
图书在版编目 (CIP) 数据

高中化学课程评价 / 江家发, 阎蒙钢编著. —长春:
东北师范大学出版社, 2005. 6
ISBN 7 - 5602 - 4272 - 3

I. 高... II. ①江... ②阎... III. 化学课—教学
评议—高中 IV. G633.83

中国版本图书馆 CIP 数据核字 (2005) 第 060751 号

☐ 责任编辑: 李敬东 ☐ 封面设计: 宋 超

☐ 责任校对: 沙铁成 ☐ 责任印制: 张允豪

东北师范大学出版社出版发行
长春市人民大街 5268 号 (130024)

销售热线电话: 0431—5687213

传真: 0431—5691969

网址: <http://www.nenup.com>

电子函件: sdcbs@mail.jl.cn

东北师范大学出版社激光照排中心制版
印装

2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

幅面尺寸: 148 mm×210 mm 印张: 11 字数: 320 千

印数: 0 001 — 5 000 册

定价: 14.00 元

前 言

《基础教育课程改革纲要（试行）》的公布标志着新一轮课程改革的开始，《普通高中新课程方案（实验）》及《普通高中化学课程标准（实验）》的颁布，将普通高中新课程改革的落实提到了重要位置，也标志着基础教育化学新课程改革的稳步推进和全面实施。

2004年9月，已有4个省（自治区）率先进入高中新课程改革的实验，2005年高中新课程改革实验区将扩大到6个省（自治区），2007年所有高中起始年级也将进入新课程改革实验。普通高中教育是基础教育阶段中的最高层次，接受普通高中教育的学生面临升入大学或步入社会这两种对人生具有决定性意义的选择。本次高中课程改革将从根本上改变教育精英主义的倾向，把培养每一名高中生成为一个具有健全人格的人作为基本任务，让每一名高中生走向成功。这为我国高中教育的发展提供了前所未有的机遇，但也为各级教育行政部门管理者、校长、教师、教研人员提出了前所未有的挑战。

本次课程改革涉及学校课程的方方面面，对课程的目标、结构、内容、实施、评价和管理进行了全面的调整，其观念之新、范围之广、力度之大，是新中国成立以来少见的，也是我国近代教育史上少见的。尤其是高中阶段，新课程对普通高中教育进行了重新定位，强调了普通高中教育的基础性，明确地指出了高中教育是在九年义务教育基础上进一步提高国民素质、面向大众的基础教育，普通高中教育应为学生的终身发展奠定基础。

新的高中课程重新界定了高中课程的目标，要求高中课程致力于：帮助学生初步形成正确的世界观、人生观和价值观；培养其对中华民族的认同感和自豪感；培养其公民意识、社会责任感和道德修养；形成终身学习的愿望和能力，使其具有初步的人文与科学素养，创新精神和实

践能力；具有强健的体魄和顽强的意志，形成积极健康的生活方式和审美情趣；初步具有独立生活的能力、职业意识、创业精神和人生规划能力；正确认识自己，尊重他人，学会交流与合作，具有团队精神；理解文化的多样性，培养初步面向世界的视野和与其他国家、民族交往的能力。为此，高中新课程着重在以下几方面进行了改革：

(1) 转换课程功能。改变单纯注重知识传授的倾向，倡导积极主动的学习态度，在获得知识与技能的过程中，学会学习，学会生存，树立正确的价值观。

(2) 统整课程结构。克服传统课程学科本位、科目过多、缺乏整合的局限，整体设计课程门类和课时比例，设置综合课，加强实践性课程，采用由“领域、学科、模块”三个层次构建的课程结构。

(3) 精选课程内容。转变传统课程内容繁、难、偏、旧和注重书本知识的倾向，提倡课程内容与学生生活、社会和科技发展相联系，精选学生终身学习和个人发展必备的内容，突出高中新课程内容的基础性、时代性和选择性。

(4) 强化课程实施。高中新课程采用了以学分为单位的自主选择模式，强调学生主动参与、探究学习和合作学习，培养学生搜集和处理信息的能力、获取新知识的能力、分析和解决问题的能力以及交流与合作的能力。

