

程书肖 编著

JIAOYU PINGJIA

FANGFA

JISHU

教育评价

方法

技术



北京师范大学出版社

BEIJING NORMAL UNIVERSITY PRESS

京師教育書苑

教育评价 方法 技术



北京師範大學出版社

BEIJING NORMAL UNIVERSITY PRESS · 2004

图书在版编目(CIP)数据

教育评价方法技术/程书肖著. - 北京:北京师范大学出版社, 2003. 12
ISBN 7-303-06841-4

I. 教… II. 程… III. 教育评估-方法 IV. G449.7

中国版本图书馆 CIP 数据核字(2003)第 116394 号

北京师范大学出版社出版发行
(北京新街口外大街 19 号 邮政编码:100875)

出版人:赖德胜

北京师范大学印刷厂印刷 全国新华书店经销

开本: 787mm×980mm 1/16 印张: 18.5 字数: 320 千字

2004 年 4 月第 1 版 2004 年 4 月第 1 次印刷

印数: 1~3 000 册 定价: 32.00 元

前 言

教育评价是现代世界上教育科学研究的三大领域之一。这三大领域是教育基础理论研究、教育发展研究、教育评价研究。教育评价的思想虽然源远流长,已有三千多年的历史,但是教育评价作为一个独立的研究领域,形成教育科学的一门重要分支学科却是近几十年的事情。教育评价是一门综合性、交叉性、边缘性的教育学科。它以马克思主义哲学为哲学基础,以教育学、心理学、教育管理学等为理论基础,以系统理论为方法论基础,以教育测量、教育统计、计算机软件等为基本工具。教育评价作为教育管理与指导的主要手段,它可以为教育管理、教育决策提供科学的信息和依据,在加强教育的宏观管理、深化教育改革以及不断提高教育质量实现教育目标的过程中,起着十分重要的作用。当前,世界各国都把发展教育、培养人才作为经济与社会发展的战略重点,一般都重视增加教育投入,开发智力资源,运用教育评价来促进教育事业的发展,促进国民素质的提高。在我国,教育评价虽然起步较晚,但发展很快,已成为教育科学研究的重要领域之一。教育评价的理论与实践越来越受到各级教育行政部门的领导、从事教育科学研究的理论工作者与广大中、小学教师的普遍重视。早在1986年,我国政府在七五报告中就明确提出了“要建立系统的教育评价和监督制度”。1995年3月18日公布的《中华人民共和国教育法》,第24条规定:“国家实行教育督导制度和学校及其他教育机构教育评估制度。”用法律的形式把教育评价作为教育基本制度肯定了下来。1997年9月在山东省烟台市召开的全国中小学素质教育经验交流会上,李岚清副总理提出“要建立符合素质教育要求的科学的评估机制”。原国家教委副主任柳斌同志也在会上指出,推进素质教育“建立新的评价制度是关键”。可见,我国现在已经把教育评价研究提到了重要日程,各级政府也开始重视教育评价的理论与实践。

教育评价方法技术是达到评价目的、实现评价目标的手段。我们的任务是过河,没有桥和船就不能过,教育评价方法技术就是达到评价目的、实现评价目标的桥和船。如果没有科学合理、通俗实用的评价方法技术,实施教育评价就有失败的危险。就教育评价的现状而言,研究、挖掘、创造科学实用的评

价方法技术,已成为提高教育评价质量的当务之急。本人多年从事教育评价理论与方法技术的学习、探讨和研究,并担任教育管理专业硕士研究生和研究生课程班《教育评价的理论与方法》课的教学工作。在多年的教育评价学术活动和教学实践中有一些看法和心得体会,尤其是教育评价方法技术方面体会较多,愿意将讲稿整理成册,供研究生课程班学员及广大中、小学校长、教育工作者、行政管理人员学习和参考。由于作者学疏才浅,水平有限,书中错误和疏漏在所难免,欢迎专家、学者、同仁、朋友批评指正。

