

关于建立普通高校动态性通用评估指标体系的设想

天津工业大学管理学院副教授 邱法宗

关键词 单一方案评估 分类评估 通用性 动态性 指标体系

一、问题的提出

1、从单一方案评估到分类评估的必要性

依据教育部《普通高等学校本科教学工作水平评估方案（试行）》，从2003年起，我国开始进行新一轮高校评估。相对于20世纪90年代进行的三次评估，本轮评估力求用一套方案和指标体系来评价所有类型的普通高校。“应该说，用统一的评估方案来代替原来同时进行的三类不同评估方案，有利于国家对新形势下高校办学质量的整体把握和宏观管理。”[1]另外，各类普通高校评估方案统一，易于操作，便于对照，同时也可以降低评估成本，体现评估简约、效率原则。

但是，进入新世纪不久，我国高等教育整体上由精英教育开始进入大众化教育阶段（提前实现教育部关于2010年高等教育“入学率接近15%”的规划指标），高等教育得到超常规、跨越式发展，高校类型、层次、规格、办学主体等呈现出越来越多样化的趋势。这样，原先设定的评估方案的单一性与高校发展的多样性的矛盾开始凸现。

作为对未来具有引导功能的高校评估，单一方案评估容易诱导不同类型高校的价值目标趋同化。前几年，出现普遍追求升高办学层次的倾向。“一些本科院校不顾自身实际，盲目争办‘综合性、研究型’大学；一些高职、专科院校积极向本科‘升格’；而一些中专学校刚刚升为高职专科，马上又开始策划升本科，导致办学定位的错位。”[2]这样，“多样化的社会需求与单一化的发展目标的矛盾，成为当前高等教育事业发展中突出的问题”。[3]最令人忧虑的是大学生就业，“在实践发展过程中，部分地方本科院校仍然沿袭传统的单一化人才培养模式，以培养研究型、学术型人才为己任，与精英大学盲目攀比，贪大求全，导致所培养的人才与社会需求相脱节。”[4]

要改变上述现状，就需要调整高等教育结构以及反映该要求的高校评价。但是，“现行评估指标是根据学术性研究型大学的标准制订的，它无助于激励竞争而且起着消极误导的作用。应当按不同类型高校制订不同的评估指标，至少应按三大类型制订三种或更多的分类评估指标，使各类高校各安其位，开展公平的竞争。”[5]“高等学校实行分类指导、分层次管理是发达国家高等教育管理的一贯做法，不同类型、不同层次的高等学校有不同的分工，不同的发展目标、重点和特色，呈现出互补关系，彼此不可替代。”[6]因此，对高校分类评估，以

取代单一评估方案的呼声越来越高。

2、分类评估是必要的，但是实施成本过高

分类评估，是高等教育多样化、可持续发展的客观要求，因而是必要的。但是，是否实施“分类评估”政策，还在于该政策是否具有可行性以及可行性程度的高低。

分类评估首先要解决高校类型的划分，因为后者是前者的必要前提。但是高校类型的划分本身非常复杂，不仅如此，对实施分类评估的成本、组织管理等要求更高。

面向评估目标的高校分类，目前学术界倾向于按照纵横两个纬度结合进行。从横向来看，按照学科门类来划分；从纵向来看，按照人才培养规格或者科研的高低层次来进行划分。

关于横向划分（类划分）主要有：以学科比例为分类标准，可划分为五种基本类型——综合类大学、文理类大学、理科类大学、文科类大学、专业类大学；以学科门类的一定集成度来进行划分，可把高校划分为单科性大学、多科性大学、综合性大学；《中国大学评价》课题组 2002 年提出的《中国大学分类新标准》主张：按照大学科（学科门）来进行划分。即参照教育部对学科门类的划分，结合各学科门类的比例，分为综合类、文理类、理科类、文科类、理学类、工学类、农学类、医学类、法学类、文学类、管理类、体育类、艺术类等 13 类[7]。这就出现了至少 3 类、5 类、13 类三种横向划分。

关于纵向划分（型划分）主要有：有的“建议把我国的高等学校分为研究型、教学研究型、教学型”[8]；有的主张除了研究型、教学型以及教学研究型大学之外，还应再加上“教学服务型大学”[9]；还有的主张“按科研规模的大小，现有大学分为研究型、研究教学型、教学研究型、教学型等 4 型。”[10]这样就出现了至少 3 种、4 种纵向划分。

