

李 加 义
顾 永 琦
惠灵顿维多利亚大学

英语形成性评估的目标和标准

【提 要】在课堂形成性评估中,教师的评估目标源于课程标准和语言能力的理论构念,明确的教学和学习目标起到重要的导向作用。此外,应明确成功达到学习目标的标准,教师对标准的理解和解释构成了学科教学知识的重要组成部分。本文通过梳理语言能力和语言标准,旨在帮助教师在课堂形成性评估中准确理解和解释目标及标准,提升学科教学知识。

【关键词】形成性评估,学习目标,语言标准,学科教学知识

【中图分类号】H319

【文献标识码】A

【文章编号】2095-9648(2021)01-0054-05

DOI:10.19502/j.cnki.2095-9648.2021.01.008

1. 引言

形成性评估(formative assessment)是指“教师、学生或同伴收集、解释和使用学生学习情况信息的过程,以便教师做出改善教学的决策”(Black & Wiliam 2009 :8)。理想的形成性评估由明确评估目的、完成评估实践和实现评估效果三部分组成(见图1),在课堂教学和评估中,形成性评估实践是由收集证据(elicitation)、解释证据(interpretation)、提供反馈(feedback)和后续行动(action)四个步骤组成的循环(Gu 2020)。在此循环中,明确的教学和学习目标起到重要的导向作用,是形成性评估的出发点和最终归宿,贯穿于评估始终(顾永琦、李加义 2020),决定收集、解释和使用信息的方式(Cowie & Bell 1999 :103),并为检查教学成功与否提供参照依据。因此,在形成性评估中,首要任务

是确定目标,只有将学习者的现有水平定位在目标坐标中,才能了解现有水平与理想水平之间的差距并探寻缩小差距的方法。

除明确教学和学习目标之外,教师应清晰了解达到学习目标的成功标准(success criteria)。在课堂语言评估中,教师对课程标准、语言能力构念以及对教与学的认识将指导其选择、设计和使用评估工具或进行实时判断,以观察学生的语言能力,明确进步方向。

近年来,国内英语形成性评估研究呈现出类型多样和主题丰富的发展特点。研究基于不同教育层次展开综述性探讨和实证研究,研究主题涵盖形成性评估的实践方法与系统、评估素养研究和评估效度研究。大多数研究都涉及学生自评、同伴互评、教师评价等评估方式,以及档案袋促学的实践手段。这些研究体现了我国形成性评估研究的最新

