

英语教学理论与实践丛书

英语教学评估

行为表现评估和学生学习档案

龚亚夫 罗少茜 编著

*The Alternatives in English Assessment
Performance Assessment and Portfolio*

人民教育出版社

英语教学理论与实践丛书

英语教学评估

——行为表现评估和学生学习档案

The Alternatives in English Assessment

*Performance Assessment and
Portfolio*

龚亚夫 罗少茜

人民教育出版社

图书在版编目 (CIP) 数据

英语教学评估: 行为表现评估和学生学习档案/龚亚夫, 罗少茜编著。—北京: 人民教育出版社, 2002

ISBN 7-107-16114-8

I. 英…

II. ①龚… ②罗…

III. 英语课—教学评估—中学

IV. G633.412

中国版本图书馆数据 CIP 核字 (2002) 第 086127 号

人民教育出版社出版发行

(北京沙滩后街 55 号 邮编: 100009)

网址: <http://www.pep.com.cn>

人民教育出版社印刷厂印装 全国新华书店经销

2002 年 12 月第 1 版 2002 年 12 月第 1 次印刷

开本: 890 毫米 × 1 240 毫米 1/32 印张: 4.25

字数: 100 千字 印数: 00 001 ~ 15 000 册

定价: 8.10 元



说 明

近两年,评价的改革越来越多地受到教师们的关注。尤其是新的国家课程标准提倡任务型教学和评价的改革,许多教师希望了解和运用新的、切实可行的评价方法。这使我们想到应该向教师们系统地介绍评价的一些基本理念和实际操作方法,以及国内外已有的研究成果,以便教师了解评价并参与评价的改革。

有些地区的教师已经行动起来,并且积累了很多经验。为此,我们编写了两本小册子:《任务型语言教学》与《英语教学评估——行为表现评估和学生学习档案》。同时,我们去浙江、江苏和深圳等地听课并采访教师、学生和家長,还拍摄了一些任务型教学和评价的课堂教学录象。

在听课和采访过程中,我们惊喜地发现,教师、学生和家長的反应与任务型教学和评价改革的理念是完全吻合的。他们的实践表明,运用新的评价方法,增加新的评价内容,如行为表现评价、学生学习档案、形成性评价等在我国中小学是完全可行的。

在此,我们要特别感谢浙江省教育厅英语教研室夏谷鸣老师、浙江省临海回浦中学任美琴老师、杭州十四中学邱峰老师、江苏第十中学胡明老师、苏州太仓高级中学嵇永宁老师、深圳南山区教研室禹明主任、南山外国语学校李红燕老师、南山实验小学徐阳杰老师等。他们不仅提供了大量的第一手资料,也增强了我们编好这两本小册子的信心。

由于时间和水平所限,书中难免有不尽如人意之处。希望教师和专家学者不吝指教。

另外,录象中的示范课也决非尽善尽美,其中有许多值得反思和研讨的问题。我们想,这些录象是供教师参考、借鉴并对自己的教学进行反思的,有助于促进我们的英语教学和评价。



序

学习的基本原则和基础教育的发展目标是要使学生成为:

1. 积极涉猎知识者: 学生能在校内和校外获取知识, 内容涉及全球热点问题和个人感兴趣的问题, 这些内容包括本地的信息以及世界各地的信息。

2. 探险者: 学生能充满自信, 不怕失败, 积极探索新的领域。他们愿意接触新观点, 采用新策略, 敢于创新, 并大胆表述自己的观点。

3. 有道德和爱心的人: 学生将具备较高的道德修养, 为人正直、讲道理、诚实、公正。他们以敏感、理解、同情、尊敬和宽容的心态来面对不同观点和文化。这些优秀品质使他们做事时充满责任感。

4. 有思想的人: 学生在做研究时运用他们的创意、组织能力和想象力, 创造性地运用组织技巧, 做出自己的决定, 解决问题, 有效地化解矛盾。他们能不断地反思自己的思维方式和学习方法。

