

高教信息动态

(2018年第17期)

教学发展中心(高等教育研究所)

2018年12月19日

编者按：客观、准确对教师评估、评价是一项世界性难题，涉及到教师专业发展、履职尽责的问题，也涉及评估结果的应用、改进等系列问题。中外大学由于办学体制、制度及文化氛围不同，评估的标准差异较大，但从其政策制定、评估实施、自我评估开展及个性化的指导帮扶上依然有许多值得学习借鉴的意义。现摘编加州大学伯克利分校和麻省理工学院对教师评估的案例，供各位领导参阅。

世界顶尖大学如何进行教师评估

——伯克利和麻省理工的案例

组织治理的情况与学术质量的高低是否相关？我们深度观察了两所顶尖的研究型大学——加州大学伯克利分校(以下简称伯克利)和麻省理工学院(以下简称麻省理工)后发现，三个关键因素有助于产生持久稳定的高绩效：首要的便是自我评估和培

养才能；其次是教师的行为受到诸多价值观和规范的约束，例如群体成员身份、“人合性”（*affectio societatis*）承诺以及相互尊重和信任等；最后是复杂的内部组织程序能够有效发挥作用，使得整合和分化并行不悖。这些因素中的每一个都有助于以协同的方式产生很高的学术质量。

组织治理如何以及在多大程度上有助于大学的绩效？为了解决这个问题，笔者深入观察了无论是在美国国内还是在世界上排名一直都非常靠前的两所顶尖的研究型大学。

大学面临的两个挑战

大学依赖于教师，而教师是一个同时存在于两个舞台上——地方性的和世界性的特定职业。一方面，尽管教师长久以来都被认定是组织松散的机构（Weick 1976），但他们在一个特定的组织内发挥作用，并且处在享有一定权力的等级权威的管辖之下。教师还被要求去应用行政程序并且要对政策有所贡献，而这些不是由他们自己规定的。另一方面，教师属于一种职业，这涉及到与他们作为雇员或公务员工作于其中的机构之外的各种价值观、技能和标准。作为专业人士，对于生产什么以及如何交付，他们依赖于学院式的自我监管，这是基于“只有专业人士可以评估其他专业人士”的观念。教什么、出版什么以及如何教、如何出版的相关性、适当性和公平性，都是由内部或外部同行、网络或者是某个领域和学科的专业协会来确定的。无论是学生还是行政管理者都不能像教师那样享受同样程度的特权以及理解哪些知识是课堂或研究中所需要的。

大学在以质量驱动为基础的竞争性市场中运行，它们既不在标准化或匿名的生产驱动型企业中运行，也不提供事先就详细规定的商品或服务。外部利益相关者对于怎样使得以及为什么一些成果是有质量的而另一些则不然有着不同的期待，但他们掌握的信息却相对匮乏。大学身处市场之中并且相互竞争，而经济社会学中发展出的质量范式（Karpik 2010）对于更好地理解塑造这些市场的原则起到了很大帮助。大学在如何开展教育、如何实施研究方面比它们的用户——学生及其家庭、研究资助机构和公司等——控制着更多的信息，也因此，存在于决策、社会化以及控制教师的集体过程中的信任在使相关各方能够进行交流方面发挥着主要作用。这些程序的准确内容及其相对重要性因不同的学系、学科或者学校而异。教师质量的判断无论如何都是与某一学科的专业价值和规范有关，这些价值观和规范由各种方式控制，包括被指导机构所控制、被不同类型的团体——高校自身在这些团体中扮演着具有影响力的角色——所控制，以及归根结底是被雇佣教师的大学所控制。机构声誉是一个成功因素，它向用户承诺会交付出高质量的产品并与他们分享。

我们的目的是澄清和解释治理在多大程度上以及为什么能够支持两所不同的顶尖研究型大学的绩效产出，而在这两所学校中上述两个挑战显得尤为紧迫。文章的分析重点将放在内部社会过程在传递前沿学术研究时有助于保持高水平机构能力的方式上，它是一种以连贯、可靠的方式在校园中广泛传播的表现。为了评估潜在的社会过程、规范和价值观，要研究的不仅仅是决策被采纳的方式，同时还需要对组织治理进行深入研究（Thompson

