

# 关注学习效果:建设全校统一的教学质量保障体系

——美国“以学生为中心”的本科教学改革研究之五

赵炬明 高筱卉

**【摘要】**本文是美国以学生为中心本科教学改革系列研究的第五篇。首先讨论了评估和评价这两个专业术语的含义以及它们的习惯用法。由于历史原因,在美国高等教育发展史中,评估主要与学校和专业认证有关,而评价主要与学习效果测评有关。通过这一分析,为读者提供一个基本历史背景。然后介绍美国高校的本科教学质量保障体系的构架与逻辑。这个体系的基本特点是:一条主线、三级管理;具体表现为一个使命(学校教育使命)、三个矩阵(通识教育矩阵、专业教育矩阵、课程教学设计矩阵)、两个辅助机构(教学支持中心和院校研究办公室)。贯穿整个体系的基本管理思想是系统管理和全面质量管理。随后介绍了该体系和校内外各种评估的关系。最后讨论了与体系构建有关的六个关键问题。最后是给中国高校的一个建议。

**【关键词】**以学生为中心的本科教学改革 教学质量保障体系 本科教学 美国高等教育 中国高等教育

本文是美国“以学生为中心”的本科教学改革(SC改革)研究的第五篇,主要讨论评价与评估问题。《论新三中心》一文提出<sup>[1]</sup>,美国的SC改革有三个特点:以学生发展为中心、以学生学习为中心,以学习效果为中心,评价评估问题主要涉及第三个特点:以学习效果为中心。这是整个SC改革的落脚点,因此十分重要。

评价评估研究有很多角度,如评估指标体系、评估流程、评估方法、内外部评估等。相关文献可谓汗牛充栋,但有两个主题研究对大学教学十分重要,研究却相对不足,且值得中国大学关注。其一是构建全校统一的教学质量保障体系(Educational Quality Assurance System, EQAS),其二是基于课堂教学的评价模式与方法创新。

任何评估,其基础归根结底都是学校的EQAS。构建全校统一的EQAS是学校各类教学评估的基础。如何使这个体系能围绕学生发展、学生学习、学习效果这三个核心问题来构建,并把SC改革精神贯穿其中,是SC改革的核心问题之一。然而在这方面,中英文文献都相对不足。

另一个重点是课堂教学评估。说到底,EQAS只是个质量管理体系,教学质量最终要落

实到每个课堂。体现在每个课堂的教学之中。教师做课程设计,不仅要设计课程内容和教学方法,还要设想好预期的学生学习效果,以及相应的效果评价方式。要在教学过程中收集学生学习的证据,用以改善课堂教学,并支持校院系的教育决策。因此,如何在课堂教学中做好教学评价设计、系统收集学生学习效果的证据,成了学习效果评价中的核心问题。在SC改革中,美国大学教师和研究者确立了一些很有创意的课堂教学评价模式和方法,值得介绍给中国的大学老师和教学管理者。

因此,笔者计划用两篇文章来分别介绍这两个方面的情况。本篇介绍学校教学质量保障体系构建,下篇文章介绍美国大学在课堂教学评价方面的创新。

需要再次指出,本文虽然研究的是美国,但目的是为中国,因此在问题选择、研究角度聚焦、内容选择和建议方面,都不一定与美国学者一致。

本文共有五部分。首先是概念和历史。重点讨论评估和评价这两个概念,目的是厘清这两个术语的含义,同时说明教学评价与学习效果运动之间的关系,从而为全文提供一个基本的历史

收稿日期:2019-03-30

作者简介:赵炬明,华中科技大学教育科学研究院教授、博士生导师;高筱卉,华中科技大学教育科学研究院博士研究生,中南财经政法大学管理学院讲师。

背景。第二节介绍美国高校的 EQAS 框架。这个框架是“理想型”框架,旨在揭示美国大学 EQAS 构建的思路和逻辑。第三节讨论如何利用 EQAS 来支持校内外评估。第四节讨论六个与 EQAS 有关的关键问题,并就此给中国高校提出一个建议。最后是一个简要总结。

### 一、概念与历史

在关于评价评估的文献中,有两个术语令人头疼。在中文是“评价”和“评估”,在英文是“assessment”和“evaluation”。这两个词都是从英文翻译过来的,由此产生的两个问题是:① 这两个词各是什么意思? ② 应该如何翻译? 人们由于分不清两者的差别,因此经常混用。

例如,“教育部高等教育教学评估中心”官方英文译名中的“评估”用的是 evaluation; 而史海燕主编的教育部统编教材《教育测量与评价》则把 evaluation 译为“评价”<sup>[2]</sup>,直接忽视了“assessment”的存在; 黄海涛撰写的《学生学习成果评估——美国高等教育质量保障研究》一书中,则把“assessment”译为“评估”。虽然作者在梳理文献时注意到了英文文献中存在“assessment”和“evaluation”这两个不同术语,但最后决定忽视二者差别,把它们都译成“评估”。<sup>[3]</sup>这个处理方式显然不会让北美学者们满意。例如,2017 年 Charles Secolsky 和 D. Brian Denison 主编的《高等教育测量、评价和评估手册》(Handbook on Measurement, Assessment, and Evaluation in Higher Education)(简称“手册”)中,编者特别指出:

在高等教育中,许多人都不加区别地使用这两个术语(指 evaluation 和 assessment—引者加)。然而我们认为它们是两个不同的领域。Assessment 的核心是收集、分析和解释与某个特定问题或结果的有关信息,而 evaluation 则涉及确定某些东西的价值和有效性,通常指一个专业项目(program)。因此,evaluation 可以包含一个 assessment 计划,并把它作为专业质量判断的依据。<sup>[4]</sup>

这段话传递两个信息,第一,这两个术语在英文文献中也经常混用,不只是中国学者的问题。第二,专业人员认为,assessment 指信息收集、分析、解释及有关的活动;而 evaluation 指以这些信息为证据,对评估对象作价值和有效性判断,尤其是在专业评估和学校评估中是如此。由于 evaluation 必然涉及相关信息收集、分析和解释,因此完整的 evaluation 中必然包括 assessment,但 as-

essment 不是 evaluation。按照这个说法,教育部高等教育评估中心的用法符合专业用法,因为评估中心的主要工作就是对学校和专业的质量和有效性作评估;而史海燕和黄海涛对这两个术语的处理可能还需要斟酌。

应当指出,上述看法并非该书两位主编的一己之见,而是美国高等教育评价评估界的共识。例如,著名的美国学习效果评价研究所(National Institute for Learning Outcomes Assessment, NILOA),用 assessment 而不用 evaluation,其目的是要告诉大家,我们研究的是如何收集、分析、解释、使用大学生学习效果的信息,但并不打算对大学或专业的工作做价值判断。我们做 assessment,不做 evaluation。

那么,这两个英文术语究竟应该如何翻译呢? 既然教育部评估中心已经把“evaluation”翻译成“评估”,为了避免混淆,本文也循此为例,把“assessment”统一译为“评价”,而把 evaluation 译为“评估”。评价指信息收集、分析和解释的活动,其目的之一是为评估提供证据;而评估指根据收集到的信息和证据,对学校和专业办学的绩效做出有效性判断。

中英文文献中都出现这两个术语的混用,说明它们语义接近,是近义词甚至同义词,很难仅靠词义来加以区分。因此,如果要想深入理解这两个术语的专业用法,则需要追溯它们的历史,即从词源学角度加以考察,看它们是怎样演变成了代表两类不同实践的专业术语。

在美国高等教育的历史发展中,评估的发展早于评价,尤其早于对学生学习效果的评价(student learning outcomes assessment, SLOA)。1910 年代由卡耐基教育促进会发动,而盛行于 1930—40 年代的院校自我研究运动(institution self-study movement)中,对高校的专业评估和学校评估就已成形。二战后到 1970 年代美国高等教育大发展,促成了质量认证(quality accreditation)的发展,于是出现大量全国性 with 地区性认证组织。当时认证的重点是两个:专业与学校的办学质量和有效性。这进一步强化了学校评估和专业评估的发展。但有必要指出,这些评估是管理导向的,主要关心学校和专业的办学效果与效益,评估主要围绕情景(context)、输入(input)、过程(process)、生产(productivity)四个方面进行,用今天的术语,即所谓“组织有效性”(institutional

effectiveness)。这里的“组织”可以是学校,也可以是院系或专业。围绕评估和管理决策而出现的学校教学与管理方面的信息收集、整理、记录、分析、报告与决策支持,成了学校管理的必要辅助条件,于是专业化的院校研究应运而生。参与校内外评估变成了院校研究的基本任务之一。<sup>①</sup>

到了1970年代评估进入成熟期。很多重要评估期刊开始出现,如教育评估研究(Studies in Educational Evaluation, 1975)、评估综述(Evaluation Review, 1977)、评估与专业规划(Evaluation and Program Planning, 1978)、教育评估与政策分析(Educational Evaluation and Policy Analysis, 1979)、专业评估新方向(New Directions for Program Evaluation, 1982)等。1976年专业化的评估组织评估研究会(Evaluation Research Society)和评估网(Evaluation Network)相继成立,1986年这两个组织再合并成为今天的美国评估协会(CHEA)。因此美国评估学界认为,以认证和管理改进为导向的评估到1970年代已经进入了成熟期。<sup>[5]</sup>

