

刘智运 编著
严新平 主审

高等学校教育评估 与督导概论



高等教育出版社

教育部 2013
年 12 月 1 日

高等学校教育评估 与督导概论

教育部 2013 年 12 月 1 日

高等学校教育评估 与督导概论

刘智运 编著
严新平 主审

高等教育出版社

内容提要

高等学校教育评估与督导是高等教育管理研究的重要领域之一,是教育管理和决策走向科学化、民主化和现代化的标志之一。就教育评估和督导这一领域研究内容的性质而言,“高等学校教育评估与督导”是一门综合性、交叉性和应用性很强的学科。就其发展历史而言,它又是一门正在发展中的新兴学科。

本书以马克思主义哲学、高等教育学和高等教育管理学为理论基础,以教育测量学、教育统计学和模糊数学、数理统计学为方法论基础,比较全面系统地论述了高等学校教育评估和教育督导的基本知识、基本理论和基本方法;结合我国高等学校开展教育评估与教育督导的实际,阐述了教育评估的涵义、历史演进及现代教育评估的特点、教育评估的指标体系与评估技术,介绍了我国高等学校开展的本科教学工作水平评估、专业评估、课程评估、教学质量评估和研究生教育评估等实施办法;并对我国高等教育督导制度的形成、教育督导的基本原则、组织建设和督导功能、机制及其发展趋势一一作了论述。

本书可作为高校开展评估与督导的参考书和培训教材,也可作为高等教育管理专业研究生、本科生的专业教材,还适用于广大评估专家、督导员、教师和教育研究人员阅读、参考。

图书在版编目(CIP)数据

高等学校教育评估与督导概论/刘智运编著. —北京:

高等教育出版社,2005.9

ISBN 7-04-017925-3

I. 高... II. 刘... III. ①高等学校-教育评估-研究-中国②高等学校-教育视导-研究-中国

IV. ①G642.0②G647

中国版本图书馆 CIP 数据核字 (2005) 第 068945 号

策划编辑	刘振天	责任编辑	刘振天	封面设计	王 睢	责任绘图
版式设计	胡志萍	责任校对	胡晓琪	责任印制	杨 明	

出版发行 高等教育出版社
社 址 北京市西城区德外大街 4 号
邮政编码 100011
总 机 010-58581000

经 销 北京蓝色畅想图书发行有限公司
印 刷 国防工业出版社印刷厂

开 本 787×1092 1/16
印 张 19
字 数 460 000

购书热线 010-58581118
免费咨询 800-810-0598
网 址 <http://www.hep.edu.cn>
<http://www.hep.com.cn>
网上订购 <http://www.landracom.com>
<http://www.landracom.cn>

版 次 2005 年 9 月第 1 版
印 次 2005 年 9 月第 1 次印刷
定 价 35.30 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 17925-00

序

高等学校教育评估是当前高等教育管理研究中的一个热门而颇有争议的领域;高等学校教育督导则是高等学校内部管理实践中一个悄然兴起而理论研究不足的领域。刘智运教授的《高等学校教育评估与督导概论》一书,把两者结合起来,并作为本书创新要点,粗看不难理解,深思则令人困惑。承嘱写字,我认为有必要对此谈谈个人学习这本著作,由粗到细、由浅入深的解读过程,以就正于作者和读者。

为什么粗看不难理解?因为教育管理部门对高等学校办学水平进行评估,是为了“要加强对高等教育的宏观的指导和管理”(《关于教育体制改革的决定》);而建立督导制度,则是为了“加强对教育工作的行政监督”(《教育督导暂行规定》)。两者的性质都是政府行为,两者的目的都是为了加强政府对教育工作的行政监督,从而两者的关系,不难理解为以评估作为督导的依据、手段,借以提高行政监督的权威性与有效性。本书将两者结合起来进行研究,顺理成章。

然而这种理解,似乎有悖于现代评估的理念。本书所介绍的现代教育评价发展的趋势之一,是从管理者的外部评价向自我评价转变。早期的评估,以管理者为评估主体,是自上而下的外部评估,被评对象处于被动等待的位置,积极性不高,甚至有抵触情绪和惧怕心理。现代教育评估,强调以自我评估为主,尊重与信任被评估者,使被评估者从被动受评转为主动合作参与,^①提高评估的质量,发挥评估的功能,达到“以评促改”、“以评促建”的目的。也就是说,评估活动不只是作为行政监督的手段。从行政监督的角度对待教育评估,正是当前中国高校评估实践产生急功近利的负面影响所在,显然不能以此来理解作者将评估与督导结合的新意与深意。

