

리더십 다면평가에서 나타난 평가자 조망*

박 영 석†

가톨릭대학교 심리학과

본 연구는 현장에서 사용된 리더십의 다면평가에서 평가자 수준(상사, 동료, 부하)별로 리더십의 평가에서 사용하는 평가조망을 탐색하였다. 150명의 경영자 후보를 피평가자로 하고 각자의 상사 1명, 동료 2명, 그리고 부하 10명이 평가자로 참여하였다. 피평가자의 리더십은 리더십 경쟁가치 모델(Quinn, 1988)에서 사용된 네 가지 리더 행동 즉 성과지향 행동, 관계지향 행동, 안정지향 행동, 그리고 혁신지향 행동의 준거에서 평가되었다. 연구 결과, 리더 행동의 평가에서 평가자 수준 간 상관이 매우 낮게 나타났으며, 그 원인이 평가자 수준에 따라 범하는 평정오류의 차이와 평가자 수준에 따라 채택하는 리더십 준거가 다르다는 결과를 관찰하였다. 상사는 성과와 혁신 조망을, 부하는 관계와 안정 조망을, 그리고 동료는 네 가지 조망을 모두 채택하는 경향이 있었다. 이 결과를 기초로 다면평가의 연구에서 심리측정적 접근뿐만 아니라, 사회-조직적 접근의 필요성을 논의하였다.

주요어 : 다면평가, 평가자 조망, 리더십, 경쟁가치모델, 사회-조직적 접근

* 본 연구는 2003년도 가톨릭대학교 교비연구비의 지원으로 이루어졌음

† 교신저자 : 박 영 석, 가톨릭대학교 심리학과, yspark@catholic.ac.kr

다면평가의 심리측정적 특성

다면평가 혹은 360도 평가는 한 사람의 평가 대상자에 대해 여러 수준에 속해있는 여러 사람이 평가하는 방법이다. 평가하는 절차는 다소 다르더라도 일반적으로 본인을 포함하여 그들과 자주 접촉하는 상사, 동료, 부하들이 평가에 참여한다. 이 평가방법은 전통적인 방법과 달리 상사를 제외하고는 피평가자가 평가자를 확인하기 어려우므로 평가가 보다 객관적일 수 있고, 평가자의 개인차를 최소화 할 수 있으며, 직무의 특성이 다차원적이므로 한 사람의 다양한 측면을 평가할 수 있는 강점이 있다.

다면평가에 대한 연구는 주로 상사, 부하, 동료 등 여러 수준으로부터 제공된 평가의 심리측정적 특성에 주로 관심을 가져왔다(Judge, Scullen, Sytsma, & Hezlett, 1998). 예를 들면, 평가 수준 내 후광효과나(예: Cooper, 1981; Lance & Woehr, 1986; Murphy & Anhalt, 1992) 평가자간 일치도를 검토하는데(예: Harris & Schaubroeck, 1988; Viswevaran, Ones, & Schmidt, 1996) 주로 초점을 두었다.

이러한 효과를 검토하는 대표적 방법이 다중특성-다중평가자(Multitrait-Multirater) 매트릭스를 분석하는 것이다. 이 방법은 피평가자의 여러 측면을 여러 수준의 평가자가 평가하는 것인데, 평정치의 변량이 피평가자의 특성이나 평정자에 기인하는 정도를 파악하는데 유용하다. 일반적으로 피평가자의 동일한 특성에 대해 평가자간 상관성이 높을 때 평가 결과가 피평가자의 특성 변량으로 기인되고(수렴타당도), 동일 평가자가 피평가자의 다른 특성을 평가했을 때 상관성이 높으면 평가결과가 평가자 효과(후광효과)로 기인된다. 따라서 다면 평가의 목적은 평가자 변량이 적고 피평가자 특성변량이 높은 것이 바람직하다.

그러나 전통적으로 평가 수준간의 일치성이 낮은 것으로 보고되고 있다. 초기에 평정척도를 사용한 연구나(예, Springer, 1953; Zedeck, Imparato, Krausz, & Oleno, 1974), 행동기준 평정척도(BARS: Behavior Anchored Rating Scale)를 사용한 연구에서도(Borman, 1974; Zedeck et al., 1974) 상사-부하, 상사-동료 간의 일치성이 낮은 것으로 보고하고 있다. 최근에 Conway (1996)가 수행한 연구에서는 20개의 다중특성-다중평가자 매트릭스를 분석한 결과 평가자 변량이 피평가자 변량보다 상대적으로 높다는 결과를 제시하고 있다.

Tsui와 Ohlott(1988)은 이와 같은 평가자간 불일치가 나타나는 원인이 평정오류(rating errors), 평가자가 접하는 정보의 차이(information differences), 그리고 평가자가 채택하는 준거의 차이(criteria differences)에서 비롯될 수 있다고 추정한 바 있다.

