

# 电子图书



信息技术的结晶

人类文明的载体

网络的基本资源

## 前言

教育测量学作为对教育现象进行定量化测定的一门教育科学，从本世纪初建立以来，已受到世界各国的普遍重视。这门学科在人才选拔、因材施教、教育评价、教育研究等方面发挥着越来越大的作用。而化学教育测量和评价则是在运用教育测量和评价的研究成果逐渐生发、凝聚而成的一门年轻学科，可以说是在借鉴或移植教育测量和评价的原理、规则来具体研究化学教育现象，以获得定量化的认识而逐渐成型的。

从我国的教育和研究实践看，在 70 年代至 80 年代初，重新恢复教育测量和教育统计等专业课程及重组研究课题以后，随着教育的深入开展，化学教育界也开始重视科学的、定量化的测量和评价研究，经过一段试验改革的实践，开始积累了大量的、宝贵的数据和经验。近几年来，已开始有研讨化学教学目标的论著问世。应当承认，化学教育测量和评价的研究与化学教育的其他研究领域相比，仍处在资料不够充分、理论上停留在经验性概括的阶段，但其建立和发展的条件业已具备。“一切知识都是从经验开始，又终结于经验。”（爱因斯坦语）任何学科的建构如同学者治学一样，总是从实践中取得经验，储宝撷英，积厚薄发，然后自成一家。基于以上认识和信念，我们在近六七年理论研究和教学实践的基础上，将不太系统的研究成果梳理成一个体系，并形成成为本书。

本书的重点在于介绍化学教育测量和评价的基本理论、方法和技术，讨论化学教育测评研究的重要课题及其研究成果和发展趋势；在写作时采用突出化学教学测评，尽力体现化学教学中“学”与“教”这一双边特征，以化学教育目标与化学教育测评标准为基点和起点，延伸、分解为实施测评的程序与方法，化学标准化考试放在本书的中间，最后以化学课堂教学评价为结尾。这种体例意在促使教育者（化学教师和研究工作者）努力从化学教育测验与测量程序中获得更多信息，以利于为改进化学教学、提高教育质量制定决策。“测验与测量之所以有用，是由于它允许而且被用来做出较好的决策。”（R.L.桑代克和 E.P.哈根语）我们认为，化学教育测量和评价是一门实践性很强的应用学科，一定要重视紧密结合化学教育实际，运用并充实、发展教育测量和评价的一般原理，要在实践研究上下功夫，努力去探求化学教育定量化测定的诸多规律性的结论，进而对改革和提高化学教育工作做出科学的决策。

全书由王建成执笔，刘知新统稿、审稿。我们尽力将文稿写得既贴近实际，又反映研究成果及发展趋势，既重视理论探讨，又注意可操作性。这一尝试的效果如何，还有哪些不足，尚祈海内同行不吝指教。

作 者

1995 年 10 月于北京师范大学

## 序

刘知新

化学教育作为科学教育的一个分支，以其特有的功能在实现第一流人才培育及普遍提高全体公民的文化科学素质这一使命中，起着重要作用。众所周知，化学教育正是在化学科学、技术与社会，以及学校教育、社会教育等这些大教育环境中，不断发展的。应当说，化学科学的进展与科技教育的发展推动了化学教育的繁荣和更新；学校教育与社会教育的扩展、改革和不断完善，为化学教育提供了培育良才的广阔天地。总之，社会的进步，科学技术的发展，教育的普及与提高，为化学教育的产生、发展和繁荣并发挥其多种教育功能提供了智能源泉与人才基础。应当强调指出：化学教育与其他学科教育一样，在大教育系统中发挥着维系事业兴衰、人才延续等多种功能。这些教育功能可概括为：简约有效地将人类的文明遗产传授给受教育者的传输功能；按社会的需要培养人才的塑造功能；用最经济的人力、物力和时间造就大批合格人才的高效功能；以及超前为社会的进步和革新培养适用人才的变革功能。

