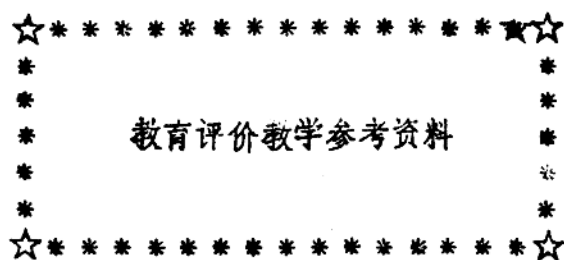


教育评价教学参考资料

刘本固 编

东北师范大学教育系

一九八六年十月



刘本固 编

东北师范大学教育系

一九八五年十月一日

说 明

一、编选《教育评价教学参考资料》的目的，在于为教育管理班的学员提供一些自学材料。

二、由于资料来源所限，还有许多有科学价值的论著内容尚未收进来。

三、为便于查阅，基本上采取节选归类的方法。不过，对内容新颖的同类材料也一并收入，且另立条目，以助思路的开拓。

四、在编选过程中，参考了有关资料，特向作者致谢。

五、因时间仓促，水平所限，难免有错误和缺点，恳请批评指正。

编 者： 李 芳

1985年10月1日

目 录

第一章 教育评价的性质与作用·····	1.1
一、教育评价的概念·····	1.1
二、教育评价的作用·····	1.10
第二章 教育评价的产生与发展·····	2.1
一、教育评价的产生、发展及现状·····	2.1
二、国外教育评价发展的背景·····	2.11
第三章 教育评价的类型与功能·····	3.1
一、依据目的划分的评定分类·····	3.1
二、自我评价·····	3.4
三、到达目标与到达度评价·····	3.15
第四章 教育评价的过程原则与方法·····	4.1
一、教育评价的过程·····	4.1
二、教育评价的原则和步骤·····	4.11
三、普通教育评价的原则·····	4.13
四、评价高等教育的渠道与方法·····	4.19
第五章 教育评价标准·····	5.1
一、评价高等教育的标准·····	5.1
二、研究学生的基本大纲和评定标准·····	5.3
三、研究教师的基本大纲和评定标准·····	5.11
四、课堂教学的评价标准·····	5.29

第六章 学力评价与道法评价	6.1
一、学力的评定	6.1
二、测验的种类和功用	6.5
三、如何衡量考试效果	6.11
四、国外中小学成绩评定的研究	6.13
五、学生的道法评价能力	6.21
第七章 教师工作评价	7.1
一、怎样评价教师的工作 (波尔特诺夫)	7.1
二、教师对教学和教育活动进行自我分析	7.16
三、教师教学工作质量的评审量表	黄云龙 7.21
第八章 学校评价研究	8.1
A. 简论学校评价	王恩民 8.1
B. 日本学校管理的评价	刘本固 8.24
C. 怎样评价一所学校	凯斯·斯密斯 8.41

第一章 教育评价的性质与作用

一、教育评价的概念

(一) 教育评价的诸种定义

1. 教育与评价

△ 导 言

一般认为教育是一个改变人类行为的过程。在这个过程中，人们通过学习将逐步改变自己的思想、感觉和行为。评价则是指我们试图判断一个预定的教育意图在教育过程中效果如何？

通常，教育中的评价可分为两大类：

① 方案评价。这是对各种类型的教育方案有效性的评价。如果它测量的是一门课程的有效性，通常就称课程评价。

② 学生评价。这是对单个学生成绩的评价。

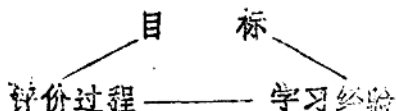
教师在教育过程中有三项基本任务：

① 制定适合于学生的教育目标。

② 创设一定的情景，为学生提供完成目标的课程和教学设计。

③ 判断这些目标实际完成的程度。

这三项任务是密切相关且相互作用的，不应孤立地看待。它们能形象地描述成一个互相关联互相影响的三角形：



其中最重要的、起决定作用的是教育目标的选择，然后是为完成目标而制定的课程和教学方法（提供学习经验），最后才是检查目标实际上是否达到的评价过程。

每一个从事教育评价的人都应该明白：确定教育目标是位于达到和测量这些目标的方法之上的、第一位的。对教师来说，在他的工作中首要的是用明确的语言讲清楚他们的教学目标。这是保证他们的教育方法实际有效的前提。也是评价过程中可检查的保证。当然，评价过程中所测量的应是教师认为在教学内容中重要的东西。

