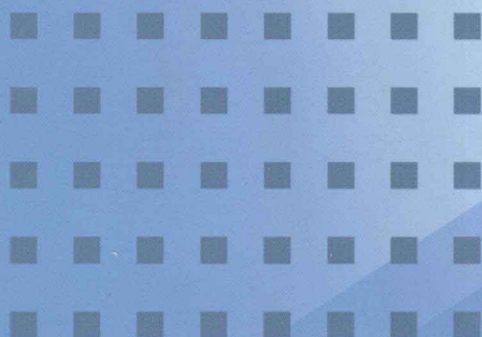


现代 学校管理评价指南

GUANLIPINGJIAZHINAN



本书编委会 编



中国科技文化出版社

第三卷

中国科技文化出版社



第一章 教师评价体系构成的基本要素

一些教育工作者将教师评价等同于课堂观察，其他一些人也将其等同于我们所用过的一些形式。我们需要改变他们对评价体系的看法，其实就是改变评价的形式和课堂观察的方法。尽管在阐述一个评价过程的结构和围绕它展开的专业对话时，评价的形式是非常重要的，但是形式不是体系的全部。一个有效的教师评价体系远比形式来得复杂，它必须包括三个必要的组成部分：

- 对教学的范围有个清楚的定义（即“是什么”），包括依据教学标准，界定满意的教学行为（“教学到什么程度才算好”）。
- 对教学进行全方位评价所需要的技能和程序（即“怎样做”）。
- 训练有素的评价人员，能根据评价程序中说明的教学依据对教师进行连贯的评价。

除了上述几点，在设计（和修改）评价体系时，还必须有各方面人员的参加（如教师、行政人员和教师协会的领导等等），以便保证评价视角的多样性。

设计教师评价体系所面临的挑战是：既要鼓励教师的专业学习，同时也要保证教学质量。多亏了近来的评价经验，如那些由全国的专业教学标准委员会（National Board for Professional Teaching Standards）开发的评价，使得教育工作者在对付这些挑战时站到一个好的起点。

在本章中，我们将探究前面提过的一个完整的评价体系所包括的三个必要的组成部分。我们将会列出保证教学质量的要求和教师专业学习所需要的条件，然后再描述通过哪些方法可以将两者相结合。

第一节 保证教学质量的本质

保证教学质量的要求与我们将在这里将要讨论的评价体系中基本要素的构成（即“是什么”，“如何做？以及训练有素的评价人员）紧密联系在一起。

一、“是什么”

教学中保证质量的核心概念是对示范性教学进行的清晰且一致的定义（更多信息，

详见第四章)。许多州已经清楚地规定了各自的教学标准(如:加利福尼亚州的教学专业标准、德克萨斯州的以学生为中心的熟练标准、以及佛蒙特州的教育工作者标准),学区都以这些标准作为他们评价程序的基础。其他学区发现在丹尼尔森 1996 年所著的《教学的结构:提升专业实践的有效方法》一书中提出的专业实践的几个组成部分对评价的构成也是有帮助的。表 5-1-1 列出了此结构的四个方面及它的 22 个组成部分。

表 5-1-1 专业实践的组成部分

一、计划和准备 第一部分(计划和准备)包括对教学内容、学生背景知识的理解以及设计教学和评价。它包括:计划和准备。它的组成部分包括: 1a、说明教学内容和教学方法 1b、说明学生的学习背景 1c、选择教学目标 1d、说明知识来源 1e、设计连贯性的教学 1f、评价学生的学习	三、教学 第三部分(教学)涉及教师在吸引学生参与学习方面的技巧,以及能使学生学习的一系列的教学技巧。它的组成部分包括: 3a、非常清楚、准确的交流 3b、运用提问和讨论的技巧 3c、吸引学生参与学习 3d、为学生提供反馈 3e、展示出教学的灵活性和反应灵敏性
二、课堂环境 第二部分(课堂环境)强调了教师在建立一个促进学习的教学环境(包括物理环境和人际环境)时所需的技能。它包括: 2a、创设充满尊重的和睦氛围 2b、建立有利于学习的氛围 2c、管理课堂的程序 2d、控制学生的行为 2e、安排空间	四、专业责任 第四部分(专业责任)强调了教师额外的专业责任,包括教师的自评和自我反省,与学生家长的沟通和参与教师专业发展进修,以及为学校 and 学区环境的建设出力。它的组成部分包括: 4a、教学的反思 4b、保持准确的记录 4c、与学生的家长交流 4d、为学校 and 学区的建设出力 4e、专业成长和发展 4f、显示出专业化倾向

为了明确“什么是好的教学”,我们有必要确定不同标准的相对重要性(“他们是同等重要的吗?”),确定教学行为的水平(“当做得非常好的时候是什么样子的?”),还要确定令人满意或示范性教学行为的标准(“什么程度、是足够好,什么程度是非常好?”)。这些教学行为的标准必须明确而不含糊,而且为公众所熟知(详见第四章关于这些问题的充分讨论)。

为了保证教学质量,学校和学区将评价标准建立在近期有关教和学研究的基础之上,以保证标准的有效性。除此之外,评价标准还必须包括教学的各个方面,而不仅仅是教师

所做的那部分工作。例如,如果一个评价体系仅仅阐明了教师在课堂教学里与学生的互动,而忽视课堂以外的发挥教学作用的重要方面,则这个体系是不完善的。我们谁能说,教师课外与学生家长的交流活动不是评价体系中的一部分呢?