中华人民共和国的化学教育，自 1949 年至今，从基础教育、职业技术教育、高等教育到继续教育，不论在规模上还是质量上，都取得了令人瞩目的成就。全国各级各类学校的广大化学教育工作者为此付出了辛勤的劳动，创造并积累了丰富的教育教学经验，这是我国和世界教育科学的珍贵财富。但毋庸讳言，由于受历史的制约和人所共知的原因，我国化学教育理论研究工作起步较晚，人员较少，机构又不够健全，对于教育实践中提出的诸多重大命题，以及国际上普遍关注的某些学术研究前沿课题，尚未从理论上给出回答，或未从理论与实践结合的高度上进行深入探研。理论来源于实践，理论一经群众掌握就会变成巨大的物质力量。化学教育理论也是如此。本人作为一名老化学教育工作者，有幸从 50 年代中期就参加中国化学会组织的有关化学与化学教育的学术研讨活动，从 1979 年至今一直亲身参与中国化学会化学教育委员会及中国教育学会化学教学研究会（1983 年起始）计划和组织召开的多次全国高、中等学校化学教育经验交流会、学术研讨会、课程和教材研讨会、化学实验教学经验交流会等。每次会议的论文均有数十篇，甚至百多篇，其中不少优秀论文已在国内期刊发表。从国际上看，IUPAC（国际纯粹化学和应用化学联合会）从 1970 年开始，已加倍努力于谋求改进世界各国的化学教育，并与 UNESCO（联合国教科文组织）协同召开过 13 次 ICCE（国际化学教育会议），出版了若干部化学教育论文集和论著。另外，世界各国的化学教育家在各自的研究领域，都笔耕不辍地为国际化学教育理论做奉献。1981 年据 64 个国家的不完全统计，各国创办的化学教育（教学）期刊就有 168 种。文苑书林，浩瀚得很！但是由于各种原因，国际上大量论文、资料难以为我国广大化学教育工作者检索、利用；国内的诸多专题研究论文和著述，似多局限于就某些论题的研究或偏重于适应教材建设的需要，对于化学教育学——化学教育理论体系的几大构成，尚未见到系统论述的著作面世。

为了建构具有中国特色的化学教育理论，反映国内外当前的研究水平，以促进我国教育改革，面向现代化，面向世界，面向未来，广西教育出版社组织出版了这套《学科现代教育理论书系》中的化学现代教育理论丛书。本丛书各册的第一作者和统稿人都是在该领域学术有成的专家。全体作者均本

着理论联系实际的原则，力求从化学教育规律来阐释和探研有关的理论与学术前沿课题。当然，作为化学教育理论著述，本丛书也完全可以作为大学后继续教育或化学教育高级学位研修用书。各册（《化学教育史》除外）论述的重点虽侧重于基础教育阶段的化学教育理论问题，但是从教育规律的普遍适用性这一层面来看，这些结论对于大学或大学后教育，以及中专等化学教育实践，可供借鉴之处当不是个别的。

本丛书共 6 册，简要介绍如下：

《化学实验论》以辩证唯物主义认识论、自然科学方法论、现代教学论为指导，论述化学实验的构成、意义和作用，剖析各类化学实验及其功能，探研化学实验与发展学生思维的关系，从宏观与微观的视角揭示化学实验及其方法论的深刻涵义。

《化学课程论》从化学课程的设计与化学教材编制的现实出发，探研不同课程论思想在化学课程开发的实践中运行与演变的规律及趋向，阐释、论述我国化学课程、教材建设中的基本经验与理论问题。

《化学教学系统论》运用系统论的观点阐发、研讨化学教学的构成要素及其相关领域的原理或范型，从多视角考察、概括化学教学系统的结构和功能及其运作圭臬。

《化学学习论》从化学学习系统与学习原理的高度探研化学学习过程、模式和方法，对化学学习能力与学习机制进行剖析，探索深入开展学科心理研究的某些基本课题。

《化学教育测量和评价》基于化学教育目标论阐发和研讨化学教育测量和评价的基本理论、方法和技术问题，对认知、情感和动作技能领域的化学教育测评等作了新的探索。

《化学教育史》以历史唯物主义和辩证唯物主义为指导，对化学教育产生的历史背景、化学史各时期的化学教育的演进，以及近、现代化学教育的发展等进行了研究、概括。以史为鉴，明古鉴今。

本丛书写作注意了：科学性，力求准确、完整、系统；新颖性，取材努力反映时代气息，体现教育改革精神；实用性，各册在介绍有关理论和研究前沿的同时，均力求结合实例给读者以解决实际问题的思路与方法。

本丛书在成书过程中得到不少同行的关心，并参阅、借鉴了不少国内外学者的研究成果，在此一并表示诚挚的感谢！衷心希望本丛书面世以后能够得到化学教育界的专家和广大读者的关注与指教，祈使这套丛书在加快、深化化学教育的改革和发展，发展大学后继续教育和活跃化学教育学术研究等方面，发挥它应有的作用。

