

高等教育评价

原理与方法

杨异军 庞德谦 邓忠良 编著

陕西师范大学出版社

高等教育评价原理与方法

杨 异 军
庞 德 谦 编 著
邓 忠 良

陕西师范大学出版社

高等教育评价原理与方法

杨异军 庞德谦 邓忠良 编著

*

陕西师范大学出版社出版发行

陕西省岐山县印刷厂印刷

*

升本787×1092 1/32 印张 7 375 字数 154千

1988年8月 第1版 1988年8月第1次印刷

印数 1—3000册

ISBN7-5613-0135-9

G·135 定价：2.50元

前 言

教育评价是现代教育科学的三大研究领域之一。世界发达国家都把教育评价，特别是高等教育评价作为实现教育目标、推动教育事业发展、增强经济竞争能力的重要手段和有力杠杆，研究十分活跃。

我国高等教育面临着世界新技术革命的挑战和改革、开放、搞活的新形势，如何利用教育评价的竞争机制，使高等学校增加压力，产生动力，形成活力，更好地实现教育目标，努力适应国家政治、经济和科技发展的需要，已引起全社会的共同关注。为此，《中共中央关于教育体制改革的决定》中明确提出：“国家及其教育管理部门要加强对高等教育的宏观指导和管理。教育管理部门还要组织教育界、知识界和用人部门定期对高等学校的办学水平进行评估，对成绩卓著的学校给予荣誉和物质上的重点支持，办得不好的学校要整顿以至停办。”教育评价将是我国今后进行教育决策的科学依据和教育管理的重要工具。因此，开展高等教育评价理论研究和实践活动，建立具有中国特色的高等教育评价制度，就成为发展高等教育事业和教育科学研究的必然要求。

我国开展高等教育评价领域的研究时间较短，目前尚未建立起系统的、完整的评价理论体系。为了给广大高教工作者和教育管理干部开展高等教育评价工作提供方便，我们在广泛收集资料、认真学习国内外教育评价理论与实践经验的基础上，吸收了最新研究成果，参阅引用了有关资料，集思广

益，博采众长，编写了这本《高等教育评价原理与方法》。所以说本书凝结着众多学者专家的智慧 and 心血，我们在此谨致谢意。

本书采用集体编写提纲，分工执笔编写的办法。各章的执笔人是：第一、二、五章，庞德谦；第三、四章，庞德谦、邓忠良；第七、八章，邓忠良；第六、九章，杨异军。本书初稿形成于1987年5月，曾内部交流，在广泛征求意见的基础上做了较大修改。修改统稿由庞德谦、邓忠良承担，杨异军最后审阅定稿。

本书编写过程中受到陕西省高教局科研处、陕西师范大学出版社以及作者所在单位的大力支持。西安交通大学石方、叶尚思同志为修改本书提出了宝贵意见，并给予多方面的帮助，陕西师范大学副校长李钟善同志审阅了全稿。在此一并表示衷心的感谢。

由于我们的水平和实践经验有限，有不少问题没有解决或解决得不够好，有待今后继续努力。本书难免有疏漏和错误之处，恳请读者批评指正。

编 者

1988年4月10日

目 录

第一章 引 论	(1)
第一节 教育评价概念的科学含义	(1)
第二节 教育评价的产生与发展	(6)
第三节 我国高等教育评价的现状与趋势	(15)
第二章 高等教育评价学的研究对象与社会功能	(20)
第一节 高等教育评价学的研究对象与学科性质	(20)
第二节 评价在高等教育历程中的功能	(29)
第三节 我国开展高等教育评价的目的意义	(33)
第三章 高等教育评价学的理论依据	(38)
第一节 辩证唯物主义的哲学原理	(38)
第二节 马列主义教育原理和我国高等教育的实践经验	(47)
第三节 以系统论为核心的现代管理科学原理	(53)
第四节 现代数学若干理论原理	(62)
第四章 高等教育评价的原则与标准	(68)
第一节 高等教育评价的基本原则	(68)

