

文章编号: 1000-5862(2018)02-0134-05

# 属性多级和项目多级评分的认知诊断模型

夏梦连<sup>1</sup>, 毛秀珍<sup>2\*</sup>, 杨 睿<sup>2</sup>

(1. 四川师范大学教师教育与心理学院, 四川 成都 610068; 2. 四川师范大学教育科学学院, 四川 成都 610068)

**摘要:** 认知诊断理论中属性多级和项目多级评分均能极大地丰富项目信息, 提高诊断结果准确度. 基于此对已有认知诊断模型扩展的属性多级和项目评分多级的认知诊断模型进行介绍, 并基于认知诊断模型和相关理论研究视角, 对当前认知诊断的发展方向提出思考和建议.

**关键词:** 认知诊断理论; 属性多极化; 多级评分项目; 认知诊断模型

**中图分类号:** B 841.7    **文献标志码:** A    **DOI:** 10.16357/j.cnki.issn1000-5862.2018.02.03

## 0 引言

经典测量理论(classical test theory, CTT)只报告测验总分. 项目反应理论(item response theory, IRT)基于被试在项目水平上的反应, 利用项目反应模型分析被试整体或多个领域上的潜在能力水平. 无论是 CTT 还是 IRT, 均未能反映被试作答项目时的心理特征或认知过程, 更不能获得被试在细粒度知识点上的掌握情况. 然而, 随着测量理论和教育事业的发展, 教育实践者越来越希望通过被试外在反应来了解其内部知识掌握状态, 通过知识链间的关系获取更恰当的学习路径和学习策略, 更好地指导教学实践. 于是, 认知诊断理论(cognitive diagnostic theory, CDT)在弥补现有理论的不足和顺应实践要求的情况下应运而生.

一般地, CDT 将被试的潜在特质分为多个相关(或独立)的维度, 并表示为多维连续向量或多维 2 分向量. 同样, 它视项目测量为多个维度或属性, 并用多个维度的属性参数来刻画项目特征. 认知诊断的任务在于挖掘项目特征、被试特质与作答反应之间的内在联系, 进而根据被试的作答反应推断其掌握了哪些属性、没有掌握哪些属性. 进而分析学习过程、探索有效的学习路径, 进行补救教学, 实现教学目的. 其中, 认知诊断模型是认知诊断方法的基础, 是决定诊断结果的准确性和有效性的主要因素之一. 然而, 测验情景包罗万象, 测量模型总是基于不

同视角去拟合测验情景. 于是, 不同诊断模型通常具备不同意义的模型参数, 采用不同方式表征作答反应过程, 也表现出不同的模型复杂度. 因此, 了解认知诊断模型特征, 不仅有助于实践者选用模型, 还是改进与发展新认知诊断模型的基础.

诊断模型的开发是认知诊断的核心, 目前已提出了大量认知诊断模型. 纵观认知诊断的应用研究可知, 逻辑斯蒂克潜在特质模型(LLTM)<sup>[1]</sup>、规则空间模型(RSM)<sup>[2]</sup>、决定性输入, “噪音”与门模型(DINA)<sup>[3]</sup>、重参数化融合模型(RUM)<sup>[4]</sup>、属性层级模型(AHM)<sup>[5]</sup>是实践中广泛采用的几类认知诊断模型. 郭磊等<sup>[6]</sup>曾指出当前认知诊断模型过于简单, 没有考虑到很多实际测验情景. 例如, 针对多项选择项目, 大部分模型没有考虑到干扰选项提供的诊断信息; 又如, 大部分模型没有考虑作答项目的多策略问题和项目采用连续计分的情况. 针对这些不足, 郭磊等<sup>[6]</sup>介绍了认知诊断模型在处理多项选择项目、多策略解题和连续性测验数据方面的进展.