1995 年 10 月于北京师范大学

## 总 序

顾明远

师范院校中有一门必修课，叫做教材教法。它是一门培养教师技能的专业课程，但是历来不受人们所重视。在一些专业学科的教师、专家们眼里，似乎教材教法不过是剖析中小学的教学大纲和教科书，教会师范生如何去上好一堂课，没有什么学术性。他们认为，上好一堂课，保证教学质量的关键主要是有高的学术水平。这是一种误解。但是这种误解不是没有缘由的。原因之一是，这些专家们不懂得，教育既是一门科学，又是一门艺术，只有高深学问，不懂教育规律，没有掌握教育教学的艺术，课就上不好，或者事倍功半。原因之二是，过去的教材教法课确实存在着不少问题，它只分析现有的教材，不对学科、课程以及教育教学的规律进行研究。因此要解决这个问题，除了改变专家们的误解以外，更重要的是研究这门学科的发展，提高学科的理论水平。我认为，师范院校的教材教法不能只分析一门课如何讲授，更重要的是要研究、分析一门科学的发展历史和现状，以及其发展的内在逻辑，结合学生的认知特点，遵循教育规律，把它组织成一门学科。学科并不等于科学。一门科学要变成学校里的学科，需要经过一番改造。改造的理论就是一门学问，本身也应该是一门学科。这门学科是跨学科的，它既要研究某门学科的科学规律，例如数学教材教法既要研究数学教学规律，又要研究教育规律，要把两者有机地结合起来，从这个意义上来讲，教材教法的名称显得落后了。因此把它改为学科教学论或学科教育学是适宜的。

讲到这门学科还有一段历史，不得不讲一讲。我国学位制度建立之初，在教育类门类中就设有教材教法作为二级学科培养研究生，授予学位。但是它的评议因为涉及文理各学科，因此分散在文理各学科评议组中。由于教材教法主要是研究学科教学的理论，文理各学科评议组的专家们认为难以对他们评议。这样这门学科的授权问题就处于无人评议状态。1983年在国务院学位委员会召开第二届博士、硕士授权点学科评议组会议期间，我向当时教育学评议组召集人刘佛年教授提出，把教材教法的硕士授权点拿到教育学组来评议，并把名称改为学科教学论，以提高对它的学术要求，从而提高它的学术地位。这个提议得到刘佛年教授的支持和学位委员会的批准，并在以后专业目录调整时把教材教法正式更名为学科教学论。从此学科教学论有了较大的发展。至今全国已有硕士授权点19个，培养了硕士研究生数百名，出版的专著也有几十部，这是十分可喜的现象。

学科名称的更改是十分容易的事，但要把它发展成一门真正的学科并非易事。当时有人提出改为学科教育学，我们认为时机还不成熟，首先要把学科的教学理论研究好。教育学是一个更广泛的概念，它涉及到教育系统内部各个领域，而学科教学论主要涉及教育系统中教学方面的理论，即使把这部分研究透彻，成为一门学科也是不容易的。当然，有的学者愿意把它称为学科教育学，如果确已研究成熟，这无疑是对教育科学发展的一个贡献。

把教材教法改造成为学科教学论是一次理论上的飞跃。教材教法过去只是教育学中的一个部分。学科教学论则变成了教育科学中的一个重要分支学科。这种飞跃有没有根据，具备不具备条件呢？1988年我在为《语文教育学》写序时就说，已经具备了必要的条件。这是因为：第一，近几十年来教学论、课程论、心理学、教育测量学、教育评价学等学科有了新的发展，它为学科

教学论的建立奠定了理论基础；第二，我国改革开放以来引进了国外的各种教学理论，开拓了我们的视野，启迪了我们的思想；第三，我国有一批长期从事教材教法研究的学者，他们在师范院校有长期的教育实践，积累了丰富的经验，并且有较高的理论修养，这是建立学科教学论的组织基础。应该说，1978—1988年这门学科的建设是有成绩的，不仅培养了众多研究生和出版了多部专著，而且学科体系基本上建立起来了。更为可喜的是不少专家都在关心这门学科的建设。得到各学科的专家的重视是至关重要的。因为学科教学论这门学科毕竟是跨学科的，文理各专业学科是它的基础。

近些年来，许多学者把学科教学论又提高到学科教育学的高度来研究，这又是一次飞跃。学科教育学不仅要研究学科的教学理论问题，而且要从教育学的基本原理出发，从培养人的高度来讨论学科教育的问题。它不仅要揭示学科教学的教学规律，还要揭示学科教学培养人的规律。学科教育学不仅要讨论该门学科如何设置课程，如何编制教材，如何选择教学方法，如何组织教学，更重要的是要分析本门学科在培养人的整体工作中的地位和作用，并从这个角度出发研究课程、教材、教法，研究它与其他课程的关系，与学校中其他教育活动的关系等等。

广西教育出版社组织全国学科教育理论工作者和实际工作者编写一套大型丛书《学科现代教育理论书系》，我认为正是时候。这刚好是十多年来的大总结，大检阅。证明学科教育学这门新兴学科已经在中国大地上成长起来。我当然不可能通览这套丛书，但是从编辑出版计划中的书目可以看到，它涉及语文、数学、物理、化学、外语等中学教学计划中的主要学科，每门学科又分教学论、课程论、学习论、实验论、教育测量和评价等专著，有的学科还著有教学艺术论及其他更细的内容，真是丰富多彩。作者群中有老一代的学科教育学专家，也有年青一代的学者。我认为，这套丛书的意义，不仅在于它总结了十多年以来我国学科教育学研究的成果，而且在于它展示了学科教育学发展的广阔前景，在于它培养了年轻一代学者。这是从教育理论战线上来讲的。至于对我国教育的实际来讲，这套丛书的出版一定有利于我国广大教师业务水平的提高，有利于教育质量的提高。我预祝出版的成功。

