

外研社·基础外语教学与研究丛书
——英语教师发展系列之二

真实性评价

——从理论到实践

王笃勤 著

**Authentic
Evaluation:
From Theory
to Practice**

外语教学与研究出版社

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS

新思维，掌握新思维，掌握新思维
——最新思维与理论成果

真实性评价

——从理论到实践

王明华 著

Authenticity
Evaluation
From Theory
to Practice

2014年1月第1版

ISBN 7-309-05411-1

外研社·基础外语教学与研究丛书
——英语教师发展系列之二

真实性评价

——从理论到实践

王笃勤 著

**Authentic
Evaluation:
From Theory
to Practice**

外语教学与研究出版社
FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS
北京 BEIJING

图书在版编目(CIP)数据

真实性评价: 从理论到实践 / 王笃勤著. — 北京: 外语教学与研究出版社, 2007.7

(基础外语教学与研究丛书. 英语教师发展系列; 2)

ISBN 978-7-5600-6832-9

I. 真… II. 王… III. 英语课—教学研究—中小学 IV. G633.412

中国版本图书馆 CIP 数据核字 (2007) 第 107482 号

出 版 人: 于春迟

责任编辑: 严雪芳

出版发行: 外语教学与研究出版社

社 址: 北京市西三环北路 19 号 (100089)

网 址: <http://www.fltrp.com>

印 刷: 北京京科印刷有限公司

开 本: 880×1230 1/32

印 张: 10.75

版 次: 2007 年 7 月第 1 版 2007 年 7 月第 1 次印刷

书 号: ISBN 978-7-5600-6832-9

定 价: 17.80 元

* * *

如有印刷、装订质量问题出版社负责调换

制售盗版必究 举报查实奖励

版权保护办公室举报电话: (010)88817519

前言

随着语言学习理论、教育测量理论、信息技术的发展以及社会需求的变化,语言学习评价也经历着变革。评价在经历了描述、判断之后进入了意义建构时期,引发了评价领域各个层面的研究,从评价理论研究到评价实践研究,从评价模式和范式研究到评价技术研究,出现了多元评价、档案袋评价、表现性评价和真实性评价等评价理念,其中最能代表评价发展方向并与外语学习理论、外语教育目标以及评价本质最为一致的就是真实性评价。面对新的时代需求,国家《英语课程标准》(基础教育阶段适用)和《大学英语课程教学要求》提出了注重形成性评价和课程考核中形成性评价与终结性评价相结合的评价理念。这些理念与真实性评价是什么关系?真实性评价如何实施?如何帮助实践外语教育的发展目标,培养学生的英语综合应用能力?又如何评定学生运用英语完成生活和工作中各种任务的能力?本书将根据笔者就读博士期间对评价的研究,在“大学英语真实性评价研究”的基础上,对各教育阶段尤其是基础教育阶段的真实性评价进行初步的探讨。

本书正文共由六章构成,分别从理论和实践两方面探讨真实性评价的内涵,真实性评价与其他评价模式之间的关系,以及在听力、阅读和语言表达能力方面如何实施真实性评价。

第一章为评价概述,论证了评价的内涵以及评价与学习和教学之间的关系。本章首先论证了评价作为一种价值判断过程的本质,论证了评价、评估和测试之间的关系,认为测试与评估作为评价数据采集的手段和过程,为价值判断提供数据支撑。本章还分析了真实性评价与形成性评价、总结性评价、档案袋评价以及表现性评价之间的关系,论述了评价在教学与学习中的诊断、促进以及导向作用,尤其是过程性评价和能力总结性评价在教学和学习中的作用。

第二章为真实性评价，从评价标准、评价内容、评价主体等方面论述了真实性评价的内涵。在对真实性评价内涵论证的基础上，本章介绍了真实性评价理念下的课堂教学，主要介绍了任务教学、项目教学和支架教学模式，并且从语言观、学习观和认知观角度论述了真实性评价与学习的关系，介绍了项目学习、探究式学习和自助自主学习模式在真实性评价理念下的应用。真实性是在真实的语言环境中通过学生完成现实生活或工作中的真实任务或准真实任务的表现对学生的语言理解能力、表达能力和综合素质进行价值判断的过程。真实性评价由于尊重学生的多元差异，顺应语言学习的基本规律，以评价价值主体（学生和社会等）的需求为评价标准，可以有效促进学生的发展，促进教学的发展，促进课程的发展。