평정오류는 전반적인 평정의 정확성을 떨어뜨리며, 개인마다 다르게 범하는 평정의 오류는 평가자간의 일치성을 낮추는 결과를 초래한다. 예를 들어, 어떤 평가자가 다른 평가자 보다 피평가자를 더 호의적으로 평가하는 관대화 오류를 범하면 평가자간의 상관성이 낮을 수 있다. 또한, 다면평가에서 피평가자의 상사가 부하나 동료보다 후광오류를 일반적으로 더 많이 범하는 경향이 있는데(Landy & Farr, 1980). 후광오류는 피평가자의 특성을 독립적으로 평가하지 못하는 경향이므로, 피평가자의 상사가 동료나 부하보다 이 오류를 더 많이 범하면 평가자간의 상관계수가 낮아진다.

범위축소(range restriction)의 문제에 의해서도 평가자간 일치도가 영향을 받는다. Tsui(1983)에 의하면 피평가자의 상사, 동료그리고 부하 순으로 평정에서 좁은 범위를 사용한다. 평정의 범위가 낮거나 서로 다르면 상관의 크기를 줄이는

경향이 있으므로 평가자간의 불일치를 유발하는 원인이 될 수 있다.

일반적으로 피평가자가 수행하는 업무는 다양한 측면을 지니고 있으므로, 평가자에 따라 피평가자가 수행하는 다양한 활동 중 일부에 친숙할 수 있다. 평가자에 따라 상이한 직무수행 활동을 관찰하고, 저장하고, 회상하는 인지적 처리과정 때문에 동일한 피평가자에 대해 평가자 마다 상이한 평가를 내릴 수 있다. 그러나 Tsui(1984)의 연구에서 평가자들 간에 동일한 피평가자에 대해 관찰한 행동이 큰 차이가 없으며, 선택적 지각의 문제는 평가자간의 일치성을 낮추는 중요한 원인이라고 보기 어렵다고 주장한 바 있다. 이러한 주장은 Harris와 Schaubroeck(1988)의 메타분석에서도 같은 결과를 나타냈다.

평가자간 불일치의 또 다른 원인은 피평가자의 효과성을 평가하는 평가자의 준거의 차이에서 비롯된 것일 수 있다. Biddle(1979)의 역할 이론(role theory)에 따르면 사람들은 특정 개인에 대해 두 가지 기대를 갖는다. 하나는 개인적 기대(personal expectation)이고 다른 하나는 직위 기대(positional expectation)이다. 후자는 어떤 직위를 수행하는 역할에 대한 기대이고, 전자는 역할을 떠나서 한 개인으로서 기대하는 내용이다. 피평가자에 대한 직위기대는 동일한 계층의 평가자들 간에 대체적으로 합의가 있지만, 개인적 기대는 동일한 수준의 평가자이더라도 서로 다른 내용일 수 있다. 또한 평가자 수준에 따라 그 기대내용도 다를 수 있다. 따라서 동일한 피평가자에 대해 평가자가 어떤 수준, 즉 계층에 속해 있는가에 따라 기대의 내용이 다르고 상이한 평가의 준거를 채택할 수 있으므로, 결과적으로 평가자간 불일치가 발생할 수 있다.

본 연구에서는 다면평가에서 평가자간 불일치가 발생하면, 그것의 원인이 앞에서 언급한 평정

오류에 기인하는 것인지, 아니면 평가자 수준에 따라 채택하는 준거나 조망의 차이에 기인하는지를 검토하는데 그 목적이 있다. 선택적 지각에 의해 평가자가 접하는 정보의 차이는 앞에서 언급한 바와 같이 평가자자간의 불일치를 설명하는 충분한 원인이라고 보기 어려우므로(Tsui, 1984; Harris & Schaubroeck, 1988), 본 연구의 관심사에서 제외하였다.

리더십의 평가 준거

리더십을 평가하는 모델은 다양하지만 리더십의 영역을 포괄적으로 다루고 있는 것으로 Quinn(1988)의 경쟁가치모델(Competing Value Model)을 들 수 있다. Quinn(1988)은 최근까지 등장한 여러 경영 이론들을 정리하여, 포괄적인 리더십 모델을 제안하였다. 생산성과 목표달성 행동(이하 '성과지향행동'이라 칭함)을 발휘하는 Market 리더, 대인관계와 팀워크 행동(이하 '관계지향 행동'이라 칭함)을 발휘하는 Clan 리더, 안정과 통제 행동(이하 '안정지향 행동'이라 칭함)을 발휘하는 Hierarchy 리더, 그리고 변화와 혁신을 추구하는 행동(이하 '혁신지향 행동'이라 칭함)을 발휘하는 Adhocracy 리더가 그것이다. 이러한 네 가지 리더 유형에서, Market 리더와 Clan 리더는 성과와 관계라는 가치에서 서로 경쟁적이고, Hierarchy 리더와 Adhocracy 리더는 안정과 혁신이라는 가치에서 서로 대립적이다. Quinn(1988)은 이러한 네 가지 유형의 리더십을 한 개인이 균형있게 발휘하는 리더가 효과적이라고 제안하고 있다.

