韓國心理學會誌 : 社會 및 性格
Korean Journal of Social and Personality Psychology
1998. Vol.12, No.2, 181~198

매개변인이 평가자간 일치에 미치는 효과: 연구결과와 연구방법에 대한 고찰

정 태 연

연세대학교 인간행동연구소

본 연구는 평가자간 일치에 대한 매개변인 접근방식을 연구결과 및 연구방법론 중심으로 고찰했다. 먼저, 매개변인(행동일관성, 관찰가능성, 그리고 사회적 바람직성)에 관한 지금까지의 연구들이 사용한 두 가지 유형의 평가자간 일치(자아-동료간 일치와 동료-동료간 일치) 및 그 일치정도를 나타내기 위해 사용된 지표(불일치, 상관, 그리고 변량)들의 차이점을 알아보았다. 또한, 평가자간 일치에 미치는 매개변인의 효과를 검증하기 위해 사용된 기존의 두 가지 접근방법(평가대상자에 기초한 방식과 성격특질에 기초한 방식)의 특성과 문제점들을 살펴보았다. 특히, 여러 매개변인들 중에서 관찰가능성만이 일관성 있게 그 효과가 있는 것으로 나타났는데, 이와 같은 매개변인 접근방식의 전반적인 실패가 방법론상의 문제점과 깊이 관련될 수 있음을 보였다. 본 연구는 이런 문제점을 해결할 수 있는 대안과 함께 특질수준에 기초한 새로운 접근방법을 소개했다. 마지막으로, 매개변인 접근방법의 활용 및 앞으로의 연구방향에 대해 논의하였다.

행동을 주체하는 개인의 특성과 그 행동이 일어나는 상황 중에서 어느 쪽이 이 개인의 행동을 결정하는데 더 중요한 역할을 하는지, 그리고 한 개인의 행동이 여러 상황에 걸쳐 변하지 않고 일관적인지에 대한 논의가 사회 및 성격심리학 분야에서 오랜 기간동안 계속되어 왔다(e.g., Kenrick과 Funder, 1988; Mischel, 1968; Mischel과 Peake, 1982; Ross와 Nisbett, 1991). 예를 들면, Mischel(1968)은 행동이 여러 서로 다른 상황과 시간대에 걸쳐 자주 변하기 때문에 성격특질로부터 행동을 예측하는 것이 불가능하지는 않더라도 매우 어렵다고 주장했다. 그는 행동과 성격특질간의 상관정도가 최고 .30이라는 연구결과를 그 증거로 제시했다. 전통적인 성격이론에 대한 이러한 비판은 여러 반론들을 지속적으로 불러 일으켰다. 가령, Funder와 Colvin(1991)의 반론에 따르면, 성격특질이론을 비판하는 사람들은 행동일관성을 검증하는데 있어서 지나치게 구체적인 행동에만 집착하여 하위수준의 분석에만 의존하고 있다는 것이다. 말하자면, 비판자들이 보는 행동의 일관성이란 여러 다른 상황에 걸쳐서 보이는 똑같은 행동만을 지칭한다는 것이다. Funder와 Colvin은 실험실 연구를 통해 개인의 구체적인 행동이 여러 상황에 걸쳐 다르더라도 심리적인 차원에서 매우 일관적일 수 있음을 보여주고 있다.

평가자간 일치(interjudge agreement) - 두 사람이 어떤 한 사람을 평가할 때 이 둘의 평가가 같은 정도 - 는 행동과 특질간의 관계에 대한 논의와 깊이 관

련되어 있다. 예를 들어, 한 개인의 행동이 여러 상황에 따라 다르며 평가자들이 이 개인의 행동을 서로 다른 상황에서 관찰한다면, 이 개인에 대한 평가자들의 평가는 서로 같지 않을 것이다. 반대로, 한 개인의 행동이 여러 상황에 걸쳐 일관적일 때, 평가자들이 보는 그의 행동은 모두 비슷할 것이기 때문에 결국 이 개인에 대한 평가자들의 평가는 일치할 것이다. 연구자들은 이러한 논리 속에서 평가자간 일치를 행동의 일관성을 나타내는 지표로 사용해 왔다(Zuckerman, Bernieri, Koestner와 Rosenthal, 1989). 평가자간 일치는 행동의 일관성을 연구하는데 많은 유용성을 가지고 있다. 한 개인의 행동이 서로 다른 상황에 걸쳐 일관적인지를 실제로 측정하는 것이 쉽지 않고, 여러 다른 상황간의 차이점을 제대로 평가할 만한 기술도 없다(Bem과 Funder, 1978; Funder와 Colvin, 1991). 또한, 평가자간 일치를 사용함으로써 시간적으로나 경제적인 측면에서도 유리한 점이 많다.

본 연구에서는 평가자간 일치에 대한 지금까지의 연구결과와 연구방법을 고찰해 볼 것이다. 먼저, 평가자간 일치에는 어떠한 유형이 있으며, 기존의 연구에서 평가자간 일치를 어떻게 측정하였는지 알아볼 것이다. 그 다음, 평가자간 일치 정도에 영향을 미치는 변인들, 즉 매개변인들의 종류와 이 매개변인들의 효과를 검증하기 위해 사용된 기존의 연구방법들의 특성 및 취약점을 살펴볼 것이다. 또한, 이러한 취약점들을 개선할 수 있는 방안을 탐색해 볼 것이다. 마지막으로, 평가자간 일치에 대한 매개변인 접근방법의 과제를 고찰해 볼 것이다.

평가자간 일치의 유형과 측정

대인지각이나 성격평가 분야에 관여하는 연구자들은 평가자간 일치를 주로 두 가지 유형으로 측정했다: 자아-동료간 일치(self-peer agreement)와 동료-동료간 일치(peer-peer agreement). 자아-동료간 일치는 어떤 특정 개인의 자기평가(self-ratings)와 그 개인에 대한 타인의 평가가 일치하는 정도를 나타낸다. 동료-동료간 일치는 제3의 인물에 대한 두 사람의 평가가 같은 정도를 나타낸다. 대부분의 연구자들이 자아-동료간 일치와 동료-동료간 일치를 뚜렷한 구분 없이 사용해 왔지만, 이 두 유형은 여러 측면에서 다르다(John과 Robins, 1993; Kenny, 1994). 먼저, 자아-동료간 일치가 의미를 갖기 위해서는 먼저 동료-동료간에 일치가 존재해야 한다(Kenny, 1994). 예를 들면, 두 명의 평가자가 제 3의 인물을 "정직한" 특질 차원에서 평가할 때, 한 평가자는 매우 정직하다고 평가한 반면, 다른 평가자는 매우 정직하지 않다고 평가해서 이 둘의 평가가 서로 다르다고 가정해 보자. 또한, 제 3의 인물은 자기 자신을 보통으로 정직한 사람으로 평가했다고 가정해 보자. 이 때, 자기평가 점수와 두 평가자가 내린 평가의 평균점수와의 일치정도를 계산해 보면, 일치수준은 매우 높을 것이다. 그러나, 실질적으로 자기평가는 어느 평가자의 평가와도 크게 일치하지 않는다. 따라서, 평가자간에 평가가 불일치할 때 자기평가와 동료평가가 일치한다는 것은 큰 의미가 없기 때문에, 동료-동료간 일치는 자아-동료간 일치의 필수조건이 된다. 또한, 이 두 유형의 일치는 서로 다른 요인에 의해 영향을 받는다. 예를 들면, 성격특질차원의 평가적 특성이 자아-동료간 일치에 큰 영향을 미친다(John과 Robins, 1993). 가령, "정직한-부정직한" 특질차원은 "외향적인-내성적인" 특질차원보다 더 평가적이다. 왜냐하면, 정직한 사람이 긍정적으로, 부정직한 사람이 부정적으로 평가되는 경향이 외향적인 사람이 긍정적으로, 내성적인 사람이 부정적으로 평가되는 경향보다 더 강하기 때문이다. John과 Robins(1993)에 따르면, 어떤 특질차원이 평가적일수록 자아-동료간 일치는 적어지는데, 이는 평가대상자(target)가 평가적인 특질차원에서 자신을 더 긍정적으로 평가하는 경향 때문이다. 반면에, 이러한 요인이 동료-동료간 일치에는 영향을 미치지 않는다.

이 두 유형의 평가자간 일치는 사회 및 성격심리학에서 여러 중요한 의미를 가지고 있다. 예를 들면, 대인지각을 연구하는 사람들은 종종 평가자의 평가가 정확한지를 판단하기 위한 기준치(criterion measure)로서 평가대상자의 자기평가를 사용해 왔으며, 대인

지각의 정확성을 재는 지표로 자아 동료간 일치를 사용하였다(Kenny, 1994). 자아-동료간 일치는 여러 서로 다른 상황에 걸쳐 행동이 일관적인지를 재는 측정치로도 사용된다(Zuckerman 등, 1989). 마지막으로, 몇몇 연구자들은 자아-동료간 일치를 개인의 성격에 특질이 존재하고 있음을 보여주는 증거로 사용해 왔다 (e.g., Wiggins, 1973; McCrae, 1982). 이 연구자들은 자아-동료간 일치가 평가자와 평가대상자가 가질 수 있는 여러 편향(bias)들로부터 영향을 받지 않는다고 주장한다. 가령, 지금까지 성격평가에 있어서 자기평가와 동료평가의 타당성에 대한 비판이 제기되었다(e.g., Shweder, 1975). 이러한 비판에 따르면, 자기평가는 자기방어나 자신을 긍정적으로 평가하려는 경향 때문에 정확하지 않고, 동료평가는 평가대상자에 대한 평가자의 고정관념에 의해 영향을 받기 때문에 부정확하다. 그러나, Wiggins(1973)나 McCrae(1982)의 주장에 따르면, 이 두 가지의 편향은 서로 독립적이기 때문에 자기평가에 있어서의 편향이 동료평가에 영향을 주지 않고 또한 동료평가에 있어서의 편향이 자기평가에 영향을 주지 않는다. 결국, 이 두 유형의 편향은 자아 - 동료간 일치수준에 영향을 주지 않기 때문에 이러한 편향들이 자아 - 동료간 일치수준을 설명할 수 없다는 것이다. 따라서, 이 연구자들은 자아-동료간 일치가 평가대상자의 성격에 특질이 존재한다는 증거가 될 수 있다고 주장한다(그러나, 이러한 주장은 사실과 다를 수 있다. 자세한 논의는 Amelang과 Borkenau(1986)를 참조).