1996年春节

## 出版说明

这套丛书，从 1991 年 3 月出版第一批第一本《数学学习论》算起，至今已有 6 个年头了。如果从 1988 年年初开始数学教学理论丛书的组稿活动算起，则有 9 年之长。如今，数学、物理、化学、语文、外语，五个主要学科的教学理论丛书，已配套成龙，每个学科 6 本共 30 本，取名为《学科现代教育理论书系》。洋洋洒洒几千万字，构成了基础学科的基本理论研究，也构成了我社的基本骨干工程和基本的教育理论出版特色。

以近十年的时间建构一整套力求具有中国特色的教育理论丛书，其间的曲折、甘苦，自然一言难尽。但从反映教改成果、服务教学改革来看，又当义不容辞。从建构教育出版社的出版个性、出版文化来考虑，更有深刻意义，有重大价值。在改革开放的新历史时期，出版社靠什么来支撑？靠什么去竞争？靠什么求发展？用什么作奉献？答案可以有很多，对策可以开列不少。但根本的应少不了这么两条：一靠骨干工程，二靠名牌精品。骨干工程是出版社的战略布局，名牌精品是出版社的灵魂生命。两者的完美结合，构成了出版社的质量、信誉、知名度和文化品位，它是出版社存在的基础，竞争的手段，持续发展的后劲，文化积累的主体，向人民奉献优秀文化的根本保证。

本着这样的认识，这样的追求，我们出版了这套丛书。当然，还有另外几套别的系列。

我们期待着读者的鉴定。

我们迎接着市场的检验。

我们也渴望着教育界、理论界的支持。

我们将一如既往地努力，千方百计奉献更多的精品，给教育，给民族，给将来。

广西教育出版社

## 本书内容提要

本书在一般教育测量和评价理论、方法与化学学科特点密切结合的基础上，全面地介绍化学教育测量和评价的基本理论、方法和技术，以化学教学测量和评价为重点，讨论化学教育测量和评价研究的重要课题及其研究成果和发展趋势。写作上既重视理论探讨，又贴近实际，增加可操作性，为读者开展化学教育测量和评价活动及进行有关研究提供有益信息。

## 第一章 化学教育测量和评价概述

化学教育测量和评价是在一般教育测量和评价的原理、方法应用于化学教育领域的基础上逐步发展起来的。化学教育测量和评价既要接受一般教育测量和评价理论的指导，又有其不同于其他学科测量和评价的特殊性质。因此，本章将首先简要介绍教育测量和评价的产生和发展的历史进程，在此基础上讨论现代教育测量和评价的基本理论和方法，然后讨论化学教育测量和评价的基本特点及研究的现状与发展趋势。本章介绍的有关化学教育测量和评价的一般理论与方法，对以后各章内容的讨论都具有指导意义。

## 第一节 现代教育测量和评价的意义

### 一 教育测量和评价的产生和发展简介

#### 1. 教育测量和评价的产生

教育测量和评价的产生与对学生学力和学业成就的检测活动的发展有着紧密的联系，其历史源远流长。早在我国的西周时期，就初步建立了学校教育制度，开始了教育测评。到公元 606 年，我国隋朝开始实行了科举制。国外学者也认为，中国古代的科举考试是教育测评的最初萌芽。但是由于种种原因，本世纪以来，我国的教育测量和评价的研究一直未能发展起来；而在西方，其中主要是在美国，自 19 世纪后半叶以来，教育测量和评价得到了迅速的发展。

19 世纪上半叶以前，学校考试一般都缺乏标准和应有的客观性，其基本方法是对学生逐个口试。随着学生人数的增多，这种考试很难继续实施了。1845 年，美国波士顿文法学校首先引入了书面考试。但这时学生成绩评定的客观标准问题仍然未能得到解决。1864 年，英国教师费舍尔（George Fisher）收集了许多学生的成绩样本，编制了第一本《量表集》，作为度量学生成绩的标准；1897 年，美国的莱斯（J.M.Rice）博士发表了他对 20 个学校的 16000 名学生所作的拼字测验的结果；1905 年，第一个智力测验量表——法国的《比纳-西蒙量表》发表。1916 年，美国斯坦福大学心理学教授特曼（L.M.Terman）发表了对比纳量表进行修订后的更完善的智力测验量表——《斯坦福量表》，首次引入了智商的概念，使智力测量有了比较科学的计算方法，标志着心理测验已达到了比较成熟的阶段。

