

# 第一章 英语课堂教学评估的作用

## **The Function of English Classroom Assessment**

### 本章内容要点

教学不是无目标、无目的过程。我们需要重新审视全面评估学生英语学习的发展和进步的方式，重新审视我们的教学目的、教学计划、教学方式。反思我们的教育是否为学生的发展提供了良好的环境和条件。以学生行为表现为本的教学评估可以使我们的学校从单纯的报告“分数”的评估模式转移到一个新的动态的评估环境中去。在这个新的环境中，学生应该成为积极的学习者和思考者。

### 第一节 英语课堂教学的评估与反思

#### **( English Classroom Assessment and Reflection )**

教学不是无目标、无目的过程。很多教师和教育家都感觉到我们应该重新确定我们的评估目标和方式，以保证学生朝着我们的教学目标进步。而要重新设计评估方式，我们必须确定期望学生达到的基本目标 (outcomes) 以及达到那些目标的衡量标准。行为表现教学评估不仅可以给学生展现他们所学到的知识和技能的机会，也可以给教师一个比较全面的反映教学效果的信息：究竟我们的学生已经达到了哪些学习目标？

目前所采用的主要学习评估方式——考试分数或排名，并不能完全说明一个学生学习的优劣。分数也不能完整而严格地“体现学生的成绩”。教学应重在内容、应用和方法而不是分数。随着我国教育改革的深入，教育机关、学校以及教师对检测教学的手段和方式采取了各种改进措施，但至今仍然摆脱不了分数的困扰。分数在很大程度上仍是衡量学生的重要手段，分数在人们心目中仍然占据着很重要的位置，它甚至关系到一个人的升学、谋职和未来的命运，受教育在某种意义上说就是获取高分。我们熟悉考试状况是数千人涌入一个大考场，坐下来焦急地等待密封试卷的拆封，紧接着是学生埋头用笔在答卷上画圈。这样一个半天或几个半天的考试结果，通常将决定这些学生的未来。这实际上给一些学生造成了非常大的心理压力，对他们的自信心和自尊心都有不同程度的伤害，不利于他们的健康成长。

教师和学生对目前这种评估方式和效果都非常关注，因为这种评估方法并不能反映他们真正学到的知识以及真实的能力，考试内容也往往不能代表学生日常的真实学习成就。实际上，终结性考试中的许多项目与学生所学和实际应用没有联系，这样教师和学生只好为了考试、为了好分数去教、去学，反反复复地做题和试卷，学生成了考试的机器，教师成了判卷专家。不幸的是，这样出来的“好成绩”不仅没有反映学生的真实水平和才能，还扼杀了学生的学习积极性和创造性。我们的基础教育应着眼于学生的整体发展，为每个学生提供学习的机会，为他们的终身持续发展奠定基础。以考试成绩决定学生的命运是不公平的。这种应试教育体制下的评估方式显然有悖于我们目前提倡的素质教育和终身教育的思想。素质教育的教学评估应更多地有学校、家庭、社会的共同参与，以便为学生提供开放的成长环境。学校、教师、家长、学生以及各界人士都应更新教育评估的理念，重视知识和能力，重视校内和校外、教师和学生、学生和学生的评估，让品格修养、生活和学习态度、社交情感、独立意识等行为表现都进入教育评估的行列。这

不仅需要理论上的进一步探索，更需要人们——尤其是教育工作者从理念上对教学评估有新的认识，从而将新的评估模式带到教学实践中去。以学生行为表现为本的教学评估就是一种新的英语教学评估模式，它可以使我们的学校从“分数”的评估模式转移到一个新的动态的评估环境中去。在这个新的环境中，学生是积极的学习者和思考者。

在这样一种评估方式中，我们根据英语教学的特点和作用，将教学和评估有机地结合起来。从理论上和逻辑上，我们知道外语教学和评估实践应培养学生的学习兴趣，进行启发式教育，使他们发展高素质行为并以不同的方式或形式表现出来，但在具体教学和评估中做起来并不是那么容易，我们还必须根据我们的国情、现状和管理体制来实施各种形式的课堂评估。

