

化学教学测量 与评价导论

Huaxue Jiaoxue Celiang
Yu Pingjia Daolun

阎蒙钢◎著

化学教师授课质量的测量与评价

科学统计与分析化学成绩的方法

化学学业成绩的测量与评价

化学教学目标的分类、编制与目标教学

化学教学测量与评价概述

北京科学技术出版社

安徽师范大学出版基金资助

(Supported by the Publication Grants of Anhui Normal University)

化学教学测量与评价导论

阎蒙钢 著

 北京科学技术出版社

图书在版编目(CIP)数据

化学教学测量与评价导论/阎蒙钢著. - 北京:北京科学技术出版社, 2004.5

ISBN 7-5304-2864-0

I. 化… II. 阎… III. 化学课-教学研究-中学 IV. G633.82

中国版本图书馆 CIP 数据核字(2004)第 003007 号

化学教学测量与评价导论

作 者: 阎蒙钢

责任编辑: 徐利明 章 健

责任校对: 黄立辉

责任印制: 严加琦

封面设计: 世纪白马

出 版 人: 张敬德

出版发行: 北京科学技术出版社

社 址: 北京西直门南大街 16 号

邮政编码: 100035

电话传真: 0086-10-66161951(总编室)

0086-10-66113227 0086-10-66161952(发行部)

电子信箱: postmaster@bjkpress.com

网 址: www.bjkpress.com

经 销: 新华书店

印 刷: 北京凯通印刷有限责任公司

开 本: 850mm×1168mm 1/16

字 数: 320 千

印 张: 15

版 次: 2004 年 5 月第 1 版

印 次: 2004 年 5 月第 1 次印刷

印 数: 1—5000

ISBN 7-5304-2864-0/G·305

定 价: 30.00 元



京科版图书, 版权所有, 侵权必究。

京科版图书, 印装差错, 负责退换。

序

通览全书可以看出，作者力求体现素质教育新形势下基础教育课程改革、考试改革的新理念和新评价观。既注重深入浅出地运用现代教育测量与评价的基本理论，全面系统地介绍化学教学测量与评价的编制原理、方法和技巧；又强调实际运用，突出重点，用较多的篇幅侧重研讨了学生学的质量（主要是学业成绩）和教师教的质量（主要是课堂授课质量）的测量与评价这两大内容。尤为有特色的是分类探索了化学课堂教学质量评价的理论与方法，并以较多的典型实例展示了有关研究成果。还把微格教学评价、说课评价等最新内容和方法纳入了课堂教学质量评价体系。同时，也对高考等各类化学考试的命题特点、编制技术、复习对策及评价方法等诸方面进行了较为深入的阐述。因而本书具有观点新颖、内容翔实、特点明显、体系完整、操作性强、便于自学与掌握等特点。本人相信《化学教学测量与评价导论》一书的出版会使广大读者从中获益，利于读者在深入开展化学教学的测评研究工作中，受到多层面的启示和帮助。

刘知新

2004年2月于北京师范大学化学系

注：刘知新，我国著名化学教育家，北京师范大学化学系教授。曾任原国家教委高等学校理科化学教学指导委员会委员，原国家教委和教育部全国中小学教材审定委员会第一、二、三届中学化学学科审查委员，中国教育学会化学教学专业委员会理事长，中国化学会《化学教育》主编等职；现兼任教育部全国中小学教师继续教育教材评审专家、化学教学专业委员会顾问、《化学教育》顾问等职。

序

化学教学测量与评价是化学教学系统的重要组成部分。“化学教学测量与评价”这门新兴学科是基于“以学生发展为本”的教育理念，以化学教学目标为依据，运用教育统计学、教育测量学、教育评价学和计算机科学的理论与技术，结合化学教学的特点与课程改革的热点，对化学教与学的效果进行比较和价值判断研究的一门学科。这门学科运作的终极目的在于对化学教学活动中的诸要素及其互动效果给予教育价值上的科学判断，进而做出正确的教育教学决策，以充分发挥测量和评价对教与学的激励、诊断、强化与发展等多种功能。

关于教育测量和评价的研究，在国外始于20世纪初，我国早在辛亥革命后就开始了这项研究，其后因历史的原因而中断，20世纪80年代才重新开展这方面的实证研究。如学校教育管理中使用的各类评估方法，部分省市试行的标准化考试，基础教育新课程实验中开展的新课程评价等。同时，出版了多部由我国学者撰写的著作。

