

无领导小组讨论的发展历程与权重分析

李璇^{1*}, 吴汶洁²

¹重庆大学公共管理学院, 重庆

²山东理工大学管理学院, 山东 淄博

收稿日期: 2025年5月13日; 录用日期: 2025年6月27日; 发布日期: 2025年7月9日

摘要

无领导小组讨论(LGD)作为一种重要的人才评估工具, 具有独特优势。本文深入探讨其发展历程, 从起源到在国内外的广泛应用与改进。在研究现状方面, 分析国内外学者对LGD技术的研究差异, 指出国内多单独研究其测评要素、信效度及考官评分一致性等, 国外则将其作为评价中心技术一部分进行更深入研究。构建无领导小组讨论面试测评要素模型, 包括五个一级要素变量和24个二级要素变量, 并运用问卷调查、因子分析和AHP层次分析法等方法, 对测评要素的重要性进行计算和分析, 得出各要素权重。最后提出建议, 如加大技术投入、合理选择测评要素、增加评价者人数并加强培训等, 以提升LGD的有效性和实用性, 为企业和组织的人才选拔、评估及培训提供参考, 推动该技术在未来人力资源管理中发挥更重要作用。

关键词

无领导小组讨论, 测评要素模型, 权重计算

Development and Weighting Analysis of Leaderless Group Discussions

Xuan Li^{1*}, Wenjie Wu²

¹School of Public Policy and Administration, Chongqing University, Chongqing

²School of Management, Shandong University of Technology, Zibo Shandong

Received: May 13th, 2025; accepted: Jun. 27th, 2025; published: Jul. 9th, 2025

Abstract

Leaderless Group Discussion (LGD) has unique advantages as an important talent assessment tool. This paper discusses its development history in depth, from its origin to its wide application and

*通讯作者。

文章引用: 李璇, 吴汶洁. 无领导小组讨论的发展历程与权重分析[J]. 社会科学前沿, 2025, 14(7): 39-47.

DOI: 10.12677/ass.2025.147585

improvement at home and abroad. In terms of the current research status, it analyzes the research differences between domestic and foreign scholars on LGD technology, pointing out that domestic scholars mostly study its assessment elements, reliability and validity, and consistency of the examiner's score separately, while foreign scholars study it more deeply as part of the evaluation center technology. A model of LGD assessment elements is constructed, including five first-level element variables and 24 second-level element variables, and the importance of the assessment elements is calculated and analyzed by using questionnaire research, factor analysis and AHP hierarchical analysis to derive the weights of each element. Finally, suggestions are made, such as increasing technical input, rationally selecting assessment elements, increasing the number of evaluators and strengthening training, etc., in order to enhance the validity and practicality of LGD, provide reference for talent selection, assessment and training of enterprises and organizations, and promote this technology to play a more important role in future human resource management.

Keywords

Leaderless Group Discussion, Modeling of Assessment Elements, Weighting Calculation

Copyright © 2025 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

无领导小组讨论(Leaderless Group Discussion, LGD)作为现代人才评估体系中的重要工具,采用非结构化集体研讨形式,通过情境化行为诱导实现对个体特质的深度解析。该方法通过模拟真实团队决策场景,系统观测候选人的领导潜质、互动策略及协同效能,其核心价值在于突破传统笔试与绩效评估的局限,能够有效捕捉人际互动中的隐性能力维度。邹聪在研究中强调,该技术通过预设议题激发被试的群体表达,尤其擅长揭示个体在团队融合与矛盾协调方面的特质,这些正是常规评估手段难以量化的核心素质[1]。

在数字化浪潮与全球化进程的双重驱动下, LGD 技术正经历多维革新。技术演进层面,虚拟仿真系统可构建高复杂度决策环境,结合人工智能的情绪识别与微表情分析技术,显著提升行为观测的客观性;应用场域层面,多元文化背景下的组织管理需求催生出跨文化效度研究,为团队异质性与效能关联性探索开辟了新维度。这种技术融合与场景扩展的双向突破,使传统测评方法获得了更广阔的发展空间。

然而,尽管无领导小组讨论在理论和实践层面取得了长足的进步,其在测评标准的科学性、技术应用的伦理问题以及跨文化适应性等方面仍存在诸多挑战。本研究旨在梳理无领导小组讨论的理论脉络,探讨其在新发展背景下的应用潜力与局限性,并提出未来研究的方向,以期组织管理、人才发展和团队建设提供更具前瞻性的理论支持和实践指导。