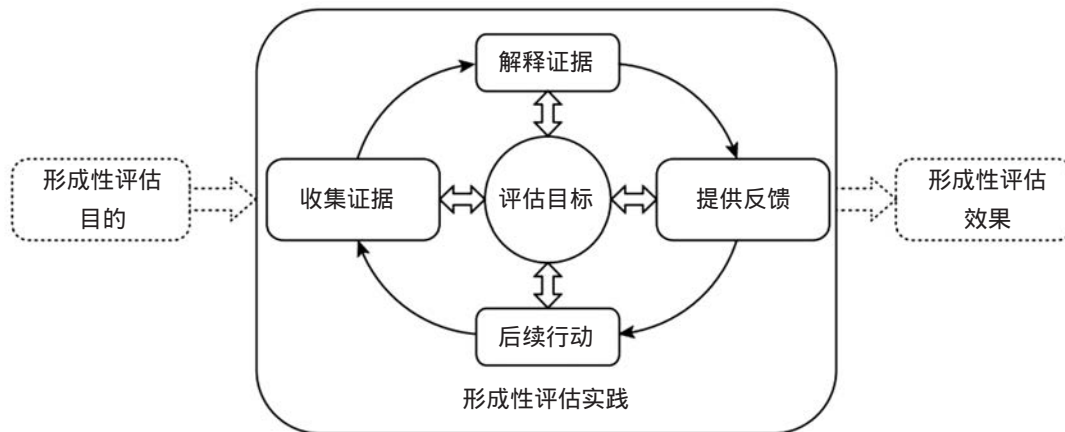


图1 形成性评估理论图

成果和发展前景,但纵观整个研究领域,笔者发现(1)我国对形成性评估的论述与研究偏重于评估的手段和应用(how),而鲜有研究重点探讨评估的目标和标准(what)(2)谈及评估目标和标准,一些教师和研究人員对于目标缺乏明确、具体的认识,对成功标准的理解模糊。在教、学、评一体化的形成性评估中,评估的目标和标准同样也是教学和学习的目标和标准。对于形成性评估,尤其是课堂形成性评估而言,如何评估、如何解释评估结果、向学生提供哪些反馈、是否或者如何采取后续行动,整个过程都离不开评估者对评估目标和成功标准的理解。因此,相较于“怎么评估”,教师应首先厘清“评估什么”,才能在教学与评估中做到的放矢,有效获取教学信息、改进教学方法、确保教学质量,帮助学生优化学习策略、改善学习方法、提高学习能力。本文旨在梳理和探讨一个“评估什么”的问题,即评估的目标和标准,以期使我国教师和研究者对语言能力和语言标准有更清晰的认识,提升学科教学知识,达到形成性评估的最终目的。

2. 语言能力标准

在英语作为外语的教学、学习和评估中,目标是什么呢?英语语言能力?那么,英语语言能力的具体体现是什么?无论教学或测评,教师和评估者都应对语言能力具有正确的看法和系统的认识。在相同语言能力范畴中,评估目的不同,所要达成的评估目标和标准也有所不同;与测试相联系时则需关注测试目的。第一,若评估目的是检查学生学习效果,评估应关注过去某一特定时间段内学生对所学知识和技能的掌握程度,目标和标准的建立应参照课程标准等。第二,若评估目的是通过检查学习者现有水平推断学生未来能否胜任继续学习,或在未来工作中使用语言的能力,评估内容往往与教学内容无关,评估目标和标准的建立需参照语言能力的理论构念和目标语言使用任务。第三,若受试者未在固定学校学习,且评估者不需要推断其未来能力,只希望了解其现有语言能力,评估学习者外语语言能力的标准就应借助语言能力理论框架。在实际测评中,评估目的通常不是单一的,因此评估者设计的测评工具可能包括多重目标和标准。例如,高考英语试题既检查考生高中

阶段对所学知识和能力的掌握程度,也对考生未来在多大程度上适合大学学习做出推断,还要了解考生对语言能力框架的整体把握情况。若学生有固定的职业方向,评估目标也包括未来语言运用领域的需求。

教师在课堂形成性评估中,评估的是什么样的英语语言能力?首先,应参照课程标准,课程标准中规定的教学和学习内容即为评估内容。其次是语言能力的理论构念,教师应跳出每堂课的目标,对语言能力的理论构念进行整体把握,将每堂课的目标对应于语言能力坐标上,并在课堂中依据情况进行随时判断。

2.1 评估构念

人们希望了解的东西(例如语言能力)往往是无形的,需对其进行理论分析以定义其本质,并设计某种具体工具尝试获取其显性的表现形式,以揭示其潜在本质。这种潜在的理论概念被称为“构念”(construct)。例如,如果目标是了解学习者的语言能力,则应从理论上分析语言能力的构成,并基于理论分析和设计测评任务,从而根据受试者完成测评任务的情况判断其语言能力的高低。构念是指人的一些假定属性,人们期望一个人在任何时候都具有或不具有定性属性,或具有某种程度的定量属性,并假设它能在测试表现中反映出来。构念在这种一般性质的陈述中具有一定的相关含义:在情况X中,具有这一属性的人将(在既定概率下)以Y的方式行事(Cronbach & Meehl 1955: 283-284)。

2.2 语言能力构念

2.2.1 交际能力

“交际能力”(communicative competence)这一概念由Hymes首次提出。Hymes(1972)认为,交际能力指准确且恰当使用语言的能力。Canale & Swain(1980)认为交际能力包含语法能力、社会语言能力、策略能力和话语能力。Bachman(1990)和Bachman & Palmer(1996)结合外语教学进一步提出最为全面、具体的交际能力模型,即“交际语言能力(communicative language ability)模型”,包括三种能力:语言能力、策略能力和心理生理机制。在此模型中,语言能力是交际能力的核心,分为结构能力和语用能力,前者侧重于语言方面,后者侧重于交际方面,这两种能力构成成功沟通的关键,