化学教育与测量评价研究工作，即将教育测量与评价的理论运用于化学教育教学，科学地、量化地开展化学教学效果的测量与评价，在我国也是从20世纪80年代开始的。如一些省市开展的对化学教学目标分类、目标教学及化学教师授课质量评价指标体系的研究；1985年《化学教育》载文介绍教育统计与测量的基本知识；1987年起在广东省试行的化学标准化高考及近年开展的化学新课程教材评价等。我国高等师范院校也相继开设了“化学教育统计”、“化学教育测量与评价”等课程。应当说，掌握现代化学教学测量与评价的理论与方法，是适应化学教学管理和考评方法日趋标准化、科学化的需要，也是素质教育新形势对化学教师现代化学教育科研能力结构的新要求。由于进行这项工作要用到教育统计、教育测量、教育评价、教育学、教育心理学、计算机科学等多学科知识，不少一线化学教师和师范生多感到很陌生。因而需要结合化学教学实际，有计划地学习有关的论著。至今，国内除了本人于1996年主编的《化学教育测量与评价》一书外，尚很少见到结合当今基础教育课程改革和教学改革撰写的这方面的著作。

阎蒙钢同志将潜心研究并从事十余年教学实践的成果整理成书，力图为构建和完善我国“化学教学测量与评价”这门新兴学科的理论体系做出奉献，是一件可喜可贺的事！

目 录

第一章 化学教学测量与评价概述	(1)
第一节 化学教学测量与评价的概念及功能	(3)
一、什么是测量	(3)
二、什么是教育统计学	(4)
三、什么是教育测量学	(5)
四、什么是教育评价学	(7)
五、化学教学测量与评价的概念	(9)
六、化学教学测量与评价的主要功能	(10)
第二节 化学教学测量与评价的基本类别	(11)
一、教师化学教学质量的测量与评价	(11)
二、学生化学学业成绩的测量与评价	(12)
第三节 化学教学测量与评价的常用方法和步骤	(13)
一、常用方法	(13)
二、基本步骤	(16)
三、注意事项	(17)
第二章 化学教学目标的分类、编制与目标教学	(19)
第一节 分类	(21)
一、教学目标与测量目标的关系	(21)
二、教学目标与教学目的、教学内容关系	(21)
三、布卢姆的教育目标分类方法简介及化学例证	(22)
四、我国化学教学目标研究现状	(26)
第二节 编制	(28)
一、编制教学目标的依据	(29)
二、教学目标的陈述方式	(29)
三、教学目标的层次	(29)
四、化学教学目标的编制步骤	(29)
五、初中化学单元教学目标编制示例	(32)

第三节	化学实验教学目标的编制与测评	(35)
	一、化学实验教学目标的分类	(35)
	二、化学实验操作技能的学习水平分类与目标制定	(36)
	三、化学实验操作技能测评的基本类别与方法	(37)
	四、高考化学实验考试目标研究	(43)
第四节	目标教学	(46)
	一、理论依据	(46)
	二、基本程序	(46)
	三、我国目标教学的实践成就与理论贡献	(47)
第三章	化学学业成绩的测量与评价	(51)
第一节	考试的素质教育功能	(53)
	一、考试与素质教育的关系	(53)
	二、考试的意义	(54)
	三、国外基础教育考试改革趋势	(54)
第二节	化学学业成绩考试的类型与评价方法	(55)
	一、常用的化学学业成绩考试类型	(55)
	二、化学学业成绩考试的评价方法	(57)
	三、化学学业成绩评价中实施多样化评价的策略	(60)
第三节	标准化考试的方法	(66)
	一、基本程序	(66)
	二、主要优缺点	(67)
	三、澄清几点误识	(67)
	四、广东省 1987 年首次化学标准化高考简介	(68)
	五、我国化学标准化考试的展望	(71)
第四节	化学常规考试的科学化	(71)
	一、科学程序	(71)
	二、操作过程	(72)
	三、自编目标参照性测验示例	(73)
第五节	考试组卷的蓝图设计	(74)
	一、什么是蓝图设计	(74)
	二、组卷蓝图的形式与内容	(75)
第六节	化学试题的题型与编制要领	(76)
	一、化学试题的分类与功能	(76)
	二、化学试题的编制要领	(84)
第七节	化学考试命题与组卷	(86)
	一、命题的意义	(86)

