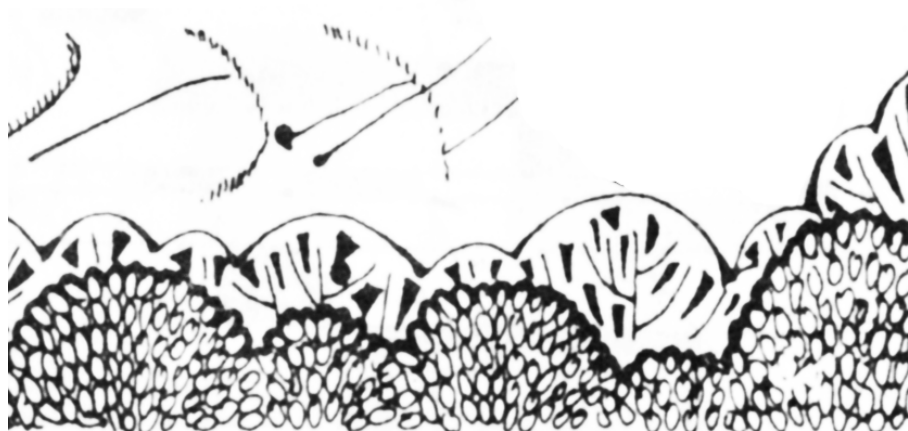


# 化学教育理论

## (五)

文强 编著



# 目 录

化学教育测量和评价 .....	1
概述 .....	1
现代教育测量和评价的意义 .....	1
其基本功能、实施程序与主要原则 .....	8
化学教育测量和评价研究 .....	13
化学教育目标与化学教育测评的标准 .....	21
化学教育目标概述 .....	21
化学教育目标的类别 .....	29
认知领域化学教育目标的学习水平分类 .....	35
化学教育目标的制订 .....	43
化学教学测试题的编制 .....	48
教学测量的基本原理 .....	48
命题计划的设计 .....	54
不同类型项目的编制 .....	60
不同学习水平项目的编制 .....	68
项目编排与试题的确定 .....	73
化学试题质量分析 .....	79
试题质量的初步评价 .....	80
试题的信度 .....	83
试题的效度 .....	89
项目质量评价 .....	95
化学测试结果的分析与评价 .....	106
学生集体测验分数的评价 .....	106
学生个体的测验成绩评价 .....	111
化学标准化考试 .....	120
标准化考试的基本理论 .....	120

化学标准化考试的实践.....	126
情感领域化学教育目标及测评 .....	133
情感领域的化学教育目标.....	133
情感领域化学教育目标测评 .....	138
学生学习化学的兴趣和态度 .....	147

## 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,这里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。这里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。

### 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定量化、

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标准，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方式,又相互配合、协调统一。教育系统的结构及其活动进程可如图 1-1 所示。教育活动进程具体体现为教师的教授进程与学生的学习进程,两者密切联系;而教育活动进程的三个主要阶段又分别与教育系统的三个子系统相对应。

总之,教育目标为教师和学生提供了一致的活动方向;教育测量和评价为防止偏离这一方向提供了有效的反馈控制手段;两者共同作用于教育过程,保证了教师的主导作用和学生主体地位的和谐统一,最后达到教育目标,实现教育系统的整体功能。

从上述对教育系统的简要分析不难看出:

(1)教育测量和评价是教育系统中不可缺少的组成部分。换言之,教育测评并非教育系统可有可无的,也不是教育系统外部强加的。只要有教育过程存在,就必然伴随着教育测评,否则,教育过程就不完整。认识到这一点,就可以增强我们开展教育测评活动的自觉性。对教育测评不需要讨论要或不要的问题,而是应当研究如何进行教育测评、使之符合教育系统功能的要求的问题。

(2)教育测量和评价的主要目的不仅是鉴定,而且还

有改进，从某种程度上说，改进比鉴定更为重要。因为教育的根本目的并不是对被教育者做出鉴定，而是要帮助他们全面发展。教育测评作为控制教育过程的主要手段，必须为改进教育过程提供必要的反馈信息，这是教育测评直接的、主要的功能；对学生学习成绩优劣、好坏的区别，与教师和学生的动机有紧密的联系，因而很重要，也是调控教育过程所需要的信息，但其毕竟不能直接而是间接作用于教育过程。认识到这一点，有助于我们积极、主动地利用和开发教育测评的全部功能，提高教育的质量和效率。

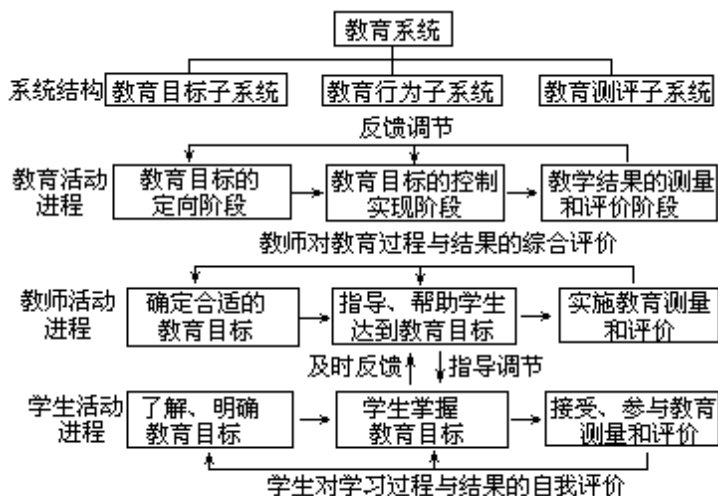


图 1-1 教育系统的结构及其活动进程

## 其基本功能、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，能够对学生的学成成就做出鉴定，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。

## 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