

PUTONG GAODENG XUEXIAO
ZHAOSHENG
QUANGUO TONGYI KAOSHI
XILIE YONGSHU

高考数学测量 研究与实践

教 育 部 考 试 中 心

高等教育出版社

□封面设计/曹 铀□

ISBN 7-04-009903-9



9 787040 099034 >

定价：16.00元

高考数学测量研究与实践

教育部考试中心

高等教育出版社

图书在版编目 (CIP) 数据

高考数学测量研究与实践/教育部考试中心. —北京: 高等教育出版社, 2001.4
ISBN 7-04-009903-9

I. 高… II. 教… III. 数学课—教育测验—研究—高中 IV. G633.602

中国版本图书馆 CIP 数据核字 (2001) 第 14668 号

责任编辑 蒋 青 责任绘图 尹 莉
版式设计 马静如 责任校对 杨雪莲

高考数学测量研究与实践
教育部考试中心

出 版	高等教育出版社		
社 址	北京市东城区沙滩后街 55 号	邮 政 编 码	100009
电 话	010-64054588	传 真	010-64014048
网 址	http://www.hep.edu.cn http://www.hep.com.cn		
印 刷	北京市鑫鑫印刷厂		
开 本	880×1230 1/32	版 次	2001 年 4 月第 1 版
印 张	12.75	印 次	2002 年 2 月第 2 次印刷
字 数	350 000	定 价	16.00 元
印 数	52001-72400		

本书如有缺页、倒页、脱页等质量问题, 请与教材供应部门联系调换。

版权所有 侵权必究

序

高考（普通高等学校招生全国统一考试），由于受我国国情及历史、文化、传统等多种因素的影响，它被赋予了特殊的社会本质功能，那就是为普通高等学校选拔新生。它的改革被视为中小学全面实施素质教育的重要措施，要按照有助于高校选拔人才、中小学实施素质教育和扩大高校办学自主权的原则进行这项制度的改革，科目设置和内容的改革应进一步突出对能力和素质的考查，这在中共中央国务院《关于深化教育改革全面推进素质教育的决定》中已作了明确规定。

在市场经济条件下，高等学校成为具有独立办学的法人实体后，招收什么样的新生，从理想条件来说应是他们自己的事情，更多的着眼点是大学需要，而不是考生在中学受教育的具体学业内容。这在美国最大的考试机构（2000多名员工）为进入大学而设计的“学能测验”就是如此，当然他们现在也逐步改革；香港考试局是1977年成立的法定机构，它主要承担中学会考和高级程度会考（即大学入学考试）。香港考试局在考试局委员会领导下开展工作，委员会28位成员分别来自中学、高等院校及工商专业界，所有成员都由香港特别行政区行政长官委任，其日常工作由170人组成的考试局秘书处执行，据2000年该局报告称，正在进行由“评核学生所学”向“为学生学习而评核”的模式转移，并已向特区政府建议易名为“香港考试及评核局”。通过上述两个例子说明，市场经济发达的地区考试体系也愈加发达，“高考”置中学教学情况于不顾是不现实的。另一方面，设有专门的考试机构执行这类考试的运作，是一种有效的考试机制，这也从一个侧面证明，在我国社会主义市场经济条件下，面对新世纪人才需求的新形势，中共中央国务院对我国高考制度改革提出的原则和要求是多么的正确。

众所周知，为贯彻落实中共中央国务院的《决定》和全教会的精神，教育部于1999年发出了教学[1999]3号文件《关于进一步深化普通高等学校招生考试制度改革的意见》，即高考制度“3+X”改革方案。它标志着面对新世纪形势要求，新一轮高考制度改革的开始。其明确规定改革的重点是考试内容，核心是更加注重对考生能力和素质的考查，并指出这是改革的重点和难点，要花大力气、长时间深入、细致地进行这项改革。经过近几年对这项改革的调研论证，在考试内容与形式上的改革试点，以及试点范围的不断扩大与推进，在这一重点、难点的改革方面已初步获得成功，取得了一定经验，也有一些新突破，从而促进了这一改革不断深入发展。但是，必须看到与建立起一套具有教育测量理论支持、丰富改革实践经验保障、科学公正地突出测量能力和素质、有助于高校选才和中学实施素质教育的高考新机制还有不小距离，而且困难也不小。

