

学校的理想装备

电子图书·学校专集

校园网上的最佳资源

学校管理操作规范  
教育教学评估方法·原则和标准



## 前 言

随着社会主义建设事业的发展，国家对人才的要求越来越高。为社会和国家造就和培养合格的人才，是各级各类学校的根本任务。一流的学校离不开科学的管理，科学的管理需要卓越的管理制度和办法。

为适应新形势下各级各类学校管理的科学、规范、高效的要求和需要，我们特约请和组织了有关专家、学者和教育管理工作共同编撰了《学校管理制度方法操作规范》全书。

在编撰过程中，我们从全国各级各类学校中精心选取了优秀和成功的学校管理规章制度和方法，基本上涵盖了学校管理的每一个部门、每个工作环节、每个管理层次，力求为每项工作的管理提供具有典型性、规范性、实用性的操作规范，为学校广大管理工作所直接借鉴和参考。

本书是集体智慧和劳动的结晶，广泛吸收了国内学校管理理论和实践的最新成果，参考和引用了有关学校管理的文献和资料，在此向有关单位和人士表示衷心地谢忱。

在编撰过程中，我们力求体系完整、资料翔实、内容丰富，反映国内学校管理理论和实践的新成果、新问题，但由于我们水平有限，加上时间和资料的限制，难免有这样或那样的缺点和不足，敬请广大读者批评指正，以便再一步修订和完善。

《学校管理操作规范》编委会  
一九九七年三月于北京

## 教学评估的方法

教学评估的方法是指评估主体为完成评估任务所采用的方式、措施和手段，受教学评估的目的和内容等因素的制约。教学评估的方法，应当根据教学工作的性质和特点来确定，一种适合于所有学校教学的评估方法和模式是不存在的；教师参与是实施任何一种教学评估方法的基本条件，也是教学评估目标得以实现的保障；任何教学评估方法或模式使用的最终目的只有一个，就是改进教学，大面积全面提高教学质量。

教学评估的方法依据不同的划分标准有不同的类别。根据是否采用数学方法可分为非数量化方法和数量化方法；根据评估的范围可分为分析评估法和综合评估法；根据评估的主体可分为自我评估法和他人评估法；根据评估的基准可分为相对评估法和绝对评估法。另外还有内差异评估法等。

### 非数量化方法

非数量化方法是指在教学评估过程中并不采用数学的方法，而只是根据观察、检查和记录等，对教学工作及质量进行定性描述和判断，故亦称定性评估法。这种方法由于简便易行而在实际的教学评估工作中常被采用。其中等级法、评语法是我国教学评估工作中经常使用的非数量化方法。

#### 1. 等级法

教学评估者依据教学评估指标体系的要求和教学质量管理的某种标准，把教学质量划分成几个等级，然后再把评估对象的教学工作质量归到某个等级之中，依此来比较和判断教学质量的高低。在我国的教学评估实践中，等级的划分主要有“三级制”（上、中、下）、“四级制”（甲、乙、丙、丁或优、良、中、差）、“五级制”（优秀、良好、中等、及格、不及格或很高、高、一般、低、很低）等。等级法是一种传统的简单的定性评估方法，其主要缺点是：等级的划分主观性较强，标准很难掌握；等级划分之后，对教学质量的归类判定受评估者主观意向的影响极大。

#### 2. 评语法

评语法就是对教学评估中的每一项指标或内容用简洁的评语来判定教学质量的方法。例如对教师的课堂教学质量用评语法来评定：教学目标明确，认知目标和情意目标具有针对性；教学内容正确，重点难点突出；教学方法恰当，善于启发学生；学生主体参与意识强，学习积极性高；基本实现了教学目标，课堂教学效果良好。与等级法一样，评语法也是一种比较传统的简单的评估教学质量的定性方法，评估结果受评估者

价值观、质量观等主观因素的影响较大，很难全面地反映教学工作的质量状况。但若集思广益、评语中肯，也不失为一种教学评估的辅助方法。

### 数量化方法

用数值表示和刻划教学评估结果的方法称为数量化方法，亦称定量评估法。随着教育统计学、教育测量学、教育实验学、教育评估学、模糊数学的发展完善和普及，在教育评估特别是教学评估领域中，数量化方法日益受到人们的重视，探讨科学可行的数量化评估方法，也成为教学评估研究的重要

课题。教学评估的数量化方法主要有“计分评估法”和“综合评判法”。

### 1. 计分评估法

计分评估法是把教学质量分解为若干指标，对每一项评估指标确定一个最高分数，评估者利用某种评估工具和方法对每项评估指标进行评分，利用这个评分对教学质量做出某种判断的评估方法。根据处理所评分数时采用的方法，计分法又可分为累积分数法和加权计分法两种。

(1) 累积分数法。累积分数法就是假设教学评估的指标或内容有几项，评估者就每一项评定一个分数  $X_i$ ，然后求出它们的总和作为评估的结果  $X$

$$X = \sum_{i=1}^n X_i$$

式中  $X_i$  一般参照人们的习惯取值而评定，即对某项指标评分时，水平很高取 0.95，水平高取 0.85，水平一般取 0.75，水平低取 0.65，水平很低取 0.50。亦可根据实际需要将上述等级标定或转化成隶属度区间。

累积分数法的不足之处在于未能区别出反映教学质量的各项指标间的重要程度，而把各项评估指标置于同等重要的位置。然而，教学实践和教学质量管理工作并非如此。一般地，把教学和教学效果作为教学评估的重要内容和指标，因而评估时应该占较大的比重，但累积分数法并不能反映这一特点。

(2) 加权计分法。加权计分法是针对累积分数法之不足而设计的一种教学评估方法。其具体做法是：假设教学评估的指标有  $n$  项，视每项指标之重要程度的不同而赋予一定的权重  $P_i$  ( $i=1, 2, \dots, n$ )，评估者对每项指标的评分为  $X_i$  ( $i=1, 2, \dots, n$ )，则教学质量评估的结果为  $X$

$$P = \sum_{i=1}^n P_i X_i$$

当式中的  $P_i=1/n$  时，加权计分法与累积分数法是一致的，即累积分数法是把各项评估指标的重要程度看作相同时的加权计分法的特殊情况。因此，相对于累积分数法，加权计分法要合理得多。在教学质量评估中运用加权计分法要注意两点：一是要科学地设计反映教学质量的指标体系的权重集合，尽量避免人为杜撰；二是在具体到某位教师的教学质量评估时，要注意反映出该教师的某些具有高品质或颇具特色的指标，充分发挥教学评估的导向功能。

### 2. 综合评判法

这里所说的综合评判法是指利用模糊数学中的“模糊综合评判模型”对教学工作进行评估的一种方法。综合评判法是最近几年我国教育评估工作者和数学工作者模糊数学应用于教育评估而形成的一种新的有效的方法。由于这种方法所需的数学知识较多，在运用的实践中许多问题又需要进一步探讨和完善，故在后面列专题研讨。

## 分析评估法

分析评估法是根据一定的评估观点，把反映教学工作质量的指标或内容分解成为几个要点部分，然后对每个要点或部分采用适当的方法分别进行评定。

对每个要点或部分进行评估时，可以采用数量化方法，也可采用非数量

化方法，还可以采用其他方法。具体采用哪种或哪些方法，要根据教学评估的实际需要和可能选定。但是，不论采用何种方法，都要力争做到客观地、正确地、真实地反映教学质量各项指标的要点和内涵，尽量地减少主观性因素的影响。

### 综合评估法

综合评估法是对教学工作的评估内容或指标的整体进行评估的方法。在非数量化方法中的等级法中，评估者采用的就是综合评估法。经验丰富的专家型的评估者多采用综合评估法，凭他们的直观对教师的教学质量做出评估，然后再分出等级。

综合评估法从表面上看似乎没有分解评估内容，实际上分解过程是在头脑中进行的。因此使用综合评估方法评估教学质量时要注意两个条件：一是评估者要有丰富的经验；二是评估者靠视觉直觉可以获得评估对象大量的信息。

分析评估法与综合评估法并不是对立的。没有分析就没有综合，综合建立在分析的基础之上；只有分析，没有综合，就会“只见树木，不见森林”。在教学评估实践中，往往是先进行分析评估，然后进行综合评估。

### 自我评估法

自我评估法是评估者根据一定的教学观和教学质量观，对自己的教学工作进行价值判断的一种方法。对于一位教师来说，对自己教学工作的自我检测就是自我评估的过程；对于一所学校来说，教学管理者对学校教学工作的质量检查，质量分析都是自我评估。

自我评估法是教学工作评估方法体系中的一种高级形式，他体现了现代教育评估从被动走向主动、从他律走向自律、从导向走向自控的发展趋势，具有也应该具有自主性、自律性、自控性三大特点。自我评估的优点在于教学活动的主体就是评估的主体，便于自我认识、自我教育、自我激励、自我控制、自我完善，积极主动地提高教学质量。其缺点是缺少外界的参照标准，很难进行模型比较和鉴别。

### 他人评估法

自我评估之外的评估称为他人评估。例如，上级教育行政部门对下属学校教学质量的评估，学校管理者对教师教学工作的评估，学生对教师教学工作的评估，以及社会舆论对学校教学质量的评估，都属于他人评估。一般地说，他人评估要比自我评估相对客观、真实一些。但是他人评估要求比较严格，组织工作繁重，花费的人力、物力、财力也比较多。

自我评估法与他人评估法各有其优缺点，在教学评估实践中，把自我评估法与他人评估结合起来使用，能够相互补充，收到更理想的效果。一般先进行自我评估，然后再进行他人评估。要特别注意的是，教学评估的目的不是为了排名次、论高低、评等级，而是为了使教学工作按照既定的目标前进，全面提高教学质量。因此，评估特别是他人评估决不能过于频繁，否则就会打乱

教学计划，影响正常的教学秩序，与教学评估的目标背道而驰。

### 相对评估法

相对评估法，又称“常模参照”评估法，是在评估对象的集合中先取一个或若干个作为标准（亦称“常模”），然后把各个评估对象与这个常模进行比较，确定其位置，或者用某种方法将所有评估对象排成先后顺序。相对评估法的特点是根据评估对象集合的整体状况确定和判断某一或某些评估对象的相对水平，其标准或常模只适用于所选定的评估对象的集合，对于另外的集合未必适用，假设评估对象集合由元素  $A_1$ 、 $A_2$ 、 $A_3$ …… $A_m$  组成，而选定的参照标准和常模是  $M_0$ ，则相对评估法可用下图表示（图见下页）。

学校教学质量管理中的“赛讲”、“评优课”活动，采用的都是相对评估法。根据学生的学习成绩，采用标准分数进行的教学评估也是相对评估。

相对评估法的主要特点是在某一评估对象集合的内部进行比较、判定或排序，无论这个集合的整体水平如何，都可以进行比较，因而其适应性强，应用面广。但是，若集合的“整体水平”较低，经过与之比较而排序在前的评估对象的教学质量水平也未必就高。因此，相对评估有从“矮子里选高个”之嫌，容易降低教学评估的标准。另一方面，相对评估的结果并不表示评估对象教学工作的实际水平，只表示其在集体中所处的相对位置。因此相对评估法容易使评估对象产生激烈竞争，若引导不好，易误入歧途。

