

【专题·基于核心素养的评价】

编者按:《义务教育语文课程标准(2022年版)》明确指出,义务教育语文课程培养的核心素养是语文课程总目标及课程价值的集中体现。本期专题精选5篇文章,从理论层面探讨基于核心素养的课程评价理念变革、评价模型构建、综合性试题特点,从实践层面探索指向问题解决过程的测试工具开发和综合性、实践性试题命题策略,以期对读者朋友深入理解和把握素养导向的评价改革有所启迪。

义务教育语文核心素养评价模型的建构与实践应用

李倩 郑国民

【摘要】依据课程标准研制语文核心素养评价模型,推进语文考试评价改革是促进基础教育发展的重要议题。参考核心素养取向的国际大型教育评价项目的实践经验,以《义务教育语文课程标准(2022年版)》为依据,建构以“情境—语文实践活动—语文课程大观念”为核心的语文核心素养评价模型,其中:情境主要包括日常生活、文学体验、跨学科学习3种类型,语文实践活动是指为了达成语文课程学习目标、学生自主开展的运用语言文字的实践行为,语文课程大观念是学生通过语文实践活动获得的对语言文字运用本质规律的认识。该模型以真实而复杂的情境为依托,借助典型且综合的语文实践活动,系统评估学生对语文课程大观念的理解,可用于命题质量评估、语文过程性评价设计与实践、语文教育质量监测与提升等。

【关键词】义务教育课程标准;核心素养;学业评价;学业质量;试题情境;过程性评价

在核心素养立意的教育改革背景下,《义务教育语文课程标准(2022年版)》(以下简称《语文新课标》)重构了语文课程内容与结构体系,强调以语文学习任务群为载体,引导学生在积极的语言实践活动中逐步发展正确的价值观念、必备品格和关键能力。课程内容与教学方式的改革,进一步推动了义务教育语文课程评价内容与方式的改变。近年来,随着教育评价改革进程的不断深入,基础教育在评价理念、评价实施路径等方面取得丰富成果。与此同时,越来越多的研究者意识到,课程评价尚存在过于重视量化评分,忽视对多种评价模式的统筹问题^[1]。就语文课程而言,学习目标的整体性与学习内容的包容性,以及学习体验的主观建构性等特征,无疑进一步提升了语文学业质量评价的设计与实施难度。在这样的背景下,在以证据为中心(evidence-centered design, ECD)的教育评价设计理论的基础上,本研究充分吸收国际教育评价改革先进经验,以《语文新课标》为根本依据,尝试构建可应用于过程性评价、学业水平考试的语文核心素养评价模型。

一、核心素养立意的国际教育评价的研究成果与实践经验

(一)价值取向:以证据为中心的科学推论

对教育评价的理解持有怎样的观念是开展教育

评价实践的起点。一直以来,对这一问题的认识有2种代表性理念,分别为以构念为中心(construct-centered design, CCD)、以证据为中心(ECD)。在价值取向方面,CCD理论重视对评价的复杂知识、技能或特质内涵与本质属性的界定^[2],而ECD理论则更为强调评价要以学生真实学习表现为基础,对其内在特质发展情况或水平进行推论的过程,有效采集与分析证据是教育评价实践的关键^[3]。近年来,因在评价设计与实施环节的创新性探索,ECD理论逐渐成为核心素养测量与评价设计的重要理论基础。

就评价设计而言,ECD理论认为评价框架至少应包括学生模型(阐述“评什么”)、任务模型(描述“怎么评”)、证据模型(说明任务模型如何为学生模型提供证据)。以此作为理论基础,教育评价研究者展开了评价模型建构与具体学科领域评价2个方面的实践探索。Pellegrino基于“评价是基于证据的推理过程”的核心理念,建构了过程性评价与终结性评价都适用的评价三角模型。该模型包括认知、观察、阐释3个要素,其中:认知主要指向证据的理论基础,评价者要明确评价对象的认知发展规律及特定学科领域的学习规律;观察集中解决如何获取与分析证据的问题;阐释重点在于如何理解与运用评价证据,即评价设计者基于证据的分析与揭示,制定教

