

聚焦新课程系列丛书

编著 魏立明 战秀琴



G 高中英语

AOZHONG YINGYU JIAOYU PINGJIA

教育评价



东北师范大学出版社

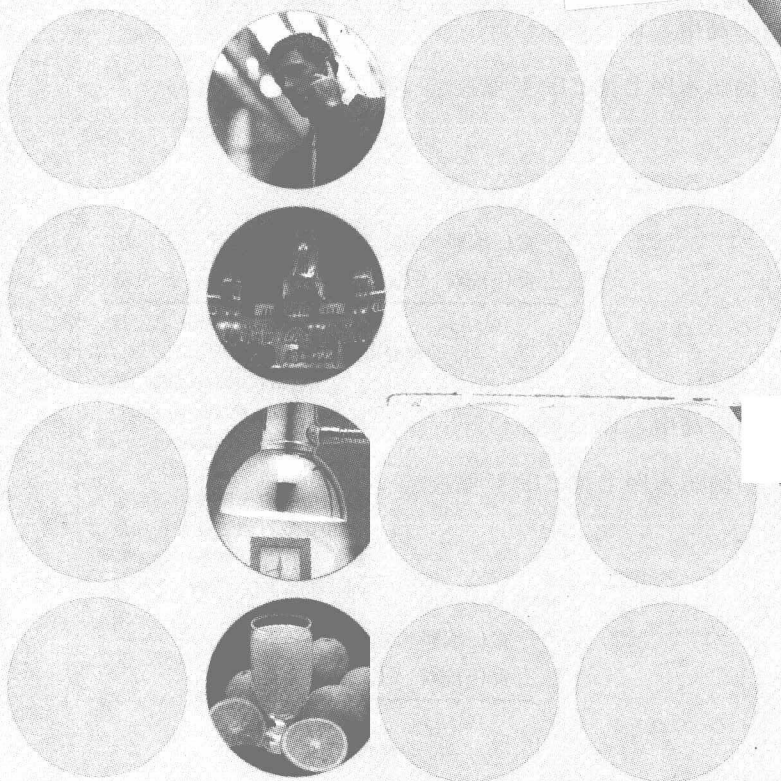
聚焦新课程系列丛书

编著 魏立明 战秀琴

G 高中英语

AOZHONG YINGYU JIAOYU PINGJIA

教育评价



东北师范大学出版社
长 春

图书在版编目 (CIP) 数据

高中英语教育评价 / 魏立明, 战秀琴编著. —长春:
东北师范大学出版社, 2005. 7

ISBN 7 - 5602 - 4271 - 5

I. 高... II. ①魏... ②战... III. 英语课—教育
评估—高中 IV. G634.412

中国版本图书馆 CIP 数据核字 (2005) 第 075108 号

☐责任编辑: 魏 昆 ☐封面设计: 宋 超

☐责任校对: 王 杰 ☐责任印制: 张允豪

东北师范大学出版社出版发行
长春市人民大街 5268 号 (130024)

电话: 0431—5687213

传真: 0431—5691969

网址: <http://www.nenup.com>

电子函件: sdcbs@mail.jl.cn

东北师范大学出版社激光照排中心制版

长春市海山印业有限责任公司印刷

2005 年 7 月第 1 版 2005 年 7 月第 1 次印刷

幅面尺寸: 148mm×210mm 印张: 7 字数: 169 千

印数: 0 001 — 5 000 册

定价: 9.00 元

前言

《高中英语教育评价》是依据英语课程改革评价理论而编撰的专题性评价研究。全书包括高中英语教育评价概述、高中英语学习评价、高中英语教材评价和高中英语教师评价四个专题。每个专题从文献综述入手，用界定、功能、原则、内容、方式和研究工具等基本概念进行理论穿线，用案例说明理论，用任务活动内化专题内的要点，用章后参考书目为读者提供文献来源和专业阅读的渠道。