二、“如何做”

为了保证一个评价体系的实用性,当学校和学区对“好的教学”做了明确界定的时候,他们还必须保证能让教师来体现这些标准。例如他们将教师与家长的交流列入了评价标准,教师怎样才能展示他们的这些技能呢?因为这项技能在课堂里看不到,学校和学区就必须设计其他的程序来评价它。

评价过程必须允许评价人员依据教师的教学质量做出合情合理的评判,学校和学区必须具备对其提供集中援助的程序,如果有必要,也要对努力提高教学的教师提供帮助。如果在学校和学区对教师提供援助并遵循预期评价过程的所有要求之后,教学行为还不能被至少少部分人接受,那么学校最终必须修改程序,使之便于管理。

怎样进行教师评价包括很多条目,从一般的程序(对新手和有经验的教师或许有所差异)到时间安排、参加评价的人员、以及评价的形式及程序。所有这些对于参加评价的所有人来说都应非常清楚,而且在评价的过程中一定要做到公平公正(关于评价程序的一些详细信息见第五章和第六章)。

三、训练有素的评价人员

评价人员必须接受充分的培训,这样才能作出准确、连贯且有根据的评判。从被评价者(教师)的立场来看,无论是谁对他们进行评价,结果都应是一致的,不受到评价者身份的影响。评价人员这一方进行精确的评价是整个评价体系有效性的重要保证。

对评价人员的培训包括以下几个重要方面:

- 第一,评价人员在行动上能够辨别评价标准的不同样本。课堂活动和教学安排组成了起码的资料,评价者应该选择哪些资料作为不同评价标准的依据?所选择的资料不应该仅仅与各种标准相关,也应该有代表性,比如,不应该仅拿出负面的依据。

- 第二,评价人员必须能依照评价标准解释教学活动。任何一个细心的教学观察者都必须意识到,一件事往往不止一种可能的解释。恰当解释一件事是对教学进行专业评判的一个重要方面。

- 最后,评价人员必须对教师的教学行为做出评判,并且对教学行为水平的描述做出解释。除此之外,评价人员还能够进行反思性的谈话,给教师提供建设性的反馈。

这些与专业学习保持一致了吗?初次看来,很难看出个所以然,因为保证教学质量的要求比较严格,似乎是不可能促进教师的专业成长的。但是在我们认为它们不协调之前,

我们应该再进一步地看一下专业学习的本质。

第二节 专业学习的本质

许多因素都有利于教师的专业学习,包括:教学实践的反思、与教师的合作以及教师的自评。

一、对教学实践的反思

对教师的专业学习来说,没有多少活动能比教师的反思实践更有效。正如舍恩(Schon, 1983)指出的那样,我们并没有从我们的经验中学到多少东西,而是从我们对经验的反思中学到了很多。反思需要提出(和回答)这样的问题“这些是对学生合理的学习期望吗?”“不同组会不会做得更好?”以及“我怎样才能知道学生们已经真正学会了这一概念”?

许多有经验的教师经常在非正式的基础上自发地反省这些问题。但是仅有很少一部分新教师会这样做,而且许多有经验的教师几乎不花什么时间在这种持续性的反思上(因此真正的学习需要这种反思)。

学校和学区在教师的评价过程中的许多方面应该包含对教师实践的反思。自评、对学习活动的描述和评论、以及对学生作业的分析都取决于是否仔细考虑到了教师的期望,以及是否实现了教师预期的目标。

二、与教师的合作

许多教育工作者能够独立地对教学实践进行反思而且能尽量投入一些时间。但是许多人发现为了这样的探究而做自我约束非常困难,而且会被即时出现的紧迫事件给排挤掉。所以建立在合作基础上的体系,尤其当那样的合作需要对实践进行反思,这个体系就很有可能比那些不建立在反思基础上的体系更有效。

有许多文件证明了教师之间的相互孤立。在经历过工作坊之后,许多教师在他们的评价表格中会写上,有机会与同事们谈论一些问题是一天里对他们最有帮助的一个方面。教学是高度复杂的活动,许多教师缺少探究这些普遍性问题和获得可能的解决方案的机会,或者缺少与同事分享新的教学方法的机会。在教育工作者为评价设计体系的时候,体系中可以包含对教师专业发展讨论(教师与教师之间、教师 and 行政人员之间)的规定。

