

# 应用多质-多评价者程序对评定误差的分析\*

李 锋 王二平 张福松

(中国科学院心理研究所,北京 100101)

**摘 要** 应用多质-多评价者矩阵分析绩效评定中两种常见误差-光环效应和自我中心效应。以足球运动员为被评价者,以关键事件技术为职务分析方法获得绩效评定特质。以45名被试观看足球比赛后评定的各职务特质间的相关为基础相关系数;两组共8名被试观看录像剪辑并采用两种评价标准作评价,计算各特质评定结果间的相关系数,并整理为多质-多评价者矩阵。结果表明,多质-多评价者矩阵可以直观显示光环效应;评价者在不同职务的个别特质上表现出自我中心效应。

**关键词** 绩效评定,光环效应,自我中心效应,多质-多评价者程序。

**分类号** B849:C93

## 1 问题

绩效评定(performance appraisal)是人力资源管理的重要手段。由于绩效评定过程的复杂性,绩效评定结果受多种因素的影响,而产生各种评定误差。为提高绩效评定质量,分析、控制各种评定误差是人事心理学研究的重要课题。

绩效评定中采用的效标测量一般包括判断性测量(judgment measures)和非判断性测量(nonjudgmental measures)两类<sup>[1]</sup>。非判断性测量主要有以下三方面的局限性:受情景因素的影响较大;表示的是行为的结果而非行为本身;某些职务找不到适合的客观指标衡量其绩效<sup>[2]</sup>。因此,判断性测量是绩效评定普遍应用的方法。然而,判断性测量可能产生各种评定误差。评定误差表现为多种形式,最为常见的评定误差主要有:宽容误差与苛严误差、趋中趋势、光环效应、自我中心效应、顺序效应。光环效应是研究较多的误差之一。表现在绩效评定过程中的光环效应是指由于评价者基于对被评价者的总体印象,或被评价者的某些突出特征影响到评价者对其他特征的评定,使得评价者无法在不同的绩效维度上判断绩效水平,表现为个体不同绩效维度上的绩效水平有过高的正相关<sup>[3]</sup>。

Cooper<sup>[4]</sup>在回顾前人关于光环效应研究和理论分析的基础上提出了减小虚光环效应的七种措施并

回顾了分析光环效应的五种方法。Landy等人<sup>[5]</sup>利用偏相关技术对绩效评定中的光环效应进行了统计上的控制。但处理的结果只是得到了一个相关减小了的特质相关矩阵,而无法得到排除光环效应的评定结果。Murphy等人<sup>[6]</sup>认为光环效应并非是无所不在的,有时观察到的特质相关比实际特质相关还要低,产生所谓的负光环效应。Lance等人<sup>[7]</sup>分析了三个光环效应的原因模型-总体印象模型、突出维度模型和区分不足模型。这引起了对光环效应的新注意,使人们认识到这种现象远比过去理解的复杂。

多质-多评价者矩阵是研究者用来分析绩效评定数据的一种方法,这种方法来源于Campbell和Fiske<sup>[8]</sup>应用多质-多法矩阵对区分效度和汇聚效度的研究。Lawler等人<sup>[9]</sup>在研究绩效评定方法中将多质-多法矩阵演化为多质-多评价者矩阵。“多质”是指被评定特质,“多法”是指不同的评定源,比如不同的评定方法。Kavanagh等人<sup>[10]</sup>将方差分析用于分析多质-多评价者矩阵。基于评价者、被评价者、评定特质三因素的方差分析模型,通过相关矩阵计算评定过程中的各种变异成分,分析汇聚效度、区分效度和评价者的光环效应。这种方法只适合于分析大型的多质-多评价者的矩阵。Kavanagh等人对多质-多评价者矩阵的分析可以得出光环效应在整个评定中的影响,表明的是整个评定的总体质量如

收稿日期:2000-11-20。

\*国家自然科学基金资助项目(79770090)。

何。但这种方法也无法排除光环效应的影响。

在管理实践中,一般有几名主管上司来对中层管理人员和专业技术人员进行评定,评价者人数不多,如果某个评价者的误差较大,则对整个评定结果影响较大。因此多质-多评价者程序中,若评定源以评价者个体出现,那么分析评价者个体的评定误差更有实际意义。通过比较多质-多评价者矩阵中的异质-同评价者的相关三角中的相关系数,判断哪个评价者具有较大的光环效应,即相关系数显著高于其他评价者的评价者有较大的光环效应;同质-异评价者的相关对角线中的相关系数则表明的是评价者在同一特质上评定的一致程度,比较这些系数可以发现可能具有“自我中心效应”的评价者。