△ 评价的定义

通常认为，评价意指判断一些事情的价值。而这种价值的判断又是建立在某些准则或价值标准之上的。在作出评价之前需要测量，评价则是一个高于测量的阶段，它包括判断被测的事情是否优质，是否有力，是否有效以及价值如何等等，因此，还必须制定作为判断基础的准则。

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。

2. 什么是教育评价

什么是教育评价？教育评价是一种新的教育成绩的考查方法。有人认为“教育评价是利用所有可行的评价技术评量教育所预期的一切效果。”（《教育评价的理论和方法》第一章）就这一概念来分析，我们可以看到，教育评价与历来的考查办法，在背景、范围和对象等方面均有明显不同。

从思想背景来看，传统教育重视记忆性考试，主要考查注入了多少内容，学生记忆了多少？后来由于科学思想特别是心理测验的发展，开始重视教育考查的客观性，就有了标准测验和客观测验。及至教育发展到重视知识、技能、态度、兴趣、情操和人格发展等全面评价的时代，根据新的价值观，就采用了教育评价法。综观以上所述考试、测验、检查和评价，都是教育成绩的考查方法，但是由于其思想背景不同，其含义也有所不同。

考试只能考查记忆、理解多少，却不能考查态度、兴趣、品德、性格、情操等方面的效果。根据现代教育理论的观点看来，考试的范围已经过于狭窄。测验只着眼于客观数字的测定，其思想与价值观并无联系。教育评价则以教育的价值观点为标准，来评量教育效果。即以达到教育目标的程度来计量教育成绩，它要求对教育和学习作全面性考查。考查的范围包括：学习方面的知识、理解、技能和认识能力；情意方面的态度、兴趣、习惯、鉴赏；个性方面的性格、道德品质等。

教育评价并不是取代其他考查方法，而是根据评价对象的需要，选择适当的方法。如考查学习效果就需用考试法，考查兴趣、态度等就需用观察法、测验法、调查法等等。

这就是说教育评价要运用一切可行的评价技术来评量教育所预期的一切效果。

张秀媛：《关于教育评价问题的概述》，摘自北京《教育研究资料》1985年第1期和第二期

3 评定与测量

① 评定的意义

所谓教育，就是实现价值和理念，即实现教育的理念与目标的有意识的活动。在有效地履行这种教育活动方面，要明确规定教育目标，引起旨在实现该目标的行为，对作为其结果而产生的进步性变化作出估价，调整教育活动，这一连串的有组织的工作就是必要的条件。这种有组织的工作，即把“教育所产生的（或是将要产生的）人的行为和人格的变化（或是变化的可能性），对照一定的价值标准进行判定的过程（冈部弥太郎主编《教育评定事典》）”就是评定（evaluation）。因此，教育中的评定，可以说就是对照教育目标对由于教育行为而产生的变化进行的价值判断。

② 测量与评定

前节已述，测量本是物理学术语，用一定的量表来衡量对象，从数量上表现出来，同测量对象所具有的价值无关。因此在测量中，如何使对象作出正确的数量上的表现是问题的核心。另一方面，评定含有价值的意味，它是对某某对象和现象的价值进行估量。亦即，教育上的测量是客观地、正确地将学生达到教育目标的程度加以数量化。评定则是以测量结果有多大的教育价值为其课题的。教育活动中所进行的种种测验可以说不是测量的手段，而是用来进行评定的手段。

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) 教育评价的基本概念

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”1988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。

1. 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。明确评价测量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。

2 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，动广学校”。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。

3. 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次数愈来愈少。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布曲线。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向平均数进行比较。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。

4 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。

张王麟：《教育评价的基本概念》，

摘自天津《教育科研资料》1984年第9期

二、教育评价的作用

(一) 教育评估工作是提高教育质量和学校管理水平的重要环节

本报讯 国家教育委员会委员黄辛白在最近召开的高等工程教育评估专题讨论会上说,为实现中央提出的高教发展的战略目标,国家在扩大高校办学自主权的同时,必需做好宏观的指导和管理,要对人才培养的数量和质量是否与社会和经济发展相适应,进行预测、评估和毕业生工作情况的信息反馈,搞好教育评估是提高教育质量和学校管理水平的重要环节,对促进高教事业的发展,必然起到重大作用。