其难点主要集中在以下几方面：

一是传统观念和习惯势力的扭转要有一个过程。自1952年建立起全国统一高考制度以来，它一直在改革，除被废除又恢复那次也算改革的影响巨大外，其近20多年来的改革从来也没有像“3+X”高考改革产生的社会影响如此深刻。上一轮高考“3+2”改革，是于1993年开始实施的高中毕业会考制度基础上的减少高考科目的改革，它把9门中学必修课中的地理、生物排除在高考之外，且比较绝对地分了文理科，即文科考语文、数学、外语加政治、历史，理科考语文、数学、外语加物理、化学。这样做的后果，对地理、生物科教学的影响不必说了，还对高中教学安排，产生过早地针对高考科目设置进行了文理分班的教学，对高中毕业生的文化知识结构产生了不良影响，这已是人所共知的事实。此次“3+X”新一轮高考改革则要逐步改革“3+2”考试模式的做法，克服其存在的主要缺陷，一要涵盖中学所有必修科目，增加高校及考生在学科性向上的选择权，二要在考试内容形式上深化改革，就不得不打破原来的备考复习模式，改革考试组织实施方法和录取方式等。因此在改革开始阶段，不易被人们迅速理解和轻易放弃传统的做法与习惯，增加了改革的难度。要达到

预期目标要求，就需要一个过程。

二是从教育测量学角度来说，也有许多理论难点需要在边改革边实践中加以研究逐步解决。例如，从理论上讲，命题专家们和考试工作者要学习研究党的教育方针体现在一个合格的高中毕业生（也涉及初中）所必备的能力和素质的内涵是什么？要继续进入高等学校学习深造所必备的基本能力和素质有哪些要求？在目前以纸笔测试为主要形式的考试中，能够测量出什么样的能力和素质？用什么样的手段与技术来达到想要测量的那些能力和素质？对这类问题好在上一轮“3+2”高考改革中已为我们奠定了一个好的基础，那时已按教育部的要求，在考查知识的基础上注重考查能力的研究与实践，比较好地处理了考查知识与考查能力的关系，并形成了一系列的研究报告或著作。但在“3+X”新一轮高考改革中，要求更加注重能力和素质的考查，就遇到了上述理论上的那些问题需要进一步深入研究，并要形成一套在理论与实践的结合上的测量理论体系，以支持考试内容与形式改革的健康发展。目前我们已得到的共识是，没有知识就谈不上有能力，知识和能力都是素质的重要组成部分，在今后的改革进程中，不能排斥单学科考试的基础地位与作用，对于应该具备且有要求的而在考试测量中无法实现的那些素质，应设置另外的机制和办法加以弥补，不能因为考试先天性的功能缺陷而影响中学全面推进素质教育。

三是从命题操作层面上讲，应该进一步研究当前在命题环节上需要深化和完善的一些概念和原则，以及它对高校选才和中学实施素质教育产生的长远影响，例如，命题掌握的内容范围是遵循但不拘泥于《中学教学大纲》，这是多年来争论的一个焦点问题；试题设计增加应用性和能力型题目，这里有个把握高中生的实际水平和现实教育水平的问题；“综合能力测试”项目引入“X”之中，它应以能力测试为主导，考查考生所学相关课程基础知识、基本技能的掌握程度和综合运用所学知识分析、解决问题的能力，首先是学科内的综合，其次才是跨学科的综合，等等。这要在构造成一张试卷时能准确恰当地体现这些原则，仅靠原来的命题经验是不能适应的，一定要在认真总结近两年来高考内容与形式改革进程中的命题实践经验基础上，上升到教

育测量学要求的高度，形成较为系统的、适应改革形势发展要求的命题理论框架体系，并在深化改革实践中不断加以丰富和完善，为这项改革的健康发展奠定较为扎实的教育测量学理论基础。

教育测量学是具有理论和实践意义的研究领域。它是以现代教育学、心理学和统计学作为基础，运用各种测试方法和手段，对教育过程、教育结果、学业成就及受教育者的能力、品格、学习能力倾向等方面进行科学测定的学科。虽然，教育测量的精确度永远达不到物理测量的水准，但准确、科学、客观、公平的测量永远是教育测量努力追求的目标。为不断接近这一目标，教育测量必须深入研究、探索测量技术及工具的科学性和标准化。而教育测量的主要工具便是测验或考试。新一轮“3+X”高考在内容改革方面的新要求，恰恰为这一领域的深入研究提供了广阔的天地。