二、命题计划的内容与形式	(86)
三、命题工作的任务与程序	(88)
四、命题的一般原则	(88)
五、命题中常见的问题与组卷技术	(90)
第八节 化学考试题库的建立	(92)
一、必要性	(92)
二、建立条件	(93)
三、建立方法	(93)
四、化学试题库示例	(93)
五、化学教师怎样建立自己的小题库	(94)
第九节 考试实施的组织管理	(95)
一、实施过程及基本要求	(95)
二、考场管理工作	(96)
三、评卷管理工作	(97)
四、综合实践活动课考试实施探讨	(99)
第十节 高中化学会考简介	(103)
一、会考的产生与现状	(103)
二、会考的目的、性质、功能及科目	(104)
三、化学会考的实施方法	(105)
第十一节 我国高考化学命题改革评述	(106)
一、高考的特点	(106)
二、高考的历史回顾	(107)
三、高考改革现状与趋势	(107)
四、高考化学分卷命题特点	(109)
五、“3+X”高考理科综合能力测试试卷及其中化学试题特点	(110)
六、高考化学重要题型与解题技巧研究	(110)
第十二节 化学考试的心理训练	(117)
一、化学考试心理的测量方法	(118)
二、化学考试心理训练的基本方法	(121)
三、考试卫生	(123)
第四章 科学统计与分析化学成绩的方法	(125)
第一节 考试分数的初步整理——图表法	(127)
一、统计表	(127)
二、统计图	(130)
第二节 考试分数的进一步整理——特征量数法	(135)
一、集中量	(135)

二、差异量	(142)
第三节 原始分数的转换与解释	(147)
一、为什么要对原始分数进行转换	(147)
二、怎样把原始分数转换成标准分数	(148)
第四节 标准分数在测量与评价学生化学学业成绩中的应用	(151)
一、在确定学生相对地位中的应用	(151)
二、在比较学生成绩优劣中的应用	(151)
三、在高考录取中的应用	(152)
四、在管理学生学习质量中的应用	(155)
第五节 相关统计与化学考试质量分析	(158)
一、相关系数	(158)
二、相关系数的显著性检验	(160)
三、考试质量分析的基本指标	(161)
四、化学考试质量分析的基本方法	(167)
五、示例	(169)
第五章 化学教师授课质量的测量与评价	(183)
第一节 化学教师授课质量测量与评价指标体系的设计技术	(185)
一、指标体系设计的基本要求	(185)
二、指标体系的设计过程	(186)
三、指标的量化技术及权数的确定	(186)
四、示例	(188)
第二节 关于建立我国化学教师授课质量测评指标体系的探讨	(193)
一、我国化学教师授课质量测评指标体系研究简介	(194)
二、化学课堂教学质量分类评价探讨	(199)
第三节 化学教师授课质量评价的实施方法	(209)
一、选择评价方法	(209)
二、评价结果的量化处理	(213)
第四节 化学教师课堂教学技能的分类与评价	(214)
一、微格教学简介	(214)
二、化学教师课堂教学技能分类的研究	(215)
三、化学教师课堂教学技能的评价	(215)
第五节 化学教师的说课评价	(218)
一、说课的内涵	(218)
二、说课的内容	(218)
三、化学教师说课评价指标体系探讨	(219)

附录

一、高中化学知识内容和测量水平双向细目表	(221)
二、调查问卷示例	(223)
三、采用弗拉南根方法估计二列相关系数的数据表	(225)
四、正态分布表	(226)

参考文献	(227)
------------	-------

化学教学测量与评价概述

学习指导

(1)掌握化学教学测量与评价的理论和方法,是适应化学教学管理和考评方法日趋标准化、科学化的需要,也是对新世纪化学教师现代教育科研能力结构的新要求。

(2)化学教学的测量与评价是中学化学教学活动的一个不可缺少的环节,但它绝不是单纯为了证明教学效果或为了评等级、排名次,其目的在于对化学教学活动所涉及的主要因素及其效果给予价值上的判断,进而调节和改进教学,使教学过程真正成为一个首尾相接的可控制系统。

