

目 录

第一章 教育评价理论的产生与发展	1
第一节 教育评价理论产生的历史背景	1
第二节 现代教育评价理论的发展趋势与特点	6
第三节 现代西方教育评价理论的主要派别	11
第二章 教育评价的意义和作用	15
第一节 教育评价的意义	15
第二节 教育评价的作用	20
第三章 教育评价原则	23
第一节 确立教育评价原则的必要性	23
第二节 教育评价原则	25
第三节 对教育评价原则的认识	32
第四章 教育评价标准概述	35
第一节 教育评价标准的结构	35
第二节 教育评价标准的设计	43
第四章 教育评价方法	49
第一节 教育评价方法评述	49
第二节 统计分析法	59
第三节 综合评判法	66
第六章 教育评价用具	83
第一节 教育评价用具的概述与种类	83
第二节 考试和测试	86
第三节 问卷、接谈、行动观察记录及其他评价用具	95
第七章 评价心理及其调控	103
第一节 概述	103

第二节	主评心理与调控·····	109
第三节	被评心理与调控·····	119
第八章	教学评价 ·····	128
第一节	教学评价概述·····	128
第二节	教育授课质量的评价·····	131
第三节	学习成绩的测验评价·····	143
第四节	学科成绩测量评价举隅——音乐、美术科成绩的测量评价·····	156
第九章	学力评价 ·····	164
第一节	学力概念·····	164
第二节	学习水平的测验评价·····	169
第三节	学生学习的实际可能性的评价·····	175
第四节	智力、特殊能力的测验评价·····	182
第十章	教育督导概述 ·····	190
第一节	教育督导的基本概念和历史沿革·····	190
第二节	建立教育督导制度的意义·····	195
第三节	教育督导的性质·····	197
第四节	教育督导的任务和对象·····	199
第十一章	机构 ·····	204
第一节	教育督导机构的设置·····	204
第二节	各级教育督导机构的职责划分和上下关系·····	206
第三节	教育督导机构与教育行政部门的关系·····	209
第四节	教育督导机构的自身建设·····	212
第十二章	督学 ·····	216
第一节	督学的职责和权限·····	216
第二节	督学的选拔和任职条件·····	219
第三节	督学的素质要求·····	220
第四节	督学的培训·····	224
第五节	督学的管理、考核和聘任·····	229

第十三章 督导	233
第一节 教育督导的分类和原则.....	233
第二节 教育督导的过程与步骤.....	236
第三节 教育督导的常用方法.....	239
第四节 教育督导报告的撰写.....	243
第五节 教育督导档案的建设.....	246
第六节 教育督导与教育评价.....	250
第十四章 对地区教育工作的督导评估	257
第一节 对地区教育工作督导的意义和必要性.....	257
第二节 对地区教育工作督导的主要内容.....	258
第三节 对地区教育工作督导的步骤和方法.....	262
第四节 在地区教育工作督导中应注意的问题.....	263
第十五章 对学校工作的督导评估	268
第一节 对学校工作督导评估的目的和类型.....	268
第二节 对学校工作督导评估的意义.....	269
第三节 对学校督导评估的步骤和方法.....	271
第四节 在学校工作督导评估中应注意的问题.....	275
第十六章 专项督导评估	282
第一节 对德育工作的督导评估.....	282
第二节 办学方向的督导评估.....	286
第三节 对学校管理的督导评估.....	289
第四节 对教育经费的督导评估.....	291
第五节 对教师队伍建设的督导评估.....	294
后 记	302

第一章

教育评价理论的产生与发展

人与动物的区别就在于思想的存在,他能在追求比以前更好的目标、效率的过程中,对自身行为的结果做出反省与评价。评价存在于人的一切有目的的活动之中。例如,一个人去商店买某种商品,常常要对商品的规格、颜色、质量进行评比,最后作出决择,在这个活动中就包含着评价。教育是一种有目的的活动,所以在一切教育活动中都含有评价。从这个意义上讲,教育评价是伴随着教育的产生而产生的,并且是随着教育的发展而发展的。纵观历史,教育评价理论的发展大致经过了古代的传统考试、近现代的科学测试和当代的科学评价三种不同阶段。

