

# 外语类教材评价标准框架的构建及其内涵解读<sup>1,①</sup>

张 莲 孙上鸚 北京外国语大学

**提 要:** 本文报告外语类教材评价标准框架的构建及其内涵解读。研究采用混合研究法, 研究步骤包括: 通过相关概念解析和文献综述初构框架, 进行大规模调研重构框架, 采用准德尔菲法修正框架及相关指标, 通过模型拟合进行量化验证, 以及提交专家组质证等。通过上述研究步骤, 构建并验证外语类教材评价标准框架。该标准框架包含六个一级指标, 分别是思想性、规范性、科学性、系统性、适用性和趣味性, 以及在一级指标基础上析出的 18 个二级指标。最后, 本文对标准框架一级指标的内涵进行了简要解读, 并对框架应用前景提出了建议和展望。

**关键词:** 外语类教材评价; 评价标准框架构建; 混合研究法; 一、二级指标

[中图分类号] H319 [文献标识码] A [文章编号] 2096-6105(2024)04-0033-08

DOI: 10.20083/j.cnki.fleic.2024.0035

## 1 引言

建设教育强国, 是全面建成社会主义现代化强国的战略先导, 是以中国式现代化全面推进中华民族伟大复兴的基础工程 (中国政府网 2023)。教材 (materials) 是教育教学活动的核心载体, 承载着为党育人、为国育才的使命和任务。立足新的历史方位, 为推动中国特色外语教育高质量发展, 建设高质量外语教材体系是基础工程中的基础。长期以来, 国内外语类教材建设取得了积极成效, 推出了一大批具有中国特色、适应时代变化的大中小学优秀外语教材, 以服务不同历史阶段国家和社会对外语人才培养的需要。但同时也存在一些亟待解决的问题, 如对新时代外语类人才培养模式变革的新理念与新要求体现不足, 价值塑造、知识学习和能力发展的融合尚不充分; 服务国家发展需求的意识仍不够凸显; 重大主题教育内容尚未在教材中得到系统化、制度化的落实; 同时, 新技术的突飞猛进也给外语类教材建设带来更多挑战等。历时来看, 这些问题的起因复杂且多元, 因此, 任何解决这些问题的构想必然都需要全面、系统的考量和斟酌, 然

而, 这些问题的产生无疑和当前缺乏一套行之有效的评价标准有关。无论是外语类教材建设机制的完善, 还是外语类教材体系的完备和外语类教材质量的提升, 都离不开外语类教材评价标准框架的构建。

教材评价 (materials evaluation) 是对教材质量和价值的科学、系统的分析和评估, 以引导教材的设计、编写和选用 (Fey & Matthes 2018; 何安平 2020; 庄智象 2006)。相关文献纵览显示, 国外教材评价标准研究起步较早, 积累了较丰富的研究成果。自 20 世纪 70 年代末至今, 先后已有五十余份教材评价标准被提出 (如 Cunningsworth 1995; Ellis 1997; Hutchinson & Waters 2002; McDonough *et al.* 2013; Seaton 1982; Tomlinson 2023)。国内教材评价标准的探索始于 20 世纪 90 年代中期 (李定建等 1995; 周雪林 1996)。在之后近三十年里, 研究者先后提出十余份教材评价标准。这些成果有效支持并推动了国内教材评价实践, 但也存在一些问题。首先, 现有标准以引介或应用国外相关标准为主, 系统循证构建标准的研究较少。国内相关标准的提

1 本文系北京外国语大学“双一流”重大(点)标志性项目“高校外语教育与外语人才标准体系研究项目”(项目编号: 2022SYLZD010)的研究成果。

出或基于国外某个或若干个标准的改编和修订,或基于研究者的个人认知,缺乏实证和循证程序,因此,这些标准是否具备本土适用性和实操性有待验证(刘道义 2020;吴先强、王祖浩 2021)。其次,在过去五年间,大中小学英语和外语类本科专业均颁布了相关课程标准、指南或教学质量国家标准等,相关理念、原则应在教材评价标准中有所体现(孙平华 2006),而现有标准尚未及时反映外语教育教学宏观层面发生的变化,对国家外语教育政策和规划的呼应尚显不足。最后,国内现有教材评价标准在评价细则上侧重对教材内容的选择、编组和呈现的评价,而对教材“教育教学属性”的评价不足,评价主体性“缺失”或主体相关性模糊(贾蕃、沈一新 2020;柳叶青 2017;潘信林 2022;吴先强、王祖浩 2021)。本研究旨在立足本土外语教育教学环境和实际,关注教材评价主体的真实感受,构建切实可行的本土外语类教材评价标准框架(简称“标准框架”),并对其可靠性、合理性和可用性进行结构验证和专家质证,以解决现存问题,弥补相关研究之不足。

