



数学教育测量与评价九讲

◇ 谢圣英 编著 ◇

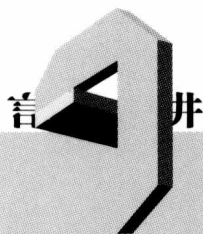
◇ 湖南师范大学出版社 ◇

2015 年度湖南师范大学出版基金资助出版

2014 年度湖南师范大学教学改革研究项目《数学教育类课程的
“参与式”教学研究与实践》成果（课题编号 121-0753）

数学教育测量与评价九讲

谢圣英 编著



湖南师范大学出版社

图书在版编目 (CIP) 数据

数学教育测量与评价九讲 / 谢圣英编著. —长沙: 湖南师范大学出版社, 2015. 9

ISBN 978-7-5648-2270-5

I. ①数… II. ①谢… III. ①数学教学—教育测验②数学教学—教学评估 IV. ①O1-4

中国版本图书馆 CIP 数据核字 (2015) 第 223843 号

数学教育测量与评价九讲

谢圣英 编著

组稿编辑 | 周基东

责任编辑 | 廖小刚

责任校对 | 施 游

出版发行 | 湖南师范大学出版社

地址: 长沙市岳麓山 邮编: 410081

电话: 0731. 88873070 88873071

传真: 0731. 88872636

网址: www.hunnu.edu.cn/press

经 销 | 湖南省新华书店

印 刷 | 湖南雅嘉彩色印刷有限公司

开 本 | 710 mm×1000 mm 1/16

印 张 | 13

字 数 | 220 千字

版 次 | 2015 年 9 月第 1 版第 1 次印刷

书 号 | ISBN 978-7-5648-2270-5

定 价 | 35.00 元

目 录

第一讲 数学教育测量与评价概述	(001)
一、数学教育测量与评价的涵义	(002)
二、数学教育测量与评价的主要内容体系	(004)
三、现代数学教育测量与评价的发展趋势	(007)
四、数学教育测量与评价的学习意义	(009)
第二讲 数学教育测量与评价的测量理论基础	(012)
一、教育测量的基本概念、要素及量表类型	(013)
二、经典测量理论 (CTT) 介绍	(015)
三、项目反应理论 (IRT) 介绍	(022)
第三讲 数学教育测量与评价的统计学基础	(040)
一、统计学概述	(041)
二、统计表与统计图	(043)
三、统计量	(046)
四、SPSS 软件的应用与教育统计分析初步	(057)
第四讲 学生数学学习的测量与评价设计	(075)
一、数学学习测量与评价的基本思想	(076)
二、数学学习测量与评价的目的	(077)
三、各类数学学习结果的测量与评价	(080)
四、学生数学学习的测量与评价的程序	(084)
第五讲 学生数学学习的测量与评价技术	(085)
一、学生数学学习测量与评价的理念	(086)
二、传统学生数学学习测量与评价的反思	(089)
三、如何命制数学试卷	(096)
四、不同题型数学题的命题要求	(098)

五、试题改编的一般方法与常见模式	(099)
六、试卷分析报告的基本框架及质量控制指标	(102)
第六讲 数学教师评价	(107)
一、数学教师评价概述	(108)
二、数学教师评价模式	(113)
三、数学教师评价的基本方法	(115)
第七讲 数学课堂教学的测量与评价	(120)
一、数学课堂教学测量与评价的相关理论	(121)
二、我国数学课堂教学评价现状与反思	(124)
三、数学课堂教学测量与评价的趋势	(127)
四、数学课堂教学评价的方法	(129)
第八讲 数学教育评价的心理及其调控	(139)
一、教育评价心理概述	(140)
二、评价者心理与调控	(142)
三、被评价者心理与调控	(147)
第九讲 国际数学教育测量与评价	(149)
一、PISA 项目简介	(150)
二、TIMSS 项目简介	(158)
三、数学教育的 PISA 与 TIMSS 比较	(166)
附录 常用统计表	(174)
附表 1 标准正态分布表	(175)
附表 2 标准正态分布分位数表	(178)
附表 3 t 分布分位数表	(180)
附表 4 χ^2 分布分位数表	(184)
附表 5 F 分布分位数表	(186)
参考文献	(198)
后记	(201)

