

# 中学英语阅读素养问卷的编制及信效度检验\*

李未来<sup>1</sup>, 康淑敏<sup>2</sup>

(1. 泰州学院教育科学学院, 泰州 225300; 2. 曲阜师范大学外国语学院, 济宁 273165)

**摘要:** 阅读素养是英语学科核心素养培育的重要内容。中学英语阅读素养评价具有鲜明的学段和年龄特点。文章依据国家课程标准, 编制了由“习惯与态度、知识与能力、文化与策略”3个维度、14个题项组成的中学英语阅读素养评价问卷, 并借助806份中学生调查数据对问卷进行了探索因子和验证因子分析, 及其信度和效度的检验。结果显示, 编制而成的中学英语阅读素养评价问卷具有较好的信效度, 可作为我国中学生英语阅读素养的测评工具。

**关键词:** 中学英语; 阅读素养; 问卷编制; 信效度

**中图分类号:** B841.2

**文献标志码:** A

**文章编号:** 1003-5184(2024)06-0559-07

## 1 引言

阅读是信息获取、知识习得的重要途径, 对于青少年而言, 尤为重要。学生阅读素养的高低直接影响着他们的学习效果和学业成就, 乃至惠及个体一生的发展。对于国家, 阅读是民族文明程度的标志, 决定着一个民族思维的深度与高度, 对于文化的传承、国家的发展有着深远的意义。近年来, 我国越来越重视“全民阅读”问题, 将推动国民阅读素养水平的提高提升到国家战略高度(李毅, 谭婷, 张睿, 2019)。阅读素养概念的内涵和外延在全民阅读的社会大背景下发展愈加丰富。在学科教学方面, 最新发布的《义务教育英语课程标准(2022年版)》也强调培养和发展学生阅读素养的重要性, 指导教师要充分认识阅读在文化传递、价值引领、思维发展、语言学习及意义理解与表达中的重要作用(教育部, 2022)。与此同时, 全社会对个体阅读素养普遍重视也引起了阅读素养评价内容趋向多元化, 不仅关注学习者智力因素, 而且愈加重视非智力因素对阅读素养的潜在影响。中学是英语教育承前启后的衔接阶段, 如何培育和评价中学生的阅读素养, 成为英语学科核心素养落地生根的关键一环。因此, 非常有必要编制符合学科核心素养要求和中学生认知特点与心理品质的阅读素养评价问卷。

## 2 相关研究述评

### 2.1 阅读素养的内涵

关于“素养”一词的定义, 《现代汉语词典(第7版)》中的解释是: “素”即本来的, 原有的; “养”即培养、修养; 素养即为平日的修养。它与“素质”一词不同在于强调后天的教育, 与“能力”一词区别于注重知识、技能、情感、态度、价值观等综合品质。显然, “素养”一词的内涵意义更加宽泛, 更能表征教育的功效和个人价值。同样, 阅读素养的概念, 也随着时代的发展而不断丰富, 已远远超越对阅读文本进行字面的解码, 更加倾向于个体的阅读能力和阅读综合品质。国外的研究机构对此展开了积极探索, 泛加拿大评估项目(Pan-Canadian Assessment Program, 简称PCAP)将阅读素养界定为“个体为实现个人发展目标、为参与社会生活, 通过理解、解释文本以及对文本做出个性和批判的回答、构建文本意义的能力”(PCAP, 2016)。经合组织(OECD)的国际学生评估项目(Program for International Student Assessment, 简称PISA)在“2018评价与分析框架”中把阅读素养的定义更新为: 阅读素养是指学生为取得个人目标、形成个人知识和潜能及参与社会活动而理解、运用、评价和反思文本的能力(OECD, 2019)。这一定义体现了阅读过程中的目标导向、

