

小学生劳动素养认知模型的建构

菅璐

内蒙古师范大学心理学院, 内蒙古 呼和浩特

收稿日期: 2024年5月27日; 录用日期: 2024年7月8日; 发布日期: 2024年7月15日

摘要

近年来, 教育领域广泛采用认知诊断, 为教育提供了全新的评价方式, 推动了教育评价的多元化、精确化和个性化发展。但在劳动教育过程中, 认知诊断的应用相对较少, 尚未充分发挥其潜力。认知诊断的引入为劳动测评带来了新的可能性, 使教师能够更加精准地制定个性化的教学方法和策略。研究以小学生劳动素养的潜在心理结构为切入点, 运用认知诊断技术, 建构并验证了小学生劳动素养的认知模型基本完备且合理, 可以作为编制其认知诊断量表的理论构念。

关键词

劳动素养, 认知诊断, 模型构建

The Comparative Study of Cognitive Diagnosis Models for Elementary Students' Labor Literacy

Lu Jian

School of Psychology, Inner Mongolia Normal University, Hohhot Inner Mongolia

Received: May 27th, 2024; accepted: Jul. 8th, 2024; published: Jul. 15th, 2024

Abstract

In recent years, cognitive diagnosis has been widely adopted in the field of education, providing a new assessment method and promoting the diversified, precise, and personalized development of educational evaluation. However, the application of cognitive diagnosis in labor education processes is relatively limited, and its potential has not been fully realized. The introduction of cognitive diagnosis has brought new possibilities to labor assessment, enabling teachers to develop more precise personalized teaching methods and strategies. Taking the potential psychological structure of

elementary students' work literacy as the entry point, this study utilized cognitive diagnostic technology to construct and validate a basic and reasonable cognitive model of elementary students' work literacy, which can serve as the theoretical framework for developing cognitive diagnostic scales.

Keywords

Labor Literacy, Cognitive Diagnosis, Model Construction

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着时代的发展,中国的教育体制正在经历一系列变革和创新。在此背景下,教育测评也面临着全新的挑战与要求。通过不断更新的教育测评理念,研究者正在创新测评工具,并致力于制定更加准确、具有更高解释性的测评手段(李心钰, 陆宏, 2023),以提高教育测评的效率和精确度。在众多测评方法中,认知诊断测评将“个体认知”和“测量”相结合,通过观察被试的反应模式来估测其知识加工过程、认知结构和思考过程等微观水平(Leighton & Gierl, 2007),从而实现对个体宏观能力水平和微观认知结构水平的双重测验(Lee & Sawaki, 2009)。通过观察学生的思维过程、解决问题的策略和反应模式(Nichols, 1994),可以深入了解学生在劳动活动中的思维过程和认知发展。认知诊断不仅能够弥补传统量化分数评价的缺陷,还可以帮助教师实现因材施教和补偿性教学(Zhan et al., 2018; 詹沛达等, 2016)。教育领域逐渐引入认知诊断方法,为教育测评注入新思路。然而,在对劳动教育进行测评中,认知诊断的应用相对较少,尚未充分发挥其潜力。认知诊断的引入为劳动教育测评带来新的可能性,使得教师能够更加精准地制定个性化的教学方法和策略(武小鹏等, 2020; 万海鹏等, 2022; 汪大勋等, 2023; 骈扬等, 2023)。

然而,在劳动教育测评中应用认知诊断仍面临一些挑战。首先,劳动教育的评价标准和指标相对模糊,难以量化和测量,这给认知诊断测评的实施带来了困难。其次,认知诊断测评的精准度和可信度直接取决于认知诊断模型的构建。然而,由于劳动教育涉及的认知过程较为复杂,构建相应的诊断模型也有一定的挑战性。最后,不同能力之间的加工机制和层级关系各异,也会对诊断建模产生重要影响。