第一节 教育评价理论产生的历史背景

一、教育测试的萌芽

教育评价理论的产生,究其历史根源来说,是源于对学生的学力检验。在古代中国,自隋炀帝置进士科始(公元606年),沿续了一千三百年之久的封建科举制度,其实就是一个关于学生学历检测、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。在十八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,对毕业生一一口试已不可能,于是,波士顿市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。由于笔试客观性、可靠性比口试高,并且节约时间,测试结果

大大优于口试。但是,因为评分易受主观偏见影响,况且试题太少,不足以反映学生所获知识与能力的全貌,因此,客观性仍大受影响。为矫正此弊,力求考试的客观化,于是测验方法随之出现。

1864年,英国格林威治医学校校长菲奢(G·Fisher)搜集了本校学生的答案与作品,进行客观评分,并依据一定的价值标准汇编成标准评分“尺度簿”(Scale—Book),试图为当时的考试提供一个可供参考的标准。但是,由于种种原因,菲奢的这一开创性工作却没有引起教育界的重视,因而未能产生多大的影响。

在教育界引起人们对测验问题极大关注的是著名的莱斯(J·M·Rice)拼字测验。1897年,莱斯发表了他对二十个学校三千余名学生所作的拼字测验研究的结果,测验表明:八年中每天花45分钟时间进行拼字练习的同每天花15分钟进行练习的成绩并没有多大的差别。这一结论尽管遭到了不少人的反对,但它引起了人们对测验问题的普遍关心,推动了教育测验问题的研究,因此,莱斯获得很高荣誉,甚至有人称他为教育测量的创始人。

二、美国的教育测验运动

前面所叙述的只是教育测验先驱们的工作,在测验水平上仅仅是用一定测定的尺度,寻求一定客观量的结果。教育测验在美国真正形成一个运动还是本世纪二十年代的事情。然而,教育测验之所以在本世纪二十年代形成一种运动,表面上是针对论文或考试不客观而发,实际上,究其原因有如下几个方面:

(1) 自然科学的急速发展

中世纪以后,西方自然科学由于在方法论上引入了测定、观察和实验,一些传统学科如物理学、化学、医学、天文学等有了长足的发展。1879年,冯特在德国莱比锡设立了第一个心理学实验室,为进一步揭示人的心理本质问题而设计了种种周密的实验方案与实现方案的各种严密的测量方法。所有这些方法论的变革促进了教育测量运动的兴起。

(2) 个别差异的研究与统计学在教育上的运用

1882年,英国高尔顿(Galton, SirF)受达尔文影响,在伦敦设立了人类学测验实验室和德国冯特的实验心理学派相对峙。德国实验心理学派通过专门研究人类的一般行为规律来进一步揭示人的心理本质,而英国人类学派则对人类个别差异寄与关心。高尔顿在统计者皮尔逊(Pearson, K)的帮助下,设计了许多统计方法。这些方法不仅对美国的人事工程思想(想把人类之能力一一量化成为数目,以为代表个人之价值)大有裨益,同时,一些教育家借鉴这些统计方法,在教育上把不同学生的学习能力与学习效果量化,并加以客观比较,促进了教育测量运动的发展。

(3)对特殊儿童的研究

十九世纪末,法国在“自由”、“平等”、“博爱”等进步思想的影响下,提出社会不仅要对身体有缺陷的儿童加以关怀,而且还要对心理、精神智能有缺陷的儿童给予相当的关怀。恰好当时巴黎一带的学校有许多学习效果低劣的孩子很成问题,比奈想鉴别这些孩子是因为懒惰、还是因为智能低下而不能适应,并且想在教育上尝试如何加以补救,因此,在1895年,比奈和亨利(Henri, V)设计了一套智力测验的方法。1905年西蒙(Simon, T)的协助下,制成了著名的比奈—西蒙智力量表。

上述三种研究的流潮,加上美国原有的“人事工程思想”,都是试图对学生的心理及学习效果、先天和后天的所有能力及性质用物理学的方法拟以数字计量,以求其客观性。本世纪二十年代,美国的教育测验运动就是在这种思想潮流下蓬勃发展起来的。

