

# 人力资源系统与公司绩效关系研究的测量误差:随机误差与系统误差

肖永松,王重鸣

(浙江大学 管理学院,浙江 杭州 310058)

**摘要:**文章对人力资源系统与公司绩效关系实证研究中影响其效应大小的围绕项目、时间和评定者测量信度的随机误差和系统误差进行评述与分析,介绍基于构思水平理解评定者间信度的思路,较好解决了实证中的数据来源问题,并提出对管理感知这一新的系统误差的研究。总结改善信度的方法和探索,讨论了人力资源系统测量中需要改进的地方和进一步的研究方向。

**关键词:**人力资源系统;公司绩效;随机误差;系统误差

中图分类号:F240

文献标志码:A

文章编号:1008-5831(2008)03-0058-05

关于人力资源系统与公司绩效关系的研究相当丰富,并获得了越来越多的正向结果,尤其在高绩效工作系统与公司绩效关系的研究方面成果颇丰。但此类研究始终伴随的方法学问题是:人力资源系统与绩效关系的效应受到人力资源系统测量信度的影响,而测量信度受多种因素影响而不易确定,实证中由信度带来的问题已经引起了系列的研究。

## 一、HR 系统测量的随机误差

Schwab<sup>[1]</sup>认为:确定构思有效的第一步是评估存在于构思测量中的测量误差。HR 系统测量的随机误差一般造成低估 HR 系统与绩效的关系效应值,而系统误差造成的影响可能需要具体分析。为了显示 HR 系统与绩效的真正关系,需要尽量控制各种误差,提高测量的信度。目前主要研究的三种误差来源是项目、时间和评定者,相应地一般用克隆巴赫  $\alpha$  系数、重测信度和评定者间信度来对几种误差来源进行可靠性检验。

### (一)项目引起的随机误差

HR 与绩效关系研究都重视由于项目导致的测量误差,几乎所有此类研究都报告了量表的内部一致性估计,通常  $\alpha$  系数高于 0.60(如 Delery & Doty, Huselid, Huselid, Jackson, Schuler)。但 Nunnally & Bernstein 建议  $\alpha$  系数最低应达到 0.80,很多研究报告的系数仍然低于此水平。为改善一致性,量表编制的主要精力都集中于项目的控制,关键点在于操作化尽量客观具体,实质聚焦,同时消除项目语义模糊性的影响。如果概念覆盖范围较广,回答者可能使用模糊性来进行合理化、重构或错误表征 HR 成分的效果<sup>[2]</sup>;即使客观回答,也会由于语义模糊性加大理解差异而影响评定者间信度。因此,根据收集的数据对量表和项目进行多轮修正是必要的,Huselid 与 Becker<sup>[3]</sup>在这方面做了很多努力,不断“升级”他们的 HR 实践测量以获得更好的 HR 测量项目。Wright<sup>[4]</sup>认为改善内部一致性的另一方面努力是探索不同的评定量表,但对于哪一种是最恰当的

收稿日期:2008-01-12

作者简介:肖永松(1972-),男,四川广元人,浙江大学管理学院博士研究生,主要从事人力资源管理研究;王重鸣(1949-),男,江苏张家港人,浙江大学管理学院教授,博士生导师,主要从事人力资

量表目前仍未达成一致。如 Huselid 与他的同事倾向于使用客观量表,用某一 HR 实践覆盖的员工比例作为指标,而 Delery 与 Doty, Snell 以及 Wright 等一般使用 Likert 量表来评定对 HR 实践使用陈述的同意程度。Wright 认为没有证据支持“使评定客观化会消除甚至是减少测量误差”的观点<sup>[4]</sup>。

## (二)时间引起的随机误差

在 HR 系统测量中,检验时间造成的测量误差研究很少,鲜有研究报告其重测信度。一方面可能因为整个管理研究中的纵向研究越来越少<sup>[5]</sup>,另一方面可能因为现实环境太复杂,不易评估从一个数据样本到下一个的任何变异是由于误差还是 HR 实践的真正改变所致。