2. 认知诊断概述及模型分类

2.1. 认知诊断概述

对个体的认知过程、加工技能或知识结构的诊断评估被称为认知诊断(Nichols & Joldersma, 2008)。认知诊断有助于人们更好地了解个体内部的心理活动规律及加工机制,实现对个体认知发展实况的诊断评估,以促进个体全面发展(涂冬波等, 2021)。认知诊断是认知心理学与心理测量学的融合(Nichols, 1994; 高旭亮, 涂冬波, 2017)。认知诊断评估的实现需要基于特定的认知诊断模型(cognitive diagnostic models, CDMs),认知诊断模型是认知诊断评估的核心环节。认知诊断模型在其计量学特征和假设方面各有不同,它们之间存在一定联系和相似性,甚至在特定条件下可以进行转换(陈秋梅, 张敏强, 2010)。

认知诊断测评中,其中最为关键的是建立学科认知模型(Lee & Sawaki, 2009),即确定认知属性并开发与之相应的量表,构建Q矩阵模型。Q矩阵理论通过量表项目分析,提取出难以直接观测到的认知属

性，通过转换邻接矩阵、可达矩阵以及 Q 矩阵，将其与内在的认知结构建立联系。

2.2. 认知诊断模型及其分类

随着数量众多的诊断模型被相继开发与使用，认知诊断测评的使用逐渐走向成熟。目前存在百种以上的认知诊断模型，这些模型主要可以分为两大类：简化模型(Reduced Model)和饱和模型(Saturated Model)。其中简化模型又分为补偿模型和非补偿模型(汪文义等，2022)。

饱和模型的理论假设更宽泛，不会严格限制属性之间的作用，并且饱和模型认为根据被试掌握不同的属性，其答对项目的概率也应该有所不同。饱和模型同时涵盖补偿与非补偿关系。

相对于饱和模型，简化的认知诊断模型在构建项目反应函数时，属性之间的组合方式会对被试者在项目上的答题概率产生影响，这些影响可以被看作是对作答行为的限制。补偿模型强调了属性之间的互相补偿，即只要掌握一个属性，就可以弥补未掌握的属性，从而有较高的正确作答率(Yamaguchi & Okada, 2020)。非补偿模型则认为，必须同时掌握题目所测量的所有属性才能正确作答，缺少任一属性都会导致答对概率较低(De La Torre, 2011)。

饱和模型在一定的约束条件下，可转换为简化的认知诊断模型(De la Torre & Minchen, 2014)。因此，简化模型可以被认为是饱和模型的一种特殊情况(高旭亮等，2018)。

3. 认知模型构建

3.1. 核心概念界定

小学生劳动评价则以小学生的劳动素养作为评价对象，遵循小学生的身心发展规律，以每个阶段的学生特点为依据，以促进小学生具备劳动素养为出发点和落脚点，通过对小学生劳动素养的国家要求进行剖析，抓住每一个要素的关键点进行针对性评价与分析，对小学生劳动素养进行深入的挖掘与衡量，并及时提供内容反馈，为小学生劳动素养指明方向。

基于林崇德先生对核心素养的定义和《义务教育劳动课程标准(2022 年版)》的基础上，本研究尝试将小学生劳动素养界定为：小学生在认知发展过程中，通过接受相应学段的教育和培养，加强对个体认知、情感、行为等方面引导，逐渐形成劳动问题解决能力。

3.2. 关键认知属性的提取和层级关系的确立

Table 1. Key cognitive attributes of elementary students' work literacy and their meanings
表 1. 小学生劳动素养关键认知属性及其含义

数量	质量	排序
A1	小学生劳动价值观	是指在劳动教育过程中，小学生形成对劳动等方面的认知和评价，是影响他们决策和行动的核心因素。
A2	小学生劳动知识 (概念性知识)	是小学阶段学生对劳动相关概念的认知和理解。
A3	小学生情绪调控策略	通过劳动教育，引导小学生在处理劳动相关问题时，调控自身情感并获得积极的情感体验的方法。
A4	小学生劳动态度	指小学生在劳动问题解决过程中，展现出来的心理和认知行为倾向。
A5	小学生劳动操作认知	是指小学生在劳动问题解决过程中，对于具体操作过程的认知能力。
A6	小学生劳动习惯	是指在劳动问题解决过程中养成良好、稳定的行为倾向。
A7	小学生劳动解决能力	小学生在面对劳动任务或问题时，能够主动思考、分析和解决的能力。

