

论院校教学评价中“专家进校”的评价失真风险及其规避

张清, 种艳艳

(武警工程大学 军事基础教育学院, 陕西 西安 710086)

摘要:“专家进校”是上级对院校进行教学评价的一个重要环节。对院校的教学评价结果是由进校专家集体讨论得出的。由这种方式得出的评价结果能否完全与院校教学事实和价值相符合,直接影响着对院校的最终评价结果。“专家进校”后展开的评价活动是对院校教学建设与发展的认知和价值判断过程,在这个过程中有产生认知和判断失真的风险。该风险源自专家团队客观存在的有限理性、集体认知的迷思和“三性一度”的认知难度等诸方面。有效规避专家评价失真,提高评价可信度,将更有利于提升院校教学评价的科学性。

关键词: 教学评价; 专家进校; 失真; 风险

中图分类号: G647 **文献标志码:** A **文章编号:** 1672-8874(2022)03-0069-05

Risks of Distorted Evaluation by Outside Experts in Teaching Evaluation in Colleges and Universities and Avoidance of Such Risks

ZHANG Qing, CHONG Yan-yan

(College of Military Basic Education, Engineering University of PAP, Xi'an 710086, China)

Abstract: Evaluation of college teaching work done by outside experts is an important part of teaching evaluation. The evaluation results of college teaching work are made by the collective discussion of those experts, and whether the evaluation results made in this way can be fully consistent with the teaching facts and values of colleges and universities directly affects the final evaluation results. Their evaluation activities involve the cognition and value judgment of teaching construction and development with a risk of cognitive and judgment distortion because of the limited rationality of the expert team, the myth of collective cognition and the difficulty of cognition of “San Xing Yi Do”. Effective avoidance of the risk of evaluation distortion and improvement of the credibility of evaluation will enhance teaching evaluation in colleges and universities.

Key words: teaching evaluation; expert entering the university; distortion; risk

对教学评价工作所包含的“专家进校”及随后的整个活动过程及其效果的反思属于教学评价的元评价研究范畴。“元评价”概念最早由米切尔·斯凯文(Michael Scriven)首次提出,认为元评价是对评价活动、评价系统或评价机制的评价^[1]。

著名评价专家丹尼尔·斯塔佛尔比姆(Daniel Stufflebeam)将“元评价”定义为“获取和运用那些评价的实用性、可行性、合理性和准确性,以及其系统本质、行为能力、诚信度、受尊重程度等描述性和评价性信息,以引导评价并向公众报

告该评价的缺陷和价值的过程”^[2]。国内有学者认为,元评价是实现评价质量和效益的评判,包括评价方案的科学性、评价活动组织实施的正确性和规范性、评价结果的合理性等;元评价的对象是原来的评价,包括评价结构、过程和结果,除了评价客体之外,还涉及参与评价的实施机构、工作人员和专家等^[3]。很显然,这里将评价活动中的工作人员和专家作为元评价的重要对象,与本文的研究对象一致。但遗憾的是,现有的研究很少关注专家对评价结果的影响。有鉴于这种研究的阙如,结合院校教学评价的实际,我们将专家参与到教学评价活动中的一个重要行动——“专家进校”对受评院校教学进行价值判断时可能产生的失真或偏差问题作为研究论题进行探讨,重点探究造成失真的原因及规避办法,以提高评价结论的信度和效度,助力于院校教学评价活动更加趋于科学性,从而更好地为院校发展服务。

一、“专家进校”在教学评价活动中的地位与作用

对院校教学的评价最终会得出相应的评价报告。在这个报告中专家们会努力和更为客观地把一所院校在一定时期内的整体教学工作情况呈现出来,除了反映院校所有的教学成效事实外,还要在基于这些事实的基础上进行相应的价值判断,主要说明院校在教学工作上有哪些优异的表现、哪些是可堪称道的、哪些是要改善的甚至哪些是要遭受批评的等等。得出这样全面的结论是组织教学评价的上级所期望的,也是院校为改进自身教学工作所期待的。上级要依据这个评价报告来了解院校教学工作的真实状态,并在此基础上实施相应的教育管理决策,包括指导院校宏观上的发展战略,中观上培养任务的赋予与调整,更细微处甚至延伸到对院校的课堂教学情况的管理与干预等。从这个意义上说,教学评价才能够实现“以评促建”“以评促管”的目的。