分类评估不能是单纯的“类”划分，也不能是单纯的“型”划分，因为现实的高校是横向“类”与纵向“型”的统一。这样，如果把“类”与“型”结合起来，就会至少出现如下基本类型组合：3 类*3 型=9 种类型、5 类*3 型=15 种类型、13 类*3 型=39 种类型的状况（不考虑 3 种以上“型”的划分）。很显然，上述最简单的划分是 3 类*3 型=9 种类型的划分，如表 1 所示。

表 1 普通高校基本类型组合表

类（横向划分） 型（纵向划分）	综合性	多科性	单科性
研究型	综合性研究型	多科性研究型	单科性研究型
研究教学型（教学研究型）	综合性研究教学型	多科性研究教学型	单科性研究教学型
教学型	综合性教学型	多科性教学型	单科性教学型

问题在于：即便对最简单的 9 种类型划分的高校进行分类评估，其评估成本也会数倍于

单一方案评估。实践中的问题是：如果对各类型高校同时评估，对组织、成本等要求相当高；如果非同时评估（比如：一年度评估一种类型），评估周期就会因为太长，而失去评估的意义。

因此，分类评估是必要的，问题的焦点在于研究分类评估的可行性。这种可行性的主要思路应该是化繁琐为简易。即应力争研究简便易行的方案。

二、动态性通用评估指标体系的基本特点

所谓“动态性通用评估指标体系”，就是吸取“单一评估方案”适应面广、简约效率的优点，又吸取“分类评估”适应、促进高校发展多层次、多样性的优点，而建立的普通高校评估指标体系。该体系用一套评估方案，但通过“动态性”设计，力求适应于所有类型的普通高校。其基本特点是：

第一，“通用性”。即该体系适用于评价各类型、各层次普通高校（目前暂不考虑成人高校）。当然，这里隐含着—个理论假设，即在对象类型繁多的情况下，建立“通用性”标准（方案）的设想有可能吗？回答是肯定的。实践中，确实存在着研究和—使用“通用标准”（或“万能标准”）的先例和可能。“通用标准”起源于技术管理。在技术管理过程中，可以发现许多个性标准之间包含一些相同的特征，如产品标准中的尺寸规格、参数系列、环境条件、验收规则和试验方法等。将这些相同特征综合在一起，制定一个在一定范围通用的标准，即为通用标准，也称共性标准。在高校评估方面，也有学者多年探索不同类型高校的可比性问题，以建立不同类型高校之间的“通用性”评估标准”。因此，存在着建立“通用评估指标体系”的可能性。

第二，“动态性”。从表面上看，“通用性”标准与原来的“单一性”标准似乎都是“一个”标准。但是，两者实际上是不一样的。“单一性”实质上只适应一类高校或者说以一类高校（研究型大学）为标准；而“通用性”本质上适应于各类型高校。而由“单一性”转变为“通用性”的关键是“动态性”。“动态性”意味着：标准尽管还是“一个”，但此标准却是可变的、动态的、弹性的，从而区别于传统的“单一性”评估标准的不变性、固定性、静态性等局限性。“动态性”意味着在集成一个体系的前提下，可以适用具体情况或者要求进行转换或者切换。可以在建立“通用性”标准的基础上，在各类型高校转换使用。

三、动态性通用指标评估体系的“通用性”

所谓评估指标体系的“通用性”，就是要在覆盖各类型普通高校的基础上，选择、筛选“通用性”指标，然后按一定规则将这些“通用性”指标加以集成。因此，在指标体系的“通用性”问题上，有两点：一是考虑高校类型的“通用性”，即指标体系应当适应于各类普通高校；

二是要考虑指标选择的“通用性”。下面着重分析指标选择的通用性。

1、“通用性”评价指标的选择与集成

按照“从抽象到具体”的思维方法建立，第一步，从最抽象（最基本）的方面看，高校评估指标应当涵盖高校教学、科学研究和社会服务三大功能。因此，各类型高校都应当涵盖三大“通用性”职能（设定为一级指标）。第二步，对以上三个方面的一级指标进行分解，筛选二级指标体系。比如：“科学研究”作为一级指标，可分解为“科技成果获奖”、“学科建设”、“项目研究”、“论文发表”等二级指标；第三步，将二级指标分解为三级指标。比如：可以将“学科建设”分解为：一级学科博士点/二级学科博士点/一级学科硕士点/二级学科硕士点/学士授予等层级。