为了解除这一困惑,有必要弄清作者关于现代督导的新理念。作者认为现代督导工作要树立“以人为本”的理念,从教师的角度出发,提倡进行换位思考,把督导工作从原来以检查、监督为主转为以鼓励和帮助为主,摒弃监督、检查心态,抱着向教师学习的态度,怀着发现好的教学典型和总结成功经验的愿望,尊重教师、平等待人,帮助教师实现自我价值。^②这样,才能将评估与督导,在现代新的理念基础上结合起来。这就是作者在本书中最重要的创新之处:“把以人为本的理念,贯穿于评估和督导的全过程,在评估与督导中,给评估对象与督导对象的人文关怀。”^③

当前,高等学校的教育评估,正在全国大张旗鼓地进行。然而教育评估的指导思想,还停留在早期的评估理念上,亟须现代评估理论指导,以提高评估工作质量,减少所引起的负

① 见本书第6页。

② 见本书第229页。

③ 引自本书“前言”。



面影响;内部教育督导,虽悄然兴起,行政监督,只在中小学层面,可能高等教育也将面临主管部门的督导,更需现代督导理论的指导,刘智运教授这本著作的出版,是及时的。

我认识刘智运教授已有 20 多年,每次有关高等教育研究的学术会议,差不多都能见到刘教授;每次会议,他都认真地准备了内容充实的论文。他专心致志研究高等教育理论的精神和实事求是解决高等教育实际问题的态度,令人钦佩。这也体现在这部著作中。本书内容既充分反映了现代教育评估与现代教育督导的理念,又收集与设计了大量可行性的案例与方案。在他的研究与教学的生涯中,上了一个新台阶;对推动高等教育的创新与发展,做出了新贡献。当然,中国的教育评估与教育督导,正在发展中,一本著作所论述的理论与实践,很难说已经是完善的定论与最佳的方案,但无疑有重要的参考价值。

是为序。

潘懋元 于厦门大学教育研究院
2004. 6. 18

前 言

我已年逾古稀,早已离休,但我的心还是关注我国高等教育的改革与发展,情系高校的精英教育和大众化教育。由于连续几年扩招,更关心高等教育质量。可以说是离而不休。我在高教战线已奋斗了近五十年了,有很深厚的感情。现在是“老骥伏枥,志在千里,烈士暮年,壮心不已。”我离休后,精神还充实,身体还可以,总想做些力所能及的事。

我受聘担任学校的教育督导和教育评估工作,已有几年了。我想在我有生之年,把这些工作实践经验上升到理论高度作些思考和总结,整理出来,供同行参考。

如果本书对今后的教育评估与督导工作能做一点微薄的贡献,我将感到欣慰。也算是长河落日圆,余晖照人寰吧!有了这个想法,要付诸行动,困难不少。在得到学校领导的支持下,我就动手准备。提笔起草,挑灯夜战了。

本书命名为“高等学校教育评估与督导概论”,这是一本教育评估督导理论与实践相结合的作品,体现了我国高等教育的特点。

我国高等教育评估从1985年由教育部第一次在镜泊湖召开的高等工程教育评估专题研讨会开始至今,只有近20年的历史。我有幸参加了当时机械工业部教育局组织的机械制造及工艺专业的评估试点工作,并在《教育研究》杂志上发表过研究论文。这些年来,断断续续地做了一些高等教育的评估工作,积累了一定的实践经验。我国高等教育评估是在实践与研究中逐渐提高的,高等教育评估制度也是在实践中逐步形成的,一开始就注意联系中国高等教育的实际,探索和形成具有中国特色的高等教育评估体系。当时主抓这项工作的是教育部高等教育司副司长王冀生教授,他起到了关键作用,使我铭记不忘。我撰写这本书时,吸收了他的一些思想,也吸收了“具有中国特色的高等教育评估制度的研究和实践”课题组的研究成果,并在此基础上有所发展和创新。正如大科学家牛顿所说的:“如果说我比别人看得略为远些,那是因为我是站在巨人们的肩膀上的缘故。”

中国是一个有着漫长教育评价实践的国家。世界上最早的制度形态的教育评价实践,可以追溯到公元606年隋朝所建立的科举制度。科举制度从建立到1905年被废止,不仅对中国的社会发展和教育实践产生了巨大影响,也为建立现代教育评价理论提供了经验基础。

从较严格的意义上讲,教育评价成为理论研究的领域,并形成相对完整的理论体系,是在20世纪30年代,其标志是美国的泰勒(W. R. Tyler)以目标为中心的评价模式的建立。

我国的高等教育评价的研究始于20世纪80年代,1985年5月,《中共中央关于教育体制改革的决定》提出“教育管理部门还要组织教育界、知识界和用人部门定期对高等教育办学水平进行评估”,使得高等教育评估实践和研究日益为人们所重视,尽管起步较晚,但起点较高,发展较快。在我国的实际运用中,评估和评价通常是被人们当作同一概念在同一个意义上加以使用的。评估与评价一样,都译自英文的evaluation。由于英文evaluation一词的原意是引发出价值,也就是作出价值判断的意思。因此,评估这一词汇的内涵就被解释为:依