然而,担负如此重任的评价报告却从来都是需要人来负责完成的,是具备相应专业能力的个人或者团队进入现场后通过个人的感觉、认知、判断和理性思考及其与院校人员沟通的基础上,经过集体讨论而得出的结论。这些个人或团队就是筹划教学评价的上级单位极其依重的评价专家和专家团队。通过他们力求客观的工作,上级才

能最终得出关于院校的客观评价结论。专家团队要想得出负责任的评价报告,就必须组队进入相应的院校展开高效的认知判断活动。这一活动在教学评价全过程中居于如此重要的地位,以至于教师们充满敬畏与尊崇地将其称为“专家进校”。这一称谓既表达出了对专家所拥有的专业权威的敬意,更表达了受评院校的教职人员对专家团队的热切期待态度。

专家凭借其专业能力与出众的认知能力进入受评院校后,以评价指标和观察点等作为标准,对院校在一定时期的教学工作(教学状态水平、建设发展情况等)进行综合考察与全面检视。其主要任务是掌握院校教学活动的基本事实和观念等,并靠多种措施(如观察、访谈、座谈、问卷、参与活动、查阅、测试等)获得全方位信息,尔后在对信息分析的基础上,专家个人初步得出对教学事实的判断,再经由专家之间的沟通交流达成初识,并形成局部或整体的意见报告,再通过专家团队集体讨论形成最终的评价报告。这个报告主要是判断受评院校“三性一度”的实现程度,并依此为依据,对照评价手册中相应条件给出明确的评价等级。这个等级常常是院校上级、院校和专家团队高度关注的结果,它关乎到各方不同的使用态度,与院校利益攸关。

二、“专家进校”的认知本质与失真风险

如前所述,“专家进校”所进行的评价过程,是事实认知与价值判断过程。这个过程是专家团队争取自己得出的报告与受评院校实际情况无限接近的过程。由于专家团队要在很短的驻校评价活动中,给出相应的认知结果,这无疑对专家和团队提出了很大挑战。因为人们对事物的认识过程,不仅与自己接触事物的方式、时空、环境、手段等密切相关,而且还会与认识主体自身的状态密切相关。这意味着人们认知事物常会受到主、客观条件的影响。因此,“专家进校”的认知过程也存在认知失真的可能性,即认知结果与院校自身的教学事实、观念事实存在或多或少的不符情形。这也意味着“专家进校”存在认知与判断上的失真风险。

(一) 专家团队有限理性的规约

从不同的视角与学科解读人的理性时,对理

性的理解就呈现出非常多样的形态。从本体论上来说,理性就是人区别于动物的本质规定性,这是人之所成为人的根本依据。从认识论上来说理性是指人们在认知事物的基础上,通过自身能动性进行分析判断而得出正确结论的能力。著名的管理学家赫伯特·西蒙(Herbert Simon)在其《管理行为》一书中提出了人的有限理性原则,意指人们在认知与判断中不是全知全能的,是会出现片面性甚至是错误的可能,这一点得到了人们的广泛认同^[4]。马克思主义哲学认识论强调,人们认识世界的过程是由相对真理向绝对真理的无限接近过程,其间充斥着谬误与错讹^[5]。如果将这些原理用于考察进校专家及其所展开的评价与认知活动,可以说,专家们对受评院校的教育教学事实和观念事实的了解与认知并不都是完全正确的,只能说是向着真理性认识方向的无限接近过程。如果专家的认知结果有时候逼近了真相,那么就给受评院校可以提出很中肯的考评报告及相应的诊断结果,这使得受评院校会从中分外受益;而当专家的认知结果貌似接近真相实为相差甚远时,他们的评价报告及诊断结果可能存在不当之处。