当前, 大部分认知诊断模型将属性划分为掌握和未掌握的情况. 事实上, 人们常会为此假设感到无奈. 因为, 属性的掌握与否常常不能一刀切, 它常常表现为一种程度而不是一个断点. 正如实践教学讲授一个知识点总包含着认识、理解和应用 3 个部分. 另外, 认知诊断分析方法对属性个数的限制常常影响属性粒度的大小. 这就使得将属性描述为掌握与否显得不合实际. 因此, 将属性掌握划分为多个等级水平, 更符合和满足人们对信息挖掘的需求.

收稿日期: 2017-11-02

基金项目: 国家自然科学基金(31400897, 11201313)资助项目.

通信作者: 毛秀珍(1980-), 女, 四川成都人, 副教授, 博士, 主要从事心理测量与评价方面的研究. E-mail: maomao\_wanli@

163.com

另外,除2级评分项目外,多级评分项目是另一类广泛使用的测验项目. Lin Haiyan<sup>[7]</sup>指出自从在K-12系统和其它高风险测验中使用多级评分项目,测量多种能力和使用多级评分项目的测验变得越来越重要.特别地,尽管多级评分项目的开发费用更高,但它与2级评分项目相比具有很多优势.例如,多级评分项目与相同项目2分化后相比,能提供更多信息量<sup>[8]</sup>. P. W. van Rijn等<sup>[9]</sup>发现,与2级评分项目库相比,多级评分项目库对极端能力被试的估计误差更小.开发出适合多级评分项目的认知诊断模型将极大地推动认知诊断的实践应用.

基于属性多级和项目多级评分,详细介绍了DINA模型、DINO模型、RUM模型和AHM模型的新发展,并分析新扩展模型的特征.基于认知诊断理论在诊断模型、参数估计、模型拟合、项目功能差异检验、 $Q$ 矩阵设计等方面对认知诊断理论的研究提供一些思路和建议.

## 1 属性多极化认知诊断模型

### 1.1 属性多极化认知诊断模型的基本概念

在传统0-1掌握模式 $Q$ 矩阵基础上,研究者们提出属性多级化的 $Q$ 矩阵思想<sup>[10-12]</sup>,以细化项目测量属性的水平层次.在多级化 $Q$ 矩阵中, $Q$ 矩阵的元素 $q_{jk}$ 代表项目测量的属性水平,取值为 $0 \sim T$  ( $T$ 表示属性的层次水平数).如 $q_{jk} = t$  ( $t = 0, 1, 2, \dots, T$ )代表项目 $j$ 测量属性 $k$ 的第 $t$ 个水平层次.相应地,被试掌握属性的水平也划分为多个层次,即 $\alpha_{ik}$ 取值为 $0 \sim T$ 表示被试 $i$ 在属性 $k$ 上的掌握水平.

### 1.2 属性多极化DINA模型

在众多CDM中,DINA模型因其简约性受到了研究者的广泛关注,近年来有大量关于DINA模型的研究<sup>[3,13]</sup>,有研究表明<sup>[14-15]</sup>:该模型具有较高的判准率.

蔡艳等<sup>[16]</sup>针对属性多极化情形,基于DINA模型提出了属性多极化DINA模型(DINA model for polytomous attributes, PA-DINA).该模型假设被试要正确答对项目,则应掌握项目所考察属性的相应水平层次.被试的知识状态与潜在反应向量具有如下关系:

$$\alpha_{ik}^* = \begin{cases} 0 & \alpha_{ik} < q_{jk} \\ 1 & \text{其它} \end{cases}, \quad (1)$$

$$q_{jk}^* = \begin{cases} 0 & q_{jk} = 0 \\ 1 & \text{其它} \end{cases}, \quad (2)$$

$$\eta_{ij}^* = \prod_{k=1}^K \eta_{ijk}^* = \prod_{k=1}^K (\alpha_{ik}^*)^{q_{jk}^*},$$

$\eta_{ij}^*$ 与传统0-1化属性下的 $\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$ 的含义基本一致,即描述被试是否达到了项目测量属性所需的水平数,于是PA-DINA模型表示为 $P(X_{ij} = 1 | \alpha_i) = (1 - s_j)^{\eta_{ij}^*} g_j^{1-\eta_{ij}^*}$ .在PA-DINA模型中, $\alpha_{ik}$ 和 $q_{jk}$ 的取值有 $L_k$ 种( $L_k$ 指属性 $k$ 的水平数为 $L$ ,  $L \geq 2$ ).多级的 $\alpha_{ik}$ 和 $q_{jk}$ 需要转换为0-1的 $\alpha_{ik}^*$ 和 $q_{jk}^*$ ,其方法见(1)式和(2)式. $s_j$ 为失误参数, $g_j$ 为猜测参数.