앞에서 언급한 바와 같이 다면평가에서 평가자의 수준에 따라 피평가자에 대한 평가준거나 조망이 다르다면, 리더십에 대한 다면 평가에서도 상사, 부하 및 동료가 평가대상자인 리더를 평가하는 기준이 다를 것이다. 즉 앞의 네 가지

행동 중 특정 행동에 준거하여 평가할 것이다.

절차

연구 문제

따라서 본 연구에서는 1) 리더십 다면 평가에서 나타나는 심리측정적 특성, 즉, 평가자의 후광효과와 평가자간 불일치를 검토하고, 2) 평가자간 불일치가 나타난다면, 그 원인이 평정 오류 때문인지 아니면 평가자 조망의 차이 때문인지를 다루고, 3) 그 원인이 평가자 조망의 차이에 있다면, 리더십 다면평가에서 평가자 수준에 따라 취하는 조망이 무엇인지를 검토할 것이다.

평가는 on-line을 통해 실시하였다. 평가자에게 E-mail을 보내어 지정된 Site에 접속하여 평가하도록 하였다. 평가자가 Site에 접속하면 피평가자의 이름과 소속을 알려주었고, 지시에 따라 평가하도록 하였다. 한 화면의 모든 문항에 대해 평가하여야 다음 페이지로 넘어갈 수 있으므로 모든 행동을 빠짐없이 평가하도록 하였다.

결 과

리더십 다면평가의 심리측정적 특징

방 법

연구참가자

피평가자는 국내 한 그룹의 여러 회사에 근무하는 경영자 후보(부장) 150명이며, 평가자는 평가대상자 별로 직속상사(임원) 1명, 부하(차장 또는 과장) 10명, 동료(부장) 2명이었다. 따라서 총 1,950명이 평가에 참여하였다.

표 1에 리더 특성에 대한 평가자간의 상관계수를 제시하였다. 리더 특성에 대한 평가자간 상관의 평균은 상사-부하 간에 .12, 상사-동료 간에 .14, 동료-부하 간에 .36이다. 이 결과는 동료-부하 간에는 비교적 평가가 일치하는 경향이 있지만, 상사-부하나 상사-동료간에는 일치도가 매우 낮음을 나타낸다.

도구

평가에 사용된 도구는 Quinn(1988)이 개발한 32 문항의 리더십 평가도구로써, Market, Clan, Hierarchy, 그리고 Adhocracy 리더의 행동을 각각 8문항으로 평가한다. 문항의 의미를 좀 더 명확히 할 필요가 있는 문항은 일부 수정하여 사용하였고, '전혀 발휘하지 않는다(1)'~ '항상 발휘한다(10)'의 11점 척도 상에서 평가되었다. 각 리더 유형을 측정하는 척도의 내적 합치도는 각각 .91, .90, .81, .89이다.

다면평가의 심리측정적 특성에서 수준간 평가의 일치성이 낮은 원인을 검토하기 위하여 먼저 평가자 수준별로 후광오류를 검토하였다. 수준(평가자) 내 리더 특성 간 평균 상관계수는 상사 평가 .82, 부하평가 .92, 동료평가 .80이다. 이 결과는 각 수준의 평가자 내에서 후광효과가 심하게 나타나는 것이다.

수준간 후광효과의 크기를 차이 검증한 결과, 부하평가의 평균상관과 상사평가의 평균상관 간에 유의한 차이가 있고($z = 3.57, p < .05$), 부하평가의 평균 상관과 동료의 평균상관 간에도 유의한 차이가 있으나($z = 4.05, p < .01$), 상사 평가의 평균상관과 동료평가의 평균상관 간에는

표 1. 리더 특성 평가 점수의 평가자간 상관계수(최소사례수 = 139)

	a	b	c	d	e	f	g	h	i	j	k
a. 성과지향행동 상사											
b. 성과지향행동 부하	.11										
c. 성과지향행동 동료	.07	.23*									
d. 관계지향행동 상사	.79*	.14	.19*								
e. 관계지향행동 부하	.00	.91*	.26*	.11							
f. 관계지향행동 동료	.01	.70*	.73*	.19*	.81*						
g. 안정지향행동 상사	.91*	.10	.11	.78*	.00	.03					
h. 안정지향행동 부하	.11	.96*	.18*	.14	.90*	.14	.13				
i. 안정지향행동 동료	.06	.20*	.94*	.17*	.25*	.69*	.10	.17*			
j. 혁신지향행동 상사	.87*	.06	.05	.76*	.00	.02	.82*	.09	.05		
k. 혁신지향행동 부하	.11	.94*	.21*	.15	.88*	.66*	.12	.93*	.18*	.12	
l. 혁신지향행동 동료	.12	.20*	.89*	.26*	.20*	.65*	.17*	.15	.88*	.18*	.23*

* $p < .05$

유의한 차이가 없다($z = 0.48, n. s.$). 이것은 부하평가가 상사평가나 동료평가보다 후광오류를 유의하게 더 많이 범하고 있음을 나타낸다.

표 2는 평가자 수준별로 네 가지 리더행동에 대해 평가한 결과의 기술통계치이고, 표 3은 평가자 수준간의 관대화 효과를 검증하기 위한 평가자 수준간 평균의 차이검증 결과이다.

