

教育评价文选

北京教育行政学院资料室

1986年3月

陈
三
平

教育评价文选

(一)

北京教育行政学院资料室

总 目



第一部分.....	1 ~ 5 6
第二部分.....	1 ~ 4 1
第三部分.....	1 ~ 8 2

目 录

关于教育评价问题的概述.....	1
教育评价.....	14
教育评价的基本概念.....	19
略论普通教育评价.....	27
关于教育评价.....	40
怎样评价教师的工作.....	46

关于教育评价问题的概述

北京教育科学研究所

张秀媛

当前在教育改革当中，关于教育评价问题越来越引起人们的重视。探讨、研究教育评价这一课题的单位和个人也越来越多。人们首先遇到的问题就是什么是教育评价？教育评价有没有科学的定义？它有什么意义和作用？这些问题都需要进一步弄清，以便为开展这一课题的研究，提供理论上的依据。

根据我们收集到的有关资料，对教育评价问题概述如下：

一、关于教育评价的概念

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。测验只着眼于客观知识的测定，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去评量教育所预期的一切效果。

二、教育评价的产生

根据现有的材料，一般认为教育评价一词，最早是由美国进步主义教育联盟于1933年提出的。三十年代初期，美国进步主义教育联盟组织了著名的“八年研究”。为研究出一套评价方法，组成泰勒（R. W. Tyler）教授为首的评价委员会。该委员会以新教育学理论为依据，接受心理学的观点，以全面发展人的才能为目标，为“八年研究”新课程实验设计了一套教育评价办法。经过八年实验证明教育评价促进了新课程实验的效果，教育评价获得成功。该委员会最后提出的史密新——泰勒报告，被后人称为“划时代的教育评价宣言”。这个报告第一次提出了评价的思想，指出对于教育效果，不能只是测定受教育者的某些能力和特征，而应该评价受教育者由于教育的作用向着教育目标的成长和发展。因此，教育评价的基本思想是：1、所

谓教育，就是使人的行为方式变化和改进的过程；2、这些形形色色的行为方式的变化，就是教育目标；3、评价教育过程，就是看这些教育目标实际上达到了什么程度；4、因为人的行为是复杂的，所以评价要从各个侧面进行，不仅要有分析，而且要有综合；5、评价的方法仅用纸和笔的测验是不够的，还须包括行为观察等多种多样的手段。

这种教育评价思想，先是被全美国采用，后又传遍西欧各国。第二次大战后，世界各国为繁荣经济，开发智力资源，普遍注意人才的培养，因而积极推行教育评价法，使教育评价发展成“评价运动”。

三、教育评价的发展

从世界范围看，教育评价的发展是一个不断追求客观性的同时逐渐向追求教育性发展的过程。这一过程迄今共经历了四个阶段。

(一)从测定到评价

十九世纪末由于实验心理学、个性差异的研究和智力测验的出现，心理测验的原理与方法应用于教育，于是就出现了教育测验。在本世纪四十年代以前，心理测验和教育测验都追求客观性，因此所有难以客观地、数量化地加以描述的事物，都被排除在测量对象之外，这样测定所抓住的只是教育上零星片断的现象，完全忽视了教育对象的全方面的人格发展。

教育评价的思想和方法正是弥补了教育测验的不足，主张全面考查教育与学习的各个方面。后来形成的教育评价运动，主要强调不仅要评价知识的掌握情况，还要评价应用、分析、综合等高层次的认识能力，以及评价兴趣、关心、价值观、态度等情意特性。

(二) 建立教育目标体系

1948年美国心理学大会，提出了一项重要任务，就是对组成教育目标的人的能力和特性进行科学分类。为只有在科学的和统一的教育目标分类学的基础上，各个进行测定、评价工作的研究者，才能有共同语言，才能深入地开展研究。

教育目标的第一级分类是认知领域、情意领域、精神运动领域。认知领域的教育目标，由与知识的掌握、理解和认知能力的发展有关的各种目标组成，该体系由布卢姆(B. S. Bloom)等人于1956年完成。情意领域的教育目标，由与兴趣、态度、价值观的形成和正确的判断能力、适应能力的发展有关的各种目标组成，该体系由克雷斯沃尔(D. B. Krathwohl)等人于1964年完成。精神运动领域的目标，由与手的技巧及其他运动技能有关的各种目标组成，这套体系至今尚未完成。

