

# 高质量教育评估中若干关键问题的思考\*

冯 晖

(上海市教育评估院, 上海 200031)

**摘要:** 教育评估是教育体系的重要组成部分,也是教育现代化的重要治理手段。高质量教育评估是高质量教育体系建设内在要求下教育评估的迭代升级。结合当前教育评估存在的问题与不足,就教育评估的评估主体、客体、标准和方法等基本要素,探讨高质量教育评估的核心内涵意蕴,包括完善多元主体参与的评估机制、实行分类评估、提高评估标准的科学性、实施定性与定量相结合的综合评估,以提高教育评估的科学性、专业性和客观性。

**关键词:** 教育评估; 高质量; 高质量教育评估; 专业化

中图分类号: G40-058.1 文献标识码: A 文章编号: 2095-3380(2024)01-0001-06

## Consideration of Some Key Issues of High-quality Education Evaluation

FENG Hui

(Shanghai Education Evaluation Institute, Shanghai 200031)

**Abstract:** Education evaluation is an important part of the education system and an important means of governance for education modernization. High-quality education evaluation is an iterative upgrade of education evaluation under the inherent requirements of building a high-quality education system. Combined with the shortcomings existing in current education evaluation, the core connotation requirements of high-quality education evaluation are explored with regard to the basic elements of education evaluation, such as evaluation subjects, objects, standards and methods, etc., including perfecting the evaluation mechanism with multi-subject participation, carrying out classified evaluation, improving the scientificity of evaluation standards, and implementing comprehensive evaluation combining qualitative and quantitative methods, so as to enhance the scientificity, professionalism and objectivity of education evaluation.

**Keywords:** Education evaluation; High-quality; High-quality education evaluation; Professionalization

习近平总书记在2023年5月29日中央政治局就建设教育强国进行第五次集体学习时强调:“教育评价事关教育发展方向,事关教育强国成败。”将教育评价的地位作用提高到前所未有的高度。教

育评估与教育评价同根同源,都是教育体系的重要组成部分,也是教育现代化的重要治理手段与动力引领,在实践中两者既有交叉也各有侧重。<sup>[1]</sup> 高质量教育评估是高质量教育体系建设内在要求下教

\*基金项目: 国家社科基金教育学重点项目“普及化阶段高校分类评价指标体系构建研究”(A1A220016); 上海市教育科学规划研究项目“上海高校本科专业结构优化与调整机制研究”(A2022012)。

收稿日期: 2023-12-19, 修回日期: 2024-01-05

育评估的迭代升级,在评估理念、原则、机制、标准、方法和工具等方面具有更深层次的内涵。<sup>[2]</sup>一般来说,教育评估是评估主体根据一定目的、依照一定标准、采用有效方法对评估客体给出价值判断并促进客体改进提高的过程,由此,教育评估包含评估目的、评估主体、评估客体、评估标准、评估方法及相应的评估结果运用等要素。从实践层面看,教育评估的核心是评估主体采用什么样的标准和方法对客体进行科学、客观的价值判断。因此,评估主体、评估客体、评估标准和评估方法是教育评估体系中最基本的四大内容,具体回答谁来评估、评估谁、评判准则与尺度、怎么评估等问题。在加强治理体系与治理能力现代化要求下构建高质量教育评估体系,需要进一步厘清评估主体、评估客体、评估标准和评估方法等方面的内涵要求,进一步提高教育评估的科学性、专业性和客观性,有效发挥教育评估的导向、鉴定、诊断、调控和改进作用。

### 一、评估主体:完善多元主体参与的评估机制

“谁来评估”是教育评估体系中的主体问题,明确评估主体是开展教育评估活动的首要问题。评估主体是指组织开展或实际参与教育评估活动的组织或个人,是评估活动的发起者或评估任务的执行者。按评估主体分,教育评估可分为政府评估、社会评估和自我评估等类型,根据政府职能转变的要求,政府评估通常委托专业评估机构开展实施。当前教育评估的主要问题之一是评估主体过于单一,主要集中在教育管理者(第二方)组织开展的评估,而且评估主体通常在评估标准、评估方法和评估过程中拥有绝对的主导权,参与评估的人员也以教育系统内的管理人员、研究人员和同行专家为主。社会组织(第三方)自主独立开展的教育评估以各类排行榜为主,以分数、等第等形式对评估客体进行座次排列,未能有效发挥教育评估的诊断、改进功能。学校(第一方)等评估客体通常只在迎接重要的政府评估时开展自我评估,以对标评估标准进行查漏补缺,还未能制度化地将自我评估作为改进办学实践、自我完善提高的内生动力。