程书肖

2003年6月于北京师大

目 录

前 言	(1)
第一章 教育评价的几个基本问题	(1)
第一节 什么是教育评价	(1)
第二节 教育评价的现实意义	(7)
第三节 教育评价与相关概念的区别	(16)
第二章 如何组织与实施教育评价	(20)
第一节 教育评价的组织	(20)
第二节 教育评价的准备	(25)
第三节 教育评价的实施	(27)
第三章 如何设计评价方案	(44)
第一节 评价方案的设计程序	(44)
第二节 指标体系的设计原则和设计方法	(51)
第三节 分配和确定指标权重的方法技术	(60)
第四章 教育评价结果的计算和分析方法	(70)
第一节 问卷调查结果的计算和分析方法	(70)
第二节 加权平均法	(77)
第三节 综合评判法	(80)
第五章 教育评价中如何进行心理调控	(94)
第一节 心理调控的意义和作用	(94)
第二节 评价者的心理调控方法	(104)
第三节 被评价者的心理调控方法	(115)
第六章 教育评价中常用的统计参数	(120)
第一节 相对数	(120)
第二节 平均数和标准差	(125)
第三节 相关系数	(138)
第七章 教育评价中的测量方法技术	(148)
第一节 测量法的特点和测量量表	(148)
第二节 对测量结果的加工处理方法	(155)
第三节 信度、效度、难度、区分度的分析 和测定方法技术	(161)

第八章 统计推断方法技术在教育评价中的应用	(175)
第一节 统计推断法的原理.....	(175)
第二节 统计推断法的参数检验.....	(183)
第三节 统计推断法的非参数检验.....	(206)
第九章 如何对教育评价方案进行再评价	(214)
第一节 对方案科学性的再评价方法.....	(215)
第二节 对方案有效性的再评价方法.....	(228)
第三节 对方案可行性的再评价方法.....	(235)
第十章 如何对评价结果进行再评价	(242)
第一节 对评价结果是否可信的再评价方法.....	(242)
第二节 对分类结果是否科学合理的再评价方法.....	(249)
第三节 对评价结果后效性的再评价方法.....	(253)
附表 1 10 000 个随机数字表.....	(256)
附表 2 正态分布表.....	(264)
附表 3 χ^2 值表.....	(270)
附表 4 t 值表.....	(272)
附表 5 F 值表.....	(274)
附表 6 符号检验表.....	(282)
附表 7 秩和检验表.....	(283)
附表 8 相关系数 r 的临界值表.....	(284)
附表 9 肯德尔和谐系数中 S 的临界值表.....	(286)
主要参考文献.....	(287)

第一章 教育评价的几个基本问题

第一节 什么是教育评价

一、如何理解评价

要搞清什么是教育评价，首先应该理解什么是评价，评价的核心和关键是什么。《辞海》中是这样说的，评价是“泛指衡量人物或事物的价值”。一般地说，我们在日常生活、工作、学习或社会活动中，经常会遇到各种各样的人物或事物，在与这些人物或事物的接触中，我们随时都会从外界接收到大量的信息和资料，对于这些信息资料，经过加工整理、分析综合等思维活动过程，就会形成某些认识，从而对人物或事物作出“是好是坏”的判断，这往往就被称之为评价。所以，所谓评价，就是指对人物或事物有无价值进行判断。当然，这里所说的“价值”并不是马克思关于商品的价值，而是哲学意义上的价值，是指事物的用途和积极作用。对于评价者来说，有用途和能起积极作用就是有价值的；没有用途和不能起积极作用就是没有价值的。这里又遇到一个实际问题：因为评价者是人，每一个人都是有意识、有思维的。不同的人，由于经历不同，经济地位不同，利害关系不同，于是形成的意识和情感也就不相同，就会产生不同的习惯、爱好、兴趣、利益和欲望，反映到各种活动领域就会有不同程度的需要和追求。因此，不同的人就有不同的价值观。你认为有用的东西，他可能认为无用或者用途不大；你认为有积极作用，他可能认为没有积极作用，甚至还可能认为有消极作用。比如对一个烟灰缸，吸烟的人可能认为有用；不吸烟的人则认为没有什么用。正因为如此，对于同一个人物或者同一个事物，不同的人可能会作出程度不同的甚至截然相反的“价值判断”。比如对于同一篇作文，不同的评判者结论往往有较大的差异，这就是由每个人的主观评价所导致的结果。