(5) 改革课程评价。高中新课程评价应淡化甄别和选拔的功能，既要关注结果，又要关注学生学习的过程，关注情感、态度以及行为的变化。建立以促进发展为目的的多元、多样的评价方式，全面反映学生的成长历程和目标达成情况。

长期以来，化学课程评价由于受应试教育的束缚，逐渐远离学生本身和真实的化学教学情境，过分强调评价的甄别与选拔功能，忽视改进与激励的功能。评价内容过于注重学习成绩，而忽视对学生学习过程的考查。评价方法也相对单一，对“质性”评价方法不够重视，未能形成健全的教师、学校评价制度，过分强调教师评价学生的形式，忽视了学校管理者、教师、学生、家长的共同参与，严重制约了化学素质教育的推进。

高中化学新课程改革给教学评价以高度的关注，倡导“立足过程，

促进发展”的教学评价思想，是评价理念、评价方法与手段以及评价实施过程的根本性变革。高中化学新课程的实施是一次教育思想、教学能力和课程管理水平的挑战。从教师的角度看，过去熟悉的教学要求和教学内容改变了，传统的教学模式被打破了，衡量教学效果的评价方式也发生了质的变化。化学教师必须转变教育教学观念，掌握新的教学内容和方法，以适应新的教学评价制度。为使广大高中化学教师尽快理解基础教育新课程改革倡导的教育理念，以时代的责任感和使命感积极投身于新课程改革的探索和创新，我们编著了这本《高中化学课程评价》。

本书共分为八章，着重从理论与实践两个层面系统地论述了高中化学新课程评价的基本理论、评价的内容和方法、教师的评价、教材的评价、课堂教学质量评价、学生学业评价以及新课程高考评价改革等热点、疑点和难点问题。本书对广大一线高中化学教师、高等师范院校化学专业学生以及化学教育研究人员深入、具体地理解高中化学新课程的评价思想，探索高中化学新课程评价方法有参考价值和借鉴意义。

本书由江家发、阎蒙钢主持编写。杨明生、张勇、徐泓、江胜根、张宏伙、李琛、周卫东等为本书提供了案例。全书由江家发统稿和定稿。书中选取了不少来源于一线广大教师的实践素材，在本书出版之际，特向这些教师表示真诚的感谢。

限于作者的学识、水平和实践经验，书中一定有疏漏之处，敬请读者批评指正。

江家发 阎蒙钢

2005 年 6 月

目 录

第一章	
课程评价及其相关理论	1
第一节 课程评价理论的建立与发展	1
第二节 课程评价的理论基础	13
第三节 课程评价的功能、类型与实施	41
第二章	
高中化学新课程评价概述	47
第一节 我国高中化学课程评价实践中的问题 ...	47
第二节 我国高中化学新课程评价改革的 重点与方向	51
第三节 高中化学新课程评价倡导的基本理念 ...	59
第四节 高中化学新课程评价的基本要求 和实施建议	61
第三章	
高中化学新课程评价的方法	72
第一节 纸笔测验法	72
第二节 档案袋评价法	80
第三节 活动表现评价法	103
第四章	
高中化学新课程教师评价	109
第一节 高中化学新课程教师的素质要求	109
第二节 建立发展性化学教师评价体系	119
第三节 高中化学新课程化学教师的专业化 发展	135

第五章

高中化学新课程教材评价	147
第一节 教材评价理论概述	147
第二节 高中化学新课程教材评价	170

第六章

高中化学新课程课堂教学评价	192
第一节 高中化学新课程课堂教学评价概述	192
第二节 化学课堂教学评价指标体系的设计	199
第三节 高中化学新课程课堂教学评价指标 体系探讨	209
第四节 高中化学新课程课堂教学评价 的实施	230
第五节 高中化学新课程说课评价	243

第七章

高中化学新课程学业评价	252
第一节 高中化学新课程学业评价的内容 和方法	252
第二节 高中化学新课程学业考试评价的 类型	261
第三节 化学考试质量分析	279
第四节 建立促进学生全面发展的学生学业 评价体系	301
第五节 高中化学新课程学习方式评价	308

