

高等学校文科教学参考书

【美】B·S·布卢姆等著

教育评价

邱 渊 王 钢 夏孝川 译
洪邦裕 龚伟民 李亚玲

华东师范大学出版社

高等学校文科教学参考书

教 育 评 价

[美]B.S.布卢姆等编

邱 渊 王 钢 夏孝川 译
洪邦裕 龚伟民 李亚玲

华东师范大学出版社

高等学校文科教学参考书

教 育 评 价

[美]B.S.布卢姆等编

邱 渊 王 钢 夏孝川 译
洪邦裕 龚伟民 李亚玲

华东师范大学出版社出版

(上海中山北路3603号)

新华书店上海发行所发行 华东师大印刷厂印刷

开本: 850×1168 1/32 印张: 17.25 插表2 字数450千字

1987年3月第一版 1987年3月第一次印刷

印数: 1—18,500本

统一书号: 7135·199

定价: 3.00元

译者说明

本书是一部由国家教委推荐的高等学校文科教学参考书。

这部名作叙述的是关于学生学习的评价方法，旨在帮助教师运用评价手段，指导学生更好地掌握教学目标的要求。它在1981年出版时，原名《为改善学习而评价》。现在我们把译稿定名《教育评价》，用意在于突出它的内容主要为介绍评价方法。

教育评价是一项在我国有待逐步系统开展的重要工作。其中基础的一环正是对于学程或教学方案的活动效果的评价。翻译出版《教育评价》这部系统介绍教学活动的现代评价方法的著作，对于我国高等师范院校增设的教育评价课程，以及在职教师的教学评价实践，均可提供一定的帮助。

参加本书译校工作的人员有：邱渊，负责全书译法设计、统校和定稿，并译序、附录等部分；王钢，第一至三章；夏孝川，第四、五章；洪邦裕，第六、七章；龚伟民，第八、九、十一章，并协助做组织工作；李亚玲，第十章。感谢邵瑞珍教授拨冗为附录（《教育目标分类学的缩本》）作了校阅。此外，协助工作的尚有陈国良、冷志强、操宏高、杨晓青等同志，恕不一一详列，均此致谢。

1986年6月

序

这是一本介绍对学生学习进行评价的当今“工艺水平”的书。它首先是为目前和将来上课的教师写的。在运用得当时，评价应该使得教师能够促成其学生的学习显著改进。改善学生的学习，这就是本书所考虑的中心问题。

本书并非讨论有关智力、能力倾向和标准成绩等各种测试。一名中常型(可以代表一般情况的)在校生，很可能每年只参加一两次这类的标准测试。本书不如说是用来讨论如何改善和正确使用这类学生每年所要参加的教师编制的五六十个小测验、进度测试以及其他考查等等。

负责大班的教师，学生多种多样，工作忙忙碌碌。他们对于教学过程要操很多心，常常难得把时间和注意力放到评价过程上面来。而且，他们很可能赶不上日益增多的有关评价技术与科学的文献。本书汇集了最佳的评价技术，其用意就在于帮助教师通过评价去改善教学过程和学习过程。

过去十年中的教学工作有一项非常重要的发展，这就是掌握性学习法得到了广泛运用。掌握性学习使得大多数学生在各级学校课程的学习中达到高水平。掌握性学习方法的功效就在于改善学生的学习，提高他们对学科的兴趣，并使他们增强对学习的信心。大多数使用掌握性学习法的教师发展了关于学生潜力的一种新观点。教师们还发现，他们自己正在把更多的时间用来处理学生的学习，而较少去管理和训练学生。

第一至第三章提出了关于教育和教育目标的一种观点，并详细描述了怎样才能利用评价使学生达到学习上的各级掌握水平。

第四至第六章是为了帮助教师注意评价的不同目的，注意开

发为课堂所用的各种评价工具的途径。教师将会找出改善自己目前所用的终结性评价的途径。关于诊断性评价和形成性评价的讨论，很可能在使用评价来改善教学和学习方面提出新颖的、大不相同的方式。

