

DDI评鉴中心 效度白皮书

1956年，DDI创始人之一道格拉斯·布雷先生在AT&T公司首次在业界应用评鉴中心。从那时起，研究者就开始对评鉴中心的效度进行研究，并记录了数千个成功的应用案例（如：Bray, Campbell, & Grant, 1977; Byham, 1970; Hunter & Hunter, 1984; Schippman, Prien, & Katz, 1990; Schmitt, Gooding, Noe, & Kirsch, 1984; Thornton & Byham, 1982）。当今企业广泛应用评鉴中心以解决各种问题，包括选才、人员配置、管理潜能的早期识别、晋升、发展、职业生涯管理以及培训。毫无疑问，评鉴中心对在职表现以及未来的绩效表现都极具预见性。

DDI 大力投入于新型评鉴中心技术的开发和提高。对录像评估、行为评分以及自动评估方法的结合运用，使我们成为该领域的权威。DDI 的众多成功案例一再证明了我们在评鉴中心的建构、内容和预测效度方面的实力。

评鉴中心实例

一项以 50 个评鉴中心项目为对象、包括 107 个效度系数的元分析 (Meta-analysis) 研究表明，评鉴中心具有很高的预测效度 (Gaugler, Rosenthal, Thornton, & Bentson, 1987)。效度系数显示评鉴中心得分与其它绩效评估方法之间的关系 (0 表示毫无关系; 1 表明完全的关联)。这些关系在各种不同用途 (如晋升、选才等) 的评鉴中心中显示出一致性。如表 1 所示，虽然个别研究结果之间会有差异，但这项综合性研究明确说明了评鉴中心的价值。

表 1: 校伪加权效度系数

	系数
总计	.37
绩效标准	
总体绩效评分	.36
总体潜能评分	.53
各项能力评分	.33
培训中的表现	.35
职业提升	.36
评鉴中心用途	
晋升	.30
早期识别	.46
选才	.29
调研	.48

DDI 曾为一家电制造企业设计了一个评鉴中心，以培养其现有的领导人，帮助他们为晋升至高层做好准备。该中心向评鉴对象提供反馈，根据其强项和弱点安排了一个长期的导师辅导制。在评鉴后的 6 至 12 个月期间，评鉴对象的主管就评鉴中涉及的相同能力对其当前工作绩效进行评分。

表 2 表明，评鉴中心的评分与工作绩效评分显著相关。

表 2: 家电制造企业

评鉴中心得分	系数
客户导向	.37**
愿景领导力	.33**
授权委责	.30*
管理工作	.38*

N = 49, *p < .05, **p < .01
相关系数根据信度准则校正

北方电信公司管理评估项目旨在确定候选人晋升到管理层的准备程度，并诊断其发展需求。为实现这个目标，DDI 与北方电讯合作，确定了相关的各项工作能力，并以此为基础设计评鉴中心。最终的评鉴中心包括多种评估技术，并为评鉴对象提供详尽的评鉴报告，让他们了解自己各项能力上的表现。为了验证评鉴中心的分数，我们从评鉴对象的同事那里收集了有关绩效标准的资料。

表 3 显示了评鉴中心得分与工作绩效数据之间大量的密切关联。

表 3: 北方电信公司

评鉴中心测定的能力	系数
客户服务导向	.24*
影响力	.22*
创新	.30*
工作适配性	.34*
Multiple R	.38

N = 61, *p < .05

相关系数根据信度准则校正

一家电信公司曾与 DDI 合作，设计一个为期两天的评鉴中心项目，其中包括的 11 个练习模拟了一家锐意进取的高科技企业副总裁“一天的真实生活”。针对高管典范涵盖的 11 项能力，公司对大约 140 名主管就各项能力进行评估。评估结果显示了评鉴对象在副总这一管理职位上的绩效潜能及其有待改进的绩效领域（实际的评估结果为有待发展、可以接受和表现优异）。

高管评鉴流程所得结果与评鉴对象上级对其在副总级别上的绩效潜能及发展需要的评估都达到了高度的一致（见表 4）。换言之，得到其主管高度评价的评鉴对象，更有可能在评鉴中心得到较高评分。

表 4: 电信公司

能力	绩效潜能	发展需要
客户服务/关注市场	94%	95%
口头沟通/简报	93%	92%
解决问题/决策	89%	91%
计划与执行	87%	87%
团队精神/授权委责	87%	86%
培养组织人才	86%	85%
积极主动	86%	83%
高管风范	85%	83%
管理绩效	83%	83%
专业知识/全球意识	81%	81%
策略远见	60%	76%

N = 140

DDI 在一间公用事业公司使用了主管级评鉴中心，以评估可能晋升到主管职位的候选人。该评鉴中心使用了 4 个情境模拟以评估候选人在与职位相关的 10 项能力上的表现。我们综合了在职时间至少一年的 33 位主管给出的工作绩效评分进行综合，与评鉴中心得出的结果极具关联性。该效度研究发现 R = 0.45 的多重相关性，这表明评鉴中心的评分与工作绩效评分之间存在一个有意义的关联。

评鉴与发展结合

高管们也反映说，在参加高管发展项目后他们收获颇丰。Dusenbury (1993) 报告了参加该项目的 64 位高管的反应，在反馈后准备了他们的发展计划。除了一位，其他所有参与者（98%）都认为，发展计划中的评鉴部“重要或非常重要”的发展工具。同时，近一半的参与者（49%）认为评鉴为“单一最重要的”的发展工具。这一比例远远高于他们对其它方法重要性的认同：培训（27%），导师辅导（7%）或技能项目（5%）。