第二节	高等教育评价标准·····	(74)
第三节	高校评价中应注意处理的几种关系·····	(90)
第五章	高等教育评价的类型与模式·····	(103)
第一节	高等教育评价的类型·····	(103)
第二节	高等教育评价的模式·····	(117)
第六章	高等教育评价方法·····	(128)
第一节	考试评价法·····	(128)
第二节	专家评价法·····	(134)
第三节	数学方法·····	(144)
第四节	系统方法·····	(153)
第七章	高等教育评价技术(一)·····	(165)
第一节	评价指标体系的设计·····	(165)
第二节	权系数集合的设计技术·····	(177)
第三节	指标体系的量化技术·····	(185)
第八章	高等教育评价技术(二)·····	(194)
第一节	评价结果的检验·····	(194)
第二节	评价误差分析与处理·····	(206)
第九章	高等教育评价的程序与组织·····	(214)
第一节	评价程序·····	(214)
第二节	评价组织和评价人员的素质·····	(221)

第一章 引 论

当今世界，正处在一个巨大的变革时期。以信息、材料、能源为代表的新技术革命的浪潮，正引导人类迅速地从工业化社会走向信息化社会。信息化社会的特点在于靠知识、技术上的突破，而不象工业化社会那样依靠能源与原材料的消耗来创造财富。面对信息化社会的严重挑战，国家的力量将取决教育水准的高低，教育的质量关系着民族的未来。因此，世界上发达国家都把发展教育、延揽人才看成是跟上时代发展的战略基点，除了普遍增加教育投资外，一个十分重要的措施就是运用教育评价的手段去实现教育目标，以增强教育的竞争能力。教育评价正在成为一股波及全球的潮流，深刻地影响着教育事业的发展。因而，系统地学习、研究和发展教育评价的理论与实践，就很自然地成为摆在我国教育工作者、教育管理部门乃至全社会面前的重要课题。

第一节 教育评价概念的科学含义

一、教育评价概念的科学含义

评价原指评论货物的价格。《宋史·戚同文传》曰：“市物不评价，市人知而不欺。”以后广泛用作衡量人物或事物的价值。那么，什么是教育评价？它的科学含义是什么？对于这个基本问题，由于评价者观察事物的角度不同，形成了

概念界说上的差异，迄今为止尚未有一个比较确切的、统一的和科学的定义。

辩证唯物主义认识论认为，评价属于认识范畴。人的认识和思维具有反映、评价和创造的特性，人类认识活动在本质上是反映、评价和创造的统一。这是因为，人类在认识活动中要如实地反映客观事物，把握客观事物的本质和规律，但又不以反映客观事物为满足，以复写、摄影为最终目的，而要在此基础上继续判别客观事物对人类生活的价值和意义。对客观事物进行评价，表现主体的意志和情感，尔后在头脑中创造客观事物的理想蓝图，以确定实践活动中所追求的具体目标。教育评价作为评价者主体对教育客体的认识运动，是评价者在认识教育客观规律的基础上，对教育客体及其属性给予价值上的判断，进而在头脑中创造教育客体的理想蓝图，并确定教育实践活动所追求的具体目标的认识过程。

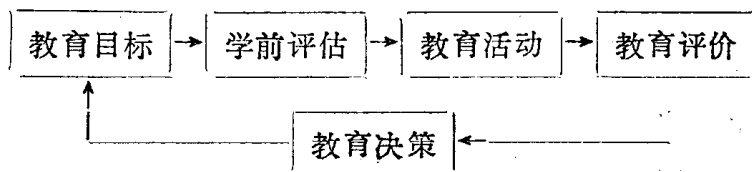
系统论的观点认为，教育评价是教育这个庞大的开放系统的信息反馈通道。它是根据评价对象的功能和特征，规定指标体系，通过对其系统状态所提供的信息进行观察、分析和比较，针对其偏离指标体系的原因和程度，制订和选取最优决策，以及采取这些决策对系统状态所产生的后果进行综合研究的过程。

经济学家认为，教育评价是教育计量与统计理论得以量化并实施的中心环节。根据教育评价的量化结果，可对教育投资进行择优分配，从而在有限的教育投资条件下，获得尽可能好的教育效益。