### 绝对评估法

绝对评估法又称“目标参照”评估法，是在评估对象的集合之外，确定一个客观标准，把评估对象与这个标准相比较，从而判断某评估对象教学工作质量的方法。在教学评估中，“标准”则是预先设定的教学质量的目标，因此，绝对评估法实质上是判断教学质量目标的达成度。假设绝对评估的元素是  $A_1$ 、 $A_2$ 、…… $A_n$ ，客观标准是  $M_0$ ，则绝对评估法可以用下图表示。

高中会考这种学科教学质量水平评估，就属于绝对评估的范畴。评估的标准是国家教委颁发的教学大纲和教科书以及会考大纲，会考试题囊括了教材和大纲的基本内容，评估的尺度是分数，满分 100 分，凡 60 分及以上者为合格。

1966 年，日本国立教育研究所的板仓圣宣提出的到达目标与到达度评估，也属于绝对评估。他根据目标的性质把目标分为“方向目标”和“到达目标”两大类。所谓方向目标，就是把期待的努力方向提出来作为目标。方向目标给定的是方向，而目标本身是向无限远处延伸的，与之相适应的评估是相对评估法。所谓到达目标就是把要掌握的提出来作为固定的目标，其本身是确定的、实在的，与之相适应的评估是绝对评估，即对照该目标来判断评估对象达到目标的程度。

由于绝对评估的标准具有客观性，因此在真实、客观、准确的完成教学工作评估之后，每个评估对象可以明确自己的教学质量与客观标准的差距，从而可以激励评估对象积极进取，具有较强的教育性。但是，绝对评估的“标准”实际上也很难做到客观，况且其历史甚短，从目标体系到测量、统计等

的技术和方法尚未完全成熟，因此绝对评估法并不是最完美的评估方法。但是，从相对评估法走向绝对评估法，是现代教育评估发展的一大趋势，应予以足够重视。

在教学评估实践中，呈现出将相对评估法和绝对评估法结合使用的趋势。前述“到达目标”与“方向目标”是分离的；所适合的绝对评估和相对评估也是对立的。事实上，“到达目标”在教学上仅限于基本概念、基础知识和基本技能的范围，这些又只是构成学生学力的外部因素，而学生的学力还有兴趣、态度、毅力等内部因素，况且内部因素更为重要。板仓圣宣提出的到达度评估，在实践上显然有把相对评估与绝对评估对立起来之不足。针对这种不足，日本女子大学教授梶田睿一进一步把到达目标分解达成目标、提高目标和体验目标三大类，每一类各包含一系列的认知领域、情意领域、精神运动领域的目标，并且有各自的测量评估规则，梶田睿一的理论使得两种分离的目标结合起来，也使得对立的两种评估方法相容互补了。在教学评估实践中，将两种方法结合使用，既能使评估对象明确自己的教学水平与质量标准间的差距，又能认清自己在所有评估对象中处于哪个等级或位置，从而全面地把握自我，寻求发展。

### 内差异评估法

内差异评估法，是把评估对象集合中的各个元素的现在和过去相比较，或者一个元素的若干侧面相比较，以判断其进步程度或所长侧面的评估方法。假设评价对象集合中的元素是  $A_i$  ( $i=1.2.....n$ )，过去的水平状态是  $A''_i$ ，现在的水平状态是  $A'_i$ ，则内差异评估法可用下图表示。

内差异评估法应用于教学评估时有两种情况。一种是比较评估对象现在与过去间的差异。如某教师过去的课堂教学质量是“良”等，现在是“优”等，说明其教学质量有明显提高。另一种情况是比较评估对象的某几个侧面，考察其所长或所短。如一位教师的课堂教学可以从教学目标、教学内容、教学方法、主体参与、教学效果五个方面来考察，考察之后可以发现某位教师在哪一方面或哪些方面较好一些，在哪一方面或哪些方面较差一些。

内差异评估法充分地照顾了个体的差异，便于形成“教有特色”的局面，有效地促进教师自身教学水平的提高，而且在评估过程中给评估对象造成压力，因而是一种很有发展潜力的评估方法。由于内差异评估法既不与客观标准相比较又不与其他评估对象比较，评估结果表明若有提高，可能会自我满足。因此，若能将内差异评估法与相对评估法结合使用，效果会更好一些。

### 模糊综合评判法

“模糊综合评判”是模糊数学所提供的解决模糊现象的评估问题的一种数学模型。它是融模糊测量、模糊统计、模糊评估于一体的综合性的数学方法，在解决多因素、多指标的教育教学评估问题中，有十分广泛的应用。

## 如何设计教学评估方案

教学评估方案是在一定的教育价值观的指导下，为评估教学目标的达成度所规定的评估对象和任务、评估的指导思想和原则、评估的指标体系以及评估的方法步骤等的复合体，是进行教学评估工作的通盘的计划和策略，是开展教学评估工作的纲领和依据。

设计教学评估方案，是教学评估组织者最基本的职能，方案管理也是教学评估的基本制度和教学评估组织者的职责。只有通过对方案的研究设计、组织实施、检查复核和总结分析，才能对教学评估的全过程进行协调与控制，有效地发挥教学评估的作用。所以，最重要的还是搞好评估方案的设计。

### 教学评估方案的内容结构

关于教学评估方案的内容结构，目前尚无统一的模式。各地所设计的评估方案的内容结构也不尽相同。在求同存异的前提下，我们认为评估方案的内容结构应该由“评估的对象和目的任务”、“评估的指标体系”、“评估的组织领导和方法步骤”等组成。

#### 1. 评估的对象和目的任务

教学评估的对象，在广义上包括所有的与教学活动有直接或间接关系的实态把握和价值判断。可以是学校教学水平的综合评估；也可以是教学条件、教学管理水平、培养水平的单项评估；还可以根据从核心到边缘影响教育目标达成度的顺序划分为六种水平的评估对象：每个受教育者，教学活动，直接规定教学活动的内容和方式（教学计划、教学大纲、教师、教科书、教学观、教学方法等），潜在地影响受教育者发展的学校社会文化背景（学生集体、教师集体、学校的整体状况等），学校物质的社会的环境条件（基本设施、校址校舍、社会环境），以学校为构成要素的包罗范围更广的教育体制（教育行政体制、学校教育在整个社会中的地位 and 职能）。总之，中小学教学评估的对象，包括凡是在某种意义上参与教育成果的一切。而作为评估方案内容的评估对象，则可能是所有对象的某一或某些对象。

评估对象不一样，评估的目的任务则不尽相同。以教学水平为对象的综合评估，目的在于对教学条件这个“投入”和培养水平这个“产出”的测量与统计，综合判断教学成果这个“转化”率，从总体上把握学校教学水平的高低。以教学条件为对象的单项评估，目的在于抓住影响教学目标达成度的这一条件因素，通过测量统计，全面地把握教学条件实态，统一人们“教育投入”的价值观。以管理水平为对象的单位评估，目的在于促进教学工作者树立“向管理要质量”、“以管理求效益”的价值观，全面提高教学管理者的素质和管理水平。以培养水平为对象的单项评估，目的在于对教学目标的达成度进行实态把握和价值判断，端正人们的教学质量观和人才观，大面积提高教学质量。

因此，在教学评估方案中，只有明确地阐述评估的对象和目的任务。才能有效地指导教学评估工作。同时，这个问题解决了，也为其他问题特别是评估指标体系的设计奠定了基础。

#### 2. 教学评估的指标体系

教学评估的指标体系是教学评估方案设计阶段的核心工作，也是评估方案的核心内容。指标体系是由评估指标系统、评估标准系统、评估计量系统组成的有机整体。

评估指标系统的本质是评估目标的某方面的规定，它是具体的、可测量的、行为化和操作化了的评估目标。评估指标系统的设计过程，实际上也是人们价值认识取得统一的过程，是把对客体的总体认识转化为对客体某些局部方面的认识过程。

评估标准系统是对评估指标系统在行为化、操作化方面的具体界定，是评估指标集上的模糊子集，常用“优、良、中、差”等等级性的模糊语言表示，表明评估指标系统对教育目标的隶属度。

评估计量系统由测量方法、统计方法、权重集等主要内容组成。评估测量方法包括验典的测量方法和“项目反应理论”等现代的测量方法，其具体规定如何根据评估标准系统，获取评估指标系统在教学目标达成度方面的有效的、可靠的信息，确保教学评估工作的客观性；评估统计方法也包括验典的数理统计方法和现代的模糊统计方法，其具体规定测量信息的整理、估计和推断，确保教学评估的科学性；权重集则是评估指标系统和评估标准系统这个直积集上的模糊关系集，其充分表明各项评估指标对教学目标达成度的影响价值有多大。

一般地，设计科学的教学评估指标体系的主要依据有四个方面。一是来自上面的、具有全局性、长期指导作用的有关方针、政策、法规和上级领导部门特别是教育管理部门对某一时期工作的具体要求及指示。前者如国家的教育宗旨，《中共中央关于教育体制改革的决定》、《中共中央关于精神文明建设的指导方针》、《义务教育法》、《关于加强学校德育工作的意见》、《中国教育发展和改革纲要》等，后者如《关于纠正片面追求升学率的若干规定》等。分析这方面的依据而设计的教学评估指标体系，能保证教学评估工作的方向性。二是“来自理论”的，主要包括马列主义的基本理论和原理、建设有中国特色的社会主义理论、教育学、教育测量学、教育统计学、教育评估学等教育科学理论和教育心理学、教育管理心理学等心理科学理论以及应用数学等数学理论。理论依据还包括科学的教学评估的经验总结。因为科学的经验总结的借鉴可以使设计工作少走“弯路”，多行“捷径”。设计教学评估的指标体系要重视理论依据，实质是要使设计工作按规律办事，确保教学评估工作的科学性。三是“来自下面的”。一方面是指当前的实情，诸如人力、物力、财力状况，评估对象的主客观条件，以及与前期评估工作相比，有哪些变化了的因素。另一方面是反映在前期评估报告之中的工作情况和实情。设计指标体系时，要重视调查分析，将“活”的实情与“死”的结论两方面的内容结合起来，确保教学评估的可行性。四是对未来情况的预测。评估指标体系作为评估方案的核心内容，是面向未来的。因此，采用科学的态度，对现有的信息进行分析，对可能发生的情况与出现的问题进行预测，是使评估指标系统具有应变性的关键。

### 3. 组织实施及其程序

设计教学评估方案的目的，在于通过对评估方案的组织、实施和管理，优化教育、教学管理及其评估过程。大面积提高教学质量。因此，在明确了指标体系这一评估方案的核心基础上，有必要对评估工作组织和实施及其一般程序做出具体的规定。

