

# 企业业绩评价行为特征分析

## ——以 ZARA 为例的一项实验研究

吴 革, 代寒冰

(对外经济贸易大学 国际商学院, 北京 100029)

**摘要:** 文章采用实验研究方法探索业绩评价者在综合评价中的行为特征, 以全球知名服饰公司 ZARA 两家门店为实验背景材料, 实验参与者为拥有一定工作经验的 EMBA 与 MBA 学员。研究发现评价者相对于非财务指标来说通常更加关注财务指标; 在给予相关战略信息时, 评价者的财务指标偏好有所减弱; 评价者信息处理能力有限, 通常关注不超过四个指标。

**关键词:** 业绩评价; 指标偏好; 战略信息; ZARA; 实验研究

**中图分类号:** F234.3 **文献标志码:** A **文章编号:** 2095-8838(2015)04-0037-07

### 一、引言

当前的企业业绩评价正沿着综合业绩评价的模式不断探索和前进。典型的综合业绩评价如卡普兰和诺顿提出的平衡计分卡, 该体系建立了业绩评价和战略实施之间的联系, 摒弃了以往业绩评价中单纯关注财务指标的做法, 把财务、顾客、内部流程、学习与成长四个层面的内容联系起来设置均衡的指标体系, 在同时注重过程和结果的情况下全面考察企业业绩。

综合业绩评价是一个综合各方面信息、参考一定标准的包含着评价者主观认识的行为过程。在综合评价模式下, 业绩评价者的主观行为所具有的特征将成为影响评价结果不可忽视的重要因素, 对其进行研究和分析对于业绩评价朝着更科学更客观方向的发展意义重大。

### 二、文献综述

由于综合业绩评价体系涉及各类指标, 一些学者研究发现, 在评价过程中评价者对这些指标的关注程度有所差别。Lipe 和 Salterio(2000) 用实验方法得出: 虽然可以

设计出符合不同情况的平衡计分卡用以更有针对性地评价某一单位业绩, 但只有共同性指标会对评价结果产生影响, 即独特性指标的作用并不明显。Roberts 等 (2004) 通过实验证明, 采用“分解——汇总”把评价过程分为两步, 先决定各指标权重再对具体得分进行加权加总, 可以使评价者在关注共同性指标的同时也关注独特性指标, 从而克服“共同性指标偏好”。Humphreys 和 Trotman (2011) 通过两个实验研究了战略信息和战略相关业绩指标在消除共同性指标偏好方面的作用, 研究发现: 当经理被提供有战略信息并且一部分指标是与战略相关时, 存在共同性指标偏好; 当所有的指标都与业绩相关但是不提供战略信息时, 该偏好依然存在; 当提供战略信息并且所有指标都与战略相关时, 共同性指标偏好消失了。

何晴 (2006) 采用了与 Lipe 和 Salterio(2000) 的实验相同的案例, 比较了在个体评价和集体评价两种不同的方式下主观偏好对结果产生的不同影响, 研究得出中国的业绩评价者在使用平衡计分卡进行业绩评价时同样存在着共同性指标偏好。吴国灿 (2010) 沿用了和 Banker 等 (2004) 类似的实验案例, 得出评价者存在共同性指标偏

收稿日期: 2015-03-01

作者简介: 吴 革 (1967-), 男, 江苏南京人, 教授, 博士生导师;