第三章为真实性评价与课程考核，主要介绍了真实性评价与课程考核的关系。在分析网络评价模式、多元评价模式和发展性评价模式的基础上论证了真实性评价模式，同时介绍了真实性评价理念下的基础英语课程考核模式、职业英语课程考核模式和大学英语课程考核模式。测试作为课程考核模式中重要的评价手段对课程考核的实施起着关键性作用，本章论证了标准化考试的不足，介绍了真实性评价理念下表现性测试的基本操作。

第四章为阅读理解真实性评价，从过程性评价和能力评价两方面介绍了阅读理解的真实性评价。过程性评价部分介绍了阅读理解评价目标的确立和阅读评价活动的设计，展示了阅读过程中诊断性评价和阅读理解能力培养活动的设计，并且通过大量的活动设计实例分析阅读理解过程性评价的实施。在能力评价部分介绍了阅读评价标准的确立和评价活动的设计，同时举例说明各种阅读理解中能力的评价方式。

第五章为听力理解真实性评价，同样是从过程性评价和能力评价两方面介绍听力真实性评价的具体操作。本章在分析真实听力内涵的基础上介绍了在基础英语教育、职业英语教育和大学英语教育阶段如何确立评价的目标，如何设计过程性评价活动，并且通过大量的真实活动展示了听力过程性评价的操作。与阅读理解中的能力评价一样，听力能力评价同样从评价标准的确立、评

价活动的设计两个方面论述了听力能力的真实性评价，读者可以从本章详细的例示中理解各种听力能力评价是如何贯彻真实性评价的理念的。

第六章作为本书的最后一章为语言能力（主要是语言表达能力）真实性评价，介绍真实性评价的操作。本章论证了语言能力的内涵，通过大量的评价活动介绍了词汇和语法的过程性评价和应用能力评价。在语法方面以语法条目为纲介绍各种时态、语态、从句、情态动词、非谓动词和虚拟语气的真实性评价方式。相信本章对语言能力真实性评价的介绍会给语言测试一定的启示。

需要说明的是，本书所述主要是本人在读博期间结合自己的研究项目所作的初步研究，由于主客观因素所限只就阅读、听力和语言表达能力的真实性评价作了初步的探讨，对于口语和写作能力的评价，仅依托于语言表达能力作了简单的介绍，对情感态度、文化意识等方面的真实性评价，则未设专门的章节研究，不妥之处敬请各位同行不吝赐教。另外，书中引用了《英语》（新标准）教材（陈琳等，外语教学与研究出版社）、职业《英语》教材（中等职业学校英语教材编写组，山东省职业教育教材编写组，高等教育出版社）中的大量资料，也引用了《大学英语综合教程》（李荫华等，上海外语教育出版社）中的部分材料，作为活动设计之用，在此对这些教材的编者和出版机构一并表示感谢。

目 录

第一章 评价概述	1
第一节 评价的内涵	1
一、什么是评价?	1
二、评价的发展	4
三、评价的种类	5
四、评价的实施	15
第二节 评价、测试与教学	18
一、课程评价、教学评价与学习评价	18
二、评价、教学与学习	20
三、过程性评价与能力评价	24
 第二章 真实性评价	35
第一节 真实性评价的内涵	35
一、真实性评价与评估及测试	35
二、真实性评价的真实内涵	38
三、真实性评价与学习风格	42
四、真实性评价与多元智能	46
五、真实性评价的原则	51
第二节 真实性评价与课堂教学	53
一、课堂教学评价范式	53
二、真实性评价范式下的课堂教学	56
三、课堂教学的真实性评价	69
第三节 真实性评价与语言学习	78
一、语言的真实	78
二、语言学习的真实	82
三、真实性评价的学习模式	88

