

第四章 教育评价的过程、原则与方法

一、教育评价的过程

(一) 教育评价的过程

教育评价的过程可分为三个阶段。第一，决定在特定教育事物的价值判断中，什么性质是重要的。第二，设计可以准确测量这些性质的程序。第三，综合各种材料和数据并形成一个价值判断。这三阶段依次阐述如下：

1 具体形成教育目标。

使教育目标最确切地具体化的实际过程，需要十分繁复的思考。因为这是一个费脑筋的、困难的任务。它要将一般教学目的的抽象描述转变成尽可能详细的教学计划，这个计划是具体的，实际的，它足以使教师在学生的行为表现中观察得到。

应该说明，对有关课程或方法的阐述，说明的是手段而不是目的，这些课程和教学方法只是我们达到预定目标的途径。

教师对教育目标的描述，应该具体为要求学生做到的行为动作事情，这些事情都是在教育过程中所产生的。这种操作主义式的教育目标叫做“可量度的结果”，简称M.O.O.

教育目标也应和儿童应该达到的最终生活价值有所区别。人的价值感，正义感、珍视他人的完整性和贡献，各种各样的美学经验等等都是所谓的生活价值，但这些对于学校教师的课堂学习目标来说，就太一般化了。因此，教师必须把注意集中在一些最终

能导致学生获得上述生活价值的行为模式上。这意味着，要从我们固有的文化环境中，从各种成人活动的广阔范围内进行选择。

但对教师来说，这个范围太广了，因此，通过选择我们就能清楚地知道哪些才是作为教师应该负责的。

规定目标的准则简要说明如下：

(1) 哲学准则。用这种准则来决定适合我们价值尺度的、值得化时间的行为模式。

(2) 机构准则。用这种准则决定哪些行为模式应由学校负责，哪些应由家庭、教会和公众负责。

(3) 心理准则。在确定目标时，我们要根据这种准则注意人的自然成熟的作用。

我们必须对以下问题作出满意的回答：

(1) 如果某个儿童正朝着预定的目标发展，在多大程度上是由于一般成长发育的作用，而不是教育的作用？

(2) 在什么样的成熟水平上，目标才能顺利达到？也就是我们的愿望要和学生能力差异相适应。

(3) 我们所培养的学生的行为的持久程度如何？

(4) 我们设计的内容受欢迎的程度如何？

在儿童教育生活的每一个连续阶段中，这些准则能帮助我们挑选较有价值的目标而摒弃那些无足轻重的。然后，这些目标应该以学生的学习行为来定义。这样，就可用这些行为来构造恰当的测量了。

2 描述或测量有关的行为性质。

这一步是创设一些情景，使预期的行为能发生、能被观察。

换句话说，最好的测验程序是：要求学生去做那些要求他们学会做的事情。

有时候，直接测量某些我们所关心的行为，不一定可行。因此，必须使用间接的测验方法。但是，使用间接测验有一种危险，因为“一个学生能做X，他就必然能做Y”的假定常会产生错误。不同的测量方法常会得出非常不同的结果。教师所要求的学习行为与测验中检查的行为越接近，教师对测验结果有效性的信心就越大，这是一条通常的规则。

目前可供教师使用的最精确的测验是标准化成绩测验。遗憾的是，用这种测验测量的行为与教师的要求完全相符是不常见的。因为有一种趋向，标准化测验的目标范围过于狭窄，它们只测量那些易于测量的行为。因此，标准化成绩测验不应该是评价学生进步或教育大纲价值的唯一标准。评价要建筑在测量所有目标的基础上，而不可仅仅以便于测量的那些目标为基础，这一点是重要的。

3. 判断的过程。

评价过程的最后阶段是将涉及到每个学生行为的材料数据集中在一起并加以概括，然后指出，某个儿童、某个班级或整个学校是否取得了令人满意的进步。

这儿出现两个问题。第一，我们用什么标准来判断一个已完成的课程所达到的水平是否令人满意？第二，如果我们发现，在一个方面的质量提高上见了成效，而在另一方面却遭致了损失，那末，为了对整个情况作出公平的判断，我们又如何来确定哪一

