

입사지원자와 일반 표준집단의 성격검사 요인점수 및 요인구조에 대한 비교연구*

정 승 철†

가톨릭대학교 심리학과

본 연구는 많은 기업에서 사용하고 있는 성격검사와 관련하여 꾸준히 제기되고 있는 응답왜곡 현상이 어느 수준까지 영향을 미치는지를 분석하였다. 연구에서 사용한 데이터는 프랑스 공공 교통 관련 기업에 입사지원자 21,349명이 5년간의 작성한 NEO PI-R 성격검사 자료이다. 분석 결과에 의하면 일반 표준그룹의 점수 비교에서 5개 요인(신경증, 외향성, 개방성, 친절성, 성실성) 모두에서 의미 있는 점수의 차이가 확인되었다. 신경증 요인과 개방성 요인의 경우, 지원자들의 점수가 일반 표준그룹의 점수보다 낮은 것으로 나타난 반면, 외향성 요인, 친절성 요인, 성실성 요인의 경우는 지원자들의 점수가 일반 표준그룹의 점수보다 높은 것으로 나타났다. 확인적 요인분석 결과, 이러한 요인별 점수의 변동에도 불구하고 5요인 구조의 안정성에 대해서는 영향을 끼치지 않은 것으로 나타났다. 따라서 인력선발이라는 특수한 상황에서 NEO-PI R 성격검사를 사용할 경우에는, 일반 표준그룹의 표준점수가 아닌 입사지원자들의 점수를 바탕으로 한 표준점수를 사용할 필요가 있다.

주요어 : 성격검사, 성격 5요인 모델, 응답왜곡, 사회적 바람직성, 구조안정성, 선발 평가

* 본 논문은 가톨릭대학교의 지원을 받아서 수행된 연구(M-2009-B0002-00317)입니다.

† 교신저자 : 정승철, 가톨릭대학교 사회과학부 심리학전공, parisjung@catholic.ac.kr, 02-2164-4548

기업들은 인력을 선발할 때 인지능력, 신체능력, 심리운동능력, 지각능력, 성격, 면접, 개인사 정보, 평가센터 등과 같은 다양한 선발 도구를 사용하고 있다(유태용, 2009; Salgado, Viswesvaran, & Ones, 2001). 최근 전 세계적으로 많은 기업들이 성격검사를 사용하는 것은 개인들의 직무성과가 능력적인 요소뿐만 아니라 동기나 성향적인 요인에 의해서도 영향을 받으며, 이러한 동기나 성향을 가장 잘 나타내는 것이 성격이기 때문이다(Borman, Hanson, & Hedge, 1997).

Salgado 등(2001)에 따르면, 1980년대 이후 지금까지 서방선진국 대부분의 국가(미국, 캐나다, 영국, 프랑스, 벨기에, 독일, 호주, 뉴질랜드, 등)에서 인력 선발에 성격검사의 사용은 꾸준히 증가하였다. Ryan, McFarland, Baron, 및 Page(1999)는 세계 20개국의 959개 기업을 대상으로 한 조사에서, 조사대상 기업의 약 21%에서 50%까지가 성격검사를 직원선발에 사용하고 있다고 보고했다. 특히, 성격의 5요인 모델 (Big 5 model of personality 혹은 Five Factor Model: FFM)은 현재 많은 연구자들의 지지를 받으며 성격관련 연구를 주도하고 있으며, 이 모델에 기초한 NEO-PI R (NEO Personality Inventory Revised) 성격검사는 많은 나라에서 가장 널리 사용되고 있는 방법이다(Marshall, DeFruyt, Rolland, & Bagby, 2005).

이러한 성격검사의 확산은 조직에서 인적자원의 역량에 대한 관심이 증가한 것도 있지만, 오랜 기간 동안 성격연구에서 가장 근본적인 문제인 성격의 요인구조에 대한 문제가 해결된 후 이어진 많은 연구들의 결과라고 할 수 있다. 과거에는 성격을 분류하는 일관된 분류 기준이 없었기 때문에 타당도가 낮았고, 그 결과 성격변인과 직무준거변인 간의 의미 있

는 관계가 있는지를 연구하는 것이 힘들었다 (Barrick & Mount, 1991). 1980년대 이후 많은 학자들과 실무자들로부터 호응을 얻고 있는 성격의 5요인 모델은 확고한 성격의 분류체계를 제공하며, 성격변인과 직무수행 간의 관계를 증명한 연구들과 인력선발 과정에서 성격변인의 유용성을 증명한 연구들이 나올 수 있는 바탕이 되었다(Barrick & Mount, 1991; Hogan & Hogan, 1989; Ones, Viswesvaran, & Schmidt, 1993; Salgado, 1997, 2002).

하지만 이러한 성격의 5요인 모델을 인력의 선발과정에 적용하는 것은 단순한 작업이 아니다. 여러 가지 요인들이 성격측정을 오염시키기 때문에 상당한 주의를 요하며, 그 중 가장 많이 제기되는 문제가 입사지원자들의 경우 자신들을 긍정적으로 보이려고 하는 경향 (positive presentation management 혹은 social desirability)에 따라 응답을 왜곡할 수 있다는 것이다(Hogan, 1998). 개인들에게 중요한 결과를 가져올 것으로 기대되는 상황에서 개인들이 응답을 의도적으로 왜곡하는 현상은 성격 검사 요인들의 점수에 영향을 미치고, 그 결과로 요인구조의 안정성을 위협할 수도 있다 (Ellingson, Smith, & Sackett, 2001). 만약 요인 점수의 변동만 생긴다면, 그 집단에 적합한 규준점수를 적용하면 해결이 된다. 하지만 요인구조가 유지되지 않는다면, 문제가 좀 더 복잡해진다. 무엇보다도 이론적인 틀에 대한 근본적인 의문이 제기될 수 있으며, 선발뿐만 아니라 향후 직무준거를 예측하는 데에 활용을 할 수가 없게 된다.

이에 따라 성격검사가 기업 등의 조직들에게서 관심을 받기 시작한 1990년 전후로 응답자의 의도적인 응답왜곡(Faking)이 성격검사 요인의 점수 및 요인구조 안정성에 끼치는 영향

에 대한 연구가 진행되어 왔다. 하지만 요인의 점수에 대한 영향과 요인 구조에 대한 영향에 대해서는 다소 상반된 결과들이 보고되었다(Montag & Comrey, 1990; Schmit & Ryan, 1993 vs. Marshall, DeFruyt, Rolland, & Bagby, 2005; Montag & Levin, 1994; Smith, 1996; Collins & Gleaves, 1998). 또한 연구 방법도 실제 장면이 아닌 상황설정을 통한 실험상황에서의 연구가 많으며, 연구 결과를 다양한 분야의 다양한 특성의 지원자들에게 일반화 할 수 있을 만큼 대규모의 입사지원자 표본을 대상으로 실행된 연구는 많지 않다. 특히 우리나라에서는 대규모의 실제 입사지원자 표본을 통한 연구는 더욱 드문 실정이며, 다른 문화권과의 동일한 성격검사 결과를 비교한 연구 또한 찾아보기 힘들다. 따라서 본 연구는 대규모의 실제 입사지원자를 대상으로 한 연구를 통해서 입사지원 상황에서 자신을 긍정적으로 보이려는 경향이 NEO-PI R 성격검사의 요인점수와 요인구조에 미치는 영향을 알아보았다. 또한 다른 나라의 연구 결과들과의 비교를 통하여 NEO-PI R 성격검사의 교차문화적 타당성을 검토하는 것에서 그 의미를 찾을 수 있다.

성격 5요인 모형

성격의 5요인 모형에 대한 연구는 1930년대에 성격을 독립적인 요인으로 구분하기 위해서 요인분석이란 연구방법을 개척한 McDougall(1932), Thurstone(1934), Guilford와 Guilford(1936) 등에 의해서 시작되었으며(Digman, 1996; Rolland, 2004), 이후 Fiske(1947), Tupes와 Christal(1961), Norman(1963), Borgata(1964) 그리고 Smith(1967) 등에 의해서 성격의 5요인 모델이 제시되었다(McAdams, 2001). 이후 80~90년대를 전후해서 많은 연구자들(Borkenau & Ostendorf, 1990; Costa & McCrae, 1988; Digman, 1996; Digman & Takemoto-Chock, 1981; Goldberg, 1993; John, 1990; McCrae, Zonderman, Costa, Bond, & Paunonen, 1996; Peabody & Goldberg, 1989)에 의해서 성격의 5요인 모델이 다양한 상황, 다양한 표본, 다양한 문화에 걸쳐서 상당히 일관성이 있다는 것이 증명되었다.