本研究以马克思主义劳动观为理论基础，结合《义务教育劳动课程标准(2022 年版)》中所提到的劳动观念、劳动能力、劳动习惯和品质、劳动精神四个维度进行解构式分析。通过对重要文献和政策的梳理，发现不同学者从不同角度对小学生劳动素养展开研究，形成了多元的观点。基于对劳动课程标准和近 10 年来国内外关于劳动教育和劳动素养文献的研究，确定了关键属性。本研究最终将小学生劳动价值观、小学生劳动知识、小学生劳动情绪调控策略、小学生劳动态度、小学生劳动操作认知、小学生劳动习惯以及小学生劳动问题解决能力确定为小学生劳动素养认知诊断的关键属性(见表 1)。

劳动价值观作为劳动教育的导向，通过引领学生的理念和观点，指导小学生劳动素养培养的方向。影响他们在意识形态层面朝着社会期望的方向不断成长。劳动价值观对劳动知识的选择和学习具有重要影响。当个体的劳动价值观与某一种知识相契合时，他们会更加主动地深入学习和探究相关知识。同时，价值观也会影响情感的表达方式。适当引导和培养小学生正确的劳动价值观可以帮助他们形成积极向上的情感体验，激发对劳动的热情和兴趣，增强劳动参与和体验的积极性。劳动知识和小学生情绪调控策略是劳动教育的理论保障，是小学生劳动素养的基础要素。二者是相互依存、互相促进的关系。小学生劳动态度是小学生在劳动价值观支配下经过长期小学生情绪调控策略体验和知识积累而逐步形成的一种相对稳定的行为或行为准则。正确的态度可以促进小学生对劳动的认识和理解，培养他们的劳动操作认知和习惯。良好的劳动习惯是建立在正确的劳动态度基础之上(俞国良，何妍，2023)。劳动操作认知是劳动素养形成的关键环节，而劳动知识和劳动操作认知是劳动习惯形成的前提条件。良好的劳动习惯可以让小学生养成认真细致的工作方式，从而增强他们的观察力和问题发现的意识。通过劳动习惯的培养，小学生逐渐掌握解决问题的方法和技巧，提高解决问题的自信心和能力，并找到更多解决问题的新思路和方法。综上，最终确定了小学生劳动素养关键认知属性间的层级关系呈收敛型(见图 1)。

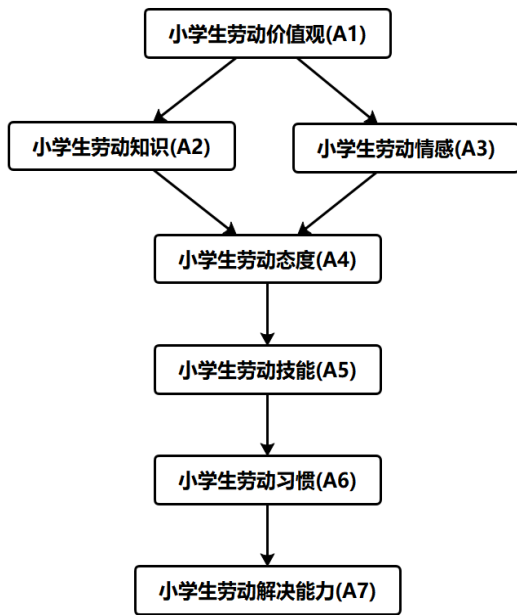


Figure 1. The hierarchical relationship of cognitive attributes of elementary students' work literacy
图 1. 小学生劳动素养认知属性的层级关系