代寒冰 (1990-), 男, 河南洛阳人, 对外经济贸易大学国际商学院硕士研究生。

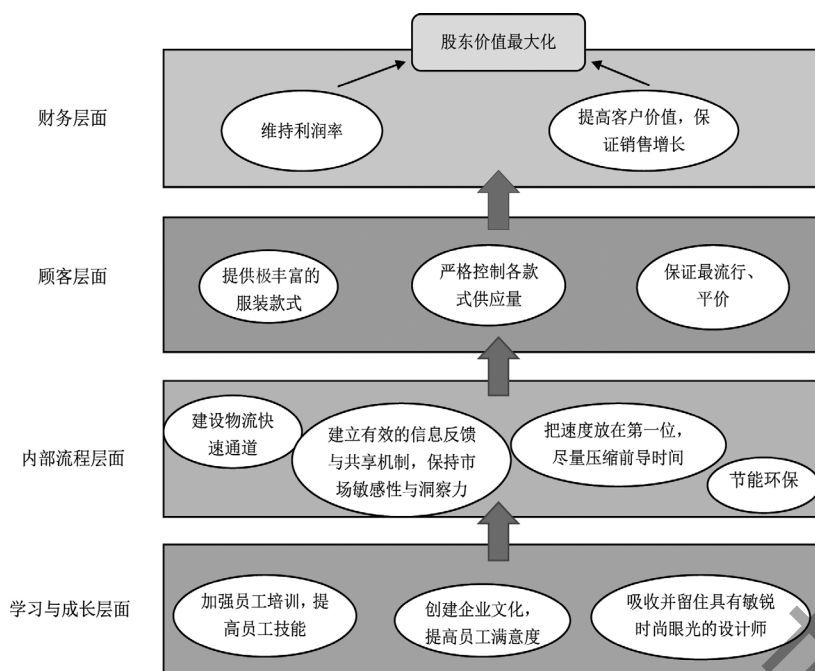


图1 ZARA公司战略地图

好的现象，而采用改进战略信息表述和分解决策两种方法均有效减轻了该偏好现象。刘俊勇等（2011）通过实验研究也对这种共同性指标偏好做了研究，指出无论是拥有或者不拥有战略信息，实验参与者在评价业务单位业绩时对共同性指标的依赖程度都高于独特性指标。

行为受心理的影响，根据心理学的认知负荷理论，人们拥有有限的对信息的认知和处理能力，因此过大的信息量对于大多数评价者来说其实是没有被充分利用的。这种信息超载会带来认知超载（Cognitive overload）的后果。这种后果在综合业绩评价这种包含了各方面大量信息的模式下可能表现得更突出。Halford等（2005）的研究指出，一个人必须同时依据多项信息做决定时，要想保证决定的质量，最多只能同时处理四个信息。Neumann等（2008）的研究表明，在实验设计中的八项指标中只有一项指标（ROI）显著解释了评价结果的绝大部分的差异。

从上述文献中可以看到，大部分文章都证实了业绩评价者对共同性指标的偏好，一些学者还探究了缓解这一偏差行为的相关方法，但是财务指标与非财务指标之间的选择偏好却并未得到讨论，战略信息的作用大小尚待验证，此外对认知超载问题也尚未达成一致意见。基于前人的研究成果，本文运用实验研究方法对企业业绩评价中评价者的突出行为特征进行进一步研究分析，设计了综合业绩评价指标体系，加入了战略地图作为重要战略信息，重点讨论评价者对于财务信息和非财务信息的偏好、有用战略信息在评价中的影响以及实际中评价者信息处理能力的大小。

### 三、研究假设与实验设计

#### （一）研究假设

财务指标截至目前已经发展得相当完善，具有成熟的指标体系。在现行会计准则下，企业的对外报告也以财务相关内容的披露为主，进一步强化了财务指标的核心地位。在指标特点上，财务指标具有较强的可度量性、对业绩衡量的直接性、表现方式的直观性以及较高的客观性。而非财务指标则大多难以直接以货币计量，计量手段存在多元性，从而就易产生主观性强的结果，同时非财务指标衡量的是企业经营过程的表现，而这些在过程上好的表现是否必然导致同样好的经营结果尚不得而知，那么这样的评价便显得具有更强的间接性。综上，本文提出假设1：

H1：其他条件不变时，相对于非财务指标而言，评价者在进行业绩评价决策时会更加重视财务指标。

评价者只有对企业或组织的战略、经营方式和特点有较好的了解才可能做出更符合实际的评价。非财务指标与企业或组织经营过程密切联系，在“过程”业绩衡量方面有着独特优势，是财务指标所不能比拟和替代的。因此非财务指标更能有效反映企业或组织的战略实施过程的情况。评价者在得到有用的战略信息之后便对企业或组织战略有更好的理解，从而在评价过程中会更加关注该战略是否有效实施，因此也就会更加关注通过非财务指标传达的战略实施的信息。基于此，本文提出假设2：