也为教学、学习和评估提供索引。

2.2.2 语言能力框架

语言能力框架是对语言使用者运用某种语言能力的一系列理论描述,是语言能力构念具体化和操作化的体现,贯穿于外语教学的各个环节中,为外语教学目标或评估目标的制定提供参照依据。

欧洲语言共同参考框架(the Common European Framework of Reference for Languages,以下简称欧框)认为与语言有关的交际能力包括三个组成部分:语言能力、社会语言能力和语用能力(Council of Europe 2001)。对于这些组成部分,欧框划分了三等六级,详细的“能做某事”描述语(can do statements)用于描述各个等级的语言使用者在不同语言使用情境和目的下完成任务的语言能力。随着语言教学理论及学习理论的发展,学界对语言能力的认识不断更新。2018年,欧框发布《新描述语指南》(*Companion Volume with New Descriptors*)除了最初强调的“行动”(“能做某事”描述语)外,更新后的框架扩充了语言能力范畴,纳入多语言和多文化、语言媒介、跨文化交际、在线交流等(Council of Europe 2018)。

欧框为中国英语能力等级量表(*China's Standards of English*,以下简称量表)的研制提供启示和借鉴(刘建达、彭川 2017)。与欧框相同,量表同样以交际语言能力模型为基础,着重强调交际中的实际运用(刘建达 2015)。量表分为“基础、提高、熟练”三个阶段,共九个能力等级,对各等级对应的英语语言能力进行描述,全面界定学习者使用英语进行交际必须达到的标准。英语学习者可参照量表定位自己的英语能力水平,制定学习目标。教师可依据量表制定教学目标和教学内容,选择适当的教学方法,选择和设计评估方法和工具等。

2.2.3 课程标准

在课堂形成性评估中,除语言能力的理论构念之外,教师的评估目标主要源于课程标准。教师根据课程标准中预设的学习目标和要求,选择和设计评价工具,采取多种具体的评价形式和手段了解学生的学习情况。课程标准使学生明确学习目标,了解要达到某一级别的标准需要在哪些方面做出努力。基于课程标准做出评价有助于保证评价标准的一致性和评价结果的客观性(Hughes 2011)。我国各学段的外语教学都建立了相应的课

程标准或教学要求,如《义务教育英语课程标准》《普通高中英语课程标准》以及《大学英语课程教学要求》等,既对我国外语教育起到了规范和指导作用,也对在教学过程中实施形成性评估提出明确指示,是教师形成性评估的主要参照指标。

2.2.4 目标语言使用任务

在语言测评(尤其是水平测试)中,目标语言使用任务(target language use tasks)是典型的测试目标,判断语言学习者和使用者将来能否胜任学习或工作的语言需要。为做出这种推断,语言测评应包括语言使用者在目标语言使用域(target language use domain)达到特定目的的语言使用任务(Bachman & Palmer 1996:44)。在语言测试的设计和开发中,测试任务应与目标语言使用域中的实际语言使用任务相关并具有代表性。此外,测试任务的真实性将决定测试任务与目标语言使用域中的真实语言使用任务的相似程度。测试开发者应把握语言使用任务与测试任务的特征,从而恰当选择和设计测试任务,确保对受试者语言能力表现的解释能够推广至测试情境外的目标语言使用域中。由于语言使用具有复杂性和多样性的特点,难以做到一一列举,但应选取目标语言使用域中具有区别性的特征。