2. 国内外研究现状

国内外对 LGD 技术的研究范畴是不同的。国外学者认为 LGD 技术是评价中心技术的一部分,所以并不是单独研究 LGD 技术,而是置于评价中心当中研究。从研究内容上看,国外学者对 LGD 的研究更为深入。不仅对 LGD 的要素、信效度研究较为深入,还把其他测评技术结合起来,进行相关性研究,试验其互相间的影响,是一种研究趋势。而在国内,通常是把 LGD 技术作为单独的一项测评技术进行研究。

2.1. 国内研究现状

根据顾大荣研究,我国无领导小组讨论面试法是在20世纪80年代末引入的,而在21世纪初逐渐纳入我国公务员遴选面试中[2]。在试验探索阶段,国家2001年颁布了全国公开选拔党政领导干部考试大纲(试行),2003年,广东省公选100名副县级干部,采用LGD测评程序,这次创新尝试是我国公选以来第一次运用LGD测评技术。在改进推广阶段,国家2004颁布了党政领导干部公开选拔和竞争上岗考试大纲,在《考试大纲》中对LGD等测评技术的实施程序进行了补充和完善,给今后的施测工作、评价工作提供了重要的参考价值。2009年9月至今是制度推行阶段,国家2009年颁布了党政领导干部公开选拔和竞争上岗考试大纲,各地各部门积极开展工作,推动无领导小组讨论的面试形式和评价体系进一步完善。

国内研究关于LGD方法的介绍性文章较多。顾大荣总结出“无领导小组讨论的优点在于情境模拟的真实性、讨论角色的平等性、正面竞争的直接性、考察要素的独特性、客观评价的准确性、考察效率的高效性[2]。缺点在于,对题目编制较困难、对评分者要求很高、评分要素定义困难、考生表现受到影响”。学者冯丽萍认为有经验的考官和没有经验的考官在评分差异上比较大,有经验的考官评分一致性高,没有经验的则低,且通过短期培训改变这种差异比较困难,但有无经验的考官在评分时间上均无明显差异[3]。彭平根等人也指出,评分者在评价合作能力、应变能力和总体印象这几个要素时,评价标准具有较高的一致性,但对决策能力等要素的评价标准的理解存在一定的偏差,启用4名考官可以达到预期目标,符合测评的有效性[4]。唐忠明在其实证研究中证明了不同的评价者在总体上其评分一致性非常好,但个别测评要素还存在差异;不同考官之间存在明显的评分差异(高低之分);评价者在评分时容易受到价值观以及职业倾向的影响[5]。董越在LGD中,分别考察主动性人格、行为对LGD个体绩效的影响,研究表明:外向性格并不能影响绩效,但主动性行为可以影响个体绩效[6]。

综上所述,国内在无领导小组讨论的研究上,除了对测评要素、信效度及考官评分一致性的深入探讨外,还关注了个体特质对讨论结果的影响。研究者们通过实证分析,逐步揭示了性格特征、行为模式等个体因素在无领导小组讨论中所扮演的角色。例如个体的主动性行为对于提升其在讨论中的绩效具有显著作用,而外向性格则并非影响绩效的关键因素。这些发现不仅丰富了无领导小组讨论的理论体系,也为实际应用中的候选人筛选提供了更为精细化的指导原则。

2.2. 国外研究现状

无领导小组讨论起源于德国选拔有指挥潜能的军官测试,后作为世界范围内广泛采用的面试形式,主要用于测评被测者动员小组成员达成共同合作的能力。最初是在20世纪70年代末和80年代初被引入到人才测评领域中的。起初,这种方法主要用于选拔高级管理人才和领导者。随着时间的推移,无领导小组讨论逐渐被应用于各种职业领域,包括销售、客户服务、研发等。同时,无领导小组讨论的形式和技巧也得到了不断的改进和优化。

在对无领导小组讨论(LGD)相关要素的深入研究中,学者Waldman的发现具有重要意义。他指出,LGD的四个要素之间的关联系数偏高,这一现象揭示了考官在实际评价过程中存在对不同测评要素区分不够的问题。深入探究其背后的原因,首先,少数关键要素在LGD评价中占据主导地位,吸引了考官的主要注意力,从而使得其他要素的区分度受到影响[7]。其次,考官的评价往往基于被试者在LGD中的整体表现为依据,而非严格依据各个具体的考评要素来进行精准评估。