PA-DINA模型在DINA模型基础上以属性具有多种掌握水平为视角进行扩展而得.它将测验考察属性的程度划分为2个及以上的水平,进而推断被试掌握属性的水平.与DINA模型2分化属性相比,它更加细致和准确地体现了被试的知识状态.同时,项目考察的属性水平也进行相应的细化,极大地丰富项目信息.然而,PA-DINA与DINA模型只能将每个项目的作答概率分为2类,即当被试达到项目 $i$ 要求的所有属性水平时,他正确作答项目的概率为 $1 - s_i$ ;若对项目 $i$ 考察的所有属性水平,被试至少在1个属性上没有达到项目考察的水平,则他正确作答该项目的概率为 $g_i$ .因而,PA-DINA没能精细地区分不同属性掌握水平的被试正确作答同一项目的概率,降低了项目的区分度.此外,根据PA-DINA的特点,分析数据时除了需要分析项目考察的属性,还需要刻画属性水平分析测验的多级化 $Q$ 矩阵,无疑都增加测验设计和编制的困难.但这些尝试和探讨,都有待更多的研究者参与进来,一步一步地完善和改进模型、思考更有效的方法和技术解决测验编制和数据分析的困难.

### 1.3 属性多极化RUM模型

除了DINA模型,蔡艳等<sup>[16]</sup>还以RUM模型为基础,提出属性多极化RUM模型(RUM for polytomous attributes, PA-RUM),即

$$P(X_{ij} = 1 | \alpha_i) = \pi_j^* \prod_{k=1}^K r_{jk}^* q_{jk}^* (1 - \alpha_{ik}^*). \quad (3)$$

对比被试属性掌握水平和多极化 $Q$ 矩阵,依据(1)式和(2)式可以获得2分化的掌握水平 $\alpha_{ik}^*$ 和项目 $q$ 向量的元素 $q_{jk}^*$ . (3)式中 $\pi_j^*$ 和 $r_{jk}^*$ 分别为项目难度参数和属性区分度参数. $\pi_j^*$ 越大说明项目越容易,一个项目仅有一个难度参数.对于未掌握某属性的惩罚参数(penalty parameter)定义 $r_{jk}^*$ ,可以看成是属性 $k$ 在项目 $j$ 中的区分度指标,其值域为 $[0, 1]$ , $r_{jk}^*$ 越小说明掌握属性 $k$ 对正确回答项目 $j$ 越重要,越能区分答对与答错该题的被试.

PA-RUM 与 RUM 模型都将失误和猜测概率定义在项目水平上. 它们既能区分不同属性掌握模式的被试正确作答同一项目的概率, 又能区分同一被试在包含相同属性的不同项目上的作答情况, 更符合测验实际. PA-RUM 与 PA-DINA 模型都以繁杂精细的测验编制和数据分析为基础, 以期获得更精细的被试和项目信息为目的. 相比 RUM 与 DINA 模型更符合实践需要, 丰富认知诊断理论的实践价值.

#### 1.4 属性多级化认知诊断模型的应用与简评

属性多极化认知诊断模型在 2 分属性水平认知诊断模型基础上扩展, 将项目和被试特征推广到多个水平层次, 能提供更为丰富和细致的诊断信息. 与 2 级评分模型相比, 虽然属性多极化模型更加复杂, 但有研究<sup>[16-19]</sup> 表明它们能获得较高的模式判准率、较高的参数估计精度和较好的稳定性, 具有广阔的应用前景.

