

实践中的多元智力理论丛书

丛书主编 霍力岩



多元智力评价的 理论与实践

霍力岩 赵清梅 著



教育科学出版社
Educational Science Publishing House



实践中的多元智力理论丛书

丛书主编 霍力岩



多元智力评价的 理论与实践

霍力岩 赵清梅 著

教育科学出版社
· 北 京 ·

策划编辑 王春华
责任编辑 王春华
版式设计 沈晓萌
责任校对 张 珍
责任印制 曲凤玲

图书在版编目 (CIP) 数据

多元智力评价的理论与实践 / 霍力岩, 赵清梅著.
——北京: 教育科学出版社, 2010. 4
(实践中的多元智力理论丛书 / 霍力岩主编)
ISBN 978 - 7 - 5041 - 4430 - 0

I. ①多… II. ①霍… ②赵… III. ①基础教育—教
育评估—研究 IV. ①G632. 0

中国版本图书馆 CIP 数据核字 (2009) 第 208415 号

出版发行	教育科学出版社	市场部电话	010 - 64989009
社 址	北京·朝阳区安慧北里安园甲 9 号	编辑部电话	010 - 64989584
邮 编	100101	网 址	http://www.esph.com.cn
传 真	010 - 64891796		
经 销	各地新华书店	版 次	2010 年 4 月第 1 版
制 作	北京金奥都图文制作中心	印 次	2010 年 4 月第 1 次印刷
刷 厂	北京人卫印刷厂	印 数	1—3 000 册
开 本	169 毫米×239 毫米 16 开	定 价	32.00 元
印 张	16.25		
字 数	257 千		

如有印装质量问题, 请到所购图书销售部门联系调换。

目 录

导论	(1)
一、从教育评价观之嬗变看多元智力评价	(2)
二、多元智力评价研究现状	(12)
第一章 多元智力评价理念研究	(16)
一、多元智力理论对传统智力测验的批评	(17)
1. 脱离儿童的生活和学习情境	(17)
2. 静态地评价儿童的发展	(17)
3. 片面地理解儿童的智力	(18)
4. 难以反映儿童的独特性	(18)
二、多元智力评价观	(19)
1. 强调情境	(19)
2. 强调发展	(20)
3. 强调多元	(21)
4. 强调独特性	(21)
三、多元智力评价体系	(22)
1. 评价目的：支持并促进儿童的独特发展	(22)
2. 评价主体：多主体全面参与评价	(24)
3. 评价内容：评价儿童多元智力的发展	(26)
4. 评价方式：多种方式展示儿童多元智力的发展	(28)
四、多元智力评价原则	(29)
1. 真实性原则	(29)
2. 情境性原则	(30)
3. 主体性原则	(30)

4. 广泛性原则	(31)
五、从与主要评价模式的比较中看多元智力评价	(32)
1. 行为目标评价模式与多元智力评价	(32)
2. CIPP 评价模式与多元智力评价	(38)
3. 外貌评价模式与多元智力评价	(42)
4. 差距评价模式与多元智力评价	(45)
5. 目标游离评价模式与多元智力评价	(47)
第二章 多元智力评价策略研究	(50)
一、多元智力评价策略概述	(50)
1. 个别化评价策略	(51)
2. 为促进理解而评的评价策略	(53)
3. 整合八种智力的评价策略	(55)
4. 合作式评价策略	(57)
5. 与学习风格相结合的评价策略	(59)
二、几种典型的多元智力评价策略研究	(63)
1. 成长档案袋：动态地记录儿童的成长历程	(63)
2. 观察评价：在儿童日常活动中收集信息	(73)
3. 记录评价：视觉化地呈现儿童的发展	(80)
4. 专题评价：评价儿童解决问题能力的有效策略	(86)
5. 访谈与讨论：在交流与互动中透视儿童的发展	(89)
6. 产品评价：评价儿童的成果	(91)
三、从与主要评价策略的比较中看多元智力评价策略	(93)
1. 自然式探究策略与多元智力评价策略	(93)
2. 三角测量策略与多元智力评价策略	(99)
3. 阐明式评价策略与多元智力评价策略	(102)

第三章 多元智力评价工具研究 (107)

一、全面评价多元智力发展的评价工具 (108)