## 2 关键概念界定及标准框架构建的依据和原则

### 2.1 教材评价、教材评审和教材评价标准

如前文所言,教材评价是对教材质量和价值的科学、系统的分析和评估,以引导教材的编、选、用。教材评价标准是评价教材质量和价值的工具,帮助评价者对教材进行综合评价(Mukundan *et al.* 2011;刘道义 2004;王晓丽 2021)。在教育领域中,“标准”一词包括定性作为参照范围,定量作为切点分数点两个维度(黄光扬 2022),所以,常见的教材评价标准既有量化打分表,也有质性陈述清单,或两者混合,各有长短。评价标准常见的别名包括清单(checklist)、原则(principle)、指南(guideline)、框架(framework)等(AbdelWahab 2013; Littlejohn 1998; Mukundan 2009; Mukundan *et al.* 2011; Nimehchisalem 2018; Williams 1983; 翟志峰、董蓓菲 2019; 赵勇、郑树棠 2006)。

与教材评价相关,并时常共现的概念是教材评

审。在《现代汉语词典》(中国社会科学院语言研究所词典编辑室 2005:1055)中,评价意为“(动词)评定价值高低,(名词)评定的价值”,如“评价文学作品”“观众给予这部电影很高的评价”等。评审意指“(动词)评议审查,(名词)担任评审工作的人”,如“评审验收”“评审文艺作品”“这次大奖赛共有九名评审”。两者既有联系也有区别(刘道义 2020)。在教材评价标准的语境下,评价和评审的不同体现在四个方面:(1)二者的参与主体有异,评价者多为教材使用主体,评审者多为教育行政主管部门者;(2)二者的标准(或原则)虽有重叠,但质量关切和侧重不同;(3)评价和评审的程序及工作方式不同;(4)二者的最终结果及影响不同(同上)。

### 2.2 标准框架构建的主要依据和原则

构建教材评价标准框架的基础是对教材及其本质属性进行定义。依前文逻辑,外语类教材是外语教育教学活动的核心载体。这一论断虽简洁,但含义丰富且深远。在哲学层面上,这一载体必然涉及外语知识的本质,外语教育的目的,文化的传承、传播与变革等。在心理学层面上,这一载体涉及二/外语习得和学习过程、认知发展、动机与情感等,自然也涉及内容选择和过程安排的科学性和合理性。在教育学层面,教材不仅是系统、规范地传递、传承和传播信息和知识的工具,更是“一个国家的‘精神国防’”,承载着“传承文化知识、培养时代人才、维护价值导向、赓续社会规范”的使命,对人的成长和发展具有深远的影响(余宏亮 2021:30)。可见,任何一个外语教材实体的出现和存在都应该是这些学科视角汇聚、碰撞、互动和协同的结果。换言之,它必然体现外语教育教学活动的目标和价值导向,融合关于语言、语言习得和学习的研究成果与共识,关注学习者作为活动主体的心理认知特点,遵循二/外语教与学的本质与规律,并基于此,规划外语知识学习和能力发展的愿景和路线,进而定义和形塑活动主体间性关系。任何特定学科、学制、学段教材的内容筛选和编排、问题和练习设计、语言使用和辅助材料的搭配,甚至印刷的风格、字体、色彩、配图的选择等,都是源自以上各个层面

的斟酌和考量 (Norton & Buchanan 2022; Tomlinson 2023)。应该说, 这些层面构成教材的基本面或教材研究的知识基础, 体现教材作为教育教学活动载体的一般属性。所谓一般属性是指教材区别于其他书籍或资料的特有属性, 具有区别性, 亦称教材的本质属性。不同科目、学制或学段等则体现教材的特殊属性。本研究意在构建基于外语类教材本质属性的评价标准框架。因此, 无论是作为领域的教材研究, 或是 (将要) 作为一门学问的教材学, 它都必定是跨学科的 (李政 2022; 束定芳 2024; 文秋芳 2023)。这或许是“教材还因‘作为一类特殊文本’极为复杂而‘不好’研究”的一个主要原因 (余宏亮 2021: 31)。二/外语教材作为应用语言学的一个次领域, 虽然积累了较多成果, 但其理论性和系统性相对于其他次领域似乎处于欠发达状态 (张莲、黄国文 2024)。