동료-동료간 일치 역시 대인지각과 성격평가에 중요한 의미를 가지고 있다(Kenny, 1991, 1994; Kenrick과 Funder, 1988). 사회적 지각이 외부 자극에 의해 이루어지는지 아니면 평가자의 내적인 요인에 의한 것인 지의 문제가 대인지각 연구분야에서 중요한 연구주제가 되어 왔다(e.g., Higgins와 Bargh, 1987). 이 때, 연구자들은 외부자극인 평가대상자가 평가자의 지각에 미치는 영향을 동료 - 동료간 일치수준으로 측정해 왔다(Bourne, 1977; Dornbusch, Hastorf, Richardson, Muzzy와 Vreeland, 1965; Park, 1986). 성격심리학에서는 개인차를 측정하는데 있어서 기본적으로 자기평가에 의존해 왔지만, 근래에는 동료평가를 보충적인 자료로 사용하기도 한다. 성격심리학에서 동료-동료간 일치는 성격에서의 개인차를 보여주는 경험적인 증거로 사용되어 왔다(Bourne, 1977; Kenrick과 Funder, 1988). 예를 들면, Bourne(1977)은 낮은 동료 - 동료간 일치수준을 근거로 개인의 특정 성격은 존재하지 않고 대신 그들이 처한 상황이 행동을 규정한다고 주장한다. 반면에, Kenrick과 Funder(1988)는 동료-동료간 일치는 존재하며, 이것이 개인간의 성격 차이를 보여준다고 반론한다.

불일치(discrepancy), 상관(correlation), 그리고 변량(variance)으로 평가자간 일치를 측정할 수 있다(Kenny, 1994; Kenny, Albright, Malloy와 Kashy, 1994). 불일치는 두 명의 평가자가 동일한 한 명의 평가대상자를 평가할 때 일치하지 않는 정도를 나타낸다. 한 평가대상자에 대한 두 평가치간의 절대적인 차이로 불일치를 정의할 수도 있고, 여러 평가대상자에 걸친 평가치들의 기하평균으로 정의할 수도 있다. 또한, 둘 이상의 평가자가 다수의 평가대상자를 평가할 때 평가자간 일치수준을 상관계수로 측정할 수 있다. 마지막으로, 평가자간 일치의 정도는 다수의 평가자가 다수의 평가대상자를 평가할 때 생기는 총변량 중에서 평가대상자들간의 차이 때문에 생기는 변량으로 나타낼 수 있다. 예를 들어, 열 명의 평가자가 여섯 명의 평가대상자를 어떤 특질차원에 대해서 평가한다고 하자. 각 평가대상자에 대한 열 개 평가치들의 평균을 계산하여 그들을 비교한 결과 평가대상자간에 유의미한 차이가 있다면, 이것은 잠정적으로 평가자간 평가가 어느 정도 일치할 수 있음을 함축한다. 이 때, 평가자간의 평가가 어느 정도 일치하는 지는 전체 변량에서 평가대상자들간의 차이 때문에 생기는 변량으로 나타낸다. 즉, 사람마다 어떤 특질을 가지고 있는 정도에 차이가 있다는 가정 하에, 평가자들이 이 차이를 정확히 지각하는 정도(평가대상자에 의해 생기는 변량)가 곧 평가자들간의 일치수준을 나타내는 지표가 된다.

평가자간 일치를 나타내는 불일치와 변량은 대인지각에 있어서 서로 다른 측면을 반영한다(Jung, 1998).

이 두 측정치의 차이점을 알아보기 위해서 먼저 대인지각에 대한 사회적 관계 모형(The Social Relations Model, SRM: Kenny, 1981; Kenny와 La Voie, 1984)을 살펴볼 필요가 있다. 이 모형은 대인지각에 대한 수학적 무선효과 모형으로서 대인지각을 평가자 효과(perceiver effect), 평가대상자 효과(target effect), 그리고 (평가자와 평가대상자의) 관계 효과(relationship effect) 등 세 가지 효과에 의한 결과라고 가정한다. 즉, 이 모형은 대인지각에 있어서의 변량을 평가자 효과에 의한 변량, 평가대상자 효과에 의한 변량, 그리고 관계 효과에 의한 변량으로 분할한다. 평가자 효과는 한 평가자가 모든 평가대상자에 걸쳐서 평균적으로 보이는 반응수준을 의미한다. 예를 들면, 어떤 평가자는 모든 평가대상자들을 전반적으로 지적이라고 평가한 반면, 다른 평가자는 그들을 전반적으로 비지적이라고 평가한다면, "지적인" 특질차원에서 평가자 효과가 있다고 말한다. 평가대상자 효과는 모든 평가자들이 한 평가대상자에 대해 내린 평가치들의 평균을 반영한다. 예를 들어, 모든 평가자들이 어떤 평가대상자를 매우 지적이라고 평가하고 다른 평가대상자를 비지적이라고 평가한다면, "지적인" 특질차원에서 평가대상자 효과가 있다고 말한다. 앞에서 언급한 것처럼, 이 평가대상자 효과는 평가자간의 일치를 나타내는 측정치이다. 관계 효과는 특정 평가자가 특정 평가대상자를 특별히 다르게 지각하는 정도를 측정한다. 예를 들면, 어떤 평가자는 특정 평가대상자를 특히 지적이라고 평가할 수 있는데 - 비록 이 평가자는 다른 모든 평가대상자들을 전반적으로 비지적이라고 평가하고 또한, 다른 모든 평가자들이 이 특정 평가대상자를 비지적이라고 평가하더라도 - 이 때 관계 효과가 존재한다. 대인지각에서 관계 효과는 평가자 효과와 평가대상자 효과를 제외한 후에 남아있는 효과이다. 심리측정의 차원에서 관계 효과는 주어진 평가대상자에 대한 평가자들의 평가가 서로 일치하지 않음을 나타내는 바, 이는 곧 평가자간의 신뢰도가 낮음을 의미한다. 이와 같이 측정에서의 비신뢰성이 대인지각의 측면에서는 평가자와 평가대상자간의 상호작용 효과를 의미한다. 그러나, 평가자간의 비신뢰성이 평가대상자의 관계 효과 이외에 불완전한 평가도구의 사용에 기인할 수도 있기 때문에, 관계 효과를 정확히 측정하는 것은 좀 더 복잡한 절차를 요구한다(Kenny, 1994 참조).[1)]

평가자간 일치를 나타내는 불일치와 변량의 차이점을 알아보기 위하여, Jung(1998)은 주어진 특질내에서 각 평가대상자에 해당되는 불일치 점수, 평가자 효과, 평가대상자 효과, 그리고 관계 효과를 측정한 다음, 불일치 점수가 세 종류의 효과와 갖는 상관정도를 계산하였다. 두 연구 모두에서 불일치 점수는 평가대상자 효과와 큰 상관이 없었다. 20개의 성격특질에 대한 이 상관들의 평균이 -.06과 -.13이었다. 이 둘의 측정치가 거의 독립적이기 때문에, 실제 연구에서 평가자간 일치를 어떻게 측정하느냐에 따라 매개변인의 효과가 달라질 수 있다. 대신에, 불일치 점수는 평가자 효과(평균 r = .37과 .35) 및 관계 효과(평균 r = .72와 .80)와 정적상관이 있었다. 뿐만 아니라, 이 두 효과가 불일치 점수와 갖는 중다상관이 두 연구 모두에서 실질적으로 1.00이었는 바, 불일치 점수는 이 두 효과의 조합으로 구성됨을 알 수 있다. 이러한 결과는 불일치와 평가대상자에 따른 변량이 대인지각에 있어서 서로 다른 측면을 반영하고 있음을 잘 보여준다.

평가자간 일치에 대한 매개변인 접근방법

지금까지의 경험적 연구에 의하면, 평가대상자를 평가하는데 있어서 평가자간 일치수준은 결코 높지 않은 것으로 나타났다(Bourne, 1977; Dornbusch 등, 1965; Kenny, Horner, Kahsy와 Chu, 1992; Kunda와 Nisbett, 1986; Ross와 Nisbett, 1991). 이러한 결과에 대한 한 반응으로서, Bem과 Allen(1974)을 시작으로 그 이후 수십년 동안 많은 연구자들은 평가자간 일치수준을 높일 수 있는 변인들 즉, 평가자간 일치에 대한 매개변인(moderator variable)들을 탐색해 왔다. 주로 연구된 매개변인들은 행동일관성(Bem과 Allen,

1) 평가에 있어서의 비신뢰성과 관계 효과와의 관련성을 지적해 준 익명의 심사위원에게 감사드린다.

