

Application of Assessment Centre in Talents Recruitment

—Case Study of Back-Up Talents Selection in a Commercial Bank

Junzhen Liu, Jianhui Zhao, Mengqiu Shen

Faculty of Human Resource Management, Business School, Nankai University, Tianjin
Email: liujunzhen@nankai.edu.cn

Received: Jan. 13th, 2013; revised: Feb. 14th, 2013; accepted: Feb. 20th, 2013

Abstract: Based on back-up talents selection case that the author participated in a commercial bank, the authors conclude the problems through relevant data analysis: the match of evaluation contents and evaluation approaches, appraiser cognition and preference of dimensions, and evaluation bias. Finally, the paper arrives at the following proposals: designing selection tools scientifically based on evaluation contents, unifying the standard through integrative training, and lowering bias through arranging the right interviewee quantity and proper intervenes.

Keywords: Commercial Bank; Assessment Centre; Case Study; Bias Control

人才招聘中评价中心技术的应用

—基于某商业银行后备人才选拔的个案分析

刘俊振, 赵剑慧, 申梦秋

南开大学商学院人力资源管理系, 天津
Email: liujunzhen@nankai.edu.cn

收稿日期: 2013年1月13日; 修回日期: 2013年2月14日; 录用日期: 2013年2月20日

摘要: 论文根据笔者主持的某商业银行后备人才选拔为个案, 在数据分析基础上, 归纳了该银行使用评价中心技术进行测评时出现的三类问题: 测评维度与测评方法的匹配设计问题、主试对测评维度的认知与偏好问题, 以及测评的偏松偏紧问题。论文在探究原因基础上, 提出科学匹配测评维度与测评方法、借由整合性培训统一主试测评标准, 以及通过降低测评强度和适当过程干预来降低偏差等建议。

关键词: 商业银行; 评价中心技术; 个案研究; 偏差控制

1. 引言

伴随人力资源管理在组织中愈来愈发挥专业性、战略性作用, 评价中心技术这种综合评价管理潜能的人才测评技术, 已经在工业、教育、政府、军队和其它组织中广泛应用, 主要领域包括选聘、职业生涯发展、培训发展、管理发展、领导力发展。

我国商业银行的迅猛发展, 需要一大批胜任能力强

的管理人员队伍, 稳健而有效的外部招聘和内部选拔是实现组织快速发展的人才保障, 评价中心技术在人才选聘、选拔中发挥着至关重要的“技术保障”作用。

2. 文献回顾

评价中心(Assessment Center), 或称评价中心技术, 作为一种综合性的人才测评技术, 以超越传统测

试和单一测试的方式,通过可观察的、具普遍意义的关键管理行为,来检视和预测被测评对象的管理能力和发展潜力。根据“情境-行为-绩效”的关联效应,测评个体的素质与潜在能力不能脱离特定的情境,因此,科学的情境设计非常关键,评价中心就是强调职位分析的情境模拟技术的应用,并结合多种方法,从多种角度来收集个体与职位相关的关键信息^[1]。在具体方法方面,评价中心综合了诸如心理测验、知识/能力测验,公文筐测验,演讲,案例分析,小组解决问题,角色扮演,无领导小组讨论,管理游戏以及情境面谈等。

评价中心起源于德国心理学家 Hartshorn 和 May 1928 年的研究以及此后的军官甄选程序,二战期间,美国战略情报局使用小组讨论和情境模拟来选拔情报人员,获得成功。之后,美国许多著名的大公司,如通用电气,IBM,福特汽车,柯达等都应用这种技术选拔人才。评价中心技术在国内应用时,其范围也在不断扩大,不但包括大量的工商企业,还包括高校教师的招聘(如乔花云 2011 年的研究)、地方政府后备干部选拔(如司林波等 2011 年的研究),等等^[2,3]。