基于上述讨论, 笔者认为, 研制外语类教材评价标准框架的主要依据应该包括: (1) 教材的概念和属性界定; (2) 相关课程标准、指南和质量标准

及其背后的二/外语习得和二/外语课程与学科教学的理论等; (3) 教材使用者 (如教师、教研员、学生等) 的感知和评价; (4) 教材研究者和团体的研判意见; (5) 教材编写者、出版者的认知和研判; (6) 教材管理部门颁布的各级各类教材管理办法。基于上述依据, 笔者提出六条研制原则: (1) 与学科和专业育人、育才目标相一致; (2) 经验性归纳和概念性演绎相结合; (3) 理论研究和实践应用相结合; (4) 国际经验参照和本土实际情况相结合; (5) 突出本质属性, 兼顾特殊属性; (6) 框架和指标话语的学术性、规范性和可理解性相结合。

### 3 标准框架构建与验证的过程和方法

标准框架的研制过程分为两个阶段: 构建与验证。整个过程包括五个步骤: 框架初构、规模性调研重构、专家论证质性验证、模型拟合量化验证以及专家质证 (如图 1 所示)。步骤的设定兼顾质性和量化的多轮综合范式, 以确保构建的标准框架和指标内容科学、规范、合理且适用。

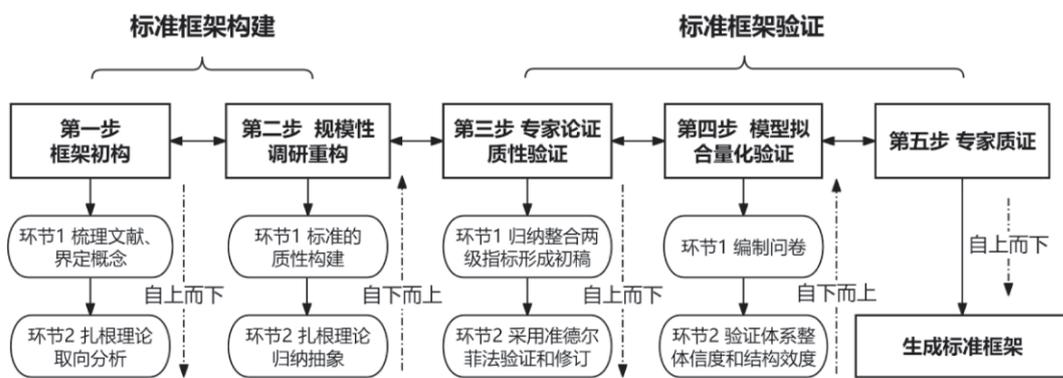


图 1 标准框架构建和验证的过程

#### 3.1 多轮综合的外语类教材评价标准框架构建过程

教材评价标准框架构建的第一步是框架初构。笔者梳理并分析了国家层面对教材的要求和外语类教材评价研究重要文献, 特别是相关核心概念和国内外 70 余份教材评价标准或框架及其历时演进的概貌。其中, 2008 年之前的国外教材评价标准, 选取 Mukundan & Ahour (2010) 综述的 48 个教材评价一览表进行分析; 2008 年之后的选取 Web of Science 数据库中高被引文献提出的教材评价标准进行分析。国内外语类教材评价标准则选取中国知

网中的北大核心期刊及 CSSCI 来源期刊中的文献进行分析。笔者用 NVivo 系统地对两组文献进行扎根理论取向的分析, 通过节点关联, 发现并建立概念矩阵, 形成核心概念类属, 并基于此构建初始的标准框架和维度。此步骤遵循自上而下的演绎逻辑。

第二步是开展规模性调研重构框架。基于第一步形成的初始框架和维度, 笔者采用半结构化访谈的形式进行标准框架的质性构建。笔者采用目的性分类抽样的方法, 在全国七大行政区不同类型的高

校和机构中开展了 192 场半结构化访谈，历时三个多月。访谈对象即潜在评价主体，包括 81 位高校外语教师、93 名学生、四位教材主编和 14 位出版行业人员。其中，15% 的受访教师也兼任教材主编或编者。访谈的教材主编和编者均为首届全国教材建设奖外语类教材奖获得者及相关领域内知名专家。当访谈数据达到信息饱和时，笔者使用 NVivo 对访谈结果进行标注和分析，之后经过多轮概括和凝练，形成 36 条描述项，并基于此，最终析出六个一级指标，即思想性、科学性、系统性、适用性、规范性和趣味性。