(3)教学测量与评价主要包括教师教的质量(主要是授课质量)和学生学的质量(主要是学业成绩)这两大内容。

(4)教师授课的质量直接决定着教学质量,所以对化学教师教学质量的测评又应以课堂授课质量为主,并以设计科学的评价指标体系作为评价标准。

(5)由于学生的学业成绩既反映了学生的学业水平,也在一定程度上体现了教师教学的能力和水平,所以学生学的质量测评又应以化学学业成绩为主要内容。

(6)考试、测验是衡量学生学业成绩的主要方式,学业成绩的测验与考试又分为常模参照性测验(用于衡量学生学业成绩相对水平)和目标参照性测验(用于衡量学生学业成绩的实际水平)两大类。因而学业成绩的考评又有相对评价和绝对评价之分。

引言

教学是学校里一项最主要的工作,是培养人才、实现教育目的的基本途径。教学过程是由相互连接的各种教学环节组成的一个统一体,其中之一就是对教学质量的测量与评价,以提供信息改进教学和对被检测对象作出某种资格证明,这是教学过程的一个十分重要的环节。教学质量的测量与评价的范围很广,包括了教学活动的各个方面,如教学过程、教学成果的测量与评价。在中学,最大量的是对教师教学水平(包括教材处理、教学设计、教学方法、教学效果等)和学生学业水平的测量与评价。

长期以来,很多学校总是把考试作为衡量教学质量的惟一方式,把一次考试、几张试卷的成绩作为一把尺子来衡量教学质量,甚至以此来决定学生的命运,因而带有很大的局限性和片面性,不能反映出全部教学的质量。现实中存在的片面追求升学率及大学新生中一些人“高分低能”的现象,以及教学工作无标准的状况,都和缺乏对教学质量进行全面的、科学的检查与评价有关。

实际上,教学是教师教、学生学的统一活动,因此对教学质量的测量评价就应该主要包括教师教的质量(主要是授课质量)和学生学的质量(主要是学业成绩)这两大方面。

本书正是从这两方面入手,分别对化学教师授课质量和化学学业成绩如何实施科学地测量与评价进行研究。又由于这种测量与评价所采用的基本工具大多是各种形式的考试和测验,因此,本书也用较多的篇幅去探讨如何实现考试与测验的标准化、科学化。

第一节 化学教学测量与评价的概念及功能

一、什么是测量

测量(measurement),通常是指人们对客观事物进行某种数量化的测定。如对物体的长度、重量、温度以及时间、空间、运动等物理特性的测量。它是一种直接测量,其操作方法已被人们共同接受,并有一些较好的量具来表现这些结果,如天平、尺子、温度计等。这种测量具有以下三个基本要素。

(一)单位

单位(unit)即计算数量的单位。如测量长度以米、厘米等单位。这种单位的特

点是等值的,如 1cm 与 2cm 之间的间距等于 5cm 与 6cm 之间的间距。

(二)参照点

参照点(reference point)即确定事物数量的计算起点。参照点有两种:一种是绝对的零点,即以“恰恰没有一点重量”、“恰恰没有一点长度”为计算起点,对其测量只要统一单位就行了;另一种是人定(为)的参照点,即相对零点,如以海平面为测量陆地高度的起点,以冰点为测量温度的起点等。最好的参照点是绝对零点,人定的参照点越接近零点越好。

(三)量表

量表(scale)是具有单位和参照点的连续体,是测量的工具,也称量具。

实施物质测量,既要求量具本身的标准化(即统一单位和参照点),又要求使用量具方法的标准化(即统一操作程序),否则就没有准确性和可靠性。

上述均为物质实态的直接测量。那么,对人类的精神现象(心理现象),如人的知识、技能、智力、兴趣、气质、性格、品德等精神特性,能否测量呢?

