

导 读

一、概述

《面向学习的测评：一种系统的方法》(*Learning Oriented Assessment: A Systemic Approach*) 一书 2016 年由剑桥大学出版社出版。本书是“语言测试研究”(Studies in Language Testing, 简称 SiLT) 系列丛书的第 45 册。作者是剑桥大学英语考评部 (Cambridge Assessment English, 以下简称“考评部”) 的 Neil Jones 和 Nick Saville。书中提出了语言测评领域一种新的测评范式——面向学习的测评。该模式具有系统性和生态性, 是将终结性评价 (summative assessment) 和形成性评价 (formative assessment) 相结合的有益探索。作者特别强调, 为了有别于其他学者对面向学习的测评的研究, 本书的正标题首字母需要大写, 简称为 LOA。

关于 LOA, 考评部官网的解释如下:

LOA 的提出是为了从传统测评方式中探索出一种新的测评范式, 该范式注重信度和效度, 首要目标是促进学生的语言学习。LOA 的系统性表现为它强调多种测评方式在多个层面以不同形式加以运转, 既包括在宏观层面上设置教育目标和成就评价标准, 又包括在微观层面上促进个人在课堂内外进行学习互动。LOA 强调所有层面的测评方式都能够并且应该保障学习的有效性和评价的可靠性。LOA 基于对定量测量维度和定性个性化功能划分, 界定了测评和教学的互补关系, 把教师放在创造学习环境的中心位

置，把正式测评（formal assessment）引入学习环境中以取得积极的影响。

(Cambridge Assessment English 2019)

本书作者提到了 LOA 六个方面的研发背景。

1. 基于测试反拨效应（washback）的影响，即外部测评（external assessment）对学校 and 课堂环境下的教与学的影响

自 20 世纪 80 年代以来，受交际法（communicative approach）对教学和测评的影响，学者们开始对反拨效应产生浓厚兴趣。当时虽然很多教师认为测评会对教学产生积极的影响，测评的内容和形式对课堂教学具有积极的引导作用，但是高风险考试也有阻碍学习的一面，例如，很多学生总结的应试技巧影响了考试的效度。事实上，考试的影响不仅涉及学生的课堂学习，而且会延伸到社会的各个方面，例如，考试会影响教育政策的制定和整个教育生态系统。考评部一直致力于研究考试是如何促进正式教育环境下的教学，并使考试对课堂教学产生积极的反拨效应。随着测试反拨效应的研究范围不断扩大，考评部思考的测试反拨效应的维度也由课堂和学校层面上升到整个社会层面。

2. 基于测评“二分法”对立的传统

一直以来，一提到测评，很多学者就会想到几组对立面：终结性评价和形成性评价，标准化评价（standardised assessment）和课堂评价（classroom assessment），以评测学（assessment of learning）和以评促学（assessment for learning）。每组的前者都具备以下特征：1）在学习结束时实施；2）基于某项教学大纲或者理论模型；3）设计基于信度和效度的统筹考虑；4）作为一种判断，经常与“评分”挂钩，是对过往学习成果的“回头看”。与之相反，每组的后者则具备以下特征：1）具

有不间断性；2) 是对学习者不断变化的需求的回应；3) 注重交流、支持和发展；4) 被贴上缺乏“信度”和“效度”的标签；5) 目的是“向前看”，引导学生前进。几组对立面的前者和后者貌似互相对立，实际上可以相辅相成、相互融合、共同促进学习，这是考评部研究的焦点，即让貌似对立的测评整合在一起，服务于学生的学习。

3. 基于《欧洲语言共同参考框架》(*The Common European Framework of Reference for Languages*, 简称《欧框》) 的确立

在欧洲，从1989年至2010年，测评改革小组 (Assessment Reform Group) 深入研究了以评促学，以提高教育水平、提升教学效果。这个目标反映到测评中就是研发新的量表，以更加准确地评估学习者的学习情况，构建解释性框架，将学习成果报告给利益相关者。在这一需求下，《欧框》应运而生。《欧框》独立于任何形式的语言测评，它是对不同程度的语言水平所作的系列描述，包括在任何一个等级上的人对于该语言应该有怎样的运用能力。就LOA而言，依托于一个解释性框架至关重要，《欧框》的诞生为LOA的提出奠定了理论基础与框架。