1904 年，美国心理学家桑代克（R.L.Thorndike）发表了《心理和社会测量学导论》，介绍了心理统计方法和编制测验的基本原理；1909 年，桑代克又编写了用于书法、拼字、作文、图画测验的标准测量工具，使教育测量开始走上了科学化的道路。因此，桑代克被称为教育测量学的鼻祖。1918 年以后，教育测量的使用范围逐渐从小学发展到中等以上学校，许多大学也开始设立了教育测量学课程。至本世纪 20 年代末，教育测量已发展到全盛期，教育测量在学力检测与教育成就的量化、客观化、标准化方面取得了很大的成绩。

随着教育测量运动的发展，其弱点也逐步暴露出来。由于当时的测验都是围绕着教科书编制的，测验内容只要求学生记忆教材的知识内容，未能包括对学生的社会态度、兴趣、情感的检查，不能反映学生的全面发展和实际需要，因而它很快引起了人们的不满和批判。

1929 年，美国经历了一场严重的经济危机，教育的社会效果问题越来越多地受到人们的关注，这就直接推动了以教育的社会价值为依据的教育评价活动的发展。从 1934 年始到 1942 年止，美国俄亥俄州立大学教授泰勒（R.W.Tyler）受卡内基基金会的资助，进行了历时 8 年的课程与评价研究，这就是教育评价发展史上著名的“八年研究”。在对以往的课程和测验设计进行了尖锐批评的基础上，泰勒提出了一套以教育目标为核心的课程和测验编制原则，试图以此把社会的要求、学生的需要反映在课程与测验中。为了把这一思想与早期的测量区别开来，泰勒和他的同事正式提出了教育评价的概念：“在本质上，评价过程乃是一种测量课程和教学方案在多大程度上达到了教育目标的过程”；即认为教育评价就是衡量教育活动达到教育目标程

度的一种活动，测量是它的手段。根据泰勒的理论进行的评价活动能够获得关于目标完成情况的信息，有助于发现存在的问题，改进教育工作。所以，这一理论受到了较为广泛的欢迎。

## 2. 教育测量和评价的发展

泰勒等人的教育测量和评价观点以及相应的“目标评价模式”是教育测评历史上第一个比较完整的、影响较大的理论。自30年代以来的几十年中，这一理论一直占据着主导地位。为了发展泰勒评价模式的实施技术，美国教育心理学家布卢姆(B.S. Bloom)等人提出了教育目标分类学理论，对完善教育评价理论起了重要作用。但随着测评实践和理论研究的发展，这一理论也逐渐暴露出一些根本性的缺陷。泰勒理论的核心是目标，但由于教育目标概念与教育结果的概念紧密相关，这就容易使人们将更多的注意力集中到对教育结果的评价上。因此，泰勒评价更多地被看作是一种对结果的总结性评价；然而，形成性评价可能有着更重要的作用。此外，人们还认识到，目标本身的合理性也需要评价，对那些预期目标以外的教育效果也不能排除在评价之外。

在对泰勒评价观进行分析和批判的基础上，又有人提出了新的教育评价观。克龙巴赫(L. J. Cronbach)修改了泰勒的评价定义，将评价定义为：“为作出关于教育方案的决策，收集和使用信息。”他强调：“评价能完成的最大贡献是确定教程需要改进的方面。”斯塔佛比姆(D. L. Stufflebeam)给评价的定义是：“为决策提供有用信息的过程。”并强调：“评价最重要的意图不是为了证明(prove)，而是为了改进(improve)。”在这一观点的基础上，他提出了著名的CIPP(Context-Input-Process-Product)模式，即把对背景、输入、过程、结果四个方面的评价结合起来的一种评价模式。受斯塔佛比姆等人的影响，美国教育评价标准委员会给评价的定义是：“评价是对某些现象的价值和优缺点的系统调查。”

## 二 对现代教育测量和评价概念的认识

从上述教育测量和评价概念的产生和发展历史看，首先是对学生学力和学业成就的评定需要，导致了教育测量的产生；初期的教育测量逐步发展为标准化测量，提高了测量的科学性和客观性；又随着社会对全面发展的学生的需要，教育测量开始走向教育评价，并从只注重结果的总结性评价，逐步发展为同时注重改进教育过程的形成性评价。西方的学者更强调评价为教育决策提供信息的功能，甚至拒绝把评价看作是一种价值判断过程，试图在更广阔的范围内给出评价的定义。

我们认为，现代教育测量和评价是在对教育目标、教育过程、教育结果（现实的或潜在的结果）以及影响教育的各种因素做出量（或质）的记述的基础上进行的一种价值判断活动。教育系统中的一切因素都是测评的对象；教育测量就是对评价对象的现状、属性与规律做出客观的、定性或定量的描