教师合作还有另一个真正的好处——获得不同的观点。大部分教育工作者发现他们的同事在某种情况下的看法肯定与他们自己的有一些不同,而且这种不同的观点能为他们提供其他解决问题的办法。教师们正发现,在面临来自于同事的反例时,坚持对一件事

的一种解释是相当困难的。所以教师合作为解释教育实践提供了一种更平衡、更准确的可能性。

三、自评和自我探究

教师们都是专业人员；他们是一件复杂工艺的实践家。教师们需要知道他们的长处和缺陷，而且他们急切想从各个方面提高他们的实践。如果为他们提供一个安全和尊重的环境，绝大部分教师都会选择在他们最需要的方面付出他们的全部努力来促进专业成长。

成人学习的基本原理表明，当他们在专业成长中运用自评和自我探究时，会比外界强加给他们专业发展的需求更有可能保持他们的学习。当人们选择了他们自己需要解决的问题、他们自己所要完成的项目时，他们会比由其他人来选择问题而投入更多的精力。

一个人选择自己今后努力工作领域的基础是自评，但可以把评价者的加入作为一个补充。特别是许多有经验的教师能够非常准确地评价他们自己的实践，相比较而言，由于只接受过职前培训新教师在这个活动的技能上要比老教师逊色。所以，评价体系应该既包括给有经验的教师自评和自主的专业成长的机会，也要包括给新教师自评方面的指导。当然，如果管理人员有充足的理由相信，教师的自评不都准确，或是教师还没有意识到那些值得注意的教学实践，那么他们就应提醒教师注意这些环节。但是教师的自评应该始终成为我们讨论的起点。

四、学习者共同体

当教师们在同一个项目上进行合作从而提高他们知识的时候，他们就为自己营造了一个学习者共同体。他们会觉得自己在一个小组中共同学习。在这样的情况下，没有差异，所有的人都平等参与活动以提高自身的认识。除了这些，这种环境给他们的尝试活动创造了一个安全的氛围，在学习者共同体中没有一个人会感觉到害怕。

为了获得合作性探究的所有益处，学校和学区需要做出有计划有步骤的努力。专业化探究的氛围不是自然而然产生的，它需要学校来创造。这样的氛围可以采取多种形式创造出来，如：为学习团体碰面提供时间，在观察期间可以安排代课教师，或者提供资金让教师可以参加专业性的会议并获得专业材料。学校可以选择性地组成一个“批判性的朋友”体系，通过这样一个体系，教师可以在一起探究并为其他人提供建设性的意见。通过这样的活动和对教师们有形的帮助，学校和学区为教师专业继续发展履行了该尽的义务。

这种义务同时也反映在学校无形的文化特征中，在一个重视专业探究的学校中，学校领导不容许在教师休息室听到任何对学生或其他教师的否定意见。教师们表现出了一种愿望，他们希望与别人分享他们开发出来的教学材料，希望与其他教师共同面对来自在学

习上有困难的学生的挑战。在这样的学校中,教师不是孤立工作的,他们的心扉是敞开的,他们可以随意交流他们的教学策略。教师评价程序应该包括这样的有利于形成教师专业探究氛围的活动,以及支持这些活动的一些方法。

五、形成性评价的作用

学习几乎都包含有形成性评价。确实如此,由于它根植于教和学的全过程,反而容易被忽视。我们可以将整个过程看成是学打高尔夫球。一个新手必须寻找一位好老师,一个可以提供指导和帮助他提高技能的老师。这个教练当然也必须是一个打高尔夫球的高手,他对这项活动的各个方面都非常熟练。除此之外,教练还应该是一个熟练的诊断家,因为仅仅为学生提供鼓励是远远不够的。如果这个教练总是让学生自己练习或者说“很好,再来一次”也是不够的,相反,教练应该提出建立在明确诊断基础上的具体建议,如“你的右肩太低了”或者“你的臀部太往左了”,或者“在你回击之前要先看准”。这就是教练在评价一个新手,也就是在提供一种旨在提高新手学习的反馈。教练不是评判新手本人,而是评价她的成绩。

这一差异非常关键,它可以帮助我们设计出更有效的评价体系来促进专业学习。举个例子,一个教学的教练必须观察全班并注意教师所提问题的类型。如果这些都是低水平的基础问题,只有一个正确答案,那么这些反馈信息对教师来说是有价值的,而且可以用来提高教师的实践。

第三节 反映“行动计划”中提到的教师专业学习的要求

正如保证教学质量的要求可以在“行动计划的要素”中得到反映一样,专业学习的环境也同样如此。

一、“是什么?”