3.3. 确立 Q 矩阵

确定了属性及其层级关系后，需要建立 Q 矩阵作为测验编制的蓝图，用以指导项目的编制。Q 矩阵

描述了测验项目与被测认知属性之间的关系，指导了测验项目如何测量特定认知属性。假设一个测验共测量了 K 个属性，那么理论上可能的项目属性组合模型就一共有 $2^K - 1$ 种(除去未测量任何属性的题目)。然而，由于属性间存在层级关系，某些属性必须以其他属性为前提，有些属性组合就不可能存在。因此，建立 Q 矩阵前，首先要确定表征属性间层级关系的邻接矩阵和可达矩阵。

邻接矩阵(Adjacent Matrix, A 矩阵)是反应属性之间是否存在直接关系的 0-1 矩阵。由横向和纵向按顺序分别标出邻接点，1 代表一个属性是另一个属性的前提条件，即直接关系，0 代表没有直接关系。其表示的是各认知属性之间仅存在对称关系，而不含间接关系和自身关系的矩阵。 A 矩阵是属性层级关系的数量化表达的形式，能够把线性的属性逻辑关系转换成用数字形式逻辑关系。见表 2。

Table 2. The attribute A matrix of elementary students' work literacy
表 2. 小学生劳动素养属性 A 矩阵

	A1	A2	A3	A4	A5	A6	A7
A=	A1	0	1	1	0	0	0
	A2	0	0	0	1	0	0
	A3	0	0	0	1	0	0
	A4	0	0	0	0	1	0
	A5	0	0	0	0	0	1
	A6	0	0	0	0	0	1
	A7	0	0	0	0	0	0

可达矩阵 R (Reachability Matrix, R 矩阵)与 A 矩阵一样对认知属性个数为 K 的模型来说是 K 行 K 列的 0-1 矩阵，属性间只要存在对称关系、反对称关系、传递关系、反传递关系和自我关系之一。因此，在 A 矩阵的基础上只需增加属性间的间接关系和自身关系即可， R 矩阵如表 3 所示。

Table 3. The attribute R matrix of elementary students' work literacy
表 3. 小学生劳动素养属性 R 矩阵

	A1	A2	A3	A4	A5	A6	A7
R=	A1	1	1	1	1	1	1
	A2	0	1	0	1	1	1
	A3	0	0	1	1	1	1
	A4	0	0	0	1	1	1
	A5	0	0	0	0	1	1
	A6	0	0	0	0	0	1
	A7	0	0	0	0	0	1

理想掌握模式是指符合逻辑的属性层级关系所组成的掌握模式。其中会有不符合逻辑、需要剔除的模式。根据 R 矩阵可得出理想掌握模式，即所有符合逻辑关系的个体知识状态(knowledge states, KS)。典型考核模式是符合逻辑的项目考核模式种类，比理想掌握模式少一种全为 0 的模式。见表 4。

模式种对属性 7 进行了两次测量，属性 8 进行了一次测量，不满足对这两个属性的多次测量(三次以

上)。因此,本研究在典型项目考核模式的基础上,就能够同时考察属性 7 和属性 5 的第 8 种考核模式增加两题。形成小学生劳动素养认知诊断测验的 Q 矩阵,如表 5 所示。

Table 4. Typical assessment model of elementary students' work literacy attributes
表 4. 小学生劳动素养属性典型项目考核模式

		1	2	3	4	5	6	7
典型项目 考核模式	A1	1	1	1	1	1	1	1
	A2	0	1	0	1	1	1	1
	A3	0	0	1	1	1	1	1
	A4	0	0	0	1	1	1	1
	A5	0	0	0	0	1	1	1
	A6	0	0	0	0	0	1	1
	A7	0	0	0	0	0	0	1