1. 阿姆斯特朗智力评价量表 (108)
2. 多元智力发展评价量表 (112)
3. 观察记录卡与观察核查表 (115)
4. 多元智力报告卡 (118)
5. 开放性问题 (122)
6. 双向反馈 (123)

二、聚焦于各种智力发展的评价工具 (126)

1. 言语—语言智力评价工具 (128)
2. 数理—逻辑智力评价工具 (135)
3. 视觉—空间智力评价工具 (139)
4. 身体—动觉智力评价工具 (141)
5. 音乐—节奏智力评价工具 (144)
6. 人际交往智力评价工具 (149)
7. 自知自省智力评价工具 (152)
8. 自然观察智力评价工具 (155)

第四章 多元智力评价个案研究 (159)

一、光谱方案 (159)

1. 光谱方案评价特点分析 (162)
2. 光谱评价带给我们的启示 (165)

二、艺术推进方案 (168)

1. 艺术推进方案评价特点分析 (170)
2. 艺术推进方案带给我们的启示 (174)

三、新城市学校 (175)

1. 新城市学校评价特点分析 (176)
2. 新城市学校带给我们的启示 (180)

四、评价历史理解项目 (182)

1. 评价历史理解项目评价特点分析 (183)
2. 评价历史理解项目带给我们的启示 (187)

五、其他评价方案	(192)
1. 学校实用智力方案	(192)
2. 多元智力理论实验学校方案	(197)

第五章 多元智力课程开发与课程评价一体化

研究	(201)
一、课程开发与课程评价一体化的“诱因”	(202)
1. “课程—教学—评价”线性范式的崩溃	(202)
2. 以“共同建构”为特征的评价观的确立	(203)
3. 大教学论与大课程论的整合发展	(203)
二、多元智力课程开发与课程评价一体化之途径研究	(205)
1. 借助“通过多元智力来教”	(205)
2. 借助“为多元智力而教”	(205)
3. 与教育目标分类学结合	(206)
三、多元智力课程开发与课程评价一体化之个案研究	(206)
1. 借助“通过多元智力来教”实现一体化 ——以光谱方案为例	(206)
2. 借助“为多元智力而教”实现一体化 ——以重点学习社区为例	(213)
3. 与教育目标分类学结合实现一体化 ——以欧汉森小学为例	(221)

第六章 多元智力评价与我国基础教育评价改革

.....	(228)
一、对多元智力评价的几点反思	(228)
1. 多元智力评价的局限性：多元文化的不彻底性	(228)
2. 以档案袋评价为例再看多元智力评价的局限性	(232)
二、多元智力评价与我国基础教育评价改革	(235)
1. 多元智力理论呼唤新型教育评价	(236)

2. 多元智力评价对我国基础教育评价改革的启示	(238)
三、多元智力评价与我国学生发展评价改革	(240)
1. 从符号到场景：在真实的情境中评价学生的发展	(240)
2. 从静态到动态：评价学生的发展过程	(241)
3. 从单维到多维：多元化地评价学生的发展 ...	(242)
4. 从统一到多样：尊重学生的差异性	(243)
主要参考文献	(246)

导 论

教育评价一词在西方过去常常使用“evaluation”一词，现在更多使用“assessment”。从语义学的角度来分析，评价（evaluate）是由词干“valu”加上词头“e”和动词性词尾“ate”组成，其中“valu”意为价值，词头“e”的意义等同于“out”，即“引出”（霍力岩，2000a）¹。因此，教育评价就是引出和阐发教育的价值或者对教育进行价值判断，强调的是一种“自上而下”的评价模式。而“assessment”来自拉丁词根“assidere”，表示“随旁就座”（Weber，2003）⁷，是指作为评价者的教师“坐在考生旁边”（李雁冰，2001），强调的是多方参与和共同建构的平等评价模式。从评价手段来看，“evaluation”下的测验（test）被看作是唯—公平有效的评价手段，而这种测验更多时候是一种标准化测试，实行正式管理，且在设计上力求差异最小化以达到对不同被测验者评价的“公平化”——强调的是统一化的量化评价手段。“assessment”则嵌入了多样化的评价手段，如观察、档案袋评价、专题评价以及半结构化的考试等，一方面强调要沟通教育干预之间的关系，另一方面也更清晰地描绘了儿童思维、学习和工作的过程，因而强调的是多样化的定性与定量相结合的评价手段。从评价指向看，“evaluation”下的测验或标准化测试“关注的是不同个体之间的差距，回答的是‘这个儿童与参加考试的其他儿童相比，发展程度如何’的问题，强调的是儿童发展所达到的最终结果”，而“assessment”反映的是“现代认知理论对某一特定儿童学习与认知过程的兴趣”，关注的是儿童个体自身的发展，强调的是儿童的学习过程与结果如何实现效果的最大化（Bowman，