评估的组织包括设置必要的组织机构和组织力量开展教学评估工作的两个方面。评估组织机构的设置没有固定的模式，可以根据评估的对象、目的，评估的规模和实施评估的现实条件，或设置管理部门与社会力量相结合的评估机构，或设置地区或同行及社会民间团体的机构，或设置同行评估的机构，或设置自我评估的组织机构，或设置教育科研工作者、教育实践工作者、教学管理者参与的三位一体的综合性的机构。组织机构的主要职能是调动和组织力量进行教学评估工作，是顺利进行教学评估的组织保障。

评估的实施包括评估实施的主要工作和评估方法。评估实施的主要工作有：广泛而又深入地进行宣传和动员，简便而又科学的测量和评分，全面而又重点地收集信息。准确而又迅速地统计和整理。评估实施的方法因评估的对象、目的，特别是评估人员的素质而定，一般地，评估实施的方法应注意四个方面的结合：一是自评与他评相结合，以自评为主；二是定量方法与定性方法相结合，以定量方法为主；三是统一性与多样性相结合，以统一性为主；四是形成性评估与前置性、总结性评估相结合，以形成性评估为主。

组织实施及其程序还包括评估结果反馈的内容、形式和对象。反馈的内容主要包括综合判断的结论、分析诊断出的问题以及本次评估的质量及指标三个方面；反馈的方式主要是评估报告；反馈的对象、或评估对象、或领导部门、或同行。结果反馈的根本目的，在于使评估对象看清自己的问题所在，有针对性的工作；在于使领导部门了解实情，科学决策；在于使同行互相借鉴，共同提高。

## 教学评估方案设计的意义

1. 设计教学评估方案，是教学评估工作的基础工程。教学评估是一项复杂的系统工程。就教学评估的步骤而言，一般可分为评估准备、评估实施、评估结果的处理三个有机联系的阶段。评估准备是评估实施之前的预备阶段，包括方案准备和组织准备两个主要内容，主要解决“为什么评”、“由谁来评”、“评什么”、“怎样评”等四个问题，而这四个基本问题则是教学评估方案的基本内容，其中解决“评什么”、“怎样评”的问题。则是设计评估方案的核心。“评什么”主要是根据评估所依据的教学目标，设计教学评估的指标系统，确保教学评估的方向性和科学性；“怎样评”主要是在“评什么”的基础上，设计教学评估的标准系统和以测量、统计和权重集为核心的计量系统，确保教学评估的可检性和操作性。而由“评什么”所决定的评估指标系统和由“怎样评”所决定的评估标准系统和评估计量系统所构成教学评估的指标体系，是设计教学评估方案的核心。因此，评估方案的设计，是教学评估这项系统工程的基础工程。

2. 设计教学评估方案，不仅是教学评估的起点，也是教学评估的依据。

教学评估过程是由方案决策、组织实施、复核检查、总结反馈等评估职能所组成的有序活动。在这一活动序列中，任何教学评估工作，若不经历方案决策阶段，那是很难开展的；即使进行若干评估活动，但由于缺乏方案也是很难奏效的。因此，“方案决策”，即在分析依据、科学预测的基础上，遵循科学的决策程序，优选和设计教学评估方案，则是教学评估过程的起点。同时，设计的教学评估方案，将使教学评估的各类机构、各类人员、各级评估的领导和管理者在工作中“心中有蓝图”，是教学评估过程中及其他后继

环节的基础。按评估方案办事，组织实施阶段的工作内容，由方案所决定，并以实现评估方案所规定的为目标；检查复核是对组织实施的监督，也是对评估方案的验证，总结提高是对评估方案、组织实施、检查复核等的总分析和总评估，也是下一个评估循环过程的依据和基础。在各个评估职能之间，都必须通过“方案信息”的反馈，进行调节控制，使其沿着评估方案所规定的轨道前后衔接、相互渗透、互相促进，向着预定的评估目标逼近。因此，从这个意义上说，教学评估由方案决策开端；对教学评估的控制，主要的和首先的也是针对方案决策的控制。

### 3. 设计评估方案，实行方案管理是教学评估管理科学化的重要标志

“向管理要质量”、“以管理求效益”是现代教学评估发展的重要趋势之一，而“方案管理”则是管理教学评估工作的主要方式和手段。评估方案和方案管理的含义不同，但是又是密切联系的。教学评估工作的方案管理，是一种管理思想，也是一种管理方式和手段，它是以评估方案为中心，为实现评估方案而使用的一系列管理步骤、方法和手段，包括评估目标管理、评估预测、评估方案的可行性分析和评估方案的决策论证等主要内容。

概括起来，设计评估方案，实行方案管理在教学评估管理中的作用主要有：一是全面贯彻党和国家的教育方针、政策和法规，提高教学质量和改善学校评估的经营管理的的基本保证。方案设计和方案管理的核心是目标，教学评估各方面的工作，无论是人、财、物、事诸方面的评估工作，还是评估信息、时间、空间的运筹工作，都是以目标为核心组织起来的。方案设计及其管理可以通过目标的定向作用，使不同性质和特点的评估工作有秩序地运行，消除违背评估方案的各种干扰。二是协调和促进教学评估各方面工作及其整体平衡的重要依据。在教学评估工作中，矛盾和冲突是经常发生的，有的属于人事和规章制度方面的，有的属于资源分配和信息传递方面的等等。这些矛盾形成的原因很多，时间有长有短，涉及到的人员有多有少，对当前评估工作的影响也有急有缓。评估方案及其管理就是通过“方案”来协调各方面的工作，消除各种矛盾，以保证教学评估工作的整体平衡。三是激励全员团结奋斗的重要条件。评估方案及其管理，能促使学校成为全员的“命运共同体”，把全员的兴奋中心集中于对评估目标的追求上，使全员认识自己在整体方案中的地位和作用，增强使命感和责任感，从而振奋精神、增强信心、团结奋斗。四是有效地控制教学评估整体工作的一种手段。评估方案的功能和控制的功能是密不可分的。没有评估方案也就没有评估过程的控制而言。评估控制意味着通过纠正偏离评估方案的评估行为，使评估活动保持既定的方向。评估控制是实现评估方案的手段，评估方案为评估控制提供依据。

总之，设计科学的教学评估方案，是教学评估的基础工程，是进行教学评估的起点和依据，也是教学评估管理科学化的需要。

## 教学评估方案设计的原则

### 1. 科学性原则

教学评估方案是进行教学评估的“蓝图”，它的设计是由评估组织者在有关人员的参与下完成的。因此，方案设计者的知识水平、想象力、胆识、态度与远见的程度，直接影响评估方案设计的完善程度。所谓坚持科学性原则，一是方案设计者要懂得教育科学特别是教育管理与评估科学，遵循教学

的客观规律；二是有求实的精神和科学的态度，严禁片面性和偏激；三是要有民主的精神，动员和组织有关人员参与方案的设计；四是要懂得科学的决策程序，遵循科学的决策程序评估选优，设计方案。

### 2. 综合性原则

教学评估方案是科学、内容、时间以及人的心理状态的结合体，本身就是有综合性。所谓坚持综合性原则，就是在总体设计的指导下，正确处理整体与局部、时间与空间、数量与质量、刚性组织与弹性结构，情感逻辑与效益逻辑之间的关系，进行综合地、整体性的调查研究和分析，设计具有综合性的教学评估方案。

### 3. 连续性原则

教学评估方案是建立在原有的评估工作的基础之上的，也是为了解决未来的评估活动而设计的。因此评估方案的设计要了解过去，把握现在、展望未来。所谓连续性原则，就是要避免离开过去的评估基础和现在的评估条件去设计未来的评估方案，也要避免迷恋过去，或只顾眼前的评估条件而去设计面向未来的评估方案。同时，连续性原则也要求指导教学评估的教育方针政策和教育目标的连续性，如果教育的方针、政策和教育目标是非连续的、多变的，教学评估方案及其实施也必然失去连续性。

### 4. 激励性原则

教学评估方案是在数量与质量、规模与速度诸方面“设计未来”“安排未来”的，没有执行者的主观努力和评估对象的认同，评估方案必然是一纸空文。所谓激励性原则，就是在设计评估方案时，必须充分考虑方案执行者和评估对象的态度，尽力排除方案可能带来的心理障碍，设计容易被评估方案执行者接受和评估对象认同的具有激励作用的评估方案。

## 教学评估方案设计的过程

在坚持科学性、综合性、连续性、激励性原则的前提下，设计教学评估方案还必须遵循科学的决策程序。

### 1. 分析依据，确定目标

这个环节概括起来，就是要求在设计评估方案时，遵循规律，吃透“两头”，面向未来，确立评估目标。所谓吃透“两头”一是认真分析党和国家的教育方针、政策、法规和上级教育行政部门的指示精神，把握教学目标的基本点，统一思想，明确教学的目的任务，弄清楚“应该做什么”的问题，以确保教学评估的方向性；二是要调查教学及其评估的实际，弄清楚过去做了些什么，做到何种程度，并认真分析研究当前评估的人、财、事、物、时间、信息等资源，以及评估的主、客观条件，使“应该做什么”具体化而带有针对性的特点，所谓“遵循规律”，是指要研究和探索教学评估的基本原理和规律，认真总结过去评估工作中有规律性的成功经验，从而解决评估工作“应该怎样做”的问题。然后，以“应该做什么”和“应该怎样做”的结合为基础，而向未来，预测评估工作发展的趋势和前景，确立教学评估的总体目标。

### 2. 集思广益，拟定方案

设计教学评估方案，不仅是教育评估机构和评估组织管理者的职能和本份，也是教育科研人员、教育督导与管理人员、教育工作者的权利和义务。

群众参与评估方案的设计，不仅是评估民主化的措施，也是评估科学化的重要内容。在设计评估方案时，切忌设计者闻争则怒，长官意志、“一言堂”、“闭门造车”、必然导致评估方案失去科学性和可行性。而发动群众、民主讨论、提供交流与方案有关的信息、可以弥补“分析依据，确立目标”环节之不足或证实其可靠程度。同时，群众参与评估方案的设计，不仅能集思广益，而且使参与者产生变他律为自律的心理效益应使评估方案的要求变成参与者理智的行动。因此，群众的参与和讨论，也是教学评估工作中心理调控的有效手段。

在组织群众讨论时，发表的第一个意见不一定是全面而科学的意见，提出的第一个评估方案也不一定是理想的方案，由一种方案立即形成的评估方案，往往也是不成功的方案。因此，评估方案设计者在组织群众讨论时，要善于把讨论中的各种不同看法归纳整理成不同的初步方案，并进行比较研究，权衡各种方案的利弊，为拟定全面地、科学地，可行的评估方案奠定基础。

