

MATE 多元教学有效性评价工具的设计与检验

邢 磊 张兴旭

摘要：通过梳理教学有效性评价的多种取向和主流评价方法，衡量有效教学相关的课程特性、教师特质、教师教学行为、学生特性以及学生学习成效等不同维度之间的关系，形成对教学有效性的综合性界定并设计开发教学有效性的多元评价框架（MATE 模型）和相应的评价工具。通过对上海交通大学 2017—2019 学年的 170 余门课程实施评估的应用实例以及对数据的定量分析，探讨评价框架和工具对教学管理和教学发展的有效性和局限。

关键词：教学有效性；教学评估；多元评价

一、教学有效性的界定与取向

有效教学的理念源于 20 世纪上半叶。从产生至今，有效教学就一直是教学研究的焦点之一，频繁出现在教育文献和实践者的视野之中。但有效教学是一个动态发展的概念，其内涵一直随着教学价值观、教学的理论基础及教学研究的范式变化而不断扩展。美国心理学会和高等教育协会前主席麦基奇曾在 1979 发表的文章中将教学有效性定义为（教师）帮助学生达成教育目标的程度^[1]。然而，高等教育界对于“好的教学”始终缺乏共识。一方面是因为对教与学本身规律的探索仍在不断深入，另一方面则是高校教学实践存在差异。纵观教学有效性评价的历史，出现过以下几种主要的评价取向：

（1）教师个人特质取向。伴随着 19 世纪心理测量运动、心理测量理论的形成和在实际中的大量应用，有效教学最初被认为是教师某些个人特质的体现^[2]。此时一般会借助于心理测量量表来评价一系列理想化的教师特质，如：热情、公平、严苛、幽默等。简言之，能不能教得好，首先要看你是否具备理想化特质。但若仅以此评判，就过于依赖天性，有失偏颇。

（2）学生学习成效取向。关注“学”的结果，以学生最终在知识、态度、技能、能力、行为、价值观、素养等方面的发展变化来衡量教学的质量。在当代高等教育质量越来越受到家长和社会关注的背景下，在专业评估和认证盛行的推动下，学生学习成效评价越来越受到重视^[3]。但对于高等教育而言，它的最大局限在于操作性不强。因为用学生的学习直接代表教学的有效性，不但需要测量的效度能够区分不同教师、不同学科、不同教学目标的差异，而且要能支持增值方式（前后测之差表征学习变化的程度），这对学生学习测量工具的要求非常高^[4]。

（3）学生认知过程取向。跟以学习成效的结果取向不同，认知过程取向关注学习有效发生的内在心理机制，认为有效学习取决于学生的学习观及先前的学习经验、学习方法（如深度学习法和表层学习法等）、动机水平等^[5]。值得注意的是，由于认知过程是内隐的心理属性，评价测量技术角度的可行性一直难以突破。然而随着新技术的发展，越来越多的测量手段不断在提升对学生能力作出判断的可行性，比如隐形性评估、学习分析等都是当前评价研究和实践最前沿的热点。

邢 磊，上海交通大学教学发展中心主任助理，助理研究员；张兴旭，上海交通大学教学发展中心培训助理，研究实习员。

(4) 教师教学行为取向。围绕“教”展开，立足于尽可能描述客观反映和还原教学过程的教学方法等行为现象，作为判断教学有效性以及帮助教师改进教学的参照。这类对教学有效性的教学行为取向具有非常好的操作性，因而在过去几十年里得到了很好的发展，出现了一系列有代表性的观点。如：美国高等教育协会1987年发表的美国大学本科教育的七条原则，得到高教界的广泛关注和认同^[6]。美国学者鲍里奇在回顾了针对基础教育教师教学行为的30年研究基础上，梳理出最重要的5种教学行为^[7]。虽然是从基础教育中总结出的规律，但和上述大学本科教育的七条原则颇有相似之处。

(5) 教师核心素养取向。教师核心素养(Competencies, 也译成能力、胜任力)取向与教学行为取向在实际的评估中都是以外显的教学行为作为观察和衡量的。不同在于素养是“一种将知识与技能、认知与情感、创造性与德性融为一体的复杂心理结构”^[8]，体现了对将有效教学简单看作在特定情境和条件下呈现的外部行为的修正。

