诉求,在上世纪80年代进入了评价时代(assessment era)^[5],但主要以问责为评价目的。在这种背景下,认识到评价的潜在教育价值,为发展教师的评价能力,美国于1990年颁布了《在教育中的学生评价的教师素养标准》(standards for teacher competence in educational assessment of students)(简称90标准)^[6]。然而,该标准并未把评价素养作为一个专门的术语,而是提出了教师在评价中应达到的标准或应掌握的知识与技能,但可视为教师评价素养发起的背景。斯蒂金斯于1991年首次提出评价素养这个术语之后,评价素养才获得真正意义上的关注。因此,本文主要关注近三十年来教师评价素养的相关研究。当前,关于教师评价素养的认识比较碎片化,对教师评价素养的发展和认识带来

作者简介: 张志红(1983-),女,山东临沂人,北京师范大学中国基础教育质量监测协同创新中心博士生,主要研究方向为学习评价、教学评价。

了一定的挑战,需要做出系统梳理,为教师评价素养的发展与研究提供启示。

一、教师评价素养是什么

关于教师评价素养是什么的问题,主要从教师评价素养的内涵及其组成去探索。内涵是理解和研究教师评价素养的逻辑起点和基础,而组成是对教师评价素养是什么的进一步解构,它关照到评价素养的组成要素以及各要素之间的关系,每个要素下又有不同的内容组成。随着对评价素养认识和研究的推进,教师评价素养是什么的问题具有了不同的视角,总体上可以分为单要素、多要素及社会文化性三种视角。

(一) 单维度视角下的教师评价素养内涵及组成

单维度视角将评价知识与技能这一维度视为评价素养,包含评价的知识、程序以及技能等。国外持这种视角的研究者偏多。

1. 教师评价素养内涵

上世纪90年代以间接界定为主要方式,对具有评价素养的教师所具有的特征进行了描述,以斯蒂金斯等人为主要代表。斯蒂金斯在1991年指出:教育决策需要以学生成绩为依据,决策与数据质量有关,教师需要能够辨别出什么是好的数据,但是教师这方面的素养还不足。基于此,他提出了具有评价素养的教师应具备的特征^[7]:有评价素养的人能够理解和区分高质量与低质量评价的作用,能够将评价知识用于学习成绩的测量中;思考并能够回答“成绩可以为学生提供什么样的评价信息以及会对学生产生什么样的影响”;也需要知道高质量评价的组成。此后,斯蒂金斯在1995年对该特征做出了进一步明晰^[8]:有评价素养的人知道好的评价和不好的评价之间的区别,并不会被有时神秘而总是令人生畏的评价技术世界所吓倒;知道正在评价什么、为什么要那样评价、怎样最好地评价、如何生成良好的学生表现样例、要知道在评价中有可能出什么错、怎样阻止

错误的发生、能够意识到不好的与不精确的评价带来的消极后果。总体而言,超越了传统的测量和统计素养,强调具有评价素养的教师能够区分和识别出高质量的评价,限定在对评价知识与技能的要求,更多表现为对学生成绩的促进性。除了斯蒂金斯,也有其他研究者做了描述。尽管表述不同,但一般强调评价要超越以往的统计测量或总结性测验。例如,谢帕德(Shepard)认为要超越单元测试和考试的总结性评价^[9];波帕姆(Popham)指出评价素养不是那些与教育者日常决定无关的大量深奥难以理解的概念与术语,具有评价素养的教育者需要理解少量的常识类的测量基本原理,而不是一大堆心理测量的新奇特术语^[10]。然而,总体上未超越斯蒂金斯的描述范围。

随着研究的进展,教师评价素养的内涵得到了明确的界定。一种是以教师为视角或主体的,体现的是“教”为中心,强调教师应该掌握什么样的评价知识与技能、能够做什么。例如,佩特诺(Paterno)指出:教师要掌握健全评价实践的基本原则,包括术语、评价方法和技术的发展与使用、熟悉评价质量标准 and 替代传统测量学的评价方法^[11];再如,波帕姆强调:理解可能影响教育决定的基本的评价观念和程序^[12]。另一种虽然也是把知识与技能作为教师的评价素养,但是关注到对学生学习的促进性。比如,韦伯(Web)对评价素养内涵的界定中就有所体现,他指出:评价素养是了解评价学生知道什么和能做什么的方法,以及如何解释这些评价结果,还要知道如何应用这些结果来提高学生的学习和项目的有效性,应该专注于利用评价信息去监控学生学习成就的进步,改善教学和教学项目的效能去帮助学生学习^[13];另外,斯蒂金斯等人于2017年提出:教师评价素养是指精确测量和报告学生学业成就的知识与技巧,使用评价过程和结果去改善学生学习^[14]。可以看出,斯蒂金斯对评价素养的认识,较之以往有了明显的不同,既关注教师知道、能做什么,又关注学生的学,同时注重