1974; Kenrick과 Stringfield, 1980; Mischel과 Peake, 1982; Paunonen과 Jackson, 1985), 관찰가능성(Kenrick과 Stringfield, 1980; Paunonen, 1989), 그리고 사회적 바람직성이다(Funder와 Dobroth, 1987; Hayes와 Dunning, 1997; John과 Robins, 1993). 이러한 매개변인들은 주로 두 가지 전통적인 방법에 의해 연구되어 왔다: 평가대상자에 기초한 접근방법(target-based approach or method)과 성격특질에 기초한 접근방법(trait-based approach or method). 또한, Jung과 Kenny, submitted는 평가대상자의 특질수준(target standing)에 기초한 새로운 접근방법을 최근에 제시하였다. 여기에서 특질수준은 어떤 평가대상자가 특정 성격특질을 가지고 있는 정도를 나타낸다. 예를 들면, 어떤 평가대상자가 전혀 성실하지 않으면 그의 특질수준은 "성실한" 특질차원에서 매우 낮고, 반대로 극도로 성실하다면 그의 특질수준은 이 특질차원에서 매우 높다. 이 절에서는 세 가지 접근방법의 특성과 각 접근방법을 통해 밝혀진 매개변인들의 효과에 대해 알아보겠다.

평가대상자에 기초한 접근방법

이 접근방법은 매개변인의 효과를 검증하는데 있어서 주어진 특정 성격특질내에서 평가대상자를 분석단위로 사용한다. 먼저, 각 평가대상자들에 해당되는 매개변인 점수를 측정하여 이들을 높은 점수집단과 낮은 점수집단으로 나눈다. 그 다음, 각 집단내에서 평가자간 일치를 계산하여 그 값을 집단간에 비교한다. 집단간 차이가 통계적으로 유의미하면 이 매개변인이 평가자간 일치에 영향을 미친다고 본다. 연구자들은 행동일관성을 이러한 접근방법으로 연구해 왔다. 행동일관성은 어떤 성격특질과 관련된 행동이 여러 상황이나 시간대에 걸쳐 변하지 않는 정도를 나타낸다. 지금까지의 연구들은 행동일관성의 매개변인효과를 분명하게 보여주지는 못하고 있다. 몇몇 연구에서는 낮은 일관성 집단보다 높은 일관성 집단에서 평가자간 일치가 더 컸다(Bem과 Allen, 1974; Kenrick과 Stringfield, 1980; Mischel과 Peake, 1982; Zuckerman 등, 1988). 그러나, 많은 다른 연구들은 평가자간 일치수준에 있어서 두 집단간의 유의미한 차이를 발견하지 못했다(Amelang과 Borkenau, 1986; Chaplin과 Goldberg, 1985; Cheek, 1982; Paunonen과 Jackson, 1985; Zuckerman 등, 1989).[2] 행동일관성과 함께 관찰가능성도 평가대상자에 기초한 방법으로 검증되어 왔다. 관찰 가능한 성격특질의 경우 이 특질과 관련된 행동이 관찰자에게 이용가능하며 다른 성격특질을 나타내는 행동과 분명히 구별된다(Kenrick과 Stringfield, 1980). 몇몇 연구를 제외하고(Cheek, 1982; Zuckerman 등, 1989), 대부분의 연구들은 높은 관찰가능성 집단에서 더 큰 평가자간 일치수준을 얻었다(Amelang과 Borkenau, 1986; Kammann, Smith, Martin과 McQueen, 1984; Kenrick과 Stringfield, 1980; Zuckerman 등, 1988).

평가대상자에 기초한 접근방법에 대해서 몇 가지 문제점이 제기되었다. 첫 번째 문제점은 이 접근방법으로 행동일관성과 관찰가능성을 검증할 때 생기는 한 오염변인과 관련이 있다. Paunonen과 그의 동료들은 이 오염변인에 대한 이론적인 근거와 경험적인 증거를 제시했다(Paunonen, 1989; Paunonen과 Jackson, 1985; Rushton, Jackson과 Paunonen, 1981). 이 연구자들에 따르면, 어떤 평가대상자가 하나의 양극특질척도(bipolar trait scale) 상에서 극단적일수록 -- 예를 들면 "부정직한-정직한" 차원에서 극단적으로 부정직하거나 극단적으로 정직한 경우 - 그의 행동은 더 관찰가능하고 더 일관적이다. 즉, 행동일관성과 관찰가능성은 평가대상자의 특질수준과 U자 형의 관계를 가지고 있다. 이 매개변인을 검증하기 위해 전체 평가대상자들을 행동일관성이나 관찰가능성 점수에서 높은 점수집단과 낮은 점수집단의 두 집단으로 나눌 때 종종 중앙치(median)를 사용한다. 이러한 경우, 특질수준에 있어서 양극단에 위치한 평가대상자들은 높

2) 한 심사위원이 지적했듯이, 평가자간 일치가 행동의 일관성을 나타내는 지표로 사용되어 왔는데, 행동일관성이 평가자간 일치에 미치는 매개변인 효과를 검증한다는 것이 비논리적인 것처럼 보인다. 그러나, 평가자간 일치에 대한 매개변인적 접근에서는 행동일관성에 개인간의 차이가 존재하고 이러한 차이가 평가자간 일치에 영향을 미친다고 본다. 따라서, 기존의 행동일관성에 대한 연구들은 행동일관성에 있어서의 개인차이를 경험적으로 검증한 연구라고 볼 수 있다.

은 점수집단에 그리고 상대적으로 중앙에 위치한 평가대상자들은 낮은 점수집단에 속하게 된다. 예를 들면, 매우 정직하거나 매우 부정직한 사람은 높은 행동일관성 집단에, 그리고 보통으로 정직하거나 부정직한 사람은 낮은 행동일관성 집단에 속하게 된다. 경험적인 연구에서 밝혀진 바처럼 자기평가와 동료평가는 어느 정도 일치한다. 따라서, 높은 점수집단의 경우 평가대상자가 갖는 특질수준의 범위(range)가 넓고 낮은 점수집단의 경우 이 범위가 좁게 된다. 또한, 평가자간 일치를 자아-동료간 평가의 상관관계로 측정할 때, 범위가 넓을수록 상관계수는 커지는 경향이 있다. 결국, 중앙치를 이용하여 두 집단을 만들 경우, 평가자간 일치는 높은 점수집단의 경우 과대평가될 수 있는 반면, 낮은 점수집단의 경우 과소 평가될 수 있어서 검증하고자 하는 매개변인의 효과가 과대 평가될 수 있다. 몇몇 기존의 연구들은 두 집단을 만들 때 이와 같은 오염변인을 통제하지 못했다 (e.g., Cheek, 1982; Kenrick과 Stringfield, 1980). 특질수준과 매개변인과의 관계에 기인한 이러한 오염을 통제할 수 있는 한가지 방법은 같은 특질수준을 가진 평가대상자들을 높은 점수집단과 낮은 점수집단으로 나누는 것이다. 이 경우 평가대상자의 특질수준 범위가 두 집단 모두 같기 때문에, 앞에서 언급한 오염변인을 통제할 수 있다.

두 번째 문제점은 평가대상자에 기초한 접근방식의 유용성과 관련이 있다. 앞서 살펴본 연구결과처럼, 이 접근방식은 행동일관성의 매개변인 효과를 증명하는데 거의 성공하지 못했다. 한가지 이유는 이 접근방식을 이용하는 대부분의 연구자(e.g., Bem과 Allen, 1974; Paunonen과 Jackson, 1985; Zuckerman 등, 1988, 1989)들이 앞서 기술한 Paunonen의 경우처럼 평가대상자의 특질수준을 하나의 오염변인으로 간주한다는 사실과 관련이 있다(Jung과 Kenny, submitted). 그러나, 평가대상자의 특질수준이 이 매개변인을 검증하는데 있어서 통제되야 할 오염변인인지, 아니면 이 매개변인을 결정하는데 큰 영향을 미치는 변인인지에 대한 경험적인 근거가 없다. Jung과 Kenny(submitted)는 평가대상자의 특질수준이 여러 매개변인에 미치는 효과를 알아보았다. 평가대상자들은 총 400명의 가상적인 인물들이었는데, 이들은 80개의 성격특질형용사와 다섯 개의 특질수준을 나타내는 부사의 서로 다른 조합으로 기술되었다. 실험참가자들은 각 평가대상자의 행동일관성, 관찰가능성 및 사회적 바람직성 등을 평가하였다. 예를 들면, 실험참가자들은 "보통으로 정직한" 사람의 행동이 "정직한" 특질차원에서 얼마나 "일관적인지", 혹은 "극단적으로 말이 없는" 사람의 행동은 "말이 없는" 특질차원에서 얼마나 "관찰가능한지"를 평가했다.

이들의 연구결과에 따르면, 평가대상자의 특질수준이 모든 매개변인들을 지각하는데 큰 영향을 미쳤다. 예를 들면, 한 평가대상자의 행동이 "정직한" 특질차원에서 얼마나 일관적인지는 그가 어느 정도 정직하냐에 따라 달랐다. 또한, 대부분의 성격특질에서 매개변인은 특질수준에 따라 비직선적으로 변했으며 비직선적인 정도는 성격특질마다 달랐다. 즉, 매개변인들은 평가대상자의 특질수준과 성격특질의 상호작용에 의해 결정되었다. 이 연구결과는 평가대상자의 특질수준이 매개변인에 큰 영향을 미치기 때문에, 특질수준이 매개변인에 있어서의 변량(variance)의 상당부분 혹은 전부를 설명할 수 있음을 시사한다. 따라서, 평가대상자에 기초한 접근방식처럼 평가대상자의 특질수준을 통제하는 것은 강력한 매개변인효과가 나타날 가능성을 크게 줄이는 것과 같다. 왜냐하면, 매개변인의 변량을 좌우하는 변인을 통제함으로써 남아있는 변량이 상대적으로 적어지고, 결국 이것은 통계적으로 유의미한 매개변인효과를 가져오는데 부정적인 요인으로 작용하기 때문이다.

