

다면평가의 구체적 내용으로 관리자 능력에 대한 본인, 상사, 부하평가의 특성과 그들간의 관계*

박 원 우**

〈目 次〉

- | | |
|-----------------------------------|--------------------------------------|
| I. 다면평가의 제 이슈 | 1. 참여자 |
| 1. 다면평가의 의미와 구성요소 | 2. 측정내용과 방법 |
| 2. 다면평가의 배경과 목적 | 3. 연구절차 |
| 3. 다면평가의 핵심이슈: 평가
주체별 평가결과의 차이 | IV. 조사결과 |
| II. 연구목적과 문제 | 1. 평가주체별 평가특성의 비교 |
| 1. 연구목적 | 2. 과대/과소평가의 가능성 분석 |
| 2. 연구문제 | 3. 평가요소의 동질/이질성에 따
른 동등화/차별화 경향분석 |
| III. 연구방법 | V. 토의 및 결론 |

인사관리에 있어서 평가의 중요성은 우리 모두가 다 아는 사실이다. 그 어떠한 이름과 내용으로 임금, 승진, 육성 등의 인사제도가 바뀐다/개선된다 하여도 인사관리의 성패는 정확하고 수용성 있는 평가제도의 존재를 전제로 하기 때문이다. 평가의 정확도와 수용성을 높이기 위한 노력이 지금까지 끊임없이 이어져 왔고, 그 내용은 평가내용(요소)의 변화, 평가/고과자의 훈련, 평가양식과 절차의 변화, 평가대상(개인과 집단)의 확산 등 여러 측면에서 나타났다. 이런 노력의 일환으로 근자에는 평가주체에 따른 평가방식의 다양화가 추구하고 있다. 거의 대부분의 조직에서는 전통적으로 상사에 의한 평가가 주를 이루어 왔다. 그런데 사회의 변화와 더불어 인사철학 뿐만 아니라 조직과 업무구조가 크게 바뀌다 보니 많은 조직에서 전통적 상사평가에서 벗어나 동료평가, 팀평가, 본인평가, 부하평가, 고객평가, 혹은 360°의 다면평가를 시도하고 있다.

본 논문에서는 다면평가의 제 이슈를 살펴본 후 다면평가의 구체적 내용으로 동일 피평가자를 대상으로한 상사, 본인, 부하평가의 특성과 그들간의 관계를 실증적으로 연구하여 국내

* 본 연구는 서울대학교 경영대학 경영연구소의 연구비 지원에 의해 수행되었음.

** 서울대학교 경영대학 기금조교수

의 경우 확산의 초기단계에 있는 다면평가의 효과적 실시에 보탬이 되고자 한다.

I. 다면평가의 제 이슈

1. 다면평가의 의미와 구성요소

1) 다면평가의 의미와 활용정도

국내에서는 다면평가 혹은 360° 평가로 불리지만 국외에서는 영어의 경우 *multirater feedback*, *360-degree feedback*¹⁾, *multisource assessment*, *multi-perspective ratings*, *full-circle feedback* 등으로 다양하게 표현되는 다면평가는 국외는 물론 국내조직에서도 급속히 확산되고 있는 대인평가방법이다.

다면평가는 한 평가대상자를 두고 다양한 수준과 측면에서 다양한 평가자가 평가하여 그 결과를 결합하는 평가방법을 일컫는다. 비록 평가절차는 조직마다 다르지만 대체로 피평가자를 잘 알고 또 그 평가결과가 피평가자에 의해 가치 있는 것으로 받아들여질 수 있는 사람에 의해 평가가 이루어진다. 보통 본인을 포함하여, 상사, 동료, 부하가 평가하지만 때로는 공급자나 고객과 같은 외부인에 의해서도 이루어진다.

Moses, Hollenbeck, 및 Sorcher(1993)는 "360° 평가"라는 표현은 내용상 결코 새로운 것이 없음에도 불구하고 마치 뭔가 새로운 혹은 신비한 발견이 있는 것처럼 오도하는 문제를 지니고 있다고 한다. 그들은 360° 혹은 다면평가는 내용상 지금까지 조직내에서 항상 존재해 왔으며 단지 근자에 새로운 이름이 붙여졌을 뿐이라고 하면서, 이는 새 병에 있는 오래된 술(*old wine in new bottle*)과 같다고 한다. 새 술이든 헌 술이든 아무튼 병이 새것임에는 틀림없고 조직내에서 이의 활용은 날로 늘고 있다.

다면평가는 Fortune 500의 조직에서는 이미 일반화된 평가방법인데, 조사(Antonioni, 1996)에 의하면 미국기업의 12% 정도가 완전한 360° 평가를, 25%가 상향식 평가를, 그리고 18%가 동료평가를 쓰는 것으로 나타났다. 같은 1996년에 *Compensation & Benefits Review*의 조사(Bohl, 1996)에서는 미국 기업 중 13%가 완전한 360° 평가를, 9%가 부분적 360° 평가를 쓰는 것으로 나타나기도 했다. 이러한 비율을 현재 계속 늘고 있는데, 다면평가를 쓰는 회사의 예로는 Amoco, AT&T, Caterpillar, Chrysler, Du Pont, Exxon, GE, GTE, IBM, Mass Mutual Insurance, Motorola, PG, Syntex, TRW,

1) 360° 평가를 영문으로 표현할 때는 *360-degree appraisal*이라고 하기보다는 *360-degree feedback*이라는 표현을 훨씬 많이 쓴다.

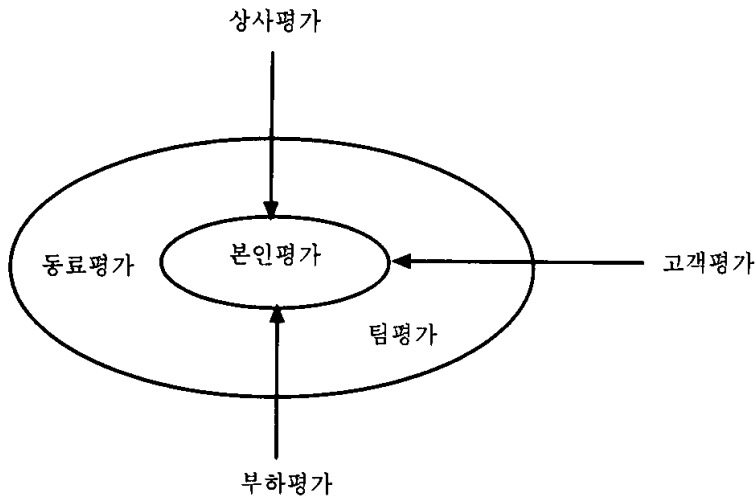
UPS, US West, Walt Disney Company, Westinghouse 등이 있다. 국내에서도 이미 많은 회사가 (예: 대한생명보험, 동양 SHL, 삼성물산, 삼성전관, 삼성화재, 성광전자, 신세계백화점, 하나은행, 한국IBM, 현대정공, LG전자, LG정보통신, SECL) 이 제도를 활용하고 있고, 근자에는 많은 금융기관과 정부의 공조직에서도 활용시도가 늘고 있다.

2) 다면평가의 구성요소

다면평가란 평가주체가 다양하다는 특성을 지니고 있는데 평가주체에 따라 평가유형을 구분하여 그 구성요소를 살펴보면 다음의 <그림 1>과 같다.

어느 평가유형이 가장 좋으나에 대한 정확한 답은 아무도 모른다. 하지만 더 좋은 즉, 타당성 있는 평가유형의 조건을 생각해볼 수 있다. 예로서 Milkovich와 Boudreau(1994)는 평가자가 세 가지 특성 즉, 관찰기회, 평가능력, 평가의욕을 지녔을 때 그 평가의 결과가 타당성이 있을 것이라고 했다.

그런데 사회와 조직의 변화와 더불어 이들 세 가지 기준을 바탕으로한 평가주체의 상대적 우위성 판단결과도 변하고 있다. 예로서 관찰기회를 보기로 하자. 전통적으로 상사에 의한 피평가자 관찰기회가 가장 많다고 생각해왔다. 그 결과 상사평가²⁾가 인사평가의 주종을 이



<그림 1> 평가주체에 따른 구분한 다면평가의 구성요소

2) '상사평가'란 용어는 '상사에 의한 평가'란 뜻도 되고 동시에 '상사를 평가하는 것'이란 뜻도 될 수 있다. 그러나 본 논문에서는 평가주체에 따른 용어구분이기엔 전자를 의미한다. 이러한 해석은 나머지 용어 (예, 부하평가)에서도 같이 적용된다.

루어왔다. 물론 아직도 여전히 상사가 부하를 관찰할 기회를 많이 지닌 것은 사실이지만, 조직 및 직무설계의 변화와 더불어 상황이 바뀌고 있는 경우도 있다. 예를 들자면, 요즘 국내 기업에서도 팀제로의 구조변화가 많아지면서 팀웍이 강조되고 있는데, 팀제 하에서는 팀장보다는 팀원간의 평가가 중시되기도 한다.

또 다른 예는, 저층구조화(flat화)되면서, 한 명의 상사가 관찰해야 하는 부하의 수가 대폭 늘어나는 경우이다. 이때는 한 명의 상사가 많은 수의 부하를 관찰하기보다는, 역으로 여러 명의 부하가 한 명의 상사를 관찰할 기회가 많아진다. 따라서 부하에 의한 상급자평가가 상사에 의한 하급자평가를 대신하기도 한다. 그럼 여기서 각 평가주체별 특성을 간단히 보기로 하자.