-----◇第一讲◇-----

数学教育测量与评价概述

数学教育测量与评价是数学教育改革与发展的风向标和助推器，它贯穿、引领和改造着我们的教育实践。数学教育测量与评价不同于数理统计、模糊数学等数学课程，它阐述了同教育测量与评价相关的数学应用理论及方法；它又不同于教育系和心理系的教育测量与评价，因为它着重论述的是同数学教育有关的测量与评价，所以它既体现了数学学科特点，又体现了师范性特色。

一、数学教育测量与评价的涵义

为了正确及更好、更有效地使用数学教育测量与评价的概念，首先，我们有必要对数学教育测量与评价的相关概念进行辨析和必要阐明。

（一）测量与评价

史蒂文斯（Stevens, S. S.）认为测量是用一定规则给实物属性指派数字或符号的过程。测量的这一定义包含三个元素：事物及属性、法则、数字或符号。任何领域内的测量都应包含上述三个元素。^① 测量是以追求客观的量化结果为目的的。^②

评价活动最早开始于中国的古代社会，用于官员的测试，从而决定其提升或降职，主要是几个领域的书面考试。评价从辞源上可以解释为“评定价值”^③，“是一种价值判断的活动，是对课题满足主题需要程度的判断。”^④

测量和评价二者的区别比较显著。测量是对事物数量特征的获得，强调数量化的方法与结果的过程；测量强调事物的客观性；测量的任务是对事物的量的认识。评价则是对教育现象客体的价值进行判断的过程，它强调定性与定量相结合的方法，具有客观性与主观性的两种属性；评价本质上是人类的一种认识活动；^⑤ 评价是人的意识对实践活动及其结果的综合反映。

（二）教育测量与教育评价

随着科技的进步与发展，测量与评价不仅仅局限于较为精准的自然科学领域，其工具和方法经改良后运用并服务于教育领域。从广义上来讲，教育测量是依据一定的法则（标准）用数值来描述教育领域内事物的属性，是事

① 张敏强. 教育测量学 [M]. 北京: 人民教育出版社, 1998.

② 马云鹏. 数学教育测量与评价 [M]. 北京: 北京师范大学出版社, 2009.

③ 中国社会科学院语言研究所词典编辑室. 现代汉语词典 [M]. 北京: 商务印书馆, 1978.

④ 陈玉琨. 教育评价学 [M]. 北京: 人民教育出版社, 1999.

⑤ 冯平. 评价论 [M]. 北京: 东方出版社, 1995.

实判断的过程。^①从狭义上来讲,它是针对学校教育影响下学生的学业、智能及其他有关教育的心理属性等各方面从量的规定性上予以确定和描述的过程。^②比如学生的学业成就、学习能力水平等。测验则是教育测量的主要工具。教育测量为教育评价提供客观的数据信息,是教育评价的基础。

教育评价的概念最早出现于20世纪三四十年代,在美国教育家泰勒主持的“八年研究”中被证实提出。泰勒认为“评价在本质上是一个确定课程与教学计划实际达到教育目标的程度的过程”。克龙巴赫把评价广义地定义为“为获取教育活动的决策资料,对参与教育活动的各个部分的状态、机能、成果等情报进行收集、整理和提供的过程。”^③斯塔弗尔比姆在1969年提出,评价是“为决策提供有用信息的过程”。这些定义都是从评价的功能上作出的,认为评价是为了更好的决策搜集和使用信息的过程。陈玉琨认为“教育评价是对教育活动满足社会与个体需要的程度作出判断的活动,是对教育活动现实的(已经取得的)或潜在的(还未取得,但有可能取得的)价值作出判断,以期达到教育价值增值的过程。”^④

(三) 教育评估与教育评价

教育评估与教育评价是相近概念,在很多场合下可通用。二者的区别在于:评估有估价的含义,模糊粗略;评价相对是较为精确的价值判断;评估多用于群体或单位的状态及效果的估价,如本科教学水平评估;评价则既有对群体和单位的价值判断,又有对个人的判断。

(四) 教育统计与教育评价

教育统计学将数理统计学的理论和方法运用于教育领域,从数量的角度研究教育现象的数量特征及其关系的一门应用学科。它的主要任务是研究如何搜集、整理、分析由教育调查和教育实验等途径所获得的数字资料,并以此为依据,进行科学推断,从而揭示蕴含在教育现象中的客观规律。^⑤教育评价中获得的大量评价信息要靠教育统计提供的方法进行处理。教育统计为

① 金婣,王钢.教育评价与测量[M].北京:教育科学出版社,2007.