성격의 5요인은 외향성, 신경증(vs. 정서적 안정성), 개방성(지성), 친절성, 성실성 등의 5가지 요인으로 구성되어 있으며, 번역에 따라 약간의 차이는 있지만 명칭에 대해서는 전체

표 1. 5요인 성격요인의 전형적 특성

	외향성	친절성	성실성	신경증	개방성
전형적 특질	수다스러운, 주장적인, 활동적인, 활기찬, 주도적인, 긍정적인, 사교적인	동정적인, 친절한, 온정적인, 신뢰로운, 착한, 우호적인, 협조적인	정돈된, 능률적인, 책임성 있는, 믿음직한 신뢰할 수 있는, 신중한, 근면한	불안한, 신경질적인, 변덕스러운, 겁 많은, 자기연민의, 불안정한, 감정적인	흥미가 많은, 상상력이 풍부한, 지적인, 통찰력 있는, 호기심 많은, 예술가적인, 독창적인

주. 출처: John & Srivastava(1999).

적으로 유사하면서도 연구자마다 조금씩 다른 면이 있다. 이것은 다섯 요인의 세부적인 분류에서는 어느 정도 이견이 있으며, 성격을 5가지 요인, 또는 5개 요인으로 나누는 것 자체에 대해서 비판적인 학자들도 있다. 그러나 성격을 나타내는 어휘적 접근(lexical approach)이 위계적(hierarchical)으로 구조화될 수 있다는 점에는 대개 동의를 하고 있다. 이 점은 사람의 특성을 나타내는 여러 가지 성격 용어들이 일상생활에서 혼재되어 사용되고 있기는 하지만, 보다 상위의 개념, 또는 더 추상적인 요인으로 묶을 수 있다는 것은 무리 없는 해석체계가 될 수 있음을 반영하는 것이다.

성격 5요인 모델에 기초한 성격검사는 인력 선발 과정에서 유용하게 사용할 수 있다는 것을 증명한 일련의 연구들(Barrick & Mount, 1991; Hogan & Hogan, 1989; Hough & Oswald, 2000)과 조직 장면에서의 준거타당도에 대한 설득력 있는 연구들(Hogan, Hogan, & Roberts, 1996; Hough, 1992; Hurtz & Donovan, 2000; Ones, Viswesvaran, & Schmidt, 1993; Salgado, 1997, 2002; Tett, Jackson, & Rothstein, 1991) 덕분에 그 사용이 괄목할만하게 증가하였다. 또한 자기보고식 설문지를 통해서 성격을 측정할 수 있게 되면서 다른 방법으로는 불가능할 수 있는 다양한 상황에서 개인들의 행동을 평가할 수 있다는 것 때문에 연구자들이 선호하고 있다(Pervin & John, 2001). DeFruyt(2001)는 이러한 5요인 모형과 같이 특질이론에 바탕을 둔 성격검사지를 인력선발과 평가 과정에 사용하는 5가지 이유를 다음과 같이 제시하였다.

- 1) 지능검사를 보완하는 증분타당도가 분명하다.
- 2) 개인을 대상으로 하든 집단을 대상으로 하든 어려움 없이 사용할 수 있다.

3) 비용이 적게 든다.

4) 채점이 표준화되어 있고, 검사결과를 제시하는 것이 자동화될 수 있다.

5) 다양한 문화와 언어권에서 타당성과 신뢰성이 검증된 검사지가 갖춰져 있다.

성격 5요인 모형과 응답왜곡

성격 5요인 이론에 기반을 둔 성격검사가 앞 절에서 설명한 것처럼 여러 가지 장점을 가지지만, 자기보고식 성격검사를 기업 장면에서 응용하는 데 있어서는 성격측정이 쉽게 왜곡될 수 있다는 비판을 피할 수 없으며(Hogan, 1998), 이는 성격을 측정한다는 것 자체에 대해서 근본적인 의문을 제기할 수 있다(Hogan, DeFruyt, & Rolland, 2006). 성격을 평가하는 과정에서는 응답자의 성격적인 특성들 이외에 평가과정의 목적과는 관계가 없는 다른 여러 요인들(심리적, 사회적, 언어적 요인 등)이 응답에 영향을 미치게 되는데(Lanyon & Goodstein, 1997), 응답자들은 설문이 지시하는 내용을 이해하지 못하거나, 설문에 무관심 혹은 비협조적으로 응답할 수도 있다. 또한 질문의 대상인 자신의 일상 행동에 대해서 잘 모를 수도 있으며, 의도적으로 자신의 긍정적 혹은 부정적 이미지를 보이려고 노력할 수 있다(Rolland, 2004). 특히, 자신이 타인의 평가를 받고, 그 평가가 중요한 결과(선발 혹은 탈락)를 가져오는 상황에서는 긍정적인 평가를 받을 것으로 기대되는 방향으로 자신의 성격을 왜곡하게 되는데, 이를 응답에서의 사회적 바람직함 경향(social desirability) 혹은 긍정적 자기표현 관리(positive presentation management: PPM)라 한다. 이러한 경향에 따른 입사지원자의 응답왜곡은 자기보고식 성격검사의 요인

점수와 요인구조 안정성에 영향을 줄 수도 있다(Barrick & Mount, 1996; DeFruiyt, Aluja, Garcia, Rolland, & Jung, 2006; Ellingson et al., 2001; Furnham, 1986; Schmit & Ryan, 1993; Zerbe & Paulhus, 1987).

조직의 관점에서는, 인력을 선발하는 상황에서 조직이 좀 더 선호하는 성격적 특성이 있을 수 있다. Caldwell과 Burger(1998)의 연구 결과를 보면, 기업들이 좀 더 선호하는 성격적 특성이 있다는 것을 알 수 있다. 이들은 기업으로부터 입사제의를 더 많이 받은 개인들의 경우 그렇지 않은 개인들보다 정서적으로 더 안정적이며, 더 외향적이며, 더 개방적이고, 더 친절하다는 것으로 나타나 대학생들이 기업으로부터 입사제의를 받은 횟수가 성격과 관련이 있는 것으로 보고하였다.

지원자의 관점에서 볼 때, 입사제의를 받기 위해서는 고용주에게 긍정적인 인상을 보이는 것이 필요하기 때문에 긍정적 자기표현 관리(PPM)는 입사지원이라는 상황적인 요구에 대한 하나의 적응 수단으로 이해될 수 있다(Marcus, 2009). 사람들은 지원하는 직업에 따라 서로 다른 성격 프로파일을 나타내며(Birkeland, Manson, Kisamore, Brannick, & Smith, 2006; Furnham, 1990; Pauls & Crost, 2005), 긍정적인 왜곡에 의한 성격 프로파일이 이상적인 종업원을 나타내는 성격프로파일(Martin, Bowen, & Hunt, 2002) 혹은 전형적인 종업원의 성격 프로파일(Mahar, Cologon, & Duck, 1995)과 일치한다는 연구 결과들을 볼 때, 자기를 긍정적으로 보이려는 경향이 단순히 개인들이 임의로 자신을 긍정적으로 보이려는 것이 아니라, 지원하는 조직이나 직업의 전형적인 성격 특성이나 이상적인 성격특성에 부합하는 방향으로 응답을 왜곡하는 것을 알 수 있다. 이

에 대해서 Hogan(2006)은 인상관리(impression management)라는 것은 사회적 관계 속에서 수용과 지위를 극대화하고 거절과 지위상실을 최소화하려는 지속적인 노력으로써, 입사지원 상황에서 이러한 인상관리에 의한 요인점수의 변동은 피할 수 없는 현상이라고 하였다.

만약 자신을 긍정적으로 보이려고 하는 경향에 의한 점수변동 때문에 성격검사의 요인구조가 변하는 경우가 발생하면, 이는 각 요인들이 상호 독립적이며 특정 요인구조를 가지고 있다는 이론적인 바탕에 대한 근본적인 의문을 가지게 되고, 예측변인으로서의 가치까지 상실하므로 실제 기업에서 인력의 선발이나 개발을 위한 개인의 평가도구로서 사용할 수 없게 된다. 이러한 이유에서 입사지원 상황이나 경력 개발 상황에서의 성격검사의 요인구조안정성과 준거관련타당도(예연타당도)에 대한 연구들이 많이 이루어졌다.