① 동료 혹은 팀평가

동료평가(peer 내지는 coworker appraisal)란 같이 일하는 수평적 관계의 동료가 평가하는 것으로서, 평가주체가 단수일 수도 있고 복수일 수도 있다. 그런데 만약 복수인이고 그 주체가 동료 팀원들이라면 팀평가가 되는 것이다.

Digital Equipment Corporation에서는 모든 팀원들이 서로를 평가하는데(Norman & Zawacki, 1991), 매일 같이 일하는 동료가 가장 정확히 평가할 수 있다는 가정에 근거하고 있다. Quaker Oats Company는 20년 이상 팀제를 써왔는데, Quaker Oats의 애완동물 사료공장(pet food plant)에서는 평가의 100%를 동료평가에 의존하고 있다. 그 결과 모든 승진과 임금관련 의사결정이 팀내에서 이루어지고 있다.

이러한 동료평가의 장점은 서로간에 건설적 견해를 제공하는 경향이 있어 결과적으로 개인, 나아가 집단의 개선이 이룩된다는 것으로서, 이때 서로 제공하는 의견이 직무와 관련하여 매우 구체적이기 때문에 실질적 도움이 된다고 한다(Ramsey & Lehto, 1994).

나아가 동료평가는 높은 정확성 때문에도 각광받는데, 군대조직내 장교들간 평가의 경우(Downey, Medland, & Yates, 1976) 뿐만 아니라, 일반 사기업의 연구(Kraut, 1975; McEvoy & Buller, 1987)에서도 승진자 예측의 경우 매우 높은 정확도를 보였다.

② 본인평가

본인평가(self appraisal)는 자신을 되돌아보고, 자신의 강·약점/장·단점을 도출해 내는 계기가 된다고 한다. 하지만 문제는 사람은 대체로 자기 자신을 높이 평가하는 현상(leniency effect)이 있고, 또 역으로 자신을 낮추는 현상(modesty effect)도 있는 등 본인평가의 결과가 타인평가의 결과와 차이를 보인다는 점이다. 따라서 자기평가는 관리상의 의사결정 목적보다는 개발지향적 목적에 적합한 것으로 취급되고 있다. 결국 본인평가는 타 평가방법과

결합하여 사용될 때 비로소 효과 있는 것이라고 하겠다.

③ 부하평가

Subordinate 내지는 upward appraisal 혹은 reverse review라고 불리는 부하평가는 구미의 경우 관조직에서 많이 쓰여왔다. 근자에는 사기업체(예: Amoco, AT&T, Chrysler, Du Pont, IBM, Johnson & Johnson, Xerox 등)에서도 많이 활용되고 있는데, 이들 중 IBM의 경우엔 이미 20여년간 이 방법이 활용되어 왔다. 그런데 아직까지는 관리자들에 그들에 대한 부하의 생각을 알려주기 위한 보조수단으로 쓰는 경향이 강하다. 일본에서도 관리직의 활성화 차원에서 일본유지, 후지사, 동경가스 등에서 처우개선보다는 관리자의 능력향상을 목적으로 실시하고 있다.

부하평가는 영역(내용)에 따라 평가의 정확도가 달라진다. 정확성이 높은 분야는 상사의 리더십, 의사소통, 권한위임, 팀조정, 부하에 대한 관심·배려 등이지만, 반대로 정확성이 떨어지는 분야는 상사의 계획력, 조직력, 예산입안력, 창의력, 분석력 등이다(Sherman, Bohlander, & Snell, 1996). 따라서 부하평가는 필요한 경우에 제한해서 써야하는데, 그 방법은 무기명이며, 복수부하의 평가내용을 결합시킬 때 비로소 의미가 있다.

④ 고객평가

고객평가(customer appraisal)를 조직내 타 부서원에 의한 내부고객평가(internal customer appraisal)와 조직 외부인에 의한 외부고객평가(external customer appraisal)로 구분하는 경우(Waldman & Bowen, 1998)도 있으나 일반적으로는 외부인에 의한 평가를 고객평가라고 부른다. 360° 평가에서 고객평가의 중요성을 강조하여 London과 Beatty(1993)는 고객평가가 (그것이 내부고객평가든 외부고객평가든 관계없이) 빠진다면 평가는 270° 평가에 불과하다고 주장하기도 하였다. 외부고객은 보통 제품과 서비스를 사용·활용하는 소비자를 뜻하지만 외부인에 의한 평가(outsider ratings)를 의미하기도 한다. 실제로 Furnham과 Stringfield의 연구(1998)에서는 컨설턴트에 의한 평가를 외부인평가로 칭하고 있다. 정리하면, 고객평가는 극단적으로는 내부고객이라는 측면에서 동료평가까지도 포함할 수 있는 표현이기에 보통은 고객평가를 외부인 평가로 인식하는 경우가 많다.

2. 다면평가의 배경과 목적

1) 다면평가의 배경

다면평가의 배경으로 과거의 학문적 등장배경과 현재의 실무적 사용배경을 구분하여 생각해 볼 수 있다. 즉, 어떠한 학문적 배경을 가지고 등장하였으며 현실에서 많이 쓰이는 이유

는 무엇인가 하는 점이다.

360-degree feedback이라고 많이 불리는 다면평가의 등장배경은 1950년대와 60년의 인간관계운동(human relations movement)으로 까지 이어진다(Waldman, Atwater, & Antonioni, 1998). 이 당시의 조직은 훗날 조직개발(OD)이라고 알려진 것처럼 각종 방법을 통해 조직과정과 의사소통의 증진을 추구했었고, 이때 나온 조직개발기법 중 많이 쓰인 것이 survey feedback이다. Survey feedback은 직무, 복리후생, 급여, 혹은 조직내 의사소통과 같은 요인에 대하여 종업원 일반으로부터의 의견조사를 바탕으로 한다. 전통적 survey feedback이 조직의 전반적 과정에 관련되어 다수에 관한 정보를 다수의 구성원에게 동시에 제공하는데 비해 다면평가 즉, 360-degree feedback은 특정개인에 관한 정보를 특정인(들)에게 제공하는 차이가 있다.

실무에서 다면평가의 사용이 확산되는 배경에는 실무자들이 믿는 두가지 가정이 있다. 하나는 익명으로 여러 사람으로부터 도출·집계된 관찰결과가 알려진 한 개인에 의해 이루어진 것 보다 더 타당성과 신뢰성이 높을 것이고, 따라서 더 의미 있고 사용가치가 높을 것이라는 믿음이다. 또 하나는 자신에 대한 이해가 증진될 수록 행동변화가 잘 일어난다는 점이다(Church & Bracken, 1997). 물론 이러한 가정들이 대체로 옳으니까 다면평가가 확산되지만, 보다 현실적이고 실무적인 배경을 살펴보면 다음과 같다(Nowack, 1993).

- 평가센터(assessment center)보다는 비용이 적게 들기 때문.
- 다면평가를 위한 실무 software가 많이 제공(출시)되기 때문.
- 계속적 개선노력이 계속적 측정을 필요로 하기 때문.
- 경력단계상 고원에 이르러(career plateauing) 더 이상의 발전이 없는 구성원에게 직무관련 피드백을 부여할 필요가 있기 때문.
- 기술변화, 경쟁적 도전, 작업집단의 다양성 증대에 직면한 구성원의 잠재적 극대화가 필요하기 때문.

2) 다면평가의 사용 목적

다면평가의 목적으로 연구자들(예, Nowack, 1993; Church & Bracken, 1997)에 의해 제시되는 내용을 모아보면 다음과 같다.

- 감독자 훈련과 관리자 개발
- 개인의 스타일과 리더십 인식
- 경력개발
- 승계계획입안
- 선발
- 문화측정
- 조직변화유발
- 니즈파악
- 고객관련 행동과 인식의 측정
- 성과평가와 성과관리시스템에 활용

위의 내용들을 정리하면 다면평가는 크게 개발(development)목적과 평가(appraisal)목적으로 대별된다. 이때 후자인 '평가목적'은 '성과관리목적'이라고 표현하는 것이 보다 이해에 도움될 것이다. 그런데 이들 두 목적 중 전자가 후자보다 훨씬 큰 비중을 차지한다. 즉, 다면평가는 평가를 바탕으로 급여나 승진과 같은 관리상의 의사결정에 쓰이기보다는 평가대상자 개인의 개발과 성장을 증진코자 쓰이는 경우가 대부분이다. 국내에서는 '다면평가' 혹은 '360° 평가'라고 호칭되어 평가의 기법으로서 강조되지만 영어로는 'appraisal'이나 'ratings'라는 표현보다는 'feedback'이라는 표현이 훨씬 많이 쓰이는 데서 현실의 주된 용도를 즉, 다면평가가 평가의 목적보다는 개발의 목적에 치중함을 이해할 수 있다.

개발과 평가(성과관리)라는 두 목적은 구체적 실행방법으로 들어가면 너무나 차이가 많다. 따라서 한 목적으로 쓰던 조직이 다른 목적으로 용도변경을 추구할 때는 많은 변화가 수반되어야 한다. 기존의 방법과 절차를 그냥 둔 채 목적만 바꾸어서는 성공하기가 매우 힘들다는 말이다. 개발과 평가목적은 실행방법상 크게 두 가지 면에서 차이가 난다(Lepsinger & Lucia, 1997).