5. 团队建造者: 学生有能力承担不同的角色, 成为思维开阔、对团体有贡献的具有合作精神的成员。

6. 成功者: 学生通过个人努力, 在各方面表现良好, 发展优良品质, 追求学业更高标准。他们能成为独立自主、不断攀登学术高峰的优秀学生并进行终生学习。

7. 个人和社会心态平衡的人: 学生在处理问题时, 能展现他们平衡的身心发展和良好的适应社会的能力。

8. 善于沟通的人: 在不同的场合, 学生能用多种语言, 准确、有效、自信地展示他们的沟通能力。

9. 懂得现代科技的人: 学生在学习时能准确、有效地运用高科技手段。



显然，目前大多数学校的考试或测验方式无法达到这样的教育目的。要使评价起到促进学生发展的目的，需要补充评价的内容，改进评价工具，增加评价方法。

提到评价，许多教师自然地想到考试和测验。实际上，评价并不简单地等同于考试。首先，评价的功能或者说评价的主要目的不在于甄别，不在于发现一个学校、一个班级学生的成绩有多少是优、多少及格或平均分数是多少。评价的主要作用在于促进学生发展，鼓励学生找到适合自己的学习方法和策略；在于让学生发现自己的长处与弱点；在于培养学生自我调节能力；在于调动学生学习的积极性；在于养成学生合作的精神，发展与人合作的技巧。

评价不仅应该有终结性的评价，同时也应有形成性的评价；不仅应该有测试型的评价，也应该有非测试型的评价。评价的内容不仅有对语言知识的检测，也应有对语言行为表现的评价，还应该有对学生的策略、情感态度的评价；不仅有教师对学生的评价，也应有学生之间的相互评价。

评价的方法和工具多种多样，限于篇幅，本书不全面讨论评价系统的改革，只重点介绍两种目前评价活动中的热点问题——行为表现评价与学生学习档案。第一章说明为什么要进行英语教学评价改革以及什么是评价；第二章主要说明采用行为表现评价的原因与方法；第三章介绍什么是学生学习档案以及如何制作和评价学生学习档案。

目 录

说明	
序	
第一章	以学校为基础的评估和评价 (1)
一、我们为什么评估 (1)
二、评估的主要类别 (2)
三、评估的理念与基本原则 (5)
四、课堂评估的步骤 (6)
五、评估的重要因素 (7)
六、学生参与的评估 (9)
七、评估需要考虑的其他因素 (10)
第二章	行为表现评估 (12)
一、基本理念 (12)
二、行为表现评估的特点 (21)
三、学生行为表现的评分方法 (24)
四、设计行为表现评估 (31)
五、实施行为表现评估中的难点 (35)
六、实行与计划行为表现评估 (37)
第三章	学生学习档案 (41)
一、理念 (41)
二、目的 (42)
三、学生学习档案介绍 (43)
四、学生学习档案的收集 (46)

五、学生学习档案的使用	(49)
六、长期使用学生学习档案	(50)
七、学生学习档案回顾	(51)
八、开始使用学生学习档案时需要注意的问题	(56)
评估术语表	(62)
参考文献	(87)
附录	(93)



第一章 以学校为基础的评估和评价

评估 (Assessment) 和评价 (Evaluation)

教学评价在广义上指的是分析评估数据,进而判断教学方案是否有效的过程。在使用评估这个概念时,往往涉及评估工具、技巧与过程。但人们常常交替使用这两个术语。依据评估数据,可以判断学生在学习过程中的进展。收集到足够的信息后,就开始评价评估数据。如:

1. 评价学生的学习进展可作为在教师、学生和家長间开展合作学习的一种方式。

2. 评价评估活动是否有效。在分析某评估活动的结果时,教师应考虑到评估本身的有效性,也就是说,评价活动和方式是否真实、全面地反映了教师想要了解的情况。教师不应只是关注学生在评估中的表现。