在具体分解的基础上，按照纵横两个方面组成指标体系。如表2所示（示例，未作全部排列）。

表2 普通高校通用性评估指标体系（部分）

标层	科学研究			教学			社会服务		
	科技成果 获奖 (1)	学科建设 (2)	项目研 究 (3)	师资队伍 (4)	人才培养 (5)	教学成果 (6)	合作研 究经费 (7)	学生就业 率 (8)	社会评价 (9)
一	国家级重 大科技奖	博士后流 动站/一级 学科博士	国家重 大基金 项目	两院院士 及其公认 大师	杰出政治家 (企业家)等	国家一 等奖	亿元级	90%以上	全国有重 要影响(或 世界有影 响)
二	国家级一 般科技奖	二级学科 博士点	国家一 般基金 项目	长江学者 或同级别	研究型人才	国家其他 级别奖	五千万 元级	80%-89%	全国有影 响(或本省 有重要影 响)
三	省部级重 大科技奖	一级学科 硕士点	省部级 重大基 金项目	博导/教 授级	高级专门人才 (1)	省部一等 奖	千万元 级	70%-79%	同类高校 “优秀”
四	省部级一 般科技奖	二级学科 硕士点	省部级 一般基 金项目	硕导/副 教授级	高级专门人才 (2)	省部其他 级别奖	百万元 级	60%-69%	同类高校 “良好”
五	省部级以 下级重大 科技奖	本科学士 授予	厅局级 项目	讲师/实 验师	技术应用型人 才(1)	省部以下 级别重大 奖	十万元 级	50%-59%	同类高校 “合格”
六	省部级以 下级一般 科技奖	其他层次	厅局级 以下项 目	其他教师	技术应用型人 才(2)	省部以下 级别一般 奖	十万元 级以下	50%以下	同类高校

2、指标分区（模块化）

对所建立的“普通高校通用性评价指标体系”（表 2）所列各项指标进行分区，划分不同性质的基本模块。仔细分析即可发现，表 2 指标体系所列第（1）-（5）项所涉及到的纵向指标与特定类型的高校有着相对确定的对应关系。比如第（5）项“人才培养”中的“研究型人才”与“研究型大学”直接对应，应用型人才与一般“教学型大学”直接对应。再比如第（2）项“学科建设”中的“博士后流动站”与“研究型大学”有着极高的关联性因此，我们把上述（1）-（5）项所对应的区域叫作“指标对应区”（对应模块），所涉及的指标可以叫做“对应性指标”。而第（6）-（9）项所列指标与特定类型的高校之间没有确定的对应关系，而是表现为一种随机性关系。因此，把上述第（6）-（9）项所对应的区域叫作“指标非对应区”（非对应模块），其所涉及的指标可以叫做“随机性指标”。

深入分析，又可发现，在“指标对应区”里，对某一特定类型的高校来说，又可分为“基本对应区”（基本对应模块）和“非基本对应区”（非基本对应模块）。比如，对“教学研究型大学”来说，其“基本对应区”是表 2 所列的“层级三、四”。“基本对应区”的确定是非常重要的，它可以避免高校定位不准，盲目“攀高”的倾向。除了“基本对应区”之外，还存在着“非基本对应区”。比如某些“教学研究型大学”（特别是一些历史悠久、实力较为雄厚的老牌高校），少数学科（专业）的研究成果甚至不亚于一些“研究型大学”，这类大学可以完成级别很高的国家级重大科技项目、一定数量的级别较高的“SCI/EI/ISTP/SSCI”的论文，获得国家重大科技奖等（参见表 3）。因此，“非基本对应区”是指是某一类型的大学在完成与“基本对应区”相对应的任务之后，完成的更高层次的任务。这样，“非基本对应区”设置在“基本对应区”之上（最高层次的研究型大学[1 型]只有“基本对应区”，没有“非基本对应区”）。

对“非基本对应区”，不能简单否定（当然也不应当一概肯定）。理论上讲各类型高校的定位应当是确定的，不应该出现“非基本对应区”。但是分析问题的出发点是应当基于现实。现阶段，我国高校不像处于相对稳定时期的美国等国家的高校“层次类别十分清楚。”[12]比如说依据美国卡内基教育促进基金会的分类，美国博士级研究型大学居于最高层次而数量只占高等教育机构的 6.6%[13]“而我国共有高等学校 1973 所，其中最高学位授予权为博士学位的学校就有 245 所”[14]，约占高等学校总数的 12.4%。这就意味着我国具有博士学位授予权的大学中，非“研究型大学中”占有很大比重（截至 2007 年，我国“985 工程”大学——研究型大学有 38 所）。另外，从整体上说，我国尚处于转型时期，高校类型相对变化较快。比如：“985 工程”高校（研究型大学），1998 年刚刚起步建设的时候，第一期工程