첫 번째 차이는 누가 평가결과를 소유하느냐는 점이다. 개발목적의 경우에는 피평가자가 소유하게 되고, 누가 그 결과치를 볼 수 있는가, 또 그 결과가 어떻게 활용되는가는 전적으로 피평가자에 의해 결정된다. 이에 반해 다면평가가 평가(성과관리)의 목적으로 쓰일 경우는 평가결과를 조직이 소유하고, 조직의 관리방식에 의거해 조직의 판단에 따라 쓰이게 된다. 주위에서 나를 어떻게 생각하는지 나만이 비공식적으로 알고 있는 경우와, 주위의 그러한 생각이 회사의 공식서류상에 존재하여 나의 급여나 승진결정에 영향을 미치는 것은 매우

성격이 다르다. 물론 평가결과가 누구의 소유가 되고 누구에 의해 무슨 목적으로 쓰이느냐에 따라 애초 평가내용도 다르고(개발목적의 경우에는 피평가자의 행동에, 평가목적의 경우에는 업적/성과에 초점을 많이 둔다), 또 실령 같은 평가내용의 경우에도 평가자들의 평가치가 달라질 수가 있다.

두 번째 차이는 피평가자가 어떤 입장에 놓이게 되느냐는 점이다. 개발목적의 경우엔 강제성이 없다. 평가결과를 보고 피평가자는 임의의 자기개발목적을 잡을 수 있고, 실령 목적달성이 이루어지지 않아도 제재가 없다. 하지만 평가(성과관리)목적의 경우엔 강제성이 있고 급여나 승진과 연계되기에 훨씬 피평가자에게 부담을 주게된다. 따라서 다면평가가 평가(성과관리)의 목적으로 쓰이는 경우엔 조직뿐만이 아니라 개인에게도 위협부담이 있으므로 철저한 준비와 관리가 있어야 한다. Coates(1998)는 다면평가가 성과관리와 연계되었을 때 자칫 잘못하면 오히려 조직내 불신과 갈등을 조장하는 등 문제가 많으므로 다면평가를 특히 급여와는 연계시키지 말 것을 주장했다.

개발이든 평가(성과관리)든 둘 다 긍정적인 목적이다. 흔히 목적을 언급할 때는 좋은 것만을 목적으로 열거하는 규범적 측면이 강한데, 현실에서 다면평가가 확산될 때는 좋은 의도만 있는 것은 아니다. 나쁜 의도라고 까지 말할 수는 없으나 긍정적이지 않은 현실적 측면의 이유도 있다. 대표적인 것이 모방과 이미지관리다(Waldman et al., 1998). 제도화 이론에서 조직은 조직간 네트워크에서의 경쟁사나 타사를 모방한다고 하듯이, 주위의 수많은 기업이 다면평가를 실시하니까 뚜렷한 목적의식 없이 동조화현상을 보이는 조직이 많다. 그런 행동이 마치 변화하는 환경에 대처·적응하는 듯한 기분을 들게 하기 때문일 것이다. 두 번째는 다면평가의 도입을 통해 개방성과 고객지향성을 지닌 조직이라는 인상 혹은 이미지를 형성코자 실시하는 경우이다. 여기서의 문제는 이들 조직의 실제 조직문화는 그렇지 않으면서 단지 그러한 인상을 형성하고자 다면평가가 쓰인다는 점이다.

3. 다면평가의 핵심이슈: 평가주체별 평가결과의 차이

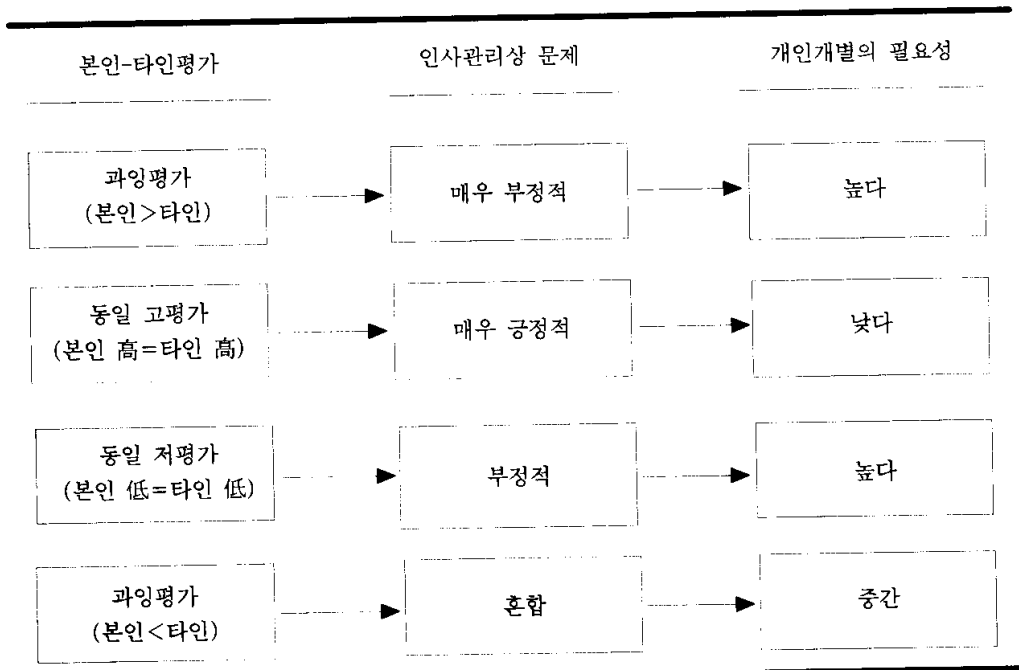
다면평가에서 가장 문제되는 것이 정확성과 수용성이다. 이들 두 가지 중에 어느 하나라도 문제가 있으면 다면평가는 성공하기 어렵다. 다면평가 결과의 정확성과 수용성이 문제시되는 것은 흔히 평가주체별로 평가결과가 다른 데에 기인한다. 피평가자 본인에 의한 평가결과와 타인에 의한 평가결과가 다른 경우가 많고, 그러다 보니 피평가자에 의해 타인평가의 결과가 수용되지 않는 문제가 발생한다. 평가주체별 평가결과의 차이에 관해서는 여러 이슈가 있다. 여기서는 우선 평가주체가 다름에 따라 평가결과가 어떻게 차이 나는지 그 현상(유형)을 보

고, 왜 그런 차이가 나는지, 나아가 어떤 요인들에 의해 평가주체별 평가결과의 차이가 영향을 받는지 알아본다. 그 후 과연 이러한 차이를 줄이는데 초점을 두어야 할지 아니면 그 차이를 인정하고 거기에서 의미를 도출해야할지 논의한다.

1) 본인-타인평가 결과차이의 유형

'평가주체별 평가결과의 차이'라고 할 때는 타인평가(예, 동료, 상사, 부하, 고객평가)간 차이도 포함된다. 물론 본인-타인평가 결과차이 만큼 타인-타인평가 차이도 중요한 이슈다. 하지만 한편으로 경우의 수를 모두 따지는 것이 번거롭고 또 한편으로 피평가자의 개발이라는 목적을 위해 피평가자에게 평가결과를 피드백 해준다는 과정에서 본인-타인평가 차이가 상대적으로 더 중요한 내용이므로 여기서는 본인-타인평가의 차이에 초점을 둔다.

본인-타인평가의 결과차이의 유형은 4가지로 구분된다(Yammarino & Atwater, 1997). 이를 도식화하면 다음의 <그림 2>와 같은데, 과연 각 유형이 인사관리 측면에서 어떻게 문제되고 어떤 유형에서 개인개발의 필요가 많고 적은지를 보기로 하자.



자료: Yammarino & Atwater(1997, p. 37)의 내용을 수정했음

<그림 2> 본인-타인평가 결과차이의 유형별 특성

기존의 연구에서는 대체로 본인평가의 결과가 타인평가의 결과보다 높게 나온다고 한다. 이러한 현상을 자신에 대한 관대화 효과(leniency effect)라고 부르는데, 이는 인사관리상 부정적인 현상으로 이런 사람은 개발될 필요가 높다. 간혹 역으로 겸손하여 오히려 남보다도 자신을 스스로 낮게 평가하는 겸양효과(modesty effect)도 있는 바, 이럴 경우엔 인사관리상 긍정적인 면과 부정적인 면이 공존하며 이런 사람 또한 개발될 필요가 약간 있다.

비록 본인평가와 타인평가의 결과가 동일하더라도 내용상 다시 두 가지 유형이 있는바, 하나는 양쪽 모두 피평가자를 높게(긍정적으로) 평가하는 경우이고, 또 하나는 자타 모두가 낮게(부정적으로) 평가하는 경우이다. 전자의 경우엔 인사관리상 아무 문제가 없고 피평가자는 이미 주관·객관적 측면 모두에서 높게 평가되었기에 달리 개발되어야 할 필요가 적다. 하지만 후자의 경우엔 피평가자가 주관적으로나 객관적으로 모두 부정적으로 평가된 사람이기에 문제가 있는 사람이며, 따라서 개발될 필요성도 높다.

2) 평가주체별 평가결과의 차이발생 이유

동일한 피평가자를 두고 평가자에 따라 평가결과가 달라지는 이유는 세 가지 즉, 이기적 편견(egocentric bias), 조직내 수준차이(differences in organizational level), 관찰기회(observational opportunities)를 바탕으로 설명할 수 있다(Harris & Schaubroeck, 1988).

