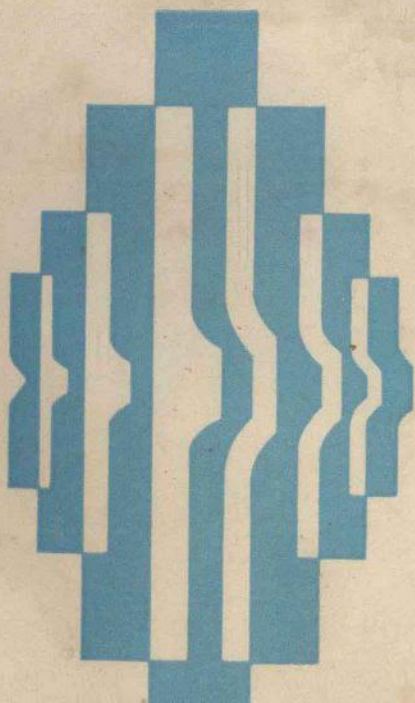


郑玉林 高笑天 编著
于文安 王文波

现代教育 评估概论

XIANDAI JIAO YU
PING GU GAI LUN



● 黑龙江教育出版社

现代教育评估概论

郑玉林 高笑天
于文安 王文波 编 著

黑龙江教育出版社

1989年·哈尔滨

现代教育评估概论

郑玉林 高笑天 编著
于文安 王文波

责任编辑：王贵桓

封面设计：陈 江

黑龙江教育出版社出版(哈尔滨市道里森林街42号)
黑龙江省教育委员会印刷厂印刷·黑龙江省新华书店发行

开本787×1092毫米 1/32·印张 8.5·字数 180千

1989年9月第1版·1989年9月第1次印刷

印数：1 — 2000

ISBN 7-5316-0960-9/G·686

定价：3.45元

序

评估作为教育信息反馈的重要环节，作为现代教育管理与指导工作的重要手段，对它的研究和应用越来越被人们所重视，有人把它与教育基础理论、教育发展并列为当代教育科学研究的三大领域，这不是没有道理的。正因为如此，掌握它的理论和方法应是现代教育管理者与指导者所必需，因而在我院教育管理与思想管理专业均开设了这门课程，这对于提高教育的管理水平具有重要意义。

目前国内外教育评估著作不少，各有其所长，郑玉林等同志撰写的《现代教育评估概论》一书，在总结我院几年来教学实践经验的基础上，概括了现行的各著作之所长，简明扼要，通俗易懂，能从教育工作的已有基础与实际需要出发，初学者可从中得其要领。它不仅可作为教育管理专业的教材，同时也可作为广大教育管理者的自学用书。

该书是郑玉林、高笑天、于文安、王文波同志理论与实际相结合的产物，是他们辛勤钻研的成果。

谨祝贺该书的出版，赞美它的诞生，相信读者们一定会从中受到教益。

李 诚 忠

一九八九年十月二十六日于哈尔滨

前 言

一所学校的业绩，不应仅以它向上一级学校输送了多少合格的新生或培养了多少高级人才来衡量，而应看它是否在全体教职员工的积极努力下，按党和国家的要求，使每个接受教育的青少年儿童在德智体美劳等方面都得到了应有的发展，是否造就出了社会主义现代化建设所需要的各级各类人才。这必然要涉及到教育的全过程和教育过程所有参与者的全部行为及其最后结果的社会价值的确定。要解决这个问题，并尽可能做到客观公正，就必须借助于现代教育评估。

现代教育评估理论与方法的研究与教育基础理论研究和教育发展理论研究并列，成为当代教育科学的三大研究领域之一。现代教育评估的研究成果作为现代教育管理与指导的主要手段，已应用于教育实践的各个方面。