### 3. 深思熟虑，果敢决策

在前面两个环节“孕育怀胎”的基础上，方案设计者再加深思熟虑，果敢决策，评估方案就“分娩脱胎”了。所谓“深思熟虑”，指对各种方案比较研究之后，在权衡利弊得失的基础上，认真分析和思考自己或他人在教学评估工作中成功的经验和失败或失误的教训，不仅要深思“把事情做好”，而且要多虑“做对的事情”，而不是“做不对的事情”；然后再决断如何进一步把“对的事情做对”。这样，孕育成熟的优化的评估方案就脱胎而出了。

### 4. 反复论证，坚持实验

教学评估方案设计出来后，应该组织本地区各级各类有关人员对方案进行论证和评价。一是本地党政主要负责人对方案的评价和审核；二是地方教育行政领导和管理人员对方案的评价和审核；三是请有关教育专家和科研人员对方案进行评价和审核；四是学校管理者和教师对方案的审核和评价。设计者在吸收各方面提出的意见和建议再次修改的基础上，选定实验标本，组建实验队伍，按照实验的工作规则和程序，开展评估实验。在实践中检验评估方案的科学性、可行性、实效性中，使评估方案日趋完善。

总之，教学评估方案的设计过程也是一个决策过程，方案设计者就是决策者。坚持科学的决策原则，遵循科学的决策程序，是设计优化的评估方案的关键。

## 教学评估指标体系设计的方法

### 1. 教育目标分类学的方法是建立评估指标系统的基本方法

坚持与目标的一致性原则，是设计评估指标系统最基本的原则。所谓教育目标分类学的方法，就是借助若干中间过渡环节，先将教学的总体目标分解为次级目标，再将次级目标分解为指标，使教学目标成为指标而且有可测性的特点。

运用教育目标分类学的方法建立评估指标系统，必须坚持“教育的——逻辑的——心理的”原则。首先要考虑的是教育上的原则，即各类别之间的主要区分应在大体上反映教学工作者对教学行为表现所做的区分，这些区分应按照教学工作者特别是教师阐述教学目标的方法来揭示，在可能的范围

内，各类别之间的界限应与教学工作者特别是教师在安排或选择教学情境时所做的区分紧密联系。其次，指标体系的分类应该是逻辑的分类，尽可能地使各种术语精确些，并前后一致，同时每一类别应有符合逻辑的细分，这种细分，在必要和有用的情况下，可再加以合乎逻辑地明确地解说和细分。再次，指标体系的建立应该同有关公认的心理学的原理和理论相一致，特别是学业成就评估、教学水平评估、教师评估的指标系统，必须符合教师心理学、管理心理学、教师心理学的理论和原理。

### 2. 内涵分析法是改善指标系统质量的基本方法

系统内指标之间的相互独立性和指标系统的整体完备性，是设计指标系统的基本原则。内涵分析的方法是保证指标的独立性与整体完备性的基本方法，也是优化指标系统，提高指标系统质量的可靠途径。内涵分析有两重含义；其一指经过分析，将在同一层次相互重叠的、冗余的指标舍去、减轻评估工作的工作量，提高评估工作的可行性；同时通过剔除冗余指标，避免因对重复指标的重复评分而加大权重，提高评估工作的科学性。内涵分析的这一过程可借助“聚类分析”和“主成份分析”等多元统计方法进行。“聚类分析”的基本思想是从评估指标系统的测值中，找出能度量指标之间相似程度的统计量，构成一个对称的相似矩阵，进一步寻求各指标之间的相似程度，按相似程度的大小，将指标逐一归类。“主成份分析”是指当初拟的指标之间存在一定的关系时，可以通过初拟指标的线性组合，构成为数相对较少的不相关的相对独立的新的指标，代替初拟指标，而每个新指标都含有尽量多的初拟指标的信息。“聚类分析”和“主成份分析”都可借助电子计算机进行。内涵分析的第二重含义是，通过分析，将初拟指标中遗漏的指标及时补上，使指标系统能够全面地反映教学目标和教学管理目标，确保指标系统的整体完备性。总之，内涵分析的根本目的在于使指标系统内部不矛盾、不重复、不遗漏，以最少的指标科学、全面地反映教学目标和教学管理目标。

### 3. 从实际出发考虑特定指标的取舍和替代，是完善指标系统，提高指标系统可行性的基本方法

所谓从实际出发，就是要从教学的现状出发，舍弃那些很难达到的脱离实际的指标；所谓从实际出发，就是从可利用的人力、物力、财力、时间、信息、量化方法等现实条件出发，考虑指标的取舍替代，确保教学评估的可行性；所谓从实际出发，就是要从评估对象的具体特点出发，使指标系统具有一定的弹性；所谓从实际出发，就是要从评估实践中提炼每一评估指标的标准和估计指标的质量，修改和完善指标系统。

在指标系统设计的实践中，一般是专业人员运用教育目标分类学的方法初拟指标系统，经过试评之后，再根据试评的信息，借助多元统计分析方法进行内涵分析，优化指标系统。而在评估实践中，根据实际情况决定指标的取舍替代，修改和完善指标系统。

## 如何设计教学评估标准

### 教学评估标准

教学评估标准，是对教育活动效果和影响进行价值判断时所采用的准则和尺度，是对评估对象在具体的评估指标上的表现进行测量的质和量的统一体。

在我国的教学评估实践中，对教学活动及质量的价值判断，是通过测量教学评估指标体系中每项指标的目标达成度来实现的，一般是根据教学评估指标体系中最低层次的设定，用评估标准先对各项末层指标的状况进行测定，再按设计规定的要求和法则逐级综合，最后整理合成评估对象教学活动及质量水平的结果。由此可见，教学评估标准是测量教学评估指标系统中末级指标所反映的评估对象某一方面的教学活动和质量状况的准则和尺度。同时，在教学评估方案中，教学评估标准作为一个集合，是在质和量的统一上评估对象的各项指标进行价值判断的准则和尺度。

#### 1. 教学评估标准的要素构成

标准作为测量事物属性的尺度和准则，有其本身的要素结构，用途不同构成也不同。如测量事物长度的标准物（公尺）是由一定的物质材料和一定的量度标志构成。教学标准作为对教学水平进行价值判断的标准物，主要由以下几个要素构成。

（1）评估的强度和频率。评估的强度，是指达到教学评估指标体系项目要求的程度或达到各种规范化行为的优劣程度，也称“定性标准”。如在教学的等级性评估中，明确规定达到什么样的要求程度为优秀（A等），或良好（B等），或合格（C等）。

评估的频率，是指达到教学评估指标体系项目要求的数量或各种规范化行为的相对次数，也称“定量标准”。如巩固率、优秀率、合格率等等。

评估的强度和频率是具体的教学评估标准内容的主要组成部分。

（2）评估的标号和标度。评估的标号，是指不同的评估强度和评估频率的标记符号。其通常是用大写英文字母（如A、B、C、D）和数字（如1、2、3、4）来表示的，也用有汉字（如甲、乙、丙、丁）来表示。这些标号没有独立意义，只表示一种分类或等级。

评估的标度，是指评估时的测定单位标准，一般反映为等级代号，是教学评估标准的辅助成份。标度既可以出现在评估标准中，也可以出现在评估表格之中。其一般用两种方法表示：一种是定性形式的，如优、良、中、合格；一种是定量形式的，如数字表示。

除评估的强度、频率及标号、标度外，评估标准的对象也是构成评估标准的要素。在教学评估标准中，标准对象是教学评估指标体系中的最低层次的评估指标。

在具体的教学评估方案中，评估标准的强度和频率及标号和标度与教学评估的标准对象之间相互联系、互为补充，共同构成一个有机整体，形成一个质与量统一的价值判断的准则和尺度。

#### 2. 教学评估标准的内容

（1）条件标准。所谓条件标准是指依据教学评估指标体系中教学条件类指标所确定的评估标准。对教学评估来说，条件评估是一项基础性的前提评

估，是对“教学投入”的评估。在教学评估中，条件评估的核心是教师素质评估，因此，条件标准又称素质标准。对教师素质的评估，在其标准制订上就要体现出教师的品德素质、文化知识素质、业务素质以及身体素质等。

(2) 过程标准。所谓过程标准，是指依据教学评估指标体系中教学过程类指标所确定的评估标准。如对教师教学过程的评估，在其标准的制订上，要体现出备课、上课、作业批改、答题辅导、教书育人等方面的内容。

(3) 效果标准。所谓效果标准，指的是依据教学评估指标体系中效果类指标所确定的标准。如对教师教学效果的评估，在其标准的制订上，就要体现出教书效果、育人效果，以及社会效果等方面的内容。

“条件”是基础，“过程”是关键，“效果”是核心，三者有机结合，就构成了一个完整的教学评估标准体系。

## 设计教学评估标准的原则

教学评估标准的设计，必须在一定的原则指导下进行，以确保教学评估标准的科学性、客观性、有效性和操作性。

### 1. 科学性原则

设计教学评估的标准，首先要科学的揭示教学评估指标的内涵和本质，并紧扣教学评估指标的要求。其次，要体现出教学发展的时代特征，紧扣“素质教育”的质量要求。第三，教学评估标准的内容、评估标准设计的方法和技术，都要以科学性为指导原则。特别是教学评估标准的内容，要积极吸收国内外教育教学改革与发展的有益成果及设计教学评估标准的已有经验，既要尊重教学活动的内在规律，又要体现出时代的风貌。如在课堂教学质量评估中，不能单纯地考虑评“教”的标准，要把学生的主体参与意识和能力放在重要位置。创造主体参与的情境，强化主体参与意识，培养主体参与能力，激发学生学习的主动性、积极性、创造性，应当成为课堂教学评估标准的主要内容。第四，在评估标准的标度上，也要有科学的求实态度，需要相应地考虑指标的权重。

### 2. 操作性原则

教学评估的标准，是测量教学工作水平及质的尺子，必须方便实用，便于操作。不管是标准的内容，还是标准的形式，都要尽可能地简洁明确、通俗易懂，尽可能地用行为化、操作化的语言或符号表示，切忌抽象化、概念化，模棱两可。同时，教学评估的标准必须客观、准确、严谨，符合教学实际，从内容、文字技巧到评分比例及相应的标度都要严格推敲，切忌草率行事。教学评估标准的等级划分及量标，也不宜过细、过于琐碎，以免增加教学评估的难度而影响教学评估的操作性。

### 3. 协调性原则

教学评估的标准要具有科学性和可操作性，必须具备协调性。首先，评估标准必须具有整体性、层次性和相关性。整体性是指评估标准必须形成一个完整的体系；层次性是指评估标准体系结构合理，有完整的层次，各层次之间不能混乱；相关性是指评估的标准体系各层次有机联系，这种联系是有序的联系，而非互不相关内容的堆砌。整体性、层次性、相关性要求评估标准体系必须具有内在的逻辑上的协调一致性。第二，教学评估标准必须具有内容和形式的协调统一性。内容是指教学工作质量标准内在的要素的总和，