표 3에 나타난 바와 같이 네 가지 모든 리더 행동에서 상사와 부하는 동료보다 더 관대하게

평가하고, 그 차이가 통계적으로 유의하다. 반면, 상사와 부하 간에는 유의한 차이가 없다. 이 결과는 상사와 부하가 동료보다 관대화 오류를 더 많이 범하거나, 동료가 엄격성 오류를 더 많이 범하는 것으로 해석된다.

이 결과들을 종합하면, 리더십 다면 평가에서 평가자 수준 간에 불일치가 존재하고, 그 원인이 평가자 수준 간의 평정오류의 차이, 즉 후광오류나 관대화 오류의 차이에 의해서도 발생할 수

표 2. 평가자 수준별 리더행동의 평가 기술치

리더행동 평가자	성과지향행동			관계지향행동			안정지향행동			혁신지향행동		
	상사	동료	부하	상사	동료	부하	상사	동료	부하	상사	동료	부하
<i>M</i>	8.28	7.94	8.32	8.08	7.77	7.93	8.11	7.80	8.02	8.04	7.67	8.09
<i>SD</i>	1.09	1.17	1.46	1.07	1.18	1.62	1.04	1.03	1.44	1.19	1.23	1.49
<i>Range</i>	7.13	7.63	7.75	6.88	8.28	7.78	5.63	6.44	7.81	7.13	7.69	7.78
<i>n</i>	146	140	147	146	139	147	146	140	147	146	140	147

표 3. 평가자 수준 간 리더행동 평가 평균의 짝지은 차이검증

리더행동 비교쌍	성과지향행동			관계지향행동			안정지향행동			혁신지향행동		
	상사	동료	부하	상사	동료	부하	상사	동료	부하	상사	동료	부하
	vs. 부하	vs. 상사	vs. 동료	vs. 부하	vs. 상사	vs. 동료	vs. 부하	vs. 상사	vs. 동료	vs. 부하	vs. 상사	vs. 동료
<i>t</i>	0.26	2.66*	3.31*	0.85	2.56*	3.01*	0.61	2.59*	2.09*	0.41	2.67*	3.43*
<i>df</i>	144	138	138	144	137	138	144	138	138	144	138	138

* $p < .05$

있음을 의미한다.

리더십 다면평가에서 나타난 평가자 조망

평가자간 불일치에서 평가자가 취하는 리더십 준거나 조망의 효과를 검토하기 위하여 다음과 같은 절차를 따랐다.

먼저, 평가자가 리더십 평가에서 성과조망과 관계조망을 사용하는지 여부를 판단하기 위하여 다음과 같은 절차를 따랐다. 1) 평가자 수준에 따라 동일한 가중치를 갖도록 부하나 동료의 자료를 각각 하나로 집합화(agggregation)하고, 2) 상사, 부하 및 동료의 평가를 성과지향 행동과 관계지향 행동별로 합하여 그 중앙치로 각각 상하 집단을 나누는 후, 3) 성과지향 행동과 관계지향 행동이 모두 상 집단이면 Hi-Hi 리더, 모두 하 집단이면 Lo-Lo 리더, 성과지향 행동이 상 집단 관계지향 행동이 하 집단이면 Market 리더, 그 반대이면 Clan 리더로 구분한다. 4) 이 네 가지 유형 중에서 Market 리더와 Clan 리더를 평가자 수준에 따라 성과지향 행동과 관계지향 행동 중 어느 것에서 유의미하게 변별하는지를 검증한다. 5) 만약 평가자가 성과조망을 가지고 리더십을 평가한다면, Market 리더가 Clan 리더 보다 과제지향 행동에서 유의하게 더 높게 평가되어야 할

것이다. 그것은 평가자가 두 가지 리더 유형을 과제지향 행동에서 잘 변별하고 있다는 것으로 해석할 수 있기 때문이다. 그러나 그 차이가 유의하지 않다면, 평가자는 두 가지 리더 유형을 과제지향 행동에서 잘 변별하지 못하는 것으로 해석할 수 있다. 6) 마찬가지로, 평가자가 관계조망을 가지고 리더십을 평가한다면, Clan 리더가 Market리더 보다 관계지향 행동에서 유의하게 더 높게 평가되어야 할 것이다.

평가자가 안정조망과 혁신조망을 사용하는 지의 여부도 같은 절차에 따랐다.