② 黄光扬.教育测量与评价[M].上海:华东师范大学出版社,2002.

③ 永野重史.教育评价论[M].东京:东京大学出版社,1986.

④ 陈玉琨.教育评价学[M].北京:人民教育出版社,1999.

⑤ 王孝玲.教育统计学[M].上海:华东师范大学出版社,2001.

教育评价提供数据处理的工具，也为评价结果的质量提供保证。需要指出的是，教育评价内容的确定、对象范围的确定、评价结论的得出与分析，却不是教育统计学的研究任务，这些问题还要依靠教育专业知识来解决。

（五）数学教育测量与评价

数学教育测量是依据一定的原理和法则，用数值来描述数学教育领域内事物的属性（数学课程设置、数学教学效果和学生的数学知识、数学能力）并进行事实判断的过程。数学教育测量同一般的测量一样，同样包含三个元素：事物及属性、法则、数字或符号。不仅如此，数学教育测量还是一种间接性的测量。^① 它的测量对象是外显行为、外在表现和特征，如数学成绩，通过测量外显性测量我们可以得到内隐的潜在特质水平，如数学能力。数学教育评价是“全面搜集和处理数学课程、教学设计和实施过程中的信息，由此作出价值判断、改进数学教育决策，从而促进学生发展的过程。”^②

数学教育测量与数学教育评价的联系与区别一如测量与评价的联系与区别。数学教育测量以数学教育评价为目的，数学教育评价以数学教育测量为基础，数学教育测量为数学教育评价提供依据，是数学教育评价信息的主要来源。数学教育测量的结果也只有通过数学教育评价才能获得实际意义，成为改进数学教育教学的依据和参考，否则就是一堆枯燥抽象的数值而已。

二、数学教育测量与评价的主要内容体系

数学教育测量与评价的对象最终是指向学生。学生的数学学习结果和数学能力发展状况是数学教育测量与评价的首要任务。建立科学合理的数学教育测量与评价体系是保证测量评价信息全面、客观的重要因素。如果按照测量与评价的对象来分，数学教育测量与评价主要包含数学课程测量与评价、数学教学测量与评价两大块。

① 马云鹏. 数学教育测量与评价 [M]. 北京: 北京师范大学出版社, 2009.

② 马云鹏, 张春莉. 数学教育评价 [M]. 北京: 高等教育出版社, 2003.

（一）数学课程测量与评价

数学课程测量与评价按对象可以分为对数学教材评价和数学课程标准评价。正如陈玉琨、沈玉顺等指出的，课程是实现教育目标的基本途径，保证教育质量的关键就是要保证课程质量。^①高凌飏主张建立建立专业中介机构并通过它构建起全国性的评价网络^②，并进一步提出教材评价的具体任务包括：对教材的文本进行静态分析和评价；对教材使用情况进行动态分析和评价；向教材评审委员会报告搜集到的静态和动态资料以及分析结果；向学生、教师、教研员、家长等报告教材的特点、质量水平和试用情况；向教材编者反馈修改意见；开展教材研究。^③数学课程标准的评价相关研究已有不少，其中主要基于对课程标准“官方文本”的解读^④和反思^{⑤⑥⑦⑧}，或者对我国数学课程标准的纵向比较研究^⑨，或者进行不同国家间数学课程标准的横向比较研究。^{⑩⑪}