方面更重要呢？

有很多问题跟上述第一个问题有关，也就是说，要考虑到客观和主观准则，所谓客观准则，是指我们的行为应与我们周围的人们认为可接受的标准有关，这些标准在知识、技能、能力、行为等方面取最低限度的形式，它们是某一给定的年龄组、作业组或其他有关社会生活单位的所有成员都应做到的。另一方面，要考虑主观准则，就是说，我们的行为和个体的潜力有关，这些准则要取我们预期可达到的最高标准形式，这是以一个人尽最大努力所表现出来的行为为依据的。因为不能期望学生在每一件事情上都能充分发展，因此，必须弄清楚，对于某个学生，我们的目的是只要他达到最低标准，还是使他的某种行为能发展到他可能的最高点，不论是哪种情况，我们的评价性判断都必须将实际效果与预期效果加以比较。

最大的困难是估计一个学生的潜力，因为在确切了解他的潜力以前，我们并不能判断这个学生是否有达到最低标准的能力，更没有任何办法知道能希望于他的最高标准是什么。

参考一些一般智能的测验，我们能够大致判断在各个不同智力水平上的学生组平均能达到的水平，然而在解释这些结果时必须非常谨慎，因为，某些影响学校成绩的特殊因素，如情感、意志等因素，它们所产生的作用，在这些测验中是完全不去测量的。

结 论

让我们重复上文中再三提到的论点来结束这一章：教育过程中的三要素是不可分割地连接在一起的。

确定在操作上可定义的教育目标；设计恰当的学习情景和材料来达到目标；测量和评价教育的结果。这三要素是如此紧密相联，因而它们也是相互深刻影响的。

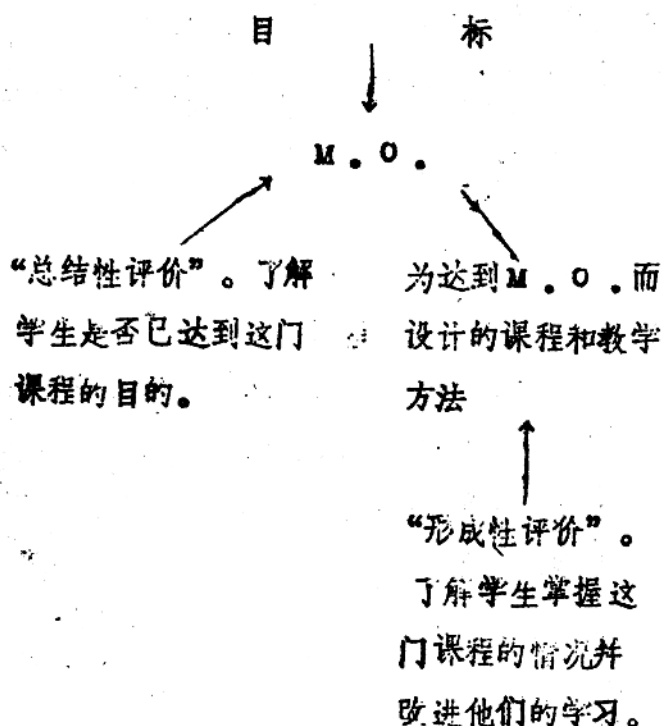
由于这个原因，所以强调应以可测量的形式仔细说明目标，它的重要意义是为了仔细选择教育手段（指要用的方法与材料），从而使这些手段不是无意识地形成目标甚至是歪曲目标的。也由于这个原因，必须将评价过程看成是一个经常的反馈，也就是说通过这些教育活动的结果来调整我们的教育手段，从而保证教学手段在现实中的有效性。

* * * *

就学校中的评价而言，有两个主要目的：第一，在学生学习过程中向他们提供反馈。因为有效地理解学生的学习成果不是轻而易举的，所以提供正确的反馈是一个好教师的主要任务之一。第二，估计学生学习成绩。这是提供成绩报告单，决定是否升级的基础，同时也有助于教师判断他的教学计划的有效性。评价的这两个目的是完全不同的，不应混淆，如“每一个测验都是一次学习经验”就是一种常见的但却是谬误的说法。

