

XIAOXUE ZIRAN
KEXUE
JIAOYU PINGJIA

全国中小学教师
继续教育教材

小学自然 科学教育评价

周若新 著



东北师范大学出版社

全国中小学教师继续教育教材

XIAOXUE ZIRAN KEXUE

JIAOYU PINGJIA

■ 东北师范大学出版社

长 春

小学自然 科学教育评价

周若新 著

(吉) 新登字 12 号

- 出版人：贾国祥
- 策划编辑：杨华云 唐东梅
- 责任编辑：张志文
- 封面设计：未 名
- 责任校对：李 群
- 责任印制：张允豪 栾喜湖

全国中小学教师继续教育教材

小学自然科学教育评价

周若新 著

东北师范大学出版社出版发行

长春市人民大街 138 号 (130024)

电话：0431—5695744 5688470

传真：0431—5695744 5695734

网址：<http://www.nenu.edu.cn>

电子函件：Chubs@ivy.nenu.edu.cn

东北师范大学出版社激光照排中心制版

长春市南关区文教印刷厂印刷

1999 年 10 月第 1 版 1999 年 10 月第 1 次印刷

开本：850×1168 1/32 印张：4.375 字数：102 千

印数：0 001 — 5 000 册

ISBN 7 - 5602 - 2469 - 5/G · 1389 定价：5.50 元

出版说明

历史将翻开新的一页，人类即将跨入 21 世纪。21 世纪是充满机遇和挑战的世纪，是一个科学技术更加发达，竞争更加激烈，社会对人的素质要求更高的世纪。提高人的素质的关键在教育，振兴教育的关键在教师，只有造就一支高素质的教师队伍，才能满足 21 世纪教育发展的要求。而建立和完善适应 21 世纪需要的中小学教师继续教育制度，则是造就高素质中小学教师队伍的根本措施。

1998 年 6 月，国家教育部师范教育司制定并印发了《中小学教师继续教育课程开发指南》（以下简称《指南》）。《指南》对中小学教师继续教育的教学内容和课程体系作了原则规定，对现阶段中小学教师继续教育提出了基本要求，这标志着我国中小学教师继续教育教学内容和课程体系的确立。

我们组织编写的这套教材是以《指南》为指导，按《指南》所规定的课程和内容要求而编写的。我们目前出版的这些教材，大部分都是《指南》中规定的必修课。根据中小学教师继续教育开展的情况，我们还将陆续组织编写出版《指南》中规定的其他教材。

在教材编写过程中，我们认真汲取了“八五”期间全国各地

开展中小学教师继续教育的宝贵经验，坚持从中小学教师队伍建设的需要和中小学的实际出发，力求反映先进的教育思想、教育理论，反映最新的学科知识发展动态、教育教学改革实践和研究成果，反映现代教育技术和先进教学方法，在确保科学性的前提下，进一步突出了教材内容的针对性、实效性、先进性和时代性，体现了中小学教师继续教育的特点和要求。