回答是肯定的,其测量所要依据的基本理论和方法就是本书所要依据的基本理论和方法——教育统计学、教学测量学和教育评价学。

二、什么是教育统计学

(一)统计学

统计学是研究统计原理和方法的科学,属于应用数学领域。从广义上说,统计学是对于事物总结信息的搜集和分析,并以此为根据进行推断的方法和理论。

统计学分为两大类:一类是与研究对象的特征密切结合的各科专门统计学,另一类是数理统计学。

统计学最初从研究人口、土地等国情开始,经过三百多年的发展,至今已经成为一门拥有众多分支和具有方法论特色的科学。

(二)教育统计学的产生与概念

教育统计学的历史并不长,到现在才 100 年左右。它是统计学和教育科学互相结合、渗透而产生的一门边缘科学。(图 1-1)

可见,教育统计学既是统计学的一个分支,又是教育科学的一个分支,但又与二者不同,所以它的定义为“教育统计学是运用统计学中的一般原理和数理方法研究教育现象的一门应用科学”。

例如,著名的世界第一条遗忘曲线——艾宾浩斯遗忘曲线的内容和规律是教育心理学范畴的,而教育统计学在其中的作用则是形象地表达了规律。

(三)应用范围

教育统计学可用于正规学校教育的研究,也可用于业余教育的研究;它可用于儿童身心发展规律的特点的研究,也可用于教育事业发展规划和速度的计算;它可给心理实验结果以统计检验,还可为教育机关对人、财、物的科学管理提供必要的数据,等等。无论哪一方面的研究,必须在充分占有数据的基础上才有可能科学分析和说明,从而掌握其一般发展趋势,探索出规律性的东西。

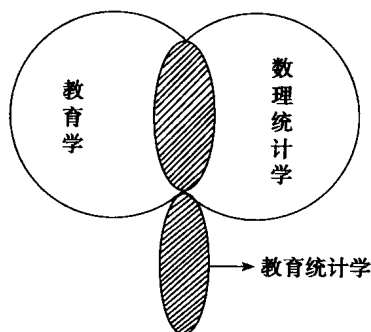


图 1-1 教育统计学是边缘科学

三、什么是教育测量学

(一)教育测量学的产生与概念

教育测量(educational measurement)作为一门科学产生于 20 世纪初西方的客观测验运动,但其源流却可追溯到我国古代的考试制度。早在公元前 840 年的西周时期,我国就初步建立了学校教育制度,实行了教育考评。我国专门以考试为主选拔人才的制度始于隋炀帝大业二年(公元 606 年)开创的科举考试制度,它通过分科考试来选取人才,采用了贴经、墨义、策问、诗赋等方法去测试考生,此可谓开教育测量之先河。欧美各国 18 世纪末、19 世纪初采用的以考试的方法选拔官吏,是从我国学去的。

教育测量虽起源于我国,但是我国古代的考试,一般是试题的数量少,知识的覆盖面窄,不足以代表所测学科知识的内容,而且评分又难以保证客观性,因此,它不是科学的教育测量。

1904 年美国心理学家桑戴克(R. L. Thorndike)编著出版了《心理与社会测量》一书,该书介绍了心理统计方法和编制测量的基本原理。其后他又编制了《书法量表》、《拼字量表》、《作文量表》、《图画量表》等标准测量工具,使教育测量走上了科学化的道路。因此,桑戴克被称为教育测量学的鼻祖。

早期的教育测量主要用于客观化、标准化地考查学生学业。20 世纪 30 年代起又扩展到考查学生的学习态度、兴趣、品德、性格以及教育工作的各个方面。因此,现代教育测量的定义应为:“教育测量就是对学生们的学习能力、学业成绩、兴趣爱好、思想品德以及教育措施上许多问题的数量化测定”。它主要用于对学生精神特性的测定。

(二)教育测量的特点

教育测量学是把教育统计学方法运用于考试分析而产生的一门学科,它主要研

究教育测验的编制、实施和测验结果的解释。教育测量和物理特性的测量一样,也具有单位、参照点、量表这三个基本要素,但它不像物理特性的测量那样可以直接测定,而是具有以下特点。

1. 教育测量一般是间接测量

这是因为人的知识、智力、才能,是人的大脑活动情况,还无法直接测量。我们只能通过学生对测验题目的反应和一些行为表现,运用推理、判断的方法,来间接地测量出他们的知识水平、智力高低和品德好坏。