正是基于上述的改革形势与背景，高考有关学科命题研究人员撰写了《高考数学测量研究与实践》、《高考物理测量研究与实践》、《高考化学测量研究与实践》、《高考生物测量研究与实践》四本书，这四本书构成了一套丛书。这套丛书都结合各自学科教学规律与特点，详尽阐释了在大力推进“3+X”高考内容改革和深入实施素质教育的前提下，教育测量的基本理论、原则在本学科的具体体现与要求，包括学科测量的特点及命题的理论基础；知识和能力测试内容和要求；题型设计的原则与方法；试卷构造的原则与难度控制；试卷效度、信度分析与区分度控制等。

这套丛书的问世，标志着这些学科的命题测量理论研究与实践经验得到了总结与提高，它不仅是对上一轮高考改革相关学科命题实践经验与理论研究成果的初步总结，而且对刚起步的新一轮高考改革命题测量理论在相关学科测验中的应用研究进行了概括和提高，必将对新一轮深化高考内容改革产生积极影响，对这些相关学科的教学与素质教育起到了良好导向作用。

这套丛书本来是理论性、专业性都很强的学术专业著作，而我并没有在这方面用多少文字，因为丛书各册中都有专门论述；多说了上面那些话占了不少篇幅，目的是说明看这套丛书时，不要脱离当前这

项考试改革的复杂繁难形势，丛书中的理论、观点将随着这项改革的深入而不断修订、完善；同时起到抛砖引玉之功效，诚望识者教正。

本册由任子朝主编。

教育部考试中心副主任 研究员

马金科

2001年3月

目 录

第一章 概论	1
第一节 测量的意义和测验的编制	1
第二节 高考的目的、性质	17
第二章 数学教育测量的基本理论	31
第一节 高考命题的理论基础	31
第二节 高考数学科测试特点	48
第三章 高考数学考试的内容与要求	57
第一节 知识的内容与要求	57
第二节 能力的内容与要求	174
第三节 非智力因素的要求	264
第四章 考试设计	270
第一节 题型设计	271
第二节 难度设计	320
第三节 试卷设计	336
第五章 测量的评价	360
第一节 特征量数与试题的分度	360
第二节 试卷的信度和效度	366
第三节 标准分数	381
第四节 试卷质量分析	388
附表 标准正态分布面积与横轴关系对照表	394

第一章 概 论

普通高等学校招生入学考试是我国影响最为深远的教育考试和社会考试,从1977年恢复高考招生制度以来,我国的高考进行了一系列重大的改革,有关高考的研究不断深入,为我国现代化建设人才的选拔和培养作出了应有的贡献。同时高考的理论研究和改革实践的探索创新,也有力地推进了具有中国特色的考试理论和考试制度的建设和发展。

有关高考的研究可以从不同角度,在不同层面展开。本书将从教育测量学的角度,应用现代心理学和学科教学的理论在考试理论与命题实践相结合的层面上探讨数学高考的特点与规律,以期促进学科测量理论研究和高考改革的深化。

第一节 测量的意义和测验的编制

一、测量、测验和评价

在教育测量学中,测量(measurement)是指根据某种标准,将实施测验的结果化为分数,用以表达受试者对所测问题了解多少的一种工作历程。测验(test)或称心理测验(mental test或psychological test),是包括多个问题所构成的用来鉴别能力或性格差异的工具,测验是测量的手段和工具。测验有多种形式,教育上使用最多的是性向测验(aptnitude test)和成就测验(achievement test)。与测验一词比较,测验结果所显示的是受试者工作表现的优劣程度;而测量的历程,则是对受试者表现以数字的多寡给予“他能做到多少”的解释。

心理测量是指依据一定的法则,用数字对人的心理特征的行为表现加以确定。教育测量(educational measurement)在本质上属于心理测

量,是心理测量的原理和方法在教育领域的应用.教育测量有广义和狭义之分,狭义的教育测量只包括学绩(成就)测量,是对学习结果——知识、技能的测量.评价指的是对事物进行价值分析和判断.教育测量是教育评价的基础,测量的结果是教育评价(educational evaluation)的依据,评价是对测量结果的解释过程,如果不依据测量结果进行价值分析和判定,测量结果就会失去意义.