为了有效地实施英语课堂评估，学校应保证提供空间、时间、技术设施等支持（如投影仪、语言实验室、录音机、录相机，甚至纸张、笔和教材）同时，还要求学校、教师和家长共同合作，一起来制定评估标准并关心评估标准的发展和运用（Darling-Hammond, 1995:253），为学生提供开放的成长环境。英语课堂行为表现评估的实施对外语教师各方面的要求也更高。他们不仅要设计和实行有效的非官方的评估标准，还需要学习有关外语学习评估知识的发展和析（Farr, 1999）。行为表现评估中的道德问题以及教师职责，不同时期的教学大纲对评估的影响，评估对大纲的影响，研究方法及数据的问题、质量等问题以及评估的发展问题都是我们需要关心的。外语教师的素质提高面临着新的挑战，因为教师自身的语言水平、专业资格、教学风格和态度对评估的恰当与否有着重要的影响。

**案例 1** 一位教师在北京一所实验小学教三年级英语。有一天她要讲“HOUSES”一课，她发现课文中所介绍的有关房子方面的知识、图片与学生实际生活中的知识相去甚远。于是她收集了

各种房子的图片以便学生有更多的直观感觉从而掌握本课的知识。虽然这位教师做了努力，课堂上学生的表现和反应并未达到她所预期的那样好。

**案例 2** 一位在北京某大学教一年级写作的教师。在教学生基本写作知识和技巧的同时，采用了创造性写作（Creative Writing）的教学模式（即在基础英语写作过程中对学生进行创造性思维的训练），目的是想提高学生写作兴趣，鼓励他们创造性地进行思维，积极主动地去观察生活、描述生活、表达思想。但是前几次作业总有三四个学生不交作业。这位教师为此大为恼火，批评他们没有正确的学习态度等等。在学生小组讨论他们的作文时，教师参与了一个小组的讨论，了解到他们从小学到中学，都是被动地跟着课本、老师走，一旦给了他们“自由空间”去想像去创造他们就不知所措，不知从何下手，反而觉得没有东西可写。教师受到很大的震动，在与同事的讨论中，逐步认识到学生表现出的这种“不适应”不仅受国情和教育体制的影响，还受个人心理因素和社会家庭环境的制约。

在案例 1 的情况下，教师应该继续她的教学计划还是考虑改变教案、探讨如何使学生的学习与他们的生活更接近有关 HOUSE 方面的词汇和话题？如果我们让学生自己画各种各样的房子，然后以他们的作品为模型来讨论我们想要学生知道的词汇或句子结构及其用法效果会不会更好？案例 2 的关键是怎么使学生从“不适应”创造性写作到“适应”创造性写作。从案例 1 和案例 2 我们可以提出几个问题并作出相关的决定：（1）教材处理问题。当教材知识和内容与我们的实际生活差别太大时，我们怎么办？（2）教学计划灵活性问题。当所准备的资料和方法进行不顺利时，教师怎么继续课堂教学活动？我们是否有备用（back-up）方案？（3）评估学生问题。当我们上课不顺利时，我们是责备学生：他们不预习，他们不认真，他们没有学习积极性，他们……还是教师自己进行反

思，从自身去找原因？(4) 反思问题。当一堂课的效果不尽人意，我们是否能从学生的角度去考虑问题，是从自身找原因，从各种社会、环境、心理去反思，还是匆忙地妄下结论而导致不公平的评价结果？