目前, 实践研究已尝试使用属性多级化模型. 例如, Chen Jinsong 等<sup>[12]</sup> 运用 PG-DINA 模型测量了 8 年级比例推理的掌握情况. 李丽<sup>[18]</sup> 在研究中学生的空间能力掌握水平时, 运用属性多级化思想, 将属性掌握水平划分为多个阶段, 对学生的空间知觉、空间记忆、空间想象和空间思维能力水平进行了精细分析. 结果表明, 学生的空间知觉能力的属性掌握水平中等偏低, 空间记忆和想象能力属性掌握水平较高, 空间思维能力大部分处于描述推理技能、结构关联推理技能水平阶段. 除了在纸笔测验上的应用, 涂冬波等<sup>[19]</sup> 在传统 CD-CAT 基础上, 首次开发设计了属性多级化的认知诊断计算机化自适应测验. 詹沛达等<sup>[17]</sup> 还进一步对属性多极化 DINO、DINA 和线性逻辑斯蒂克模型进行重新参数化. 由此可见, 属性多级化的认知诊断模型符合实践发展需求, 在提供被试分类信息的同时更有助于教师开展补救教学.

## 2 多级评分认知诊断模型

2 级和多级评分项目是教育、心理、医学及社会测量等领域广泛使用的 2 种计分方式的项目. 然而, 大部分认知诊断模型, 包括属性多级化认知诊断模型都局限于分析 2 级评分项目. 认知诊断理论在实践中必将面临多级评分项目, 因而在已有模型的基础上, 扩展或开发新的适合于分析处理多级评分项目潜在特质模型是研究者们关注的焦点之一.

#### 2.1 多级评分 DINA 模型

涂冬波等<sup>[20]</sup> 借鉴累积类别反应函数的思想将

0-1 评分的 DINA 模型扩展到多级评分 DINA 模型 (记为 P-DINA 模型), 其概率反应函数为  $P(X_{ij} = t | \alpha_i) = P^*(X_{ij} = t | \alpha_i) - P^*(X_{ij} = t + 1 | \alpha_i)$ ,  $P^*(X_{ij} = t | \alpha_i) = (1 - s_{ji})^{\eta_{ij}} g_{ji}^{1-\eta_{ij}}$ , 其中  $P^*(X_{ij} = t | \alpha_i)$  表示累积类别得分概率, 即表示被试得分  $\geq t$  ( $t = 0, 1, 2, \dots, K-1$ ) 分的概率, 于是被试得  $t$  分的概率  $P(X_{ij} = t | \alpha_i)$  等于相邻 2 个类别得分概率之差. 在 DINA 模型中项目的猜测和失误参数相应地推广到每个得分上的猜测和失误参数. 例如,  $s_{ji}$  表示被试掌握了项目  $j$  考察的所有属性, 但在得  $t$  分上失误的概率;  $g_{ji}$  表示被试未掌握项目  $j$  考察的所有属性, 但在得  $t$  分上猜对的概率.

类似于等级反应模型, P-DINA 模型假设被试得 0 分及以上的概率为 1, 得  $K$  分以上的概率为 0; 当满分为 1 时, P-DINA 模型便简化为 DINA 模型, 因此 P-DINA 模型既能处理 0-1 评分的数据资料, 又能用于多级评分资料的数据处理, 还适用于混合数据资料.

#### 2.2 多级评分 DINO 模型

目前, 认知诊断理论主要应用在对知识掌握情况和技能的测验方面, 而在对潜在特质 (如性格、情绪、动机等) 测量中的应用非常少. 针对传统心理学量表的数据分析, 期望通过运用诊断理论和方法以获得隐藏的、有意义的潜在信息. J. L. Templin 等<sup>[21]</sup> 提出了 DINO 模型. 考虑到心理特质的补偿性特征, 吴方文等<sup>[22-23]</sup> 以 DINO 模型为基础借鉴等级反应模型中累积类别反应函数的思想, 将 0-1 评分的 DINO 模型拓展到多级评分的 P-DINO 模型:  $P(X_{ij} = t | \alpha_i) = P^*(X_{ij} = t | \alpha_i) - P^*(X_{ij} = t + 1 | \alpha_i)$ ,  $P^*(X_{ij} = t | \alpha_i) = (1 - s_{ji})^{\omega_{ij}} g_{ji}^{1-\omega_{ij}}$ .