## 2 方法

研究通过对足球爱好者的访谈,运用关键事件技术对前锋、前卫、后卫、守门员 4 种职务进行职务分析获得绩效评定指标。以得出的绩效指标为评定量表中的评价特质,采用 Likert 六点式量表,量表上的数字 1 和 6 分别代表绩效的最差和最好水平,其余数字介于其间的绩效水平。每个被评定特质放在单独的一张表上,每个特质下的被评价者(队员)随机排列,对于每份量表随机排列特质顺序,并且在排列顺序时不区分被评价者的职务(场上位置),以排除顺序效应。每张表上标明特质名称以及根据职务分析整理结果得出的每个特质的评价标准。

2.1 被试 足球爱好者。

2.2 评定材料 中央电视台第五套节目直播的甲 A 联赛(重庆隆鑫 vs. 四川全兴)上半场的比赛实况;复制的甲 A 联赛录像剪辑(约 90 分钟)。

2.3 程序 以直播实况为评定材料的被试在直播比赛上半场结束之后进行评定;以录像剪辑为评定材料的被试随机分为两组,一组两人使用歪曲的评价标准(将绩效特质所包含的内容人为改变后形成新的评价标准作为歪曲标准);另一组 6 人使用正常标准。被试重复三次观看录像剪辑(每次约 90 分钟),每次看完之后根据场上运动员的表现以及量表要求做出评定。

2.4 数据分析 根据以直播实况为评定材料被试的评定结果计算各职务下对于每个被评价者的特质之间的相关;根据以录像剪辑为评定材料被试的评定结果计算评定特质间的相关系数,并整理为多质-多评价者矩阵。

## 3 结果

根据半开放式访谈中 41 名球迷对 4 种职务的分析(对每一职务回答以下 4 个问题:该位置在一支球队中的作用,设置该位置的目的是什么;在该位置上应该如何发挥才能发挥其作用,达到其目的,什么样的行为有助于达到该目的,什么样的行为会有碍于达到目的;结合具体实力来说明,具体分析你记忆中的某些球员的表现并做出评价;请被访者对前面提到的种种绩效行为或绩效要素按照对该位置的重要性排序)的记录整理出凡是两人以上提到的各职务的绩效行为,按照频次和重要性对各职务的绩效行为进行归类、合并并概括出各职务的绩效特质。

表 1 后卫绩效特质

绩效特质	该特质包含内容(评定标准)
技术娴熟,基本功扎实	长传准、头球好、脚头硬,远射、任意球技术好
战术意识,默契配合	卡位准、配合默契,组织能力强,易控制比赛节奏
心理素质	遇事不急躁,及时化解场上矛盾
场上作风,精神风貌	自信,敢拼抢,跑动积极,注意力集中
对抗能力	冲撞合理,不无畏犯规,善于抢点
抢断技术	出脚快,果断,不漏人和球

表 2 前卫绩效特质

绩效特质	该特质包含的内容(评定标准)
技术娴熟,基本功扎实	脚法好,传球准确到位,远射准,任意球技术好
战术意识,默契配合	配合默契,创造机会,助攻积极,站位好,不粘球
心理素质	遇事不急躁,不无畏犯规
场上作风,精神风貌	敢于拼抢,跑动积极,自信
对抗能力	下脚准、狠、果断,有冲击力
视野开阔,组织攻防	协调,调度能力强,易控制比赛节奏

表 3 前锋绩效特质

绩效特质	该特质包含的内容(评定标准)
技术娴熟,基本功扎实	射门脚法好,易摆脱后卫,善于远射、突破,左右脚均衡
战术意识,默契配合	善于与队友配合,跑位积极准确,为队友创造机会,不粘球
心理素质	心理状态稳定,射门冷静、沉着
场上作风,精神风貌	积极跑动、穿插,有信心、有激情,敢于突破
对抗能力	有冲击力
捕捉战机,创造机会	把握机会,门前抢点、争顶、就地反抢,射门意识好
门前冷静,果断	临门一脚不慌,射门果断,善于补射

