

序 摇言

新世纪，我国社会的政治、经济、文化、科技与教育正在全方位地转型，在原有历史背景下发展起来的我国师范教育面临着严峻的挑战。虽然在过去的半个世纪的时间里，师范教育为我国的基础教育提供了庞大而稳定的师资队伍，为我国社会与教育的发展奠定了重要基础，但历史发展到今天，伴随着知识经济的扑面而来与教育国际化的发展趋势，面对着 严峻的严峻挑战，我国基础教育对教师的要求已由数量上的扩张转向质量上的提高，带有强烈计划经济色彩的，以终结性、封闭性、垄断性为特征的我国师范教育已远远跟不上世界发展的潮流，取而代之的将是以终身教育思想与教师专业发展理论为依据的教师教育。教师教育的理念将教师一生的成长与发展视为一个持续发展的过程，并以教育的专业化、一体化、开放性、终身性为自己的重要特征。

在新的世纪里，我国教师教育如何突破传统的思维定式，构建既符合我国国情又能融入世界潮流的教师教育体制，将成为我国教师教育发展的关键。而构建新的教师教育体制最终要落实在课程改革上。纵观中外教育改革的历史，也无不把教师教育课程改革放在突出的位置。近年来，世界许多国家特别是一些发达国家，他们无论是反思本国教育的弊端，还是对教育发展提出新的目标和要求，往往都是从改革教师教育课程入手。通过改革教师教育课程，调整人才培养目标，改变人才培养模式，提高人才培养质量。教师教育课程改革是整个教师教育改革的核心和主导，它集中体现着时代教育思想和教育的价值观。

新的基础教育课程改革突出课程的整体性和综合性，这就要求我们的教师教育课程改革要审时度势，调整专业设置，建立全新的课程体系。在保证一般文化课和学科专业课程占主体的同时，加强教育学科课程、教育技能课程、教育实践课程和教育专业课程的比重，以使未来教师具有扎实的基础知识和专业化水平，精通现代教

育技术，增强他们的适应能力。

在教师教育的改革中，教材改革起着举足轻重的作用。教材是新的教师教育理念的集中体现，是新的课程体系的载体。因此，为了将我省的教师教育改革推向新的高潮，为了提高从教者的专业素质，特别是教育专业素质，我们组织编写了这套教师教育专业系列教材。

该套教材共 5 本。教材涉及教育哲学、教育史、教育经济学、教育社会学、教育管理学、教育心理学等方面。

该套教材有以下三个特点：

① 内容新颖。能够反映本专业最新的有价值的研究成果，这些成果既有宏观的教育学、心理学理论，也有微观的实验研究。

② 重点突出。突出教育学、心理学经典的有代表性的研究，而不是面面俱到。这样可以使读者在有限的时间里，把握教育学、心理学的主要内容。

③ 理论与实际相结合。一是心理学的理论与教育学的理论相结合，它不是单纯的学习心理学，也不是单纯的教育学，而是二者的有机结合；二是心理学的理论与基础教育改革相结合，达到心理学为基础教育服务的目的；三是心理学理论与教学方法相结合，突出心理学理论向教学方法的转化，而重在方法。

总之，这套教材既体现了心理学的理论基础，又突出了它的实用性和可操作性，能在理论和实践中架起一道桥梁，消除中小学教师和学生以及其他爱好者把理论运用于实践时遇到的迷惘和困惑。

该套教材的阅读对象是高等师范学校的学生、中小学和幼儿园教师以及其他教育学和心理学爱好者。

韩清林

二〇〇四年 九月 十四日

前 言

教育评价是教育科学研究领域的三大课题之一，研究范围和内容日臻扩大和深化，学科地位也愈益被重视，但学科建设尚不完善。由于其学科具有较强的应用性特点，需要学科专家及广大的教育实际工作者共同构建和完善。