H2：其他条件不变时，较之于没有获得相关战略信息的业绩评价者，拥有相关战略信息的业绩评价者在进行业绩评价决策时会给予非财务指标更多关注。

个体对信息的认知和处理能力是有限的，过大的信息量会引起认知资源的分配不足，带来认知超载。而综合业绩评价如平衡计分卡一般会运用16~28个指标。Halford等（2005）认为一个人在保证决策质量的前提下最多只能处理四个信息。Neumann等（2008）的实验结果则指出，在同时处理八项业绩指标时，只有一项指标显著影响了评价者做出业绩评价决策。这都证明评价者信息处理能力有限，但就评价者能具体处理几项信息的问题上并未达成一致结论。虽无直接说明，但从平衡计分卡的指标体系中可以看出其隐含着评价者可以同时关注多个指标的假设。在我国，国有企业绩效评价体系也包含多项指标，隐含着同样的假设。因此本文认为虽然评价者处理信息的

表1 两分店业绩表现

指标	A店			B店		
	预期值	实际值	实际优于预期 (%)	预期值	实际值	实际优于预期 (%)
财务信息						
1. 销售增长率	14%	17.15%	22.50	15%	16.11%	7.40
2. 毛利率	46%	51.28%	11.47	44%	44.02%	0.05
3. 净利润增长率	16%	20.10%	25.63	15%	15.32%	2.13
4. 成本利润率	54%	60.74%	12.48	52%	53.98%	3.81
5. 投资回报率	10%	12.31%	23.10	10%	10.13%	1.30
非财务信息						
1. 顾客满意度	80%	81.00%	1.25	80%	96.00%	20.00
2. 顾客年均光顾次数	17	17.12	0.07	17	20.98	23.41
3. 提供款式数量(种)	5 400	5 560	2.96	5 500	6 210	12.91
4. 库存周转率(次)	13	13.56	4.31	15	16.85	12.33
5. 员工满意度	75%	80.63%	7.51	75%	94.00%	25.33

能力有限但并非仅限于只处理一项相关信息, Halford 等的结论更显合理。然而研究该问题的文献相对较少, 本文在 Halford 等 (2005) 的基础上提出假设 3:

H3: 无论是信息简略还是信息详细, 评价者主要基于不超过四项的业绩评价指标来做出评价决策。

## (二) 实验设计

### 1. 实验案例材料介绍

在本次实验设计的案例中, 假定实验参与者(即评价者)是 ZARA 公司的中国区高层管理者, 负责对 ZARA 在北京地区的两家分店的经营业绩作出评价。为了对假设做出验证, 本实验将实验参与者随机分成两组来进行, 分别是甲组(信息简略组)和乙组(信息详细组)。

以下是实验参与者所获取的共同说明材料: “ZARA 公司是主要针对年轻人提供女性服装、男性服装以及男女箱包、饰品, 并同时兼售童装的全球四大服装零售商之一。ZARA 从 2006 年开始在中国大陆开设门店以来, 截至目前已经在中国大陆开设了超过 130 家门店。ZARA 公司可谓全球零售时装业的巨头, 它开创了精准时尚模式……ZARA 在中国大陆的分店全部采用直营店的方式经营, 各门店均由西班牙总公司直接开设。在我们的这次调查研究中涉及 ZARA 在北京的两家分店。A 店位于王府井商业街, 开设于 2011 年; B 店位于国贸世贸商业中心, 开设于 2007 年。这两家分店所在的地方均是北京繁华的商业区, 交通便利, 人流量大, 充满时尚气息, 也都是年轻人所喜爱和经常光顾的地方。”