第七章描述编写测试题目和为特殊测试的目的挑选试题的一般技术。第八章到第十一章是环绕教育目标分类学组成的。对于各级教育和大多数学科中所见的各类目标，要编制出可作推理依据的评价工具。这几章就是为这种编制工作提供模型和技术的。为便于参考，附录中列有认知和情感目标分类的摘要。

本书虽然首先是为在校教师而写的，但它也是为师资培训方案和院系系的研究生学程中的学员们设计的。读者不应当把它当作一本要从头读到尾的书。更确切地说，它是教师或学员在遇到有关评价的各种疑难问题时可以查阅使用的一本手册——无论是在一学年的教学循环开始时或在循环中的各阶段，其中也包括最后的总结阶段。

读者应该先对全书的概念熟悉起来。这样在学年中的各段时间里，你要试图解决一定的课程、教学以及评价问题时，就可以再回过头来更仔细地阅读有关章节。

对于学习有关编制测试的各门课程的学生来说，本书也应当是有用的。既然全书与教育目标分类学有关，测试编制者将会从这里找到与他们要评价的目标和行为有联系的大量评价技术。书中的详细目录，将使测试编制者能够找到他们所需要的模型和例解。

但是就全书而言，它是一本适用于多种教学场合和学习场合的手册。我们希望师资培训机构里的学生将能在求学生涯的前期就得到本书。这样他们将会发现，在学习课程论的课程时，可以用它作为参考；在学习评价的课程时可以从中找到例解和模型；在学习方法类的课程时，它有助于把教和学的方法同自身学科范围内的适宜目标尽力联系起来，而且各项评价程序将会帮助他们更

加明了目标的意义和重要性。在他们进行实习教学时，他们将能发现本书对于改善他们的评价工作实践是有帮助的。最后，我们希望他们发现本书是如此之有用，因而在他们教学生涯中面临各项选择时，仍然继续使用本书。这类选择可以涉及课程和教学方法、评价程序的改善以及保持和发展一种关于教育、掌握性学习和评价的观点。

我们相信，熟练的教师最后会超越本书所提供的限度。——这时候我们希望本书已不再为人们所需要，或者说将会有新的书可供使用。在本书得到使用时，我们希望，书中把明细规格表转换成学习经验和评价程序的细节，并没有将教育之大道鸿猷的充分意义淹没掉。这些都是操作工序，用意在于要使更多的教师能够实现那个把他们吸引到教育事业中来的动人美梦——他们的学生在教育上得到了最充分的发展。

本杰明·S·布卢姆

乔治·F·马道斯

J·托多斯·黑斯廷斯

目 录

序.....	(1)
第一章 一种教育观.....	(1)
教育的功能：挑选还是发展.....	(1)
评价地位的变化.....	(3)
作为一种变革过程的教育.....	(5)
树立目标.....	(7)
确定目标的依据.....	(10)
学科的结构与学习过程的结构.....	(13)
总结.....	(21)
第二章 制定和挑选教育目标.....	(22)
教育目标的定义.....	(22)
教育目标的重要性.....	(23)
教学目标与两种成果模型.....	(29)
表述得当的教学目标的特征.....	(40)
制定教育目标的策略与根据.....	(54)
练习.....	(69)
第三章 为掌握而学习.....	(71)
导言.....	(71)
掌握性学习策略的各种变量.....	(74)
掌握性学习的一种策略.....	(85)
掌握性学习策略的成果.....	(93)
练习.....	(98)

第四章 终结性评价	(100)
终结性评价与形成性评价的区别.....	(100)
终结性测试的基本特点.....	(102)
编制终结性测试的一般步骤.....	(110)
为特定目的制订测试.....	(127)
修订测试的一般步骤.....	(137)
数据处理设备在修订测试中的使用.....	(151)
评分.....	(156)
总结.....	(171)
练习.....	(172)
第五章 诊断性评价	(174)
引言：诊断性评价的作用.....	(174)
对入学准备程度的评价.....	(177)
决定对学生的适当分置.....	(180)
在诊断性评价中使用能大量供应的测试.....	(200)
对一直不能从教学中获益的原因的诊断.....	(223)
总结.....	(226)
练习.....	(226)
第六章 形成性评价	(228)
导言：系统性评价的类型.....	(228)
反馈与改正的必要性.....	(229)
对学习单元的分析.....	(231)
形成性测试的编制.....	(241)
学生对形成性评价的使用.....	(249)
教师对形成性评价的使用.....	(259)
练习.....	(267)
第七章 试题的编写与挑选	(270)
导言：自发性评价与系统性评价.....	(270)
本章的中心议题.....	(271)