\* 基金项目: 国家社会科学基金“香港语言教育政策实施现状与对策研究”(BEA180110)。

通信作者: 李未来, E-mail: 2029308106@qq.com。

理解过程,以及与阅读文本深层含义的互动过程。在“PISA 2025 外语评估框架”中,阅读评估同样注重阅读的认知过程,但更侧重于语言知识和阅读能力的协调与运用,所以在阅读测试中采用基于学生初测表现而进行阅读任务分配的自适应评估模式(OECD,2021)。语言知识的熟悉度(包括单词认知、词义提取和句法分析)是文本意义建构的重要支撑,能有效控制低级别的无效阅读过程,因为阅读策略从母语向目的语的迁移也必须基于一定的语言知识。可见,语言知识是阅读素养评价中占有相当比重。国际教育成就评价协会(IEA)的国际阅读素养进展研究项目(Progress in International Reading Literacy Study,简称 PIRLS)在“2021 阅读评价框架”中阐述阅读素养为理解和运用社会需要的或个人认为有价值的书面语言形式的能力,读者能够从各种文本中建构意义,并通过阅读进行学习、参与阅读群体的互动与阅读赏析(IEA,2019)。PIRLS 的阅读素养定义更强调阅读目的和阅读情境,使阅读成为日常生活中不可或缺的一部分。隶属于 PIRLS 的在线阅读评估(ePIRLS)提出阅读能力不能仅限于文本信息的获取,而要包含信息的检索、理解、整合和评价的动态过程(邹建平,贺卫东,2021)。

## 2.2 英语阅读素养评价的研究现状

从国际上看,阅读素养评价研究呈两个方面的发展趋势:一是由国际知名教育评价机构(如 PISA, PIRLS, PCAP 等)领衔开展的阅读素养研究,主要对不同国家学生的阅读素养进行比较研究,进而对教育政策制定提供数据参考。此类评价除了包含阅读试题外,常常还有关于学习者背景知识的调查,涉及家庭文化素养和经济能力,以及当地的教育政策等。这些机构对阅读素养概念的发展和评测体系的架构贡献巨大,但对个体心理因素对阅读的影响缺乏足够关注。二是由专家、学者开展的研究,主要聚焦于学习者心理品质对阅读表现的影响,如阅读动机,阅读习惯,阅读态度,阅读兴趣等(Adrian, Kevin, & Roger, 2007; Skehan, 1991; Tseng & Gao, 2021),这是目前阅读素养研究的一个重要转向——人们逐渐认识到个体的心理品质在阅读素养培养过程中的重要作用,这在以英语为第一语言的国家表现尤为明显。此类研究对影响阅读的因素探索较为广泛,既有语

言认知方面,也有心理倾向方面,但整合众多因素到单个问卷调查之中的研究相对较少。事实上,阅读是读者和作者、文本三者之间的互动、交流和理解的心理过程(Colin, 2004; Eva, 2002)。因此,阅读素养评价不可仅停留在测试成绩的单一维度,而是基于语言、认知和心理等多维度并且持续进行动态调整的过程(Marilyn & Peter, 2014)。在阅读素养评价的整体框架下,具体的评价内容和方式常依据评价对象的不同而有所调整,如年龄特征、语言能力和文化背景等。