以上材料是甲组实验参与者除具体的业绩指标表

现外所能获得的全部信息。对于乙组实验参与者, 实验为其提供更为丰富的增量信息, 即 ZARA 的战略信息<sup>①</sup>及战略地图(见图 1)。

两个分店的业绩采用同样一套指标体系来展现, 该套业绩指标体系分别包括 5 个财务指标及 5 个非财务指标共 10 个具体指标。其中财务指标分别为销售增长率、毛利率、净利润增长率、成本利润率、ROI(投资回报率)。非财务指标依据平衡计分卡的三个非财务方面并结合 ZARA 的战略而设计, 顾客层面包括顾客满意度和顾客年均光顾次数, 内部流程层面

包括提供款式数量和库存周转率, 学习与成长层面为员工满意度指标。其中的顾客年均光顾次数、提供款式数量以及库存周转率, 是本文根据 ZARA 的经营特点和策略而特别设定的指标, 能够与 ZARA 公司的经营战略相匹配, 更好地衡量该公司经营过程中战略的实施程度和效果。

为保证实验尽量接近现实, 表 1 中两家分店业绩的“实际值”均以 ZARA 母公司 Inditex 集团公司 2013 年度财报披露数据为基础并加以适当调整。“预期值”则是以 ZARA 近三年的平均业绩为基础对两分店各自进行了略微调整后的结果。总体上看 ZARA 两家分店的表现均好于目标值, 但具体表现有所不同, 本文通过设计, 使得 A 店在财务指标上表现好, 而在非财务指标上表现一般; 相反, 使得 B 店在非财务业绩上表现好, 而在财务业绩的表现一般。

本文选取对外经济贸易大学国际商学院会计系的 EMBA 和 MBA 学员作为实验参与者, 下文研究中所采用的数据全部来自于该实验。此次实验共发放 120 份问卷, 回收后经检查, 全部为有效样本。在 120 位实验参与者中, 共有 66 位男学员, 占总人数的 55%; 全部实验参与者的平均工作年限约是 9.3 年, 其中最长工作时间为 18 年, 最短工作时间为 3 年。对实验参与者的行业经验调查显示, 大部分实验参与者对于服装行业的了解处于一般水平, 满足实验要求。

### 2. 变量设置

为了满足检验假设 1 和假设 2 的要求, 本文以信息含量作为影响业绩评价者对指标偏好的因素, 分两种状态,

<sup>①</sup>由于篇幅限制, 此处不予详细表述。ZARA 公司的战略主要是推行“多款、少量、平价、快速”的经营模式, 这四个方面成为该公司在市场上的强大竞争优势。

表2(a) 两分店得分描述统计结果

实验组别	分店名称	样本量	均值	标准差	极小值	极大值
甲组	A店	60	81.90	4.64	73	92
	B店	60	75.08	6.08	64	88
乙组	A店	60	81.89	4.64	73	91
	B店	60	78.38	5.11	69	90

表2(b) 两分店得分T检验结果

实验组别	假定	对等方差的Levene检验		对等均值的T检验		
		F	P值	t	df	P值(双尾)
甲组	等方差	5.31	0.02	6.91	118	0.000
	非等方差			6.91	110.37	0.000
乙组	等方差	0.20	0.66	3.95	118	0.000

表3(a) 甲乙两组各自分差绝对值的描述统计结果

组别	样本量	均值	标准差	极小值	极大值
甲组	60	7.92	4.92	1	22
乙组	60	5.15	3.69	1	16

表3(b) 甲乙两组分差绝对值T检验结果

假定	对等方差的Levene检验		对等均值的T检验		
	F	P值	t	df	P值(双尾)
等方差	5.24	0.02	3.49	118	0.001
非等方差			3.49	109.40	0.001

即信息含量高与信息含量低。实验参与者作为案例中涉及公司的高级管理人员对下属两家分店的业绩表现进行评价,打分的分布从0(表现极差)到100(表现极好)不等。由于受指标运用偏好的影响,因此不同的参与者对两家分店的打分会存在差异,那么两家分店得分的分差便可以代表在指标运用上的偏好。对于信息详细的影响也可以依据分差大小的变化来考察。另外,对于那些在评价过程中有显著影响的指标,其影响最终也表现在分差上。因此本文选取分差作为因变量,用Y来表示A店得分减去B店的得分,若分差为正,则表明评价者偏好财务指标,若为负则反之。