若干初步考虑·····	(274)
试题类型与编写各类试题的特定规则·····	(279)
有关试题编写与挑选的一般建议·····	(309)
试题库·····	(318)
关于试题编写的提要清单·····	(324)
练习·····	(324)
第八章 知识目标和领会目标的评价技术·····	(332)
导论：知识目标和领会目标的地位·····	(332)
知识目标·····	(332)
领会目标·····	(345)
练习·····	(359)
第九章 应用目标和分析目标的评价技术·····	(361)
原理和通则的应用·····	(361)
对应用的测试·····	(370)
分析·····	(390)
对分析的测试·····	(395)
练习·····	(414)
第十章 综合目标和评价目标的评价技术·····	(416)
综合·····	(416)
对综合的测试·····	(420)
评价·····	(435)
对评价的测试·····	(441)
练习·····	(473)
第十一章 情感目标的评价技术·····	(474)
导言：课程中的情感·····	(474)
忽视情感目标的原因·····	(475)
评价情感目标的重要性·····	(478)
表述情感目标·····	(480)
教育目标分类学：情感领域·····	(484)

情感领域中的评价问题.....	(494)
评价情感成果的方法.....	(496)
情感目标和评价技术的资料来源.....	(517)
总结.....	(525)
练习.....	(526)
附录：教育目标分类学的缩本.....	(528)

第一章 一种教育观

教育的功能：挑选还是发展

许多世纪以来，世界各地的教育强调了一种选拔功能，教师与行政人员的许多精力都用于确定在教育计划的每个重要阶段应淘汰的学生。公共教育体系的顶峰被认为是进入或完成大学学程。因此，在100个接受正规教育的学生中，只有约10%的学生被认为是由于天资或教养而能适合高等教育的严格要求。教育工作者对于在教育体系不同阶段所淘汰的90%的学生兴趣极微。

在这种十分浪费的做法的背后，隐藏着这样一种观点：只有很少的人真正能完成中学或进入并完成高校学程。教育的基本任务被假设为：鉴定少数可以进入并完成中学学程，然后再接受高等教育的人。这种挑选过程中的一些重要决定尤其早在儿童12岁之前就作出，其作用是剔除大多数工人阶级出身的儿童，而为专业人员的孩子提供特殊利益。尽管有人可以就这些儿童的遗传差异问题进行辩论，但是事情越来越明显，在不同社会阶级的儿童之间，可观察的差异实际在于以下方面：标准语言的发展水平；想得到尽可能多的教育的动力；乐意为教师表扬或为了长期目标而工作的程度；还有就是是否能以最少的反抗去接受学校所规定的学习任务。

儿童的这些特性，很少有人怀疑能被适当的家庭与学校的条件显著地改变。此外，有越来越多的证据表明，这些特征的早期改变会影响以后学校的学习（Bloom, Davis & Hess, 1965; Bronfenbrenner, 1974; Kennedy, 1978; Weikart, Bond, & McNeil, 1978）。

与这种为了选拔目的而使用学校的观点十分不同的见解是：

教育的基本功能是使个人获得发展。根据这种见解，学校的中心任务是：发展能使学生在一个复杂社会中有效地生活的那些特性。其基本假设是：通过教育手段能够发展英才；学校的主要资源应当用于增进每个人的工作能力，而不是用于预测或选拔英才。

许多高度发达的国家正迅速出现这样一种形势，10%以上的劳动力将需要某种形式的高等教育。经济体制中的变化，以及对更多受过高级训练的工人的需求，在很大程度上支配了这种动向。这种动向在某种程度上也是生活更加富足，以及人口中更广泛的阶层希望提高社会与经济地位的结果。几乎在世界上每一个国家里，对于各级教育水平的需求都超出了目前可得到的设施与机会。这种需求的一种征兆是：除了少数例外情况，几乎每个国家公共资金中的最大的单项支出是用于教育的，而且增加支出的要求正在迅速增强。