二、测验的分类

测验的类型,可按不同的分类标准进行区分.例如,从用途上区分,一般可分为成绩测验、学能测验、人格测验等;从分数标定和解释方法上区分,可分为目标参照性测验和常模参照性测验;从测验方法上区分,可分为笔试、口试和操作测验;从测验方式上区分,可分为综合式测验和分离式测验;从测验规模上区分,可分为个别测验和团体测验;从测验要求上区分,可分为难度测验、速度测验;等等.测验的分类方法和种类都很多,我们只讨论与高考性质有关的几种测验.

(一) 诊断性测验

这类测验,是在测量学生掌握某一部分教育内容的情况时使用,是一种具有诊断性质的测验.通过这种测验可及时看到学生的学习情况,直接获得教学的反馈信息,调整教学的进度和教学的方式、方法.这种测验一般在课堂上进行,学生的情况,通常比较相近.测验的安排,多为主讲教师根据本人教学的需要自行命题,灵活性比较大.学生的成绩,主要作为了解教学情况之用,不一定作为衡量水平之用.这种测验的范围比较小,内容比较集中,目标单一.试题往往只围绕一个核心进行命题,关键是能否抓住考查内容的重点和难点,抓住了才能促进教学.在试题的设计上,要十分注意其诊断作用,即:通过试题的解答,能清楚地反映学生对某一特定的知识点和教学方法是否理解和掌握,有时还得反映其熟练程度和深刻程度.一般说来,不宜采用综合性试题.

诊断性测验主要特点在于:第一,诊断性测验一般注重于与诊断相关的目标,对每一特定的目标需要包括大量题目,每个题目之间只有很小的差异.第二,测验题目依据于对成功学习特殊技巧的详细分析以及

常见的学习错误的分析研究。第三,题目难度一般较低,重在确定学生所犯学习错误的类型以及学习困难根源所在。第四,诊断性测验一般限于课程教学中有限部分内容,且通常按若干部分的测验分数与测验记录来分析,很少用于测验全部内容。

(二) 成绩测验

学生在一定阶段的时间内,完成了某一教学大纲或教材的学习,为了检测他们的学习成绩,往往对其进行成绩测验。例如,每学期的期中、期末,甚至升级、毕业测验,都属于这类测验。这是一种检查学习进度的测验,测验内容为所考查的阶段内的教学内容,试题的命制,可看作对所考核的教学内容进行抽样。这种抽样,不应认为是随机抽样,而应该认为是:能比较好地反映教学内容的全貌的抽样。因此,必须对教学内容和教学要求作深入的分析,弄清其中的主次轻重,命题时应把握住基本的和重点的内容,而且在设问上也应考虑不同的层次要求,使之通过测验能较好地反映学生对所学内容的掌握程度,包括学到手的知识之多寡,以及理解和掌握的程度。如果采用百分制来评定成绩,得60分也就应该是相当于掌握了60%左右的教学内容。如果试题的知识覆盖面太少,或者覆盖的多为次要部分,则难以实现这样的测验目的。在试题的难度控制上也有同样的问题,太浅、太深都会使测验失效。这类测验,通常是按班级进行,而且多数情况下,由教师根据所教的内容自行命题。因此,校际之间的测验成绩,往往难以比较,也难以用它来评估学生,因为它缺乏统一的标准。

(三) 学能测验

这类测验,旨在测量考生完成某项任务的能力倾向,这些能力倾向,在很大的程度上带有潜在性,即在测验时,考生也不一定已具备这些能力,而只是具备发展这些能力的基础和倾向。这类测验的命题并不完全根据以往的教学内容来命题,更主要的依据是对学能结构的分析,对以往的教学内容,作为命题的取材,往往是选用那些可作为学能基础的内容,在要求上也不一定局限于教学要求这个层次上。这类测验往往带有选拔的性质,应有较好的区分度,使能力倾向的大小和优劣能得到较为细致的区分。

各类入学测验,其实质都可作为学能测验来对待.成绩测验与学能测验的基本区别在于:前者着重的是现时所达到的水平,后者所看重的是未来的发展倾向和趋势.从这一角度看,我国普通高等学校入学统一测验(简称高考),由于历史的原因,在高中毕业会考制度建立之前,它既有水平测验的特征,又兼有学能测验的选拔性质.随着高中毕业会考制度的实施和成熟,高考学能测验的性质日显突出,能力和学习潜力的考查,在高考中,正在得到加强.