& Tuden 1967)。紧要的是，在应对三个严峻挑战之时，教师和机构内部其他层面之间能够进行多大程度的协调努力，以减少——如果无法避免——组织错误的风险：

- 为单个成员（即学术专业人士）的自主性和创造性留出足够的空间，从而在需要交付集体产出的时候能够激励他们相互协调合作以生产知识。

- 在维持共同体身份和局部性公民身份的同时，允许他们产出非固定化的、非重复的结果。

- 能够使分化和整合在很大程度上相互兼容并且可持续。

两所研究型大学的田野调查

麻省理工和伯克利在国际上一直处于大学学术排名的最前列，二者在许多方面有所不同，但在剩余的其他方面却非常相似。

伯克利是公立的，而麻省理工是私立的，两者均由非学术性的董事会管理。麻省理工的校长和伯克利的校长均由监事会成员任命。麻省理工向其董事会报告，而伯克利则向加州州立大学系统负责行政和财务的主席报告。在这两所大学中，校长转而又分别任命一名执行副校长（副主席）和一名教务长。伯克利教师是加州公务员，他们的薪水和社会福利是根据州政府当局监管的规模来确定的。

在二者所覆盖的学术领域方面，两所大学都包含了广泛的科学领域。然而，麻省理工在法律或医疗职业教育与研究，或是在教育和公共卫生方面不太活跃，而伯克利除了医学院，其他学科则均有涉猎。在学术劳动和领域分工方面，麻省理工由 5 个学院、30 个学系和大约 300 个研究中心组成。伯克利则更加复杂，14

个学院被分为 130 个学系，然而却只有 82 个研究单位，远远少于麻省理工。

麻省理工和伯克利的收入模式也不同。作为私立学校，尽管麻省理工没有从州政府的补贴中获益，但它比伯克利更加富有，而州政府的补贴占到伯克利总预算的 12%。自 20 世纪 90 年代（那时州政府的补贴占伯克利总预算的 48.2%）以来，州政府的补贴急剧减少，使得伯克利的预算结构越来越接近于麻省理工。然而，麻省理工获得的捐赠基金（2012 年为 101.5 亿美元）是伯克利（30 亿美元）的三倍多，并且支出（22 亿美元 vs. 17 亿美元）比伯克利多出约 35%。就萨尔米（Salmi）的良性三角而言，虽然二者的资源都很丰富，但是麻省理工要比伯克利好得多。

在人力资源方面，两者的情况也都十分相似：2012 年，尽管伯克利国际教师的比例要高得多（31% vs. 麻省理工的 7%），但教师总人数（麻省理工 1805 人；伯克利 1736 人）则相差无几。非终身制教师的比例（麻省理工 26%；伯克利 24%）和各种类别的顶尖学术荣誉的数量（例如麻省理工共获得 33 个诺贝尔奖，而伯克利有 28 个）大致相同。最大的区别在于学生的数量和结构：麻省理工（10220 名）只是伯克利（32910 名）的三分之一。麻省理工专注于研究生教育（60%），外国学生集中（3200 人，占总数的 30%），而伯克利作为州立大学必须为大量的本科生服务（占总数的 60%），因此，招收的外国学生比例少得多（4926 人，占总数的 13%）。伯克利的本科生比例和相对较低的收入也部分解释了他们的高师生比（19: 1 vs. 麻省理工的 6:1）和