据一定的标准对客观事物进行观察并作出价值判断的过程。

时至今日,我国高等教育评价的领域正在不断扩大,并逐步形成了制度,对促进学校管理规范化、促进高等教育改革,起到了积极的作用。由于在高等教育评估中,遇到许多新情况和新问题,仅仅借鉴国外的评价理论和经验,已不足以解决我们的实际问题。因此,要研究和建立有中国特色的高等教育评估理论体系的任务显得十分紧迫,本书也是在这种形势下所作的尝试。

本书的创新之处有以下几点:

1. 把“以人为本”的理念,贯穿于评估和督导的全过程。在评估与督导中,给评估对象和督导对象以人文关怀。现代教育评估最大特点是评估者和被评估对象是平等协商关系,要相互合作与沟通。更加重视评估的诊断作用与形成性功能,更加重视改进(improvement),而不是鉴定(identification),已成为现代教育评估的发展趋势和最重要的特点。高校教育评估重点是促进学校管理的规范化和教育质量的提高,培养高素质创新型人才。

2. 把教育评估与督导结合起来。在评估督导过程中,一定要坚持求真务实的原则,要深入实际,实事求是,敢讲真话,敢揭真情,不徇私情,为学校决策者提供符合客观规律的意见和建议。

3. 在评估与督导的过程中,要努力探讨科学的评估督导方法,深入研究评估理论,提高评估督导的实效。

4. 积极探索构建中国特色的高等教育评估和督導體系。本书只作了初步尝试,还有待深入研究。

本书在编写过程中,参考了很多有关教育评估和督导方面的书籍和论文,并引用了其中一些材料,在此,谨向作者致以诚挚的谢意。

本书由武汉理工大学主管教学的副校长、博士生导师严新平教授主审。特请我国高等教育学创始人、博士生导师潘懋元教授作序。书中有关数学模型和模糊数学评判等内容特请武汉理工大学理学院数学系朱勇教授、张小柔教授审核。作者对他们的支持表示深深的感谢!

由于本人水平有限,撰写的时间仓促,本书肯定有不少错误和缺点,望读者不吝批评指正。

刘智运

2004年4月于武昌马房山

目 录

第一章 教育评估的产生和发展..... 1	第三节 课程评估的内容..... 150
第一节 教育评估的概念..... 1	第四节 课程评估的指标体系..... 153
第二节 历史背景及发展趋势..... 3	第五节 课程评估的实施..... 165
第二章 教育评估理论..... 7	第六节 毕业设计(论文)评估..... 167
第一节 教育评估的目的和功能..... 7	第九章 教育质量评估..... 170
第二节 高校教育评估理论..... 9	第一节 教育质量评估的目标及原则..... 170
第三节 高校教育评估的基本原则..... 13	第二节 课堂教学质量评估..... 172
第三章 教育评估方法..... 18	第三节 教师评估..... 176
第一节 高校教育评估方法..... 18	第四节 学生评估..... 184
第二节 评估指标体系的设计..... 24	第十章 中西方教育督导制度的
第三节 评估信息的获取和处理..... 33	历史演进..... 199
第四章 国外高校教育评估之比较	第一节 我国古代的教育督导..... 199
研究..... 53	第二节 我国近代的教育督导..... 201
第一节 国外高校教育评估简况..... 53	第三节 西方的教育督导..... 204
第二节 比较研究及启示..... 66	第十一章 中国教育督导制度的
第五章 我国高校教育评估制度的	新发展..... 214
建立与完善..... 71	第一节 高校教育督导悄然兴起..... 214
第一节 我国开展现代教育评估概况..... 71	第二节 教育督导的基本概念及其性质..... 215
第二节 高校教育评估的调整与完善..... 75	第三节 完善督导机制,强化督导功能..... 218
第三节 高校教育评估实施程序..... 77	第四节 加强教学督导工作的建议..... 219
第六章 高等学校综合评估..... 85	第十二章 构建中国特色的高校教育
第一节 高校评估的特点及评估目的	督導體制..... 222
与作用..... 85	第一节 组织建设..... 222
第二节 高校的合格评估..... 87	第二节 教育督导的基本原则..... 225
第三节 高校的水平评估..... 90	第三节 高校教育督导工作的创新..... 230
第七章 高校学科专业评估..... 113	附录..... 236
第一节 概述..... 113	1. 普通高等学校教育评估暂行规定..... 237
第二节 专业合格评估(专业鉴定)..... 114	2. 中美大学校长谈高校评估..... 240
第三节 专业的办学水平评估..... 119	3. 开展高等学校教育评估工作初探..... 243
第四节 研究生教育专业综合评估方案..... 139	4. 自评:促进了学校的发展与建设..... 249
第八章 课程评估..... 148	5. 武汉理工大学本科教学工作水平自
第一节 课程评估的意义和特点..... 148	评估方案..... 253
第二节 课程评估的类型和模式..... 149	主要参考文献..... 293