近几年来，教育评估问题已引起我国教育界的普遍重视。许多教育工作者对现代教育评估的理论与实践问题进行了广泛的探讨和研究，取得了可喜的成果。但是，由于这项工作刚刚起步，普及程度不够，许多教育工作者对教育评估的基本理论与实践问题还存在一些模糊认识，甚至是误解。这就难免影响到教育评估工作的效度和信度。为澄清人们对教育评估的模糊认识与误解，普及现代教育评估的基本理论，推动教育评估工作的顺利开展，我们编著了这本《现代教育评估概论》，供各级各类教育行政管理干部和广大教育工作者学习、参考。

在本书编著过程中，我们从本国教育评估工作的实际出发，吸收国内外的研究成果，力求以简明、通俗的阐述为广大读者提供现代教育评估的基本原理和方法基础。并结合实际工作的需要，编辑了一些教育评估模型和具体材料，以供借鉴。

本书编著的具体分工是：郑玉林一第七、十章；高笑天一第一、二、三、四、五、九章；于文安一第八、十一章；王文波一第六章。全书由高笑天、王文波统稿。

在本书编著过程中，我们参考和引录了国内外的一些文献资料，黑龙江省教育行政干部学院院长李诚忠教授为本书撰写了序言，在此一并表示谢意。

编 著 者

1989年7月

目 录

前言	(1)
第一章 教育评估的历史与发展	(1)
第一节 教育评估的历史渊源	(1)
第二节 教育评估的发展趋势	(7)
第二章 教育评估的意义与作用	(14)
第一节 现代教育评估的实质	(14)
第二节 现代教育评估与相关活动的关系	(17)
第三节 现代教育评估的作用与功能	(21)
第三章 教育评估的分类	(25)
第一节 根据教育评估目的分类	(25)
第二节 根据教育评估内容分类	(30)
第三节 根据教育活动主体分类	(33)
第四节 根据其他因素分类	(34)
第四章 教育评估的指标体系	(37)
第一节 教育评估指标体系的分类构成	(38)
第二节 制定教育评估指标的依据	(42)
第三节 对教育评估指标的基本要求	(45)
第四节 教育评估指标制定的步骤和方法	(49)
第五章 教育评估常用的检验工具	(54)
第一节 考试	(54)
第二节 测验	(62)
第三节 问卷	(65)
第四节 谈话	(67)

第五节	材料审查	(69)
第六节	工作考核或达标检查	(69)
第七节	行为观察记录	(70)
第六章	教育评估过程	(73)
第一节	教育评估的规划与准备	(73)
第二节	教育评估计划的实施	(80)
第三节	材料的分析与评估	(82)
第七章	教育评估的指导思想和基本原则	(86)
第一节	教育评估的指导思想	(86)
第二节	教育评估的基本原则	(92)
第八章	教育评估的量化分析法	(100)
第一节	指标量化法	(101)
第二节	量化评定法	(104)
第三节	数理统计法	(106)
第四节	统计推断法	(129)
第五节	综合评判法	(142)
第六节	S—P 表综合评估法	(151)
第九章	教育评估的定性分析法	(158)
第一节	等级法和比较排列法	(159)
第二节	评定法	(160)
第三节	总结法	(163)
第十章	教育评估方法的应用方式	(165)
第一节	相对标准评估与绝对标准评估	(165)
第二节	分析评估与综合评估	(167)
第三节	自我评估与他人评估	(168)
第四节	量标评估	(170)
第十一章	学校教育评估简介	(171)

第一节	教学评估	(171)
第二节	学力评估	(174)
第三节	思想品德教育工作评估	(176)
第四节	学校管理工作评估	(182)
第五节	教师评估	(186)
附:	评价心理及其调控	(189)
第一节	概述	(189)
第二节	主评心理与调控	(191)
第三节	被评心理与调控	(198)

附表

第一章 教育评估的历史与发展

第一节 教育评估的历史渊源

评估是对实践活动过程和结果的社会价值进行判断和估量的过程。它是一切有目的的社会实践活动的重要组成部分。任何一种社会实践活动，从其目的确定到活动的实施，以至最后结果的确认，都与评估密切联系着。评估的正确与否，影响着社会实践活动全过程的发展及其活动目标的实现。