응답왜곡에 의한 요인점수의 변동

사회적 바람직성 경향을 통해서 응답할 경우, 사람들은 사회적으로 인정되는 행동들보다 돋보이게 하려고 하고 사회적으로 인정되지 않는 행동이나 의견과는 거리를 두는 방식으로 응답을 하게 된다(Furnham, 1986). 입사지원 상황에서 이러한 사회적 바람직성 경향이 5요인 성격검사의 각 요인의 점수에 미치는 영향을 연구한 논문들을 살펴보면, Topping과 O'Gorman(1997)의 연구에서는 입사지원자들의 신경증 요인의 점수가 비입사지원자들의 점수보다 더 낮은 것으로 나타났으며, 외향성 요인, 친절성 요인, 성실성 요인의 점수는 더 높은 것으로 나타났다. Costa, McCrae, 및 Rolland (1998)의 연구는 신경증 요인과 개방성 요인에

서는 입사지원자들의 점수가 더 낮은 것으로 나타났고, 외향성 요인, 친절성 요인, 성실성 요인에서는 입사지원자들의 점수가 더 높은 것으로 발표하였다. 또한 Viswesvaran과 Ones (1999)의 통합분석 연구에 따르면, 사회적 바람직성에 따른 응답조건에서 외향성 요인, 정서적 안정성 요인(신경증의 반대 개념), 성실성 요인, 친절성 요인, 개방성 요인의 성격 5 요인 모두에서 점수의 증가가 있는 것으로 나타났으며, 특히 정서적 안정성과 성실성 요인이 가장 크게 영향을 받은 것으로 나타났다. 이후 Birkeland 등(2006)의 통합분석 연구에서는, 판매직, 비판매직, 판매와 비판매 혼합직, 관리직, 비관리직, 관리와 비관리 혼합직 등 다양한 종류의 직무 전체를 고려하였을 때, 외향성 요인, 정서적 안정성 요인, 성실성 요인, 개방성 요인에서 입사지원자들의 점수가 비입사지원자들의 점수보다 더 높은 것으로 나타났다. 이상의 연구들을 종합하면 표 2와 같다.

표 2의 내용을 살펴보면, 외향성 요인, 정서적 안정성(신경증) 요인, 성실성 요인의 경우는 모든 연구에서 공통적으로 같은 방향으로의 점수 변화가 있다는 것을 알 수 있다. 반면 친절성 요인의 경우는 일부 연구에서 점수

가 증가하였다는 것을 보여주었지만 다른 연구에서는 의미 있는 결과를 보여주지 않았다. 개방성 요인의 경우, Costa 등(1998)의 연구에서는 점수가 감소하였지만, 다른 두 연구에서는 점수가 증가한 것으로 나타났다. 표에는 제시하지 않았지만, Sandal, Musson, Helmreich, 및 Gravdal(2005)의 연구결과를 보면, 우주비행사를 대상으로 한 연구에서 Topping과 O'Gorman(1997)의 연구결과처럼 외향성 요인, 성실성 요인, 친절성 요인에서는 점수의 증가가 나타났고, 신경증 요인에서는 점수의 감소가 나타났다고 보고했다. 이들의 연구에서 주목할 점은 연구 대상이 NASA의 우주비행사들이라는 점이다. 즉 우주비행사라는 직업에서 개방성 요인의 점수가 높다는 것은 다른 직업에서와 달리 바람직한 것으로 간주되지 않는다.

같은 맥락에서 Pauls와 Crost(2005)의 연구는 관리자에 지원한다는 지시사항을 주었을 때와 간호사에 지원한다는 지시사항을 주었을 때의 점수의 변화 방향이 서로 다르게 나왔다고 보고하였다. 관리자에 지원한다는 지시사항을 준 그룹에서는 외향성 요인, 정서적 안정성 요인, 성실성 요인의 점수는 아주 크게 증가하였고, 친절성 요인과 개방성 요인은 의미

표 2. 응답왜곡에 의한 5요인 성격검사 요인점수의 변동

	외향성	정서적안정성 (신경증)	성실성	친절성	개방성
Topping & O'Gorman(1997)	+	(-)	+	+	
Costa, McCrae, & Rolland(1998)	+	(-)	+	+	-
Viswesvaran & Ones(1999)	+	+	+	+	+
Tsaousis와 Nikolaou(2001)	+	(-)	+		
Birkeland et al.(2006)	+	+	+		+

있게 감소한 반면 간호사에 지원한다는 지시 사항을 준 그룹에서는 외향성 요인, 정서적 안정성 요인, 성실성 요인, 친절성 요인의 점수는 증가하고 개방성 요인의 점수는 감소하였다. 즉, 연구 참가자들은 관리자 조건에 대해서는 친절성 요인과 개방성 요인의 점수를 감소시켰고, 간호사 조건에 대해서는 친절성 요인의 점수는 증가시키면서 개방성 요인의 점수는 감소시켰다. 이는 사람들이 지원하는 직업의 전형적인 특성이나 이상적인 특성에 맞추어 응답을 왜곡한다는 것을 의미한다. Furnham(1990)의 연구에서, 도서관 사서, 은행가, 광고 기획자에 지원한다고 가정하고 성격검사를 작성하라고 지시한 결과 실험참가자들은 도서관 사서는 고립된 내성적 성격의 소유자로, 은행가는 남성적이면서 관리적이며 관습적인 경험적 사고자로, 광고 기획자는 예술적이고 정이 많으며 사교적인 직관적 외향적 성격의 소유자를 나타내는 방향으로 점수를 왜곡하였다. 마찬가지로, Krahe, Becker 및 Zollter(2008)은 실험참가자들이 기자라는 직업에 대해서 외향성 요인을 연관시킨다고 보고하였다. 즉 사람들은 입사지원 상황에서 하나의 모형을 통해서 점수를 왜곡하는 것이 아니라 지원하는 직업에 따라 서로 다른 모형을 통해서 점수를 왜곡하는 것이다.

선발상황에서 성격검사의 구조안정성

성격검사와 같은 심리검사는 단일 요인인 경우는 드물고 일반적으로 다차원구조로 되어 있다. 따라서 성격검사의 경우 이론에서 제시하는 요인에 부합되게 문항들이 일관되게 모여지는지를 알아보는 것과 이러한 요인들 간의 상대적인 독립성을 검토하는 것이 필수적

이다(Rolland, 2001). 이러한 요인의 구조는 다양한 상황에 걸쳐서 안정적이어야 하고, 좋은 성격검사는 개인이 자신을 좋게 혹은 사회규범적으로 바람직하게 보이도록 하는 의도에 의해 크게 영향을 받지 않아야 한다(McAdams, 2001). 하지만 사회적 바람직성 경향에 따른 입사지원 상황에서의 응답왜곡은 일반적인 현상이며, 시간적인 안정성을 지니면서 일정한 패턴을 나타낸다(Hogan, Barrett, & Hogan, 2007).

결국, 이러한 사회적 바람직성 경향에 따른 응답방식은 성격특질의 측정에 있어서 외생분산(the extraneous variance)의 원인이 될 수도 있으며, 이로 인해서 성격검사의 구성요인 타당도가 영향을 받을 수 있다(Holden & Fekken, 1989). 또한 응답왜곡 조건에서의 요인 간 상관관이 상당히 높게 나타났으며, 이는 실제 입사지원자들의 요인 간 상관보다 큰 것으로 나타났다(Ellingson et al., 2001; Zickar & Robie, 1999), 이것은 성격검사가 본래 측정하고자 하는 것과는 다른 추가적인 공통 요인이 개입된 것으로 설명될 수 있다. 또한 Sisco와 Reilly(2010)는 사회적바람직성 척도 점수가 낮은 그룹의 공변량 변화가 사회적바람직성 척도 점수가 높은 그룹의 공변량 변화보다 더 작으며 성격의 5요인 모형과 일치하는 요인구조에 더 가깝다는 것을 보여주었다. 따라서 성격검사가 가지고 있는 본래의 요인들 간의 관계성에 변화가 있을 정도로 응답왜곡이 크게 나타난다면 구조안정성이 위협받을 수 있다.

입사지원 상황에서의 긍정적 혹은 부정적 응답왜곡의 영향을 검증하기 위해 주로 대학생들을 대상으로 가상의 입사지원 상황을 가정하고 이루어진 연구들은 응답왜곡이 성격검사의 요인구조의 안정성을 위협할 수 있는 것

으로 보고했다(Ellingson, Sackett, & Hough, 1999; Ellingson et al., 2001). 하지만 실제 상황에서 입사지원자를 대상으로 한 연구에서는 자신을 긍정적으로 보이려는 현상이 성격검사의 요인 점수의 수준은 변화시키지만 구성요인타당도나 예언타당도에는(Barrick & Mount, 1996; Marshall, De Fruyt, Rolland, & Bagby, 2005; Ones, Viswesvaran, & Reiss, 1996; Rolland, 2003, 2004; Schmit et al., 1995; Viswesvaran & Ones, 1999) 영향을 미치지 않다고 한 연구들이 지배적이다. 이러한 상반된 결과는 실험상황과 달리 실제 상황에서는 각 성격 요인 간의 독립적인 관계를 왜곡할 만큼 극단적인 긍정적 혹은 부정적인 응답의 왜곡은 없으며, 신뢰도 또한 영향을 미치지 않으면서 점수의 변동이 있기 때문이다. 따라서 본 연구의 표본이 실제 입사지원자를 대상으로 한 연구이기 때문에, 요인구조의 안정성은 유지될 것으로 가정할 수 있다.