但随后二十年的两个发展改变了这个趋势。一是始于1980年代的SC改革运动,二是始于1990年代的高校问责制(accountability)运动。SC改革要求高校从关注教师教学转移到关注学生学习,问责制提出高校要把学生学习效果作为评价学校工作的主要考核指标。到1990年代,大多数州都要求把学生学习效果当作学校绩效的主要证据,甚至把评估结果和政府拨款挂钩。这样一来,学生学习效果就成了评估的关键指标,学习效果评价就被提到前所未有的高度,而学习效果的认定与测量则成了评估的关键。

但是,大学生学习效果认定与测量几乎完全是一个新的领域。什么是大学生学习效果?应该包括哪些内容?如何兼顾大学生的四大发展(生理发展、心理发展、社会发展、职业发展)?如何兼顾短期、中期、长期效果?各种学习效果如何检验?如何把学习效果达成和评价活动整合到学生的整个大学学习过程中,合理有序地安排在通识教育、专业教育,甚至课堂教学中?如何组织、管理、领导、推动这一活动?如何和原有的评估活动(专业认证和学校认证等)联系起来,构成系统、形成整体?所有这些都是原来没有遇到过的新问题。人们知道应该做,但不知道如何做。于是,人们把这些与学习效果认定和测量的问题都归结为

大学生学习效果评价问题,而且用assessment这个词来表征这类活动。于是,评价(assessment)这个词就和“大学生学习效果”紧密联系起来,形成固定用法(outcome assessment),甚至固定为术语“大学生学习效果评价”(student learning outcomes assessment, SLOA)。也就是说,从词源角度看,这种搭配就是一种约定俗成,没有什么语义上的道理可寻。

还要注意,SLOA和专业及学校的有效性评估非常不同,SLOA主要与大学教学和教师发展有关,学校管理只是辅助性支持系统。这就使得SLOA的范围远超出认证评估的范围,直接延伸到大学教学和教师发展领域。由此可见,SLOA有两个维度:教学维度和管理维度,而且以教学维度为主。这样一来,SLOA就和如火如荼的SC改革直接联系起来,成了SC改革的一部分。

为了解决SLOA问题,在美国国家教育研究所(NIE)和美国高等教育学会(AAHE)主导下,于1985年秋举办了第一届全国高等教育评价研讨会(First National Conference on Assessment in Higher Education)。这个会议后来被认为是高等教育评价运动(assessment movement)诞生的标志。从1989年起,美国高等教育协会又开始组织每年一度的评价论坛(Assessment Forum),以此汇集专业力量。同年,第一份评价领域的专业研究刊物《评价快讯》(assessment update)发行。这些活动也得到了各州政府的支持,于是在各地出现了一些专门的评价研究机构,如印第安纳大学的评价研究所(Assessment Institute)、南卡罗莱纳州高等教育评价网(South Carolina Higher Education Assessment Network)、华盛顿评价小组(Washington Assessment Group)等。这些发展构成了所谓的“评价运动”。

评价运动在1990年代后进入快车道。据调查,1987年只有55%的高校建立了评价机构,到1993年这个数字已经达到98%。2000年美国至少有2000人专门从事评价工作,同时积累了大量可观的文献。1996年舒尔曼(Lee Shulman)继续接任博耶(E. Boyer)出任卡内基教育促进会主席,他和博耶一样,大力倡导大学教学学术研究,但他在博耶的“scholarship of teaching”的后面加了一个“learning”,成了“scholarship of teaching and learning”(SoTL),这个变化表明,博耶时期,学习还没有得到足够重视,但到了舒尔曼时期,学

习已经引起了重视。目前这个表达成了“大学教学学术”的标准表达方式。除此之外,他还发起了一个名为卡内基学院(Carnegie Academy)的大学教学研究项目,这个项目直接推动了 SLOA 的研究和实践。总之,到 21 世纪初,SLOA 已经变成了影响美国高等教育健康发展的重大关键问题。明确提出这个问题的是当时的美国教育部部长斯佩林斯(Margaret Spellings)。

2006 年美国教育部部长斯佩林斯领导的一个委员会发表了一个题为《考验领导力:描绘美国高等教育的未来》的报告。报告建议,对高等学校的问责和认证,要从过去注重投入要素(inputs),如校园面积、设施设备、教师人数、学校经费等转向注重产出(outcomes)和工作业绩(performance),如学生发展和学业增值(value-added)、研究与社会服务成绩等,要用平实易懂的方式向社会报告,并且还要求评估结果可以进行校际比较,以便学生、家长、政府官员、捐助者等利益相关者能做出选择。<sup>[6]</sup>这个报告直接导致了学校和专业的评估重心从组织有效性转向学生学习效果,进一步推动了 SLOA 运动。为了应对这个挑战,美国著名高等教育学者乔治·库(George Kuh)和伊利诺斯大学前校长艾肯伯雷(Stanley Ikenberry)联合创办了著名的全国学习效果评价研究所(National Institute for Learning Outcomes Assessment, NILOA)。NILOA 通过联络全国学术力量、发布新闻和研究报告、推荐最佳实践、举办学术交流活动等方式在全国 SLOA 研究与实践中发挥了领导作用。经过十年努力,该所已经成为了美国最重要的 SLOA 研究组织。<sup>[7]</sup>也就是说,经过 30 年发展,SLOA 成了 SC 改革中一个专门的实践和研究领域。<sup>②</sup>

在这股重视学习效果的热潮中,有人甚至提出了“以结果为导向的教育”(outcomes-based education, OBE)的口号。这个口号旨在强调学习效果的重要性。但它有两个问题,一是它太简单,忽视了 SC 改革的目的是和焦点。SC 改革的基本目的是促进学生发展,焦点是要把教学的中心从传统的“教师教”转变到“学生学”。因此不能只提学习效果。二是 SLOA 本身还面临一些根本性的学理和技术问题,因此也不宜将学习效果绝对化。<sup>③</sup>自 1987 年以来,SLOA 的研究基本上是围绕着四个问题展开:① 学生学习成果认定。即确定什么是学生学习成果?包括哪些内容?为什

么它们值得关注?② 评价的角度与方法,即学习成果可以从哪些角度测量?用什么方法检测和评价?③ 结果利用,即如何把评价结果用于教学改进和学校工作改进?④ 推广问题。如何让更多学校和老师接受 SLOA 从而全面推广 SLOA?前两个问题主要涉及教学维度,后两个问题主要涉及管理维度。下面分别介绍。

首先是学习成果认定,学习成果认定主要涉及两个问题,一是大学生学习成果应该包括哪些内容?二是各种学习成果应通过什么教育方式来实现?又如何评价?从 SLOA 提出之后,尤其是《斯佩林斯报告》发表后,在学术界引起了很大的担忧甚至抵制。学界的最大担忧是,SLOA 是个指挥棒。如果 SLOA 只关注学生的当前发展而忽视其长远发展,只关注就业需要而忽视基本素质,只关注有形需要而忽视精神发展,只关注可测量的学习效果而忽视尚不可测量的学习效果,这样的 SLOA 对学生发展、对大学教学、乃至对整个高等教育,都将是一场大灾难!这些问题在美国学术界引起了一场广泛的大讨论。仔细分析会发现,这些问题实际上与大学教育目标相关:即大学究竟要培养什么样的人。联邦政府与州政府“吹糠见米”的态度,确实让坚信大学应培养“伟大灵魂”的学校和学者担心。

在这场关于教育目标的讨论中,最被热议的两个话题恐怕要算通识教育(general education)和全人教育(holistic education)了,甚至引发了关于文理教育(liberal arts)与美国未来的大讨论。这就是这些年美国高教界会如此热烈地讨论通识教育、全人教育、大学使命等基本问题的主要原因。随着新世纪的到来,21 世纪人才需要和培养目标也成了讨论的一部分。经过这些讨论,美国高等教育界大体达成五点共识:① 在教育目标上,大学应该坚持兼顾学生长期发展和短期需要、基本素质培养和职业能力训练、社会能力发展与公民素质培养等教育理念。② 面对当代社会需求多样性和高等教育多样化,应该允许各校根据自己的实际情况和本校学生的实际需要来决定自己的教育使命和教育目标,不要“一刀切”。③ 所有教育目标都应当通过适当的课程设置和教育途径来落实,因此对所有课程和教学活动都要精心设计,合理安排,使之形成统一的教育教学体系,以确保所有教育目标的实现。④ 所有教育教学都应当关心学生发展、学生学习和学习成果,因此

在设计教育教学计划时,应当明确标示出预期的学生学习成果,以及将通过何种方式实现教学目标,并用什么方法来评价学生学习成果。务需把教育目标和学生学习效果有效联系起来,避免目标与效果的脱节现象。⑤ 学校有责任构建全校统一的教学质量保障体系,并且定期收集相关信息,以便随时监控学校和院系的教学情况。同时学校要为落实这些教育目标、课程体系、课程教学、评价方式等提供必要的物质资源和组织保障。凡此种种,就构成了美国在 SC 改革时代大学教学质量保障体系建设的思想基础。

