

中小学 STEM 学习中高层次思维测评 模型构建与应用

首新¹, 胡卫平², 刘念³

(1.重庆师范大学 科技教育与传播研究中心, 重庆 401331;
2.陕西师范大学 现代教学技术教育部重点实验室, 陕西 西安 710062;
3.重庆市中山外国语学校, 重庆 404500)

[摘要] 发展高层次思维已成为 STEM 跨学科教学的重要目标, 但如何评估 STEM 学习中的高层次思维, 仍没有可行的工具。本研究采用问卷调查、专家咨询等研究方法, 运用探索性结构方程模型(ESEM)、层次分析(AHP)、拟合 Rasch 模型等手段, 建立测评模型、确定指标权重、开发测评工具。数据分析显示, 中小学 STEM 学习中高层次思维测评工具具有良好的信度和结构效度。正式测量发现, 小学阶段学生 STEM 学习中高层次思维整体呈现波动增长, 性别、地域等因素影响较小。测评模型对评估中小学 STEM 教育项目具有应用价值。

[关键词] STEM 教育; 高层次思维; 测评; Rasch 模型; 义务教育阶段

[中图分类号] G434 **[文献标志码]** A

[作者简介] 首新(1987—), 男, 重庆梁平人。讲师, 博士, 主要从事科学教学心理研究。E-mail: 346532216@qq.com。胡卫平为通讯作者, E-mail: weipinghu@163.com。

一、研究背景

发展高层次思维(Higher Order Thinking)是 STEM 跨学科学习区别于传统学科知识和技能导向学习的重要特征。随着跨学科课程改革的不断推进以及核心素养培养目标的确立, STEM 教育理念下的高层次思维培育在世界范围内得到了高度重视。那么, 何为高层次思维呢? 早期 Bartlett 认为, 高层次思维过程是个体为弥补先验知识不足而进行改补、推断、重新解释的过程^[1]。这一界定带有浓郁的教育哲学意味, 是从整合的视角看待新知识对先验经验的“间隙填充”。Resnick 认为, 若问题解决过程没有固定的规则路径, 需要应用多个判断标准, 且常常涉及不确定性, 那么, 高层次思维必定参与问题解决过程^[2]。Lewis 等人对 Resnick 的观点进行了心理学视角的诠释, 认为高层

次思维是个体面对复杂情境探求信息的过程中, 记忆中的已知信息与寻求到的新信息之间相互影响或(和)重新组合、延伸整合后达成某一目标的过程^[3]。Goodson 认为, 当个体面临的情境是不熟悉的, 要解决的是不确定的、或两难的问题, 高层次思维便被激活了, 其表现是能进行正确的解释、决策以及形成创造性的产品^[4]。从外在表现描述学习效果为高层次思维培养找到了突破口, 导致教育目标从识字、记忆等发展到了推理、判断、创新等积极的心智建构^[5]。Richland 认为, 高层次思维是一种关系结构映射(A Relational Structure Mapping)^[6], 反映出学科内容与能力的映射关系: 跨学科的信息分离和整合发展出更复杂的关系, 学习这些关系(如相似、不同、因果关系等)必定涉及问题解决、推理等高层次思维。在当前 STEM 跨学科整合趋势下, 这一观点回应了高层次思维与学科内

基金项目: 国家教育考试科研规划 2019 年度课题“大规模计算机交互式科学创新能力测评研究”(课题编号: GJK2019033); 重庆师范大学 2019 年博士启动基金项目“面向核心素养培育的科学高阶思维及其发展机制研究”(项目编号: 19XWB006)

容之间的关系,二者表现出映射结构。

测评高层次思维由来已久。Ross等人较早依据布鲁姆教育目标认知过程中的分析、综合、评价层级,开发了筛选4—6年级天才学生的高层次思维量表(The Ross Test of Higher Cognitive Processes)^[7],但是后来由于这一认知过程理论饱受争议,因此,使用逐渐较少。为回应学校教育检验高层次思维的需求,Plaw等人设计了创造—批判思维测验(Creative-Critical Thinking Styles Test)^[8],但该量表将批判性思维与创造性思维相对立,这一观点仍有待进一步商榷。国内也进行了相关研究,如杨翊等人也基于布鲁姆教育目标认知过程构建了高阶思维测评框架^[9],但理解、应用阶段是否蕴含高层次认知过程,目前仍有争议;姜玉莲等人利用数据驱动的因素分析显示,高层次思维包括创新性思维、决策能力、元认知与反思性评价、问题解决能力、同伴情感支持、自我效能感、批判性思维、自我调节学习等8个要素^[10]。目前,国内研究者多关注某一特定思维(如复杂问题解决、创造性思维等)的测评研究,专门用于中小学STEM学习活动中测评高层次思维的工具在已有文献中很少提及。