由于时间仓促，加之中小学教师继续教育教材建设尚处在起步阶段，缺乏足够的经验，缺憾之处在所难免，恳请广大读者不吝赐教，并在研究和探讨方面与我们进行更多的合作。

希望本教材能对广大中小学教师完善自我，提高自身素质，顺利地跨入 21 世纪，助一臂之力。

东北师范大学出版社

1999 年 7 月

目 录

第一章 教育评价的产生和发展 / 1

- 第一节 教育评价的历史和发展 / 1
- 第二节 教育评价理论的发展趋势 / 5
- 第三节 我国教育评价的发展历史 / 8

第二章 教育评价的意义 / 11

- 第一节 教育评价的概念 / 11
- 第二节 教育评价的类型 / 16
- 第三节 教育评价的意义和作用 / 21

第三章 教育评价的原则 / 24

- 第一节 确立教育评价原则的必要性 / 24
- 第二节 教育评价应遵循的原则 / 25

第四章 教育评价标准的制定/30

- 第一节 制定教育评价标准的要求/30
- 第二节 教育评价标准的设计/33
- 第三节 教育评价评定标准的制定/40

第五章 自然学科的特点与目标/46

- 第一节 自然学科的特点/47
- 第二节 自然学科的目标/50

第六章 自然学科教学的过程/57

- 第一节 自然学科教学过程的本质/57
- 第二节 传授间接经验的自然学科教学过程/64
- 第三节 获得直接经验的自然学科教学过程/71

第七章 自然学科教学的评价/76

- 第一节 教师授课质量评价的基本指标/76
- 第二节 自然学科教学评价指标体系的探讨/82

第八章 小学生学习《自然》效果的评价/89

- 第一节 学生科学精神、态度与价值观的评价/91
- 第二节 学生科学知识与技能的评价/99
- 第三节 学生科学方法与能力的评价/106

第四节 学生科学行为与习惯的评价/112

第九章 小学自然学科教育评价改革/119

第一节 素质教育评价研究的深入发展/119

第二节 自然学科学生学习效果评价探索/124

第一章

教育评价的产生和发展

小学教育是以全面提高国民素质为根本目的的，是以学生的自然属性为基础，重视学生成长与发展的实际需要的基础教育。为了使教育适应时代生产的社会化和科学技术的进步，以提高学生素质为核心的教育改革正在深入进行。因此，在教育和教学的过程中，必须对自身行为的结果作出反省与评价，以追求更好的目标和效率，从而更有力地促进学生的成长与发展。而教育评价作为新兴的教育科研领域，已受到世界各国的重视。对教育评价的深入研究，建立多学科教学的评价体系，应该成为实施素质教育的重要组成部分。

第一节 教育评价的历史和发展

一、教育测验的萌芽

人类社会有教育活动，便有考评其成效的方法。我国的《学记》有这样的记载：“此年入学，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。”这可说是当时考评教育活动的的方法和标准。公元606年，隋炀帝时代开始了封建科举制度。所谓科举，就是分科进士，即设科考试，根据考试成绩去录用官

吏。唐代每年举行一次，明清则三年举行一次。沿续了 1300 年之久的封建科举制度，其实就是一个关于学生学历检测、评价的制度。虽然，科举制缺乏对人全面的、科学的考查，并存在着许多弊端，但是，对世界各国公务员的录用考试均产生很大影响。

在 18 世纪以前的西方各国，由于学校尚未普及，考试的方式主要是口试。1702 年，美国剑桥大学首开学校考试笔试之河。1845 年，美国初等学校普及，学生数激增，于是，波士顿市率先在美国以笔试取代口试，考查毕业生。由于笔试客观性、可靠性比口试高，而且节约时间，测笔结果大大优于口试。但是评分易受主观偏见影响，而且不足以反映学生所获知识与能力的全貌，因此，客观性大受影响。西方国家于 19 世纪以后陆续建立了文官考试制度。英国于 1855 年建立了“文职人员委员会”，采取竞争性的公开考试来招募文职人员。美国也于 1883 年规定担任公职必须经过考试。意大利在宪法里规定：为了升入各种、各级学校或在其校毕业，以及获得就业资格，均须经过国家考试；国家行政机关的官员均由考试选拔录用。日本受我国科举制的影响，在明治维新以后相当长的历史时期中，高等文官考试均仿照科举的办法。因此，可以说教育测验起源于中国的科举考试制度，但因中国长期封建制度的束缚而未得到应有的发展，却在西方的国家得到了发展。

二、教育测验的开拓期（1900—1915 年）

这一段时间是教育测验方法的探索与初步的发展时期。1904 年桑代克发表了《精神与社会测验学导论》。这是一个在测验学史上划时代的巨著，标志着测验运动的开始。桑代克在书中称，“凡是存在的东西都有数量”，“凡是有数量的东西都可以测量”，引起人们对测验方法进行探索。1905 年法国的比纳（A. Binet）与其助手西蒙（T. Simon）发表了《诊断异常儿童智力的新新方法》，

