

◎ 王斌华 著

*Xuesheng Pingjia*

# 学生评价：

>>> >>> >>>

## 夯实双基与培养能力

*Hangshi Shuangji Yu  
Peiyang Nengli*

上海教育出版社

*Xuesheng Pingjia*

# 学生评价：

>>> >>> >>>

## 夯实双基与培养能力

*Hangshi Shuangji Yu  
Peiyang Nengli*

◎ 王斌华 著

上海教育出版社  
SHANGHAI EDUCATIONAL PUBLISHING HOUSE

图书在版编目(CIP)数据

学生评价: 夯实双基与培养能力 / 王斌华著. —上海:  
上海教育出版社. 2010. 1

ISBN 978-7-5444-2696-1

I. ①学... II. ①王... III. ①学生—教育评估—师资  
培训—教材 IV. ①G449.7

中国版本图书馆CIP数据核字(2009)第236570号

学生评价: 夯实双基与培养能力

王斌华 著

上海世纪出版股份有限公司出版发行  
上海教育出版社

易文网: [www.ewen.cc](http://www.ewen.cc)

(上海永福路123号 邮政编码: 200031)

各地新华书店经销 江苏常熟新骅印刷厂印刷

开本 700×1000 1/16 印张 19.5 插页 2

2010年1月第1版 2010年1月第1次印刷

印数 1—4,000本

ISBN 978-7-5444-2696-1/G·2099 定价: 40.00元

(如发生质量问题, 读者可向工厂调换)

# 前 言

这是一本专门探讨学生评价的著作,旨在帮助教育理论工作者、在职教师、即将成为教师者学习和掌握学生评价的基础知识与基本技能,了解和把握学生评价的发展态势、最新理念和实施方法。

我国是考试的故乡。早在我国奴隶社会的西周时期就建立了“考校”与“考选”制度。在随后的几千年间,我国的考试制度遥遥领先于世界各国。然而,在现代意义上的“教育测验”和“学生评价”方面,我国却远远落后于西方发达国家。直到改革开放后的 20 世纪 80 年代,学生评价的概念、理论和方法才开始进入我国。最近几年,一系列国外有关学生评价的著作不断进入国内,但是总体而言,仅仅停留在译介的层面。目前,我国从事学生评价研究的学者专家为数不多,至今还没有创建具有中国特色的学生评价理论体系。

随着教育理论的深化和教育实践的拓宽,学生评价已经成为当今世界各国共同关注的热点、重点和难点。学生评价是学校教育的中心工作之一,是为教和学服务的,其根本目的是促进教师的教和学生的学。与学校评价、课程评价和教师评价一样,学生评价是教育评价领域不可或缺的重要组成部分,直接关系到学校的办学质量。学生评价也是一项系统工程,将极大地影响教学观、学习观、教师观和学生观。学生评价更是一门科学,具有专门且独特的理论与方法。为了确保学生评价的客观和公正,教师必须接受专门培训,否则难免在学生评价过程中走弯路、碰钉子。

古人云:“成也萧何,败也萧何。”学生评价也是如此,用之得当受益匪浅,用之不当则贻害无穷。比如,应试教育背景下的学生评价必然囿于少数应试科目和应试技巧,使学生沦为“应试机器”,导致“高分低

能”，而素质教育背景下的学生评价，旨在夯实学生的“双基”（基础知识和基本技能）和培养学生的能力（创新能力和实践能力）。

“师者，所以传道授业解惑也。”（《韩愈·师说》）笔者在撰稿过程中，试图努力体现下列三个特点。

第一，阐述学生评价的基本原理和实施办法。如：学生评价的发展脉络；学生评价的理论基础、重要概念、主要作用和基本原则；学生评价目标的基本范畴、主要类型、重要特征和制定方法；学生评价的实施办法、操作要求和注意事项；学生评价管理的阶段、任务和要求，等等。