在检验假设3时,本文以实验参与者对十项指标的重要性排序作为自变量,这十项指标分别被每位实验参与者赋予1(最重要)到10(最不重要)中的一个序位,并且不能重复。然后使用逐步回归来研究有多少项指标在业绩评价过程中对指标运用的偏好产生了显著影响。

### 3. 关于信度和效度的说明

实验问卷的设计研究者常常需要考察该问卷的信度和效度。对于本文来说,在财务维度与非财务维度的刻画与反映方面,本文使用的是可以通过具体数字对研究对象

进行准确描述和反映的显性指标,这些指标分别是依据财务指标与非财务指标的定义而设置,具有明确的归属与区分,因此无需进行信度和效度检验。

## 四、实验结果与分析

本文对于实验数据的分析主要采用SPSS17.0作为统计工具,进行了描述统计分析、参数检验、方差分析以及逐步回归分析。

### (一) 指标运用偏好一般性表现分析

根据甲组(信息简略组)和乙组(信息详细组)的打分情况,首先进行描述统计分析。

从表2(a)可以看出,甲组A店与B店的得分均值分别为81.90分、75.08分,甲组对A店的评分明显高于B店。从极大值与极小值来看,A店的得分都高于B店。以上数据表明,甲组对于在财务业绩上表现好的A店给予了更高的评价,而实际上两分店在财务和非财务业绩上各有优势,这就说明评价者更侧重财务指标。乙组A店得分均值为81.89分,B店得分均值为

78.38分,B店的得分均值较之于被甲组打分时更高。均值上,乙组对A店的评分仍明显高于B店。另外从极值来看,虽然差距都有所缩小,但A店的得分仍然均高于B店。以上数据表明,乙组(信息详细组)同甲组(信息简略组)一样,对于在财务业绩上表现好的A店给予了更高的评价。

进一步通过T检验来验证评价者对于两分店的打分均值是否具有统计意义上的显著差别。结果如表2(b)所示。

从表2(b)中可以看出,甲组方差齐性检验的F统计量的值为5.31,对应的P值为0.02,小于显著性水平5%,因此两总体的方差有显著差异,对应表中第二行数据:t统计量的值为6.91,双尾P值小于0.01,因此A店与B店得分的均值在0.01的显著性水平上存在着显著性差异,结合具体数值可知A店的平均得分显著高于B店。这表明甲组显著偏好于财务业绩表现好的分店。乙组方差齐性检验中F统计量的值为0.20,对应的P值为0.66大于显著性水平5%,因此两总体的方差没有显著差异。相应的,对等均值假设的检验表明:t统计量的值为3.95,双尾P值小于0.01,因此A店与B店得分的均值在0.01的显著性水平上存在着显著性差异,结合具体数值又可以得知A店

的平均得分显著地高于B店。这个结果表明乙组也显著偏好于财务业绩表现好的分店。

综合分析,无论是甲组(信息简略组)还是乙组(信息详细组),对于财务业绩好的分店均给予了明显更高的评分,这就说明评价者在业绩评价过程中更偏好于财务指标。即在其他条件相同的情况下,相对于非财务指标而言,评价者进行决策时会更加重视财务指标。从而假设1得到了验证。

#### (二) 增量战略信息对指标运用偏好的影响分析

本实验的分组完全遵循随机的原则,且为大样本实验(各组样本量均为60),那么各组实验者本身的个体特征(如个人行为、思维特征)虽有不同但不会出现集中分布而引起实验系统性偏差。经过这样的控制之后,组间分差绝对值的差异(若显著)便主要由各组实验条件的不同,即信息量的差别来决定。因而根据本文的思路,在其他条件都相同时,若战略信息的加入有助于提高评价者对于非财务指标的关注,那么加入了战略信息的组(乙组)中两分店的分差就会较没有战略信息的组(甲组)要小,即两分店的得分会更加接近。由于分差可能为正也可能为负,为了更方便地考察分差的变化,本文对分差采用绝对值进行描述。