Lance从不同角度出發,提出了“情景的特定性假说”,强调评价效度的波动应从情景本身去探究原因,而非凭借预先设定的观念进行解释[8]。这一观点为后续研究提供了新的思路 and 方向。Gatewood等人运用情景评价的方法对LGD的信度展开研究,结果显示评分者一致性信度相较于以往有了显著提升,这

表明采用情景评价的方式有助于增强 LGD 的信度[9]。而在 Gaugler 和 Thornton 的研究中, 分别尝试使用 3、6、9 个要素进行评价, 结果清晰地表明, 当使用 6 个测评要素时, 行为观察的准确性达到最佳状态, 无论是使用 3 个要素还是 9 个要素, 都无法取得与 6 个要素相媲美的理想结果[10]。

综合国外众多学者的研究成果, 可以得出以下几点总结。首先, 在 LGD 的测评要素数量上, 3~6 个要素相对合理。若要素数量过少, 无法全面、深入地考查被试者; 反之, 若要素过多, 则会增加考官区分不同要素的难度, 进而影响评价的精准度。其次, 对每组评价者的人数进行分析发现, 2 人组的评价者信度最低, 而 6 人组的信度最高。这可能是因为人数较多的评价者小组能够通过充分的讨论和交流, 对被试者的表现进行更全面、客观的评估, 从而提高信度。

2.3. 文献述评

综合来看, LGD 技术具备诸多优势。其多人同时施测的模式, 有效降低了时间成本, 提高了测评效率。同时, 在效度方面, LGD 的表现也优于传统的笔试和面试方法, 其设计的情景能够紧密贴合工作实际, 为被试者提供一个模拟真实工作环境的平台, 使其能够充分展示自身的能力和潜力。国内关于无领导小组讨论的研究, 历经试验探索、改进推广到制度推行三个阶段, 逐步推动了该面试形式及评价体系的完善。在研究内容方面, 既有对 LGD 优点如情境模拟真实性、讨论角色平等性等的总结, 也客观指出其在题目编制、评分者要求等方面的不足。此外, 对 LGD 在不同行业、职位选拔中的应用差异也进行了积极探索, 为其在人力资源管理领域的广泛应用提供了科学依据。国外研究则聚焦于 LGD 的测评要素数量、评价者信度等方面, 得出 3~6 个要素相对合理、6 人组评价者信度最高等结论。整体而言, 国内外对无领导小组讨论的研究各有侧重且不断深入, 为这一面试形式的优化与发展奠定了坚实基础。

然而, 与国外在 LGD 领域的深入研究和广泛应用相比, 我国在 LGD 的理论研究和实际应用方面仍处于相对滞后的状态, 还有很大的提升空间和研究价值, 需要进一步投入精力进行深入探索和实践应用, 以推动我国在该领域的长足发展。

3. 无领导小组面试测评要素

3.1. 模型构建

根据学者魏玉华的面试测评维度表[11], 以及黄欣的面试测评要素定义[12], 归纳出本文无领导小组讨论测评要素。并根据成熟量表及专家咨询分析, 最终将本文无领导小组讨论面试的测评维度分为能力水平、知识水平、态度和价值观、个人品质、动机/内驱力五个一级要素变量; 以及创新决策能力等 24 个二级要素变量(如表 1)。

Table 1. Leaderless Group Discussion assessment dimension scale
表 1. 无领导小组讨论测评维度表

一级要素	二级要素	一级要素	二级要素
能力水平	创新决策能力	个人品质	积极乐观
	问题处理能力		富有自信
	沟通协调能力		机敏灵活
	统筹管理能力		客观冷静
	任务执行能力		耐心礼貌
	自我管理 ability		科学严谨
知识水平	基础知识	动机/内驱力	诚实守信
	专业知识		公平公正
	特殊知识		自我价值实现

续表

态度和价值观	工作态度 集体意识 服务意识 责任意识	职业兴趣 事业成就感
--------	------------------------------	---------------

3.2. 数据分析

通过问卷调研的方式，将最终形成的无领导小组讨论面试测评要素库整理成附件一的问卷，并采取方便抽样的方式，通过网络发放问卷，并剔除未参与无领导小组面试的数据，从剩下的样本中，研究不同测评要素在总体面试打分中的重要性。