2001)²³³。

20 世纪 80 年代后期的这种概念演变，也标志着教育评价理念的转变。教育评价尤其是对学生发展的评价不再是仅仅基于一系列量化数字的纯量化评价，而是将定性评价也融入其中；不再仅仅是“一次定位”的终结性评价，而是将能把握学生发展过程的形成性评价也纳入其中；不再仅仅由教师或一些外部评价机构来掌控对学生发展的评价，而是将评价看作是一个师生在平等协商与全面参与基础上的共同建构过程；不再仅仅关注主流文化，而是将评价看作是多元文化的共同参与，从而使我们对学生的评价更加公平、真实与客观。

伴随着教育评价概念的演变，人们对教育评价的四大目的逐渐达成共识：①为教师进行教学与课程设计提供信息支持，使评价与课程开发成为不可割裂的过程，使教师能够把握学生发展的契机，并最终支持学生的学习和发展；②诊断那些需要特别帮助的儿童，如阅读困难的儿童，并为他们的发展提供有价值的建议和引导；③进行课程评价和督导，以进行及时的改进和修正，确保课程对儿童发展所起的积极作用；④提高学校的绩效。以此为基础，世界基础教育改革近年来一个重大的研究变化就是关注如何促进课程开发与课程评价一体化。可以说，课程开发与课程评价一体化的研究趋势体现了整个社会科学研究的发展走向：从哲学的角度看，后现代主义哲学重新审视定性研究的价值，帮助评价走出了“定量决定一切”的迷宫；从管理学的角度看，评价中浓厚的“管理主义”色彩被削弱，被评价者在评价过程中的主体作用得到进一步体现；从心理学的角度看，建构主义心理学的兴起，特别是近年来出现的整合皮亚杰、布鲁纳和维果茨基学说的研究趋势，倡导个人建构和社会文化建构相结合；从教育学的角度看，课程改革的进一步深入以及可持续发展概念的引入，使人们对儿童和课程的可持续发展有了更迫切的需要。多元智力理论正好迎合了这一发展走向，成为众多学校教育评价改革的指导思想之一，也成为促进课程开发与课程评价一体化的“推手”。

一、从教育评价观之嬗变看多元智力评价

1. 教育评价观之嬗变

(1) 教育评价即教育测量

这一评价观始于 19 世纪后半期，终于 20 世纪 30 年代（卢乐山，林崇

德,王德胜,1995)³⁴⁵,认为评价在本质上是以测验(testing)或测量(measurement)的方式来测定学生对知识的记忆状况或某项特质(李雁冰,2001)。

20世纪初,随着自然科学的发展,各种统计方法、测量技术也得到了相应发展,一批教育学家、心理学家开始把这些技术运用于教育测验领域,取得了一定成果。如美国心理学家桑代克(E. L. Thorndike)的《心理与社会测量导论》,系统介绍了统计方法及编制测验的基本原理,并提出了著名的论断:凡是存在的东西都有数量,凡是有数量的东西都可测量。法国的比奈-西蒙智力量表,采用智力年龄的方法计算成绩,解决了不同年龄被试相互比较的标准问题,使测验结果得到明确的解释。美国斯坦福量表首次引用了德国人施太伦(L. W. Stern)的智商概念,使心理测验达到了较为成熟的阶段。在心理测验的基础之上,教育测验也迅速发展起来,主要包括智力测验、学历测验和人格测验三种不同性质的测验。据统计,到1928年止,已有标准心理测验和学历测验3000多种。所有这一切都为教育测验的客观化、标准化创造了条件。与此同时,第一次世界大战后发展到顶点的工商业领域的“科学管理运动”,也对学校教育产生了深刻影响:学校被视为工厂,学生被视为原料和产品,教师成为加工者。学生产品是否符合需要、教师教学有何成效、学校教育是否成功,似乎都可以通过“测量”来检验,这显然为教育测验的发展提供了客观条件。教育测验也由此而得以兴盛。