以往研究成果对明晰高层次思维结构、测量模型等具有参考价值,但仍存有较大的完善空间。第一,高层次思维结构确立方法有待精准化。大多数研究通过分析已有测评模型、解读课程标准等手段建立高层次思维结构,这容易忽略文化、学生学习特征的影响,以及期望课程与实际课程的差异^[11]。第二,高层次思维的学科化测评框架有待明晰。研究者普遍认为,高层次思维具有领域特殊性(Domain-specific Construct)^[12],学科内容对其有关键作用。因此,基于学科内容建立测评框架是未来研究的趋势,这可以了解学生在具体课程方案中的高层次思维状况,以便改进课程。第三,大规模高层次思维测评理论有待丰富。虽然高层次思维评价得到空前重视,但一些大型国际测评项目如PISA、TIMSS目前并没有单独的评估高层次认知过程。因此,PISA(2021)拟单独评估15岁青少年的创造性思维,这为建立大规模高层次思维测试模型奠定了基础。

为弥补上述高层次思维测评研究的不足,方法上,本研究通过数学建模确立面向STEM学习的高层次思维测评模型(STEM Based Higher Order Thinking Assessment Model),以提高研究过程的标准化程度,增强测评的规范性和实用性。测试任务设计过程中,以STEM跨学科内容为情境设计试题,确保对STEM学习过程具有诊断、调节和监督功能,进而以评价结果“返哺”STEM学习过程,促进学生高层次思维可持

续发展。

二、研究过程和方法

首先,依据文献分析,同时结合小学、初中科学课程标准(2017、2011),以及STEM跨学科学习特征,析出高层次思维测评指标,解决测评指标仅仅依靠课标解读或简单移植的弊病;其次,进行专家咨询,确定指标权重,解决测评指标之间动态关系不明的问题;最后,遵循测评指标及其权重,依据情境化试题设计过程,开发测评工具,并进行初步检验,验证测评工具的信效度。

(一)确定测评指标

目前,从STEM学习情境中测评高层次认知过程的工具并不多见。以前研究者普遍认可的高层次思维要素包括创造性思维、批判性思维、问题解决过程、决策思维、元认知等,我们拟利用STEM学习中的具体内容和行为来间接表征上述要素。同时,我们也考虑了STEM学习中的科学探究策略、学习动机、态度等思维倾向,因为以合作探究为主的STEM活动中,同伴间的社会建构对个体高层次思维发展具有重要作用^[13]。

在具体操作过程中,我们参考了国内外高层次思维通用框架、领域化测评框架中的认知过程条目。如“Kaufman领域创造力量表”中的日常、学术、科学题目^[14];Borich评估美国中小学生学习高级思维的“高阶思维和问题解决核查表”^[15];“Halpern批判性思维评价”中涉及STEM学科的题目^[16];改编大学生批判性思维倾向问卷^[17]、问题解决能力评分指标^[18]、典型行为创造性思维能力测验^[19]、高阶思维结构调查问卷^[20],加入STEM课堂背景,删除与中小学生学习不相关的题目;截取“中学生社会问题解决能力量表”中的社会科学问题^[21]。同时,还参考成熟的“加利福尼亚批判性思维倾向问卷”“Torrance创造性思维测验”中有关学习行为的测量指标。另外,还借鉴了“科学学习动机量表”^[22]、“对科学学习的态度量表”^[23]、“物理实验元认知水平问卷”^[24]中有关高层次思维倾向的观测指标。最终,初步形成117项测评条目。

利用调查样本对测评条目进行降维,析出测评指标。调查对象来自某直辖市451名中小学生学习,其中,男生占55.4%、女生占44.6%;农村学生占59.8%,城镇学生占40.2%。使用SPSS 22、MPLUS 7.4进行数据处理。

(二)确定指标权重

结合专家咨询意见,使用AHP法梳理指标的权重。咨询对象是高等院校15名科学教育、学科课程与教学的专家和科学教育学博士。邀请专家对测评指标

按照经验进行两两比较 (i, j), 形成重要性比例标度 (a_{ij} , 1-9 的数值, 1 表示 a_i 与 a_j 同样重要, 3 表示 a_i 比 a_j 略微重要, 5 表示 a_i 比 a_j 重要, 7 表示 a_i 比 a_j 明显重要, 9 表示 a_i 比 a_j 绝对重要; 2、4、6、8 表示两个相邻判断的中间值)。根据 a_{ij} 值形成专家判断矩阵, 计算特征向量和最大特征值, 并进行一致性检验, 最终通过层次总排序得出测评指标权重值。