我们所追求的评价，是科学的评价，是在人们的主观认识最大程度地符合客观现实的情况下所作出的价值判断。因而，这种评价的结果是比较客观、比较合理、比较公正的。做到这一点的前提条件是要求评价者有共同的

价值标准。在实际评价工作中，我们的做法是在进行评价时适量减少一些在价值标准上有过大争议的评价目标和评价指标，并对保留的评价指标进行详尽地表述，以便作出评价者均可接受的价值方面的约定，从而求得评价标准的大体一致。换句话说，主观评价反映客观实际的前提条件，首先应该是大家在价值观的较多层次方面相近。有的学者提出，在评价中要求评价者有共同的价值观，这是理想化的提法，实际上是做不到的。不同的人有不同的价值观，这是客观现实，做到完全一致是不现实、不可能的。但是经过大家讨论、协商、论证，作出大家均可接受的共同的价值约定却是现实、可能的。其次，在评价活动中评价者要依法行事。以大家公认的评价标准为依据，不能感情用事，不能过分强调自己个人的价值观。只有这样所作出的价值判断，才能比较客观公正、合情合理，因而是比较科学的。由此可以看出，现实中所谓的“科学评价”本身就已经包含有不科学的成分，因为所谓客观的评价是以评价者认可的评价目标、评价指标、评价标准为依据的。这并不意味着评价者在价值观上取得了完全的统一，因此，所谓科学的、客观的只能是相对而言，并不是绝对的。

评价活动包含两层关系：第一层是评价主体与评价客体的关系。在评价活动中，评价者是主体，被评价者是客体。因此，进行评价首先要搞清楚谁是评价者（评价主体），谁是被评价者（评价客体）；第二层是价值主体与价值客体的关系。无论何种形式的评价，其结果可以表达为：谁（什么）是有价值的（或无价值）。这里的谁（什么）就是价值客体。在谈论有无价值时，必须明确指出对谁来说是有价值（或无价值）的。这里的谁就是价值主体。同一个客体，对不同的主体来说价值并不一定相同。比如，同一个烟灰缸对吸烟者与不吸烟者来说价值是不同的。评价就是要揭示价值主体与价值客体的关系，即价值关系。价值的实质是主体与客体之间的一种特殊关系，即客体满足主体需要的关系。对于主体的需要，客体的属性与功能能够满足，客体就是有价值的；对于主体的需要，客体的属性与功能不能满足，客体就是无价值的。如果客体的属性与功能对主体的利益和需要构成损害，那么客体就是有负价值的。在马克思主义者看来，所谓价值，就是客体与主体需要之间的一种特定的关系。马克思指出：“价值这个普遍的概念是从人们对待满足他们需要的世界的关系中产生的。”价值的通俗解释是指事物的用途和积极作用。即有用途有积极作用就认为有价值；无用途无积极作用就认为没有价值。价值的完整含义应该是指主体的需要和客体的属性与功能二者的统一。即主体的需要，正是客体的属性与功能所满足的就是有价值的；主体的

需要，如果客体的属性与功能不能满足，即二者不能统一，则无价值。因此，我们在进行评价时必须注意以下三点：第一，把握价值主体的需要；第二，把握价值客体的属性与功能；第三，以价值主体的需要来衡量价值客体的属性与功能，以便判断价值客体是否满足价值主体的需要。在评价活动中，评价标准就其实质而言，就是评价主体所把握的、所理解的价值主体的需要。在评价活动中，评价客体与价值客体往往是统一的，但评价主体与价值主体则不一定是统一的。有时是统一的，即二者相重合；有时是部分重合的；有时是分离的，即评价主体完全是第三者，是站在第三者的立场上来评价价值主体的需要是否能得到价值客体的满足以及满足的程度（即价值的大小）。例如在自我评价中，评价主体与价值主体是统一的，评价主体与评价客体也是统一的，评价客体与价值客体也是统一的。但督学对学校进行评价时，评价主体与价值主体是分离的。因为在评价活动中，评价主体是督学，评价客体是学校，价值客体也是学校，但价值主体是国家、政府和人民。价值主体的需要是指国家、政府和人民对学校的需要。一所学校办的好坏，质量高低是看其是否满足国家、政府和人民的需要。