在对各职务绩效特质的概括中,凡是属于队员个人素质的指标,如后卫的“速度,转身灵,高大,判断力”等,以及含义过于概括或模糊或难以通过行为

进行判断的指标,如前锋的“灵感、创造性,个人魅力,天分”等,均未被概括入绩效特质。而“心理素质”,“场上作风、精神面貌”虽然不是具体行为,但单独将“心理素质”作为绩效表现的人数较多,使得该特质很突出;“场上作风,精神风貌”应属于“关系绩效”,是与职务奉献、人际促进有关的绩效<sup>[11,12]</sup>,因此将“心理素质”、“场上作风、精神风貌”作为单独绩效特质。

表 4 守门员绩效特质

绩效特质	该特质包含内容(评定标准)
基本技术	扑球稳、技术简洁、实用,能处理各种球,球性好
把握出击时机	出击时机准,活动范围大,站位恰当
心理素质	心理素质稳定,沉着冷静,状态稳定
场上作风,精神风貌	敢扑,精神集中,自信,处理球果断
组织后卫攻守	与后卫配合好,考虑全局,组织进攻

据 45 名被试观看现场直播后对 7 名后卫,9 名前卫,4 名前鋒(因守门员人数太少而未作评定)评定的结果计算各职务下对于每个被评价者的特质之间的相关,若某被评价者的特质之间的相关系数多数大于 0.40 则剔除该被评价者的数据,用剔除某些被评价者之后的数据计算各职务的各特质间的相关系数。

后卫、前卫、前锋特质相关计算中分别剔出 1 名、4 名、1 名被评价者。研究中所开发的评定量表是在评定项目下评运动员,每一评定项目单放在一张纸上并且随机排列,每一评定项目下的被评定者也随机排列,这样就排除了顺序效应对评定结果的影响,表 5、6、7 中显示的后卫、前卫、前锋各职务下各特质间的相关,是各评价者平均评定结果间的相关。Kenny 等人<sup>[13]</sup>通过数学模型分析认为平均评定特别是大样本的平均评定的确能够减小相关偏差。研究样本均大于 30,因而所得平均评定结果具有较小的相关偏差,因而将各职务下各特质之间的相关作为特质之间相关的基础值是合理的。也就是

说显著大于相应表中相关系数的评定结果则具有光环效应。

据两组共 8 名被试重复 3 次看完录像剪辑后对 45 名后卫、42 名前卫,33 名前鋒(每次评 15 名后卫、14 名前卫、11 名前鋒)的评定结果计算每种职务下评定特质之间的相关系数并整理为多质-多评价者矩阵(如附录 1 后卫多质-多评价者矩阵)。表中两竖线间的三角内表示的是异质-同评价者相关系数,通过不同三角内数据的比较可以确定具有较大光环效应的评定者;表中黑体数字表示同质-异评价者相关系数,通过比较这些数据确定评定者是否存在自我中心效应。

将各职务下的异质-同评价者相关系数以及各特质间的基础相关系数(表 5、6、7)转换为 Fisher  $z$  值,并进行配对  $t$  检验,结果如表 8 所示。

表 5 后卫特质相关系数

特质	1	2	3	4	5
1 技术娴熟,基本功扎实					
2 战术意识,默契配合	0.33				
3 心理素质	0.11	0.23			
4 场上作风,精神风貌	0.27	0.16	0.28		
5 对抗能力	0.34	0.32	0.19	0.27	
6 抢断技术	0.38	0.26	0.20	0.42	0.37

表 6 前卫特质相关系数

特质	1	2	3	4	5
1 技术娴熟,基本功扎实					
2 战术意识,默契配合	0.35				
3 心理素质	0.36	0.34			
4 场上作风,精神风貌	0.28	0.34	0.26		
5 对抗能力	0.26	0.35	0.34	0.49	
6 视野开阔,组织攻防	0.49	0.43	0.38	0.36	0.29

表 7 前锋特质相关系数

特质	1	2	3	4	5	6
1 技术娴熟,基本功扎实						
2 战术意识,默契配合	0.48					
3 心理素质	0.39	0.40				
4 场上作风,精神风貌	0.29	0.44	0.37			
5 对抗能力	0.36	0.38	0.34	0.25		
6 捕捉战机,创造机会	0.47	0.39	0.43	0.50	0.36	
7 门前冷静,果断	0.36	0.29	0.53	0.18	0.43	0.42