达到 34 所（1998-2003），第二期工程（2004 开始）达到 38 所，2008 年以后，第三期工程规划达到 43 所。因此，设置“非基本对应区”是符合目前我国各类高校发展变化较快的现实的。最后，从充分利用资源角度看，可以发挥一般高校（有学者称之为“中间高校”）在某些学科和方面的优势、特长和积极性。但是，要着重指出：某些高校对“基本对应区”所对应的任务投入不足，却盲目往上“攀高”，这是应当着重加以解决的。为此，应当对“基本对应区”设置较高的评估权重，而降低“非基本对应区”评估的权重。详见表 3。

表 3 高校通用性评价指标体系分区（模块化）

指标	科学研究					高校教学		社会服务	
	科技成果获奖 (1)	学科建设 (2)	科研项目 (3)	师资队伍 (4)	人才培养 (5)	教学成果 (6)	合作研究经费 (7)	学生就业/创业率 (8)	社会评价 (9)
一级	指标	指标	指标	指标	指标	指标	指标	指标	指标
二级	指标	指标	指标	指标	指标	指标	指标	指标	指标
三级	指标	指标	指标	指标	指标	指标	指标	指标	指标
四级	指标	指标	指标	指标	指标	指标	指标	指标	指标
五级	指标	指标	指标	指标	指标	指标	指标	指标	指标
六级	指标	指标	指标	指标	指标	指标	指标	指标	指标

不同模块的设置，既是建立“通用性”标准的需要，也是适应各类高校“动态性”评估的需要。

四、通用指标评估体系的“动态性”问题

所谓“评估指标体系的动态性”，就是在将通用性评估指标体系应用于各类型普通高校评估的过程中，要处理好模块、指标、权重的关系，使评估具有良好的可比性、平移性。解决的主要环节有：

1、权重的设置

在通用性评价指标体系分区（模块化）的基础上，应当对各区赋予恰当的权重，为不同类型高校绩效的比较和数据的归一化处理打好基础。权重具体设置见表 4。

表 4 高校通用性评价指标体系权重赋值表

指标分区	指标对应区						指标非对应区 (Z)		
	基本对应区 (X)			非基本对应区 (Y)					
指标类	指标	指标	指标	指标	指标	指标	指标	指标	指标
重要性	重要	较重要	一般	重要	较重要	一般	重要	较重要	一般
权重取值范围	0.2-0.3	0.1-0.15	0.05-0.1	0.1	0.07	0.03	0.2-0.3	0.1-0.15	0.05-0.1
权重取值合计 $\Sigma (0.4+0.2+0.4) = 1$	0.4			0.2			0.4		

对上表所赋权重设置的解释是：在总权重为 1 的情况下，“指标对应区”的指标刚性较强、也更重要一些，因此权重设置应当高一些（0.6），相对来说，“指标非对应区”的指标权重设置低一些（0.4）。在“指标对应区”中，“基本对应区”权重（0.4）是“非基本对应区”（0.2）权重的 2 倍。这样设置权重的依据，主要针对近些年许多高校不顾社会需求和自身条件，一味“攀高枝”（研究化、博士点化、综合化）的倾向，通过设置较高的权重，来约束和鼓励各类型高校科学定位、各安其位，在对应层次上争创一流高校，但也适当考虑我国高校类型和层次的历史和现实状况。

某个普通高校评估总分计算的一般公式是：

$$\sum (0.4\bar{X} + 0.2\bar{Y} + 0.4\bar{Z})$$

（式中， \bar{X} ——基本对应区算术平均值； \bar{Y} ——非基本对应区算术平均值； \bar{Z} ——非对应区实测平均值。）

2、动态性测许

在设置权重的基础上，可以利用统一的通用性评价指标体系对各类型普通高校绩效进行测评。

其基本环节是：第一，归类。将某高校归于特定的基本类型组合（表 1 所示）。第二，利用普通高校通用性评价指标体系测值（表 2、表 3 所示）。在测值时，要注意基本对应区、非基本对应区、指标非对应区取值方式的不同。第三，按照 $\sum (0.4\bar{X} + 0.2\bar{Y} + 0.4\bar{Z})$ 的公式进行合并计算（表 4 所示）。

