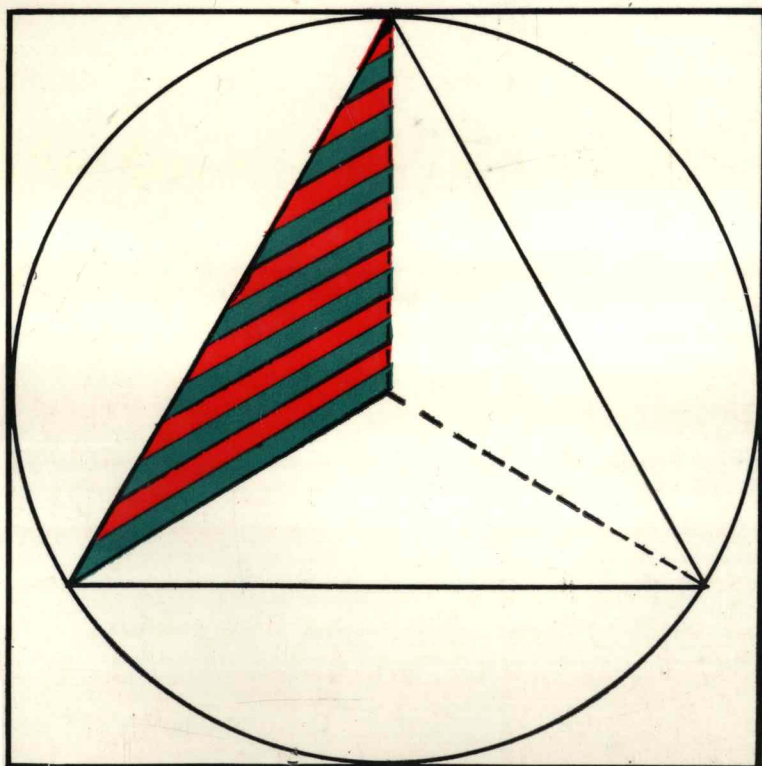


# 數學教育測量與評價

SHU XUE JIAO YU CE LIANG YU  
PING JIA

汪德營主編

數學教育叢書



南海出版公司

# 数学教育测量与评价

汪德营 李玉琪主编

南海出版公司

1992.年·海口

# 数学教育测量与评价

---

主 编:江德营

---

装帧设计:英华

---

南海出版公司出版发行  
山东省惠民县印刷厂印刷

---

850×1168毫米 32开本 8印张 198千字  
1992年5月第1版 1992年5月第1次印刷  
印数:1—4100

---

ISBN7—80570—682—4/G·190

---

定价:4.00元

## 序

当今的世界，已进入第三次科学技术革命的新时代。知识激增、信息高度发达，已成为我们这个伟大时代的重要特征。随着科学技术的飞速发展和现代科学的数学化趋势的增强，社会对人才智能提出了更高的要求，也引起了数学教育思想的深刻变革，促进了数学教育的发展。

我国数学教育家曹才翰先生指出：“在国际、国内的教育领域中，数学教育始终是最活跃的一个学科。”国际上，从六十年代的“新数学”、七十年代的“回到基础”，到八十年代的“问题解决”，数学教育的发展一日千里；在国内，学术组织林立，专业会议频繁，数学教育的研究呈现出一派生意盎然的兴旺景象。

数学教育学是研究数学教育规律的科学，是教育学、心理学、数学、哲学、思维学和数学史等科学的交叉学科。近年来，数学教育学已初步形成为包括教学论、学习论、课程论、方法论等多个分支学科的庞大学科群。

从1987年起，我们山东省高等师范院校数学教育研究会建立了教学论、方法论、学习论、思维论和数学史等专题组，开展了多种形式的学术研究和学术交流活动，在国内各种刊物上发表了一批论文，并出版了研究会论文集。在深入开展学术研究的同时，我省

各高等师范院校相继开设了有关必修和选修课,促进了我省数学教育的发展。

为了适应数学教育的发展,满足高等学校和中学教师继续教育的需要,为了反映数学教育的新思想、新观点,我们研究会组织编写了《数学教育丛书》,这是我省从事数学教育研究和教学的同志们的一项研究成果。

数学教育学是一门新兴科学,正处在建设和发展阶段,它的理论建构和学科体系还有待于深入研究。因此,编写这套丛书只是对建立数学教育学的一种尝试。我们期望得到读者的批评和建议,以便进一步修改和完善。