从国内看,英语阅读素养在阅读素养研究的基础上侧重与英语学科相联系,体现英语学习的学科特点。《普通高中英语课程标准(2017 年版 2020 年修订)》重新优化了原有的课程目标,提出了由“语言能力、文化意识、思维品质、学习能力”四要素构成的英语学科核心素养(教育部,2020)。《义务教育英语课程标准(2022 年版)》从“语言能力、文化意识、思维品质、学习能力”四个方面构建了课程目标结构(教育部,2022)。语言能力的发展带动着其他三要素的发展;反之,文化意识、思维品质和学习能力的发展推动着语言能力的提升。在此基础上,英语阅读素养根据技能特点衍生出不同的评价维度,如阅读策略、阅读态度等。各维度互相关联与支撑,形成有机的整体。国内不少学者对英语阅读素养评价开展积极的研究,如:王蔷和敖娜仁图雅(2015)提出,英语阅读素养应由英语阅读能力和英语阅读品格两大要素构成,其中阅读能力包含解码能力、语言知识、阅读理解和文化意识;阅读品格包含阅读态度和阅读体验;李明远和彭华清(2016)在英语阅读素养的内涵上提出了文本的预测能力、学生的推断能力和概括能力,以及与文本作者的思维互动;郭宝仙和章兼中(2016)从外语阅读能力角度构建了“两维度、三因素、六层次”的结构框架体系;张金秀(2018)从分级阅读的实证研究中发现有效阅读量、阅读频度和教师指导对学生阅读素养的影响;杜文博和马晓梅(2018)从认知诊断评估角度,构建了英语阅读诊断模型;程晓堂和陈萍萍(2019)依托大数据,从“语言知识、文化知识、阅读策略、阅读习惯”四个方面分析了英语阅读能力标准。在评价量表研制方面,王淑花(2012)编著的《中国学生

英语理解能力量表的构建及验证研究》,以及刘建达、吴莎(2019)编著的《中国英语能力等级量表研究》在英语阅读能力等级评价方面提出了具体的参照指标,对今后英语阅读素养评价研究具有重要的理论与实践意义。

总体而言,国内外英语阅读素养相关研究的视角多样、观点新颖,但在英语阅读素养评价问卷编制方面的研究却少有涉足,尤其是针对中学生的英语阅读评价研究则更少。即使在英语作为第二语言的香港地区,也缺乏中学阶段的英语阅读素养评价问卷。鉴于此,研究依据国家课程标准,尝试编制符合英语学科核心素养要求和中学生认知特点的英语阅读素养评价问卷。

### 3 问卷编制过程与方法

#### 3.1 英语阅读素养评价维度分析

中学英语阅读素养评价问卷的编制主要基于以下两方面的思考。一是基于英语学科的核心素养要求以及阅读教学实践发展需求。英语学科的核心素养涵盖“语言能力、文化意识、思维品质、学习能力”四个维度,超越了传统的学科知识和能力的意蕴,既包括掌握学科知识的基础性学科素养,又包括合理运用学科知识与方法解决实际问题的能力素养以及尊重客观规律、追求真理的严谨态度与科学精神的品格素养(康淑敏,2016)。以阅读理解为主要评价内容的传统评价方式,已无法适应信息化时代英语阅读教学的需要。此外,随着全民阅读理念的推行,英语阅读素养的评价也势必跳出课堂的束缚,在场域上延伸至学校、家庭、社会等不同的阅读群体空间。评价维度应该涉及阅读行为、阅读习惯、阅读态度、阅读品格等因素。再者,英语作为了解世界文化、传播中国文化的语言工具,文化意识、思维品质也是阅读素养评价中的重要维度。二是充分考虑中学生的英语学习特点和阅读方式。中学生认知活动的自觉性明显增加,他们的观察力、注意力,以及记忆、思维等心理要素迅速发展。英语学习的内容和评价方式,自然需要符合青少年的语言认知特点。中学生阅读能力发展的重要部分就是对字词概念的理解与运用,他们在知识经验基础上能够较好地分析抽象概念和理解语篇信息。另外,中学英语阅读属于“学术阅读”阶段,以探究性阅读为主,侧重通

过阅读培养学生的逻辑思维和归纳总结能力,以及对篇章、文体、背景等知识的指导;在中学阶段,英语阅读材料以研究型、论证型内容为主,主要包括对文章主题、段落、层次、技巧的理解,学会接纳不同的观点、养成开放的心态,促进认知能力发展(香港课程发展议会,2017)。因此,英语阅读素养评价问卷题目的设计必须充分考虑上述因素。