第八章

高中化学新课程下的高考改革评价	323
第一节 考试评价的素质教育功能	323
第二节 我国高考改革述评	325
第三节 高中化学新课程高考改革动态	334

参考文献	340
------------	-----

第一章 课程评价及其相关理论

新一轮基础教育课程改革正在如火如荼地展开，新理论、新观念、新思维、新方法扑面而来，在普通教育工作者看来，所有这些新事物似乎是层出不穷，令人应接不暇。其实，任何表面上看起来纷繁复杂的事物都能寻找到其主要脉络，所以不管新课程改革有多复杂，只要能抓住其基本的理论根据，一切困惑和疑问便会迎刃而解。作为开篇，我们在第一章将首先介绍课程评价的相关理论。

第一节 课程评价理论的建立与发展

简单了解考试、教育测量与课程评价的发展历史及其重要事件，可以让我们“以史为鉴”，更好地理解相关概念，开展课程评价活动。

一、中国考试制度对课程评价的贡献

1. 中国是考试制度的发源地

中国是世界四大文明古国之一，历史悠久，文化灿烂。中国是考试的故乡，是考试制度的发源地。据有关文献记载，我国在进入“文明时代”之后，学校即已出现。《孟子》一书中说：“夏曰校，殷曰序，周曰庠，学则三代共之，皆所以明人伦也。”这里的校、序、庠都是指学校。学校要衡量士子的学习成绩和文化知识水平需要考试，任官要判别智愚也要通过考选。无论是学校育才，还是政府选官，都要有一定的标准和方法来判断优劣与智愚。于是，在西周时期就建立了“考校”与“考选”制度。现在学者一般认为，考试制度在西周已见端倪，西周选仕是

我国考试制度的萌生阶段^①。在我国最早的教育专著《学记》中记载了距今 3000 多年前的西周时期的教育制度和考试制度，其曰：“古之教育，家有塾，学有庠，术有序，国有学。比年入学，中年考校，一年视离经辨志，三年视敬业乐群，五年视博习亲师，七年视论学取友，谓之小成；九年知类通达，强立而不反，谓之大成。”这段话的大意是说：古代的教育制度，20 户人家设一私塾，500 户的县设一学堂，12500 户的行政区设一学校，国都设一大学。大学每年招收学生，每隔一年考查一次，第一年考查学生分析课文的能力和志趣；第三年考查学生的专业思想是否巩固，同学之间能否相亲相助；第五年考查学生的知识是否广博，对教师是否尊敬；第七年考查学生研究学问的本领和识别朋友的能力，合格的就叫作“小成”；第九年考查学生对于学业是否能触类旁通，见解和行动是否能坚定不移，合格的就叫作“大成”。

可见，早在 3000 多年前的西周社会，学校教育就初步建立了考试评价的内容、标准和程序的框架。学校通过对士子的“考核”，将优秀者一级一级地往上报，直至“献贤能之书于王”。“选”出来的“贤者”与“能者”，可授予官职，这是“乡学”。至于中央的“国学”，虽然都是贵族子弟，但也要通过“考校”分出高低级别，“因其才而用之”。总之，西周社会发展具有重教、尚贤、选仕、授官等特点，孕育出的早期考试评价制度对后来的学校考试和用人考试，特别是对科举考试制度的建立，产生了重要而深远的影响。

春秋战国时期是我国社会处于大变革、大动荡的时代，原来的“考校”任官制度也逐渐被破坏。就教育而言，学在官府也渐被打破，私家讲学之风盛行。而私人设教，注重于平日的考察，所以大教育家孔子对他的弟子，尤其是“七十二贤人”，哪个品行最好，哪个学业最好，哪个长于聚敛，哪个长于外交，哪个长于斗勇，他都了如指掌。就任官制度而言，除了一部分官位仍然从国家或地方办的学校中加以“考核选拔”外，相当一部分官位是通过向社会“择善而用之”。何者为善？在《国语·齐语》中述及齐桓公用人的标准，即“有居处好学、慈孝于其父母、聪慧质仁、发闻于乡里者”，号召大夫们把这类人“举荐”出来。

① 杨学为等主编. 中国考试制度史资料选编. 合肥：黄山书社，1992

由此可见，春秋战国时期在选官制度与评价人才标准方面，不像西周时期那样有普遍的“考校”制度，但“学而优则仕”的思想仍然是存在的。此外，我们看到，春秋战国时期国家在人才评价方面较侧重于对士子的日常考察，其评价人才的标准似乎更加道德化与多元化。