연구목적 및 가설

이상의 논의를 바탕으로, 본 연구는 우선 실험상황이 아닌 실제 입사지원 상황에서 성격검사의 요인 점수가 일반 표준집단과 비교할 때 어떠한 경향성을 띠는가를 알아보고, 다음으로 이러한 요인 점수의 변화가 요인 구조에는 어떠한 영향을 주는가를 알아보고자 한다.

선행 연구들을 종합할 때, 성격검사의 요인 점수와 관련해서는 입사지원 상황에서 성실성 요인, 외향성 요인, 친절성 요인의 점수는 증가하고 신경증 요인의 점수는 감소한다는 것을 알 수 있었다. 반면 개방성 요인의 경우 일관된 결과를 보여주지 않았다. 따라서 본

연구의 대상자가 특정한 직무에 국한되지 않고 여러 가지 직무가 혼재되어 있기 때문에, 본 연구에서는 선행연구들의 일반적인 결과에 근거해서 다음과 같이 가설을 설정하였다. 개방성 요인과 관련해서는 일관되지 않은 선행 연구들의 결과로 인해서 탐색적인 수준에서 분석을 하고자 한다.

가설 1: 입사지원자들의 성격 요인의 점수는 일반 표준집단의 성격 요인 점수에 비해서 자신을 긍정적으로 표현할 수 있는 방향으로 점수의 증가 혹은 감소가 있을 것이다.

가설 1-1: 입사지원자의 신경증 요인 점수는 표준집단의 점수보다 더 낮을 것이다.

가설 1-2: 입사지원자의 외향성 요인 점수는 표준집단의 점수보다 더 높을 것이다.

가설 1-3: 입사지원자의 친절성 요인 점수는 표준집단의 점수보다 더 높을 것이다.

가설 1-4: 입사지원자의 성실성 요인 점수는 표준집단의 점수보다 더 높을 것이다.

요인구조의 안정성과 관련해서는 긍정적인 응답왜곡에 대한 연구자의 의도적인 지시가 있는 몇몇 연구들에서는 구성요인 타당도가 영향을 받은 것으로 나타났지만, 실제 입사지원자를 대상으로 한 연구에서는 요인구조의 안정성이 유지되는 것으로 나타났다. 따라서 실제 입사지원자를 대상으로 하는 본 연구에서는 다음과 같이 가설을 설정하였다.

가설 2: 입사지원자들의 성격 요인 점수에 변화가 있더라도, 성격의 5요인 구조는 유지될 것이다.

연구방법

연구대상

본 연구는 공공교통 관련 프랑스 기업에 지원할 때 인력선발 평가절차를 밟은 21,349명(여성 3,894명, 남성 17,455명)을 대상으로 이루어졌다. 평균연령은 26세이며, 연령대는 17세에서 53세까지 다양했다. 지원자들은 다양한 학력수준(중졸 ~ 대졸)과 사회·경제적 수준을 나타냈고, 이들이 지원한 분야(지하철기관사, 버스기사, 안전요원, 관리직, 유지·보수직, 서비스 등) 또한 다양했다. 더불어, 이러한 지원자 집단과의 점수 및 요인구조 안정성의 비교를 위해서 프랑스를 비롯한 3개국의 NEO-PI R 표준집단의 자료를 활용하였다: 프랑스 표준집단($N = 801$: 남자 334명, 여자 467명, Costa, McCrae, & Rolland, 1998), 미국 표준집단($N = 1,000$: 남자 500명, 여자 500명, Costa & McCrae, 1992), 그리고 스페인 표준집단($N = 2,000$: 남자 1,000명, 여자 1,000명, Costa & McCrae, 1999).

측정도구

성격의 측정을 위해서는 프랑스어 판 NEO-PI R 성격검사(Costa et al., 1998)를 사용하였다. NEO-PI R 성격검사는 리커트 5점 척도의 240개 문항으로 이루어져 있으며, 5개 성격요인 및 30개의 하위요인을 측정한다: 신경증(불안성, 분노-적대성, 우울성, 수줍음, 충동성, 취약성), 외향성(온정성, 사교성, 주장성, 활동성, 흥분추구, 긍정적 정서), 개방성(공상에 대한 개방성, 미적 개방성, 감성에 대한 개방성, 행위에 대한 개방성, 생각에 대한 개방

성, 가치에 대한 개방성), 친절성(신뢰성, 정직성, 이타성, 순응성, 겸손함, 동정심), 성실성(자기유능감, 질서, 충실성, 성취의지, 자기관리, 숙고).

분석방법

긍정적 자기표현 경향에 따른 점수 차이의 유의성을 검증하기 위해서는 t -검증과 코헨의 d (Cohen's d)를 사용하였으며, 요인구조의 안정성 검증에는 AMOS를 통한 확인적 요인분석과 프로크루스테스 요인분석(orthogonal procrustes rotation) 방법을 사용하였다. 코헨의 d 는 두 집단의 평균을 비교할 때, 그 차이의 중요성을 나타내주는 지표로 활용된다(0.66을 넘으면 두 집단의 차이가 크다고 할 수 있으며, 0.33보다 작으면 그 크기가 작다고 할 수 있다).

본 연구에서는 요인구조 안정성의 비교를 위해 확인적 요인분석을 위해 사용된 구조방정식의 모형과 프로크루스테스 요인분석 방법은 NEO-PI R을 가지고 수행한 선행 연구의 결과와 비교를 위해서 기존 연구에서 사용한 방법과 동일한 분석방법을 사용하였다. 우선 구조방정식모형을 통한 확인적 요인분석방법을 실행하는 데에 있어서 McCrae 등(1996)이 제시한 8가지 모형들 중에서 complete oblique 모형과 complete orthogonal 모형을 적용하였다. 그림 1에서 보는 바와 같이 complete oblique 모형은 NEO-PI R의 5개 요인과 30개 하위요인들이 모두 서로 상관되어 있다는 모형이며 complete orthogonal 모형은 요인 간에는 서로 상관되어 있지 않고 하위요인 간에만 모두 상관되어 있다는 모형이다. 이 모형들은 이론적으로는 NEO-PI R의 5개 요인과 30개 하위요인이 서로 독립적이지만 실제 자료에서 이들

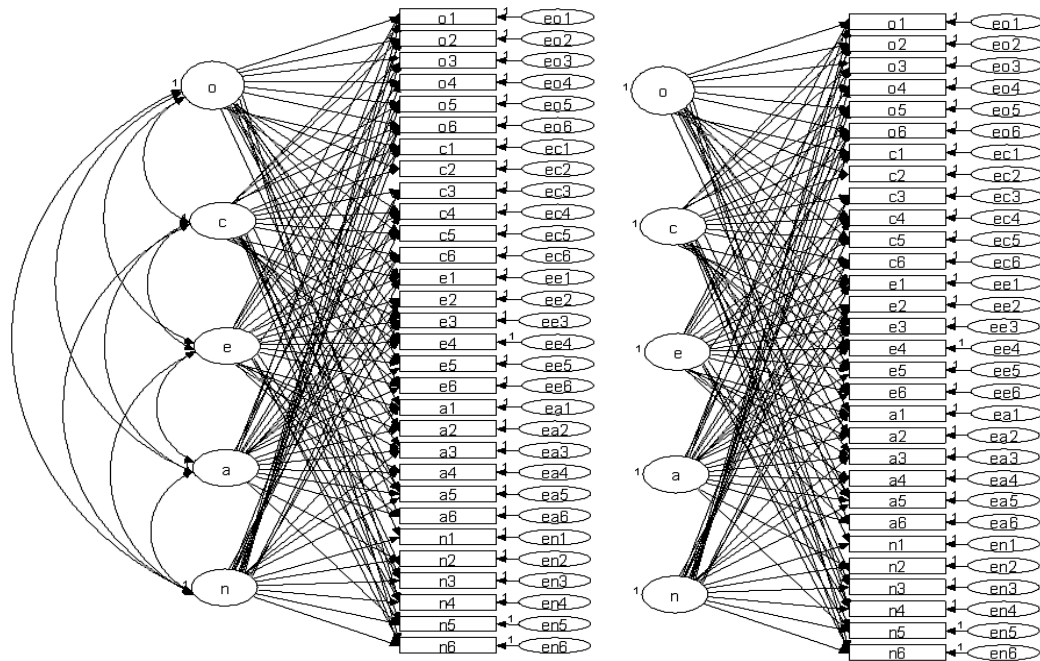


그림 1. McCrae 등(1996)이 제안한 Complete Oblique Model(좌)과 Complete Orthogonal Model(우)

간의 상관이 '0'인 경우는 없었기 때문에, 실제자료에 적용되는 모델은 이러한 상황을 반영해야 한다는 것을 의미한다(McCrae et al., 1996). NEO-PI R의 요인구조와 관련하여 이상의 두 모델을 비교한 연구들은 결과에서 큰 차이가 없다고 보고하였다(Aluja et al., 2004; Terracciano, 2003). 따라서 본 연구에서는 실제 입사지원자 집단의 자료를 선행연구와 같은 방법으로 분석하였을 때에도 같은 결과를 나타내는지에 대해서 검토해볼 필요가 있다.