DINO 模型和 DINA 模型非常相似. 它们的不同在于 DINA 模型中被试掌握项目考察的所有属性正确作答项目的概率就等于 1 减去失误参数, 而 DINO 模型假设被试只要掌握至少 1 个项目  $j$  考察的属性, 其正确作答项目的概率就等于 1 减去失误参数. 因此, 它们主要区别在于, 在 DINO 模型中  $\omega_{ij} = 1 - \prod_{k=1}^K (1 - \alpha_{ik})^{q_{jk}}$  描述被试是否至少掌握了 1 个项目  $j$  考察的属性, 在 DINA 模型中  $\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$  表示被试是否掌握了项目考察的所有属性.

#### 2.3 多级评分 AHM 模型

基于等级反应模型, 祝玉芳等<sup>[24]</sup> 在 AHM 中使用对数似然比作为判别被试知识状态的方法, 即

$$LL(X_\alpha, V_\beta) = \sum_{j=1}^J \left| \log\left(\prod_{t=0}^{f_j} P_{\alpha_{jt}}^{x_{\alpha_{jt}}}\right) - \log\left(\prod_{t=0}^{f_j} P_{\beta_{jt}}^{v_{\beta_{jt}}}\right) \right|. \quad (4)$$

当  $LL(X_\alpha, V_\beta)$  中期望项目反应模式  $V_\beta$  达到最小时,与之对应的为被试  $\alpha$  的属性掌握模式. 其中,  $X_\alpha$  是被试  $\alpha$  ( $\alpha = 1, \dots, N$ ) 的观察得分向量,  $V_\beta$  ( $\beta = 1, 2, \dots, R$ ) 是  $\beta$  类期望得分向量 ( $R$  代表期望项目反应模式数).  $f_j$  ( $f_j \geq 1, j = 1, 2, \dots, J$ ) 是第  $j$  个项目的满分 ( $J$  是项目个数). (4) 式中  $P_{\alpha_{jt}}$  与  $P_{\beta_{jt}}$  是根据等级反应模型计算所得的在项目  $j$  上得  $t$  分的概率  $x_{\alpha_{jt}}$  和  $v_{\beta_{jt}}$  的值非 0 即 1. 若被试  $\alpha$  在第  $j$  个项目上的得分为  $m$ , 记  $X_{\alpha_j} = m$ , 则  $x_{\alpha_{jm}} = 1$ , 而当  $t \neq m$  时  $x_{\alpha_{jt}} = 0$  ( $0 \leq t \leq f_j$ ); 若  $V_\beta$  的第  $j$  分量为  $m$ , 即  $V_{\beta_j} = m$ , 则  $v_{\beta_{jm}} = 1$ , 而当  $t \neq m$  时  $v_{\beta_{jt}} = 0$  ( $0 \leq t \leq f_j$ ).

#### 2.4 多级评分认知诊断模型的应用

针对多级评分项目,研究者通常将多级评分项目得分转化为 0-1 评分,然后采用 2 级评分模型分析数据.这无疑是精细信息转化为粗略信息,从而导致误差增加.例如,吴方文等<sup>[23]</sup>利用 P-DINO 模型和 DINO 模型对抑郁症状评估数据作比较时发现,与 DINO 模型相比, P-DINO 模型不仅在整体上拟合更好,而且还能提供更详尽的剖面信息,并为医生的诊断分析和选择治疗措施时提供有效指导.此外,蔡艳等<sup>[25]</sup>在 0-1 评分的 CD-CAT 基础上,拓展出适合多级评分项目的 CD-CAT 及选题策略.