从表3(a)可以看出,甲乙组分差绝对值的均值分别为7.92分、5.15分,前者大于后者;另外,从极大值大小来判断,甲组的极大值为22分,大于乙组的极大值16分。从分差的原始数据(文章中未列出)可以看到,甲组分差极小值为-8分,而乙组分差极小值为-6分,从而也说明乙组的分差范围小些。并且从表3(a)标准差一栏中可以看出,乙组的标准差小于甲组,其分差绝对值的分布更为集中。这初步验证了本文关于有效信息含量对于业绩评价中指标偏好的假设。此外,本文还对两组分差的绝对值进行了T检验,结果如表3(b)所示。

从表3(b)中可以看出,方差齐性检验的F统计量的值为5.24,相应P值为0.02,小于显著性水平5%,因此两总体的方差存在显著差异。对应观察表3(b)中第二行数据,对等均值检验的结果显示:t统计量的值为3.49,对应P值小于0.01,从而在显著性水平为1%的条件下两总体的均值具有显著差异。再结合实际数值,说明甲组的分

表4(a1) 甲组回归模型方差分析

变差来源	SS	df	MS	F	P值
回归	1 883.04	4	470.76	50.18	0.000
残差	515.94	55	9.38		
总计	2 398.98	59			

表4(a2) 甲组回归模型系数

解释变量	未标准化系数		标准化系数	t	P值	共线性统计量	
	B	标准误差	Beta			容差	VIF
(常量)	18.72	1.55		12.07	0.000		
销售增长率	-2.21	0.20	-0.77	-11.14	0.000	0.82	1.23
净利润增长率	-0.69	0.14	-0.32	-4.79	0.000	0.91	1.10
毛利率	-0.39	0.15	-0.18	-2.66	0.010	0.86	1.16
顾客年均光顾次数	0.33	0.14	0.16	2.44	0.018	0.97	1.03

表4(b1) 乙组回归模型方差分析

变差来源	SS	df	MS	F	P值
回归	1 303.17	4	325.79	51.52	0.000
残差	347.82	55	6.32		
总计	1 650.98	59			

差绝对值显著大于乙组的分差绝对值。由于实验的其他条件都相同,因此以上结果表明有效的增量战略信息在评价者的评价过程中起到了显著作用,非财务指标的重视程度得到了显著提高,从而使得B分店的得分与在没有战略信息的甲组评价相比时有显著提高,从而使两店的得分差距有显著缩小,尽管此时A店的得分仍显著( $P < 0.01$ )高于B店<sup>②</sup>。假设2在此得到了验证。

#### (三) 评价过程中所运用指标的数量研究

在考察究竟有多少个指标在评价者的评价过程中产生了显著影响时,本文采用逐步线性回归的方法,将分差作为因变量,将十个业绩指标相对排序的序位作为自变量进行研究。该部分的研究依然将研究过程分为两个部分:首先对信息简略的甲组对指标的排序情况进行逐步回归;然后对信息详细的乙组对指标的排序情况进行逐步回归。由于篇幅限制,仅列出逐步回归最终结果。

本文构造基本线性回归模型如下:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + e_i \quad (k=1,2,\dots,10; i=1,2,\dots,60) \quad (1)$$

其中 $Y_i$ 代表因变量,即两分店得分的分差。 $X_{ki}$ ( $k=1,2,\dots,10; i=1,2,\dots,60$ )代表自变量,即各业绩指标。

甲组最终回归模型 $R^2$ 为0.79,调整 $R^2$ 为0.77,估计

②见本节第一部分“指标运用偏好一般性表现分析”结果。

表4(b2) 乙组回归模型系数

解释变量	未标准化系数		标准化系数	t	P值	共线性统计量	
	B	标准误差	Beta			容差	VIF
(常量)	2.84	2.13		1.34	0.187		
销售增长率	-0.75	0.20	-0.34	-3.67	0.001	0.44	2.27
库存周转率	0.75	0.18	0.37	4.23	0.000	0.50	1.98
顾客年均光顾次数	0.42	0.17	0.20	2.52	0.015	0.59	1.71
净利润增长率	-0.35	0.17	-0.16	-2.06	0.045	0.64	1.55