两汉时期，我国的考试制度正式建立。汉代的“察举制”，是一种承前启后的任官制。所谓“察举”，就是察其贤能，举以授官。由此而产生了“对策”和“射策”两种策试（考试）方法。策试一般是“上书言事”和“口试、答辩”相结合。“射策”是事先出好多道题目，由学生随机抽取一道加以作答。“对策”一般是根据主考官提出的一个问题，用笔试方式加以作答，有时也用“口试”方式加以进一步考察。不过，在西汉时期，举“孝廉”则不要策试即可授官。到了东汉时期，举“孝廉”也要加笔试，即先“举”后“试”，以观其能。然而，两汉时期的学校已具相当规模。多数官办的学校设有各种考试科目，考得好的可“补官”或进一步深造，考得差的“留级”，有的还要“补考”。据《文献通考·太学》中记载，汉代有一时期学校及社会盛行考试，不仅太学和郡国学里的学生要考试，就是太学里的老师——博士，也是要事先通过一定方式的考试评价后方能选用，即所谓“始试而后用，盖欲其为人之师范，则不容不先试其能否也”^①。

魏晋南北朝时期，由于长期战乱，疆土分裂，政局不稳，以致文教不兴。学校考试制度与方法也没什么重要发展。只是这一时期，在任官制度上，淡化了“察举制”，实行“九品官人法”，即九品中正制，其特点是以中正为中心，论品定级，选拔和升降官吏。具体办法是各州设大中正，郡设小中正，对其管辖区域内的人物，按家世门第高低，才德优劣，以上上、上中、上下、中上、中中、中下、下上、下中、下下九等划分品第，并逐级向吏部推荐。吏部根据中正报告，按品授官。中正还有权对管辖人物依其言行修著或德义亏缺情况分别升降进退。九品中正制度实行之初，尚有选拔人才的作用，但各级中正官后来被世家豪族占据，因而九品中正制度便逐步演化为“门阀制度”^②。九品中正制选拔

① 杨学为等主编. 中国考试制度史资料选编. 合肥：黄山书社，1992

② 廖平胜主编. 国家公务员考试大词典. 北京：中国人事出版社，1996

人才主观性较强，较少使用考试方法。总之，这个历史时期的考试制度和人才选拔评价制度并无多少可称道之处。

二、中国科举制度的世界地位

中国古代科举考试制度对世界教育、政治和文化，尤其是考试制度产生过重大的影响。法国大革命时期的资产阶级启蒙思想家伏尔泰（Voltaire）曾经对中国的科举制度倍加赞扬，他说：“人类精神，肯定想象不出比这样的政府更好的政府。在这个政府里，重要的衙门彼此统属，任何事情都在那里决定，而其成员，都是先经过几场严格的考试的。”^① 欧美各国采用考试办法选拔官吏，是 18 世纪末 19 世纪初从我国学去的。

所谓科举制度，就是“开科取士”的制度，即国家设立许多科目，通过定期的统一考试以选拔人才的一种制度。科举制度始于隋炀帝大业二年（606 年），直到清光绪二十九年（1903 年）举行最后一科进士考试后宣告结束。科举考试制度在我国历史上绵延将近 1300 年。

隋炀帝大业二年创进士科取士，当时只用策试方法考时务策一门。但考试还不是唯一的依据，取士伴有“行卷”，即考生将自己平时作的优秀诗文，主动地献给有关当权者或知名人士甚至主考官，求得权威人士的青睐和帮助，以利录取。行卷和考试是影响考生及第与否的两大因素。

唐承隋制，并对科举制度大加发扬，使科举制度日趋完备。唐高祖李渊即位便开科取士，到太子李世民登基，便成定制。当时开考科目很多，其时常贡科，包括秀才、俊士、明经、进士、明法、明算、明字、武举等。另有特设科，包括道举（玄学）、童子（10 岁以下能通一经）、一史（考《史记》）、三史（考《史记》、《汉书》和《后汉书》）、开元礼（通习开元时期的礼仪）、三礼（通习《礼记》、《周礼》和《礼仪》）、三传（考《左传》、《公羊传》和《穀梁传》）等几十种不同的考试科目^②。