美国的教育测验运动的发展可分以下三个时期:

(1)1904年—1915年的开拓期。这一段时间是方法的探索与初步的发展时期。在美国心理学家卡特尔研究的基础上,1904年桑代克发表了《精神与社会测验导论》。这是一本在测验学史上划时代的巨著,标志着教育测验运动的开始。桑代克在书中宣称:“凡是存在的东西都有数量;凡是有数量的东西都可测量”。在这一时期曾修订了比奈智力测验量表;奥比斯(Otis, A. S.)设计了团

体智力测验。在学历测验方面,有桑代克的书法量表与斯通(Stone, C. W.)的标准算术测验。这些充分表明这一时期的研究重点在于将测验客观化和标准化。

(2)1915年—1930年的兴盛期。这一时期对桑代克提出的信条(凡是存在的东西都有数量;凡是有数量的东西都可以测量),不但在技术方面努力求得正确应用,而且不断提高到理论上加以证实。这一时期教育测验已发展成三种不同性质的测验,即学历测验、智力测验和人格测验。在学历测验方面,从桑代克发表书法量表到1928年为止,据统计,已有标准心理测验和标准学历测验三千余种之多。在智力测验方面,自比奈—西蒙智力测验传入美国后,经辜鲁满(Kunlman, F)及斯坦福大学推孟(Terman, C·M)的相继修改、加以量化,并引用德国斯特恩(Stern, W.)智能商数(IQ)。在第一次世界大战中,美国军队首次用此量表对美国士兵进行测验。另外,在职业指导和升学指导方面还有性向测验(Aptituae Test)和诊断测验,可谓五花八门、丰富多彩。在人格测验方面,教育家们基于“不能正确把握学生个性,所进行的教育,犹如在黑夜中射箭”这一教育信念。1921年,华纳德(Fernald, G·G)着手试作人格测验,1924年至1929年,哈芝红(Hartshorne, H.)等人组织了人格教育委员会,着手研究人格测验工具,并不断加以改进,使之相当精密。

(3)1930年—1940年教育测验运动的批判期。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、最富有成果的学历测验也不能测得学历的全部领域。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,这些又都是教育的重要方面。1931年,塞蒙兹(Symonbs, P·M)发表了《人格与行动的诊断》一书,主张人格测量应用评定法(rating method),问卷法(Questionnaire method)、交谈

法(Interview method)、轶事记录法(anecdotal record)等,从而在思想观念上否定了单纯的人格测量法。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

美国教育评价理论的产生与发展,不但有上述教育自身发展的要求,同时,也有其深刻的社会历史背景。

美国在十九世纪末二十世纪初,经济发展比较顺利,因而美国的一些思想家便认为美国是一个完美的社会,是可以自由发展的国家。但是,在1929年,美国发生了一次规模巨大的经济危机,事实使美国很多人的头脑开始清醒了。1933年,罗斯福实行了“经济的社会化政策”,这个政策强调政府与产业的协调体制。随着这种经济上的限制与干涉的出现,大批青年没有就业机会,只能涌向中学。他们根本没有上大学的要求与兴趣,但是,美国的中学在大学入学考试这根指挥棒的作用下,课程设置主要是为升大学服务,不适应整个社会和失业青年的需要,这就使学生与学校课程之间发生了尖锐的矛盾。在这种情况下,美国的一些教育家成立了进步主义教育联盟(Progressive Educational Association,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,以全面发展人的才能为目的,设计了一套新的课程,并选了七所大学、三十多所中学进行实验。实验的方法是参加实验的三十多所中学一律用新课程、新方法授课,学生毕业之后可以直接升入参加实验的十三所大学。在这十三所大学里,每所都有一半学生来自参加实验的所谓进步大学,另一半学生是来自未参加实验的所谓传统学校。他们边实验边评价,为了进行评价,还专门成立了一个评价委员会,评价委员会的负责人是泰勒教授。从1933年开始,实验了八年,到1940年提出一个报告,美国教育史称之为“八年研究”。在八年研究的报告中详细地阐述了课程设置的原则和评价的原则,并且根据评价原则论述了试验的结果。八年研究