教师每天都在积极不断地评估学生、评估自己、评估教材、评估教法，甚至评估整个教育体制。有很多课堂情形需要教师对学生行为进行观察和分析，然后做出或帮助他人做出有利于教学和学生的决定。英语学习的评估依赖于多种信息，单纯的课堂口头问答或书面考试不能全面反映学生的学习成果。其他因素，如学生的课堂行为、他们对学校或对他们自己的态度、学习英语的目的、学习习惯、学习方式和策略也很重要 (Genesee & Upshur, 1998:3)。当教学内容与各种技能和生活知识相关时，教师应以学生实际表现出的技能水平和所得成果作为评估的依据，即教师应观察学生的实际行为和作品，然后做出判断。这种评估方式叫做行为表现评估 (Performance Assessment)，亦称“真实评估”(Authentic Assessment)，“新评估”(New Assessment) 和 3Ps (performance, portfolio and product)。

行为评估重视创造与思考能力的开发，这些能力的培养靠长期的学习和实际运用来完成。它是自然学习环境中的一部分，应在个体参与学习的情景中“轻松”地进行，应该尽可能在日常活动过程中自然地、无意识地进行 (Gardner:1993)。通过观察一项学生活动(如演讲)或一件学生已经完成的作品(如宣传画 posters)，观察他对这一项目的理解和学习的愿望、他和其他同学一起参加项目的表现来判断一个学生的技能 (Airasian, 1991:252; Gardner, 1993)。通过课堂行为表现评估和不断观察学生的外语学习和进步，教师能不断了解学生的需求和能力以及教学策略的有效性。

## 第二节 英语教学课堂评估介绍

### ( An Introduction to English Classroom Assessment)

近年来,美国的很多中小学校和教育工作者都在调整他们的教学评估观念,尝试各种评估方式,来了解学生的学习情况。Darling-Hammond等的 *Authentic Assessment in Action* (1995) 就是通过对五个学校的个案分析,介绍“真实”行为表现评估。各学科也有相应的行为表现评估理论和标准出台,语言学科也不例外。如O'Malley等的 *Authentic Assessment for English Language Learners* (1996)。Andrew D. Cohen的 *Assessing Language Ability in the Classroom* (1994),还有 Genesee 和 Upshur 的 *Classroom-based Evaluation in Second Language Education* (1998) 对语言和英语课堂评估都有详尽的介绍和描述。Norris 等的 *Designing Second Language Performance Assessments* (1998)和 Campbell 等的 *How to Develop a Professional Portfolio* (2001) 对外语行为表现评估以及学生学习档案和教师教学档案的建立方法都有具体的描述。确实如此,当课程与各种语言技能和作品相关时,教师应观察学生的实际行为和作品,然后做出判断。在行为表现评估过程中,教师应努力去理解学生所做的努力并对他们的努力程度和进步表现做详尽的记录。教师应尽量使评估“真实”地反映学生所做的努力和其真实的行为表现,教师通过不断收集学生的资料,与学校、社区和家庭进行交流,设计实际学习机会,让家长、其他教师以及学生自己参与评估,帮助学生健康地成长和发展。

课堂行为表现评估不是单维的、平面的,它是一个多维的、立体的画面。每一面和每一个角度都代表完全不同的学习方式和创造方式(Gardner, 1993)。

以下我们先从形成性评价与传统评价方式的对比来看一看行为表现评估的特点。

英文对照表

<b>Classroom Assessment in Comparison</b>		
<b>Dimension</b>	<b>Traditional Assessment</b>	<b>Performance Assessment</b>
Purpose of Test	Summative: Teacher wants to find out what the students can remember about the course material so that a mark can be determined. It lets the teacher “sum up” what the students have learned, or to make judgements	Formative: Teacher wants to find out what the students understand about the course material so that instruction can be improved. It lets the teacher “form” a better opinion of what the students have learned, or to make diagnoses
Focus	Product of instruction	Process of instruction
	Teacher-made assessment	Student-made assessment
	Memorization is important	Critical thinking is important
	Single or narrow	Multiple or broad
	Quantity over quality	Quality over quantity
	Test itself is not a learning activity	Test itself is a learning activity
Feedback	Marks based on test are final	Marks based on task are not final but evolving
	Quick results but less complete	Slow results but more complete