3. 评价学习活动的有效性,如评估结果是否说明学习活动有效。

4. 评价教学方案的有效性,如此教学方案应做哪些修改。

一、我们为什么评估

● 促进学习

1. 教师依据评估,了解学生对知识的掌握程度以及学生完成学习任务的经验,以便调整教学方案,适应学生个人和整体的需求。



2. 教师依据评估, 引导学生进步, 了解学习是否按计划进行并及时发现问题。

3. 教师依据评估, 给学生提供反馈, 进而鼓励学生, 帮助学生增强信心、提高学习效率。

◆ 报告进步

1. 向家长提供关于学生进步和成绩的信息。

2. 向学生提供前一学习阶段的进步情况。也可为同一学校不同教师提供学生的学习情况。

◆ 提供综合性信息

1. 提供学生学习的整体表现的概况——知识、技能、学习策略和态度。这可以成为学生毕业时终结性成绩记录的基础。

2. 帮助学生了解自己是否具备完成课程的条件, 以便对未来的学习选择做出决定。

◆ 改进教学方案

为评价和调整学校的教学方案提供信息, 帮助教师确定哪个教学方案和学习方案有效、学生如何进步以及哪些方面有待改善。

◆ 反馈

就学生学习活动的准确性和有效性给教师提供反馈, 帮助教师确定如何改进教学方法。评估也反映了学生的进步是教师和学生应共同履行的义务。

二、评估的主要类别

评估有多种分类方法, 我们可以从以下四个方面来确定评



估的类别:

(一) 评估的过程

根据评估过程,可以把评估分为起始性(initial)、形成性(formative)和终结性(summative)评价。如:

1. 起点评价(Sizing-up Assessment)

起点评价一般在学期或学年刚开始时进行,目的是为了了解学生当时学业的状况,以便确定教学计划或分班。

2. 形成性评估(Formative Assessment)

作为教学和学习过程不可分割的一部分,此种评估向学生提供反馈,以便学生改进学习方法、巩固学习成效。还可以帮助教师了解学生的学习情况,反映学生进步的整体情况,确定下一步的教学计划及学生的学习计划。该种评估可通过多种方式进行,如:学生学习档案、座谈、访谈、自我评价和同伴评价等。

3. 终结性评估(Summative Assessment)

经过一段时间的学习后,为了解学生的成绩而展开的评估。与形成性评估相比,此种评估通常采用更为固定的形式,如包括听、说、读、写各方面技能的期末测试和毕业考试。

(二) 评估的功能

评估有各种目的和功能。有的评估目的在于了解学生掌握知识的能力,发现他们的长项与弱项,这类评估就是诊断性的。有的评估是为了选拔学生,如高考和中考,这样的评估就是选拔性的。比如诊断性评估(Diagnostic Assessment),旨在帮助教师了解学生对知识的掌握情况和能力,发现学生存在的问题及其性质、范围,进而设计学习活动来满足学生的需



要。诊断性评估可通过多种方式进行,如:精心设计的测验、课堂教学中向学生简单提问以便了解学生的情况。

(三) 评估的内容

评估的内容可以是多方面的,有的以课程标准为评估的内容,有的以能力为标准。

1. 以课程标准为基础的评估 (Curriculum-Based Assessment)

以结果或标准为基础的评估是根据规定的标准来测量学生的成绩。此种评估不假设能达到标准的学生人数。该种评估应依据课程标准对语言知识和语言能力的描述。

2. 以学业为基础的评估 (Achievement-Based Assessment)

此种评估通常是以与年级相关的标准来报告学生的成绩,可分为优、良、及格等。此种评估针对学生的知识和技能掌握情况提供整体评价。应让学生在考试前知道评估的标准、考试的范围及目标。

3. 以能力为基础的评估 (Competence-Based Assessment)