① 이기적 편견

이기적 편견을 설명하는 방법이 여러 가지 있는데, 대체로 본인평가가 타인평가에 비하여 어느 쪽으로 치우치는 경향을 지닌다는 것이다. 가장 대표적인 설명은 자기방어심리가 발동하여 본인자신에게 더 높은 점수를 부여하게 된다는 점이다(Holzbach, 1978; Steel & Ovalle, 1984). 흔히 이를 관대화 오류(leniency error)라고 하는데, 이러한 현상이 발생 되면 본인평가 결과치의(분포상) 범위가 줄어들게 되고(restriction of range), 그 결과 본인평가치와 타인평가치간의 상관계수는 낮아지게 된다. 이때 타인평가치들간(예: 동료-상사평가, 상사-부하평가)의 상관계수는(범위의 제약이 없으니까) 본인-타인평가치의 경우보다 더 높게 나타난다. 이처럼 방어심리의 발동으로 영향이 한 쪽(점수를 높이는 쪽)으로 미치는 경우에는 범위제약의 수정(correction of range restriction)을 통해 편견의 문제를 어느 정도 해소할 수도 있다.

두 번째 설명은 조절변수의 영향으로 본인평가지 점수의 치우치는 현상이 양쪽(높게되는 쪽과 낮게 되는 쪽)으로 나타난다는 것이다. 예로서 자긍심이 조절작용을 하는 경우인데

(Baird, 1977; Kay, Meyer & French, 1965), 자긍심이 강한 사람은 본인평가시 점수를 높게 부여하지만, 자긍심이 낮은 사람은 오히려 낮게 부여한다는 것이다. 이런 경우엔 범위제약을 수정할 때 그 조절변수까지 고려해야만 문제가 해결된다.

세 번째는 귀인이론(attribution theory)을 적용하는 설명이다(DeVader, Bateson, & Lord, 1986; Jones & Nisbett, 1972). 본인평가의 경우는 높은 성과를 자기 탓으로 낮은 성과를 환경 탓으로 돌리는 반면, 타인평가의 경우에는 역으로 높은 성과를 환경 탓으로 낮은 성과를 피평가자 탓으로 돌린다는 것이다. 그 결과 본인평가의 결과와 타인평가의 결과간에 상관계수가 낮게 된다는 설명이다. 이러한 경우에는 범위제약의 수정이 문제해결에 도움을 주지 못한다.

② 조직내 수준차이

조직내의 수준차이와 관련해서는 두 가지 설명이 있는데, 하나는 평가자가 조직내 어느 수준에 있는가에 따라 성과를 재는 평가요소에 부여하는 가중치가 다르다는 견해다(예, Klimoski & London, 1974; Zammuto, London, & Rowland, 1982). 이럴 경우 개별 요소에 대한 평가는 유사하게 나와도 전체적 평가결과는 다르게 나온다. 또 다른 설명은 조직내에서의 수준에 따라 평가자의 성과에 대한 정의와 측정 그 자체가 다르다는 견해다(예, Borman, 1974; Landy, Farr, Saal, & Freytag, 1976). 이럴 경우엔 개별요소에 대한 평가치와 전체요소에 대한 평가치 모두가 평가자에 따라 달라진다.

③ 관찰기회

평가주체별 평가결과의 차이를 설명하는 세 번째 내용은 평가주체마다 피평가자를 관찰할 수 있는 기회가 다르다는 점이다. 구체적으로는 동료들이 상사보다는 피평가자를 관찰할 기회가 많고, 상사는 상대적으로 관찰기회가 적어 상사에 의한 평가가 잘못될 수 있다(예, Latham & Wexley, 1982; Lawler, 1967). 따라서 상사-본인평가간 상관계수가 동료-본인평가간 상관계수보다 낮게 나타난다고 한다.

3) 평가결과의 차이를 유발하는 영향요인

평가주체별 평가결과치의 차이는 여러 요인에 의해 유발된다. <표 1>은 Yammarino와 Atwater(1997)가 제시한 것을 중심으로 하여 타연구자들에 의해 제시되는 영향요인을 결합·정리한 것인데, 이들 요인 중 일부를 부연설명하면 다음과 같다.

직무 유형에 따라 평가결과치 차이가 달라진다는 것은 예로서 생산직이나 서비스직은 관리직이나 전문직보다도 직무특성상 정형화된 정도가 강해서 평가주체가 달라도 비슷하게 평가

〈표 1〉 평가주체별 평가결과치 차이 유발요인

구분	영향요인		
인구통계학적 요인	· 성별 · 연령	· 근속년수 · 직위	· 소수민족여부
개인적 특성	· 지능 · 능력 · 성취여부 · 통제위치	· 분석력 · 인지복잡성 · 자이인식 · 자긍심	· 기억력 · 양심 · 대인지향성 · 자기감시
직무경험	· 성공/실패	· 직무피드백	· 직무유형
인지과정		· 인식도표, 믿음, 기대 · 정보와 피드백 사용	
상황	· 압력 · 비교정보 · 평가경험	· 정치적 영향 · 정보의 구체성 · 평가목적	· 친밀감 · 유사성
평가도구	· 평가형태	· 평가척도	

될 가능성이 높다는 것이다(Campbell, Dunnette, Lawler, Weick, 1970; Cascio, 1987). 따라서 평가주체간 평가결과치의 상관관계가 생산직과 서비스직에서 높게 나타난다.

다면평가가 어떤 목적으로 쓰일 것인가에 따라 평가주체별 평가경향이 크게 달라진다. 개발목적일 때는 평가가 비슷하게 나오다가 공식적 성과평가의 목적이라면 이야기가 달라지는데, 성과평가의 경우 친구들은 높게 점수를 주는 반면 경쟁자들은 냉담하게 반응하는 것으로 나타났다(O'Reilly, 1994). Farh와 Werbel(1986)연구에 의하면, 학생을 대상으로한 실험의 경우에서, 평가의 목적이 참여도 조사에서 학점산정용으로 바뀌자 평가의 정확도가 매우 떨어지는 것으로 나타났다. 이처럼 다면평가의 결과는 평가목적에 의해 크게 영향을 받으므로, 다면평가를 실시하려는 조직은 다면평가를 통해 얻으려는 목적을 분명히 하고, 그러한 목적을 이룩하기에 적합한 과정/절차를 형성(개발)시켜야 한다. 어떤 조직은 목적과 방법에 대한 확립과 준비 없이 360° 평가가 마치 무슨 만병통치약이나 되는 듯 무조건 실행하고 보는 우를 범하곤 한다. 조직이 추구하는 목적이 명확히 설정되어 구성원에게 전달되지 않으면, 서로의 목적이 다른 관계로, 구성원의 좌절과 실망이 뒤따르게 되어 결국에는 360° 평가의 실패를 유발하게 된다.

평가도구 또한 평가결과치의 차이를 유발하는데, 평가형태와 척도를 생각할 수 있다. 평가형태란 평가내용을 여러 영역으로 구분하여 평가하는 경우(dimensional)와 전체적으로 결합하여 평가하는 경우(global)로 구분되는데, 평가자에 따라 평가영역별 가중치가 다를 수 있으므로 영역구분의 경우가 전체적인 경우보다 평가주체간 평가결과치의 상관관계가 높게 나온

다. 평가척도란 행동(behavior)을 재는 척도나 특성(trait)을 재는 척도냐는 구분 즉, 평가의 대상에 따른 구분이다. 평가주체가 달라도 행동에 대한 평가치는 차이가 적게 나오지만, 특성에 대한 평가치는 많은 차이를 보인다는 것이다. 이런 연유에서 행위기준평가척도(BARS)가 객관성이 있는 평가방법으로 강조되고 있다.

4) 평가결과의 차이를 보는 견해: 차이감소 對 차이인정

다면평가의 결과 평가주체의 차이에 따라서 평가결과가 다르게 나타나면 대부분의 사람들은 문제가 있다고 생각한다. 따라서 가급적 그 격차를 줄이는 방법을 찾으려고 한다. 하지만 모든 사람이 그렇지 않다. 어떤 사람은 평가주체에 따라 평가결과상 차이가 난다는 것 자체를 문제시하기보다는, 그것을 인정하고 거기에서 의미를 도출해 내야 한다고 주장한다. 전자의 경우는 측정 그 자체에 초점을 두고 정확한 측정치를 도출하는데 목적이 있는 심리측정가의 관점이고, 후자는 측정은 과정상의 내용이고 그 결과를 어디에 어떻게 활용하느냐에 초점을 두는 인사관리자의 관점이라 할 수 있다(Tornow, 1993). 이들 견해차를 요약하면 다음의 <표 2>와 같다. 어떤 관점을 취해야 하느냐는 것은 입장과 목적에 따라 달라진다고 본다. 중요한 것은 어느 한쪽 관점·입장에서 무조건 타입장을 비판·무시해서는 안된다는 점이다.

<표 2> 다면평가시 결과차이를 보는 관점의 차이

관점	360° 평가를 보는 견해	초점을 두는 것	평가주체간 평가결과의 다양성에 대한 견해
심리측정가	결과	객관적 현실	차이감소
인사관리자	과정	주관적 현실	차이인정 + 의미도출

II. 연구목적과 문제

1. 연구목적

전통적 상사중심의 평가에서 다면평가로의 변화가 일어나는 요즘 인사관리 담당자의 입장에서 궁금한 것은 여러 평가유형 중 어느 것이 더 정확하고 수용성이 클 것인가 하는 점이다. 설령 결합하여 사용하는 다면평가를 실시한다 하더라도 이들 평가유형간의 관계를 알아야만 어떻게 결합하여 쓸까하는 답이 나올 수 있는 것이다. 즉, 평가유형별 특성과 평가유형

간 관계를 알아보는 것이 무척 중요하다.