Rifkin and Heine (1993) 也报告了一些高管评鉴与发展项目参与者的正面看法。在完成发展中心项目后的 18 个月之内，取得的积极成果包括：

- 成为更有效的管理者。
- 确定提高技能的具体方法。
- 更清晰地了解优势和发展需求。
- 得益于有效的发展，薪酬收入增加。

一家涂料、油漆和玻璃的制造商与 DDI 合作，为其 100 多名高潜力领导者制定了一个“高管发展流程”

(EDP)。在他们参加完一整天的评鉴中心、建立起发展计划并积极参与发展活动后，该组织观察到许多积极的成果，其中包括：

- 90%的参与者认为自身有很强的发展动力。
- 68%的参与者认为在知识或意识上有高度改善。
- 53%的参与者认为有大幅度的技能提高。
- 73%的参与者表示，该发展流程非常有价值。
- 该高管发展流程的效度系数相当高，而且在普遍可接受的标准内（如，主管评分与评鉴中心表现间的系数为.28）。
- 95%的参与者非常接受反馈意见，认为这准确反映了其真实的技能；96%的人表示，他们将使用评鉴中心反馈，以发展和制定改善行动计划。

在对该高管发展流程参与者进行的为期两年的后续追踪研究中，评鉴中心的结果显示，以发展为目的的 11 项能力中，有 7 项取得了显著提高。

DDI 如何创建有效的评鉴中心

评鉴需要实行严格的标准以确保准确性。DDI 在创建和管理评鉴中心的过程中谨从既定的准则。事实上，我们已经公布的评鉴中心程序和伦理准则已获“评鉴中心方法国际大会”的认可。以下几部分可以勾勒出 DDI 推进评鉴中心实效性方面的努力：

工作分析

- 请熟悉目标职位的人（即任职者及其主管）提供有关在该职位上取得成功所必需的工作活动的行为性事例。更为理想的情况是，随后请具代表性的工作内容专家对这些工作活动的频率和重要性进行评分。
- 工作分析师基于工作活动中描述的相关行为的多个例子，确定出一套初步的能力。
- 工作分析师根据工作内容专家对重要性的评分，选出最终的能力。该评分问卷包括具代表性的工作活动，以及定义能力的重要行为（关键行动）。

情境模拟设计

- 练习或模拟的设计旨在引出接近工作期望的行为，并反映工作活动中的重要内容。判断模拟相关性的一个方法，就是比较工作活动问卷中和情境模拟中的活动。
- DDI 将练习集中于几个能力。我们并不会设计一个模拟情境以衡量一个能力，而是精心设计以引导出与特定能力相关的行为。
- 情境模拟的设计中也包括一些提醒和暗示，以便引出与定义每项能力的关键行动相关的行为。这确保了评鉴对象能够在练习中展示能力所包含的所有重要方面。

评鉴师标准

- 评鉴师受到大量的培训，以辨识评鉴对象展示的与被考查能力相关的行为。这包括有关能力评估的一般培训以及针对各项练习的特别培训。
- 对一项能力评分涉及个人判断。DDI 为评鉴师提供行为事例以指导其客观判断，并召开校验会议以建立和确认标准。
- 记录下情境模拟中的行为事例后，评鉴师会评估相关的关键行动。基于对关键行动的评分、不同关键行动的相对重要性，以及评鉴对象在实现该能力目的方面的整体成果，评估师可运用全面判断来对能力评分。

技术的统一性

• DDI 采用多种形式，不断查验我们内部评鉴师之间的共识，包括双重打分、质量检查和正式的信度分析。如果评鉴师一致同意观察到的行为反映了同一能力上相同的水平，则他们有证据表明，该情境模拟可以导出与该能力相关的行为。

- DDI 不断检查评鉴练习中关键行动和能力评分的分布。如果大部分评鉴对象的能力评分呈正态分布，则证明该练习导出了一系列与该能力相关的行为。

- 如果评鉴师没有观察到任何关键行为，他们的评分为“0”。如果一个模拟设计的频率分布中出现大量的“0”，则表明该模拟没有导出关键行动。这种情况下，DDI 会调整情境模拟的设计，以确保它能充分衡量目标能力。

参考文献

- Bray, D.W., Campbell, R.J., & Grant, D.L. (1977). *Formative years in business: A long-term AT&T study of managerial lives*. New York: Wiley.
- Byham, W.C. (1970). Assessment centers for spotting future managers. *Harvard Business Review*, 7, 150–160.
- Development Dimensions International. (1980, 1989). *Guidelines and ethical considerations for assessment center operations*. Pittsburgh, PA: Task Force on Assessment Center Guidelines.
- Dusenbury, D. (1993, March/April). *Getting the most from assessment*. Presentation at the 21st International Congress on the Assessment Center Method, Atlanta, GA.
- 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.
- Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Rifkin, J., & Heine, D. (1993, March/April). *The impact and payoff of development centers*. Presentation at the 21st International Congress on the Assessment Center Method, Atlanta, GA.
- Schippman, J.S., Prien, E.P., & Katz, J.A. (1990). Reliability and validity of in-basket performance measures. *Personnel Psychology*, 43, 837–859.
- Schmitt, N., Gooding, R.Z., Noe, R.A. & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407–422.
- Thornton, G.C., & Byham, W.C. (1982). *Assessment centers and managerial performance*. New York: Academic Press.