的报告中强调指出,评价对于试验的成功起了积极的、重大的作用。当然,报告是用事实说明实验是成功的。

下面的表1是他们证明试验成功的一个论据。

表1 新旧课程优劣之评价比较

教 育 目 标	平 均 分	标 准 差
市民信仰 < 进步学校 传统学校	123.05	14.88
	108.68	15.54
独创性活动 < 进步学校 传统学校	86.7	50.5
	32	36.9
协作的活动 < 进步学校 传统学校	65.3	46.0
	31.1	37.5
记忆的活动 < 进步学校 传统学校	43.0	40.4
	49.2	44.1

上表系“八年研究”比较进步学校和传统学校所用新课程和旧课程之结果。从表中可看出进步学校使用新课程的学生比传统学校使用旧课程的学生,在社会、政治、经济、文化等方面的信仰与态度远为进步与开化。独创性活动、集体性活动,进步学校的学生比传统学校学生无论在质或量上都有显著的差别,只有记忆一项,传统学校优于进步学校,而且,差距之小在统计学上无意义存在。因此,可以证明:进步学校的新课程优于传统学校的旧课程。

八年研究中评价委员会的工作是有创造性的。他们提出的评价原则和方法是系统的、完整的,是教育评价理论和方法的开端。

第二节 现代教育评价理论的发展趋势与特点

一、教育评价理论发展的两个阶段与特点

教育评价的发展可以划分为两个阶段。

第一个阶段是从 1940 年到 1957 年,这个阶段的特点是:

(1)教育评价理论和方法得到了进一步充实和发展

美国八年研究报告中关于教育评价的理论和方法被人们接受之后,人们普遍认为,教育评价是教育领域中的重要课题。但是,为了使评价能够达到预期的目的,关键是要明确每种教育活动的目标,为此,很多人着手研究教育目标问题。在此期间,布鲁姆建立了目标分类学。布鲁姆的目标分类学是关于认知领域的,其中包括知识的掌握、理解和认识能力的发展。建立目标分类学的任务是 1948 年提出的,直到 1956 年才完成。与此同时,有人研究情意方面的教育目标(包括兴趣、态度、价值观的形成和正确的判断能力、适应能力的发展)。这个目标体系由克雷斯沃尔于 1964 年完成。

(2)教育评价活动尚未普及

这个时期教育评价活动仅限于研究教育的专家学者和从事教育工作的实际工作者,尚未引起政府与群众的普遍重视。

第二阶段,从 1957 年到现在。这个阶段的特点是:

(1)教育评价活动逐步普及

这个时期教育评价由教育工作者所关心的课题逐步变成了各级政府和广大群众所关心的课题。1957 年苏联人造地球卫星发射成功之后,美国人分析他们自己落后的原因在于教育。教育落后的状况不仅震动了统治集团,也震动了美国的公众,于是美国中等教育遭到了舆论的猛烈抨击。舆论界纷纷质问:“美国学校为什么产生不出科学家?”“为什么美国的中等教育如此拙劣?”据当时调查洛杉矶初级中学一万一千名学生的情况表明,30%的人认不出钟表上的时刻;18%的人不知道一年有几个月;9%的人不知道七角伍分钱可以买三分钱一张的邮票多少张;5%的人不知道七十的一半是多少;4%的人不认识字母表。公众要求学校提出证据证明学校办学的成效,随后,政府当局大规模地开展评价工作,有全州性的或全市性的评价活动。由此,评价由专家的课题转入政府

部门的教育措施。

(2)由早期注重“终结性评价”方法到现代注重“形成性评价”方法

首先,过去的教育评价为选拔适合教育的儿童服务,而现代教育评价则注重创造适合儿童的教育。其次,现代教育评价从早期重视终结性评价(Summative evaluation)逐步转入更重视形成性评价(Formative evaluation)。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价对象作出某种资格证明,比如通过鉴定,对某一学校的办学能力加以认可。形成性评价与此不同,它是通过揭示存在的问题、反馈有关信息,以促进工作的改进。因此,在对教育方案、教育机构、团体的评价中,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 模式的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。