涉及人的心理与行为的研究是极其复杂的，教学活动更是如此，上述取向都有其存在的意义和价值，但都有被诟病的缺点。因此，将教学有效性的不同取向整合到同一框架下，多角度地去获取教学过程、教学行为的全貌，将使教学评估更为客观、科学。

二、多元教学有效性评估模型与工具的设计与开发

评估工具的开发始于评估框架的搭建，我们基于教学有效性的几种取向，在文献基础上设计了教学有效性多元评价模型。该模型整合了教学行为、教师特质、课程特质、学生特质、学生满意度和学习成效六个维度。教学有效性评价另一个重要的方面就是评价方法。对教学有效性的不同取向，直接决定了教学评价方法的选择。当代主流的高等教育教学评价方法按普及程度有学生评教、同行评议、教师自评。

为了避免单方面评价所造成的偏差，我们借鉴企业360度环评技术路线，采取兼而有之的混合策略，集学生评教、同行评议、教师自评于一体。这样即形成“三元六维”的评估模型(见图1)，最核心的是当中的教学行为，除了大家较为熟悉的学生满意度、教师特质、学习成效等维度，还纳入了课程特质和学生特质。课程特质这个维度是为了在课程比较时能够区分不同课程所存在的差异。学生的特质主要是了解学生的学习方式、状况，进而帮助老师更好地找到匹配的教学改进方法。

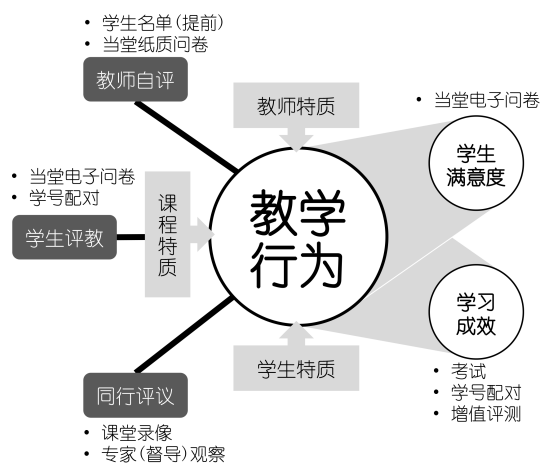


图1 多元教学有效性评估模型

该模型能够很好地提供一个系统考查教学有效性的框架，而且充分考虑了其他影响评价结果可比较性和普适度的维度，如课堂规模、课程难度等决定课程间差异的课程特质，为评估结果兼容总结性和形成性用途提供了保障，也为教学管理和教学发展工作提供理论指引。但涉及具体的实施落地，该评估方案的缺陷就是需要较多的投入。

三、评估模型应用

该评估工具目前已经在上海交通大学推广使用，截至2019年秋季学期共覆盖283门课程。其间也面临了很多困难和挑战，但随着具体的实践和反思，问题正在不断地被优化。其一，学生问卷体量较大，为保证填答率，我们采取发小礼品的方式来鼓励学生认真填答，邀请院系领导参

与，录制学生问卷填答说明，从行政约束的角度使学生重视，也能消除学生的评价顾虑。其二，目前评估的大部分工作是纯人工完成，尤其是评估报告的生成，占用较多工作量，未来如在更大范围内推广应用，这种评估形式可能会承载不了。为破解此困境，我们目前已经在做评估网站，明年全自动化的版本将会上线。

该评估模型更深层次的应用体现在评估报告的解读。评估结束后可以给到老师课程层面的评估报告，帮助老师更好地审视和改进教学。对于院系而言，评估报告可以应用在多个方面，如做本科评鉴、专业认证、专业院系课程质控系统或者是完成学校硬性的评估任务等。例如，新建了一个专业，需要摸底教学质量，就可以对一组课程进行评估，除了个人报告外，还可以提供整门课程的团队报告，为院系进行教学改革、提升整体教学质量提供一个有力的抓手。目前，MATE 评估使用量不断增长，由此可见，大家对教学评估的重视程度与日俱增。对于任何报告，进行专业的解读是关键，这样才能正确地把控未来改进的方向和范围。正所谓磨刀不误砍柴工，不要急于给课程质量或教学能力下一个结论，而是要花时间和心思去研读报告，从报告所呈现的数据中去发现个人教学中所存