(三) 形成测评工具

依据测评指标及其权重, 设计情境化试题考查 STEM 跨学科背景下的高层次认知过程。测试题设计包括背景材料、知识内容、认知过程等三个要素。背景借鉴了 PISA(2015)划分的个人、地区/国家、全球等三层级背景, 是带有科学意义的现象、观点或规范。知识内容具有跨学科性, 旨在建立认知过程的起点, 以此形成思维能力主导、必备知识引导、认知过程参与、情感态度支持的心理表征过程。另外, 知识内容还充当唤醒某一认知过程的作用。

采用如下程序形成测评工具: (1) 修订^①、自编题目, 组成试卷初稿。(2) 邀请 1 名科学教育专家、2 名科学教育研究生、2 名高级科学教师评估题目与测试指标的一致性、题目难度, 以及术语表述、语言陈述等方面的问题。研究者本人也结合自身专业判断对有争议的题目进行进一步修订。(3) 建立 Rasch 模型进行 3 轮测试, 完善测评工具的信效度。(4) 利用工具进行大规模测试, 评估 STEM 学习中的高层次思维。

表 1 被试分布信息

时间	学校位置	2 年级	4 年级	6 年级	男生	女生
第 1 轮	城市	42	48	40	67	63
第 2 轮	城市	41	44	37	61	61
第 3 轮	城市、农村	84	92	87	132	132
合计		167	184	164	260	256

评估测评工具信效度的调查对象来自某直辖市 4 所普通小学 2—6 年级 516 名学生, 见表 1。采用

表 2 探索性因素分析提取不同公因子时模型拟合情况

公因子	χ^2	df	χ^2/df	TLI	CFI	BIC	AIC	SRMR	RMSEA(90% CI)
8	1235.633	895	1.380	0.931	0.951	64005.364	62048.866	0.026	0.03(0.026 0.034)
9	1070.089	810	1.321	0.941	0.963	64327.43	62025.907	0.022	0.027(0.023 0.032)
10	1176.65	852	1.381	0.931	0.954	64166.458	62035.419	0.024	0.03(0.026 0.034)

表 3 提取 9 个公因子两种模型(ESEM、CFA)拟合指标的比较

模型	χ^2	df	χ^2/df	TLI	CFI	AIC	BIC	SRMR	RMSEA(90% CI)
ESEM	1176.538	852	1.381	0.931	0.954	62035.419	64166.458	0.024	0.03(0.026 0.034)
CFA	1760.399	1188	1.482	0.912	0.918	62312.216	63079.39	0.043	0.034(0.030 0.037)

①修订的题目来自 PISA、TIMSS 公布的科学公开题目、NEAP 科学测试题目、英国科学教育质量监测题目等。

Winsteps3.72.3 软件分析测评工具信效度, 应用 SPSS 22 软件分析 STEM 学习中高层次思维状况及影响因素。

三、研究结果

(一) 中小学 STEM 学习中高层次思维测评指标

首先进行正态分布、信度和效度检验。偏峰度检验显示, 偏度系数和峰度系数大部分小于 1, 基本符合正态分布; Cronbach's Alpha 信度系数 0.921; 采用 KMO 和 Bartlett 球形检验, 显示 KMO 值 0.917, $\chi^2 = 21060.949$, $df = 6786$, $p < 0.01$, 初步判定可进行探索性因素分析。

按照因子载荷大于 0.45 的标准, 逐步删除载荷小于 0.3, 或者仅有 3 个以下指标构成的公因子项, 最后删除 66 个条目, 保留 51 个条目, 此时 Cronbach's Alpha 信度系数 0.960, Guttman 半分信度 0.900, KMO 和 Bartlett 球形检验显示, KMO 值 0.959, $\chi^2 = 9140.169$, $df = 1275$, $p < 0.01$, 问卷信度较好。最终, 共抽取 9 个公因子, 其共同性大于 0.5, 累计方差贡献率达到 54.751%。

再利用 MPLUS 7.4 进行探索性因素分析(EFA), 验证抽取 9 个公因子的合理性。MPLUS 进行 EFA 也可以提供模型拟合指标, 从而判断模型拟合情况, 当抽取 8、9、10 个公因子时, 拟合指标见表 2, 各指标的含义可见相关文献^[25]。可以发现, 抽取 9 个公因子时, 模型拟合指标较好。