国家《基础教育课程改革纲要》（试行）指出，要“改变课程评价过分强调甄别与选拔的功能，发挥评价促进学生发展、教师提高和改进教学实践的功能”。新课程要求开展评价活动，要打破传统的评价观。在评价目的上，由侧重甄别和选拔转向侧重发展，从关注结果的评价转向关注诊断性评价和过程性评价；在评价方法上，强调多样化，注重把质性评价与量化评价结合起来，以质性评价统整量化评价；在评价主体上，强调评价主体多元化，注重自评和互评的作用，强调评价对象的参与性，改变过去那种自上而下、评价对象被动接受评价结果的状态，等等，对我国广大的中小学教师和教育工作者提出了新的要求和挑战。顺应课程改革，不仅要求广大教师及教育工作者掌握教育评价的一般技术、方法，还要更新教育和评价理念。为此，本书设置了三个领域供教师学习，一是教育评价的基础、原理方面，了解教育评价的内涵及基本原理。二是技术和方法方面，掌握教育评价的一般技术和方法，为开展教育评价活动提供技术支持。三是教育评价的诸实践领域方面，包括课堂教学评价、教师评价、学生评价、学校评价，探讨教育评价的实际应用。

本书编写中力求使内容平实，强调方法技术的可操作性，站在理论前沿，贴近改革实践。为便于教师学习，采取了如下编写体例：每章前设计“本章学习提要”，概括本章的主要内容和观点；每章正文后一方面提供重点“思考题”，另一方面对书中的重点或尚存争议的观点进行“相关问题导读”，为读者深入学习提供更广泛的信息支持；各评价领域章末设“附录”，主要提供一些经验性

的资料或案例，提高本书的应用价值，便于学习者借鉴、采用。

本书由史晓燕主持编写，并负责全书的修改与统编。参加编写的人员为：史晓燕（河北师范大学）第一章、第七章；雷彦兴（河北师范大学）第二章、第三章；王贺兰（河北省教科所）第四章；张琴秀（山西师范大学）第五章；杨雅清（河北省教科所）第六章。

本书编写中引用、参阅了有关专家的研究成果及有关的教育实践经验案例，谨向原作者致谢！同时由于我们的理论水平和实践经验所限，书中难免疏漏、失当，敬请读者斧正！

编摇者

圆园园年 猿月

目录

第一章 教育评价基础	(员)
第一节 教育评价概念	(员)
一、评估与评价	(圆)
二、教育价值与教育评价	(猿)
三、教育测量与教育评价	(源)
四、对教育评价的不同理解	(缘)
第二节 教育评价的历史沿革	(远)
一、教育评价的源流	(苑)
二、教育测量运动	(愿)
三、现代教育评价的产生	(员)
四、现代教育评价的发展	(猿)
第三节 发展性教育评价	(圆)
一、发展性教育评价思想基础	(圆)
二、发展性教育评价的基本特点	(圆)
三、国家基础教育课程改革与发展性教育评价 ...	(圆)
第二章 教育评价的基本原理	(猿)
第一节 教育评价方案的设计	(猿)
一、教育评价方案的含义	(猿)
二、教育评价方案设计的基本内容	(猿)
三、教育评价方案设计的意义	(猿)
第二节 教育评价的程序	(猿)
一、教育评价的准备	(猿)
二、教育评价的实施	(圆)
三、评价结果的处理和反馈	(圆)
第三节 教育评价的再评价	(猿)
一、教育评价再评价的意义及内容	(猿)
二、再评价的标准	(猿)
三、再评价的实施	(猿)

第三章 教育评价的方法与技术	(远)
第一节 评价指标体系的建立	(远)
一、教育评价指标体系	(远)
二、设计教育评价指标体系的原则和步骤	(远)
三、教育评价指标体系的设计	(远)
第二节 评价信息的收集方法	(远)
一、观察法	(远)
二、访谈法	(远)
三、问卷法	(远)
四、测量法	(远)
五、文献法	(远)
六、其他方法	(远)
第三节 教育评价信息的处理与整合方法	(远)
一、价值判断的一般方法	(远)
二、教育评价信息的定性处理方法	(远)
三、教育评价信息的定量处理方法	(远)
四、教育评价结果的整合方法	(远)
五、计算机在教育评价中的应用	(远)
第四章 建立发展性课堂教学评价体系	(远)
第一节 课堂教学评价的意义及类型	(远)
一、课堂教学评价的意义	(远)
二、课堂教学评价的类型	(远)
第二节 课堂教学评价观的转变	(远)
一、建立新型的教学观	(远)
二、树立现代的学生观	(远)
三、树立全面的教学质量观	(远)
四、树立以发展为本的课堂教学评价观	(远)
第三节 发展性课堂教学评价的内容	(远)
一、学生的学习表现	(远)
二、教师的教学行为	(远)
三、教师的基本素质	(远)
第四节 课堂教学评价的方法	(远)

