

# 从教育测验到教育评论相关概念及其关系的辨析

刘尧

(浙江师范大学教育评论研究所, 浙江 金华 321004)

**摘要:**人类对教育进行事实判断和价值判断,走过了从教育测验到教育评论的历程,形成了教育测验、教育统计、教育测量、教育评估、教育评价、教育评论理论,并且这一学科群落已经形成。但由于这些理论和学科群落并不成熟,许多人对该学科群落的相关概念及其关系认识不清。因此,很有必要从多角度对该学科群落的相关概念及其关系进行辨析,并对该学科群落建设进行有益的探索。

**关键词:**教育测验;教育测量;教育评价;教育评论

**中图分类号:**G40-058.1   **文献标志码:**A   **文章编号:**1672-8742(2007)03-0025-06

近10年来,由于科研和教学的需要,我阅读和研究了与教育评价学相关的教育测验、教育统计、教育测量、教育评估和教育评价等方面的著述,同时,开展了教育评论学研究工作<sup>[1]</sup>。教育测验从隋炀帝(公元606年)设立科举制度开始,在我国延续了1300年。教育统计、教育测量是20世纪初在西方国家兴起随后传入我国的,教育评估、教育评价是20世纪30年代在美国兴起,70年代末传入我国。这些学科的理论研究在我国进展不大,学科成熟度不高,与我国广泛开展的教育测评活动很不适应。我国各省市都有考试院(中心),也成立了许多评价(评估)机构,广泛展开了对教育的评价活动,各种测验、考试充斥了学校及社会各方面。而与此相应的教育测验、教育统计、教育测量、教育评估、教育评价理论很不成熟。尤其值得重视的是混淆了教育评价与教育评估,并把手伸向教育测验、教育测量甚至教育评论,不同程度影响了相关学科的研究和发展。同时,也难以有针对性的指导教育评价实践。鉴于此,本文对从教育测验到教育评论相关概念和学科群落进行简述,并从医学和价值论角度辨析它们的关系。

基金项目:本文为浙江省新世纪“151”人才基金项目与高等学校中青年学科带头人基金项目的研究成果之一。

作者简介:刘尧(1960-),男,陕西西安人,教授、硕士生导师,浙江师范大学教育评论研究所所长,研究方向为高等教育学、教育评论学和教育评价学。

## 一、从教育测验到教育评论相关概念简述

教育测验:即为考试(examination),它是对人的知识、技能和能力以及其他某些心理特征的测量。考试是教育评价的一种手段。考试是个系统,它由主考、被考(考生)、试卷(试题的集合及其评分标准)、考分四要素组成。考试的过程就是主考组织编制试题,通过试题作用于被考,将正确答案、评分标准与被考的反应(答案)相比较,然后给出考分的过程。考试从宏观上可分为社会考试和教育考试两大类。教育考试又可分为校内考试和校外考试两类。按考试分数的解释与使用,可分为常模参照性考试和标准参照性考试。按考试对象分,有学业成就考试、学术能力倾向测验、人格测验。按测验人数分,有个别测验、团体测验。按测验方式分,有纸笔测验和非纸笔测验。其它可分为难度测验、速度测验、总结性测验、诊断性测验以及大规模测验、教师自编测验等。<sup>[2]</sup>

教育统计(学)(educational statistics):运用数理统计学的原理和方法研究教育现象中的数量关系的一门应用学科。它通过数据的分析和处理,准确地掌握教育情况,为制定教育方针、政策提供科学依据,为开展教育科学研究,探索教育规律,制定教育方案,检查教育效率提供科学方法,也为教育管理和教育评价提供有效的工具。它分为两大类:其一是描述统计学,研究简缩数据和描述这些数据。其二是推断统计学,用数据进行统计检验,统计分析和非参数统计法作出决策。前者只是分析某一样本的特点,而后者则是通过抽取母体的一个样本,并以此样本为根据来分析母体的特点。它是在教育问题的研究中,描述样本,分析样本,分析样本与总体的关系,从而进行教育科学推论的一种原理和方法。其研究方法一般有:大量观察法、统计分组法、综合指标算法、统计资料检验法、统计资料图示法等。<sup>[3]</sup>

