

大五人格特质与无领导小组讨论测评得分的关系研究

陈培峰, 魏玉华

重庆大学公共管理学院, 重庆
Email: iseese@126.com, 1163386854@qq.com

收稿日期: 2020年9月1日; 录用日期: 2020年9月12日; 发布日期: 2020年9月22日

摘要

无领导小组讨论结构效度的研究越来越关注情境特征与特质维度的交互作用, 而人格特质的作用有待进一步明确。模拟应聘情境, 对35名被试进行大五人格测试后, 随机分组进行无领导小组讨论实验。结果显示, 1) 被试的人格特质得分与LGD测评得分之间不存在显著相关水平, 相关系数最高为0.21。但人格特质与LGD维度得分及被试对自我评价存在关联; 2) LGD测评得分与被试在讨论中的行为表现具有较高水平的相关, LGD维度间的相关性低于行为绩效与LGD维度之间的相关性; 3) 被试的特定表现与组内成员的评价及LGD测评得分联系紧密。同时, 通过群体决策的角度探讨LGD的评分过程和影响因素, 为该领域研究结构效度提供了新的视角。

关键词

无领导小组讨论, 人格特质, 结构效度

Research on the Relationship between Big Five Personality Traits and Evaluation Score of Leaderless Group Discussion

Peifeng Chen, Yuhua Wei

School of Public Administration, Chongqing University, Chongqing
Email: iseese@126.com, 1163386854@qq.com

Received: Sep. 1st, 2020; accepted: Sep. 12th, 2020; published: Sep. 22nd, 2020

Abstract

In recent years, more and more attention has been paid to the interaction between situational features and trait dimensions, while the role of personality traits needs to be further clarified. Af-

ter the big five personality test, 35 subjects were randomly divided into groups for leaderless group discussion. The results showed that: 1) There was no significant correlation between the scores of personality traits and LGD, and the highest correlation coefficient was 0.21. However, personality traits were associated with LGD dimension scores and subjects' self-evaluation; 2) There was a high level of correlation between LGD scores and participants' behavior performance in discussion, and the correlation between LGD dimensions was lower than that between behavior performance and LGD dimensions; 3) The specific performance of subjects was closely related to the evaluation of members in the group and the scores of LGD. At the same time, this paper discusses the scoring process and influencing factors of LGD from the perspective of group decision-making, which provides a new perspective for construct validity research in this field.

Keywords

Leaderless Group Discussion, Personality Traits, Construct Validity

Copyright © 2020 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)具备较好的效标关联效度和预测效度(Gaugler, Rosenthal, Thornton, & Bentson, 1987; Arthur, Day, Mcnelly, & Edens, 2003), 被广泛应用于人员招聘、选拔、考核、晋升等人力资源管理实务, 但 LGD 存在着结构效度不佳的问题(Arthur, Woehr, & Maldegen, 2000; Lance, Newbolt, Gatewood et al., 2000)。许多研究发现, 活动后维度评定的变异中, 占主导地位的并不是我们想要测的维度, 而是被视为方法的活动(Bowler & Woehr, 2006; Sackett & Dreher, 1982), 这一现象甚至动摇了评价中心的测量构想(Lievens & Conway, 2001; Woehr & Arthur, 2003)。为了解释和应对行为的跨情境不一致性, 活动中心取向的研究认为, 受测者跨情境(活动)的特异性表现, 实际上代表着我们想要测量的真分数变异, 而不是测量方法偏差(Gibbons & Rupp, 2009)。交互作用取向的研究转而将注意力放在受测者行为表现及其影响因素上, 认为不同的活动可能给受测者呈现了不同的情景要求(Howard, 2008; Neidig & Neidig, 1984), 从而引发了受测者在相同维度上不尽相同甚至是很不一致的行为表现, 维度(特质)与活动(环境)共同的交互作用最终导致维度变异降低而活动变异增高。