教育科学理论认为，教育评价是一种有系统的寻找收集资料并参照合理的衡量标准，对教育客体作出价值判断，以

便协助决策者在诸种可行的途径中择一而行的动态过程。

以上对教育评价概念所作的定义虽不尽相同，但我们可以看出，它们对教育评价的本质特征有几点共同的认识。第一，评价是一个过程，即是一种特殊的、有目的的继续性活动，其间包含一系列的步骤和方法；第二，这一过程是一种有系统的、正式的过程，其活动一般是确定目标→收集资料→分析资料→形成判断→作出决策；第三，评价的中心永远是对评价对象的系统状态及其功能给予价值上的判断以供决策者在几种途径中择优而行；第四，教育评价不是教育历程的终点站，并非代表教育活动的结束，教育评价的主要目的在于分析教育得失和诊断教育困难，作为实施补救教育的依据。这一特征，通过图1-1可以清楚地反映出来。



通过以上分析，我们可以初步将教育评价概念的科学含义概括为：根据正确的教育价值观的目标函数，采用科学方法与途径，多方面搜集有关的事实性资料，科学地判断对象的系统状态与功能属性及其转化为主观和客观效用的过程。当然这仅仅只是概括了教育评价概念的基本含义，并非是一个确切的定义。因为教育评价尚处于形成发展阶段，它还有待于实践的不断完善。

二、教育评价与相近概念的区别

为了正确开展教育评价工作，还必须划清评价与测量、测验、评定、评估、鉴定以及教育研究等相近概念的区别。从表象看，评价、测量、测验、评定、鉴定、评估以及教育研究都以教育客体为研究对象，但是由于研究问题的目的、方法、途径和角度等不同，从而使这些概念的内涵有很大的差异。

评价不同于测量。区别在于，测量是事实的描述，它是根据量尺或其它测量工具，用数字描述被测对象的性质。例如依据比西量表（Binet—Simon Scale）测量的结果，可以得出智力商数（IQ），用这个商数来描述个人的智力特质，这个历程可称智力测量。评价则是根据某一标准，对所测量到的数量给予价值上的判断。例如一个学生数学成绩测验获得80分，这个分数代表他成绩优异呢？中等呢？或低劣呢？单独从分数去看，无法获知，只有根据一定的参照目标或个人的能力去评判，才能确定。评价必然建立在测量的基础上，但测量不能代替评价。正象丈量土地面积，这是测量不是评价，结合地形、土质、地理位置、水利条件等，才能给予土地价值的判断，这才是评价。

评价也不同于测验。测验重在数量的测定，因而较具有客观性，而评价则要求质量并重，主观的与客观的方法兼用；测验重在“获得”事实，而评价则重在“解释”事实，作出价值判断；测验是一种单一的活动，而评价则为一系列的活动。从目的追求上区分，测量的目标在于证明量的多少，而评价的目标在于通过对事物的了解，进而求得对成绩的肯定以及对缺点的改进。

评价同评定、评估、鉴定等概念也有很大区别。评价

(Evaluation) 是通过系统地搜集信息, 对客体作出价值判断的过程, 其对象主要是教育制度、教育政策、学校、教师、学生、课程、科研成果等; 评定 (Assessment) 则用客观的测量方法对个体特征作出判断, 在教育领域中, 评定主要用来测量教与学的数量与质量; 评估 (Appraisal) 用于对客体的内在潜力、以往工作成绩或未来发展作出判断, 诸如对国家或大系统的教育水平与发展战略、学校的办学方向、办学水平和办学条件等所作的估价; 鉴定 (Accreditation) 则由校外的鉴定机构或其它组织对学校工作或专业予以承认和认可。由此, 我们可以看出, 测量、测验等活动只是评价的一种手段, 而评估、评定、鉴定等活动则是评价的重要形式。

此外, 我们还应该注意教育评价与教育研究两个概念的区别, 因为这两者有许多相似之处, 容易被人们混为一谈。两者都是在进行严肃认真的科学探讨; 两者都使用测量的方法; 两者都对其所收集的资料进行系统的分析; 两者都将活动的经过和结果形成正式的报告等等。尽管有如此之多的貌似之处, 但它们在本质上却有很大区别: 第一, 两者研讨问题的目的不同, 教育研究是通过研究去获得某种“结论”, 而教育评价者则是通过评价去获得某种“决定”; 教育研究者的兴趣在于通过研究现象抓住本质, 而教育评价者的兴趣是通过对现象与本质的了解, 以便形成“决策”用以指导人的行为。第二, 两者探讨结果的“可类化性”不同, 教育研究者的兴趣在于识别教育上有关的各种变化因素之间究竟有何种关系, 理想的结果应该是研究结论可以普遍类化到各种相似的情景之上, 研究结论的类化性越高, 则认为研究的成