教育测量(学)(educational measurement):研究如何测量学生的学业成绩、学习能力、兴趣爱好、思想品德以及教育效益等问题的一门教育学科。它运用各种测量手段和数理统计方法等数学工具,根据教育目标或教学计划的要求,研究如何测量和评价教育效果,如何编制测量计划,各种类型试题的优缺点,衡量一个学业成绩测验好坏的数量指标,如何进行测验结果的统计分析,测验分数的解释和评价等。其主要内容包括:测量原理、项目分析、效度、信度、标准化测验、测验的设计和编制、测验的常模、测验的实施和评分、测验结果的解释等。<sup>[4]</sup>

教育评估(educational appraisal):是一种有系统地去寻找并搜集资料,对评估对象作预测性、估计性的评判,以便协助教育决策者从若干种可行的策略中择一而行的过程。在我国教育评估与教育评价不作严格区别,在实践中具体运用时,不同的范围和场合有不同的习惯用法。如高等教育中多用评估,在督导部门也称督导评估,而在普通教育领域多用教育评价。但也有学者认为,两个概念有一定区别。评估有估价的含义,较为模糊粗略,而评价相对而言是较为精确的价值判断;评估多用于群体的状态及效果的估价,而评价则既有对群体的价值判断,也有对个人的判断。教育评估的

一个重要目的就是为教育决策提供重要依据,国家教育部门要根据教育评估的结果,及时调整教育的决策。教育评估的结果对国家、对学校都是一份咨询材料,而不是行动纲领。<sup>[5]</sup>

教育评价(学)(educational evaluation):以教育为对象,研究各类教育目标与相应的教育效果之间的关系,并给予一定的价值判断的学科。它的研究范围基本上可以划分为:宏观教育评价(研究教育思想评价和教育制度评价)和微观教育评价(研究学校教育系统内部各类教育思想和教育活动的评价,如学校办学思想、学校领导管理质量、教师教育和教学质量、学生德智体全面发展质量的评价等)。它的基本范畴是:其一教育评体:教育评价的对象。其二教育评域:教育评价的范围。其三教育评模:教育评价的模式。其四教育质评:教育质的规定性的评价。其五教育量评:教育量的规定性的评价。其六教育度评:教育质和量变化的度的评价。其七教育评标:教育评价的指标体系。其八教育评类:教育评价的类型。<sup>[6]</sup>

教育评论(学)(educational review):教育评论学是研究教育评论现象和教育评论活动规律的科学。教育评论学是与教育理论、教育史并列的一门教育学科。教育评论学的学科体系是建立在范畴体系和概念体系之上,并给予科学的揭示和阐述,使这些概念和范畴形成一个有机整体。同时,还要运用这些概念、范畴及其固有的内在联系,采取科学的方法论证教育评论活动中最普通、最一般的科学原理,从而揭示出整个教育评论活动及其形成和发展的基本规律。教育评论学的学科性质:首先,教育评论活动是一种教育科研活动,它的直接结果是鉴别、评判教育文化的优劣,进而对教育界产生社会舆论导向作用;而教育文化本身及其所产生的附着于受教育者身心的精神产品,都属于意识形态范畴。可见,教育评论学属于研究意识形态的科学。其次,教育评论学的主要任务在于探索和揭示教育评论活动的本质规律,建立完整的教育评论学学科体系。所以,教育评论学也属于基础理论学科。再次,教育评论学的研究对象决定了教育评论学涉及整个教育领域,内涵十分丰富。也可以从不同视角,针对不同范畴建立多部门、多层次的分支学科(如:教育理论评论学、教育史评论学、教育行政评论学、教育家评论学等),所以教育评论学又是一门综合性的学科。<sup>[7]</sup>