表 8 异质-同评价者相关系数  $t$  检验值

评价者	1	2	3	4	5	6	7	8
后卫	2.37*	3.29**	12.0**	2.54*	-0.52	1.12	8.22**	9.55**
前卫	1.71	3.06**	5.37**	2.81**	2.72*	3.04**	8.43**	3.80**
前锋	0.51	3.22**	3.02**	8.77**	0.44	3.80**	7.13**	7.93**

注: \*表示  $p < 0.05$ ; \*\*表示  $p < 0.01$ 。

表 9 异质-同评价者相关系数 Fisher  $z$  值平均数

评价者	1	2	3	4	5	6	7	8
后卫	0.41	0.61	0.80	0.40	0.27	0.35	0.72	0.64
前卫	0.44	0.54	0.72	0.58	0.48	0.49	0.70	0.62
前锋	0.44	0.64	0.62	0.89	0.43	0.56	0.74	0.81

据表 8 中  $t$  值及其显著性水平可知对于后卫评价者 2、3、7、8,对于前卫评价者 2、3、4、6、7、8,对于前锋评价者 2、3、4、6、7、8 的异质 - 同评价者相关系数与基础相关系数在 0.01 水平上存在显著性差异。

将各职务下的异质 - 同评价者相关系数转换为 Fisher  $z$  值并求平均数,结果如表 9 示。

表 9 表明对于后卫评价者 2、3、7、8,对于前卫评价者 2、3、4、6、7、8 对于前锋评价者 2、3、4、6、7、8 的评价结果平均值较大,根据研究假设这些评价者在不同职务下具有较大的光环效应。

应用 Excel 中的 Fisher 函数将同质-异评价者相关系数转化为 Fisher  $z$  值并在 SPSS 统计软件包中进行 Fisher  $z$  值间的  $t$  检验,如附录 2 后卫同质-异评价者相关系数  $t$  检验值。

根据研究假设,歪曲组被试 7、8 与正常组被试间的同质-异评价者相关系数与正常组被试间的同质-异评价者相关系数存在显著性差异;而被试 7、8 间的同质-异评价者相关系数与正常组被试间的同质-异评价者相关系数不存在显著性差异,但表中数据并不完全支持该假设。对附录 2 中数据的进一步的统计结果表明对于后卫这一职务评价者 3、8 之间的同质-异评价者相关系数与其他 74% 的同质-异评价者相关系数至少在 0.05 水平存在显著性差异,这说明评价者 3、8 之间对评定特质的理解相对于整体存在不一致。根据相同的分析,对于前卫这一职务,评价者 1、2(与其他 52% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),3、4(与其他 59% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),4、5(与其他 52% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),6、8(与其他 52% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异)之间对评定特质的理解相对于整体存在不一致。对于前锋这一职务,评价者 2、4(与其他 52% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),3、6(与其他 63% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),4、7(与其他 56% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异),3、5(与其他 52% 同质-异评价者相关系数至少在 0.05 水平存在显著性差异)之间对评定特质的理解相对于整体存在不一致。

统计不同职务下同质-异评价者相关矩阵中相关系数  $\geq 0.40$  的数据所占百分比,结果如表 10 所示,百分比表示不同评价者对某一评价特质评价的一致性水平,百分比越高一致性越高。

表 10 异质-异评价者相关系数的数量分析 (%)

特质	1	2	3	4	5	6	7
后卫	18	54	14	54	32	46	
前卫	43	39	32	21	7	39	
前锋	46	75	61	71	25	21	32

将正常组与歪曲组各职务下不同特质上的同质-异评价者相关系数进行配对  $t$  检验,结果如表 11 所示。数据表明,对于后卫评价者在特质 5 上表现出自我中心效应;对于前卫,评价者在特质 1 上表现出自我中心效应;对于前锋,评价者在特质 3 上表现出自我中心效应。

表 11 两组间不同特质上同质-异评价者相关系数  $t$  检验值

特质	1	2	3	4	5	6	7
后卫	-1.03	-1.34	-0.99	-0.67	-2.08*	-0.32	
前卫	2.71**	-1.32	-0.55	-0.21	-0.52	-1.05	
前锋	0.41	-0.35	-2.03*	0.06	-0.75	-0.52	-1.25