对总结性评价和形成性评价的平衡性使用。

简而言之,知识与技能视角下教师评价素养的内涵有了明确的界定,将教师评价素养往前推进了一步。在界定取向上,走向关注促进学生的学习。但是,局限于知识与技能,总体而言具有明显的工具主义取向。

2. 教师评价素养结构组成

只是看内涵,还无法确定具体包含哪些内容。通过对其组成的梳理,有助于进一步明确该视角下的教师评价素养到底是什么,把握其中的具体成分。

总体而言,知识与技能这个维度下包含多个要素或指标。对教师评价素养组成最早的探索,当属 90 标准提出的七个具体的要素^[15],然而斯蒂金斯认为其包含的范畴不足,建构了另七个要素^[16]。实际上,二者具有内容上的一致性,更多关注教师总结性评价能力。布鲁克哈特(Brookhart)指出 90 标准有些过时了,未考虑形成性评价理念及基于标准的变革背景,他在 90 标准的基础上提出了十一个要素,增加了形成性评价的内容^[17]。斯蒂金斯于 2006 年在已有的基础上又提出了五个要素^[18],而且提供了具体的操作策略,同时考虑到形成性评价与总结性评价的作用,强调对二者的平衡性使用,将学生视为评价的重要主体。随后,又将五个维度做了调整,更加强调学生在评价中的参与性,并突出了促进学生学习的指向性,促使学习的效果最大化^[19]。可以看出,知识与技能视角下的教师评价素养的要素从主要关注总结性评价转变为聚焦于如何利用评价帮助学生学习。

从布鲁姆目标分类学看,这些要素都可以归在知识与技能这一维度之下,为教师评价素养的发展提供了可操作化的参考。虽然切入点和表述存在着差异,但是存在着共性。例如,强调目标的清晰性、方法的适切性与多样化、评价的有效设计与管理、评价结果的交流、评价的伦理性等,在价值取向上逐步具有了促进学生学习的理念。

(二) 多维度视角下的教师评价素养内涵及组成

多维度视角下的界定体现了教师评价素养是由多个维度组成的,关注到了教师的情感,认识到了教师在评价中的身份以及是谁的问题,凸显了教师的自身角色、自主性和自我反思的重要性。

1. 教师评价素养内涵

多维度视角下教师评价素养的界定包含了多个维度,不再仅局限于知识与技能单一维度。国外从教师评价素养的多维度界定主要局限于近期的少数研究者。例如,德卢卡(Deluca)提出:教师评价素养是在基于标准的框架内,结合评价实践、理论与信念以支持教与学^[20];再如,帕斯托(Pastore)指出:教师评价素养是由一系列相关的知识、技能和倾向组成的,可以用来在课堂和学校设计与实施相关恰当的评价^[21]。而我国学者倾向从多个维度去界定,比如赵雪晶指出:教师评价素养是指在教学过程中从事各种评价活动所需的素质和能力,包括教师对于评价的正确认识和价值取向、源于教学需要的评价冲动和意识、与评价相关的知识摄取和在教学实践中评价技术的应用^[22];又如,孟丹将教师的学生学业评价素养界定为:是指教师在完成学生学业评价这项工作所应具备的素质与修养,它不同于素质,包括教师的学生学业评价态度要素、教师的学生学业评价知识要素以及教师的学生学业评价技能要素的总和^[23]。

知识与技能是不可或缺的成分,但是融入了教师倾向或态度等成分,有利于推动教师对评价角色的确立和身份的自我认同,是评价成为专业属性的重要特征。不再只是把教师视为评价的被动执行者,而是具有了自主性与能动性。因此,和单要素视角下侧重于知识与技能的界定相比,多维度视角下的界定更具有了价值理性。