形式是把内容诸要素统一协调起来的内部结构或内容的表现方式。有什么样的内容就有什么样的形式，形式也具有能动地制约内容的功能，评估标准的内容和形式应该是辩证统一、协调一致的。同时，遵循协调性原则设计教学评估的标准时，要参照国家或上级教育部门颁布的有关教育教学的标准，要与之保持协调一致；同一教学评估方案中各项指标的评估标准要格调一致，评估标准的上、下限也要保持一致，尽量避免相互矛盾、彼此脱节；评估标准的等级与等级之间的档次、间隔要协调统一，不重叠或彼此距离不等。

## 设计教学评估标准的程序

教学评估标准的设计，是一项科学性强，技术性要求严格的工作，遵循规范的设计程序，是提高教学评估标准品质的前提。设计教学评估标准的程序是“拟定——修订——试行”。

### 1. 拟定

(1) 要成立教学评估标准拟定小组。拟定小组的成员要由富有现代教育科学知识、具有教育教学及评估工作经验，比较了解教育、教学实际情况的有关专家、教师及领导组成。

(2) 拟定小组要对教学评估方案的内容、指标体系及其内涵进行认真细致的调查研究，充分吸取某些有益的研究成果。

(3) 拟定小组要科学的分解教学评估标准的对象，寻找可作为测量评估标准对象尺度和准则的主要因素。要特别明确的是，评估标准的对象是教学评估指标体系最低层次上的指标，每项标准对象都包含若干项因素。其中有些因素可以作为测量标准对象的尺度，有些因素则不具备作为测量尺度的条件。可作为测量尺度的因素往往数量较多，没有必要把其全部作为测量的尺度，择其主要因素即可。

(4) 科学的确定标号和标度。标号是区分教学评估结果程度层次的符号，要按区分评估结果的精细程度确定；标度是区分教学评估结果幅度的代号，其与标号密切相关。如把数字 100 作为评估结果的满分值，若设五个标号(A、B、C、D、E)作为区分评估结果的层次的标志，每个层次间的标准幅度平均为 20；若用叙述性语言描述作为标度，最好按标号的小数设计叙述要点，以便于区分程度层次。这样，经过拟定组讨论，在意见基本一致的情况下，确定专人执笔，起草教学评估标准的草案，形成教学评估标准的初稿。初稿由拟订小组成员集体审阅、讨论，及时修正。

### 2. 修订

修订工作是保证教学评估标准科学性和可操作性的重要步骤。

(1) 要把由拟定组起草的教学评估标准的初稿交给教学评估专家、实际的教育教学工作者、教学研究人员、教学管理及领导者审议、讨论，对草案的方向性、科学性、全面性、操作性、鉴别性进行探讨和论证，充分听取他们的意见和建议。在条件许可的情况下，还可以把对象扩大到学区及学生，广泛听取有关人士的意见，多角度、多侧面地获取反馈信息，以便于集思广益。特别要珍视反对意见，切忌“闻争则怒”。

(2) 拟定组要对所征询的信息进行综合分析讨论，取长补短，修订草案。特别是对一些争议较大的项目评估标准，可以采用研讨会的形式充分探讨和交流，形成共识。

### 3. 试行

经过修订的教学评估标准，可以成为较为成型的标准，定稿试行。为慎重起见，定稿试用之前，评估组织者还应聘请有关专家讨论修订稿中是否还存在着某些问题，再试行。试行的步骤是，先用教学评估标准去评估和测量一些评估对象，获取评估结果，并对其进行效度、信度和鉴别力等方面的分析，或者用经验的方法进行验证。再根据验证的结果，对那些达不到质量要求的评估标准进行修改，对那些质量实在太差的评估标准要重新拟定，达到质量要求后，方可正式使用。这里要强调的是，因为是试行，试行的对象样本应当按照随机抽样的规则抽取样本，力求使试行样本具有代表性。

## 教学评估标准的类别和设计方法

教学评估的标准，可按不同标志分成若干种类。从教学评估标准能测量到教学工作水平及质量高低的程度的表示形式，可分为用数字与符号表示的评估标准和用语言描述表示的评估标准。以数字和符号表示测量结果的评估标准，一般用自然数列作为区分测量结果的标度，可以用整数（如规定 A 等为 100~85 分、B 等为 84~75 分、C 等为 74~60 分、D 等为 59~0 分），小数（如 A 等为 1.0~0.85，B 等为 0.84~0.75，C 等为 0.74~0.60，D 等为 0.59~0）表示，有时也可以整数与小数合用，还可以用标准分数表示测量结果的标度。用语言描述的评估标准，是一种用描述性的语言作为区分测量结果标度的评估标准。用来作为标度的语言可以是程度性语言（如完全达到规定要求的 A 等，绝大部分达到规定要求的为 B 等，大部分达到规定要求的为 C 等，达到规定要求的为 D 等），也可以是叙述性语言（如双基目标明确，有思想教育要求、有培养与发展能力与智力的措施）。按照测量的尺度，教学评估的标准，可以分为类别标准、等级标准、等距标准、比率标准和隶属度标准。按照标准的参照点，又可以分为绝对标准和相对标准。

在我国的教育和评估实践中，常用的评估标准主要有评语标准、期望行为式标准，隶属度式标准和分段式标准四种，其设计方法介绍如下。

### 1. 评语式标准的设计

所谓评语式标准，是用类似“评语”式的语言对教学评估指标体系中每项指标的评估标准进行阐述。一般来说，这种标准有期望评语式标准和积分评语式标准两种类型。

（1）期望评语式标准的设计。所谓期望评语式标准，是对教学评估指标体系中每项评估指标的标准采用期望、理想式的语言加以描述，并对所描述的要点按照一定的规则赋值，然后按等级逐级评分。

这种方法的特点是，只有设计出所期望的最理想的最高等级（上限）作为评估标准，以这个最高等级的达成程度进行评分。由于其他等级没有具体的评估标准，只能根据最高等级的要求推及，其分寸较难把握。期望评语式评估标准多用于鉴定“合格”的评估领域。陕西省“两基”（基本普及九年义务教育，基本扫除青壮年文盲）督导评估验收采用的就是这种标准体系。

（2）积分评语标准的设计。所谓积分评语式标准，是先将教学评估指标体系的每项指标按其重要程度确定一个分值，并将这个分值随每项指标的具体化分解而分配到每一项评估要点中去，使这些具体的项目要点的分值总和，与该项指标的总分相等。科学、合理地分解指标要点，适宜、恰当地分

配分值，是设计积语分评式教学评估标准的关键。

## 2. 期望行为式教学评估标准的设计

所谓期望行为式评估标准的设计，是以期望的最理想要求与行为要求为最高等级，逐级而下划分，以最不期望的行为要求为最低等级，从而设计出教学评估标准系统的方法。这种设计方法的优点是：评估标准构成一个完整的等级系列，每个等级中都有相应的反映该评估指标状况和水平的定性描述和定量数值，便于将评估对象与评估标准相对照，以按其符合某等级标准的符合程度定等计分，具有较强的操作性。期望行为式教学评估标准的设计模式与实例如表所示。

设计期望行为式教学评估标准时，首先要明确教学评估指标体系中各项具体指标的内涵，全面分析，深刻了解并掌握该项指标所要反映的具体内容及其深度与广度，使评估标准的等级内容明确、清晰。其次，在明确评估指标内涵要点的基础上，选定最适合表现该指标内涵期望要求的关键性的行为特点，并用相应的最恰当的行为词语表示出来。这种反映或表现某项指标内涵期望要求的关键行为的词语，应尽可能地避免使用过于刺激性的词语，力求其科学、稳妥，具有客观性可接受性。

## 3. 隶属度式教学评估标准的估计

所谓隶属度式标准，是运用模糊数学的隶属函数为标度来设计教学评估标准体系的方法。隶属度式评估标准，就其内容而言，仍是评语式等级标准，不过是采用隶属函数为标度，通过评估对象教学目标的成适度取值于 $[0, 1]$ 来判定等级的评定值。

以隶属度式教学评估标准的设计有分段法和全域法两种。分级法是规定教学评估各项指标的各个等级隶属度的范围(或点)，如A等为1.0或 $1.0 \sim 0.85$ ，B等为0.7或 $0.84 \sim 0.60$ ，C等为0.4或 $0.59 \sim 0$ 。全域法与分段法相对，其不规定某一教学评估、指标各个等级的隶属度的范围(或点)，每个评估等级都可以在 $[0, 1]$ 全域范围内选择，以表示评估对象所属等级及其程度的高低，如某指标达成度属于A等的隶属度为0.4属于B的隶属度为0.8，属于C等的隶属度为0.1等。在我国的教学评估实践中，一般采用分段法设计隶属度式教学评估标准系统。

## 4. 分段式教学评估标准的设计

所谓分段式评估标准，指的是将教学评估的每项指标分为若干个等级，把指派到该项指标的分数(含权重)等距离地分配到各个等级中，然后再把每一等级的分值划分为若干个小档次。这种教学评估标准的制定方法，就是分段式标准方法。这种方法的具体步骤是先确定等级数，然后再选用适切的标号(如A、B、C，或优、良、中，或甲、乙、丙等)。接着，再将该项指标的分数等距离地分配到各个级别中，每一级别的分数再划分为若干个小档次(参下表)，例如，把教学方向的分值定为6分，按优、良、中、差分为四个等级，进一步将每一等级按等距把分值分配成6 4.8 3.6 2.4，再把每一等级的上限下限之差作为该等级的评估幅度。这种幅度可以相应地再分配成几个档次，如4.9、5.4、6这样三个档次，也可以分为4.9~5.2、5.3~5.6、5.7~6这样三个档次。

教学评估标准制定的分段式标准方法是一种比较简易的方法。其优点在于：它的分值的分档较细，能够有区别地、较好地反映出被评估对象间的差异，使用起来也比较简便。使用这种制定标准的方法存在的问题是：由于它

的分档较细，往往存在着小数，这样就给实际评估与计分统计增添了一定的麻烦，较为繁琐一些。

## 如何设计教学评估的评定量表

### 数值量表法

所谓数值量表法，指的是通过对数学评估标准体系中的每一个等级标准都标上分数而形成的一种评定量表的方法。这种量表在形式上一般采用5点法、7点法和3点法。典型的5点法划分为：非常同意；同意；无所谓；不同意；很不同意。

它表示了五种反应的连续体。这种量表常常是将答案的定序尺度用单纯的数字表示出来，回答者可在标线的任何一点（数字）上划上标记以表示自己的评定。例如，当评定项目为“教学中是否尽可能使学生运用多种感官学习”时，可将回答的尺度设计为数值量表的制定，相对来说比较简易，通常情况下是在相应等级的评语标准下，标上一个等级分数即可。

这种形式的评定量表，由于它具有量化的特点，不仅使它的结果更容易被制成量表，便于输入电子计算机进行现代技术手段的资料统计与处理，而且使用起来也比较简易、方便。因此，数值量表成为实际教学评估中，最为流行的评定量表之一。制定数值量表时需要注意的是，要将第一等级中所标的数字等距离拉开，以便有所区别。使用数值量表进行评估时，需要适当考虑并尽可能地避免可能出现的中间化趋势。