习近平总书记在2023年中央政治局第五次集体学习会的讲话中指出:“要紧扣建设教育强国目

标,深化新时代教育评价改革,构建多元主体参与、符合我国实际、具有世界水平的教育评价体系。”将“多元主体参与”列为新时代教育评价体系的首要目标特征,体现了多元主体的有效协作配合是推进国家治理现代化的有效路径。价值判断是教育评估的本质,而价值是客体属性与主体需要之间的关系,客体属性是形成价值的基础,主体需要是度量价值的基准。同时,价值是多元的,不同主体的视角立场、利益观、价值倾向不同,对评估客体的认识也是“横看成岭侧成峰,远近高低各不同”,自然会产生不同的价值判断结果。只有多样化的评估主体对评估客体进行全方位、立体化的考量,评估结果才会全面、真实、客观。

完善多元主体参与的评估机制,多样化的评估主体除了需要直接参与价值评判,更需要有效参与评估方案的设计,包括制定评估标准、设计评估方法等环节。例如教师和学生是学校的管理对象和教育对象,也是学校教育教学实践活动的直接利益相关者。从内部评估看,学校需要对教师的教育教学情况进行评估,教师对学生的学习情况进行评估,学生也可以通过评教等方式对教师进行评估。从外部评估看,教育主管部门在组织开展学校办学实践评估、教育教学质量评估时,教师和学生有亲身体会,具有重要的发言权,因此他们应该也是评估学校的重要主体。另外,学校培养的学生进入更高学段(层次)学校继续学习或者走向社会时,是否适应高学段(层次)学习要求或者用人单位的岗位要求,高学段(层次)学校的教师、行业企业、社会用人单位以及学生家长都有直接感受,他们也应成为评估学校的主体。因此,政府部门、学校、教师、学生、行业企业、用人单位等都应该在评估方案设计和评估过程中充分发挥他们的协同治理作用。同时,由于利益相关者并非都是专业人员,他们的视角与观点不可避免地存在非专业因素,如何达成专业性与大众性的统一、专业正义与权益公平的平衡,是有效协调、整合多元主体参与的基础性难题。

完善多元主体参与的评估机制,还需要将外部评估与自我评估有机结合。高质量教育评估以评估客体为中心,突出评估客体的主体地位和作用,不但在制定评估方案时需要充分听取评估客体的

意见建议,尊重评估客体的价值观与诉求,更需要在评估过程中高度重视评估客体的自我评估,使评估主体与客体进行有效沟通,从而全面深入了解评估客体的个性特色和真实情况,指出其优势与不足,并共同探讨分析其优态劣势的成因以及发展趋势,充分调动评估客体的主观能动性,帮助促进评估客体反思、整改、提高,有效实现“对客体评估”向“为客体评估”的转变,激发评估客体从“要我评”转向“我要评”,有效发挥教育评估的效能和作用,促使每个评估客体都有出彩机会并从中获益。

## 二、评估客体:实行分类评估

“评估谁”是教育评估体系中的客体问题,明确评估客体也是评估工作的根本问题。评估客体即指被评估的对象,是根据评估目的进行考量的教育事件或活动,是评估任务的承受者。教育评估客体的规模、跨度范围很大,宏观上可以是整个国家的教育成就、教育发展水平等,中观上可以是一所学校的综合实力、办学绩效等,微观上可以是一名学生的学习成绩、一位教师的教学科研水平等。评估客体是由人员集合、事物集合和时限范围共同组成的三元结构,其中人员集合指评估客体的人员范围;事物集合指评估客体的事物范围,例如评估高等学校的办学水平,其包括学校的办学理念与定位、落实党的全面领导、人才培养目标、立德树人成效、服务国家战略和地方经济社会发展的贡献、办学声誉与特色等各个方面,也可单项评估其中某一方面的工作;时限范围指评估客体在时间上的规定性。评估客体的多样化决定了其人员范围、事物范围和时限范围具有很大的不平衡性。