摇摇摇摇一、直接观察法	(员源)
摇摇摇摇二、间接观察法	(员源)
摇摇摇摇三、参与观察法	(员源)
摇摇摇摇四、问卷调查法	(员源)
摇摇摇摇五、访谈法	(员源)
摇摇摇摇六、苏格拉底式评定	(员源)
摇摇附录摇高效、高质课堂评价的基本标准	(员源)
第五章摇建立促进教师发展的评价体系	(员源)
摇摇第一节摇新时期的教师评价观	(员源)
摇摇摇摇一、教师质量观	(员源)
摇摇摇摇二、教师评价观	(员源)
摇摇第二节摇教师评价的内容	(员源)
摇摇摇摇一、教师工作资格	(员源)
摇摇摇摇二、教师工作过程	(员源)
摇摇摇摇三、教师工作绩效	(员源)
摇摇第三节摇教师评价的步骤及方法	(员源)
摇摇摇摇一、教师评价的步骤	(员源)
摇摇摇摇二、教师评价的方法	(员源)
摇摇第四节摇教师评价的新进展	(员源)
摇摇摇摇一、发展性教师评价的由来	(员源)
摇摇摇摇二、发展性教师评价的特点	(员源)
摇摇摇摇三、发展性教师评价的实施	(员源)
摇摇附录摇教师评价实施方案	(员源)
第六章摇建立促进学生发展的评价体系	(员源)
摇摇第一节摇现代学生评价观	(员源)
摇摇摇摇一、对现代学生评价观的理解	(员源)
摇摇摇摇二、对现代学生评价功能的理解	(员源)
摇摇第二节摇学生评价的指标及标准	(员源)
摇摇摇摇一、一般性发展目标	(员源)
摇摇摇摇二、学科学习目标	(员源)
摇摇第三节摇学生评价的实施	(员源)
摇摇摇摇一、明确评价的指导思想	(员源)

摇摇摇摇二、现代学生评价实施的原则	(页码)
摇摇摇摇三、现代学生评价的具体实施	(页码)
摇摇第四节摇学生评价新方法	(页码)
摇摇摇摇一、表现性评价	(页码)
摇摇摇摇二、学生成长记录袋评价	(页码)
摇摇附录摇青岛中小学生不再“怵”考试	(页码)
第七章摇建立促进学校发展的评价体系	(页码)
摇摇第一节摇学校评价的特点及功能	(页码)
摇摇摇摇一、学校评价的特点	(页码)
摇摇摇摇二、学校评价的基本功能	(页码)
摇摇第二节摇学校评价的新理念	(页码)
摇摇摇摇一、强调学校特色	(页码)
摇摇摇摇二、鼓励学校自我评价	(页码)
摇摇摇摇三、重视学校发展过程	(页码)
摇摇摇摇四、注重学校评价的生成性	(页码)
摇摇摇摇五、倡导评价过程的开放性	(页码)
摇摇第三节摇学校评价的内容	(页码)
摇摇摇摇一、学校办学水平评价	(页码)
摇摇摇摇二、学校发展评价	(页码)
摇摇第四节摇学校评价的步骤和趋向	(页码)
摇摇摇摇一、学校评价的步骤	(页码)
摇摇摇摇二、学校评价的趋向	(页码)
摇摇附录摇上海市闵行区中、小学办学水平评估方案（修订稿）	(页码)
主要参考文献	(页码)

第一章 摇摇教育评价基础

本章学习提要

摇摇教育评价是教育科学的重要研究领域，也是现代教育管理的重要研究课题。对教育评价的概念表述不同，反映了对教育评价本质及目的的不同理解，有的着眼于效果、有的着眼于信息、有的着眼于方法等等，我们认为教育评价是按照一定的教育价值标准，对教育活动及其相关因素进行系统描述，并作出价值判断的活动。

教育评价有其产生发展的过程，自从有了学校活动便有了对教育评价的要求，吸引了许多学者及教育工作者研究如何开展教育评价，从教育测量运动到美国的“八年研究”，又到“第四代教育评价理论”的出现，教育评价不断走向成熟。