定义示范性实践——它的教学行为水平,特别是用以获取这一定义的过程——可以提高专业学习的基本原理。当教育工作者考虑什么样的教学才是“好的教学”的时候,以及怎样在各种环境下将其展示出来的时候,当他们争论教学行为的水平以及在不同情境下是否可以同样运用这些行为时,他们就在合作的环境中参与了专业讨论,也会经常反思自身的教学实践。

这种专业谈话的环境,对于专业性的冒险行为来说是安全的。当教师们在考虑一个评价文件的草案以及争论这些评价标准是不是在所有的情境下都生效时,或者争论这些对于多种层次的教学行为的描述是否适用于教学环境时,他们可以各抒己见而不需要惧

怕有对他们不利的后果。因为这样的讨论,没有分数等级也不是走过场式的。

这些讨论总是促使教师进行自评和自我反思。如果叫一个教师去描述不同层次的教学行为,而他不问自己诸如“我是怎样做的?”“在这个标准上我达到了哪个水平?”,那他是不可能真正做到这些的。除此之外,当教师在考虑描述教学的不同组成部分以及各部分的因素,并将自己的教学效果和实践与其他老师相比较时,他们就会与同事切磋教学技能、从同事那里学到新的教学策略。这样的对话是富有成效的,他们以教学质量为重点,为参与者的专业学习提供了很大的帮助。

二、“如何做?”

在教师评价体系为教师专业学习服务的多种多样的因素中,程序所起的作用最大。从这方面来说,教师候选人获得国家专业教师标准委员会的高级资格证书的经历很有意义。实际上,每个经历该过程的教师,甚至那些不是很成功的教师都反映说,他们因为努力通过认证而成了更棒的教师。

这样的现象变成了一个设计评价体系方面的问题:“我们应该怎样为教师评价设计能帮助他们进行专业学习的体系?在评价的过程中,我们能让教师做什么呢?这些作为评价过程的一部分给教师带来专业上的回报并促进其成长。”作为评价过程的一部分,学校或者学区会要求教师提交他们班上一个单元的学习材料,主要包括该单元的教学活动,还要提交一些学生依据该活动而做的作业样本,然后写一份对这个活动以及对学生作业的简短评论。如果评价者问教师的问题要求教师反思和深思熟虑,或要求教师必须挑选学生的作业来证明学生的学业成绩,那么教师必定会从该过程中学到很多。如果他们必须回答这样的问题“这样的活动是怎样加深学生对概念的理解的?”或者“学生的作业样本反映了他们怎样的理解水平?”教师们就会对自身的活动以及对学生个体参与活动的经验进行反思。在要求教师呈递这一单元的教学活动以及学生的作业时,评价体系并不是给他们带来额外的“负担”:教师设计这样的活动以及收集学生的作业都是日常进行的。换句话说,活动代表的就是教育测验服务机构的玛丽·伯尔曼所说的“正常收获”。当然,教师不会总是系统地反思学生的作业或根据这些作业分析教学活动或一个较大的单元,他们也不会将他们的反思写成一个简短的评论。然而,一年叫他们这么做一次,这不是一个不合理的要求,而且这样的反思还可以提高教师的思维习惯。

另外,评价体系的设计者们面临着一个挑战,就是规范评价过程,让教师们有确定的事情可做,而且他们所做的事情都能促进其专业学习。这些必要的活动当然包含行政人员的“听课”。但是评价体系也包含自评和合理安排描述不同教学水平图表的谈话。评价体系要求教师提供有关学生学习的具体依据,还要求教师提供课堂职责以外的样本(如与学生家长的交流、保持正确的记录、或者服务学校或者学区的项目,例如参与委员会等)。

如果一个评价体系设计得好,教师在整个评价过程中就会以积极主动的姿态出现,并且能从中学到很多。

三、训练有素的评价人员

评价体系中用来定位教师 and 培训评价人员的程序也体现出有用的专业学习。当教育者们在谈论评价标准时,当他们在检查一些教学依据(如:录像带、评语)和共同决定这些教学行为处于何种水平的时候,就是在彼此进行有价值的专业讨论了。从这样的讨论中达成的一致意见可以使整个学区对“好的教学”有个共同的认识。

第四节 把保证教学质量 and 专业学习整合起来

学校和学区怎样将保证教学质量的要求和教师专业学习的要求整合到一个体系中呢?什么样的评价体系能满足这两者的要求,能考虑到当前教学方法中某些方面的失调呢?这样的一种综合体系的基本特征包括:差异的评价方法、专业性探究的氛围、以及精心设计的评价活动。

