

基于行为的绩效评定方法的研究进展*

马成功 王二平 林平

(中国科学院心理研究所, 北京 100101) (北京机械工业学院人文社科系, 北京 100012)

摘 要 从 20 世纪 60 年代至今, 基于行为的绩效评定一直是各类组织的主要绩效评定方法。基于工作结果的非判断性绩效测量有其难以克服的局限, 各种基于工作行为的判断性绩效测量方法相继开发出来。行为指标的筛选成为研究的主要问题之一。组织公民行为、区分作业绩效与关系绩效等研究, 向基于行为的绩效评定提出了新的问题。

关键词 关键事件技术, 组织公民行为, 关系绩效, 作业绩效。

分类号 B849:C93

绩效评定是人力资源管理中的最棘手的问题之一。它既是人力资源管理的一个必不可少的重要环节, 又因为绩效评定结果的不确定性而导致使用者经常对其持不信任态度, 所以 Cascio 把绩效评定说成是人力资源管理中的“阿基里斯的脚后跟”(意指绩效评定是人力资源管理的最薄弱之处)^[1]。多年以来, 工业与组织心理学家从理论研究和实践方面一直在寻找更有效的绩效评定方法, 研究是什么原因导致绩效评定过程中出现的不确定性, 如何才能有效地加以避免。本文将从以下几个部分来介绍有关的文献综述: 研究基于行为的绩效评定方法的意义; 绩效行为指标的获得; 关于不同作业特征职务的行为评定方式的研究; 绩效行为指标内涵的扩展——组织公民行为 (organizational citizenship behavior, OCB) 和关系绩效; 喜欢因素对绩效评定影响的研究; 绩效评定的实验室研究。

1 研究基于行为的绩效评定方法的意义

基于行为的绩效评定方法的研究开始于 20 世纪 60 年代, 经过 20 年的研究后, 在 80 年代就成为绩效评定方法的主流。至今, 仍然是管理实践和理论研究的热点。

Cascio 把绩效评定定义为对个人或群体与工作有关的优缺点的系统描述^[1]。绩效测量主要分为下述两种类型: 非判断性测量和判断性测量^[2]。非判断性测量包括生产数据和人事数据 (如事故率、离职率、缺勤次数等); 判断性测量侧重于对个体行为的评价。非判断性测量的重点不是行为, 而是行为的结果。显然, 在行为和结果之间有某种程度的交叉, 但两者在内涵上是不同的。由于行为的结果常常受到个人控制之外因素的影响 (如工作环境、经济形势等), 再者, 在许多职务中, 根本就没有良好的客观绩效指标, 因此, 作为绩效标准, 它们通常作用不大。因为绩效评定的目标是评定个人绩效, 以控制任职者无法控制因素的影响。正是由于非判断性测量存在这些缺点, 研究者和管理者将绩效评定的重点放到工作绩效的判断性测量上。判断性测量能够比较有效地观察和评价个体的行为。然而, 由于判断性测量依

收稿日期: 2002-01-09

* 国家自然科学基金资助项目 (批准号: 79770090)。

赖人来做出判断,容易出现与判断过程相联系的某些判断误差。因此要充分发挥判断性测量的作用,就必须对那些与工作绩效紧密相关的重要行为进行深入地分析。

从20世纪60年代开始,行为主义心理学家开始对如何增强绩效评定的有效性这一问题产生兴趣,他们试图寻找和描述与职务有关的员工行为,并且认为这种与职务有关的行为应该是可被观察、可被定义、可被测量的^[3]。他们把基于行为的绩效维度定义成与绩效相关的行为和特征,并且可以通过关键事件技术(critical incident technique, CIT)来获得^[4]。并在70年代和80年代开发出颇具影响的行为测量方法:如行为锚定等级评价法(BARS)和行为观察量表法(BOS)等方法。这些方法的一个共同的特征是:通过关键事件技术进行职务分析,从而获得与绩效有关的行为指标,再采用等级评价方法将这些行为指标合成为评价工具。其中,行为锚定等级评价法被认为比其他的绩效评定工具有更好的和更公平的评定结果,因为它是把关于特别好或特别差绩效的叙述加以等级化,来建立行为锚定等级体系,并且每个等级尺度上都附着相应的关键事件描述。当然,使用行为锚定等级评价法比使用其他的绩效评价方法要花费更多的时间^[5]。