4. 基于对“学习”和“二语习得”的理解

建构主义 (constructivism) 强调学习者的主动性，认为学习是学习者基于原有的知识和经验生成意义、建构理解的过程，而这一过程常常是在社会文化互动中完成的。建构主义对教学和学习具有重要的意义，而社会建构主义 (social constructivism) 则进一步强调学习的社会性和合作性。建构主义，尤其是社会建构主义，让人们深思“学习”的本质，而“二语习得”作为“学习”的一种特例，也需要探讨学习的社会属性，这也促进了LOA的研究。

5. 基于对“教育状况改变”的认识

Dewey (1916) 指出，教育的目标是使个体能够继续他们的教育。对于正式教育下的学习者而言，他们要学习的内容不仅仅是课程教学

目标要求的知识，更重要的是学习者要学会改善性情、态度，掌握实际的学习技巧，这对于学习者而言意义重大，因为学习是终身的。从建构主义角度而言，如果能够将课程、教学、测评和社会融合在一起去促进学习，那么教育的目的性和连贯性就可以得到保障，课程学习目标的达成和学习者终身学习能力的养成就有实现的可能，而《欧框》的研制就是这种可能的一个范例。基于《欧框》，LOA 致力于培养学生的学习能力，促进其终身学习，力图为教育状况的改变提供新的思路。

6. 基于移动技术的发展

自 20 世纪中叶以来，信息技术快速发展，移动设备应运而生。在此背景下，人们的信息存储与处理能力以及知识的获取能力显著增强，学习者的学习和移动设备开始挂钩。这为 LOA 的提出奠定了技术基础。

二、内容简介

本书一共九章。第一章总体引介 LOA，介绍了全书的编排结构，并着重介绍以“任务”（task）为驱动四个“世界”——“个人”“教育”“社会”“测评”。

第二章主要探究 LOA 的发展根源，介绍了 LOA 在不同背景下的发展情况，并列举了一些面向学习的方法。LOA 的提出是为了应对剑桥英语测试所面临的新问题和新使命。LOA 反思了测评在社会和学习中所扮演的角色，梳理了不同学者对 LOA 的定义。

为了阐释不同层次的测评是如何共同促进学习的，第三章至第五章提出了 LOA 涉及的三个基本问题：什么是学习？什么是语言学习？要学习什么？第三章从个体认知的角度介绍了建构主义和社会建构主义，为 LOA 提供了较为合理的理论模型。本章强调任务和教师在鼓励学习

者互动方面的重要性，指出这种做法可以提高课堂评价的效果。

由于 LOA 与学习紧密相关，语言学习是 LOA 的目标，同时也是学习的中介，所以第四章将主题聚焦到教育的特殊案例——语言。通过运用《欧框》这一基于行动的案例，本章主要讨论学习者语言能力的发展。此外，本章还引出第一语言和第二语言的学习过程，以加深读者对社会—认知 (socio-cognitive) 模式下的学习的整体理解。

第五章聚焦正规教育背景下语言学习的组织，讨论学习的三个方面：社会期待的结果、需要掌握的技能（社会—认知建构）和学习的内容（课程内容）。本章“技能” (skill) 一词在第一语言学习中被解释为初级机制 (entry-level mechanism)，而在第二语言学习中则被解释为高级能力 (higher-order competence)。

第六章到第八章探讨大规模测试在学习中的角色和 LOA 在课堂中扮演的角色，以及二者的结合问题。第六章介绍了考评部等机构实践的大规模测试，探讨了一系列保障大规模考试信度和效度的重要条件，比如水平测试中的能力参照、量表建构、用于估计题目难度的项目反应理论 (item response theory)、用于测试建设的题库 (item banking)。这些都有助于验证 LOA 增强学习的积极影响。