평가대상자에 기초한 접근방법으로 행동일관성의 매개변인효과를 검증할 때 제기되는 세 번째 문제점은 이 매개변인을 측정하는데 있어서의 신뢰도와 관련된다. 행동일관성에 관한 대부분의 연구들은 평가대상자의 행동일관성을 측정하기 위해서 자기평가를 사용해 왔다. 평가대상자들이 어떤 특질차원에서 얼마나 일관적인지는 자신들이 가장 정확히 평가할 수 있다고 연구자들은 가정한다(Bem과 Allen, 1974; Kenrick과 Stringfield, 1980). 그러나, 이와 같은 자기

평가의 신뢰성에 대한 증거가 거의 희박하다. Paunonen과 Jackson(1985)은 평가대상자의 행동일관성에 대한 자기평가와 동료평가의 상관관계를 보고하였다. 총 20개의 성격특질에 대한 상관계수들은 -.23에서 .28 사이에 있었으며, 그 평균은 .04였다. 최근에 Krueger, Ham과 Linford(1996)도 자기평가와 동료평가간에 실질적으로 아무런 상관이 없음을 보였다. 18개의 성격특질형용사에 대한 상관계수의 평균은 .03이었다. 또한, 여러 매개변인 중에서 행동일관성에 대한 평가가 가장 신뢰롭지 못한 것으로 나타났다 (Jung과 Kenny, submitted). 이처럼 행동일관성에 대한 평가가 신뢰롭지 못할 때, 이 변인이 평가자간 일치수준에 강력한 매개효과를 미칠 것이라고 기대하기는 어렵다. 평가자간 일치는 다수의 평가자가 - 가령 평가대상자 자신과 동료 - 관련되어 있다. 따라서, 관련된 다수의 평가자들이 평가대상자의 행동일관성을 거의 비슷하게 평가할 때에만 행동일관성이 평가가간 일치수준에 영향을 미칠 수 있다.

성격특질에 기초한 접근방법

성격특질에 기초해서 매개변인의 효과를 검증하는 접근방법은 성격특질간 분석으로서 분석단위는 성격특질이다. 이 접근방법으로 어떤 매개변인을 검증하고자 할 때, 먼저 각 성격특질에 해당되는 매개변인 점수와 평가자간 일치수준을 측정한다. 그 다음, 성격특질을 분석단위로 하여 매개변인 점수와 평가자간 일치수준과의 상관계수를 구한다. 이 상관계수가 통계적으로 유의미하면 검증된 매개변인이 평가자간 일치수준에 신뢰할 만한 영향을 미친다고 본다. 관찰가능성은 앞서 언급한 평가대상자에 기초한 방식 이외에 성격특질에 기초한 방식으로도 연구되어 왔다. 몇몇 연구들을 제외하고(Funder, 1980; Paunonen, 1989), 성격특질간 분석방법을 이용한 대부분의 연구들은 성격특질의 관찰가능성이 높을수록 평가자간 일치수준이 증가함을 보고했다(Funder와 Colvin, 1988; Funder와 Dobroth, 1987; Hayes와 Dunning, 1997; John과 Robins, 1993; Kenrick과 Stringfield, 1980). 또한, 성격특질의 사회적 바람직성이 갖는 매개효과도 성격특질에 기초한 접근방법으로 연구되어 왔다. 어떤 성격특질의 사회적 바람직성은 그 특질을 가지고 있는 사람에 대한 타인의 호감이나 긍정적인 평가정도로 나타낸다(Hayes와 Dunning, 1997). 지금까지의 경험적인 연구결과에 따르면, 사회적 바람직성이 어떠한 형태의 매개적인 효과가 있는지는 매우 불분명하다. 사회적 바람직성은 평가자간 일치수준과 정적이며 선형적인 관계(Colvin, 1993; Funder와 Colvin, 1988; Funder와 Dobroth, 1987; Hayes와 Dunning, 1997), 종모양의 비선형적인 관계(John과 Robins, 1993), 혹은 둘간에 어떤 관계도 없는 것으로 밝혀졌다(Funder, 1980).

성격특질에 기초한 접근방법에 대해서도 몇 가지 문제점이 제기되었다. Paunonen(1989)은 어떤 성격특질차원의 관찰가능성을 양극특질척도로 측정할 때 이 척도의 양극단에 위치하는 성격특질을 모두 고려해야 한다고 주장했다. 일반적으로 양극단의 성격특질은 관찰가능성에 있어서 서로 다르기 때문에, 한 성격특질차원의 관찰가능성에 대한 지표로서 두 성격특질의 관찰가능성 점수를 대표할 수 있는 값(가령 평균값)을 사용해야 한다는 것이다. 예를 들면, "조용한-말하기를 좋아하는" 특질차원의 관찰가능성을 9점척도를 사용하여 측정한다고 가정하자. 이 때, "조용한"의 관찰가능성 점수가 6.50이며 "말하기를 좋아하는"의 점수가 8.40이라고 하면 이 특질차원의 관찰가능성 점수는 두 관찰가능성 점수의 평균인 7.45가 될 것이다. 만약 어느 한 성격특질만의 관찰가능성을 측정하여 그것을 이 특질차원의 관찰가능성 지표로 이용한다면 편파가 생길 수 있다. 관찰가능성에 대한 몇몇 기존 연구들이 이와 같은 오류를 범했다(Funder, 1980; John과 Robins, 1933, 연구 2; Kenrick과 Stringfield, 1980; Zuckerman 등, 1988). Paunonen의 주장은 단극특질척도(unipolar trait scale)를 사용하여 관찰가능성을 측정한 연구들은 이와 같은 문제점이 없음을 시사한다. 가령, 위의 예에서 "조용한" 특질차원에 대한 관찰가능성은 "조용한"의 점수인 6.50을 사용하는 것은 문제가 없음을 함축한다.

Paunonen(1989)의 이런 주장은 몇 가지 가정 하에

이루어지고 있다. 먼저, 양극특질차원의 관찰가능성이 평가대상자의 특질수준과 선형적인 관계를 가지고 있다는 가정이다. 이 가정이 틀리지 않는 한, 양극단의 두 성격특질에 해당되는 관찰가능성의 평균치는 이 특질차원의 관찰가능성을 대표하기에 적절하다. 그러나, 이 특질차원의 관찰가능성이 특질수준에 따라 복잡하게 비선형적으로 변한다면, 양극단의 관찰가능성만을 고려하는 것은 이 특질차원의 관찰가능성을 측정하는데 충분치 못하다. 단극특질척도에 기초하여 성격특질차원의 관찰가능성을 측정할 때도 마찬가지이다. Paunonen이 가정한 것처럼, 한 성격특질차원의 관찰가능성이 특질수준에 상관없이 변하기 않고 일정하면, 하나의 특질수준에 해당되는 관찰가능성만으로도 이 성격특질차원의 관찰가능성을 나타내는데 충분하다. 그러나, 관찰가능성이 특질수준에 따라 선형적이거나 비선형적으로 변할 경우 하나의 특질수준만 고려하는 것은 이 특질차원의 관찰가능성을 측정하는데 충분하지 않다. Jung과 Kenny(submitted)의 연구결과는 Paunonen의 이러한 가정을 지지하지 않았다. 앞서 언급한 것처럼, 대부분의 성격특질에 있어서 관찰가능성이나 사회적 바람직성은 특질수준에 따라 복잡하게 비선형적으로 변했다. 이러한 결과는 지금까지 성격특질간 분석방법을 이용한 많은 연구들이 매개변인들을 측정하는데 있어서 오류를 범했을 수 있음을 시사한다.

매개변인이 성격특질과 특질수준의 상호작용에 따라 변할 때(e.g., Jung과 Kenny, submitted), 매개변인을 성격특질간 분석방법으로 검증하기 위하여 성격특질에 해당되는 매개변인 점수를 측정하는 것은 쉬운 일이 아니다. 왜냐하면, 어느 특질수준에서 매개변인을 측정하느냐에 따라 각 성격특질의 매개변인 점수가 달라질 뿐만 아니라, 이 점수에 있어서 성격특질간의 차이도 달라지기 때문이다. 단극특질척도를 사용하여 매개변인를 측정할 때 세 가지 측정치가 있을 수 있다(Jung, 1998): 극단적 측정치(extreme measure), 규준적 측정치(normative measure), 그리고 평균적 측정치(average measure). 극단적 측정치는 기존의 많은 연구자들이 사용한 측정치로서, 각 성격특질차원에서 가장 높은 특질수준에 해당되는 매개변인 점수이다. 규준적 측정치는 극단뿐만 아니라 그 외 다수의 특질수준에 해당되는 매개변인 점수의 평균치이다. 매개변인이 특질수준에 따라 변할 때, 규준적 측정치는 극단적 측정치와는 달리 이러한 변화를 잘 반영해 준다. 마지막으로, 어떤 성격특질에 대한 평가자간 일치를 측정하는데 관련된 평가대상자들의 매개변인 점수를 구한 다음, 이 점수들을 평균하여 얻은 값이 평균적 측정치이다. 예를 들면, "정직한" 특질차원의 관찰가능성이 가장 낮은 특질수준(1)에서부터 가장 높은 특질수준(5)까지 9점 척도 상에서 각각 6.7, 6.3, 5.0, 6.9, 그리고 7.7이라고 가정하자. 이 경우, 극단적 측정치는 특질수준 5에 해당되는 7.7이며 규준적 측정치는 다섯 개의 특질수준에 해당되는 점수들의 평균치, 6.52이다. 또한, 어떤 특정 연구에서 특질수준이 1, 3, 그리고 4에 해당되는 평가대상자를 이용하여 평가자간 일치수준을 측정했다면, 평균적 측정치는 이들의 관찰가능성 점수 6.7, 5.0, 6.9 의 평균치인 6.2이다.