1) 样本特征

样本的人口学变量特征如表 2：

Table 2. Characterization table of demographic variables of the sample
表 2. 样本人口统计学变量特征表

类别		人数	占比
性别	男	71	47.3
	女	79	52.7
年龄	18 岁及以下	2	1.3
	19~22 岁	46	30.7
	23~25 岁	98	65.4
	26 岁及以上	4	2.6
学历	博士研究生	6	3.9
	硕士研究生	96	64
	大学本科	46	30.6
	专科及以下	2	1.3

本研究采用李克特 5 点计分法以及方便取样的原则，同时兼顾性别、年龄、学历等因素，并采取自愿的方式，共发放网络问卷 167 份，其中有无领导小组讨论经历的问卷为 150 份，数据主要来源者学历为大学本科及硕士研究生。

2) 因子分析

对各一级要素及二级要素所涉及题项进行分析，所得结果如下。各级要素得分情况见表 3，可靠性检验见表 4，KMO 与 Bartlett 检验见表 5。

Table 3. Scoring of elements at various levels
表 3. 各级要素得分情况

一级要素	平均得分	二级要素	平均得分
能力水平	3.997	创新决策能力	4.05
		问题处理能力	3.75
		沟通协调能力	3.83
		统筹管理能力	3.79
		任务执行能力	4.24
		自我管理 ability	4.32
知识水平	3.83	基础知识	3.91

续表

态度和价值观	4.405	专业知识	3.72
		特殊知识	3.86
		工作态度	4.23
		集体意识	4.50
		服务意识	4.38
个人品质	4.306	责任意识	4.42
		积极乐观	4.54
		富有自信	4.38
		机敏灵活	4.51
		客观冷静	3.94
		耐心礼貌	4.16
		科学严谨	4.03
		诚实守信	4.28
		公平公正	4.34
动机/内驱力	4.37	自我价值实现	4.39
		职业兴趣	4.43
		事业成就感	4.29

Table 4. Reliability testing
表 4. 可靠性检验

Cronbach 的 Alpha	基于标准化项目的 Cronbach 的 Alpha	项目个数
0.937	0.939	24

Table 5. KMO and Bartlett test
表 5. KMO 与 Bartlett 检验

Kaiser-Meyer-Olkin 测量取样适当性。		0.879
Bartlett 的球形检定	大约卡方	2415.340
	df	465
	显著性	0.000

量表的科隆巴赫 α 系数为 0.937, 大于 0.8, 说明其信度, 良好; KMO 值为 0.879 大于 0.8, 说明这个量表适合进行因子分析。根据分析结果和专家意见, 结合内容效度和结构效度, 作此划分。

3.3. 要素权重分析法

基于问卷调研所得数据和维度划分, 运用 AHP 层次分析法对各级各项要素的重要性进行计算, 为面试设计中不同测评要素所占权重提供参考。

层次分析法, 简称 AHP, 是指将与决策总是有关的元素分解成目标、准则、方案等层次, 在此基础上进行定性和定量分析的决策方法[13], 其主要步骤如下:

- 1) 建立递阶层次模型。将问题条理化、层次化, 构造出目标层、准则层和方案层, 其中准则层中所需层数视具体问题而定。
- 2) 构造各层次中的所有判断矩阵。在准则层中, 依照决策者主观判断或相应数据为各准则之间的相对重要性进行打分, 引用数字 1~9 及其倒数作为标度, 1、3、5、7、9 分别代表因素 i 与因素 j 相比同等重要、稍重要、明显重要、强烈重要、极端重要, 偶数代表相邻判断的中间值, 以此构造出一个类似下表的判断矩阵(见表 6)。

Table 6. Example of a judgment matrix
表 6. 判断矩阵示例

因素	a	b	c	d
a	1	1/4	1/5	1/5
b	4	1	1/2	1/2
c	5	2	1	1
d	5	2	1	1

- 3) 一致性检验。由于判断矩阵生成过程带有主观性，其内部数据有可能存在相对矛盾的情况，因此需要一致性检验，当 $CR = CI/RI < 0.1$ 时，认为判断矩阵的一致性可接受。
- 4) 计算权重。基于判断矩阵，可以采用几何平均法、算术平均法、特征向量法或最小二乘法计算不同要素的权重。