关于效果评价的角度与手段,即从什么角度和用什么方法来检测和评价学生学习效果,这是从学理和技术角度思考 SLOA。下一篇文章将专门介绍美国关于“学习效果”概念的讨论和在 SC 改革中创造的各种课堂教学评价方式和方法。这里只想强调两点:

一是课堂教学评价的创新和发展,主要是由脑科学、认知科学、大学生发展研究和学习科学的进步引起的,这些进步揭示了一个传统大学教学几乎完全忽略的领域,即大学生是如何发展和学习的。过去大学老师和大学管理者几乎从来没有认真思考过这个问题。因为在传统教学模式中,教学主要被理解为知识传输,学习则主要被理解为知识记忆,学习效果评价的主要方式是考试,重点是检查学生记忆的完整性和准确性,或许还要加上基于演绎逻辑的有限的理解。但新的学术进步指出,学习是一个知识建构过程,在这个过程中,学生的大脑和智力、情感、态度都会发生变化。从知识建构的角度看,传统教学模式显然不利于大学生的学习和发展。在此背景下,“以学生为中心”的新教学模式应运而生。新模式认为教学要以学生学习为主,要能促进学生发展(包括生理发展、心理发展、社会能力发展和职业能力发展)。在教学中,学生主动学习的效果远远好于被动学习的效果,知识的有用性、情景的真实性、任务的挑战性、场所的社会性、过程的互动性等都有助于激发学生主动学习。在教学中,高度重视学生的知识建构、自律、活动、经验、情景、合作、积累、重复、个体差异等,都可以有效提高学习效果。活动产生经验、经验改变大脑。在教学中,要注意活动(activity)的重要性,要让学生动起来。有效的学习必然表现为学生行为的改变,因此学习效果测量的重点不是知识记忆,而是学生行为(performance)

变化。<sup>[8]</sup>所有这些研究,都对老师如何教、学生如何学、学习效果如何评价等提出了新的看法,指出了新的方向,也对 SLOA 构成了新的挑战。

二是由于相关研究和实践才刚刚开始,当前的 SLOA 研究还不成熟。无论是从学理角度还是从实践角度看,都还处于初级阶段,甚至尚未达到轮廓初现的程度。2018 年 NILOA 发布了一个十年总结,评价当前美国 SLOA 的发展状况:

(虽然 NILOA 已经做了大量工作,也取得了很多成就,)但在评价领域中,还有大量工作有待完成。对高等教育的价值以及接受高等教育是否值得的问题仍然存在,公众对这项高等教育信心的下降令人担忧。人们对学生的学习质量、高校应该如何最好地记录学生成绩、如何能用学生学习的证据来改进实践,并以此支持学生成功等问题,都还存在疑问……美国高等教育尚未能提供可操作的学习证据,用以指导学习和教学方法的改变,从而提高学生和学校的表现水平。美国大学尚未能以有意义的方式将控制成本的努力与学生学习成果联系起来。高技术含量的创新和新的高等教育提供者层出不穷,但它们大多数都缺乏足够的证据来证明其创新的有效性。已有的几种认证方式之所以能够幸存,仅仅是因为尚未发明其他可被接受的替代方法,能对教育的质量和学校的品质(integrity)做出让人理解的判断。<sup>[9]</sup>

这个判断有点令人泄气吧? 还有更严厉的! 美国著名高等教育学者、曾两度担任哈佛大学校长的博克(Derek Bok),他长期关注美国本科教育,从 2006 年至 2017 年写过三本讨论美国本科教育的著作。第一本是《回归大学之道》(Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More),直译应为《我们那些成绩不佳的本科高校:对学生学了多少以及为什么他们应该学得更多的坦率考察》。该书 2006 年出版,当时博克 70 岁左右;第二本是《大学的未来》(Higher Education in America),直译应为《美国的高等教育》。2013 年出版,此时博克已年近八旬。此书实际是一个业内老人对美国高等教育做的一个全面述评;第三本是《The Struggle to Reform Our College》,直译应为《为改革我们的本科教育而斗争》。2017 年出版,当时他已 85 岁。或许他没想到自己还能再写一本书,于是写了这本书。由此可见他心心念念的仍然是本科教育。

这三本书中有一个不变的批评是,美国大学未能很好测量与评价学生学习,因而阻碍了其他方面的改善和改革。<sup>[10]</sup>注意第三本书标题中的“Struggle”这个用词,这个词有“挣扎求生”的含义,我相信博克绝不是随便使用这个词的,这个词表明了他对美国本科教育的心态。我认同博克的判断。我甚至认为,在美国 SC 改革的所有领域中,SLOA 是最短的短板!

以上这些思考和判断足以使我们对当前 SLOA 的发展有一个清醒认识。然而仔细观察会发现,真正的障碍是,对大学学习的基础研究尚不能提供一个相对清晰和完整、足以支持对大学生学习效果评价进行合理预测和方法设计的理论框架。换言之,与教育技术领域中的情况一样<sup>[11]</sup>,基础研究的薄弱阻碍了应用与实践领域的进步。因此,当前的研究应当高度重视大学生学习领域的基础研究。没有这个领域的重大进展,SLOA 很难取得突破。笔者认为,所有人都要注意这个现实的局限性,不要对 SLOA 抱有过度乐观的态度和不切实际的希望。SLOA 在学理和技术方面,都还有很长的路要走。

三是评价结果利用。即如何把 SLOA 结果用于改进教学和学校工作。用于改进教学属于教学维度,用于改进学校工作属于管理维度。在美国大学中,这两项工作分别由两个部门负责,即教学支持中心和院校研究办公室。从逻辑上讲,前者是基础,后者是支持保障。因此,重点应当放在前者而非后者。但在实践中,这两个机构彼此平行,几乎没有交集。这对整体推进 SLOA 工作非常不利。此外,由于院校研究属于学校管理系统,因此“近水楼台先得月”,在资源和人员方面往往得到更多重视;而教学发展中心属于教学系统,远离行政中心,因此经常面临的情况是训练不够、资源不足和人手不足。<sup>[12]</sup>

不仅如此,由于教学和管理两者的关注重点不同,因此会提出不同的评价标准,而且这些标准还相互打架。例如,出于对问责制的回应,学校管理者最关心的是学生的学分完成数、毕业率、就业率等终极指标。如果把这些指标往下压,就可能迫使院系管理者也开始追求考试通过率、毕业率、就业率等,这就会对老师们产生压力,迫使他们不得不在课堂教学评价上“放水”。而如果老师们坚持学习效果标准,就可能影响考试通过率、毕业率和就业率,从而对院系和教师的绩效评价产生负

面影响。在 2016 年的一个调查中,就有学者指出,把大学教学质量和学位完成率联系起来是“误导”。<sup>[13]</sup>由此可见两种评价之间的争议之大。另一个问题是教学活动的多样性和统计数据的一致性的矛盾。随着 SC 改革的发展,大学的教学活动日益多样化,学习效果也呈现多样性的特点,但在数据收集方面,为便于统计却要求信息一致。这个实践多样性和管理统一性的矛盾如何解决也是一个关键问题。这个矛盾将直接影响信息系统的设计和收集。

这些冲突实际上提出了一个更为严肃的问题:SLOA 究竟应该是为教学还是为管理?显然这已经不仅是一个学理问题而是一个实际问题。无论如何,结论应该很清楚,如果老师们在教室里做不好,学生没有培养出来,管理层再怎么评价也没用,因此大学的 SLOA 应当重心下沉,更多关注教学维度,关心教师的教学改进和教室里的学生学习,而绝不可本末倒置。

第四是 SLOA 推广问题。这可能是个美国问题,不具有普遍性。由于美国的 SLOA 是由外部力量开始推动的,而且是自上而下的,因此很多学校和老师感到 SLOA 侵犯了他们的学术自由,于是出现了普遍的抵制和抱怨。直到 30 年后老师们认识到 SLOA 可以帮助自己改进教学,学校认识到 SLOA 可以帮助提高学校声誉之后,这种情况才有所改变。2018 年 NILOA 发布了第三份全国调查报告《评价很重要:记录真实学习的趋势》。这份报告指出,调查表明,已经有更多的学校和教师认识到 SLOA 的价值,开始接受并参与学校的 SLOA 工作,做好课堂教学评价。同时该报告也指出,SLOA 仍然面临很多阻力,还需要争取更多老师参与。<sup>[14]</sup>

以上是关于评价评估这两个概念的辨析,以及美国 SLOA 运动发展的简要介绍,希望能为读者理解美国的 SLOA 提供一个基本的背景介绍。

## 二、构建全校统一的本科教学质量保障体系(EQAS)

美国高等教育界在教育目标、课程体系、教学过程、评价方式等方面大体达成五点共识,要求学校为此提供必要的物质资源和组织保障。

这一节将具体说明美国高校是如何构建教学质量保障体系,以满足这五点共识要求的。先说明一点,本节介绍的体系只是一个理想型(ideal type)。关于理想型,马克斯·韦伯说过,提出理