第一章 教育评估的产生和发展

第一节 教育评估的概念

教育是一种有目的的活动,所以一切教育活动中都含有评估,即对目的的实现的程度的评估。在我国,评估与评价通常被当作同一概念在同一个意义上使用,评估就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。我们认为,评估是人类社会生活中的一项基本活动,对于我们自己当前或过去的活动,人们无时无刻不在做出好与坏、成功与失败、长处和不足、有效与无效的判断。但是,这些判断活动并非都是有意识、并运用系统科学方法进行的。只有对那些认为是重要的活动进行有意识的评估,并运用科学系统的评估方法进行的评估,我们才称之为评估,这样的评估才有积极意义。在我国教育评估领域里,就词语而言,评估与评价一样,都译自英文的 evaluation。原意是引发出价值,也就是作出价值判断的意思。

教育评估是现代教育研究的三大领域之一。但如何定义教育评估,到目前为止,尚无一个被普遍认同的科学界定。

国外学者对教育评估有几种看法:

1. 泰勒(Tyler, R. W.)在著名的“八年研究”(1933—1940)报告中,首先提出教育评价的概念。认为,“评价过程在本质上是确定课程和教育大纲在实际上实现教育目标的程度的过程。”^①

2. 1966年,斯塔弗尔比姆(Stufflebeam, D. L.)在对泰勒评价理论提出异议的前提下,主张“教育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判断,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。”^②

3. 1975年,比贝(Beeby, C. E)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。”^③

4. 1979年出版的日本辰见敏夫编的《教育评价小辞典》中,对教育评价的解释是“依据教育目标和一定的价值观判断学生的学习成果、课程及教育活动状况符合目的性即实现目的的过程”。

5. 1981年美国教育评价标准联合委员会对教育评价进行了综合性界定,他们认为“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,为决策提供依据的过程。”^④

以上这些对教育评价内涵的界定是评估专家在教育评估实践和理论发展过程中提出的见解,虽有不同之处,但也不乏相同之点。为我们理解教育评估提供了依据和启示。相同之点都认

① 瞿葆奎主编:《教育论文集·教育评价》,人民教育出版社1989年版,第263页。

② 王致和主编:《高等学校教育评估》,北京师范大学出版社1995年版,第18页。

③ 同上,第19页。

④ 钱在森等主编:《普通教育评价原理与方法》,辽宁大学出版社1992年版,第5页。



为教育评价是对教育活动效果的价值判断过程。值得特别指出的是 1981 年美国教育评价标准联合会给教育评价所下的定义是需要给予重视的。“因为它是综合各家之长的一种较为全面的界定,也是综合了考试评价的形成和发展,反映教育评价的实质和功能的界定。它侧重于评价的信息作用,同时兼顾了教育目标与价值判断。”^①

我国学者根据我国教育评估的实践和理论研究,本着“面向世界”、“洋为中用”的精神,在学习国外评估理论的基础上,结合我国国情,对教育评估作出自己的阐释。把教育评估的内涵界定为“根据教育的目的和标准,采用科学的态度和方法,对教育工作中的活动、人员、管理和条件的状态与效绩,进行质和量的价值判断。”^②对这一定义应把握四层含义:一是抓住教育目的和标准,一般是指教育的总目标,各级各类学校的目标,各种教育活动的目标,各科教学的目标,教育改革的目标等;二是标准,是指某一次评估所依据的评估指标体系所规定的评估内容和标准;三是采取的态度和方法,是指实施教育评估时要采取实事求是的态度,选用规范的评估方法以及调查、实地考察、测量统计等评估工具和手段;四是明确评估的对象和范围以及教育活动的全部环节和领域。

总之,我认为,教育评估是根据教育目标和标准对教育工作做出价值判断并改进工作的过程。这里主要涉及以下几个基本理论问题:

(1) 教育评估与教育测量:教育评估是在教育测量的基础上发展起来的,教育评估与教育测量的本质区别就在于引入了统一的、科学的评估标准,可以在教育测量的基础上做出正确的价值判断,所以,教育评估的本质是价值判断。

(2) 声誉评估与发展评估:教育评估的目的是改进工作,这并不意味着仅仅应当重视发展评估(即“形成性评估”),因为声誉评估(即“终结性评估”,包括合格评估和选优评估)从根本上说也是为了改进工作。

(3) 定量评估与定性评估:定量评估与定性评估是相互联系的,应当相互结合。定量评估是定性评估的基础,所以应当高度重视定量评估。但是,定性评估是反映事物本质的,应当努力在定量评估的基础上做出科学的定性评估。脱离定性评估,一味地追求定量评估,搞所谓的“数字游戏”是不妥的。