山东省高等师范院校  
数学教育研究会理事长

汪德营

1990年12月

# 前 言

教育测量与教育评价是现代教育科学研究的重要领域,数学教育测量与评价是其中的一个分支。近年来,随着我国数学教育的蓬勃发展,数学教育测量与评价的理论在改革教育思想、教学内容和教学方法等方面的重要作用正日益展现出来,它的发展已受到人们的广泛关注。进一步研究它的理论和方法,对于促进数学教育的发展和提高数学教学质量无疑具有重大意义。

数学教育测量与评价是一门新兴学科,它的理论体系尚处在不断发展和完善之中,许多问题还有待于进一步研究和探讨。我们编著这本书,正是适应这种形势发展的需要,希望能对这门学科的发展起到一定的促进作用。在本书第一到第四章中,主要研究数学教育测量与评价的基本理论和方法;第五、六、七章专题讨论数学学习评价、数学课堂教学评价和数学教材评价等问题;在第八章中,简要介绍了数学教育评价中数据处理的微机软件系统。本书可作为高等师范院校和中学教师继续教育有关课程的教材,也可以供广大数学工作者、数学教学研究人员和数学教师参考。

本书编写分工是:汪德营(第一章),姜玉武(第二章),郑强(第三章),马希林、郑强、田建国(第四章),李玉琪(第五章与第七章),卞惠林、于骏(第六章),赵岗(第八章)。

此外,徐云、侯兆群、史庆祝、刘培进、崔德悦、伊萍和朱习军等

同志也参加了部分稿件的撰写。全书由汪德营、李玉琪同志设计并修改定稿。

在本书编写过程中,我们参阅了曹才翰、孙瑞清等许多名家的著作,在此谨向他们表示最诚挚的谢意。

由于水平有限、时间仓促,书中缺点错误在所难免,恳请广大读者批评指正。

编 者

1992年2月于济南

# 目 录

<b>第一章</b>	数学教育测量与评价概述	(1)
§ 1.1	数学教育测量评价及其特点	(1)
§ 1.2	数学教育测量评价的产生与发展	(6)
<b>第二章</b>	数学教育测量与评价的作用及原则	(16)
§ 2.1	数学教育测量与评价的作用	(16)
§ 2.2	数学教育测量与评价的原则	(19)
<b>第三章</b>	数学教育评价的方法	(29)
§ 3.1	数学教育评价的一般程序	(29)
§ 3.2	定性评价与定量评价	(46)
<b>第四章</b>	评价信息的统计检验与分析	(53)
§ 4.1	统计检验	(53)
§ 4.2	相关分析与回归分析	(72)
§ 4.3	灰色系统关联度分析	(88)
<b>第五章</b>	数学学习评价	(95)
§ 5.1	数学学习评价概述	(95)
§ 5.2	数学学习评价的指标体系	(102)
§ 5.3	学习成就测验	(113)
§ 5.4	数学学习的评价方法	(121)
<b>第六章</b>	数学课堂教学评价	(136)
§ 6.1	数学课堂教学评价的意义	(136)
§ 6.2	数学课堂教学评价的设计与分析	(142)
<b>第七章</b>	数学教材评价	(176)
§ 7.1	数学教材评价的意义和原则	(176)
§ 7.2	数学教材评价的指标体系和评价程序	(178)



**第八章 数据处理的微机软件系统**..... (183)

    § 8.1 应用程序概述 ..... (183)

    § 8.2 程序功能图及使用说明 ..... (184)

    § 8.3 源程序及打印格式 ..... (194)

**附表**..... (218)

**参考书目**..... (243)

# 第一章 数学教育测量与评价概述

## § 1·1 数学教育测量评价及其特点

### 一、数学教育测量及其特点

要把教育的效果、效率像对物理现象的测量一样,以数量化的形式表示出来,就要进行教育测量。教育测量是对教育现象进行定量测定的一门教育科学,主要研究对教育或训练效果进行测量的原理和方法。一般性的教育测量是指对学习结果——知识、技能的测量。教育测量起源于中国古代的考试,但那时的考试还不是科学的教育测量。本世纪初,美国心理学家桑代克等人把心理统计与测量的基本原理、方法用于教育,才使教育测量走上科学化的道路。目前,教育测量在人才选拔、教育评价、教育研究等方面发挥着重要的作用。

数学教育测量是教育测量的一个分支,是以教育学、心理学、测量学的理论为基础,研究如何根据数学教育目标和教学计划的要求,运用各种测量手段和统计方法,按照一定的规则,用数字对数学教学效果和学生知识和数学能力加以描述和确定的过程。