采用探索性结构模型(ESEM)和验证性因子分析(CFA)进行验证性因素分析。ESEM 吸纳了探索性和验证性因素分析、结构方程模型的特点, 兼具有目的的探索和验证功能。抽取 9 个公因子时, ESEM 和 CFA 模型拟合情况见表 3。可以发现, ESEM 拟合较好。这和两个模型的假设有关, CFA 会限制某些因子载荷为 0 而高估其中一个或几个因子的载荷, 以及因

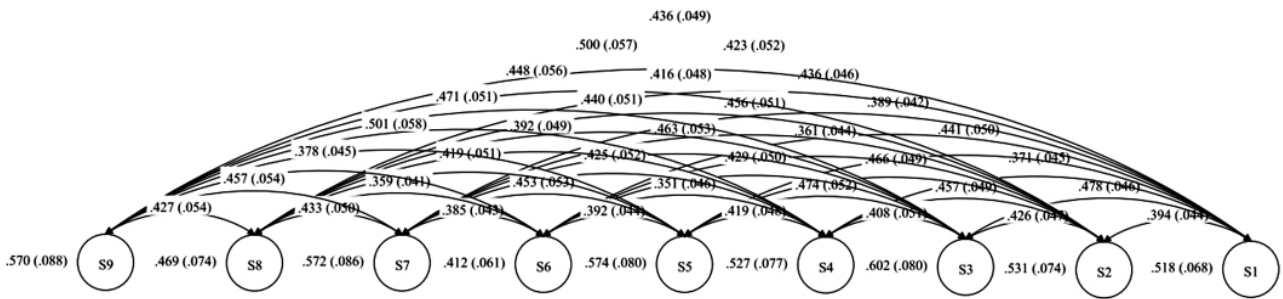


图1 一阶CFA分析因子结构

子与外部变量的关系,因此,拟合指标相对不太理想。但总体而言,提取9个公因子能很好地说明高层次思维的结构。

利用ESEM和CFA抽取9个公因子所包含的条目基本一致,稍有不同的是,ESEM使某些条目在几个公因子中出现了跨载荷的情况,且ESEM提取的某些条目的载荷较低。对于抽取结果有差异的、ESEM因子载荷低于0.3的以及跨载荷的条目,我们通过查看条目内容判断是否删除、修订以及变更。因为CFA分析显示,因子载荷都在0.5以上,说明前期EFA分析结果较为可靠。最终,修订条目A36、A41、A28、A8、A40、A50、A2、A12,再进行验证性因素分析,如图1所示,模型拟合良好。

最终,中小学STEM学习中高层次思维测评指标为9个:第一个命名为“问题解决”,阐述的是学生的探究学习过程;第二个命名为“元认知与反省”,阐述的是STEM学习的计划、组织、信息解释、反思等过程;第三个命名为“创造性思维”,解释了从多角度、应用创新的方法解决问题;第四个命名为“批判性思维”,阐述的是STEM学习中对不同观点的论证、评价以及交流的能力;第五个命名为“决策思维”,阐述的是STEM学习中的推理、合理选择、适时分析与预测的过程;第六个命名为“自我监控”,解释了STEM学习中自我管理过程;第七个命名为“迁移与应用”,阐

述的是将STEM学习中形成的行为规范、知识、方法等迁移到真实情境或其他领域;第八个命名为“科学方法”,阐述的是STEM学习中形成的思路、标准、规范、技能、模式等;第九个命名为“自我调节”,阐述的是STEM学习中面对不合理因素的自我反应。具体说明见表4。

(二)中小学STEM学习中高层次思维测评指标权重

中小学STEM学习中高层次思维测评指标的重要性各不相同,为了客观呈现重要性比例标度(a_{ij}),邀请15位专家评分,取平均值建立判断矩阵,权重向量见表5,同时计算出指标权重、最大特征根 λ_{\max} 和检验系数CR。

综合中小学STEM学习中高层次思维测评指标及其权重值,可以构建测评模型的数学方程式:

$$Y = \sum_{i=1}^9 W_i \times R_i$$

公式中,Y为STEM学习中高层次思维能力的综合评价值, W_i 为第*i*项指标的权重, R_i 为学生在第*i*项测评指标中的分数。Y值越大,表明高层次思维能力越强。