以上述一级指标为基础，笔者结合相关文献和数据分析，继续挖掘、解析、界定各一级指标的内涵，生成 18 个二级指标；同时，将一、二级指标与教材实体本身的六个基本维度（即目标、内容、结构、语言、编绘和印刷）关联起来，使标的和标的物之间形成闭环。以一级指标“思想性”为例，思想性并不是一个空洞、虚幻的评价标的，而是应该体现在一本（套）教材具体的目标设定、内容筛选、结构编排、语言使用、版式设计（或文图）编绘和实物印刷中。限于篇幅，仅以“思想性”的二级指标之一“反映人类文明进步成果”为例进行阐释。

一本（套）优秀的外语类教材应该在上述六个维度上，以不同方式、不同程度反映人类在科学技术、社会制度、文化艺术等方面的进步和发展。比如，通过特定语篇的选取，练习、活动乃至配图的设计等，传递特定的信息和知识，构建特定的意义和价值，帮助学生认识到，文明是多彩、平等和包容的，正因为如此，才有了人类文明交流互鉴的价值、前提和动力，并最终理解“文明交流互鉴是推动人类文明进步和世界和平发展的重要动力”这一命题的重大意义。至此步骤，标准框架初稿已成。整体来看，这是一个自下而上进行归纳、抽象的过程。

### 3.2 标准框架的综合验证

第三步是专家论证和两轮质性验证。笔者将标准框架的一、二级指标及内涵解读提交专家论证会进行论证，并基于论证意见进行了修改和优化，形成了标准框架的修订稿。随后，笔者采用准德尔菲法（quasi-Delphi method，亦称专家调查法），邀请领域内十余名专家各自单独审读标准框架及其内涵解读，进行两轮质性验证和修订（Linstone & Turoff 1975；Powell 2003），最终形成此标准框架一、二级指标的再修订稿（见表 1）。

表 1 外语类教材评价标准框架一、二级指标

评价标准框架		教材实体的六个维度					
一级指标	二级指标	目标	内容	结构	语言	编绘	印刷
思想性	融入正确的“三观”教育	✓	✓	✓	✓	✓	✓
	体现先进的育人理念	✓	✓	✓	✓	✓	✓
	反映人类文明进步成果	✓	✓	✓	✓	✓	✓
科学性	使用准确、规范的语言、知识和信息	✓	✓	✓	✓	✓	✓
	遵循学科学习规律，适应学习者的心理认知过程和特点	✓	✓	✓	✓	✓	✓
	采纳先进的语言教学理论	✓	✓	✓	✓	✓	✓
系统性	内容结构的逻辑性	✓	✓	✓	✓	✓	✓
	学习阶段的关联性、接续性和衔接性	✓	✓	✓	✓	✓	✓
	教材本体和配套资源体系化	✓	✓	✓	✓	✓	✓
适用性	适合学习者语言学习的社会文化环境，具有文化适宜性	✓	✓	✓	✓	✓	✓
	难易度、容量适当，利学便教，教学环境和情境适应性强	✓	✓	✓	✓	✓	✓
	满足学生多元发展需求，具备开放性和延展性	✓	✓	✓	✓	✓	✓
规范性	落实相关国家课程标准和质量要求	✓	✓	✓	✓	✓	✓
	符合学科话语和学术规范	✓	✓	✓	✓	✓	✓
	物理外观（编绘、印刷等）符合国家有关标准	✓	✓	✓	✓	✓	✓

(待续)

(续表)

评价标准框架		教材实体的六个维度					
一级指标	二级指标	目标	内容	结构	语言	编绘	印刷
趣味性	激发积极情感, 促成高效学习	✓	✓	✓	✓	✓	✓
	引导、培育学科志趣	✓	✓	✓	✓	✓	✓
	教育性与审美性、愉悦性统一	✓	✓	✓	✓	✓	✓

第四步是标准框架的模型拟合量化验证, 分为整体信度验证和结构效度验证两个环节。量化验证意指标准构建的有效性验证, 是此类研究的标准研究范式中的重要一环。基于 Bachman & Palmer (1996) 的测试有用性 (test usefulness) 概念, 验证围绕标准框架的信度 (reliability)、结构效度 (construct validity)、真实性 (authenticity)、交互性 (interactiveness)、影响力 (impact) 及实用性 (practicality) 六个方面展开。