一、差异性

教师职业和其他行业一样,有着独特的生命周期。他们的工作是繁杂的,熟练的教学实践需要有足够的时间和支持来实现。但是一旦教师达到职业熟练的某个特定的水平,专业学习就会表现出与早期经验不同的方式,开始以自我为导向。如果教师在教学技能上有所下降,教学技能降低,或他们的教学滑落到某个勉强够格的水平之下,那么他们也同样能从高水平的帮助和集中的援助中获益。这些情况表明评价过程中所用的程序在教师不同的职业生涯阶段可以有所不同。

所以,学校和学区必须根据教师的专业需要设计出有差异的评价体系。许多学区已经设计出了有三个“层次”的评价体系:(1)针对新手(见习期教师或者非正式的教师),(2)针对有经验的教师(职业的或获得终身资格的),(3)针对一些需要集中援助或行动计划的教师(一些有经验,但是因为无教学能力接下去要被辞退的教师),这三个层次的评价有着独立的活动和不同的时间期限(见第八至第十章对这三个层次评价的详细阐述)。

二、专业学习的氛围

学习氛围,无论对于在校学生还是成人来说,都不是自发形成的,它必须有计划地,在实践中加以弘扬。很大程度上,所有教育工作人员在行政人员的领导下都对建立和维持这样的氛围起着一定的作用。其因素包括以下几方面:

·专业性探究的合作氛围。在复杂的教学活动中,没有现成的专家,所有的参与者都必须向其他人学习。合作性的探究氛围要求所有的教师和行政人员都应期望他们在评价过程中所做的工作能得到专业性的回报。学校领导始终关注学生的学习质量,在这样的情景下,学校的每个人都“统一在一起”来提高学生的学习成绩,他们的集体努力可以被看成像合作一场音乐会一样。

·(对新教师)有支持和援助的精神。这种支持和援助的精神应该通过辅助或引导的项目得到制度化。即使在一个非正式的体系中,这样的精神也值得提倡,而且可以避免那些恶劣行为的出现,如分配新教师去教最差的学生,或叫新教师做最多的准备,而且还没有一个固定的教室等等。这种专业性的“侮辱”(琳达·达林—哈蒙德)是不可原谅的,其结果会带来不必要的压力,还会导致高度的人员消耗和冷嘲热讽的氛围。

在一个有效的评价体系中,有经验的教师会主动邀请新同事来听他们的课,而且在课后会主动与新教师谈论教学策略,帮助他们准备评价过程中出现的一些事情,这些都反映出了支持和援助的精神,使新教师不会在他们的职业生涯中感到孤单。

·(对于获得终身资格的教师而言)关于教学能力和持续的专业成长有两种设想。一旦教师进入职业状态(成为了合格教师),他们就是专业团体的合格成员。明确接受在能力和持续的专业成长上的这两种设想,就是对这种地位的证明。第一种有关能力的设想表明,除非被告知相反的情况,教师的行为至少要达到一个满意的水平。它传递了这样一种观念,就是教师的工作永远不成问题(即生计问题)。这种假设可以删除评价过程中的一些让教师产生焦虑的成分,而且可以让教师在一个放心的环境中去尝试。

但是,这种有关能力的设想,就它本身而言,是不充分的,它必须辅之以持续的专业学习。第二种设想表明,持续的专业成长是每个教师的职责,否则教师的专业知识就会停滞不前。这两种设想必须兼顾,没有后者,前者就会滋生骄傲情绪,没有前者,后者就会失去信任度。

在评价体系的情境中,这两种设想不仅对评价过程中规定的活动起作用而且还对活动方式有影响。对于活动而言,特别是如果一个学校采取了自我专业成长的一些方面作为专业评价的一部分,那么就必须认识到这样的活动能为教师的专业学习提供机会。为了完成这些活动,所有的教师必须提高他们的实践能力。除此之外,对于能力的设想表明,评价的过程也是教师进一步提高他们技能的一个机会。

这两种设想以及探究的氛围形成了教师进行专业尝试的安全氛围。教师们都知道,即使在一次正式的评价观察课上,他们也敢于尝试新的教学策略并接受反馈;或者他们也可以呈递一份教案或者学生的作业,作为他们个人成长档案袋的一部分,这些正是他们想要与资格较深的同事讨论的。这不是“做秀”,而是在切实地提高自己。

三、作为评价过程一部分的活动,可以促进教师的专业学习

一个完整的评价要求在评价过程中能涉及到教学的各个方面。但是在设计上要做出这样的规定:其中一些活动必须要比其他的活动更能促进专业学习。

通过要求自评、参与某个领域的小组合作、而且通过成长档案袋的活动对个人实践进行反思,这样的评价体系可以促进教师的专业学习。不管教师的教学技能怎样,只要参与了这些活动,这些活动都会指导和帮助他们的专业成长。即使没有评价体系,如果教师们简单地做一下这些事情,他们的实践也会得到提高。在这些活动成为评价体系中的一部分后,许多学区确保它们能落到实处。