第一种类型的评价是为调整教学过程提供反馈所必须的，叫做“形成性评价”。这在课堂的正常教学中几乎每天，每时、每刻都在发生，如通过正式的或非正式的复述、提问，就是了解学生理解教材的程度以及提高他们的技能所必不可少的。

第二种类型的评价是为记录存档而对学生成绩进行估价，叫做“总结性评价”。它的目的不是向学生提供反馈来改进学习，而是为作出各种决定提供资料。这种类型的评价只注意前面讲过的M.O.，也就是只注意教师所试图达到的目标。下面的图解可说明这两种类型的评价的作用。



举个例子来讲清图解中所提出的几点。设想你在英国文学的教学中有一个目标：“认识短篇小说的结构型式。”以下是被你所操作主义化了的M.O.：“给一篇短篇小说，指出高潮的所在，主要情节发展和人物性格塑造的途径。”为此决定选用四个不同作家的作品。这只是因为它们是说明上述几个问题的好材料，而不是因为它们本身的特殊文学价值。我们也可以从一百篇可为我们目的服务的其他作品中选择。换句话说，我们是使用这几篇小说，而不是讲授这几篇小说。这些作品规定在课堂里讨论和分析，通过口头提问或书面作业的不断评价来了解学生的理解情况。最后，还要看清所有这些活动是否已导致用M.O.所表达的行为，这就必须再给学生一篇陌生的短篇小说，对他们的能力作一次总结性评价。

前文已强调了以下这一点：一个对学生承担着某些责任的教师必须明白，他是否确实在履行这些责任，对所用的任何材料和方法的最后测验都是为了了解学生在达到目标方面成功的程度究竟如何，如果没有对最后结果的理解，那么达到目的的手段也不能具有针对性。事实上如果不明课程目标的成功程度应是什么，那么教好这门课程有较大把握是没有什么基础的。

教育目标的形成

教师在形成教育目标时必须牢记几个原则：

- 1 目标必须和教育相适应。就是说，这些目标必须是通过学校的教育过程所能达到的。

2 目标必须符合实际。就是说，学生在校学习期间一定有着朝着目标取得进步的可能性。

3 从学校整体的角度来讲，目标还必须适应学生和学校整体的需要。

“教育目标”的表述也要和它们在学生中实行的不同目的相适应。如为了达到向学生传递信息的目的，教育目标的说明就要毫不含糊，并且是可能做到的。

M.O. 是由于评价过程的需要而产生的。只有精确定义要评价的东西，才能实现有效的测量和评价。为达到这个要求，下一章“目标和M.O.”中将介绍如何根据准则来叙述M.O.。

最后应该说明：

1 教师不必对每一课的目标都以行为性的措词来表达。教师应当特别需要说明的目标是：

- (a) 学生行为的主要变化。
- (b) 识别学习阶段的关键性步骤。

2 教育目标的叙述和M.O. 不允许有伸缩性。如以下的情况就是不允许的：

- (a) 目标和M.O. 由教师制定后又允许任意改变。
- (b) 在教学过程中频繁产生新的教学目标，并任意加进教学计划。
- (c) 在评价过程中随意增删预定的目标。