성격특질에 기초한 접근방법으로 매개변인을 검증할 때 위의 세 가지 측정치 중에서 어떤 것이 가장 적절한 측정치인가? 평균적 측정치가 다른 두 측정치보다 더 적절한데, 그 이유는 다음과 같다. 몇몇 이론적 연구와 경험적 결과에 따르면(Colvin, 1993; Funder, 1995; Kenny, 1991), 평가자간 일치정도는 평가대상자가 누구냐에 따라 달라질 수 있다. 즉, 표집된 평가대상자가 다르면 얻어진 평가자간 일치수준도 다를 수 있다. 또한, 평가대상자에 따라 매개변인 점수도 달라질 수 있는데, 왜냐하면 그들의 특질수준이 다르면 이에 상응하는 매개변인 점수도 변하기 때문이다(Jung과 Kenny, submitted; Paunonen과 Jackson, 1985; Paunonen, 1988). 따라서, 평가자간 일치수준과 매개변인 점수가 평가대상자에 따라 동시에 모두 변하기 때문에, 동일한 평가대상자에 기초해서 이 두 변인을 측정하는 것이 가장 타당하다. 즉, 평가자간 일치를 측정하는데 관여한 평가대상자를 이용하여 매개변인도 특정되어야 한다. 비록 규준적 측정치가 특질수준에 따른 매개변인의 변화를 잘 반영하긴 하지

만, 이 측정치는 평가대상자의 변화에 상관없이 변하지 않기 때문에 성격특질간 분석방법에 의한 매개변인의 검증에 적절치 못하다.

위의 세 측정치들간의 상관관계를 알아보기 위하여 Jung(1998)은 두 개의 경험적인 연구를 실시하였다. 첫 번째 연구에서 실험참가자들은 비디오를 통해 서로 모르는 남녀가 4분 정도 대화하는 것을 본 후, 그 중 한 명을 20개의 성격특질차원에 대해 평가하였다. 두 번째 연구에서는 실험참가자들과 상당히 친숙한 공인(public figures)들을 - 가령, 정치인, 연예인, 운동선수 등 - 평가대상자로 사용하였다. 실험참가자들은 이 평가대상자들을 앞의 연구에서 사용된 성격특질차원에 대해서 평가하였다. 그 결과, 극단적 측정치와 규준적 측정치가 평균적 측정치와 갖는 상관정도는 매개변인에 따라 다를 뿐만 아니라 연구마다 달랐다. 예를 들면, 첫 번째 연구에서는 세 개의 모든 매개변인(행동일관성, 관찰가능성, 그리고 사회적 바람직성)에 있어서 극단적 측정치와 평균적 측정치 간에 유의미한 상관이 없었다. 그러나, 두 번째 연구에서는 행동일관성과 사회적 바람직성의 경우 극단적인 측정치가 평균적 측정치를 잘 반영했지만(r = 각각 .57과 .68), 관찰가능성의 경우는 그렇지 않았다(r = -.01). 이와 같은 결과는 기존의 연구에서 사용된 극단적 측정치가 매개변인을 검증하는데 가장 타당한 평균적 측정치를 안정적으로 반영하지 못하고, 따라서 매개변인 검증에 적절하지 않음을 시사한다.

Jung(1998)은 성격특질간 분석방법으로 행동일관성, 관찰가능성, 그리고 사회적 바람직성의 매개변인 효과를 평균적 측정치를 사용하여 검증하였다. 평가자간 일치는 동료-동료간 일치수준으로 측정되었으며, 불일치와 변량의 두 측정치가 사용되었다. 평가자와 평가대상자 사이의 친숙도가 낮은 조건에서는 관찰가능성만이 변량 측정치에 통계적으로 유의미한 매개변인 효과를 보였다. 즉, 성격특질이 더 관찰가능할수록 평가자간 일치정도는 증가하였다. 그러나, 어떠한 매개변인도 불일치 측정치에는 유의미한 매개변인 효과를 미치지 못했다. 평가자가 평가대상자를 잘 알고 있을 때의 결과는 위의 결과와 매우 달랐다. 관찰가능성의 변량 측정치에 대한 매개변인 효과는 없었으며, 그 대신 행동일관성과 사회적 바람직성은 불일치로 측정된 평가자간 일치에 유의미한 효과를 나타냈다. 특히, 사회적 바람직성은 평가자간 일치와의 관계에서 선형적인 경향과 U자형의 비선형적인 경향성을 동시에 가지고 있었다. 즉, 성격특질이 사회적으로 바람직하거나 혹은 사회적 바람직성에서 극단적일 때 평가자간 불일치는 더 적었다. 이 연구결과는 사회적 바람직성의 매개효과에 대한 John과 Robins (1993)의 예측 즉, 성격특질이 사회적 바람직성에서 극단적일수록 평가자간 일치수준은 낮을 것이라는 예측을 지지하지 않았다. 또한, 이 연구결과는 평가자간 일치를 어떻게 측정하느냐가 매개변인 효과가 달라질 수 있음을 잘 보여준다.

특질수준에 기초한 접근방식

일반적으로 어떤 성격특질에 해당되는 매개변인 점수는 평가대상자의 특질수준에 따라 변한다(Jung과 Kenny, submitted). 이와 같은 사실은 평가자간 일치에 미치는 매개변인의 효과를 검증할 수 있는 새로운 방법을 가능케 해준다. 이 방법은 한 성격특질내에서 평가대상자를 분석단위로 사용한다. 각 평가대상자에 대한 평가자간의 일치수준이 그 평가대상자의 특질수준에 해당되는 매개변인 점수와 공변(covary)하는 지를 알아봄으로써 매개변인의 효과를 검증한다. 특질수준과 매개변인과의 관계 및 특정 평가대상자의 특질수준을 앎으로써 이 평가대상자에 해당되는 매개변인 점수를 결정할 수 있다. 매개변인 점수가 평가자간 일치수준과 통계적으로 유의미한 상관관계가 있을 때, 이 매개변인이 평가자간 일치수준에 신뢰할 만한 영향을 미친다고 본다. 특질수준에 기초한 접근방법은 앞에서 살펴 본 평가대상자에 기초한 접근방법과 마찬가지로 각 성격특질내의 평가대상자를 분석단위로 사용한다. 그러나, 이 두 접근방법간의 근본적인 차이점은 특질수준을 어떻게 취급하느냐에 달려 있다. 특질수준에 기초한 접근방법은 매개변인의 효과를 평가대상자의 특질수준에 기초하여 검증하지만, 기존의 평가대상자에 기초한 접근방법은 평가대상자

의 특질수준을 하나의 오염변인으로 간주하여 통제한다.

특질수준에 기초하여 매개변인 효과를 검증하는 방법을 예를 통해 알아보자. 사회적 바람직성이 "정직한" 특질차원에서 평가자간 일치수준에 미치는 영향을 알아보기 위해서, 먼저 이 특질차원에서 사회적 바람직성이 특질수준에 따라 어떻게 변하는 지를 측정한다. 다수의 평가자가 어떤 평가대상자를 평가하고, 가령 그 평가치들의 평균을 이 평가대상자의 특질수준을 나타내는 지표로 이용한다. 결국, 이 평가대상자의 특질수준에 해당되는 사회적 바람직성을 측정할 수 있다. 또한, 이 평가대상자에 대한 평가자간 일치수준을 측정해야 하는데, 가령 평가치들의 표준편차로 평가자간 일치수준을 나타낼 수 있다. 마지막으로, 평가대상자들에 걸쳐서 평가자간 일치수준과 사회적 바람직성 점수와의 상관관계를 검증함으로써, 사회적 바람직성이 "정직한" 특질차원에서 평가자간 일치에 미치는 영향을 알 수 있다.

Jung(1998)은 특질수준에 기초한 접근방법을 이용하여 매개변인들이 평가자간 일치수준에 미치는 효과를 알아보았다. 이 효과는 평가자와 평가대상자가간에 친숙도가 낮은 조건 및 높은 조건에서 각각 검증되었다. 각 평가대상자는 20명의 평가자로부터 평가를 받았는데, 20개의 평가치들의 평균으로 평가대상자의 특질수준을, 그리고 이 평가치들의 표준편차로 평가자간의 일치정도를 측정하였다. 또한, 주어진 특질수준에 해당하는 매개변인 점수는 Jung과 Kenny (submitted)의 연구결과에 근거해서 측정하였다. 매개변인이 평가자간 일치에 미치는 효과는 높은 친숙도 조건에서 더 컸다. 낮은 친숙도 조건에서, 평가대상자의 행동이 일관성일수록 혹은 평가대상자가 극단적으로 바람직하거나 극단적으로 바람직하지 않을수록 평가자간 불일치가 감소하였다. 그러나, 관찰가능성과 사회적 바람직성의 매개변인효과는 매우 약했다. 높은 친숙도 조건에서는 모든 매개변인이 평가자간 일치에 상당한 효과를 미치고 있었으며, 특히 행동일관성과 사회적 바람직성에서의 극단성이 가지는 매개변인 효과가 컸다.