从已有的文献来看, 早期的研究只局限于整体测评得分、人格各维度与人格特质的相关关系, 缺乏从群体决策的角度来验证两者之间的关系(Collins, Schmidt, Sanchez et al., 2010; Meriac, Hoffman, Woehr et al., 2008), 但目前群体互动评价的过程是研究 LGD 内在机制新视角(Kniffin, Wansink, Griskevicius, & Wilson, 2014)。因此本研究通过组织真实的 LGD, 强调 LGD 群体互动性指标的观察, 将人格特质与环境的交互作用作为研究重点, 以期发现人格因素在 LGD 中的具体作用, 不仅为无领导小组讨论结构效度的研究和应用增添依据, 还对更好地发挥人格测评与其他测评手段的协同作用具有重要意义(赫寅, 2014)。

2. 研究方法

2.1. 被试

从 146 名完成了大五人格测验的重庆大学学生中随机抽取 35 名学生作为实验被试。其中男生 17 名, 女生 18 名; 平均年龄 22.37 ± 1.71 岁; 有 26 名同学参加过学校组织的 LGD 模拟面试。

2.2. 材料

2.2.1. 人格特质的测量

采用中国大五人格问卷简式版(CBF-PI-B)作为测量工具, 该问卷包含 40 个条目, 内部一致性系数在 0.764~0.814 之间, 平均为 0.793, 重测信度平均为 0.742, 各因子与完整版量表对应因子的相关均在 0.850 以上(王孟成, 戴晓阳, 姚树桥, 2011)。

2.2.2. LGD 实验材料

讨论题目: 以测评目的和测评要素为出发点, 选取 LGD 常用的多选排序型题目, 以某城市招聘大型国际会议志愿者为背景, 要求被试在 15 项招聘要求中选出最为重要的 6 项, 排序并说明理由。

测评要素: 选取已有研究中经检验信效度较高的测评要素, 分别为综合分析能力、组织协调能力、语言表达能力、团队合作能力、主动性和自信心六项。

前测材料: 在开展讨论之前, 要求被试简单写下自己的观点和理由。

后测材料: 讨论结束之后, 要求被试对讨论过程和自己扮演的角色, 以及对其他组员进行评价。主要搜集被试的信息分享程度、角色变化、观点变化情况, 以及对其他组员在领导、组织、观点启发和人际等方面的印象。

2.3. 程序

2.3.1. 人格数据的获得

在自习室、图书馆、公选课等学生比较集中的情境随机发放人格问卷, 简单告知问卷的调查目的和内容, 要求被测者如实回答, 数据仅作为研究之用。并强调要留下联系方式以备后续研究进行联系。问卷采用 A4 双面打印, 问卷完整填写后现场收回。对回收的问卷进行录入和整理。

2.3.2. LGD 的组织与评价

实验场地摆放条形会议桌, 被试与评委分布于会议桌两端, 并提前摆放编码牌。

根据人格测试中被试留下的联系方式, 随机抽取 35 名, 分成 5 个小组, 每组 7 名成员, 每组成员互不认识, 成员之间以编号互称。首先主持人宣读指导语并告知此次活动的目的, 实验后会提供专业的测评报告以及建议, 同时会根据表现给予相应的被试费。正式实验过程分为五个阶段: 1) 准备阶段: 熟悉讨论题目及规则, 填写前测材料, 整理发言提纲, 限时 10 分钟; 2) 个人观点陈述, 主持人宣读发言环节要求, 并且告知发言环节开始, 进行独立发言, 顺序自行决定, 每人一次机会, 每人限时 2 分钟; 3) 自由讨论阶段, 要求小组成员对招聘要求按照顺序达成一致, 并说明理由, 限时 25 分钟; 4) 小组代表发言, 小组成员通过讨论后, 选举代表, 进行总结发言, 限时 3 分钟; 5) 讨论结束阶段, 被试填写后测材料, 对自己及他人在讨论过程中的表现进行评价。

评委由三名资深 LGD 评分专家和两名人才测评专业研究生组成, 为了增加评委评分的准确性, 对评委进行行为观察培训及参照系培训。评委按照评分标准对每个被试在综合分析能力、组织协调能力、语言表达能力、团队合作能力、主动性和自信心 6 个维度上进行四级评分, 并给出简单地评分理由。然后, 将每个被试在每个维度上的得分进行平均, 作为最终的测评得分。

2.4. 结果分析

2.4.1. 评分者一致性信度检验

LGD 的评价方式为多个评委对多个被试同时进行多个维度的评价, 常用评分者一致性系数(肯德尔和谐系数 W)作为信度衡量指标。本研究中, 5 名评委对所有 35 名被试评分的一致性系数见表 1, 六项测评