从上述教育测验的发展历史中,我们可以发现教育测验是遵循着客观化与标准化的脉络而发展的,并且最终导向了纯量化阶段。自然科学的研究成果在这一阶段得到了充分的应用,以致将教育对象也等同于了纯自然的东西。因而,虽然当时教育测验在一定程度上适应了教育发展的需要,并因其在某种程度上克服了口试的弊病而得以推广,然而,将复杂的教育活动、教育对象简单化,并试图用纯客观与量化的手段来加以测量,本身就具有很大的局限性与缺陷:①教育测验就是对照统一的精确指标,严格测定学生的各种心理品质及其学业成绩,而教育评价工作的中心则是编制各种统一规划的测验量表,从而使测验尤其是智力测验成了当时衡量学生的万能标准;②教育测验主要用于甄别和选拔那些“适合于教育”的学生,完全掩盖了发展各异的学生的独特个性;③教育测验主要由教师或专门的评价人员以及校外的专门机构来组织和实施,并且,评价者在评价中扮演的角色也仅仅是“测量技术员”;④评价内容主要涉及学生的语言和逻辑数理智力,并且,

对学生这两个领域的测验也多局限于诸如对知识的记忆情况等一些较低级心理技能的测验；⑤主要运用测验和测量等纯量化的评价方式，以智力测验、学历测验和人格测验的形式开展，并且，主要以结果为导向，将评价看作是对学生的一次次测验，将测验的结果看作是衡量学生智力发展水平或学习成就的唯一标准。正因为如此，这个时期在评价历史上也被称为是“测量时代”。

（2）教育评价即描述教育目标的实现程度

随着教育测验的不断发展，人们日益认识到教育测验的一些不足，而美国 1929 年开始的大规模的经济危机凸显了学校培养目标与社会发展不相适应的问题。在此背景下，美国进步教育协会在七所大学和三十余所中学开展了声势浩大的教育改革实验，为期八年，即为教育评价史上著名的“八年研究”。美国俄亥俄州州立大学教授泰勒（Ralp W. Tyler）受邀主持了该项工作。泰勒认为，如果教育的目标是高级智慧技能，那么就on应该针对这些技能开展评价，而以往的课程和测验都以教科书为中心，只要求学生记诵教材的知识内容，不能说明人的高级智慧技能，十分片面。基于以前测验的弊病，泰勒提出了一套以教育目标为核心和依据的课程和测验的编制原则，并正式提出了教育评价的概念。泰勒认为，“教育评价是一个确定实际发生的行为变化的程度的过程。评价过程实质上是一个确定课程与教学计划实际达到教育目标的程度的过程”（朴圣玉，2008）。也就是说，实施教育评价，首先必须分析教育应达到的目标，再将这个目标作为评价教育效果的标准，并借助评价促使教育活动向预定的教育目标逼近。依据泰勒的观点，教育目标是课程评价的出发点和依据，是进行课程评价的决定因素。选择目标、表述目标的过程也是教育评价的一部分，只有清楚、明确地陈述了目标，才能说明课程与教学是否成功。泰勒把教育方案的目标用学生行为化的成就来表示，并把这一行为目标作为对教育进行评价的主要依据。预定的目标决定了教育活动，同时也规定了教育评价就是找出实际活动与教育目标的偏离，从而通过信息反馈，促进教育活动尽可能地逼近教育目标。因此，泰勒模式有时被直接称作教育评价的“行为目标模式”。泰勒模式注重目标的思想合乎理性的原则，因为有了教育目标的导向作用，教学与评价在实施起来都更有计划，更易于操作，也更易于被人们所接受。由此，泰勒关于教育评价的理论和方法受到广泛欢迎。在 20 世纪四五十年代，不少人着手发展泰勒的评价模式，并研究每种教育活动的教育目标，其中比较有代表性的研究是布卢