(二) 专家团队集体认知的迷思

专家们为了减少上述基于有限理性而造成的失真风险时,在评价实践中常常会组织专家组研讨,用这种集体认知的方式来减少个人理性的局限性,将个人认知失误降低到最低的限度。俗语“三个臭皮匠顶个诸葛亮”就非常形象地说明了集体思考比个人思考存在显著优势,这也就是我们常说的集思广益的效应。基于此原理,上级单位会邀请数十余名专家组成评价团队担负此项重任,而且会采取灵活的人员分工与组合,目的都是为了发挥团队集体认识与思考的优势,从而达到对一线情况的全面掌握。然而,对这类临时组建起来的专家团队的期望真的会与结果达到一致吗?显然未必。因为这种临时专家组本身的特点对认知结果会产生影响。有学者就对此表示了担忧。“因为专家组具有临时性,导致他们没有一个稳定的团体来从事某类高等教育评估的专门活动,评估无论在理论上还是实践上都无法形成体系;一些评估者也无法潜心于评估的研究,得出结论真理性成份受质疑。”^[6]不只如此,专家组的认知还受到团队构成的影响,如年龄结构、专业结构、职业结构、从业时间结构和学缘结构等都会成为影响集体认知结果的重要因素。另外,专家组也

会出现过分依赖或沉迷过去经验的现象,陷入管理学家彼得·圣吉(Peter Senge)所谓的“向经验学习的错觉”当中。明知运用经验有其时空限制,但却一味突破有效范围,从而造成认知失误^[7]。有学者还认为:“在集体评价中专家间的从众行为、趋中效应、暗示效应和威望效应等都对评价结果产生消极影响,造成认知误差。”^[8]这些担忧与质疑不无道理。由是观之,用团队评价的方式尽管克服了个人思考与认知的短板,却也常常被其自身固有的不足所困扰。因此,如果对临时专家团队的集体认知与思考能力过度自信,无疑是导致进校评价失真的一个隐匿性因素。

(三) “三性一度”的认知难度

审核性评价持有两个评价标准:一个是共性的评价标准;另一个是院校依据共性评价标准制定的个性化评价标准。后者就是人们通常所说的“用自己的标准来评价院校自己”的依据。专家团队在开展评价工作时,除了用这两个标准作为依据外,还找出一个将标准化繁为简的工具,这个工具就是所谓的“三性一度”,即考察受评院校建设的科学性、先进性和发展性及所承诺的建设目标达成度。如果把院校的个性化评价标准视作一把尺子的话,那么“三性一度”则可以被看作尺子上的刻度,进校专家用它来测量并读取评价结果。实际上,从科学性、先进性、发展性和达成度这四个刻度来读取院校的实际表现是一件非常不容易的事情,对专家提出了极大挑战。

首先,“三性一度”具有主观性,将其作为衡量依据,得出的测量结果及其准确性受人质疑。“一千个人就有一千个哈姆雷特。”不同的人理解科学性、先进性、发展性和达成度本身得出的结论会千差万别。尽管考量这四个方面时专家会持有一些共识,但专家要达到对它们趋同和一致的认识却是几乎不可能的。对这四个方面的个性化理解所带来的差异性常态,而一致性通常则是不易得的。在此种情况下,专家们用它们来测量受评院校的实际表现时就会得出差异的结果。在这个结果基础上,如何采取措施和采取什么样的措施对专家的意见达成一致具有挑战性。

其次,“三性一度”本身认知难度非常大,不易得出客观结论。专家对“三性一度”理解各异,除了与主体的主观认知水平相关外,从认知客体来说,“三性一度”本身存在认识上的难度,也会造成专家们的理解各异。就拿“达成度”来说,