教育的目的在于引起人的行为、知识、能力的变化,而人的知识、能力是通过学习、实践获得的,由于个别差异的存在,同样的学习和实践,各人获得知识的多少、能力的强弱会有各种差异,而数学教育测量就是要确定学生的数学知识和数学能力增长变化的情况,表明其现状及差异等,并把这些结果用数量描述出来。要完成这一测量任务,必须采取一定的手段,并按一定的规则进行。按照教育目标分类学理论编制的各种测量,就是测量人的知识、能力的最基本的手段,而按标准化考试的规则编制和实施的测验,可以使这种手段更加完善、结果更加科学。因此,数学教育测量主要是研

究考试的学问,它主要以学生的数学学业成绩的考查与评定作为自己的研究对象。数学教育测量研究的主要内容是各种数学成绩的测验,也就是测量学生所学数学知识与技能的系统方法,其具体内容包括各类数学试卷的设计与编制、试题与试卷的分析及评价、测验的实施、评分及分数的合成等。

由于教育现象的测量和物理测量相比,不但困难多,而且精度差,因而数学教育测量具有其自身的特点。

### 1. 数学教育测量方法的间接性

物理测量大都是直接测量,但对学习的效果、学生的行为变化的测量却无法用任何一种量具直接获得,而必须用间接的方法,即通过测量学生对所学知识的掌握程度来了解学生的水平、能力等。数学教育测量主要通过学生回答编制好的反映教育目标的数学题目来实现对学生的数学知识、能力的测定,由学生正确回答数学题目的数量或观察其操作的行为来确定其数学知识和技能水平的高低。在实际测量中,无论考试题目编拟的质量和数量多么适当和科学,操作实践多么似真和全面,也不可能包括测量目标的各个方面。所以,学生对测量目标的达到程度或对数学知识的掌握情况、能力水平等,只能通过有限的测试题目的回答情况进行推测,这就表明了数学教育测量是间接测量,所得到的数据是一种近似值。

### 2. 数学教育测量因素的不稳定性

物理测量的对象是物,而且一般来说被测物体本身的变化是微小缓慢的,因而测量结果比较稳定。但是,数学教育测量的对象是人,影响学生知识、能力变化的因素比较多,如不同的时间、不同的环境、不同的条件及学生本身的思想、情绪、身体等状况,都可能引起其知识、能力的变化,这就给测量带来了一定的困难。另外,数学教育测量多是通过考试实施的,而考试内容是由教学内容和教学目标决定的,因此,数学教育测量的因素多而且不够稳定。

### 3. 数学教育测量结果的相对性

物理测量的结果有相对与绝对之分。但是在数学教育测量中，由于无法找到学生知识、能力的“绝对零点”，因而测量的结果都是相对的，学生的考试成绩只有在某种标准的比较中才有意义。例如，一名学生的数学成绩是90分，物理成绩是85分，孤立地比较两个分数是无法确定其意义的，只有把它们放到各学科的群体中去参与比较，才能确定分数的意义。

## 二、数学教育评价及其特点

当前，世界上教育科学的研究包括教育基础理论研究、教育发展研究和教育评价研究三个方面，其中教育评价是一个重要的内容，而且随着科学技术的发展，教育评价会越来越重要。那么，什么是教育评价呢？所谓教育评价，就是对照数学教育目标提出的评价标准，运用科学的评价方法，系统地收集信息、对教育过程及教育效果作出价值判断，并把判断的结果反馈于教育实践，为教育决策提供科学依据的过程。

数学教育评价是教育评价的一个分支，它是对照数学教育目标提出的评价标准，运用系统科学和统计方法系统收集信息，对数学教育过程及教学效果作出价值判断，并把判断的结果反馈于数学教育实践，为提高数学教育质量提供科学依据的过程。

数学教育评价的实施一般分三个阶段：一是确定评价对象；二是收集分析各种有关信息；三是做出评价结论，并将结果反馈给被评价对象。数学教育评价的内容有数学教学评价、数学学习评价、数学教材评价等；评价中要收集的信息主要是评价对象的目标、计划、实现的过程、结果和效益；使用的方法有测验法、调查法、统计法等；评价种类按目的不同可分为诊断性评价、形成性评价、总结性评价，按评价者的不同可分为自我评价、同行评价等。

数学教育评价有以下特征：

1. 数学教育评价是一个比较复杂的“过程”

数学教育评价比传统的考试方法科学得多，它避免了很多入

为主观因素的干扰,得出的结论是客观的、真实的,它追求的是确定教育目标的主观要求和客观要求的统一、评价技术的科学性和教育性的统一。这种过程中包括一系列的步骤和方法都是科学和可行的,从而向人们揭示出哪些做法是符合教育规律的,以便继续发扬。