(3)在评价过程中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,而现代教育评价理论相对而言则更重视反映各类人员的意见,特别是注重听取被评价者(包括被评价团体、机构有关人员)的自我检查、自我评价的意见,并且把被评价者自我评价的意见与评价者(包括外部团体、机构组织)的检查、评价有机地结合起来。通常现代教育评价活动主张把被评者的自我评价作为整个评价过程的预评价阶段,这样不仅有利于被评价者自己发现问题,从而改进工作,同时也有利于在被评价者自我评价的基础上,由专家们作出的结论为被评价者接受和理解,防止产生对立情绪。

(4)在评价方法上注意定性与定量的结合、实验设计与非实验设计的结合

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,对当时的测量理论的发展有重要影响,使定量方法一时受到人们极大的重视。

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)。因而,近年来,国外的教育评价理论又开始重视定性分析方法和非实验的自然观察法,并且在教育评价活动中力图不片面追求定量法、实验法,努力把定性分析与定量分析、实验设计与非实验设计方法有机地结合起来,根据实际情况灵活地加以运用。

(5)新的科学方法的引用

近年来,由于系统科学方法论的异军突起,人们逐渐开始借鉴系统科学方法论来开展教育评价。如在评价某一个教育现象或教育过程时,把它们放到一个更大的系统中加以评价,这样就更加具有科学性。另外,更重要的是模糊数学和电子计算机的引入。这两种方法在处理教育领域中大量的模糊现象(如知识结构、品德、智力、能力、情感、意志等)时显示出巨大的生命力和渗透力,极大地促进了关于教育、教学质量评价量化(即质的量)问题的研究。

(6)立体评价与全方位评价

在八年研究中主要是研究课程设置对教学质量的影响,因而评价的内容集中于学生的信仰、创造力、协作精神以及记忆能力的发展水平。现在的教育评价,不仅是考察教育的结果——学生的学习质量和发展水平,而且要考察整个教育过程的诸方面,比如对学生的评价(包括品德、智力、知识、体质等)、教师的评价(包括师德、素质、工作量等)和学校管理者的评价等。不仅如此,近年来,教育评价的范围由教育、教学的微观领域又进一步扩展到了教育、教学的宏观领域。如课程设置的科学性,教育结构的合理性,社会

对人才需求(质与量)的适宜性,师资队伍的建设,教育科研的学术水平等等,展开了一场教育的立体评价和全方位评价。

(7)对评价结果持慎重态度

所谓教育评价往往是根据对一组具有共同特征的对象普遍适用的标准去评价某一特定的对象,如某一特定的个人、学校和教育方案。但是,由于对象与对象所处的环境不同、条件不同等等,因此,由着眼于评价对象共性方面获得的评价结论究竟能说明什么问题,还需要进行全面的分析综合之后才能下断语。鉴于上述原因,近年来,西方不少教育评价工作者都提出了评价结论的慎重解释问题,甚至有人就根本反对给评价对象下结论。他们提出,在一次评价以后是否需要作出正式报告,须根据被评者的愿望而定,评价者只需把你见到的事实及各种人员的想法告诉被评价者就可以了。对评价结果持慎重的态度,是现代教育评价的又一特点。

(8)教育评价理论体系逐渐完善

由于国外近三十年来教育评价活动的普及,教育评价理论也逐步完善起来,相继提出了比较系统的教育质量评价原则、评价方法及评价程序等,并试图着手建立教育质量评价体系。

目前,教育评价工作已经引起了许多国家的重视。1984年12月20日,美国联邦政府教育部公布的高等教育问题报告中指出:“要设计并实行科学评价学生知识、能力、技能的系统方法”。

日本在第二次世界大战之后受美国的影响较大,他们在教育的各个领域里开展了教育评价工作,已出版的教育评价专著很多。

二、促进教育评价理论发展的几个原因

当前,各国对教育评价工作的普遍重视是有一定的原因的。其原因概括起来有以下几个方面:

(1)教育实验和教育改革需要教育评价的发展。当前,世界各国都面临着新技术革命的挑战,新的技术革命要求教育必要改革。既然要搞教育实验和教育改革,就有一个如何判断一项新措施功效的问题,这就需要一套科学的评价技术,为教育决策提供科学的