教育是人类特有的有目的的活动，所以，评估问题也存在于一切教育实践活动之中。随着教育实践活动的发展，教育评估已逐渐形成一门独立的学科，成为当代教育科学研究的三大领域之一。

作为一门独立学科的教育评估，其形成经历了漫长的历史。

教育评估作为教育实践的组成部分，是随教育的产生而产生，伴随着教育的发展而发展的。但是，教育评估作为一种有目的的活动过程，则是在教育实践发展到一定阶段的产物。

纵观教育评估发展的历史，它大致经历了经验评定、量化评价、科学评估三个阶段。

一、 经验评定阶段

教育评估的经验评定阶段始于社会人才选拔制度的形成

时期。当时的教育评估活动是应人才选拔时进行政治、道德审查,学识、能力认定以及体质等方面的考查需要而进行的。这阶段的教育评估缺乏客观标准,评估主体往往以自己已有的经验,对被评估者进行评定。这一阶段经历了漫长的历史。

在世界教育史上,早在父系氏族社会时期,氏族或部落中接纳青年人为社会成员而进行的“青年礼”一检验他们的各方面本领和面临困难、危险时持何态度以及有无摆脱和克服困难的决心与信心,以确定他们能否成为氏族成员一就是一种原始的经验式的教育评估活动。

在中国,西周时期就有了“乡兴贤能”制度,定期对有关人员进行考查,以选拔人才。《周礼》中就有“三年则大比,考其德行道艺,而兴贤能者”的记载。到春秋战国时期,经验式的评估活动进入了学校。《礼记·学记》中说:“比年入学,中年考校;一年视离经辨志,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”。这是中国最早以考试、考查方式进行教育评估的记载。

西汉武帝时期,继承并发展了“乡兴贤能”的选士制度,建立了察举制。察举制设三科考试,分别选拔不同的人才。设“方正贤良科”为中央选拔有文学才能的人才;设“孝廉科”为地方选拔官员;设“童子科”以鼓励和选拔天才儿童。到三国魏文帝时,又以“九品中正制”取代了察举制。以考试为选士的手段逐渐代替了“乡举里选”。九品中正制设“孝廉”、“秀才”两科,“孝廉”考经,“秀才”试策。进一步发展了以考试为手段的早期教育评估活动。

具有完整体系的正式的教育评估活动始于中国古代的科举制。科举制是中国封建社会设科考试,用以选拔官吏的制

度。科举制开始于隋朝，延续至清末。科举制是九品中正制的发展，是以考试为主要手段的有关教育、学力检验、测定、评价的教育评估制度。它具有等级划分、考核标准、实施办法、选择目标等一整套体系。

科举制实行各地预选与全国统考结合。唐代规定，考生须经地方考试合格才有资格参加全国统考。清代，读书人未经考试者称童生，经县、府、院试合格者称生员（秀才），经乡试合格者称举人，经会试合格者称贡士，最后经殿试合格者称进士。

科举制也是一套政治条件与文化水平并重，兼及身体状况的，以综合测评为基础的人才选拔制度。唐宪宗时规定：

“迹涉流狂、兼亏礼教，或曾任州府小吏有一事不合清流者，虽薄有辞艺，不得申送”。元代亦规定，患废疾者，不得应试。

以科举制为背景的教育评估曾对中国封建社会统治阶级选“贤”举士发挥了重要作用。但是，到了明、清时期，科举制度日趋腐朽，以八股文取士所带来的考场舞弊现象日趋严重，加上西方新教育制度的冲击，到清朝末期，科举制走上了绝路，于1906年被完全废除。清政府又制定了新的教育制度，以科举制为背景的教育评估体系也随之取缔，代之以适应新教育制度的教育评估体系。但是，从评估的方法、手段来看，仍是以经验评定为特点的。

在西方，早期的教育评估也同样是与人才选拔、学力评定、教育评定紧密联系在一起的。其方法手段，也是以各种考试为主。在十八世纪以前的西方各国，由于学校尚未普及，学校考试主要是口试。1702年，英国的剑桥大学首先以笔试代替口试。此后，随学校普及，各国纷纷效仿，并相继