第二，探讨学生评价的发展趋势和最新动态。如：替代性评价革命、真实性评价、表现性评价、基点评价、增长评价、学生参与式评价、分层考试，等等。

第三，厘清学生评价的困惑和误解。如：考试、教育测验与学生评价有什么区别？学生评价目标、教材编制、教学实践和学生评价之间是什么关系？考试成为“指挥棒”的根源在哪里？学生评价是紧扣教材还是紧扣学生评价目标？为什么要反对偏题和怪题，而不反对难题？识别答案、再现答案与建构答案有什么区别？真实性评价与表现性评价有什么异同？如何改革评价结果的反馈？等等。

但是，坦率地说，笔者在撰写书稿的过程中，头脑中不时闪现一些新的困惑，譬如：

- 日本也是“考试强国”和“应试大国”，其高考竞争的激烈程度并不亚于我国，但为什么能培养出那么多诺贝尔奖获得者？
- 目前我国的学生评价仍然主要采用两千多年前我国发明的考试方法和 20 世纪 30 年代美国发明的客观试题，这显然已经无法满足当代学生评价的要求，但为什么至今无法突破？
- 国外在学生评价过程中广泛采用真实性评价、表现性评价、学习档案袋评价、项目研究法等，但为什么这些方法在我国困难重重、寸步难行？
- 我国古代发明的考试制度曾引领世界几千年，现在中国和平崛起，能否在“后学生评价时代”有所作为，再创辉煌呢？

笔者为上述困惑感到苦恼。也许,这些困惑将成为笔者日后的关注重点,或者成为笔者未来的研究方向。

华东师范大学课程与教学系

教授、博士生导师

王斌华

2009年8月31日

# 目 录 Contents

## 第一章 学生评价历程/1

### 第一节 考试制度/1

- 一、战国时期的《学记》/2
- 二、汉代太学/4
- 三、汉代的察举选士/5
- 四、科举制度/7
- 五、科举制度的消亡/15
- 六、考试制度传入西方/18

### 第二节 教育测验/19

- 一、心理测量的发展/19
- 二、教育测验的发展/27
- 三、教育测验进入中国/33

### 第三节 学生评价/34

- 一、“八年研究”/34
- 二、替代性评价革命/38
- 三、学生评价进入中国/41

---

## 第二章 学生评价基础/46

### 第一节 学生评价的理论基础/46

- 一、泰勒原理/46
- 二、行为主义学习理论/51
- 三、建构主义学习理论/57
- 四、多元智能理论/63

## 第二节 学生评价的重要概念/65

- 一、科学-实证主义学生评价观与人文-自然主义学生评价观/65
- 二、教育评价、学校评价、课程评价、教师评价与学生评价/67
- 三、安置性评价与诊断性评价/69
- 四、终结性评价与形成性评价/70
- 五、基点评价与增长评价/71
- 六、真实性评价/72
- 七、表现性评价/74
- 八、绝对评价与相对评价/78
- 九、标准参照考试与常模参照考试/78
- 十、口头考试与书面考试/79
- 十一、信度、效度、难度与区分度/80
- 十二、偏题、怪题与难题/84
- 十三、单元评价主体与多元评价主体/86

## 第三节 学生评价的主要作用/89

- 一、导向作用/89
- 二、反馈作用/89
- 三、诊断作用/90
- 四、促进作用/90
- 五、激励作用/91

## 第四节 学生评价的一般原则/92

- 一、科学原则/92
- 二、发展原则/92
- 三、全面原则/93
- 四、全员原则/93
- 五、全程原则/93
- 六、区分原则/94
- 七、多样原则/94

## 八、民主原则/94

---

### 第三章 学生评价目标/96

#### 第一节 学生评价目标的范畴/97

- 一、布卢姆的教育目标分类法/97
- 二、加涅的教育目标分类法/98
- 三、教育目标分类法的影响/99

#### 第二节 学生评价目标的类型/103

- 一、不同类型的学生评价目标/103
- 二、关于行为目标与表现性目标的争论/107
- 三、行为目标与表现性目标之间的失衡/111
- 四、行为目标与表现性目标的评价/114