有关评价中心的有效性,研究颇丰,而且普遍性表明该技术的作用超过其他方法。赫克的对比试验发现,任意选拔经理的正确率为 15%,经管理部门提名的正确率为 35%,经管理部门推荐并结合评价中心测评其准确率可达到 76%^[4]。Woodruffe (1990)研究表明,选择心理学家作为测评人,其评价者之间的信度低于训练有素的直线管理者,而且,无领导小组讨论对于候选人集体领导技能的评价非常有效,其信度随着小组人数的增加而有所提高,效度在 0.15~0.85 之间^[5]。Moser (1995)和 McDaniel 等(1994)的研究发现,经严格设计的结构化面试,其效度要明显高于非结构化面试,效度在 0.44~0.51,而且结构化面试的有效性在用于个人面试时要优于小组面试^[6]。国内学者也在持续关注评价中心的应用以及有效性问题,如陈民科、王重鸣(2002)关注构思效度问题^[7],王博等(2010)关注金融企业评价中心的结构效度问题等^[8]。

3. 个案描述

个案涉及的银行是中国较大规模的股份制商业银行,在人力资源管理日益重要的今天,引入科学的

评价中心技术来储备后备人才队伍,逐渐摆脱传统上单纯依靠组织考察与任命的识人和用人机制。

3.1. 个案银行后备人才选拔程序

首先,界定后备人才,确定后备人才的胜任模型。经过前期的访谈、分析与咨询互动,确定该家银行后备人才范围为目前担任团队主管、部门经理、分行行长等职位中具有更高一级胜任潜能的候选人。胜任模型维度主要包括六个方面:远见分析、团队领导与影响、开拓创新、变革管理、执行力、组织协调。

其次,根据胜任模型设计测评方案,编制测评题目,并实施测评。整体测评分三部分:纸笔测试、无领导小组讨论和结构化面试。

(1) 纸笔测试

包括五个方面的测试:知识测试、能力倾向测试、管理风格测试、个性测试、文件筐测试,详见表 1。

(2) 无领导小组讨论

每次测试 6~7 人,测试时间为 40 分钟。测试与评价维度为:主动性、团队领导力与影响、组织协调、合作意识、分析判断、人际敏感性。

(3) 结构化面试

行为描述型面试,每名候选人面试 40 分钟,测评与评价维度为:进取心与心理动力、分析判断、组织协调、影响力、社会成熟度、口头表达、执行力。

第三,对测试数据进行整理和统计分析。

最后,整理测试报告,提交银行高管决策。

3.2. 个案银行后备人才的基本信息

候选人来自该银行各部门,它们竞聘同一类工作

Table 1. Contents and elements in paper and pencil test
表 1. 纸笔测试的内容与要素

测试工具	主要测试要素	测试时间
知识测试	组织协调/冲突解决/适应/应变/大局观/创新	100 分钟
能力倾向测试	分析判断/逻辑推理/言语理解/书面表达	120 分钟
管理风格测试	心理能量指向/收集信息方式/决策依据/对待生活和工作方式	30 分钟
个性测试	进取心/责任心/社会成熟/情绪稳定/果敢性/动力来源	90 分钟
文件筐测试	授权/决策/组织协调/时间分配	90 分钟

职务,所有的候选人均需经过全部的评价中心技术测试,最后综合权衡知识测试、无领导小组讨论、结构化面试的成绩来决定候选人之间的相对排序。

笔者作为测评人员主持参加了 46 名候选人的测试工作(见表 2),并在事后用 JMP4 对获得的相关数据进行了统计分析,指出本次银行后备人才选拔中存在和需要改进问题。

这 46 名候选人的纸笔测试工作已经提前完成,在接下来的四天测试中,第一天上午分组安排他们参加无领导小组讨论;从第一天下午到第四天下午安排候选人的结构化面试。

4. 问题及原因

4.1. 问题

4.1.1. 一些测评方法对一些测评维度(指标)测量效果差

首先,可以从针对相同胜任维度的不同测评方法得分的相关性分析得出这一结论。该个案中,分别使用无领导小组讨论和结构化面试两种测试技术对同一候选人的如下三个胜任维度进行了两次测量:影响力、组织协调能力、分析判断能力。从表 3 两种方法得分的相关性分析看,针对这三个维度的两种测试分数相关程度弱,说明其中有一种方法不能用于上述三个维度的测量,也说明在设计阶段选择该方法测量上述三个维度是不科学的。

其次,从无领导小组讨论各测量维度对测试分数的影响(详见表 4)来看,也能得出类似结论。表 4 反映

Table 2. Basic information of 46 candidates attended tests
表 2. 参加测试的 46 名候选人的基本信息