### 2. 数学教育评价的核心问题在于“价值判断”

美国教育评价专家斯克里芬认为教育评价有两大功能:一是形成性功能,是指对仍在发展中或未定型的教育或教学方案进行价值性判断,它要求评价者处于“参与者”的地位搜集资料,并对某一教育或教学活动的各方面给予价值性判断;二是概括性(又称总结性)功能,是指对已经完成的教育方案进行价值性判断,它要求评价者处于公正的地位搜集资料,对某数学教育活动给予价值性判断。价值性判断的目的不仅要对数学教育或数学教学的效果作出成功或失败的判断,更重要的是使评价的结果有助于提高数学教育的质量。

### 3. 数学教育评价的直接对象是数学教育过程及其效果

教育目标是教育的核心问题,这个目标又是依赖若干个子任务的完成而共同实现的。数学教育评价以数学教育目标为总目标,通过教育评价不断修正和调整教育过程,促使整个教育过程逐步逼近教育目标。

## 三、数学教育测量与评价的联系及区别

数学教育测量是对学生的数学学业成绩、学习能力、兴趣爱好等进行数量化的测定,而数学教育评价则是对学生的数学学习成绩和发展水平进行价值判断的过程。因此,数学教育测量与数学教育评价二者既有联系又有区别。

### 1. 数学教育测量与数学教育评价的联系

一方面,数学教育测量是数学教育评价的直接基础。数学教育测量通过考试来测定学生的学业成绩,为认识教学过程中的各种

因素提供了客观的依据,提供了数学反馈信息。例如测量学生智力水平的高低,测量学生的态度、性格、心理品质等,这些测量都具有客观性和标准化的特点,在此基础上可以进行分析、鉴别,从而做出价值判断,进而提出改进和发展的意见及建议。由此可知,数学教育测量的结果是进行数学教育评价的重要依据和直接的基础,如果没有数学教育测量所提供的资料,数学教育评价是无法进行的。

另一方面,数学教育测量的结果只有通过数学教育评价才能获得实际的意义,否则便成为一堆抽象枯燥的数字。

## 2. 数学教育测量与评价的区别

数学教育测量强调数量化,要求把数学教育过程中的某些信息尽可能量化,以形成测量的数据和资料;而数学教育评价除包括这些测定的数据和资料外,重在事实的解释、诊断和价值判断,量与质并重,特别从质上考虑问题。譬如,某班甲、乙两名学生的期中考试数学成绩都是70分,从测量的角度看,它们是一样的,而从评价的角度看就可能大不相同,甲生由60分上升到70分并因概念不清失分而得70分;乙生则由80分下降到70分并因计算错误失分而得70分。因而判断甲生有进步或概念不清,乙生学习退步或计算不准确。由上看出测量的结果——分数是测量问题,而分析、判断甲、乙两名学生的知识、能力水平则是评价问题。

数学教育测量是一种单一的活动,只注意事实资料数据的获取,注意量的多少或所处的位置;而数学教育评价是一种系列的活动,它不仅注意数量,更注意对事实的分析、判断,主要在于提出改进和发展意见及建议,以促进教学工作、提高教学效果。

## § 1·2 数学教育测量评价的产生与发展

### 一、国外教育测量与评价的产生与形成

就世界范围考察,教育评价形成于美国,起源于教育测量。

## 1. 教育测量的产生

1879年德国心理学家冯特在莱比锡大学建立了世界上第一所心理实验室,进行了测量个别差异的尝试。美国的卡特尔在莱比锡大学的心理实验室里,学习了实验心理学、人类学和统计学,回国后也建立了心理实验室,对大学生进行了能力心理的测验,以实验的方法来研究个别差异。但卡特尔测验没有说明测验分数与学习成绩之间的关系,比奈——西蒙智力量表真正开创了智力测验的新时代。1904年卡特尔的学生、教育测量学的创始人之一桑代克发表的“心理与社会测量学导论”,成为测量史上划时代的巨著。桑代克在书中明确提出了“凡存在的东西都有数量,凡有数量的东西都可以测量”的著名的教育测量思想,并把心理统计和测量的基本原理及方法用于教育,这都标志着教育测量的产生。