建立教育目标分类学本身就包含了客观性和教育性这两方面的追求。分类研究可分两方面：一方面是教育学上的，即对课程和教育目标的研究；另一方面是心理学上的，即对能力、价值观和技能的掌握和学习、内化过程等课题的研究。这两方面的研究成果，以教育上的有效性、逻辑性和心理依据为准则，进行综合、整理，以行为目标的形式设定出最具体的目标，然后按照发展顺序，层层上达构成体系。

布卢姆等人建立的教育目标分类学，已成为现代教育测定和评价的重要基石。已被很多国家的教育家采用。

(三) 从“心理测定性”的评价到“教育测定性”的评价

这个转折的标志，就是提倡绝对目标标准的测定和评价。相对目标标准的测定和评价，以客观性为首要的追求目标，对于个人学力，

往往不是就其本人的实际水平来测定或评价，而是用该人与他人的相对位置来表示。这种以正态分布为依据的测定和评价，的确有其客观性强的优点，但也有缺乏教育性的问题。

从四十年代的评价运动开始，人们就在追求评价的教育性，但只有在教育目标分类学建立起来之后，这种追求才有了实现的可能。

格拉泽 (R. Glaser) 1963年发表文章指出相对目标标准的不足，提出在今后的学校教育中，应该着重于绝对目标标准的主张。这个主张立刻受到欢迎。这方面的理论研究和实践研究，很快在世界各国蓬勃地开展起来。

相对目标标准评价和绝对目标标准评价的性质区别，正如有人指出的，前者是表现在正态分布中的、一贯而稳定的、个人和个人之间的差异，所以它在原则上属于“心理测定性的 (Psychometric)”；而后者以到达目标的形式，设定所期待的、通过教育活动来实现的学力内容，并且就每个人来测定这些目标是否达到。换言之，它关心的是个人内部的发展，所以它属于“教育测定性 (edumetric)”。

四从总括性评价到形成性评价

斯克里芬 (M. Scriven) 于1967年，在论述教育课程计划的发展时，曾提出有两种类型的评价。一种是总括性评价，即在教学完成之后，对教学成果进行的评价。从这个意义上说，以往的教育评价都是总括性评价。另一种叫形成性评价，它是建立在人们对于评价的反馈机能和教育决定上的重要意义的认识上的，它在教学过程中进行，目的是形成适合于教育对象的教学。形成性评价由斯克里芬提出，经布卢姆运用于教学活动中，经过丰富、发展，目前已成为教育评价的重要内容。

以上四个阶段就是世界范围的教育评价运动发展的过程。总的发展趋势是从相对评价向到达度评价、从总括性评价向形成性评价转变。

(待续)

(选自《教育研究资料》1985年第1期)

关于教育评价问题的概述(续)

北京市教育科学研究所

张秀媛

四、教育评价的现状

近二、三十年来,国外对教育评价问题已有相当多的理论研究和实践经验。随着教育规模和普及程度的飞速发展,教育评价问题的研究和应用日益广泛深入。尤其是六十年代更出现了较大的转折,形成了教育评价运动。要求教育评价从主要为筛选合格的儿童服务,转变到为教学、为促进受教育者的发展服务。这是一股世界性的潮流。日本的到达度评价就是这种潮流的表现之一。

1、日本提出到达度评价的背景

1948年日本制定的学校教育法规定,学习成绩的评定办法,实行五级相对评价。这种评分方法,要求按照正态分布规律把全班或全校学生成绩分成两头小中间大的五级。五级的比例,一般都按1:2:4:2:1这样的分配打分。实践证明五级相对评价不利于学校教育的发展,不能反映学生的学习质量,反而造成激烈地竞争和层层筛选的“考试地狱”。因为硬性制造的正态分布往往不符合班级的实际情况,五级相对评价的成绩并不表示学生掌握知识技能的实际水平,只表示他在班里属于第几名。这样,师生们首先关心的不是掌握知识和技能,而是考试的名次。考试规模越大,名次越重要,统考的分数经标准化以后,就成为指导学生升学就业的依据。一次次考试就是一次次竞争,一次次预选。无形之间,教育目的变成不是为了受教育者