(4) 教育评估的本质是价值判断(证明)与教育评估的目的是改进工作(改进)是一致的,改进工作是建立在准确的价值判断的基础之上的。所以,我们在实际工作中不能把二者对立起来,也不能简单地说“改进”比“证明”更重要。当然,如果把评估工作仅仅停留在价值判断上,搞所谓的“证明”,忽视在价值判断的基础上指出进一步改进工作的方向,那也是不对的。因为“证明”并不是评估的最终目的,评估的真正目的是“改进”,从这个意义上说,“改进”比“证明”更重要是正确的。

(5) 在教育评估活动中,评估主体与评估客体是一对矛盾。长期以来,我国教育评估的主体一直是政府,高等学校只是评估对象,处于比较被动的地位,而且长期忽视社会用人单位对高等学校的外部监督作用,这种状况应当迅速改变。尤其是当前我国高等教育发展、改革的重点已经从以政府为主体和以宏观领域为主向以高等学校为主体和以微观领域为主转移,与此相适应,我

① 钱在森等主编:《普通教育评价原理与方法》,辽宁大学出版社 1992 年版,第 5 页。

② 王汉澜主编:《教育评价学》,河南大学出版社 1998 年版,第 15 页。



国应当尽快建立以“以内为主,内外结合,以外促内”为主要特征的高等学校教育、教学质量保障体系,突出高等学校自主进行的教育、教学质量保障活动的基础地位,在这个基础上加强政府对高等学校整体办学水平及其教育教学质量的宏观监控和社会、特别是社会用人单位对高等学校教育教学质量的外部监督。这就是我们应建立的以高校自我评估为基础的,与外部评估相结合的高等学校教育评估制度。

在教育评估中,我们也常常采用教育测量的方法,对被评估对象特别是学生采取抽考,抽查毕业论文或考卷等方法。但教育评估与教育测量还是有区别的。这是两个既有联系又有区别的概念。教育测量主要是定量地描述所需实现的教育目标或目标实现的程度,是以多或少的数量表示的。而教育评估是对教育活动、教育系统工作情况进行分析判断,它是对评估对象作定性的表述,如合格不合格、好不好或分优、良、中、差等级。二者不同,但又有联系,不是对立的。教育评估与教育测量的产生与发展是分不开的,教育测量是教育评估的基础,但教育评估更反映事物的本质。

第二节 历史背景及发展趋势

国外许多学者提出,教育评估最早起源于中国。但在现代意义上的教育评估则产生于美国的30年代,是教育测量运动发展的产物。而现代教育评估在我国真正兴起,应该说是20世纪80年代以后的事。

一、教育测试的产生

教育评估是随学校教育的出现而产生的。自实施系统教育的机构——学校产生以来,就有了对学生入学、升级、学习情况、毕业的认定,就有了对教师教学质量的考核与评定,因为这是学校管理运作上不可或缺的要害。也就是,不管以什么样的形式,测定也好,评估也好,总是要对学生的学历以及教育结果进行检查确认的。在我国《学记》中就有记载:“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”。又如日本创立于7世纪律令制下的大学,就规定有旬试和年终考试两种考试。我国古代,自隋炀帝置进士科举开始(公元606年),延续了1300余年之久的封建科举制度,其实就是一种关于学生学历检测、评价的制度。但是,它缺乏对人全面的、科学的考察,并且存在着许多弊端,后被废止。

在18世纪以前的西方各国,由于学校尚未普及,学校考试主要是口试。1702年,英国的剑桥大学首先以笔试替代口试,开西方学校考试笔试之先河。1845年,美国初等学校普及,学生数量激增。对毕业生口试已不可能,于是波士顿教育委员会率先在美国相继以笔试取代口试,考察该市所属学校的毕业生、笔试优于口试。但在实践中,人们发现笔试也有不少弊端,因为评分易受主观偏见的影响,客观性仍受很大影响。为了确保测量的客观性,纠正弊端,出现了新的评估方法和评估体系,这就是教育测量运动。

二、教育测量运动

19世纪末到20世纪20年代,欧美一些学者在当时的实验心理学、统计学与智力测验发展



成果的影响和推动下,为排除以往评估笔试题编制的片面性和评分方法上的主观经验性,对学校中的考试进行了积极的改革,逐渐形成为教育测量运动。教育测量运动的发展历程经历了三个阶段:

初始阶段(1904年—1915年):首先是用笔试代替口试。应该说,从口试转向笔试,确实使评估向客观化迈出了一大步,使评估和测量手段更加灵活。但是笔试或书面考试,也存在着不足,即其试题内容仅为能够用文字在纸上表现出来的知识性教育成果,而无视对学生情感、意志等非智力因素方面发展的检测。此外,成绩评定即评分的客观标准问题尚未解决。这一时期的研究重点在于将测验客观化和标准化。