第三章 真实性评价与课程考核	110
第一节 真实性评价与课程考核模式	110
一、评价模式的发展	110
二、真实性评价理念下的课程考核	126
三、真实性评价在课程考核中的角色	136
第二节 真实性评价理念下的测试	138
一、标准化测试与真实性评价	138
二、表现性测试与真实性评价	142
 第四章 阅读理解真实性评价	153
第一节 阅读理解概述	153
一、什么是阅读	153
二、什么是理解	154
三、影响阅读理解的因素	156
第二节 阅读理解过程性评价	157
一、阅读评价标准的确定	157
二、阅读评价活动的设计	163
三、阅读过程性评价示例	176
第三节 阅读理解能力评价	184
一、评价标准的确立	185
二、评价方式的选择	190
三、评价活动的设计	195
四、阅读理解能力评价活动例析	196
附录4-1 普通中学英语阅读能力目标	213

第五章 听力理解真实性评价	215
第一节 听力理解概述	215
一、听力的内涵	215
二、听力理解真实性评价的内涵	220
第二节 听力理解过程性评价	224
一、根据听力目标确定评价标准	224
二、听力过程性评价活动的设计	230
三、听力过程性评价活动举例	246
第三节 听力能力评价	253
一、评价标准的确定	253
二、评价方式的选择	259
三、评价活动设计的原则	264
四、听力能力评价活动举例	265
附录5-1 英语听力分级标准	272
第六章 语言能力真实性评价	275
第一节 语言能力过程性评价	275
一、语言能力的内涵	275
二、语言能力评价的过程	276
三、语言能力过程性评价活动示例	292
第二节 语言能力评价	304
一、评价标准的确立	305
二、评价方式的选择	308
三、语言能力评价活动示例	310
附录6-1	323
附录6-2	324
参考文献	326

第一章 评价概述

评价作为一种认识活动渗透于人类生活的各个方面。人们时时都在对自然、社会、教育、他人以及自己周围的一切进行评价。在生活、教学和学习中,评价无处不在,例如:“他获得了很高的评价”,“应试教育真是害人”,“你的英语真棒”,“他这个学生很勤奋”,以及“四、六级考试效度很低”等等。那么,到底什么是评价?评价与测试又有什么不同?外语教学的学习评价应如何开展?本章将就评价的内涵作初步的探讨。

第一节 评价的内涵

一、什么是评价?

(一) 评价:一个价值判断过程

评价存在于生活的各项活动之中,是人们认识活动的一部分,尤其是价值认识活动不可缺少的一部分。但是,人们对评价概念的认识却不尽一致。

有人认为评价是确定课程是否达到预定目标的方式(Tuckman, 1979: 1),是“通过不同的渠道收集学生的相关资料,将这些收集到的资料与预先设置的标准进行比较,并基于这些信息作出决策或判断的过程”(Montgomery, 2004: 8),是“对相关信息的系统收集与分析借以促进课程的发展,评定课程的效度以及参与者的态度”(Brown, 1989: 223)。但是,更多的人认为评价是一种价值判断(Stufflebeam, 1974: 73; 冯平, 1995; 鲁子问, 王笃勤: 2005: 1),也就是说,评价是一种特殊的认识活动,是揭示世界(个人、社会、自然)的价值,建构价值世界的认识活动(李雁冰, 2002: 30),评价的目的在于揭示主体与客体之间的价值关系(冯平, 1995: 33),亦即价值客体满足主体需要的程度。

那么什么是价值呢?有的认为价值是情感的表达,是欲望的涵项(培里, 1962: 324; 罗素, 1982: 123);有的认为价值是事物本身的一种属性,是客体满足主体需要、满足人和社会需要的属性(杜齐才, 1987: 9),

评价所要表达的就是这种属性；有的认为事物本身没有价值，价值取决于主体的需要，评价不过是主体需要的表达。但是现代认识论认为，价值是主体和客体之间的一种特殊关系，是客体属性对主体需要的满足，是客体属性满足主体需要的一种现实效应（陈新汉，1995：81；冯平，1995：31）。价值不依赖于评价而存在，评价所做的是揭示价值关系，而不是创造价值。

评价有评价主体和评价客体之分，由于评价中价值关系的存在，因此又包含价值主体和价值客体。评价的主体即实施评价的人，评价的客体即被评价的对象，但却不是指人，而是指价值主体和客体之间的关系。评价中的价值不是以实体的形式存在，而是以价值主体的需要的形式存在，即价值主体实际是学习者的需求、社会对学生的需求，而评价所揭示的是价值主体的需求与价值客体的属性及功能之间的关系（冯平，1995：33）。