唐代科举考试不仅科目多，当时“制举之名多至八十有六，凡七十

① 郑日昌，漆书青等. 考试的教育测量学基础. 北京：高等教育出版社，1990

② 王志东. 科举故事百则. 海口：海南出版社，1991

六科”（《困学纪闻》），而且各科考试内容与方式有不同的规定。例如，明经科考试以儒家经义为主，以强记博诵理解为旨；明经科考帖经，即从经书中抽出一句，命考生将上下文默写出来；还考墨义，即对经文字句的简单作答，相当于我们现在的填空、默写和简答题。进士科考时务策，后来又加考“杂文”、“诗赋”等文体。明经与进士两科最受世人青睐，所以竞争也最激烈。明经约十人取一，进士约百人取一。明字科，也就是“明书科”，考试内容主要为《说文解字》、《字林》等语言文学或文字学的经典著作。明算科考《九章算术》、《周髀算经》等算学内容。明法科的考试内容主要是关于法律、法令方面的条文规则。武举始于唐代武则天长安二年（702年），考试内容既有文化知识考试，也有技能考试，只是唐代的武举未成定制。

唐代的科举考试，其考生通常有三类，一是“生徒”，即国子学、太学及各地学校的学生，只要经过学校考试合格，就可以直接到尚书省参加“省试”。二是不在校的读书人。通过所在郡州考试的读书人称为“乡贡”。“乡贡”考试合格后再到尚书省参加“省试”。被送去参加省试的考生，可称为“举人”。“省试”考中进士后，还须经过吏部考试——“选试”，合格后，才能奏请皇帝授予官职。三是有某种专长的知名人士。经人推荐给朝廷后，由皇帝特意下诏考试。

唐代以后的各个时期，科举考试在考试科目、内容、方式等方面虽然不尽相同，但就科举考试制度整体而言，基本上未出唐代的范围。五代纷争时期，科举考试基本上是沿袭唐制。元代的科举制度一波三折，断断续续。明代科举考试再度兴盛，实行四级考试制度，即院试、乡试、会试和殿试。在院试以前有县试和府试，及格者称“童生”。童生参加省、府所在地的地方书院考试，即院试，及格者称“生员”，生员与国子监的监生有参加乡试的资格。乡试三年一考，由皇帝钦命的主考官主持，及格者称“举人”。全国举人在京都参加由礼部主持的京都会议，考中者称“贡士”。贡士再到奉天殿参加皇帝亲自主持的殿试（策试），然后分“三甲”出榜，均称“进士”。“一甲”三名，其中第一名为“状元”，第二名为“榜眼”，第三名为“探花”。明代科举最大的特点：一是多级考试，层层选拔，组织有序；二是实行“八股取士”，就是以《四书》、《五经》中的文句做题目，依照题义阐述其中的义理，并

限制为用一定的格式、体裁、语言、字数做应试文章。至于清朝科举，基本上沿用明代旧制。由于种种原因，科举制度最终在清朝后期寿终正寝。

科举制度在中国历史上延续了近 1300 年，形成了较完备的考试制度和独特的考试方法，并且积累了很丰富的经验，可供后人借鉴。其中包括“锁院”制度、“糊名”制度、“誊录”制度、复试制度、多级选拔制度、回避与人闹制度等，这些制度均为的是追求考试的公平性和真实性。尽管科举考试有许多消极的侧面，如考试内容僵化、不重视科学技术等，但用考试的方法来选拔人才，可使许多平民百姓通过自身努力跻身于上层统治集团。据统计，明清两代进士中平均有 42.9% 的及第者出生于从未有过功名的家庭^①。不可否认，科举制度作为一种考试和人才选拔制度，不仅对我国古代政治、文化、教育、军事起了重要的促进作用，而且对民族的融合和国家的统一也起到了一定的作用。