표 4는 리더 유형별 성과지향 행동과 관계지향 행동의 다면 평가의 평균 및 차이검증 결과이다. 상사는 성과지향 행동에서, 부하는 관계지향 행동에서, 동료는 두 가지 행동 모두에서 Market 리더와 Clan 리더를 유의하게 구분하지만, 상사는 관계지향 행동에서, 부하는 성과지향 행동에서 두 가지 리더 유형을 유의하게 구분하지 못한다. 특히, 상사는 관계지향 행동에서도 Market 리더를($M = 8.5$) Clan 리더($M = 8.2$)보다 더 높게 평가하는 경향을 보이고(그림 1), 부하는 성과지향 행동에서 Clan 리더($M = 8.7$)를 Market 리더($M = 8.2$) 보다 더 높게 평가하는 경향을 나타낸다(그림 2). 이러한 역전 현상은 상사가 리더십 평가에서 성과조망을, 부하가 관계조망을 매

표 4. 성과지향 행동과 관계지향 행동에서 각 리더에 대한 다면 평가

리더유형		Hi-Hi	Hierarchy	Adhocracy	Lo-Lo	F	사후비교+
평가준거		리더	리더(H)	리더(A)	리더		(H 대 A)
		(n=52)	(n=11)	(n=18)	(n=57)		
성과지향행동	상사	8.7(0.7)	9.1(0.6)	7.8(0.9)	7.7(1.3)	11.00**	p < .00
	부하	9.3(0.5)	8.2(0.7)	8.7(0.6)	7.4(1.7)	23.05**	p < .19
	동료	8.7(0.7)	8.1(0.7)	7.4(0.9)	7.3(1.2)	18.26**	p < .04
관계지향행동	상사	8.5(0.7)	8.5(0.7)	8.2(0.5)	7.4(1.3)	13.33**	p < .20
	부하	9.0(0.6)	7.2(0.8)	9.0(0.5)	6.8(1.7)	37.48**	p < .00
	동료	8.7(0.5)	7.2(0.4)	8.1(0.7)	6.8(1.1)	44.98**	p < .00

**p < .01

우 강하게 취하고 있는 것으로 해석된다.

표 5은 리더 유형별 안정지향 행동과 혁신지향 행동의 다면 평가의 평균 및 차이검증 결과이다. 상사는 혁신지향행동에서, 동료는 안정지향 행동에서 Hierarchy와 Adhocracy 리더를 유의하게 구분한다. 또한 부하는 안정지향 행동에서, 동료는 혁신지향 행동에서 두 가지 리더를 유의하게 구분하는 경향이 있다.

즉, 상사는 혁신조망에서 부하는 안정조망에서, 그리고 동료는 이 두 가지 조망에서 리더십을 평가하거나 그 경향이 강하게 나타난다.

표 4와 표 5의 결과를 종합하면, 리더십 다면 평가에서 상사는 성과와 혁신, 부하는 관계와 안정, 동료는 성과, 관계, 혁신, 그리고 안정 조망을 사용한다고 할 수 있다.

표 5. 안정지향 행동과 혁신지향 행동에서 각 리더에 대한 다면 평가

리더유형		Hi-Hi	Hierarchy	Adhocracy	Lo-Lo	F	사후비교+
평가준거		리더	리더(H)	리더(A)	리더		(H 대 A)
		(n=52)	(n=11)	(n=18)	(n=57)		
안정지향행동	상사	8.7(0.7)	8.4(0.7)	8.0(0.8)	7.5(1.1)	18.75**	p < .19
	부하	8.7(0.8)	8.8(0.7)	8.0(0.8)	7.3(1.7)	12.57**	p < .12
	동료	8.4(0.8)	8.2(0.6)	7.5(0.8)	7.2(1.1)	16.31**	p < .04
혁신지향행동	상사	8.7(0.7)	7.5(1.1)	8.4(0.8)	7.3(1.4)	16.69**	p < .04
	부하	8.9(0.7)	8.3(0.7)	8.5(0.8)	7.2(1.8)	13.45**	p < .65
	동료	8.4(0.8)	7.3(0.9)	8.0(0.7)	6.9(1.4)	18.54**	p < .09

**p < .01

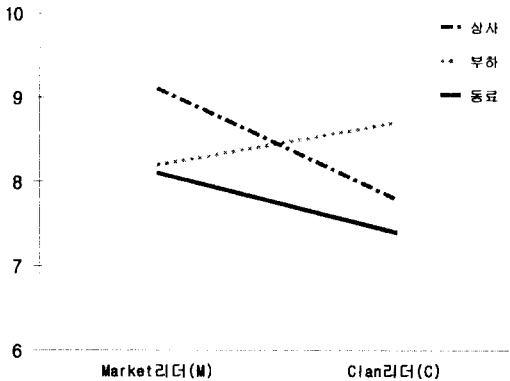


그림 1. 성과지향 행동에서 Market 리더와 Clan 리더에 대한 다면 평가



그림 2. 관계지향 행동에서 Market 리더와 Clan 리더에 대한 다면 평가

논 의

본 논문의 목적은 리더십 다면평가에서 평가자 수준, 즉, 상사, 부하, 그리고 동료에 따라 평가의 불일치가 발생하면 그것의 원인을 탐색하는데 있었다. 특히 본 연구에서는 그러한 불일치가 존재하면, 그것의 원인이 평정오류 뿐만 아니라 평가조망에 의한 것인지를 검토하고, 평

가조망의 효과가 있다면 리더십 다면평가에서 평가자 수준에 따라 취하는 평가조망을 확인하고자 하였다.