想型的目的是为了显示研究对象的状态和逻辑,以便对其进行分析和讨论。理想型不指任何具体对象,也不是要显示其完美。理想型只是一个思维工具,用来帮助理解和研究对象。但理想型也不是虚构的,有大量真实存在的事实来保证其现实性和可靠性。就本文讨论而言,建议大家参考卡内基·梅隆大学(Carnegie Mellon University, CMU)和北科罗拉多大学(University of Northern Colorado, UNC),也可以参考 NILOA 在其网站上推荐的最佳实践案例学校,以及《高等教育测量、评价、评估手册》的有关章节。美国大学的 EQAS 系统是一个从学校使命到课堂教学构成的系统,由学校、学院、学系三级,乃至全校所有教师来共同建设和维护。此外,学校通常还设有教学支持中心和院校研究办公室来帮助支持和维护这个系统。下面具体说明。

### 1. 校级:学校使命与通识教育矩阵。

构建全校统一的教学质量保障体系,学校一级有两项任务。一是制定学校的教育使命,二是决定本校通识教育的学习内容和效果评价方式。

每个学校都会因其历史、传统、环境和学生的不同而选择不同的教育哲学、教育理想和教育目标。这些都应该准确地反映在学校的教育使命中,而且学校使命一定要让所有利益相关者都能清楚了解,以便大家形成合力,共同实现学校的使命。制定学校教育使命,是构建全校统一的教学质量保障体系的第一步,也是最为关键的一步。制定学校使命中最困难也是最重要的一点,是要明确学校的定位。就本科教学而言,就是要说清楚,我是一个什么样的学校,我要培养什么样的人。下面是卡内基·梅隆大学(CMU)的学校愿景和使命<sup>[15]</sup>:

“愿景:卡内基梅隆大学将通过在教育、研究、创新和创业方面的持续创新,对社会产生变革性的影响。

使命:为学生创造一种变革性(transformative)的教育体验,聚焦深入的学科专业知识;学会解决问题;获得领导能力、交流沟通和人际交往技能;享有个人的健康与福祉(well-being)。

培育一个变革型的大学社区,使其致力于(A)吸引和留住多样化的世界级人才;(B)创造一个合作的环境,使思想能自由交流,使研究、创造力、创新和创业精神得以蓬勃发展;(C)确保每个人能够充分发挥其潜力。

通过与传统大学校园边界之外的合作伙伴合作,以

一种变革的方式对本地、全国和全球社会产生影响。”

这样的愿景和使命表明了 CMU 是一所什么样的学校,它的重点发展体现在哪些方面,以及它将通过何种方式来实现其愿景和使命。特别要注意使命的第一段,这是它的教育使命。其中有 4 个分号,表明它的教育目标具体分为四块,一是深入的学科专业知识;二是问题解决能力;三是领导、沟通与人际交往能力;四是学生的身心健康。值得注意的是,CMU 特别强调培养学生的问题解决能力和领导、沟通、合作能力。作为一所以培养工程师为主的学校,这两种能力显然是其学生未来事业成功的基本保证。既然将其写入了学校使命,那么,学校所有专业计划和课程教学,都必须把这两种能力的培养和训练纳入其中。也就是说,这两点是以工程师培养为主的卡内基梅隆大学本科教育的基本特色。作为一个高度选择性的教育使命,体现了 CMU 作为工程师学校的特点及其对学生教育的承诺。

相比 CMU,以培养社会领袖和文理教育闻名的哈佛学院的使命就迥然不同,下面是哈佛学院的使命:

使命:哈佛学院的使命是为我们的社会培养公民和公民领袖。我们通过致力于文理教育(liberal arts and sciences education)的变革力量来做到这一点。

从课堂开始,学生们从接触到新思想、新理解方式和新认识方式,即开始一段心智转变的过程。通过一个多样化的生活环境,学生与学习不同主题、来自不同生活阶层、认同不断演变发展的人生活在一起,他们的心智转变不断深化,从而为社会转型创造了条件。由此出发,我们希望学生们开始了解他们想用自己的天赋和才能去做什么,评估他们的价值观和兴趣,学习如何才能最好地为世界服务,从而开始改变他们的生活。

愿景:哈佛学院将为 21 世纪的寄宿制文理教育确定标准。我们将致力于创造和持续创造条件,使所有哈佛学院学生都能体验到一个在心智、社会与个人方面的转变的无与伦比的教育历程。<sup>[16]</sup>

哈佛学院使命的第一段是教育目标,哈佛学院要通过文理教育培养公民领袖。第二段是教育过程,通过让学生接触新思想、新知识、新理解、新认识方式,通过寄宿制让学生接触到学科多样化、社会群体多样化和学习过程多样性,从而在这些多样化的学习生活过程中,重新认识自己,完成自

已在心智、社会理解和个人品质方面的转变,为学生未来参与社会发挥领导作用做好准备。第三段是愿景。哈佛学院要成为21世纪寄宿制文理教育的标杆。

在这个使命表述中,对哈佛学院是一个什么学校,它将如何培养学生做了明确说明。因此这些声明也都会彻底贯彻到它的所有专业、课程教学和学生事务之中。

由此可见,制定一个学校教育使命,对学校的教育目标、价值取向、教育过程、结果期望都做出了明确说明,这是构建学校教学质量保障体系的第一步,也是最重要的一步,因为它为学校的教育教学确定了目标、价值、过程和重点,使全校上下得以围绕学校教育使命共同努力。

校级工作的第二项是制定全校的通识教育(general education)计划。注意,这里的通识教育不是泛指的通识教育,而是某个具体学校的教育使命所决定的所有学生都应当接受的通识教育。因此,各个学校的通识教育会有所不同,以体现不同学校的教育使命。

关于通识教育,美国顶尖高校有两种不同模式。一种是哈佛大学为代表的全面通识教育模式,另一种是以哥伦比亚大学为代表的经典阅读模式。哈佛模式是在人类所有知识中选出一些特别重要的知识,然后由本领域优秀学者系统讲解给学生听。哈佛大学政治伦理学教授桑德尔讲授的《正义课》就是一个例子。<sup>[17]</sup>哥伦比亚大学的经典阅读模式是由哥伦比亚学院的老师们从人类历史长河中选出一些经典名著(章节),然后由老师带着学生一起阅读,让学生从中体会原著作者是如何思考的。也即是说,经典阅读模式的重点是思维训练(mind training)。由赫钦斯在芝加哥大学倡导的伟大名著(great books)阅读模式也属于这类模式。这两个模式对美国高校通识教育都有很大影响。哈佛模式认为,未来社会的领导者必须对所有人类知识有精到的认识和理解,才能在未来发展中领导社会前进,因此哈佛模式强调全面与精到。而哥伦比亚模式认为,历史表明,那些对人类社会有重大影响的伟大思想者通常并不具有全面而精到的知识,他们的伟大恰恰是因为他们独特的思维角度和思维方法,研究与学习其思维角度和思维方法更为重要。这是哥伦比亚模式强调思维训练、反对知识拼盘的主要原因。<sup>④</sup>在我看来,一个人要做出伟大贡献,或领导社会前进,

完备知识和思维训练同样重要,因此以培养未来社会领导者为己任的一流大学,其通识教育似乎应在这两种模式中找到某种平衡点,而不是有所偏废。不言而喻,不同的教育哲学确实会产生不同的通识教育模式,我们应当根据自己的条件和需要有所扬弃、有所改造、择善而从、优化创新。

对以大众教育为己任的多数高校,通识教育的基本目的就是帮助学生丰富知识,训练思维,为学生未来的职业发展、进入生活乃至终身学习和长远发展做好准备。由于这类学校通常也找不到那么多优秀学者来辅导学生做经典阅读。因此以丰富知识、开拓视野、兼顾思维训练的哈佛模式在这些学校里获得了广泛认同,成为大部分美国高校通识教育的主流模式。

显然,面对这些模式和思考,不同学校会根据自己的教育目标,设计本校的通识教育模式与教学方法。例如,CMU就把问题解决能力、领导合作交流能力特别挑了出来,作为本校教育的重点。总之,制定适用于所有学生的通识教育是学校一级的责任。

另外,新世纪以来科学技术飞快发展,知识更新速度加快,使得学会学习、掌握基本技能比知识积累更重要,这两点已成为培养学生长期发展能力的关键。于是,在美国就出现了通识教育应注重21世纪基本技能学习和训练的看法。2007年美国高校联合会(AAC&U)提出了美国高校毕业生应该具有四种基本知识和7种基本技能。四种基本知识是:①人类文化知识和自然界及物理世界的知识;②智力与实践技能;③对个人与社会的责任;④整合性学习(integrative learning)。7种基本技能是:写作、审辩思维(critical thinking)、定量推理、口头交流、跨文化技能、信息素养、伦理推理。2015年AAC&U又增加了4种基本技能:分析性推理、研究技能和做研究项目的的能力,跨专业学习与整合能力,把课堂知识用到课堂之外的能力,公民责任和公民参与能力。不仅如此,AAC&U还为这11种的学习效果测量开发了专门的评价量表(rubrics)。<sup>⑤</sup>

上述研究获得了广泛的认可,美国许多高校就围绕这些知识和技能,结合本校特点,开发自己的通识教育模块。显然,这些知识和技能都不是一门课可以完成的,需要所有课程合作,共同达成教育目标。于是有人便创造了一种叫课程地图(curriculum mapping)的方法<sup>[18]</sup>来协调所有通识