育教学决策^[4]。王慧君等构建了证据导向的高校混合学习课程学业评价框架,主要包括识别证据、创设证据、采集证据、鉴定证据、分析证据、解释证据、使用证据^[5]。美国教育考试服务中心(ETS)以美国共同标准(Common Core State Standards)为参照,研制了基于认知的评价系统,并在英语语言艺术领域评价中首次提出“基于关键实践的评价”,即针对不同学习主体、学习主题、语言材料而设计的情境性语言实践活动系统^[6]。张娜等以ECD理论中的概念评价框架为依据,研制了合作问题解决能力的评价蓝图^[7]。通过梳理上述实践成果发现,ECD的优势至少表现在4个方面:第一,适用于对内涵丰富、结构复杂的认知技能、素养的评价^[8];第二,突出真实情境的重要价值,强调动态交互式的评价任务,有效采集学生复杂的外在表现;第三,基于复杂数据分析方法与学生模型建立复杂的关联,提升评价的有效性;第四,可广泛应用于终结性测验和教学过程性评价,有利于解决评价功能单一化问题,实现评价促进学习、评价改进教学的功能。

(二)理论转型:根植于真实社会文化的核心素养

明确评价构念的基本内涵是教育评价设计的关键环节。评价构念是评价者对“什么是学习者内在特质”这一问题的理论假设与核心观点。一直以来,语言学、心理学与学习科学的理论发展,持续革新着语文课程设计理念与教学实施模式,也深刻影响着评价构念的阐述。

20世纪中后期,人类语言文字实践研究的视角从描画“现象”转向“心智”解释^[9],即强调从信息加工视角探究语言学习的内在认知机制。具体来说,认知主义理论集中探究人类语言学习所经历的关键环节,并以线性逻辑阐释语言学习的加工过程^[10]。以阅读素养为例,成熟读者通常要经历信息检索与定位、筛选与整合、分析与推论、鉴赏与评价等认知环节。毋庸置疑,上述研究打开了人类语言实践研究的黑箱,也为评价构念的建构、评价框架的设计提供了重要的理论基础,但研究者也逐渐认识到过度内向性探究的弊端,即忽视社会文化环境因素对语言实践发生与发展的影响并不可取。以Hymes为代表的语言学家认为语言的本质属性并非“符号集合”^[11],而是社会文化实践的产物。同时,Vygotsky也更为关注认知发展与社会文化的内在关联,即有意义的学习行为和知识运用都应在特定文化环境中进行^[12]。综合来看,上述观点均表明语言并非是脱离外在环境的抽象符号,而是扎根于社会文化背景中的意义系统,对语言实践的评价应充分考虑个人

环境及其互动关系。在这一观念的影响下,社会文化理论逐渐成为国际语文教育评价构念建构、模型设计的重要理论基础。PISA将阅读素养视作“为实现个人目标、丰富个体的知识与潜能,参与社会活动的根本目的而理解、运用、反思评价并参与文本的过程”^[13];PCAP认为阅读素养即“通过理解、解释文本,并对文本内容作出个人的、批判性的回应,从而理解世界、参与社会的过程”^[14]。

(三)评价模型:拓展内在构成,调整互动关系

在价值取向与理论转型的共同作用下,教育评价模型设计发生结构性变革,具体表现在“评价模型应容纳哪些要素”“评价模型内部要素之间呈现怎样的互动关系”2个方面。

就评价模型构成要素而言,基于个体与社会文化环境互动理念,应丰富“学习主体—学习材料”二维评价模型。以阅读素养评价为例,阅读通常被视作个体与文本的单向互动过程。基于此理念,阅读素养评价模型多采用二维矩阵模式^[15],一是不同类型的文本材料,二是学习者理解文本材料所经历的认知过程。然而,真实情境中的阅读素养是读者、文本、社会文化情境、实践活动等相关要素综合作用的结果。据此,丰富与拓展阅读素养评价模型的内在构成要素成为必然趋势。如RfU项目从语言文字、句式语法、篇章结构、概念整合推理、社会背景及交际等多个维度界定阅读素养,建构了容纳文本(类型、来源、数量)、情境、认知发展模式、认知技能等多维度评价模型^[16]。