在按照 $\sum (0.4\bar{X} + 0.2\bar{Y} + 0.4\bar{Z})$ 的公式对某所高校进行合并计算时，步骤：

首先，测算 \bar{X} 、 \bar{Y} 、 \bar{Z} 平均值。

计算基本对应区算术平均值 \bar{X} 。分别测出（统计） X_1 、 X_2 、 X_3 …… X_n 的实际数值。然后，

根据 $\bar{X} = (X_1+X_2+X_3, \dots, X_n) / n$ 的公式进行算术平均值计算。例如：某教学研究型大学对应于层级三、四区的科技成果获奖、学科建设、研究项目、论文发表、师资队伍分别得分是：80、70、60、40、90 分。则 $\bar{X} = (80+70+60+50+90) / 5 = 70$ （分）。同理，测算出非基本对应区 \bar{Y} 的分值。比如：某研究教学型大学 $\bar{Y} = 80$ 分；指标非对应区 \bar{Z} 值 = 60 分。

其次，根据所赋权重进行归一统计测算（满分为 100 分）。

比如：将某研究教学型大学测算的 \bar{X} 、 \bar{Y} 、 \bar{Z} 的数值，代入

$\sum(0.4\bar{X} + 0.2\bar{Y} + 0.4\bar{Z})$ ，即可测算出某研究教学型大学总分值是：

$$\sum(0.4 \times 70 + 0.2 \times 80 + 0.4 \times 60) = 68 \text{ 分}$$

其三，测算时的变通。

“型”的变通。为了保持一定的弹性、张力性，本文所主张的“普通高校通用性评价指标体系”，将研究型、研究教学型、教学型分别对应 2 个层级（研究型大学对应一、二层级；研究教学型大学对应三、四层级；教学型大学对应五、六层级）。这种对应层级也可做有条件的变通。比如：有学者主张将研究型大学细分为“研究 1 型、2 型”、研究教学型大学细分为“研究教学 1 型、2 型”、教学型大学细分为“教学 1 型、2 型、3 型” [18]。如果各取 2 个亚类型，则每个亚类型对应的就是一个层级区。

“类”的变通。比如：有聿者经过测算，“平均每一名自然科学全时人员获得的科研经费是社会科学全时人员的 6.2158 倍。这就使取 0.4244 万元与自然利学家获取 2.6378 万元具有同等难度。” [19] 这说明，文科类高校（学科、专业）与理工类高校（学科、专业）的“科研经费”指标，在统计测算时，应做归一处理，即做好“类”的变通处理。

“区”（模块）的变通。如前所述，最高层次的研究型大学（比如：研究 1 型）只有“基本对应区”，没有“非基本对应区”。因此，在对研究型大学（研究 1 型）具体测算时，公式

$\sum(0.4\bar{X} + 0.2\bar{Y} + 0.4\bar{Z})$ 应该变通为 $\sum(0.6\bar{X} + 0.4\bar{Z})$ ，然后进行测算。

“指标”的变通。由于各类型高校有很大区别。比如：艺术类、体育类高校，不但与工科类院校有根本区别，就是与文科类院校也有很大区别。“艺术学校上钢琴课，每个学生一架钢琴，一个老师就不能指导 40 个学生的 40 架钢琴，但是其他一些文科专业就可以一个班 40 个学生” [20]。因此，在指标数量、权重确定的前提下，具体选择的指标，可以做适当变通处理。

总之，本文初步提出的“普通高校动态性通用评估指标体系”，是一种框架性的构想、思路。还有待于做进一步的探讨。

参考文献

- [1]曾洁,彭安臣.高校本科教学质量评估指标体系的多样化[J].高教发展与评估(鄂),2006(1).
- [2][10]杜彦良等.关乎成与败,得与失——对不同类型高校定位与人才培养模式的思考[N].中国教育报(京),2004-07-09.
- [3][5][7][16]潘懋元,吴玫.高等学校分类与定位问题[J].复旦教育论坛(沪),2003第1卷(3).
- [4]詹星.关于地方本科院校人才培养目标的思考[J].黑龙江教育(高教研究与评估),2006(6).
- [6]陈方红.从社会学视角审视高等教育过渡的影响[J].教育与职业报(京),2007(8).
- [8][17]翟帆.高等学校要分类指导分层次管理[N].中国教育报(京),2003-3-10.
- [9][12][18]武书连.再探大学分类[J].中国高等教育评估(沪)2002(4).
- [11][14]刘献君.建设教学服务型大学——兼论高等学校分类[J].教育研究(京)2007(7).
- [13][19][20]武书连等.挑大学选专业-2004 高考志愿填报指南[M].中国统计出版社(京),2004.
- [15]周玉.解读哈尔滨工业大学办学传统与特色[N]光明日报(京),2006-04-27.

(文见《中国高等教育评估》2009年第2期)