《国家基础教育课程改革纲要》提出了课程改革的目标，提出了建立发展性教育评价体系的要求，树立发展性的教育评价理念应面向未来，以人的发展为根本目的，坚持多元价值观、体现民主协商精神、关注目标的非预期效应和效果、注重评价标准的全面性和对个性内容的反映、强化过程性的评价等等。

第一节 摇摇教育评价概念

教育评价被认为是与教育基础理论、教育发展并存的教育研究的三大课题之一，在如火如荼的教育改革与实践中，教育评价更是因其具有的功能备受关注。但到目前对教育评价概念却还没有一个统一的界定，本节旨在通过对一些与教育评价相关概念的理解及一些对教育评价具有代表性的认识，来了解教育评价的本质。

一、评估与评价

评估与评价是否可以通用，或者说哪一种提法更合理，这是在
进行教育评价研究中首先涉及的问题。其实两词均译自英文 ~~evaluation~~
~~evaluation~~，即价值判断的意思，有人提出用于较易确定的价值判断时用
评价，用于较复杂的价值判断时则宜用评估。甚至提出按中文的理
解，似乎评估不够严谨精确，而评价则反之。

在实际使用中我们也发现两词混用的情况。一般来看，在理论
研究中使用评价一词居多，对已出版的一定数量的国内教育评价著
作统计(含译本)表明，使用“评价”一词的占到了 ~~80%~~；已发表
的论文，使用“评价”的也占了绝大多数。而在政府或教育行政
部门下发的官方文件中则使用“评估”一词为多，如《普通高等
学校教育评估暂行规定》、《普通中小学校督导评估工作指导纲
要》、《示范性普通高级中学评估验收标准(试行)》等等，~~1995~~
年 ~~9~~月正式颁布的《中华人民共和国教育法》中规定：“国家实行
教育督导制度和学校及其他教育机构教育评估制度”。虽然一些学
者一直在区分和辨别评估与评价，但在实际使用中，从对评估与评
价的意思表达来看并没有什么本质的区别，无论是理论界多用评
价、还是实践中常用评估应视为习惯使然。本书使用了评价一词，
一方面是由于习惯的原因，另一方面也在于作者较赞同刘本固先生
的解释：

第一，从国际交往看，“评价”这一概念已被各国所接受和使
用，是大家认可的。

第二，从我国的语言实践看，人们习惯用“评价”这一概念。
现在人们常说对某某历史人物或大事件的评价，而不说评估。

第三，从词义看，两个词的含义有所不同。“评估”是评论估
计的意思，并不含有价值的意义，而“评价”则具有评定价值程
度高低的意思。英文“~~evaluation~~”一词也不宜译作“评估”，应译
为“评价”。^①

① 刘本固：《教育评价的理论与实践》，浙江教育出版社，~~1994~~年版，第 ~~144~~ 页。

二、教育价值与教育评价

从上面区分评估与评价中，不难发现，评价与价值有着十分密切的关系。《辞海》认为：评价泛指衡量人物或事物的价值。按照马克思的观点，价值反映着主客体的关系，是客体对主体需要的满足。评价是关于价值的判断，我们对人物或事物作出主观的“好”与“坏”的判断，往往就被称为“评价”，这里的好与坏就是“价值的判断”。

提及价值判断自然又会提起与之相对的事实判断。

事实判断是对事物的现状、属性与规律的客观描述。如我们说，某个学生已经掌握了四则运算的规则，就是事实描述。事实判断的基本要求是它的客观性，即真实地反映事物的本来面目。

价值判断则是在事实描述的基础上根据评价者的需要和愿望对客观事物作出的评判。很显然这其中有很大的主观成分，因为这个评判与评价者的价值观相联系。而每个不同的评价者（具有主观能动性的人），却有着不同的经历，政治、经济地位存在差异，所获得的经验、形成的意识、情感等各不相同，各有自身的习惯、爱好、志趣等，也就是说评价者之间的价值观千差万别，他们可能会对同一件事物、同一个人作出完全不同的甚至截然相反的“价值判断”。相对于事实判断的客观性，价值判断的主观性、不一致性则是绝对的。