设计和编制数值量表时，各刻度或数值之间的间隔应该保持相等，否则会使回答者发生困惑，不好回答，并会增大某些间距大的类别选答的可能性，降低类别小的类别的选答可能性。在项目问题设计时，还应积极地运用积极与消极的陈述或正面与反面的陈述，以便考察评估者回答问题的一致性。

修改数值量表时，一般需要检验各项目标与总分之间的相关值，由于这一相关值往往高估需要用下式进行校正。

$$r = \frac{r_{Ti} \sigma_T - \sigma_i}{\sqrt{\sigma_T^2 + \sigma_i^2 - \sigma_{Ti} \alpha \sigma_T}}$$

式中： $r_{Ti}$ ——项目与总分数之间的相关； $\sigma_T$ ——总分的标准差； $\sigma_i$ ——项目分数的标准差使用数值量表进行评估时，需要适当考虑并尽可能地避免可能出现的中间化趋势。

现在希望您能提供有关您的一些背景资料，再次衷心感谢您的合作！

您的姓名\_\_\_\_性别\_\_\_\_年龄\_\_\_\_教龄\_\_\_\_职称\_\_\_\_职务\_\_\_\_在此学校的工作年数学校总人数\_\_\_\_学校为哪一级重点\_\_\_\_所教年级（1，2，3，4，5，6）\_\_\_\_现在所教最大班级的学生人数现所教课程（1，语文；2，数学；3，政治；4，自然；5，英语；6，音体美\_\_\_\_

您成为一名教师的原因（1. 父亲或母亲是这种职业；2.

受一位老师的影响；3. 是我自己的理想；4. 自然而然；5. 别无其他选择）

\_\_\_\_\_

### 图示量表法

所谓图示量表法，是指将教学评估标准体系的量标设计成图样，通过对图示尺度的划号标记确定评估的结果，并用图样表示教学评估结果的方法。

图示量表是以图示的形式形象直观地表示教学评估的结果，便于同教学质量的常模式评估对象之间的相互比较。但是，图示量表法标线上刻度的设计比较困难，在教学评估中把握其准确度也比较困难，特别是对“中间状态”更难把握。因此，在设计教学评估的图示量表时，要特别注意标线刻度的数量及其区分的界线问题，一般请专家设计。用图示量表法设计的教学评估定量表的模式，如下图所示的实例。

图示量表和前述的数值量表统称总加量表，是由美国评估测量学家李克特(Likert, R.A) 1932年列制的，故又称“李克特量表”。图示量表是用标线来表示量表总数的，数值量表是用单纯的数字表示量表点数的。无论是使用图示量表的形式，还是数值量表的形式，各刻度或各数字之间的间隔应该相等；在设计项目时，还应积极地运用积极与消极的陈述和正面与反面的陈述，并重视量表的试测与修订工作。

### 谓词量表法

谓词是动词和形容词的合称。谓词量表法主要是应用一组有明显等级区别特征的谓词(或谓词性短语或形容词)序列法标明评估指标的各项具体要求的实现状况而形成的一种评定量表的方法亦称形容词量表法。根据量表的具体内容可以把谓词量表分成两种：一种是在提出的问题项目之后，用一组系列性的形容词提供各种不同程度的答案，评定者在自己同意的形容词或短语上画圈，打勾或做上其他记号，以表示自己的评定。这种形容词量表称为“形容词评定量表”。另一种是形容词或短语构成了量表的全部成分，它们都是描述教学某个方面的形容词，评估者从这些形容词中画出符合评估对象的那些形容词即可。这种形容词量表称为“形容词检查量表”或“形容词核对量表”。制定这种评定量表的方法也比较简易，根据不同的指标项目要求选取一组基本上有差距的谓词或谓词性短语等，作为不同等级的标准要求即可。需要注意的是，这组有序列的词语本身一定要有较为明显的彼此之间有相同间隔的序列词语。但由于在具体的制表过程中，有时有可能一时难以找到完全符合标准要求的词语，因此，也允许有一部分彼此之间的差别不十分明显的评定词语存在。但这部分词语起码有比较容易辨别的细微差别。谓词量表存在的问题是：由于有时有的评定语缺乏等距离的间隔，评估时有可能形成偏上或偏下的现象。

形容词评定量表的设计与编制的核心是选择出质量较高的量表项目，根据不同项目的要求，选取一组基本上有距离差异的形容词或短语作为不同评估等级的标准。同时，由于教学评估量表中用到的形容词量比较固定，因而也是比较容易确定的。例如评估标准的等级要求可用下列词语来表示：极好、很好、好、尚好、不好；优、良、中、差；极重要、重要、较重要、不重要、不知道；非常同意、同意、无所谓、不同意、极不同意等。这种评定量表的特点是：使用方便，操作简单，只要在自己同意的栏目中圈画上规定的记号即可。用形容词量表法编制的教学评估量表的模式如“对教师职业态度的自我评定表”(根据黄天中《人事心理学》，三民书局1976年版，第162—163页的“工作满意度问卷”改编)。

形容词核对量表设计的重点是在所选用形容词的编制与筛选上。在设计初期，需要编拟出大量的有关形容词，越多越好。然后聘请有关教学专家和

教学评估专家进行筛选，也可以在试用中不断发现和剔除不适合的形容词。形容词核对量表所选用的形容词数目的范围一般是十几个至几百个。

### 行为量表法

所谓行为量表法，是根据教学评估指标体系的要求，把教学评估标准系统制定成一定的行为表现要求而形成的一种评定量表方法。行为量表的特点是：通过对评估对象的行为观察进行评定，评定结果较为可靠；评估时有一定的具体化、行为化的标准要求，可以避免因评估标准过于空泛而使评估结果出现某种程序上的偏斜。但是，行为量表的编制工作繁重，难度也较大；对评估者的素质要求较高，必须具备观察评估对象教学行为表现的能力与技巧。

行为量表设计的一般程序是：首先，量表设计者对就要评估的教学现象进行比较深入的分析研究，从中分解出最能体现这一教学现象的教学行为及其有效与无效性。其次，将这些具有一定代表特征的有效与无效行为分类，并用通俗易懂，流畅简洁的语言加以描述。最后，将这些语言描述条目归入一个统一设置的量表，做出一定的教学评估评定量表的标度。

## 如何搜集教学评估信息

获取评估信息是教学评估中的一项重要工作，只有获得评估所需要的大量信息，才有可能得出比较科学的教学评估结论。教学评估信息的获取一般从教学测量和教育统计两种渠道搜集，而其基本方法较多。经常应用的主要有资料法、观察法、调查法包括问卷法、测验法、谈话法、个案法、实验法、统计分析法等等。这些方法均有各自的特点和应用范围，但在使用上，又往往是互相联系、互相配合的。所以，在教学评估中，信息获取的基本方法是联合运用的；在运用中又是综合了抽样技术、分类测量、位次测量和量化技术。

### 抽样法

教学评估的对象和范围（总体）是相当广泛的，要对所有评估对象及其属性进行测量是很困难的，也没有必要。为了使测量及其信息的统计推断正确可靠，就必须遵循抽样的原理和程序，抽取反映评估对象基本属性和特征的具有代表性的样本。而样本的代表性，不仅涉及抽样的程序、方式，而且还涉及到样本的容量。

抽样法指的是通过对从被评估对象的总体情况中抽出的能够代表总体情况的部分样本的观察分析，推断出被评估对象的总体情况的方法。

#### 1. 抽样的一般程序

“确定抽样总体 确定样本数 确定取样范围 抽取样本 评估样本的代表性”，这是抽样的一般程序。

首先，根据本次评估的目的和任务，明确所要测量属性的范围，确定取样的对象，即抽样总体。

第二，在抽样总体内收集界定其全部个案，并予以编号，形成抽样框架，明确总体内抽样的范围。

第三，根据抽样框架内个案的差异程度及其本次评估所要求的测量精确水平，确定合算的最佳样本数。

第四，采用科学的抽样方法及其组合，从总体中抽取样本，确定测量的对象。

第五，运用计算法、比较法等方式对抽取样本的代表性进行评估，尽量减少抽样误差。

#### 2. 抽样方式的样本容量

（1）抽样的基本方式。抽样的基本方式分为两大类：随机抽样和非随机抽样。

若总体中每个个体被抽取的机会是均等的，则称为随机抽样，随机抽样必须遵循一个原则，一是总体中的每个单位都有同等的被抽中的机会。抽取应当是完全客观的，不能依据某个人的主观意志加以选择。二是所选择的人或单位彼此之间没有牵连，每个人或单位的选择都是独立的。它的特点主要是：

从总体中随机抽出容量为 $n$ 的一切可能样本的平均数（ $\bar{X}$ ）之平均数  $E(\bar{X})$  等于总体平均数（ $u$ ）即

$$E(\bar{X}) = u \quad (6 \cdot 15)$$

容量为  $n$  的样本平均数在抽样分布（某一统计量的分布）上的标准差（ $\bar{\sigma}_x$ ），等于总体标准差（ $\sigma$ ）除以  $n$  的方根。即

$$\bar{\sigma}_x = \frac{\sigma}{\sqrt{n}} \quad (6 \cdot 16)$$

从正态总体中，随机抽取容量为  $n$  的一切可能样本平均数的分布也是正态分布。

虽然总体不呈正态分布。若样本容量  $n$  较大，反映总体平均数  $u$  和标准差的样本平均数的抽样分布，也接近于正态分布。

随机抽样的抽样分布的上述特点，决定了它是教学评估测量中科学的、也是常用的抽样方式。

非随机抽样是根据主客观条件而主观选择样本的方式，又称判断抽样。这种方式虽然有省人、省时、省物、易实施的优点，但科学性较差，不能保证样本的代表性，故在教学评估中很少应用。

（2）抽样的基本方法和样本容量。随机抽样的主要方法有单纯随机抽样、机械或等距抽样、分层抽样、整群抽样和阶段抽样，各种方法都有其相应的样本容量计算公式。

单纯随机抽样（简单随机抽样）。如果评估总体中每个评估对象（个体）被抽到的机会是均等的（即抽样的随机性），并且在抽取一个对象之后总体内的成分不变（抽样的独立性），这种抽样方法称为单纯随机抽样。

抽样的随机性原则可通过抽签法，随机数码表法或摇号机摇号法来保证。抽签法是指先将评估总体中的每一个个体都编上号码，再将每个号码写在标签上，将标签充分混合后，从中抽取几个签，与号码相应的个体的集体就成为样本。随机数码表法是借助随机数码表随机抽取样本的方法。

抽样的独立性和保证，对于有限的评估总体要用放回抽样方法，对于无限总体则不要求放回。在教育评估的测量抽样中，放回抽样往往不可能，因此在应用时即使是有限总体，若总体中的个体数较多，而且总体内的个体数是样本容易数的 10—20 倍时，也可视为无限总体，实行不放回抽样。