Table 5. The attribute Q matrix of elementary students' work literacy
表 5. 小学生劳动素养认知诊断测验 Q 矩阵

item	A1	A2	A3	A4	A5	A6	A7
1	1	0	0	0	0	0	0
2	1	1	0	0	0	0	0
3	1	0	1	0	0	0	0
4	1	1	1	1	0	0	0
5	1	1	1	1	1	0	0
6	1	1	1	1	1	1	0
7	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1
9	1	1	1	0	0	0	0
10	1	1	1	1	1	1	1
考察次数	10	8	8	6	5	4	3

成功构建的 Q 矩阵用以指导“小学生劳动素养认知诊断量表”的编制。编制的小学生劳动素养认知诊断量表由两大部分组成。首先是参与者的人口学信息部分,内容包括性别、年龄、班级、所在年级学段以及家庭独生子女情况等。其次是小学生劳动素养认知诊断量表的主体部分,考察了关键认知属性,共设计了 10 道题目,每道题目采用 5 级评分制度。

4. 认知模型验证

本研究采用文献回顾法和专家评定法初步确定了小学生劳动素养的认知模型。为了保证认知模型的完备性与合理性,对其做进一步的验证。[Tatsuoka \(1995\)](#)提出访谈并分析学生的真实答题过程是验证和改进认知模型的最佳方式。本研究在修改过程中采用了认知访谈法对 18 名小学生进行了口语访谈。通过与研究对象进行深入交流和讨论,可知本研究所构建的认知属性及属性之间的层级关系基本合理。学生在完成劳动任务时展现的认知过程与所构建的认知模型基本吻合。针对认知模型中存在的差异以及需要进

行修正的部分，需要经过下一步的实证验证，进而进行具体分析。

4.1. 模型拟合检验

对于认知模型的实证验证通常可以从两个方面进行，Tatsuoka (1995)指出，可以通过建立难度对项目属性的回归来验证认知属性的完备性。Leighton 等人通过建立属性掌握概率对量表总分的回归来验证层级关系的合理性。

4.1.1. 诊断工具与对象选取

- (1) 实测工具：“小学生劳动素养认知诊断量表”。
- (2) 实证验证通过线上线下同时发放的方式，发放问卷 815 份，经过数据筛选，最终获得有效数据 789 份。其中，学段一、学段二、学段三的学生分别为 305 名、363 名、121 名；男生 332 名，女生有 457 名。利用 R 语言，SPSS 以及 Excel，对实测数据进行分析，对小学生劳动素养的掌握情况进行诊断。

4.1.2. 认知模型属性完备性检验

使用 R 语言 ltm 程序包(Latent Trait Models under IRT)对项目难度进行估值，其值越大，表示项目越难，一般在[-3, 3]取值。结果如表 6。然后，使用 R 语言中的 CDM 程序包，以难度值为自变量，认知属性为因变量，进行项目难度对属性的回归分析，见表 7。

Table 6. The estimated values of test item difficulty parameters
表 6. 测验项目难度参数估计值

项目	考核模式	项目难度
1	1,000,000	-1.368
2	1,100,000	-1.935
3	1,010,000	-1.519
4	1,110,000	-1.745
5	1,111,000	-2.446
6	1,111,100	-1.902
7	1,111,110	-1.920
8	1,111,111	-0.632
9	1,111,111	-0.704
10	1,111,111	-1.254

Table 7. The regression of difficulty on attributes
表 7. 难度对属性的回归

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate	F
DINA	0.955	0.911	0.870	0.200	22.19***

注：* $p < 0.05$ ；** $p < 0.01$ ；*** $p < 0.001$ 。

结果表明，项目难度对小学生劳动素养七个属性的回归方程 F 检验显著性水平 $p < 0.01$ ，说明回归方程显著。小学生劳动素养的认知属性与项目难度的相关为 0.955，其解释量为 91.1%，效应值为 0.870。根据康春花等人(2013)提出的当解释量在 60%以上时，可以认为所确定的认知属性充分可靠。因此可以得出结论，本研究提出的小学生劳动素养认知属性基本完备且较为可靠，能够较好地预测项目难度。