美国学者格兰兰德（Gronlund, 1905-1972）为评价进行了定义：评价是在量或质的记述的基础上进行价值判断的活动。

评价越测量（量的记述）或非测量（质的记述）均价值判断

顺着这样的思想，我们不妨给评价下个定义：

评价是依照一定的标准对客观事物进行观察并作出价值判断的过程。

那么，怎样认识教育评价呢？教育作为一种有目的、有意识的社会活动，也同样反映着主客体需要之间的关系。这里的客体是教育活动本身，而主体则是教育个体及社会，由于人们的价值取向不同，在实现个体价值与社会价值的过程中可能会偏重一方，即只重个体价值的实现或社会价值的实现，甚至将两者对立起来，我们强

调进行教育评价要考虑二者的有机结合，不仅要满足个体的发展需要还要满足社会发展的需要。换句话说，教育评价就是在进行教育活动时满足个体及社会需要程度的价值判断。

基于以上认识，我们如此界定教育评价：

教育评价是按照一定的教育价值标准，对教育活动及其相关因素进行系统描述，并作出价值判断的活动。

理解这一概念应明确：

第一，教育评价的本质是作出价值判断。

第二，一定的教育价值标准是教育评价的依据。

第三，事实的系统描述是教育评价的基本过程。

三、教育测量与教育评价

评价的本质是价值判断，而价值判断是以事实判断为基础的。测量则是依据一定的标准对某一事物的属性所作的事实判断，进而赋值的过程。可见，测量与评价存在十分密切的联系，也可以说他们之间的关系就是价值判断与事实判断的关系。

测量虽没有一个统一的概念，但其具有三个基本的要素和一个核心的本质。三个基本要素包括：标尺，即依据什么来测量；对象，即对什么进行测量；数量，即测量结果的表现形式。一个核心的本质就是其客观性。测量包括物质测量和精神测量，物质测量是一个纯客观的过程，不因测量者不同而存在差异。但教育测量本质上讲属于精神测量，是一种间接的度量，将心理测量的原理和方法应用于教育领域。如学习质量的测量，不可能像度量物质那样只要依一定的标准便能直接获得测量结果，它需要通过考试、观察、问卷等工具通过分析，推断结果，尽管这其中会有误差，但如果除去允许存在的误差，不同测量者测量的结果应是相同的。这也同样反映了教育测量具有的客观性的本质特征。

教育评价则是要在教育测量的基础上作出价值判断，对同一事物，同样的测量结果，由于价值观不同，评价结果会大相径庭。所以，我们认为教育测量与教育评价存在如下的联系：

第一，教育测量是教育评价的基础，教育测量所获得的结果是教育评价所需信息的主要的、可靠的来源，是对教育的状态和价值

进行客观判断的前提。

第二，教育测量的结果只有通过教育评价这个环节才能获得实际意义，否则，便成了一堆抽象而枯燥的数字。

四、对教育评价的不同理解

一直以来，对教育评价概念理解不一，表述更是种类繁多。较有影响的几种观点^①：

(员) 着眼于效果，强调通过评价判断教育目标或教育计划实现的程度。泰勒(裁德烈·厄宰扬)在“八年研究”《史密斯泰勒报告》中指出：“评价过程在本质上是确定课程和教学大纲在实际上实现教育目标的程度的过程。”

(圆) 着眼于信息，强调通过评价收集资料，为教育决策服务。克龙巴赫(悦燥遭群燥，蕴附扬)《通过评价改进课程》一文中指出：“所谓教育评价，是为获取教育活动的决策资料，对参与教育活动的各个部分的状态、机能、成果等情报进行收集、整理和提供的过程。”“一个搜集和报告对课程研制有指导意义的信息过程。”

(猿) 着眼于方法，强调评价是成绩考查的方法或调查的方法。斯塔弗尔比姆(杂塔泽尊比，阅随扬)认为：“评价不应局限于评判决策者所确定的教育目标所达到预期效果的程度，而应该是收集有关教育方案实施全过程及其成果的资料，为决策提供信息的过程。”协作组认为：评价是“对当时方案中发生的事件以及方案结局的系统考查——一种导致帮助改进这个方案或其他同样总目的方案的考查”。

(源) 着眼于过程，强调评价收集信息的过程、提供决策依据的过程、判断效果的过程、教育优化的过程以及价值判断的过程等。比贝(拜尊，悦随扬)把评价定义为“系统地收集信息和解释证据的过程，在此基础上作出价值判断，目的在于行动”；得雷斯(孛随尊燥)认为，“所谓评价，就是决定某种活动、目的及程序的价值过程。这个过程，分为目的明确化、收集有关合适的情报、