此外,改变评价模型内部构成要素之间的关系是教育评价变革的重要方向。受现实因素影响,评价模型内部要素一直以来只呈现了简单的线性对应关系。以阅读素养评价为例,由于要综合考虑评价的操作条件与现实因素(如测试效率、年度稳定性等)^[15],PISA与NAEP等大规模测试项目延续了已有的“文本类型—认知要素—任务”一一对应的模式。换言之,评价任务指向了学生阅读某一特定文本时所经历的某一项认知能力,这与阅读的真实发生规律存在明显矛盾。针对该类问题,研究者尝试在评价模型内部呈现要素之间的复杂互动关系,如CBAL项目的评价模型包含认知加工、认知表征、关键实践和学习进程等核心要素,其中学习进程将关键实践细化为不同的学习任务,每个任务涉及的不再是零散的技能点,而是特定的读写活动和认知技能组^[6]。从真实的语言实践活动出发,各个任务之间可能涉及某些相同的读写活动,调用同样的认知技能。为此,CBAL评价模型中的认知技能通常以组合方式出现,即认知技能与关键实践活动、认知表征模式及评价任务等维度分别呈现多

对一关系,且各个关键实践、认知表征模式、评价任务之间还存在认知技能交叉、重叠的情况^[17]。

整体来看,核心素养取向的教育评价有其基本逻辑,即立足“基于证据开展有效推理”的思想,尊重核心素养的本质属性、形成与发展规律,建构多维评价模型。系统梳理国际教育评价变革趋势,其根本目的是服务我国语文课程评价体系建设。在建构义务教育语文核心素养评价模型时,至少须关注以下3个方面:(1)以学生真实语文学习规律为根本依据,筛选评价模型的关键构成要素;(2)丰富评价模型的内在构成维度,重视复杂情境与实践的重要价值;(3)依据核心素养表现的复杂性与多样性,调整评价模型内在构成维度之间的内在关联。

二、义务教育阶段语文核心素养评价模型的建构

(一)构建评价模型的基本原则

核心素养立意的教育评价力图借助思想观念、设计思路、实施方法的革新,消解评价对教学实践的负向回冲效应,引导教师、学生将更多的精力放在对学科现象与问题的深度探究,进而实现评价促进学习、改进教学的目的。据此,语文核心素养评价框架的建构应处理好评价与课程标准、学生真实语文学习和语文教学实践3个方面的关系。

首先,以课程标准为根本依据,准确把握核心素养的内涵与特征。课程标准对教学实践、考试评价具有统领意义,评价模型应与课程理念、课程目标等相关要求保持内在一致。构建评价模型的重要前提是准确把握核心素养的本质属性和特点,明确其形成与发展机制。从构成要素来看,核心素养既包含可量化、可分解的知识和能力,又包含高度关联、内隐性的思维品质、价值观念、情感态度、文化修养等。持续性发展是核心素养结果形态的外在表征。素养并不是固体性、可以占有的东西,而是一种不断发展的动态过程^[18]。这意味着不能仅从“结果”角度确定具体学习内容,还要关注学生对自我发展的持续性投入,即是否能自主发掘有价值的学习问题,是否能对有价值、有意义的问题开展持续性研究等。此外,实践活动是核心素养发展与显现的必要条件。在理解人类实践活动的复杂性时,研究者关注到真实情境、开放任务、实践共同体等方面,这些基本要素不仅成就了核心素养的发展,也促成了核心素养的外显形态。

其次,以语文学业质量为重要参照,落实以学生为中心的理念。学业质量是教育评价设计与实施的重要指南。语文学业质量以核心素养为主要维度,

是结合课程内容对“真实学习者”应具备的语文学业成就的整体刻画。同时,语言文字运用情境、语文实践活动也是组织与呈现语文学业质量的关键要素,这为解答“从哪些维度评价,以及设计怎样的任务评价核心素养发展水平”这一关键问题提供了明确方向。换言之,学业质量能够让评价者明确学生会怎样的情境下、面对不同类型的学习内容与材料、具有怎样的外在表现。在这样的前提下,评价模型的建构要从“学生有哪些典型表现”“命题要引出学生哪些典型表现”等方面进行综合考量。