要素的评分者一致性系数较高, 均达到及其显著水平($\rho < 0.01$), 表明评委的评分具有可靠性。

Table 1. Score coefficient consistency of judges

表 1. 评委的评分一致性系数

| 评分维度 | W | $\chi^2 (n = 5)$ | ρ |
|--------|---------|------------------|--------|
| 综合分析能力 | 0.43*** | 73.5 | 0.00 |
| 组织协调能力 | 0.76*** | 129.47 | 0.00 |
| 团队合作能力 | 0.55*** | 188.32 | 0.00 |
| 语言表达能力 | 0.66*** | 111.67 | 0.00 |
| 主动性 | 0.67*** | 113.00 | 0.00 |
| 自信心 | 0.81*** | 136.95 | 0.00 |

注: *代表 0.05 水平显著; **代表 0.01 水平显著; ***代表 0.001 水平显著。下同。

2.4.2. 被试人格特质与 LGD 测评分数的相关性分析

根据大五人格量表简版的计分方式, 将 146 名大学生的人格得分与王孟成(2011)对 1221 名大学生样本的测量结果进行对比, 发现两个样本的得分具有较高的一致性(见表 2)。从 146 名大学生中随机抽取的 35 名参加 LGD 实验的被试在人格分布上也具有代表性; 算出 35 名被试人格得分的 Z 分数, 作为人格得分。

Table 2. Personality score and norm reference of subjects

表 2. 被试样本的人格得分及常模参照

| | 实验样本($n = 35$) | | 调查样本($n = 146$) | | 大学生常模($n = 1221$) | |
|-------|------------------|------|-------------------|------|---------------------|------|
| | M | SD | M | SD | M | SD |
| 神经质 N | 22.83 | 6.38 | 26.12 | 6.84 | 26.20 | 7.36 |
| 外向性 E | 31.63 | 6.26 | 28.64 | 6.89 | 30.58 | 7.08 |
| 严谨性 C | 36.86 | 6.14 | 33.77 | 6.53 | 33.13 | 6.57 |
| 开放性 O | 31.34 | 5.25 | 28.68 | 5.32 | 32.20 | 6.40 |
| 宜人性 A | 35.89 | 4.75 | 35.40 | 5.48 | 37.16 | 5.93 |

为了探索被试在 LGD 各维度的得分与人格特质之间的关系, 进行了相关分析, 得到如表 3 所示的相关矩阵。该矩阵包括了 LGD 各维度得分之间的相关、大五人格维度之间的相关以及人格特质与 LGD 维度之间得相关。LGD 各维度之间具有较高的相关性, 未达到显著水平的相关主要来自团队合作维度。大五人格之间达到显著水平的主要来自神经质和开放性与其他维度的相关。而人格特质与 LGD 得分之间的相关系数都比较低, 最高达到 0.217, 均未达到显著水平。

Table 3. Correlation analysis between personality traits and scores of each dimension of LGD

表 3. 人格特质与 LGD 各维度得分的相关分析

| | 综合分析 | 组织协调 | 团队合作 | 语言表达 | 主动性 | 自信心 | C | A | O | E |
|------|------|--------|--------|-------|--------|--------|---|---|---|---|
| 综合分析 | | 0.51** | 0.19 | 0.38* | 0.39* | 0.44** | | | | |
| 组织协调 | | | 0.48** | 0.11 | 0.47** | 0.42* | | | | |
| 团队合作 | | | | 0.27 | 0.17 | 0.37* | | | | |
| 语言表达 | | | | | 0.46** | 0.40* | | | | |

Continued

| | | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|--------|--------|------|--------|---------|
| 主动性 | | | | | | 0.62** | | | | |
| 神经质 N | -0.05 | 0.20 | -0.02 | -0.10 | -0.13 | -0.01 | -0.43* | 0.08 | -0.35 | -0.60** |
| 严谨性 C | 0.18 | -0.06 | -0.17 | 0.01 | -0.01 | -0.01 | | 0.09 | 0.57** | 0.32 |
| 宜人性 A | -0.06 | -0.02 | -0.14 | -0.21 | -0.20 | -0.21 | | | -0.05 | 0.17 |
| 开放性 O | 0.14 | -0.09 | -0.12 | -0.05 | 0.06 | -0.02 | | | | 0.49** |
| 外向性 E | 0.17 | -0.20 | -0.22 | 0.19 | 0.12 | -0.16 | | | | |