课程的培养目标、教学方式与和效果评价方式,以期达到最佳整体效果(见表1)。

在表1中,最上面一行是各门课程,左面是预期的学习效果,中间是考察方式和评价方法。比如课程1是欧洲文化导论,目标是培养学生的写作能力、审辩能力、口头交流能力、合作学习能力等。于是课程设计了期中论文、课堂讨论、小组报告、小组研究项目等作业方式和考核方法。如果学校想特别加强学生的写作能力训练、问题解决能力、合作交流能力,就会要求所有课程必须安排写作任务、研究项目和小组项目作业,以培养学生的这些能力。

通识教育效果通常不必进行跨校比较。但为了了解本校通识教育在全国的位置,也可以对通识教育效果进行可比较的评价。美国主要有两类可用于校际比较的通识教育效果评价方法。一是用AAC&U开发的11种技能评价量表。评价结果可用于校际比较。二是用美国教育考试服务中心(ETS)开发的专项考试(subject test),包括英语、音乐、商科、商管、经济、政治科学、心理学、社会学、生物学、化学、数学、物理学等18种。<sup>[19]</sup>这些考试类似于我们的大学外语考试和大学计算机考试,其成绩也可以进行校级比较。因此,有些学校规定,通识教育模块结束后,所有课程除了任课老师的学习评价外,还要参加AAC&U方式评价或ETS考试评价,以便所得结果可用于学校评估和专业评估。由于ETS考试很贵,会产生额外费用,因此采用ETS考试的学校较少。

通识教育由学校通识教育部负责。在设计通识教育矩阵时,一个重要的原则是,所有任课教师都要参加。集体讨论,共同设计出最佳课程组合、教学方式和学习效果评价方式。可以看出,学校通识教育目标对整个通识教育模块的课程设计、每门课程的教学设计及效果评价方式设计,都会有巨大影响。例如,CMU的通识教育模块,可能会要求所有课程都必须包含旨在培养学生问题解

决能力和领导交流合作能力的项目或作业。

表1实际上是一个矩阵,因此本文称之为“课程矩阵”。使用课程矩阵的好处是,可以在统一的教学目标下,建立起各门课程的协调关系,使各门课程一致发力,共同培养学生。这种方法对基本能力的系统训练尤为重要,任何基本能力都是不可能通过一门课程达成的,因此需要在不同课程中反复训练同一种能力,实现多样化重复。只有这样,这些能力才能深深改变学生大脑,成为他们的第二天性(second nature),真正变成他们的能力,达到改变学生行为的目的。

如果一个学校专业太多,各专业领域之间差异过大(如文科、工科、医科、农科),不易制定全校统一通识教育模块的话,也可以采用学部制模块,即每个学部自行制定自己的通识教育模块。学部制模块的好处是可以较好把通识教育模块和专业教育模块衔接起来。但要防止缩小学生知识视野,否则将不利于学生未来发展。

2. 学院:专业(program)和专业教育矩阵。<sup>⑥</sup>

教学质量保障体系的第二级是学院,学院要对本学院所有专业的质量负责。

美国的专业管理和美国的高校管理体制密切相关。美国高校管理的典型模式是“三级管理、院为基础”。在学校管理中,三级指校、院、系三级。三级管理是说学校一级主要是行政机构,有人事权和财权,但在学术事务上管理权限有限;学院一级是学术行政机构。作为行政机构,它独立的人事权和财权;作为学术机构,它有很大学术权力,决定本院的所有专业教育。学系是学术机构,是同一学科教师的共同体,它只有学术权力,没有独立的人事权和财权。因此在校院系三级中,学院是最重要的一级。

美国大学的学术管理主要依靠学院,学校越大越是如此。学校像中央政府,学院像地方诸侯。学校和学院之间的力量对比决定了不同学院的管理模式和风格。有的学校是学院强、学校

表 1 通识教育矩阵

|        | 政治科学     | 高等数学     | 经济学      | 无机化学     | 可比较评价方式  |
|--------|----------|----------|----------|----------|----------|
| 写作能力   | 期终论文     | 专题论文     | 期终论文     | 期终研究报告   | AAC&U 量表 |
| 审辩思维能力 | 课堂讨论     | 专题论文     | 课堂讨论     | 报告讨论     | AAC&U 量表 |
| 口头交流能力 | 小组报告     | 专题论文报告   | 小组报告     | 小组交流     | AAC&U 量表 |
| 定量分析能力 |          | 期终考试     |          | 统计分析     | AAC&U 量表 |
| 合作交流能力 | 小组研究项目   | 课程小组项目   | 小组研究项目   | 小组研究项目   | AAC&U 量表 |
| 问题解决能力 | 小组研究项目   | 小组研究项目   | 小组研究项目   | 小组研究项目   | AAC&U 量表 |
|        | ETS 专题考试 | ETS 专题考试 | ETS 专题考试 | ETS 专题考试 |          |

弱,如同邦联制;有的学校是学院弱、学校强,如同联邦制;如果两种情况都有,就形成混合体制。如在美国大学中,医学院、法学院、工学院、管理学院是传统的强势学院,这些学院在人事、财务、学术事务方面有很大的独立性,与学校的关系通常是邦联制。如 UCLA 的安德森管理学院甚至成了全校第一个不要政府经费、实行完全自主经营的学院。有的学院很弱,如神学院、社会工作学院等,这些学院严重依靠学校资源支持,因此它们和学校之间便形成联邦制。在美国大学中,只有少数学院由于历史原因,可以既高度依赖学校资源,又有很强独立性,如哈佛大学的哈佛学院、哥伦比亚大学的哥伦比亚学院、耶鲁大学的耶鲁学院,这些学院强势是因为它们就是大学的根基。总之,学院是美国大学的学术与行政管理的基本单位。

在教学上,专业(program)对学院来说就像自己的产品。学院有权决定自己的专业,也对自己的产品负责。学院要保证筹集到足够的经费和资源来支付所有专业的开支。因此,学院必须保证自己的产品足够好,能吸引到足够的学生,收到足够学费和经费来维持学院开支,因此学院的专业设置和专业质量对学院来说是性命攸关的大事。还要注意,如果由于某专业因招不到足够学生而被撤销,该专业的老师在原则上不受教师终身制保护。随着专业撤销,学院可以将他们辞退。因此专业质量的好坏直接关系到本专业所有教师的饭碗。专业好,则大家待遇高;专业差,则大家待遇低;专业撤销,则大家失业。<sup>⑦</sup>这个利害关系对本专业教师参与专业质量建设与维护的积极性有重大影响。

学院一级教学质量保障的重点是专业质量。为了确保专业质量,学院首先要确保专业培养的学生能满足社会需求,最好是毕业生不仅能顺利就业,而且还很有竞争力。竞争力标准有二,一是就业单位,二是起点工资。如果毕业生就业单位好,而且起点工资高,且一生难求,学院就会享有很高的社会声誉,优秀的老师和学生会纷纷慕名而来,还可以吸收到很多捐赠,使学院声誉与资源供求形成正循环,这样的学院当然就前途无量了。

因此学院在确定专业、制定专业计划、确保专业质量方面通常会下很大功夫。首先是专业设置,是否开办某个专业对学院来说是一个大决定,这意味着它要开一条新的生产线。新专业从开办到成熟通常需要十年,这也意味着一大笔投资。

为此,需要做全面深入的专业论证,尤其是在资源投入和产出方面要算细账。

开设新专业的原则是“谁提议,谁出钱”。政府要办政府出钱;学校要办学校出钱;学院要办学院出钱,大家要办大家出钱。办新专业首先要做市场分析,分析内容包括毕业生市场规模,毕业生质量标准、同一专业学生的市场供求以及其他学校的竞争情况、学生来源和学费支撑情况等。

就专业质量而言,最关键的是要确定专业质量标准,即该专业学生在知识、技能和职业伦理等方面的具体要求和标准。然后根据专业质量标准,组织相关教师共同制定本专业的专业矩阵(见表 2)。与通识教育矩阵类似,表 2 的最上面一行是该专业要开设的所有课程。左边是各类知识、技能和价值观明细。中间填写各门课程在各种知识、技能、价值观方面能做的贡献。

表 2 专业矩阵

|    |    | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 |
|----|----|----|----|----|----|----|----|----|----|
| 知识 | K1 | 3  |    | *  | *  | *  | *  |    |    |
|    | K2 | 2  | *  |    | *  |    | *  | *  | *  |
|    | K3 | 1  |    | *  |    |    |    |    |    |
|    | K4 |    |    |    |    | *  |    |    |    |
| 技能 | S1 | 2  | *  | *  | *  | *  | *  | *  | *  |
|    | S2 |    |    | *  | *  |    | *  |    |    |
|    | S3 |    | *  |    |    |    | *  |    | *  |
|    | S4 | 5  | *  | *  |    | *  | *  |    |    |
| 伦理 | V1 |    | *  |    | *  |    |    | *  |    |
|    | V2 | 2  |    |    |    |    | *  |    | *  |
|    | V3 | 2  | *  |    | *  | *  |    |    |    |
|    | V4 |    |    | *  |    |    |    | *  |    |

注: \* 表示贡献,数字可以表示程度级别或特定要求。

制定专业矩阵的过程有三步:第一步是确定专业质量标准。由院长指定专人或聘请专业公司做专业质量标准调查,调查中也会组织本院相关教师参与讨论,直到专业质量标准得到本院学术委员会认可。