除根据评估目的确立学习目标以外,教师和评估者还应清楚成功达到学习目标的标准。换句话说,形成性评估所要衡量的不只是学生掌握哪些知识和技能,还应衡量学习效果如何。Andrade & Heritage (2017)提出执行标准(performance criteria)用于衡量学生在语言运用过程中的表现,明确学习者成功执行一项语言任务时需要完成的任务类型以及完成程度,产出标准(product criteria)指学习者所完成的某一特定任务的质量或特征,以表明学习目标的实现情况。通常,执行标准用于为量规(rubric)或核对清单(checklist)提供产出标准,将学习目标操作为课堂可观察的评估任务。换句话说,教师和学生可利用课堂上的量规和核对清单检查学生是否和在何种程度上达到学习目标。在将学习目标转化为量规、清单和工作表(worksheet)之前,应明确学习目标在执行标准中如何体现。此外,教师需考虑学生在不同程度地成功执行特定任务时所表现出的特点。在课堂中,教师通常使用评估工具获取学生是否以及在多大程度

上实现了学习目标,而表现标准和产出标准是连接学习目标与评估工具之间的桥梁。

3. 构建和使用语言标准

3.1 构建标准

语言标准是对语言学习和语言表现预期结果的明确陈述,对语言学习者和使用者在各个层次上所应具备的能力进行详细描述。Lambert(1993)提出研制统一的语言能力标准是一个国家制订学习与使用外语语言政策的关键,构建统一的英语语言能力标准是外语教育的重要工作。标准的具体化是课程改革的整体趋势,具体的标准不仅为监测和评价提供清晰的参考点,还有助于做出一致的解释和判断。

语言能力标准的构建需考虑目的、表述形式以及理论基础与研制方法。(1)纵观世界各国语言能力标准,其目的分为:描述语言学习目标,作为考试级别或定级尺度;评定不同人员的语言水平(韩宝成 2006)。(2)语言能力标准的表述大致有两种方式:一是使用完整句子进行表述(如 21 世纪外语学习标准);二是采用“能做某事”描述语的形式(如欧框和量表)。(3)在理论基础与研制方法方面,标准的起草是将自上而下和自下而上的方法相结合,并参照类似目的和类似情境的现有标准,是经验、定性和定量研究的结果。自上而下的方法通常由对构念的理论分析和/或对学生在不同层次上的需求分析驱动。例如,欧框、量表、欧洲语言测试者协会语言能力标准和加拿大语言能力标准等均以交际语言能力理论为理论基础,聚焦完成语言活动的的能力维度;美国的外语能力标准从语言知识和语言技能两方面对语言能力进行描述。自下而上的模式依赖于从业者和利益相关者对每个层次所需的能力和任务所做出的判断。在实践中,大多数标准的构建都经历反复修订、公开协商和实证研究的过程。

3.2 解释和使用标准

在形成性评估中,制定目标和标准只是明确了起点。对教师而言,如何解释和使用标准用于判断和反馈才是真正的挑战。课程标准是出发点,语言能力的理论构念是教师对语言本质的整体把握,语言能力框架是语言能力构念的具体化,目标语言使用任务是教师对学生未来语言需求的认

识。每一次评估的目标取决于教师的专业判断。标准本质上是高度抽象的构念,人类的判断具有主观性和任意性,因而再具体的标准也会出现不同的解释。教师在课堂判断过程中通常使用三种标准:明确或规定的标准;潜在的标准;元标准或明确和潜在标准的使用规则(Sadler 1985)。教师决策会受以往判断经验影响,使用的标准是从情境和任务嵌入的经验中提取的隐性标准。这意味着对参与基于标准的评估过程的教师和学习者来说,在解释标准、做出有效和一致判断方面的培训是必要的。虽然理想中的绝对准确、公平和一致的评估难以实现,但不准确、不公平和不一致的评估是难以实现其作用的。

在实际教学和评估实践中,教师需对课程标准、语言能力构念、目标语言使用任务以及对学科本身具备清晰的认识,指导其设计和选择评估工具。教师的这种专业素质称为学科教学知识(pedagogical content knowledge,简称 PCK),即“教学的学科内容知识”和“帮助他人理解学科的示范及阐述方法”以及“对引起特定主题学习难易的因素的理解”(Shulman 1986:9)。在课堂形成性评估中,教师的学科教学知识在各方面发挥积极作用:(1)了解学习的总体目标以及学生在每堂课中应达到的具体目标;(2)计划或即时决定如何使用最合适的工具衡量学生现有学习水平;(3)对所取得的结果进行解释并就下一步的教学和学习做出判断;(4)向学生提供反馈,提出缩小学生现有水平与理想目标之间差距的方法;(5)为学生提供缩小差距的机会和任务。如果明确教学与评估的目标和标准是形成性评估的第一步,那么教师的学科教学知识则是形成性评估赖以成功的基本保证。