此种评估规定学生应达到的语言行为表现 (performance) 标准,进而确定学生是否具备实际运用语言的能力。像其他以标准为基础的评估一样,评估的标准应明确具体、切实可行,教师和学生评估前应充分理解这些标准。

(四) 评估参照的标准

评估可以是为了确定学生在一个群体中所处的位置,这是常模参照的评估。如果评估只确定学生是否达到某个目标或标准,而不做横向比较,那就是标准参照评估 (Criterion-referenced)。而以常模为参照的评估 (Norm-Reference Assessment) 则是依据学生的表现和其他组里学生表现的比较来



评估学生成绩(学生在组中的位置)。小组可以是一个班、学校内或一个学区所有学校相同年级的学生。此类评估通常根据标准化试题进行比较。常模由学生的平均表现来确定。

评估还有许多其他的分类方法。上述几种方法也并非相互排斥,而是相互交叉的。比如,形成性评估和终结性评估都可以是以课程标准为基础的。有时一谈到形成性评价,人们就想到非测试的评价,如学生学习档案等。实际上,形成性评价也可以采用测试型评价等多种评估工具。

三、评估的理念与基本原则

1. 评价的主要目的不是为了让学生了解自己在班级、年级中的位置,而是为了学生的发展。评价不是为了甄别,而是为了学生的成功。评估旨在帮助学生学习和提高教学质量。

2. 评价是多元的。不应把学生看做是被动的“评价对象”。学生也应是评价的主体。学生将直接参与评估,以便理解、思考、评价自己的学习活动。学生的参与对于评价的成功与否关系极大。评价是民主参与、协商和师生交流的过程。

3. 既要评价学生学习的结果,也要关注学生的学习过程和思维过程。师生都应能通过评价活动,进行自我反思。学生可以通过评价调整自己的学习策略,教师应通过评价活动调整自己的教学方法。

4. 评价应该有持续性的形成性评价。在使用各种评价手段时,教师应关注诊断性评价的信息。

5. 评价应能反映出学生运用英语的实际能力,而不仅仅是知识的掌握。学生应能通过完成语言学习任务,表明自己掌握的知识结构和具备的能力。



6. 评价应公平、公正、公开。应使师生清楚地了解评价的标准和目标。采用评估的全体学生应了解评估的目的、过程和意义。

7. 评价不仅是测量与预测,更多的应是描述(description)与解释(interpretation)。评价人员不仅应包括语言专家,也应有心理学专家与社会学专家。评估内容应该包括情感、策略、认知等多方面。

四、课堂评估的步骤

◆ 教师的非正式评估 (Teacher Informal Assessment)

教师随时对学生的成绩进行评估,如在课堂讨论中,在学生写作中,以及在课堂教学的各个环节中。每当教师提问或学生分组活动时,学生的回答和在活动中的表现都可以作为评价学生成绩的依据。

在形成性评估中,非正式评估必不可少,因为它可以随时给学生提供反馈。非正式评估和正式评估相辅相成,其作用应是同等重要的。但它不应作为唯一的评估方式,而且需要教师具备良好的职业技能。

例如,学生运用所学语言描述一些物体的不同特点时,教师应评估学生的下列能力:

1. 能否运用所学的语言结构和词汇来描述日常物品的形状和构造。
2. 能否运用自己的语言来解释物品间的不同点和相似点。

◆ 观察 (Observations)

观察包括非正式的和正式的。非正式观察如:学生某一方面语言行为的表现;正式观察如:采用标准化评价标准。观察



随时进行，而且需要按照系统的方法进行，以保证其客观性。教师要明确评估内容，如：根据想要评估的特定目标，在某一阶段只观察学生连贯表达的能力，重点做这方面的记录。这时可以暂时忽略其他方面的情况。不必追求面面俱到。