第二步是制定专业矩阵。院长先把专业质量标准发给所有相关教师,由他们决定是否参与该专业教学。如果缺少核心课教师,就要进行招聘。

学院的收入是固定的,教师越多大家收入越低。因此,开办新专业原则上以依靠学院现有教师为主,尽可能避免招聘新教师,以提高现有资源利用率。美国大学教师皆有明确的工作量。通常研究型大学教师每学期上 2 门课,博士大学 3 门、本科院校 4 门,社区学院 5 门。<sup>⑧</sup>教师需要填满自

自己的工作量,在一般情况下,需要填满工作量的教师们都会积极参与新专业计划制定。如果一个教师的工作量总是不满,那他就应该考虑走人了。

填写专业矩阵时要注意三点:① 所有专业要求都要被适当覆盖,并能确保满足专业要求;② 对需要强化的专业技能,要保持有足够的训练,安排多门课程参与,采用多样化训练模式。例如对新闻专业,写作(S1)是基本技能,于是所有课程都要安排写作作业,以保证学生毕业时能有很好的写作能力;③ 要彼此协调,如果是电力专业,专业课(C6)是电磁场,前面的数学课就必须有足够的微分方程和复变函数内容,否则后面的电磁场就没法教了。如此等等。

采用矩阵方法的好处是在所有课程之间建立起合理关系,不仅要确保达到专业标准,还要尽量避免不必要的重复。因此,当教师们接到专业质量标准后,要提出自己的课程能为专业培养做哪些贡献,然后大家一起填写,经过反复讨论,形成专业矩阵。重要的是,通过讨论,所有任课教师都能弄清楚自己课程的教学目标是什么、应该如何教,以及与左邻右舍的关系。也就是说,教师的课程设计在专业设计阶段就已经开始了。显而易见,制定课程矩阵,教师全员参与是关键。

为了应对日后的专业评估,编制专业矩阵时要为所有专业质量标准制定相应的评价方式和方法,即所有专业质量标准,都要明确用什么方式评价,要达到何种水平等。在美国很多专业都有专业协会,这些协会通常下设一个教育委员会。该委员会负责制定本行本专业的教育标准。如果专业协会已有专业标准,就采用协会标准,如果没有,就要自己制定。这里要注意,制定的专业评价标准一定要所有参与教师达成共识,这一点极为重要。因为相关共识对确保专业教学整体质量,对指导每个教师的教学有直接影响。北科罗拉多大学的黄茂树校长告诉我,该校护理专业毕业生都要参加一个护士资格考试,通过才能就业。这个考试由美国护理行会举办,学院可以通过分析考试答卷来看哪些课程教得好,哪些课没教好,并以此决定教师的聘任和待遇。

第三步是指定一个专业协调人(program coordinator),负责协调和监督整个专业教学的执行。执行过程可能出现的所有问题,包括课程衔接、教学资源保障、质量监控等问题,都由专业协调人负责协调。因此最理想的专业协调人是那些

懂专业、善协调、会管理的人,其学术水平倒在其次。专业协调人由院长选择,对院长负责,院长对专业质量负责。专业协调人相当于产品经理;而院长则相当于分厂厂长。院长要对学院所有专业的质量负责,责任重大,需要对专业和对管理都有一定经验。因此美国大学的学院院长通常是公开招聘,即使是内聘也要经过公开招聘程序,以便使聘任的院长能得到全院教师的认可和支持。公开招聘的另一个好处是,可以通过引进新院长带入优质资源和新的经验,从而改善本院质量。

3. 学系:课程矩阵和课堂教学质量保障。

教师按专业组织在一起,就形成了学系。学系不是一级行政单位,没有独立的人事权和财权,把教师按专业组织起来主要是方便他们彼此切磋学术和相互帮助。学系是一个学术单位,一个按小专业组织起来的教师群体。如果某校把大专业放在学系(如数学系),那么,这个学系就相当于学院;而本文所说的学系相当于它的教研室。教研室是组织和研究教学的机构,这就是学院所辖学系的责任。换言之,学系的责任是支持教师教学和监督课堂教学质量。

当教师们从学院里领了课之后,就要设计和准备课程并实施教学。在《聚焦设计:实践与方法》一文中已介绍过美国大学课程教学设计方面的内容,这里不再赘述。其中最重要的是教师要根据专业要求和课程要求编制课程矩阵。见表 3。

表 3 课程矩阵

| 一般目标 | 具体目标 | 教法  | 学法  | 评价  |
|------|------|-----|-----|-----|
| A    | 描述   | T1  | L1  | O1  |
|      | 分解   | T2  | L2  | O2  |
|      | 鉴定   | T3  | L3  | O3  |
| B    | 确定   | T4  | L4  | O4  |
|      | 计算   | T5  | L5  | O5  |
|      | 组装   | T6  | L6  | O6  |
| C    | 检验   | T7  | L7  | O7  |
|      | 讨论   | T8  | L8  | O8  |
| D    | 撰写   | T9  | L9  | O9  |
|      | 演示   | T10 | L10 | O10 |

来源:赵炬明《聚焦设计:实践与方法》。<sup>[20]</sup>

由于所有与课堂教学有关的事务都集中在学系,从教室安排、课程安排、教学资源到教师教学评价、处理学生意见等,都是学系的事,因此在美国大学,系主任是一个“官小事多”的岗位,教师们大都不愿意当系主任。一些美国高校的应对之策是,系主任由教师轮流做,或者请年高事少的资深

教授来做,越是优秀的学系越是如此。反过来,一般高校中,系主任还是一个很诱人的岗位。

就课堂教学管理而言,系主任有两个重要职责。一个是教学支持,如教学安排、教学资源保障、帮助教师解决教学问题、处理学生学习问题等。另一个是给教师安排工作和做年度教师评价,系主任负责给教师安排课程以保证他们完成教学工作量,同时对教师的教学工作做年度评价。所以,课程教学质量保障是由系主任负责的。

4. 教学支持中心和院校研究办公室:培育学校教学质量文化和收集分析教学信息

教学支持中心的主要任务是,通过教师教学咨询和培训,推广新的教学模式和教学方法,也包括新的学习效果评价方法,从而协助构建全校统一的课堂教学评价模式和文化。

有的学校把这个机构叫“教学支持中心”,有的学校叫“教师发展中心”,或类似的名字。不论称谓如何,其核心功能只有一个,即为教师提供课程教学方面的咨询和培训,帮助教师提高教学水平,推动全校的教学改革。因此它既提供教学支持,也帮助教师发展。由于 SC 改革,美国大学教学研究人员和教师不断研究开发出大量新的教学模式和教学方法。而各校教学支持中心则通过教学咨询和教师培训,推广传播这些新知识和新经验,提高本学校教学质量。

帮助教师学习新的教学评价理念,掌握新的教学评价方法,对促进课堂层面的学习效果评价意义重大。近年来,教学评价越来越重心下沉,不仅从管理角度转向教学角度,而且强调课堂学习效果评价。这种背景下,教学支持中心开始得到重视。尤其是它在推广与创新先进课堂教学评价方法方面的工作,显得越来越重要。由于教学支持中心是在教学第一线推广和规范本校的课堂教学评价模式和方法,因而在构建全校统一课程教学评价体系方面的作用举足轻重。笔者在“研究与培训”部分将具体介绍这些方面的情况。

另一个机构是院校研究办公室。它属于学校行政机构的一部分,其基本任务是收集学校各方面信息,以支持学校管理决策。具体到教学质量保障,院校研究办公室的任务有二:一是系统收集和分析全校教学方面的信息,以支持本校教学方面的决策;二是在外部认证机构来做学校评估或专业评估时,代表学校做好信息支持方面的工作,提供相应的帮助。

这两个机构本应协同工作,共同协调教学维度和管理维度在信息收集和分析方面的冲突。但在实际工作中,这两个机构各自在自己的领域按自己的规范工作,很少彼此协调,甚至互不往来。这对建立统一的质量保障体系十分不利。这个问题应该引起学校注意。

以上是 EQAS 构建的基本情况,其中有四点值得注意:

第一是管理思想。教学质量保障工作说到底是一种质量管理。因此 EQAS 也需要有明确的管理指导思想。很明显,美国的这个体系深受系统管理和全面质量管理思想的影响。前者主要表现在对整个体系的系统设计和系统管理,使得整个系统各部分得以在统一目标下相互协调,统一发力。后者强调所有环节、全流程的全员参与式管理。对大学教学质量这样的复杂系统管理,采用全面质量管理模式确实是最好的管理模式。

第二是结构特点。如果问 EQAS 的特点是什么?最简单的回答是:一条主线,三级管理,具体表现为一个使命、三个矩阵、两个辅助机构。全面质量管理思想最主要的体现之一是矩阵法的广泛运用。三级质量管理的基本工具都是矩阵法(mapping),这说明矩阵法在美国大学质量管理中作用重大。矩阵法的优点是协调所有相关课程,使它们在教学目标、教学内容、教学方法和评价方法之间保持协调,以最大限度发挥教学的整体性作用。矩阵法协调,可以保证学习过程中必要的多样化重复,这对学生的知识掌握和基本技能训练大有好处。一条主线外,这个系统还包括教学支持中心和院校研究办公室这两个重要的辅助机构,前者在推广和普及统一的学习效果评价方式方法、建立全校统一的质量文化方面发挥作用;后者则在系统收集与分析学习效果信息,以支持学校教学决策方面发挥作用。两者分别在效果评价的教学维度和管理维度发挥作用。