#### 3.2 初始问卷编制

参照国家课程标准和已有的研究成果,在条分缕析多方影响因素后,研究从6个维度、20个评价指标界定了英语阅读素养评价框架,即:语言知识(词汇、语音、语法)、阅读技能(猜词义、拼读、流畅度、阅读速度)、阅读策略(信息提取、语篇理解、多元思维、方法选择)、文化意识(感知、体验、鉴赏、自觉)、阅读习惯(兴趣、细节、自我评估)、阅读态度(频率、时间、阅读量)。调查问卷设计包含了被试对象的年级、性别、年龄等变量因素;20个题目以陈述句形式呈现,采用“李克特五级量表”划分方式,每一题包含“完全符合、较好符合、基本符合、勉强符合、完全不符合”五种选项。

#### 3.3 初始问卷预试与数据收集

初始问卷预试的主要目的是检验学生对题项语义的理解、答题时间,以及问卷内容和结构是否合理且符合实际等。在预试前,研究对部分中学英语教师和学生进行了访谈和咨询,从教和学两个角度探寻英语阅读素养评价的合理方式,并对题项的表述进行适当修改。研究选取江苏省M中学学生为被试,被试年龄分布主要集中于12~15岁,在最近一次的年级统一英语测试中,他们的英语成绩呈正态分布。预试阶段共发放问卷450份,收回问卷440份,其中无效问卷10份。有效问卷440份按七、八、九三个年级随机被平均分为A、B两组:A组(220份)进行探索性因子分析检验(EFA),B组(220份)进行验证性因子分析检验(CFA)。探索性因子分析后,初试问卷的20个题目被删减至14个。然后,采用优化后的问卷和《中国学生英语等级量表》中的阅读理解能力3~4级量表,同时对同地区的H中学的学生进行问卷调查,此次调查收回两类有效问卷各366份。两组调查数据进行相关性分析检验。《中国学生英语等级量表》是经过权威部门验证,信

效度能得到保证,可作为该评价体系效度验证的重要参照。数据分析工具采用软件 SPSS 26.0 和 Amos 23.0 的版本。

## 4 研究结果

### 4.1 题目分析

题目分析旨在检测问卷题目的趋同性和区分度。研究采用预试样本中 440 份有效数据进行分析。结果显示,所有题目的标准差在 0.949 ~ 1.163

之间,满足题目得分差异性的要求;所有题目得分与总得分的相关系数有统计学意义( $r = 0.583 \sim 0.740, p < 0.01$ )。然后,研究将数据按总分进行降序排列,取前 27% 作为高分组,后 27% 作为低分组,进行了数据正态性检验,由于数据不符合正态分布,所以采用了非参数秩和检验,结果显示,所有题目得分在高、低分组上均有统计学意义  $p < 0.01$  (见表 1)。

表 1 高低分组差异非参数秩和检验及题总相关分析

题目	$M \pm SD$	$z$	$r$	题目	$M \pm SD$	$z$	$r$
1	$4.01 \pm 1.019$	-9.745	0.583**	11	$3.75 \pm 1.125$	-11.692	0.684**
2	$3.98 \pm 0.949$	-11.036	0.672**	12	$3.80 \pm 1.036$	-10.504	0.640**
3	$3.89 \pm 0.950$	-9.768	0.609**	13	$3.88 \pm 1.027$	-11.436	0.711**
4	$3.71 \pm 0.999$	-10.496	0.658**	14	$3.80 \pm 1.060$	-11.750	0.673**
5	$3.45 \pm 1.134$	-10.607	0.583**	15	$3.85 \pm 1.060$	-10.379	0.645**
6	$3.48 \pm 1.163$	-11.095	0.570**	16	$3.85 \pm 1.055$	-10.113	0.621**
7	$3.82 \pm 1.086$	-10.563	0.631**	17	$3.96 \pm 1.066$	-10.285	0.636**
8	$3.85 \pm 1.100$	-11.070	0.659**	18	$3.76 \pm 1.073$	-11.199	0.664**
9	$3.70 \pm 1.061$	-11.206	0.665**	19	$3.74 \pm 1.108$	-10.977	0.625**
10	$3.68 \pm 1.130$	-12.268	0.740**	20	$3.74 \pm 1.089$	-11.510	0.710**