Test Task	Involves written work	Often involves non-written work
	Uses paper, pen or pencil	Uses multimedia
	Non-realistic tasks	Authentic tasks
	Formal assessment	Informal assessment
Classroom Management	Intrusive (interrupts class routines)	Integrated (part of class routines)
	One chance for students to show what they have learned	Many chances for students to show what they have learned
	Revision and improvement missing	Revision and improvement present
	Aids during tests are few or none	Aids during tasks are many
Test Results	Norm (students are measured against another, larger group of what are considered to be normal test-takers)	Criterion (students are measured against specific things that are considered to be important or desirable to learn)
Decision-making	Teachers and administrators decide design of assessment	Students and parents assist in design of assessment
Assumptions	Students learn best when they are held accountable	Students learn best when they are given the chance to grow

中文对照表

课堂评估对比		
对比面(内容)	传统的评估方式	行为表现评估
测试的目的	总结性、鉴定性的:教师根据学生记诵教材知识与内容的多少来给出学生成绩。它使教师“总结或鉴定”学生已经学到了什么	形成性的:教师想搞清楚学生对所学教材的知识与内容是如何理解的,以便更好地指导教学。它有助于教师对学生所学的知识“形成”较好的认识,帮助教师进行诊断
侧重点	教学的最终结果	教学的过程(促使教师自我反思)
	教师评估	学生自我评估(做学习的主人)
	依赖或强调记忆	培养批判性思维
	评估手段和内容单一、有限	评估手段和内容多样化、广泛
	侧重于定量分析	侧重于定性分析
	测验本身不是一项学习活动	测验本身也是一项学习活动
反馈	测试成绩是最终的,可及时获取信息,但信息不完全	基于完成学习活动或任务而取得的成绩不是最终的,但是可以发展变化

测试任务	通常是书面性的	经常是非书面性的
	运用纸和笔	运用多种媒体
	非现实性的	真实的
	正式评估	非正式评估
课堂管理	侵入性的(打断了课堂教学的固有程序)	溶入性的(课堂教学固有程序的一部分)
	只有一次让学生展示他们的能力和所学知识的机会。	可以有多次让学生展示他们的能力和所学知识的机会。
	没有改正和提高的机会	允许改正和提高
	测试中不能寻求任何帮助	可借助多渠道的帮助来完成任务
测试结果	以预先设定的目标为基准,对学生达成目标的程度进行评价,通常是学生与学生之间的比较	评价的标准是看学生在学习中是否学到了真正重要的和有用的东西
评估设计者	教师和管理者决定评价的内容和标准	学生和家长协助评估的设计
假设	学生在考试中获得的成绩越好,会学得越努力	当学生得到发展自我机会时,他们会学得最好

(翻译 高越)

以学校为基础的各项英语教学评估和评价应该:

1)促进学习。教师根据评估,了解学生对知识的掌握程度,汲取学生开展学习任务的经验,以便计划教学方案,适应学生个人和学生整体的需求;引导学生进步,规划实施学习,了解学生学习是

否按计划进行（如：教师依照教学大纲规定以及学生表现而设计的特定教学目的是否达到）并及时发现问题；给学生提供反馈，进而鼓励学生 帮助学生增加信心、提高学习速度。评估是描述性的（descriptive）而不是判断性（judgmental）的过程。

2)报告进步。向家长提供关于学生进步和成绩的信息；评估提供学生进入另一学习阶段的转型信息。

3)提供终结性信息。评估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，这些可以成为学生中小学毕业时终结性成绩记录的基础；评估帮助学生对将来的学习选择做出决定，并发现自己是否具备完成课程的条件。

4)改进教学方案。评估为评价和调整教学方案提供信息，帮助教师确定哪种教学方案和学习方案有效、学生如何进步以及哪些地方有待改善。

5)起到反馈作用。评估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善教学方法。评估也反映了学生的进步是教师和学生共同履行的责任和义务。