## (三)评定者引起的随机误差

### 1. 单一多评定者之争

HR 系统与绩效关系研究中,近些年引起系列研究争论的是关于评定者间信度,主要集中于 Gerhart, Wright, Snell 与 Huselid, Becker 间进行的单一多评定者之争。Gerhart, Wright, Snell<sup>[6]</sup>在 12 个大公司(超过 4 万员工,包括能源、银行等 9 个行业)中挑选 44 名 HR 经理对 10 项人力资源实践进行评定(评定 HR 实践覆盖各类岗位员工的百分比),发现评定者间一致性一般低于 0.50,存在基于评定者的显著测量误差。于是结论认为:基于单一回答者的 HR 实践测量,使 HR 实践与绩效关系的点估计显著受较低的评定者间信度所影响而低估真实关系。认为应使用多个评定者以增加 HR 实践测量的一致性。Huselid 与 Becker<sup>[3]</sup>认为上面研究的结论不可靠,因为 Gerhart 等的研究缺陷在于最大化了评定者间信度问题:选取的公司规模太大,在产业、民族、商业单位来看是异质的,而研究者只聚焦于其中的小规模样本,使 HR 难以准确描述整个公司的实践;没有选择关键的信息提供者,增加其他不掌握情况的评定者反而降低信度;使用的问卷项目在用词上混淆了 HR 政策与实践。他们认为选取一个关键信息提供者比从多个不了解情况的回答者能得到更精确的数据。Gerhart, Wright<sup>[7]</sup>对 Huselid 等的批评进行了解释,吸收了有用部分,认为其指出的问题会对信度有小的影响但不影响根本结论。针对以上争论与进一步的研究建议,Wright, Gardner<sup>[4]</sup>用另外 3 套样本控制相应变量进行了 3 项研究。(1)研究 1:16 个大公司内小样本;管理/专业/技术多类工作岗位;高级 HR 经理与直线经理各 2 名作为评定者;报告各类员工被 15 项 HR 实践覆盖的百分比。(2)研究 2:银行业内 3 种规模的银行各 350 家;具体的 3 类工作岗位;公司 HR 主管与岗位工作人员做评定者;回答 HR 实践测量的 7 点 Likert 量表。(3)研究 3:一家食品服务公司内 33 个服务单位(平均 500 员工);每个单位内都有具体的 6 类工作岗位;每单位 HR 主管和每类工作的 20% 员工对 11 项 HR 的实践评定。

在逐一控制了行业、规模(多样性)、工作岗位后(研究 3 被称为评定者间信度估计的最佳设计),发现从相同企业单位来的相同工作的员工间平均的评

定者间信度很低( $ICC(1) < 0.2$ ),而且员工与 HR 主管的回答相关很小( $R = 0.24$ )。而多评定者累加的信度达到 0.71(单项目)和 0.88(量表)。于是认为:个体评定者,即使是了解情况的现任职者,会提供关于 HR 实践的高度不一致的信息,评定者间一致性没有达到合理水平。研究结果支持“HR 测量单回答者信度低”的结论。Wright 等<sup>[8]</sup>用与研究 3 相似的样本与设计进行了研究,结果支持上述结论。

针对评定者引起的误差问题,Wright 等<sup>[4]</sup>提出 3 点控制方法:(1)要求多评定者(增加评定者比增加项目更有用);(2)使用最知情的评定者(关键信息提供者);(3)不超过每个评定者的信息加工能力,应窄化聚焦于一个工作群体、单个商业单位或一个区域的信息(相应就需要多个评定者)。

Becker 与 Huselid<sup>[9]</sup>提出:在大规模多产业 SHRM 研究中使用多评定者实际上难以获得足够的回答率。并列出了 3 项采用多回答者设计的研究由于只有很低的回答率而实际上联合使用了单回答者设计和多回答者设计。