### 3 总结与展望

认知心理学和心理测量学为认知诊断理论的产生、发展和应用提供了理论依据、方法和技术支持.认知心理学关于认知过程的理论为细化测验特征、提取和分析被试和项目属性、建立  $Q$  矩阵提供了诊断分析依据.在此基础上,心理测量学通过概率模型去反应实际中被试知识状态、项目特征与被试作答反应之间的关系,并基于数理分析的原理、方法和技术实现概率模型的参数估计,进而获取项目测量学特征和被试认知结构的信息.由此可见,认知诊断概率模型的建立是诊断分析的核心和基础.优良的诊断模型不仅能合理揭示被试作答项目的认知加工过程,还能体现测验过程中的随机误差,准确可靠地反应被试和项目信息.诊断模型的建立不仅要考虑模型的复杂度、参数的可解释性,还要考虑参数估计的可能性、准确度和稳健性.因而,改进已有模型、探索新的认知诊断模型一直都是广大研究者关注的重点.

针对 2 级评分项目,研究者提出了大量认知诊

断项目.随着研究深入,针对多项选择的干扰项问题、解题的多策略问题和连续计分的认知诊断问题,开发了众多认知诊断模型.针对属性多级和项目多级的情况,本文介绍了认知诊断模型的进展.首先,属性多级化将 2 分属性掌握情况扩展到多分类属性掌握水平,更加符合认知过程的发展,也能获得更为丰富和细致的诊断信息;其次,多级评分项目的认知诊断模型极大地拓宽了传统认知诊断模型的应用范围,使认知诊断不再局限于分析多项选择、判断正误、填空等项目类型.实际上,测验实践越来越多地采用如操作测验、表现性测验等多级评分的项目类型.因而,认知诊断模型在属性多级化和项目多级评分方面的发展,极大地推动了认知诊断理论的实践,体现其实践价值.

认知诊断模型未来研究趋势不局限上述提及的测验因素,针对属性不同权重、题组项目、反应时间、反应风格、项目内容预知、作弊等情况都可以提出或改进已有认知诊断模型.除了诊断模型的研究、诊断模型的模型、项目和个人拟合指标,诊断测验的等值、诊断分析中  $Q$  矩阵的效度和数理分析,诊断项目的项目功能差异检验,甚至认知诊断理论相关的理论、技术和问题在计算机化自适应测验背景下的情况等,都值得研究者们进行长期深入地研究.

### 4 参考文献

- [1] Fischer G H. The linear logistic test model as an instrument in educational research [J]. *Acta Psychologica*, 1973, 37(6): 359-374.
- [2] Tatsuo K K. Rule space: an approach for dealing with misconceptions based on item response theory [J]. *Journal of Educational Measurement*, 1983, 20(4): 345-354.
- [3] Junker B W, Sijtsma K. Cognitive assessment models with few assumptions and connections with nonparametric item response theory [J]. *Applied Psychological Measurement*, 2001, 25(3): 258-272.
- [4] Hartz M C. A Bayesian framework for the unified model for assessing cognitive abilities: blending theory with practicality [J]. *American Journal of Gastroenterology*, 2002, 95(4): 906-909.
- [5] Leighton J P, Gierl M J, Hunka S M. The attribute hierarchy method for cognitive assessment: a variation on Tatsuo's rule-space approach [J]. *Journal of Educational Measurement*, 2004, 41(3): 205-237.
- [6] 郭磊,苑春永,边玉芳.从新模型视角探讨认知诊断的发展趋势 [J]. *心理科学进展*, 2013, 21(12): 2256-2264.