第三是市场机制。这在学院对专业的管理方面表现尤为明显。还可以看到,美国的三级管理中,专业管理是重点。专业管好了,下面的课程教学管理就顺理成章了。而在学院的专业管理中,市场机制的作用明显。

第四是脱节问题。纵观整个系统,仍然可以发现一个潜在危险,即三个矩阵之间会不会出现脱节?尤其是通识教育与专业教育之间的脱节,以及专业计划与课程教学之间的脱节。仅从这个

体系看,如何控制这个脱节问题,似乎还没有很好的方法。实践上,这种脱节确实存在。总之,三大矩阵之间脱节是可能的,尤其是通识教育和专业教育之间的脱节。如何防止脱节,仍然是一个未能很好解决的问题。学校教学质量管理者对此不可掉以轻心。因为当人们专注于各自的工作领域时,忘记整体是常见的现象。

### 三、教学质量保障体系和教学评估

上一节介绍了 EQAS 的基本情况,本节讨论它与各类评估的关系。教学评估可以分校内评估和校外评估两类。先讨论校外评估。

美国大学的校外评估主要有两个,一是由各种认证机构发起的对学校教学质量的认证评估(institution accreditation),相当于我国的学校评估。二是由各个专业协会或组织发起对学校各个专业(program)进行的专业评估(program evaluation),相当于我国的专业评估。这两类评估都属于以合格评估为主,兼顾咨询评估。也就是说,主要是检查被评单位是否达标。不能达标的学校可能关门,不合格的专业可能停办,因此认证评估对学校 and 专业的命运有重大影响。如果达标合格了,认证专家通常还会根据在评估中发现的问题,提出一些改进性建议,但这已经属于专业咨询性质的活动了。因此对认证评估而言,合格审查远比专业咨询重要。在中国,为了全面发挥评估专家的作用,教育部高等教育评估中心区分了两类评估——合格评估和审核评估。对未经合格评估的学校,首先进行合格评估。对已经通过合格评估的学校,进行审核评估。审核评估的任务是为国家把关,为学校提供咨询。但在实践中,专业咨询的功能远大于合格审查。

在认证和评估中,外部专家要依靠学校提供数据,这是院校研究办公室的任务。院校研究办公室的数据只能来自学校的 EQAS。如果这个系统缺失,学校平时又没有收集有关数据,那就只能临时发动各部门收集整理数据。其结果一定会严重扰民,更为严重的是,临时拼凑的数据都不能为日常教学改进和管理决策服务。如果建立了 EQAS,平常就能系统收集数据,一旦评估开始,只需把数据整理出来交给评估专家即可。所以,院校研究办公室在教学决策支持和认证评估支持方面的工作高度依赖 EQAS。

首先,建立一个有效的 EQAS 需要三个条件:一是要有一个覆盖全校所有院系的教学信息

收集系统;二是要以统一方式填报信息,信息要符合三一原则,即一个信息、一个定义、一套操作,这样填报的信息才能对数据进行统计和分析;三是各院系要设专人负责填报,以保证填报信息的准确性。填报管理信息,满足这三个条件问题不大。但学生学习效果和效果评价的信息填报则可能碰到困难。各个学科的教师以不同的方式教学,用不同的评价方法,而且还在不断创新。如何解决信息管理统一性和教学实践活动多样性之间的矛盾,这是摆在 EQAS 建设面前的一大问题。到现在为止,我没有看到美国在这方面有什么妙法,这也为中国同行的创新留下了空间。

我的建议是采取分类处理的方法,即把所有信息按学校评估和专业评估分为两类。专业评估较多涉及教学过程,很多形成性评价信息至关重要,可以按专业或按课程建立分数据系统,各分系统按自己的教学和评价特点来构建,由各学院自行管理。评估专家只要到这些学院索取这些专业或课程系统的信息即可。由于评估者和被评估者都是同行专家,他们可以对这些信息做出专业评价,并不需要统一要求。至于学校评估,由于评估的重点是机构有效性,基本不涉及教学过程,因此可以提供终结性评价信息为主,如招生人数、学生完成学分数、毕业生人数等信息。如果评估专家要求深入到某些专业,他们可以到专业学院去查看专业数据系统的信息。这样,就可以缓解教学多样性和管理统一性的矛盾。尽管笔者建议不要试图建立包罗万象的大系统,但专业或课程分系统仍应有一致的结构,以便必要时可以收集同类信息做统计分析。这样构建 EQAS,显然需要教学支持中心和院校研究办公室之间的密切合作,前者从课堂教学角度提出问题,后者则从信息收集、统计与分析角度提供解决方案。

第二个重要问题是信息的收集和保存。无论何种系统,尤其是专业和课程系统,都需要以教师为主收集信息,而且并非所有信息都是电子信息,很可能包括其他类型的信息(如实物)。这时学院和学系的支持和保障作用就非常关键。如果教师缺少积极性,收集的实物无处保存,就不可能形成完整的专业和课程信息收集。尤其是当前,教学评价界认为应以多种方式记录学生学习效果,鼓励老师用档案袋方法(portfolio),非电子信息的收集和保存问题就显得非常突出。

第三个问题是规范化收集问题。即使是课堂

教学过程的信息,也存在信息收集的科学性和规范化问题。这时教学支持中心有责任为教师提供足够的学习效果评价方法指导,使得信息的收集更加科学和规范。学校研究所要在信息规范性方面为教学支持中心提供专业支持。

只有发展出有效的 EQAS 系统,且平时做好信息收集和保存工作,才能在认证和评估中做好信息服务。校外评估如此,校内评估也如此。EQAS 是一切教学评价评估的基础。

#### 四、六个重要问题和一个建议

以上简要介绍了美国 EQAS 的情况。本节讨论 EQAS 建设的六个重要问题,并给中国高校提一个建议。

第一是领导问题。建设这样一个教学质量保障系统,学校领导的决心至关重要,这是他们对本科教学的承诺是否兑现的一个检验。美国的经验表明,学校领导的作用是第一位的。要建设这样一个系统,需要对原有管理制度做很大改变,要投入很大的人力、财力和物力,还需要学校领导投入很大精力,甚至可能是数年持续不断的努力。凡此种种,都是考验。引用管理学的一句名言就是:领导是关键!

第二是正确对待评价技术的局限。目前关于大学学习和大学学习效果评价的理论和方法都还不成熟,还有很长的路要走,因此学校需要接受这个不完美的现实。我们只能按照已知的原理和已有的实践,努力朝所能设想的最好方向努力。持续探索,不断改进。要勇于实践,勇于探索,不要害怕失败。学校各级领导对此也要有正确态度:支持探索、宽容失败,惩罚不作为。

第三是实行系统管理和全面质量管理,坚持全员参与。EQAS 是基于质量的系统管理和全面管理,因此在整个体系建设中,要注意规划的系统性,把全面管理和全员参与作为指导思想。尤其是在通识教育矩阵和专业矩阵编制过程中,务必要求所有有关部门和所有任课教师都参加矩阵编制,使所有部门所有相关人员都了解自己的任务和职责,并有机会提出自己的看法和建议,能了解与左邻右舍的关系以便让彼此能及时协调,防止出现脱节现象。

第四是教师激励问题。美国的经验表明,在 SC 改革中,教师需要投入更多的时间和精力来从事教学改革。因此学校对教师付出要积极支持,对他们付出的额外时间和精力给予补偿,不要视

为理所当然。凡是漠视教师额外劳动的学校,SC 改革都不可能进行下去。学校领导应该认识到,没有广大教师的参与,SC 教学改革不可能成功。学校应该拿出相应的政策和资源,激励教师参加改革。传统教学模式已经存在多年,形成了巨大的组织惯性。SC 这样的改革,将是对组织惯性的巨大挑战。组织研究表明,如果一项活动只有不到 5% 的成员参加,属于启动时期;从 5% 到 60%,属于推广时期;当 60% 以上的成员都参加之后,组织就进入自动自发时期。由于组织惯性,启动期阻力最大,最需要额外激励和推动;推广时期次之,自动自发时期就已经不大需要额外激励和推动了。目前中国大多数高校的 SC 改革都处于启动时期,因此特别要注意提供额外激励和上层推动的问题。

第五是重视教学支持中心和院校研究办公室建设。美国经验表明,这两个辅助机构非常重要。尤其是教学支持中心,在整个改革期间,它将是所有新思想、新方法、新实践的发祥地。开展 SC 改革,首先应建立强大的教学支持中心,给予足够的政策支持和资源支持,把它变成 SC 改革的引爆点和动力站。要让所有参加 SC 改革的教师都能在这里得到支持和帮助、找到资源和知音、感受到决心和热情,从而坚定把 SC 改革进行到底。院校研究办公室主要是在信息收集与分析技术方面为教学支持中心提供帮助,同时负责教学信息的收集,为学校教学决策提供支持。没有这个重要的技术支撑,不可能建成强大有效的 EQAS。