4.1.3. 属性层级关系合理性验证

验证层级关系需要进行属性掌握概率对总分的回归分析，计算属性掌握概率则需要使用最优拟合模型。首先将通过 R 语言中的 CDM 包进行模型拟合检验。评估认知模型的拟合情况是通过比较相对指标和绝对指标来评估，以确定认知诊断模型的拟合程度。在认知诊断中，通常会先以绝对性指标为基础，来评估每个模型的拟合优度和复杂度，这些绝对性指标可以帮助确定单个模型的表现。接着，通过相对性指标来比较不同模型之间的表现，以确定最优模型。

相对指标检验在同等条件下，根据不同模型的相对拟合统计量指标确定更合理的模型。常用的相对拟合指标包括-2LL (-2 Log-Likelihood)、AIC (Akaike's Information Criteria)和 BIC (Bayesian Information Criteria)。AIC 是在基于-2LL 的基础上引入惩罚项得到的指标，旨在提高模型的拟合度(极大似然)同时限制模型参数的数量，以降低过度拟合的可能性。涂冬波等人(2012)指出，BIC 不仅考虑了待估参数的影响，还考虑了样本容量对拟合指标的影响，有效防止模型精度过高造成的模型复杂度过高。选择最佳模型时，通常选择这几类指标值最小的一类。

绝对指标检验考察假定模型与测试数据的拟合程度，包括三项具体指标：一、模型总体拟合度的指标 MX^2 ，用以评估所有项目实际观测值与模型预估值之间偏离程度的平均值；二、MADQ3 指标，反映项目残差间的相关性(Q3)的平均绝对值；三、SRMSR 指标，即标准化残差的均方根。上述绝对指标值越接近 0，则模型与数据的拟合度越高(Ravand & Robitzsch, 2019; 单昕彤等, 2014)。

在日常使用中，不同的模型在维度个数、属性关系、属性掌握表征、Q 矩阵完备性和多重策略方面存在一定差异(范士青, 刘华山, 2015)，若选择不恰当的模型，将会对诊断结果的精准性和可靠性产生影响(范婷婷等, 2023)。选取饱和模型 G-DINA (De La Torre, 2011)与 4 个简约模型 ACDM (De la Torre, 2011)、RRUM (Hartz, 2002)、DINA (Haertel, 1989)、DINO (Templin & Henson, 2006)进行认知诊断分析，其中 DINA 和 RRUM 是非补偿模型，DINO 和 ACDM 是补偿模型。除此之外，还选取了混合模型和 Seq-GDINA，共选取 6 个代表性的模型进行评估，构建多模型比较框架，对小学生劳动素养进行跨模型比较分析，以筛选出诊断性能最优的模型。这些模型在认知诊断评估中得到了较为广泛的应用，其参数设置各有特点，且属性间关系涵盖了补偿和非补偿两种类型。同时，跨模型比较还可以揭示各种模型在特定问题上的优缺点，为模型的改进和发展提供指导。见表 8。

Table 8. Comparison of fit statistics between different cognitive diagnostic models and test data
表 8. 不同认知诊断模型与试测数据的拟合统计量比较

模型	-2LL	AIC	BIC	χ^2	SRMSR	100 * MADRESIDCOV	MADaQ3
DINO	22375.32	22551.32	22962.35	243.41	0.077	0.769	NAN
DINA	21886.44	22062.44	22473.47	237.63	0.054	0.584	0.056
ACDM	21598.85	22046.85	23093.10	239.02	0.052	0.565	0.043
RRUM	22018.77	22466.77	23513.02	243.04	0.067	0.689	0.051
GDINA	20625.18	24795.18	34533.73	242.80	0.045	0.453	0.035
SeqG-DINA	21886.40	22104.40	22613.52	237.82	0.054	0.581	0.055

Table 9. The regression analysis results of the probability of mastering attributes on the total test score
表 9. 属性掌握概率对测验总分的回归分析结果

Model	R	R ²	Adjusted R ²	Std. Error of the Estimate	F
DINA	0.805	0.651	0.648	2.865	208.3***