^① 马永霞主编：《教育评价》，当代世界出版社，圆园零年缘月第员版，第员怨~圆园页。

决策三个阶段。评价所追求的目的便是为达到目标最有效地去灵活使用手中的资源。”

（缘）着眼于价值，强调教育评价的关键在于价值判断。如桥本重治（日）认为：“评价是与教育的目标和价值有明确关系的概念，是按照教育目标和价值观对学生的学习成果及教育计划的效果等进行测量的过程。因此教育评价概念的重点在于以教育目标为标准价值判断。”

（远）1955年（美）教育评价标准联合委员会对教育评价界定：“教育评价是对教育目标和它的优缺点与价值判断的系统调查，为教育决策提供依据的过程。”

以上提法虽各有侧重，着眼点不同，但综合来看，教育评价具有如下的内涵：

第一，教育评价是一种活动过程，有明确的目标。

第二，教育评价的活动过程中，包含着一系列的步骤和方法，需要采取诸多的手段和技术来收集信息，分析、处理信息。

第三，教育评价是依一定的价值观，将评价对象实际状态与预定目标相比较。

第四，教育评价的目的是要通过价值判断，取得评定结果，为教育决策提供依据。

其实，以上对教育评价概念的不同理解，也反映了西方教育评价流派对教育评价本质的不同认识。分析其认识论基础是对教育现象的不同理解，是把教育现象看做是一个可控的、可量化的过程，还是看做一个处处充满了随机性、选择性的不可控过程。是对教育评价功能的不同理解，开展教育评价是要为评价对象鉴别优劣、分等还是为了使其通过评价这样一个手段得到改进，提高教育质量。换句话说，是强化教育评价的管理功能还是强化教育评价的服务功能的问题。

第二节 教育评价的历史沿革

对于教育评价发展的历史阶段有着不同的划分方法，有人将教育评价分为萌芽、产生、专业化、发展阶段；有人将教育评价分为传统考试、教育测验、教育评价三个时期；还有人将教育评价分为

萌动时期、探索时期、发展时期、兴盛时期。我们在此通过对教育评价源流、教育测量运动、现代教育评价的产生、现代教育评价的发展几个阶段的探究来揭示教育评价的发展历程。

一、教育评价的源流

评价活动是人类社会生活中必不可少的，自从有了人类社会，评价活动就存在了。教育评价活动，则始于学校教育出现，从学生成就测验开始。迄今，学术界对教育评价的源头仍存在不同的观点。较为典型的有两派：

一派以日本、中国学者为代表。

认为教育评价的最初萌芽源于中国古代的考试制度。

中国是考试的故乡，通过考试选拔贤才是历朝历代的基本选才用人之策。尤其是始于我国隋朝的科举制度更是影响深远。一般认为隋炀帝在公元 605 年增设进士科，开科取仕，即通过科考来选拔官吏是科举制度建立的标志。科举制度以才能取人，在中国历经 1300 年之久，其制度齐全、方法完备，对世界各国公职人员的选拔、录用产生了极大影响。日本的公职人员高级职务考试借鉴了中国科举考试的基本思想和方法；第一个建立“文官”（公务员）制度的英国采取竞争性的公开考试招募文职人员；西方国家建立的国家公务员制度的共同特点就是公开考试、择优录用。

另一派以西方学者为代表。

认为教育评价萌芽最初源于西方，以笔试代替口试为标志。

19 世纪初，伴随学校教育的发展，学生人数日益增多，美国学校普遍感到口试、面试的选拔考核学生方式难以实施，且有失标准，客观性较差。1844 年英国剑桥大学以笔试代替口试进行入学考试，首开西方考试的先河。1853 年在美国著名教育家贺拉斯·曼（Horace Mann）的倡导下，美国的波士顿文法学校首先引入笔试，这样，美国的初等学校、中等学校渐渐将笔试的考试方法普及开来。

其实，不论是中国选拔官吏的科举制度，还是以笔试替代口试、面试的考试形式，都不是现代意义的教育评价，只能说是教育评价的源流。基于此，我们还是认为教育评价的源流在中国更有说