#### (1) 外语类教材评价标准框架的整体信度

信度验证分为先导调研和大规模调研两步。笔者将表 1 中的二级指标与教材实体的六个维度进行对照, 使二级指标进一步具象化, 生成三级指标, 亦即验证问卷的核心题项。问卷采用李克特五级量表, 包括应答者基本信息和 91 个题项。笔者通过“问卷星”平台投放 40 份电子问卷进行先导调研, 最终回收有效问卷 37 份, 占比 92.5%。问卷应答者来自东北、华北、华东、华南和西北五大区的高校,

平均年龄为 46.5 岁, 均为经验型和专家型教师。运用 SPSS 统计软件对问卷数据进行信度检验, 结果显示 Cronbach's  $\alpha$  系数为 0.977 ( $>0.7$ ), 表明问卷内部具有高度一致性。

根据先导调研结果和反馈, 笔者对问卷进行微调和优化, 并在“问卷星”平台进行定向投放, 问卷应答者来自七大区的 70 余所学校, 共回收有效问卷 1087 份。统计结果显示, 问卷总体信度 Cronbach's  $\alpha$  系数为 0.985 ( $>0.7$ ), 内部一致性很高。综合先导研究和大规模调研结果可知, 问卷总体信度系数分别为 0.977 和 0.985, 表明问卷的内在连贯性和一致性很强, 具有很高的信度。

#### (2) 标准框架的结构效度

笔者采用 Amos 29.0 版本结构方程模型软件对问卷数据做验证性因子分析 (confirmatory factor analysis, 简称 CFA), 并对标准框架各指标的测量结构进行模型拟合度计算, 结果见表 2。

表 2 外语类教材评价标准框架各指标模型拟合度指数

指标	CMIN/DF	$p$	GFI	AGFI	RMSEA	RMR	CFI	NFI
参考值	$\leq 5$	$> 0.05$	$\geq 0.90$	$\geq 0.90$	$\leq 0.08$	$\leq 0.10$	$\geq 0.90$	$\geq 0.90$
思想性	6.024	0.000	0.914	0.888	0.068	0.021	0.937	0.925
科学性	5.168	0.000	0.936	0.914	0.062	0.012	0.949	0.938
系统性	2.615	0.000	0.978	0.968	0.039	0.009	0.985	0.977
适用性	5.342	0.000	0.925	0.903	0.063	0.016	0.935	0.921
规范性	9.682	0.000	0.943	0.903	0.089	0.028	0.935	0.928
趣味性	4.208	0.000	0.951	0.934	0.054	0.011	0.967	0.957

根据表 2, 每个一级指标下的二、三级指标的拟合度结果表明, 六个一级指标 (潜在变量) 的理论构想均得到了数据支持, 即 CMIN/DF、GFI、AGFI、RMSEA、RMR、CFI 和 NFI 等指标均已基本达到参考范围, 所有指标的标准化回归系数均在

0.50 以上。测量六个一级指标 (一级潜在变量) 和 18 个二级指标 (二级潜在变量) 的 91 个题项与该标准框架整体呼应关系良好。综上表明, 标准框架一、二、三级指标体系模型稳定、可靠。

在第五步中, 笔者将标准框架提交项目专家组

进行质证和确认。专家组从国内外语类教材的编、选、用、评的实际情况出发,进行了两轮深入的讨论和质证,最终确定了一级指标的呈现顺序,即思想性、规范性、科学性、系统性、适用性和趣味性。这一排序体现了新时代中国特色外语教育的价值取向和发展理念。这一步骤的完成为该标准框架后续在外语类教材评价工作场景的具体应用奠定了基础。在一定程度上,该标准框架不仅可以指导外语类教材的评价工作,也为提升中国外语类教材质量提供了科学依据。

## 4 标准框架一级指标内涵的解读

限于篇幅,笔者仅就标准框架一级指标的内涵进行简要解读,关于二、三级指标的详细解读将另文探讨。

### 4.1 思想性

教材建设是国家事权。教材建设关乎国家利益,必须用正确的思想贯穿始终。外语类教材处于人文交流和意识形态的前沿阵地,与“三观”的树立、社会主义核心价值观的培养、国际视野的开拓、家国情怀的养成息息相关,同时肩负着多项时代重任,必须秉承先进的育人理念,不断落实教育的新要求,反映人类文明进步成果,实现文化传承。