的发展，而是为了使他们通过考试了。人们把这种教育称为“考试主义教育”。

此外，五级相对评价的最大弊病，就是抹煞师生的积极性。因为种种原因，学生的成绩可能长期固定在某一等级上，这样学生就看不到自己的进步，成绩差的就心灰意冷，不求上进。同时，教师也无意努力去提高全体学生的水平。因为相对评价总是否定或看不出这种努力的成果。总之无论师生如何共同努力，成绩始终是两头小中间大，因此，大大地影响了师生的积极性。由于放任自流的结果，出现了许多问题。虽然日本文部省针对暴露的问题对成绩评定方法作了一些调整，但是由于仍然坚持相对评价的原则，中小学里的竞争和筛选有增无减，由它造成的问题也随着竞争的激化，而日趋严重。

2. 关于到达目标与到达度评价的理论

面对“考试主义教育”的恶果，急需求得问题的解决。

1966年日本国立教育研究所的研究人员板仓圣宣，发表了题为《正确的学力评价》一文，试图从探讨教育目标与教育评价的关系入手，解决中小学的教育评价问题。板仓把教育目标分为“方向目标”与“到达目标”两大类。第一次提出了“到达目标”这一概念。他认为在学校教育中，特别是在义务教育阶段，教育目标应该从“教师如何组织所传授的教学内容”的角度提出，通过深入研究教学内容来设定到达目标，组织合理的体系。使每个儿童都能切实地掌握知识和技能。同时他主张应当采用绝对评价，通过评价来了解学生掌握教学内容的程度。

板仓提出的到达目标的主张，引起教育科研人员的重视。御茶水女子大学的中内敏夫教授把到达目标的概念引进教育学的范畴，进行

了整理和深化。他认为要彻底改革相对评价，首先必须彻底改革现行教学大纲和指导要录所遵循的方向目标的理论，代之以到达目标的理论。中内认为到达目标不是来自教育基本法，而是来自教育实践，来自儿童的生活逻辑同教材逻辑之间的矛盾运动。因此，它不象现行教学大纲的目标那样，按学科、学年规定死了的条目，而是同儿童发展相适应的、既是可变的又是一贯的东西。其次，到达目标可归结为概念、法则、课题、方法、知识等等。所以成绩通知单上应记载该学科各方面的到达目标。最后，相对于方向目标的评价手段——能力测验来说，到达目标的评价手段是学力测验。能力测验的目的是了解儿童可以学会什么，学力测验的目的在于了解可以教会儿童什么。前者测验范围超大纲，后者仅仅对传授过的知识进行测验。

中内敏夫的研究为七十年代兴起的中小学成绩单改革运动提供了理论基础。

东京都立大学教授坂元忠芳，在总结了前一阶段的实践之后，进一步发展了到达目标理论。他认为前一阶段在实践中所拟定的到达目标都限于概念、知识和技能的范围。这些只构成儿童学力的外部，学力还有其内部，即学习态度、倾向、意志等等，在某种意义上这些内部因素更为重要。他认为具体化为到达目标的知识、技能可以数量化，但构成学力内部的因素却几乎不可能进行数量化的评价，因此，他对于相对评价一般持不排斥态度，只反对中小学的五级相对评价。他要求到达目标既应体现学科系统的知识、技能，又体现儿童认识能力的发展水平。

日本女子大学 田馨一针对前一阶段到达目标理论上的缺陷，结合布卢姆等人的教育目标分类学，提出了三分类的到达目标理论。梶

田把到达目标分为达成目标、提高目标、体验目标三大类，每一类各包含三个领域：认知领域、情意领域和精神运动领域。每个领域都包括一系列目标，并各有其测定和评价的要求。

梶田把促进每个儿童在各方面都最大限度发展的教育称为修得主义教育（意为全班都达到学习要求）。他认为修得主义教育的到达目标，必须具备以下几个性质：

(1)到达目标应表示出必须掌握的和必须发展的最低限度的内容。

(2)到达目标应该是有系统的，以便能切实地集中地反映出中小學生身心发展的过程和整个课程的结构。

(3)到达目标体系必须包括测定和评价时期、方法和观点各不相同的范围广泛的目标。

(4)到达目标是学生学习的目标，也是教师指导的目标，教师有责任使每个学生都达到目标。

梶田、坂元等人的理论工作，使到达目标与到达度评价的理论日益完善。他们的理论对中小学教育的到达度评价的实践，起了重要的指导作用。

3、到达度评价的具体实践

在理论研究的同时，日本各地到达目标与到达度评价的实践研究也开展得比较活跃。开展最活跃最普遍的是京都地区。根据这个地区的报告和有关文章，他们的实施办法大致可概括为以下四个阶段。