本书采用认识论中的评价理念研究英语学习评价，研究评价中如何把握价值主体的需要，即学生的需要，也就是学生要成为社会所需要的英语人才对英语的需求是什么，对英语学习的需求是什么；把握价值客体的属性与功能，即语言的属性与功能，语言学习的属性与功能，评价的属性与功能；用价值主体的需要衡量价值客体的属性和功能，判断价值客体是否能够满足价值主体的需要，也就是说语言学习、语言教学、测试评价等能否满足学生的需求，能否满足社会的需求。本书将致力于研究学习如何满足价值主体的各种需要，如何评价需求的满足程度，如何评价社会所需要的各种能力。

评价活动总是从主体出发，以满足主体需要与否来判断客体对主体的意义。这表明，评价活动总是与一定的人、一定的团体、一定的社会利益相联系。外语学习以满足学生个人发展以及社会发展对英语人才的需求为目标，个人发展以及社会发展的需求也就成为外语学习中价值判断的标准。评价的这一特点要求现代教育评价必须贯彻真实性原则，实施真实性评价。

（二）评价、测试与评估

“这个人英语口语很好”，“他的英语太差了，连说明书都看不懂”，“他

的英语很好，都过六级了”。当对一个人的外语水平进行评价时我们是如何作出这些价值判断的呢？我们是通过亲耳听其说话、与其交流等来判断其口语如何，还是根据其口语考试成绩？当然，对其是否过六级的判断显然根据的是测试成绩，但是有关“连说明书都看不懂”的评价却是根据现实生活中任务完成的情况来评价的。可见，评价所依据的信息很多，有时借助观察，有时借助测试，有时借助任务。由此来看，评价与测量不同，与测试也不同。

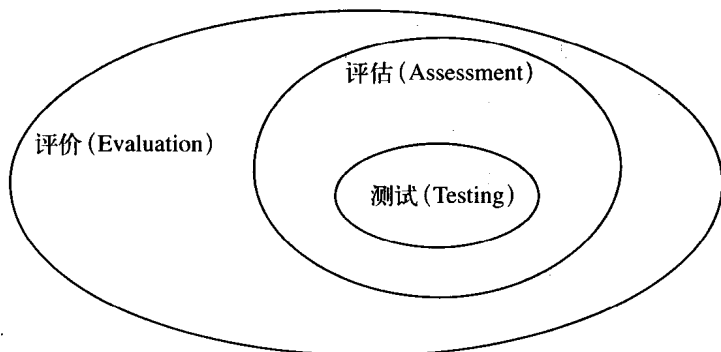
测试 (tests and examinations) 是我们最熟悉的测量工具，是以让学生完成规定的问题来测试其知识与技能的一种方式。测试可以为评价和评估提供数据。

评估 (assessment)，亦称评定，指评价信息的收集过程 (Hart, 2004: 1-2)。语言学习中，评估指对学生个体语言的学习过程、结果、能力等各方面的数据收集，包括需求分析、水平的测量、课程目标的达成信息以及任务操作与完成信息 (Lynch, 2003: 1)。Hart (2004) 认为，评估是收集学生相关信息的过程，所关注的是“我们如何发现学生正在学习什么”，而评价是对评估信息进行解释和判断的过程，所关注的是学生是否正在学习他们被期待学习的东西。可以看出，评估指通过多种渠道收集有关学生学习过程以及语言能力、态度、参与、认知发展等方面的信息，借以促进学习和教学，而评价是在评估基础上分析评估数据对教学和学习的总体的价值作用。

对于评价与评估两个概念的内涵，不同的学者看法不一，尽管有人主张评价与评估可以交替使用 (Montgomery, 2004: 8)，但是大部分学者认为，评价与评估所指不同，评估过程是评价过程的一部分，是评价的基础，也就是说，评价是依据评估数据所作的价值判断。但同时学习中的评价又发生在评估的过程中。当人们根据已有信息发现存在的问题时，就会采用需求分析的方法，探求并实施解决问题的方案，从而又导致进一步的评估与评价。评价与评估是一个交互进行的过程。