基于行为的绩效评定在企业的管理实践中有着重要的意义。首先,基于行为的绩效评定能够尽量客观而准确地描述员工在一个既定时期内对组织的贡献,以及在工作中表现出来的高效率行为和低效率行为,从而帮助企业判断应当做出何种晋升或工资方面的决策。其次,基于行为的绩效评定能够给员工提供实质性的反馈,为了提高未来绩效水平,员工需要了解他在过去工作中的低效率行为有哪些,知道如何去改进。再者,基于行为的绩效评定能帮助确定培训课程的目标,通过针对行为的培训来帮助员工获得组织所期望的行为。

评定行为的目的是为了能够更好地提高那些与绩效有关的行为表现,在此基础上的培训是员工绩效能够不断提高的重要保证。在80年代主要的培训理论几乎都是在行为理论的基础上发展起来的,这些培训理论包括:行为角色模型(Behavioral Role Modeling),基于计算机的指导,模拟教学法(包括案例研究,角色扮演,无领导小组讨论和商业游戏),以及行为调节和行为的自我管理^[6]。基于行为的绩效评定方法能够寻找到那些组织和职务本身所要求的行为,这些有效的行为就可以作为员工培训的行为目标。

2 绩效行为指标的获得

绩效评定的目标不是区分职务差别,而是区分人的绩效差别,尤其是区分从事同一职务的人的绩效差别。那么,如何从众多员工行为中筛选出那些能够反映出这种差别的行为指标就显得尤为重要。行为一致性模型(behavioral consistency models)认为过去的行为是将来行为的最好的预测源^[7]。因此,有效的行为指标往往是由那些对职务及其要求最为熟悉的人对一些代表好绩效和差绩效的关键事件的行为描述。

关键事件技术(CIT)是Flanagan在1949年《人事评价的一种新途径》(“A new approach to evaluating personnel”)一书中提到的方法。通过关键事件技术,可以从行为的角度系统地观察和描述实际职务的绩效和行为^[3]。Cascio认为某一职务的关键事件是指在有效工作和无效工作之间造成差别的行为,它是由了解职务的观察者对那些与有效工作和无效工作有关的行为进行报告^[1]。关键事件可以通过个体访谈、个体观察和问卷3种方式来获得。通过任职者对职务活动的分析,来确定那些导致工作成功和失败的行为,从而形成一个行为列表,再

让专家从中选出与特定的绩效维度相关的行为。分析者可以通过经验方式或因素分析的统计方法来归类某些内容相似、彼此相关的行为，从而形成特定的绩效维度。

行为指标的获得一般通过图 1 的流程来实现。通过职务分析获得的行为指标，经过实践证明有效后，就可以作为绩效标准，再将这些绩效标准合成为绩效评定的工具，进行绩效评定。

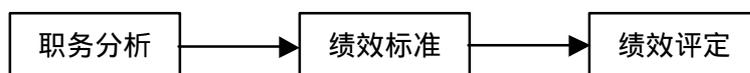


图 1 绩效行为指标的筛选

当运用基于行为的绩效评定方法测量绩效时，应该用足够的维度来完整地覆盖职务的所有关键行为。这种基于行为的绩效评定维度直接来源于对职务内容和职务行为的分析。分析职务内容的目的在于建立工作职责，包括确定完成工作所需的知识、技能和能力，以及当前的工作环境和条件。分析职务行为的目的在于获得员工履行工作时那些可以被观察、定义和测量的特定行为，通过熟悉职务的任职者和直接上级来确定与绩效有关的工作事件和所必须的关键行为。

3 关于不同作业特征职务的行为评定方式的研究

从基于行为的绩效评定方法诞生之初，研究者就试图去创造一种最佳的绩效评定工具，使它适用于所有的职务。研究者不断地研究出图表评价法（graphic rating scale）、交替排序法（alternative ranking method）、配对比较法（paired comparison method）、强制分布法（forced distribution method）、关键事件法（critical incident method）、行为锚定评价法（Behaviorally Anchored Rating Scale, BARS）、行为观察法（Behavioral Observed Scale, BOS）等不同的绩效评定工具，并且比较各种工具的优缺点，进行不断的改进。但是，由于不同职务的作业特征是不同的，在工作中会体现出不同的关键行为，因此通过职务分析就会获得不同的绩效评定维度。那么，对于不同的绩效评定维度是否存在着“最佳”的评定方法呢？

周智红在一项关于销售员的绩效评定方法的研究中发现，由于销售员职务本身具有上级和同事无法观察其行为表现的特征，所以对于销售员的部分绩效评定维度而言，采用自评的迫选量表方法能够得到较为客观的绩效评定结果^[8]。

Conway 发现同事在评定整体绩效时更注重关系绩效，上级在评定整体绩效时更注重作业绩效^[9]。这说明，对关系绩效进行评定时，采用同事评定的方法可能比上级评定更有效；对作业绩效进行评定时，采用上级评定的方法可能比同事评定更为有效。