第七章主要探讨 LOA 在课堂层面关注的两个问题：如何生成“交互”从而促进学习和如何获取“交互”的证据以最大限度地增强其对未来学习的积极影响。本章揭示了课堂中 LOA 的本质，包括特定领域的学习技能 (domain-specific learning skills)、可推广的学习技能 (generalisable learning skills)、以学习为中心的活动 (learning-centred activities) 及以内容为中心的活动 (content-centred activities)。

基于前两章，第八章将大规模测试和课堂评价结合起来建构了一个综合模式。该模式可以帮助读者理解在共同解释框架内宏观层面（外部测试）和微观层面（课堂评价）的证据，从而确保这两个层面的评

价都聚焦共同的目标和标准。

第九章阐述了全面实施 LOA 涉及的问题，如教育政策制定、形成性评价实践的经验教训、技术的关键作用等，以及实施 LOA 的大体步骤。

三、本书特色

本书具有三大特色。

特色一：探索了大规模测试和课堂评价的共性与互补性，更新了读者对测评的理解。大规模测试以其总结性特征，通过对学习者学习成果的判断、描述、记录和报告，可以衡量学习者在一段时间内的学习成效。课堂评价具有形成性特征，能够通过教师对学习者的情感、态度和策略的观察、记录和反思来展现学习者的日常表现和进步。LOA 是一种综合性思维，它提供了一种系统性的评估模型，在某种教育背景下的多个层面上运行，并具有不同的表现形式 (Jones *et al.* 2013)。它既包括宏观层面上制定的教育目标和评价结果，也包括微观层面上个体在教室内外进行的学习互动 (Jones 2014)。基于《欧框》，LOA 在宏观层面为外部测试提供交流技巧、主旨、话题、功能、任务等方面的参考，指导外部测试实施，产生成就记录；在微观层面制定 LOA 大纲指导，课堂评价。课堂评价作为内部测评，主要围绕 LOA 活动以获取非正式的成就记录。这些记录和外部测试的成就记录共同为 LOA 提供了促进学习的证据。外部测试的成就记录既要能够诠释《欧框》和 LOA 大纲，又要起到对课程成就的外部终结性测试进行监控的作用。通过将这两种测评结合，借助二者的优势，LOA 可以产生有价值的结果：促进学习、提升测评的解释力度 (如图 1 所示)。

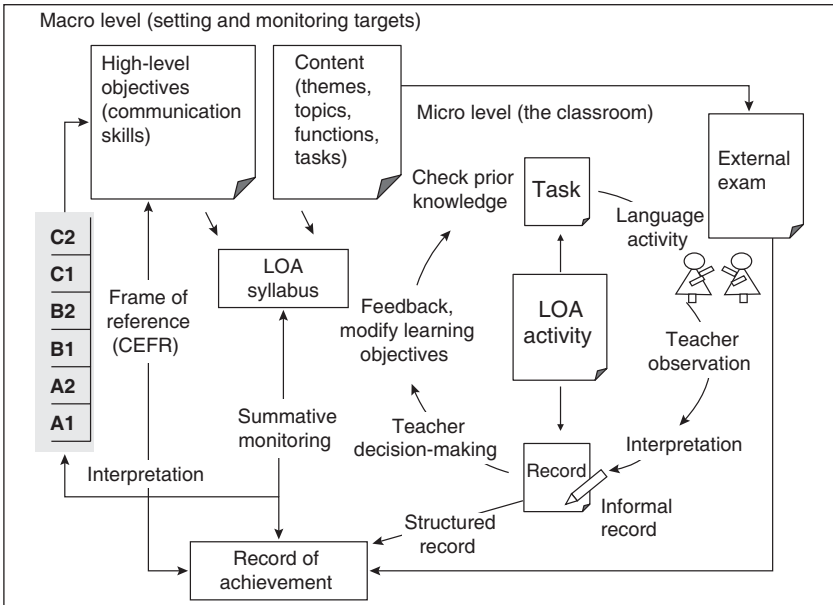


图 1 LOA 实践模型 (Jones and Saville 2016: 13)