兴盛阶段(1915年—1930年):这一时期教育测量日益活跃,激起人们的普遍热情。其原因除了教育实践的需要外,还有两点:一是19世纪物理学、医学、天文学等自然科学的研究,由于使用了数量测定法而获得了快速发展,这样的发展给人文科学以极大的刺激和影响;二是受19世纪后半期个人差异心理现象研究和统计理论、方法发展的影响与推动。其中影响最大的是美国心理学家桑代克。作为教育测量运动的开拓者,他在《心理与社会测量导论》中提出了一个著名的论断:“凡是存在的东西都有数量,凡是有数量的东西都可以测量”。这一论断推动了教育成果的测量研究。此后陆续编制出测量标准、标准测验。嗣后,在欧美、日本等兴起了利用教育测量进行教育研究之风。由此可见,教育测量是以心理学、测量学、统计学为指导原理的。从一定意义上讲,教育测量运动是现代教育评估的原点,教育评估必须是客观的,有信度的。

批判阶段(1930年—1940年):随着教育测验运动的不断发展,人们逐渐认识到,教育测验尽管能使考试客观化、标准化,并能把人的能力换算成数字,甚至个别差异的程度也可以量化,但它毕竟不能测出人的全部,即使是研究最多、最富有成果的学力测验也不能测得学力的全部领域。在人格测验方面,单纯的测验是无法把握的,如社会态度、兴趣、情绪、情感、鉴赏力等。然而,这些都是教育的重要方面。从此,教育测量运动逐步过渡到教育评价的时期。

三、八年研究与教育评价概念的诞生

美国一些教育家成立了进步主义教育联盟(Progressive Educational Associator,简称PEA)。他们一方面批判旧的教育,一方面自我反省,寻找新的出路。这个联盟的负责人叫艾钦。他们以新的教育理论为依据,接受心理学的观点,以发展人的才能为目的,设计了一套新的课程,并选了7所大学、30多所中学进行实验。从1933年开始,实验进行了8年,到1940年提出了研究报告。这项工作是有创造性的,他们提出的评价原则和方法是系统的、完整的,是教育评价理论与方法的开端。

在W. R. 泰勒(W. R. Tyler)及其同事所进行的“八年研究”中,为了评价其成果的需要,他们正式提出了教育评价的概念,并提出了一个评价模式(这个模式被后人称为泰勒模式),这标志着教育评价已经形成。泰勒模式所关注的是:评价就是测定目标已达到的程度。就是根据预定的目标,通过系统地有目的地收集资料来判断实际教育活动是否达到了预定的教育目标或判断达到目标的程度。^①因此,目标本身的构成、目标的表现形式及其测量方法成为评价研究的重

^① 周朝森:“教育评价理论的新探——美国第四代教育评价述释”,《教育研究》1992年第2期。



点,与泰勒模式相关的教育评价领域有两项重要的工作:一是标准化测验及其推广;二是布卢姆(B. S. Bloom)等人进行的目标分类学研究及其成果。

泰勒模式提供了教育评价理论的最初形态,也提供了评价的方法论。他认为,应该从各方面对评价对象进行评价,评价不仅是分析的,也是综合的,不能只局限于书面的测验。特别是关于教育目标方面的研究更是反映了教育评价的本质——价值判断,因为教育目标从本质上讲是价值的一种表现形式。通过实践,人们又提出了一些新的评价模式,这些模式中,较有影响的有斯塔弗尔比姆(L. D. Stufflebeam)的 CIPP 模式及美国加利福尼亚州立大学评价中心提出的 CSE 模式等。这些模式的提出既丰富了教育评价理论,也为教育评价理论进入一个新的发展阶段进行了理论反思和实践的准备。特别是应答模式更成为所谓的“第四代评价”的先声。这个时期的教育评价的目的在于总体上还是为了判断是否具有某种价值,并强调价值判断是教育评价的关键工作。在一定程度上我们可以认为 CIPP 模式以及应答模式等都是为了作出正确价值判断在收集评价信息范围与方法上所作出的改进。当然,也应当看到这些评价模式的应用范围已经突破了教育的课程和学习评价范围,而进入教育其他领域。

此外,斯里克文在《评价方法论》中,将评价方法进行分类所提出的总结性和形成性评价也是这个阶段的重要工作。这说明研究者已经注意到评价实践中评价活动之间的差异,并上升到理论的高度。

20 世纪 80 年代到现在。这个时期评价理论的主要特征是评价主体与价值主体分离,价值主体与价值客体同一,评价主体价值中立,强调教育评价为其评价的对象服务。

四、现代教育评价发展趋势及特点

从以上教育评价产生发展的历程中,我们可以看到其变化、发展的趋势和特点。国内外学者将其基本方面概括为:

(一) 在教育评估的目的和功能的选择上,从注重“终结性评价”方法向注重“形成性评价”方法与终结性评价方法相结合转变

终结性评价重评估结果,形成性评价重评价过程。所谓终结性评价就是对活动结果的评价。它的目的在于对被评对象作出某种资格证明,比如通过评估,对某一所学校的办学水平加以认可。形成性评价与此不同,它是揭示被评对象在运行过程中存在的问题,反馈有关信息,以评促改,以评促建。因此,现代教育评价已从注重评价为分等、证明服务,转向为改进教育工作服务。正像 CIPP 模式创始人斯塔弗尔比姆强调的那样,方案评价最重要的目的不是证明而是改进。

(二) 在教育评估的对象、范围上,从学生的学力和课程扩展到教育的全部领域

在美国教育评估的“八年研究”中,主要是研究课程设置对教学质量的影响,因而评价对象,集中评价学生的学习成绩,并在此基础上对教育计划和课程编制的优劣得失作出判断。而现代教育评估则把教育的全部领域或者说与教育活动相关的各个领域作为评估对象。如课程设置的科学性、教育结构的合理性、社会对人才需求(质和量)的适应性、师资队伍的建设、教育科研的学术水平、办学条件的达标率等等,展开了一场教育的全方位评价。

(三) 在教育评估的方法和技术上,从重视定量方法向注重定量方法与定性方法结合转变

在教育测量运动初期,桑代克提出的“凡是存在的东西都有数量;凡是有数量的东西都可以测量”这一著名的观点,对当时的教育测量理论的发展有重要影响,使定量方法一时受到人们的



极大重视。同时,在自然科学的影响下,运用实验设计的方法(如在评价设计中运用对照组进行对比)以及客观资料(如各种统计资料)的方法也占据了主导地位。近年来,随着现代教育评价活动的深入,人们愈来愈发现,片面追求定量的、实验设计的方法,往往导致评价者对那些一时难以定量的或由各种条件限制、实验难以反映确切情况的那些因素的忽视。然而,这种因素对教育活动来说又具有极为重要的影响。离开这些因素的评价,教育评价则是不全面的(比如,情感、情绪、鉴赏力等)。所以,现代教育评价在选择评价方法时,强调在有效使用定量分析的基础上,重视定性分析,使二者有机地结合。

(四) 在评价方式上,从重视管理者的外部评价向重视自我评价转变,充分强调自我评价方法的作用

早期泰勒主持的教育评价活动,主要反映了管理者的意见,通常以管理者为评估主体,对被评估者实施自上而下的评估,这种评估对被评估者来说是来自外部的评估,他完全处于被动等待评估的位置。因此,往往积极性不高,甚至有抵触情绪和惧怕心理。现代教育评估的民主化进程,动摇了这种做法,要求重视被评估者在评估活动中的地位和作用,强调以自我评估为主。这是从尊重、信任被评估者的民主要求出发,达到调动被评估者的自觉性、积极性,使他们从被动接受评估转为主动合作参与评估的目的。通常把被评估者的自我评估作为整个评估过程的预评估阶段,即教育评估的基础,这样不仅有利于被评估者自己发现问题,从而改进工作,也有利于专家作出公正客观的评估结论,更好的提高评估的工作质量和可信度。

第二章 教育评估理论

第一节 教育评估的目的和功能

一、教育评估的目的

在每一次评估启动之前,操作者都必须先回答这样一个问题,即该次评估活动要达到什么目的。不管是哪个领域的评估,如教育评估,不管是哪一类评估,如政府评估,都应该有明确的目的。

马克思主义根据意识与存在的关系,第一次对人类活动的目的性作出了科学的说明。马克思指出,人的活动是有目的的,即在活动结果产生出来之前,先在头脑里以观念的形式构造出来。对此我们可以理解为,人们进行任何活动的目的都是由人们所规定的。当然,这不是说活动目的是人在头脑中主观臆造出来的。马克思深刻指明这是客观见之于主观的东西,是主客观的统一。这正是人类活动与动物活动的根本区别。动物的活动只是一种本能的生命活动。目的,对于评估活动及具体到教育督导评估活动来说,同样是第一重要的。随着管理实践与理论的发展,目的也成为管理学的一个基本范畴。任何管理活动都有目的。其目的至少有两层意义:一是管理活动的终极目的是实现管理客体的使命和需求;二是管理职能和手段的选择、发挥和使用的目的。而后者是为前者服务的。因为一般管理理论认为,管理活动的功能(或称职能)及手段与管理活动的目的有着密切的关系,在一定意义上,确定目的就是选择功能。为此,我们对教育评估活动目的确定,通常也是从评估功能的选择入手的。