2.4.3. 被试 LGD 测评得分与行为表现的相关性分析

LGD 是一个群体决策过程, 参与其中的个体更清楚该过程中发生了什么。讨论结束之后, 被试根据讨论过程对小组中其他参与者的表现进行了评价。将被试的 LGD 维度得分与获得的评价进行相关分析发现(见表 4), 被试的各种表现之间存在较高的一致性。具体来说, 被试表现出的被认为可靠的言行与具有组织作用的行为、具有领导作用的行为、对讨论具有贡献的行为之间具有较为密切的联系, 而与被试表现出的亲和行为关系不大。被试表现出的这些行为与 LGD 各位的评分之间也具有较高的相关性, 除了亲和行为外, 都达到了极其显著的相关水平。

Table 4. Correlation analysis between behavior performance and scores of each dimension of LGD

表 4. 行为表现与 LGD 各维度得分的相关分析

| | 综合分析 | 组织协调 | 团队合作 | 语言表达 | 主动性 | 自信心 | 组织表现 | 亲和表现 | 领导表现 | 贡献表现 |
|------|--------|--------|--------|--------|--------|--------|--------|-------|--------|--------|
| 可靠表现 | 0.44** | 0.45** | 0.38* | 0.51** | 0.49** | 0.44** | 0.74** | 0.42* | 0.74** | 0.60** |
| 组织表现 | 0.49** | 0.66** | 0.48** | 0.44** | 0.54** | 0.50** | | 0.37* | 0.91** | 0.75** |
| 亲和表现 | 0.19 | 0.31 | 0.07 | 0.09 | 0.03 | -0.22 | | | 0.27 | 0.16 |
| 领导表现 | 0.57** | 0.68** | 0.54** | 0.46** | 0.57** | 0.60** | | | | 0.83** |
| 贡献表现 | 0.58** | 0.61** | 0.58** | 0.39* | 0.46** | 0.47** | | | | |

评委与组员在多大程度上揭示了个体在讨论中的真实表现呢? 被试对自己的表现也进行了自评。相关分析发现(见表 5), 被试在讨论过程中感觉自己的观点是否得到了充分表达、自己的观点在讨论过程中是否有所变、在讨论中自己扮演的角色(发挥的作用)是否有所变化以及小组最终达成的一致结论与自己最初的观点是否一致等与组员对自己的评价具有一定的相关性, 观点表达充分的个体更容易被认为具有领导力、具有组织作用、言行可靠及对讨论有贡献, 观点发生变化的个体更容易被认为具有亲和力, 讨论中角色发生变化的个体更容易被认为具有组织作用和领导力。被试的自评与 LGD 维度得分具有较少的关联, 主要表现为观点表达的越充分, 越容易在组织协调、主动性和自信心维度上得到高分, 而角色变化更容易在团队合作维度上得到高分。

Table 5. Correlation analysis of self-evaluation, other evaluation and dimension scores of LGD

表 5. 被试自评与他评及 LGD 维度得分的相关分析

| | 综合分析 | 组织协调 | 团队合作 | 语言表达 | 主动性 | 自信心 | 可靠表现 | 组织表现 | 亲和表现 | 领导表现 | 贡献作用 |
|------|-------|-------|--------|-------|-------|--------|--------|--------|-------|--------|--------|
| 观点表达 | 0.28 | 0.35* | 0.33 | 0.31 | 0.42* | 0.44** | 0.52** | 0.57** | 0.12 | 0.50** | 0.54** |
| 观点变化 | 0.05 | 0.10 | 0.27 | 0.27 | 0.12 | -0.03 | 0.32 | 0.21 | 0.38* | 0.25 | 0.30 |
| 角色变化 | 0.22 | 0.28 | 0.49** | 0.17 | 0.08 | 0.10 | 0.33 | 0.39* | 0.18 | 0.49** | 0.30 |
| 结果一致 | -0.12 | 0.20 | 0.29 | -0.13 | -0.09 | 0.10 | -0.07 | 0.10 | -0.02 | 0.03 | 0.15 |