评估标准单一也是当前教育评估存在的主要问题,通常对评估客体采用“大而统”的指标体系,“一把尺子量天下”,过于强调评估客体的共同特性,忽视评估客体的个性特点。教育对象是一个个活生生的人,尤其在教育普及化阶段,需要多样化的学校办学实践,只有尊重学生个体天性特征的多样性与个体成长发展过程的多样性,才能更好地满足社会对人的多元化需求。因此,理想的教育评估应该是“一人一标准,一事一标准”,但在评估实践中,鉴于评估成本的客观限制以及评估客体相互比

较的需求,为了提高评估指标体系的精细化以及对评估客体的适切性,需要有一个“折中”方案,即实行分类评估,对不同类别的评估客体采用差异化的评估指标体系,并且“分赛道”实施评估。

分类评估的概念源自于高等学校本科教学工作水平评估。针对教育部第一轮本科教学工作水平评估中指标体系单一、对所有高校采用同一个评估方案等问题,我国学者和教育行政部门提出了分类指导、分类评估的思想。<sup>[3]</sup>随着社会各界对教育评估科学性要求的持续提高和教育评估理论的发展,分类评估的外延已经从本科教学评估拓展至学科、专业、教师、学位论文等各类评估客体,即对不同类型的评估客体采用不同的评估标准、不同的评估方案,以提高教育评估的针对性和有效性。2020年10月,中共中央、国务院印发的《深化新时代教育评价改革总体方案》(以下简称《总体方案》)提出要“推进高校分类评价,引导不同类型高校科学定位,办出特色和水平”,“改进高校教师科研评价。根据不同学科、不同岗位特点,坚持分类评价”。可见,分类评估已是评估方法论的范畴,对各种评估客体具有普适性,旨在更好地满足评估客体多样化的需求,提高教育评估的科学性、客观性与有效性。

实施分类评估的前提是对评估客体进行合理分类,即根据评估客体的特征与属性,将其划分为若干个子集。<sup>[4]</sup>评估客体的属性往往是多元的,多元属性构成了一个多维空间,每个评估客体对应于多维空间中的一个点。按照评估客体的某个属性划分类别,实质上是在 $N$ 维空间的某个维度上进行分割,形成若干个 $N-1$ 维空间。因此,评估客体属性的多样性势必导致分类方式和分类结果的多样性,而且同属一类的评估客体仍不同质,即根据某个属性进行分类时,每一类中的评估客体的其他属性也不尽相同,也就是说,分类并非将相同的客体、而只是将相近的客体聚拢在一起。所以,对评估客体进行分类时,需要根据评估目的和导向合理选择分类标准,如果分类后的评估客体仍然差异较大,必要时还需进行二次、三次分类,例如根据工作侧重将高校教师分为教学为主型、科研为主型和教学科研型,再根据科研内容分为基础研究为主型、应用研究和科技创新为主型、科技成果转化为主型

等,以提高评估标准的针对性和适切性。

实施分类评估的关键是构建共性与个性相协调的评估指标体系。共性是指评估客体的共同特征,体现了教育的内在规律。例如评估高校整体办学能力时,由于立德树人是中国特色社会主义高校办学治校的本质要求与价值诉求,高校肩负着人才培养、科学研究、社会服务、文化传承创新和国际交流合作的重要职能,因此评估指标体系为引导高校推进立德树人根本任务落实落细,需要将其融入高校五大基本职能实现的各环节、各方面。而个性则是从多维度、多选项的方式体现多元评估标准,反映不同类别评估客体的差异性。例如同一项指标在不同类别评估客体中的观测点不同,如师资队伍结构指标,在研究型高校中的观测点主要是具有博士学位的教师占比、高被引科学家占比等,在技能型高校中的观测点主要是双师型教师占比、国家级和省部级技术技能大师占比等;同一项指标在不同类别评估客体中的权重不同,如科学研究指标,在研究型高校中的权重要高于技能型高校。