注:\*\* $p < 0.01$

### 4.2 效度分析

#### 4.2.1 内容效度

内容效度是指问卷题目对相关概念测量的适切性、评价维度和评价指标的一致性,以及题项表述是否能涵盖评价内容的完整性。内容效度主要采用专家判断的方式。研究过程中邀请了英语教学方面的专家、中学英语教研员和一线英语名师对问卷进行了检验,对发现的问题进行了适当修改。最后,参与鉴定的所有专家一致认为,中学英语阅读素养评价问卷的所有题目符合 6 个维度、20 个指标的要求,能充分反映中学生英语阅读素养的全部内涵。

#### 4.2.2 结构效度

结构效度是指测量题目与测量变量之间的对应关系,主要目的是探寻问卷的结构是否合理,测量方法有两种:探索性因子分析和验证性因子分析。

探索性因子分析的目的是通过测量因子载荷,探索因子对潜在评价维度的贡献度,进而决定对部

分因子进行删减。研究数据采用预试对象中的 A 组 220 份调查数据。首先,在分析之前,研究进行了因子分析的可行性检验。结果显示,KMO 统计量为 0.877,大于 0.7;巴特利特球度检验的  $p$  值为 0,表明数据适合进行探索性因子分析。其次,研究利用主成分分析法,因子旋转采用最大方差法。首次旋转后,提取了四个公因子,累积解释变量为 56.397%,并且多个变量均有共有因子,相关度比较接近。经过多次因子旋转,最终删除了 6 个题项,保留了 14 个题项。研究最终提取了三个公共因子,分别命名为:习惯与态度、知识与技能、文化与策略。三个因子的累积贡献率达 61.748% (见表 2)。旋转后各成分的相关度高,题目的负荷值在 0.507 ~ 0.911 之间。因子命名与题目内容较为符合,反映出因子名称所表达的范围 (见表 3),表明问卷有较好的结构效度和内容效度。

表 2 解释的总方差

成分	初始特征值			旋转载荷平方和		
	总计	方差百分比	累积%	总计	方差百分比	累积%
1	5.873	41.948	41.948	3.616	25.827	25.827
2	1.534	10.959	52.907	2.587	18.475	44.302
3	1.238	8.840	61.748	2.442	17.445	61.748

表 3 评价体系的维度与内容

题号	维度	题项内容	负荷值
1	习惯与态度	我能每天坚持阅读英语课外读物。	0.695
2		我会经常做英语课外阅读笔记。	0.649
3		我乐意通过英语阅读了解更多的人和事。	0.729
4		我会积极参与学校组织的阅读活动。	0.761
5		我会认真完成老师布置的阅读任务。	0.735
6	知识与能力	我能够预测语篇的内容发展和人物关系。	0.740
7		我能够根据语篇提示推断作者的情感和态度。	0.774
8		我能借助构词法知识推测语篇中生词的含义。	0.679
9		我能利用表格、思维导图等工具梳理语篇信息。	0.636
10	文化与策略	我能够在理解语篇内容基础上表达个人看法。	0.507
11		我会根据需要采取不同的英语阅读方法。	0.897
12		我喜欢阅读有关中国文化的英语课外读物。	0.525
13		我了解英美国家的语言文化和风俗习惯。	0.911
14		我会在阅读中有意识地比较中西方文化的异同。	0.765

验证性因子分析的目的在于检验问卷的理论结构模型与实测数据的拟合程度。研究借助 M 中学被试的 B 组 220 份调查数据,采用 AMOS 23.0 统计软件,依照旋转后采用 3 个公共因子和 14 个题项,重新整合、导入数据,进行验证性因子分析,分析结果见图 1。运算结果显示,显著性  $p$  值为 0;CMIN/DF = 2.363, RMP = 0.062, GFI = 0.889, AGFI = 0.843, NFI = 0.879, IFI = 0.926, RMSEA = 0.079。综合各项指标,发现该模型与数据拟合度较好,问卷有很高的结构效度。