본 연구에 사용된 프로크루스테스 요인분석 방법은 입사지원자 집단의 NEO-PI R 성격검사를 요인 적재량 행렬 추출 방법에서는 주성분분법을, 요인 행렬을 회전시키는 방법으로는 Varimax 사용하여 프랑스 표준 집단 및 미국 표준 집단을 비교 하는 방식을 취하였다. 이러한 방법은 미국과 유럽의 성격 연구자들이

요인구조의 일치도를 분석하면서 가장 보편적으로 사용되고 있는 구조방정식 모형의 한계성을 알게 된 후, 대안적으로 활용되고 있는 방법이다. 프로크루스테스 요인분석 방법은 구조방정식 모형처럼 요인구조를 재확인하기 위해 활용되는데, 이 방법은 요인 간의 직교성은 유지하면서 요인부하의 상대적인 위치에 영향을 주지 않고 요인공간에서의 축들의 위치를 정렬하며 또한 목표 요인구조를 얻도록 실험 데이터를 강제하거나 왜곡하지 않는다(Terracciano, 2003). 따라서 이러한 방법은 이론적인 요인 구조를 검증하는데 유용하며, 각 요인의 일치도(congruence)를 함께 제공할 수 있어서 구조방정식 모형을 이용한 확인적 요인분석이 제공하지 못하는 요인별 일치도 정보를 얻을 수 있다(McCrae et al., 1996).

결 과

요인점수의 비교

자신을 긍정적으로 보이려는 경향(PPM)이 요인 점수에 미치는 영향을 살펴보면(표 3 참조), 5개 요인 모두에서 의미 있는 차이를 보이는 것으로 나타났다. 이러한 점수 차이의 중요성 정도를 보면, 신경증 요인과 성실성 요인은 코헨의 d 가 각각 1.21과 1.36인 것으로 나타나서 다른 요인들의 경우보다 그 차이의 크기가 상당히 큰 것으로 나타났다.

점수 차이의 방향과 관련해서는, 신경증 요인과 개방성 요인은 입사지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 낮은 것으로 나타난 반면, 외향성 요인, 친절성 요인, 성실성 요인은 입사지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 높은 것으로 나타났다. 또한 지원자 집단에서 각 요인의 왜도를 보면, 신경증과 개방성 요인의 경우(각각 0.164, 0.193) 분포가 왼쪽으로 치우쳐 있고, 외향성, 친절성, 성실성 요인(각각 -0.12, -0.126, -0.293)은 분포가 오른쪽으로 치우쳐 있다는 것을 보아 전반적으로 점수의 이동이 있다는 것을 알 수 있다.

본 연구의 가설 1에서, 신경증 요인의 경우에는 지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 낮고 개방성 요인을 제외한 나머지 요인들의 경우에는 지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 높을 것으로 기대하였는데, 분석 결과에서도 가설에서 기대한 것과 같은 결과가 나왔다. 이러한 결과를 통해서 본 연구의 가설 1은 검증되었다고 할 수 있다. 개방성 요인의 경우 지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 낮은 것으로 나타났다.

하위 요인의 점수를 비교 분석한 결과, 전반적으로 요인의 점수가 차이가 나는 방향으로 하위 요인의 점수도 차이가 나는 것으로 나타났지만 개방성 요인의 하위 요인의 경우 그 차이가 일관적이지 않은 것으로 나타났다. 공상에 대한 개방성, 미적 개방성, 감성에 대한 개방성, 가치에 대한 개방성 하위 요인의 점수는 프랑스 기준 집단의 점수가 입사지원자 집단의 점수보다 높은 것으로 나타났고, 행위에 대한 개방성과 아이디어에 대한 개방성 하위 요인의 점수는 입사지원자 집단의 점수가 프랑스 기준 집단의 점수보다 더 높은 것으로 나타났다(표 4 참조).

표 3. 지원자 집단과 프랑스 기준 집단의 요인 점수 비교

	지원자 집단				프랑스 기준 집단		t -검증		코헨의
	평균	표준편차	왜도	첨도	평균	표준편차	t	p	d
신경증	69.5	18.6	0.164	0.354	92.2	23.2	-33.5	0.000	1.21
외향성	121.9	15.0	-0.12	0.498	110.0	18.3	21.9	0.000	0.79
개방성	109.5	14.2	0.193	0.427	116.5	19.1	-13.5	0.000	0.49
친절성	130.6	14.9	-0.126	0.429	121.2	16.8	17.5	0.000	0.63
성실성	137.6	15.8	-0.293	0.597	115.9	20.3	37.9	0.000	1.36

표 4. 지원자 집단과 프랑스 기준 집단의 하위요인 점수 비교

		지원자 집단				프랑스 기준 집단		t-검증		코헨의 <i>d</i>
		평균	표준편차	왜도	첨도	평균	표준편차	<i>t</i>	<i>p</i>	
신경증	불안성	13.6	4.416	0.105	0.010	17.8	5.438	-26.2	0.000	0.941
	분노-적대성	10.2	4.037	0.419	0.522	14.3	5.224	-27.9	0.000	1.004
	우울성	11.3	4.554	0.316	0.241	15.4	5.594	-24.7	0.000	0.890
	수줍음	13.3	4.233	0.188	0.126	16.0	4.729	-17.4	0.000	0.625
	충동성	13.4	3.884	0.081	0.168	16.9	4.501	-24.5	0.000	0.882
	취약성	7.7	3.588	0.581	1.289	11.9	4.920	-32.1	0.000	1.155
외향성	온정성	25.0	3.659	-0.582	0.883	22.4	4.407	19.3	0.000	0.695
	사교성	20.9	4.250	-0.428	0.455	17.4	5.418	23.0	0.000	0.829
	주장성	16.3	4.238	0.064	0.014	14.5	4.631	11.9	0.000	0.428
	활동성	20.5	3.804	-0.149	0.177	19.0	4.427	10.2	0.000	0.367
	흥분추구	17.9	4.297	-0.097	-0.050	16.6	4.737	8.6	0.000	0.308
	긍정적 정서	21.3	3.952	-0.284	0.351	20.1	4.808	8.5	0.000	0.305
개방성	공상에 대한 개방성	14.8	4.340	0.325	0.256	18.8	5.159	-25.7	0.000	0.925
	미적 개방성	17.3	4.804	-0.089	-0.124	18.8	5.691	-8.1	0.000	0.292
	감성에 대한 개방성	19.4	3.804	-0.066	0.085	21.1	3.974	-11.9	0.000	0.427
	행위에 대한 개방성	18.1	3.506	-0.192	0.116	17.4	3.857	5.9	0.000	0.211
	아이디어에 대한 개방성	20.0	4.481	-0.272	0.145	19.5	5.220	3.0	0.003	0.108
	가치에 대한 개방성	19.9	3.189	0.058	0.215	21.1	3.750	-9.8	0.000	0.351
친절성	신뢰성(타인)	18.1	4.657	-0.184	-0.136	17.6	5.355	3.2	0.002	0.114
	정직성	22.3	4.522	-0.494	0.213	20.3	5.059	12.3	0.000	0.443
	이타성	25.7	3.231	-0.541	0.912	23.1	3.636	22.7	0.000	0.815
	순응성	20.4	4.118	-0.151	0.061	17.7	4.636	18.0	0.000	0.647
	겸손함	20.9	4.131	-0.217	0.106	20.7	4.273	1.1	0.258	0.041
	동정심	23.2	3.459	-0.311	0.565	21.8	3.703	11.2	0.000	0.402
성실성	자기유능감	22.5	3.390	-0.252	0.385	19.3	3.835	26.6	0.000	0.956
	질서	21.6	3.733	-0.205	0.209	18.3	4.996	24.6	0.000	0.885
	충실성	25.8	3.514	-0.576	0.620	22.9	4.049	23.0	0.000	0.827
	성취의지	22.7	3.224	-0.278	0.356	19.1	4.423	30.8	0.000	1.109
	자기관리	24.0	3.534	-0.414	0.866	19.4	4.991	35.9	0.000	1.292
	숙고	20.9	4.053	-0.342	0.306	16.9	4.786	26.8	0.000	0.965