본 연구의 결과를 요약하면 다음과 같다. 리더십 다면평가 결과 나타난 심리측정적 특성으로, 1) 평가자 수준 간에 평가의 일치도가 매우 낮았다. 동료 평가와 부하 평가 간에 평가가 일치하는 경향이 있지만, 상사평가와 부하평가, 동료평가와 부하평가 간에는 일치도가 매우 낮았다. 2) 각 평가자 수준 내에서 평가자 특성 간에 후광효과가 매우 높게 나타났다. 특히 부하평가에서 이러한 경향은 더욱 심하게 나타났다. 3) 동료 평가 보다 상사평가와 부하평가에서 관대화 오류가 더 심하게 나타났다.

다면 평가 결과 나타난 평가자 조망에 있어서, 1) 성과준거와 관계준거를 비교했을 때, 상사는 성과 준거에서 피평가자를 평가하는 반면, 부하는 관계준거에서 평가하였으며, 동료는 성과준거와 관계준거를 모두 평가에 사용하였다. 2) 안정준거와 혁신준거에서 비교하였을 때, 상사는 혁신준거에서, 부하는 안정준거에서, 그리고 동료는 두 가지 준거를 모두 사용하여 평가하는 경향이 강하게 나타났다.

이러한 결과를 종합하면, 리더십 다면평가에서 평가자 수준에 따른 평가의 불일치는 평가자 수준별로 평정오류와 평가조망이 다르기 때문이라고 볼 수 있다.

리더십 다면평가의 심리측정적 함의

본 연구의 결과는 전통적인 직속상사에 의한 평가 보다 동료와 부하가 참여하는 다면 평가가 더 객관적으로 리더십을 평가할 수 있음을 시사한다. 본 연구에서 평가자 수준간의 불일치가 크게 나타난 것으로 미루어, 어느 한 수준(일반적

으로 직속상사)에 의한 평가로는 평가대상자인 리더의 행동에 대한 다양한 측면의 정보를 수집하기 어렵다는 것이다. 따라서 리더십의 객관적인 평가를 위해서는 다면평가가 더 효과적이다.

리더십의 평정오류는 일반적인 수행 평가와 마찬가지로 평정척도를 평가의 도구로써 사용하는 한 피하기 어렵다(Landy & Farr, 1983). 평가자 자신이 평가대상자를 평가한 값을 확인할 수 있으므로, 의도적으로 평가대상자를 관대하게 또는 엄격하게 평가할 수 있다. 예를 들어 본 연구에서도 상사와 부하가 동료보다 관대하게 평가하는 경향이 유의하게 높았다. 아마도 동료평가에서 엄격한 평가가 나타난 것은 평가대상자가 그들의 경쟁자이기 때문일 것이다. 이러한 평정오류를 최소화하여 리더십 평가의 객관성을 확보하려면, 평정척도에 의한 방법보다 행동체크리스트(Behavior Checklist)가 유용할 것이다. 이 방법은 평가자에게 리더십 평가차원별로 행동목록을 제시하고, 피평가자가 평소에 발휘하는 가장 대표하는 행동(들)을 선택하도록 하는 것이다. 행동의 선택에 따라 사전에 결정된 척도 값을 부여하여 피평가자에 대한 각 차원의 평가점수를 얻는 방법이다. 이 방법은 평가자가 부여한 평가 값을 자신이 알 수 없으므로, 평정척도를 사용하였을 때 발생하는 오류를 최소화 할 수 있을 것이다.

리더십 다면평가의 사회-조직적 접근의 필요성

본 연구의 결과, 리더십 평가에서 성과준거와 관계준거를 평가자 수준 간에 비교하였을 때, 상사는 성과 준거를, 부하는 관계준거를, 그리고 동료는 두 가지 준거를 모두 사용하였다. 또한 혁신준거와 안정준거를 평가자 수준 간에 비교하였을 때, 상사는 혁신준거를, 부하는 안정준거

를, 그리고 동료는 두 가지 준거를 모두 사용하는 강한 경향을 보였다.

이 결과는 평가자와 피평가자 간의 역할관계에 따라 리더십 평가에서 리더십에 대한 조망을 다르게 갖는 것으로 해석된다.

예를 들어, 피평가자의 상사로서 부하의 리더십을 평가할 때에는 부하가 얼마나 성과지향 행동과 혁신지향 행동을 많이 발휘했는가라는 시각에서 부하들을 주로 변별하지만, 부하로서 상사의 리더십을 평가할 때에는 이와 달리 상사가 얼마나 관계지향 행동과 안정지향 행동을 잘 발휘 했는가라는 시각에서 상사들을 주로 평가한다. 반면, 피평가자의 동료로서 동료의 리더십을 평가할 때에는 이 네 가지 시각을 모두 사용한다.

이러한 결과는 다음과 같이 두 가지로 해석할 수 있다. 하나는 상사, 동료, 그리고 부하가 평정에는 리더십에 대한 네 가지 조망을 모두 가지고 있지만, 리더십을 평가할 때에는 자신과 피평가자와의 관계에 따라 조망이 축소되는 것이라고 볼 수 있다.