### 三、评估标准:提高评估标准的科学性

评估标准是衡量评估客体优劣、强弱程度的准则与标尺,体现了评估目的,反映了价值导向与价值标准,通常以评估指标体系的形式表达。评估标准是评估工作的核心,也是影响评估科学性的重要环节。评估指标体系描述了评估客体整体的主要特征,关乎信息采集内容与采集方式,并对应价值判断工作。教育事件或活动等评估客体往往是多目标、多层次、多因素的复杂系统,评估指标体系一般是遵循导向性、完备性、可比性、独立性、可测性、简约性等原则,通过对评估目标依次逐级分解而构成的一个完整、多级的结构化体系。

《总体方案》提出要“扭转不科学的教育评价导向”“提高教育评价的科学性、专业性、客观性”,《总体方案》还明确提出,要“遵循教育规律,系统推进教育评价改革”“充分发挥教育评价的指挥棒作用,引导确立科学的育人目标,确保教育正确发展方向”。因此,科学性是教育评估(评价)的根本价值取向与教育评估有效性的基础,缺乏科学性的教育评估不但不利于教育的健康发展,而且还会导致教

育发展的异化和决策产生偏差。“教育评估的科学性是教育评估活动的合规性与合目的性的统一。科学的评估是指教育评估者在一定条件下所做出的合规性与合目的相统一的对客体价值的衡量。”<sup>[5]</sup>构建符合教育规律和教育目的的评估标准,将社会主义办学方向、对教育规律的认识和评估客体的特征属性转化为比较科学、客观的评估指标体系,是提升教育评估科学性、有效性的重要基础。

评估标准单一化是当前教育评估的主要弊端,一方面是指对不同的评估客体采用统一的评估标准,缺乏针对性和适切性;另一方面是指仅以唯一的指标来表征评估客体某方面的属性,导致“五唯”顽疾。实施分类评估、构建多元标准是解决评估标准单一化的有效方法,能以更精细化、多元化的指标描述评估客体的多样化与个性化特征。“五唯”是评估标准单一化、片面化的典型表现,构建评估指标体系既要全面又要简约,处理这对矛盾的“折中”办法是遵循主要性原则,通过提炼与概括评估客体有限的主要特征属性作为指标,力求比较全面客观地反映评估客体的真实面貌,当极端化地只用一项指标时就成了单一、片面的评估,更是不科学的评估。诚然,“唯”的内容(分数、文凭、论文、帽子、职称、奖项等)确实是评估客体的内涵属性,破“五唯”不等于去“五唯”,“不唯”不等于“不要”,破除“五唯”在于击破“唯”,构建多维度、多选项的评估指标,鼓励评估客体自设指标,避免简单化、极端化、绝对化,从而实现主导价值与多元价值的有机统一,使每个评估客体都有出彩机会。

不当具体化、虚假精确性、以数量代替质量也是评估标准中的常见问题。例如评估高校学科贡献度时,采用发表论文数、出版著作数、科研获奖数、高水平论文数等若干量化指标。又如某地在评估中小学生思想品德时,设置爱党爱国、遵纪守法、文明礼貌、乐于助人、诚实守信等若干指标,每项指标采用具体事件的积分方式,如旷课一节扣2分,迟到、早退一次扣1分,帮老师搬东西、扶老人过马路等做好事每次加1分。但是评估客体的许多内在属性是隐性的、抽象的、非结构化的,更多的是一种主观认知与感受,将其简单化、片面化地转化为显性的、具体的、结构化的指标体系,而且以客观具象的

方式进行表征,以“失真”的指标体系开展评估,必然导致表面上的合理性掩盖实际上的不合理性。

制定科学的评估标准是世界性、历史性、实践性难题,评估标准不仅要真实客观地刻画评估客体的特征,更要契合教育规律与教育目的,但教育的事理往往充满了矛盾与对立,例如升学率体现了教育教学质量,过度追求升学率又让教育过程丧失了人性;既要注重对教师的考核,又要激发教师的活力;既要促进学生的个性发展,又要加强规范管理;等等。教育评估标准正是在这种矛盾对立中把握方向、寻找平衡、引导发展。同时,制定评估标准还需要处理好当前与长远的关系、规模与质量的关系、总量与均量的关系、周期性评估中评估时段前的历史存量与评估时段内的发展增量的关系等,使每项指标都与教育目标、评估目的保持一致,既真实客观地反映评估客体的本质,又引导评估客体科学发展。