介绍了第一个智力量表——“比纳·西蒙量表”。1908年比纳对其智力量表进行了修订，把测量项目增加到59个，并采用了智力年龄的方法计算成绩。1911年比纳对其智力量表又作了第二次修改，增添了许多新测验，废除了一些旧测验。为此，比纳被称为智力测量的鼻祖。在学历测验方面，1908年，美国教育家斯通(C. W. Stone)提出了以小学生为对象的测试算术的方法。同时，美国有许多规模大的学校系统完成对各种教育活动进行综合测量，测量采用了最新的客观测验方法，测量内容为算术、拼字、书法和英语作文，来衡量教育质量。早期学校系统的测量主要采用标准参照测验，1920年出现了常模参照测验，这种测验开始用于个体行为水平，后来发展到用于团体的测量。这些充分表明这一时期的研究重点在于测验客观化和标准化。

三、教育测验的兴盛期（1915—1930年）

在20世纪的20年代后期和30年代，大学介入了教育测量领域的研究，并指导当地的测量活动。不但在技术方面努力求得正确应用，而且不断提高到理论上加以证实。在这一阶段，教育测验已发展成3种不同性质的测验，即学历测验、智力测验和人格测验。在学历测验方面，从桑代克发表书法量表到1928年止，据统计已有心理标准测验和标准学历测验3000余种之多；在智力测验方面，自比纳—西蒙智力测验传入美国后，经辜鲁满(F. Kuhlman)及斯坦福大学推孟(C. M. Terman)的相继修改，加以量化，并引用德国石登(W. Stern)智能商数(IQ)；在人格测验方面，1921年，华纳德(G. G. Fernald)着手试作人格测验；1924年—1929年，哈芝红(H. Hartshorne)等人组织了人格教育委员会，着手研究人格测验工具，并不断加以改进，使之相当精密。

在这一时期，测量和评价被认为是同义词，评价被当作以测验结果来对学生行为进行分类。评价主要是为了满足当地的信息

需要。

四、教育测验的批判期（1930—1945年）

——教育评价概念的诞生

随着教育测验运动的不断发展，人们逐渐认识到教育测验的弱点和不足——它不能测得人的全部。无论是知识测验还是人格测验，都只能做片断的测定，不能全部了解人格与知识的发展过程；测验只是注意于客观的信度，不足以说明效度；教师为测量成绩所采用的学业测验的方法，其实质就是教科书中心主义。对于学生的态度、兴趣、情感等方面人的素质靠单纯的学业测验是无法把握的。

1929年，美国教育家泰勒（W. R. Tyler）在改进学校测验方法的实践中提出，测验应以一定的教育目标为指导。教师应首先找出课程目标，根据学生应掌握的内容、方法，用行为将课程目标表述出来，制定出一系列教育目标。这样的测验不仅能检查学生的记忆能力，还能检查学生的实际动手能力和解决问题的能力。于是泰勒将“评价”概念应用到教育上，以同“测验”相区别。他指出，教育“评价”就是这样一种过程，即发现学生学了什么和这些东西所具有的价值。

同年，美国发生了严重的经济危机，经济萧条使许多青年无法就业而只能进学校。但是美国的中学主要是为学生升入大学服务，不适应失业青年的需要，这就使学生与学校课程之间发生了尖锐的矛盾。在这种情况下，美国的一些教育家成立了进步主义教育联盟，为寻找新的出路而进行了一项课程内容实验研究，从1932年到1940年历经8年完成，史称“八年研究”。为了评价其实验研究成果，美国组成了以泰勒为领导的评价委员会。通过这场研究，泰勒及其同事正式提出了教育评价的概念，即教育评价就是衡量教育活动达到教育目标的程度，测验是它的手段。他们

还提出了评价的原则和方法,其基本思想是:(1)教育是改变人的行动方式的过程;(2)教育目标是各种行动方式的变化;(3)教育能够根据对该目标实际完成情况分析、评价;(4)人的行动是复杂的,所以应从各方面对学生进行综合评价;(5)评价方法应采取包括观察行动在内的更为广泛的方法。可以说泰勒的教育评价概念的提出,是教育评价理论和方法的开端。