### Summary

Teaching is not a random process; likewise, assessing what we teach should not be left to chance. Performance-based assessment can act as a lever to move schools away from the delivery-of-facts model to a new paradigm in which students are active learners and questioning thinkers. A variety of English classroom assessments promote reflection on both a teacher's teaching performance and a student's learning performance as a means to develop and improve themselves constantly (Genesee, 1998:4). We need all kinds of information such as student achievement, student behavior, student attitudes toward school or themselves, student

goals and needs concerning the outcomes of foreign language learning, student work habits, learning styles and strategies to assess our students and to improve the chances of educational success. Our students' progress is a shared responsibility between teachers and learners.

## 第二章 课堂评估的概念、 目的、种类以及原则

### **The Definition , Purposes , Types , and Principles of Classroom Assessment**

#### 本章内容要点

课堂评估是一个收集、综合和分析信息的过程；是了解学生的各项技能发展水平和发展潜力等信息的过程。它通过收集学生日常学习、教师教学以及课堂学习气氛信息，帮助教师了解每个学生的学习和需要而做出相应的决定，提高课堂效率。根据不同的评估目的，我们相应的有诊断性评估、指导性评估、正式评估以及主观性评估。评估指导原则是改进外语的教与学；评估内容包括学生需要和能力、社区对外语学习的价值观和态度、当前外语教与学的理论、学生目前语言水平以及教师能力。

评估一般来说是指具体的评估工具和过程。评价在广义上指的是分析评估数据的过程，进而成为判断教学方案是否有效的过程。根据评估数据，还可以判断学生在学习过程中的进步情况。评价学生的进步情况的过程应该是教师、学生和家长之间开展合作学习的过程。评价评估活动是否有效，教师应考虑到评估本身的有效性，而并非只是学生在评估中的表现。评价学习活动的有效性，我们应观察评估结果是否说明学习活动有效；评价教学方案

的有效性，我们应看看我们的教学方案应做哪些改善。

评估项目作为教学方案的一部分，与教学紧密相关。整个项目与课堂外的因素息息相关。在设计教学方案时，教师必须明了学校评估原则和措施。各年级学生可以通过一定时间的学习达到某一评估结果中的成绩。因此，教师在了解学生目前成绩级别和需要后，必须有选择地决定哪些结果对于教学方案是重点。这些结果将成为改善教学方案和评估活动的基础。

## 第一节 为什么评估？

### ( Why Do We Assess? )

课堂评估是一个收集、综合和分析信息的过程，是了解学生的各项技能发展水平和发展潜力等信息的过程。它可以达到两个目的：一是为学生个人提供有益的反馈；二是为学生所在的学校和社区提供有用的资料 (Gardner, 1993) 这样有利于学校、家长和社区的合作，帮助学生进步。它重视从日常表现中提取信息，通过收集学生的日常学习表现，教师教学以及课堂气氛等信息，帮助教师了解每个学生的学习和需要，从而做出相应的决定，提高课堂效率。课堂评估是一个整体的概念，不仅可以是教师对学生或是学生对教师的评估，也可以是学生对学生、学生对家长、教师对家长的评估。在教学实践中，我们非常熟悉的测试（测验）只是课堂评估技巧中的一种。这里，我们将对课堂评估作一个较全面的讨论，重点讨论行为评估。

在课堂里，由于学生的程度和能力的差异，教师需要作出各种各样的决定以适合学生的不同要求，所以他们评估学生的目的也就多种多样。有些是对学生学习的评估，有些是关于学生的个人性格和社交方面的评估，有些是对个人进步的评估，另一些则是对集体学习进展和集体行为的评估。所以相应地有以下几种主要评估种类：诊断性评估 (Diagnostic Assessment)、形成性评估 (Forma-

tive Assessment)、终结性评估 (Summative Assessment)、起点评估 (Sizing Up Assessment)、教学评估 (Instructional Assessment) 以及正式评估 (Official Assessment)。