### 2. 基于构思水平的分析

管理研究方法中,由多特征—多方法矩阵发展而来的三角测量思路要求数据来源多渠道<sup>[5]</sup>,但在实证研究过程中,往往又难以获得适当的回答率。针对以上学者的观点和争论,Jeffrey B. Arthur 与 Trish Boyles<sup>[10]</sup>认为:单一多回答者设计并非非此即彼,相应条件下两种设计都有可行性。可以从 HR 系统结构成分的构思水平出发理解使用单一多回答者设计的恰当性。确定构思来源的水平,并决定在怎样的条件下累加个体水平数据来建立集体水平构思,对构思效度非常重要<sup>[1]</sup>。当测量匹配理论时才有构思效度,当构思水平与测量水平不一致时,发生错误指定问题<sup>[11]</sup>,如盲目累加个体水平测量去表征单位水平构思;使用单位水平构思推断低水平关系。

Jeffrey B. Arthur 与 Trish Boyles<sup>[10]</sup>总结了 HR 系统的 5 个成分:HR 原则、政策、规划/项目、实践、气氛,他们都可以表征组织水平特征,但这 5 种成分自身的水平并不一致,因此其代表组织水平的方式也不一样。前 3 个是整体的、组织水平的构思,是全域性概念,后 2 个是来源于个体水平的集体水平构思。Kozlowski 与 Klein<sup>[12]</sup>认为:产生并显示组织水平构思没有必要考虑跨水平的加工问题,能直接观察或通过文档记录获取信息,也可由掌握情况的关键信息提供者提供数据来源;而来源于个体水平的集体水平构思需要通过多个个体来源数据累积来获取集体信息,如通过组成方式可以以个体数据的平均值代表集体水平构思。

根据构思水平确定数据来源方式将有利于数据来源的可靠性。Jeffrey B. Arthur 认为,HR 实践与气氛属个体水平形成的集体水平构思,须从有经历与感知的多个个体搜集数据。HR 政策属于组织水平构思,最好由高层 HR 经理作为关键信息提供者。HR 原则也属于组织水平构思,最好由高级经理人或创建者作为关键信息提供者。HR 规划(如 360 度绩

效反馈等)属组织水平构思,一个关键信息提供者是恰当的,但由于此方面信息加工比较复杂,往往需要收集多回答者数据,并确定这些信息提供的一致程度;重要的是需要减少模糊性和回答者的信息加工

要求。如果对较大复杂组织进行总结性或要求复杂测算数据的回答,误差肯定增大<sup>[13]</sup>。Jeffrey B. Arthur总结了HR系统成分基于水平的框架(表1)。

表1 HR系统成分基于水平的框架

Hr system construct component	Proposed unit property of construct component	Proposed source level of component (who to ask)
HR system construct component	Global	Org-level business leader(s) acting as informant(s) or respondent(s)
HR policies	Global	Org-level HRM leader(s) acting as key informant(s)
HR programs	Global	Org-level HRM leader(s) acting as key informant(s)
HR practices	Shared or configural	Lower-level managers and employees acting as respondents
HR climate	Shared or configural	Lower-level managers and employees acting as respondents

注:此表改自 Jeffrey B. Arthur<sup>[10]</sup>

基于多水平的思路分析,当HR系统成分属于组织水平时,Huselid所主张的使用一个关键信息提供者是恰当的,因关键信息提供者能对组织水平变量进行总结而得到可靠一致的评定。但当成分属个体水平形成的集体构思时,必须要有多个有经历的个体数据,如果只用“关键人物”来评定,会犯基于水平的误指定问题,更可能以感知代替实际。以上研究都是对HR实践的评定,因此多回答者设计是合理的,能得到比单回答者设计更高的信度。

战略人力资源管理(SHRM)研究中所产生的一些问题,是由于HR系统构思处于组织水平,而其中不同成分有不同的来源水平,但SHRM对HR系统结构测量的数据来源一般是个体资料提供者(大多是HR经理),由此潜在的错误是研究者将基于个体水平的感知或态度的测量关系归因于客观的组织水平的HR项目或实践的观察<sup>[10]</sup>。