姆 (B. S. Bloom) 的教育目标分类学。

作为评价史上第一个较为完整、影响颇大的理论体系, 泰勒模式自问世以来便成为指导教育评价实践的主要理论, 在评价领域中雄踞领导地位近 30 年之久。泰勒模式的优点在于: ①合乎理性。泰勒模式可以被概括为教育目标、教育过程与评价三者的闭环结构。在这一结构中, 目标既是教育活动的依据又是教育评价的准则; 教育过程一方面贯彻落实教育目标, 另一方面也是评价的对象和内容; 评价即依据目标对教育过程进行评价, 找出教育活动与预定目标的偏差, 并最终促成目标的实现。可见, 该模式脉络清晰, 逻辑性强, 易于被人理解、接纳和采用。②目标的可测性。泰勒主张应将目标“细目化”“可操作化”, 在他看来, 只有明确具体的目标才能作为评价的根本标准。目标所陈述的结果明确具体, 便能协助教师确定教学内容, 选择适宜的教学策略, 并借助目标来评价和促进学生的行为进步, 最终把行为目标转化成学生的行为变化 (郑为川, 1993)。泰勒模式也存在如下一些不足: ①难以考察非预期效应。教育活动除了可以收到预期效应外, 还能产生非预期效应, 其影响有时甚至更大。泰勒模式是以预定的目标去评价实际达到的教育效应, 将非预期效排除在外。②目标自身的合理性无法考究。泰勒模式没有提供评价目标自身的方法, 也没有提供用来判断评价目标与评价结果之间差异的标准。③模式的刻板性。泰勒模式用预先设定的目标和明确的评价程序来评价教育过程, 将本应生机盎然的教育活动限定在严格的框架之中, 缺少了生机与活力, 抹杀了个性与特色。

(3) 教育评价即为教育方案的决策提供信息支持

泰勒模式在历史上占据了很长一段时期的统治地位之后, 开始遭到众多的批评和反对。在这一时期, 评价人员开始质疑仅以教育目标作为评价出发点和最终归宿的合理性, 也开始对评价中诸如“对已经确定的教育目标本身是否需要进行评价, 即是否需要价值判断?”“判断自身是否应该成为评价的一项基本活动?”“判断是否需要标准? 如果需要标准, 能否建立科学、客观而‘价值中立’的标准?”等一系列问题进行思考。如斯塔弗尔比姆 (Stufflebeam) 认为, “教育者需要一个较广义的评价定义, 而不只是局限于确定目标是否达到”, 评价应“有助于更好地执行和改进我们的方案”。作为替代泰勒评价的定义, 斯塔弗尔比姆将评价定义为“为决策提供有用信息的过程”。他强调“评价最重要的意图不是为了证明, 而是为了改进”, 这一看法不久就受到了教育界的广泛支持。1985 年, 他进一步提出: “评价

是一种划定、获取和提供叙述性和判断性信息的过程。这些信息涉及研究对象的目标、设计、实施和影响的价值及优缺点，以便指导如何决策，满足教学效能核定的需要，并增加对研究对象的了解。”斯塔克（Stake）主张评价应该包括描述和判断两个方面，而且不论是描述还是判断，都要搜集被评价对象如教育计划或方案的先在因素、过程因素和结果因素。他认为，要使评价结果能真正产生效用，评价者必须关心这一活动的所有参与者的需要，并通过信息反馈，使活动结果能满足各种人的需要。斯塔克强调评价的“多元现实性”，他认为以前的预定式评价对正在进行的方案缺乏敏感，难以察觉学生在与教师和其他同学接触中获得的收益，也难以反映对方案抱有的不同观点。克龙巴赫（Cronbach）把评价定义为是收集和利用信息以便作出关于某一教育方案的决策，认为评价并不只是调查某一教育过程是否有效，而是要确定其中需要改进的方面。根据克龙巴赫的观点，评价的重点应该放在教育过程之中，对教育决策进行必要的改进，而不只是关心教育过程结束之后目标达到的程度。克龙巴赫的观点在教育评价界产生了广泛影响。比贝（Beeby）认为评价是“系统地收集信息和解释证据的过程，在此基础上形成价值判断，目的在于行动”。他指出评价是在收集系统而非零散的资料，并将这些资料加以精心整理和解释，引入评判性的思考。在这里，比贝首次提示了教育评价的本质，即价值判断。他强调评价要对教育活动的价值作出判断，包括对教育目标本身作出判断，使评价活动有助于决策的科学化，指导实际工作。比贝关于“价值判断”的论断，深化了教育评价的内涵（王琰春，2005）。