#### 第三节 学生评价目标的特征/119

- 一、纲领性/119
- 二、整体性/122
- 三、系统性/123
- 四、严密性/127
- 五、现实性/128
- 六、超前性/128

#### 第四节 学生评价目标的制定/129

- 一、特尔斐法/129
- 二、参照法/130
- 三、筛选法/130
- 四、聚类法/131
- 五、增添法/131

---

### 第四章 学生评价现状/133

#### 第一节 七种不正常的现象/133

- 一、注重应试科目的评价,忽视其他科目的评价/133
- 二、注重书本知识的评价,忽视其他目标的评价/134
- 三、注重书面形式的评价,忽视其他形式的评价/134
- 四、注重容易量化的评价目标,忽视不易量化的评价目标/134
- 五、注重死记硬背,忽视能力培养/135
- 六、注重终结性评价,忽视形成性评价/137
- 七、注重分数与名次,忽视起点与进步/138
- 第二节 一种“优势”/140
- 第三节 一根“指挥棒”/143
  - 一、高考“指挥棒”的舞动/143
  - 二、高考改革任重道远/146
- 第四节 三种严重后果/157
  - 一、缺乏创新能力/157
  - 二、缺乏实践能力/161
  - 三、埋没杰出人才/166

---

## 第五章 学生评价重心/169

- 第一节 学生评价重心的转移/169
  - 一、“双基”教育的学生评价重心/169
  - 二、应试教育的学生评价重心/170
  - 三、素质教育的学生评价重心/171
- 第二节 基础知识的评价/177
  - 一、基础知识的界定/177
  - 二、基础知识的意义/178
  - 三、基础知识的形成/179
  - 四、基础知识的评价/180
- 第三节 基本技能的评价/182

- 一、基本技能的界定/182
- 二、基本技能的意义/183
- 三、基本技能的形成/183
- 四、基本技能的评价/184
- 第四节 创新能力的评价/187
  - 一、创新能力的界定/187
  - 二、创造能力的评价/187
  - 三、发散性思维的评价/191
  - 四、批判性思维的评价/195
- 第五节 实践能力的评价/198
  - 一、实践能力的界定/198
  - 二、实践能力的意义/198
  - 三、实践能力的形成/199
  - 四、实践能力的评价/202

---

## 第六章 学生评价方法/204

- 第一节 分层评价/204
  - 一、分层评价的界定/204
  - 二、分层评价的形成/205
  - 三、分层评价的意义/209
  - 四、分层评价的实施/211
  - 五、分层评价的注意事项/214
  - 六、分层评价实例/214
- 第二节 分层考试/216
  - 一、分层考试的界定/216
  - 二、分层考试的原则/217
  - 三、分层考试的试卷模式/218
  - 四、分层考试的计分/219

## 五、分层考试实例/220

### 第三节 客观性考试/224

- 一、客观性考试与主观性考试/224
- 二、客观性考试的优势与弊端/226
- 三、提高客观性考试的客观性/228

### 第四节 学生参与式评价/230

- 一、学生参与式评价的界定/230
- 二、学生参与式评价的意义/230
- 三、学生参与式评价的形式/230

### 第五节 学习档案袋评价法/233

- 一、学习档案袋评价法的界定/233
- 二、学习档案袋评价法的意义/233
- 三、学习档案袋评价法的发展/234
- 四、学习档案袋评价法的特征/236
- 五、学习档案袋的设计/237
- 六、学习档案袋的类型/238
- 七、学习档案袋的评价方法/239
- 八、学习档案袋评价法的注意事项/242

### 第六节 项目研究法的评价/242

- 一、项目研究法的界定/242
- 二、项目研究法的意义/243
- 三、项目研究法的特征/244
- 四、项目研究法的设计/245
- 五、项目研究法中的教师角色/247
- 六、项目研究法例题/248
- 七、项目研究的评价方法/250
- 八、项目研究法实例/251

### 第七节 掌握学习法的评价/253

- 一、掌握学习法的界定/253

- 二、掌握学习法的意义/253
  - 三、掌握学习法的特征/253
  - 四、掌握学习法的实施/254
  - 五、掌握学习法实例/256
  - 第八节 情感倾向的评价/257
    - 一、情感倾向评价的范围/257
    - 二、情感倾向评价的意义/257
    - 三、情感倾向的评价方法/258
    - 四、情感倾向评价的注意事项/264
- 