第二节 教育评价理论的发展趋势

一、教育评价理论发展的两个阶段与特点

教育评价理论的发展可以划分为两个阶段。第一阶段为1946—1957年;第二阶段为1957年到现在。

(一) 第一阶段的特点

1. 教育评价活动尚未普及

这个时期教育评价活动尚未引起政府与群众的普遍重视,仅限于研究教育的专家学者进一步研究。

2. 教育评价的理论与方法有了新的发展

泰勒的教学评价概念提出后,人们普遍认为教育评价是教育领域中的重要课题。但是为了使评价能达到预期的目的,关键是要明确每种教育活动的目标,为此,很多人着手研究教育目标问题。于是,就产生了教育目标分类学。在此期间,布鲁姆(B. S. Bloom)和他的同事建立了目标分类学,它对完善学习和教育评论理论起了重要作用。建立目标分类学的任务是1948年提出的,直到1956年才完成。布鲁姆的目标分类学是关于认知领域的,其中包括知识的掌握、理解和认识能力的发展。与此同时,有人研究情意方面的教学目标(包括兴趣、态度、价值观的形成和正确的判断能力,适应能力的发展)。这个目标体系由克雷斯沃尔于1964

年完成。

（二）第二阶段的特点

1. 教育评价活动逐步普及

这个时期教育评价由教育界所关心的课题逐步变成各级政府、社会、教师所关心的课题。例如，1957年苏联人造地球卫星发射成功，在美国引起了震动。美国人分析他们自己落后的原因在于教育。随后，政府当局大规模地对学校开展评价工作。由此，教育评价由专家的课题转入了政府部门的教育措施。

2. 评价的范围扩大

20世纪60年代以前，绝大多数评价都是针对学生学习的。自60年代中期以来，评价范围扩大到了对教育内容和方法等的评价，即扩大到了中观和宏观层次。随着教育事业的发展以及评价理论和方法的逐步成熟，中观和宏观层次的评价问题将更加受到重视。

3. 由早期注重“终结性评价”方法到现代注重“形成性评价”方法

过去的教育评价为选拔适合教育的儿童服务，而现代教育评价则注重创造适合儿童的教育。其次，现代教育评价从早期重视终结性评价逐步转入更重视形成性评价。所谓形成性评价，是通过揭露存在的问题，反馈有关的信息，以促进教育工作的改进。而终结性评价，则是对活动结果的评价，它的目的在于证明对评价对象作出的鉴定。因此形成性评价更具有为改进教育工作服务的作用。

4. 重视在评价过程中的自我评价

随着教育评价的发展，强调被评价者进行自我分析和自我认识，并且把被评价者自我评价的意见与评价者的检查、评价有机地结合起来，使评价结论尽可能取得一致。通常现代教育评价活动主张把被评价者的自我评价作为整个评价过程的预评阶段，这

样不仅有利于在被评价者自我评价的基础上，由专家们作出的结论为被评价者接受和理解，防止产生对立情绪。

5. 重视定性和定量评价的结合

近年来，随着现代教育评价活动的深入，人们越来越发现，教育现象全部应用数学方法来量化评价是不可能的，当然也无此必要，而一些定性方法仍然是必要的。所以国外的教育评价理论又开始重视把定性分析与定量分析的评价方法有机地结合起来，根据实际情况灵活地加以运用。

6. 教育评价理论体系逐步完善

随着国外 40 多年来教育评价活动的普及，教育评价理论也逐步完善起来，相继提出了比较系统的教育质量评价原则、评价方法及评价程序等。1984 年，美国联邦政府教育部提出“要设计并实行科学评价学生知识、能力、技能的系统方法”。日本在第二次世界大战之后，在教育领域开展了教育评价工作，出版了很多教育评价专著。前苏联 1984 年提出要“制定出评价教师和整个学校工作的有科学根据的标准，提高他们对客观评价学生知识的责任感”。同时，还强调“评价学校与教师的工作首先要根据学生知识的深度和巩固性、思想品质与道德品质，作好走向生活与参加劳动的准备程度”。苏联一些学校的领导人认为：对学生的评价应是全面和经常的，不仅对知识，而且还要对学习的勤奋程度进行客观的评价。这是时代的要求，是日益增长着的教育的社会作用的要求。