特色二：构建了“四个世界”，并将四者以“任务”为驱动形成环状，以服务于“学习”的目的。这“四个世界”为“个人”“教育”“社会”“测评”，分别代表个人认知的发展、课程学习、社会技能和专门语言技能的获取、测评任务的构建。“个人”运用“任务”发展个体的认知；“教育”运用“任务”设计课堂练习、组织正式教学；“社会”运用“任务”评价技能表现；“测评”以“任务”为测量基础。通过“任务”驱动实现了两种方式的“学习”：在现实世界中通过参与“任务”实现自然习得（natural acquisition）和在正式教育环境中通过接受教育实现学习。通过“四个世界”，个人、教育、社会、测评构成了一个连贯的有机体，体现了LOA的系统性和生态性。“四个世界”也体现了LOA的社会认知性，例如，个人为了获得进入社会所需要的技能，可以通过接受教育获得认知的发展，实现学习的进步。LOA对学

习的诠释体现了教育的本质：学习不只发生在课堂，而应贯穿学习者的一生（如图 2 所示）。

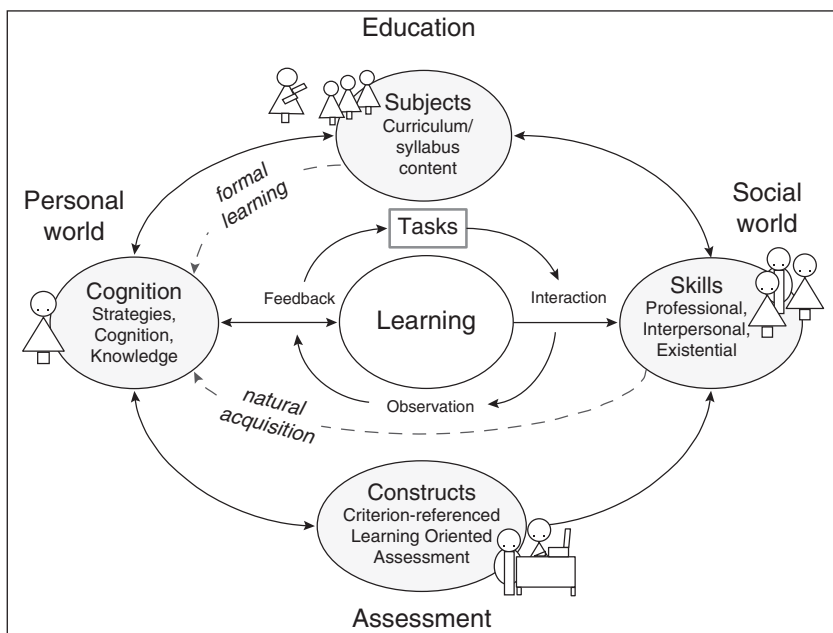


图 2 以“任务”为中心连接起来的“四个世界”（ Jones and Saville 2016: 5 ）

特色三：LOA 的实施设想涵盖了一系列的利益相关者，在一定程度上使 LOA 更具可行性和可操作性。LOA 认为，考试机构应该负责制定测试的总体原则；政策制定者对相关政策的批准和实施应起到关键作用；学习者和教师是 LOA 实施的关键参与者，学习者应该积极主动参与到学习中，教师的主要责任是创造良好的学习环境、观察学习过程、反馈学习成效；教育管理者也应该参与到 LOA 中，为 LOA 的实施提供必要的软硬件设施。

总之，本书理论性强，囊括了大量与语言测评相关的议题、原则与实践。语言测评理论及概念涉及项目反应理论、题库、过程描述

(processing accounts)、复杂理论 (complexity theory)、基于频次的描述 (frequency-based account) 等。书中探讨了大规模测试与课堂评价目标和标准的整合问题, 提供了 LOA 视域下新时代的课堂评价模式。

