

怎样评价教育质量

ZENYANG
PING
JIAJIAOYU
ZHILIAN

湖北教育科学研究所

编 者 的 話

教育质量评价是当今世界上教育科研的一个重大课题，也是我国教育事业发展中亟待解决的问题。正确地评价教育质量，不仅是加强教育管理的重要内容，也是检验教育改革的科学依据。随着世界新技术革命的兴起和我国经济建设的发展，社会对人才质量又提出了新的要求，因此，开展教育质量评价的研究工作更具有十分重大的理论意义和现实意义。

长期以来，由于缺乏科学的教育质量观和准确的评价标准，在我们的教育工作中出现了不少的这样或那样的偏差和失误，一些学校和教育行政部门往往把升学率的高低，有的甚至把主观印象或感想作为评定教师或学校工作的依据，严重影响了教育事业的健康发展。为了解决这个问题，近年来不少同志围绕教育质量评价问题，开展了一些研究和实验活动，取得了一定的进展。一、逐步树立了比较正确的质量观，弄清了一些基本概念，有的同志还提出了比较系统的教育质量评价原则、评价方法和评价程序，并试图着手建立教育质量评价的体系。二、教育质量评价的研究领域已涉及到教育过程的各个方面、各个环节。对怎样评价学生、怎样评价教师、怎样评价学校领导等，都进行了探讨。在教学、教育领域，对德、智、体育各方面也都提出了评价原则和方法。而且还从对教育过程的评价扩大到对教育行政管理工作

的评价。三、开展了关于教育质量评价量化问题的研究。引进了教育统计学、系统工程等方法，研究如何将质量问题数量化，以利于比较精确地衡量教育质量。

由于教育质量评价是一项十分复杂的工作，这一研究在我国才刚刚起步，还有一系列问题需要进一步探讨。

第一，如何根据“三个面向”的要求，树立全面的质量观。衡量一个受教育者，怎样才是合格的，总的原则是什么？这些问题还研究得不够。

第二，要建立什么样的质量指标体系才能全面地科学地衡量教育质量。升学率不能做为衡量教育质量的综合指标，那么，综合性指标应该是什么？如何建立？这个问题没有解决。

第三，怎样根据各级各类学校培养目标和办学条件的差别，建立各自不同的评价标准。重点学校和非重点学校、普通教育与职业教育、基础教育与专业教育，评价标准不应是相同的，这是如何从实际出发评价教育质量的一个重大问题，但至今很少涉及。

第四，教育质量评价是不是都可以量化？知识掌握程度可以量化，能力强弱程度能不能量化？智育水平可以量化，德育水平能不能量化？不能量化的，怎么办？

我们编印这一本小册子的目的，一方面是为了让广大教育工作者了解近几年来我国对这一问题研究的进展情况，另一方面也是为了推动这一研究工作的深入开展。我们希望广大教育工作者能在现有的基础上，结合各自的实践经验，对这一事关教育大局的重大课题进行广泛深入的研究和讨论，为建立我国科学的教育质量评价制度作出贡献。

编者 一九八五年三月

目 录

(一)

- 教育评价的基本概念.....张天麟 (1)
- 关于教育评价.....肖垠 (10)
- 教学结构量质论.....孙治权 曹明 (29)
- 全面衡量教育质量与人才.....秦阳 赵华 (48)
- 素质结构的综合化.....秦阳 赵华 (48)
- 简论学校评价.....王恩民 (53)
- 浅谈学校评价的几个问题.....唐文 (77)
- 开展我国高等教育评价工作...:.....沈阳师院教科所 (84)

(二)

- 教学质量是教学管理的
一项基本建设.....康尔珪 赵立志 (91)
- 试拟教师教学工作质评审量表.....黄云龙 (106)

(三)

- 学校教学质量可以
计算出来.....郝新生 费贵麟 (117)
- 教学效果计量评价方法初探.....戴梅竞 (131)
- 关于《教学效果计量评价方法》
的几个问题.....戴梅竞 (138)
- 关于教学质量评价问题的初步探索.....吴锦骝 (114)