평가자가 기대하는 바대로 피평가자가 역할을 수행하면, 피평가자의 행동은 평가자에게 이득이 되고, 기대와 다르게 수행하면 손해가 될 수 있으므로 평가장면에서는 이해관계가 개입될 가능성이 많다. 예를 들어, 단위조직의 책임을 지고 있는 상사는 피평가자인 부하가 성과를 내는데 많이 기여하였을수록 자신에게 유리하고, 그러한 책임이 상대적으로 적은 부하는 상사가 부하의 육성과 구성원간의 관계증진에 많은 노력을 했을수록 자신에게 도움이 된다. 따라서 자신의 이해에 따라 상사는 성과조망에서 리더십을 평가하고, 부하는 관계조망에서 평가하는 조망의 축소 현상이 일어날 수 있다. 반면, 동료는 이러한 역할에 따른 이해관계에서는 벗어나 있고, 오

히려 경쟁자인 동료들 엄격하게 평가해야 하므로 조망이 축소되지 않을 수 있다. 장재윤(1999)의 연구개발관리자에 대한 다면 평가 연구에서, 연구개발관리자가 팀원결속 및 부하 육성에 더 주의를 기울일수록 상사평가에서 높은 점수를 받지 못하는 결과를 관찰한 것도 평가장면에서 평가자의 역할에 따라 이해가 개입된다는 것을 보여주는 것이다.

같은 맥락에서 평가자 수준에 따른 안정 조망과 혁신 조망의 사용도 이해할 수 있을 것이다. 일반적으로 상사는 조직의 경쟁력을 높이기 위해 구성원들에게 도전적인 목표를 설정하고 이를 달성하는데 필요한 부하들의 능력과 태도의 변화를 요구한다. 반면 부하들은 상사들이 요구하는 이러한 변화가 자신에게 안정적인 직장생활을 위협하는 것으로 지각될 수 있다. 이러한 이해관계 때문에 리더십의 다면 평가에서 상사는 혁신조망을 취하고 부하는 안정조망을 취하는 경향이 있으며, 상대적으로 이러한 면에서 피평가자와 이해관계가 없는 동료는 두 가지 조망을 모두 사용할 수 있다.

다른 하나는 경력단계가 높아짐에 따라 역할이 변하고, 그에 따라 리더십에 대한 조망도 변화하기 때문으로 해석할 수 있다. 본 연구의 피평가자의 직위는 부장이며, 상사평가는 상위 직위의 임원, 동료평가는 같은 직위의 부장, 그리고 부하평가는 하위 직위의 차장이나 과장이었다. 차과장의 단계에서는 리더십에 대한 관계조망과 안정조망을 갖고, 부장이 되면 이에 더하여 성과조망과 혁신조망을 갖지만, 임원이 되면 성과조망과 혁신조망의 리더십만을 갖는 것으로 해석할 수도 있다. 부장은 조직 내 '중간 계층인(man in the middle)' 으로서 상하로부터 다양한 기대를 받기 때문에(Roethlisberger, 1945) 임원보다도 더욱 다양한 조망을 가질 수 있다. 요약하면,

평가장면이 아니더라도 조직 내에서 직위와 역할에 따라 리더십에 대한 조망이 다를 수 있다는 것이다.

이러한 두 가지 가능성을 검토하기 위해서는 동일한 평가자가 상사, 부하, 동료의 역할에서 리더를 평가하는 결과를 관찰할 필요가 있을 것이다. 이러한 검토는 실험실 연구에서도 가능할 것이고, 현장에서 여러 계층을 대상으로 다면평가가 시행되면 한 평가자가 자신의 상사, 동료, 부하를 평가해야 하므로 이를 검토할 수 있을 것이다. 또한 종단적 연구를 통해 한 개인의 직위나 역할이 변하면 리더십에 대한 조망이 변화하는지를 검토할 필요가 있다.

최근에 윤소연과 김범준(2003)은 일련의 연구를 통해 타인들이 평가한 다면평가의 정보가 평가자의 평정과 평정의 수정에 미치는 효과를 다루었다. 이 연구에서 자기평가나 상사평가보다 동료나 부하평가의 정보가 평가자의 평정과 평정의 수정에 더 많은 영향을 미치는 것으로 나타났다. 이 연구는 다양한 정보를 한 개인이 통합하는 과정을 다룬 것으로 사회심리학의 핵심 분야 중 하나인 대인지각의 원리를 다면평가에 적용하는 것이다.

산업 및 조직 심리학에서 다면평가에 대한 연구는 주로 평가의 심리 측정적 속성을 다루어 왔지만, 이들의 연구나 본 연구는 다면평가에서 사회 심리학이나 조직 심리학적인 접근의 필요성을 시사하는 것이다. 평가자의 역할과 조망, 평가자와 피평가자의 관계가 다면평가에 미치는 효과, 다면평가 정보의 수용에 영향을 미치는 조직 문화 및 개인의 특성, 다면평가와 절차정의, 다면평가 정보의 통합과정 등과 같이 다면평가에서 조직 및 사회 심리학적 접근이 필요하다.