(三)中小学STEM学习中高层次思维测评工具的开发

编制题目是工具开发的重点,遵循以下原则:严

表4

9个公因子表征的含义

潜变量	名称	内 涵	支持因子数
S1	问题解决	发现问题、判断信息、计划并实施、解释与分析等过程	9
S2	元认知与反省	自我反思、判断、评判以获得学习更好发展的过程	8
S3	创造性思维	思维的流畅性、灵活性和独创性,注重科学概念、方法、情感的迁移	6
S4	批判性思维	对现象、结果、他人的观点等进行适宜的判断,开展基于证据的论证	5
S5	决策思维	基于证据的判断和分析,并能兼顾现有结果开展新的预测	6
S6	自我监控	不断进行监察、自我反馈、评价,进而进行改进和完善的过程	5
S7	迁移与应用	将学习行为、知识、技能等应用在新情境中	3
S8	科学方法	寻找证据、分析数据、解释信息的途径和手段	5
S9	自我调节	跟踪学习效果、对比学习目标,进而有意识地进行调整	4

表 5

测评指标判断矩阵

指标	问题解决	元认知与反省	创造性思维	批判性思维	决策思维	自我监控	迁移与应用	科学方法	自我调节	权重
问题解决	1	3	2	3	1/3	5	5	3	3	0.1955
元认知与反省	1/3	1	1/3	1/2	1/5	2	1/3	1/2	2	0.0518
创造性思维	1/2	3	1	1	1/2	4	4	6	4	0.1590
批判性思维	1/3	2	1/1	1	1/2	4	3	4	4	0.1347
决策思维	3	5	2	2	1	5	3	5	5	0.2569
自我监控	1/5	1/2	1/4	1/4	1/5	1	1/2	1/2	1/1	0.0342
迁移与应用	1/5	3	1/4	1/3	1/3	2	1	3	2	0.0773
科学方法	1/3	2	1/6	1/4	1/5	2	1/3	1	2	0.0539
自我调节	1/3	1/2	1/4	1/4	1/5	1	1/2	1/2	1	0.0367

注： $\lambda_{\max}=9.6971$, $CR=0.0596<0.1$ 。

格遵循测评指标和权重,题目反映某一高层次认知过程,题目具有 STEM 跨学科背景,内容基于每个年级学生所学知识,题目之间相互独立,题目以开放题为主,并建立良好的评分标准。工具最初包括 39 个题目,建立部分赋值 Rasch 模型,根据拟合度、题目难度—被试能力图(Item-Person Map)、等级概率值等指标对部分题目(共 10 个)进行修订,最终,正式测评工具包括 39 个题目(10 个选择题,29 个主观构造题)。选择题采用 0、1 记分,针对开放题,根据部分赋值 Rasch 模型提供的等级概率曲线对评分标准进行修订,部分赋值 Rasch 模型的优势是可以针对每个题的等级评分进行参数估计,从而更加直观地判断评分的合理性。修订之后,开放题记分 0、1、2,或 0、1、2、3,或 0、1、2、3、4。

正式测评显示题目质量良好:测试工具的 Alpha 信度系数为 0.94,所选被试能力信度为 0.77,题目难度—被试能力图显示,被试能力绝大多数都有试题群与之对应,说明题目有较好的代表性,避免了天花板效应和地板效应,被试能力和题目难度基本呈正态分布。但也发现,某些能力水平的被试并没有合适的题目匹配,如 Rasch 值为-0.59~-0.21 间的被试(即题 MI3 与题 PR6 之间)没有靶向题目,从经典测量理论来看,需要进一步增加难度值在 0.55~0.65 之间的题目。

进一步利用 ESEM 模型验证 39 个题目与测试指标之间的关系,当公因子设定为 9 时,显示模型拟合良好($\chi^2/df=1.418$, $TLI=0.938$, $CFI=0.965$, $SRMR=0.021$, $RMSEA=0.031$),说明题目的载荷与理论预设(如指标权重)具有高度一致性。少部分题目出现了跨因子现象,即某一题目可能涉及两个或两个以上高层次认知过程,这正是 ESEM 模型的优势所在,虽然我们设计一个题目时尽量只关注一个指标,但高层次思

维指标之间本身就具有相关性,可进一步对跨因子的题目进行修订。

(四) 中小学 STEM 学习中高层次思维总体状况及差异

1. 中小学 STEM 学习中高层次思维状况

测试工具包括 5 个链接题,利用链接题可将各年级的高层次思维测评结果固定在同一量尺上进行发展状况分析。如图 2 所示,二到六年级学生在 STEM 学习中的高层次思维差异显著, $F(2,249)=19.580$, $p<0.01$,整体呈现上升趋势,在四年级有所回落,但二、四年级无显著差异($p=0.646$)。这一结果与其他高层次思维测量研究基本一致,如董奇等人发现儿童发散思维能力在一至三年级呈现上升趋势,在四年级开始下降,五年级后又逐渐上升^[26]。