## 第二节 评估的基本种类 ( General Types of Assessment)

### 1. 诊断性评估 (Diagnostic Assessment)

诊断性评估是指对问题的诊断 ( Airasian, 1991:5)。学生在课堂学习的过程中不仅在学习上会遇到困难,如听不懂、走神等,还有情感或社交方面的问题,如当天的心情、对老师的喜爱程度、与同学是否发生了冲突等。所以,作为教师应该首先找出问题所在,记录其频率,从理解的角度选择解决方法。学习进步不仅体现在学生每月测试的分数上,还有学生对某一主题的项目完成记载以及教师与学生家长的交谈结果。通过所获得的评估信息和结果,教师得到反馈,学生产生学习动力。诊断性评估帮助教师了解学生的知识掌握情况和能力,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。评估可通过多种方式进行如精心设计的测验、课堂教学中对学生的简单提问和回答。

### 2. 形成性评估 (Formative Assessment)

形成性评价作为教学和学习过程不可分割的一部分,描述了学生进步的整体情况,提供学生反馈来巩固学习,帮助教师了解学生的学习情况和确定下一步的教学计划及学生的学习计划。这种评估方式可以通过多种方式进行,如:对学生学习研究报告的评论、座谈、采访、测验结果的分析等。

### 3. 终结性评估 (Summative Assessment)

终结性评估是在一段学习后，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：一个学期的中间或期末，学生经过集中复习，在固定的时间内完成一套试题等。

### 4. 起点评估 (Sizing Up Assessment)

起点评估一般在学期或学年刚开始时进行，目的是为了了解学生，为良好集体的建立打下基础。起点评估要求教师在短时间内对每个学生的性格有初步了解，促进课堂上的交流。同时发现有突出特点的部分学生，让他们在今后的教学中协助教师的工作，共同作出相关决定。

### 5. 教学性评估 (Instructional Assessment)

课堂如同一个复杂的社会场所，学生和教师都是这个小小社会中的成员，在不同层面上交往。教师如同这个社会的领导人，带领着学生一起建设和营造它，使它成为一个良好的社会和学习环境。要做到这些，我们不仅需要秩序、纪律和合作，还需要正面指导。教师的决定和指导是为了建立和维持一个和谐、稳定、进取的课堂环境。所以教学性评估包括：备课评估——教什么，什么时候教，用什么材料，实施评估——把握课堂进度，对准备好的教案和计划好的课堂活动作相应的调整。

### 6. 正式评估 (Official Assessment)

正式评估指要求教师完成上级教务部门交给的成绩评估任务的职责，如：评分、评语、家长座谈等。

Airasian(1991:8)将起点评估、教学性评估、正式评估 3种课堂评估进行了比较：

	起点评估	教学性评估	正式评估
目的	给教师提供学生特点方面的情况	计划指导性活动, 监控教学的进展	执行官方任务如: 评分、排名次等
时间控制	开学后第一、二周	学年的每一天	学年期间定期进行
收集信息方法	主要是日常观察	有目的, 有计划地观察学生作业	正式考试、文章、报告、测验和作业
收集信息类型	认知、情感、心理	主要是认知和情感	主要是认知
记录方式	教师心中有数, 很少书面记录	书面备课计划, 监控信息, 无书面记录	教师有正式记载, 学校存档

上述几类评估并非彼此排斥, 有时也很难而且没有意义去分辨诊断性和形成性或其他评估。每当教师采取其中任何一种评估形式时, 都应让学生明确评估的目的、评估是如何进行的、以及评估的标准是什么。

评估也有以下四种具体类别:

1) 以结果或标准为基础的评估

以结果或标准为基础的评估根据课程大纲规定的标准测量学生成绩。此种评估不假设能达到标准的学生人数。评估和报告应依据课程标准对语言知识和语言能力的描述进行。如《普通高中英语课程标准》(征求意见稿) 六级写作的标准目标描述):

- (1) 能用恰当的格式写便条和简单的信函;