---

瞿葆奎主编：教育学文集·教育评价，人民教育出版社1989年版，第160页。

瞿葆奎主编：教育学文集·教育评价，人民教育出版社1989年版，第164页。

瞿葆奎主编：教育学文集·教育评价，人民教育出版社1989年版，第301页。

瞿葆奎主编：教育学文集·教育评价，人民教育出版社1989年版，第298页。

陈玉琨编著：教育评估的理论与技术，广东高等教育出版社1987年版，第26—28页。

述，这种描述的结果是对测评对象的某些属性分配数值；教育评价就是在教育测量的基础上，对评价对象做出价值判断。

这里，我们已经将教育测量和评价看作是两个密切联系、又相互区别的概念。测量和评价是同一活动的两个不同阶段。测量是一种事实描述过程，只对事物的属性赋值；而评价是一种价值判断过程，是在测量的基础上对事物的好与差、对与错、优与劣的判断。不过，测量只能是在评价需要的指导下的事实描述，否则，测量没有任何意义；而评价必须是在测量基础上的价值判断，否则，评价结果就是空中楼阁，毫不可信。正是由于教育测量与教育评价之间的密切联系，我们将两者合在一起，统称为教育测量和评价，或简称为教育测评。

教育测量是对测评对象进行客观的描述，尽可能地反映测评对象的本来面目，因而使测评活动带有客观性；然而，尽管教育评价是在客观事实的基础上进行的，但评价结论必然要与评价者本身对评价对象的主观认识密切相关，必然要反映评价者的主体需要和愿望，这就使教育测评活动不能不带有评价者的主体意识。可见，教育测量和评价实际上是一种客观性与主体性高度统一的活动。

### 三 教育测量和评价的地位和作用

为了更全面和深入地理解教育测量和评价的意义，有必要再从系统分析的角度，看教育测量和评价在教育系统中的地位和作用。

由于教育活动是“社会、心理和控制三方面的统一”，这三方面的因素可以构成教育系统的三个最基本的子系统，即目标系统、行为系统和控制系统。

教育目标是教师（教育者）和学生（被教育者）通过教育活动应当协同达到的教育质量，是指导教师和学生进行教学活动的基本依据。教育过程各个阶段教育目标的有机组合，就构成了教育目标系统，是教育目的的具体化。目标系统体现了教育系统中社会方面的成分，从方向、任务和内容三方面，决定了教师的教和学生的学。

行为系统是指教师和学生通过心智和信息的相互作用而实现教育目标的过程以及影响这一过程的各种心理因素、教育手段等所构成的整体。它包容了教育系统中心理方面的成分。教师和学生为实现教育目标而进行的相互作用就是通过这一系统而实现的。

控制系统是指由控制教育过程朝着实现教育目标方向发展所必需的各种教育手段和方法的最佳组合。该系统包容了教育系统中控制方面的成分。由于一切控制过程都是按照一定的目标对系统在状态空间中的各种可能的状态进行选择，使系统达到或趋近于被选择状态的过程；在选择之前，首先要确定在不同时间的教育系统的空间状态，而这一过程，实际上就是教育测量和评价过程。所以，教育测量和评价可以看作是该系统中的主要因素。

上述三个子系统的有机结合，即构成了一个完整的教育系统。教育活动过程则可以看作是由这三个方面的因素相互作用而构成的既有区别而又相互联系的三个阶段，即目标的定向阶段、目标的控制实现阶段和目标教学结果的测量和评价阶段。在不同的活动阶段中，教师和学生既有各自不同的活动

---

[苏] 尤·克·巴班斯基著，吴文侃等译：论教学过程最优化，教育科学出版社 1984 年版，第 6 页。



## 第二节 教育测量和评价的基本功能、实施程序与主要原则

### 一 基本功能

#### 1. 定向功能

教育目标体现社会需要和学生全面发展的需要；教育测评是以教育目标为依据，判断教育系统的功能是否实现。判断教师教学水平的高低，主要看他是否有效地帮助学生达到教育目标；判断学生学习成绩的好坏，也是主要看他达到教育目标的程度。这就促使教师和学生在进行教育活动时，必须以教育目标为准绳，保证教育过程朝着目标指引的方向发展。

#### 2. 鉴定功能

通过教育测量和评价，能够对学生的学习成就做出鉴定，选拔社会所需要的各种人才；也能够对教师的教育水平做出鉴定，为人事决策提供依据。

#### 3. 诊断功能

通过教育测量和评价，调查了解或验证教育过程中可能存在的各种问题，并诊断问题存在的原因，为制定解决问题的策略提供依据。

#### 4. 调节功能

教育测评所获得的大量信息，不仅可以用于选拔人才和人事决策，还能够用于对影响教育系统的各种因素之间的协调，使之更恰当地相互配合，以利于教育过程的优化。根据测量和评价信息，教师可以调整自己的教学内容、方法、程序，以利于学生达到教育目标；学生可以调整自己的学习方式，提高学习效率；教育管理部门可以制定管理措施、提高管理水平。