## 第七章 学生评价管理/265

- 第一节 准备阶段/266
    - 一、制定评价方案/266
    - 二、填写双向细目表/266
    - 三、编制试题/268
    - 四、编排试卷/269
  - 第二节 实施阶段/270
    - 一、防治舞弊/270
    - 二、试卷批阅/277
    - 三、评价计分/279
  - 第三节 反馈阶段/285
    - 一、试卷分析/286
    - 二、试卷讲评/286
    - 三、成绩反馈/288
- 

## 主要参考文献/296

# 第一章

## 学生评价历程

中国发明了考试,曾遥遥领先于世界几千年。然而,现代意义上的“教育测验”和“学生评价”却先后发轫于英国和美国,我国始终处于引进和借鉴他人研究成果的阶段。

——笔者

自从人类开办了学校,就有了学生评价,其历史可以追溯到奴隶社会。早在我国奴隶社会的西周时期(公元前 11 世纪—公元前 771),就建立了“考校”与“考选”制度,战国时期的教育名篇《学记》详细阐述了当时的考试制度,隋唐以后逐步形成的科举制度使我国的考试制度达到了相当高的水准。之后,考试制度传入西方国家。

1900 年前后,西方国家兴起了一场“教育测验”(educational testing)运动,主张采用测验的方式来测定学生知识记忆的能力和某些特质。该运动直接导致标准化测验的广泛使用。1920 年前后,教育测验被引入我国。

1930 年前后,西方国家提出了“学生评价”(student assessment)的概念。在随后的几十年间,学生评价理论不断深化,学生评价实践不断拓宽,先后出现了一系列新的评价理念和新的评价方法。20 世纪 80 年代前后,学生评价被介绍到我国。

### 第一节 考试制度

考试的故乡在中国,中国是考试制度的发源地。孙中山曾经说过:

“英国行考试制度最早,美国行考试才不过二三十年,英国的考试制度就是学我们中国的。”<sup>①</sup>中国是最早采用考试选拔人才的国家,在中国奴隶社会和封建社会的几千年间,考试制度达到了相当高的水准。

作为一大发明,考试的历史地位以及对世界作出的重要贡献绝不亚于中国古代的“四大发明”(造纸术、印刷术、指南针和火药)。笔者认为,考试理应当之无愧地被列为中国古代的“第五大发明”。

### 一、战国时期的《学记》

《学记》是我国战国时期《礼记》49篇文章中的一篇。它是先秦时期儒家教育经验和教育思想的总结,其作者和完成年代至今尚无定论。《学记》不仅是我国教育史上最早论述教育、教学的教育文献,而且是世界教育史上最早出现的体系完整的教育论文。《学记》全文1229字,详细阐述了教育目的、教育年限、考试频率、考试目标等内容,其文字精练,构思严密,见解深刻。

《学记》详细记载了我国西周时期已经建立的考试制度:“古之教者,家有塾,党有庠,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。夫然后足以化民易俗,近者说服而远者怀之,此大学之道也。”

这段文字的大意是:古代的教学,闾<sup>②</sup>中有塾,党<sup>③</sup>中有学堂,术<sup>④</sup>中有学校,国都有大学。大学每年招收新生入学,学生每隔一年都要接受学业成就的考查。入学第一年考查学生辨析圣人经典的能力和志向目标,第三年考查学生对学业的敬业态度和团结友爱的群体精神,第五年考查学生学识的广博程度和亲近师长的态度,第七年考查学

① 孙中山研究学会孟庆鹏编:《五权宪法之讲解——1921年4月4日在广东省教育学会的演说》,载《孙中山文集》(下),团结出版社1997年版,第548页。

② 闾:古时二十五家为一闾。

③ 党:古代的地方户籍编制单位,五百家为一党。

④ 术:同“遂”,古时一万二千五百家为遂。

生研读学问的能力和择友的本领,合格者可以称为“小成”。到第九年,学生在学术上触类旁通,行为、意志和见解能够坚定不移,不受外力左右,合格者可以称为“大成”(孔府大殿为“大成殿”——作者注)。达到了这种境界,就足以教化大众百姓,改变不良习俗,使近处的人心悦诚服,远方的人向往归附,这就是大学的目标。