3、在评价过程中，说明M.O. 的一条主要基本原理就是

针对教和学的过程，并保证对结果的评价是可靠和有效的。

竹本 朱怡华译：《教育与评价》，摘自上海《教育科研情况交流》1983年第3期

(二) 教育活动及评定的过程

评定是在连续不断的教育活动中，对照教育目标对教育效果作出价值判断，有效地进行教育指导的一连串的反馈活动。据此，该过程可用图5-1来表述。

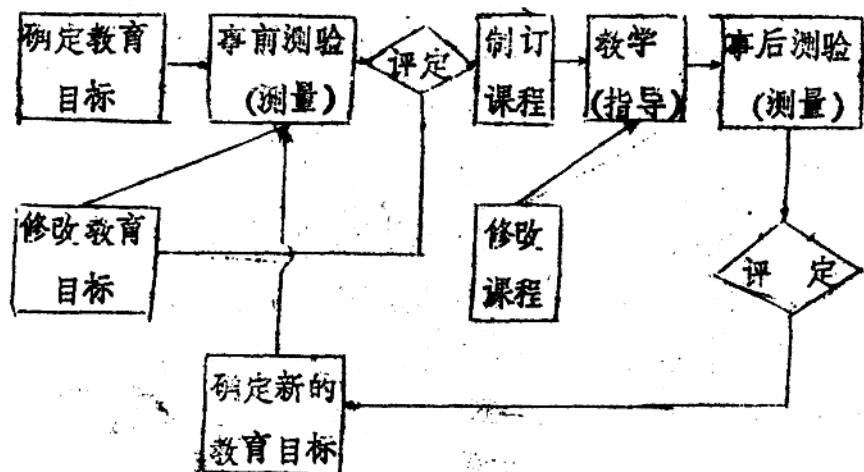


图5-1 教育活动的过程

1 确定教育目标 教育既然是有意识有计划的活动，那么所有的教育活动都是受教育目标规定的。因此，教育活动的第一步就是明确地确定教育目标。在这种教育目标中，既有促进智力发展、社会意识发展这一类极其抽象的目标，也有乃至非常具体的教育目标，形形色色，标准不一。要确定哪一种标准的教育目标，是同评定计划密切相关地一起决定的，但是一般地说，最好给教育目标下一个可以进行操作的定义。亦即教育目标应当用适合于分析的、适合于结合儿童的具体的行为的、适合于指导的具体的行为概念来表述。

2 事前测验与评定 把握并评定学生的现状，即用教育目标来衡量儿童的智力发展、学力水平和准备究竟处在怎样的水平，不仅对下一步的学习计划的设计，即编制课程是需要的，而且对研究教育目标的效度来说，有时也是需要的。这种评定也叫做诊断性评定。

3 学习经验的设计（编制课程） 根据诊断性评定设计教育目标，使教育目标具有具体的方向。不过，课程的所有指导内容要同教育目标相一致。同时指导方法也必须遵循教育学、心理学的原则。因此，不言而喻，教师应当具有学科的专业知识，还要求具有教育心理学的素养。然后按照课程的设置，根据各班级学生的能力的现状进行指导。

4 事后测验及评定 教师要根据各班级的情况编排课程，据以开展教育活动。因此，教育活动的评定也是教师责无旁贷的一项重要任务。事后评定要弄清各门课程预期的目标达到了怎样

的程度？如果预期目标未达到，那么课程的薄弱环节在哪里？怎样调整课程等等。评定结果如果对学生的学习活动和课程的改进毫无助益的话，那么这种评定就毫无价值。

大桥正夫编 钟启泉译：《教育心理学》（日）

上海教育出版社1980年2月

二、教育评价的原则和步骤

无论广义的教育评价或狭义的教育评价，都必须遵循下列四项原则：

（一）教育评价必须是为了儿童的发展：

教育评价的根本目的就是要促进儿童健康、全面的发展。因此，一切教育评价都必须遵循这一基本原则。

（二）教育评价必须做到评价的基准明确：

既然要搞评价，就必然要有一个评价基准。而这个评价基准必须是明确的、具体的、可以检查的。否则就不会得出科学的结论。

（三）教育评价的基准必须争取教师群众的同意。

（四）教育评价的基准不是固定不变的，要不断地检验，不断地修正。

教育评价的过程大体上可分为下列四个步骤：

（一）信息的输入：

教育评价必然开始于观察、测试、累加记录、问卷、面谈个案调查等办法，获取教育过程的各种信息。在教育实验中，还必

须注意实验组与控制组的各种不同条件,以及在不同条件下的各种信息。

(二) 信息的处理:

取得教育过程的各种信息后,通常要进行统计技术的处理。如有两列变量,就要计算两列变量有无相关,相关强弱,以及各种因素分析等。

(三) 原因的分析:

对各种信息进行统计技术的处理后,还要运用各学科的研究成果进行原因分析。例如从统计学上求出两列变量存在着相关,但相关不等于因果关系,这就需要从心理学、社会学等学科的角度寻找其是否存在着因果关系,又是什么样的因果关系等。

(四) 确认、报告、传送、决定:

教育评价的最终目的是为教育改革提供科学依据。因此,最后的步骤就必须是经过检验、确认,写出书面报告,提供给有关部门,作为决策的科学依据。

以上仅就国外教育评价的基本概念做了一个简单、肤浅的介绍。因本文资料多来自于外文,误译、误解在所难免,尚望读者多加指教。至于狭义的、广义的教育评价的技术问题以及国际教育成就评价协会(I E A)的活动概况,因篇幅所限,不再赘述。

张天麟:《教育评价的基本概念》,摘自天津
《教育科研资料》1984年第9期

三、普通教育评价的原则

“任何运动形式，其内部都包含着本身特殊的矛盾。这样特殊的矛盾，就构成一事物区别于他事物的特殊的本质。”⑧教育也毫不例外地有它自身的特殊规律，教育者对受教育者所进行的全部教育活动，只有遵循教育的特殊规律，才能收到良好的教育效果。为了使人们真正按教育规律办事，必须加强教育评价工作，而为了搞好评价工作，又需确立评价的原则。

(一) 客观性原则

客观性原则是指在进行教育评价时，必须采取客观的实事求是的态度，不能主观臆断或掺杂个人感情。

客观性原则是由党的思想路线和教育评价本身的性质决定的。党中央一再指出，在各项工作中，都必须实事求是。教育评价工作也应毫不例外。教育评价是根据教育目标对所实施的各种教育活动效果及学生学习质量和发展水平进行科学的判定。这种判定如果是客观的，可以推进教育事业的发展；如果判定得不客观就会妨碍教育事业的发展。评价学校、评价教师、评价学生，都需遵循客观性原则。评价需要根据由教育目标而确定的评价标准来进行。标准一旦确定，任何人都不能随意改动。确定标准时，不能为了照顾某个评价对象，把不应列入的条件列入，也不能为了排斥某一评价对象把不应列入的条件列入。那种在评价过程中，随意确定标准随意脱离既定的教育目标改动评价标准（无论是提高标准或降低标准）的做法都是错误的。特别是对学生学习质量和发展水平的评价，尤其要重视客观性原则。因为青、少年学生

正处于发展时期，可塑性很强，对评价是否客观非常敏感。如果评价是客观的，会使他们增强信心从而奋发向上；如果评价不是客观的，含有偏见，成见或个人感情色彩，就会挫伤他们的积极性，甚至会使他们失去前进的信心，而自甘暴弃。

(二) 一致性原则

一致性原则是指进行教育评价时，必须用一致的标准。无论是对集体的评价，还是对个人的评价，不能对甲采用一种标准，对乙采用另一种标准。

在评价工作中坚持一致性原则是由我国社会主义教育目的的统一性决定的。党的十一届六中全会在《中国共产党中央委员会关于建国以来党的若干历史问题的决议》中，明文规定了“用马克思主义世界观和共产主义道德教育人民和青年，坚持德智、体全面发展又红又专、知识分子与工人农民相结合，脑力劳动与体力劳动相结合的教育方针。”这个方针，是指导评价我国各级各类学校教育活动的出发点和依据。

一致性原则是有层次性的。在进行具体评价时，各级各类学校都有各自的标准，但在具体执行时，又必须遵循一致性原则。从我国中、小学来看，由于它是普通教育，它的具体任务是为上级学校输送合格的学生，为整个社会输送优良的劳动后备力量。这是对每一所中、小学共同的要求，也是评价每所中、小学教育质量的一致标准。国家颁布的教学计划、教学大纲则是评价教师的教学工作质量和评价学生的学习质量的一致标准。遵循一致性原则，对学校、对教师、对学生进行评价，才能区分评价对象的

好、坏优、劣；使被评价者知道自己在群体中的位置，从而发扬长处、补救短处，保持优点，克服缺点，激励自己，积极向上。

我们强调教育评价，要遵循一致性原则，并不否认由于学校设备、教学条件、校长和教师水平的不同，在学校与学校，教师与教师、学生与学生之间所存在的差异。这里所说的一致性原则一方面是指教育目标是一致的；另一方面也是指在同一范围内，对相同的对象必须用统一的标准。