2. 教师评价素养结构组成

内涵只是体现了教师评价素养的多维度组

成,而对于维度下的要素组成以及要素之间的关系却没有体现,通过对教师评价素养结构组成的认识,有助于形成更为具体的感知。

在结构框架上看,多要素结构组成的研究以我国为主,以郑东辉和赵雪晶为主要代表,国内大多研究也对其作了参考。郑东辉借鉴斯宾塞等人的胜任力冰山模型,建构了包含评价知识、评价技能和评价态度三个维度的结构模型^[24]。三个维度之间的结构关系相互连贯和一致,有什么样的观念就需要有什么样的知识与技能。通过该模型可以看到,评价知识包含外部和内部评价知识两个要素,评价知识一部分藏在水底下,一部分在水面上,这是因为评价知识具有显性和隐性之分;评价态度包含外部和内部态度两个要素,位于冰川的最下层;而技能包含六个要素,则是在态度与知识的双重支持下完全浮于水面的能力表现。赵雪晶建构了教学评价素养的伞形图,包含与内涵一致的四个相互联系又独立的维度,分别为评价意识、评价态度、评价知识与评价技能^[25]。把评价态度视为“伞把”,是评价活动实施的主心骨,决定着评价行为。评价意识包含无意识、潜意识、有意识三个要素,把评价意识视为伞尖,和态度都属于人的意念,只有具备了评价意识才有可能使得评价行动充分展示出来。评价知识包含理论、应用和个性化的知识三个要素,评价技能包含设计和运用评价方案、利用评价反馈和解释评价结果四个要素,评价知识和评价技能是教师评价行为的两个外显特征,共同构成了“伞面”。评价知识水平和技能的大小会决定教师是否能顺利完成预设的评价任务,但评价态度和评价意识决定教师的行为选择,即决定着教师是否愿意完成某些评价任务。

可以看出,教师评价素养中的各维度下都有相应的多个要素组成,且各维度之间是一致的,存在着密不可分相互制约的关系,需要整体推进和发展。

(三) 社会文化视角下教师评价素养内涵及组成

单维度和多维度视角下对教师评价素养的界定,多是从教师为主体出发,忽视了学生的主体性以及其他影响因素。社会文化视角下的教师评价素养考虑到了师生之间的协商与交互作用,也考虑到了其他社会文化因素的影响。

1. 教师评价素养内涵

当前,研究者开始基于社会文化视角去重新界定教师评价素养,落脚点以学习为指向,基于评价主体间以及主体与环境间的关系进行探讨。

以下属于典型的社会文化视角下的教师评价素养界定。比如,有研究者指出:教师评价素养发展不仅仅是一种评价知识的积累,而是一套复杂的、与环境相适应的相互关联的能力的发展。评价素养依赖于认知特征、情感和信念系统以及社会、文化和制度的影响^[26];再如,威利斯(Willis)强调:评价素养是动态的、依赖于情境的社会实践活动,在评价的启动、发展和实践过程中,教师与学习者相互沟通、协商课堂和文化知识,以实现学生的学习目标^[27];又如,研究者以促进学生学习的评价为主旨,探索了促进学生学习的教师评价素养:促进学生学习的教师评价素养就是做出高度基于情境的、公平的、一致的与可信赖的决策,为学与教提供信息,以支持学生和教师的专业学习。教师的目的是培养学生和其他利益相关者的能力和信心在评价、学习和教学活动中发挥积极的作用,为更有效的学习提供必要的支持^[28]。

基于以上定义可以看出,社会文化视角下的教师评价素养具有文化性、动态性、交互性及情境性。其中,师生在评价中的共同责任性以及交互性是非常重要的特点,主体之间的沟通与协商成为教师评价素养发展的关键桥梁,最终指向学生的学习。