4. 结语

在以英语为外语的课堂形成性评估中,评估的目标和标准主要来自两个方面:课程标准和语言能力的理论构念。对于评估目标和标准的理解与解释直接构成了教师学科教学知识的一部分,也决定着形成性评估的成败。在教学、学习和评估过程中,目标明确才能有的放矢,形成性评估的主要动力即是瞄准这些具体目标,利用多种评估手段,为教师和学生提供详细和丰富的学习信息,使学生明确现有水平与目标的差距以及如何进步,

为学生提供充足的进步机会和方法,达到形成性评估的最终目的。

参考文献

- Andrade, H. L. & Heritage, M. 2017. *Using Formative Assessment to Enhance Learning, Achievement, and Academic Self-regulation* [M]. New York, NY: Routledge.
- Bachman, L. F. 1990. *Fundamental Considerations in Language Testing* [M]. Oxford: Oxford University Press.
- Bachman, L. F. & A. S. Palmer. 1996. *Language Testing in Practice: Designing and Developing Useful Language Tests* [M]. Oxford: Oxford University Press.
- Black, P. & D. Wiliam. 2009. Developing the theory of formative assessment [J]. *Educational Assessment, Evaluation and Accountability*(21) 5-31.
- Canale, M. & Swain, M. 1980. Theoretical bases of communicative approaches to second language teaching and testing [J]. *Applied Linguistics*(1) :1-47.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* [M]. Cambridge: Cambridge University Press.
- Council of Europe. 2018. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion volume with new descriptors* [EB/OL]. [2020-06-30]. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>.
- Cowie, B. & Bell. B. 1999. A model of formative assessment in science education [J]. *Assessment in Education* (6) : 101-116.
- Cronbach, L. J. & Meehl, P. E. 1955. Construct validity in psychological tests [J]. *Psychological Bulletin* (4) 281-302.
- Gu, P. Y. 2020. *Classroom-based Formative Assessment* [M]. Beijing: Foreign Language Teaching and Research Press.
- Hughes, A. 2011. Towards a personal best: A case study for introducing ipsative assessment in higher education [J]. *Studies in Higher Education*(36) 353-367.
- Hymes, D. 1972. On communicative competence [A]. Pride, J. B. & J. Holmes (eds.). *Sociolinguistics* [C]. Harmondsworth: Penguin, 269-293.
- Lambert, R. 1993. Foreign language competency [A]. Moore, S. J. & C. A. Morfit (eds.). *Language and International Studies: A Richard Lambert Perspective* [C]. Washington D.C.: National Foreign Language Center Monograph Series.
- Sadler, D. R. 1985. The origins and functions of evaluative criteria [J]. *Educational Theory*(3) 285-297.
- Shulman, L. S. 1986. Those who understand: Knowledge growth in teaching [J]. *Educational Researcher*(2) 4-14.
- 顾永琦、李加义 2020 形成性评估的效度[J]《外语教育研究前沿》(3) 34-41。
- 韩宝成 2006 国外语言能力量表述评[J]《外语教学与研究》(6) 443-450。
- 刘建达、彭川 2017 构建科学的中国英语能力等级量表 [J]《外语界》(2) 2-9。
- 刘建达 2015 基于标准的外语评价探索[J]《外语教学与研究》(3) 417-425。
- (李加义:惠灵顿维多利亚大学语言学与应用语言研究学院博士生;
顾永琦:惠灵顿维多利亚大学语言学与应用语言研究学院院长,博士生导师)
- 收稿日期 2020-11-03
通讯地址 新西兰惠灵顿维多利亚大学语言学与应用语言研究学院