## 二、HR系统测量的系统误差

### (一)内隐绩效理论

Gardner等<sup>[14]</sup>实证研究了内隐绩效理论对HR—公司绩效研究中HR测量系统误差的问题:以高级HR经理与直线经理以及学生为被试,给出高、低绩效公司两场景,要他们报告各种HR实践的普及程度和HR功能的效果。结果发现:都倾向于对高(低)绩效公司评定更好(弱)的HR实践;工作经验更加剧这种偏见;工作性质间没有显著差异。而这种好、弱的评价偏见都造成对HR实践与公司绩效关系的过高估计。

由于一般的管理研究情景都是管理者对公司绩效和胜任力的回溯,因此,由内隐绩效理论带来的HR系统与绩效关系效应的系统性高估更可能存在。

### (二)管理感知效应

高于平均水平效应(the above-average effect)指人们总是倾向于高估自己的能力。模糊效应(the

ambiguity effect)指评定模糊能力时,比评定不模糊能力更倾向高估;评定模糊能力或维度时,自己比独立评定者评定更高。在个体被试研究中,这两种效应的研究有较多文献。其解释理论包括动机理论(高估保护自尊)、信息加工理论(关注自己的努力与能力而忽视外部信息)、自我表征理论(有意操纵归因以建立正面形象)。

Thomas C. Powell<sup>[2]</sup>把个体研究中的以上两个效应放大到管理情景中,相应提出两个命题:(1)管理者系统高估自己公司的胜任力;(2)相对于非模糊性胜任力,管理者更高估自己公司的模糊性胜任力。每个命题派生出5个假设,以上12个假设在个体研究中大多都有较强的实证支持,但都没有在管理者对公司胜任力归因的情景中得到检验。Lovallo与Powell<sup>[15]</sup>的研究部分支持以上两命题:让一群独立评定者评定了10个组织特质的模糊性;然后让1480个加拿大公司的管理者评定自己公司在这10个特质上的竞争性强度。结果管理者评定的等级顺序与胜任力模糊性评定间相关达0.94,感知的绩效归因与胜任力模糊性相关达0.87。

HR作为公司取得竞争优势的胜任力之一,在管理者的回答中也当然会受这些管理感知效应的影响。因此,管理者对自己公司的HR系统的评定会被系统化高估。此高估将使HR—公司绩效关系变得更模糊:可能形成高HR与低公司绩效的关系。

## 三、讨论

基于构思水平的思路,实证中HR系统问卷应尽量单维,项目尽量具体聚焦,分别由最合适的一个人或多人来填写;如果不能做到单维,也应尽量将组织水平的成分和由个体发展而成的集体成分分开而指定相应的回答者。如果将各成分汇总在一套问卷里,除了会由于相同填写者而犯同源偏差(高的假相关)外,更会由于构思测量混乱而造成误归因,影

响评定者间信度。

对项目控制的研究很多,关键是体现构思并消除语义模糊性,实质具体聚焦操作化。另外,为了让HR 实践研究真正体现实践性而非回答者的感知,有专家提出应调查HR 实践的特征而不仅是对HR 活动的描述,这开辟了一个形成新量表的思路,难点在于对特征的概括可能难于统一,会增大语义模糊性而影响信度,同时增加了项目编制的难度。目前绝大部分研究仍然使用“是或否”的方式探索人力资源实践的有无,这样的数据容易获得和比较,但在信度方面很容易出问题。Huselid<sup>[16]</sup>与Guest等<sup>[17]</sup>使用了实践覆盖的范围测量;仅Truss<sup>[18]</sup>测量了HR 实践的强度。在大量对HR 实践效能的研究普遍忽视HR 实践执行效果的情况下,努力开发并推行强调实践执行的特征量表对提高研究有效性有较大意义。

关于时间问题研究很少,由于不能分清时间效应与处理效应,需要考虑管理情景中重测的合适时间范围标准或控制机制(如等样本搜寻);或者需要探索新的研究时间问题的方法。无论怎样,都应加强HR 实践的纵向研究,深入揭示HR 与公司绩效的具体关系和模式。由时间问题而引起的系统误差也值得深入研究。