目前世界上有两个国家完成中等教育的年龄组比例尤为突出。在美国，75%的18岁青年中学毕业；在日本，将近90%的18岁青年完成了中等教育。在今后十年内，很可能会有相当数量的国家，在该年龄组中至少有五分之四青年完成中等教育。

简而言之，将来大多数青年会受到至少是中学毕业的教育。当大多数青年完成中等教育（或者在更远的将来完成二年或四年的高等教育）时，用来确定将给谁提供教育机会的选拔程序、预测以及其他判断便文不对题了。教育必须日益关心所有儿童与青年的最充分的发展，而学校的责任将是寻找能使每个学生达到他可能达到的最高学习水平的学习条件。

但教育并不随着中等或高等教育的完成而告结束。现代社会的变化性质，要求正规以及非正规的教育能终身进行。在熟练水平或熟练水平以上的大多数工人的职业生涯中，必须多次重新学习工作。希望前进到更高水平（根据任务或责任的复杂程度而言）工作的工人，必须学习所要求的新的技能以及新的职责。人们不再按“这将是学习者最后一次学习”的设想，去为任何专业或职业

计划某一课程。大多数职业界现在认识到了继续学习的需要，并且正在寻找途径，在工人职业生涯的不同阶段上提供教育机会。

然而，继续学习并不只限于提高职业能力。人们必须学习怎样做一个有效的父母、公民以及艺术的消费者(或生产者)，学习怎样理解世界的变化以及这些变化对于我们每个人的影响。不管是通过学校或是通过其他手段，教育必须被看成一种永不完全停止的过程(Faure et al., 1972)。

在这些新的、十分易变的情况下，如果教育要为青年与社会服务的话，显然学校与教师就必须学会用新的方式进行工作。虽然本书所论的是评价而非课程与教学，但这里可以清楚地看到，不仅评价而且教学，如果要适应它们的新要求的话，都必须经历深刻的改革。本章的任务是叙述某些关于教育(以及评价)的新看法，在我们看来，这些看法是与新的教育任务相一致的。

评价地位的变化

在世界各地，教育被认为是一整套学习任务；随着人们从正规学校教育的第一年走向最后一年，这些学习任务被假定为越来越难。也已有这样的假设：随着学生沿着这种教育阶梯上升时，具备必要的天赋才能或习得的技能与态度，从而成功地越过更高的阶梯的人将会越来越少。许多世纪以来，教育被看作象一个金字塔，在它的底部，较年轻的年龄组的全部或大多数成员都能入学，而只有极少数的人才能达到顶点。

考核被用来决定可以让谁进入高一级水平。作为这种过程的一部分，考核的结果与教师的判断已变为一种划分等级的制度，每年或更为频繁地把学生加以区分。

因此，教育被看成是具有固定的课程、按年级划分的成套学习任务以及一组混合的学习者。他们在这个系统里，在每个主要的时间单元上，都在被划分类别。考核或其他评价程序是用来对

每个学生在教育系统中的价值与前途作出决定，尤其是一些关键的、而且常常是不可逆转的决定。这些决定与划分等级常常影响到一个学生的全部生涯。

管理这种系统的人们——教师与行政人员，主要并不关心这种系统对于每个学生的学习或良好生活所带来的后果。尽管在每个社会中，对教育系统时时有种种议论，但很少采取协调的行动，多半只是对整体结构的某一微小部分作点更改。在这种教育系统中，确实产生了一小部分能够成功地越过所设置的困难的个人——而且其中某些人确实为社会作出了非常重要的贡献。

这种系统对于不成功的学生的影响并不是教师与行政人员所主要关心的事情，而在开始接受教育的人中，大部分在这一系统的某个阶段上是不成功的。常用的分类系统是把每个年级或层次中的学生分数，划分为接近正态分布（如A,B,C,D以及F）。由于这种系统从一个年级（或水平）到下一个年级有高度的延续性，我们的研究便发现：一些学生在每个年级都得到A等或B等，而其他学生则一次又一次地地被通知为D等或F等（Bloom, 1964; Hicklin, 1962; Payne, 1963）。这样划分的结果是，使一些学生确认他们是有能力的、良好的、合乎要求的；同时使另一些学生认为他们能力不足，是差的、不合乎要求的。（当然，这些都是从这种教育系统的观点来看的。）这种不断地贴标记的做法对个人的教育不可能带来益处，它对于许多学生的自我观念则可能具有不利的影响。身体上（以及法律上）被学校体系束缚了10年或12年，并在此期间反复得到消极的等级，这必然对人格与性格发展具有严重的不利影响（Bloom, 1976; Kifer, 1977; Stringer & Glidswell, 1967）。