注：* $p < 0.05$ ；** $p < 0.01$ ；*** $p < 0.001$ 。

通过 CDM 包择优选出的 DINA 模型对学生属性掌握概率进行估计, 计算每位学生的试测验总分。然后进行属性掌握概率对试测验总分的回归分析, 以探究学生不同属性掌握水平与其在测验表现之间的关系。见表 9。

回归方程显著性检验 $p < 0.001$, 表明回归方程是有效的。回归模型的系数 R 为 0.805, R^2 为 0.651, 说明被试劳动素养属性掌握可以解释总分 65.1% 的变异, 这说明认知属性层级关系是合理的, 同时也说明测验具有良好的结构效度。

5. 结语

认知属性的掌握概率反映了学生对认知属性的掌握程度, 是属性掌握模式和知识状态的显性表现。通过对属性掌握概率与量表总分进行回归分析, 可以验证认知模型的层级关系。本研究的小学生属性掌握概率与量表总分之间存在显著相关($p < 0.001$)。回归模型的系数 R 为 0.805, R^2 为 0.651, 小学生劳动素养属性掌握概率可以解释总分的 65.1% 的变异, 超过了 60%。可以认为小学生劳动素养认知属性之间的层级关系是合理的, 同时也验证了量表具有良好的结构效度。认知模型的潜在因素得分能够较好地解释量表总分的变异, 说明认知模型的结构是合理的。所构建的小学生劳动素养认知模型基本合理且完备, 表明该认知模型可以作为进一步编制“小学生劳动素养认知诊断量表”的理论构念。

署名单位

本文以内蒙古师范大学为第一署名单位。

基金项目

本文获得“内蒙古师范大学基本科研业务费专项资金资助(supported by The Fundamental Research Funds for the Inner Mongolia Normal University)”, (编号: 2022JBXC002), 本文系“2022 年度内蒙古师范大学研究生科研创新基金资助项目(Graduate students' research & Innovation fund of Inner Mongolia Normal University): 基于认知诊断理论的小学生劳动素养评价指标体系构建(项目编号: CXJJS22015)的部分成果。

参考文献

- 陈秋梅, 张敏强(2010). 认知诊断模型发展及其应用方法述评. *心理科学进展*, 18(3), 522-529.
- 单昕彤, 谭辉晔, 刘永, 吴方文, 涂冬波(2014). 项目反应理论中模型-资料拟合检验常用统计量. *心理科学进展*, 22(8), 1350-1362.
- 范士青, 刘华山(2015). 常见的认知诊断模型及其比较. *教育测量与评价(理论版)*, (7), 4-9.
- 范婷婷, 孙波, 曾用强(2023). 英语阅读测试中不同认知诊断模型比较研究. *中国考试*, (5), 53-61.
- 高旭亮, 涂冬波(2017). 参数化认知诊断模型: 心理计量特征、比较及其转换. *江西师范大学学报(哲学社会科学版)*, 50(1), 88-104.
- 高旭亮, 汪大勋, 蔡艳, 涂冬波(2018). 认知诊断模型的比较及其应用研究: 饱和模型、简化模型还是混合方法. *心理科学*, 41(3), 727-734.
- 康春花, 辛涛, 田伟(2013). 小学数学应用题认知诊断测验编制及效度验证. *考试研究*, (6), 24-43.
- 李心钰, 陆宏(2023). 教育评价理念的演变及测评工具的审视与应用. *现代教育技术*, 33(4), 74-82.
- 骈扬, 余胜泉, 李晟(2023). 基于认知诊断的有效作业精准化设计. *中国远程教育*, 43(8), 13-24.
- 涂冬波, 蔡艳, 戴海琦, 丁树良(2012). 一种多策略认知诊断方法: MSCD 方法的开发. *心理学报*, 44(11), 1547-1553.
- 涂冬波, 蔡艳, 丁树良(2021). *认知诊断理论*. 北京师范大学出版社.
- 万海鹏, 陈思睿, 余胜泉(2022). 基于认知诊断的学科核心素养测量路径构建. *电化教育研究*, 43(11), 92-99.