이러한 Jung(1998)의 연구결과들은 평가자간 일치에 관한 연구에서 몇몇 함의를 가지고 있다. 먼저, 행동일관성의 매개변인 효과가 크게 나타난 점에 주목할 필요가 있다. 앞서 언급한 것처럼, 평가대상자에 기초한 접근방법은 행동일관성의 효과를 분명하게 보여주지 못했다(Amelang과 Borkenau, 1986; Chaplin과 Goldberg, 1985; Cheek, 1982; Paunonen과 Jackson, 1985; Zuckerman 등, 1989). 이는 평가대상자의 특질수준을 고려하는 것이 행동일관성을 검증하는데 중요함을 시사한다. 또한, 특질수준에 기초해서 얻은 관찰가능성의 효과는 성격특질에 기초한 방법으로 얻은 것과는 사뭇 다르다. 낮은 친숙도 조건에서는 성격특질에 기초한 접근방법이, 높은 친숙도 조건에서는 특질수준에 기초한 방법이 관찰가능성의 효과를 증명하는데 더 효과적이었다. 이는 관찰가능성을 어떠한 방식으로 검증하느냐에 따라 그 효과가 달라질 수 있음을 잘 보여준다. 마지막으로, 특질수준을 기초로 한 분석결과도 성격특질을 기초한 분석결과와 마찬가지로 사회적 바람직성의 매개효과에 대한 John과 Robins (1993)의 예측을 - 극단적으로 바람직하거나 비바람직할 때 평가자간 일치가 적을 것이라는 예측 - 지지하지 않았다.

이러한 매개변인 효과의 대부분은 평가대상자의 극단성(target extremity)에 기인하는 것으로 나타났다. 평가대상자의 극단성은 그의 특질수준이 성격특질척도의 중앙으로부터 벗어난 절대거리로 나타낸다. 어떤 사람이 "정직한" 특질을 측정하는 9점척도 상에서 7점을 얻었다면, 그의 극단성은 7 - 5 = 2이다. Jung (1998)의 두 연구 모두에서 평가대상자가 주어진 성격특질차원에서 극단적일수록 평가자간 불일치는 더 적었다. 평가대상자의 극단성 효과는 사회적 범주화(social categorization)에 있어서의 전형성(prototypicality) 개념으로 설명될 수 있다. 많은 연구자들은 성격특질들이 사회적 행동을 구분 짓는 자연적인 범주로서 그 경계가 불분명하고 다만 그 개념에 가장 적합한 전형(prototype)이 있다고 생각한다(Barsalou, 1985; Borkenau, 1990; Buss와 Craik, 1980; Cantor와 Mischel, 1977, 1979). Borkenau(1985)는 극단적 행동이 성격특

질의 전형성을 결정하는 핵심 요인임을 밝혔으며, 기존의 연구들은 평가대상자의 행동이 더 전형적일수록 평가자간 일치가 높을 수 있음을 시사한다(Buss와 Craik, 1983; Cantor와 Mischel, 1977, 1979; Cantor, Mischel과 Schwartz, 1982). 즉, 사람들은 전형적인 행동에 대해 그 행동과 관련된 성격특질을 보다 쉽고 분명하게 판단할 수 있으며 또한, 이러한 행동의 평가에 있어서 평가자간 일치가 더 높을 수 있다.

특질수준에 기초해서 매개변인의 효과를 검증해 본 결과, 대부분의 매개변인들이 평가자간 일치에 효과가 있는 것으로 나타났다. 앞서 언급한 것처럼, 성격 심리학자들은 평가자간 일치를 성격특질이 존재한다는 증거로 생각해 왔다(Bourne, 1977; Kenrick과 Funder, 1988). 그렇다면 매개변인의 효과가 존재한다는 사실은 곧 성격특질이 보편적으로 존재하는 것이 아니라 어떤 한정된 상황에서만 존재한다는 증거로 볼 수 있는가? Bem과 Allen(1974)에 따르면, 사람들간에 자신에게 중요한 성격특질은 같지 않다. 또한, 사람들은 자신에게 중요한 성격특질차원에서 행동적으로 일관적이지만, 중요하지 않은 성격차원에 대해서는 일관적이지 않을 뿐만 아니라 그와 같은 성격특성을 가지고 있지 않을 수 있다는 것이다. 이런 주장에 기초해서 보면, Jung(1998)이 밝힌 행동일관성의 매개변인 효과는 어떤 주어진 성격특질이 특정 사람에게만 존재한다는 증거로 볼 수 있다. 그러나, 매개변인 효과가 반드시 성격특질의 보편성을 부정하는 증거로만 볼 수는 없을 것 같다. 왜냐하면, 개인의 행동이 그 개인의 성격뿐만 아니라 행동이 발생하는 환경과의 상호작용에 의해 야기된다고 할 때(Ross와 Nisbett, 1991), 평가자간 일치가 낮은 평가대상자의 경우 환경이 그 사람의 행동에 더 큰 영향을 미친다고 볼 수 있기 때문이다. 또한, 평가자간의 일치가 낮은 평가대상자는 일반적으로 주어진 성격차원에서 덜 극단적이라는 점은(Jung, 1998) 이들에게 주어진 성격특질이 부재한다는 것을 지지하지는 않는다. 뿐만 아니라, 평가자간의 일치만으로 성격특질의 존재여부를 결정짓는 증거로 여기는 것은 무리일 수 있다. 왜냐하면, 성격심리학자들이 생각하는 것처럼 평가자간 일치가 단순히 평가대상자의 행동일관성에 의해서만 결정되는 것이 아니기 때문이다(Kenny, 1991).

매개변인 접근방식의 과제

대인지각 분야에서 매개변인에 관한 대부분의 연구들은 주로 평가자가 평가대상자의 성격을 평가할 때 그 평가가 얼마나 일치하는 가에 관련된다. 이러한 매개변인 접근방법은 몇몇 측면에서 더 검증되어야 할 필요성이 있다. 예를 들면, 매개변인에 관한 비교문화연구가 거의 없기 때문에, 매개변인의 지각 및 효과가 문화간에 차이가 있는지를 알 수 없다. 또한, 평가자간 일치에 대한 매개변인 접근방법을 대인지각의 다른 측면들에 응용할 수 있는지 그 가능성을 타진해 볼 필요가 있다. 따라서, 본 절에서는 매개변인 접근방법에 있어서의 앞으로의 과제들을 고찰해 보겠다.

매개변인지각에 있어서 문화적 차이

다른 문화권에 속하는 평가대상자를 판단하는데 있어서 평가자간 일치에 관한 연구는 주로 얼굴표정에 나타나는 정서, 거짓말에 대한 탐지, 동안(baby-facedness), 그리고 신체적 매력에 대한 지각과 관련된다(Albright 등, 1997). 예를 들면, 얼굴표정에 나타나는 정서에 대한 지각은 한 문화내에서 뿐만 아니라 문화간에서도 평가자간 일치가 높은 것으로 나타났다(e.g., Gordon, Zukas와 Chan, 1982; Krauss, Curran과 Ferlinger, 1983). 거짓말에 대한 탐색도 문화간에 있어서 비슷한 외부단서를 이용하여 이루어지며(Bond, Omar, Mahmoud와 Bonser, 1990), 동안과 관련된 성격특질을 지각하는데 있어서도 문화간에 일치도가 높았다(McArthur과 Berry, 1987; Zebrowitz와 Montepare, 1992). 신체적 매력에 대한 지각 역시 문화적으로 매우 비슷했으며(Cunningham, Roberts, Barbee, Druen과 Wu, 1995), 그런 지각에 있어서 평가자간 일치가 높았다(Mart와 Harling, 1985; Thakerar와 Iwawaki, 1979; Bernstein, Tsai-Ding과 McClelland, 1982; Morse, Cruzen과 Reis, 1976). 평가자가 다른 문화권에 속하

는 평가대상자의 성격을 평가할 때에도 유의미한 평가자간 일치도가 있었다(Albright 등, 1997). 요약하면, 지금까지의 연구결과에 따르면, 대인지각에 있어서 문화간에 큰 차이가 없고 평가자간에 높은 일치도가 있었다.

그러나, 평가자간 일치에 대한 매개변인을 지각하는 것은 문화간에 차이가 있을 수 있다. 모든 사회체계는 그 사회를 구성하는 개인의 행동이 규범(norm)에 부합되도록 영향력을 행사한다(Berry, Poortinga, Segall과 Dasen, 1992, p. 43). 또한, 그 사회의 가치체계는 사회적 규범과 매우 밀접한 관계를 가지고 있어서, 개인의 행동이 규범에 부합되면 사회적으로 바람직한 것으로, 어긋나면 바람직하지 않은 것으로 판단된다. 결국, 사회적 규범이 다르면 사회적으로 바람직한 행동의 기준이나 성격특질의 사회적 바람직성에 대한 판단도 달라질 가능성이 크다. 예를 들면, 집단주의적 문화에서는 사교성이나 다른 사람과의 조화, 개인간의 안전한 사회적 관계유지에 가치를 부여하는 반면, 개인주의적 문화에서는 성취, 경쟁심, 자립, 평등 등에 가치를 부여한다(Triandis, 1995). 사회적 바람직성 뿐만 아니라 행동의 일관성에 대한 지각도 문화간에 차이가 있을 수 있다. 예를 들면, 사회체계가 그 구성원들로 하여금 규범을 지키도록 행사하는 사회적 압력의 수준에 있어서 문화간에 차이가 있을 수 있다. 가령, 어떤 문화권의 사회는 다른 문화권의 사회보다 사회적 규범을 엄격하게 실행하고 규범으로부터의 이탈행위에 대한 처벌이 더 심할 수 있다(Pelto, 1968). 이와 같은 요인들은 사회적으로 바람직하지 않은 행동에 대한 귀인에 영향을 미칠 수 있고(Newman, 1993), 이러한 행동을 행위자의 내부성향에 귀인 하느냐 아니면 외부 상황적 요인에 귀인하느냐에 따라 그런 행동의 일관성에 대한 지각도 달라질 수 있다.