注: \*表示  $p < 0.05$ ; \*\*表示  $p < 0.01$ 。

## 4 讨 论

关键事件技术是侧重于职务行为重要性的职务分析方法。足球运动中运动员的表现直观地显现在赛场上,因此关键事件技术适合描述其职务行为并进一步确定其绩效特质。此外,在概括绩效特质时是以职务行为为基础,按照行为的性质异同和频次概括的,并且在访谈时问题也集中在职务绩效表现上,因而量表中所用的职务绩效特质能够代表职务的绩效特征。在量表制定过程中,将各职务评定特质随机排列并且将各特质下的被评价者随机排列,这样主要是为了排除顺序效应。特质顺序排列的随机性可以使得相邻特质之间的较高相关得以相互抵消。

由表 9 可以看到对于后卫这一职务,评价者 2、3、7、8 的评价结果平均值较大,这种偏差在对各职务的任职者的评定中表现的完全一致。据假设则认为评价者 2、3、7、8 具有较大的光环效应。由表 8 中数据可知评价者 2、3、7、8 的评定结果在 0.01 水平上与期望值存在显著性差异,根据对基础相关系数的分析可知评价者 2、3、7、8 对于后卫的评价结果存在光环效应( $p = 0.01$ )。根据相同的分析可知,对于前卫,评价者 2、3、4、6、7、8 的评价结果存在光环效应( $p = 0.01$ );对于前锋,评价者 2、3、4、6、7、8 的评价结果存在光环效应( $p = 0.01$ )。此外可以看出,

评价者 2、3、7、8 的评定结果在后卫、前卫、前锋三种职务上均存在光环效应 ( $p = 0.01$ ); 而评价者 1、5 评价结果对于三种职务均不存在光环效应 ( $p = 0.01$ )。这说明, 评价者是影响光环效应的重要因素之一。而在数据分析过程中发现评价者对高水平运动员的评价结果之间的相关高于对一般运动员的评价结果之间的相关, 这说明被评者也是影响光环效应的一重要因素。这一结果与 Murphy 等人<sup>[14]</sup>的结论相一致。

由前文分析可知对于后卫这一职务评价者 3、8 之间对评定特质的理解相对于整体存在不一致。对于前卫这一职务, 评价者 1、2、3、4、4、5、6、8 之间对评定特质的理解相对于整体存在不一致。对于前锋这一职务, 评价者 2、4、3、6、4、7、3、5 之间对评定特质的理解相对于整体存在不一致。此外由表 10 可知, 对于后卫这一职务评价者在特质 1、特质 2 上存在较低的一致性; 对于前卫这一职务, 评价者在特质 4、特质 5 上存在较低的一致性; 对于前锋这一职务, 评价者在特质 5、特质 6、特质 7 上存在较低的一致性。总之, 个别评价者之间对于评价特质的理解在总体上存在差异; 所有评价者之间则在个别特质理解上存在较大的差异。

由表 11 可知, 对于后卫这一职务两组评价者之间在特质 5 上存在显著性差异; 对于前卫这一职务两组评价者在特质 1 上存在显著性差异; 对于前锋这一特质两组评价者在特质 3 上存在显著性差异, 这些差异表明歪曲组与正常组评价者之间对该特质的理解的不一致。而在实验实施时的确对于两组被试给予不同的评价标准, 这说明实验结果在个别特质上与预期相一致(歪曲组与正常组评价者在整体上存在显著性差异), 也就是说歪曲组评价者表现出自我中心效应。值得指出的是, 两组间差异未表现在整体上而是表现在个别特质上, 这有可能是因为给出的其他特质的歪曲的标准本身在评价者的理解上不存在显著性差异。总之, 在不同职务的个别特质上评价者表现出自我中心效应。

以上的分析中认为被评价者是影响光环效应的因素之一, 而本研究中选择被评价者时并未考虑这一点。另外, 评价者的自我中心效应没有体现在整体特质上而体现在不同职务的个别特质上可能是由于给出的歪曲特质标准在评价者的理解上并未达到与一般解释的显著性差异, 因而在今后研究中应对歪曲标准进行预测验以确认给出的标准与正常标