在谈到评估工作的目的和方法时,黄辛白说,通过评估工作,应该使各级各类学校都能进一步提高积极性、创造性,肯定成绩,找出差距,不断提高工作水平,以保证教育质量和教育事业的发展;通过评估,应该注意促进重点学科、专业的建设。现在,要加强对这项工作的宣传,提高高校干部和师生对评估工作的认识,吸引他们参加这项工作;要使各级各类学校都积极支持并主动地、逐步地开展这项工作,而不是把它看作被动的受检查;要进行学校、学科、和单项教育评估的研究和试点。教育评估是一项既有学术性,又有政策性的复杂的工作。黄辛白希望参加筹备这项工作的同志解放思想,开拓视野,实事求是开展这项工作,切忌一哄而起,搞形式主义;要认真研究国内外教育评估的经验,经过试验,逐步完善,以建立一整套有中国特色的教育评估制度。

这次评估工作会议是最近在黑龙江镜泊湖举行的。来自全国

三十八所高校及七个部委、五个省(市)高教厅(局)的同志参加了会议。大家认为,建立高等工程教育评估制度是教育改革在新形势下的必然要求,也是落实中央关于教育体制改革决定而办的一件实事。会议初步探讨了我国建立高等工程教育评估制度的目的、作用、理论和方法,初步交流、论证和汇总了分层次评估高等工业学校办学水平的指标体系和实施方法,并就近期如何开展高等工程教育评估的研究和试点工作广泛交流了意见。

(本报记者 刘仁镜)

《中国教育报》1985年8月3日

(二) 高等教育评价的作用

在高等学校中,要对高等学校的工作作出客观的评价,以便在各高等学校之间,在一所院校的各系之间,通过评价,使教育行政部门全面掌握各院校的实际情况,同时也使学校看到自己的成绩和不足,积极地改进工作,但教育评价是一件很复杂的工作,要作好这项工作,首先要明确总的指导思想。

高等教育是以我国社会主义的政治经济为基础,又要为我国社会主义政治经济服务的。我国现在已进入了全面开创社会主义现代化建设的新时期。1983年全国高等教育会议上提出,新时期高等教育工作总的指导思想是:贯彻党的十二大精神,适应以经济建设为中心的社会主义现代化建设的需要。教育行政部门和每一所高等院校,都应自觉地用这个根本要求来衡量工作,并做出正确的评价。然而,多年来,由于缺乏科学的教育评价,因而对高等学校的全面工作,教学、科研水平等都没有一个客观的

衡量尺度。这种情况难以为学校和领导机关提出真正可靠的改进工作的依据。实践证明，一次客观的教育评价，有多方面的意义，它不仅为高等教育改革提供基本依据，而且也是高等教育管理的基本环节，是保证高等教育事业与政治经济发展相适应的重要手段。因此，认真进行高等教育评价，是推动改革，提高高等教育质量的需要，是不断提高高等教育管理水平的需要，是高等教育适应新时期总任务的需要。

沈阳师范学院教育科学研究所：

《认真开展关于高等教育评价的研究》，摘自《教育科研信息》1985年第一期

(三) 教育评价的意义和作用

教育评价是对教育进行有效管理的一项重要工作。它早已广泛地存在于各种实际的教育活动之中。教育行政部门对自己管辖的学校常常划分为哪些是办得好的或比较好的，哪些办得一般或比较差的。这是上级教育部门对下属教育单位的评价；学校间的互相参观学习，取他人之长，补自己之短，这是校际之间的相互评价；在一所学校里，开展各项教育活动之前，需要研究、讨论计划是否妥当，这种研究讨论是可行性评价；在执行计划过程中，一方面经常检查计划落实的情况，另一方面又要及时发现原计划不足之处，积极进行修改、补救。这种活动实际上是实施性评价；在一项教育活动完成之后，总结经验教训，肯定成绩，找出问题，明确今后前进方向，这是总结性评价。教师选择教学方法、检查效果、评定学生成绩，无不包含着评价工作。所以，教育评价工