在的问题，去挖掘整个团队教学的规律和共性短板。下面从一些维度进行剖析和解读。

(1) 课程特质差异。如图 2 所示，从阅读量、作业量、难度系数可以了解课程设置情况，通过与整门课程的均值对比，可以更具具体地知晓课程的阅读量是多还是少，课程难度是否把握合理。从团队的数据中，也可以挖掘出更深层次的信息，如不同类型课程的特点。图中三角形和圆形表示的是专业基础课和专业必修课，环形和菱形是通选和选修课，从这一组数据中可以看出课程对于学生的强度。整体而言，课程特质的维度能够很好地区分课程间的差异，显示出一些课程的规律。

(2) 教师特质的差异。20 世纪 20 年代高校开始使用问卷以来，教师特质出现在很多学校的评估问卷中，很多学校现在也使用教师特质来做评估。从人岗匹配的角度来分析，这是合理的举措。因为一开始我们在不知道教师实际教学行为的情况下，可以根据其性格特质去预测未来是否能胜任这个岗位，总是会有一些和老师个人特质相关的变量，如文献中所提到的幽默、正直、热情。

(3) 学生认知过程的差异。“学生认知”这一维度的设定借鉴了约翰比格斯的认知量表，

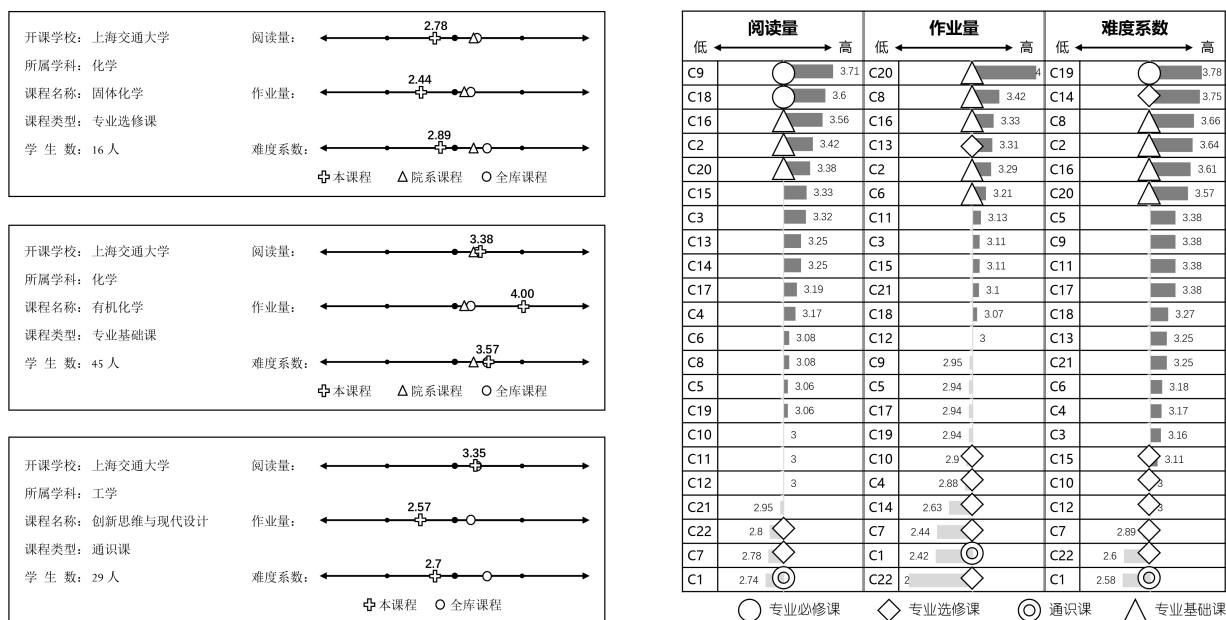


图 2 课程特质的差异

是测量学生学习过程当中的状态，有表层学习和深度学习两个象度，划分为四个象限。X轴越高表示学生深度学习越多，Y轴越高表示学生表层学习越多。图3所列出的三门课，学生的分布都不一样。图中每个点代表一个学生，对于老师而言，最大的意义在于能够看到班级学生所处的学习状态，集中反映出的某种趋势，是很好的反馈信息。同时院系报告中包含关于学生认知的分布图，图3中右下方散点图分