## 二、从教育测验到教育评论相关概念的关系

上面我们对教育测验(即考试)、教育统计、教育测量、教育评估、教育评价、教育评论的概念作了简明的介绍,下面我们从相互作用的角度看它们之间的关系。

### (一) 教育评价与其他概念的关系

教育评价与教育测验的关系:教育测验是对人的知识、技能和能力以及其他某些心理特征的测量,是教育评价的一种手段,确切地说它是教育评价获得评价信息的工具之一。教育评价获取信息的测验法就是来自于教育测验。

教育评价与教育统计的关系:教育统计是运用数理统计的原理和方法研究教育

现象中的数量关系,通过数据的分析和处理确切地掌握教育情况,也是教育评价的一种手段和工具。教育评价中获取信息的统计法就来自于教育统计。

教育评价与教育测量的关系:教育测量通过对学生知识的掌握、智能的发展、思想品德的变化、体质的状况以及教育活动的各个方面的测定,为教育评价提供价值判断的依据。另外,教育测量的结果只有通过教育评价才能获得实际意义,成为改进教育工作有参考价值的信息。<sup>[8]</sup>

教育评价与教育评估的关系:教育评价与教育评估是非常相近的两个概念,在许多场合是通用的。评价是精确的价值判断,而评估是估计的、预测性的价值判断。在我国从政府的教育政策文件到学者的文章论著,以及教育实践都把评估与评价混用,高等教育多用评估,其他多用评价。

教育评价与教育评论的关系:教育评论是对教育文化的反思和展望,它的对象是以教育文化中的精神文化和制度文化为主的,以真善美作为评论标准的观念活动。虽然它与教育评价只一字之差,但它与教育评价分属于教育科学的不同学科,有着显著的区别。比如,教育评价是一种管理活动,而教育评论则为观念活动;教育评价由评价机构按严密的程序实施,而教育评论则由教育评论家个人采用灵活多样的方式方法实施;教育评价有明确的指标体系并要求量化,而教育评论无需指标体系更不用量化;教育评价的对象是具体的可控的,而教育评论的对象是抽象的,不可控的;教育评价的结果多用于改进工作,而教育评论的结果则多用于改变观念;等等。这不是说教育评价与教育评论一点关系都没有,它们的关系表现在教育评价的结果可以成为教育评论的论据影响教育评论,而教育评论的结果可以作为教育思想观念影响教育评价所依据的教育价值观。<sup>[9]</sup>

(二) 从教育测验到教育评论相关概念的区别

1. 从医学的视角看教育测验到教育评论相关概念的区别

如果我们把教育比作一个“病人”的话,教育测验就是用来测量“病人”病症的医疗器械(工具);教育测量就是医生用医疗器械为“病人”进行检查;教育统计就是医生向“病人”问诊;教育评估就如同医生对“病人”会诊(诊断);教育评价就如同医生对“病人”实施治疗;教育评论就如同保健医生对“病人”进行保健指导或对健康的人实施保健教育一般。可用图 1 表示:

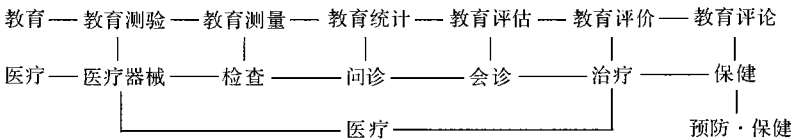


图 1 从医学视角看教育测验到教育评论相关概念区别

从上图类比可以看出,这些概念是有根本区别的。尤其强调的是从医疗器械到治疗都属于医疗的范畴,而保健则属于预防保健范畴,这两个范畴在医学上是有显著区别的,所以类比结果可以间接表明,教育评论与教育评价等在教育科学上是属于不同