最后,尊重语文教学实践规律,体现典型教学理念与实践经验。一直以来,在设计评价模型与评价工具时,评价对教学的导向作用往往容易受到关注。但是,通过梳理教育评价内涵会发现,并非评价决定教学,而是教学决定评价。评价是根据学习目标与学生学习规律,运用多样化的评价方法,系统收集与分析学生学业表现的相关信息,进而推动学生发展性学习并提升教学质量的过程。从该内涵出发,语文核心素养评价应与语文教师日常教学实践保持高度一致,这种一致性应表现在目标的确定与内容的选择上,更应体现在实践模式层面,即充分体现语文教学实践的内在规律与典型经验。近年来,项目式学习、专题学习、研究性学习、任务型教学、体验式活动链等教学实践模式备受关注。虽然上述教学实践成果的实践路径不同,但设计理念殊途同归,即都强调学生采用自主实践与合作探究的方式对语文学科的本质问题开展持续探究。换言之,核心素养取向的语文教学重在引导学生对某一领域相关问题或现象的持续钻研,学生要通过解决一连串问题,完成对核心问题的探究。因此,评价模型的建构应充分尊重教学实践规律,融入典型的教学理念与实践经验,为核心素养取向的教学变革提供有力支持。

(二)评价模型的基本要素

以ECD模型为理论基础,依据《语文新课标》建构义务教育语文核心素养评价模型。该评价模型包括评价观念建构、评价任务设计、评价任务验证3个组成部分,可以集中回应评什么(要获得哪些方面的证据)、怎么评(运用哪些适宜的方式获得学生表现相关证据)2个关键问题。其中:评价观念建构模块主要解决“评什么”问题,即从评价角度界定语文核心素养的基本内涵与测评指标体系,并在此基础上进一步明确情境与语文实践活动的类型、语文课程大观念的基本范畴;评价任务设计模块集中解决“怎么评”的问题,即从具体任务设计的角度确定运用哪些情境,设计怎样的语文实践活动,综合考查学生对

语文课程大观念的理解;评价任务验证模块重点解决“评价任务设计是否适宜”的问题,以语文学业质量为工具,检验评价任务是否紧扣语文核心素养的基本内涵、是否符合相应学段学生语文学业成就的基本特点与发展规律。

由于《语文新课标》已经清晰界定了语文核心素养、语文学业质量的基本内涵,以下仅阐述情境、语文实践活动和语文课程大观念3个基本要素的操作性定义与基本构成要素。

1. 情境

真实而富有意义的情境是语文核心素养形成与发展的必要条件。有研究者认为,情境主要是指评价任务或问题的发生背景或依赖的因素、条件,具体表征为语言材料、图片、表格,以及任务涉及的具体条件或事件^[19],其功能在于调动学生对测试任务的参与程度,赋予学生完成任务、解决问题目的、立场、条件或影响因素,为系统而全面地了解学生的真实思维过程和学习成果、保证评价结果有效性等提供外在支撑。在本研究中,情境主要包括日常生活、文学体验、跨学科学习3种类型。日常生活情境集中指向真实而具体的日常生活场景,具体表现为发生在社会公共场所、家庭生活的事件或实际问题等。文学体验情境侧重指向个人情感体验得以形成与发展的文学阅读场景,如借助不同作品体悟社会主义先进文化、革命文化、中华优秀传统文化,以及积极当代文学文化生活等。跨学科学习情境则指向综合运用多学科学习经验与学科思想的复杂活动场景,如社会问题调查、专题研究、社会实践活动策划与组织等。

2. 语文实践活动

一般来说,“活动”往往兼具名词与动词的性质,具体指为达成某种目的而采取的行动。在本研究中,语文实践活动是指为了达成语文课程学习目标,学生自主开展的运用语言文字的实践行为。与教学情境中的语文实践活动不同,它是对日常具体的语文实践行动抽象化的结果,超越于具体的学习主题、学习材料和学习内容,是独具语文课程特色的“元学习活动”。根据《语文新课标》的具体要求,评估模型中的语文实践活动主要包括阅读与鉴赏、表达与交流、梳理与探究。阅读与鉴赏是指对文本的局部与整体形成合理理解,能从思想内容、表达方式、组织结构等不同角度进行探究与评价,并能将阅读经验与体验运用于问题解决;表达与交流是指呈现和记录事件、描绘事物、解释原因、表达自己个人的思考和体验,并根据现实目的与需要进行有效交流;梳理与探究是指对积累的语言材料、语言实践经验的整理、分类与深度探