### 4.2 规范性

外语类教材是国家外语教育政策与规划的载体。教材的编写和出版应符合国家课程标准、学科话语体系和学术规范,以及出版相关法规,还应体现国家对教材治理的立场与成效。教材目标与内容要与国家外语教育方针等紧密契合;教材语言应符合外语学科话语和学术规范,融入中国特色话语体系;有助于实现标准化教学和教学资源均衡分配。插图与版式设计积极健康,体现合规与美的有机结合。

### 4.3 科学性

外语类教材的语言要准确、规范、地道,选材要丰富、可靠、严谨,对语言知识、文化知识和学习策略等的呈现,要科学地组织、分级和排列,符

合学科与学术规范。教材编写要基于正确的语言观,遵循语言学习规律,关注学习者个体差异中的心理因素,充分考虑学习者的认知水平和特点,采纳先进的外语教学理论与方法。

### 4.4 系统性

教材内容结构体现语言、知识和能力发展的客观规律,编排要具有整体性和逻辑性,风格统一;内容安排时序得当、前后关联、相互照应;教材的单元、册次、教育阶段与学段之间应衔接连贯,并且与其他学科之间保持关联。教材应配有各类教辅资源,与教材本体高度契合、相辅相成。

### 4.5 适用性

外语学习过程也是文化输入和传承的过程,教材内容选择和编排应体现社会文化适宜性,体现全球视野,兼具本土意识;教材难易度设定适中,容量适当,利学便教,能切实有效地引导和帮助学生自主学习、助力教师教学,而且能满足学生的多元发展需求。教材应能够实现教学目标、充分体现教学性,并具有较强的教学环境和情境适应性。

### 4.6 趣味性

教材能够引导学生通过多种方式进行学习,关注现实和学生生活实际,激发积极情感,增强主观能动性,促成高效学习;同时引导、培育学生的学科志趣和未来发展潜力。教材的语言要流畅、优美且富有感染力,以美养德;融入艺术元素,帮助学生以审美的方式认知世界,以美启真;教材还应提升视觉审美体验,引导学生形成美的感知和创造能力,以美育美。

## 5 结语

本研究立足本土外语教育教学环境和实际,旨在构建切实可行的本土外语类教材评价标准框架,并对其可用性进行了系统验证,解决了现有标准研发中存在的问题。经过五轮综合研究,笔者构建了外语类教材评价标准框架理论模型,包含六个一级指标和18个二级指标,并对相关一级指标的内涵

进行了简要解读。该标准框架的模型拟合量化验证结果表明其具有稳定性、可靠性和可用性。

标准框架的构建不仅为外语类教材的编、选、用、评提供了科学、系统、可靠、可用的参照和支持,而且在一定程度上,为外语教育育人、育才的意义、内容、路径和方法提供了线索和蓝本。同时,该标准框架的构建也是中国特色外语教育知识体系、话语体系和学术体系创新发展的组成部分,因此,进行全面、系统、扎实的研究重要且必要。