我们可以观察学生单独、一对一或分组执行任务时的表现。教师和学生可以通过一个任务或游戏来评估学生分类、记录和描述的能力。例如，当学生两人一组时，一个学生描述一堆物品中的某一个物品的形状，另一个学生把它找出来。他们从一个袋子里摸出若干物品，然后描述物品的相同和不同之处。或者一个学生描述袋子里的一个物品的特性，让另一个学生猜。

在上述例子中，教师可以评估学生的下列能力：

1. 语言表达的准确性和流利程度。
2. 能否运用所学语言描述物品特征。

五、评估的重要因素

◆ 公正

为了评估的公平，评估活动应有效、可靠，适于所有学生。评估进行的情景和状态，以及学生开展任务时的个人经验，对于某些学生，尤其是儿童来说，可以产生重要影响。

◆ 效度

有效性指的是评估任务是否能测量出所要达到的评估目标及准确的程度。例如：以考察学生语法知识的形式来检测学生写作水平的测验就不会是有效的写作评估。即便测验涵盖已经教过的知识和技能，如果测验没有被正确处理过，也不会是有效的。又如：一篇考察科学学科短文的测验要求学生具备科技



方面的词汇、知识以及相关的阅读技巧,这对于母语不是英语的学生来说,要考察阅读技巧也不会是有效的测验。

◆ 信度

可靠性是指在不同时间的各种评估活动的结果的一致性。如果同一学生几次测试的结果差别很大,那么这些测试就缺乏信度。评估活动应范围广泛,以便产生稳定的结果,使教师正确地评估学生的一种技巧在多种情况下的使用和掌握情况。例如:为评估学生的写作能力,做出最终评价,学生应至少写三篇作文;而依据15分钟的测验就可以评估学生的拼写能力。对于儿童来说,为了评估他们的成绩,教师应采取多种评估活动。

◆ 可行性

教师必须考虑评估在课堂中的可行性,并在可行、可靠和有效三者间寻求平衡。

◆ 偏见

有些偏见在任何评估时都会存在,所以应注意避免偏见并使之最小化。教师评估时,要考虑所有学生的家庭背景、生活经历、特性、知识、技能及心态,因此要采用多种评估方式。

在设计评估任务时,教师要注意:

1. 考虑到学生的不同基础。在现有基础上,给他们提供更多取得成功的机会。

2. 考虑到学生的不同经验和兴趣,因为这些因素不但影响他们解决问题的方法,而且影响到他们看待问题的角度。

◆ 调节

不同教师进行的评估互相比较的过程就是调节。调节十分



重要的一个方面就是教师在日常教学中随意地讨论教学过程或自己的教学方案。如我们会互相问：“您是怎么理解……？”或“您是怎么评估……？”调节的另一重要方面是教师正式运用同一评估任务、标准评估任务、同一打分方法、流水作业，或讨论评估过程。当不同班级或小组学生的表现互相比拟时，调节的重要性就体现出来了。如果要比较学生的终结性成绩，教师就要把不同学生的成绩结果标准化，并决定哪些评估活动更为重要。

六、学生参与的评估

◆ 自我评估(Self-assessment)

要求学生自我评估可以提高学生学习的主动性和积极性，促进学生对自己的学习进行反思，并能帮助学生掌握评估技术，增加教师的评估信息。为准确评估自己，学生应了解评估自己成绩的依据。自我评估可以很随便，如回答“告诉我你怎么……”这样的问题；也可以很正式，如：学生分析自己的成绩和知识掌握情况。学生可以用检查表来给自己排名，或依据标准衡量自己的表现。例如：我可以完成这项任务，但还不能完成那项任务。

通过自我评估，学生可设定目标，并更清楚地认识到自己的优势与不足。教师可得到诊断性信息。所有学生都应该参与自我评估，而且应及早参加。当学生熟悉评估技术，可以客观描述自己的表现时，他们将能够用事实来说明自己的成绩。这些事实可以是作业样本、录音磁带、录像、小作文和正式记录等等。

◆ 同伴评估 (Peer-assessment)