그러나 동일한 피평가자를 두고 여러 평가주체가 조직내의 다른 수준에서 평가한 후, 그 결과를 비교한 연구는 아직 드물다. 비록 국외의 연구에서 몇편있으나 아직 확실한 결론을 내리기 어려운 시점이고, 국내에서는 이러한 주제로 출간된 논문이 아직 없다. 따라서 본 논문은 다면평가의 구체적 내용으로 여러 평가주체가 동일 대상을 평가하였을 때 나타나는 평가주체별 결과의 특성과, 그들간의 관계를 알아보려고 한다. 앞서 다면평가의 제이슈를 살펴 보았을 때 논의되었듯이 평가주체별 평가결과의 차이가 바로 다면평가의 핵심이슈이다. 이 핵심이슈를 국내의 실증자료를 바탕으로 연구하는 것이다.

2. 연구문제

본 연구를 통해 알아보고 싶은 내용 즉, 연구문제는 다음과 같다.

연구문제 1: 상사, 본인, 부하 중 누가 관리자를 더 높게 혹은 낮게 평가할까?

연구문제 2: 상사, 본인, 부하 중 누가 과대/과소평가를 많이 혹은 적게 범할까?

연구문제 3: 평가자들이 과연 평가요소가 비슷하면 점수를 비슷하게, 평가요소가 다르면 점수를 다르게 줄까? 즉, 평가요소별 차별화를 얼마나 할까?

앞서 기존연구를 정리하면서 제시했던 것처럼, 흔히 본인이 평가할 때는 과장되는 경향이 있다고 한다. 하지만 과연 누가 많게/적게 점수를 주는지 알아 볼 필요가 있다고 본다(연구문제 1). 그런데 점수 절대치의 높낮이와 점수분포는 차이가 있다. 즉, 점수대(점수의 위치)는 달라도 점수의 분포경향은 같을 수 있고, 역으로 평균치는 같아도 점수분포는 다를 수 있다. 따라서 평가주체간 차이에 따른 평가결과간 상관관계를 봄으로써(연구문제 2) 과대/과소평가를 적게하는 즉, 보다 일관성 있게(혹은 평가결과의 변동폭이 좁게) 평가하는 주체가 누구인지를 알아낼 수 있다. 이상 두 가지 연구문제에 대한 답은 평가요소별로 그 현상이 달라질 수 있으므로 평가요소별로 경우를 구분하여 살펴볼 필요가 있다.

비록 같은 사람을 대상으로 평가하더라도 무엇을 평가하느냐가 즉, 평가요소가 다르면 평가결과가 달라질 수 있다. 그러나 후광효과(halo effect)나 그 역의 경우(horn effect)가 발생하면 비록 평가내용이 달라져도 동일/유사한 결과가 나온다. 평가요소의 차이에 따른 평가결과치의 차별화가 필요한데, 과연 상사, 본인, 부하 중 누가 그 차별화를 잘하고 못하는지 연구문제 3을 통해 알아본다. 상식적으로 보았을 때 평가주체가 다르더라도 동일 피평가자의

동일요소를 평가함에 있어서 점수가 비슷하게 나와야 하고, 평가주체가 같더라도 상이한 평가요소에 대해서는 점수가 상이하게 나와야 한다. 만약 그렇지 않다면 평가상의 오류가 발생한 것이다. 이는 바로 집중타당도와 판별타당도의 문제이다.

이상 세 가지 연구문제에 대한 답을 찾게되면 회사와 관리자는 과연 무엇에 근거하여 구성원을 평가하고 육성해야할지를 판단하는데 도움이 될 것이다.

Ⅲ. 연구방법

1. 참여자

다면평가를 연구함에 있어 두 평가주체간(예, 상사-본인, 동료-본인) 결과를 비교한 연구는 많았지만 세 평가주체간의 비교연구는 70년대 중반에 시작되었으나, 아직 많지 않다. Klimoski와 London(1974)은 상사-본인-동료에 의한 평가를 비교연구했는데, 이때만 해도 상향식 평가는 언급되지 않을 때였다. 그 뒤 상향식 평가가 유행하고 중요하게 인식되었다. 따라서 부하에 의한 상향식 평가를 연구에 포함시켜 비교분석하게 되었고, 상향식을 고려하다 보니 피평가자가 관리자가 되었다. 상사-본인-동료로 구성된 연구는 일반하급직을 대상으로 할 수 있으므로 표본수도 쉽게 늘릴 수 있고, 자료수집도 용이하다. 하지만 상사-본인-부하로 구성된 연구는 본인이 관리자여야 하므로, 평가주체의 직급이 높아져, 표본수도 많이 확보하기 어렵고 자료수집도 용이하지 않다. 그 결과 세 평가주체간 비교연구의 경우 대부분 의 기존연구는 부하 대신에 동료를 넣고 있는데, 부하를 넣는 경우엔 표본수가 적은 흠을 지니고 있다.

본 연구에서는 다면평가의 취지를 살리고자 상사와 본인평가 외에 부하에 의한 평가를 포함시켰다. 그런데 관리자에 대한 동료평가는, 보통 관리자들끼리는 같은 부서에 있지 않는 관계로, 관찰기회부족에 따른 평가정확도의 저하가 우려되어 제외하였다.

앞서 평가결과에 영향을 미칠 수 있는 제영향요인을 살펴보았다. 기존의 연구에서는 많은 경우 학생을 활용하여 (그들에서 서로 다른 역할을 부여한 후) 다면평가의 현상을 연구하였는데, 역할연기(role play) 그 자체가 평가결과에 영향을 미칠 수도 있기에 본 연구에서는 외생변수의 영향을 최대한 줄이고자 실제 기업의 임직원을 대상으로 한 현장연구를 실시하였다.

동일한 피평가자(관리자)에 대한 세 주체(상사-본인-부하)의 평가결과를 비교해야하고, 평가내용과 방법이 같아야만 결합과 비교가 가능하기에 본 연구는 동일조직내 구성원을 대

상으로 연구하였다. 연구대상이 된 조직은 종업원 2,000명 정도의 증권회사(이하에서는 A사라고 칭함)였다. 구체적으로 A사 팀점장 108명의 능력을 본인, 상사, 복수의 부하가 평가하도록 하였다. 108명의 관리자는 본사팀장 38명과 지점장 70명으로서, 해외근무자와 출장중인자를 제외한 국내근무 단위조직 책임자급 관리자의 100%가 참여한 것이었다. 관리자별 부하의 수는 팀이나 지점의 구성원 수에 따라 달랐다. 대체로 5~20명 정도였는데, 본 조사에 응한 부하의 수는 1,570명이었다. 상사평가는 관리자의 직속상사인 임원에 의해 이루어졌다. 본 연구에 참여한 구성원 수는 임원, 관리자, 부서 구성원을 합쳐서 1,700명 정도였다.

2. 측정내용과 방법

1) 관리자의 능력요건

관리자의 능력요건을 규정하기란 쉬운 일이 아니다. 업종, 조직규모, 근무지역 등 제반변수가 관리자의 능력요건 결정에 영향을 미치기 때문이다. 본 연구는 관리자의 능력요건을 파악하는데 목적이 있는 것이 아니라 능력을 평가한 결과에 초점이 있다. 하지만 평가대상 즉, 관리자의 능력요건이 무엇인지 객관적으로 정립되지 않은 상태에서 평가주체를 상사, 본인, 부하로 구분하여 평가하여 본들 의미가 없으므로 (서로 다른 것을 관리자의 능력요건으로 인식할 수 있으므로), 관리자의 능력요건이 기정립된 기업을 대상으로 연구하였다.

본 연구의 대상이 A사에서는 1996년 사내의 전문가집단의 작업을 통해 구성원의 능력요건을 도출한 후, 그 내용을 바탕으로 인력개발을 추구하고 있다. A사에서는 구성원의 공통능력요건과 함께 계층별(일반사원과 관리자)로 추가능력요건을 도출하였는데, 관리자(A사에서는 팀점장이라고 함)의 능력요건은 다음 <표 3>과 같다.

이들 능력요건은 임직원 대상의 설문조사와 인터뷰를 바탕으로 도출된 내용(능력요건)을 세 차례(관리자 대상 두 차례와 임원대상 한 차례)의 workshop을 거쳐 마련된 것으로 해당 회사(A사)내에서의 정확도와 수용성은 충분히 검증된 것이다. 모든 구성원에겐 네 가지의 태

<표 3> A사 관리자의 능력요건

조직능력	전사최적화능력	팀점전략수립능력	팀점활성화능력
개인능력	과제해결능력		대인관계능력
	업무지식	정보 서비스능력	창조적 학습능력
태도(기본자세)	긍정적 사고	일등정신	고객지향성 정도의식

도(기본자세)와 다섯 가지의 기본적 개인능력이 필요하며, 관리자에겐 이들 아홉 가지에 추가적으로 조직능력이라 하는 세 가지가 더 요구된다는 것이다. 물론 <표 3>에 적힌 능력요건은 세부능력요건을 묶어서 표현한 용어으로써, 각 영역(category)마다 세부능력요건이 있는데 그 예는 다음과 같다.