合理评定多科考试成绩问题的探讨·····	宋岭梅 (156)
班级之间考试成绩的评价方法·····	张敏强 (161)

(四)

如何评价课堂教学的质量与效果·····	章蓓春 (166)
试论课堂教学的常规标准·····	孙小林 (170)
武汉市小学教学工作质量考核标准 及办法 (讨论稿) ·····	武汉市第一教育局 (177)
谈用《教学评议卡》 评价课堂教学效果·····	梁桢 齐世凯 (183)
对学生品德考查量化问题的探讨·····	林昌华 (198)
用考核学生学习效果的方法 来检查和评定教师教学质量的设想·····	程洪健 (211)
怎样评价教师的教学效果·····	蒙万融 李楼瑞 (218)
到达目标与到达度评价·····	万勇 (224)
怎样评价教师的工作 (节 译) ·····	(苏) 波尔特诺夫 余强基译 (259)

教育评价的基本概念

张 天 麟

教育评价是当前世界上教育科研的三大领域之一。随着科学技术和教育事业的发展，教育科学出现了多种分支，如：教育哲学、教育社会学、教育经济学、教育心理学、教育管理学、教育统计学、教育未来学、教育工艺学、比较教育学等。但综合起来，在教育科学研究中，大体上可分为三大领域：即基础理论研究；教育发展研究；教育评价研究。

教育评价是本世纪六十年代新兴的教育科研领域，开端于美国，但很快受到各国政府和学术界的重视。最近，日本首相中曾根康弘提出“教育改革七条设想”1983年12月，第二条是：纠正按偏差值进行对初中毕业生的升学指导办法……彻底改变轻易以测验成绩的好坏进行不合理的升学指导”。实际上就涉及到教育评价问题。

一、评价与测定

“评价”这个词不仅用于教育，而且广泛地用于各个方面。评价不同于测定。测定是事实的描述；评价则是给予价值上的判断。评价必然建立在测定的基础上，但测定不能代替评价。例如，丈量土地的面积，这是测定，不是评价；结合地形、土质、地理位置等才能给予价值上的判断，这才是

评价。明确评价与测定这两个概念的联系和区别，对研究教育评价有着重要意义。

教育评价当然是以教育为评价对象，对其效用给予价值上的判断，目的是为教育决策提供科学依据。教育评价既不同于传统的考试，也不同于科学的测试。考试和测试仅能测定事实情况而不能给出价值上的判断。

二、传统的考试

考试主要指采用口头或书面回答问题的方法来检查受试者的知识和能力。考试是一种古老的制度，几乎从人类有历史记录以来，就有考试制度。统治阶级常常用考试的办法为统治阶级选拔统治人才。

使用考试做为选拔人才的方法，我国是最早的国家之一。早在西周时代就采用以射选士。《礼记·射义》说：“古者，天子以射选诸侯、卿、大夫、士。射者，男子之事也，因而锦之以礼乐也”，又说：“天子之制，诸侯岁献贡士于天子，天子试之于射宫”。从西周经秦、汉、魏、晋、南北朝至隋、唐，基本上以选为主，以考为辅。隋文帝开皇七年（581年）“制诸州岁贡三人，工商不得入任”（马端临文献通考）。隋炀帝大业二年（606年）始置进士科，试策问，这是以考为主的科举制度的开端。这在当时是一大进步。康有为也说：“任官先试，莫我先焉”。一直到清光绪三十一年（1905年），“谕立停科举，以广学校”。科举考试制度在我国持续了一千二百余年。

西方国家从十九世纪以来，陆续建立了文官考试制度，

英国于一八五五年建立“文职人员事务委员会”，采取公开竞争性考试招募文职人员；美国于一八八三年成立“文职人员事务委员会”，规定必须经过考试才能担任公职（西方国家一般分政务官与事务官，政务官由选举产生，事务官通过考试选拔）。还有些国家，将考试制度明确规定在宪法中，如现行意大利宪法第三十三条第五款写道：“为了升入各种和各级学校或在各该学校毕业以及获得就业资格，均须经过国家考试。”第九十九条第二款写道：“除法律有规定的情况外，国家行政机关之官吏均由考试选拔录用”。