① 本文得到教育部规划一般项目“中国英语教材评价标准研究”(项目编号:2023GH-YB-ZH-Y-02)的资助。

### 参考文献

- AbdelWahab, M. M. 2013. Developing an English language textbook evaluative checklist [J]. *IOSR Journal of Research & Method in Education* 1: 55-70.
- Bachman, L. F. & A. S. Palmer. 1996. *Language Testing in Practice* [M]. Oxford: Oxford University Press.
- Cunningsworth, A. 1995. *Choosing Your Coursebook* [M]. London: Macmillan Education.
- Ellis, R. 1997. The empirical evaluation of language teaching materials [J]. *ELT Journal* 51: 36-42.
- Fey, C.-C. & E. Matthes. 2018. Textbook quality criteria and evaluation [A]. In E. Fuchs & A. Bock (eds.). *The Palgrave Handbook of Textbook Studies* [C]. New York: Palgrave Macmillan. 157-167.
- Hutchinson, T. & A. Waters. 2002. *English for Specific Purposes: A Learning-Centered Approach* [M]. Shanghai: Shanghai Foreign Language Education Press.
- Linstone, H. A. & M. Turoff (eds.). 1975. *The Delphi Method: Techniques and Applications* [C]. Reading: Addison-Wesley.
- Littlejohn, A. 1998. The analysis of language teaching materials: Inside the Trojan horse [A]. In B. Tomlinson (ed.). *Materials Development in Language Teaching* [C]. Cambridge: Cambridge University Press. 179-211.
- McDonough, J., C. Shaw & H. Masuhara. 2013. *Materials and Methods in ELT: A Teacher's Guide* (3rd Ed.) [M]. Hoboken: John Wiley & Sons.
- Mukundan, J. 2009. *ESL Textbook Evaluation: A Composite Framework* [M]. Köln: Lambert Academic Publishing.
- Mukundan, J. & T. Ahour. 2010. A review of textbook evaluation checklists across four decades (1970-2008) [A]. In B. Tomlinson & H. Masuhara (eds.). *Research for Materials Development in Language Learning: Evidence for Best Practice* [C]. London: Continuum. 336-352.
- Mukundan, J., R. Hajimohammadi & V. Nimehchisalem. 2011. Developing an English language textbook evaluation checklist [J]. *Contemporary Issues in Education Research* 4: 21-28.
- Nimehchisalem, V. 2018. WH questions in book evaluation: Why, how, when, and who? [A]. In M. Azarnoosh, M. Zeraatpishe, A. Faravani & H. R. Kargozari (eds.). *Issues in Coursebook Evaluation* [C]. Boston: Brill Sense. 1-10.
- Norton J. & H. Buchanan (eds.). 2022. *The Routledge Handbook of Materials Development for Language Teaching* [C]. London: Routledge.
- Powell, C. 2003. The Delphi technique: Myths and realities [J]. *Journal of Advanced Nursing* 41: 376-382.
- Seaton, B. 1982. *A Handbook of English Language Teaching Terms and Practice* [M]. London: Macmillan.
- Tomlinson, B. (ed.). 2023. *Developing Materials for Language Teaching* (3rd Ed.) [C]. New York: Bloomsbury Academic.
- Williams, D. 1983. Developing criteria for textbook evaluation [J]. *ELT Journal* 37: 251-255.
- 何安平, 2020, 第七章 第一节 教材评价的意义和作用[A].载刘道义(编),《英语教材论》[C].北京:人民教育出版社. 363-366.
- 黄光扬(编), 2022,《教育测量与评价(第三版)》[C].上海:华东师范大学出版社.
- 贾蕃、沈一新, 2020,“产出导向法”理论框架下《国际人才英语教程》评价研究:教师视角[J],《外语教育研究前沿》(3): 19-26.

- 李定建、邹林大、曹守福、梁素兰, 1995, 高校教材工作评估指标体系初探[J], 《江苏高教》(S2): 53-54。
- 李政, 2022, “教材学”的建构何以可能: 基于教材媒介性的分析[J], 《全球教育展望》(1): 65-75。
- 刘道义, 2004, 浅议英语教材的评价标准[J], 《教育实践与研究》(12): 14-17。
- 刘道义(编), 2020, 《英语教材论》[C]。北京: 人民教育出版社。
- 柳叶青, 2017, 从实体思维到实践思维: 当前教材评价研究的新趋势[J], 《课程·教材·教法》(12): 24-30。
- 潘信林, 2022, 教材评价主体之间的差异及其整合[J], 《课程·教材·教法》(12): 74-79。
- 束定芳, 2024, 外语教材学: 研究目标、内容与方法[J], 《外语界》(4): 33-41。
- 孙平华, 2006, 论我国中小学英语教材的评价标准[J], 《课程·教材·教法》(4): 49-55。
- 王晓丽, 2021, 《教科书评价学》[M]。广州: 广东教育出版社。
- 文秋芳, 2023, 构建大学外语教材编写理论体系[J], 《外国语》(6): 2-11。
- 吴先强、王祖浩, 2021, 从聚焦知识到关注主体: 近三十年来美国科学教材评价工具的发展与思考[J], 《比较教育学报》(4): 150-163。
- 余宏亮, 2021, 通向根脉与面向未来: 建构教材学的基础、逻辑与方略[J], 《华东师范大学学报(教育科学版)》(2): 30-39。
- 翟志峰、董蓓菲, 2019, 美国教材评价标准的指标和方  
法——以《优质教材工具》为例[J], 《全球教育展望》(5): 91-104。
- 张莲、黄国文, 2024, 中国应用语言学研究 70 年: 回顾与反思[A]。载课题组(著), 《外国语言文学学科知识体系创新研究》[M]。北京: 外语教学与研究出版社。40-63。
- 赵勇、郑树棠, 2006, 几个国外英语教材评估体系的理论分析——兼谈对中国大学英语教材评估的启示[J], 《外语教学》(3): 39-45。
- 中国社会科学院语言研究所词典编辑室(编), 2005, 《现代汉语词典》(第 5 版)[Z] 北京: 商务印书馆。
- 中国政府网, 2023, 习近平主持中央政治局第五次集体学习并发表重要讲话[OL], [https://www.gov.cn/yaowen/liebiao/202305/content\\_6883632.htm](https://www.gov.cn/yaowen/liebiao/202305/content_6883632.htm)(2024 年 11 月 10 日读取)。
- 周雪林, 1996, 浅谈外语教材评估标准[J], 《外语界》(2): 60-62。
- 庄智象, 2006, 构建具有中国特色的外语教材编写和评价体系[J], 《外语界》(6): 49-56。