二、什么是教育评价

从上述对评价含义的理解中，我们已经看到评价的核心是对人物或事物进行价值判断。没有“价值判断”不能叫做评价，只能称之为描述或议论。进行评价就必须进行价值判断。教育评价就是对教育现象中的人物或事物进行价值判断。教育现象十分复杂，涉及的面非常广泛。因此，教育评价的对象涉及整个教育领域。不仅包括学生、教师、行政人员、校长，还包括学校管理、教育课程、教育计划、教育质量、领导素质、办学条件、办学水平等。可以对某个学校进行评价，也可以对一个地区的教育水平进行评价，还可以对一个国家的教育方针、教育体制、教育战略、教育改革进行评价。总而言之，从宏观到微观教育领域内的方方面面都可以作为教育评价的对象。对教育现象进行价值判断应该搞清教育价值的含义是什么。教育就其价值来说，人们常常把它分为个人价值和社会价值两种类型。所谓教育的社会价值，是指教育在满足国家，尤其在满足占这个社会统治地位的阶级、群体需要方面的程度；所谓教育的个人价值，是指教育在满足个体（其中主要是受教育者）需要方面的程度。在奴隶社会、封建社会和资本主义社会，教育的这两种价值往往是矛盾的。在我们社会主义国家，这两种价值从本质上来说是一致的。因为在我国，社会对教育的需要集中地反映或者说代表了个人对

教育的需要，因而教育的社会价值与教育的个人价值是可以统一起来的。例如，我国的教育事业要加速发展，要提高教育质量，就需要有高质量的教师。因此，加强对教师的教育与培训是国家的需要，人民的需要，也是教师个人的需要，社会价值和个人价值二者是统一的。当然，这两者之间也会产生一些矛盾，但这种矛盾属于人民内部矛盾，一般情况下是可以协商解决的。比如，现在各级学校对教师的学历有一定的要求。提高教师的学历有利于教学质量的提高，是学校所需要的。但是，学校的教师进修需要有计划地进行，总会有先有后。这时，教师个人的需要与学校的需要就可能产生矛盾，这种矛盾当然属于人民内部矛盾，可以通过协商解决。

我们是社会主义国家，不同于西方资本主义国家。在进行教育评价时也必须要有我们的特色。应该特别指出，在我国开展教育评价活动应该以社会评价为主。也就是说，我国的教育评价就其本质来说是一种以社会的价值观为依据的评价。我们强调社会价值，并不是否定个人价值的重要性。我们所反对的是过分强调个人价值，反对用个人价值来否定社会的价值。这就要求我们在作出价值判断时，首先应该把国家利益放在优先位置，在判断有无用途、有无积极作用时，一定要坚持把社会的需要作为依据。比如，为了使教育评价更好地促进教育为社会主义建设服务，必须坚持把社会的实际需要当做评价学校教育的依据，把反映社会实际需要的教育目标作为得出评价结论的基准。我们判断一个学校的办学水平高低，从根本上来说就是看它在满足社会实际需要方面的能力和水平；判断一个学生是否为好学生，就是看他是否达到了德智体等全面发展的目标。不能只看学习成绩，不看思想品德和身体素质。在我们社会主义国家，只有摆正个人和国家的位置，才能比较容易达到共识，形成比较一致、比较完备的价值约定，从而达到价值判断的一致性、客观性和正确性。