许多新开发的评价体系要求教师进行自评,设立自己的专业成长目标,以及与同事一起参加研究小组来探讨他们共同感兴趣的话题。(这样的要求让教师进行自评,反思自己的教学实践,并进行自我探究、团体合作和专业性讨论。)所以,除了“课堂观察”之外,还要求教师呈递反映他们教学技能的依据,这些依据可以是计划形式的文件,也可以是学生作业的样本(附教师评语),以及其他一些反映专业性的依据(如,与学生家长的交流、对学校及学区的贡献等等)。收集和挑选这些依据需要对教学实践有较为深刻的反省,评价人员应将这些反映给行政人员,要求教师参与专业性讨论。

尽管起初我们看来,保证教学质量和教师专业学习的要求不可调和,但是,事实上,他们是相辅相成的,但是肯定需要精心的设计。这两个概念本身可以共存,甚至可以相互促进。(见表5-1-2的总结)

表5-1-2 在教师评价中将保证教学质量和专业学习整合起来

条目/程序	保证教学质量	专业学习
教学的定义,包括教学行为的水平	• 清楚,毫不含糊 • 以研究为基础 • 因地制宜	发展的过程促进共同理解
评价的技术和程序	• 信息来源规定所有的评价标准 • 评价人员必须遵循评价程序,包括预期的过程 • 程序必须公平公正	设计要最大限度帮助教师的专业学习和对实践的反思
教师和评价人员的培训	• 评价人员应有根据地进行持续性的评价 • 提高评分者一致性效度	培训过程本身能统一意见并促进共同理解

第二章 教师评价标准

一系列的评价标准作为任何一个评价体系的基石,是学校或学区进行教师评价的基础。我们怎样来评判一个教师的教学行为?确定评价标准与完善评价体系并不是一回事,但是,如果没有一系列清晰、明确的评价标准来共同规定什么样的教学才是“好的教学”,那么这样的评价体系就是不完整的。

那么,优秀(或者适当)的教学应该包括哪些方面呢?我们通过什么样的标准来定义优秀的教学呢?当一名优秀的教育者在社区中享有崇高声望时,当太多学生要求转入优秀教师的班级时,当一个教师受到他的同行高度赞誉时——基于什么样的基础,人们可以做出这样的评判呢?学生在这些班级中感到开心吗?他们参与一些激动人心的活动吗?他们通常能在州评价中获得好成绩吗?

当一个教师在学生和家长中获得了“教学水平低”的名声,那他的名声建立在何种基础之上呢?这就反映出了矛盾的或者不公平的等级标准,或者反映出了教室内师生互动的一种偏好。这或许是教师对一些概念和程序解释不清楚的结果,因此,学生不能很好地理解所学的内容。或许,该教师这种名声是多年来形成的,因为他的学生总是在州的评价中表现很差。很明显,当学生、家长以及行政人员对教师教学进行评价时,他们将评价建立在教师行为表现的基础之上。这些评判是非正式的,但却得出了评价结果。

第一节 输入还是输出?

教学标准规定了教师应该掌握什么知识和他们应该在职业活动中做些什么。人们对这个概念通常从以下两个基本方面来理解:根据教师所做的事或者是他们所获得的最终结果。前者被称为“输入”——教师工作的细目反映了这项工作的复杂性。(这些工作细目不应该成为核对教师不同教学行为方式的一个清单,但是他们表明了满足学生高水平学习的教学的不同方面。)另一方面,后者可以看成是一种“输出”——教师在他们的工作中获得的结果,例如,把学生所要学习的知识或学生的表现,作为衡量教师教学质量的一种指标。

这两种方法,前者在教学评价中的运用有着更长的历史。但是,直到近年来,随着社会对教师和学校责任的呼声越来越高,许多学区(一些学区应他们的州立法机关或教育委

员会的要求)在对教学进行定义时开始将测量学生的学习收录进来,并且运用到教师评价中。

本章分别探究了“输入”和“输出”这两种方法,并指出了它们两者在教师评价运用中的长处和缺点。

第二节 “输入”,即教师所要做的

在传统的教学行为中,“好的教学”的定义就是——一系列的评价标准——指教师在他们的专业实践中所要做的,以及全天的工作任务(甚至是晚上所要做的工作)。在过去的20年中,对“好的教学”的定义逐渐发生了变化,从注重教师本身的教学行为发展到重视运用教学方法来促进学生对知识的理解。

一些州在教师资格审定的程序中掺入了教学标准。佐治亚州在教师行为评价方法(TPAI)上成为领路人(1980),北卡罗来纳州、佛罗里达、康涅狄格和其他各州紧随其后,正如马德琳·亨特(1982)和其他研究者描述的那样,有效的教学是这种传统的一个例子,这种传统在20世纪80年代赢得了全国广泛的关注。