4.2.3 效标效度

效标效度是指通过实测数据与效标之间关系的分析,检验问卷编制质量的一种指标。研究的效标效度检验采用相关法。研究利用优化后的评价问卷和《中国学生英语等级量表》中的阅读理解能力 3 ~ 4 级量表同时对同地区的 H 中学进行调查,分别收集两组调查数据,每组有效问卷 366 份,并利用 SPSS 26.0 软件对其进行了相关分析。研究理论假设为:中学生的英语阅读素养和阅读理解能力呈正相关。结果显示,英语阅读素养和阅读理解能力的相关系数为 0.867,  $p$  值为 0,两者呈高度正相关,由此得出,优化后的评价问卷具有较好的效标效度。

4.3 信度分析

信度分析的目的在于检验样本数据的真实可靠性。研究借助 SPSS 26.0 软件,利用 H 中学的 366

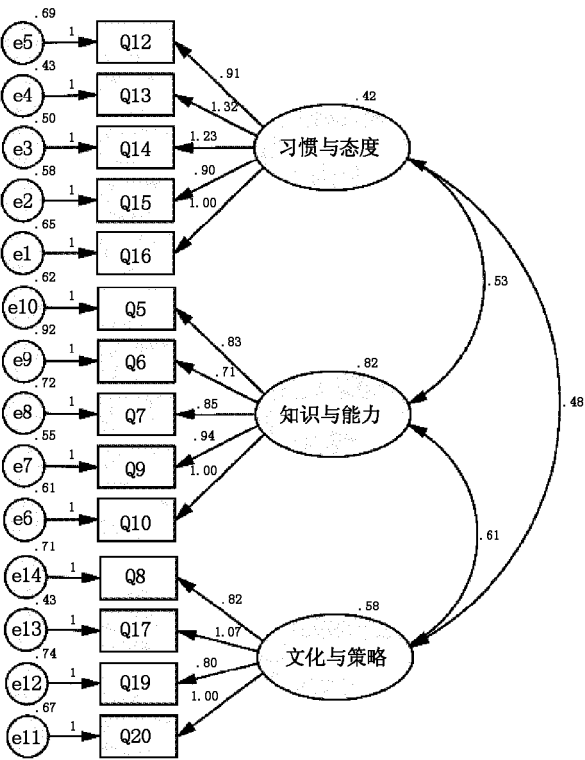


图 1 中学生英语阅读素养问卷结构模型

份调查数据,计算问卷各维度及总问卷的克隆巴赫内部一致性系数。结果显示,基于标准化项的克隆巴赫  $\alpha$  系数为 0.903,大于 0.7,并且各维度的  $\alpha$  系数在 0.763 ~ 0.818 之间,证明各维度与总问卷内部

一致性水平较高,具有较高的可信度(见表4)。

表4 问卷的信度系数

维度	题目个数	Cronbach's $\alpha$ 系数
习惯与态度	5	0.763
知识与技能	5	0.810
文化与策略	4	0.818
总问卷	14	0.903