科举制度也对世界文化发展和考试制度的建立起过重要的影响。可以说，中国古代科举制度不仅仅是中华民族的文化遗产，而且是世界的文化极品之一。它曾经对西方文官考试制度的建立和现代教育测量与评价的发展，起到重要的作用。事实上，欧美各国从早期开始就一直对中国的科举制度感兴趣。例如，意大利旅行家马可·波罗曾于 1275 年沿着丝绸之路千里迢迢来到中国，在元朝任官达 17 年之久，对中国的文官制度非常熟悉，在他的《马可·波罗游记》一书中作了详细的介绍，并大加赞扬，使西方人有机会了解中国科举考试制度的文明与精华。英国在 1570~1870 年期间，出版的有关中国文官制度和政治制度的书籍达 70 多种，大量介绍与研究中国的科举制度，极力主张英国政府实行中国式的文官考试制度。终于在 1855 年英国建立了文官考试制度。其后美国、意大利等国家也相继建立了文官考试制度。诚如孙中山先生指出的：“现在各国的考试制度，差不多都是学英国的。穷流溯源，英国的考试制度，原来还是从我们中国学过去的。”^②

中国科举考试虽然对世界的文官考试制度有过重要的作用，但缺乏

① 刘虹. 中国选士制度史. 长沙: 湖南教育出版社, 1992

② 杨学为等主编. 中国考试制度史资料选编. 合肥: 黄山书社, 1992

从教育学、心理学、统计学、哲学等科学视角的反思与改进，因而，还不是现代意义的教育测量与评价学科。目前我国的教育测量与评价理论，主要是辛亥革命后，从西方传入我国，并与我国深厚的考试文化、考试实践相结合，形成了有中国特色的考试评价制度，促进了我国教育测量与评价学科的建设发展。

三、课程评价理论的建立与发展

1. 教育测量和教育评价概念的建立

欧美国家早期的学校教育考试，大多使用口头测验。英国剑桥大学直到1720年才开始使用笔试，这比中国要迟1000多年。美国于1840年在波士顿市第一次在全市范围的学校中进行书面考试，从而进一步认识到实施统一的书面测验具有许多优点。欧美国家推广笔试虽然对形成标准化测验的思想有一定影响，但对教育测量科学的诞生实际上并未产生实质性影响。对教育测量科学的建立有直接影响的，实际上是欧美国家一些学者为追求学校成绩评定客观化与标准化而开展的理论研究和实践探索。

早在1864年，英国一位名叫费舍（G. Fisher）的教师，曾用心广泛地收集学生的书法、拼写、算术、文法、作文、历史、自然、图画等不同水平的作业样本，然后选择有代表性的样本汇编成册，并为每个水平的作业样本评定一种分数，以示优劣，形成《作业量表集》，作为教师以后评定学生各科成绩的参照标准。美国学者赖斯（J. M. Rice）也是主张用统一的测验去考查、比较、评定各校学生的成绩。他从1894年开始，用若干年时间，编制了算术、拼写、语言等测验，曾有数万名的学生接受过他的测验。据此，赖斯开展了一些相关的教育研究项目，在会议或刊物上发表研究论文，引起了许多对教育测量感兴趣的学者的重视。还有来自英国的个别差异研究、德国的实验心理学研究、法国的异常心理及智力测验研究、美国本土化的智力测验研究等一些研究力量，共同对教育测量学科的诞生起过孕育和催生的作用。美国心理学家桑戴克（E. L. Thorndike）集欧美各国的研究与思想于一身，于1904年出版了闻名于世的论著《精神与社会测量导论》一书。在这本书中，他提出一个著名的假设：“无论什么东西，只要存在的，总存在

于数量之中。(Whatever exists at all exists in some amount.)”并首次较系统地介绍了教育统计方法及编制测验的基本原理。这本著作被公认为教育统计学、教育测量学、教育评价学等学科的第一本著作，它标志着以科学理论（包括统计学、教育学、心理学、哲学等）为指导的教育测量理论的诞生。对教育统计、教育测量、教育评价等学科的诞生与发展具有划时代的意义。

最早呼吁人们将注意力从“测验”转向“评价”的是美国的教育家泰勒(R. W. Tyler)。20世纪30年代初，美国进步主义教育协会在全国挑选了30所中学，进行课程改革研究，历时8年(1933~1940)，又称“8年研究”。泰勒教授担任了“8年研究”评价委员会主任，为新课程实验设计了一套有利于教育测验和考试的教育评价方法。因此，现代教育评价的观点诞生于20世纪30年代的“8年研究”。从20世纪30年代到50年代末，泰勒的评价思想一直在教育评价领域占主导地位。