三、教育评价的定义

因为教育评价是一门新兴学科，它的理论体系、指导思想、评价模式和评价方法技术等都处于不断发展、不断完善的过程中。因此，各国教育评价专家、学者对教育评价的含义及本质属性的理解和论述并没有形成共识。由于他们的理论流派不同，评价模式不同，看问题的立场观点不同，探索的角度不同，实验的方法不同，看问题的侧重点不同，于是就形成了不同的评价观，对教育评价给出了不同的定义，至今尚未形成一个统一的为大家所公认的科学定义。

1. 以美国教授泰勒 (R. W. Tyler) 为代表, 认为教育评价就是衡量教育目标在实际上达到程度的过程。他们把侧重点放在了教育目标上, 认为教育评价就是衡量教育活动达到教育目标程度的一种活动。因此, 在进行教育评价时, 首先要确定教育目标, 然后在分析应达到的教育目标的基础上设计评价方案, 确定评价指标和评价标准。然后根据评价方案的要求对教育活动的效果进行分析比较, 看看在多大程度上达到了目标。早在 20 世纪 40 年代初泰勒在总结“八年研究”实验教育评价的成就时就指出: “根据八年研究的基本思想, 设计教育评价方案的出发点是注重教育效果的价值观, 从分析应达到的目标, 再根据目标来评价教育效果。”同时认为, 教育评价本身能够促进教育措施向理想的目标接近。这种评价观点特别注重教育效果, 是以教育目标为中心来判断教育活动实际上达到教育目标的程度。因此, 对被评对象能够鉴别好坏和区分优劣, 能够激励人们发扬优点, 克服缺点; 发挥优势, 改变劣势, 采取有力措施向理想的目标前进, 以便达到或接近教育目标。正因为如此, 由泰勒教授建立的以教育目标为中心的行为目标评价模式, 在教育评价领域中占有相当重要的地位, 并在一定时期里产生过深远的影响。但是泰勒的评价观点和评价模式也有一定的缺点和不足。首先, 它以教育目标为依据来判断教育活动的效果。那么, 教育目标的合理性又怎么判断呢? 其次, 过分强调以教育目标为中心, 过分强调教育结果, 会引导人们把注意力集中在教育结果的评价上, 容易忽视过程评价, 忽视在教育活动过程中及时总结经验教训, 不断对照目标找差距, 及时采取措施改进工作, 随时调节自己的行动, 向着理想的目标前进。

2. 以克龙巴赫 (L. J. Crombach) 等人为代表, 认为教育评价就是为教育决策提供信息和依据的过程。1963 年, 克龙巴赫在《通过评价改革课程》一文中提出了“评价的中心不应仅是目标, 而更应当是决策”的思想。他们认为评价者不但应当关心规定的目标, 检验这些目标达到的程度, 而且更应当关心谁是决策人, 他们作了什么决策, 根据什么样的准则作决策等。持这种观点的人特别强调评价的信息作用。他们认为教育评价是为教育决策提供信息和依据的过程, 教育评价的中心应当突破目标的局限, 使评价为决策服务。另外, 他们认为评价者也不应当过分重视教育结果, 而应当把注意力集中于教育活动过程。教育评价不能仅仅注意结果评价, 而应当重视过程评价。他们特别强调过程评价的作用要比只关心结果的终结性评价重要的多。因为过程评价是在教育活动的进程中不断对照评价方案找差距和发现存在的问题, 能够及时调节自己的教育行为并对不合理的教育目标、评价指标、评

价标准进行修正，能够采取有效的措施，沿着正确的方向朝理想目标前进。这就弥补了以教育目标为中心为依据对教育效果进行评价的泰勒评价模式的不足之处。这种评价观的贡献是把评价视为获取信息的过程，并把评价信息和评价结果作为教育决策的依据。