世界上还有许多国家，如德国、美国、澳大利亚等，也都十分注重教育评价的研究。教育评价已成为当代世界教育科学研究的三大领域之一，国际上已组成了“国际教育成就评价协会”（简称 IEA），开展世界性的教育评价和研究工作。

第三节 我国教育评价的发展历史

一、我国教育评价发展的进程

教育测验的思想起源于中国古代，它对现代教育评价发展有着较大的影响。但是由于我国长期封建制度的束缚，未能很好地得到发展。20世纪初，正值西方教育测验运动兴起，其理论也传入中国，有的学者也开始研究并撰写了有关教育评价的论著，但30年代后由于日本对我国的侵略，中华民族处于危难之中，教育评价的研究也就中断了，美国泰勒的8年研究成果也未能及时被中国政府或教育界人士所认识。从解放到“十年动乱”，由于众所周知的原因，我国的教育评价未能得到充分的研究，教育评价工作显得十分落后。

80年代以来，为了迅速提高我国学校的教学质量，根据教改实践的需要，教育测量和教育评价研究才逐步得到了恢复和兴起，并迅速发展。

党的十一届三中全会以来，我国在拨乱反正基础上，坚持推行全面改革和对外开放，开创了社会主义现代化建设的新局面，这使得教育事业也需要不断改革和发展。由于教育评价具有系统搜集信息、对教育工作过程及成果质量进行科学评定并作出价值判断、为科学管理与决策提供依据等功能，因而日益受到教育界以及社会的密切关注。

1985年，中共中央在关于教育体制改革的决定中指出了对教育要进行评价的问题。1986年，六届全国人大四次会议通过的《关于第七个五年计划的报告》中也提出：“要加强教育事业的管理，逐步建立系统的教育评价和监督制度。”各级教育部门都日益

重视教育评价，把它作为加强宏观指导和管理的重要手段。

学校的领导和教师，在教改实践中也深切认识到教育评价的重要性。感到端正办学思想，进行教育课题研究，提高教学质量，检验各种改革措施，都离不开教育评价，促使我国教育评价工作蓬勃发展。

二、近 20 年来我国教育评价研究的成就

在国家教委的领导下，近些年来经过教育评价理论工作和教师的共同努力，我国教育评价的理论研究与实践操作的研究取得了很大成绩。

1. 扩大了教育评价方面的对外交流活动

由于评价研究和实践的需要，国外的研究成果逐渐被介绍进来。如 1983 年，加拿大派专家来华东师范大学作“教育评价”的专题学术报告，促进了我国教育评价研究的开展；同年，教育部邀请了国际教育成就评价委员会（简称 IEA）前任主席、瑞典斯德哥尔摩大学教授胡森及教育评价专家海德曼等人来北京讲学；1986 年，美国教育评价专家布卢姆教授来华东师范大学举办“教育评价专题学术报告会”；1987 年在北京大学召开了第一次中美教育评估研讨会，讨论了教育评价的概念、意义，并着重讨论了课程评估、学生学习质量评估等；1989 年又召开了第二次中美教育评估研讨会，讨论了教育评估如何发挥促进教育质量提高和学校管理的作用。这些年来，北京、上海、福建等地的有关高校曾多次举办了关于教育评价的理论和技術、测量与统计等讲习班以及研讨会。1984 年我国正式参加了 IEA 组织，同时，教育部确定河北、山西、北京、天津等省市参加 IEA 组织实施的第二次自然科学学科的教育成就评价研究活动。

这一时期，教育评价的研究与国内外交流进一步加强，全方位、多层次地开展了各种类型的教育评价实践活动，对教育评价