<표 4> A사 능력요건의 세부사항(예)

일등정신	대인관계능력	팀점전략수립능력
<ul style="list-style-type: none"> · 도전의식 · 승부근성 · 목표달성의식 · 선점/선도의식 	<ul style="list-style-type: none"> · 금욕인 매너 · 스트레스 내성 · 감성활용 · 적극적 경청 · 포용성 	<ul style="list-style-type: none"> · 전략적 사고 · 방향성/역할선정 · 변화관리

2) 능력의 측정방법

위에 설명된 관리자의 능력요건은 설문지를 통해 측정되었다. A사에서는 앞서 설명한 관리자의 능력요건이 도출된 후, 본인과 타인평가가 가능하게끔 세부능력요건마다 복수의 문항을 내·외부 전문가집단에 의해 개발하였다. 설문지 개발과정에서 세 차례의 workshop을 거치면서 300여명을 대상으로 pilot test를 실시하였다. 그 과정에서 요인분석을 통해 설문문항의 타당성을 높였고, 신뢰도 검사(반분신뢰도 검사, 측정-재측정 검사, 및 문항분석)를 통하여 매우 높은 수준의 신뢰도를 확보하였다. 그 결과 Cronbach α 의 경우 거의 대부분 항목이 .80을 넘었다. 이러한 과정을 통해 최종적으로 95문항으로된 설문지가 만들어졌는데, 이 설문지는 개발과정에서 조직내 구성원의 충분한 참여와 의견반영이 있었기에 측정내용과 방법에는 충분한 타당성과 구성원 수용성이 있다. 동일한 설문지가 본인평가, 상사평가, 그리고 부하평가에 사용되었는데, 이때 피평가자를 나타내는 주어만 다를 뿐(예로 본인평가의 경우에는 '나는', 그리고 타인평가의 경우에는 '그는'으로 문장이 시작됨) 내용상 같았다.

설문지의 95개 문항 모두는 5점척도에 입각하여 응답케 하고 있다. 이들 95개 문항은 43개의 변수(그 예는 <표 4>에 제시되었다)를 측정하는 것이고, 이들 43개 변수가 묶여져서 <표 3>의 12가지 평가요소를 나타낸다. 설문응답의 결과는 피평가자 개인별 전체점수를 계산할 때는 43개 변수의 개별값을 더한 수치(따라서 평가치의 범위는 43~215임)를 쓰고, 평가요소별로 파악할 때는 5점척도의 평균치(따라서 범위는 1~5임)를 활용했다.

3. 연구절차

설문조사와 분석은 1996년 7월과 8월에 걸쳐 이루어 졌는데, 설문지 배포 및 회수는 A사의 인력개발부서가 주관했고, 자료입력과 분석은 내·외부 전문가집단이 수행했다. 설문조사의 결과가 구성원에게 피드백 되었는데, 해당자들은 그 내용이 자기변화와 발전계획수립에 기초 자료로서 매우 유익하다는 반응을 보였다. A사에서는 이 설문조사(본인, 부하, 및 상사평가)의 결과를 바탕으로 각 구성원별 인력개발의 방향 및 내용파악, 나아가 인사관리의 자료로 활용하고 있다.

IV. 조사결과

1. 평가주체별 평가특성의 비교

평가주체별 개별평가요소와 그들의 결합치인 총점에 대한 평균과 표준편차, 그리고 제평가주체별 평가요소간 상관관계가 <표 5>에 제시되어 있다. 연구문제 1에서 제기되었듯이 상사, 본인, 부하 중 누가 관리자의 능력을 더 높게 혹은 낮게 평가하는지를 알아보기 위해 분산분석, t-검사, 빈도분석을 활용했고 그 결과는 <표 6>에 나와있다.

<표 6>의 윗부분에 나와있듯이, 상사, 본인, 부하의 평가가 개별평가요소 세 가지의 경우 개인능력을 제외하고 조직능력과 태도면에서 뚜렷이 (유의한 수준에서) 서로 달랐다. 그 결과 종합점수인 총점에서도 세 주체에 의한 평가결과가 서로 유의하게 차이나는 것으로 나타났다. 본인평가가 가장 높았고, 그 다음은 부하평가, 그리고 상사에 의한 평가점수가 가장 낮은 즉, 본인>부하>상사평가의 순으로 나타났다.

세 평가주체의 평가결과차이를 동시에 비교한 분산분석에서 유의한 차이가 나왔지만 세부적으로 누구에 의한 평가가 어떻게(얼마나) 높고 낮았는지를 파악하기 위해 두 평가주체씩 짝을 지어 비교·분석하였다. T-검사와 빈도분석을 하였는데 그 결과 또한 <표 6>에 같이 수록되어 있다. 상사-본인평가간 비교와 상사-부하평가간 비교에서도 결과는 마찬가지였다. 즉, 개인능력의 경우를 제외하고 태도, 조직능력, 총점에서 유의한 차이가 있는 것으로 나타났다. 그러나 본인-부하평가간 결과는 사뭇 달랐다. 태도측면에서는 평가결과간 유의한 차이가 있었으나 나머지 측면에서는, 나아가 총점에서도 유의한 차이가 없었다. 즉, 본인평가와 부하평가의 결과는 거의 같게 나타났다.

〈표 5〉 평가주체별 평가요소의 평균, 표준편차, 및 상관관계

평가주체 및 평가요소	평균	표준편차	1	2	3	4	5	6	7	8	9	10	11
1. 총점	149.532	27.172											
2. 태도	3.669	.493	.785**										
3. 개인능력	3.635	.478	.789***	.941***									
4. 조직능력	3.336	.871	.960***	.594***	.591**								
5. 총점	162.861	11.091	.184	.111	.142	.186							
6. 태도	4.077	.290	.104	.095	.118	.090	.877***						
7. 개인능력	3.635	.327	.216	.135	.159	.220*	.919***	.687***					
8. 조직능력	3.716	.263	.165	.106	.140	.163	.915***	.755***	.784***				
9. 총점	160.137	12.104	.387**	.461***	.445***	.305***	.116	.095	.110	.118			
10. 태도	3.930	.299	.441**	.501**	.478***	.359***	.135	.152	.091	.138	.951***		
11. 개인능력	3.651	.311	.336**	.410**	.398**	.259**	.152	.072	.217	.125	.952***	.850***	
12. 조직능력	3.673	.298	.371**	.439***	.423**	.292**	.060	.080	.019	.083	.964***	.908***	.869***

주) N=108. 대각선으로 놓인 직사각형내의 수치는 집단타당도계수를, 삼각형내의 수치는 판별타당도계수를 나타냄.

*p<.05 **p<.01 ***p<.005

〈표 6〉 평가주체간 평가결과치 비교표

분석내용	총점	태도	개인능력	조직능력
상사-본인, 부하평가 결과간 분석분석: 유의도	.000	.000	.937	.000
상사평가의 결과: 평균	149.532	3.669	3.635	3.336
본인평가의 결과: 평균	162.861	4.077	3.635	3.716
부하평가의 결과: 평균	160.137	3.930	3.651	3.673
상사와 본인평가 결과간 T-검정: 유의도	.000	.000	.992	.000
상사평가치>본인평가치의 경우: 빈도	36	23	61	43
상사평가치=본인평가치의 경우: 빈도	-	-	-	-
상사평가치<본인평가치의 경우: 빈도	72	85	47	65
상사와 부하평가 결과간 T-검정: 유의도	.000	.000	.708	.000
상사평가치>부하평가치의 경우: 빈도	36	35	52	39
상사평가치=부하평가치의 경우: 빈도	-	-	-	-
상사평가치<부하평가치의 경우: 빈도	72	73	56	69
본인과 부하평가 결과간 T-검정: 유의도	.069	.000	.680	.238
본인평가치>부하평가치의 경우: 빈도	54	69	48	53
본인평가치=부하평가치의 경우: 빈도	-	1	1	3
본인평가치<부하평가치의 경우: 빈도	54	38	59	52

N = 108

상사평가치와 본인평가치의 비교에서 특이한 것은, 대체로 본인평가치가 훨씬 높았지만, 개인능력의 경우엔 상사가 본인보다 높게 평가한 경우가 그 반대의 경우보다 61 대 47로 오히려 많았다는 것이다. 그런데 평균치는 양측 모두 3.63이었으므로 상사가 높게 평가한 경우라 하더라도 본인보다 월등히 높은 점수를 부여하지는 않았음을 알 수 있다. 그에 반해 본인이 상사보다 높게 평가한 경우엔 본인이 스스로에게 상당히 높은 점수를 부여하였다는 것이다.

본인-부하평가의 경우엔 비록 부하들이 관리자의 개인능력을 높게 인정해주는 경우가 더 많았으나, 조직능력의 경우엔 높게 보는 쪽과 낮게 보는 쪽이 거의 같았고, 태도의 경우에는 오히려 낮게 보는 쪽이 높게 보는 쪽보다 더 많았다. 그 결과 종합치(총점)에서는 관리자를 높게 평가한 경우와 그 반대의 경우가 54 대 54로 완전히 일치하였다. 〈표 6〉에 나타난 결과를 다시 정리하여 평가요소별 세 평가주체간 평가결과치의 관계를 나타내면 다음 〈표 7〉과 같다. 태도에서는 세 평가주체간 유의한 차이가 있으면서 본인평가시 관대화 현상이 있었으나, 개인능력 평가에서는 세 평가주체간 차이가 없는 것으로 나타났다.