资料：测验、测量、考试、评价等术语的相互关系

测验是引起某种行为的工具。通常是指运用某些仪器、试题来引起人们的某种行为，从而测定人们的某种特征，它是进行数量化分析和科学推断的前提和手段。测量比测验的含义要更为广泛，测量不仅包括着运用仪器、试题来测定事物的质量与特性，而且包括运用调查、观察等方法来测量事物的质量与特性，不过在以往的测量书中，测验与测量常互相通用。

测验与考试也不尽相同，我们平时所说的考试，通常只凭教师自己的经验去出题评分，带有主观随意性，而测验是经过较细致的科学分析才编制出题目，在测验的程序和评分方面也有较严格的要求。考试一般用于考核学生的学业成绩，测验不仅用于此，还用于心理特征的测量。

当然，离开教育测量学的观点去谈考试与测验，二者在有些场合没有区别，如期中考试与期中测验。但有些场合不能混用，如“高考”不能说成“高测”，“心理测验”不能说成“心理考试”；二者出现的时间也不同，“考试”一词于东汉时出现，而“测验”一词直到20世纪初在中文里才有。

一般而言，测验的应用较广，它不含淘汰的意思，其目的在于了解情况，如小测验和时事测验，而考试用于比较正式的场合，二者可以不加区别。

测量与评价既有联系，也有区别，测量是对客观条件的一种“铁面无私”的数量化描述，重在量的测定，有较大的客观性，但若对其结果不进行评价就是无意义

的东西。所以测量的结果是评价的重要依据，但又不是唯一的依据。因为评价比测量的范围要大，评价需要质量并重，主客观兼容，依据更多的方法和资料进行，如对学生毕业成绩的评价，除依据测量所得分数外，还要依据学生的平时作业、课堂回答、以往的学习成绩等情况。

2. 课程评价及其相关概念

在教育部颁发的《基础教育课程改革纲要（试行）》中，使用了“课程评价”一词，这是新中国成立后首次在国家发布的课程文件中这样使用。那么，什么是课程评价呢？

首先，课程评价的定义随着课程理论和教育实践的发展而呈现出阶段性特征，在一定的历史时期，大家对课程评价有着较为一致的看法。按照美国学者古巴和林肯的观点，课程评价的由来和形成大致分为四个时期^①。

第一个时期是测验和测量时期，从 19 世纪末到 20 世纪 30 年代。在这个时期，教育评价、课程评价的概念还未出现，盛行于教育界的是教育测量、心理测验这样一些概念。人们将对学生的评价等同于对学生进行智力测验，当时很有影响的如比奈编制了筛选智力落后儿童的智力量表，桑戴克编制了第一批标准化的教育测验。

第二个时期是描述时期，从 20 世纪 30 年代到 50 年代。20 世纪 40 年代，泰勒总结了 30 年代进行的“8 年研究”的经验和成果，发表了课程史上具有里程碑意义的重要著作《课程与教学的基本原理》。在这本著作中，泰勒正式提出了评价的概念，并提出了评价原理。正如泰勒描述的那样：“人们普遍接受用教育评价来替代测验。‘8 年研究’提醒教育工作者，教师在教一门课时，通常都寻求达到若干个教育目标，但凭一个测验分数是不能客观地概括教学结果的。通过使用问卷、观察、产品样本和测验，都可以评定学生在每个主要目标上进展的情况。这已成了一种普遍的评价观。”^②在泰勒的努力推动下，评价一词取代了考试、测验这些提法。这个时期普遍认为评价就是客观地描述教育结果与