标准误为3.06, 回归模型方差分析及系数见表4(a1)、表4(a2)。

从表4(a1)、表4(a2)中可以看出, 对甲组的数据进行逐步回归后共有四个变量进入了模型, 依次为: 销售增长率、净利润增长率、毛利率、顾客年均光顾次数。这四个变量对分差的解释能力超过了70%(调整 $R^2=0.77$ ), 模型的拟合度很高, 其中销售增长率的解释能力最强。回归方程的显著性上, F统计量的值为50.18, 对应的P值小于0.01。因此最终的回归方程显著, 四个解释变量联合起来对分差有着显著( $P<0.01$ )影响。此外, 四个变量之间并不存在严重多重共线性。在回归参数的显著性检验(T检验)除最终模型中毛利率的系数( $P=0.010$ )、顾客年均光顾次数的系数( $P=0.018$ )之外, 所有方程参数的显著性均在0.01的水平上显著。从整体上来看, 各个模型的参数均在5%的显著性水平上显著。

通过以上检验, 可以得出, 共有四个指标在信息简略组评价者的评价过程中对结果产生了显著( $P<0.01$ )影响。在本文设定显著性水平5%的条件下, 没有更多的变量被纳入模型中, 因此对于信息简略组, 假设3得以验证。

乙组最终回归模型 $R^2$ 为0.79, 调整 $R^2$ 为0.77, 估计标准误为2.52, 回归模型方差分析及系数见表4(b1)、表4(b2)。

从表4(b1)、表4(b2)中可以看出, 在5%显著性水平上, 对乙组的数据进行逐步回归以后依然是只有四个变量进入了模型中, 依次为: 销售增长率、库存周转率、顾客年均光顾次数、净利润增长率。我们注意到由于获得信息不同, 评价者对指标的关注度也有了调整, 库存周转率作为非财务指标在无详细战略信息的甲组中没有表现出对分差的显著影响, 而在乙组的评价过程中一跃成为第二强的解释变量。顾客年均光顾次数也超过净利润增长率成为了第三强的解释变量。这说明在非财务指标中这两项指标对案例中ZARA公司的战略反映程度最高, 最具相关性。进一步来看, 这四个变量联合起来对分差的解释能力也超过了70%(调整 $R^2=0.77$ ), 模型的拟合度很高, 销售增长率的解释能力依旧最强。回归方程的显著性检验上, F统计量的值为51.52, 对应的P值小于0.01。因此最终

的回归方程显著, 四个解释变量联合起来对分差有着显著( $P<0.01$ )影响。回归模型中的变量之间同样也不存在严重的多重共线性。在回归方程参数的显著性检验(T检验)方面: 逐步回归过程中, 在5%的显著性水平上对于最终模型, 除了常数项未通过显著性检验外, 其他四项指标均通过了显著性检验。

通过以上分析, 可以得出, 在信息详细组评价者的评价过程中仍然是共有四个指标对结果产生了显著( $P<0.01$ )影响。在5%显著性水平的条件下, 依然没有更多的变量被纳入模型中, 只是对于信息详简不同的两组, 其对分差产生显著影响的指标变量不同而已。因此对于信息详细组, 文章的假设3也得以验证。

### 五、结论及对策建议

本文研究了综合业绩评价过程中财务指标与非财务指标受重视水平以及信息相对充分条件下影响业绩评价的指标的个数, 通过对比不同信息含量情形及要求实验者主观排序, 运用方差分析法和逐步回归法, 对于平衡计分卡背景下企业业绩评价者指标运用的决策特征进行了分析, 得出以下结论:

第一, 通常情况下, 相对于非财务指标来说评价者总是倾向于更加关注财务指标。第二, 给予充分相关的战略信息有助于业绩评价者提高对非财务指标的关注程度。在实验中可以看出在其他条件没有发生变化的情况下, 拥有详细相关战略信息的乙组对两分店打分的分差显著小于甲组对两分店打分的分差。第三, 在信息相对充分的条件下, 实质上影响业绩评价者做出评价决策的指标数量不多于四个。