(四) 速度测验和难度测验

在各类测验中,对解题速度都有一定的要求,因为任何测验都得在规定的时间内解答一定数量的试题,这本身就表明了速度的要求.所谓速度测验,指的是测量考生完成某一任务的速度之测验.速度的要求往往是在能力要求的基础上提出的,突出速度的测验,在能力要求上往往比较集中单一.例如,打字速度的测验,数和式的四则运算的速度测验,解方程的速度测验,作图速度的测验,等等.对于大规模的综合测验,一般较难把速度要求放在首位.速度测验的试题,其题量往往比较大,对速度也往往提出定量化要求.例如,计算速度测验,通常也会规定一小时要完成多少个算式的计算任务.

难度测验主要是考查考生掌握知识的深度和水平,试题的数量不多,但其中有些很难的试题,只有少数考生能够解答,虽然考生有机会接触所有的题目,但并不是每个考生都能完成全卷.

速度的快慢往往反映了思维的灵活性和敏捷性,因此,在一些大规模的测验中,往往含有较多的考查项目和试题,其用意不在于要求考生都能全部完成,主要的目的是借助它测量考生的解答速度,从一个重要的侧面考核考生的能力,提高测验的区分度.而难度测验则更注重思维的深刻性,反映了思维的强度和深度.

(五) 目标参照性测验

目标参照性测验,也称为目标参考性测验,或者尺度参考性测验.凡是评分时参照某一事先规定好的尺度标准和目标,并用评出的分数作为考生成绩的反映和标识,这样的测验统称为目标参照性测验.一般说来,单科(或单项)的成绩测验和水平测验多属目标参照性测验,而

多科的水平测验,则往往介于目标参照性测验与常模参照性测验之间。

目标参照性测验所给的分数,并不考虑其他考生的情况,只是相对于给定的标准加以评定。只有达到规定的目标标准或尺度标准才算通过,或称及格、达标。分数的解释离不开具体的试题,孤立的一个分数也反映不出考生成绩在全体考生中所处的地位。不同学校、不同时间的测验成绩,缺乏可比性。在目标参照性测验中,可能全体考生都达到及格,也可能全部不及格。社会上许多任职资格测验,也都属于目标参照性测验。通过测验,并达到及格,才能领取从事该项工作的证书或执照。

(六) 常模参照性测验

常模参照性测验,也叫常模参考性考试。在这种测验中,反映考生成绩的分数与目标参照性测验不同,它不是以事先制定的尺度标准为参照点,而是参照某个常模来反映考生的分数。这里所说的常模是指:某一考生群在测验中的成绩(通常,用该群考生的卷面分数的平均分与标准差来刻画这个成绩)。也就是说,常模参照性测验中,反映考生成绩的分数是结合其他考生的成绩来给出的,从所给的分数便可看出考生在某群考生中的地位,是优是劣,一目了然。这样做,对于以选拔为目的的大规模测验,无疑是必要的。这是因为:选拔一个考生,就必须将他与其他的考生进行比较,分辨优劣,才能决定取舍,不能光看他一个人的测验卷面分数。这个道理显而易见。事实上,在常模参照性测验这个概念及其给分方法提出之前,人们已经明白这个道理,也已懂得用排名次的方法解决这个问题。即是说,用序数来反映考生的测验成绩。

在常模参照性测验中,测验的内容和要求,同样必须根据测验的目的和用途,给予明确的规定。命题时,同样要力求准确和稳定地反映这种规定,试题不能忽易忽难,波动太大。同时,命题时,必须更好地了解考生的实际,提高试题的区分度,使测验结果的卷面分数(通常称之为原始分),能尽可能切实地反映考生的成绩,这是建立“常模”的基础。