#### 四、评估方法:定性定量相结合

“怎么评估”是教育评估的方法性问题,评估方法直接影响评估结果的信度与效度,关乎评估目的能否有效达成。教育评估先后经历了“测量—描述—判断—协商”的四代演变,既是评估理念的发展,也是评估方法的变化。教育评估的类型与方法有很多,根据评估的时间点不同,可分为诊断性评估、形成性评估和总结性评估;根据评估主体不同,可分为内部评估(自评)与外部评估(他评);根据评估的基准(参照点)不同,可分为绝对评估、相对评估和个体内差异评估;根据评估功能与内容不同,可分为结果评估、过程评估、增值评估;等等。无论何种类型的评估,就评估过程中有关实证分析和思辨方式而言,教育评估方法必然涉及定量评估或定性评估。

定量评估源于实证主义,是指运用数学、统计学工具,收集处理评估客体资料,通过数量化的分析与计算,对客体价值作出判断。它有助于一些概念的精确化,提高评估的区分度,降低评估的主观性与模糊性,增强评估的说服力,因此定量评估强调结果,适用于选拔、甄别等评估目的。而定性评估源于解释主义,是指通过观察调查、访谈互动、资

料分析等方式,收集分析评估客体信息,以归纳、概括、诠释等方式对客体给出鉴定或写实等文字描述,揭示和解释客体的性质或程度。它突出客体的发展过程和个体独特性,把握客体的特征,形成对客体优缺点及其成因的全面完整认识,因此定性评估强调客体发展过程,适用于诊断、促进等评估目的。定量评估具有确定性、高分辨性和实施简便的特点,目前一些教育评估中,如各类排行榜多采用定量评估的方式,指标体系以有限的简单化、易获取的量化指标为主,缺乏立德树人、人才培养内涵。引导方向不科学的评估理念与强化定量的评估方法,在一定程度上误导了大众对学校、学科、专业等评估客体整体质量和水平的认识,也在一定程度上助推了“唯”某些定量指标的顽疾。

辩证唯物主义认为,任何事物都是由“质”与“量”构成的统一体。质是对事物进行定性描述,量是对事物进行数量表达。美国学者格朗兰德曾给出著名的评价公式,即“评价=测量(量的记述)或非测量(质性记述)+价值判断”。<sup>[6]</sup>在评估实践中,有的客体属性可以量化,有的则不能量化或不宜量化,如教育理念、办学方向、品德修养、学生情感等,而且数量仅是质量的表现形式之一。定量评估强调数据背后的“事实客观性”,突出客体的统一性与可测性等规律性特点,但缺失人文要素,而且数量指标选取、计量模型等都对结果敏感,导致评估结果的不确定性。定性评估主观性较强,评价者的个人背景以及与客体之间的关系可能会对结果产生影响,而且要求评价者具有专业知识和经验。定量评估与定性评估并不是简单的对立关系,二者具有较强的互补性和相辅相成之处。

高质量的教育评估将定性评估与定量评估有机结合,用来全面分析客体的本质,切实提高评估的科学性、客观性与公信力,并且可以实现有效反馈,帮助评估客体“知其所以然”,促进客体整改提高。2014年《莱顿宣言》提出合理利用科学评价指标的十条原则,其中位列首位的是“量化评价应支持而非取代定性评价”。2018年中共中央办公厅、国务院办公厅印发《关于分类推进人才评价机制改革的指导意见》,指出要“坚持共通性与特殊性、水平业绩与发展潜力、定性定量评价相结合”,2020

年教育部、科技部印发《关于规范高等学校SCI论文相关指标使用,树立正确评价导向的若干意见》,指出要“采用定性与定量相结合的综合评价方式,突出创新质量和实际贡献”,表明在人才评价、学术评价、教育评估等工作中,相关量化指标不是评估评价学术水平与创新贡献的直接判据,更不是唯一依据,要审慎选用量化指标,采用定性与定量相结合的方式,引导评估评价工作突出科学精神、创新质量、服务贡献。