高中英语教师是英语基础教育队伍中最活跃的因素之一。由于高中英语教材的备课任务相对繁重，高中学生的认知能力和需求相对较高，高中英语教师在所处的专业环境的竞争压力也就相对较大。高中英语教师亟须在教育理论和学科知识方面充电，这一点毋庸置疑。但是在诸多学科理论中选择切入点并非易事。教育评价在教育理论中占有举足轻重的地位，因为评价在整个教育活动中起导向作用。评价不仅“指挥”社会教育需求的导向，同时指挥教师专业发展的导向。从这个意义上来讲，掌握了教育评价的理论，如同掌握了专业发展的航向。因此，本书所提供的教育评价专题研究内容，不仅可以为高中英语教师搭建专业理论发展的平台，也可以成为大学、初中和小学英语教师，各级教育行政部门的教研员，教学管理者和教材编写人员继续教育的教材，还可以作为英语教师备课和撰写论文的案头参考。

参加本书各章的撰写者是东北师范大学外国语学院首届本硕连读的研究生。这些学生将成为教育部直属师范大学向国家重点高中首批输送的“新生代”英语教师。与传统的初出茅庐的新手教师不同的是：他们不仅经历了本科的教育、教学实习，还在研

究生学习期间加固了专业实习环节，所实习的学校均为国内教育质量卓著的重点高中；他们不仅广泛地“接触”了英语学科的教学理论，还系统地学习了教育学的基本理论。《高中英语教育评价》一书的撰写，源自两位硕士研究生导师的“创意”，即本书的两位主编，东北师范大学外国语学院魏立明教授和北华大学外国语学院的战秀琴教授。在如何培养从校门到校门的教育专业硕士研究生的问题上，她们形成了一个共识：在“火线上”培养研究生的研究能力。通过专题研究项目，使教育专业硕士研究生的教学与科研相结合，锻炼其动笔和动手能力，使他们的研究成果尽早转化为社会教育的精神财富，为他们今后的科研生涯打下基础。这是两位导师的共同心愿。

如今书稿完成了，但从本科到研究生的过渡，仅仅只有一年的时间。这些学子从未尝试过如此大面积的笔头耕耘，写作历程颇艰。编者案头疾书，主编字斟句酌，几易其稿。待交付稿件时，师生均已感到身体透支，但收获的喜悦自然溢于言表。在欣慰之余，我们仍不敢掉以轻心：因为项目和出版档期的紧迫，错误和不当之处在所难免，因此还有望读者的赐教。

《高中英语教育评价》一书的出版得益于下列学术同仁和人士的关心与帮助：东北师范大学研究生院十分重视各学科教育硕士专业研究生的培养；外语学院张绍杰院长和李增副院长对英语教育硕士的科研水平寄予关切，并从培养体制方面为研究生提供了发展空间。

本书编者在写作中参阅了国内外教育评价领域中诸多学术同仁的观点和理论成果。作为文献引证，各章作者尽可能标明了资料出处。但凡所引证的作者，对本书均作出了重要贡献，在此一并致以诚挚的感谢。

编 者

2005年6月

目 录

第一章 高中英语教育评价概述	1
第一节 教育评价的概念.....	1
第二节 教育评价的历史沿革及其 发展.....	9
第三节 教育评价的类型与功能	18
第四节 英语教育评价的原则	25
第五节 高中英语教育评价的内容 与方法	28
第二章 高中英语学习评价	49
第一节 高中英语学习评价的内容 与方式	49
第二节 高中英语学习评价的价值 取向	63
第三节 测试性与非测试性评价	77
第四节 个案分析	87
第三章 高中英语教材评价	112
第一节 与教材评价相关的概念.....	112
第二节 教材评价理论评介.....	121
第三节 英语语言教学大纲与教材.....	128
第四节 高中英语新教材评介.....	133
第四章 高中英语教师评价	165
第一节 高中教师评价概况.....	165
第二节 高中英语教师评价内容.....	169

第三节	高中英语教师评价的方式 与手段·····	172
第四节	高中英语教师发展评介·····	194
第五节	高中英语教师反思·····	199
第六节	个案分析·····	205

第一章

高中英语教育评价概述

第一节 教育评价的概念

一、评价的含义

评价在生活中随处可见。例如，评价一个物品的等级，或评价一个人的社会地位。词典中给评价所下的定义为“评论货物的价格，今泛指衡量人物或事物的价值”。评价一词在英文中为 evaluate。该词由一个前缀 e-加上 value（价值）这个词。由此可见，评价一词与价值有着密切的关系。前缀 e-起到一个引出的作用。因此，评价也可以理解为“引出和阐发价值”。从本质上说，评价是对价值判断的过程。