- [7] Lin Haiyan. Item selection methods in multidimensional computerized adaptive testing adopting polytomously-scored items under multidimensional generalized partial credit model [J]. Dissertations & Theses-Gradworks, 2012, 5(4): 392-403.
- [8] Samejima F. Graded response model of the latent trait theory and tailored testing [C]// Clark C K. Proceedings of the first conference on computerized adaptive testing. Washington, DC: Government Printing Office, 1976: 5-17.
- [9] Van Rijn P W, Eggen T J H M, Hemker B T et al. Evaluation of selection procedures for computerized adaptive testing with polytomous items [J]. Applied Psychological Measurement 2002, 26(4): 393-411.
- [10] Karelitz T M. Ordered category attribute coding framework for cognitive assessments [EB/OL]. [2017-09-13]. 10.13140/RG.2.1.4628.9440.
- [11] de la Torre J, Lam D, Rhoads K et al. Measuring grade 8 proportional reasoning: The process of attribute identification and task development and validation [C]. The Annual Meeting of the American Educational Research Association. Denver: CO, 2010.
- [12] Chen Jinsong, Jimmy D L T. A general cognitive diagnosis model for expert-defined polytomous attributes [J]. Applied Psychological Measurement 2013, 37(6): 419-437.
- [13] Torre J D L. DINA model and parameter estimation: A didactic [J]. Journal of Educational & Behavioral Statistics, 2009, 34(1): 115-130.
- [14] Cheng Ying. Computerized adaptive testing: new developments and applications [EB/OL]. [2017-08-19]. <https://www.ideals.illinois.edu/bitstream/handle/2142/82159/3337733.pdf?sequence=2&isAllowed=n>.
- [15] Rupp A A, Templin J. The effects of  $Q$ -matrix misspecification on parameter estimates and classification accuracy in the DINA model [J]. Educational & Psychological Measurement 2008, 68(1): 78-96.
- [16] 蔡艳, 涂冬波. 属性多级化的认知诊断模型拓展及其  $Q$  矩阵设计 [J]. 心理学报, 2015, 47(10): 1300-1308.
- [17] 詹沛达, 边玉芳, 王立君. 重参数化的多属性诊断分类模型及其判准率影响因素 [J]. 心理学报, 2016, 48(3): 318-330.
- [18] 李丽. 基于属性多级化模型的中学生空间能力认知诊断研究 [D]. 南昌: 江西师范大学, 2016.
- [19] 涂冬波, 蔡艳. 基于属性多级化的认知诊断计算机化自适应测验设计与实现 [J]. 心理学报, 2015, 47(11): 1405-1414.
- [20] 涂冬波, 蔡艳, 戴海琦, 等. 一种多级评分的认知诊断模型: P-DINA 模型的开发 [J]. 心理学报, 2010, 42(10): 1011-1020.
- [21] Templin J L, Henson R A. Measurement of psychological disorders using cognitive diagnosis models [J]. Psychological Methods 2006, 11(3): 287-305.
- [22] 吴方文. 补偿型的多级评分认知诊断模型的开发及其在抑郁症状评估中的应用 [D]. 南昌: 江西师范大学, 2015.
- [23] 吴方文, 涂冬波, 刘明矾. 分离型的多级评分认知诊断模型开发及其应用研究 [J]. 心理科学, 2017, 40(1): 209-215.
- [24] 祝玉芳, 丁树良. 基于等级反应模型的属性层级方法 [J]. 心理学报, 2009, 41(3): 267-275.
- [25] 蔡艳, 苗莹, 涂冬波. 多级评分的认知诊断计算机化适应测验 [J]. 心理学报, 2016, 48(10): 1338-1346.

## Cognitive Diagnosis Models Under Polytomous Attributes and Polytomous Item

XIA Menglian<sup>1</sup>, MAO Xiuzhen<sup>2\*</sup>, YANG Rui<sup>2</sup>

(1. School of Teacher Education and Psychology, Sichuan Normal University, Chengdu Sichuan 610068, China;

2. School of Educational Science, Sichuan Normal University, Chengdu Sichuan 610068, China)

**Abstract:** Polytomous attributes and polytomous item in cognitive diagnosis theory can greatly enrich the item information and improve the diagnosis accuracy. Firstly, polytomous attributes and polytomous models are introduced which are extended based on cognitive diagnosis model. At last, based on the new models, some future research directions and suggestions are put forward to cognitive diagnostic.

**Key words:** cognitive diagnostic theory; polytomous attributes; polytomous item; cognitive diagnosis model

(责任编辑: 冉小晓)