第六是加强基础研究。与所有科学事业一样,没有扎实的基础研究,就不可能有强大的应用发展。SC 改革也是如此。在 SC 改革中,对大学生发展和大学生学习的基础研究的水平,直接推动或制约着新思想、新理念、新方法、新实践的产生和发展。由于 SC 改革是由新的学术进步引发的,因此基础研究对实践的影响巨大。目前基础研究落后已经严重约束了实践的发展,因此应加强对大学生发展和大学生学习方面的基础研究。

最后是给中国高校的一个建议:要学习不要复制(learn but not copy)。本文介绍美国大学的教学质量保障系统是因为它有几个明显优点,值得中国大学学习借鉴。首先它是在 SC 改革中诞生的,体现了 SC 改革“以学生发展为中心、以学生学习为中心、以学习效果为中心”的根本理念。在大学使命、通识教育、专业教育、课堂教学设计

等教学基本环节上,都力求体现“三中心”原则,而且已经把它们系统化和实践化,这一点难能可贵。二是这个系统体现了系统管理和全面质量管理思想。从质量管理角度看,像高校教学质量管理系统这种规模巨大、涉及面广、参与者众的系统,系统管理和全面质量管理是最好也最有效的质量管理方法。<sup>[21]</sup>三是美国大学创造的教学质量管理方法如矩阵法、“一条主线、三级管理;一个使命、三个矩阵、两个辅助”等,都是非常具有启发性和实用价值的方法;四是它正在美国大学教学质量管理中发挥积极作用。

笔者深知,美国大学和中国大学在社会环境、历史传统、制度体系、工作过程、工作方式、价值选择等方面存在巨大差别。因此,中国大学应当根据自己的情况,学习借鉴美国的 EQAS,用以支持自己的 SC 改革,而不是简单复制,原样照搬。学习之后,也可能采用类似构架,但这应该是建立在其真正适合本校实际基础上,而不是仅仅因为它在美国取得了成功。“要学习不要复制”。如果能在学习借鉴的基础上有所改进,有所创新,能为 EQAS 建设提供具有自身特色而又行之有效的中国架构,当然更是笔者所乐见的了。

## 注 释

- ① 美国大学评估与院校研究发展参考:赵炬明. 院校研究与现代大学,《高等工程教育研究》,2003 年第 2—3 期;超越评估(上)——中国高等教育质量保障体系建设之设想,《高等工程教育研究》,2008 年第 6 期。
- ② Ewell P. An Emerging Scholarship: A Brief History of Assessment[M]//Trudy W. Banta and Associates edited. Building a Scholarship of Assessment. Jossey-Bass, 2002:3-25; Ewell P, Cumming T. Introduction: History and Conceptual Basis of Assessment in Higher Education[M]//Cumming and Miller, D, edited. Enhancing Assessment in Higher Education. Styles Press, 2017: Chapter One; R. Shavelson. Measuring College Learning Responsibly. Stanford University Press, 2010.
- ③ 赵炬明. 聚焦设计:实践与方法(上),《高等工程教育研究》第 3 期第 31 页。关于 SLOA 面临的学理和技术困难,参考 Charles Secolsky & D. Brian Denison (ed.) Handbook on Measurement, Assessment, and Evaluation in Higher Education. Routledge.
- ④ 此节的撰写要特别感谢香港城市大学协理副校长、原哥伦比亚学院副院长程星,和他关于美国通识教育的讨论启发了我。
- ⑤ AAC&U 网站. Talking Points: AAC&U 2009 Member Survey Findings; Recent Trends in General Education Design,

Learning Outcomes, and Teaching Approaches, 2016; Trends in Learning Outcomes Assessment (2016)。

- ⑥ 此节要特别感谢北科罗拉多大学学术副校长黄茂树先生、该校蒙福德管理学院院长 John Alexandra 分享他们关于美国高校管理的知识和经验。这些知识和经验构成了本节的主要内容。在他们二人领导下,蒙福德管理学院 2006 年获得著名的博德利奇国家质量奖,成为获得该项美国质量大奖的少数高校之一。该奖是美国最高质量奖,每年颁发一次,分若干领域,教育领域每年只有一个机构获奖。
- ⑦ 美国大多数大学的教师聘任采用个人合同制,一人一个合同。因此好专业和差专业老师的工资相差一倍,是常有的事,同一学校不同学院老师的工资差两三倍也有可能。参见赵炬明:美国大学教师管理研究,《高等工程教育研究》,2011 年第 5 期。
- ⑧ 按卡内基分类标准,如果一所学校每年授予的博士学位超过 50 个,则属于研究型大学;不足 50 个者属于“博士大学”(doctoral institution)。

## 参 考 文 献

- [1] 赵炬明. 论新三中心:概念与历史——美国 SC 本科教学改革研究之一[J]. 高等工程教育研究, 2016(3):35-56.
- [2] 史海燕. 教育测量与评价[M]. 北京:北京师范大学出版社, 2016:2.
- [3] 黄涛. 学生学习成果评估——美国高等教育质量保障研究[M]. 北京:教育科学出版社, 2014:25,95.
- [4] Charles Secolsky, D Brian Denison. Handbook on Measurement, Assessment, and Evaluation in Higher Education [M]. New York:Routledge, 2012.
- [5] R Lance Hogan. The Historical Development of Program Evaluation: Exploring the Past and Present [J/OL]. On-line Journal of Workforce Education and Development, 2007(4).
- [6] Pellings commission. A Test of Leadership: Charting the Future of U. S. Higher Education [EB/OL]. <http://www.ed.gov/about/pubs/intro/index.html?src=gu>.
- [7] George D Kuh, Stanley O Ikenberry. NILOA at Ten: A Retrospective[R]. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA), 2018.
- [8] 赵炬明. 打开黑箱:学习与发展的科学基础[J]. 高等工程教育研究, 2017(3-4);赵炬明. 聚焦设计:实践与方法[J]. 高等工程教育研究, 2018(2-3).
- [9] George D Kuh, Stanley O Ikenberry. NILOA at Ten: A Retrospective[R]. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA), 2018:4-9.
- [10] Derek Bok. The Struggle to Reform Our Colleges [M]. Princeton University Press, 2017: 181.

- [11] 赵炬明. 助力学习: 学习环境与技术——美国“以学生为中心”的本科教学改革研究之四[J]. 高等工程教育研究, 2019(2):7-25.
- [12] Intentional Futures. Instructional Design in Higher Education: A Report on the Role, Workflow, and Experience of Instructional Designers[EB/OL]. <https://intentionalfutures.com/insights/portfolio/instructional-design/>; Chronicle of Higher Education. Instructional Designers in Higher Ed: Changing the Course of Next-generation Learning[EB/OL]. <http://results.chronicle.com/LP=1341?CIGNNGSLENSP?elqTrack=true>.
- [13] Rayane el se (Feb. 22, 2016), “Higher Ed Insights Results of the Fall 2015 Survey”, Inside Higher Eds 网站.
- [14] Jankowski N A, Timmer J D, Kinzie J, Kuh G D. Assessment that matters: Trending toward practices that document authentic student learning[R]. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). 2018.
- [15] <https://www.cmu.edu/about/mission.html>.
- [16] <https://college.harvard.edu/about/mission-and-vision>.
- [17] 桑德尔. 公正——该如何做是好? [EB/OL]. [2019-03-03]. <http://open.163.com/special/justice/>.
- [18] National Institute for Learning Outcomes Assessment. Mapping learning: A toolkit[R]. Urbana, IL: University of Illinois and Indiana University, 2018.
- [19] ETS, About the ETS? Major Field Tests[EB/OL]. <https://www.ets.org/mft/about/>.
- [20] 赵炬明. 聚焦设计: 实践与方法(上)——美国“以学生为中心”的本科教学改革研究之三[J]. 高等工程教育研究, 2018(2):30-44.
- [21] 布兰顿·戈弗雷. 全面质量管理[M]//约瑟夫·朱兰, 等. 朱兰质量手册(第五版). 北京: 中国人民大学出版社, 2003:396-435.

## Align with Learning Outcomes: Construct a Unified Educational Quality Assurance System

—Studies of the SC Undergraduate Education Reform in the USA (5)

Zhao Juming, Gao Xiaohui

**Abstract:** This article is the fifth in a series of studies on student-centered college education reform (SC reform) in the United States. This paper starts with a discussion of the meaning of the two professional terms—evaluation and assessment—and their customary usages. Due to historical reasons, in the American higher education, history, the term of “evaluation” mainly associate with the accreditation at institutional and program levels, while the term of “assessment” mainly associates with learning outcome appraise. This brief illustration is to provide readers a historical background of the two terms. Then the article introduce the framework and logic of educational quality assurance system (EQAS) in American colleges and universities. The basic characteristics of this system are as follows: one main line, three levels of control, and it is manifested as one mission (institutional education mission), three matrices (general education matrix, professional education matrix, instructional design matrix), two auxiliary units (instruction support center and institutional research office). The basic management ideas that run through the whole system are system management and total quality management. Then the article explains the relationship between the system and various evaluations inside and outside the institute. After that, six key issues related to the construction of the system are discussed. Finally, the author offers a suggestion to Chinese colleges and universities who are interesting to construct a similar educational quality assurance system.

**Key words:** The SC Undergraduate Education Reform; educational quality assurance system; undergraduate education; American Higher Education; China’s Higher Education (责任编辑 姜嘉乐)