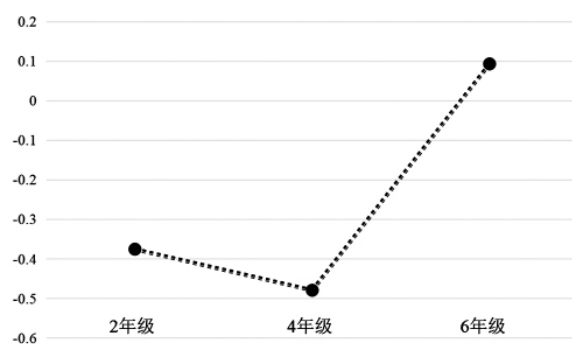


图 2 小学阶段 STEM 学习中高层次思维水平状况

2. 中小学 STEM 学习中高层次思维差异分析

采用 2(性别:男、女) \times 2(地域:城市、乡村)两因素组间设计、单因变量方差分析法(Univariate)分析性别、地域对 STEM 学习中高层次思维的影响。结果见表 6,性别、地域的主效应不显著($p>0.05$),说明男女之间、城市和乡村之间学生 STEM 学习中的高层次思维无显著性差异,但总体而言,女生略高于男生(-0.189>-

0.346)^①,城市略高于农村(-0.242>-0.295)。另外,性别、地域以及二者交互作用的效应值(η_p^2)都很小,如性别差异仅能解释STEM学习中高层次思维1.5%的变异,地域差异仅能解释0.1%的变异,这也说明男女生、城市和乡村之间差异不大。这主要是由于本研究样本年龄跨度较大,而中小学生的认知经验处于不断修正和发展之中,高层次认知过程在不断完善,虽然在各年龄段有所差异,但总体而言,男女生都表现出不断向前发展,导致差异较小。

表6 两因素(性别、地域)方差分析结果

方差来源	偏差平方和	自由度	F值	Sig.	η_p^2	Pe
性别	1.644	1	3.770	0.053	0.015	0.490
地域	0.147	1	0.337	0.562	0.001	0.089
性别×水平	1.113	1	2.552	0.111	0.010	0.356

四、讨论与分析

(一)中小学STEM学习中高层次思维测评指标及其权重清晰

研究结果显示,虽然验证性因素分析显示因子载荷均大于0.5,但采用探索性结构模型(ESEM)发现,某些条目出现了跨载荷的现象,这显示了ESEM的分析优势,同时也说明了STEM学习中的高层次思维要素之间存在紧密联系。进一步采用二阶因子模型(Second Order Confirmatory Factor Analysis)分析发现,决策思维(S5)是一个跨载荷指标(载荷分别是0.440、0.491),这和决策思维本身所蕴含的复杂的高层次认知过程不无关系。针对预期目标,在一定约束条件下,从诸多方案中选其一,并付诸实施,称为决策^[7]。从其定义可以推论,决策思维与问题解决、创造性思维、批判性思维、科学方法等指标有强烈关系。其他指标之间也有较强的关联,只是没有达到因素分析的显著水平,对测量指标进行相关分析,发现各指标间的Pearson相关系数在0.710~0.908之间。

指标权重确定过程中显示出了专家观点与数据驱动结果之间的差异。层次分析显示,决策思维、问题解决、创造性思维、批判性思维拥有较大权重,这与大多数研究者的观点基本一致,都认可这四个指标属于高层次思维。但由于STEM学习活动本身的独特性,增加了科学方法、迁移与应用、自我调节等指标,但有些指标并没有得到专家的认同,如有专家认为自我调节、自我监控不应纳入STEM学习中高层次认知过程,但STEM学习过程多以小组合作、相互交流、自主

探究的方式进行,基于同伴互动的自我调节和监控在合作学习中显然尤为重要。最终,综合专家意见确定的某些指标权重较小。

综上所述,通过建立测评指标、确定指标权重等过程,我们将STEM学习中的高层次思维进行了适宜的操作化解和降维,进而建立了中小学STEM学习中高层次思维测评模型,这一模型为后续建立测评工具提供了理论基础和操作准则。