学科的,也是有显著区别和不可混为一谈的。

## 2. 从价值论角度看教育测验到教育评论相关概念的区别

唯物主义的价值论认为,价值判断和事实判断是人们把握客观世界的两种不同方式。两者反映的对象、意义和参加的心理成份都有着质的区别。事实判断是以客体的本质和规律为对象的,它探讨客体“是怎样的”、“是什么”,探讨事物的现象、本质和规律等实体属性,即以理性认识这种抽象思维形式反映客体的本质及其规律。而价值判断是以客体与主体需要的关系为对象,它探讨客体的价值属性,即客体的社会意义——对社会需要意味着什么,它以多种意识形式——认识(价值与事实判断)、情感(对价值的态度体验)、意志(对价值的自觉保证)等诸种形式的综合来反映客体与主体需要的关系。<sup>[10]</sup>

从上面的分析可以看出,教育评价对评价的对象不仅要借助教育测验、教育统计、教育测量进行量的测定和分析,而且要进行质的分析,把所有考查绩效的材料和分析综合起来,在多大程度上达到了预期的目的作出价值判断。从这里可以看出,教育评价行为是建立在事实判断基础上的价值判断。教育评价中的事实判断在于诊断与鉴别教育的性状,价值判断则是对诊断鉴别而得的结果或结论,按其价值取向进行合乎目的的推断和判断;事实判断是价值判断的基础,价值判断是事实判断的目的性追求。而教育评论是依据一定的教育价值观和教育价值取向,对教育的一种价值判断。从教育测验到教育评论是一个由事实判断向价值判断的变化过程,我们不能说教育测验丝毫没有价值判断,如在测验什么不测验什么的问题上是有价值判断的;同样,我们也不能说教育评论就无点滴的事实判断,教育评论需要通过事实判断的真凭实据作论据。对上述分析,我们可以以事实判断为纵轴,以价值判断为横轴,用一条抛物线把从教育测验到教育评论的事实判断与价值判断的变化表示出来。见图 2 所示:

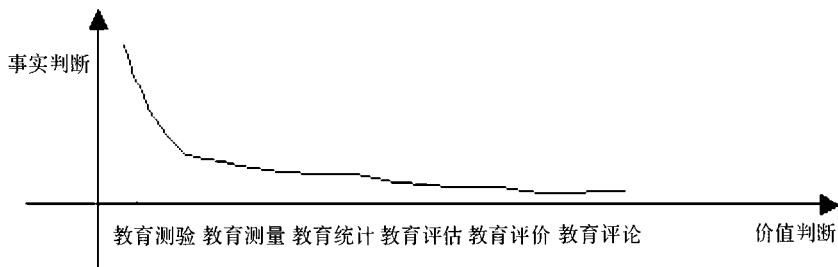


图 2 从教育测验到教育评论的事实判断与价值判断的变化

## 三、对教育测验到教育评论及其学科群落研究的建议

教育评价学经过几十年的发展已初步形成了学科体系,尽管概念、体系有待完善,但已为下一步的研究打下了坚实的基础。与此相应的考试学(研究教育测验的学科)、教育统计学、教育测量学、教育评估学、教育评论学都已有相应的发展,形成了一个共生的教育评价学学科群落。进入新世纪,要使这个生态群落得到良好的发展,我

认为首先应对各个学科的研究范围有个明确的划分,尤其是教育评价学要有合理的定位,不要再把考试学、教育统计学、教育测量学的内容纳入自己的势力范围,要让他们以独立的学科更好地为教育评价学服务。另外,教育评价学还要把属于教育精神文化、制度文化类的抽象的、教育评价无力评价的范围还给教育评论学,让教育评论学有更快的发展。同时,要把教育评估与教育评价的关系处理好,依我看也可以将两者合二为一,形成统一的学科。其次,各学科要明确自己的研究对象,形成概念体系和学科体系,把一些已成熟的概念和理论约定俗成稳定下来,逐步形成相对稳定的学科知识和概念体系。第三,要处理好理论与实践的关系,加强理论研究,使其适应实践的要求,要积极总结实践经验并升华为新的理论,丰富和发展教育评价学学科体系。第四,要处理好引进与消化吸收的关系问题。无论是从国外引进的教育评价学理论,还是从教育科学界以外引入教育的评价理论都要结合我国的教育理论和教育实践,进行修正使其“本土化”和“教育化”。