二、教育评价的概念

教育评价是现代教育的一个重要课题。在进行教育评价之前，首先要弄清楚什么是教育评价。国内外学者就如何定义教育评价提出了许多有见地的观点和主张，但是迄今为止，尚未形成对于教育评价的科学的、统一的定义。为了深入理解教育评价，首先要了解有关教育评价的几个代表人物及观点。

1. 泰勒 (Tyler, R. W.) 的评价观

拉尔夫·泰勒是美国教育心理学家和课程论专家。由于在教育评价方面作出的贡献,他被尊称为“教育评价之父”。泰勒首次提出了教育评价的概念,并创建了第一个教育评价模式——行为目标模式。这一模式在美国统治评价领域达三十年之久,至今还有深远影响。

泰勒的评价观着眼于效果,强调通过评价判断教育目标或教育计划的实现程度,并提出了一套以教育目标为核心依据的课程和测验编制的原则,借以把社会的要求、学生的需要反映到课程与测验中。泰勒认为,评价过程在本质上是确定课程和教学大纲实现教育目标的程度的过程。同时,评价过程也是确定行为发生实际变化的程度的过程^①。为了把这一思想与早期的测量区别开来,泰勒正式提出了教育评价的概念。在20世纪40年代初,他将教育评价解释为“确定教育目标在实际上被理解到何种程度的过程”。

2. 克龙巴赫 (Cronbach, L. J.) 的评价观

自20世纪40年代以来,泰勒的评价理论一直占据着主导地位。由于泰勒的观点是找出结果与目标的差距。因此,人们更多地把泰勒的评价看成一种对结果的终结性评价。

而克龙巴赫认为,形成性评价的作用远比终结性评价重要得多。他指出,所谓教育评价,“是作为一个搜集和报告对课程研制有指导意义的信息的过程”^②。1963年,克龙巴赫在《通过评价改革课程》(Evaluation for Course Improvement)一文中提出了“评价的中心不应仅是目标,而更应当是决策”的思想。克龙巴赫这一着眼于信息,强调通过评价收集资料,为教育决策服

① 瞿葆奎. 教育学文集·教育评价. 北京:人民教育出版社,1989.263

② 瞿葆奎. 教育学文集·教育评价. 北京:人民教育出版社,1989.82

务的观点一经提出,就受到不少西方教育评价专家的重视,在教育评价界产生了广泛的影响。

3. 斯塔弗尔比姆(Stufflebeam, D. L.)的评价观

斯塔弗尔比姆同样看重评价的形成性作用。他强调:“评价最重要的意图不是为了证明,而是为了改进。”^①为此,在1969年,他对评价作出这样的定义:“为决策提供有用信息的过程。”^②

4. 豪斯(House, E. R.)和斯克里芬(Scriven, M.)的评价观

豪斯和斯克里芬主张把评价定义为一种对优缺点或价值的评价,或者把评价定义为一种既有描述又有判断的活动。这种评价观点认为,教育评价的着眼点在于教育现象的价值。也就是说,评价应当着重判断教育活动、教育过程和教育结果所产生的效益,看其是否具有价值。据此,可以认为教育评价就是对教育现象进行价值判断。

5. 美国教育评价标准委员会的评价观

该委员会于1981年对教育评价进行了综合性界定。这一组织认为:“教育评价是对教育目标以及它的优缺点进行价值判断和系统调查,为教育决策提供依据的过程。”^③这一评价的界定,在很大范围内被当作教育评价权威性的定义。

6. 国内学者的评价观

20世纪80年代以来,教育评价在中国得到了迅速的发展。结合我国国情以及教育的发展现状,国内学者对教育评价的定义也进行了阐释。有的学者认为:“教育评价是对教育的社会价值作出判断的过程。”^④也有人指出:“教育评价是按照一定社会的