别代表所有学生和所有课程的平均分布情况，从点的集中趋势可以看出一组课程的学生学习状况如何。

(4) 教学行为的差异。MATE 评估的教学行为包含 5 个维度。维度模型的构建借鉴了经典的美国本科教育七法则理论和关键教学行为的理论。相对而言，这两个理论被认为是比较符合有效教学的原理。如图 4，从这些教学行为的维度切入，可以看到课程和课程之间的差别。

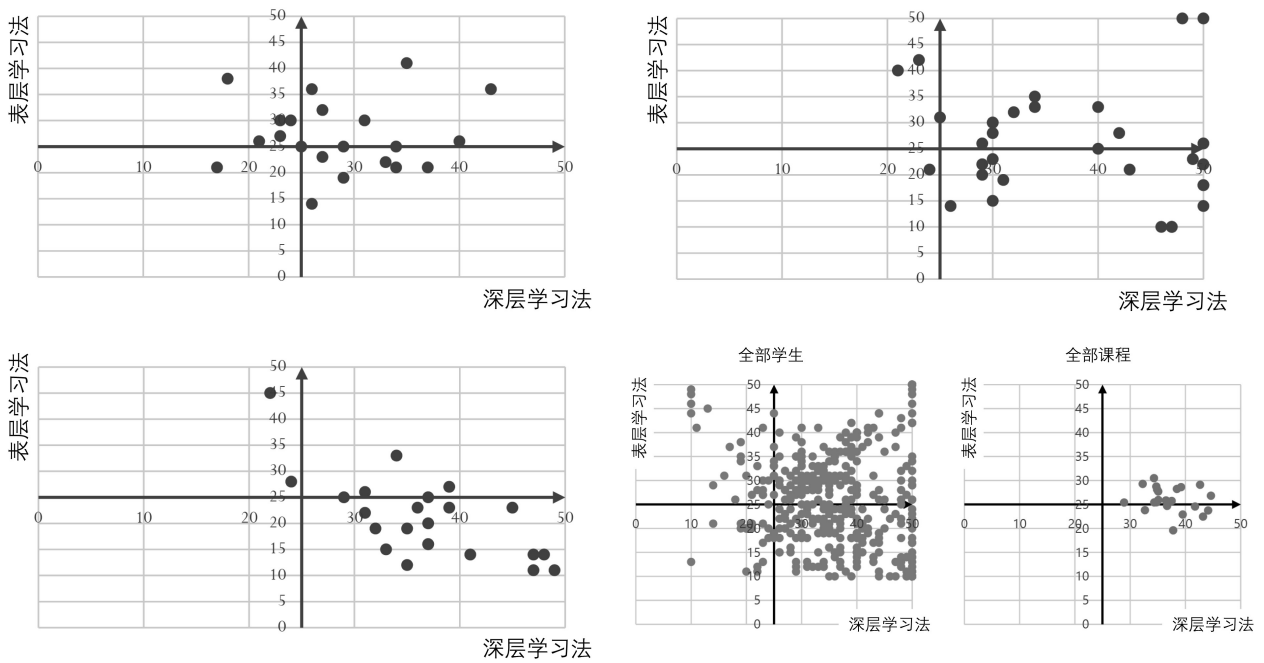
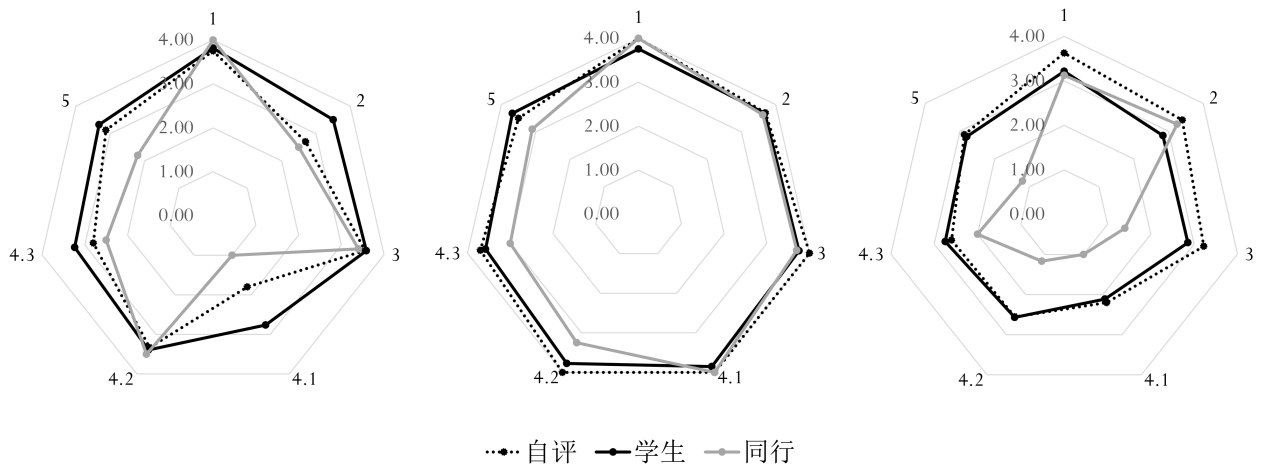