综上所述,搞清楚从教育测验到教育评论相关概念及其关系,对理清教育评价学学科群落内各学科的关系,促进这个学科群落的发展,为丰富和发展教育科学,加快教育事业的发展有着重大的理论和实践意义。

#### 参考文献

- [1] 刘 尧.关于教育评论学之我见[J].教育科学论坛,1995,(3):9~13.
- [2][4][6] 李 冀.教育管理辞典(第二版)[M].海口:海南出版社,1997:216,167.
- [3] 曹延亭.教育统计学基础[M].沈阳:辽宁人民出版社,1984:2~5.
- [5] 许茂祖.高等教育评估理论与方法[M].北京:中国铁道出版社,1997:16~27.
- [7] 刘 尧.教育评论学[M].北京:中国文联出版社,2000:7~23.
- [8] 侯光义.教育评价概论[M].石家庄:河北教育出版社,1999:58.
- [9] 刘 尧.教育评论与教育评价的区别[J].青岛科技大学学报,2002,(3):55~58.
- [10] 张玉田,等.学校教育评价[M].北京:中央民族学院出版社,1987:113.

(收稿日期:2007-01-10;编辑:谈 园)

Zhu Zhonghua

Page 17

**Abstract:** First, this paper analyzes the locality advantage of local universities and the character of talents cultivation. Then, taking as an example the universities in north Jiangsu, this paper discusses the role of development of local universities in building a well-off society in an all-round way.

**Key words:** local universities; well-off society; north Jiangsu; role

## Exploring the Group Cultivating Mode of Science and Technology Innovation Ability for University Students

Cao Xiaohua, Wu Qing, Li Wenfeng

Page 22

**Abstract:** This paper sums up the experiences and methods of cultivating innovation ability for university student in Chinese and foreign universities. With the consideration of the situation in our country, this paper puts forward a group cultivating mode as a major organization form, which is oriented by the regulations, organized in the form of the extracurricular innovation group, and demands setting up opening innovation labs and stable out-of-school practice bases, and strengthening practice.

**Key words:** university student; science and technology innovation; cultivating mode; innovation group

## Related Concepts from Educational Test to Education Review and Analysis of Their Relationships

Liu Yao

Page 25

**Abstract:** Mankind experiences the process from educational test to educational review in carrying out fact judgement and value judgement on education, and forms a discipline group of educational test, educational statistics, educational measurement, educational appraisalment, educational evaluation, and educational evaluation theories. However, since these theories and the discipline group are immature, many people fail to have a clear understanding of the related concepts and their relationships. Therefore, it is necessary to differentiate the related concepts and their relationships from a multi-view point, and to explore the construction of the discipline group.

**Key words:** educational test; educational measurement; educational evaluation; educational review

## Internal Quality Assurance in Higher Education Institutions: Student Evaluation of Teaching

Bi Jiaju

Page 31

**Abstract:** Internal quality assurance system is being initiated in many higher institutions in China. One of the important items in this system is the student evaluation of teaching. The paper intends to show an overall picture of student evaluation of teaching in foreign countries and Hong Kong mainly for faculty and staff members of quality assurance in higher education institutions in China. The main topics include the quality requirements of teaching, design of the questionnaires, and operation of the survey process. The results of some researches on reliability and validity are summarized, and the importance of nurturing and enhancing school quality culture is emphasized.

**Key words:** educational quality assurance; student evaluation of teaching; university management