较低的学生人均支出（51656 美元 vs. 麻省理工的 216007 美元）。

这些数据挑战了只有私立研究型大学才能迈向顶峰的观点。尽管与麻省理工相比，伯克利的资源更少、任务约束更多，但是它始终保持着在全球的声誉。

伯克利这种制度化了的共治原则是美国大学运行模式的一个体现。相当大的权威和权力被正式分配给教授会，这是由教师全体选举产生的代表机构，它对各种各样的政策和项目——无论是严格意义上的学术性的还是非学术型性的——都享有发言权。它的意见来自各个层级，从学系到学院再到学校，甚至是掌握管辖权的州一级。教师是具有加州政府雇员地位的专业人士，他们在一个公共的和高度层级化的系统内工作。同时，他们共同管理整个加州系统的关键学术决策。

麻省理工既没有教师协会也没有教授会。教师的权力存在于一个相当独特的治理结构中，该结构由一个强大的委员会系统和十一个常设教师委员会所支持，它们是由同行、教师、学生、行政管理者以及专业人员选举产生的。教师政策委员会的当选主席既是教师代表，同时也是由学校高级领导组成的学术委员会的成员之一并享有表决权。该主席有权查看人事方面的保密数据、在所有的晋升和终身教职决策中投票以及查看机密的财务信息和其他信息。

从对所收集的有关麻省理工和伯克利的材料分析中得出的一个主要结论是，就组织治理和学术质量的生产二者之间的关系

而言，两所研究型大学的功能实际上是相同的。两者进行动员的社会过程和发挥作用的方式都创造出一种具有三个特征的模式：

- 独特性和才能是评估学术质量的最终标准。评估是一项常规工作，它关乎一个特殊的问题：学术质量。才能被认为是质量的支柱。

- 文化价值观和社会控制过程建立起了强大的局部群体认同感。

- 整合和分化的互动过程在组织和专业之间以及在机构的不同单位和层级之间保持着高水平的兼容性。

作为制度优先的学术质量评估

从人力资源的角度来看，大学有许多选项可供选择以应对各自的教师群体。在频谱的一端，质量的生产即使不被有意忽略那也可能被疏忽，官僚程序被用来控制承诺和绩效；在频谱的另一方面，学术质量被提到首要位置，个人报告会被他人阅读、绩效也会被详细评估，这些都是根据大学团体所共享的各种规范来进行的。绩效评估尤其是个人的研究质量被认为不仅对其个人而且对整个本地共同体都是十分重要的。这是管理学术资源的主要标准，也是在伯克利和麻省理工的校园里所发生的。

评估的首要性

在观察伯克利和麻省理工如何评估学术质量时，有两个事实是令人吃惊的：这是一个由教师进行自我评估的连续过程。

评估在个人的整个学术生涯中一直持续着，它似乎永远都不会结束。从被招进学校到退休离开学校，无论是初级教授还是资深教授，终其一生都需按要求提交报告。对于研究中心、学系、

学院这些次级单位来说，情况也是一样的，它们的绩效每 6-7 年就要被评估一次，而且当它们面临困境之时，评估甚至会更加频繁。一个在本学科的全 国排名中连续 30 多年排名最高的学系在 18 年的时间里被评估了 4 次。一位具有广泛国际认可度的教授发现，在他作为正教授积极服务的 30 年里，他写了八份关于自己学术产出的详尽报告。通常在 被招为新人后的 6 至 8 年内，一名助理教授必须提交至少两份中期报告，然后才会被考虑是否能够晋升为终身副教授。即使是大学教授——极少数被精心挑选出的正教授才可能达到的最高位置——每 2-3 年也得撰写评估报告。报告必须要详细列出教学、研究和服务方面的成就，以及未来几年的研究项目。似乎没有教师或者学术子单位敢略过甚至轻视这一要求，被评估已成为日常学术生活中不可分割的一部分。评估是一件严肃的事情，而不仅仅是官僚式的形式主义或工作负担，评估备忘录和评估建议都会被写得非常专业。从公平的角度来说，在伯克利和麻省理工，没有人真正表示出对评估的原则、过程和结果的不满，质疑评估会被认为是不合适、不友好甚至是自大的。