- 汪大勋, 肖清文, 谭青蓉, 蔡艳, 涂冬波(2023). 非补偿的非参数化多策略认知诊断方法: NCNPMSC 方法开发. *心理科学*, 46(4), 971-979.
- 汪文义, 郑娟娟, 宋丽红, 胡海洋(2022). 认知诊断模型参数估计算法比较. *江西师范大学学报(自然科学版)*, 46(3), 291-299.
- 武小鹏, 张怡, 张晨璐(2020). 核心素养的认知诊断测评体系建构. *现代教育技术*, 30(2), 42-49.
- 俞国良, 何妍(2023). “德智体美劳”五育与心理健康教育的关系. *中小学心理健康教育*, (1), 4-10.
- 詹沛达, 边玉芳, 王立君(2016). 重参数化的多分属性诊断分类模型及其判准率影响因素. *心理学报*, 48(3), 318.
- De La Torre, J. (2011). The Generalized DINA Model Framework. *Psychometrika*, 76, 179-199. <https://doi.org/10.1007/s11336-011-9207-7>
- De La Torre, J., & Minchen, N. (2014). Cognitively Diagnostic Assessments and the Cognitive Diagnosis Model Framework. *Psicología Educativa*, 20, 89-97. <https://doi.org/10.1016/j.pse.2014.11.001>
- Haertel, E. H. (1989). Using Restricted Latent Class Models to Map the Skill Structure of Achievement Items. *Journal of Educational Measurement*, 26, 301-321. <https://doi.org/10.1111/j.1745-3984.1989.tb00336.x>
- Hartz, S. M. (2002). *A Bayesian Framework for the Unified Model for Assessing Cognitive Abilities: Blending Theory with Practicality*. University of Illinois at Urbana-Champaign.
- Lee, Y. W., & Sawaki, Y. (2009). Application of Three Cognitive Diagnosis Models to ESL Reading and Listening Assessments. *Language Assessment Quarterly*, 6, 239-263. <https://doi.org/10.1080/15434300903079562>
- Leighton, J. P., & Gierl, M. J. (2007). Defining and Evaluating Models of Cognition Used in Educational Measurement to Make Inferences about Examinees' Thinking Processes. *Educational Measurement: Issues and Practice*, 26, 3-16. <https://doi.org/10.1111/j.1745-3992.2007.00090.x>
- Nichols, P. D. (1994). A Framework for Developing Cognitively Diagnostic Assessment. *Review of Educational Research*, 64, 575-603. <https://doi.org/10.3102/00346543064004575>
- Nichols, P. D., & Joldersma, K. (2008). *Cognitive Diagnostic Assessment for Education: Theory and Applications*.
- Ravand, H., & Robitzsch, A. (2019). Cognitive Diagnostic Modeling Using R. *Practical Assessment, Research, and Evaluation*, 20, 1-12.
- Tatsuoka, K. K. (1995). Architecture of Knowledge Structures and cognitive Diagnosis: A Statistical Pattern Recognition and Classification Approach. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively Diagnostic Assessment* (pp. 327-359). Lawrence Erlbaum Associates, Inc.
- Templin, J. L., & Henson, R. A. (2006). Measurement of Psychological Disorders Using Cognitive Diagnosis Models. *Psychological Methods*, 11, 287-305. <https://doi.org/10.1037/1082-989X.11.3.287>
- Yamaguchi, K., & Okada, K. (2020). Variational Bayes Inference for the DINA Model. *Journal of Educational and Behavioral Statistics*, 45, 569-597. <https://doi.org/10.3102/1076998620911934>
- Zhan, P., Jiao, H., & Liao, D. (2018). Cognitive Diagnosis Modeling Incorporating Item Response Times. *British Journal of Mathematical and Statistical Psychology*, 71, 262-286. <https://doi.org/10.1111/bmsp.12114>