2. 教师评价素养结构组成

社会文化视角下教师评价素养的结构框架,

不仅考虑到了教师评价素养的多维度组成,并将社会文化因素纳入结构框架中,因此具有了生态性和系统性的特点。

典型代表是许(Xu)建构的实践中教师评价素养框架(teacher assessment literacy in practice),评价素养要素包含评价知识库、评价观念、社会文化背景、评价实践、教师学习、评价身份构建(或重构)等^[29]。该框架较为系统,不仅考虑到了教师评价素养的各个维度,在评价的专业角度上将教师视为评价者(assessor),还将教师专业学习纳入进来,并纳入了社会文化因素。卢尼(looney)指出知识与技能已经无法满足对教师评价能力的需求,建议进行重构^[30]。他认为当前教师的角色更多关注教师的功能,即能够做什么,具有工具主义取向。鉴于此,他从教师身份认同的视角建构了教师评价素养结构,具体组成维度包含“我”知道、“我”的角色、“我”是自信的、“我”相信、“我”的感受,同时强调评价身份受到社会文化因素制约,具有不稳定性 and 动态性;帕斯托(Pastore)构建了教师评价素养的三维度模型,将社会文化因素纳入了该模型,具体包含:观念知识维度、实践维度、社会情感维度、课堂背景、专业实践、专业智慧以及国家教育政策^[31];另外,还有富尔默(Fulmer)建构的模型^[32],虽然他依然把评价素养视为评价知识,但是这个模型实际上包含了评价素养的其他维度,例如价值、观念、实践,并且将宏观、中观和微观因素纳入到该模型中,对我们建构教师评价素养模型依旧具有重要的启发意义。从具体组成上看,教师评价素养不仅具有多维度性,还受到具体的社会文化因素的影响,并越来越强调教师作为评价者的身份。

基于以上分析,可以更为直观地看出评价素养的组成维度以及影响因素。意味着在不同的文化背景下评价素养有不同的特征,需要基于本土化特征及需求进行解读和重构。评价者身份的提出,也凸出了评价的专业属性和教师

的评价自主性。社会文化因素国外近几年开始逐渐强调,然而,这种视角在国内是比较匮乏的,需要学习和借鉴。

二、如何对教师评价素养进行测量

教师评价素养测量,有利于更为直观和具体地把握教师评价素养是什么、包含什么、到了哪里。

首先,从测量方法上看,研究者主要采用基于内容或情境的试题、量表题以及量规的方式。最早对教师评价素养的测量聚焦于评价知识与技能,主要采用测验题,大多以90标准为测量依据。目的是考查教师体现在90标准上的水平,分析教师评价素养的优势与弱势,从而为评价素养水平的提升提供干预和决策依据。第一个测量为普莱克(Plake)于1993年开发的TALQ(teacher assessment literacy questionnaire)^[33],它是基于一定内容的选择题。后来,其他研究者以TALQ为蓝本修订或开发了类似的测量工具,例如CALI(Classroom Assessment Inventory)^[34]、ALI(Assessment Literacy Inventory)^[35]、修订的ALI^[36],都属于对错选择题,其中修订的ALI是基于情境的选择题。这些测量工具中的题目数量都是35题,每个维度对应着5个题目,主要是关于总结性评价或统计测量的题目。整体上看,利用这些工具所获得的测量结果存在一致性,例如一致性系数比较低。到目前为止,以上工具仍是主要的评价素养测量工具。另外,德卢卡(Deluca)依据国际上的主要评价标准,通过主题编码的方式发现了评价目的、评价过程、测量理论、评价伦理四个要素,然后依据美国2015标准和当前特别受关注的教育主题,开发了情境题,设计了5个情境,每个情境下有4个题目^[37]。题目同时关注到总结性评价和形成性评价,但是这些情境题的选项不分对错,只要能够解决其评价目的实现教学就可以,主旨是为了促进学生学习。另外,在评价素养的测量中,量表题主要用于其他维度的测量,这

在我国比较常见,但是在测量的设计和实施上存在着不科学的问题。也有研究者开发了量规(rubric),帮助教师发展评价能力。量规属于过程性的评价工具,每个评价维度都有不同等级的表现性描述,为教师评价发展提供了可参考的质量标准。量规可用于对教师评价素养做出跟踪以及对其发展轨迹加以了解,研究证实了量规对于教师评价素养的发展是非常有效的^[38]。目前量规在其他专业领域使用的较多,例如医学教育中,量规可用于支持护理教育学习,追踪知识技能等维度随时间的变化和发展。然而,量规在教师专业发展中很少使用,尤其是在评价素养发展中使用更少。对于教师评价素养现状的问题,除了量化研究方法之外,一些研究者也开始探索质性研究方法,例如个案叙事、日志、访谈及观察等方法。

其次,从测量工具的质量上看,工具信效度不够高,更多偏向于总结性评价和知识与技能的测量,题目的代表性不好。得出这样的结论,一方面是研究者使用的文献代表性还有待探讨,但也确实反映了传统测量存在的问题。戈奇(Gotch)对1991-2012这一阶段的36个教师评价素养测量进行了分析^[39],他发现内容效度、分数稳定性、内部一致性都不满足要求,即整体心理测量属性较低。卢尼对已有的测量工具题目进行分析后,发现仅有两个工具考虑到了形成性评价或为了学习的评价(assessment for learning)。同时,卢尼发现超过四分之三的题目是聚焦于教师的评价知识、技能和实践的,他强调需要对教师作为评价者的身份进行重构,不能仅对知识与技能进行测量,在调查量表中需要涵盖教师的信念、信心或自我效能等维度^[40]。德卢卡对八个工具做了分析^[41],也发现仅有两个工具考虑到了为了学习的评价。主要原因在于这些测量工具所依赖的标准落后了,大多是基于90标准而做的测量,而90标准是在形成性评价改革之前开发的,它未考虑到新时代背景下形成性评价和为了学习的评价理念。