从选拔人才看，考试制度有一定作用，这是无可否认的。关键是考试的内容和考试的方法。传统的考试制度是不科学的，不公正的，主观随意性很大，弊端很多。把这种考试制度用于学校教育，就更加不公正、不合理、不符合教育规律。

三、科学的测试

把科学方法引入到测试中来还是近代的事情，约有一百来年的历史。早期的先驱者是英国的科学家高尔顿。高尔顿的研究领域很广，在这方面的主要贡献是把统计学的理论应用于测试。他发现大学考试的分数表现为常态分布曲线，即大多数有向平均数或常态曲线中心集中的趋势，离开中心而趋向两端的分数愈来愈少。由此，他认为人的能力也是常态分布，正确的测试分数也必然表现为常态分数曲线。他还提出了统计学上重要的量数之一——相关。而决定测验的正确性、可靠性和因素分析法等是这一发现的直接后果。他的学

生皮尔逊为了正确计算相关系数，提出了一个目前还在使用的数学公式。其后，美国心理学家卡特尔收集了大量数据，进行了一系列研究，更加强调数量化、等级法、评定法等，并且创造了“心理测验”这一术语。

科学测试运动首先是从智力测验开始的。一九〇五年法国心理学家比奈、昂利、西蒙共同发展了一种智力测验，用较复杂的方法测量高级心理能力。智力测验迅速地发展起来，目前国外的智力测验有许多不同的种类，为个人测验（一次试验一人）、团体测验（同时测试多人）、速度测验（以在限定时间内做出正确反应的数目决定分数）、才能测验（以成功地完成的作业难度决定分数、语言测验要求以语言回答问题）、作业测验（包括那些非语言反应、如把不同形状的小球按照特定顺序穿成一串）。

智力测验进一步应用于考查学生学习成绩的方法称之为成就测验。一九〇八年美国教育家斯通恩第一个提出客观化测试算术的方法，全部十二问，从易到难排列，时间十五分钟，对象为小学六年级学生。其后，逐渐发展为美国标准化测验。一九二三年出版的斯坦福成绩测验，是可以综合使用于全部课程的第一个主要的标准化成绩测验。这一测验简单易行，分数评定客观。它从有代表性的一千五百名儿童得出了各种年龄儿童的国家标准，这样就便于任何个人同平均数进行比较。这些标准以后又做了多次修改，使之更确切地结合各种地区的差异性。

除了以上的智力测验和成就测验外，第三类的主要测验是个性鉴定测验，这类测验使用多种不同设备，用以测量个

人的兴趣或其他方面的性格。

四、教育评价的基本概念

本世纪六十年代新兴起的教育评价是科学测试运动的进一步发展，但它与测试有本质上不同，两者既有区别又有联系。测试不同于评价之处，首先在于测试仅仅是一种手段，为了获得教育过程的各种信息，除了测试外，还可采取多种多样的办法，如观察、累加记录、问卷法等。其二，测试仅仅是查明事实，如何看待这种事实，则有待于对这种事实的价值上的判断。总之，有测试不一定有评价，而为了进行评价又常常采用测试的办法。

什么是教育评价，目前还很难规定一个确切的、统一的、科学的定义。笼统地说，所谓教育评价就是以教育为对象，对其效用给予价值上的判断。从评价的对象上看，又可区分为广义的教育评价和狭义的教育评价。

广义的教育评价以教育的全部领域为对象，它涉及教育的一切方面：教育与社会；教育与经济；教育与政治等。例如教育体系、教育目标、课程设置、教育内容、教育方法、教育管理、教师活动、师生关系等，这些都属于广义的教育评价的范畴。当前，我国正在进行教育改革，改革必然要有各种各样的实验方案，它们的效果如何，如何与控制组相比较，如何进行统计处理，如何运用其他学科的研究成果评价各种实验方案，以便为教育决策提供科学依据，都需要科学的评价技术。