2. 教育测量的度量单位是相对的,且不等距

物理特性的度量单位一般是绝对的,1m就是1m,无论在何地都相同(当然也是有条件的,具有一定的相对性,如在海平面上1kg重的物体,在高山顶上就会不到1kg重)。

在教育测量中,一个学生在一次班级测验中得70分,可能是这个班较高的分数,而又可能是别的班和外校较差的分数。这是因为不同班级考试题目难易不同,学生成绩分布情况不同,因而分数的价值并不相等。同样,一个考分为80分的学生的知识量并不等于在同一测验中得40分的学生知识量的2倍。因此,教育测量的度量单位是相对的。一个学生的学业成绩好坏、智力高低和能力大小等,都是就其所在团体的整个分数序列、行为序列中的地位来说的,其测量的分数单位并非绝对的。同时,分数单位又非等距的,如56分和58分、98分和100分,同样差2分,但性质不同,后2分的取得可能在难度上要大于前者。虽然在教育测量中所使用的单位并非绝对等值,但教育测量必须有单位,则是绝对的。

3. 教育测量的参照点是相对的

教育测量中所用的参照点,可以说都属于人定的参照点。而人的心理现象是没有绝对零点的,即使一个学生考试得了零分,也不能说他的化学知识为零。所以参照点是相对的。另外,教育测量中用的参照点还有一个很大的限制,即从参照点计算的分数,不能以“倍数”的方式解释,如甲生的化学成绩是100分,乙生是50分,我们则不能说甲生的化学知识掌握程度是乙生的2倍。

教育测量应用的参照点有许多种,有以编制测验者推测的零点作为参照点的,有以一个年级或一个年龄段的学生作业的平均分数为参照点的。由于测量应用的参照点不同,所测得的分数也不相同。为了对这些不同的分数进行比较,必须进行分数的转换。这也是不同考试的原始分数非但不能直接相加,也不能直接加以解释的原因。

4. 教育测量要控制的变量很多

如题目编制本身、测试过程、被试者的心境、内部条件等。所以教育测量要达到科学化,首先要实现标准化。

5. 教育测量要制作标准化量表

要测量某种事物,就需要先有一个具有单位和参照点的连续体,然后用这个连续体去测量某种事物,以确定该事物的数量,这个连续体就叫量表。教育测量中所使用的量表多以文字试题(考卷)的形式出现,也有以图形、符号、操作要求形式出现的。

教育测量常用的量表有:百分量表(常用于测量学生学业成绩)、年龄量表、等级量表、T量表等。

(三)教育测量的种类

教育测量的种类随分类标准不同而有所区别。

1. 以测量的对象来分类

(1)学业成绩测验:这类测验用于测量学生的知识、技能,如数学测验、化学测验等。也有人将此类测验称为教育测验,这是本书主要研究的内容。

(2)智力测验:这类测验用于测量学生的观察、识记、分析、判断、推理等思维活动的能力,如迷津测验、填图测验。

(3)人格测验:这类测验用于测量学生智力以外的心理的各种特性,如性格测验,兴趣、情感测验等。

(4)特殊能力测验:这类测验用于测量学生某种特殊能力的程度,如音乐能力、绘画能力测验等。

2. 以测验结果的参照点来分类

(1)常模参照测验:其目的在于鉴别或比较被试者之间的个别差异,也就是确定每个被试者的成绩在集体中所处的位置,比较所用的参照点就是测验的常模。所谓常模,就是指该考生集体的平均分、标准差和成绩分布状态。某人成绩在常模之上则高于平均水平,在它之下则低于平均水平。入学统考、高考、化学竞赛一般都具有这类常模参照测验的性质。

(2)目标参照测验:其目的在于检验被试者是否完成规定的教学目标(见第二章),以及完成的质量如何。这种测验的参照点则是事先制定好的教学目标或根据它浓缩了的测量目标,如摸底测验、期中期末考试以及不计入高一级学校招生考试录取总分的毕业会考等,都属于目标参照性考试。

这两种测验在功能和参照点方面虽然明显不同,但这种划分也不是绝对的,常模参照测验的结果也能反映某一学科部分的(而且是主要部分的)教学目标的完成情况,而目标参照测验一般也能反映出被试者在总体中的位置。近年来也有将这两种类型的测验结合进行的尝试。

四、什么是教育评价学

(一)教育评价学的产生与概念

对教育评价(educational evaluation),我们并不陌生,在教育工作中常用的考核、甄选、检查、总结、评比、奖惩、视导等都含有教育评价的因素,是教育评价的原始的、初级的形式,只是未用其名而已。

在 20 世纪 30 年代才第一次出现“教育评价”的概念,它是在 20 世纪初出现并盛