最后,从教师评价素养测量维度上看,国

内主要从多个维度进行测量,国外主要聚焦于评价知识与技能这一维度。在国内,教师评价素养多维度的测量在学位研究中较为普遍。例如,郑东辉从评价知能与评价态度两个维度进行了测量^[42],这里的评价知能主要是从CALI改编而来的。赵雪晶参考了美国90标准以及TALQ、CALI等工具,从评价知识、评价技能、评价态度与评价意识四个维度进行了测量。国内其他的研究,大多也借鉴了基于90标准开发的工具、以及郑东辉和赵雪晶的研究框架和工具。国外专门就教师评价素养多个维度进行测量的其实并不多,主要是德卢卡从评价方法、评价自信以及评价学习和培训进行了测量^[43]。国外大多借鉴基于90标准开发的工具,依旧把教师评价素养视为知识与技能,有部分学者将其与观念和实践等进行关系研究,但是这些研究对于教师评价素养的测量研究还是提供了重要启示。例如,斯奈德(Snyder)对教师评价素养(知识与技能)与评价观念和现实之间的关系进行了研究^[44]。还有一些研究不再以评价素养为话语体系,开始专注于某一维度或维度之间关系的研究,并且测量题项也关注到了形成性评价和总结性评价的平衡,其实对教师评价素养的测量也具有重要的参考意义。例如,布朗(Brown)开发的观念工具专门对教师的评价观念进行了测量,包含改善教学、学生问责、学校问责和无关四个因子,前两个因子属于为了学习的评价的内容,后两个因子属于对学习的评价的内容^[45];再如,詹姆斯(James)对教师评价实践与评价价值两个要素及其关系进行了测量^[46];另外,迪恩(Deen)从评价能力、评价实践和评价价值三个要素及其关系进行了测量^[47]。虽然以上不是以评价素养为视角进行研究的,但是每个维度却是评价素养的重要组成,使得教师评价素养的测量更加聚焦和深入,为未来测量教师评价素养提供了重要启示。

综上,教师评价素养的测量方法和要素灵活多样化,但是,现有的测量工具整体上存在

着心理测量属性比较低以及题目对当前的背景和评价需求代表不足的问题。在测量维度选择上,国内大多倾向于测量教师评价素养的多维度,国外对教师评价素养的测量大多将其视为知识与技能。国外也有一些测量不再落在评价素养上,可以考虑几个维度关系测量,也可以考虑只就某个维度做深入的测量。但是,当前已有研究者将评价素养模型进行了推进,不仅包含了多要素,并纳入了社会文化因素,这是为了适切评价素养的发展趋势。然而,对于教师评价素养进行全貌式的测量,实际测量与理论研究之间存在着滞后性,在未来研究中应对其进行推进,也是未来研究需要做出的突破。

三、讨论与展望

教师评价素养是与社会文化背景紧密相关的。需要立足我国独特的社会文化背景,并结合当前的评价改革和诉求,对教师评价素养的内涵与要素进行重识,在此基础上,探索适合我国教师评价素养发展的策略,并开发有效的测量工具。

(一) 重识我国社会文化背景下教师评价素养内涵

从教师评价素养的内涵上看,研究逐步关注到了评价素养对学习的促进性、多要素性以及社会文化属性。当前我国正处于学习为本的评价变革与传统考试文化双重背景的影响下,需要立足我国特殊背景去重新认识教师评价素养。

首先,需要考虑评价素养以促进学生学习的指向性。当前国际上已进入促进学生学习的时代,我国自2001年实施新课程改革以来就倡导实施以学习为本的评价,核心素养的新时代追求促进深度学习、高阶思维能力发展,更需要以评促学。当前,我国也针对评价中的痼疾提出有针对性的“四个评价”,为教师实施促进学习的评价提供了基调和指南。但是,我国的传统考试有存在的需要和理由,不能完全