① 李雁冰. 课程评价论. 上海: 上海教育出版社, 2002

② 泰勒. 课程与教学的基本原理. 施良方译. 北京: 人民教育出版社, 1994

教育目标的一致性程度，可以根据教育结果与教育目标之间的差距来调整课程教学，而描述教育结果不能仅仅依靠测验和考试。

第三个时期是判断时期，从美国因 1957 年苏联卫星上天而发起的教育改革始，持续到 20 世纪 70 年代。这个时期，课程研究者对第二个时期的评价概念进行了几点反思，一是如果将教育目标作为评价的标准，那么教育目标本身能否免于被评价，教育目标本身的合理性如何保证。也就是说，是否应对教育目标本身进行价值判断。二是以教育目标作为评价的标准是否合理，对于超出教育目标的教育结果如何评价。三是仅仅评价教育结果是否足够，我们是否应该关心教育的过程。因此，这个时期对于课程评价的定义发生了一些变化，主要是提出评价中应该包含价值判断的成分，其中既包括对教育结果进行价值判断，也包括要对教育目标、评价标准进行价值判断，同时还强调了教育过程本身具有的内在价值。

第四个时期是建构时期。古巴和林肯在批判前三个时期的评价基础上，提出了第四代的评价概念。他们认为：“评价在本质上是一种通过协商而形成的心理建构，因此，评价应坚持价值多元性的信念，反对管理主义倾向。”^①其主要特征是以“新的、文化人类学的研究范式取代旧有范式”，以质性方法替代传统的定量的方法。

我国教育理论工作者从 20 世纪 80 年代后才开始引进西方的课程理论，并根据我国教育实践提出了课程评价的定义。例如，陈侠认为，“课程评价是一个客观的过程，它要应用科学的工具来确认和解释教与学的内容和过程的效果，衡量它们的有效程度，以便为课程的改进作出有根据的决策^②。”这个定义基本上是属于古巴和林肯所指的描述时期的评价概念，因为陈侠先生的《课程论》是新中国成立后第一本课程研究的专著，由于研究资料的限制，当时借鉴的主要是泰勒的课程评价理论。其后，廖哲勋提出课程评价是对课程作价值判断^③。施良方提出课程评价是指研究课程价值的过程，是由判断课程在改进学生学习方面的

① 张华. 课程与教学论. 上海: 上海教育出版社, 2000

② 陈侠. 课程论. 北京: 人民教育出版社, 1989

③ 廖哲勋. 课程学. 武汉: 华中师范大学出版社, 1991

价值的那些活动构成的^①。这两个定义显然属于古巴和林肯所说的判断时期的课程评价观。

我们再从逻辑的角度来分析一下课程评价这个词。课程评价这个词是由“课程”和“评价”这两个词合成的，也就是说，课程评价的概念内涵是由课程的概念内涵和评价的概念内涵相交而成的，而课程这一概念有几十种不同的定义，因此课程评价的概念也就十分复杂了。施良方曾把课程的定义归纳为六种类型，即“课程即教学科目”，“课程即有计划的教学活动”，“课程即预期的学习结果”，“课程即学习经验”，“课程即社会文化的再生产”，“课程即社会改造”。显然，相应于这六种课程的概念，课程评价的含义也是相差很大的。如果认为课程是教学科目，那么课程评价势必是针对学校中具体科目的评价；如果把课程看作有计划的教学活动，看作学生应学习的学科总和及其进程和安排，那么课程评价就是指对课程计划、大纲、教材、教案等文件课程的评价；如果把课程理解为预期的学习结果，那么课程评价就是对学生学习结果的测量和评价（如泰勒）；如果把课程看作学习经验，那么课程评价的范围就要大得多，不仅包括对学习内容、学习结果的评价，还包括对学习方式、学习过程的评价。总之，只要课程的概念不统一，课程评价的概念也就不可能严格确定。

如何理解课程评价，关键是要看如何理解课程的概念，对课程概念的理解决定着对课程评价概念的理解。我国有学者认为课程评价有广义和狭义之分^②。

狭义的课程评价“特指对课程计划、课程标准、教材在改进学生学习方面的价值作出判断的活动或过程，一般包括对课程目标体系的评价、对课程计划的评价、对课程标准的评价、对教材的评价等核心内容，它的实施一般是由受过专门培训的评价人员借助于专门的评价方法和技术而进行的。”

广义的课程评价“即教育评价，是指按照一定的价值标准，通过系

① 施良方. 课程理论. 北京：教育科学出版社，1996

② 周卫勇主编. 走向发展性课程评价：谈新课程的评价改革. 北京：北京大学出版社，2002