본 연구자는 예비적인 자료로서 Jung과 Kenny (submitted)의 연구를 한국대학생을 대상으로 반복 검증하였다. 그 결과, 매개변인들에 대한 지각에 있어서 몇 가지 차이점을 발견하였다. 특질수준과 매개변인과의 관계를 평가하는데 있어서 행동일관성의 경우 문화간에 차이가 가장 컸다. 즉, 이러한 평가에 있어서 문화간의 유사성을 나타내는 상관계수는 행동일관성의 경우 .43, 관찰가능성의 경우 .71, 그리고 사회적 바람직성의 경우 .70이었다. 또한, 사회적 바람직성에 대한 지각에 있어서도 차이점이 있었다. 예를 들면, 두 문화권의 대학생들은 모두 긍정적 특질차원(가령, "정직한")과 중립적 특질차원(가령, "말이 없는")에서 덜 극단적인 사람(가령, 매우 정직한 사람)을 극단적인 사람(가령, 지나치게 정직한 사람)보다 사회적으로 더 바람직한 것으로 지각했다. 그러나, 한국대학생의 경우 이러한 경향성이 훨씬 더 강하게 나타났다. 이는 문화가 매개변인 지각에 영향을 미칠 수 있고, 결국 매개변인의 효과가 문화간에 다를 수 있음을 시사해 준다. 따라서, 매개변인에 대한 문화간 차이에 대한 좀 더 체계적인 후속연구가 필요하다.

친숙도(acquaintance)가 매개변인효과에 미치는 영향

평가자간 일치에 대한 연구에서 자주 논란이 되는 주제는 친숙도가 평가자간 일치에 미치는 영향이다. 직관적으로 볼 때, 평가자들이 평가대상자를 잘 알수록 평가자간 일치가 높을 것같다. Funder의 실제적 정확성 모형(The Realistic Accuracy Model: RAM, 1995)도 이러한 직관적인 생각과 같은 예측을 하고 있으며, 경험적인 측면에서도 이 예측은 많은 지지를 얻었다(Colvin과 Funder, 1991; Funder와 Colvin, 1988; Norman과 Goldberg, 1966; Paulhus와 Bruce, 1992; Paunonen, 1989). 그럼에도 불구하고, 또 다른 많은 연구들은 높은 친숙도가 평가자간 일치의 증가를 가져오지 않음을 보여주고 있다. 예를 들면, Kenny 등(1994)은 사회적 관계모형(SRM)을 이용하여 친숙도 효과를 검증한 연구들을 종합적으로 살펴보았다. 그 결과, 높은 친숙도가 몇몇 조건에서는 평가자간 일치수준을 증가시키지 않았으며, 심지어 어떤 경우에는 그 일치수준을 감소시켰다. 특히, 이들이 살펴본 12개의 장기종단연구와 최근에 발표된 Park, Kraus와 Ryan(1997)의 장기종단연구에서도 높은 친숙도가 평가자간 일치수준을 증가시키지 않았다. 이와 같은 연

구결과들을 종합해 볼 때, 친숙도가 평가자간 일치수준에 미치는 영향은 여전히 불분명하다.

앞서 살펴본 것처럼, 지금까지 친숙도에 관한 연구는 주로 친숙도가 평가자간 일치에 미치는 영향에 한정되어 있다. 그러나, 평가자간 일치에 대한 매개변인들에 대한 연구가 활발한 지금, 매개변인의 효과가 친숙도에 따라 달라질 수 있는지에 대한 연구도 필요하다. 높은 친숙도 조건에서 매개변인의 효과가 강하게 나타날 것이라는 가정 하에서, 대부분의 연구자들은 평가자와 평가대상자가 서로 잘 알고 있는 상황에서 매개변인들을 검증해 왔다. 그러나, 매개변인 효과가 높은 친숙도 조건에서 반드시 더 강한 것은 아니다. 예를 들어, Paunonen(1989)은 평가자간 일치가 친숙도와 관찰가능성의 상호작용에 의해 결정됨을 보였다. 낮은 친숙도 조건에서 평가자간 일치수준은 성격특질의 관찰가능성이 높을수록 증가하였지만, 높은 친숙도 조건에서는 이와 같은 관찰가능성의 매개효과가 나타나지 않았다. 또한, Jung(1998)도 Paunonen의 연구결과와 마찬가지로 낮은 친숙도 조건에서만 성격특질의 관찰가능성의 매개효과를 얻었다.

앞서 살펴본 두 연구(Jung, 1998; Paunonen, 1989)가 친숙도에 따라 매개변인의 효과가 달라질 수 있음을 보여 주지만, 이 연구들은 몇몇 한계점들을 가지고 있다. Paunonen(1989)은 관찰가능성만 검증함으로써 다른 매개변인의 효과가 친숙도에 따라 어떻게 달라질 수 있는지에 대한 해답을 주지 못한다. Jung(1998)은 두 연구를 통해 여러 매개변인들의 효과와 친숙도와의 관계를 다루었지만, 두 연구간의 차이가 친숙도 이외의 다른 요인에 기인할 수도 있다. 예를 들면, 두 연구에서 서로 다른 평가대상자를 사용하였기 때문에 두 연구간의 차이가 평가대상자의 차이 때문일 수도 있다. 따라서, 이러한 문제점을 극복하고 친숙도와 매개변인 효과와의 관계를 보다 분명하게 제시해 줄 수 있는 장기종단연구들이 필요하다.

매개변인 접근방식의 활용

매개변인 접근방법은 평가대상자의 성격을 평가하는데 있어서의 평가자간 일치뿐만 아니라 대인지각의 상호성(reciprocity)이나 초지각(meta-perception)을 연구할 때에도 이용할 수 있다. 상호성은 평가자와 평가대상자가 서로를 똑같이 평가하느냐의 문제이다. 예를 들면, 평가자와 평가대상자가 서로를 매우 지적이라고 평가한다면, "지적" 특질차원에서 높은 상호성이 있다고 말할 수 있다. 상호성에 관한 연구결과들을 살펴보면(Kenny, 1994), 호감(liking)에 있어서 두 사람간의 상호성이 상당히 높은 반면, 성격특질차원에 있어서는 상호성이 없거나 매우 약한 것으로 나타났다. 그럼에도 불구하고, 몇몇 연구(e.g., Park과 Judd, 1989; Reno와 Kenny, 1992)들은 상호성에 있어서 성격특질간에 차이가 있을 수 있음을 시사해 준다.

초지각은 다른 사람이 나에 대해 어떻게 생각하는가에 대한 나의 지각으로서, 초지각의 정확도는 나의 초지각이 다른 사람의 나에 대한 실제 지각과 얼마나 일치하는가의 문제이다. 예를 들면, 어떤 사람이 나를 지적이라고 평가한다고 내가 생각하는 것처럼 실제로 그 사람이 나를 지적이라고 평가한다면, 나의 초지각은 정확하다고 말할 수 있다. 이러한 초지각의 정확도는 성격특질마다 다를 수 있다(e.g., Kenny와 Depaulo, 1993). 예를 들면, 어떤 성격특질차원에 대한 상대방의 나에 대한 평가는 다른 성격차원보다 여러 비언어적 단서(nonverbal cues)를 통해 더 잘 표출될 수 있다. 특정 상황에서는 비언어적 단서에 대한 지각이 정확하기 때문에(Depaulo와 Rosenthal, 1982), 결국 초지각에 대한 정확도에 있어서 성격특질간에 차이가 있을 수 있다. 따라서, 이와 같이 대인지각의 여러 측면에서 성격특질간의 차이를 가져오는 요인이 지금까지 연구해 온 매개변인들의 특성과 관련이 있는지 연구해 볼 수 있겠다.

결 론

많은 연구자들은 지난 수십년 동안 매개변인이 평가자간 일치에 미치는 영향을 연구해 왔지만, 대부분의 매개변인의 경우 이들이 어떠한 영향을 미치는지 여전히 불분명하다. 한가지 이유는 매개변인을 부정확하게 측정한데 기인할 수 있다. 가령, 성격특질을

분석단위로 사용한 연구들은 성격특질에 해당되는 매개변인 점수를 측정하기 위하여 한 두개의 극단적 특질수준만을 고려하였다. 그러나, Jung과 Kenny (submitted)의 연구가 보였듯이 이러한 극단적 측정치가 매개변인 효과를 검증하는데 적절히 않을 수 있다. 또 다른 이유는 매개변인을 검증하기 위하여 사용된 방법론상의 문제에 기인할 수 있다. 예를 들면, 기존의 평가대상자에 기초한 연구방법은 행동일관성의 매개변인효과를 잘 밝히지 못했지만, 특질수준에 기초한 연구(Jung, 1998)는 행동일관성이 강력한 매개변인임을 분명히 보여준다. 결국, 매개변인의 효과는 매개변인의 측정과 그 효과의 검증방법에 크게 좌우될 수 있다. 앞으로 매개변인의 접근방법을 확장하는 시도가 필요하며, 특히 최근에 개발된 특질수준에 기초한 접근방법의 타당성과 유용성을 검증할 필요가 있다.