3. 以豪斯 (E.R.House) 和斯克里芬 (M.Scriven) 为代表认为教育评价就是对教育现象进行描述和价值判断。他们主张“评价是一种既有描述，又有判断的活动”，“是一种对优缺点和价值的评估”。这种主张的着眼点在于对教育现象及其效用进行价值判断，也就是对教育活动过程所产生的效益和结果进行判断，看其是否具有价值。无论是过程评价还是终结性评价，其核心内容都是对教育现象进行价值判断。它们的依据是教育目标和评价方案这是不言而喻的。教育评价也不是毫无目的为了评价而评价，它应该为教育决策提供可靠的信息和依据。这正是我们进行教育评价的目的，也是开展教育评价活动之前就应该明确的。因此，他们认为教育评价只定义为对教育现象进行描述和价值判断就足够了。

4. 美国教育评价标准联合委员会的评价观。1981年，美国教育评价标准联合委员会综合了各种评价观点，给教育评价下了一个综合性的定义：教育评价是对教育目标和它的优缺点与价值判断的系统调查，为教育决策提供依据的过程。这个定义集中了上述三种观点的基本内涵，既弥补了泰勒对教育评价解释的不足，又强调了价值判断，并且肯定了教育评价应该为教育决策服务。

5. 我国教育评价专家、学者的评价观。20世纪80年代以来，教育评价在我国得到了迅速的发展。许多教育评价研究者认为我们是社会主义国家，与西方社会制度不同，不能完全照搬美国的评价观。应该紧密结合我国的国情，从教育实际出发，对教育评价作出科学的解释。于是各种定义纷纷提出：有人认为教育评价是对教育的社会价值作出判断的过程；有人认为教育评价是指按照一定的价值标准对受教育者的发展变化及构成其变化的诸种因素所进行的判断；有人认为教育评价是按照一定社会的教育性质、教育方针和政策所确立的教育目标，对所实施的各种教育活动的效果、完成教育任务的情况以及学生学习成绩和发展水平进行科学判定的过程；有人认为教育评价是根据一定的教育目标和教育价值观对教育系统的结构和功能，对教育活动过程和它的后果以及与之有关的事物进行价值判断，从而为人的自我完善和教育决策提供信息和依据的过程；有人认为教育评价根据一定的教育价值观或教育目标，运用可行的科学手段，通过系统地搜集信息、资料和分析整

理,对教育活动、教育过程和教育结果进行价值判断,从而为不断自我完善和教育决策提供依据的过程。

分析国内外专家、学者对教育评价所给的定义和描述就会发现,尽管他们在提法和描述上有许多差异,但从不同侧面反映出教育评价具有以下特点:

第一,教育评价是一个过程,而且是一个有严格程序的、连续的系统活动过程。这个活动过程是有计划、有目的地进行的。一般要经过确定评价目标,搜集信息资料,综合分析资料,形成价值判断,指导今后行动等程序。

第二,教育评价以教育目标或一定的教育价值观为依据。

第三,教育评价是一种价值判断活动,它始终以对评价对象的功能、作用、状态给予价值判断为核心。

第四,教育评价以科学的评价方法技术为手段。

第五,教育评价为人的自我完善和教育管理、教育决策提供有用的信息和依据,指导行动,为管理和决策服务。

对教育评价下一个科学的定义目前还难以达到共识。笔者认为所谓科学的定义并不在于文字表述的长短,关键是能抓住这个概念的基本内涵和本质特征。当然,在表述上应该简明扼要,便于记忆。至于它的功能和作用则不一定写入定义之中。从本节一开始对评价定义的解释中已经看出,评价是指衡量人物或事物的价值,而价值是指事物的用途和积极作用。要想知道事物有无用途,有无积极作用,就要依据一定的目标和标准进行价值判断。现代教育评价的对象已经扩大到了整个教育领域,即一切教育现象都是教育评价的评价对象。因此,教育评价的定义自然而然地应当表述为“对教育现象进行价值判断”。这种表述不仅符合教育评价的字面解释,而且抓住了教育评价的关键和核心,并且叙述简练,便于记忆。