## 2. 教育评价的形成

教育测量的形成和发展经历了较为漫长的过程。自从数理统计方法运用于考试分析而产生了教育测量学以来,人们比较多地把注意力集中在考试方法的研究上,发现口试不能对学生考同样的题目,失去了同一考核标准,而且题目也少;笔试题目虽然统一了标准,题目仍偏少。为了克服这二者的缺点,在美国逐渐形成了一个力求客观化的教育测量运动,大样本(一二百道题)的多选题盛行起来。由于片面强调客观标准化,导致学生侧重于机械记忆和以教科书为中心的学习方式,学生的兴趣、爱好、创造性得不到重视,因而出现了进步主义联盟(P. E. A)对传统教育的批评,并开始了以课程改革为中心的实验研究,这就是1933年到1940年的“八年研究”。这项工作由泰勒和史密斯组织,由七所大学和三十多所中学参加,成立了以泰勒为中心的教育评价委员会,他们运用教育学和心理学的思想、新方法,以全面发展学生的能力为目标,设计了一套教育评价方法。其内容不仅包括确定课程的教学目标以及实现该目标的条件,也包括满足客观性并具有一定信度和效度

的测量方法,这时教育评价体系已基本形成。

在泰勒之前,人们认为测量就是评价,两者无明显区别,测量的功能主要是判断学生个体之间的差异,为学校和社会选拔学生提供依据,而泰勒第一次将评价运用于教育过程,使评价成为课程改革的有机组成部分,测量仅是评价的一个步骤或工具。

“八年研究”的实践表明,教育评价有利于促进教育目标的实现,对于培养新型人才是有成效的。进行课程改革学校的毕业生,在学习兴趣、思维能力、适应能力等方面比未进行改革的学校好,这是教育评价效果的初步体现。1940年,教育评价委员会提出的“八年研究”报告中阐述的教育评价的原则和方法标志着教育评价理论和方法的形成。此后,教育评价在美国教育界中开始制度化。

1957年11月,苏联发射了第一颗人造地球卫星,而美国于1958年2月才发射出第一颗人造地球卫星,此事对美国刺激很大。追究其失去科技优势的原因,美国认为他们的基础教育水平低于苏联,美国的公立学校及其教育制度由此受到指责,并促使美国政府开展了大规模的数学课程改革,也促进了教育评价的进一步发展。为了与苏联竞争,美国政府又增拨教育补助款,进行深入的改革,美国参议院还决定申请补助款的学校事后要向政府提出经费使用情况及其成效的评价报告。同时,州、市政府也都开展了州、市性的教育评价活动,甚至各乡镇当局也仿效推行一种“教育绩效(成绩及效果)责任制”运动,纷纷要求对课程和学生学业成绩进行评价。这样,教育评价的活动在美国普遍开展起来,不仅教育评价的方法和理论在实践活动中趋于完善,而且传到日本、苏联及欧洲各国,特别是第二次大战后,许多国家竞相采用。

六十年代以来,教育评价得到了迅速发展,各种教育评价的理论和方法陆续出现。例如普罗维司的改革评价模式理论、库克的系统方法评价方案理论、泰勒与帕姆的标准参照测验理论等等,不仅教育评价成了推动教育改革的一种强有力手段,而且教育评价形



成了独立的学科体系,成为教育科学的重要领域之一。

日本在战后推行五级相对评价,评定学生的优劣不是看他们是否达到目标,而是看其在班级或学校中的相对地位(即名次),这种评价方式大大挫伤了师生的积极性,给日本教育及社会带来了严重的后果。1966年日本教育心理学家板仓圣宣提出了到达度评价,其核心是判定学生是否朝着教育目标的方向发展,以及实现教育目标的程度,而不再强调五级相对评价的筛选功能。

教育评价在西欧各国也普遍开展起来。由于教育的发展,许多新课程、新教法不断创新,因而必须有新的评价方法来判断各种教育活动的价值。此外,现代教育制度具有政治、经济、社会的多种功能,更需要客观地、科学地评价教育整体,才能决定政府的预算和教育经费的分配,因而教育评价已成为教育事业中不可缺少的一环。

1975年5月27日在瑞士的日内瓦成立了国际教育评价协会(International Association for Educational Assessment,简称IAEA),常设机构设在巴西的波茨瓦纳城。会议接纳了62个国家和地区的教育评价机构和个人为会员,规定每年一次年会,现已召开了十七次。我国于1983年参加该协会。

IAEA的宗旨是促进和协调各国教育评价活动的发展,在不影响各自教育和文化自主权的前提下,进行国际间的教育研究和评价工作,合理使用教育评价技术,以求提高教育质量。该协会已对毕业考试与选择、教师作用评价、职业教育评价、教育成果评价、学校成果评价、标准化考试、教育与职业发展的评价等议题进行了讨论,还编辑出版了《教师评价指南》期刊,教育评价现在已成为国际上评价一国教育发展水平的重要内容之一。

教育评价当前已受到各国政府及其教育主管部门的高度重视,例如美国联邦政府教育部于1984年12月20日公布的高等教育问题报告中指出:“要设计并实行科学评价知识、能力、技能的系