Lee 认为在绩效评定中并不存在“最佳方法”，她建议对于不同作业特征（task characteristics）的职务应该采用不同的评定方式和程序^[10]。Guion 和 Gibson 认为这很可能会开拓出绩效评定的新时代^[11]。Lee 认为应该把绩效评定的方式与作业特征相匹配，这样才能够为评定者提供有关的观察信息，从而提高绩效评定的有效性。她按照职务的手段-结果转换过程所需知识的程度，把职务分为两种作业特征类型，一种是手段-结果高转换过程的职务（如记者），一种是手段-结果低转换的职务（如推销员）。对于前一种职务，Lee 建议采用行为评定的方式，对于后一种职务，应注重对结果的评定。

那么，对于同一职务而言，是否也存在类似的现象呢？Binning 和 Barrett 认为绩效区是由行为单元和结果单元组成^[12]，如图 2 所示。结果（如销售额）是组织价值的体现，行为（如销售技能）则是实现组织价值的手段。所以，不同的结果会通过不同的行为来实现。这就意味着对某一职务而言，最好的描述绩效的方式是体现出这一职务的行为和结果的关系。通常，合成效标模型多采用结果来描述绩效指标，而多维效标模型多采用行为来描述绩效指标。因此，对于同一职务而言，选择绩效评定方式时需要考虑到作业特征间的差异。

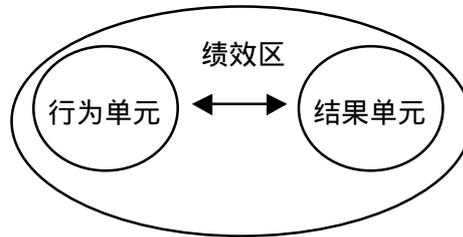


图 2 绩效的行为—结果模型

4 绩效行为指标内涵的扩展——OCB 和关系绩效

由于绩效行为指标主要是通过职务分析的方法来获得，而职务分析所关注的行为都是与作业特征联系紧密的行为指标，所以，传统的绩效评定中所采用的绩效行为指标大多是与作业绩效有关的行为指标。到 20 世纪 80 年代初，Organ 等人明确提出了组织公民行为（OCB）的概念，到 90 年代初，Borman 和 Motowidlo 提出了关系绩效（contextual performance）的概念，这些研究扩展了绩效行为指标的内涵，把那些职务说明书未明确规定的，但是会对组织绩效产生重要作用的行为上升到绩效标准的高度，从而使传统的绩效评定由强调“个体—职位”相匹配的个体层次，向强调“个体—组织”相匹配的组织层次转变。

Katz 和 Kahn 主张有效能的组织具有 3 种基本行为形态：(1) 组织必须吸引并留住员工于组织系统；(2) 确保员工以可信赖方式实现组织特定角色要求；(3) 员工必须有创造性与自发性的行为，其表现超越角色规范。若一个组织仅依规定和规章行事将是一个非常脆弱的社会系统，组织需要有合作帮助、改善建议、利他行为与其他形态的公民行为。Bateman 和 Organ 在 1983 年正式将 Katz 和 Kahn 所提出的第三种组织成员的角色外行为称之为“组织公民行为”（OCB）。组织公民行为是一种个人绩效的特定形式，它是员工个人对其组织所做的贡献，它超出了工作说明书规定的范围，通常未为组织另予报偿*。Podsakoff 认为组织公民行为可以对组织绩效提供 3 个方面的贡献：(1) 组织公民行为可以减少组织对于维持目的的稀缺资源的需要，例如：职工相互间的公民行为可以缓解团队内部的冲突，因此减少了在管理冲突活动中的时间花费；(2) 组织公民行为可以通过制造一个有吸引力的工作环境来增强组织吸引和保留优秀人才的能力；(3) 组织公民行为可以增强组织绩效的稳定性^[13]。整体而言，功能性的组织公民行为将促进组织的效能。

Organ 对 OCB 进行了重新定义，认为它不仅仅是职务外行为，而且是类似于 Borman

* 洪振顺（台）. 组织公正对组织公民行为影响之研究——信任关系之观点. 硕士论文，1998：38-39

和 Motowidlo 提出的关系绩效^[14,15], OCB 也能够对组织的社会和心理环境提供维持和增强的作用。从而把组织公民行为和关系绩效的内涵统一了起来。