服力。

中国的科举制度是西方国家公务员制度建立的源流，似乎没有太大的争议。卡特总统当政时的美国人事总署署长艾伦·坎贝尔教授说：“当我接受联合国的邀请来中国向诸位讲关于文官制度的时候，我是感到非常惊讶的。因为在我们西方所有的政治学教科书中，当谈到文官制度的时候，都把文官制度的创始者归于中国。”而有文字记载的中国采取考试的形式及笔试的方法时间甚为久远。

我国最早的也是世界最早的教育专著《学记》中说：“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。”意为，学生到了规定的年龄入学，国家每隔一年考查他们学业及操行成绩一次：第一年考查学生析句分段的能力和学习的志向；第三年考查学生是否专心学习和对周围的人是否和睦相处；第五年考查学生学识是否广博，同老师是否亲密无间；第七年考查学生研究学问的本领和识别朋友的能力。符合标准的就叫做小成。第九年，做到认识事物能触类旁通、闻一知十和政治上成熟，以及立场坚定不移。符合这一标准的就叫做大成。^①至于考试形式，我国在公元前119年汉武帝设立太学之后就已经以笔试为主了。比英美国率先用笔试替代口试的时间要早了许多。^②

二、教育测量运动

考试是学校及教师对学生的学业成就检定的基本方法，书面考试对西方的测验方法有了很大的推进作用，但并未能很好地解决考试的客观标准问题。1899年，英国格林威治医学院教师乔治·费舍（George Fiske）为了对自己学校的学生试题答案和作品进行客观评分，收集了大量的学生成绩样本，区分优劣，汇编成了成绩量表，称做《量表集》，它表示依据一定的价值程度进行评分的标准的尺度或量表。虽然费舍的这一尝试性研究并未在当时得到教育界

① 高时良：《学记记注》，人民教育出版社，1986年版，第1页。

② 王孝玲编著：《教育评价的理论与技术》，上海教育出版社，1993年1月第1版，第1页。

的注意，但被 18 年后的桑代克（~~转援转援转世转转~~）等开拓教育测量运动时作为依据，给予了极大的重视。费舍被称为致力考试客观化的先驱者。

美国的莱斯（~~转转转转转~~）进行的拼字测验，在当时引起了人们对考试问题的极大关注。他进行测验的目的是要辨别学生学力的个人差异，了解学校教育的成果。1879 年，莱斯发表了他对 10 所学校的 1500 名学生进行拼字测验的结果。测验结果表明：愿年中每天用 15 分钟进行拼字练习的学生与每天用 5 分钟进行拼字练习的学生的测验成绩并没有多大区别。这一结论立时引起了轰动，而且将人们的目光引向了对测验方法的关注。后人称莱斯为客观测量的先驱、教育测量的创始人。

对学校教育和评价方法产生深刻影响的是心理学的基本原理及方法。1889 年，冯特（~~转转转转转~~）在莱比锡大学首创心理实验室。在实验心理学的研究中，心理学家们摸索出了一套测量的方法，对教育测量的发展有很大的影响。

英国差异心理学的创始人高尔顿（~~转转转转转~~）1869 年创建了人类测量实验室，在 1890 年又在伦敦大学创办优生学实验室，最早应用统计处理心理学研究资料，以数量代表所测得的心理特征差异。美国心理学家卡特尔（~~转转转转转~~）在莱比锡大学冯特心理学实验室从事个体差异研究，他重在测量基本的身体活动或感觉——运动反映，使心理测验走出了实验室，为应用心理学的发展开辟出一条新的道路。

如果说以上心理学家所作的是心理及教育测量尝试的话，那么 19 世纪末到 20 世纪 20 年代教育测量进入了一个新的阶段，这一阶段也被称为测量运动阶段。美国的心理学家桑代克堪称教育测量运动的开拓者。

桑代克首创用实验法研究动物心理，认为动物学习的基本方式是通过尝试错误而偶然获得成功，从而在刺激与反应之间形成连结。他又将“连结”这一概念推广到人类心理，进而研究人类学习。1903 年，桑代克发表了《心理与社会测量导论》，被高度评价为测量学上划时代性的著述，并被称为教育测量形成的标志。书中介绍了统计方法及编造测验的基本原理，特别是提出了“凡存在