표 5. 프로크루스테스 요인분석 결과: 프랑스 기준 집단 vs. 지원자 집단

		N	E	O	A	C	일치도 (하위요인)
신 경 증	불안성	0.76	-0.03	0.09	-0.04	-0.17	0.97
	분노-적대성	0.63	-0.14	-0.03	-0.40	-0.20	0.98
	우울성	0.75	-0.13	0.03	-0.10	-0.27	0.99
	수줍음	0.76	-0.11	0.00	0.06	-0.12	0.98
	충동성	0.51	0.27	0.00	-0.27	-0.34	0.97
	취약성	0.59	-0.21	-0.08	-0.08	-0.51	0.96
외 향 성	온정성	-0.24	0.62	0.14	0.30	0.29	0.98
	사교성	-0.20	0.65	0.00	0.07	0.02	0.99
	주장성	-0.46	0.25	0.19	-0.35	0.32	0.96
	활동성	-0.13	0.43	-0.04	-0.18	0.41	0.98
	홍분추구	0.07	0.51	0.21	-0.37	-0.09	0.96
	긍정적 정서	-0.09	0.64	0.27	0.05	0.11	0.96
개 방 성	공상에 대한 개방성	0.14	0.07	0.56	-0.11	-0.33	1.00
	미적 개방성	0.14	0.11	0.73	-0.01	0.10	0.98
	감성에 대한 개방성	0.28	0.42	0.52	-0.08	0.11	0.98
	행위에 대한 개방성	-0.32	0.14	0.42	-0.05	0.00	0.92
	생각에 대한 개방성	-0.25	0.05	0.67	-0.02	0.34	0.93
	가치에 대한 개방성	-0.27	0.16	0.34	0.22	-0.10	0.87
친 절 성	신뢰성	-0.34	0.29	0.12	0.41	0.00	0.99
	정직성	-0.06	0.08	-0.08	0.66	0.18	0.98
	이타성	-0.12	0.51	0.11	0.41	0.36	0.98
	순응성	-0.19	-0.01	0.02	0.65	0.17	0.96
	겸손함	0.03	0.01	-0.02	0.60	-0.08	0.91
	동정심	0.18	0.37	0.17	0.46	0.23	0.94
성 실 성	자기유능감	-0.35	0.20	0.09	0.02	0.65	0.99
	질서	-0.07	0.07	0.05	0.00	0.67	0.91
	충실성	-0.18	0.22	0.00	0.24	0.67	0.95
	성취의지	-0.11	0.25	0.07	-0.06	0.69	0.99
	자기관리	-0.36	0.21	0.07	0.14	0.66	0.94
	속고	-0.28	-0.15	0.04	0.22	0.63	0.96
일치도(요인)		0.97	0.96	0.95	0.97	0.96	

표 6. 프로크루스테스 요인분석 결과: 미국 기준 집단 vs. 지원자 집단

		N	E	O	A	C	일치도 (하위요인)
신경증	불안성	0.77	-0.04	0.02	-0.01	-0.15	0.99
	분노-적대성	0.65	-0.10	-0.06	-0.40	-0.18	0.98
	우울성	0.76	-0.12	-0.03	-0.10	-0.25	0.99
	수줍음	0.75	-0.13	-0.07	0.07	-0.11	0.99
	충동성	0.53	0.29	-0.03	-0.23	-0.32	0.99
	취약성	0.59	-0.20	-0.11	-0.10	-0.50	0.96
외향성	온정성	-0.23	0.58	0.13	0.39	0.28	0.96
	사교성	-0.19	0.64	0.01	0.15	0.02	0.99
	주장성	-0.43	0.30	0.24	-0.31	0.33	0.97
	활동성	-0.13	0.44	-0.03	-0.12	0.42	0.92
	흥분추구	0.11	0.56	0.21	-0.29	-0.06	0.97
	긍정적 정서	-0.07	0.63	0.27	0.15	0.12	0.98
개방성	공상에 대한 개방성	0.20	0.09	0.55	-0.08	-0.32	0.99
	미적 개방성	0.20	0.11	0.71	0.05	0.11	0.98
	감성에 대한 개방성	0.33	0.42	0.49	0.01	0.13	1.00
	행위에 대한 개방성	-0.28	0.15	0.44	-0.02	0.00	0.96
	생각에 대한 개방성	-0.20	0.05	0.69	0.03	0.35	0.95
	가치에 대한 개방성	-0.25	0.14	0.35	0.24	-0.11	0.76
친절성	신뢰성	-0.34	0.24	0.13	0.44	-0.02	0.99
	정직성	-0.09	0.00	-0.11	0.67	0.15	0.97
	이타성	-0.12	0.45	0.08	0.48	0.34	0.97
	순응성	-0.21	-0.09	0.00	0.64	0.14	0.98
	겸손함	0.01	-0.07	-0.05	0.60	-0.10	0.94
	동정심	0.17	0.30	0.12	0.52	0.22	0.92
성실성	자기유능감	-0.36	0.19	0.10	0.06	0.64	1.00
	질서	-0.08	0.06	0.04	0.04	0.67	0.95
	충실성	-0.20	0.18	-0.01	0.28	0.66	0.96
	성취의지	-0.11	0.25	0.06	0.00	0.69	0.98
	자기관리	-0.37	0.19	0.08	0.18	0.65	0.96
	속고	-0.30	-0.18	0.04	0.21	0.61	0.98
일치도(요인)		0.98	0.98	0.95	0.96	0.98	

요인구조의 비교

NEO-PI R 성격검사에 대한 지원자 집단의 요인구조와 프랑스 표준집단의 요인구조를 비교한 프로크루스테스 요인분석의 결과(표 5 참조)를 보면, 모든 요인의 요인일치도 계수가 1에 가까운 것으로 보아 두 집단의 요인구조 일치도가 상당히 높은 것으로 나타났다. 하위 요인들의 일치도 또한 개방성 요인의 하위요인인 가치에 대한 개방성(0.87)을 제외하면 모든 하위요인의 요인 일치도가 높은 것으로 나타났다.

미국 표준집단의 요인구조와의 비교 결과를 보면(표 6 참조), 프랑스 표준집단과의 비교 결과처럼 두 집단의 요인구조 일치도가 상당히 높은 것으로 나타났으며, 하위요인들의 일치도 또한 개방성 요인의 하위요인인 가치에 대한 개방성(0.76)을 제외하면 모든 하위요인의 요인 일치도가 높은 것으로 나타났다. 이러한 결과를 통해서 지원자 집단의 요인구조가 일반 집단의 요인구조와 다르지 않다는 것

을 알 수 있으며, 또한 문화적 배경이 다른 집단의 요인구조와도 다르지 않다는 것을 알 수 있다.

확인적 요인분석 결과 또한 지원자 집단의 요인구조가 5개 요인과 30개 하위요인으로 구성되어 있는 모델을 지지하는 것으로 나타났다(표 7 참조). 자신을 긍정적으로 보이려는 경향(PPM)이 요인들의 점수에 영향을 미쳤음에도 불구하고(표 3 참조), 요인 간 상호상관을 가정한 모형(Complete Oblique Model)과 요인 간 상호 독립적이라는 것을 가정한 모형(Complete Orthogonal Model) 모두 만족스러운 모형적합도를 보이는 것으로 나타났다. 특히 자신을 긍정적으로 보이려는 경향이 없는 것으로 가정되는 미국, 스페인, 프랑스의 표준 집단의 결과들보다 더 높은 모형적합도를 보임으로써, 자신을 긍정적으로 보이려는 경향에 의한 요인들의 점수 변화가 NEO-PI R 성격검사의 5요인 구조의 안정성에는 영향을 미치지 않은 것으로 나타났다. 이상의 결과를 종합할 때, 본 연구의 가설 2는 검증되었다고

표 7. 확인적 요인분석 결과

모델	표본집단	GFI	AGFI	NFI	CFI	RMSEA
Oblique Model	미국 표준집단	0.903	0.855	0.885	0.906	0.062
	스페인 표준집단	0.936	0.904	0.931	0.942	0.051
	프랑스 표준집단	0.895	0.843	0.882	0.906	0.063
	RATP 지원자	0.959	0.939	0.945	0.946	0.042
Orthogonal Model	미국 표준집단	0.902	0.857	0.883	0.904	0.061
	스페인 표준집단	0.923	0.887	0.914	0.925	0.057
	프랑스 표준집단	0.888	0.838	0.871	0.896	0.066
	RATP 지원자	0.949	0.925	0.928	0.929	0.048

주. GFI: Goodness-of-Fit Index, AGFI: Adjusted Goodness-of-Fit. Index, NFI: Bentler-Bonett Index or Normed Fit Index, CFI: Comparative Fit Index, RMSEA: Root Mean Square Error of Approximation

할 수 있다.

논 의

본 연구에서는 지난 20여 년간 학술적 연구와 조직에서의 활용 모두에서 괄목할만한 성장을 한 성격검사를 인력선발 상황이라는 특정한 상황에 사용할 경우에 고려해야 하는 문제들 중에서 긍정적 자기표현 경향(PPM) 혹은 사회적바람직성 경향(social desirability)에 따른 성격검사의 요인 점수의 변동 문제와 요인 구조의 안정성 문제에 대해서 대규모의 실제 입사지원자를 대상으로 살펴보았다.