图 3 学生认知过程的差异



- 1. 教师讲课清晰 2. 教师确保围绕教学目标开展充分高效的教學 3. 教师对学生寄予高的期望并确保成功
- 4.1 教师发展学生合作 4.2 教师鼓励学生主动学习 4.3 教师尊重个体差异性 5. 教师给予学生及时的反馈

图 4 教学行为的差异

不同的折线代表不同的评价主体，将教师自评、学生评价、同行评议置于同一图表中，能够很好地识别出在哪些教学行为上三个评价主体的认知是一致的，从而找出优势或共性短板。如果评价主体间的评估差异较大，我们则认为在该维度上评价者存在认知偏差，需要进一步去了解从而找到可以解释此种现象的深层次原因。传统评教也包含督导或同行评议，但问题在于不同评估的维度不一样，无法实现将数据进行整合和对比。这也正是 MATE 评估的优势所在。在一个报告中把这些数据很好地汇聚在一起，方便使用评估结果给实践反馈。

(5) “教学有效性” 综合评估指标。做任何评估都有其目的和出发点，所以在评估实施前必须明确通过此次评估拟解决的问题是什么。

不同院系使用 MATE 评估的目的会有所差异，可能是为了晋升考核作参考，也可能仅是希望通过全面评估来探索提升教学质量的方式。考虑到用途不一，MATE 评估尽可能收集全面的数据，使院系依据需求去设定符合期待和应用场景的计算权重。如图 5，MATE 提供了四种不同的权重。通过实践我们发现，当权重设定不同时，课程之间也有一定的差异。如图 6，实折线代表，不管何种权重设置，多元主体对课程的评价是较为一致的，反映出评价的稳定性，评价者对这门课程有较为一致的认知。虚的折线说明，当只依赖学生评教时，对教师或课程的评价相对较低。如果按照对等的权重综合计算不同评价者的分数，结果就会发生改变，过低或过高的情形会有所改变。由此来看，