2.4.4. 被试人格特质对 LGD 测评得分的作用机制分析

人格特质与 LGD 测评得分之间相关分析没有发现显著的相关性(见表 3), 进一步对人格特质与组员评价及被试自评进行相关分析, 也没有发现显著的相关性。考虑到变量水平及分布的影响, 继续采用卡方检验的方法进行统计, 发现, 人格特质与多个变量的卡方检验达到了显著水平。

首先, 人格特质与 LGD 维度得分存在关联: 宜人性影响主动性得分($\chi^2 = 12.737, n = 6, \rho < 0.05$), 宜人性得分高, 则主动性得分低; 外向性影响团队合作得分($\chi^2 = 7.921, n = 3, \rho < 0.05$), 外向性得分高, 则团队得分也高。

其次, 人格特质与被试对自己的评价存在关联: 开放性影响被试观点的充分表达($\chi^2 = 11.045, n = 4, \rho < 0.05$), 开放性高的被试认为自己更充分的表达了自己的观点; 严谨性影响被试观点的改变($\chi^2 = 6.628, n = 2, \rho < 0.05$), 严谨性高的被试报告自己的观点在讨论中较少发生改变。

3. 讨论

通过对 35 名被试的人格特质测量、LGD 评测、前后测、互评和自评, 发现被试的人格特质得分与 LGD 测评得分之间没有显著水平的相关; LGD 测评得分与被试在讨论中的行为表现具有较高水平的相关; 被试的特定表现与组内成员的评价及 LGD 测评得分联系紧密。这些发现与已有研究具有一致性, 同时也暗示了解与解释 LGD 测评体系的特殊机制。

3.1. 人格特质与 LGD 测评得分的关系

从现有的研究来看, 有部分研究报告了人格与评价中心测评维度得分的相关性(Love & Dearmond, 2007; Waldman, Atwater, & Davidson, 2004), 例如, 大五人格外向性与 LGD 总分存在显著相关, 责任感和外向性与团队合作具有显著相关(赫寅, 2014)。然而, 本研究中, 相关性系数最高达到 0.21, 但是并没有达到显著水平, 可能跟样本的数量有关, 也可能跟人格特质维度及评价中心测评维度的不同选取有关。更多的研究则发现人格特质与评价中心测评维度得分的不存在相关性或相关性不高(Collins, Schmidt, Sanchez Ku et al., 2010; Hoffman, Kennedy, Lopilato, Monahan, & Lance, 2015; Meriac, Hoffman, Woehr, & Fleisher, 2008)。研究者认为特质(如人格)和 AC 维度的预测效度是独立的, 人格和 AC 维度来自不同的领域, 预测的也是工作绩效的不同方面, AC 更多的测到了与活动有关的变异, 两者可以联合使用(Goffin, Rothstein, & Johnston, 1996)。

3.2. 行为绩效与 LGD 测评得分的关系

评价中心结构效度的研究越来越关注面试过程中被试的行为表现。在本研究中, 通过组员互评的方式对被试的多种绩效表现进行评价。发现得到小组成员认为言行可靠、发挥了组织作用和领导作用以及行为表现推动了小组讨论的被试, 在测评维度上也获得了评委的肯定。表达充分的被试也获得了行为绩效和评分上的肯定, 自我角色变化的被试容易被认为具有组织和领导才能, 而观点变化容易被认为具有亲和力。从本研究结果的表 3 和表 4 来看, LGD 维度间的相关性低于行为绩效与 LGD 维度之间的相关性, 而行为绩效之间的相关性最高, 由此可以推测, LGD 维度的结构效度好于行为绩效的结构效度。因此, 本研究的 LGD 测评维度虽然和行为绩效存在相关性, 但并不能因此认定这是方法效应的结果。这个结果也说明, 通过行为列表和行为尺度等方法可以提高评分的信度(Day & Sulsky, 1995), 但是对于结构效度的提升效果则有待观察。