作为一个完全的陌生人——专家库里的专家，突然进入一所素昧平生的受评院校后，却要在很短的时间内判别出这个院校实际建设状况与其提出的个性化评价标准之间的距离是多少，这显然是压力很大的难题。而且每个专家对发现的这个距离即使能够达成共识，那么是用定性的结果来表达还是用定量的结果来表达，才能够让院校心服口服呢？这就再一次拉高了问题的难度。另外，就“先进性”来说，要鉴别一所院校发展的先进性，多数情况下是在有一个参照物的基础上得出的，或者说是与其它同类院校相比较而得出的。如果同类院校数量比较多，那就要求专家对这数量众多的同类院校发展了如指掌，在此基础上方能判断受评院校的先进性与否。显然，要掌握众多同类院校的发展水平需要专家付出更多的才智与精力方可，不然得出的结论则是虚妄和不可信的。由此来看，专家们在判断院校发展的先进性时，就再度陷入困难。

三、“专家进校”评价失真风险的规避

由前述分析可见，进校专家在认知、判断进而形成评价结论等一系列活动过程中，都会存在造成失真的系列因素。但是，我们如果一味只强调认知存在的困难与失真时，则未免会堕入不可知论的泥潭，从而造成对整个教学评价活动意义的质疑，这显然在理论上是极其有害的，在实践上是画地为牢的。我们认为，克服专家认知过程中的困难因素，帮助专家接近或达成认识的真理，这才具有建设性意义。相对于受评院校来说，进校专家常常是陌生的“他者”，对院校从不熟悉到稔熟需要一个过程。要使这个过程达到预期目的，专家团队理应做好如下几方面的工作：

(一) 周全筹划专家的驻校时间

观察过往专家驻校时间，就会发现这个时间一般接近两个星期左右。在这段时间内专家们按照计划繁忙地“扫描”院校的各个角落，他们分工协作在各个方向，采取各种方法来获取相应信息，以便把握院校的发展状况。在这个过程中，专家们面临着十分繁重的任务。因为他们作为受评院校的陌生人，想要短时间内了解院校建设与发展的真实情况，获得相对全面真实的一线信息非常不易。认知心理学认为，对于陌生对象的熟

知程度通常和接触其时间长短之间存在密切关系。彼此相处与接触的时间越长，双方才能达到由陌生、熟悉再到熟稔的程度，认知客体的特征和其它各类信息方能渐次暴露出来，尔后经过主体的辨析与加工，才能获得对方的较为真实的认知结果^[9]。有鉴于此，就要严密筹划专家组驻校时间。既要考虑专家们有足够的认知时间，又要顾及到专家驻校成本和防止干扰院校正常教学。参照教育部《普通高等学校本科教育教学审核评估实施方案》中规定专家组驻校时间为4-5天（这个时间是以完善的线上评估平台作为支撑的），我们认为可适当延长至10日左右为宜，这样可以保证专家们能够与院校的师生交朋友，参与他们的教学活动，体验他们的日常生活，从而将评价这一工作变成为充满日常生活气息的交流活动，有助于专家更为深入地了解院校的教学常态和院校师生的精神世界，对院校的理解与认知就会更加深刻，这样可以很好地解决“三性一度”所存在的认知难度，而且还会避免专家进校后常常被视为“浮光掠影”和“蜻蜓点水”式的刻板印象，也能够较好地规避专家认知失真的风险。

(二) 严格遴选进校专家

显然，作为评价主体的专家团队，自身条件是影响评价效果的另一关键要素。团队中的每个专家的职业操守、专业水准和信息素养等影响着进校后的评价效果。除此之外，还应当重视专家团队的队伍结构搭配，这是影响主体评价教学效果的一个重要因素。

首先，依照教学评价活动的需要，建构合理的专家能力模型。借鉴过往院校教学评价的经验，结合受评院校教学评价的特色及发展性追求，细化专家能力清单，主要将职业操守、专业水平、从业经验、交流沟通和信息利用等作为一级指标，然后再逐项细化，形成进校专家的能力模型雏形，召集有经验的评价专家对能力模型进行研究论证。在此基础上，征询受评院校或教育管理单位的意见建议，并认真修改，最终形成科学合理的进校专家能力模型。筹划评价的上级单位依此模型，在院校、部队和其它教育机构遴选专家，形成专家团队，赴院校从事教学评价活动。