① 瞿葆奎. 教育学文集·教育评价. 北京:人民教育出版社,1989.298

② 瞿葆奎. 教育学文集·教育评价. 北京:人民教育出版社,1989.301

③ 瞿葆奎. 教育学文集·教育评价. 北京:人民教育出版社,1989.345

④ 陈玉琨著. 教育评估的理论与技术. 广州:广东高等教育出版社,1987.2

教育价值、教育方针和政策所确定的教育目标，并对所实施的各种教育活动的效果、完成教育任务的情况以及学生学习成绩和发展水平进行科学判断的过程。”^① 这些阐释对我国教育评价的研究和具体实施起到了推动作用。

7. 广义的教育评价

广义的教育评价是用系统的方法搜集和分析信息资料，在此基础上对教育的社会价值作出判断，对其改善和发展给予指导的过程。广义的教育评价涉及教育活动的一切方面和教育系统的各个环节。其评价的对象已由学科的教育、教学过程和结果扩展到更广阔的领域，如对学生、教师、学校管理者的评价等，又如对教育方针、教育制度、教育结构、课程设置、师资队伍、教育发展规划、教育科研的学术水平等方面的评价。教育评价内涵的日益丰富和概念的不断外延，说明教育评价的产生和发展适应了当代教育建设和发展的需要。

8. 狭义的教育评价

狭义的教育评价是根据教育目标对学生通过教育产生的行为变化作出价值判断，为改善和优化教育提供依据的过程。其基本特点是，以分数为标准，以学生为对象，任务仅仅限于学习和生活领域内，对反映学生成长和发展状况的教育行为给予价值上的判断，以促进教学过程的变革。

无论是广义上的教育评价，还是狭义上的教育评价，我们都能够看出教育评价参照了之前对于评价一词词义的解释。从直观上讲，教育评价就是对教育价值进行判断的过程。但是这种判断的过程是建立在对教育现象进行分析的基础之上的，其范围是教育领域中的一切教育现象或者活动，目的在于提高教育质量。

尽管各国学者之间对于教育评价的概念存在着一些分歧，但

^① 张玉田等编著，学校教育评价，北京：中央民族大学出版社，1998.19

是其中也有共同点。首先,教育评价本质上是一种价值判断活动。教育评价并非单一性的活动,而是一个活动过程。这个过程是系统的、综合性的,并由多项组成,如确定目标、获取资料、分析资料、形成判断、决策指导等。此外,教育评价的活动过程是以一定的教育价值观为依据和准则的,而不是盲目地、无依据地进行评价。而且,进行教育评价,须要采取一切可行的科学手段来进行。

根据上述共同点,可以初步把教育评价界定为“根据一定的教育价值观或教育目标,通过系统的、科学的方式搜集信息,分析解释,对教育现象进行价值判断,为优化教育,并为教育决策提供依据的过程”。

三、与教育评价相近的概念

1. 教育评估

教育评估与教育评价是非常相近的两个概念,在许多情况下通用。事实上,评价和评估,从认识论的角度分析,其本质都是精神对物质,意识对存在的一种反映,是价值意识朝向客体对象的精神活动。在教育领域中,二者均指从一定社会的角度出发,按照一定的教育价值观来考查和评定评价对象的社会价值,判断其好与坏,功与过,对与错及其表现的程度。

但是,这两种概念亦有差别。有的学者将“评估”定义为评价与估量,即评估是一种模糊定量的评价。在教育界,尤其在西方,仍倾向于把客观的方法(比如测验的方法)对学生个体的评价称为评定(assessing),而把对学生个体以外其他复杂性程度更高级的评价称为评估。评价相对而言是较精确的价值判断。评估多用于对群体或单位的状态及效果的估价,而评价则既有对群体或单位的价值判断,又有对个人的判断。

当然,在国内外文献中,评价、评估、评定等词语模糊不

清,甚至混用的情况俯拾即是。术语混淆的原因之一可能是对英语中 evaluation 的译法不同。如上所说,评价与评估不同,评估在英语中常用 appraise,用于对教育机构和教育方案等的评价。由于评估的对象涉及的因素较多,复杂程度也较高,因此,对它们的评价按严格、精确的方法进行还有困难。目前,主要是依靠定性 with 定量结合,客观统计资料与主观描述资料并重的手段。“评估”这一术语表达了这一事实。

2. 教育评定 (Educational Assessment)