〈표 7〉 평가요소별 평가주체간 평가결과의 관계

상사		개인능력	
상사<부하≤본인	상사<부하<본인	상사=부하=본인	상사<부하=본인

2. 과대/과소평가의 가능성 분석

평가주체에 따라서 과연 과대/과소평가의 가능성이 누구에게 얼마나 있는지 알아보기 위해 상관관계분석과 빈도분석을 활용했다. 〈표 5〉에 있듯이 평가주체 및 평가요소간 상관관계를 보았다. 타인(상사-부하)평가간 상관관계계수(모두 $p < .005$ 수준에서 유의함)에 비하여 본인-타인(상사-본인, 부하-본인)평가간 상관관계계수는 크기가 훨씬 적었고 통계적으로 거의 유의하지 않았다. 이는 타인 즉, 상사와 부하평가의 경우는 서로 맥을 같이하여 공히 관리자의 평가대상 내용이 좋으면 높게, 낮으면 낮게 평가한 경향이 강한데 비해 본인평가는 그렇지 못한 것으로 나타났다. 즉, 본인평가의 경우에는 객관적 현상과는 달리 좋아도 낮게, 나빠도 높게 평가한 과소/과대평가의 경우가 있었다는 것이다.

평균점수로서 총점만 보면 본인평가의 점수가 가장 높아서 관대화 오류가 발생한 것처럼 보인다. 그러나 짝을 지어 비교하면서 상대방 보다 높게 평가한 빈도(〈표 6〉참조)를 보면 상사-본인평가의 경우엔 본인이 상사보다 높게 평가한 경우가 그 반대의 경우보다도 두배나 많았지만(72 대 36), 본인-부하평가의 경우엔 본인이 부하보다 높게 평가한 경우와 그 반대의 경우가 54 대 54로 동일했다. 따라서 무조건 관리자들이 본인평가시 점수를 높게 준다고 즉, 관대화 오류를 범한다고 볼 수는 없다. 하지만 전체적으로 보았을 때, 본인평가의 경우 본인에게 점수를 낮게 주는 경우에는 약간 낮게, 점수를 많이 주는 경우엔 많이 높게 주는 것으로 나타났다. 다시 말하면 본인평가시 겸손은 약하게, 이기심은 강하게 발휘하는 것으로 나타났다. 따라서 약간의 겸양효과(modesty effect)와 많은 관대화 효과(leniency effect)가 발생했다고 하겠다.

3. 평가요소의 동질/이질성에 따른 동등화/차별화 경향분석

동일 피평가자를 대상으로 평가할 때, 평가자가 달라도 동일한 평가요소는 같게 평가해야 되고, 평가자가 같아도 이질적 평가요소는 다르게 평가해야 한다. 이러한 동등화/차별화의 정도를 Campbell과 Fiske(1959)가 제창한 다속성-다측정 매트릭스(multitrait-multimethod matrix)방법을 사용하여 집중타당도와 판별타당도로 살펴보았다. 본 연구에서는 평가(측정) 도구(method)가 다른 것이 아니라 평가주체(rater)가 달랐으므로 다속성-다주체 매트릭스

(multitrait-multirater matrix)라고 해야할 것이다. <표 5>의 상관관계 매트릭스에서 대각선으로 놓인 직사각형내의 수치는 상이한 평가주체에 의한 동일평가요소의 평가결과간 관련성을 나타내는 집중타당도이고, 삼각형내의 수치는 동일평가주체에 의한 상이한 평가요소의 평가결과간 관련성을 나타내는 판별타당도이다.

결과를 보면, 판별타당도계수는 높은 반면 집중타당도계수는 대체로 낮았다. 판별타당도계수는 낮아야 타당도가 높고, 집중타당도계수는 높아야 타당도가 높은 것이므로, 결과를 해석하면 같은 평가주체하에서 상이한 평가요소를 두고 그 차이를 구분(차별화)함이 미약했고, 다른 평가주체가 같은 평가요소를 두고 그 유사성을 같게 보려는 동등화의 정도 또한 상사-부하평가의 경우를 제외하고는 약했다. 즉, 같은 것을 같게 보는 동등화는 조금 있었으나, 다른 것을 다르게 보는 차별화는 약했다는 것이다. 이는 평가자에 따라서 측정상의 차이 (혹은 심하게는 오류)가 있다는 것이다. 과연 누구의 평가에서 오류가 상대적으로 더 혹은 덜 발생했는지 살펴보아야 할 것이다.

집중타당도계수 중에서 상사-부하평가간 집중타당도는 매우 높은 것으로 나타났지만 본인이 결부되었을 때 즉, 상사-본인평가의 경우와 부하-본인평가의 경우에 집중타당도가 낮게 나타났다. 이는 앞서 과대/과소평가의 여부를 파악하려는 부분에서도 설명된 것으로, 관리자 본인의 평가시에 타인과 동일한 내용(평가요소)을 평가하더라도 그것을 보는 관점 혹은 부여하는 가중치(weight)가 다름을 의미한다.

판별타당도를 나타내는 상관계수는 모두 유의하게 높은 것으로 나타났다. 하지만 계수의 크기가 서로 차이나기에 평가주체별로 비교해 보는 것도 의미 있을 것이다. <표 5>에서 판별타당도계수에 해당하는 것을 비교가 쉽게 평가주체별로 다시 다음의 <표 8>에 옮겨 수록하였다.

부하평가의 결과를 보면, <표 8>에서의 상관계수가 고르게 .85를 넘고 있어, 평가요소간에 평가의 차별화가 거의 이루어지지 않고 있음을 나타낸다. 이에 비해 본인평가의 경우는, 상

<표 8> 평가주체별 판별타당도계수

평가요소	관리자	부하평가	본인평가
태도-개인능력	.687***	.850***	.941***
태도-조직능력	.755***	.908***	.594***
개인능력-조직능력	.784***	.869***	.591**

p<.01, *p<.005

관계수의 크기가 부하평가의 경우보다 모두 낮아, 평가요소간 평가의 차별화가 그래도 부하 평가 보다는 상대적으로 약간 있음을 나타낸다. 상사평가의 경우 태도와 개인능력의 비교에 선 평가요소간 차별화가 거의 없었지만(.941), 조직능력을 태도나 개인능력과 결부시켰을 경우엔 (.941에 비해 훨씬 적은 .594와 .591이었으므로) 차별화가 있음을 알 수 있다. 즉, 상사들은 부하인 관리자들의 조직관리능력을 평가함에 있어서는 그들의 태도나 개인능력과 차별화하여 보는(평가하는) 경향이 있음을 의미한다.

V. 토의 및 결론

이상 도출된 결과를 종합·정리하면 <표 9>와 같다. 기존의 연구에서 나타났듯이 본 연구에서도 본인평가의 결과가 타인(상사 및 부하)평가의 결과보다 대체로 높게 나타난다는 것이 입증되었다. 그리고 본인평가가 가끔 부하평가와 비슷하게 나오더라도 상사평가는 본인 또는 부하평가보다 낮은 경향이 있다는 것이 새롭게 도출되었다. 상사평가가 본인평가보다 낮게 나올 수 있는 것은, 상사-본인-부하평가의 결과간 상관분석에서 나타났듯이, 상사평가는 일관성이 있는 반면(즉, 좋으면 높게, 나쁘면 낮게 평가하는 경향이 있는 반면) 본인평가는 일관성이 적기 때문이다.

평가요소별로 구분하여 평가주체별 평가결과를 비교했을 때, 개인능력 측면에서는 세 주체

<표 9> 조사결과의 종합

연구문제	결과
평가주체별 평가특성 비교	총점: 상사<부하<본인 태도: 상사<부하<본인 개인능력: 상사=부하=본인 조직능력: 상사<부하=본인
과대/과소평가 여부	상관계수: 상사-부하(大), 상사-본인(小), 부하-본인(小) ➡ 본인평가가 타인평가와 차이남. 즉, 본인평가의 내용에 정확도가 떨어짐. 약간의 겸양효과와 많은 관대화 효과가 존재함.
평가요소의 평가동등화와 차별화 경향	이질주체의 동일요소의 동등화: 동일주체의 이질요소간 차별화: 상사-부하: 높음 상사: 약간 있음-조직능력 구분 상사-본인: 낮음 본인: 모든 요소에서 약함 부하-본인: 낮음 부하: 모든 요소에서 약함

간 유의한 차이가 없는 것에 비해 태도 측면에서는 평가주체간 d의한 차이가 있었다. 즉, 조직내에서의 위치(상사로서의 임원, 관리자 본인, 부하)가 다름에 따라 관리자의 태도는 서로 달리 평가하지만 개인능력은 같게 평가한다는 것이다. 개인능력은 겉으로 드러나는 즉, 객관성이 있는 것을 평가하니까 평가주체간 차이가 없었지만, 태도(기본자세)는 겉으로 드러나지 않고 개인 내부에 잠재되어 있는 즉, 객관화되어 있지 않는 것을 평가하니까 평가주체간 평가결과에 차이가 있는 것으로 판단된다. 조직능력은 부하와 본인의 경우는 같았으나 (즉, 유의하게 차이하지 않았으나), 상사의 평가는 사뭇 달랐다. 관리자의 조직능력을 평가하는 상사의 기준과 내용이 본인이나 부하의 그것과는 다름을 의미한다.