(1)目标分析

①分析该学科的本质和目标。这是第一次目标分析，包括分析学科内容和结构，明确学年内的目标，这些目标分为三大类：

第一类可直接设定为到达目标的，能进行客观而科学的评价的目

标，主要是认知领域，技能领域的目标。

第二类是难以规定为到达目标，难以客观评价的目标，主要是情意领域的目标，需要通过认知领域、技能领域的教学加以发展。

第三类虽不是到达度评价的对象，但对保证儿童发展有意义，可设为体验目标如手工劳动、游戏之类。

②确定基本教学事项，明确各教学单元的目标。这是第二次目标分析，主要结合教材内容对第一次目标进行细分化，确定它们的具体内容。

③制定计划。第三次目标分析，包括教学计划和评价计划，要根据诊断学生学力基础，确定教学进度和起点，形成教学结构，规定教材目标。

(2)诊断性评价

评价计划的第一步，是和教学计划相配合的，是在目标分析与教学之间进行的评价。目的在于：①了解学生在教学之前是否已到达了目标，如果达到了，就调整计划。②了解是否到达了前提要求（已有知识基础），如果没有到达前提，就要补充学习，使之获得学习的前提。③对于到达前提的学生，还要了解是否具有丰富的经验和信息，对于没有经验和信息的，要教给他们。④在了解了以上情况的基础上对教学计划进行调整或修改，使之符合学生的实际情况，然后进行教学。

(3)形成性评价

这是评价的第二步，是在教学过程中进行的。教学过程就是按计划一步一步地落实第三次分析后的目标，在这个过程中要防止单纯传授知识、技能，忽视发展认识能力的倾向。

形成性评价的实施步骤:

①测定原到达目标到达与否。

②对未到达的进行再教学。

③对再教学的进行再测定,对仍未到达目标的学生采取“诊断”和“治疗”措施。

④对极少数治疗无效者,继续进行诊断和“治疗”。

⑤对于已到达目标的学生,使之进行巩固、深化、发展学习。即使之记住或回忆学过的内容,同时从科学史、技术史以及社会等各方面的关系入手,深化和发展学习的内容,提高学习兴趣,改善学习态度等。

⑥根据评价反馈信息,对下单元的教学计划进行调整。

(4)总结性评价

这是评价的第三步,在学期末或学年末进行,对基本学力掌握、巩固的程度以及综合能力进行评价,明确下学期或下学年的补救措施和进一步教学应注意的问题等。

以上到达度评价的基本实施过程,可以明确从目标的具体化开始,到制定计划时的诊断性评价、教学过程中的形成性评价、学期终了时的总括性评价,三个评价步骤环环扣紧,每一步都伴随着评价、反馈和调节,形成了一个有机的体系,以保证所有学生都能得到最低限度以上的发展。

近二十年来,教育评价的指导思想出现了一个大转折,这个转折概括成一句话,就是过去是要“选拔适合于教育的儿童”,如今是要“创造适合于儿童的教育”。前者就是过去的评价,把教育方式视为固定不变的,从而尽力选拔、筛选适合于这种教育方式的儿童,所以

在方法上一直偏重于相对评价。后者即今天的教育评价，它的目的在于全面地、最大限度地开展和促进每一个儿童的发展可能性，所以它力图从一切可能的教育方式中选择、发现，以便创造出最适合于儿童的教育方式来。为达此目的，才提出到达目标与到达度评价，才有形成性评价——为形成适合于儿童的教育的评价。

综合上述的评价都是为改进教学服务。如何进一步提出为促进儿童个性发展的评价，还有待于今后的研究和发展。（全文完）

参考文章：

①《教育评价的理论和方法》

李聪明著（台湾）幼狮书店1972年10月版

②《到达目标与到达度评价》

万勇著《外国教育资料》1983年6期

③《教育评价之方向》

赤木爱和著（日）《儿童心理》1982年3月号

（选自《教育研究资料》1985年第2期）