究,进而加深对语言文字及其运用规律的深入理解。

根据语文实践活动的不同特点,郑国民等对学生语文核心素养的关键表现要素进行提炼(见表1)。以阅读与鉴赏为例,该过程中学生的关键表现要素包括整合信息、整体感知、理解阐释、推断探究及赏析评价。其中:整合信息是指学生能够有效捕捉文本的重要信息,并在不同信息之间建立合理的关联与整合;整体感知强调对文本内容、意蕴、语言等的直观感受和体会;理解阐释要求学生进入文本的更深层次,在文本意义的情境中解释具体词句的内涵和作用、阐释文字背后的意义;推断探究强调以文本内容为基础,结合个人生活与阅读经验,形成合理的发现与结论;赏析评价要求学生超越文本,结合自身经历和生命体验欣赏文本,进而能客观地审视、思考文本的思想内容,评价文本的表现形式。上述5个要素又可概括为3类:第一类主要指向文本信息的定位与关联,具体表现为整合信息;第二类主要指向文本内容与意义的理解与建构,具体表现为整体感知、理解阐释、推断探究;第三类主要指向对文本内容与形式的有效性、来源可靠性的反思与评估,具体表现为赏析评价。需要说明的是,上述3类语文实践活动是相互关联的有机整体,在评价任务设计中既要关注语文实践活动的独特性,又要加强不同类型语文实践活动的综合性。

表1 语文实践活动中“语文核心素养”的关键表现要素^[20]

语文实践活动	关键表现要素
梳理与探究	筛选与提炼、归纳与分类、比较与抽象、搜集与组合、发现与再造
阅读与鉴赏	整合信息、整体感知、理解阐释、推断探究、赏析评价
表达与交流	陈述与叙述、描绘与表现、解释与分析、介绍与说明、应对与调整

3. 语文课程大观念

随着核心素养这一育人目标的提出,对“什么是课程内容”提出了再认识的要求。从国际教育变革经验来看,以“大观念”或“关键主题”作为课程内容成为新趋势,即追求“活动”与“经验”的综合作用,围绕“少而精”的学习内容或主题,借助学科实践方式促进学生学科学习经验的积累与梳理。在评价模型中,“大观念”是语言文字运用的关键概念、原则与规律,并非具体的知识或技能,是学生通过语文实践活动逐渐获得的对语言文字运用本质规律的认识。

本研究以《语文新课标》中“语文学习任务群”的学习要求与教学提示为依据,通过专家访谈与问卷调查,梳理并提炼出义务教育阶段语文大观念。以拓展型学习任务群“整本书阅读”为例,本研究提炼的“大观念”主要包括:(1)根据阅读目的恰当选择阅读方法和辅助性资源(如目录、作品背景等),有助于准确把握文本内涵;(2)阅读是个性化过程,对文本的理解因个人经验、视角、观念、社会文化环境不同而不同;(3)阅读不同主题、不同类型的文本,能够帮助个体更好地了解自我、家庭和社会文化,并与其建立积极的互动关系。需要强调的是,语文课程内容与学习资源永远是挂一漏万的。语文课程大观念的梳理与提炼无法追求全面覆盖,评价者可根据具体的评价对象、课程内容、评价材料特征,不断丰富与拓展语文大观念的具体内容。

三、义务教育语文核心素养评价模型的实践应用

(一)建构命题质量元评估框架,有效落实语文命题改革理念

以国家课程方案与课程标准为根本依据,提升考试命题的科学性与有效性是深化教育评价改革的关键环节。然而,在考试命题实践中,不同命题主体对课程标准核心理念与具体要求的理解与落实程度存在一定差异。以初中学业水平考试为例,不同区域的语文试题在命题理念、测试形式、测试内容等方面均存在较大差异:一些地区注重情境导向的命题理念,强调社会生活实际与语文学习的关联;一些地区则较为关注对语法、修辞、字音字形等语言知识的考查。面对现实情况,以义务教育语文核心素养评价模型为基础建构语文命题元评估框架,有利于开展不同区域、不同学段的语文命题质量评估工作,有利于缩小语文命题质量的区域差异。具体而言,参照素养评价模型中的评价构念建构、评价任务设计和评价任务验证3个模块,可以搭建“评价构念—评价框架—评价工具—评价导向”命题质量的元评价模型与指标体系,重点评估构念内涵界定的合理性、评价框架指标筛选的准确性、评价工具设计的有效性与导向性等相关问题。