在开发《实践(Praxis)Ⅲ——对新教师资格认证的课堂行为评价》中,教育考试服务中心(ETS)运用了这些成果。开发的过程融合了大量的参考文献、专家组工作、工作分析和试验以及现场测试。最近ASCD的《教学框架:一个新教学体系的作用》(丹尼尔森,1996)一书就是建立在《实践Ⅲ》的基础上,并与由州际新教师评价和支持协会(INTASC)描述的示范性实践的基本原则相联系。这本书发展了《实践》确定的教学技能,包含了有经验教师的工作,而且该书还反映了全国专业教学标准委员会(NBPTS)对教学的研究。在其他人早期研究基础上构建的这个框架是遵循这些方法的最新成果,而且很多学区已经将它作为一种组织机构,来界定有效教学和构建他们评价体系的标准。

一、由“输入”模式引发的问题

当地教师评价体系的开发者在界定有效的教学实践时,面临着许多复杂的问题。这些问题包括重要的假设、教师对标准的认同程度和具体的水平。

1. 假设

任何教学技能的列表(以及相对应的各行为水平的描述)都反映了对教和学的假设,而且反映了什么才是值得学习的以及如何去学。所有这些都已经在逐渐改变。在界定有效教学方面,一些早期成果反映了一个相当机械的“传递信息”的教学观点。尽管信息的

获得通常是教育的一个重要目标,但是过于集中关注这一点,使得其日益与现有的学习理论相抵触,而且与各州和专业组织提出的内容标准中体现出来的学习类型相违背。

换言之,教学技能的列表必须反映现行的最佳学习理论以及学生必须学习什么。例如,全国数学教师理事会(NCTM)的标准强调数学中解决问题的能力,这种技能不能通过死记硬背的记忆来习得。因此,学校要实施那些标准,必须通过结合可能提高学生解决问题能力的教学方法来界定教学。

换个角度来看这个例子:假设你的学校或地区不认可全国数学教师理事会(NCTM)标准。那么界定你们学校良好教学的标准可能不强调认知研究——NCTM 标准是以认知研究为基础的,或者学校可能对它有不同的解释。因此,无论你的学校或地区在标准或教育研究发现中处于什么立场,认识到教学说明的理论基础是重要的,知道内容和学习理论依据的假设也不例外。

2. 认同

检验任何教学技能的一个重要的测验是看它能否被受其影响的教师所接受,这反映了实践者对他们工作的了解。因此,学校和地区在对教学进行定义时,需要为广泛的讨论提供机会。这个过程既可以使标准合法化也可以说明其适合的对象以及需要承担的义务。当这样的讨论由一定范围的教师和管理者参与时,它就会产生人们乐于接受的文件了。

当新教师加入早有一系列评价标准的学校时,他们就没有机会为开发标准做出贡献。然而,如果他们知道评价过程中听取了同事的心声,而且当评价体系再次修订时,他们也能提供意见,那么他们可以确信行政管理部门并不是独断地选择评价标准的。

3. 具体水平

教育工作者可能会在一般和具体的各种水平上组成教学技能列表。例如,州际新教师评价和支持协会(INTASC)确定了 10 条基本原则和 53 条行为指标。全国专业教学标准委员会(NBPTS)确定了 5 条主要基本原理,这 5 条基本原理进一步详细阐述了已经开发的评价的每个水平和学科的标准。加利福尼亚教学专业标准组织制订出了 6 条标准,32 个要素。教育考试服务中心(ETS)的《实践Ⅲ》围绕 4 个领域和 19 条标准构建,丹尼尔森(1996)《教学框架:一个新教学体系的作用》一书将教学组织成 4 个领域和 22 个部分。

教学技能的列表应该足够详细,便于促进关于教学实践的详尽交流,但不能过于琐碎,使人们无论牢记在心,例如,路易斯安那州教学评价的标准曾经包含 145 条独立的标准,尽管它有巨大的价值,但它包括了太多的标准;不便于大多数人的运用。

4. 冗余

形成评价体系基础的教学技能列表应该避免冗余。也就是说,列表应该包括教学每

个重要方面,但是只能是一次。无可否认地,教学是非常复杂的,技能是相互交迭缠绕的。但是确定它的不同方面还是可能的。可能之处在于,教学的不同任务应该相互分离。

5. 反馈选项

详尽的评价标准让指导者、训练者和督导员能就教学的具体方面提供反馈,而且使教师能够为改进他们的实践确立目标。一些教学技能的列表全面地界定教学,它们根据不同的标准从不同的角度来描述什么是整体的努力。虽然这样的列表加强了教学各个不同方面的联系,但是他们对教学的具体方面实施评价和提供反馈不太有利。

教学行为的水平。界定教学仅有一系列教学技能是不完善的,除非它包括对不同行为水平的描述。这些描述有利于界定等级量表在操作水平上的区分点。例如,如果一个评价体系在行为等级上确定了3个区分点(如“有待改进”、“满意”和“优秀”),那么各个区分点对每个标准有什么实际含义呢?课堂里哪些具体活动会使教师行为,例如“管理课堂程序”被认为是“不令人满意的”?