建立了考试制度，使以考试为主要方法手段的教育评估不断完善。但是，从考试内容看，十九世纪后期之前西方国家所采用的笔试，同我国历史上的考试一样，也都是一种编文体的考试。通过这种考试对教育进行评估，突出地反映了经验评定的特征。

考试的方法，特别是书面考试的方法进入学校教育，使学校全面考查学生的学力成为可能。但是，无论是中国的考试还是西方的考试，都很难比较全面、准确地测验出学生所学知识的真实情况和实际能力。至于说对学生的思想品德、行为表现、技能技巧、身体状况等进行全面、准确的质量评定就更难了。其主要原因是这些考试都缺乏评定成绩的客观标准。对这些考试成绩的评定，都是根据评估者的个人经验和主观上的好恶来进行的。基于对经验评定式的教育评估的这种认识，在十九世纪后期，西方一些国家的教育家们开始了教育的客观性评价的研究。从而把教育评估的实践与理论推向了一个新的阶段。

二、量化评价阶段

为解决考试成绩评定的客观标准问题，十九世纪后期，西方的教育家们掀起了一场教育测验运动。

1864年，英国格林威治学校的教师菲奢收集了许多学生考试的试卷与答案进行客观评分，并依据一定的价值标准把这些成绩汇编成标准评分的量表，试图为当时的考试提供一个可供参考的标准。但由于某种原因，他的研究没有引起人们的注意。

在教育界引起人们对测验问题极大关注的是莱斯的拼字测验。1897年，莱斯发表了他对二十所学校三千余名学生所

作的拼字测验研究的结果。测验表明，八年中，每天花45分钟进行拼字练习的效果同每天花15分钟进行练习的效果并没有多大差别。这一结论引起教育界的强烈反响和人们对教育测验的普遍关心。

这一时期的教育测验之所以能引起人们的普遍重视，除教育内部的原因外，德国人温特的实验心理学研究（1979）、英国人葛尔登的人类个别差异的研究和法国人比纳的儿童智力测验研究等，也都起到了重要的影响作用。特别是比纳在1895年设计的一套智力测验方法和比纳在得到西蒙协助下于1905年制成的智力测验量表，对测验运动的兴起具有重大意义。

1904年，美国心理学家桑代克发表了《精神与社会测验导论》一书，标志着测验运动的开始。

桑代克在《精神与社会测验导论》一书中宣称：“凡是存在的东西都有数量，凡有数量的东西都可以测量”。在这一信条的指导下，欧美教育界对教育效果的判定采用了科学量化的成绩考核，把追求客观性作为最高宗旨。此时期的教育评估活动，无论是在内容上，还是在方法上，都受到测验研究的影响。随着各种标准化测验被编制出来，客观的测量方法迅速地从对学生的测验发展为对学校 and 教师教学效率的测定。

教育测验运动时期的研究成果，同以前的编文体考试相比，是教育评估活动的一个极大进步。但是，由于过分强调数量化、客观化，而把所有难以数量化、客观地加以描述的事物都排除在测验对象之外。这样一来，测验所抓住的是教育上的一些零散、片断的现象，忽视了教育对象全面人格的发展。美国一些教育家深深感到教育测验的诸多不足，

于1930—1940年间兴起了对教育测验的批判运动。同时他们另辟歧径，开创了教育评估运动，于是，教育测验运动被教育评估运动取而代之。

三、科学评估阶段

教育评估作为一门学科，其理论形成于美国，它是从批判测验运动开始的。

1933年，在美国进步主义联盟会长艾钦领导下，成立了以泰勒为主的评估委员会。他们以新的教育理论为依据，接受心理学观点，以全面发展人的才能为目标，设计了一套新的课程，在七所大学和三十多所中学中进行了一场为时八年的课程改革与评估的研究，获得了良好效果。1940年，评估委员会提出了一个报告，详细地阐述了课程设置的原理和评估的原理，并根据评估原理论述了实验结果。这个研究在美国教育史上称之为“八年研究”。