三、测验的理论

在测验中,检测的事物属于心理现象。为使测验能有效地进行,必

须对这些心理现象(包括智力、学习、知识、技能、情感等)的本质和可检测性,乃至量化的问题展开深入的理论研究,并在其指导下探索测验的操作方法和施测程序.半个多世纪以来,有关这方面的研究与探索取得了长足的发展与进步.其中与我国高考的实际有较密切关系的是智力理论、教育目标分类理论和标准化考试理论.下面将其摘要介绍如下:

(一) 智力理论

高考是属于心理测量的范畴,侧重测量学生的学科知识和智力发展水平,因此智力理论是其重要的基础.智力理论是指心理学家对人类智力组成的系统解释.关于智力理论,虽然心理学家都同意将智力解释为心理能力,但对智力的内涵或组成成分,智力究竟是包含一种能力,还是包含多种能力等问题,却一直是见仁见智.教育心理学家张春兴从立论取向的角度把智力理论分为两大类,每一类又分为不同的种.第一类是心理测量取向的智力理论,包括智力二因论、智力群因论和智力结构论三种.第二类是多维取向的智力理论,包括多维智力论和智力三维论两种.下面简要介绍其中的三种.

1. 智力二因论

智力二因论(two-factor theory of intelligence)最早是由英国心理学家斯皮尔曼(Spearman, C. E.)在1904年提出的.他认为人类的智力包括两种因素:其一为一般因素(general factor),简称G因素(G-factor).智力的一般因素,其心理功能是表现在一般性的活动上.其二为特殊因素(specific factor),简称S因素(S-factor).智力的特殊因素,其心理功能只表现在特殊性的活动上.斯皮尔曼认为,人类的智力差异,既表现在一般因素上,也表现在特殊因素上,只是两种因素之间的关系,并不是一定的.1963年,斯皮尔曼的弟子,美国心理学家卡特尔(Cattell, R. B.)将智力分为两类:其一为流体智力(fluid intelligence),流体智力主要通过对空间关系的认知、机械式记忆、对事物判断反应的速度等方面表现.其二为晶体智力(crystallized intelligence),晶体智力主要通过对语文词汇和数理知识的记忆表现.

2. 智力群因论

美国心理测量学家瑟斯顿(Thurstone, L.L.),不同意斯皮尔曼的智力二因论的观点,认为智力是由一些基本心理能力(primary mental ability)组成,这些基本能力是:(1)言语理解(verbal comprehension);(2)语词流畅(word fluency);(3)数字运算(number);(4)空间关系(space);(5)联想记忆(associative memory);(6)知觉速度(perceptual speed);(7)一般推理(general reasoning)七种基本能力。

3. 多维智力理论

多维智力理论是由美国心理学家加德纳(Gardner)倡议的,他强调人类的心理能力至少应该包括以下七种不同的智力:(1)语言智力(linguistic intelligence);(2)数理智力(logic-mathematical intelligence);(3)空间智力(spatial intelligence);(4)音乐智力(musical intelligence);(5)体能智力(bodily-kinesthetic intelligence);(6)社交智力(interpersonal intelligence);(7)自知智力(intrapersonal intelligence)。

前面提到高考既是成就测验,也是性向测验.智力测验是对人的一般性向进行测量的过程,因此对智力测验有指导意义的智力理论对高考尤其是高考考试目标和学科能力要求的确定也是有指导意义的。

(二) 教学目标分类学理论

教学目标(instructional goal)是在教学之前,预期教学活动结束之后,学生从教学活动中学到些什么,是知识与技能,还是态度与观念.在教学目标的研究中,教学目标的分类问题尤其引人注目,从教学目标的分类方式看,在各家教育目标分类理论中,在理念上有共识者,一般都认为在学校不同学科的教学活动中,所期望达到的教学目标,不外乎下列三方面:(1)认知方面,指从教学中学到知识性的行为;(2)情意方面,指从教学中学到感情与意志性的行为;(3)技能方面,指从教学中学到动作技能性的行为.虽然在理念上具有共识,但在教学目标实践上,却出现两种取向:早期的教学目标分类是将认知、情意、技能三方面的教学目标分别列举,因而在教学之后的教学评价时,也分别评价.新近的教学目标分类,则趋向于将认知、情意、技能三方面的教学目标做统合性的处理,因而在教学之后的评价时,也采取统合的评价方式.从