性别	人数	竞聘前职务类型	人数
男	34	管理职位	28
女	12	技术职位	18

注: 46 名候选人在该银行工作时间均为 6-8 年, 大专或大本毕业, 年龄在 28-32 岁之间

Table 3. Correlation analysis of same assessment dimension scores using two tests (P = 99%)

表 3. 同一评价维度两种不同测试分数的相关分析(P = 99%)

共同评价维度	面试与无领导小组讨论得分的相关系数
分析判断能力	0.192
组织协调能力	0.236
影响力	0.337

出无领导小组讨论能够较为准确的测量主动性、组织协调能力和团队领导意识与影响力、分析判断能力;而对人际能力、合作性则不能有效测量。这说明在前期的测评维度与测评方法匹配设计方面,严谨性不够。

4.1.2. 主试对面试各维度重要性的认知与偏好存在差异

面试考察的 7 个方面维度大体上应该有着相似的权重,但由于主试的认知、专业、经验以及偏好,他们对这些评价维度的相对重要性认知还是存在很大差异。从表 5 的面试评价维度分数与面试总分数的相关分析可以发现,笔者担任面试主持,在测评候选人时,对面试总分影响最大的前 3 个维度依次是:分析判断能力,影响力和口头表达能力。

4.1.3. 测评结果出现较为明显的“偏松”“偏紧”误差

经过 Tukey-Kramer 单项 ANOVA 分析发现,面试时间安排对面试分数存在统计上的影响。表 6 中 B1 表示第一天下午的面试, A2 表示第二天上午的面试,

Table 4. Influence analysis between LGD scores and the overall tests scores (P = 99%)
表 4. 无领导小组讨论的评价维度对测试分数的影响(P = 99%)

评价维度	相关系数
主动性	0.787***
分析判断能力	0.669**
组织协调能力	0.787***
团队领导与影响力	0.752***
合作性	0.157
人际能力	0.359

Table 5. Influence analysis between structured interview scores and its dimensions (P = 99%)
表 5. 结构化面试评价维度对面试分数的影响(P = 99%)

评价维度	相关系数
进取心与心理动力	0.414
分析判断能力	0.843***
组织协调能力	0.497
影响力	0.737**
社会成熟度	0.486
口头表达能力	0.716**
执行力	0.374

Table 6. Tukey-Kramer analysis between interview time schedule and interview scores
表 6. 面试时间对面试分数的 Tukey-Kramer 分析

Abs(Dif)-LSD	A4	B3	A2	B2	A3	B4	B1
A4	-1.54791	-1.35112	-1.27255	-1.17969	-1.19791	-0.91457	0.21316*
B3	-1.35112	-1.43308	-1.35451	-1.26165	-1.28208	-0.99874	0.13120*
A2	-1.27255	-1.35451	-1.43308	-1.34023	-1.36065	-1.07731	0.05263*
B2	-1.17969	-1.26165	-1.34023	-1.43308	-1.45350	-1.17017	-0.04023
A3	-1.19791	-1.28208	-1.36065	-1.45350	-1.54791	-1.26457	-0.13684
B4	-0.91457	-0.99874	-1.07731	-1.17017	-1.26457	-1.54791	-0.42017
B1	0.21316*	0.13120*	0.05263*	-0.04023	-0.13684	-0.42017	-1.43308

依此类推。分析后发现第一天下午面试的平均分数明显低于以后几天的面试平均分数，这说明第一天面试主试评价的“偏紧效应”以及后几天评价的“偏松效应”。

4.2. 原因

分析造成上述问题的原因，主要为：

1) 在设计阶段，测评方法与测评维度之间欠缺科学设计与匹配。对于一些胜任素质维度到底应该采用何种测评手段来测量，某种测评方法对于哪些维度的测量更为有效，缺乏严谨的分析和设计，从而造成一些测评工具不能有效测量一些维度，造成“缺失”或“污染”。

2) 对测评主持人的培训内容、培训结构和培训严谨性不够。这样，造成主持人对各评价维度的相对重要程度、选拔标准认识不一、把握不一，因而测评结果就容易受到主持人自身专业(或职业)、经历、做事风格以及个人偏好的影响，如专业/技术出身的主持人偏好于选择那些技术能力强，分析问题缜密的候选人，尽管他们在管理能力与领导素质方面可能存在不足。