（二）数学教学测量与评价

数学教学测量与评价按对象可以分为学生评价、教师评价和学校评价。其中，学生评价主要指学生数学学习测量与评价，教师评价则主要指数学课程教学测量与评价。

学生数学学习测量与评价的目的是全面了解学生的数学学习情况，激励

① 陈玉琨，沈玉顺，代蕊华，等．课程改革与课程评价[M]．北京：教育科学出版社，2001．

② 高凌飏．基础教育教科书评价体系的构建问题[J]．华南师范大学学报（社会科学版），2002（6）．

③ 高凌飏．基础教育教科书评价：理论与工具[M]．北京：人民教育出版社，2002．

④ 郑毓信．《数学课程标准（2011）》的“另类解读”[J]．数学教育学报，2013（1）：1-7．

⑤ 何小亚．全日制义务教育阶段数学课程标准（实验稿）刍议[J]．数学教育学报，2003（1）：45-49．

⑥ 张奠宙．对《全日制义务教育数学课程标准》理念部分的意见[J]．数学通报，2005（12）：1-4．

⑦ 史宁中．《数学课程标准》的若干思考[J]．数学通报，2007（5）：1-5．

⑧ 孙卫红，陈朝东．中国数学课程标准对估算要求的变化探析[J]．数学教育学报，2013（5）：27-31．

⑨ 孙名符，谢海燕．新高中数学课程标准与原教学大纲的比较研究[J]．数学教育学报，2004（1）：63-66．

⑩ 徐文彬，杨玉东．英国国家数学课程标准的确立与变革及其启示——中英国家数学课程标准比较研究之一[J]．数学教育学报，2002（3）：48-51．

⑪ 刘长明，孙连举．中美两国数学课程标准中初中学段“空间与图形”领域的内容标准之比较[J]．数学教育学报，2002（4）：49-52．

学生的学习热情,促进学生的全面发展。目前,在义务教育阶段,对学生数学学习的测量与评价主要侧重以下四个方面:

- (1) 数学知识和技能。
- (2) 数学思考。
- (3) 数学问题解决。
- (4) 数学情感态度。

数学课堂教学测量与评价对于改进数学教学特别重要。数学课堂教学测量与评价要立足于促进教师和学生的发展,通过多种测量与评价方式,获得真实的测量与评价信息。测量与评价的目标主要包括以下几个方面:^①

- (1) 课堂教学目标是否明确、适当,是否遵循课程标准和教学大纲的要求,并能够根据实际需要作出适当的调整。
- (2) 教学目标是否关注学生的全面发展。
- (3) 教学内容是否围绕教学目标的选取,并契合学生的承受能力和发展需求。
- (4) 教学方法的选择是否遵循教学内容与学生实际的要求,并能够提高教学效率和学生学习兴趣。
- (5) 学生的参与度和参与面是否足够深广。

此外,数学教学测量与评价还应包括数学教育环境和条件的测量与评价^②,这实际是属于学校评价的范畴。

综上,数学教育测量与评价的主要内容体系,我们可以用如下框架图来表示(如图1-1所示)。

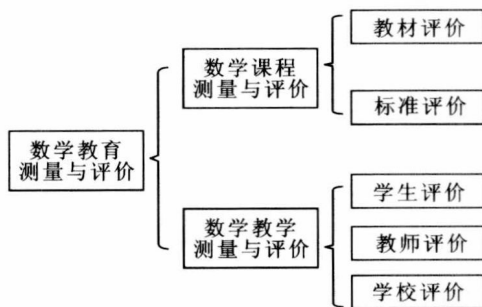


图1-1 数学教育测量与评价的主要内容体系(按测量与被评价者分)

① 马云鹏. 数学教育测量与评价[M]. 北京: 北京师范大学出版社, 2009.

② 傅海伦. 数学教育发展概论[M]. 北京: 科学出版社, 2001.

三、现代数学教育测量与评价的发展趋势

数学教育测量与评价是历史性与时代感的统一。理解数学教育测量与评价事件、构建教育测量与评价的意义既需要具有历史视野，又需要把握时代精神的发展趋势。

（一）从相对评价转向绝对评价及个体内差异评价，并使三者有机结合

相对评价是从被评价者群体之中，选取一个或若干个对象作为基准，将余者与基准做比较，排出名次、比较优劣的评价法。相对评价主要用于学业成绩评价，一般采用百分制或等级制计分。相对评价便于学生在相互比较中判断自己的位置，激发竞争意识。

绝对评价是在被评价者群体之外，预定一个客观的或者理想的标准，并运用这个固定的标准去评价每个对象，主要用于评价既定学习目标达成情况。这种评价的标准不受被评价者所在团体状况的影响，评价结果只与被评对象自身的水平相关。

个体内差异评价也就是自我评价，是把每个被评价者的过去与现在进行比较，或者把个人有关侧面进行横向比较。个体内差异评价法比较充分地照顾到学生的个性差异，力图减轻被评价者的压力。但是，它只是使被评价者与自身状况进行比较，既不是按照一定客观标准进行评价，亦无被评价者间的相互衡量，容易导致信度降低，学生自我满足，因此常与绝对评价、相对评价结合使用。