### 作者简介

- 张莲, 北京外国语大学英语学院教授、博士生导师。主要研究领域: 教师认知与发展、课堂话语分析、二语写作、教材评价。电子邮箱: zhanglian@bfsu.edu.cn
- 孙上鸮, 北京外国语大学英语学院在读博士研究生。主要研究领域: 应用语言学、外语教材、二语语音能力发展。电子邮箱: sunsy2021@bfsu.edu.cn

## Development and validation of *Vocational English Skills Standards*: Background, procedures, and features

BIN Ke, YANG Mingying, ZHANG Qiran & ZHANG Fang ..... 18

**Abstract:** The development of *Vocational English Skills Standards* (hereafter the *Vocational Standards*) is a crucial undertaking in implementing the Ministry of Education's 1+X Certificate System. It represents a new achievement in the development of proficiency scales of English for Specific Purposes. This paper elaborates on the development and validation of the *Vocational Standards* by outlining its objectives and principles, the construction of frameworks and descriptors, and the results of the two-stage validation. This paper highlights three key features in the development of the *Vocational Standards*: alignment with real workplace demands and implementation of industry-education collaboration; application of scientific methods to explore a localized approach to the development of English for Specific Purposes scales; and emphasis on precise expression to enhance its applicability in vocational English teaching. It also contends that the development and validation of the *Vocational Standards* can serve as a reference for strengthening the vocational English assessment system.

**Keywords:** vocational English skills; standards; development; validation; features

## Theoretical basis and significance of developing *Vocational English Skills Standards*

ZENG Yongqiang, HU Miao, WANG Weifang & ZHANG Fang ..... 25

**Abstract:** *Vocational English Skills Standards* (hereafter the *Vocational Standards*) was initiated in 2019 by Beijing Waiyan Online Digital Technology Co., Ltd to implement the Ministry of Education's policy of 1+X Certificate System. The *Vocational Standards* was developed through a rigorous process of development and validation in accordance with the national policies, and with the Competence Measurement (COMET) model as its theoretical basis. It includes five work fields, each containing several typical tasks, which are divided into elementary, intermediate, and advanced levels to align with distinct ability and content requirements and reflect a logic path for vocational development. The *Vocational Standards* emphasizes cultivating national commitment and an international perspective, enhancing intercultural communication skills, prioritizing real-world workplace relevance to improve foreign communication competencies, combining study with application to strengthen English language thinking, and fostering autonomous learning for lifelong growth. The *Vocational Standards* is widely applicable to teaching and to building an integrated "Vocation-Course-Competition-Certification" evaluation system.

**Keywords:** 1+X certificate; vocational education; vocational English skills; standards

## Developing a frame of reference for foreign language materials evaluation: The process and interpretation

ZHANG Lian & SUN Shangying ..... 33

**Abstract:** This paper reports on the development and interpretation of a frame of reference for foreign language materials evaluation (FRef\_FLME). Employing a mixed-methods approach, the study encompasses several steps including concept analysis and a related literature review to establish the initial framework, large-scale surveys for the refinement, a quasi-Delphi method for its further revision, and model fitting for quantitative validation. The process resulted in the FRef\_FLME which comprises six primary indicators—*intellectual depth*, *standardization*, *scientific rigor*, *systematic coherence*, *practical applicability*, and *engagingness*—and eighteen subsequent secondary indicators. Finally, the paper briefly interprets these primary indicators and offer suggestions and perspectives on the use and future application of the FRef\_FLME.

**Keywords:** foreign language materials evaluation; FRef\_FLME development; mixed-methods research; primary and secondary indicators