针对本文研究的结论, 综合前人的研究以及结合实验设计中的思考, 在此提出改善综合业绩评价效果的对策建议: 第一, 制定清晰的企业战略是做好综合业绩评价工作的前提。第二, 评价体系的设计要合理公正, 充分反映企业战略, 格式呈现要清晰明了。第三, 业绩评价者需要深入了解企业各方面的工作。第四, 给予评价者多维、充分的业绩信息, 同时更加强调整战略信息。

本文是一项基于实验的研究, 研究的进行需要选择具

体的一家企业来作为背景,选取的ZARA公司是传统行业(服装业)中发展较为成熟的一家企业,对于社会整体而言具有一定的代表意义,尤其对于传统行业中处于成熟期的企业而言代表性较高。但是文中所述评价者的行为特征是否具有行业独特属性或者发展阶段的特性本文尚未探究,因此不能保证研究结论能够适用于所有的行业以及一家企业所有的发展阶段。对于当前新的经济时代中出现的其他新兴行业以及大多数的初创期的企业,本文的结论尚需留待进一步的探究。此外,尽管本文的实验尽量模拟

实际工作环境,但是人为控制的实验环境与真实工作环境之间存在差异性是不可避免的,比如缺乏监督机制和诸多的利益约束就可能会造成一些评价者随意填写问卷,而在实际工作中评价者通常会出于自身责任而表现得更加认真。然而本文在遵循成本效益原则下,考虑到问卷内容并不复杂且给予作答的时间相对充分,并且实验参与者能够认真地予以配合,从而认为该实验能够相对有效地对所研究的命题做出考察,因此达到了文章的研究目的。但并不排除存在对实验进行优化改进的可能。

#### 主要参考文献:

- [1] 何晴. 通用性指标偏好及其实验研究——平衡计分卡在多部门企业中的应用[J]. 清华大学学报, 2006, S1(21): 60-69.
- [2] 刘俊勇, 孟焰, 卢闯. 平衡计分卡的有效性: 一项实验研究[J]. 会计研究, 2011, (5): 36-43.
- [3] 吴国灿. 基于战略的业务单元业绩评价研究——多重业绩指标应用的视角[D]. 成都: 西南财经大学, 2010.
- [4] Banker, R.D., Chang, H., Pizzini, M.J. The Balanced Scorecard: Judgmental Effects of Performance Measures Linked to Strategy[J]. The Accounting Review, 2004, 79: 1-23.
- [5] Halford, G.S., Baker, R., McCredden, J.E., et al. How Many Variables Can Humans Process[J]. Psychological Science, 2005, 16(1): 70-76.
- [6] Humphreys, K.A., Trotman, K.T. The Balanced Scorecard: The Effect of Strategy Information on Performance Evaluation Judgments[J]. Journal of Management Accounting Research, 2011, 23: 81-98.
- [7] Lipe, M.G., Salterio, S.E. The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures[J]. The Accounting Review, 2000, 75: 283-298.
- [8] Neumann, B.R., Roberts, M.L., Cauvin, E. Financial and Nonfinancial Performance Measures[J]. Cost Management, 2008, 22(6): 5-14.
- [9] Roberts, M.L., Albright, T.L., Hibbets, A.R. Debiasing Balanced Scorecard Evaluations[J]. Behavioral Research in Accounting, 2004, 16: 75-88.

## The Analysis of Behavior Features in Enterprise Performance Evaluation

—— An Experimental Study on ZARA

WU Ge, DAI Han-bing

**Abstract:** The purpose of this paper is to explore performance evaluators' behavior features in comprehensive evaluation by experimental research method. The world famous clothing company ZARA's two stores were used as background materials in the experiment, EMBA and MBA students who have certain working experience as participants. The study shows that evaluators usually pay more attention to financial measures; when given relevant strategic information, evaluators' preference on financial measures has been weakened; and evaluators usually focus on no more than four measures because their information processing abilities are limited.

**Key words:** performance evaluation; indicators preference; strategic information; ZARA; experimental study

(责任编辑 杨亚彬)