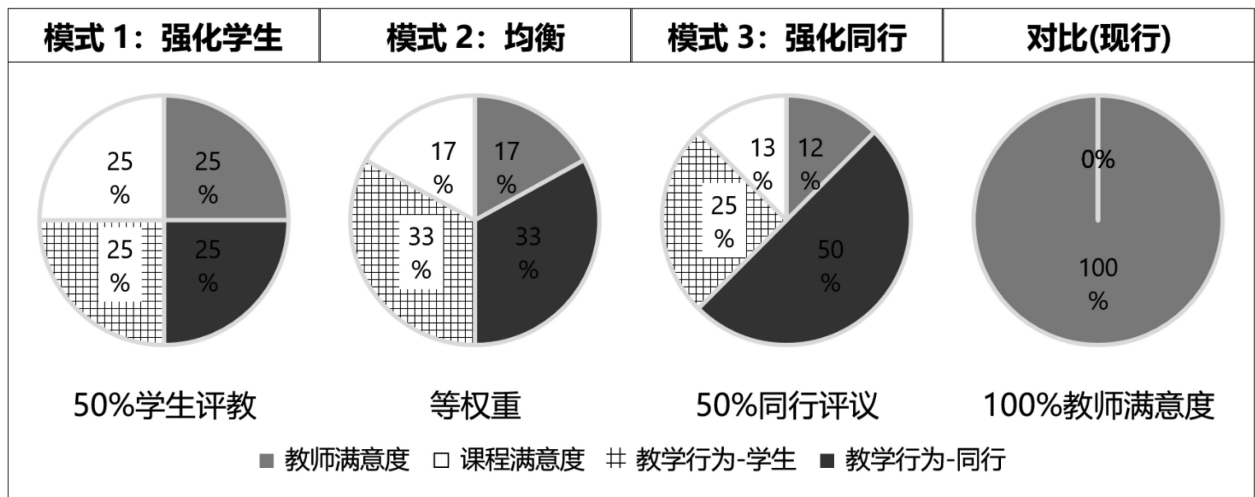


图 5 四种不同的权重分布

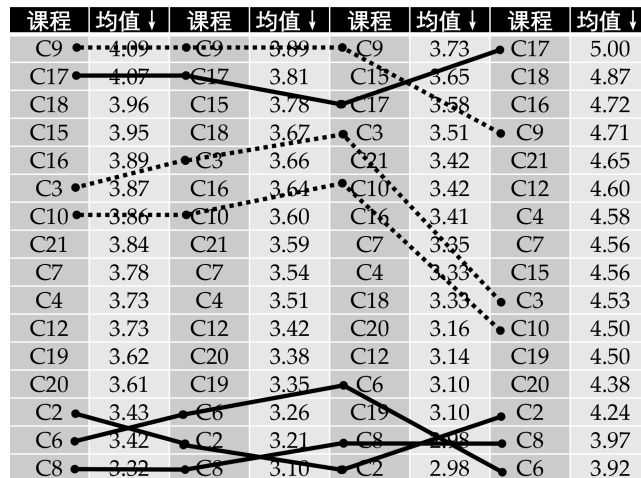


图 6 不同权重评估结果分布图

从不同的视角来看，对课程的评价结果是有偏差的，应该根据具体的应用场景来选择评价权重。

四、未来展望

综上所述，多元教学有效性评估作为一种新兴的、全面的评价工具，能够很好地缓解目前高校对教学职能日益重视与教学评价手段匮乏之间的矛盾。但其未来发展也面临着很多挑战，比如：随着使用量扩增，纯人工服务将无法负载大批量评估；院系在推进教学评估的过程中，需要专业性的指导；如何为学校、职能部门、教师本人解决评教问题，获得充分的关注和认可。基于存在的问题，我们未来有三大努力的方向。其一，2020年MATE网站在线化将实现，并全面推广到全校教师使用，网站的开发将缩小MATE评估的成本，也能够覆盖更多教师。其二，MATE 2020年工作目标，配合院系定制质控方案。以院系为主体，将学科差异考虑进去，为院系提供定制的方案。从成本角度考虑，可以兼容院系已有的资源，如督导系统，避免两条线重复建设。这就需把数据、算法、流程与学院原有的评价指标做匹配和定制。当然院系评估也面临着学生填答问卷负荷重的问题。为避免影响填答有效性，我们将开发简版问卷，只保留关键指标。其三，在高校里推进教学评估并非易事，评估工具的顺利落地应用还需要高校职能部门的一些政策或要求来保障。未来MATE评估结果将被纳入人事晋升中，作为人事晋升教学方面的评估抓

手。教师可自愿申请参评，用于不断改进个人的教学水平，同时也是为长远的晋升和个人发展作准备。

参考文献：

[1] Mckeachie W J. Student Ratings of Faculty: A Reprise[J]. Academe, 1979, 65(6):384-397.

[2] 孙炳海, 申继亮. 美国教师评价的发展历程与评价模型研究述评[J]. 比较教育研究, 2009(5): 73-76.

[3] Ewell P T. An emerging scholarship: A brief history of assessment[M]// Banta T W E. Building a Scholarship of Assessment[C]. Jossey-Bass, 2002.

[4] Wieman, C. A Better Way to Evaluate Undergraduate Teaching[J]. Change: The Magazine of Higher Learning, 2015(47):1, 6-15.

[5] 迈克尔·普洛瑟, 基思·特里格维尔. 理解教与学[M]. 北京: 北京大学出版社, 2007.

[6] Chickering A W, Gamson Z F. Seven principles for good practice in undergraduate education[J]. Biochemical Education, 1989, 17(3):140-141.

[7] 加里·鲍里奇. 有效教学方法(第七版)[M]. 南京: 江苏教育出版社, 2014.

[8] 张华. 论核心素养的内涵[J]. 全球教育展望, 2016, 45(4): 10-23.

[基金项目: 全国教育科学“十三五”规划2018年度教育部重点课题“循证的高校教师教学核心素养框架研究”, 编号: DIAI80403]

[责任编辑: 陈立民]