与描述目标实现程度的泰勒模式相比，关注为教育决策提供服务的教育评价模式至少有如下几个方面的进步：①第一次将“价值判断”引入教育评价，并把它作为评价工作中的关键。评价不再只是根据预定目标对结果的描述，预定目标本身也需要进行价值判断，需要考虑社会的政治、经济、文化因素和学生发展的内部矛盾等并加以判断。既然目标并非评价固定不变的标准，那么评价就应当走出预定目标的限制，过程本身的价值也就成了评价的有机组成部分。②认为教育目标是动态发展着的，并且意识到教育方案在达到或未达到预期目标时有可能会取得非预期目标，而教育评价应该考察的是教育计划方案所达到的实际效果而不是预期效果。自此，教育评价开始关注实际教育活动的全部效果。③开始关注评价的“多元现实性”。评价研究者们日益认识到不同教育活动参与者具有不同的需要，教育评价要能为其提

供更多的信息,尽可能满足参与者的多种需要,并且评价者要能根据社会需要和个人需要,调整教育目标,缩小教育目标与社会现实之间的差距。④评价者的责任不仅是“测量技术员”和“描述者”,测量和描述学生达到预期目标的程度,而且还要帮助制定一定的判断标准与目标,对教育目标的合理性进行判断,对学生的个人需要进行判断,并且能根据需要调整教育目标,能够运用所获得的系统信息改进教育方案。⑤重视对过程的评价,强调完整、全面地评价教育方案的整个实施过程,强调在教育方案实施的过程中不断进行评价,从而不断发现和修正教育方案中的不当之处,这种过程评价思想与我们今天强调的形成性评价是一致的。

将教育评价看作是为教育决策提供信息支持,有助于实现教育活动的科学化,有助于达到教育评价的科学化。然而,这种评价观也存在如下一些问题:①没有涉及教育目标以及教育评价的多元文化性。其对教育目标进行的所谓的“价值判断”缺乏多元文化的根基,因而,这种价值判断会因为其对某些文化的关注或忽视而导致教育实现方式以及实现程度上的不公平。②整个评价活动仍然主要围绕教育目标来开展。在评价中关注的重点是对教育目标合理性的判断,是教育目标的动态发展性以及对整个教育方案的改进,而很少关注学生作为能动个体在教育目标设置以及整个教育评价活动中的主体地位。③较少考虑教育评价的方法论问题。

(4) 教育评价即基于全面参与的共同心理建构

20世纪60年代末70年代初,随着对课程改革运动的深刻反省,传统的评价方式也受到猛烈冲击,人们日益呼唤新型的评价理论和评价范式。1972年11月,在英国剑桥大学的丘吉尔学院(Churchill College),14位学者会聚一堂,讨论评价理论和范式的创新问题。会议认为,传统的评价范式源于教育研究中一直占优势的实验或心理测量传统,适用范围狭窄,不能解决评价中所遭遇的复杂问题。因此,会议建议采用新的、文化人类学的研究范式取代旧有范式,指出评价不是对预期的教育结果进行测量,而是要对整个方案,包括前提假设、理论推演、实施效果以及困难问题等进行全面而深入的研究。质性研究范式由此推广开来,并开创了一代新的评价理念,即古巴和林肯的所谓第四代评价。第四代评价为了打破以往评价中的“管理主义倾向”,采取了“回应模式”:以相关人的各种要求为评价的出发点,以“全面参与”和“民主协商”的方式形成共同的心理建构。在具体方法上,第四代评价更强调在自然环境中,用质性研究方法,使各方人士通过各种形

式的对话达成共识。在此特别要指出的是，第四代评价突出了过去作为被评价者的学生在评价过程中的参与者身份。在这种评价中，评价者不再是评价的控制者，而是一个中介、一个条件提供者、一个创造者。古巴和林肯为评价过程开出了这样一份“程序单”：第一步，在评价开始时，各有关方面订立协议，明确各方的权利与义务；第二步，做好深入现场、获得信息的安排；第三步，确定优先协商的问题；第四步，协商；第五步，形成报告。第四代评价体现了“共同构建”“全面参与”与“多元价值”的评价思想。“共同构建”即把评价看作是所有参与评价活动的人们共同建构的过程，认为评价并不是外在于人的纯客观过程，而是参与评价的所有人特别是评价者与其对象双方交互作用、共同建构统一观点的过程。评价结果也是所有参与者共同作出的，是其交互作用的“产物”。这样大大提高了评价对象在评价活动中的主体地位。“全面参与”主张所有参与评价的人，不管是评价者，还是评价对象，都是平等、合作的伙伴，是具体、完整的个人而非传统评价中的试验者与试验对象，他们都应有机会表达自己的观点，他们的“尊严、人格与隐私”也应得到充分的尊重。“全面参与”观点的最大贡献在于它把评价过程的控制特点与评价对象的伦理要求成功地协调起来，通过调动评价对象的主体性来弥补试验控制对对象伦理上的侵犯。“多元价值”指在评价中人们的价值标准各不相同，是多元的。