3.4. 权重计算

- 1) 建立无领导小组讨论测评要素的递阶层次模型(如图 1)，该模型的目标层为无领导小组讨论测评要素，准则层包含二级测评要素，方案层则为不同评价者。

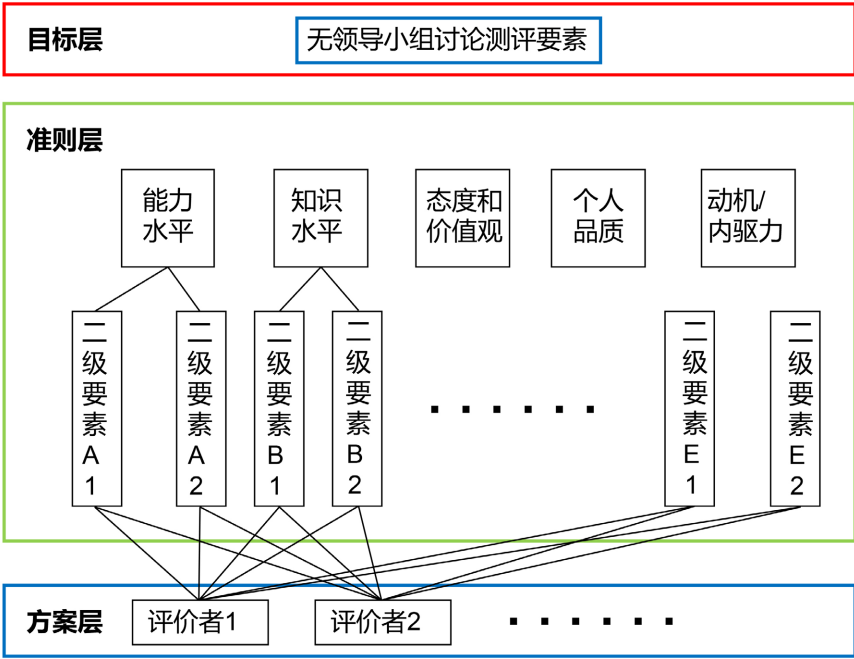


Figure 1. Recursive hierarchical model of the elements of a leaderless group discussion assessment
图 1. 无领导小组讨论测评要素递阶层次模型

- 2) 构造各级要素判断矩阵。分别对各一级要素量表中所有项目求均值，得出能力水平、知识水平、态度和价值观、个人品质、动机/内驱力五个一级要素变量；分别对各一级要素下的所有项目求均值，得出 24 个二级要素变量。将一级要素变量、二级要素变量的平均值作为进行同级要素重要性比较、构造判断矩阵的依据。除少数极端值外，绝大部分要素平均分之间的差异在 1 以内，据此当两个要素之差($i - j$)在 0.0625 以内时，标度为 1；当 $|i - j| \geq 0.0625$ 时，标度为 $[(|i - j| - 0.0625)/0.125 + 1]$ 向上取整所得整数($i - j > 0$)或其倒数($i - j < 0$)。

3) 一致性检验。由于本文判断矩阵严格按照客观数据生成，不会出现同一判断矩阵内不同要素相对重要性之间的相互矛盾，故无需一致性检验。

4) 权重计算。采用算术平均法，对判断矩阵进行按列归一化后，再求各行的平均值，最终生成不同测评要素的权重如表 7：

Table 7. Table of weighting results
表 7. 权重计算结果表

一级要素	权重(%)	二级要素	权重(%)
能力水平	21.27	创新决策能力	3.85
		问题处理能力	14.29
		沟通协调能力	16.42
		统筹管理能力	19.54
		任务执行能力	27.24
		自我管理能力	18.67
知识水平	13.24	基础知识	37.86
		专业知识	42.61
		特殊知识	19.53
态度和价值观	31.03	工作态度	10.05
		集体意识	37.70
		服务意识	22.60
		责任意识	29.65
		积极乐观	16.54
个人品质	23.89	富有自信	19.75
		机敏灵活	15.23
		客观冷静	12.68
		耐心礼貌	9.54
		科学严谨	10.21
		诚实守信	8.36
		公平公正	7.69
		自我价值实现	44.29
		职业兴趣	16.98
		事业成就感	38.73
动机/内驱力	10.57		