果价值越大，与此相反，评价则把结论重点放在某一特殊的教育现象上，它没有将其结果普遍类化到其他情景的意图。第三，两者所重之价值也不相同，研究者旨在追求科学的真知，而评价者所关心的仅是决定某种教育现象的价值。

第二节 教育评价的产生与发展

一、教育评价的思想渊源

教育评价的思想源远流长。据美国知名的教育评价专家 P. H. Bosc, W. T. Popham 及 B. R. Worthen 等人研究，认为有系统的正式教育评价活动起源于中国隋代开始的科举制度，它是用一套完整的考试方法来挑选人才的活动。我国不少学者认为，倘若以对教育活动考核为基点，那么中国教育评价的历史还可上溯到更遥远的西周时期，那时的奴隶制国家已建立了管理与考核学生成绩的制度。《礼记·学记》中说：“比年入学，中年考核，一年视离经辩志，三年视敬业乐群，五年视博习亲师，七年视论学取友，谓之小成；九年知类通达，强立而不反，谓之大成。”这是最原始的教育评价指标，那时的学校教育就已含有“选士”、“取士”的目的。《礼记·王制》篇说得很清楚，对于大学中学习好的，学校呈报负责国家教育的大乐正，“大乐正以善于莠而开诸司马，曰进士。司马辩论官材，论进士之贤者，以告于王而定其论。论定，然后官之。任官，然后爵之。位定，然后禄之。”这里需要注意的有两点，第一，隋唐以后科举中出现的秀才、举人和进士称号，就是从西周时

期的秀士、造士、进士称号衍变来的；第二，西周开始对进士进行的由司马负责的辩论考核活动，已带有浓厚的人才质量评价色彩。到了汉代，选士制度从学校中分离出来而成为选士的察举制度。察举制的选举科目主要有贤良方正、孝廉和秀才、明经等科，贤良方正科多由中央直接挑选，皇帝亲自策试授以官职，为朝廷召致具有特殊才能的人；孝廉等科则是经过乡举里选，为地方选拔人才。察举制开创了布衣卿相的新局面，如牧豕海上的公孙弘，由于对策第一，几年之间，官至宰相，封侯。所以察举制在当时是有积极作用的，并且察举制使取士制度化了，选举加策问，这对隋唐科举制度的产生是有影响的。到了隋唐以至清末，又以科举制代替了汉以后创立的九品中正选士制度。隋炀帝大业二年（公元606年）始置进士科，这是我国教育制度上的一项具有历史意义的重大转变。科举是分科选举的意思，如唐代科举常行的科目就有秀才、明经、进士、明法、明书、明算等六科，最常行的仅有明经、进士两科。科举制的特点是通过逐级考试的办法来挑选人才。它与察举制的区别，从选择的标准看，选举重德望，才学次之，而科举则全凭考试；从形式上看，选举是州郡守举其所知，由朝廷策试，而科举则是州郡逐级考选，最后送朝廷考试。科举制与九品中正制的区别，主要在选举权和任用权上。九品中正制选拔人才表面上通过中央吏部铨选，实际上选人权完全掌握在地方上担任中正的世族手里。科举制虽然也由地方州县考选人才，但要到中央参加考试，而且考试的结果只能给予学衔，还要经过吏部考试才能放官。很明显考选和任用之权都由地方转到中央了。科举制不仅适应了封建王朝政治统治的需要，而且使其本身

在考生来源及考试办法、考试科目、内容与方 法、录取程序、人数与及第后的待遇等方面，形成了一套较为完备的考选制度。特别需要指出的是，科举制逐步建立起了衡量知识才能的客观标准。以进士科而论，策论是有关治世的才能，诗赋杂文是关于文艺、应用文的知识和能力，经义是对于儒家思想的理论水平。尽管这些内容不能不带有封建性，而且在当时杜绝关节也很难做到，但要通过这种考试毕竟需要具有一定的文化水平和治世才能。事实上，在唐代及第的进士中确实得到了各方面的人才，如名相房玄龄、大诗人白居易、大文学家韩愈等都是进士出身。所以，科举制在我国古代是起过积极作用的。它在我国沿用1300多年，形成了一整套测验学力选拔人才的完整制度，这种考试选拔官吏的办法对于当时欧亚各国的文官考试也有较大影响，这是举世公认的。当然我们在看到它的积极作用的同时，还应看到科举制度对中国古代文化教育的消极影响。