准在评价者整体理解上是否存在显著性差异。

## 5 结 论

(1) 多质-多评价者矩阵可以直观显示评价者的光环效应; 评价者与被评价者是影响光环效应的重要因素。

(2) 通过分析多质-多评价者矩阵可以看出评价者在不同职务的个别特质上存在自我中心效应。

## 参 考 文 献

- 1 Landy F J, Farr J L. The measurement of work performance: methods, theory, and application. New York: Academic Press, 1983. 67 ~ 77
- 2 Cascio W F. Applied psychology in Human Resource Management (5<sup>th</sup>ed). Upper Saddle River: Prentice Hall, 1998. 64 ~ 68
- 3 Landy F J, Farr J L. Performance rating. Psychological Bulletin, 1980, 87(1): 72 ~ 107
- 4 Cooper W H. Ubiquitous halo. Journal of Applied psychology, 1981, 90(2): 218 ~ 244
- 5 Landy F J, Vance R J, Barnes - Farrell J L, Steele J W. Statistical control of halo error in performance ratings. Journal of Applied Psychology, 1980, 65(4): 501 ~ 506
- 6 Murphy K R, Jako R A, Anhalt R L. Nature and consequences of halo error: a critical analysis. Journal of Applied psychology, 1993, 78(2): 218 ~ 225
- 7 Lance C E, Lapointe J A, Stewart A M. A test of the context dependency of three causal models of halo rater error. Journal of Applied Psychology, 1994, 79(1): 332 ~ 340
- 8 Campbell D T. Convergent and discriminant validity by the multitrait - multimethod matrix. Psychological Bulletin, 1959, 56(2): 81 ~ 105
- 9 Lawler E E. The multitrait - multirater approach to measuring managerial job performance. Journal of Applied Psychology, 1967, 51: 369 ~ 381
- 10 Kavanagh M J, Mackinney A C, Wolins L. Issues in managerial performance: multitrait - multimethod analysis of ratings. Psychological Bulletin, 1971, 75(1): 34 ~ 49
- 11 Van Scotter J R, Motowidlo S J. Interpersonal facilitation and job dedication as separate facets of contextual performance. Journal of Applied Psychology, 1996, 81(5): 525 ~ 531
- 12 Motowidlo S J, Van Scotter J R. Evidence that task performance should be distinguished from contextual performance. Journal of Applied Psychology, 1994, 79(4): 475 ~ 480
- 13 Kenny D A, Berman J S. Statistical approaches to the correction of correlational bias. Psychological Bulletin, 1980, 88(2): 288 ~ 295
- 14 Murphy K R, Anhalt R L. Is halo error a property of the rater, rateres, or the specific behaviors observed? Journal of Applied psychology, 1992, 77(4): 494 ~ 500

附录1 后卫多质-多评价者矩阵

	RA1			RA2			RA3			RA4			RA5			RA6			RA7			RA8								
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3			
RA1	T1	49	31	T1	58	31	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25
	T2	58	31	T2	32	60	T2	22	26	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	31	33	T3	44	49	T3	22	26	T3	13	17	T3	21	07	T3	34	18	T3	34	18	T3	34	18	T3	34	18	T3	34	18
RA2	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA3	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA4	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA5	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA6	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA7	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07
RA8	T1	22	26	T1	18	13	T1	10	21	T1	26	25	T1	25	34	T1	26	25	T1	26	25	T1	26	25	T1	26	25	T1	26	25
	T2	26	16	T2	13	17	T2	21	07	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18	T2	34	18
	T3	16	41	T3	41	47	T3	22	26	T3	13	17	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07	T3	21	07

注:表中数字均省略个位0, 两线之间三角内的数字表示异质-同评价者相关系数, 黑体数字表示同质-异评价者相关系数, RA 表示评价者, RA 后的数字表示评价者序号。