表5-2-1就是行为水平的一个例子(出自丹尼尔森1996年出版的《教学框架:一个新教学体系的作用》)。像这样的图表能够帮助美国的学校和地区修订他们的教师评价体系。教师和督导员都发现对行为水平的不同描述是有用的,因为这样的描述侧重教学实践,并指出需要做哪些改变才能达到下一个水平。

表5-2-1 领域2:课堂环境

组成2a:创建一个尊重和谐的环境 行为水平				
要素	不合要求	基本达到	熟练	优秀教师与
学生的互动	的互动是消极的、没有意义的、挖苦性的或者不适合学生的年龄或文化背景。学生对教师是不尊敬的。	常是合适的,但可能反映出偶尔的矛盾、偏爱或者漠视学生的文化背景。学生只对教师表示一点点尊敬。	友好的,并显示普遍的热情、关心和尊重。这样的互动是适合于发展的,符合文化标准。学生尊敬教师。	现出真挚的关心和尊重。学生作为一个个体尊重教师,远超过教师本身的角色。
学生之间的互动	学生之间互动的特征是冲突的、挖苦的或者羞辱的。	学生并不对其他学生展示消极行为的。	学生之间的互动通常是有礼貌的、尊重人	学生个人对其他学生展现真挚的关心。

行为水平允许对教学进行的讨论不只针对个人。也就是说,如果评价者引用一次课堂观察中的事件作为某个行为水平的依据,那么语言是为调和对话服务的。而且如果教师不同意评判,他可以提供不同的依据,这个依据可能产生不同的评判。以表格中的行为水平作参考,教师可能很少认为督导员的评判是个人攻击。当然,运用明确标准的评判正是建立在有依据的基础上的。

在关于教育工作者如何准确描述行为水平的讨论中,学校或地区的教育工作者本身

对评价标准形成了一种共同的理解，而且通过他们创造的实践例子，参与到了重要的专业对话。但是，在所有情境中对行为水平的描述并不都是采用相同的方法。例如，一个有 18 位学生的幼儿园教师和一个有 300 位学生的高中乐队教师相比，教师了解学生的程度（他们个人的兴趣、他们的文化背景以及他们对学科的准备），的确有所不同。（这并不是指高中乐队教师不应该了解她的学生和他们运用乐器的技能水平。但是包含的学生数量使这样的教师了解个体学生的知识水平变得困难。）

灵活地将行为水平运用于教学评价还有另外一个原因：因为有很多因素影响，所以行为的变化是可以预料的。例如，大多数教师会在一月份展示出比十月份更高水平的行为——因为他们已经了解他们的学生，并且确定了课堂惯例。同样地，教师教某一个年级时间越长，而且是在同一个学校或地区（因为教师对课程以及学校和文化非常熟悉）时，教师的行为水平就越高。因此，我们应该把行为水平看作教学的行为水平，而不是看作教师的行为水平，并且要着眼于行为判断。

例如，如果一个教师教五年级教了 10 年，他已经形成了很多适应那种情境的技巧和策略。然而，如果他重新分配去教二年级，那么他不得不改变很多策略并学习一些新的技能。二年级学生不同于五年级学生，课程显然也不同。同样地，一位在某所学校教了几年书的教师可能会发现，如果她被调到镇上另一所学校，她需要不同的技巧。同样不可改变地，几年来一直专注于教生物的高中科学课教师，如果要教化学，大概花相当多的计划时间。为了在新情境中表现良好，一位尽责的教师会投入大量的时间和精力，即使是在有利的条件下，教师仍然需要时间适应这样的改变。当教学背景发生变化时（水平、内容或者环境），一位有经验的教师可能就会变成新手。

二、其他的测量问题

界定什么是良好的教学，评价者必须解决权重、标准的情境以及“分数”组合的问题。

1. 重要性

所有的评价标准都同等重要么？还是一些评价标准比其他的更重要些？例如地区的教育工作者会认为构建一个尊重和谐的环境比为学校或地区作出贡献更重要吗？或者，有些评价标准在一些情境下比在其他情境下更重要？例如，让小学教师了解一些有利于他们教学的资源是重要的，但相对于高中教师来说，这样的要求也很重要？如果从日常事务和程序的质量上对新教师进行评价，这对他们而言是重要的，那么更有经验的教师是否也应该采用这些技能呢？我们将在第九章对这些问题进行充分的讨论。

2. “分数”组合

在某方面优秀的行为能够弥补其他方面欠缺的行为么？换句话说，评价者可以“平