입사지원이라는 상황에서 지원자들의 긍정적 응답왜곡 의도에 따른 요인 점수의 변동과 관련한 분석 결과, 입사지원자 집단의 신경증 요인과 개방성 요인의 점수는 프랑스 기준 집단의 점수보다 더 낮게 나왔고, 외향성 요인, 친절성 요인, 성실성 요인의 점수는 더 높게 나왔다. 이러한 결과는 입사지원자들이 자신이 지원한 직무(혹은 직업)에서 더 좋은 평가를 받기 위해 성격검사에서 응답을 의도적으로 왜곡하는 현상을 나타내는 선행 연구들의 결과와 맥락을 같이 한다고 할 수 있다.

개방성 요인의 경우 지원자 그룹의 점수가 기준 집단의 점수보다 더 낮은 것으로 나타났는데, 이러한 결과는 우선 RATP라는 조직이 공기업이고, 실제 업무도 지하철기관사, 버스 기사, 안전요원, 관리직, 유지·보수직, 서비스 직 등과 같이 개방성의 발휘가 덜 중요한 직무이기 때문에 지원자들도 그에 맞추어 응답하는 과정에서 나왔다고 할 수 있다. 개방성 요인의 하위 요인의 결과를 보면, 다른 요인에서는 요인의 결과와 하위 요인의 결과가 같

은 방향으로 일관되게 나온 반면, 개방성 요인의 경우는 하위 요인 중에서 공상에 대한 개방성, 미적 개방성, 감성에 대한 개방성, 가치에 대한 개방성의 4개 하위 요인은 지원자 집단의 점수가 기준 집단의 점수보다 더 낮은 것으로 나타났고, 행위에 대한 개방성과 생각에 대한 개방성의 경우는 지원자 집단의 점수가 기준 집단의 점수보다 더 높은 것으로 나타났다(표 3 참조). 이렇듯 개방성의 하위 요인들은 다른 요인의 하위 요인들에 비해서 이질성이 높기 때문에 지원하고자 하는 직무에 따라서 그 결과가 달라질 수도 있다. 또한 조직의 입장에서 볼 때, 개방성이라는 것은 항상 긍정적이라고 볼 수만은 없는 특성이다. 공상에 대한 개방성, 미적 개방성, 감성에 대한 개방성 그리고 가치에 대한 개방성의 경우는 조직의 입장에서 오히려 생산성과 효율성을 저해할 수 있는 요인으로 인식될 수도 있다. 따라서 조직에 입사를 희망하는 입사 지원자들의 경우, 이러한 요인들에 대해서 소극적으로 응답하는 경향이 나타날 수 있고, 결과적으로 입사 지원자들의 점수가 일반 기준 집단의 점수보다 더 낮게 나올 가능성이 있다. 반면, 창의성이 많이 요구되는 직무를 지원한 사람들의 경우는 그 반대의 결과가 나올 수도 있다. 결국, 실제 조직의 상황에서 사용할 때에 이러한 개방성 요인의 경우는 좀 더 조심성 있게 다루어야 할 것으로 보인다.

입사지원자 그룹에서의 성격검사 요인 점수의 변동이 요인구조에 미치는 영향에 대한 분석 결과에서는, 가설에서 기대한 대로 요인구조의 안정성이 유지되는 것으로 나타났다. 이것은 성격검사를 예언변인(독립변인)으로 활용하는 데에는 아주 중요한 연구 결과이다. 이러한 결과는 NEO-PI R 성격검사가 개인들이

의도적으로 응답을 왜곡하려는 의도를 가지는 인력 선발 상황에서도 검사의 중요한 기본 요건인 요인구조 안정성을 유지하는 것을 의미한다. 또한 미국 기준집단과의 비교에서 요인구조의 일치도가 매우 높은 것으로 나타나서 교차문화적인 타당성도 동시에 만족하였다.

본 연구에서 분석한 결과를 좀 더 일반화시키자면, NEO-PI R 성격검사가 입사지원이라는 상황적 특성에 따라서 응답이 의도적으로 왜곡되어 단순한 점수의 증가나 감소만 있을 뿐 요인 구조의 변형이 없기 때문에, 상황에 적합한 새로운 기준점수를 수립하면 된다. 따라서 본 연구에서는 자신을 긍정적으로 보이려는 경향에 따른 점수의 변화를 반영하는 새로운 기준점수를 제시하고자 한다(표 8 참조). 점수의 차이가 심했던 신경증 요인과 성실성 요인의 경우 기준점수도 크게 변화가 있는 것

을 볼 수 있다. 예를 들어, 프랑스 기준 집단에서는 신경증 요인의 점수가 100이면 sten 점수가 6으로써 중간 수준이지만, 지원자 집단에서는 신경증 요인의 점수가 100이면 sten 점수가 9로써 아주 높은 수준이 된다.

본 연구의 결과에 대해서 사회적 바람직성 혹은 자신을 긍정적으로 표현하려는 경향에 의한 것이 아니라 표본의 특이성 즉, 특정 기업에 지원한 특정한 부류의 사람들에게 집중되었기 때문이라는 의문을 가질 수도 있다. 하지만 본 연구가 21,349명이라는 큰 표본을 대상으로 이루어졌으며, 연령대, 학력 수준, 사회경제적 수준, 지원 분야 모두가 다양하다는 것을 고려하면, 이러한 결과는 표본의 특이한 특성 때문에 발생한 우연한 결과라고 간주하기 어렵다.

본 연구의 한계점으로 우선 성격검사를 작

표 8. 기준 집단과 지원자 집단의 기준점수의 비교

		아주 낮음		낮음		중간		높음		아주 높음	
		1	2	3	4	5	6	7	8	9	10
		2.30%	4.40%	9.20%	15%	19.10%	19.10%	15%	9.20%	4.40%	2.30%
지 원 자 집 단	신경증	<33	<43	<51	<60	<69	<78	<88	<98	<109	109<
	외향성	<91	<100	<108	<115	<122	<129	<136	<144	<152	152<
	개방성	<82	<89	<96	<102	<109	<116	<123	<131	<140	140<
	친절성	<101	<109	<116	<123	<131	<138	<145	<153	<160	160<
	성실성	<104	<114	<123	<130	<138	<146	<153	<161	<168	168<
표 준 집 단	신경증	<50	<60	<69	<79	<91	<104	<116	<130	<141	141<
	외향성	<70	<82	<93	<102	<112	<119	<128	<136	<144	144<
	개방성	<75	<87	<99	<108	<117	<126	<136	<144	<155	155<
	친절성	<86	<95	<105	<113	<122	<130	<138	<145	<152	152<
	성실성	<73	<85	<95	<106	<117	<126	<135	<147	<156	156<

주. 기준점수를 제시하는 방법에는 여러 가지 방법이 있겠지만 본 연구에서는 'sten'을 사용한다.

성한 사람들이 조직에 입사한 후의 직무성과와의 관계를 밝히는 연구로 발전되지 못했다는 점을 들 수 있다. 성격검사를 인력선발에 사용하기 위해서는 구성요인 타당도만큼이나 예언타당도도 중요하다. 성격검사가 인력선발, 인력배치, 직무성과, 그리고 역량개발 등에 있어서 구체적이고 통합적인 정보를 제공할 수 있을 때, 성격검사의 유용성은 더 배가될 수 있다. 하지만, 21,349명 중에서 정식 계약까지 성공한 지원자의 비율이 10% 정도 수준이고, 입사 초기에 내부 이동도 잦아서 이들을 지속적으로 추적해서 예언타당도 연구를 진행하는데에는 현실적인 어려움이 많아서 실행하지 못하였다. 따라서 향후 진행되는 연구에서는 좀 더 세밀한 연구 설계를 통하여, 선발단계에서 확인된 요인구조의 안정성을 바탕으로 실제 직무에서의 성과 및 경력경로와의 관계까지 분석하는 통합적인 연구모형을 제시하는 것이 필요하다.

또한 본 연구의 방법론과 관련하여, 구조방정식모형과 프로크루스테스 요인분석을 사용함에 있어 좀 더 엄격한 기준을 사용하지 못했다는 것이다. 두 집단에서의 불변성을 검증하기 위해서는 다집단 확인적요인분석을 사용하여 측정불변성검증(Measurement Invariance test)을 할 수도 있지만, 비교 대상인 미국 기준집단, 프랑스 기준집단, 스페인 기준집단의 원자료를 얻을 수 없었기 때문에 기존 연구의 확인적 요인분석의 결과 비교만 실시하였다. 이와 같은 이유로 프로크루스테스 요인분석에 있어서도 기존 연구결과들과의 비교를 위해 비교집단에서 사용한 방법론을 동일하게 실시하였다.