评价与评估的不同还表现在评价所依据的不全是评估的信息，还包括测试结果，而且评估的信息有时并不用于评价。测试信息可以用于评估和评价，而有时测试只为甄别和筛选，不存在价值判断。

图1-1 评价与评估关系图



二、评价的发展

从19世纪末至今，评价已经经历了四个阶段，进入了建构评价时期。19世纪末至20世纪30年代，没有真正意义上的评价，评价等同于测量或测验，即通过测验或测量的方式测定学生对知识的记忆状况或某项特质。20世纪30年代后随着泰勒八年的“课程与评价”研究，评价进入了“描述时期”，主要是根据预定教育目标对教育结果进行客观描述，描述教育结果与教育目标的一致性程度，所依据的除学生的考试成绩以外，还包括问卷、观察以及产品样本等数据信息。评价取代了考试和测验，被视为一个过程，而不再只是一两次测验。

20世纪50年代，评价进入“判断时期”，随着人们对泰勒行为主义、操作主义评价观的质疑，各种评价模式如雨后春笋般出现，人们开始关注对已经确定的目标的评价，以及是否需要已知目标作价值判断。评价被视为根据一组加权（weighted）的目标标准收集和综合有关现象资料，以形成比较的或价值的判断，并说明资料收集方式、加权数和目标选择的合理程度（李雁冰，2002：52）。本阶段的评价既有描述又有判断，过程成为评价的有机组成部分，出现了“形成性评价”、“目标游离评价”以及“内在评价”等一些新的评价理念，评价的价值判断本质得以确认，确认了评价的过程性特征。

20 世纪 60 年代末，评价进入“建构时期”，随着对课程改革运动的深刻反省，人们意识到前面的三种评价存在着不同程度的管理主义倾向，忽视了价值的多元性，过分依赖科学范式，因此开始寻找新的评价方式，第四种评价开始出现。传统的评价是预定式的评价，先陈述目标，然后依据目标收集数据，最后判断目标的达成程度。在这种评价体制中，评价者始终处于管理者的位置，决定着评价的范围与任务，决定着评价的对象，学生在评价中始终处于被控制的地位，其主体性没有得到应有的尊重。随着建构主义学习理论和价值论的发展，学生在评价中的主体地位得到越来越多的认可，人们开始意识到评价的建构特征，评价中评价者与被评价者通过协商共同完成心理建构。学生在评价中不仅是评价的客体，同时处于主体的位置。评价的基本方法不是量化研究而是质性研究。

不同的评价理念反映了人们不同的价值取向。描述时期评价过程是将教育结果与预定教育目标相对照的过程，根据预定目标对教育结果进行客观的描述，体现的是行为主义的学习理念和技术理性的哲学思想；判断时期确认了评价的价值判断本质，体现了“实践理性”的哲学思想；与描述和判断时期的评价不同，建构时期的评价坚持评价价值多元的信念（李雁冰，2002：56），提倡整合了自主和责任的解放理性，把评价视为评价者与被评价者协商进行的共同心理建构过程，学生成为评价的参与者和评价的主体，评价从主体上采用质性研究方法。正是在这一时期产生了表现性评价、真实性评价、档案袋评价和替代性评价等评价理念。

三、评价的种类

随着评价的发展衍生出很多不同的评价概念，主要有形成性评价、总结性评价（常作终结性评价）、过程性评价、发展性评价、多元评价、替代性评价、直接评价、表现性评价、真实性评价、诊断性评价以及成长记录袋评价等等，这些评价之间既有差异，又存在众多的相关，从不同的侧面揭示评价的特征。由于很多评价都是从一个侧面论证评价应该遵循的原则和可采用的评价方式，这些评价理念有很多相通之处，这里就形成性评价、总结性评价、表现性评价和真实性评价作简单的介绍。

(一) 形成性评价

1. 什么是形成性评价

形成性评价与终结性评价相对提出，一般概念下的终结性评价就是测试，而形成性评价则指在教学和学习过程中对信息的收集、整合和解释的过程，借以促进学习和教学（Bryant et al, 2002），指在教学过程中为了获得有关教学的反馈信息，改进教学，使学生对所学知识达到掌握程度所进行的系统性评价，即为了促进学生掌握尚未掌握的内容进行的评价（田颖，2003：4）。形成性评价用于提供有关学生目前的理解和技能发展情况以便制定下一步的学习和教学方案（Harlen, 1996：130）。