#### 5. 激励功能

教育测评是对教师的劳动成果和学生的学习成果的一种鉴定与评价，能够给人带来精神上的满足，也会对人产生压力或动力，提高教师和学生的教与学的热情，激励他们以更多的精力投入教育活动。

#### 6. 教学功能

教育测量和评价除了能够为协调和控制教育过程提供信息之外，本身也可以作为帮助学生达到教育目标的一种有效手段。测评的内容一般都是教学的重点；在测评过程中，学生对这些重点内容会进一步地记忆、思维、强化，从而会巩固和发展已有的学习成果。由于激励功能的存在，教育测评的这种教学功能有时甚至比其他的教学手段更加有效。

#### 7. 交流功能

只有在参加教育测量和评价活动的有关人员测评的标准有一致认识的前提下，测评活动才能顺利展开，并获得预期的成果。统一认识的过程就是一个同行交流的过程。在总结教育经验时，利用教育测量和评价对教育活动的全面、系统和客观的描述、评价，有利于使经验更为客观、科学，因而也更有说服力，更便于经验的交流和推广。

### 二 实施程序

#### 1. 准备阶段

(1)确定测量和评价目的。即解决为什么进行测评的问题。测评目的是测评活动的出发点。虽然教育测量和评价兼有多种功能，但教育的具体目标不同或是在教育过程的不同阶段，测评的侧重点并不相同。测评对象的确定和测评方法的选择，显然必须符合测评目的的需要。

(2)设计测量和评价指标体系。就是按照测评所依据的标准和要求，具体规定将哪些因素作为测评对象，同时确定各个被测评因素的相对重要性程度；测评因素构成了测评指标的内容，测评因素的相对重要性程度则表现在指标的权重系数上。所有测评指标及其权重系数的集合，即构成了测评的指标体系。指标体系的设计不仅是准备阶段，也是整个测评过程中的关键性工作。

(3)选择测量和评价方法。即根据测评目的和设定的指标体系的具体要求，选择搜集信息的方法和分析信息得出评价结果的适当方法。方法确定之后，就可以设计出测评的具体方案、步骤，准备所需要的测量和评价工具、表格、资料等，确定参加测评工作的人员并使他们明确自己的职责。

#### 2. 搜集测量和评价信息阶段

按照设计好的指标体系和测评方案，系统、全面地搜集所需要的信息资料。教育测量和评价中常用的搜集信息资料的方法有测验法、观察法、谈话法、问卷调查法、个案研究法等。

#### 3. 分析、判断阶段

经过第二阶段所得到的信息资料，一般都是比较庞杂、分散的，难以从中直接得出评价结论。因此，需要运用教育统计学原理，尽可能地利用电子计算机等现代化手段，对原始信息资料进行去粗取精、去伪存真的分析、处理、加工，在此基础上，形成综合判断，获得评价结果。

#### 4. 测量和评价结果的综合利用阶段

按照测评目的组织实施的测评活动，其评价结论主要应当满足预先的需要；由于教育测量和评价具有多种功能，其结果可以综合利用。一般来说，教育测量和评价结果可用于教育目标系统的优化、教学过程的改进、提供教育管理决策的依据和对测评过程及测评方法的改进等几个方面；还可以用作其他相关教育测量和评价活动所需资料的一个部分。

### 三 主要原则

教育测量和评价的主要原则，就是在进行教育测评活动时所必须遵循的最基本的行为准则。

#### 1. 目的性原则

测评活动必须有明确的目的，不能为测评而测评。测评的标准、方法、步骤都要根据测评目的的需要来确定。测评目的越明确，测评者和被测评者就越容易在测评目的指导下统一认识、协调合作，测评活动的效果就越好，效率就越高。一定要防止那种无明确目的和计划，或只凭长官意志、主观臆断的滥评现象。

测评活动的目的要正确。教育测量和评价作为教育系统的重要组成部分，以寻求教育活动的最优化方式作为自己的出发点和归宿。因此，一切教育测评活动的最终目的，都要有利于教育目标的实现。为此，设计测评方案 and 实际进行测评活动时，要注意使学生积极参与测评过程，指导和帮助他们进行自我测评，利用测评结果激发他们的学习积极性，并培养学生良好的学习习惯和方法。测评要有利于教师改进教学过程，提高教学质量，还要注意保护和提高教师的教学积极性。