3.3. 人格特质在 LGD 测评中的作用

考虑到本研究样本量偏小, 采用非参数卡方检验, 发现了个别测评得分和行为绩效在个别人格维度

上存在差异。宜人性影响“主动性”得分, 外向性影响“团队合作”分数。另外, 人格特质也影响个体在小组讨论中的行为表现。开放性和严谨性分布影响个体的观点表达和观点改变。从整个研究来看, 个体的观点表达影响了组员和评委对自己多方面的评价和评分, 而观点改变虽然影响亲和力和表现, 但是对测评得分却没有明显的作用。

这些结果是容易解释的, 从个人特质维度的界定上来说, 在小组讨论中, 外向性的个体自信、积极, 愿意主动表达自己的想法, 对实际团队中个人的绩效贡献和内部关系建设贡献等方面有良好的促进作用 (Barrick, Mount, & Judge, 2001), 外向性的个体也更容易获得群体中的社会地位(周颖萍, 2004), 宜人性的个体能够通过更好的沟通、协商解决问题(Graziano, Jensecampbell, & Hair, 1996), 但是本研究中却发现了更低的主动性得分, 这可能跟小组讨论中的竞争有关, 和谐人际导向意味着更多的放弃, 表现出较少的“侵略性”。

4. 研究结论

通过人格测试和 LGD 测评, 对人格特质与测评得分的关系以及人格特质在 LGD 测评中的作用进行了探讨, 主要得到了三点启发。首先, 虽然人格特质与 LGD 测评维度描述以及测评方式存在较为自然的内在联系, 但是本研究并没有发现人格特质与测评维度以及行为绩效之间的明确联系, 人格特质与测评维度对未来工作绩效的预测效度是独立的, 递增效度现象可能代表了其对工作绩效不同方面的预测。其次, 人格特质与测评维度之间的低相关以及行为绩效与测评维度之间的高相关, 在本研究中, 并不直接支持 LGD 或评价中心结构效度的活动中心论, 同时, 也说明了在提高 LGD 可靠性的努力上, 单纯提高评分信度还是不够的, 效度的提高才是根本。最后, 从群体决策的角度探讨评价中心测评维度与情境特征的交互作用具有重要意义, 通过领导力、领导地位、信息分享等群体互动的过程性指标来理解 LGD 的评分过程和影响因素能够进一步挖掘 LGD 的根本机制。

参考文献

- 赫寅(2014). 大五人格测验与无领导小组讨论的结果相关性研究. 硕士论文, 北京: 北京交通大学.
- 王孟成, 戴晓阳, 姚树桥(2011). 中国大五人格问卷的初步编制 III: 简式版的制定及信效度检验. *中国临床心理学杂志*, 19(4), 454-457.
- 周颖萍(2004). 浅析大五人格特质和社会群体地位的关系. *齐鲁师范学院学报*, 19(6), 41-43.
- Arthur Jr., J.W., Woehr, D. J., & Maldegen, R. (2000). Convergent and Discriminant Validity of Assessment Center Dimensions: A Conceptual and Empirical Re-Examination of the Assessment Center Construct-Related Validity Paradox. *Journal of Management*, 26, 813-835. [https://doi.org/10.1016/S0149-2063\(00\)00057-X](https://doi.org/10.1016/S0149-2063(00)00057-X)
- Arthur, W., Day, E. A., Mcnelly, T. L., & Edens, P. S. (2003). A Meta-Analysis of the Criterion-Related Validity of Assessment Center Dimensions. *Personnel Psychology*, 56, 125-153. <https://doi.org/10.1111/j.1744-6570.2003.tb00146.x>
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next? *International Journal of Selection and Assessment*, 9, 9-30. <https://doi.org/10.1111/1468-2389.00160>
- Bowler, M. C., & Woehr, D. J. (2006). A Meta-Analytic Evaluation of the Impact of Dimension and Exercise Factors on Assessment Center Ratings. *Journal of Applied Psychology*, 91, 1114-1124. <https://doi.org/10.1037/0021-9010.91.5.1114>
- Collins, J. M., Schmidt, F. L., Sanchez Ku, M., Thomas, L., Mcdaniel, M. A., & Le, H. (2010). Can Basic Individual Differences Shed Light on the Construct Meaning of Assessment Center Evaluations? *International Journal of Selection & Assessment*, 11, 17-29. <https://doi.org/10.1111/1468-2389.00223>
- Day, D. V., & Sulsky, L. M. (1995). Effects of Frame-of-Reference Training and Information Configuration on Memory Organization and Rating Accuracy. *Journal of Applied Psychology*, 80, 158-167. <https://doi.org/10.1037/0021-9010.80.1.158>
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-Analysis of Assessment Center Validity. *Journal of Applied Psychology*, 72, 493-511. <https://doi.org/10.1037/0021-9010.72.3.493>