总之，现代数学教育测量与评价的发展趋势之一就是相对评价转向绝对评价及个体内差异评价，并使三者有机结合。在实际的数学教育测量与评价中，往往是不同形式的评价方法综合使用，从而提供更全面的评价信息。数学教育测量与评价由单一注重基础知识与基本技能的评测转向知识与技能、数学思考、数学问题解决、情感与态度的结合。更多地开始关注培养学生积极的情绪情感体验和良好的心理品质；注重对个体发展独特性的认可，帮助

学生接纳自己，拥有自尊自信。

（二）由重结果评价向重过程评价和结果评价相统一，也即从重鉴定功能转到重改进功能的发挥

重结果评价主要有终结性评价。终结性评价也可称为总结性评价。它是某一阶段的教与学结束后为检验效果而进行的评价，一般在学期、学年或某一门课程结束之后进行，比如期末考试、结业考试、毕业考试、升学考试等都属于这一类型，它们主要用于不同层次的教育决策，对于改进教与学没有太大的意义，如是否毕业、是否给予升学资格等。

重过程评价主要指形成性评价。形成性评价是由斯克利芬（G. F. Scriven）提出来的，是指在教育活动运行过程之中所进行的阶段性、过程性的评价，目的在于了解教育过程中存在的问题和改进的方向，及时修正或调整计划。布鲁姆第一次将形成性评价运用于教学活动中，深入分析形成性评价在改进学生学业成绩方面的价值。数学教育不仅要使学生掌握数学基础知识与基本技能，同时还要学会数学思考，去发现、解决生活中的问题。形成性评价的更重要的作用就是能通过多种评价方式去促进学生在数学学习过程中不断取得各方面的进步。

传统的数学教育测量与评价主要目的在于甄别和选拔，尤其是各个阶段的总结性评价，将学生按考试成绩分等，张榜公之于众，甚至一考定终身。于学生而言，考得好就得意洋洋，考得不好就垂头丧气，并由此一步步地失去学习数学的信心和乐趣。于教师而言，测量与评价的结果直接与奖惩晋升挂钩，使得教师为学生的考试成绩忧心忡忡、神经紧绷，教师压力之大，学生夜以继日的题海大战也就可想而知。随着人们对教育问题研究和认识的深化，由重结果评价向重过程评价和结果评价相统一，这是现代数学教育测量与评价的另一必然发展趋势。这本质上是将测量评价的功能从侧重甄别转向侧重发展，也即从重鉴定功能转到重改进功能的发挥。

（三）更加重视元评价，注重评价过程的反馈和监督，强调测量与评价方法的完善和改进，突出评价结果的全面解释和慎重处理

现代数学教育测量与评价的发展趋势还表现在更加注重元评价，即对评价本身的评价和反思。其中主要包括对评价过程的反馈监督，评价方法的完

善和改进，评价结果的解释处理。

除高等教育的其他教育领域，我国目前教育评价实践中普遍比较缺乏系统的元评价内容和规范的元评价过程，尤其像中高考这类大规模的正式教育测量与评价中，尚没有特别规范的元评价过程。同时，如果我们分析元评价的相关研究著作也不难发现，我国研究者对某类教育评价（如高效课堂教学质量）展开的分析，并不是对某一具体的教育评价实践活动进行评价，而事实上却属于元评价的研究范畴，而不是元评价的实践活动。^① 现代数学教育测量与评价的发展趋势将引领我们对评价过程加强反馈和监督，对测量与评价方法的改进侧重定量与定性相结合、实证与人文相结合，并注重引入新的测量与评价方法。比如，台湾地区评价方法中除纸笔测验外，还配合单元目标采用实测、讨论、口头回答、观察、作业或专题研究等方法。数学测验本身也在发生改变，测验逐渐减少了选择题、是非题、填空题等“标准化试题”的使用，取而代之的是要求学生表述自我思维过程与认知特征的问题，尤为突出的是使用“非规范”型试题；测评中所使用的问题不再仅仅是纯数学题、“人造”的应用题，还包括许多具有实体现实背景的现实问题（背景资料多选自报刊、书籍等真实材料）。此外，对评价结果的解释和使用将更加全面慎重，特别是某些高风险测量和评价的结果。

四、数学教育测量与评价的学习意义

第一，学习数学教育测量与评价，可以帮助我们理解教育测量与评价的基本原理和价值取向。

教育测量与评价不只是一个对教育进行价值判断的技术程序，它在本质上是一个哲学问题。任何教育测量与评价技术都不是普遍有效的，任何教育测量与评价都体现了特定的教育价值观，尽管教育测量评价者未必对此有清醒的自觉。教育价值观的根本问题是：通过教育培养什么样的人？对“培养

^① 霍国强. 我国教育元评价的实践缺失及对策思考 [J]. 教育发展研究, 2012 (22): 21 - 25.