(二)提升语文过程性评价的科学性,促进教师评价素养发展

过程性评价是语文课程评价的重要组成部分,其功能在于了解学习目标达成程度、发现学生学习过程中的困难与问题,并通过有效反馈干预教学进程、提升教学效率。与终结性评价不同,教师是设计与实施过程性评价的关键角色,教师对评价理念的

感知与理解会直接影响过程性评价的实施效果和學生核心素养的发展。当前,教师依托课堂教学开展的过程性评价面临诸多困境,如整体关照不足、目标定位不准、评价技术不熟练等^[21]。据此,在日常教学中,义务教育语文核心素养评价模型能为过程性评价的设计与实施提供技术支持。

在教学过程中,教师可以运用评价模型解决“评什么”“如何评”“结果怎么用”3个问题。首先,发挥评价模型的“蓝图”作用,帮助教师系统规划评价实践。借助评价模型,语文教师评价观至少应实现2个方面的突破:一要具备“构念—目标”意识,以语文核心素养的内涵为基础,建构过程性评价的目标体系;二要充分认识语文核心素养的评价必须依托情境、语文实践活动与语文课程大观念等相关要素。其次,以评价模型为根本依据,编制过程性评价工具。语文核心素养无法直接观测,需要通过观察学生在真实的言语实践活动中的认知行为表现来推断其语文核心素养的发展水平。就评价任务的设计而言,不同语文实践活动涉及的关键表现要素能帮助教师明确应着重关注学生的哪些具体表现,进而设计富有针对性的评价工具或测试任务,从而提升过程性评价的科学性与有效性。最后,精准诊断学生语文核心素养发展水平。分析学生在学习领域的优势与弱势,对教学目标的设定与选择具有重要意义。评价模型具有较强的可解释性,教师可以通过分析测试题目或评价任务的内容与数据结果,明确学生与语文学习目标之间的差距,了解学生的薄弱环节,进而将教学重点放置于“修复问题”方面。同时,教师还可以在不同阶段的过程性评价结果之间建立联系,进一步了解语文核心素养发展的个性化需要,以更好地开展课堂教学实践。

(三)优化语文教育质量监测,提升教学改进效率

了解区域教育发展现状、提升区域教育教学水平,对我国教育均衡发展目标有着重要的价值。运用义务教育语文核心素养评价模型,区域可以对不同年级语文教学效果、学生学习质量开展追踪评估。具体来说,运用上述评价模型至少可以达成2个方面的评估目标:其一,围绕单一维度进行跨年级比较,了解不同年级学生语文核心素养发展模式的差异;其二,对同一年级不同时期的学生进行纵向比较,了解区域语文教育教学质量的稳定性。在此过程中,运用评价模型可积累丰富的刻画区域语文教育质量、反应语文教育优势与不足的过程性数据,为有效诊断和改进区域语文教学、形成正确的语文学

习与教学观念提供客观依据。

此外,在运用评价模型开展教育质量监测的基础上,分析学生语文学业表现特征与规律,可有效检验区域内学生达成国家课程标准的程度与水平。具体来说,借助评价模型中的具体指标,可以了解区域内不同学校、不同群体学生的语文学业表现的优势与问题。以“语文实践活动”为例,通过借助相关统计分析方法,系统描绘区域内学生在“梳理与探究”“阅读与鉴赏”“表达与交流”的发展类型特征。综上,基于义务教育语文核心素养评价模型开展监测,不仅有利于提升语文教育与课程标准的内在一致性,还有利于深入挖掘区域语文教学发展过程中存在的过程性问题。

总之,在课程与教学持续变革的当下,建构语文核心素养的评价模型是促进课程理念落地的关键环节。同时,我们也必须意识到,真实情境中学生语文核心素养的发展、语文教学实践必然是复杂多样的。如何运用核心素养评价模型充分体现上述特征,充分落实“评价促进学习”“评价即为学习”的核心理念,仍需要深入探索。在未来的研究中,还应进一步聚焦评价模型验证与应用探索,不断提升评价模型自身的科学性与有效性。

参考文献:

[1] 崔允漷. 试论新课标对学习评价目标与路径的建构[J]. 中国教育学刊, 2022(7): 65-70, 78.