摒弃和脱离,有必要发掘总结性评价中的积极成分与形成性评价协同发挥作用,这是体现教师评价素养水平非常重要的体现。

其次,应考虑到教师评价中师生之间的协商性及其他因素的影响。在传统评价中,教师把学生视为被动的接受者,学生缺少参与性和话语权,这是控制和管理为目的的评价。倡导对学生赋权赋能,鼓励学生积极参与到评价中,建构主体间交互式的评价共同体,成为我国教育及评价改革的诉求。在评价中,师生之间的协商、建构关系,既符合促进学生学习理念的诉求,也符合第四代评价的要求。另外,教师评价素养除了受制于师生之间的关系,也受到教师自身的学习与反思能力、课堂环境、学校组织与文化氛围以及大的国家文化的影响。因此,在认识教师评价素养时,需要结合微观、中观及宏观等因素去洞悉教师评价素养内涵。

最后,也要考虑教师评价素养的组成要素。除了通过评价知识与技能这一要素来体现教师评价素养之外,也需要考虑到评价态度、情感或观念等要素是教师评价者身份的重要体现。

(二) 探索适合我国教师评价素养专业发展的有效策略

研究证实教师评价素养具有可塑的特征。教师参与持续的评价学习可以显著提高其评价素养,然而教师评价素养发展的现实支持却远远不足。当前,尤其是形成性评价素养的不足是世界教育系统中的问题。而我国深受考试文化的影响,教师形成性评价素养的不足显得更为突出,那么如何去发展教师的评价素养呢?

第一,为教师评价素养的发展开发评价标准,包含内容标准和质量标准。国际上教师评价素养的发展与测量离不开对评价内容标准的参考,例如美国的90标准以及2015标准就有相当大的影响力,其他多个国家也专门开发了评价内容标准或者作为一个模块嵌入到教师专业标准中。然而,我国缺乏教师评价专业标准的开发和制定,多是通过融合的方式体现在教师专业标准或课程标准中。例如,我国《中学

教师专业标准》把教育评价作为教师专业的一个要素,仅包含三个比较笼统的指标,对教师来说难以起到有效的参考作用。目前评价是我国教育改革的关键与首要任务,虽然提出的“四个评价”为实施教育评价定下了基调和方向,但是如何去理解和有效实施依然是难题。需要依据“四个评价”,平衡形成性评价和总结性评价去开发标准,从而为实践提供具体化的参考,可以参考国际的做法开发评价内容标准。鉴于我国教师评价素养整体水平偏低,可以基于内容标准开发质量标准,以量规的形式呈现。质量标准或量规,一方面可以帮助教师进行过程性的自评,随时了解自己到哪里了,也可以追踪自己的发展轨迹,同时也可以成为专业培训和测量的依据。另外,我国教师入职还缺乏评价素养的要求,如果在教师资格考试或录用能够将评价标准作为依据,在某种程度上也会提高教师评价素养。

第二,需要为教师提供有针对性和个性化的评价学习机会。研究证实职前教师 and 在职教师之间、不同职业阶段群体之间存在着显著性差异。因此,需要努力探索有效的方式,对教师提供有针对性和个性化的教育。然而,在教师评价素养的发展中,并未将职前和在职评价教育做出系统的规划和支持。职前教师教育中,评价教育是教师未来进行评价实践的基础,有利于提高其评价的效能感。需要为职前教师提供系统的评价理论课程及评价实践的机会,加强与基础教育学校的连接,促进其专业准备。国内大多师范生的评价课程,大多属于一般意义上的教育测量与评价概论方面的理论知识,缺乏针对具体课堂评价实践的体验和指导,对教师评价素养的发展效果并不明显。因评价是发生在具体情境中的,需要在具体教学中提供有针对性的指导,这样才有助于评价内化和迁移应用。而对于在职教师,大多教师教育或培训项目,很少把评价作为专项,只是在其他培训中有所涉及,并且偏向于理论层面,对于教师评价培训实践的指导更是少之又少。近似实

践的方法在研究中得到证实是有效的方式^[48],需要为教师提供真实的或接近真实的评价环境,为教师提供评价任务,在完成真实的任务过程中为其提供有针对性的指导。

第三,关注对教师评价观念的引领。评价观念是评价素养的重要组成,指导教师采用什么样的评价、筛选和过滤建构什么样的评价知识和理念。观念是评价实践的重要影响因素,它对评价实践可能起到促进或阻碍的作用,观念与实践之间可能一致,也有可能不一致。因此在对教师评价素养进行发展时,不仅需要关注其评价知识与技能水平的提高,更需要注重对教师评价观念的引领。评价改革的有效实施和推进,除了外部因素的影响,关键在于教师在观念上是否接受和认可、是否愿意去实施。建议未来的评价专业发展中,了解和关注教师发展的观念因素。