## 5 讨论

研究编制的中学生英语阅读素养评价问卷包含3个评价维度,14项具体评价指标。各维度和指标相互拟合,结构清晰,问卷具有良好信效度,为中学生英语阅读评价提供了有效的测量工具,具有一定的理论意义和实践价值。首先,从理论上讲,英语阅读素养是一个多维、发展的概念。随着社会的发展和阅读介质的拓展以及阅读方式的变化,其内涵和外延将不断丰富。同样,英语阅读素养的培育,虽然是一种循序渐进的过程,但不是遵循从单维到多维的发展轨迹,而是呈多元并蓄、不断积累和循环往复的发展态势。所以,英语阅读素养的评价内容和方式也需相应调整。研究中阅读素养评价维度的构建正是在语言评价理论的指导下,把个体心理倾向纳入了评价框架,其评价的多维性和动态性特征进一步验证了二语阅读过程中心理评价机制与语言认知机制相结合的必要性。其次,从实践上看,中学英语阅读素养评价问卷不仅可以调查一定群体的阅读素养水平,更重要的是可以作为阅读量化测评分析时的参照依据,如通过与阅读素养水平值或各维度的具体得分的相关性分析,进一步查明学生阅读成绩变化后的深层原因,尤其是心理因素对阅读表现所产生的影响。此外,中学英语阅读素养评价问卷的实践运用,可以根据学生的个性特征和实际需求,适时添加相应的背景信息或调整个别指标,使之更加适合不同条件下中学生英语阅读素养的评价。

## 6 研究结论

英语阅读素养评价问卷编制是一项复杂的、动态的系统工程,需要考虑英语本身的学科特点、阅读素养的发展现状,还要涉及不同学段学生的年龄特征。研究以英语阅读素养为主题,以中学阶段为研究范围,参阅了国内外的相关研究,依据课程标准和中学生的认知特点,编制了中学英语阅读素养评价问卷,并得出如下结论:一是英语阅读素养评价要注重阅读过程的评价,以及在此过程中阅读者非智力因素的影响,其评价内容既涉及到第二语言的习得规律,也包含阅读者母语的阅读品质。二是英语阅读素养评价要考虑到不同地区、不同学段的具体实际。国外的阅读素养评价研究成果可以借鉴,但不

可生搬硬套。大学阶段和小学阶段的英语阅读素养评价也与中学阶段有较大差异。三是研究所编制中学生英语阅读素养评价问卷(包含3个维度和14项评价指标)具有较好的信度和效度,能客观地反映中学生的英语阅读素养,可作为我国中学英语阅读素养的评价工具。

## 参考文献

- 程晓堂,陈萍萍. (2019). 基于大数据的英语阅读能力培养及测评体系构想. *外语电化教学*, 41(4), 40-44, 60.
- 杜文博,马晓梅. (2018). 基于认知诊断评估的英语阅读诊断模型构建. *外语教学与研究*, 62(1), 74-88, 160-161.
- 郭宝仙,章兼中. (2016). 构建我国中小学外语阅读能力的结构框架体系. *课程·教材·教法*, 36(4), 23-26.
- 教育部. (2020). *普通高中英语课程标准* (2017年版2020年修订). 北京:人民教育出版社.
- 教育部. (2022). *义务教育英语课程标准* (2022年版). 北京:北京师范大学出版社.
- 康淑敏. (2016). 基于学科素养培育的深度学习研究. *教育研究*, 38(7), 111-118.
- 李明远,彭华清. (2016). 英语阅读教学中学生学科核心素质的培养. *教学与管理*, 33(10), 58-60.
- 李毅,谭婷,张睿. (2019). 提升中小学生学习素养的实证研究:基于群文阅读与单篇阅读教学的对比. *教育研究与实验*, 37(4), 31-37.
- 刘建达,吴莎. (2019). *中国英语能力等级量表研究*. 北京:高等教育出版社.
- 王蕾,敖娜仁图雅. (2015). 中小学生外语阅读素养的构成及教学启示. *中国外语教育*, 8(1), 16-24, 104.
- 王淑花. (2012). *中国学生英语理解能力量表的构建及验证研究*. 北京:知识产权出版社.
- 香港课程发展议会. (2017). *从阅读中学习:迈向跨课程阅读*. [https://www.edb.gov.hk/attachment/tc/curriculum-development/major-level-of-edu/secondary/curriculum-guides-documents/SECG\\_booklet\\_6B\\_ch.pdf](https://www.edb.gov.hk/attachment/tc/curriculum-development/major-level-of-edu/secondary/curriculum-guides-documents/SECG_booklet_6B_ch.pdf).
- 张金秀. (2018). 运用分级阅读培养中学生英语阅读素养的实证研究与启示. *课程·教材·教法*, 38(7), 73-80.
- 邹建平,贺卫东. (2021). 国际阅读素养进展研究项目中的在线阅读评估:框架、特点及启示. *上海教育科研*, 41(5), 68-74.
- Adrian, S. C., Kevin, W. L., & Roger, G. S. (2007). Method and psychological effects on learning behaviors and knowledge creation in quality improvement projects. *Management Science*, 53(3), 437-450.
- Colin, H. (2004). Postmodern principles for responsive reading assessment. *Journal of Research in Reading*, 27(2), 163-173.
- Eva, G. (2002). Toward dynamic assessment of reading: Applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading*, 25(3), 283-298.