贯穿整个大学的评估活动是连续的。许多委员会不停地开会，许多学者被动员起来去担任从学系一级到学校一级等各个级别的评估者。花在评估任务上的时间对一个外部观察者来说是惊人的。从第一年的 9 月初到来年的 5 月末，正教授们可能要花去总计大约三个整月的时间来阅读报告并撰写评估备忘录。参加委员会会议是一项被授予的重要职责。教师作为评估员对他人进行评估是很正常的，甚至在副教授一级，每年也要在 4-6 个不同的

评估案例中担任评估员。对同侪的学术能力做出判断是正常工作生活的一部分。报告会由学校不同级别的不同委员会进行检查和复核，从学系层面到教师委员会或教授会层面，以及到做出最后决定的教务长和校长层面。同时，他们还会在晋升到更高的职业地位或是延缓晋升方面发挥作用。当与晋升无关时，他们可以影响经济福利——更高的薪资水平、更好的养老金水平、获得讲席教授资格——以及工作条件——获得设备、聘请助理的预算、旅行费用、暑期工资等。对于学系来说，评估对于学校提供的财务支持和教学时段等参数很重要。

评估意味着自我评估。伯克利的教师由伯克利的教师评估，麻省理工的教师也是由他们的同事评估。虽然外部人士也参与到自我评估中并且享有发言权，但却没有最终决定权。评估过程由学校内部人员主导，它涉及到不同层次上的许多人——没有任何一个人甚至单一委员会可以垄断这一任务的整个过程。

在伯克利，学系从它某一个专业领域开始主动提出评估需求。例如，子单位需要为一个新职位招人而提出评估需求，如果学校领导同意，就会设立一个候选人名单，包括自发申请者以及由学系成员提名者，因为这些成员通过他们的学科网络早已知晓了目前学术市场上有哪些优秀后生。候选人的各项记录和人品会被仔细地审查，因为各学系一再坚持宁缺勿滥，他们希望招募到“最好的”。

因此，他们试图保持其招聘简介足够全面，以吸引任何预料之外的在其他子领域很有能力的候选人，而不是最初瞄准好的某个人，这样才不会错过招聘到一个杰出学者的机会。一份简短的

潜在候选人名单会被列出，这些人将被邀请在学校参观几日，面对面地接触教师，并根据他们正在进行的学术工作举办研讨会或讲座。负责招聘的子单位接着会从中选择两三个名字，并在系里的全体教师会议上提出。经过投票选出最终的一个名字后，系主任就会与他/她所要报告的学院院长联系，以便在一份非常详细的报告和详略适当的论证支持下提出选择意见。到了学院层面，院长和一个专门的学术委员会将会重新审查该报告，并对系里的建议提出意见。如果没有人否决，那么系里的建议报告连同院长自己的保密意见都将送到学校一级的教授会，再由来自各学科的学者在教授会中设立一个专门委员会并起草一份报告。整份文件接下来将被提交给学校一级的执行副校长和教务长，他们同样也是专业学者。虽然从理论上讲，学校各层级可能不会同意最初的建议人选，但证据表明，只有不到3%的情况会做出否定决定并推翻最初的建议。从最底层——一个学系的某一专业领域——到最高一级的学校管理者，花费的所有时间都旨在确保在形成一个有足够说服力的评估案例中所有人都已非常谨慎并且综合考量了所考虑的人选。

在麻省理工情况亦然，尽管有一些形式上的差异，但是同样的方法也被用于招聘新人、评估正在被考虑是否应晋升的教师以及评估单个研究或教育单位及其项目的学术质量。在很大程度上，学校内部的教师控制着整个评估程序，从最开始一直到最后的结束。然而，在麻省理工和伯克利，外部参与者也确实为评估做出了贡献，比如，由本国或者国际同行撰写的推荐信通常会被附加到评估文件中；为了评估学系或研究单位，包含外部学术成