“八年研究”的结果使研究者们得出这样的结论：当时的课程与测验是教科书中心主义。测验的内容只是记诵教材的知识内容。这是极为片面的，并不能反映学生个性的全面发展。为此，他们提出了一套以教育目标为核心和依据的课程与测验的理论及方法。试图以此把社会对学生的要求反映在课程与测验之中。为把这一思想与早期的教育测验区别开来，泰勒和他的同事正式提出了教育评估的概念。

泰勒的教育评估思想主要就是注重教育效果的价值观，认为教育评估就是衡量实际活动达到教育目标的程度，而测验仅仅是评估的手段之一。测验评定只有与其他评定手段结合使用，才能正确而全面地评估教育目标是否实现。否则，测验或考试使用不当，则会使教育走上不正确的道路。

泰勒的教育评估理论，既强调学生的学习效果、智能、身体状况、思想品德、志趣和理想的评估，也强调学校、教师、课程、教材的评估。从而打破了过去只重考试分数的简单做法，开始把教育评估推到了一个比较科学的阶段，使教育评估理论日臻完备。

随着教育规模的扩大、教育普及程度的提高和教育理论与实践的不断发展，教育评估的应用和研究也日益广泛、深入。第二次世界大战后的许多国家都不同程度地借助于教育评估来引导教育实现理想的目标。联合国科教文组织把是否具备相当程度的教育评估技术水平作为衡量一个国家教育水平的尺度。至使各国教育家对教育评估的研究给予了高度重视，推动了教育评估理论与实践的发展。到了二十世纪六十年代，美国的教育家们从美国和苏联科技发展的对比上，发现了本国教育的落后状态，开始对泰勒的教育评估理论进行重新审查。提出：如果教育评估以教育目标为中心和依据，那么，教育目标的合理性又依据什么来判断；教育活动除达到预期目标外，还会产生各种非预期的效应与效果，这些效应与效果要不要评估等问题。从而使教育评估的研究进入了更广阔的领域和更深的层次。随之，教育评估流派大量涌现，各种教育评估模式相继问世，各种研究教育评估的杂志和专著大量出版，各种教育评估的研究机构纷纷创立，形成了一场教育评估运动，使教育评估活动成为一个专业领域出现在教育科学的理论和教育实践中。

第二节 教育评估发展的趋势

教育评估理论是随着教育实践的发展和社会对教育评估

的需要与变化而不断趋于系统化、完备化、科学化的。现代教育评估与其他教育科学学科相比，还不太成熟。但随着现代社会发展对教育提出的更高要求和教育理论与实践深入发展的推动，它已成为教育科学研究的重要领域，受到广泛的重视。目前，教育评估不仅仅是专家学者的研究课题，而且是社会各界，特别是国家政府所关注的学科。许多国家的政府把它纳入了教育行政管理的议事日程。现代教育评估作为一门独立的学科，已在广泛的教育实践活动中得以普遍的应用和研究。统览国内外教育评估的发展，具有以下趋势：

一、教育评估的目的由选拔适合教育的学生转向选择适合学生的教育

早期的教育评估，一般来说是为选拔人才服务的。在学校中，无论是对哪一级学生进行的学力评估，都是在证明他们是否取得了进入下一阶段学习的资格，或证明他们有无结束前一阶段学习的资格。实质上是为了选拔适合教育的学生而进行的资格鉴定。对教育机构、教育团体、教育方案等方面进行的评估，也只是为了对这些方面的等级区分进行证明。因此，早期的教育评估单纯重视终结性评估，强调对活动结果的判定。现代教育评估理论承认终结性评估在教育管理中的不可替代的作用。但是，它认为，终结性评估对教育过程中有害于教育目标实现的行为和对最终实现的结果与教育目标相悖的现象无能为力；对已经结束了的教育过程，它只能是“事后诸葛亮”。鉴于这些问题的存在，现代教育评估强调对教育全过程的评估。既重视终结性评估，也重视目标本身的评估和形成性评估。目标的评估可以使人们把握教育活动的可行性和方向性；形成性评估则可以在教育过程中