其次，注重专家队伍结构建设，形成搭配合理的进校专家团队。防止评价失真应在全面认知上下功夫，因此就要从多角度审视受评院校，力争给受评院校画出“全图”，这些都客观上需要进

校专家的多样性来实现。为此,要关注专家团队的专业结构。教学评价是一项专业性很强的活动,既需要有精于此项工作的评价专家作为核心成员,也需要有教学专家、教育研究专家、教学管理专家和数据分析的专家作为基本成员,在专业上实现专博结合和强强联手,让专业的人干专业的事。要关注进校专家的来源结构。专家团队应当既有代表教师队伍的专家,也要有院校中高层管理岗位的管理专家,还应当有上级教育行政部门的管理专家,甚至还可以吸纳教育研究机构的专家。构建团队时还应当考虑年龄上的新老构成、从事教学评价中的老专家与新专家构成、现职专家与退休专家的构成等。总之,通过关注专家团队的结构,在数量和质量上提升评价主体素质,使评价认知趋于全面,防止认知片面和失真。

(三) 严谨实施线上线下评价

当前,通过线上评价是评价理念的新发展,也是教学评价在手段上的创新。因为它使教学评价更加灵活,不再受时间和空间的限制。同时,也使得更多专家参与到教学评价中来成为可能。由于时空的灵活性,给专家带来极大的便利,也为筹划单位在挑选参与评价的专家时提供多种可能。

首先,专家进校前细致研读院校提交的数据。研读院校的自评报告、教学状态数据分析报告、毕业生跟踪调查报告和用人单位跟踪调查报告等信息,甚至还应当更进一步细致地判断细节。如,人才培养效果与培养目标的达成度,办学定位、人才培养目标和部队需求的适应度,教师和教学资源条件的保障度,教学和质量保障体系运行的有效度,学员和部队需求侧的满意度等,基于此对受评院校“三性一度”进行初步的判断,在坚持全面系统、问题导向、动态考察和关注个性特点的原则上给院校勾勒出“写意画”。

其次,进校后验证数据。通过数据判断院校建设状况毕竟具有间接性,其间因数据统计、数据撷取和数据解读不周,也会产生基于数据判断的失真风险。所以,专家进校后要认真将先前的数据判断与院校现场对接起来进行验证校对。通过观察、访谈、听课甚至查阅档案等方式,将数据与实际情况相互参照印证,以确保得出真实可靠的判断结论。特别是对其中一些数据变化较大的相关指标,重点要看相应数据指标项上具体实施了哪些积极有效的措施,哪些政策制度没有落

实到位,以此判断该数据是受评院校的亮点,还是数据失真或建设不足造成的。

(四) 严密规划专家团队的工作规程

专家进校后要做好工作筹划,科学安排认知判断的流程,使评价活动开展后能有条不紊地进行,通过高效务实的工作过程保证评价主体判断的真实性。

充分准备,产生预先判断。专家团队进校后展开考察评价的基本方式是分工合作。专家一旦受领任务并明确各自的分工后,在进校评价前逐步做足功课。首先,要学习掌握与考察内容相关的政策、制度、规定及相关指示精神和其它背景知识,做到心中有数。其次,通过受评院校教学大数据与同类院校分析比较,结合院校日常表现,专家对该院校建设水平有个初体验和预判断。再次,通过院校每年上报的教学基本信息和年度自评报告,结合上级机关的专项检查讲评对院校建设状态作出初步判断。

考察核准,形成初步结论。专家团队进校后要多观察、多思考、多记录,多问几个为什么,多深入一线,多与教师学生交流;既要紧盯问题,查清原委,又要顾及其它指标来作相应考察;既要注意从观察分析中发现问题,也要关注之前上级日常检查时指出的问题及其整改情况;既要扫描院校的短板和不足,也要注意发现其优势和特色。

综合分析,形成结论建议。专家在考察、核实、核准基础上进行综合分析,对自己分工的板块建设水平形成基本判断,并向专家团队提出结论等级意见。在此基础上,通过专家团队的充分讨论,各自署名投票表决,产生每个指标的结论等级,并在此基础上整合得出受评院校的最终评价等级。

参考文献:

- [1] SCRIVEN M. Critical Thinking: It Definition and Assessment[M]. Norwich England: Centre for Research in Critical Thinking and Edge press, 1997: 224.
- [2] STUFFLEBEAM D. Meta Evaluation: Concept, Standards, and Uses[C]// Educational Evaluation Methodology: the State of the Art. Baltimore: Johns Hopkins University press, 2015: 125.
- [3] 侯光文. 教育评价概论[M]. 石家庄: 河北教育出版社, 1996: 466.