实施定性与定量相结合的教育评估,通常可以采用组合与融合等方式。所谓组合方式是指以串行或并行模式组合使用定性描述与定量判断。在指标体系上,可以由客观指标与主观指标共同构成,能以可靠性高、共识度高、导向性好的数量直接表征客体属性的可以用客观指标,难以用数量直接表征的则用主观指标;在评判方式上,对客观指标采用计量评价,对主观指标采用同行评议,或者无论是对客体的总体评判还是对各项指标的评判,既给出等第、分数等形式的量化结果,同时也给出写实性叙述,用定性的描述来补充说明与解释定量结果的含义;在评估主体上,多位评判者分别使用定性或定量的方式分析评判同一客体或同一指标,相互印证和检验评判结果的客观性。

融合方式是整合定性分析与定量分析的一种综合评判范式,充分利用定性评估与定量评估的优势,克服两者的不足与局限性,给出综合评判,如层次分析法(AHP)、模糊综合评价法(FCE)等体现了定性定量相结合的思辨分析方式。定量评估依赖数据或者给出数值结果,但是数据的背后缺乏深入解释和理解;而定性评估可以提供深入理解和诠释,但又难以统计分析。融合评估的一种模式是对同一项指标分别采用定量与定性的方式独立评判,兼顾数量与质量,再合成汇总其结果,例如学科评估中的发表论文指标,既对高水平论文数、高被引论文数等数量指标进行计量评价,也对代表性论文的质量进行同行评议。融合评估的另一种模式则是在评判过程中融合定量与定性的交替思维,定性分析的学术性、专业性与综合性为定量分析提供视角和基调,充分应用定量分析中的数据为定性分析提供支撑、依据与辅助,例如本科专业评估中,参考

专任教师数量与结构、课程体系结构与学分学时分配等数据,结合座谈访谈、看课听课、资料查阅等情况,吸纳定性评估与定量评估中的一方优势弥补另一方劣势,克服两者独立存在的不足,重构评判过程,融合同行专家的智慧,构成综合认知,从而得到更加全面深入的评估结果。

#### 参考文献

- [1] 冯晖. 教育评价与教育评估辨析[J]. 上海教育评估研究, 2022, 11(5):26-30.  
FENG Hui. Discrimination between Educational Evaluation and Educational Assessment [J]. Shanghai Journal of Educational Evaluation, 2022, 11(5):26-30.
- [2] 冯晖. 高质量教育评估的内涵要义与推进策略[J]. 上海教育评估研究, 2023, 12(2):1-7.  
FENG Hui. Connotation Elements and Promoting Strategy of High-quality Education Evaluation [J]. Shanghai Journal of Educational Evaluation, 2023, 12(2):1-7.
- [3] 田锋. 关于“分类指导、分类评估”的研究综述[J]. 高教发展与评估, 2010, 26(3):32-37.  
TIAN Feng. Review of the Studies on “Guidance by Classification and Assessment by Classification” [J]. Higher Education Development and Evaluation, 2010, 26(3):32-37.
- [4] 冯晖, 王奇. 高等教育分类评估的若干问题[J]. 高教发展与评估, 2013, 29(1):34-40.  
FENG Hui, WANG Qi. Some Problems in the Classified Evaluation in Higher Education [J]. Higher Education Development and Evaluation, 2013, 29(1):34-40.
- [5] 杨晓江. 教育评估的科学性与科学的教育评估[J]. 教育研究, 2000(8):33-36.  
YANG Xiaojiang. The Scientificity of Education Evaluation and Scientific Education Evaluation [J]. Educational Research, 2000(8):33-36.
- [6] 张建新, 刘康宁. 质性研究方法在教育评估中的应用探索[J]. 上海教育评估研究, 2014, 3(2):6-10.  
ZHANG Jianxin, LIU Kangning. Exploration of Qualitative Research Methods in Education Evaluation [J]. Shanghai Journal of Educational Evaluation, 2014, 3(2):6-10.