第二节 教育评价的现实意义

一、教育评价是现代教育科学研究的三大领域之一

现代世界上教育科学研究的三大领域是教育基础理论研究、教育发展研究、教育评价研究。教育评价的思想虽然源远流长,如果把古代考试视为教育评价的渊源和萌芽,那么可以追溯到3 000多年前。但是作为一个独立的研究领域,形成教育科学的一门重要分支学科却是近几十年的事情。当前,

世界各国都把发展教育、培养人才作为经济与社会发展的战略重点，一般都重视了增加教育投入，开发智力资源，运用教育评价来促进教育事业的发展，促进国民素质的提高。在我国，教育评价的研究虽然起步较晚，但发展很快，已经成为教育科学研究的重要领域之一。在1990年我国成立了全国普通教育评价专业委员会，已召开过十几届教育评价学术研讨会。1994年成立了中国高等教育学会高等教育评估研究会，也曾多次召开高等教育评估学术研讨会。“八五”期间，国家教委对教育评价的理论与方法分别从高等教育和普通教育两个领域立项开展学术研究。我国教育评价工作者，从目标制定、方案设计、信息收集处理、评价组织实施、结果获得与解释以及评价模式的构建等方面，就其理论与方法技术的有关问题进行了深入探讨，并取得了累累硕果。十余年来，各级各类教育杂志及有关报刊发表的教育评价文章，据粗略统计已过千篇，教育评价著作已有几十部。1998年还出版了我国第一部《教育评价辞典》（陶西平主编，北京师范大学出版社，1998年9月）。

就高等教育评价而言，1985年11月国家教委颁布了《关于开展高等工程教育评估研究和试点工作的通知》，以后又确定了在北京市、上海市、机械电子部、煤炭部等单位首先进行试点。这些单位分别开展了高校办学水平评估、专业评估和课程评估等。国务院学位办还先后组织了财政学、货币银行学、国际金融、辩证唯物主义与历史唯物主义、马克思主义哲学史等专业硕士和博士学位授予质量检查与评价活动。教委科技司曾组织对科研机构 and 重点实验室进行评估。教委高教司还组织了委属高校进行有关校园、学生学习和生活环境的检查评估及办学水平的评估活动。除此之外，全国各省市也都开展了各种各样的评估活动。这些实践活动，对于我国高等教育评价理论的形成起了极大地促进作用。近年来，全国高等学校还开展了课程评价、学科评价、教学评价、教师评价、教学管理评价、德育评价、能力评价等各种类型的评价活动。不少学校都有选择地进行了试点。

普通教育评价研究开展得比较普遍。尽管在全国各地发展并不平衡，但全国普通教育评价专业委员会的分会已遍及全国（台湾和西藏还没有成立分会）。从1986年开始，每年召开一次全国性教育评价学术研讨会，各地专家、学者进行学术交流和教育评价实践经验介绍，对教育评价的理论问题、实践问题、方法技术问题进行探讨和研究。每年的研讨题目和研究重点都有所变化，特别是注意紧密联系我国教育实践情况和教育改革热点问题开展研究探讨。1989年重点研讨了评价指标体系及评价中定性、定量的辩证关系；

1990年的研讨重点是如何发挥教育评价的导向功能和坚持社会主义方向。在学术会上还讨论了如何把握普通教育评价的生长点问题。提出了评价理论的深入研究，学生质量综合评价研究，评价方法技术研究，教育评价制度研究等问题。当年科研成果77项，专著9本；1991年学术会研究重点是教育评价的功能和效益问题。教育评价效益问题研究具有很强的现实意义，为大家所关注。

“八五”规划国家教委给全国普通教育评价专业委员会下达了一个重点课题：普通教育评价基本理论与方法研究。下属五个子课题：

- (1) 研究我国普通教育评价的现状；
- (2) 研究具有中国特色的普通教育评价基本理论；
- (3) 研究符合中国实际的评价方法技术体系；
- (4) 研究符合中国实际的普通教育评价模式；
- (5) 研究符合中国情况的德育评价。