评价的目的，象现有的教育系统最常用的那样，基本上是把学生分等与分类。它被用来区分失败的（D或F等）学生、成功的（A或B等）学生以及过得去的（C等）学生。学校内通常使用的测试以及其他形式的评价，对于改进教学作用甚微，而且难以确保所

有(或几乎所有)学生去学会那些被学校认为是重要的任务,达到教育过程的终极目标。

本书旨在提出对评价及其在教育中的地位的一种更加广泛的看法。我们主要关心的是用它来改进教与学。简略地说,我们的看法包括:

1. 评价是一种获取和处理用以确定学生水平和教学有效性的证据的方法。

2. 评价包括了比一般期末书面考试更多种类的证据。

3. 评价是简述教育的重要长期终极目标与教学任务目标的一种辅助手段,是确定学生按这些理想的方式发展到何种程度的一种过程。

4. 评价作为一种反馈——矫正系统,用于在教学过程中的每一步骤上判断该过程是否有效;如果无效,必须及时采取什么变革,以确保过程的有效性。

5. 最后,评价作为教育研究与实践中的一种工具,用于查明在达到一整套教育目的时,可供选择的程序是否同样有效。

作为一种变革过程的教育

对我们来说,教育是一种改变学习者的过程。根据这种观点,我们期望每个教育方案、教程以及单元能给学生带来某种(或几种)重要的变化。学生在一个单元结束时的情况应不同于在单元开始前的情况。在一个单元教育完成后的学生应当不同于未受过该单元教育的学生。尽管在中学开始与结束时学生自身的某些差异确实可以归因于成熟、生长与各种经验的影响,但我们这里关心的只是在于教育所产生的并且归根到底是由学校、课程与教学所决定的那些变化。

正是在这一点上,我们关心着手段与目的。毫无疑问,通过特定的教师,通过学生、教师与学科内容之间的相互作用过程,

通过经验的特殊的组合，能够显著地影响每一个学生。出自研究的目的，清理这些形形色色的过程与经验，以便确定什么事物对每个学生产生了影响，这可能是重要的。然而，我们对于评价的兴趣，只是试图描述、评估并在某种程度上影响所发生的变化，而不是分析引起这些变化的所有过程。因此，我们希望把评价（其基本目的在于描述与影响变化）的作用与研究方法论的作用加以区别。后者可能寻求对于原因与结果的理解，或者详尽分析被认为是引起学习者变化有重要意义的变量。据我们看来，评价乃是系统收集证据用以确定学习者实际上是否发生了某些变化，确定学生个体变化的数量或程度。

在提出这些教育过程的观点时，我们正在表明自己的信念（有许多证据支持这种信念）：教育能够使学习者发生重要的变化。这并不是说所有学习者会完全按同样的方式变化或者变化到同样的程度。这也不是说所有教师、课程以及学校将会同样有效地改变他们的学生，或者会以同样的方式改变学生到同样的程度。

如果教育的作用是使学习者发生变化，那么人们必须确定哪些变化是可能的，哪些变化是令人满意的。教师与学生的每一种相互作用都是建立在师生双方对于某些变化的可能性与合意性的某种隐含的信念基础之上的。忽视这些基本问题实际是不可能的。每种教学行动以及每种学习行动都来自对于这些问题的回答。不希望表述自己教育目标的教师只是回避了对这些问题用语言作出公开回答，而他们的行动及其与学生的相互作用便是隐含的回答。

然而，用语言对终极目标予以规定，并不能确保隐含的终极目标是与申明的终极目标相一致，也不能确保申明的终极目标将会实现。这就是评价的用途之一——把学生变化的实况与对企求中的变化的陈述方式联系起来。