3) 测评时间安排不合理造成测评偏差增大。不合理主要源于两个方面：一是总的测评时间长，中间缺乏分享讨论和反馈控制。该银行人才测评持续 4 天时间，第一天对全部面试主持人进行了集中培训和分享讨论，从第二天开始，各小组只顾忙于本小组的面试任务，相互间并没太多沟通、讨论，从而产生了上述“偏松”“偏紧”误差。二是测评强度大，该次银行后备人才选拔每天测评时间均在 10 个小时左右，面

试的高强度安排增加了面试主持人的“疲劳偏差”。

5. 建议

5.1. 科学设计测评维度与测评方法之间的匹配

首先，要建立科学的胜任素质模型，明确关键的胜任素质维度。针对该商业银行而言，不能对所有的职位都进行系统性的职位分析，否则容易使工作步入僵化。建议在横向的工作族(Job Family)层面(如风险管理类、经营业务类、稽核监督类、行政管理类、技术类等等)以及纵向的职级层面(如辅助、专员、主管、经理、总监、行长等)划分银行中全部职位，在此基础上，进行工作族分析和胜任素质分析，从而明确银行中每一工作族、每一层级工作的胜任素质模型、测评维度以及测评标准。

其次，科学选择测评的方法以及测评方法之间的组合。出于成本的考虑，在选择测评方法组合方面，应让各自的方法在测评维度方面各有侧重，减少重叠。比如对候选人的主动性和进取意识的测量，使用无领导小组讨论测试效度更高；对分析判断能力的测量使用结构化面试效度会更高；对于组织协调能力的测量，使用无领导小组讨论比结构化面试的效度更高；对于影响力或团队领导能力(它是一种综合指标，反映一种综合能力)，采用结构化面试和无领导小组讨论两种测试一起则具有较好的预测性。

5.2. 借由整合性培训使测评主持专家统一测评的维度与标准

测评主持专家的专业或职业、行业经验、个性以及行事风格会自然而然带到测评过程中来，从而造成

测评结果差异。因此，首先要选择那些系统掌握心理学测量知识，熟悉银行业务知识，具有良好面试技能，社会成熟度高以及职业素养高的人作为评价中心的测评专家。其次，对测评专家的整合培训非常重要，培训内容涉及银行知识、职类划分、测评维度、测评工具应用技巧等方面，从而使测评主持专家在把控测评标准方面尽量一致，避免评价的随意性和在候选人之间相对比较。

5.3. 适当降低测评强度并增加过程干预来减少测评偏差

测评时间过长会大大影响测评的质量，因此，首先应适当降低测评强度，并安排休息休闲活动，这样会有助于测评专家对测评工作保持适度热情，降低测评工作造成的压力，从而提高测试的科学性与有效性。其次，针对整个测评时间长的情况，应将长的时间自然地划分为若干小段时间，并在测评间隔期间，

适当安排所有测评专家进行有关问题回顾与讨论，以帮助专家发现自身在评价时可能出现的偏差，从而利于过程的控制。

参考文献 (References)

- [1] 王小华, 车宏生. 评价中心的评分维度和评分效果[J]. 心理科学进展, 2004, 12(4): 601-607.
- [2] 乔花云. 评价中心技术在高校教师招聘中的应用[J]. 科技管理研究, 2011, 22: 5.
- [3] 司林波等. 评价中心技术在地方政府后备干部选拔中的应用——以秦皇岛市农业局为例[J]. 中国人力资源开发, 2011, 7: 45-48.
- [4] 凌文铨, 方俐洛. 人员测评[M]. 北京: 中国纺织出版社, 2003.
- [5] C. Woodruffe. Assessment centers. London: IPM, 199.
- [6] A. Howard. A reassessment of assessment centers: Challenges for the 21st century. Journal of Social Behavior & Personality, 1997, 12(5): 3-4.
- [7] 陈民科, 王重鸣. 评价中心的开发程序与构思效度[J]. 人类工效学, 2002, 8(2): 27-30, 34.
- [8] 王博等. 评价中心结构效度的多元概化理论分析研究[J]. 心理学探新, 2010, 30(5): 66-71, 83.