此课题分别由全国各教科所和专业评价人员进行探讨研究。根据课题研究结果研究人员发表了一批学术论文和专著，对教育评价理论与方法的进一步研究起到了推动作用。

1994年学术会的研讨重点是“教育评价面临的挑战和今后的生长点”。当时提出的问题是，我国由计划经济向市场经济转轨对教育评价提出了挑战：(1) 教育思想和教育观念必须转变；(2) 教育体制的转变及价值取向多元化；(3) 穷国办大教育；(4) 教育改革的深入发展；(5) 市场经济下的新问题如何评价，比如资产评价、社会化服务评价等新问题；1995年研讨会的重点是德育评价和评价模式研究；1996年的中心议题是“教育评价在深化教育改革中的地位和作用”，尤其是在实施素质教育中的地位和作用如何。当年在湖南汨罗召开了一次素质教育评价学术研讨会。一方面学习汨罗的先进经验，另一方面又进行了学术交流与研讨；1997年研讨的重点仍然是“素质教育评价问题”，在山东寿光召开了学术会。会上确认今后的研究重点是评价思想、评价主体、评价效度。与会者认识到应试教育向素质教育转变是一个系统工程，必然会带来评价思想、评价主体、评价模式的彻底改变。评价效度也是今后应该重视的一个重要问题。

随着我国教育督导制度的恢复与发展，督导评价也很快地发展起来。各省市自治区均成立了督导（学）室，负责实施教育督导与评价。我国现阶段的教育督导是指县以上行使教育督导职权的机构根据本级人民政府、教育行政部门和上级教育督导机构的决定或根据有关的方针、政策、法规组织实施

对下级人民政府的教育工作、下级教育行政部门和学校的工作进行监督、检查、评估、指导，并向被督单位，同级或上级人民政府，教育行政部门，上级教育督导机构通报和报告督导结果，提出意见或建议，亦可向社会公布结果。教育督导的根本目的是保证党和国家教育方针、政策、法令、法规的全面贯彻执行，保证教育目标的实现，促进教育改革和教育事业的顺利发展，提高教育质量，增强社会效益，完善教育管理体制，使教育更好地为社会主义现代化建设服务。

督导评价是教育督导工作的一个重要方面，是科学地进行督导工作的重要工具。没有科学、合理、公正、客观的督导评价，要想搞好教育督导工作是不可能的。正因为如此，教育督导评价受到了高度重视。1998年，教育部督导办副主任白景龙同志在教育督导室主任会上提出“建立对中小学校教育评估制度，是教育督导部门中心工作之一，各省都要把这项工作列入重要工作内容，加强工作的力度”。教育部督导办主任郭振有同志在《中国教育报》1998年7月28日的文章中指出对普通中小学、对县级教育行政部门和对县级人民政府教育工作等三个方面的督导评估构成了我们对地方基础教育工作督导评估工作的基本框架。督导评价（估）是教育评价的一部分，是教育评价的理论与方法在教育督导领域的科学运用。因此，督导评价的现状和发展也从一个侧面反映了教育评价的现状和发展。对教育督导评价理论与方法技术的研究，也是教育评价研究的一个重要方面。

教育评价在西方已有一百多年的历史，目前已发展到了专业化的阶段。在美国高校，1984年就有117个侧重于教育评价的专业点，他们有权授硕士、博士学位。各种专门研究教育评价的杂志也大量涌现。在美国有《评价研究》《评价评论》和《教育评价和政策分析》等多种杂志，其原因就在于教育评价学是教育科学中一门应用性很强的分支学科。

二、教育评价是一门综合性、交叉性、边缘性教育学科

教育评价是一门综合性、交叉性、边缘性的教育学科。它以马克思主义哲学为哲学基础；以教育学、心理学、教育管理学为理论基础；以系统科学为方法论基础；以教育统计、教育测量、计算机软件等为基本工具。教育评价学科的发展反过来又给相关学科提出了新的研究课题，推动了相关学科的发展。

1. 教育评价与马克思主义哲学

马克思主义哲学是无产阶级世界观，是对自然界、人类社会和思维发展