什么样的人”的回答受三种基本理性的支配，即“技术理性”、“实践理性”与“解放理性”。不同理性支配下，教育评价的内涵产生本质区别。^①“技术理性”是通过合规律（规则）的行为而对环境加以控制的人类的基本理性。假如教育评价受“技术理性”的支配，则必然把评价的目的视为对教育实践的有效控制与管理。“实践理性”是建立在对意义的“一致性理解”的基础上，通过与环境的相互作用而理解环境的人类基本理性。假如教育评价受“实践理性”的支配，则必然把评价的目的视为对教育实践的理解，强调评价者与被评价者的交互作用，强调对过程的关注。“解放理性”是人类指向“解放”和“权力赋予”的基本理性。假如教育受“解放理性”的支配，它必然把评价的过程视为评价者和被评价者通过协商而进行意义建构的过程，评价的根本目的是解放人的主体性。教育评价是实现相应的教育理想的基本途径。合理的教育评价应与所追求的教育价值观有着内在的一致性。

通过学习数学教育测量与评价可以帮助我们对教育测量与评价的基本原理和价值取向不断反思，恰当处理工具价值与内在价值、物质价值与精神价值、共同价值与多元价值之间的关系，最大限度地实现相应的教育理想，这就是测量与评价的理性追求和基本使命。

第二，学习数学教育测量与评价可以帮助数学教育工作者熟悉如何运用数学教育测量与评价获得的信息，对教学进行测量与评价，从而指导数学教学实践。

对于任何一位一线数学教育工作者而言，学习数学教育测量与评价最现实的好处或者优势莫过于熟练地对学生的学、教师的教进行测量与评价，从而根据所获信息有针对性地改善数学教与学。学习数学教育测量与评价，可以帮助我们深入了解测量数据结果的实际意义，特别是它所解释的学生数学知识、数学能力和数学学习情感态度的发展状况和趋势，为提高数学教学效率、有效指导学生数学学习掌握方向和依据。

总之，数学教育测量与评价是数学教育领域中一个至关重要的组成部分。它不仅能够为教育决策部门提供参考依据，从而制定教育方针政策，完善教育改革具体方案；它还能够为学校 and 教师提供数据信息，从而科学地制定课程教学计划，合理地安排教学的内容和进度；它还能够帮助家长和学生本人

^① 李雁冰. 课程评价论 [M]. 上海: 上海教育出版社, 2002.

更好地了解学习效果，及时调整或优化学习方法，以便取得良好的数学学业成就并全面发展。

第三，学习数学教育测量与评价还可以帮助数学教育研究者顺利地阅读运用统计测量方法进行定量分析的各类科研报告和文献，我们可以从中间接地学习国内外先进的研究成果，顺利开展教育测量与评价研究。

数学教育测量与评价离不开对数据的处理和解释，因此，量化的技术和方法一定是必不可少的。特别是现代电子计算机技术的高速发展提高了数学教育测量与评价的精度和效率。在数学教育测量与评价的定量分析中，需要大量的运用统计工具处理数据，加之数学教育研究者特殊的数学修养和计算机知识，使计算机日益成为数学教育测量与评价的有力工具。学习数学教育测量与评价可以帮助数学教育研究者跟上时代的步伐，学会使用各类计算机软件进行统计分析和数据管理，学会从各类科研报告和文献中汲取养料，从而顺利开展教育测量与评价研究。

简言之，数学教育测量与评价的学习主要为我们提供一种为教学和科研服务的工具。特别是，它可以针对性地提高数学教育工作者的工作效率（学制改革、课程建设、教学效果考核、教法学法比较和学习成绩评定与比较等）。

.....◇第二讲◇.....

数学教育测量与评价的测量理论基础

.....

数学教育测量是依据一定的测量理论和规则，运用一定的测量工具对数学教育教学现象进行数量化描述的过程。因此，扎实稳定的教育测量理论是测量工作的前提基础。