该权重计算结果较为繁琐，在实际操作中可以因具体需求和便利原则对权重进行调整。使用时需在评价者就各二级要素打分后，将得分与所属的一级要素权重相乘，最终进行加总，得出该评价者所得分数。

4. 结论与建议

无领导小组讨论作为一种人才测评技术，其在未来的人力资源管理中的作用将会越来越重要。通过不断地优化和改进，无领导小组讨论可以更好地满足企业和各类组织在选拔、评估和培训人才方面的需求。

为了进一步提升无领导小组讨论(LGD)的有效性和实用性，以下提出几点建议。首先，企业和组织应加大对 LGD 技术的投入，包括培训专业的面试官、更新和完善测评要素库以及引入先进的测评软件系统等。通过这些措施，可以确保 LGD 技术在实施过程中更加标准化、规范化和科学化。

其次，针对 LGD 测评要素的选择，建议企业和组织根据自身的实际需求进行合理配置。虽然 3~6 个

测评要素相对合理,但具体要素的选择应结合招聘岗位的特点和要求进行。例如,对于高层管理岗位,可以更加注重被试者的领导能力、决策能力和团队协作能力等方面的考察;而对于技术岗位,则可以更加关注被试者的专业知识、创新思维和问题解决能力等方面的表现。

此外,为了提高 LGD 的信度,建议增加评价者的人数,并加强对评价者的培训和管理。通过增加评价者的人数,可以确保对被试者的表现进行更全面、客观的评估;而加强对评价者的培训和管理,则可以提高评价者的专业素养和判断能力,从而降低评价过程中的主观性和误差。

最后,针对我国在 LGD 领域的理论研究和实践应用相对滞后的现状,建议学术界和业界加强合作与交流,共同推动 LGD 技术的创新与发展。通过举办学术会议、研讨会等活动,可以促进学术界和业界之间的信息共享和经验交流;同时,通过合作开展研究项目 and 实践应用案例,可以加快 LGD 技术的推广和应用进程,为我国的人力资源管理事业做出更大的贡献。

参考文献

- [1] 邹聪. 无领导小组讨论法在领导素质测评中的应用[J]. 现代商贸工业, 2021, 42(2): 75-76.
- [2] 顾大荣. 无领导小组讨论在公选领导干部中的应用[J]. 人力资源管理, 2011(11): 63-65.
- [3] 冯丽萍. 专家与新手评价者在无领导小组讨论及角色扮演的比较研究[D]: [硕士学位论文]. 广州: 暨南大学, 2011.
- [4] 彭平根, 丁彪, 苏永华. LGD 在选拔企业中高级管理人才方面的实证研究[J]. 心理科学, 2002(5): 576-579+640.
- [5] 唐忠明. 无领导小组讨论应用模式与评价者研究[D]: [硕士学位论文]. 苏州: 苏州大学, 2003.
- [6] 董越. 无领导小组讨论中的个体绩效与主动性的关系研究[D]: [硕士学位论文]. 武汉: 华中科技大学, 2010.
- [7] Waldman, D.A., Atwater, L.E. and Davidson, R.A. (2004) The Role of Individualism and the Five-Factor Model in the Prediction of Performance in a Leaderless Group Discussion. *Journal of Personality*, **72**, 1-28.
- [8] Lance, C.E., Newbolt, W.H., Gatewood, R.D., et al. (2000) Assessment Center Exercise Factors Represent Cross-Situational Specificity, Not-Method Bias. *Human Performance*, **13**, 323-353.
- [9] Gatewood, R., Thornton, G.C. and Hennessey, H.W. (1990) Reliability of Exercise Ratings in the Leaderless Group Discussion. *Journal of Occupational Psychology*, **63**, 331-342. <https://doi.org/10.1111/j.2044-8325.1990.tb00534.x>
- [10] Gaugler, B.B. and Thornton, G.C. (1989) Number of Assessment Center Dimensions as a Determinant of Assessor Accuracy. *Journal of Applied Psychology*, **74**, 611.
- [11] 魏玉华. 基于 CTT 与 GT 的无领导小组讨论构念效度研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2021.
- [12] 黄欣. G 省初任检察官遴选方案设计[D]: [硕士学位论文]. 广州: 华南理工大学, 2012.
- [13] 邓雪, 李家铭, 曾浩健, 等. 层次分析法权重计算方法分析及其应用研究[J]. 数学的实践与认识, 2012, 42(7): 93-100.