참 고 문 헌

Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X., & Winquist, L. (1997). Cross-cultural consensus in personality judgments. *Journal of Personality and Social Psychology, 72,* 558-569.

Amelang, M., & Borkenau, P. (1986). The trait concept: Current theoretical considerations, empirical facts, and implications for personality inventory construction. In A. Angleitner & J. S. Wiggins (Eds.), *Personality assessment via questionnaires: Current issues in theory and measurement* (pp. 7-34). Berlin: Springer-verlag.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 629-654.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review, 81,* 506-520.

Bem, D. J., & Funder, D. C. (1978). Predicting more of the people more of the time: Assessing the personality of situations. *Psychological Review, 85,* 485-501.

Bernstein, I. H., Tsai-Ding, L., & McClelland, P. (1982). Cross- vs. within racial judgments of attractiveness. *Perception & Psychophysics, 32,* 495-503.

Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (1992). *Cross-cultural psychology: Research and applications.* New York: Cambridge University Press.

Bond, C. F., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189-204.

Borkenau, P. (1990). Traits as ideal-based and goal-derived social categories. *Journal of Personality and Social Psychology, 58,* 381-396.

Bourne, E. (1977). Can we describe an individual's personality? Agreement on stereotype versus individual attributes. *Journal of Personality and Social Psychology, 35,* 863-872.

Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90,* 105-126.

Buss D. M., & Craik, K. H. (1980). The frequency concept of disposition: Dominance and prototypically dominant acts. *Journal of Personality, 48,* 379-392.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of personality and Social Psychology, 35,* 38-48.

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp.

3-52). New York: Academic Press.

Cantor, N., Mischel, W., & Schwartz, J. (1982). Social knowledge: Structure, content, use, and abuse. In A. M. Isen & A. H. Hastorf (Eds.), *Cognitive social psychology*. New York: Elsevier North-Holland.

Chaplin, W. F., & Goldberg, L. R. (1985). A failure to replicate the Bem and Allen study of individual differences in cross-situational consistency. *Journal of Personality and Social Psychology, 47*, 1074-1090.

Cheek, J. M. (1982). Aggregation, moderator variables, and the validity of personality tests: A peer-rating study. *Journal of Personality and Social Psychology, 43*, 1254-1269.

Colvin, C. R. (1993). "Judgable" people: Personality, behavior and competing explanations. *Journal of Personality and Social Psychology, 64*, 861 -873.

Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology, 60*, 884-894.

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology, 68*, 261-279.

DePaulo, B. M., & Rosenthal, R. (1982). Measuring the development of sensitivity to nonverbal communication. In C. E. Izard(Ed.), *Measuring emotions in infants and children* (pp. 208-247). New York: Cambridge University Press.

Dornbusch, S. M., Hastorf, A. H., Richardson, S. A., Muzzy, R. E., & Vreeland, R. S. (1965). The perceiver and the perceived: Their relative influence on categories of interpersonal perception. *Journal of Personality and Social Psychology, 1*, 434-440.

Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality, 48*, 473-493.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*, 149-158.

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology, 60*, 773-794.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*, 409-418.

Gordon, I. E., Zukas, M., & Chan, J. (1982). Responses to schematic faces: A cross-cultural study. *Perceptual and Motor Skills, 54*, 201-202.

Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology, 72*, 664-677.

Higgins, E. T., & Bargh, J. A. (1987). Social cognition and social perception. *Annual Review of Psychology, 38*, 369-425.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality, 61*, 523-551.

Jung, T. (1998). *A new look at moderator variables of agreement: The role of target standing.* Manuscript for doctoral dissertation, University of Connecticut.

Jung, T., & Kenny, D. A. (submitted). *The role of target standing in measuring the moderators of agreement.*

Kammann, R., Smith, R., Martin, C., & McQueen, M. (1984). Low accuracy in judgments of others' psychological well-being as seen from a phenomenological perspective. *Journal of Personality, 52,* 107-123.

Kenny, D. A. (1981). Interpersonal perception: A multivariate round robin analysis. In M. B. Brewer & B. E. Collins (Eds.), *Scientific inquiry and the social sciences: A volume in honor of Donald T. Campbell* (pp. 288-309). San Francisco: Jossey-Bass.

Kenny, D. A. (1991). A general model of consensus and accuracy in interpersonal perception. *Psychological Review, 98,* 155-163.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: The Guilford Press.

Kenny, D. A., & DePaulo, B. M. (1993). Do people know how others view them?: An empirical and theoretical account. *Psychological Bulletin, 114,* 145-161.

Kenny, D. A., & La Voie, L. (1984). The Social Relations Model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 142-182). Orlando, FL: Academic Press.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the Big Five. *Psychological Bulletin, 116,* 245-258.

Kenny, D. A., Horner, C., Kahsy, D. A., & Chu, L. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62,* 88-97.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. *American Psychologist, 43,* 23-34.

Kenrick, D. T., & Stringfield, D. O. (1980). Personality traits and the eye of the beholder: Crossing some traditional philosophical boundaries in the search for consistency in all of the people. *Psychological Review, 87,* 88-104.

Krauss, R. M., Curran, N. M., & Ferlinger, N. (1983). Expressive conventions and the cross-cultural perception of emotion. *Basic and Applied Social Psychology, 4,* 295-305.

Krueger, J., Ham, J. J., & Linford, K. M. (1996). Perceptions of behavioral consistency: Are people aware of the actor-observer effect? *Psychological Science, 7,* 259-264.

Kunda, Z., & Nisbett, R. E. (1986). The psychometrics of everyday lie. *Cognitive Psychology, 18,* 195-224.

Maret, S. M., & Harling, C. A. (1985). Cross-cultural perceptions of physical attractiveness: Ratings of photographs of whites by Cruzans and Americans. *Perceptual and Motor Skills, 60,* 163-166.

McArthur, L. Z., & Berry, D. S. (1987). Cross-cultural agreement in perceptions of babyfaced adults. *Journal of Cross-Cultural Psychology, 18,* 165-192.

McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology, 43,* 293-303.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W., & Peake, P. K. (1982). Beyond Déjà Vu in the search for cross-situational consis-

tency. *Psychological Review, 89,* 730-755.

Morse, S. T., Gruzen, J., & Reis, H. (1976). The "eye of the beholder": A neglected variable in the study of physical attractiveness. *Journal of Personality, 44,* 209, 225.

Newman, L. S. (1993). How individualists interpret behavior: Idiocentrism and spontaneous trait inference, *Social Cognition, 11,* 243-269.

Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality structure. *Journal of Personality and Social Psychology, 4,* 681-691.

Park, B. (1986). A method for studying the development of impressions of real people. *Journal of Personality and Social Psychology, 51,* 907-917.

Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology, 56,* 493-505.

Park, B., Kraus, S., & Ryan, C. S. (1997). Longitudinal changes in consensus as a function of acquaintance and agreement in liking. *Journal of Personality and Social Psychology, 72,* 604-616.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816-824.

Paunonen, S. V. (1988). Trait relevance and the differential predictability of behavior. *Journal of Personality, 56,* 599-619.

Paunonen, S. V. (1989). Consensus in personality judgment: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology, 56,* 823-833.

Paunonen, S. V., & Jackson, D. N. (1985). Idiographic measurement strategies for personality and prediction: Some unredeemed promissory notes. *Psychological Review, 92,* 486-511.

Pelto, P. J. (1968). The difference between "tight" and "loose" societies. *Transaction, April,* 37-40.

Reno, R. R., & Kenny, D. A. (1992). Effects of self-consciousness and social anxiety on self-disclosure among unacquainted individuals: An application of the Social Relations Model. *Journal of Personality, 60,* 79-94.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology.* New York: McGraw-Hill.

Rushton, J. P., Jackson, D. N., & Paunonen, S. V. (1981). Personality: Nomothetic or idiographic? A response to Kenrick and Stringfield. *Psychological Review, 88,* 582-589.

Shweder, R. A., (1975). How relevant is an individual difference theory of personality? *Journal of Personality, 43,* 455-484.

Triandis, H. C. (1995). *Individualism and collectivism.* Bolder, CO: Westview Press.

Wiggins, J., (1973). *Personality and prediction: Principles of personality assessment.* Reading, Mass.: Addison-Wesley.

Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don't all look alike: Individuated impressions of other racial groups. *Journal of Personality and Social Psychology, 65,* 85-101.

Zuckerman, M., Bernieri, F., Koestner, R., & Rosenthal, R. (1989). To predict some of the people some of the time: In search of moderators. *Journal of Personality and Social Psychology, 57,* 279-293.

Zuckerman, M., Koestner, R., DeBoy, T., Garcia, T., Maresca, B. C., & Sartoris, J. M. (1988). To predict some of the people some of the time: A reexamination of the moderator variable approach in personality theory. *Journal of Personality and Social Psychology, 54,* 1006-1019.

Effects of Moderator Variables on Agreement: Review of Empirical Findings and Research Methods

Taeyun Jung

Research Institute for Human Behavior, Yonsei University

This paper considered the moderator variable approach to agreement, with an emphasis on previous empirical findings and research methods. Differences both between two types of agreement (self-peer agreement and peer-peer agreement) and among three measures of agreement (discrepancy, correlation, and variance) were discussed. In addition, several problems with two conventional methods (target-based and trait-based methods) for testing moderators (behavioral consistency, observability, and social desirability) were identified. Review of literature on the moderators suggests that there is no convergence in empirical findings of the effects of the moderators except for observability. It was argued that this general failure of the moderator variable approach is due at least partially to problems with methods used. A recently developed method based on target standing was also introduced. Finally, application of the moderator variable approach and direction of future research were discussed.