(三) 全面性原则

全面性原则是指在确定或运用评价标准时要全面，不可片面；在评价过程中，要听取各方面的意见，不能只听某一方的意见，甚至某几个人的意见。

全面性原则是由教育工作的复杂性和全面发展的方针决定的。因为教育工作任何一项成果都是由多种因素促成的。其因素有全体教职员工的思想觉悟、科学文化知知识水平、业务能力、专业知识、主观努力程度、工作的客观条件以及协作精神等等。在确立评价标准时，必须包含各方面的因素。在具体运用评价标准时，更不可片面地强调构成评价标准的某一个因素或某几个因素而忽视、抛弃其它因素。

按照这个原则，在评价我国中、小学时，就要从它所承担的双重教育任务完成的情况来看。即：不仅要看它的升学率高低，而且要看它为我国社会主义四化建设输送的劳动后备力量是否优良。在工作岗位上表现得如何。那种单纯地以升学率高低为标准，对学校进行奖惩是不对的。

、我们强调教育评价的全面性，并不是要求在评价时不分主次对构成评价标准的各方面因素等量齐观。恰恰相反，教育评价的全面性要求评价学校、教师、学生时，既要看主要方面，也要看次要方面，而且要求既不要把主要的方面当成次要方面，又不要把次要方面当成主要方面。

(四) 单项评价与综合评价相结合的原则

单项评价是指对评价对象从某个侧面进行的评价；综合评价是指对评价对象进行完整的、系统的评价。

这个原则的确立，是由我国学校教育活动客观存在的多层次性决定的。其层次从行政系统来看，有教学、思想政治教育、体育卫生、生产劳动、总务等各方面活动。从党群系统来看，有党、共青团、少年先锋队，学生会以及教育工会等各方面的活动。这两个系统中的每一项活动，又无不包含着无数具体的活动。如教学活动，不仅包含各学科的教学活动，也含着大量的校内与校外的各种活动。既或是一个学科的教学，也无不包含着更为具体的活动。如：语文教学，就有识字、写字、解词、文章分析、朗读、背诵、作文等各方面的教学活动。评价工作是需要与各层次的教育活动同步进行的。它要判定各层次教育活动的效果，以便改进这方面的工作。因而需要有单项评价。通过单项评价，不仅能为教师改进某一方面的工作提供依据，而且能使学生在评价中及时看到自己学习的成败，从而自觉地控制学习发展趋势。

教育活动是一个极其复杂的活动，虽然多层的活动都具有一定的相对独立性，但他们又是相互连结在一起的一个有机整体，

只有协调起来，密切配合才能把青年一代培养成“有理想、有道德、有文化、守纪律的劳动者。”^③因此，在对各层次中的教育活动的单项评价的基础上必须对教育活动的有效性进行综合评价。如：对学生的评价，则必须对其德智体全面发展的情况进行完整的系统的评价。对教师的评价，则不仅看他的教学效果，也要看他班级辅导工作情况，不仅看他是否有比较渊博的知识、也要看他是否认真研究并掌握教育科学、懂得教育规律，更要看他是否具有高尚的道德和良好的情操。只有这样的评价，才能保证我国社会主义学校教育的方向性，才能使我国学校教育各层次的活动相互适应，使各种教育活动呈现良好状态。

但是，我们应当看到，综合评价并不是单项评价结果的相加。在一定的意义上讲，单项评价只是进行综合评价的基础或手段。综合评价是单项评价的发展，只有通过综合评价才可能判定今后一个时期内被评价对象的发展趋势和倾向，从而使被评价者明确今后的努力方向。

(五) 评价与指导统一的原则

评价是按照一定的原则、标准对评价对象已完成的行为作出肯定或否定的判定，使被评价者从中受到启发和教育。如评选先进学校，评选优秀教师、评优秀学生等。这是对他们已完成的行为的褒奖，以使人们学有榜样，赶有方向；上级对学校工作的批评，学校对教师、学生不良行为的惩处，是令其吸取教训，使人们引以为戒。那么，怎样才能使受褒奖者继续进步，努力发扬优点，使受到批评教育或受到惩处者能够积极改正错误呢？这就需要指导。