第四代评价突破了以往教育评价中的诸多局限性：①尊重“多元价值”，为来自不同文化背景、持有不同价值标准的人提供了更公平的参与乃至成功机会。②倡导“全面参与”与“民主协商”，突破了传统评价中的“管理主义倾向”，挑战了评价者与被评价者之间的传统关系，使传统上处于评价对象地位的被评价者也成为了评价的主体，得以公平、全面地参与整个评价过程，并且，在评价中享有真正的话语权，而传统上享有评价垄断权的评价者也不再是评价的控制者，成为评价过程中的条件提供者与创造者。③强调评价是一个共同建构的过程，认为评价是评价主体双方交互作用、共同建构统一观点的过程，充分强调对个体主体性的尊重，而对于达成共识的寻求，更反映着一种深刻的民主意识，即评价主体双方的平等与合作，同时，也突出体现了评价的过程性与动态性。④反对教育评价对于科学范式的过分依赖，认为评价中的价值问题是无法回避的，根本不可能做到纯客观与中立的评价；反对过分依赖量化评价而忽视质性评价，认为评价应该突破自然科学的范畴，而注入人文科学的内核；提倡在自然的背景下实施评价，使

评价成为自然的学习生活的一部分，而不是在严格控制的环境下实施的标准化测试，从而使教育评价更具有人性化，也更能体现评价对象的独特性与复杂性。

纵观教育评价观的发展历史，我们可以发现其体现了如下明显的发展趋势。第一，评价价值的多元化。从早期智力测验的统一标准，到以目标为统一评价尺度的泰勒模式，到以社会政治、经济、文化以及个体发展特点和需要为依据来评价教育目标合理性的第三代评价，再到明确提出价值多元化的第四代评价，教育评价的价值问题日益明朗化，也日益多元化。第二，评价主体的多维化。评价活动一度由评价者所掌控，被评价者处于被动地位，完全没有参与的权利，之后被评价者的主体地位日益凸显，评价者和被评价者日益成为评价活动的双方主体，通过民主协商、全面参与，在评价过程中共同建构统一的观点。第三，评价导向的过程化。早期智力测验以测验结果作为评价的唯一标准，泰勒模式也以教育目标为评价的出发点和最终归宿，它们都以结果为导向，而关注价值判断的第三代评价和基于全面参与的第四代评价则日益关注教育评价的过程性，日益将教育评价自身作为一个过程，也日益关注评价活动的每一个环节。第四，评价功能的多样化。从早先智力测验主要用以选拔适合教育的儿童，到泰勒时代主要用以考察教育目标的达到情况，再到考察教育目标合理性以改进教育、创造适合儿童的教育第三代评价等，评价的功能更加多样，也更具有人文色彩。第五，评价内容的全面化。早先的教育评价只评价学生的学习能力，后来发展到评价课程，再进一步发展到对教育活动的方方面面作全方位的评价，评价的内容更为宽广和全面。第六，评价方式综合化。从纯量化的智力测验到第四代评价，评价手段从各种客观的标准化测量，发展到提倡观察、交谈等描述性评价方法，再进一步发展到广泛收集信息，进行解析论证，作出价值判断，教育评价方式经历了从定量到定量与定性相结合的发展历程。

2. 从教育评价观之嬗变看多元智力评价

1983年，哈佛大学发展心理学家加德纳将智力定义为“是在某种社会或文化环境的价值标准下，个体用以解决自己遇到的真正难题或生产及创造出有效产品所需要的能力”（Gardner, 1993b）⁶⁰。1999年，他对智力进行了进一步界定，认为“智力是个体处理信息的生理和心理潜能，这种潜能可