员的临时委员会会被建立起来。大多数时候他们都被认为是必须的一一处理一个评估案例不能没有外部意见的参与。然而，外部判断只是作为附加信息来源被对待的。谁应该参与评估以及他们的判断在多大程度上影响着评估和自由裁决的内容，仍掌握在每个评估步骤的内部评估者手中。一般来说，观察表明只有在少数情况下，外部参与和判断确实会对结果产生影响，这可能发生在外部人员被用来干苦活儿的情况下，以摆脱那些不被认可但由于某些原因又难以淘汰的候选人。在任何情况下，最终裁决还是由内部人士的判断决定。分包给第三方依旧是例外，它被认为是违反传统的，是作为一个戏剧性的事件表示出对学校内部那些有影响力的教师和学术委员会学识的不信任，甚至被当作是在评估承载了过多内部张力时的一次绝望尝试。观察结果显示没有出现被评人十分自满的特别离谱的案例。发生在局部共同体内部并且由共同体进行的自我评估，虽然不能免受情绪的影响，但也没有打开通往不正常偏见的大门，比如政治博弈、基于友谊的偏袒或者是失控的专业赞助等。比如说，如果在招聘中出现这种“失败”，它将在中期阶段时因为其后果例如在终身教职方面以及学系的地位而变得十分明显，从而导致严重地失去自信。各学系一直坚持要求不能对招聘或晋升评估结果处理不当。

内生质量标准

当需要晋升和招聘时，学术质量如何确定？科学业绩建立在什么样的标准之上？在这里，伯克利和麻省理工之间的相似之处也很突出。

除了少数例外，个人学术成果（出版著作的数量、科研论文的数量、专利等）的定量指标在评估教师个人的绩效方面并不发挥重要作用。基于发表数量的指标在硬科学中比在软科学中会被更多地考虑，但尽管如此，它们不会产生影响。事实上，观察结果一致显示，没有指标会将发表数量与质量相关联。一方面，报告那些低质量的科学出版物只会引起他人的非难并且根本无济于事；而另一方面，堆积如林的出版物也不会被认为是培养质量声誉的实际方式。没有程序规定在某一学科中要完成所要求的出版数量，这也不是评估者认为被评人需要遵守的标准；也没有正式的清单规定哪些期刊是有价值的、哪些期刊在评估时应被丢弃。当然，一个学科内的每个人都知道哪些期刊是最好的以及哪些是自己研究领域的顶尖出版者。然而，实际上并没有指标明确表示需要多少得分点来满足评估者的要求。除了出版物的数量，其他定量标准也可能出现在报告中，例如研究被引指数或学生对教学活动满意度的问卷等。这里再一次强调，这些指标需要限制性地谨慎使用。当评估者认为被评人确实不寻常以及当他们教学得分很低或是被引指数很高时，他们会吸引一些注意力，但这些指标得分只是在众多信息中提供了额外的信息来源，并且被认为只是解释和判断的问题而已。质量或价值评估不是以简单的数豆子的形式进行的，适当的量化指标是伯克利和麻省理工自我评估中必不可少的一部分，它使得内部评估者减少了对外部人员的依赖，因为大多数定量标准都是由第三方通过某种方式所确定的。从所设定的质量角度来看科学期刊排名的话，所谓的“A期刊”被定为最高级别，它们代表了被整个学术界、国家或国际协会正