依据。

(2)各国教育投资的效益评定需要教育评价的发展。当前,各国为发展教育,逐渐增加教育投资,这样就普遍存在一个教育投资的效益评定问题。因此,以评定教育投资的效益为背景,教育评价受到了普遍的重视。

(3)教育基础理论的发展推动了教育评价理论的改进。目前,世界上较有影响的教育理论是美国布鲁纳的结构课程论和前苏联赞可夫的教学与发展论。两者都涉及到只有正确的评价,才能不断提高教育质量,从而推动教育评价理论不断完善。

(4)人民群众对传统学历社会的不满促进了教育评价的改革。今日的资本主义社会正在为学历社会所困扰,群众日益对学历社会产生不满情绪。在日本流行一句谚语:“留给子孙的仅仅是学历”,“学历已经成为一种财富”。显然,传统的教育制度不培养人的真正才能,而给人的不过是一张于社会无用的学历。为了解决这一社会问题,广大群众纷纷要求改革现存的对学生的评价制度,使教育评价更公正、更科学、更客观、更合理。

当今,在各国教育评价工作蓬勃发展的形势推动下,国际上已组成了国际教育成就评价协会(简称 IEA)。这个组织自成立以来,已组织了几次国际性的有关教育评价的比较研究。其目的是对世界一些国家和地区的中、小学的现状做出科学的评价,找出足以说明国家间自然科学教育、数学教育等之所以有差异的原因,特别是课程设置所起的作用。

我国于 1984 年正式加入了国际教育成就评价协会。我们参加这个协会的研究活动,可以为我国的中、小学教育的改进提供必要的信息,也可以从国际的研究经验和成果中取得有益的借鉴。

第三节 现代西方教育评价理论的主要派别

1966 年以来,在美国教育评价理论领域涌现出了许多流派,

一时出现了繁荣景象。为了对这个时期教育评价理论有个概括的了解,我们对这些流派择其要者作些简要的介绍^① :

一、行为目标模式

行为目标模式的代表是泰勒的理论。行为目标模式与心理中行为学派有重要联系。它把教育方案、计划的目标用学生的特殊成就来表示,并把这一行为目标当作教育过程和教育评价的主要依据。根据这一模式,教育评价就是判断实际活动达到目标的程度。有人曾经用图1表示目标、教育过程 and 评价三者的关系。

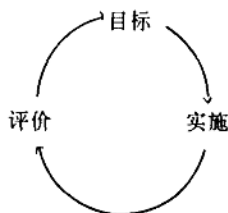


图1

预定的目标决定了教育活动,同时也规定了评价就是找出实际活动偏离目标的程度,从而通过信息的反馈,促使实际工作能够尽可能地逼近目标。行为目标是与系统分析法最为接近的理论,由于它结构紧凑,因而在过去很长时间内在教育评价理论中占有重要的地位。