과대/과소평가 여부의 조사에서는, 상사나 부하로부터 인정받는 (높게 평가받는) 사람이 본인평가지 중간정도나 낮은 수준으로 스스로를 평가하는 경우가 종종 있는 것으로 나타났다. 즉, 본인평가에서는 겸손함 때문이거나 혹은 스스로 설정한 높은 수준의 목표 때문에 자신을 평가절하 하는 경우가 있다. 그런데 이와는 반대의 이유로 인한 과대평가 또한 흔히 초래되었다. 따라서 본인평가의 결과가 타인(상사 혹은 부하)평가의 결과와 결부되었을 때엔 그 상관관계가 낮게 나왔다.

평가요소의 동등화와 차별화 조사에서 상이한 평가주체가 동일평가요소를 얼마나 같게 평가하느냐는 집중타당도의 경우에서(〈표 5〉 참조), 동일요소에 대한 부하-상사평가간 상관계수 즉, 집중타당도가 부하-본인평가간 상관계수 보다 높게 나온 것은 기존의 연구결과(예, Mount, 1984)와 맥을 같이 한다. 동일요소에 대하여 상사평가와 부하평가의 결과가 매우 유사하다는 것은 관리자가 그들(상사와 부하)의 평가에 입각하여 자기의 문제를 발견하고 발전방향을 도모하면 무리가 없다는 것을 의미한다.

집중타당도계수에 비해 판별타당도계수가 높게 나온 것 또한 기존의 연구결과(예, Klimoski & London, 1974)와 같다. 부연하면, 이질평가요소간 평가차별화 경향에서 부하와 본인은 평가요소간 차별화를 안하는/못하는 것으로 나타났다. 즉, 부하는 상사인 관리자를(본인 또한 자신을) 평가할 때 평가요소별로 구분하여 달리 인식하기보다는, 상사의 (관리자의 경우엔 본인의) 능력을 전체적으로 인식하는 경향이 크다고 할 수 있다. 이에 반해 상사는 평가시 평가요소에 따라 차별화 하는 경향이 본인이나 부하나 본인보다는 상대적으로 큰 것으로 나타났는데, 특히 조직능력을 태도나 개인능력과 잘 구별해내는 것으로 나타났다. 이를 종합하면 관리자의 조직능력은 상사평가를 기준으로 하고, 태도와 개인능력은 제주체별 평가결과를 결합하여 활용하는 것이 바람직한 것으로 판단된다.

참 고 문 헌

- Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. *Organizational Dynamics*, 25(2), 24-38.
- Atwater, L. & Waldman, D. (1998). Accountability in 360 degree feedback. *HRMagazine*, 43(6), 96-101.
- Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-other agreement: Does it really matter? *Personnel Psychology*, 51, 577-598.
- Baird, L. S. (1997). Self and supervisor ratings of performance: As related to self-esteem and satisfaction with supervision. *Academy of Management Journal*, 20, 291-300.
- Bohl, D. L. (1996). Minisurvey: 360-degree appraisals yield superior results, survey shows. *Compensation & Benefits Review*, September/October, 16-19.
- Borman, W. C. (1974). The rating of individuals in organizations: An alternate approach. *Organizational Behavior and Human Performance*, 12, 105-124.
- Campbell, J. P., Dunnette, M. D., Lawler, E. E., & Weick, K. E., (1970). *Managerial behavior, performance, and effectiveness*. New York: McGraw-Hill.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychologica Bulletin*, 56, 81-105.
- Cascio, W. F. (1987). *Applied psychology in personnel management*. Englewood Cliffs, NJ: Prentice-Hall.
- Church, A. H. & Bracken, D. W. (1997). Advancing the state of the art of 360-degree feedback. *Group & Organization Management*, 22(5), 149-161.
- Coates, D. E. (1998). Don't tie 360 ° feedback to pay. *Training*, 35(9), 68-74.
- DeVader, C. L., Bateson, A. G., & Lord, R. G. (1986). Attribution theory: A meta-analysis of attributional hypotheses, In E. Locke (ed.), *Generalizing*

- from laboratory to field settings* (pp. 63-81), Lexington, MA: Lexington Books.
- Downey, R. G., Medland, F. F., & Yates, L. G. (1976). Evaluation of a peer rating system for predicting subsequent promotion of senior military officers. *Journal of Applied Psychology*, 61(2), 206-209.
- Farh, J. & Werbel, J. D. (1986). Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. *Journal of Applied Psychology*, 71(3), 527-529.
- Fox, S. & Dinur, Y. (1988). Validity of self-assessment: A field evaluation. *Personnel Psychology*, 41(3), 581-592.
- Furnham, A. & Stringfield, P. (1994). Congruence of self and subordinate ratings of managerial practices as a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*, 67, 57-67.
- Furnham, A. & Stringfield, P. (1998). Congruence in job-performance ratings: A study of 360 ° Feedback examining self, manager, peers, and consultant ratings. *Human Relations*, 51(4), 517-530.
- Garavan, T. N., Morley, M., & Flynn, M. (1997). 360 degree feedback: Its role in employee development. *Journal of Management Development*, February-March, 134-148.
- Harris, M. M. & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Hoffman, R. (1995 April). Ten reasons you should be using 360-degree feedback. *HRmagazine*, 82-85.
- Holzbach, R. L. (1978). Rater bias in performance ratings: Superior-, self-, and peer ratings. *Journal of Applied Psychology*, 63, 579-588.
- Jones, E. E. & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. A. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valines, & B. Weiner (eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94), Morristown, NJ: General Learning Press.
- Kay, E., Meyer, H. H., & French, J. R. P. (1965). Effects of threat in a

- performance appraisal interview. *Journal of Applied Psychology*, 59, 445-451.
- Klimoski, R. J. & London, M. (1974). Role of the rater in performance appraisal. *Journal of Applied Psychology*, 59(4), 445-451.
- Kraut, A. (1975). Prediction of managerial success by peer and training-staff ratings. *Journal of Applied Psychology*, 60(1), 14-19.
- Landy, F. J., Farr, J. L., Saal, F. E., & Freytag, W. R. (1976). Behaviorally anchored scales for rating the performance of police officers. *Journal of Applied Psychology*, 61, 458-557.
- Latham, G. P. & Wexley, K. N. (1982). *Increasing productivity through performance appraisal*. Reading, MA: Addison-Wesley.
- Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology*, 51, 369-381.
- Lepsinger, R. & Lucia, A. D. (1997). 360° feedback and performance appraisal. *Training*, September, 62-70.
- London, M. & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management*, Fall, 353-373.
- McEvoy, G. & Buller, P. (1987). User acceptance of peer appraisals in an industrial setting. *Personnel Psychology*, 37(4), 785-798.
- McGarvey, R. & Smith, S. (1993). When workers rate the boss. *Training*, 30, March, 31-34.
- Milkovich, G. T. & Boudreau, J. W. (1994). *Human resource management*. Burr Ridge, IL: Irwin.
- Milliman, J. F., Zawacki, R. A., Norman, C., Powell, L., & Kirksey, J. (1994). Companies evaluate employees from all perspectives. *Personnel Journal*, November, 99-104.
- Moses, J., Hollenbeck, G. P., & Sorcher, M. (1993). Other people's expectations. *Human Resource Management*, 32(2-3), 283-297.
- Mount, M. K. (1984). Supervisor, self, and subordinate ratings of performance and satisfaction with supervision. *Journal of Management*, 10, 305-320.
- Mount, M. K., 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, 557-576.
- Norman, C. A. & Zawacki, R. A. (1991). Team appraisals-team approach. *Personnel Journal*, September, 101-103.
- Nowack, K. M. (1993). 360-degree feedback: The whole story. *Training & Development*, January, 69-73.
- O'Reilly, B. (1994). 360° s feedback can change your life. *Fortune*, 130(8), 93-100.
- Ramsey, M. L. & Lehto, H. (1994). The power of peer review. *Training and Development*, July, 38-41.
- Redman, T. & Snape, E. (1992). Upward and onward: Can staff appraise their managers? *Personnel Review*, 21, 32-46.
- Salam, S., Cox, J. F., Sims, H. P., Jr. (1997). In the eye of the beholder: How leadership relates to 360-degree performance ratings. *Group & Organization Management*, 22(2), 185-209.
- Santora, J. E. (1992). Rating the boss at Chrysler. *Personnel Journal*, May, 38-45.
- Sherman, A. W. Jr., Bohlander, G. W., & Snell, S. A. (1996). *Managing Human Resources*. Cincinnati, OH: South-Western.
- Steele, R. P. & Ovalle, N. K. (1984). Self-appraisal based upon supervisory feedback. *Personnel Psychology*, 37, 667-685.
- Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource management*, 32(2-3), 221-229.
- Waldman, D. A., Atwater, L. E., & Antonioni, D. (1998). Has 360 degree feedback gone amok? *Academy of Management Executive*, 12(2), 86-94.
- Waldman, D. A. & Bowen, D. E. (1998). The acceptability of 360 degree appraisals: A customer-supplier relationship perspective. *Human Resource Management*, 37(2), 117-129.
- Yammarino, F. J. & Atwater, L. E. (1997). Do managers see themselves as others see them? Implications of self-other rating agreement for human

resources management. *Organizational Dynamics*, Spring, 35-44.

Zammuto, R. F., London, M., & Rowland, K. M. (1982). Organization and rater differences in performance appraisals. *Personnel Psychology*, 35, 643-658.