式或非正式地采用的规范和标准。无论如何，一所大学在选择哪些期刊比其他期刊更好或更有声望从而确定质量标准时，并不会想要去控制或发挥主导作用。

在麻省理工和伯克利，定性判断占据着上风。伯克利和麻省理工的自我评估在很大程度上由基于定性因素的判断所组成。内部人员自己使用的绩效标准通常是相当隐蔽或难以捉摸的，他们依靠共享的直觉、经验和相互信任。例如，文学和科学学部的某一学系在美国排名非常靠前，并且得到了全校的尊重。它的全职教师由 50 名成员组成，而且招聘的人才大多数都是美国最顶尖的博士生。它涵盖了很多领域和专门化的子领域，并且在其中的几个领域中，它在全国范围内根据研究方向确定了基调，这些研究方向主要是在个人基础上进行的。它的规模和分散性可能乍一看不利于组织凝聚力、形成共同偏好或紧密合作，特别是出于学术人力资源的目的，但事实上，招聘和正在进行的教师评估是以一种相当统一和集体的方式处理的。推荐信被忽略，因为它们只标示着科学上和社会上的联系而不是真正的评估内容，而真正的评估内容才是实际上最重要的事情。著作——正教授或讲席教授 3-4 本、助理教授的第一本有望出版的手稿——为评估质量和绩效提供了基本的参考，特别是在人文社科领域。发表数量并不是支配性的标准。就晋升而言，个人在科学网络中的位置——通过加入学术团体、参加会议、得奖等而受到关注——也会被考虑进去。晋升副教授的职位时要求候选人被认可为国家科学界的一员；晋升全职教授则必须已获得崇高的国际声誉。教学和机构贡

献不会产生很大的差别，尽管每个人都接受如下观点：从授课或是提供某些服务的角度来说，必须满足他人最基本的期待。

共同的标准被局部地命名为原创性，它确实在影响质量判断方面发挥着作用（Lamont，2009）。无论在什么领域和相关子领域，原创性指的是在更广泛的学科共同体的主流中挑选出一部分作品和出版物，这些作品和出版物因其所涵盖的主题或所使用的方法而具有原创性。从常规科学的角度思考是才能的标志，并且它不仅是众多其他质量标准中的一个。除了探索新的途径并获得国际学术界的认可之外，原创性意味着研究应该对子学科的范式议程产生影响。“常规科学”综合症通常被认为是风险的主要来源。例如，一些学者可能在未来利用他们过去已经涉及到的领域、理论和方法，即使他们在评价很高的期刊上发表自己下一个成果的可能性仍然很高，但这是不够的，甚至不仅被视为是特定学术的潜在威胁，而且还被视为是整个机构声誉的潜在威胁。为了避免科学附加值下降的风险，伯克利和麻省理工都放弃了基于科学开发的学术剖析，而赞成通过探索取得成就的学术创新。学系不赞成按照一般研究议程进行的学术工作，为了表示其迈入科学界上层的雄心，它采用了与本国或全球顶尖部门相类似的描述评估方式，并与其进行比较、竞争。

定性判断为何要求如此多而全面的谨慎关怀，这个问题并非不重要。虽然它可能被解释为过去的遗产、一种当地传统和微文化的表达或一种被广泛分享的财产，一些精英院校也借此来显示他们的卓越，但在伯克利和麻省理工，定性判断满足了特定的实用功能需求。虽然质量标准被用于回顾性的绩效评价，但它们也

用显著有效的方式提供持续的评估，以预测未来能从一名教师身上产出什么预期的附加价值。过去由外部标准评估的绩效结果是不够的，主要原因是：现在处于知识生产的最前沿并不意味着将来还会是这样。

对原创性的追求产生了顶级知识生产者之间的制度差异动态。在人文和社会科学中，原创性仍然是一个相当难以捉摸的使用标准；在前沿研究的概念仍是首选的自然科学领域，它并不是可以自给自足的。这些概念是相似的，因为它们都呼吁引进有点不同寻常并且不符合现有类别的人，各个院校在雇用他们做新老师的时候，在其身上都下了很大赌注。在这两种情况下，系里准备下注的是对学者及本部门附加值下降的保障。这在追求较低风险研究的人中显然是不存在的，他们也因此不能维持住使他/她自己的研究以及他/她的学系与一般科学之间产生不同的东西。