这一术语含有评比鉴定和评级等的意思,在美国被专门用于对人员的评定。它包括评级(正式的或非正式的)、考试和考核等在内的各种活动^①。但在实际上,它常常与评价混用。由于美国国会通过 1965 年中小学教育法案时给教育评价以合法性,所以在美国逐渐倾向用“评价”一词。但是评定与评价混用的现象仍比较普遍。

评价和评定都包含旨在用以描述某些属性的数量的测量活动。两者也都包含获得这些测量的程序。这些测量可能包括测验,也可能包括诸如等级量表(rating scales)这类不十分客观的量表。人们倾向于用比较概括的方式来使用评价,包括以更大程度接受主观判断的更广泛的测量。此外,在评价的对象不是人(或人的群体),而是教学历程或教学方法时,评价用得更多。因此,评定通常更多地用于其程序涉及较客观的量表,以及量表用来测量个人的属性时的情境。

3. 教育测量 (Educational Measurement)

教育测量是按照一定的规则用数字来报道教育领域内事物的属性,对教育现象加以描述和确定的过程。因此可以说,教育测量的对象是教育现象的属性。教育测量的本质是对教育现象获得

① 瞿葆奎. 教育学文集·教育评价. 北京: 人民教育出版社, 1989. 743

数量化的认识。概括地说,其本质是事实判断。而教育评价是价值判断过程。一般来说,事实判断是回答事物是什么的问题,而价值判断则是回答事物意义的问题。事实判断是价值判断的基础。价值判断若离开事实判断这个基础,评价是达不到预期目的的。

教育测量强调客观性。因为测量强调的是数量化,所以要尽可能地减少判断的作用。测量者要尽可能排除主观因素的影响。不同的测量者对同一事物的测量,除允许的误差外,其结果应当是相同的。而教育评价则要求质与量并重,主观与客观的方法兼用。

教育测量为教育评价提供量化资料。教育评价必须依靠教育测量,但教育测量并不等于教育评价。教育测量是依据一定的标准对教育某一方面作出事实判断,并进行赋值的过程。测量的数量化结果,如果不依据测量的目的进行分析、解释和评价,就是无意义的东西,所以测量包含有一定的评价。评价则是突出价值观,充分重视对问题的分析与评断。它要根据测量的结果,评断是否达到了教育的目标和要求。测量的结果是评价的重要依据,而不是评价的唯一依据。评价须要依据更多的方法和更多的资料来进行。总之,测量是评价的基础,若没有测量,就没有可进行评价的素材。评价是对测量结果的进一步判定。若只有测量,而不进行评价,则无法确定该教育事实属性的价值如何,是优还是劣。测量是一种单一事实的量的获得行为,而评价则是较为复杂的一系列活动过程;测量的目标在于证明量的多少,而评价的目的在于经由对事实的了解,进而求得对缺点的改进。

4. 教育测验

在教育过程中,我们有很多测验,如单元测验、期中测验和期末测验。一个人从小学读到初中,从中学读到大學,从每一级学校毕业进入更高一级的学校,都必须经过毕业测验和升学测

验。尽管人们对测验都有亲身经历和深切体会,但很少有人考虑它的确切含义。根据测量学的理论,所谓测验,就是测量一个行为样本的一种程序。测验所测量到的是行为,是被测者对测验题目所作的反应。测验是一种程序,测验的编制、施测和评分等都是在实施活动中确定的规则,例如:试题是根据测验的目标,经过系统的分析和选择而组合起来的;在实施测验中,对所有的被测者在相同条件下施测相同的试题;阅卷评分时,也要根据事先确定的原则,保证评分者之间取得一致意见。

从对测验含义的分析可知,测验是狭义的间接的测量。从某种意义上说,它还是测量的一种工具,通过分析测量和评价的不同,就可以明显地看出评价与测验之间的本质区别。

5. 教育评比

教育评比是我国教育界普遍使用的一种经验型管理手段。评比活动较为盛行,名目也繁多。所谓评比,就是通过横向比较,评定出高低或优劣。这种评比一般都停留在直接经验的描述上,主观随意性大,缺乏可靠的客观性、科学性。因此,它远不是科学的教育评价。必须看到:教育评价的根本目的,并不只是求得横向“比”出孰高孰低,而是要从总体上判断学校教育工作的实际水平,找出其与目标的差距,以调控计划,改进工作。