(二)中小学STEM学习中高层次思维测评工具的信效度良好

研究结果显示,利用Rasch模型进行三轮测试题修订之后,试题质量提高,测评工具信效度良好。在测评工具修订过程中反映出某些学科课程内容计划与学生思维发展水平存在不一致。进一步使用Rasch模型建立STEM学习中高层次思维进阶层级发现,高年级学生处于高水平思维层级的占62.34%,比例明显提高;题项分析发现,这些高水平学生通常能够熟练运用该阶段的知识处理事物之间的相互关系,并且能够清晰地描述、推断其中的缘由,如探索铁制大门的下部分为何更容易生锈时,他们能够较好地考虑铁生锈的影响因素、铁门下端和上端在接触人和物体上的差异等。高水平学生比例过高,很可能源于教科书内容安排不适合学生思维发展水平,因为我们设计的测试题是基于学段课程内容的,因此,测试结果能反映出学生思维水平与学科课程内容之间的关系。总体而言,中低学段学生高层次思维呈现正态分布,但随着年级升高,STEM学习中高层次思维水平呈偏正态分布,可以推断,随着学生认知水平的发展,高年级的某些学科内容并没有与之相适应。

利用经典测量理论进一步验证了试题与测评指标的对应关系。某些题目的跨因子现象一方面说明了分解STEM学习中高层次思维的复杂性,另一方面也提示我们需要进一步修订测试题。根据试题设计原则,题目之间应该保持相对的独立性,但由于指标之间具有较高的相关性,也不能完全接受数据驱动的结果。鉴于此,后续研究将抽取跨因子题目进行具体分析,结合理论对某些题目进行修改,以进一步提高试题质量。

(三)中小学STEM学习中高层次思维发展趋势

研究结果显示,中小学STEM学习中高层次思维总体呈现上升趋势,但在小学中段有所回落,这可能有两方面的原因:第一,随着学生认知水平的快速发

^①数据为原始Rasch值,(UMEAN=0,USCALE=1)。

展,过往的认知经验已经不能解释现象或阻碍了问题解决过程,因此,认知结构可能面临重构,小学中段恰好处于这一阶段^[26],表现出不能理解知识之间的联系,进而影响了测评表现。第二,逐渐感受到的学业压力抑制了某些高层次思维要素的发展。传统的学校教育要求学生接受知识、巩固知识,而对于知识、信息缺乏批判性的思考和创造性应用^[28],这限制了高层次认知过程的发展。受制于自身的认知局限和外部的学业压力,最终导致了小学中段STEM学习中高层次思维的下滑。

研究结果显示,中小学STEM学习中高层次思维无显著的性别差异,男生略低于女生。但后续建立思维进阶层级发现,不同思维水平下的性别差异显著,水平越高,男女生之间的差异越大。这提示教师在进行STEM教学时,应根据思维发展的性别差异合理安排小组活动,如鼓励女生充分表达观点、推陈出新的同

时,也要引导她们进行反思;另一方面,要鼓励男生发表对问题的看法,引导他们进行创造性的设计和论证。

五、结 语

通过文献分析、问卷调查、专家咨询等,并结合实证数据分析,本研究构建了中小学STEM学习中高层次思维测评指标及其权重,形成了测评模型。九个测评指标是:问题解决、元认知与反省、创造性思维、批判性思维、决策思维、自我监控、迁移与应用、科学方法、自我调节,其大致权重分别为19.55%、5.18%、15.90%、13.47%、25.69%、3.42%、7.73%、5.39%、3.67%。以此测评指标及权重开发了测评工具,通过项目反应理论和经典测量理论修订后,工具具有较好的信度、结构效度。针对我国目前如火如荼开展的各级各类中小学STEM教育项目,亦可使用本研究的测评指标、测评工具评估“发展学生高层次思维”的STEM教学目标。

[参考文献]