假若能够证明其合理性和彻底性并说服负责评估的人员，创新就会得到很多的关注。对于申请初级教师职位的人来说，过去的成就如博士论文和由其衍生出的一些出版物，被认为是质量检验的第一关。然而，真正产生影响的却是现在及未来几个月中正在进行或将要考虑发起的新研究项目。在与招聘人员会面并被介绍给全体教职员工时，申请者的相关信息是从附在申请文件后的纸质版资料以及口头报告中收集的。只要符合系里所主张的学术特色或风格，创新的潜力就会产生影响，至于其是否与现有的子领域、教授或研究单位已经做过的研究相匹配，都是无关紧要的。作为学术共同体的学系必须有意愿支持创新，并将其与自己对于独特性的定义相融合，这是考虑到它在整个专业科学议程的未来

发展中可能达到的高度并将在局部扮演一个主要角色。从这个角度来看，在招聘新人或授予终身教职时，定性标准被认为是在不确定的学术环境中减少风险的最佳方案。

保持创新潜力的一个主要问题是要去刺激那些到了职业生涯中期的终身制老师勇于改变方向，并保持敢于冒险的心态。竞争性的资金有助于教师朝该方向努力，但会存在使那些尚未被广泛探索的领域变得荒芜的风险。尽管事实上它很难做而且需要一个人对风险承担能力充满信心，但当杰出的研究人员开发出自己的方法来决定什么时候他们应该变更项目时，这种情况就会发生。比如，其中一种方法是创建一个新的启动项目并做出决定：一旦它开始在给定主题上比研究小组做得更好时，后者必须做出改变。然而，帮助处于职业生涯中期的老师——他们收入丰厚并且从良好的工作保障中受益，但可能有点疲惫——进行创新的一种更普遍的方式就是及时以定期审查为契机，去评估他们做了什么以及还能做些什么。定期审查还重申了对每位初级教师和资深教师的期望。

作为制度对话的评估

质量的本质在于被称为“对话”的社会过程中所阐述的判断。伯克利和麻省理工的自我评估都发挥着重要的潜在作用。它提供了制度对话的舞台和机会，以促进学术质量的所有权，并将其作为组织各个层级及其各个部分共同的优先事项。如果仅仅把它理解为在招聘、评估和教师晋升时的一种方式或管理工具，就将忽视掉镜子的另一面：这个过程在参与评估的大学内部的许多组成部分之间产生了凝聚力。

制度对话发生在学系一级。建立在构成学术成就质量和高绩效之上的共识并不是显而易见的。各个子领域和子学科是共存的。教师携带着不同的标准以及关于全面性、相关性和质量的定义，全体一致和达成共识不是自发的。对话过程有助于开诚布公，以明确地进行辩论、分享不同的观点，从而纠正偏见、防止互相逃避或情绪冲突。伯克利和麻省理工的学术自我评估涉及一系列激烈的辩论和互动，这些都是有关想应聘和晋升的申请人的实际情况的。这些任务需要全体会议讨论、小组委员会审查和报告以及系内的集体投票，每个评估案例本质上都是被当作针对特定受众来处理的。他们还要求在将推荐意见上报给学院和学校、评议会以及学术行政人员时，要附上清楚明确的论证内容。

在伯克利的人文社科学系里，教师间的密集讨论和频繁交流能够在从宽泛的学科共同体主流中挑选出原创性作品方面产生共同愿景。事实上，在被观察的学系里，所有教师——无论他们的子领域是什么——都需要阅读即将接受评估的老师的几篇出版物。尽管评估处于首要地位，但它也只是学科共同体成员要完成的诸多任务之一。13个常设的内部委员会涵盖了所有感兴趣的话题。主席任命六名副主席并与他/她自己一起组成执行委员会，六位副主席接着又任命他们所领导的各专门委员会的成员。此外，教师要参加每月举行的学校一级的会议，这些会议一般会有大约一半的常驻教师参加。在任何给定的时间，无论其职务大小，四分之三的老师都要列席参加至少一个内部委员会，而且每几年就轮换一次。然而，资源和决策的公有化并不能消除系里的尊卑秩序，相反，它只是使之更加委婉。虽然正式的民主程序不