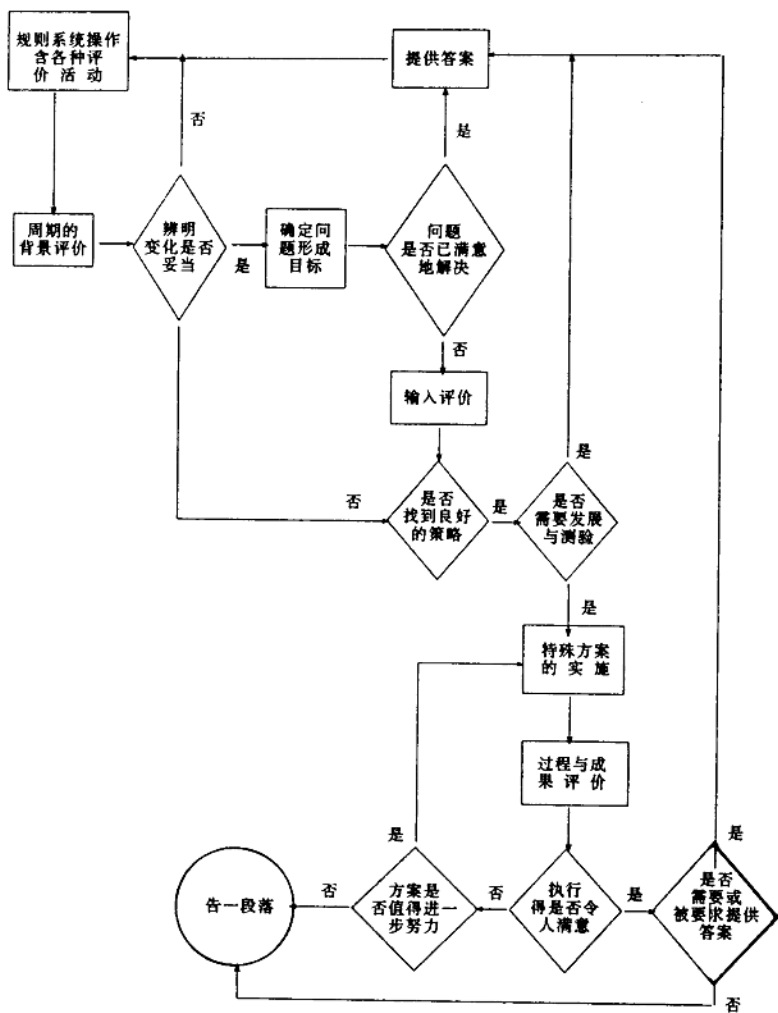
二、CIPP 模式

CIPP 模式也称决策类型模式。CIPP 模式评价事实上是由背景评价(Context)、输入评价(Input)、过程评价(Process)和成果评价(Product)四种评价组成。CIPP 评价的基本程序可用图2表示(见下页):

由图2我们可以看出,CIPP 模式已经突破了泰勒的框架,在许多地方有了新的发展,其中最主要的是目标的合理性与可行性受到了充分的重视,泰勒模式的缺陷在这里得到了克服。当然,在CIPP 模式中,目标仍然是一个重要的概念。因此,我们可以说,

^① 引自陈玉琨、赵永年:《教育评价大趋势与对我国评价制度中若干问题的探索》一文。

CIPP评价流程图



CIPP 并非是与泰勒完全对立的模式,它们之间有着若干共同之处,甚至人们可以把泰勒模式看作是 CIPP 模式的特例:在需要解决的问题与输入条件相对稳定的情况下(比如基础课的教学),CIPP 评价就转化为泰勒的行为目标。美国评价学者豪斯(E. R. House)就曾经指出,重视定量目标是它们共同的特征。

三、目的游离(Goal Free)模式

目的游离模式是由斯克里芬提出的。他认为,实际进行的教育活动除了收到预期效应外,还会产生各种“非预期效应”,或者叫“副效应”。这种非预期效应的影响有时是很大的。他进一步认为,根据预定教育目标进行评价,往往只注意目标规定的预期效果,非预期效果很难得到反映。因此,为了降低评价规定的预期效果,非预期效果很难得到反映。因此,为了降低评价活动中方案、计划制定者主观意图的影响,不能把方案制定者的预定的活动目的告诉评价者,以利于评价者收集全部的关于方案的成果信息。这种不受预定活动目的影响的评价是自由的评价。这一评价与泰勒模式、CIPP 模式最大的区别是在于:作出评价结构的依据,准绳不是方案制定者预定的目标,而是活动参与者的意图。这种模式已从评价活动主要反映管理者、决策者的意愿转向反映“群众”的意愿,因而它具有更大的“民主性”。

四、对手(adversary)模式

对手模式是为了揭示方案正反两方面长短得失,而采取准法律过程评委会审议形式的一种评价模式。对手模式十分重视听取教育方案和教育活动的争议意见尤其是反对的意见。这对各方面的情况能得到充分的反映提供保证。一般地说,这一评价的基本特点是它充分反映了各类人员“多元的”价值认识,是一种依靠人们直觉与经验的评价。

现代教育评价理论还有一些其他流派,本书不再加以介绍,但是,从上述几个主要派别的介绍中,我们不难看出,教育评价理论的发展前景是令人乐观的。