关于评定者问题,系列实证似乎偏向多回答者设计更能减少误差,但在实证中由于低回答率而难以施行,往往是单一多回答者设计相互结合。以往学者在这方面的争论并没有一种很明确的理论思路作为支持而主要从实证来检验。在Boselie, Dietz 与Boon<sup>[19]</sup>的HR 实践与绩效关系研究的综述中,共检验了1994-2003年共104篇文献,其中63个研究中使用了单一评定者设计(其中26项使用了一个人力资源管理人员作为评定者);41个研究使用了多评定者设计,可见存在测量信度上的较大分歧。一些研究者建议根据研究问题来选择被试,Wright<sup>[20]</sup>认为最好由高级经理来评价总的人力资源的有效性,Paul 与Anantharaman<sup>[21]</sup>认为由员工来对单个人力资源实践进行回答。Jeffrey B. Arthur 与Trish Boyles 基于多水平理论从构思水平确定数据来源的角度对这一问题进行了分析,一定程度上解释了以上争论,确立了一种决定单一多评定者设计恰当性的思路,对HR 系统实证研究具有现实意义。

SHRM 研究与组织文献中不同地位回答者对同一成分的评定显示出系统性的差异是相当普遍的(Kumar1993),这种系统性的差异(如工人与管理者报告的HR 原则、政策)有时其实并非对HR 成分存在与否的个体评定,而可能实际上表征了不同的构思,如工人与管理者的不一致或冲突,这样的情况下如果简单使用多回答者设计,实际上并不能增加信度,甚至可能降低信度。但如果将不同个体看作是对HR 成分不同侧面反映的话,多评定者设计是有效的。因此,根据HR 系统成分的构思水平决定数据来源的思路确定对不同HR 系统成分的回答者选择是恰当的。

然而,Jeffrey B. Arthur 与Trish Boyles 基于构思

水平的思路只是一般地使用了构成方式(composition)简单考虑个体水平数据向集体水平变量的累积,没有深入分析构成方式(compilation)以及在HR 系统测量中数据累积方式的具体条件和模式,而这对于完整地基于构思水平思路来理解数据来源和信度以及进行具体操作具有重要意义,值得结合多水平理论进一步研究。

对HR 系统测量中的系统误差研究很少,主要表现为评定者。内隐绩效模型在HR-公司绩效关系中得到了实证支持,结果总是高估HR 与公司绩效的关系。而管理者感知的平均之上和模糊两个效应对HR 与公司绩效关系的影响不确定:当公司高绩效时,管理感知导致HR 与公司绩效关系的高估;当公司低绩效时,管理感知减弱HR 与公司绩效的关系。

关于在管理情景中的系统误差还需要进一步的实证支持,特别是管理感知的两种效应。基于两种效应的三种理论解释角度考虑,笔者认为一方面将个体背景的两个效应放到组织背景下有一定的合理性,至少都涉及到个体感知以及对自己的保护;另一方面,在组织背景下,管理者对公司胜任力的系统高估的感知应该会受到组织承诺、组织归属感等因素的影响,可以假设组织归属感强的管理者更容易产生这些效应。因此加强这方面的实证研究,弄清管理情景下的效应模式,会促进对系统误差的研究。如果这些系统影响都存在,那他们各自的效应相对强度以及交互的影响最终会高估还是低估HR 与公司绩效的关系,也值得进一步研究。

尽管目前管理研究以问卷调查方式收集数据为主,但为避免HR 测量中的这些信度问题,应该使用多种数据收集方式:如用档案记录核实问卷回答的准确性,深度访谈,二手、三手资料补充等方式。否则,一定程度上说,再怎么对问卷以及实施过程进行控制,都不能很有效地解决信度问题,因此,采用多方法,重视定性分析的作用,这也正是管理研究方法发展的趋势。

#### 参考文献:

- [1]SCHWAB. Construct validity in organizational behavior[J]. Research in Organizational Behavior, 1980(2): 3-43.
- [2]THOMAS C. POWELL. Causal ambiguity, management perception, and firm performance[J]. Academy of management review, 2006,31(1): 175-196.
- [3]HUSELID, BECKER. Comment on "Measurement error in research on human resources and firm performance: how much error is there and how does it influence effect size estimates?" by Gerhart, Wright, Mcmahon and Snell[J]. Personnel psychology, 2000(53): 835-854.
- [4]WRIGHT, GARDNER, ETC. Measurement error in research on human resources and firm performance: additional data and suggestions for future research[J]. Personnel psychology, 2001(54):875-901.