附录 2 后卫同质 - 异评价者相关系数  $t$  检验值

	R12	R13	R14	R15	R16	R17	R18	R23	R24	R25	R26	R27	R28	R34	R35	R36	R37	R38	R45	R46	R47	R48	R56	R57	R58	R67	R68
R13	0.03																										
R14	0.25	0.31																									
R15	0.67	0.46	0.23																								
R16	0.37	0.56	0.21	0.01																							
R17	0.75	0.90	0.21	1.63	1.07																						
R18	0.00	0.02	0.33	0.40	0.75	0.79																					
R23	2.59*	2.32	2.31	2.32	2.98*	2.02	2.31																				
R24	1.11	1.43	1.86	1.40	2.70*	0.94	1.57	1.16																			
R25	0.59	0.50	0.48	0.26	0.22	1.37	0.59	3.41*	1.71																		
R26	0.39	0.38	0.29	0.12	0.16	0.73	0.35	2.80*	2.03	0.03																	
R27	1.64	1.32	1.41	1.50	1.39	0.84	0.94	1.44	0.04	1.61	2.28																
R28	1.64	2.03	2.18	1.64	3.91**	1.44	2.03	1.02	0.89	2.19	2.98*	0.63															
R34	1.50	2.15	3.52*	2.03	2.36	1.69	2.06	0.74	0.43	1.84	1.70	0.29	0.27														
R35	0.44	0.41	0.10	0.15	0.10	1.64	0.52	2.79*	1.37	0.38	0.17	1.24	1.87	1.80													
R36	1.22	1.37	1.79	1.92	1.28	1.02	1.13	0.76	0.04	1.35	1.20	0.07	0.35	0.40	1.30												
R37	1.31	1.25	1.75	1.65	1.17	1.07	1.01	1.07	0.08	1.55	1.13	0.06	0.50	0.41	1.43	0.16											
R38	8.05**	10.3**	7.75**	7.18**	5.33**	9.50**	5.99**	0.80	2.25	4.55**	3.15*	2.61*	1.74	2.76*	8.40**	2.67*	4.37**										
R45	0.14	0.15	0.01	0.12	0.13	0.74	0.20	2.06	1.26	0.41	0.23	0.98	1.77	1.92	0.06	1.13	1.47	4.18**									
R46	1.92	2.17	2.99*	2.05	2.56*	2.01	2.04	0.05	1.33	2.58*	3.68**	1.32	0.88	1.17	2.00	1.02	1.42	0.89	3.17*								
R47	1.45	1.60	2.17	1.54	1.64	1.13	1.16	1.13	0.00	1.66	2.06	0.08	0.63	0.37	1.36	0.04	0.17	3.26*	1.46	2.30							
R48	1.34	2.24	2.51*	1.44	3.27*	1.15	1.76	1.57	0.28	1.65	1.56	0.17	1.40	0.74	1.68	0.22	0.10	4.48**	1.45	1.92	0.35						
R56	0.42	0.43	0.53	0.21	0.19	1.13	0.50	2.04	1.74	0.01	0.03	1.29	1.84	2.70*	0.26	1.81	1.67	4.02**	0.60	3.35*	1.81	1.53					
R57	0.78	0.82	1.48	1.23	1.06	0.74	0.96	1.06	0.16	1.89	1.05	0.11	0.56	0.50	1.16	0.20	0.10	2.48	2.96*	1.48	0.18	0.00	3.67				
R58	0.50	0.59	0.93	1.06	0.93	0.17	1.05	1.61	0.67	1.15	0.59	0.46	1.03	1.11	1.46	0.67	0.57	4.30**	0.76	1.41	0.57	0.57	1.01	0.59			
R67	1.17	1.50	1.78	1.55	1.46	1.00	1.22	0.68	0.16	1.26	1.31	0.16	0.30	0.26	1.25	0.27	0.24	2.31	1.17	1.02	0.17	0.36	1.67	0.26	0.69		
R68	1.70	2.04	4.07**	2.26	2.32	2.13	2.11	0.49	0.80	2.53*	2.20	0.63	0.11	0.89	2.03	0.81	0.96	1.97	3.67**	1.01	0.99	1.19	5.91**	1.39	1.43	0.65	
R78	2.09	2.57*	4.32**	2.63*	2.57*	2.96*	2.77*	0.39	0.79	2.68*	1.86	0.73	0.27	0.88	2.82*	0.86	1.71	2.51*	4.95**	0.53	1.10	1.54	3.63**	1.69	2.11	0.71	0.37

注: R 表示相关系数, R 之后的两个数字分别表示评价者序号, \* 表示  $p < 0.05$ ; \*\* 表示  $p < 0.01$  水平上显著。

## THE MULTITRAIT-MULTIRATER APPROACH TO ANALYSING RATING BIASES

Li Feng , Wang Erping , Zhang Fusong

( *Institute of Psychology, the Chinese Academy of Sciences , Beijing 100101* )

### Abstract

This research used the multitrait-multirater matrix to analyze two kinds of bias: halo effect and egocentric effect. Professional football players were rated in this research. The performance traits of four different positions were obtained using Critical Incident Technique. The basic correlation coefficient was the correlation among position traits rated by 45 subjects after watching one football match. Eight subjects who were divided into two groups watched part of one cassette while part of them changed their rating standard. The correlation coefficients among the rating results of each trait were computed and then sorted into a multitrait-multirater matrix. The result showed that halo effect could be identified explicitly in multitrait-multirater matrix, and the raters showed egocentric effect in some specific traits of different positions.

**Key words** performance appraisal ,halo effect ,ego-centric effect ,multitrait-multirater matrix.