Borman 和 Motowidlo 指出, 职务绩效除了包括作业绩效 (task performance) 以外, 还应该包括关系绩效 (contextual performance)^[15]。作业绩效是指任职者通过直接的生产活动、提供材料和服务对组织的技术核心所做的贡献, 它主要受任职者的经验、能力以及与工作有关的知识等因素的影响。关系绩效不是直接的生产和服务活动, 而是构成组织的社会、心理背景的行为, 它可以促进其中的作业绩效, 从而提高整个组织的有效性, 比如自愿承担本不属于自己职责范围内的工作, 帮助同事, 并与之合作完成作业活动, 它主要受任职者的人格因素的影响。Borman 和 Motowidlo 认为, 关系绩效和作业绩效同等重要, 两者都是职务绩效必不可少的组成部分。Van Scotter 和 Motowidlo 将关系绩效细分为敬业精神 (job dedication) 和人际促进 (interpersonal facilitation)^[16]。

Borman 和 Motowidlo 进一步区分了作业绩效和关系绩效, 认为: (1) 作业绩效依不同职务而变化, 而关系绩效在不同职务之间是稳定和类似的; (2) 作业绩效是有明确规定的职务内行为; (3) 作业绩效的有效预测源是认知能力, 而关系绩效的有效预测源是人格^[17]。

关系绩效的提出, 使人们对绩效的结构有了更加清晰的认识, 丰富了绩效行为指标的内涵, 但同时也增加了基于行为的绩效评定的难度, 因为组织公民行为和关系绩效的评定更容易受评定者主观因素的影响。社会认知心理学研究发现, 人们在进行评价时会受到情境因素的影响, 人能够实用地采取加工策略以适应当时情境的需要^[18]。

从 1993 年至今, 尽管在关系绩效的研究领域开展了许多研究, 但是在关系绩效概念的内涵和维度的确定上, 一直存在诸多分歧。另外, 在跨文化研究方面, 也发现不同的文化对关系绩效的理解也存在差异。

表 1 中西方在组织公民行为维度上的差异

Organ (1988) 西方人的 OCB 维度	Farh, Early, Lin (1997) 中国人的 OCB 维度
公民道德(civic virtue)	认同组织(identification with company)
利他(altruism)	协助同事(altruism towards colleagues)
尽职(conscientiousness)	敬业守法(conscientiousness)
运动精神(sportsmanship)	人际和谐(interpersonal harmony)
殷勤有礼(courtesy)	保护公司资源(protecting company resources)

樊景立在研究中发现 OCB 存在文化差异, 中国人的“认同组织”、“协助同事”、“敬业守法”维度与西方人的“公民道德”、“利他”、“尽职”维度相似, 不同之处在于, 西方人的 OCB 维度还包括“运动精神”和“殷勤有礼”, 而中国人的 OCB 维度中还包括“人际和谐”和“保护公司资源”^[19]。梁开广在研究中发现, 中国企业的主管在评价下级的“工

作表现”时,往往包括了对关系绩效的评价*。周智红认为,我国组织中目前的绩效评定易受感情和人际关系等因素的影响,关系绩效往往被赋予更大的权重^[20]。

由于传统的绩效行为指标的获得是来源于职务分析,而职务分析在西方文化背景下所获得的是作业绩效,而不是关系绩效,那么,在中国的文化背景下,用职务分析的方法是否能获得关系绩效的指标呢?

5 绩效评定的实验室研究

从70年代以来,在绩效评定研究领域,针对录像资料进行分析的研究逐渐多了起来,并逐渐发展成为一种绩效评定的研究范式。Borman 在一项有关评定准确性培训的研究中,先为评定者们放映一部关于员工实际工作情况的录像带,然后要求他们对这些员工的工作绩效做出评定^[21]。

Guion 和 Gibson 认为绩效评定中大多数认知方面的研究都有必要在实验室里进行^[11]。这些研究问题包括:(1) 评定者对被评定者的喜欢,如何影响绩效评定的准确性。例如: Cardy 和 Dobbins 通过让演员扮演教师的各种角色,再让大学生观看录像后进行评定的方式来进行实验室研究^[22]。(2) 群体成员和群体行为上共有的差异不会影响评定过程。例如: Feldman 等人的研究。(3) 延迟性评价会导致更多的晕轮效应,但同时要比观看录像后立即评价要准确。例如: Murphy 和 Balzer 等人的研究^[22]。(4) 以前的绩效水平会对当前的绩效评定产生影响,并且后来的绩效水平会对前期的绩效印象产生偏差。Hakel 对实验室研究绩效评定持肯定态度,但是他认为这么好的研究方法并未给实际应用带来可行的研究结果。这一见解推动了关于实验室研究概化推广的研究进程^[23]。