- IEA. (2019). *PIRLS 2021 Assessment Frameworks*. Boston: TIMSS & PIRLS International Study Center.
- Marilyn, K., & Peter, K. (2014). *Literacy and Language in East Asia*. Singapore: Springer.
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. Paris: OECD Publishing.
- OECD. (2021). *PISA 2025 Foreign Language Assessment Framework*. Paris: OECD Publishing.
- PCAP. (2016). *Pan – Canadian Assessment Program (PCAP) 2016 Assessment Framework*. Toronto: Council of Ministers of Education, Canada.
- Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition*, 13(2), 275 – 298.
- Tseng, W. T., & Gao, X. (2021). Individual differences in second language learning: The road Ahead. *English Teaching & Learning*, 45(3), 237 – 244.

## Development, Reliability and Validity of a Questionnaire for English Reading Literacy in Middle Schools

Li Weilai<sup>1</sup>, Kang Shumin<sup>2</sup>

(1. School of Educational Sciences, Taizhou University, Taizhou 225300;

2. College of Foreign Languages, Qufu Normal University, Jining 273165)

**Abstract:** Reading literacy is an important part of the cultivation of English core literacy. The evaluation of middle school students' English reading literacy has distinct characteristics of learning stage and students' age. Referring to the national curriculum standards, this research firstly develops a questionnaire for English reading literacy in middle schools, with 6 dimensions and 14 items. The 806 data from two middle schools were used to validate the questionnaire. The questionnaire was analyzed through exploratory factor analysis, confirmatory factor analysis, reliability and validity test. The results shows that the questionnaire has high reliability and validity, which can be used as an evaluation instrument to assess students' English reading literacy at an intermediate level in China.

**Key words:** middle school English; reading literacy; evaluation questionnaire; reliability and validity

(上接第 558 页)

## The Impact of Parameter Estimation Error on IRT Linking Methods with Polytomous Items

Wang Shaojie<sup>1</sup>, Zhang Minqiang<sup>2</sup>, Huang Feifei<sup>3</sup>, Liu Ying<sup>4</sup>

(1. School of Education, Guangdong University of Education, Guangzhou 510303; 2. School of Psychology,

South China Normal University, Guangzhou 510631; 3. School of Educational Science, Guangdong

Polytechnic Normal University, Guangzhou 510665; 4. School of Teacher Education,

Guangdong University of Education, Guangzhou 510303)

**Abstract:** The information – weighted characteristic curve methods have shown excellent performance in IRT linking with dichotomous items. However, few researches explore the effect of parameter estimation error on test linking. This paper extends the information – weighted characteristic curve methods to IRT linking with polytomous items and explores the effects of parameter estimation error, ability differences, and test length on linking through simulation studies. IRT linking performance was evaluated using indices related to characteristic curves and errors. The results indicated that the information – weighted characteristic curve methods performed slightly better than the traditional characteristic curve methods, while other new methods performed as well as the traditional methods. The linking performance was better when the parameter estimation error and ability differences were smaller, and the test was longer. The bias and variance tradeoff provides a new direction for test linking and equating.

**Key words:** Parameter estimation error; polytomous item; IRT linking; information weighted; characteristic curve methods