形成性评价具有过程性评价、表现性评价、真实性评价等众多评价的特征。研究发现，形成性评价有利于促进学生语言技能的提高，有利于学生综合素质的提高。形成性评价具有诊断促进功能、反馈激励功能、反思总结功能和记录成长功能。

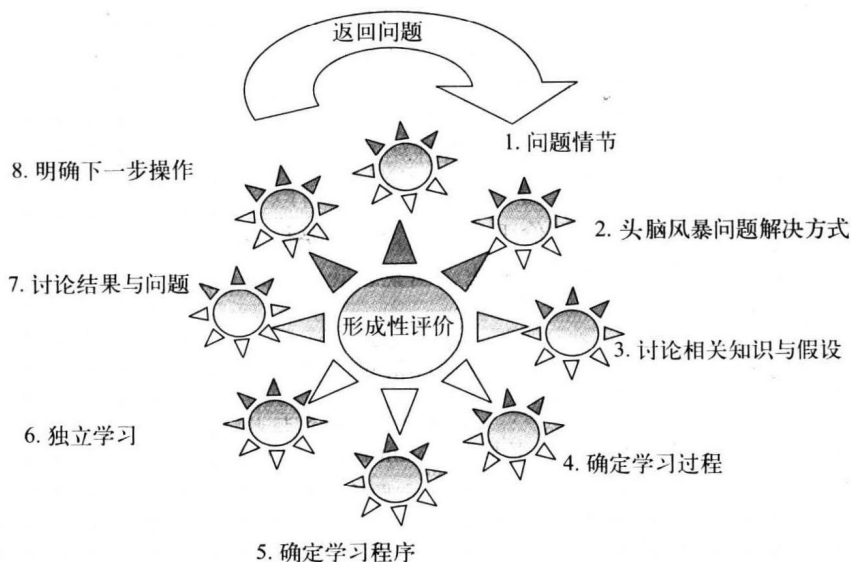
在所有的评价方式之中，形成性评价方面的研究最为普遍，多数都聚焦于如何使形成性评价更加有效。多数专家认为，形成性评价作为教学和学习的有机组成部分是一种信息收集、整合和解释的过程，目的在于在学习发生的同时促进学习，评价以发展为目的，而不是以划分等级为目的（Bryant et al, 2002：5）。形成性评价可以揭示学生的进步情况，为下一步的学习和教学提供有益的参考。因此，根据形成性评价的理念，在课堂教学的过程中，教师应该经常与学生开展对话和讨论，以了解学生的学习状况，获取有益的反馈，借以促进学生的学习和课堂教学。

研究发现，形成性评价能够更好地测量学生的实际操作能力，更加尊重学生的主体性和创造性（孙晓敏，张厚粲，2005：646），不仅能够用于评价高水平认知和技能，而且还可以用来评价非认知的结果，诸如自我定位，同他人一起工作的能力以及社会意识。

2. 形成性评价的操作

根据 Biggs & Watkins（2001），课堂教学中的形成性评价可以通过教师与学生的对话帮助学生界定问题、寻找解决问题的方式、实施学习过程以及评价学习结果（参见图 1-2）。

图1-2 形成性评价的操作程序



(Bryant et al, 2002: 4)

3. 形成性评价的常用工具

形成性评价的常用工具包括观察、座谈、问卷、测试和成长记录袋（亦称档案袋）等。

观察是指对学生学习行为表现的观察，常指课堂观察。观察前教师要根据课堂教学目标确定观察的内容，选择观察的对象与观察的方式，拟定观察量表，以保证获取有效的信息，促进课堂教学的开展和学生的学习行为。

座谈包括与学生的个别交谈和组织学生开展学生会议。通过座谈可以了解学生的感受、观点和态度，获取通过其他渠道难以获取的信息资源。座谈同样需要教师提前选择座谈的中心议题，选择适当的记录方式，选择倾听的对话姿态。

问卷可以用于调查学生的知识、观点、态度以及意识，也可以调查学生的行为。但是，调查问卷的主观性比较强，信度容易受到影响，因此常常影响调查的质量。