(三) 开发适合我国本土化的教师评价素养测量工具

当前国际上主要的测量工具是基于 90 标准开发的,有些落后,我国也未脱离于 90 标准的束缚。需要在已有研究基础上,基于我国文化背景,结合新时代评价理念的要求,对教师评价素养进行有效测量。以往更加关注的是教师的知识与技能以及总结性评价能力,而当下需要从更多的要素去研究,关注促进学生学习的

评价。在测量方法上,采用自我报告、情境题、量规等不同方法相结合的方式,多角度获取教师评价素养信息。在测量对象上,现有研究主要考虑教师本人,而忽略了学生的感知,学生作为评价的“消费者”,结合学生的视角有助于做出更加客观的判断。在题目设计上,不仅考虑总结性的评价,也需要考虑到形成性的评价,有利于关注教师所拥有的评价目的观是单一的还是综合的。在评价素养维度上,不仅就教师评价素养进行综合研究,也可以就某个维度做有针对性性的深入研究,或选取某几个维度做关系研究。另外,可以聚焦于评价素养和背景

以及学业成就关系的研究,这方面的研究在国际上还非常的稀少。还可以就某一种评价素养做出研究,例如表现性评价素养、反馈素养、自我评价素养等等。在研究对象上,当前大多基于中学教师进行研究,也可以拓展到小学教师评价素养的研究,对小学而言,一般是低利害性的评价,探讨教师是如何发挥评价促进学生学习的有着非常积极的研究意义。

参考文献:

- [1] (新西兰) 约翰·哈蒂. 可见的学习——最大程度地促进学习 (教师版) [M]. 金莺莲, 等译. 北京: 教育科学出版社, 2015: 274.
- [2] Stiggins R J. Relevant Classroom Assessment Training for Teachers [J]. Educational Measurement Issues and Practice, 2005, 10 (1): 7-12.
- [3] Popham W J. Why Assessment Illiteracy Is Professional Suicide [J]. Educational Leadership, 2004, 62 (1): 82-83.
- [4] Schafer W D, Lissitz R W. Measurement Training for School Personnel Recommendations and Reality [J]. Journal of Teacher Education, 1987, 38 (3): 57-63.
- [5] 李雁冰. 走向新的课程评价观 [J]. 全球教育展望, 2001 (1).
- [6] [15] AFT, NCME, NEA. Standards for teacher competence in educational assessment of students. Educational Measurement Issues and Practice [J]. 1990, 9 (4), 30-32.
- [7] Stiggins R J. Assessment literacy [J]. Phi Delta Kappan, 1991, 72 (7): 534-39.
- [8] Stiggins R J. Assessment literacy for the 21st century [J]. Phi Delta Kappan, 1995, 77 (3): 23.
- [9] Shepard L A. The role of assessment in a learning culture. Educational Researcher [J]. 2000: 29 (7): 4-14.
- [10] Popham J W. Needed: A dose of assessment literacy [J]. Educational leadership, 2006, 63 (6): 84-85.
- [11] Paterno J. Measuring success: A glossary of assessment terms [EB/OL]. [2020-10-22]. <http://www.angelfire.com/la2/buildingcathedrals/measuringassessment.html>.
- [12] Popham, James W. Assessment Literacy Overlooked: A Teacher Educator's Confession [J]. The Teacher Educator, 2011, 46 (4): 265-273.
- [13] Webb N L. Assessment literacy in a standards-based urban education setting [C]. New Orleans, Louisiana: American Education Research Association Annual Meeting, 2002.
- [14] [19] Chappuis J, Stiggins R J. Student-Involved assessment for learning [M]. (7th edition). New York: Pearson Education Inc, 2017: 2, 11.
- [16] Stiggins R J. Evaluating Classroom Assessment Training in Teacher Education Programs [J]. Educational Measurement: Issues and Practice, 1999, 18 (1): 23-27.
- [17] Brookhart S M. Educational Assessment Knowledge and Skills for Teachers [J]. Educational Measurement Issues & Practice, 2011, 30 (1).
- [18] Stiggins, Richard J. Classroom Assessment for Student Learning: Doing It Right—Using It Well. Upper Saddle River, N. J.: Pearson Education, 2006. Print.
- [20] DeLuca C, Klinger D A. Assessment literacy development: Identifying gaps in teacher candidates' learning [J]. Assessment in Education: Principles, Policy & Practice, 2010, 17 (4), 419-438.
- [21] [31] Pastore S, Andrade H. Teacher assessment literacy: A three-dimensional model [J]. Teaching and Teacher Education, 2019, 128-138.
- [22] [25] 赵雪晶. 我国中学教师教学评价素养研究 [D]. 上海: 华东师范大学.
- [23] 孟丹. 中学教师的学生学业评价素养研究 [D]. 沈阳: 沈阳师范大学, 2018.
- [24] [42] 郑东辉. 教师评价素养发展研究 [D]. 上海: 华东师范大学, 2009.
- [26] [29] Xu Y, Brown G T L. Teacher assessment literacy in practice: A reconceptualization [J]. Teaching and Teacher Education, 2016, 58: 149-162.
- [27] Willis J, Adie L, Klenowski V. Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform [J]. The Australian Educational Researcher, 2013, 40 (2): 241-256.
- [28] Alonzo D A. Development and Application of a Teacher Assessment for Learning (AFL) Literacy Tool [D]. Sydney: UNSW.
- [30] [40] Looney A, Cumming J, Fabienne V D K,