狭义的教育评价以学生为对象，专指在学生的学习和生

活领域中对学生的成长给予价值上的判断。通常所说的教育评价常指狭义的教育评价。如何科学地、公正地、全面地评价学生的成长？这是世界性的问题，也是我国教育改革中急待解决的问题。因为这将有助于全面提高教育质量，克服单纯片面追求升学率的现象。

五、国外教育评价发展的背景

教育评价所以在本世纪六十年代开始兴起，并迅速得到各国政府和学术团体的重视，成为教育科研的三大领域之一，有其历史背景：

（一）一九五七年苏联卫星上天，美国朝野震动，一致认为美国科技所以落后于苏联，原因之一是美国的教育落后。一九五八年美国联邦议会通过了国防教育法，拨出巨款加强理科和外语教学。既然大量投资于教育，就产生了投资效益问题。因此，以测定教育投资的效益为背景，教育评价受到了普遍的重视。

（二）教育实验和教育改革需要教育评价的发展。当前，世界各国正面临着新技术革命的挑战。科学技术的进步对教育产生三个后果：

1、学习过程趋向于代替教学过程；

这包括着强调学习的连续性，在学习过程中需要与动机的重要性，学习内容的选择，学习的层次，教育内容与环境的关系等。

2、把技术吸引到教育体系中去；

它的涵义并不是说把新的技术强加于传统的教育体系，

在传统的教育体系中单纯地增添电子计算机之类的东西。它意味着将促使人们重新考虑现存的教育体系，从根本上改革传统的教育体系。

3、利用人民的力量：

新的技术革命要求从质量、数量上提高全体公民的教育水平。而传统教育体系却存在着巨大的浪费，从一九六〇年到一九六八年这段时期内，全世界入学人数每年上升4.5%，而每年教育费用的相对增加则为11.7%。教育经费上的增加同教育活动实际上所应取得的成果之间存在着一种尖锐的矛盾。据世界银行的资料看：如果按现有比例并在相同的结构和管理条件下继续增涨，发展中国家的国民生产总值最低每年必须增涨7%，才能保持一九七五年水平的国民生产总值中的教育分配，而根据一九六〇年至一九七五年间的趋势判断，这似乎是极不可能的。总之，尖锐的教育资源不足和传统教育体系的巨大浪费又造成世界性的教育危机。所以，从本世纪七十年代以后，又兴起一门新的学科，这就是关于非正规教育的研究，它强调依靠人民的力量对现存教育体系进行根本性改革。

总之，新的技术革命要求教育必须改革。既然要搞教育实验和教育改革，就有一个如何判断一项新措施的 功效问题，这就需要一套科学的评价技术。为教育决策提供科学依据。

(三)、教育基础理论的发展推动教育评价的改进：

当前世界上较有影响的教育理论是美国布鲁纳的结构课程论和苏联赞可夫的教学与发展论。两者都涉及到只有正确

的评价，才能推动人们不断前进。正确评价儿童对提高教育质量有着重要的作用。

(四) 人民群众要求教育民主化的斗争促进教育评价的改革

在日本流行着一句谚语：“留给子孙的仅仅是学历和学力”，“学力已经成为一种财富”。教育社会学的研究已经指出：传统的教育制度对社会结构起到一种分层作用。人民群众日益对学历社会产生不满，资本主义社会正在为学历社会所困扰。为了解决这一问题，国外学者提出了各种方案，如终身教育等。其次，就是改革现存的对学生的评价制度，使教育评价更公正、更科学、更客观、更合理。

六、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、个案调查等方法，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，做为决策的科学依据。

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（IEA）的活动概况，因篇幅所限，不再赘述的。

——天津市教科所《教育科研资料》总第9期

关于教育评价

山西省教育科学研究所 肖 根

在教育行政管理中，不论是派员下去调查和视导，或者是听取汇报和经验介绍，看所属单位上报的各种材料（总结、计划、专题经验等等），或者是检阅学校的成绩展览，时时处处都有一个对教育和教学工作的评价问题。学校和教师要求对他们的工作进行评价，教育行政部门也应该对他们的工作做出正确评价。教育评价工作做得好，还是做得不好，在很大程度上都影响着教育实际朝什么方向发展，必须给予高度重视。