单纯随机抽样中样本容量的确定因测定目标和总体的有限性而不同。

机械抽样（等距抽样）。把总体中的所有个体按照一定的标志排列编号，然后以固定的顺序和间隔取样的方法，称为机械抽样。在教学评估测量中，可将评估对象总体单位  $N$  按照一定的标志进行排队编号，并将  $N$  划分成相等的几个单位，使  $K=N/n$ ，然后随机抽取  $i, i+K, i+2K, \dots, i+(n-1)K$  共  $n$  个个体组成样本。

机械抽样比单纯随机抽样能够保证抽到的个体在总体中的分布比较均匀，抽样的独立性较强；单纯随机抽样比机械抽样能够保证总体中个体被抽到的机会均等，即抽样的随机性较强；单纯随机抽样和机械抽样二者也可以结合使用。机械抽样中样本容量的确定可按单纯随机抽样的公式进行计算。

机械抽样方法一般不适用于大容量的评估总体；当评估总体的个体类别间悬殊时，机械抽样抽取的样本常常缺乏代表性；机械抽样的间隔接近评估总体中个体类别分布的间隔时，常常形成周期性的偏差。这些都是教学评估测量中应用机械抽样需要注意的基本方面。

分层抽样。按与评估内容有关的因素指标等标志先将评估总体加以分组（分层），然后根据样本容量与总体的比率，从各层中进行单纯随机抽样或机械抽样的抽样方法，称为分层抽样。例如，对评估对象为  $N$  的总体中，

拟取  $n$  个个体为样本，可根据一定的标准将  $N$  个个体分成优 ( $N_1$  个)、良 ( $N_2$  个)、中 ( $N_3$  个)、差 ( $N_4$  个) 四层，然后从各部分 (层) 中用单纯随机抽样或机械抽样的方法，各抽  $\frac{n}{N}$ ，即从优、良、中、差等中分别抽取  $N \times \frac{n}{N}$  ( $i=1, 2, 3, 4$ ) 个个体，组成一个评估测量样本。

运用分层抽样抽取教学评估的测量样本时，要尽力缩小各层组内的差异，增大层组间的差异；同时层组的划分也不宜过细，以免层组内个体数目过少而无法抽样；再次划分层次的标准必须明确，以免混淆或遗漏。

**整群抽样。**分层抽样划组层为类，其作用是尽量缩小总体，使总体的变异减少，而抽取的基本单位仍然是总体中的个体。整群抽样是将评估对象总体的各个个体划分成若干群。然后以群为单位从中随机抽取一些群而组成样本的方法。即整群抽样划组层为群，群的作用是扩大单位，抽取的单位不再是总体单位而是群。例如要测量某市某年级数学高考成绩，可以以学校为单位进行抽样。

整群抽样的主要缺点是样本分布的均匀性较差，误差也较其他抽样方法大。为了弥补这种缺陷，增强样本对总体的代表性，可以与分层抽样相结合，例如，先按一定的标准把全地区所有学校分成几部分，然后再根据本容量与总体中个体的比率，从各部分中抽取若干学校，组成整群样本。

**阶段抽样。**当评估所要测量的总体很大时，在实践中常采用阶段抽样。首先将评估总体分为  $A$  组，每组包含  $B_i$  个单位。从  $A$  组中随机抽取  $a$  组，再分别从抽中的  $a$  组的各组中随机抽取  $n_i$  个单位，构成一个样本，这种抽样方法就是阶段抽样中的两阶段抽样，其中总体单位数  $N=B_1 + B_2 + B_3 + \dots + B_A$ ，各组的单位数  $B_i$  可以相等，也可以不相等；样本单位数  $n=n_1 + n_2 + \dots + n_a$ ，各组抽取的样本单位数可以相等，也可以不等。多阶段抽样的原理与两阶段抽样的原理相同。在阶段抽样中，由于每一阶段的抽样都会产生误差，所以阶段越多，误差越大，经多阶段抽取的样本的代表性越差。因此运用阶段抽样时，要特别谨慎，尽量提高各阶段抽样的精确度，严格控制整群、分层、机构、单纯随机抽样的误差限度。防止误差传递造成阶段抽样的失败。

**目的抽样。**目的抽样是根据特定的目的，有针对性的随机抽取样本。虽然从广义上说，目的抽样也是随机抽样的一部分，但它与一般的随机抽样却有所不同。它强调抽样的针对性与目的性，而不是泛泛地任意抽取样本。这种抽样方法在实行时，先是要根据一定的抽样目的与需要，挑选出符合抽样目的与需要的对象，然后再在这些已挑选出的对象中进行抽样。如调查一个学校优秀教师的情况，事先要先把那些符合优秀教师条件的人挑选出来，然后再从这些人中抽样。目的抽样由于是在一定的符合抽样目的的范围进行，所以用目的抽样法抽出来的样本都具有一定的代表性，可以免去其他随机抽样的任意性与偶然性因素，便于集中精力，取得调查的实际效益。因此，目的抽样就成为抽样法搜集教学评估资料信息的一种行之有效的重要办法。

## 查阅资料法

### 1. 查阅资料法的概念

所谓查阅资料法，指的是通过查阅评估对象的有关书面材料，来分析、

了解和把握其教学的有关状况，以便进行教学评估的方法。查阅资料在教学评估中占有十分重要的地位。当教学评估的对象确定之后，按照教学评估指标体系的要求，首先要明确评估因素或未级指标的内涵；然后全面、真实、详细地收集内涵要求所需的资料、文件和统计资料、报表等。并对教学质量、教学研究、师资水平、教学条件、教学管理水平，以及涉及的有关资料，尽可能收集齐全。通过检索、整理、阅读、统计以至浓缩，从而获取定性和定量的教学评估信息。

## 2. 查阅资料的内容

教学评估中所要查阅的资料是较多的，其基本内容主要有：

(1) 教师类资料。教师是教学活动的主体，是教学活动的计划者和组织者，因此，教师类资料的查阅在教学评估活动中十分重要。关于教师类资料，主要有教师情况卡片和教学情况表。教师情况卡片一般有两种。一种是教学用卡，有卡片形式，也有表格形式，它通常掌握在教务部门，作为了解教师水平情况与教学工作表现情况以及效果考评时使用。这种教师情况的记录材料，是教学评估资料收集的主要内容。另一种是掌握在学校人事部门的人事档案卡，它是了解教师全面情况的最基本的原始记录材料。教学情况表一般也有两种。一种是教学活动记录册，主要记载的是教师的教学情况、授课的出勤情况、教学科研情况等。另一种是教学进度表，主要记录教师在某一学期或学年的教学计划安排情况，实际教学内容，以及教学进度等。教学情况表所记录的内容，是教学评估资料收集的重点内容之一。除此之外，教师的课时计划、教案、教学试卷、试卷分析、工作总结，以及有关鉴定等，都可成为教学评估的有益参考资料。

(2) 学生类资料。学生是学习的主体，是教学发展的主体，是教学效果的体现者，因此，学生类资料是教学评估活动不可缺少的资料。学校中可供查阅的学生类资料主要有学生情况卡，学生出勤簿、学生成绩册等。此外，学生的作业、作品、考卷、其他情况检查表等，都可以作为对学生情况进行评估的有益参考资料。

(3) 其他类资料。除教师类资料和学生类资料外，一些与教学有关的资料，如教学设备资料，图书情报资料、教辅工作资料，教师工作变动情况资料、教师科研资料、学生奖学金获得情况的资料等经常性资料和一些专题调查与实验研究资料，都可以作为教学评估资料收集和查阅的资料。

## 3. 资料查阅的具体方式

在教学评估中，由于需要查阅的资料的量较查阅的面要广泛，使教学评估的工作量增大。为了用较短的时间查阅较多的文献和资料，要采用相应的查阅方式。

(1) 带题求解法。在教学评估中，查阅资料的目的是为了获得进行价值判断的评估信息。因此，带着教学评估指标体系及评估标准的内涵，有针对性的分工查阅，可以提高查阅资料的效率。

(2) 牛食鲸吞法。对于那些对教学评估价值较高的资料，要象牛吃草一样反刍，反复阅读，反复查找，力争弄清楚；对于那些对教学评估有一定参考价值的资料，可以象鲸吞食那样，只求知其大概。“牛食”与“鲸吞”结合，可以避免肤浅和寡闻，同时又提高了资料查阅的效率。

在实际查阅资料的过程中，牛食鲸吞法和带题求解法应结合使用。

## 4. 查阅资料的注意事项

资料法是具体的教学评估实施中的一种重要的搜集资料的方法。通过对这些资料的具体分析与比较，可以对评估对象的教学情况做出相应的判断，便于对评估对象做出确切的评判。使用查阅资料法搜集教学评估的信息时，要注意：

（1）目的要明确。在查阅以前，要拟定资料查阅的范围和类别，并根据教学评估指标及标准的内涵有目的的查阅。对那些与教学评估关系不密切的资料，要舍得放弃。资料、文献要愈查愈少，愈阅愈薄。

（2）分工查阅，分类整理。在查阅资料的过程中，要根据资料的性质和类别有针对性的分工查阅，并作到边查阅、边分类、边统计。

（3）分析核证。对资料中出现的矛盾信息，要具体分析，进行核证，必要时可采用开座谈会，个人专访、实测、问卷等其他方法予以核实和验证。切忌照抄照搬，掺入伪信息，影响教学评估的效度和信度。

此外，对教学评估资料的处理，不要孤立地从某一资料出发，做出轻率的、绝对的解释，而要全面、客观地对待资料；对查到的重要的资料信息，也要注明出处，以便在发现疑问时查找，从而提高教学评估工作的效率。

## 观察法

所谓观察法，是教学评估者根据教学评估指标体系和标准体系的要求，有目的、有计划地对评估对象在自然状态或受控状态下的行为表现进行系统、深入的观察，以获取教学评估信息的方法。它是由观察获得信息，经思维进行加工，做出评定判断的完整过程；是各级各类学校教学评估中应用很广的一种方法。观察法虽然易受评估者的态度、情感等主观因素的影响，但是依靠这种方法往往可以获得许多有价值的第一手资料，而这些信息是运用其他方法特别是量化方法难以得到的。观察法所获得的大量的感性材料可以为理性分析提供充实的事实依据，帮助评估者进行正确有效的评估判断。

### 1. 观察法的种类

常用的观察法主要包括自然观察法、选择观察法和控制观察法等。

（1）自然观察法。自然观察是在教学的自然常态下对评估对象表现出的言行进行的观察，是一种不加任何控制的日常观察，它涉及的时间长、范围广、对象大，容易观察到评估对象某些内心自然流露情绪状态及行为习惯。自然状态下的言行是评估对象真实的思想反映，因此，用自然观察法获得的教学评估信息，具有较强的真实性。自然观察法又可以分为两种：直接观察与间接观察。直接观察是直接利用评估者的感官去获取评估对象有关的资料信息；间接观察是借助于某些视听设备（如照相机、录相机、录音机）的帮助去获取评估对象的有关信息资料。