et al. Reconceptualising the role of teachers as assessors: teacher assessment identity [J]. *Assessment in Education Principles Policy and Practice*, 2017: 1-26.

[32] Fulmer G W, Lee I C H, Tan K H K. Multi-level model of contextual factors and teachers' assessment practices: an integrative review of research [J]. *Assessment in Education Principles Policy & Practice*, 2015, 22 (4): 475-494.

[33] Plake B S, Impara J C, Fager J J. Assessment competencies of teachers: A national survey [J]. *Educational Measurement: Issues and Practice*, 1993, 12 (4): 10-12.

[34] Mertler C A. Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference? [J]. *Competence*, 2003: 29.

[35] Campbell C, Murphy J A, Holt J K. Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers [C] //annual meeting of the Mid-Western Educational Research Association, Columbus, OH. 2002.

[36] Mertler C A, Campbell C. Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory [C]. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada, 2005: 996-1007.

[37] Deluca C, Lapointe - Mcewan D, Luhanga U. Approaches to Classroom Assessment Inventory: A New Instrument to Support Teacher Assessment Literacy [J]. *Educational Assessment*, 2016, 21 (4): 248-266.

[38] Edwards, Frances. A Rubric to Track the Development of Secondary Pre-service and Novice Teachers' Summative Assessment Literacy [J]. *Assessment in Education: Principles, Policy & Practice*, 2017, 24 (2): 205-27.

[39] Gotch C M, French B F. A Systematic Review of Assessment Literacy Measures [J]. *Educational Measurement Issues & Practice*, 2014, 33 (2): 14-18.

[41] Deluca C, Lapointe - Mcewan D, Luhanga U. Teacher assessment literacy: a review of international standards and measures [J]. *Educational Assessment Evaluation and Accountability*, 2016, 28 (3): 251-272.

[43] Deluca C, Valiquette A, Coombs A, et al. Teachers' approaches to classroom assessment: a large-scale survey [J]. *Assessment in Education: Principles, Policy & Practice*, 2016: 1-21.

[44] Snyder, Mark R. Investigating the Link Between Current Classroom Teachers' Conceptions, Literacy, and Practices of Assessment [D]. Philadelphia: the Temple University, 2007.

[45] Brown, Gavin T L. Teachers' conceptions of assessment: implications for policy and professional development [J]. *Assessment in Education Principles Policy and Practice*, 2004, 11 (3): 301-318.

[46] James M, Pedder D. Beyond method: assessment and learning practices and values [J]. *Curriculum Journal*, 2006, 17 (2): 109-138.

[47] Deneen C C, Fulmer G W, Brown G T L, et al. Value, practice and proficiency: Teachers' complex relationship with assessment for learning [J]. *Teaching and Teacher Education*, 2019, 80: 39-47.

[48] Ayalon, Michal, Wilkie, Karina J. Developing Assessment Literacy through Approximations of Practice: Exploring Secondary Mathematics Pre - service Teachers Developing Criteria for a Rich Quadratics Task [J]. *Teaching and Teacher Education*, 2020, 89: 1-14.

(责任编辑: 林再峰)