由此可见,采用录像分析的实验室研究范式能够帮助人们更深刻地理解绩效评定过程。不过, Murphy 和 Balzer 等人采用的演员模拟教师的授课过程的研究范式还仅限于个体层次,无法观察到职务之间的协作关系,也无法对多种职务间的异同进行横向比较,并且研究成本较高,录像资料准备的周期较长^[22]。

参考文献

- [1] Cascio W F. Applied psychology in human resource management. 5th ed. USA, Prentice-Hall, 1998. 58-79
- [2] Landy F J, Farr J L. The Measurement of Work Performance. New York: Academic Press, 1983. 27-90
- [3] Henderson R I. Performance Appraisal. 2nd ed. Reston, Va: Reston, 1984. 127
- [4] Smith P C, Kendall L M. Retranslation of Expectations: an Approach to the Construction of Unambiguous Anchors for Rating Scales. *Journal of Applied Psychology*, 1963,47: 149-155
- [5] Keaveny T, McGann A. A Comparison of Behavioral Expectation Scales and Graphic Rating Scales. *Journal of Applied Psychology*, 1975, 60: 695-703
- [6] Wexley K N. Personnel training. *Annual Review of Psychology*, 1984, 35: 519-551
- [7] Zedeck S, Cascio W F. Psychological issues in personnel decisions. *Annual Review of Psychology*, 1984, 35: 461-518
- [8] 周智红, 王二平. 多维效标测量的构成. *心理学报*, 2001, 34 (1): 75-81
- [9] Conway J M. Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 1999, 83: 3-13

* Liang K G. Fairness in Chinese organizations. Doctoral dissertation, 1999

- [10] Lee C. Increasing Performance Appraisal Effectiveness: Matching Task Types, Appraisal Process, and Rater Training. *Academy of Management Review*, 1985, 10: 322-331
- [11] Guion R M, Gibson W M. Personnel selection and placement. *Annual Review of Psychology*, 1988, 39: 349-374
- [12] Binning J F, Barrett G V. Validity of personnel decisions: A conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, 1989, 74: 478-494
- [13] Podsakoff P M, MacKenzie S B. The impact of organizational citizenship behavior on organizational performance: A review and suggestions for future research. *Human Performance*, 1997, 10: 133-152
- [14] Organ D W. Organizational citizenship behavior: it's construct clean-up time. *Human Performance*, 1997, 10(2): 85-97
- [15] Borman W C, Motowidlo S J. Expanding the criterion domain to include elements of contextual performance. In: Schmitt N, Borman W C ed. *Personnel selection in organizations*. San Francisco: Jossey_Bass, 1993
- [16] Van Scotter J R, Motowidlo S J. Interpersonal facilitation and job dedication as separate facets of contextual performance. *Journal of Applied Psychology*, 1996, 81: 525-531
- [17] Borman W C, Motowidlo S J. Task performance and contextual performance : the meaning for personnel selection research. *Human performance*, 1997, 10(2): 99-109
- [18] 朱新秤 . 社会认知心理学研究的新进展 . *心理学动态* , 2000 , 8 (2) : 74-80
- [19] Farh J L. Impetus for action: a cultural analysis of justice and organizational citizenship behavior in Chinese society. *Administrative Science Quarterly*, 1997,42: 421-444
- [20] 周智红 , 王二平 . 作业绩效和关系绩效 . *心理学动态* , 2000 , 8 (1) : 54-57
- [21] Borman W C. Format and training effect on rating accuracy and rater errors. *Journal of Applied Psychology*, 1979,64: 410-412
- [22] Murphy K R, Balzer W K. Systematic distortions in memory-based behavior ratings and performance evaluations: consequences for rating accuracy. *Journal of Applied Psychology*, 1986,71: 39-44
- [23] Hakel M D. Personnel selection and placement, *Annual Review of Psychology*, 1986, 37: 351-380

THE DEVELOPMENT OF BEHAVIOR-BASED PERFORMANCE APPAISAL APPROACH

Ma Chenggong, Wang Erping

(Institute of Psychology, The Chinese Academy of Sciences, Beijing 100101)

Lin Ping

(Department of Humanity and Social Sciences, Beijing College of Mechanical Industries, Beijing 100012)

Abstract : Behavior-based performance appraisal has focused in the research on performance appraisal since the 1960' s. There are some invincible limitations in nonjudgmental measures based on job outcome, so different kinds of behavior-based judgment measures have been successively developed. Selection of behavior items has become the central issue of the approach. Studies such as organizational citizenship behavior, differentiation of task performance and contextual performance has put forward new issues to behavior-based performance appraisal.

Key words : critical incident technique, organizational citizenship behavior, contextual performance, task performance.