作为古代教育评价的另一形式——教育视导制度，在我国也源远流长。早在汉代，视学活动已在我国出现，那时太学教师（五经博士）除了在校教书，有时还要视察地方的政教。视学活动发展到清末时形成了相当完备的制度，国家设有视学官，专事巡视京外学务，并制定有《视导官章程》。

总之，中国开世界教育评价之先声是毋庸置疑的。但是，教育评价同古代自然科学、哲学等科学的产生与发展是多源头的一样，很早就以各种不同的形式，在世界其他国家也出现了。古代希腊就有关于教师考核的记载，英国等西方国家实行教育视导制度的历史也比较久远，在美国19世纪初兴起的教育测量运动则揭开了现代教育评价的帷幕。

二、现代教育评价的产生与发展

现代教育评价作为一门科学，形成于美国，至今约有200年的历史，大体经历了四个发展时期。

(一) 19世纪初—20世纪30年代，称之为萌芽期，或者叫作测量运动阶段。这一时期，美国出现了心理测量和教育测量。它以客观地确定个体所具有的能力、知识或理解水平为目的，对教育结果的判定采取了科学量化的成绩考核法。1845年，美国波士顿的教育工作者，首次使用对学生进行测验的方法，来评价学校教学效果，测验形式一改以往单纯的口头回答问题的形式，转向使用较为客观的书面论文形式，从此开始了长达百年的使用标准化测验的历史。1887年到1898年，被誉为“教育科学研究之父”的约瑟夫·赖斯为了解学习时间是否被有效的利用，设计了一套学业成绩测验，学生在测验中取得的成绩，成为他判断的标准。赖斯发现，有些学校每星期花在拼写单词上的时间多达200分钟，而有些学校则少至10分钟，但这些学校的学生在测验中取得的成绩并无显著差异。1904年，阿尔弗雷德·比奈设计了一套心理测验量表，其主要目的在于从普通班级中发现那些智力发展迟钝的学生，这套测验方法为以后的智力测验打下了基础。这一时期的评价活动，主要有以下几个特点：

1. 评价和测量是两个可以互换和通用的概念，实际上，测量一词较评价更为常用。正象桑代克1904年在《精神与社会测量学》一书中所倡导的，“凡存在的东西都有数量，凡有数量的东西都可以测量”。用数学方法检查教育效果，无疑是教育评价史上的一大进步。

2. 评价和测量都与科学探究方法紧密联系在一起。

3. 评价和测量的主要焦点集中于被测对象的个体差异，而很少直接联系造成这种差异的教学方案和课程。

4. 受当时工业界流行的科学管理思潮的影响，科学化、标准化的测量成为教育评价追求的目标和主要手段。

以测量为基本手段的评价徘徊了100来年，终于在20世纪30年代有了重大突破。

(二) 20世纪30年代——50年代，称为泰勒期，或者叫做测验批判阶段或评价阶段。教育测量虽然不失为一种积极的教育管理手段，但却具有致命的弱点。因为对于教育这个极为复杂、相当漫长而又涉及诸多因素的过程现象，是很难统计和测量精确的，尽管美国在1928年时的测量法达3000种之多，仍不能把诸如创造兴趣、鉴赏能力、实践技术和社会态度等要素反映出来。因此，从1930年起，美国开始了对教育测验的批判。当时，有个叫进步主义联盟（P·E·A）的组织，它由许多受杜威思想影响的教育家组成。该组织在会长艾钦领导下，以泰勒教授为主，在7所大学和30所中学开始了以进步主义高中的经验与传统高中的经验之成效问题为背景的教育实验。这一实验于1932年提出，历时6年，直到1940年才完成，因此教育史学家们把这次实验叫做“八年研究”。在这次实验中，泰勒教授提出了全新的教育评价概念，并从他的两项研究成果中抽象出了科学的评价原则。他的理论的出发点是注意教育效果的价值观。他首先分析教育应达到的目标，然后用这个目标评价教育效果。他认为，通过教育评价，就能促使教育措施向理想的目标逼近。在实验过程中，他们不仅把应达到的目标和实验效果联系起来，同