[2] MESSICK S. The interplay of evidence and consequences in the validation of performance assessments[J]. Educational Researcher, 1994, 23(2): 13-23.

[3] RICONSCENTE M M, MISLEVY R J, CORRIGAN S. Evidence-centered design[M]//LANE S, RAYMOND M R, HALADYNA T M. Handbook of test development. New York: Routledge, 2016: 40-63.

[4] PELLEGRINO J W. Assessment as a positive influence on 21st century teaching and learning: A systems approach to progress[J]. Psicología Educativa, 2014, 20(2): 65-77.

[5] 王慧君, 赵紫薇, 李宇婷. 基于证据的学业评价: 观点、框架与实践路径[J]. 中国考试, 2022(2): 64-72.

[6] DEANE P, SABATINI J, FENG G, et al. Key practices in the English language arts: Linking learning theory, assessment, and instruction[R]. Princeton, NJ: Educational Testing Service, 2015.

[7] 张娜, 李峰. 合作问题解决能力评价的本土化研究: 基于以证据为中心的设计模型[J]. 中国考试, 2019(8): 59-65.

[8] 袁建林, 刘红云. 核心素养测量: 理论依据与实践指向[J]. 教育研究, 2017, 38(7): 21-28, 36.

[9] 何伟, 王连柱. 语言学流派及语言观的历史嬗变[J]. 外语学刊, 2020(2): 8-20.

[10] 丁炜. 语言学习的社会文化观: 关于语言学习的一种新取向[J]. 广西师范大学学报(哲学社会科学版), 2010, 46(4): 137-141.

[11] HYMES D H. On communicative competence [M]//PRIDE J B, HOLMES J. Sociolinguistics: Selected readings. Harmondsworth: Penguin Books, 1972: 269-293.

[12] VYGOTSKY L S, COLE M, JOHN-STEINER V, et al. Mind in Society: The development of higher psychological processes[M]. Cambridge, MA: Harvard University Press, 1978.

[13] OECD. PISA 2018 assessment and analytical framework [EB/OL]. (2019-04-26) [2022-09-10]. https://www.oecd-ilibrary.org/education/pisa-2018-assessmentand-analytical-framework_b25efab8-en.

[14] Council of Ministers of Education, Canada. PCAP 2016 assessment framework [EB/OL]. [2022-09-10]. <http://www.cmec.ca/docs/pcap/pcap2016/PCAP-2016Assessment-Frame-work.pdf>.

[15] SABATINI J, PETSCHER Y, O'REILLY T, et al. Improving comprehension assessment for middle and high school students: Challenges and opportunities[M]//SANTI K L, REED D K. Improving reading comprehension of middle and high school students. New York: Springer International Publishing, 2015: 119-151.

[16] SABATINI J, O'REILLY T, DEANE P. Preliminary reading literacy assessment framework: Foundation and rationale for assessment and system design[R]. Princeton, NJ: Educational Testing Service, 2013.

[17] DEANE P, SONG Y. The key practice, “discuss and debate ideas”: Conceptual framework, literature review, and provisional learning progressions for argumentation [R]. Princeton, NJ: Educational Testing Service, 2015.

[18] 岳秋云, 董宏建. 素养本位的教育: 为何及何为[J]. 教育研究, 2022, 43(3): 35-46.

[19] VAN DEN HEUVEL-PANHUIZEN M. The role of contexts in assessment problems in mathematics[J]. For the Learning of Mathematics, 2005, 25(2): 2-9, 23.

[20] 郑国民, 李宇明. 义务教育语文课程标准(2022年版)解读[M]. 北京: 高等教育出版社, 2022: 228.

[21] 宋词, 郑东辉. 学教评一致性的课堂实践困境与突破[J]. 当代教育科学, 2018(11): 22-26.

【作者简介】李倩(1987-),女,北京师范大学中国基础教育质量监测协同创新中心讲师;郑国民(1967-),男,北京师范大学文学院教授,博士生导师(北京 100875)。

【原文出处】摘自《中国考试》(京), 2023. 1. 51~60

【基金项目】2018年度教育部人文社科青年研究项目“促进学习的高中语文核心素养评价与监测研究: 以优秀诗文为例”(18YJ880042)。