(一) 从评估功能的内容上对评估目的选择

教育评估的功能是多方面的,如检查鉴定功能、反馈指导功能、激励功能、预防功能、补救功能、应用推广功能等。但就其功能的内容而言,最根本的是监督鉴定与指导改进两个方面,其他功能都可以看作是这两个功能的延伸。比如反馈功能就是观察、监督检查鉴定功能的延伸。因此,在评估中,功能的选择主要是考虑以哪一个功能为主导功能的问题,而不是二者择一的问题。因为两个功能是有一定联系的。特别是指导、改进、发展功能的实现是需要以检查、鉴定为前提、为条件的。严格地讲,只能是以哪一个功能为主,或者为最终目的选择确立的问题。如评估整个学校的办学水平与评某一个专业在最终目的上是不一样的。这就要依评估内容而定。

(二) 从评估功能作用的领域选择与确定其目的

就教育整体而言,教育评估是为了端正教育思想,提高教育质量,这可以说是教育评价的总目的。但就教育不同领域活动的评估而言,其要求不同,具体的评估目的也不同,尽管它们的最终目的也是为提高教育质量服务的。目前,教育评估研究把教育诸领域的评估的具体目的概括地分为四个方面:

1. 指导教育教学活动的目的。包括基于教师等对学生施以教育教学的指导,为了选择决策



编制更加科学有效教育教学计划而实施的评估,在学期末及学年末以教育教学计划的恰当与否,实施结果实效如何,今后怎样改善等为评估目的终结评估,以考察一所学校的办学水平为目的的综合评估。

2. 为学生学习服务的目的。以学生对学习的自我调整为目的的自我评估,也就是学生自己评估自己的学习状况,并进行自我调控。学习心理学指出,将学生的学习结果交给学生评估,这对学生来说是最好的学习,也能调动学生学习的主动性。在这个意义上,这一评估目的,现在普遍受到重视。这种评估最好贯穿于学生学习的全过程,是注重过程的形成性评估。比如,在日常教学中,学生对教师提问的回答,学生作业批改情况的不断反馈等。当然对学生自我评估也不能采取放任自流的态度,任课教师和教育管理者也要加强指导和组织工作。我们不能单靠学生的自我评估,因为学生也很难客观准确地评估自己,外部(包括教师和学校)对学生的检查性评估还是必要的。

3. 为管理服务的目的。通过科学的评估,可以区分优劣、明辨是非,可以使教育工作者知道自己的工作成效,还可以使社会上的人士了解学校的成就和不足。通过评估,使每前进一步都是朝着既定目标迈进,保证达到目的。目前,评估目的的选择重心已移向对管理行为、管理活动的改进上。

4. 为调查研究服务的目的。这是教育评估的新的目标。评估已引起政府和社会上的广泛关注,评估的权威性和可信度越来越高,做好评估的调查研究和改进工作非常重要。美国成立了许多社会中介机构,从事教育评估。我国政府也很重视教育评估,正在建立评估制度和中介机构。此外,开展评估活动也需要开展调查研究,如何做好调查研究,收集到可靠的信息和资料,是至关重要的。例如,一个国家,一个地区或教育主管部门,为了解其实施的教育计划取得了哪些成效,还有哪些不足而进行的评估,都具有调查研究的成分。

总之,确定评估目的是评估程序的第一步。它的重要性无可置疑。简单地说,它决定着所进行的评估采用什么样的评估标准,也决定着评估工具、手段的选择,影响着评估的组织实施的具体做法。评估目的是一个纲,纲举目张。

二、教育评估的功能

教育评估的功能是指教育评估活动本身所具有的能引起评估对象变化的功用和能力。它通过教育评估活动与结果作用于评估对象而体现出来。其功能的内容取决于评估活动的结构及运行机制。

在教育教学活动、学校管理过程中,教育评估有着不同的类型。不同类型的教育评估所具有的主要功能是不同的,而教育教学活动进程、学校管理进程不同阶段所需要的教育评估的类型也是不同的。系统理论告诉我们,教育活动是一个由教育目标、教育教学活动开展、教育评估所组成的系统过程。学校管理也是一个由教育决策、决策执行、反馈控制活动所组成的系统。不管在哪一个系统中,作为具有反馈控制功能的教育评估,在整个教育活动中,都占有非常重要的地位。它是教育过程中一个基本环节。也是学校管理过程中的一个基本步骤。从整个教育活动、学校管理活动来讲,如果没有教育评估工作,只有教育目标、教育教学活动,只有决策及决策执行活动,实践中必然会使教育教学活动缺乏现实基础和科学依据。缺乏检查指导,必然会使学校管理的决策缺乏科学论证,执行缺乏检查监督,最终影响教育目标的实现,降低学校管理的效率和