끝으로 본 연구의 한계점은 리더유형을 다면평가의 총점으로 구분하였다는 것이다. 본 연구

에서 검토한 네 가지 리더 유형을 다른 방법의 의해 구분할 수 있다면 조망의 차이를 더 명확하게 검증할 수 있을 것이다. 예를 들어, Zaccaro, Mumford, Connelly, Marks, & Gilbert(2000)가 사용한 Scenario- Response 방법에 따라 리더의 능력을 측정하여 리더의 유형을 구분하고, 다면 평가에서 평가된 결과를 기초로 평가자 수준에 따라 사용하는 리더십 평가조망을 검토할 수 있을 것이다. 그러나 현재 이 방법으로 리더의 유형을 구분할 수 있는 도구가 개발되어 있지 않아 이러한 시도를 할 수 없었다. 그러나 본 연구에서 다면 평가의 총점으로 리더의 유형을 구분하였을 때에도, 평가자 수준에 따라 리더 유형을 유의하게 구분하는 준거가 다르다는 결과를 관찰하였다. 이것은 리더십 평가에서 피평가자와 평가자의 관계에 따라, 리더십 다면평가에서 평가자가 취하는 조망이 다르다는 것을 시사하는 것이다.

참고문헌

- 윤소연, 김범준 (2003). 다면평가 정보가 수행평정 및 평정의 수정에 미치는 효과. *한국심리학회지: 산업 조직*, 16(2), 153-180.
- 장재운 (1999). 연구개발 관리자에 대한 다면평가의 일치도 및 수행과의 관계. *한국심리학회지: 산업 및 조직*, 12(1), 31-53.
- Biddle, B. J. (1979). *Role theory: Expectations, identities, and behaviors*. New York: Academic Press.
- Borman, W. C. (1974). The ratings of individuals in organizations: An alternative approach. *Organizational Behavior and Human Performance*, 12, 105-124.
- Conway, J. M. (1996). Analysis and design of Multitrait-multirater performance appraisal studies. *Journal of Management*, 22, 139-164.
- Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin*, 90, 218-244.
- Harris, M. M., & Schaubroeck, J. (1988). A meta analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater and level effects in 360-degree performance ratings. *Personnel Psychology*, 51, 557-572.
- Lance, C. E., & Woehr, D. J. (1986). Statistical control of halo: Clarification from two cognitive models of the performance appraisal process. *Journal of Applied Psychology*, 71, 679-685.
- Landy, F. J., & Farr, J. L. (1983). Performance rating. *Psychological Bulletin*, 87, 72-107.
- Landy, F. J., & Farr, J. L. (1983). *The Measurement of work performance: Methods, Theory, and Applications*. New York: Academic Press.
- Murphy, K. R., & Anhalt, R. L. (1992). Is halo error a property of the rater, ratees, or the specific behavior observed? *Journal of Applied Psychology*, 77, 494-500.
- Quinn, R. (1988). *Beyond rational management: Mastering the paradox and competing demands of high performance*. San Francisco: Jossey-Bass Publishers.
- Roethlisberger, F. J. (1945). The foreman: Master and victim of double talk. *Harvard Business Review*, 23, 238-298.
- Springer, D. (1953). Ratings of candidates for promotion by co-worker and supervisor. *Journal of Applied Psychology*, 37, 347-351.

- Tsui, A. S. (1983). The quality of judgemental ratings by four rater sources. Presented at the 91st annual conference of the American Psychological Association, Anaheim, CA.
- Tsui, A. S. (1984). A role-set analysis of managerial reputation. *Organizational Behavior and Human Performance*, 34, 64-96.
- Tsui, A. S., & Ohlott, P. (1988). Multiple assessment of managerial effectiveness: Interrater agreement and consensus in effectiveness models. *Personnel Psychology*, 41, 779-803.
- Viswevaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557-574.
- Zaccaro, S. J., Mumford, M. D., Connelly, M. S., Marks, M. A., & Gilbert, J. A. (2000). Assessment of leader problem-solving capabilities. *Leadership Quarterly*, 11(1), 37-64.
- Zedeck, S., Imparato, N., Krausz, M., & Oleno, T. (1974). Development of behavior anchored rating scales as a function of organizational level. *Journal of Applied Psychology*, 59, 249-252.
- 1 차원고접수 : 2003. 8. 28.
수정원고접수 : 2003. 10. 23.
최종게재결정 : 2003. 11. 22.

Rater Perspectives in Multi-Rater Leadership Ratings

Park, Young-Seok

The Catholic University of Korea

The purpose of this study was to investigate the rater levels(supervisor, peers, and subordinates) perspectives in leadership ratings. Ratees were 150 executive candidates, and raters were 1 supervisor, 2 peers, and 10 subordinates of each ratee. Ratee's leadership was evaluated by four styles of leadership from the competing values model(Quinn, 1988): Market, clan, hierarchy, and adhocracy. Results revealed low correlations between rater levels, and the causes of the low correlations were both the different magnitude of the rating errors between rater levels and their different perspectives. Supervisors adopted achievement and innovation, subordinates adopted relation and stability, and peers adopted all of the perspectives. The need for social-organizational approaches in multi-rater leadership ratings were discussed.

key words : Multi-rater Ratings, Leadership perspectives, Competing values model