从两者的目的上看,评比活动的主要目的是评比优秀者,设立先进典型,以此来推进工作。然而教育评价的主要目的,是通过评定价值来改进工作,从而促进教育目标的实现。如果评价的目的局限在评选优秀者,这在一定意义来说削弱或失掉了教育评价的本质意义。

从客观性上看,评比活动的条件一般注重政策性。评比的条件往往随着某个时期的主要任务而变动,所以具有不稳定性。而评价标准的确定一般要求比较规范,往往须要设计出评价指标体系,稳定性也比较强。此外,评价要求在一定的理论指导下进

行,因此科学性和规范性都比较强。评比和评价之间有一定联系。在实际操作过程中,可以把两者结合起来进行。结合的前提是评比之前应认真地开展评价活动,以便判断优劣,为评优创造有利条件。

第二节 教育评价的历史沿革 及其发展

一、教育评价的起源

(一) 思想和制度的起源

教育评价的思想和制度最早起源于中国。早在西周时代就有以射选士的做法,并建立了国学和乡学。公元606年,隋炀帝时代开始了封建科举制度。所谓科举,就是分科取士,即设科考试,根据学科考试的成绩去录用官吏。此科举制度一直延续到20世纪的清朝末期,即光绪三十一年废除科举。科举制度的特点是:制度齐全,方法完备,等级分明,标准严格。科举制度形成了整套完备的体系。从考生条件、考试科目、考试程序到阅卷方法、录取规程等,都有一套标准与措施。科举制度也是集竞争考试和考核测评于一体的制度。它以封建主义的价值观为准绳,把政治审查、学力测评与健康要求结合在一起,进行综合测评。在封建社会中,这套严格的科举制度在封建统治阶级笼络和控制知识分子方面起了重要的作用。虽然由于种种原因它没能得到发展,但作为一种最初萌芽,它在教育评价史上占有重要位置。

中国的“科举制度”曾对欧亚其他国家产生了一定影响。据研究,英国学校废除口试代以笔试,以及英国的“文官制度”等形式,都是受到了科举制度的影响。中国的科举制度虽然在检测、评价考生的学力中起过重要作用,但它无法对人进行全面、

公正、科学的考查与评价，其自身又存在着许多弊端。19 世纪末 20 世纪初，我国“废科举”，“兴学堂”，但是在学校教育工作中仍沿袭传统的考试制度。到了 20 世纪 20 年代后期，虽有人民教育家陶行知先生倡导过教育评价，但实践的范围有限，研究也难深入。20 世纪 70 年代末，尽管我国的社会制度和教育性质早已发生了质的变革，西方的教育评价理念也引进到国内，但是多元教育评价的理论与方法只停留在口头上，传统的考试方法仍在评价中占主导地位。

（二）理论的起源：教育测验

教育评价理论源于西方对学生学力的检测。这场运动的核心是测验的客观化及标准化问题。从 19 世纪后半期到 20 世纪 30 年代，这一阶段又称测验运动阶段。学生个体学力测验的客观化、标准化问题是这一阶段的中心问题。测量运动阶段大体包括以下四个时期：

1. 萌芽期（19 世纪上半叶—20 世纪初）

19 世纪的工业革命给西方社会带来了经济和技术的迅速发展，社会结构也随之发生了改变，心理保健、社会生活、社会道德观念以及社会机构均发生了很大的变化，同时出现了边沁（Bentham，英国伦理学家）的自由竞争哲学。

1864 年，为了力求考试客观化，英国格林威治医学校校长菲舍试图为客观评分提供一个参考标准，于是公布了与习字、拼字、数学、圣经以及其他科目的质量水平相对应的标准对照表，并为每种标准列举了实例，作出说明，采用五分制评分的标准，即评分标准对照表。菲舍的工作事实上是用科学方法研究教育质量问题的最初尝试。虽然他的试验在当时未引起多大的反响，但其成果成为 50 年后桑戴克（Thorndike，E. L.）等人开展教育测定的重要依据。

1879 年，德国心理学家冯特（Wundt，W.）在莱比锡设立