2020年博士学位获得者发表学术论文数进行统计时发现,仅达到学校基本要求的学位获得者比例超过三分之一。学位授予标准,从学校到导师,理应是递增从严的趋势。学校制定基本标准,学科可根据建设实际情况适当提高。导师则可在上述基础上根据课题组或所在团队发展的情况,提出更高要求。建议学位授予单位要强化导师队伍建设,明确育人权责,规范导师指导行为,支持导师严格学业管理,严把质量标准,将导师履职尽责、培养质量等纳入导师考核评价体系,发挥导师作为第一责任人的积极性和主动性。

(三) 引导研究生聚焦学位论文质量

博士学位论文是博士研究生在学期间拓展的学术视野、打下的理论基础、具备的科研能力以及锻炼的写作水平的综合反映,是博士研究生培养质量的最终评价对象。我们提出博士学位论文的阶段性成果标准,目的还是为了确保博士学位论文质量。学位授予单位及学院、导师要在研究生的教育指导过程中,将学位论文质量标准、写作指导、科学精神和学术诚信等作为研究生培养的重要内容,组织研究生认真学习主管部门关于研究生教育特别是学位申请与授予方面的法规制度,切实把质量要求落实在学习研究过程中,不断强化贯彻规章制度的自觉性和执行力。学位授予单位可考虑试行“三盲”评阅(隐去作者、导师姓名,以及阶段性成果内容),将公开评阅信息

编入学位论文,排除非学术因素干扰,引导研究生和评审专家更多关注学位论文本身质量,同时完善评阅专家选定标准和办法,注意聘请的本学科专家最好是“小同行”专家,强化同行专家在学位论文学术评价中的关键作用。

参考文献:

- [1] 徐靖.论高等学校学位授予标准中的否定性条款[J].学位与研究生教育,2020(2):5-15.
- [2] 教育部、科技部.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[EB/OL].(2020-02-18)[2021-11-13].http://www.gov.cn/zhengce/zhengceku/2020-03/03/content_5486229.htm?ivk_sa=1023197a.
- [3] 教育部办公厅.关于开展清理“唯论文、唯帽子、唯职称、唯学历、唯奖项”专项行动的通知[EB/OL].(2018-11-08)[2021-11-13].http://www.moe.gov.cn/srbsite/A16/s7062/201811/t20181113_354444.html.
- [4] 中共中央、国务院.深化新时代教育评价改革总体方案[EB/OL].(2020-10-13)[2021-11-13].http://www.gov.cn/gongbao/content/2020/content_5554488.htm.
- [5] 国务院学位委员会第六届学科评议组.一级学科博士、硕士学位基本要求:全2册[M].北京:高等教育出版社,2014:1-8.

(责任编辑:赵惠君)

(上接第73页)

- [4] 西蒙.管理行为[M].北京:北京经济学院出版社,1991:19.
- [5] 肖前.马克思主义哲学原理[M].北京:中国人民大学出版社,1994:45.
- [6] 圣吉.第五项修炼:学习型组织的艺术与实务[M].上海:上海三联书店,2002:25.

- [7] 黄爱华.高等教育评估主体及其行为研究[J].现代教育科学,2004(3):82.
- [8] 熊光慈.高等教育评估中的偏差研究[D].天津:天津大学管理学院,2006:82.
- [9] 马特林.认知心理学:理论、研究和应用[M].北京:机械工业出版社,2016:121.

(责任编辑:赵惠君)