[5]SCANDURA. Research methodology in management: current

- practice, trends, and implications for future research [J]. *Academy of management journal*, 2000, 43(6): 1248 - 1264.
- [6] GERHART, WRIGHT, MCMAAN, SNELL. Measurement error in research on human resources and firm performance: how much error is there and how does it influence effect size estimates? [J]. *Personnel psychology*, 2000(53): 803 - 834.
- [7] GERHART, WRIGHT, MCMAHAN. Measurement error in research on the human resources and firm performance relationship: further evidence and analysis [J]. *Personnel psychology*, 2000(53): 855 - 872.
- [8] WRIGHT, GARDNER, MOYNIHAN, ALLEN. The relationship between HR practices and firm performance: Examining Causal order [J]. *Personnel Psychology*, 2005(58): 409.
- [9] BECKER, HUSELID. Strategic human resources management: Where do we go from here? [J]. *Journal of Management*, 2006(32): 898 - 925.
- [10] JEFFREY B. ARTHUR, TRISH BOYLES. Validating the human resource system structure: A levels - based strategic HRM approach [J]. *Human Resource Management review*, 2007, 17, 77 - 92.
- [11] ROUSSEAU. Issues of level in organizational research: Multi - level and cross - level perspectives [M] // CUMMINGS L STAW B. *Research in organizational behavior*. Greenwich, CT: JAI Press. 1985: 1 - 37.
- [12] KLEIN, KOZLOWSKI. Multilevel analytical techniques [M] // Klein K J, KOZLOWSKI W J. *Multilevel theory research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey - Bass, 2000. 512 - 553.
- [13] COLLINS, SMITH. Knowledge exchange and combination: The role of human resource practices in the performance of high Technology firms [J]. *Academy of Management Journal*, 2006, 49, 544 - 560.
- [14] GARDNER, WRIGHT, GERHART. The HR - firm performance relationship: Can it be in the eye of the beholder? [C]. Paper presented at August Academy of Management Meetings, Chicago, IL, 1999.
- [15] LOVALLO, POWELL. Ambiguity bias and strategy theory [C]. Paper presented at the Strategic Management Society Conference, Paris. 2002.
- [16] HUSELID. The impact of human resource management practice on turnover, productivity, and corporate financial performance [J]. *Academy of management journal*, 1995, 38(3): 635 - 672.
- [17] GUEST, MICHIE, ETC. Human resource management and corporate performance in the UK [J]. *British journal of industrial relations*, 2003, 41(2): 291 - 314.
- [18] TRUSS. Complexities and controversies in linking HRM with organizational outcomes [J]. *Journal of management studies*. 2001, 38(8) 1121 - 1149.
- [19] BOSELIE, DIETZ, BOON. Commonalities and contradictions in HRM and performance research [J]. *Human resource management journal*, 2005, 15(3): 67 - 94.
- [20] WRIGHT, MCMAHAN, SNELL. Comparing line and HR executives' perceptions of HR effectiveness: services, roles and contributions [J]. *Human resource management*, 2001, 40(2): 111 - 123.
- [21] PAUL, ANANTHARAMAN. Impact of people management practices on organizational performance: analysis of a causal model [J]. *International journal of human resource management*. 2003, 14(7): 1246 - 1266.

## Measurement Error in Research on the Relations between Human Resources and Firm Performance: Random Error and Systematic Error

XIAO Yong-song, WANG Zhong-ming

(School of Management, Zhejiang University, Hangzhou 310058, China)

**Abstract:** This paper reviews random error and systematic error which result from items, times and raters in research on the relation between human resources and firm performance and introduces level-based approach to analyze the problem of data - collection, as well as a new kind of systematic error. Then it summarizes some approaches to improve reliability. Some limitations and future researches are discussed.

**Key words:** HR system; firm performance; random error; systematic error

(责任编辑 彭建国)