教育评价是一门专门的学问。对于它的理论和方法，这里不可能进行全面论述，教育行政管理中需要特别注意的地方，有以下一些问题。

外国关于教育评价问题的提出

教育评价一词，在欧美一些国家的教育实践中，是和成绩考查方法紧密联系在一起的。

对教育效果的判定，在十九世纪后半期，所采用的笔试，是一种论文体的考试，还不能比较全面地测验出学习的所学知识。对这一考试办法的改进，美国在二十世纪初盛极一时的教育测验，曾经起了很大的作用。他们在那个时期搞

的教育测验研究及其实践，受到了几个方面的影响：一个是德国人温特的实验心理学研究（1879年）和英国人葛尔登的人类个别差异研究（1882年）的影响；一是比纳1895年设计的一套智力测验方法的影响，还有比纳在得到西蒙的协助下于1905年最初制成的智力测验量表的影响。美国在这个时期搞的教育测验，固然在知识考试的标准化方面比之过去是一个进步，但要测出教育的全面效果，把一个受教育者的社会态度、思想品质、实际技能、兴趣和创造能力、分析和判断能力等，都变成数字予以量化，也很没有把握。二十世纪四十年代，美国的一些教育家深感教育测验对现实的许多不适应，又出现了对于教育测验的批判。

1918年，美国许多受杜威思想影响的教育家们在芝加哥开会，组织了进步主义教育联盟（*P·E·A*组织），提出教育的目的在于生活，在于儿童本位，对于那种仍然偏重于记忆的考试和测验，表示了不能容忍。1933年，*P·E·A*组织的会长艾钦和他的协同研究者，就以对教育测验进行批判的姿态，开始了一个所谓的“八年研究（1933—1940年）”，他们认为，教育测验有很大的片面性，不能反映受教育者各方面的实际学习效果；还认为教师以教科书为中心所采取的考试方法和测验方式，使受教育者易于养成被动学习的态度，不符合学生的发展和社会的需要。他们在对课程内容进行改革实验时，特别重视提出一套有别于教育测验的成绩考查方法。以泰勒为主的评价委员会，在“八年研究”中专门从事了这一工作。

泰勒的教育评价思想认为，教育不能只是知识的灌输，

要以全面发展受教育者的才能为目标，用一套新的教育评价方法，对教育效果进行全面判定。在进行教育评价时，并不排除测验方法的一定使用，但把评定的方法，还有其他能够用以考查教育效果的方法，都综合地利用起来，以便全面评价教育目标的是否实现，达到了什么程度，原订目标有无进一步修订的必要。泰勒等的研究结果认为：教育考查的方法对于实现一定的教育目标，有着极大的影响：如果使用不恰当的成绩考查方法（如激烈的升学竞争考试），往往会促使教育走上不正确的方向；而单纯地注重知识教育，可能就是一种严重后果。他们的研究，不仅把教育评价同教育目标紧密地联系起来，而且设计了一种叫做“问题情境测验”的方法，以考查学生的创造性思考，逻辑思维分析和解决实际问题的能力，并注重了全面地、一贯地针对德智体诸方面的要求进行教育评价，注重了对学生成绩的分析 and 针对性地指导，注重了学生的自我评价和互相评价。这样的教育评价，既涉及了对于学生的学习效果、智能、身体健康状况、思想道德水平、志趣和理想的评价，也涉及了对学校、教师、课程、教材的评价。这样的评价，打破了过去考试只注重于分数、计算均分、排名次的简单做法，而注重于学习成绩的全面考查和分析，无疑是一个进步。美国的“八年研究，把这样一套教育成绩考查的指导思想和方法要求，命名为教育评价，后来得到了很多国家的赞可，现已为很多国家所采用。联合国教科文组织的委员们，也都以各国教育人员（包括教育行政人员）有无相当的教育评价能力，作为评定一个国家教育效能的一种要求。教育评价方法已经引起很多国家重