- Gibbons, A. M., & Rupp, D. E. (2009). Dimension Consistency as an Individual Difference: A New (Old) Perspective on the Assessment Center Construct Validity Debate. *Journal of Management*, 35, 1154-1180. <https://doi.org/10.1177/0149206308328504>
- Goffin, R. D., Rothstein, M. G., & Johnston, N. G. (1996). Personality Testing and the Assessment Center: Incremental Validity for Managerial Selection. *Journal of Applied Psychology*, 81, 746-756. <https://doi.org/10.1037/0021-9010.81.6.746>
- Graziano, W. G., Jensecampbell, L. A., & Hair, E. C. (1996). Perceiving Interpersonal Conflict and Reacting to It: The Case for Agreeableness. *Journal of Personality and Social Psychology*, 70, 820-835. <https://doi.org/10.1037/0022-3514.70.4.820>
- Hoffman, B. J., Kennedy, C. L., Lopilato, A. C., Monahan, E. L., & Lance, C. E. (2015). A Review of the Content, Criterion-Related, and Construct-Related Validity of Assessment Center Exercises. *Journal of Applied Psychology*, 100, 1143-1168. <https://doi.org/10.1037/a0038707>
- Howard, A. (2008). Making Assessment Centers Work the Way They Are Supposed to. *Industrial and Organizational Psychology*, 1, 98-104. <https://doi.org/10.1111/j.1754-9434.2007.00018.x>
- Kniffin, K. M., Wansink, B., Griskevicius, V., & Wilson, D. S. (2014). Beauty Is in the In-Group of the Beheld: Inter-group Differences in the Perceived Attractiveness of Leaders. *The Leadership Quarterly*, 25, 1143-1153. <https://doi.org/10.1016/j.leaqua.2014.09.001>
- Lance, C. E., Newbolt, W. H., Gatewood, R. D., Foster, M. R., French, N. R., & Smith, D. E. (2000). Assessment Center Exercise Factors Represent Cross-Situational Specificity, Not Method Bias. *Human Performance*, 13, 323-353. https://doi.org/10.1207/S15327043HUP1304_1
- Lievens, F., & Conway, J. M. (2001). Dimension and Exercise Variance in Assessment Center Scores: A Large-Scale Evaluation of Multitrait-Multimethod Studies. *Journal of Applied Psychology*, 86, 1202-1222. <https://doi.org/10.1037/0021-9010.86.6.1202>
- Love, K. G., & Dearmond, S. (2007). The Validity of Assessment Center Ratings and 16 PF Personality Trait Scores in Police Sergeant Promotions: A Case of Incremental Validity. *Public Personnel Management*, 36, 21-32. <https://doi.org/10.1177/009102600703600102>
- Meriac, J. P., Hoffman, B. J., Woehr, D. J., & Fleisher, M. S. (2008). Further Evidence for the Validity of Assessment Center Dimensions: A Meta-Analysis of the Incremental Criterion-Related Validity of Dimension Ratings. *Journal of Applied Psychology*, 93, 1042-1052. <https://doi.org/10.1037/0021-9010.93.5.1042>
- Neidig, R. D., & Neidig, P. J. (1984). Multiple Assessment Center Exercises and Job Relatedness. *Journal of Applied Psychology*, 69, 182-186. <https://doi.org/10.1037/0021-9010.69.1.182>
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and Assessment Center Dimensions: Some Troubling Empirical Findings. *Journal of Applied Psychology*, 67, 401-410. <https://doi.org/10.1037/0021-9010.67.4.401>
- Waldman, D. A., Atwater, L. E., & 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. <https://doi.org/10.1111/j.0022-3506.2004.00254.x>
- Woehr, D. J., & Arthur, W. (2003). The Construct-Related Validity of Assessment Center Ratings: A Review and Meta-Analysis of the Role of Methodological Factors. *Journal of Management*, 29, 231-258. <https://doi.org/10.1177/014920630302900206>