会给予地位高的教师任何正式特权，然而所有讲席教授都会列席一个或几个内部委员会，并对决策产生很大的影响。他们对系里的各种环境比资历较浅的同事更加熟悉。虽然资深教授的道德权威依然强大，但年轻教授的观点不会被忽视，这样的相互宽容有力地促进了共同价值观的建立，比如可以更加具体地了解原创性的意义。

在自然科学系的走廊里，贴出来的海报上写着一周中每天来访问的外部同事的情况。这些同事已被邀请参加每一个小组研讨会，以讨论新成果和持续的研究进展。每个教授都有一个涉及某一特定子学科的研究小组，它由十到二十名本科毕业生和研究生组成。他/她自己通常花大量时间在校外参加会议、担任顾问等。虽然任何小组的工作可能与该系的其他小组的工作没有明显的共同科学特性，但大多数教师实际上每周至少参加两次小组研讨会。当然，这些研讨会为科学交流提供了机会，它们也构成了制度对话的平台。它们为潜在的社会过程提供了动力，例如互相共享信息、关注内部其他同事所做的事情、共同筛选和理解国家与国际科学界的发展，以及建立起对新出现的社会问题的认识，这些问题可能代表着资助机构、基金会或公司提供捐赠或慈善事业的机会。这些研讨会实际上是制度对话的机会，在此背景下部门语言被共同构建起来，这也澄清了关于进一步招聘目的和研究议程的学术质量标准的意见。整个校园都会出现同样性质的其他机会，例如，同一学系的资深甚至初级教师在教师俱乐部的一个公共桌上每周共进两次午餐，甚至一周三次，在那里他们不仅分享

朋友的世俗八卦，还会就其所在系的共同科学利益问题进行辩论、交换信息并提出解决方案。

制度对话过程在学科和部门之间以及部门和学校顶层之间都非常活跃。质量标准随不同学科而异，学科甚至子学科的认识论状况并不完全相同，它们的科学环境不会以相同的速度和相似的科学议程发展。鉴于其学科多样性，在研究型大学中分享对“质量意味着、暗示着什么”的共同理解是一项真正的挑战。要避免的错误是将其定义为由被分割的一个个学科组成部分的简单相加。为了在不仅仅将其定义委托给各个学科的情况下达成对这种质量多样性的共识愿景，需要进行相当密集的制度对话以克服隔阂和相互间的无知。它需要开放性来理解与某一教师所在学科不同的其他学科所支持的学术质量的替代意义和隐含意义。

一个系要做出赞成某一候选人并给予其特定职位的结论和论证，如果没有从头开始进行重新分析，那也至少要由学院和学校一级的专门委员会和学术行政管理者完全自主地从专业上进行重新审查。这些委员会的成员不仅仅是该职位所在学科或学系的学者，外部人员——即不属于同一单位或领域的同事也会参与到该过程中。参与是秘密的，混合学科背景和单位制是惯常的做法，但这并不意味着该过程将被视为两套对立标准之间的战场，它也不意味着委员会的最终建议会因为其成员的背景或是支持候选人的学系评估时的风格而从一开始就是可以预测的。在某些例外的情况下，比如说当系里的意见被否定时，一些教师可能会表现出强烈的情绪。支持推荐意见的学院院长、系主任甚至会进行游说，并请学校一级的学术行政管理者重新考虑该评估结果。

然而，咄咄逼人的游说被认为是一种不恰当的政治形式，就质量上的优劣和制度权威的节制来说，它违反了公平公正的基本学术原则。这些原则和制度规范约束着系统中所涉及到的各方，在这个系统里，相互依赖性的重要而持久的，无论是在机构的学术结构内部还是在组织和专业之间都是如此。在这样的问题中公开原因是有害的。

原文载于：Thoenig J C, Paradeise C. Organizational governance and the production of academic quality: Lessons from two top US research universities[J]. 黄颖，郭展宏，谭越（北京大学教育学院硕士生）译

送：校领导、各学院（系、部）、各部门、直（附）属单位。

联系电话：（029） 87082051 87080244 pgb@nwsuaf.edu.cn