#### 2. 科学性原则

教育测量和评价的全过程都要科学、合理，所采用的测评指标、方法、

步骤等都有一定的理论或实践基础，由此而得出的评价结果也是科学、合理、符合实际的。

要以教育目标作为测评的客观标准或确定测评标准的依据；根据教育目标确定的测评指标体系一旦确定就不能随意改动；对同一类测评对象要采用同一测评标准，而对不同的测评对象必须采用不同的测评标准；测评过程搜集的信息要尽可能及时、准确、丰富、全面；对测评信息的统计、分析，要尽可能采用有一定理论和实践依据的、合乎逻辑的方法；要注意教育测评活动中客观性和主体性统一的特点，尽可能防止主观因素或错误的价值观对评价结果的干扰和影响。

### 3. 实践性原则

教育测量和评价要有益于教育实践活动，其本身也能够经得起教育实践的检验，并在教育实际中切实可行。

在满足目的性原则和科学性原则的前提下，教育测量和评价活动不应该增加教师和学生太多的负担，测评的方法、手段要与实际的教育环境和条件相称。要尽可能少地占用教师和学生的时间、精力及物力、财力，但要尽可能多地获得测评的实际效果，即实现测评活动本身的最优化。实践性原则还要求，在相同的条件下，用相同的测评方式所得到的测评结果是相同的，也就是，测评活动具有可重复性，这样的测评结果才是真正可靠的。这实际上是教育测评的科学性原则在教育实践中的具体体现。

目的性原则保证了教育测评对教育系统的有效性；科学性原则保证了教育测评活动本身的可靠性；实践性原则保证了教育测评的可行性。因此，教育测量和评价活动必须遵循这三个最基本的准则。

### 第三节 化学教育测量和评价研究

#### 一 化学教育测量和评价的特点

化学教育测量和评价的基本性质与一般教育测量和评价是一致的。由于化学学科在其学科内容上区别于其他学科，在化学学科教育目标上必然有其特殊的要求，因此，以化学教育目标为依据的化学教育测量和评价在测评的对象、内容和标准上也表现出一定的特点，并要求有与之相适应的测量和评价方法。

我们认为，可以从以下几个方面来认识化学教育测量和评价的特点。

##### 1. 对学生记忆化学知识水平的测量和评价

化学是一门以实验为基础的学科。很多化学知识都是前人通过无数次的实验，在大量实验事实的基础上综合概括而成的科学结论。尽管现代化学已经发生了重大变革，化学理论也有了很大的发展，但是，化学学科还不能像数学、物理等学科那样有比较强的概括性，化学学习仍然离不开对大量化学事实的记忆。特别是在中学教育阶段，对很多重要的化学事实及定律、学说等很难让学生从根本上加以理解，对这一类知识的学习，必须依靠准确、牢固的记忆；学生能够接受的化学理论，也必须以充分记忆的化学事实为基础。因此，化学学习对学生记忆化学知识的水平有着特殊的较高要求。

根据化学学科的这一特征，化学教育测量和评价需要精心选择那些有代表性的测评目标，以便能够全面地测评学生记忆化学知识的范围及其牢固程度。

##### 2. 对学生理解和运用化学知识水平的测量和评价

由于化学是研究物质的组成、结构、性质以及变化规律的一门学科，其研究对象已经深入到物质的微观层次，往往是人的肉眼不能直接观察、其他感官不能直接把握的。为了让学生能够适应现代化学从基本上是描述性的过渡到推理性的，从主要是宏观的过渡到微观的变革，在化学教学中，必须要求学生能够突破一定的思维障碍，学会根据宏观实验现象或化学事实把握其微观化学本质，真正理解化学知识；同时，学会经过由宏观到微观的抽象概括过程运用化学知识解决化学问题。

根据化学学科的这一特征，要求化学教学测量必须注意测量的深度。要为学生提供展开抽象思维活动的问题情境，测评他们在由宏观到微观的思维层次上理解和运用化学知识的水平，而不能将测评范围局限于化学知识的再认或回忆的水平上。同时，也要考虑到学生思维发展的阶段性，不要超出他们的思维发展水平，以免挫伤他们的学习积极性。

##### 3. 对学生化学实验操作技能水平的测量和评价

化学是一门实验科学，培养学生的实验操作技能是化学教育不可缺少的重要目标。因而，对学生化学实验操作技能水平的测评必然是化学教育测量和评价的基本特征之一。实验操作技能属动作技能领域，其测评方法有着不同于认知领域的特殊要求。长期以来，由于忽视实验操作技能测评或是用纸笔测验代替实验技能操作考评，在一定程度上导致中学化学教学中违背基本规律、忽视实验教学、教学质量下降的恶果。

所以，化学教育测量和评价必须从化学教育的基本规律出发，认真探研测评学生实验操作技能水平的切实可行的有效方法。

##### 4. 对学生化学学习兴趣与态度的测量和评价