（2）选择观察法。选择观察法是在某一特定的时间和场合内对评估对象有选择地观察而获取评估信息的方法，亦称“时间样本观察法”或“情境样本观察法”。它的主要目的，是要获取被观察对象的某一行为目标的具体情况资料。其方法是选择或设置某一特定的情境场景，从中观察被观察者的行为活动、具体的反映。如要评估学生的智力情况，就可以在具体的竞赛中、日常游戏中，乃至科研、实验中，巧作安排进行观察。例如，要观察某教师的教育教学水平，就可选择某一堂课，通过听课观察教师的教学情况，或选择教学的某一环节（如教学准备工作）进行有重点的观察。这种观察方法对

时间和范围作了限制，便于突出重点，具有较强的针对性。

(3) 实验观察法。实验观察法是在确定的观察范围内，通过严密的组织和控制，促使评估对象对某种条件要求做出反应，以获取所需要的教学评估资料的方法，亦称“特定事件观察法”或“情景样本观察法”。

观察法还可以按照观察者与被观察者之间的关系分为参与性观察与非参与性观察；按照观察的时间分为定期观察和长期观察；按照观察的主体分为自我观察和他人观察等等。

## 2. 观察法的步骤

运用观察法获取教学评估的信息资料，一般分为准备、观察、整理三个阶段。

观察的准备阶段的主要工作是：大略调查和试探性观察；确定观察的目的、中心范围；设计观察计划，进行合理分工；确定观察的手段、次数、密度、记录方法和标准。在观察的准备阶段，其中心工作是设计观察记录表。一般来说，设计观察记录表主要要考虑以下几个方面：

(1) 明确“究竟观察什么”。许多观察之所以失败，原因之一就是没有十分明确具体的观察目标。观察者必须明确对所观察行为的定义与分类，给出所观察行为的相当明确而不含糊的操作性定义。

(2) 确定观察的“行为单元”。所谓“行为单元”，即观察记录中所用的行为成分的大小。“行为单元”的大小，与评估的目的、评估的阶段以及观察者的经验等都有关系，设计者应根据不同的观察任务，灵活运用。

(3) 考虑观察者进行推论的程度。一般来说，观察之后都要求观察者做出较大程度的推论，大多数的教学评估要求做出较高程度的推论。在设计观察的程序和内容时，应使推论程度保持在适中的水平上。

(4) 设计观察记录表时，还要考虑不同的观察取样所产生的影响。下面是用于评估教师与学生在课堂教学中相互作用的观察记录表。

实施阶段即实施观察计划的过程，包括人员观察和工具观察。此阶段的主要工作是要做好观察记录，把观察到的评估对象的实际的行为表现，如实地写到观察记录表上。

整理阶段主要是对观察到的评估信息，进行整理、统计、分析，推论并根据教学评估指标体系和标准体系的要求，行价值判断和评定。

观察者姓名\_\_\_\_\_

被观察教师姓名\_\_\_\_\_

观察事项：教师与学生相互作用

说明：请每隔 15 秒钟扫视一下教室，并使用下面的代号进行记录

- (a) 教师对全班提问；
- (b) 教师对某对个学生提问；
- (c) 请学生做出反应/反应；
- (d) 不要求学生做出反应；
- (e) 消极对待学生的反应；
- (f) 其他。

## 3. 注意事项

(1) 要有充分准备。教学评估资料搜集中所使用的观察法，同一般的、日常生活意义上的观察有所不同，是一种科学意义上的观察，它必须根据评估的需要，确定观察的目的、制订观察的计划、提纲，明确观察的中心、内

容和范围，选择适当的观察方式方法。观察者应当训练有素，能够准确地从纷繁的各种表现中捕捉有价值的信息资料。

(2) 既要注意观察全面情况，又要集中注意力观察好有代表性的焦点情况，保证观察内容与观察目的的一致性。

(3) 认真地做好观察记录。观察所获得的教学信息要全面真实的记录，最好记录在较规范的“观察记录表”上。记录的方法还有量表法，录音录像、计算机储存等。

(4) 观察中要尽可能的克服主观意志的干扰；也要尽量保持观察者的自然常态，以免的信息的干扰。

## 谈话法

所谓谈话法，是指评估者通过与评估对象面对面的谈话，以口头信息沟通的途径直接获取教学评估信息的方法，亦称访谈法、访问法、交谈法、面谈法、接谈法等。在教学评估工作中，谈话法是一种相当简便而传统的方法。这种方法，不受评估对象文字理解水平的限制，通常适用于任何文化程度的评估对象。由于在交谈过程中评估者与评估对象面对面的直接接触，可以对问题的含义及回答不清楚的地方，通过进一步的询问和解释得到澄清，并且容易深入了解，认识评估对象心理活动的状态和某些真实想法，获得其他间接、定量方法难以获得的信息。同时，评估者可以在与评估对象的交谈中，有效地控制谈话的情境，提高问题的解答率。但是，这种方法与其他类型的定性评估方法一样，容易受主观因素的干扰和影响；另外，此种方法比其他方法更耗费时间和人力。

### 1. 谈话法的种类

谈话法按照评估者的提问及谈话对象的回答结构方式，可以分为不同的类型。这里所说的“结构方式”就是指谈话的设计者事先对问题及其回答方式规定的程度。“结构谈话”是一种有指导性的、正式的、事先决定了提问的问题项目及回答的可能性的谈话方式，谈话的过程结构严密、层次分明，具有固定的谈话模式，整个谈话过程，谈话对象犹如作了一次口头的问卷。

“无结构谈话”则是一种非指导性的、非正式的、自由提问和自由问答的谈话形式，谈话过程结构比较松散，层次交错，气氛活跃，没有一个固定模式。

### 2. 谈话法中的问题设计

谈话法是评估者通过跟学校管理者、教师、学生及有关人员对话、交谈或讨论，从而直接收集教学的有关信息，并以此为依据进行教学评估的一种方法。其中的“谈话”并非日常生活中的谈天或者一般的交谈，它有一套科学的编制设计原则。

(1) 要依据谈话的类型来设计问题及其相应的结构与形式。

(2) 设计问题时要充分考虑影响谈话成功与否的三个要素。

信息的可及性，评估者需要搜集的信息，谈话对象是否能够完全提供。

问题的认知因素，即所提问题，谈话对象是否能够明白和理解，尽量避免那些具有双重含义或记性要求甚高的问题。

谈话对象的动机，即要注意那些与社会规范与要求相关的问题以及可能引起防御反应的问题。

(3) 根据不同情境、谈话对象的特点及评估的目的要求，运用不同形式

的问题设计。在谈话设计中经常采用的问题形式主要有直接问题和间接问题，开放式问题和闭合式问题。

直接问题要求谈话对象直接表明自己的看法、观点和态度，间接问题要求谈话对象进行一定的推论。谈话问题是采用直接方式还是间接方式，与问题和目的之间的关系有关，即谈话的目的是询问一些问题，还是要测量某些概念。一般来说，当谈话双方没有很多共同语言、对概念缺乏相同理解，问题可能产生较大的情绪障碍及谈话对象的意见和观点受到社会规范的较大影响时，以间接问题为宜。直接问题可以避免复杂的推理，或在没有风险的条件上使用，但要求主试和对象双方有共同语言。

开放式问题是让谈话对象自由回答的一类问题；闭合式问题是限制了回答的方式和内容的问题，包括一些强迫性的选择题。在谈话问题的设计中究竟采用开放式问题还是闭合式问题，依谈话的目的、谈话对象的信息水平、意见状态、情况知晓的程度而定。如果谈话既要了解谈话对象的态度或特征，还要了解他所持意见的基础、原因及强度，宜采用开放式问题。一般而言，谈话对象对问题情况了解不多，或对问题的意见较明确，宜采用闭合式问题；如果谈话者对谈话对象的情况不太清楚，宜采用开放式问题。

(4) 谈话问题的设计要注意问题的组织与编排。由广泛、一般问题到具体问题，由较大的问题到较小的问题，问题过渡要自然。谈话初期的问题应起到说明前景、激发兴趣的作用；在谈话对象的兴趣已被激发起来而没有感到疲劳时进行关键内容和重要问题的交谈；容易引起不愉快或疑惑的问题应尽量排在后面。对于复杂的教学评估目标应多用多重问题或复合性问题，并注意问题陈述形式的随机性。

### 3. 运用谈话法的基本要求

在教学评估中，运用谈话法搜集教学评估的信息时，要做到以下几点：

(1) 准备充分。在正式谈话之前，要明确谈话的目的，制订谈话计划，设计谈话问题。特别要注意的是谈话法对谈话主持者的素质要求是很高的，因此，对于谈话的内容、提出的问题、发问的方式、问题的措词及先后的顺序，谈话主持人事先要周密思考，妥善安排。并尽可能多的掌握谈话对象的一般情况，如个性特点、兴趣爱好、经历特长以及教学方面的有关情况。

(2) 把握好谈话的气氛，努力创造一个和谐、友好、轻松、愉快的良好气氛。作为谈话主持者，语气、语调、语速都要适中，态度要大方、自然，不可出现特别的情绪状态，特别要注意防止出现有意或无意流露的倾向，以免有暗示、导向作用，妨碍谈话的客观性。如果谈话对象出现拘谨情况，宜采用迂回的办法打破僵局。对于在谈话中出现的某些保密心理，应尊重对方的意见，切勿生硬地催促或强迫回答。如果出现不利于交谈的局面，经努力仍不能缓解，应暂缓进行，不要急于求成。

(3) 紧密围绕计划进行。在谈话过程中，如果出现谈话对象的回答比较散漫、甚至离题现象时，要及时地予以纠正，使谈话及时回到主题上来，紧密围绕计划进行有效地交谈。

(4) 迅速而准确的记录。谈话法是以口头语言的直接交流为主的，若不及时记录谈话的内容，容易增加谈话的主观性，甚至歪曲谈话的原意。因此，要准确迅速记录谈话的内容。一般情况下，只记录一些重要的内容要点，有时当面记录会影响回答者的谈话情绪，对这种不便于当面记录的内容，采用事后追记。在某些必要的情况下，在征得回答人同意之后，可利用现代录音

录像设备记录谈话内容。在教学评估的交谈中，一种有效的办法是事先对回答的各种可能性作充分的估计与预测，并将每一类问题排编成数字代号，以便速记。

教学评估中的谈话法除了上述谈话者与谈话对象二者之间面对面的直接交谈外，还可以用开座谈会的形式进行集体讨论与交流，以便于交换看法，集中问题。但集体交谈时有可能出现谈话对象出于某种顾虑而言难尽意或有所保留，主持者要注意启发引导，并采取相应的措施打消谈话者的顾虑。