上述文字表明,我国早在西周时期就采用了考试这种方式,并且形成了一套完整且规范的考试制度。

#### 1. 明确规定修业目标

学生的修业目标分为“小成”和“大成”两个阶段目标,“知类通达,强力而不反”被作为学生完成修业的最高标准。

#### 2. 明确规定修业年限

学生修业年限为9年,前后分为7年和2年两个阶段。第一阶段“合格者”方可进入第二阶段学习,直至全部考试合格才准予毕业。

#### 3. 严格规定考试频率

在修业的9年期间,每隔一年只考查学生一次,而不是每年考试、每月考试、每周考试或随意考试。控制考试次数或降低考试频率可以减少考试泛滥对学生学习的干扰,可以让学生充分合理地安排学习时间,培养和发展自己的兴趣、爱好、特长和志向,可以督促教师随时提高教学质量,而不是等到考试以后才发现教学中存在的问题。更重要的是,可以提高考试的权威性和严肃性,确保社会、学校、教师、学生和家长对考试的敬畏和重视。这个道理非常浅显,就如同人们会问:奥运会为什么如此神圣,如此受人尊崇?奥运会奖牌的含金量为什么如此之高?答案是:因为奥运会每4年才举办一次。相对运动员的运动生命而言,4年的时间是那么漫长、那么遥远。况且,运动员的最佳竞技状态和巅峰时期很难延续4年以上。

#### 4. 严格规定考试内容

考试内容既要考查学生“离经辨志”、“知类通达”等学业成就,还要考查学生“敬业乐群”、“博习亲师”、“论学取友”的道德操守。前者强调学生掌握学业知识的能力,后者侧重学生人格养成、社会实践的能力。

众所周知,教育的根本任务是培养一个完整的人,而不是单纯注重一个人的知识掌握。因此,考试不仅应该考查学生的知识,还应该考查学生的品行、能力等诸多方面。

## 二、汉代太学

太学是汉代设在京师的全国最高教育机构。西汉时期,汉武帝力图改变汉初“师异道,人异论,百家殊方,指意不同”的混乱局面,在思想文化界首开“罢黜百家,独尊儒术”的政策,确立了儒家思想的正统与主导地位,使得“大一统”的专制思想成为一种主流意识形态。由于汉初人才选拔和人才使用不当,官吏主要来自官僚子弟或捐资买官,汉武帝得不到贤才辅佐。尽管他日夜操劳,“尽思极神”,但在治理国家方面并没有取得预期的效果。为了解决人才问题,汉武帝采纳了董仲舒设立太学的主张,开始在长安建立太学,开设《易经》、《诗经》、《尚书》、《礼记》、《公羊传》、《穀梁传》、《左传》、《周官》、《尔雅》等课程。太学的教师被称为博士,均为精通儒经的权威学派的名儒,其主要职责是掌经教授。此外,汉代太学博士担负提供咨询、参与议政或巡访地方的重任,这是旧时博士官制的延续。太学的学生称“博士弟子”,或简称“弟子”,东汉时称“诸生”或“太学生”。

从汉武帝到王莽,太学的人数逐渐增多。汉武帝元朔五年(公元前124),太学中只设五经博士,置博士弟子50名,汉昭帝时增至100人,汉宣帝时有200人,汉元帝时达1000人,汉成帝时增至3000人。王莽秉政后,为树立声望,并笼络广大儒生,在长安城南兴建辟雍、明堂,又“为学者筑舍万区”,博士弟子达万余人,太学规模之大,前所未有。

汉代太学不仅在促进文化发展方面发挥了重要作用,而且接受了董仲舒《对贤良策》的建议,尊崇“学而优则仕”,把考试作为培养人才和选拔人才的主要手段,并建立了相当系统、严格的考试制度。

当时的考试制度有两个功能:一是甄别和选拔人才,把学生的考试成绩和等级与其仕途的官级、俸禄挂钩;二是督促学生学习,通过考查学生的学业,督促和激励学生勤奋学习,专心攻读,刻苦钻研,兼通几