- [1] BARTLETT F C. Thinking: an experimental and social study [M]. London: Allen & Unwin, 1958.
- [2] RESNICK L B. Education and learning to think[M]. Washington DC: National Academy Press, 1987.
- [3] LEWIS A, SMITH D. Defining higher order thinking[J]. Theory into practice, 1993, 32(3):131-137.
- [4] GOODSON L, ROHANI F. Higher order thinking skills: definition, teaching strategies, assessment[EB/OL]. (2002-04-17) [2020-01-23]. http://cala.fsu.edu/files/higher_order_thinking_skills.pdf.
- [5] 钟志贤.促进学习者高阶思维发展的教学设计假设[J]. 电化教育研究, 2004(12):21-28.
- [6] RICHLAND L, SIMMS N. Analogy, higher order thinking, and education[J]. Wiley interdisciplinary reviews cognitive science, 2015, 6(2):177-192.
- [7] ROSS J D, ROSS C M. Ross test of higher cognitive processes: administration manual [M]. Novato, CA: Academic Therapy Publications, 1967.
- [8] PIAWA C Y. Building a test to assess creative and critical thinking simultaneously[J]. Procedia-social and behavioral sciences, 2010, 2(2):551-559.
- [9] 杨翊, 赵婷婷. 中国大学生高阶思维能力测试蓝图的构建[J]. 清华大学教育研究, 2018, 39(5):59-67.
- [10] 姜玉莲, 解月光. 基于ESEM的高阶思维结构测量模型研究[J]. 现代远程教育研究, 2017(3):96-106.
- [11] 李化侠.小学生统计思维测评模型构建[J].教育研究与实验, 2018(2):77-83.
- [12] KUNCEL N R. Measurement and meaning of critical thinking.The National Research Council's 21st Century Skills Workshop, Irvine, California, February 10-11, 2011[C/OL].[2020-06-29].https://atecentral.net/downloads/209/Kuncel_Measuring%20Critical%20Thinking_Paper.pdf.
- [13] 祝智庭, 雷云鹤. STEM教育的国策分析与实践模式[J]. 电化教育研究, 2018, 39(1):75-85.
- [14] KAUFMAN J C. Counting the muses: development of the kaufman domains of creativity scale (K-DOCS)[J]. Psychology of aesthetics creativity & the arts, 2012, 6(4):298-308.
- [15] BORICH G D, TOMBARI M L. Education psychology: a contemporary approach[M]. 2nd ed. New York: Longman, 1997.
- [16] HALPERN D F. Thought and knowledge: an introduction to critical thinking[M]. 5th ed. New York: Psychology Press, 2014.
- [17] 刘义. 大学生批判性思维研究:概念、历史与实践[D]. 武汉:华中科技大学, 2010.
- [18] 余林. 课堂教学评价[M]. 北京:人民教育出版社, 2007.
- [19] 骆方. 中学生创造性思维能力测评问卷的编制——一个典型表现测验[D]. 北京:北京师范大学, 2003.

- [20] 姜玉莲. 技术丰富课堂环境下高阶思维发展模型建构研究[D].吉林:东北师范大学,2017.
- [21] 杨颖. 中学生社会问题解决能力的结构、特点与功能[D].北京:北京师范大学,2009.
- [22] TUAN H L, CHIN C C, SHIEH S H. The development of a questionnaire to measure students' motivation towards science learning [J]. International journal of science education,2005, 27(6): 639-654.
- [23] HONG Z R, LIN H S, WANG H H, et al. Promoting and scaffolding elementary school students' attitudes toward science and argumentation through a science and society intervention[J]. International journal of science education, 2013, 35(10):1625-1648.
- [24] 物理实验元认知水平问卷[EB/OL]. [2019-11-09] <https://www.wjx.cn/jq/2159007.aspx>.
- [25] 王孟成. 潜变量建模与 Mplus 应用[M]. 重庆:重庆大学出版社,2014.
- [26] 董奇. 儿童创造性思维发展心理[M]. 杭州:浙江教育出版社,1998.
- [27] 原道谋. 决策思维初探[J]. 政治学研究,1986(10):60-64.
- [28] 吴亚婕,姜姗姗. 学生学习过程、批判性思维与学业成就的关系研究[J].电化教育研究,2015,36(12):90-97.

Construction and Application of Higher Order Thinking Evaluation Model in STEM Learning in Primary and Secondary Schools

SHOU Xin¹, HU Weiping², LIU Nian³

(1.Research Center for Science Education and Communication, Chongqing Normal University, Chongqing 401331; 2.Key Laboratory of Modern Teaching Technology, Shaanxi Normal University, Xi'an Shaanxi 10062; 3.Zhongshan Foreign Language School, Chongqing 404500)

[Abstract] Developing high order thinking has become an important goal of STEM education, but there is still no feasible tool to evaluate it. This study adopts questionnaire survey and expert consultation, and uses Exploratory Structural Equation Modeling (ESEM), Analytic Hierarchy Process (AHP) and Rasch Model to establish evaluation model, determine index weight and develop evaluation tools. Data analysis shows that evaluation tools for high order thinking in STEM learning in primary and secondary schools have good reliability and structural validity. According to the formal measurement, higher order thinking in STEM learning of primary school students shows fluctuating growth, with less influence of gender and region on it. The evaluation model has practical value for evaluating STEM education projects in primary and secondary schools.

[Keywords] STEM Education; High Order Thinking; Evaluation Model; Rasch Model; Compulsory Education Stage