本书写作体例合理, 便于读者理解。本书按照“什么是学习”“要学习什么”“要测评什么”的思路进行撰写, 阐释了 LOA 的基本内容, 指出了教师、学习者、测评专家等应该如何利用 LOA 促进教学、学习和测评。每章开头都引用了 Dewey 的观点道破本章主旨, 大部分章节结束时都会加以总结, 便于读者理解。

本书也有以下不足之处: 1) 没有给出 LOA 的确切定义, 容易使读者产生困惑; 2) 究竟如何在现实世界中实施 LOA, 本书似乎没有给出可操作的范例; 3) 第七章的“7.5 Learning-oriented assessment: An ecological model”放在第八章似乎更合适, 因为第八章探讨的是形成性评价和终结性评价的整合问题; 4) 第二章和第八章末尾缺少概括, 与其他章节的写作风格不一致。

四、本书对国内读者的意义

本书系统地介绍了剑桥英语视域下的 LOA, 对国内读者具有一定的参考价值。为了建构 LOA, 本书梳理了 LOA 的研究根源, 探讨了学习和语言学习的内容、学习的目标, 以及大规模测试和课堂评价在学习中所扮演的角色, 并把大规模测试与课堂评价加以整合, 提出了 LOA 的实施方案。作者的写作思路清晰、逻辑性强、论证严密、内容丰富, 为国内研究者提供了较为翔实的文献资料。

本书探讨的 LOA, 深化了我们对终结性评价和形成性评价二者互补关系的认识, 有助于我们理解如何基于这两种测评收集促进学习的证据、提高测评的质量。本书把促进学习的证据分为定性证据

(qualitative evidence) 和定量证据 (quantitative evidence)。定性证据指的是水平维度 (horizontal dimension) 的日常教学和学习, 探讨“如何帮助每位学习者获得进步”; 定量证据指的是垂直维度 (vertical dimension) 的测试, 探讨“学生进步了多少”。本书认为水平维度和垂直维度有互补, 也有重合: 定量证据可以帮助每位学习者鉴定技能水平, 定性证据可以帮助学习者获得定量测评体现的学习成果。这种阐释更新了我们对“二元对立”的认识, 为我们开展终结性评价和形成性评价的整合研究提供了思路。

本书倡导的 LOA 具有系统性、生态性、社会认知性, 有助于我们理解学习的本质, 同时有助于外语教师转变教学理念、创新教学方法、实施更加有效的教学与测评。阅读本书, 我们会认识到, 学习并非简单的“教与学”“学与测”, 而是一种社会认知行为。学习需要基于行动, 在测评“任务”驱动的“交互”中才能更有效地进行; 学习需要突破课堂的时空限制, 把学习者的认知发展与他们在社会中的发展相统筹, 培养让学习者受益终身的性情、态度和技能。这些观点对于国内语言政策的制定、课程大纲的设计、教学内容的组织具有启发意义。

阅读本书, 我们认为以下问题特别值得读者思考与讨论:

1. 在中国教育背景下如何定义 LOA ?
2. 怎样实施 LOA ?
3. 如何确保利益相关者, 尤其是教师达到 LOA 的要求?
4. 如何判断测评任务具备 LOA 的特点?

辜向东 李玉龙

2019年1月12日

参考文献

- Cambridge Assessment English. (2019) Learning Oriented Assessment. <https://www.cambridgeenglish.org/research-and-validation/fitness-for-purpose/LOA>
- Dewey, J. (1916) *Democracy and Education*. New York: The Free Press.
- Jones, N., Saville, N. & Hamilton, M. (2013) A systemic view of assessment within an educational context, paper presented at ILTA/AAAL symposium, Dallas.
- Jones, N. (2014) *Multilingual Frameworks — The Construction and Use of Multilingual Proficiency Frameworks*. Cambridge: Cambridge University Press.
- Jones, N. & Saville, N. (2016) *Learning Oriented Assessment — A Systemic Approach*, Studies in Language Testing, Vol. 45. Cambridge: Cambridge University Press.
- Li, Y. & Gu, X. (2018) Review of *Learning Oriented Assessment: A Systemic Approach*. *Innovations in Education and Teaching International* 55(4): 497–498.