결론적으로 기존의 연구들이 주로 실험실 상황에서 설정에 의해서 실행되었거나 실제

장면에서 실행된 연구라도 표본의 규모가 작은 연구가 대부분이었기 때문에, 본 연구는 기존의 연구들을 다양한 직무에 지원한 대규모 실제 입사지원자를 대상으로 좀 더 명확하게 증명하였다는 것에 차별성이 있다고 할 수 있다. 최근 우리나라의 많은 대기업들이 성격검사를 인사선발에서 활용하고 있는데, 그러한 자료를 활용한 연구는 찾아보기 힘들다. 따라서 향후 본 연구의 한계점은 보완하면서 우리나라에서의 대규모 실제 입사지원자 표본을 대상으로 한 연구를 실행하게 된다면 본 연구에서 발견된 내용들을 좀 더 발전시킬 수 있을 것으로 본다.

참고문헌

- 유태용 역. (2009). *산업 및 조직심리학*(9판). 서울: 시그마 프레스.
- Aluja, A., Garcia, O., & Garcia, L. F. (2004). Replicability of the three, four and five Zuckerman's personality super-factors: Exploratory and confirmatory factor analysis of the EPQ-RS, ZKPQ and NEO-PI-R. *Personality and Individual Differences*, 36(5), 1093-1108.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1-26.
- Barrick, M. R., & Mount, M. K. (1996). Effects of Impression Management and Self-Deception on the validity of personality constructs. *Journal of Applied Psychology*, 81(3), 261 - 272.
- Birkeland, S. A., Manson, T. M., Kisamore, J. L.,

- Brannick, M. T., & Smith, M. A. (2006). A Meta-Analytic Investigation of Job Applicant Faking on Personality Measures. *International Journal of Selection and Assessment*, 14(4), 317-335.
- Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5 - Factor model of personality. *Personality and Individual Differences*, 11, 515-524.
- Borman, W. C., Hanson, M. A., & Hedge, J. W. (1997). Personnel selection. *Annual Review of Psychology*, 48, 299-337.
- Caldwell, D. F., & Burger, J. M. (1998). Personality characteristics of job applicants and success in screening interviews. *Personnel Psychology*, 51, 119-136.
- Collins, J. M., & Gleaves, D. H. (1998). Race, job applicants, and the Five-Factor Model of Personality: Implications for Black psychology industrial/organizational psychology, and the Five-Factor theory. *Journal of Applied Psychology*, 83, 531-544.
- Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self reports and spouse ratings on the NEO-PI. *Journal of Personality and Social Psychology*, 54, 853-863.
- Costa, P. T., & McCrae, R. R. (1992). *Professional manual for the NEO Personality Inventory (NEO PI-R) and NEO Five Factor Inventory (NEO-FFI)*. Odessa: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1999). *Inventario de Personalidad NEO revisado (NEO PI-R) e Inventario NEO reducido de cinco factores (NEO-FFI) (NEO PI-R, Revised NEO Personality Inventory and NEO Five-Factor Inventory (NEO-FFI))*. Madrid: TEA Editions.
- Costa, P. T., McCrae, R. R., & Rolland, J. P. (1998). *Manuel de l'Inventaire NEO PI-R*, Paris: ECPA.
- DeFruyt, F. (2001). Faut-il prendre en compte la personnalité dans les décisions de recrutement et de sélection: Le rôle particulier de la dimension «Conscience». In C. Lévy-Leboyer, M. Huteau, C. Louche, J.-P. Rolland, *La psychologie du travail: RH, les apports de la psychologie du travail*, Paris: Editions d'Organisation.
- DeFruyt, F., Aluja, A., García, L. F., Rolland, J. P., & Jung, S. C. (2006). Positive Presentation Management and Intelligence and the Personality Differentiation by Intelligence Hypothesis in Job Applicants. *International Journal of Selection and Assessment*, 14(2), 101-112.
- Digman, J. M. (1996). The Curious History of the Five-Factor Model. In J. S. Wiggins(Ed.), *The Five-Factor Model of Personality: Theoretical Perspectives*, 1-20, New York: Guilford.
- Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate Behavioral Research*, 16(2), 149-170.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of*

- Applied Psychology*, 84, 155-166.
- Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology*, 86, 122-133.
- Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and Individual Differences*, 7, 385-400.
- Furnham, A. (1990). Faking Personality Questionnaires: Fabricating Different Profiles for Different Purposes. *Current Psychology: Research & Review*, 9(1), 46-55.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26-34.
- Hogan, J. (1998). Personality and Job Performance. *Human Performance*, 11(2/3), 125-127.
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92(5), 1270-1285.
- Hogan, J., & Hogan, R. (1989). To measure employee reliability. *Journal of Applied Psychology*, 74, 273-279.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socio-analytic perspective. *Journal of Applied Psychology*, 88(1), 100-112.
- Hogan, R. (2006). *Personality and the fate of organizations*. Mahwah, NJ: Erlbaum.
- Hogan, R. T., DeFruyt, F., & Rolland, J. P. (2006). Validité et intérêt des méthodes d'évaluation de la personnalité à des fins de sélection: une perspective de psychologie appliquée aux problématiques des entreprises. *Psychologie française*, 51(3), 245-264.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality Measurement and Employment Decisions: Questions and Answers. *American Psychologist*, 51(5), 469-477.
- Holden, R. R., & Fekken, G. C. (1989). Three common social desirability scales: Friends, acquaintances, or strangers. *Journal of Research in Personality*, 23, 180-191.
- Hough, L. M. (1992). The 'Big Five' Personality Variables-Construct Confusion: Description Versus Prediction. *Human Performance*, 5(1/2), 139 - 156.
- Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future - Remembering the past. *Annual Review of Psychology*, 51, 631-664.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big five revisited. *Journal of Applied Psychology*, 85(6), 869-879.
- John, O. P. (1990). The 'Big Five' factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (66 - 100). New York: Guilford.
- John, O. P., & Srivastava S. (1999). The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research*, 2nd Edition, 102-138. New York: Guilford.
- Lanyon, R. I., & Goodstein, L. D. (1997). *Personality Assessment*, 3rd Edition, New York: John Wiley and Sons.
- Krahe, B., Becker, J., & Zolter, J. (2008).

- Contextual cues as a source of response bias in personality questionnaires: The case of the NEO-FFI. *European Journal of Personality*, 22, 655-673.
- Mahar, D., Cologon, J., & Duck, J. (1995). Response strategies when faking personality questionnaires in a vocational selection setting. *Personality and Individual Differences*, 18, 605-609.
- Marcus, B. (2009). 'Faking' From the Applicant's Perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection and Assessment*, 17(4), 417-430.
- Marshall, M., DeFruyt, F., Rolland, J. P., & Bagby, R. M. (2005). Socially desirable responding and the factorial stability of the NEO-PI-R. *Psychological Assessment*, 17, 379-384.
- Martin, B. A., Bowen, C. C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and Individual Differences*, 32, 247-256.
- McAdams, D. P. (2001). *The Person, 3rd Edition*. New York: Harcourt College Publishers.
- McCrae, R. R., Zonderman, A. N., Costa, P. T., Bond, M. H., & Paunonen, S. P. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70(3), 552 - 566.
- Montag, I., & Comrey, A. L. (1990). Stability of major personality factors under changing motivational conditions. In J. W. Neuliep(Ed.), *Replication research in the social sciences*, 253-262. Newbury Park, CA: Sage.
- Montag, I., & Levin, J. (1994). The five-factor personality model in applied settings. *European Journal of Personality*, 8, 1-11.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of Social Desirability in Personality Testing for Personnel Selection: The Red Herring. *Journal of Applied Psychology*, 81(6), 660-679.
- Ones, D. S., Viswesvaran, C., & Schmidt, E. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679-703.
- Pauls, C. A., & Crost, N. W. (2005). Effects of different instructional sets on the construct validity of the NEO-PI-R. *Personality and Individual Differences*, 39, 297-308.
- Peabody, D., & Goldberg, L. R. (1989). Some Determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology*, 57(3), 552-567.
- Pervin, L. A., & John, O. P. (2001). *Personality: Theory and Research, 8th Edition*. New York: John Wiley & Sons.
- Rolland, J. P. (2001). Validité interculturelle du modèle de personnalité en cinq facteurs. *Psychologie Française*, 46(3), 231-249.
- Rolland, J. P. (2003). *NEO PI-R: Inventaire de Personnalité-Révisé, Additif*. ECPA: Paris.
- Rolland, J. P. (2004). *L'évaluation de la personnalité*. Sprimont, Mardaga.
- Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection

- practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52(2), 359-391.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European community. *Journal of Applied Psychology*, 82, 36-43.
- Salgado, J. F. (2002). The Big Five Personality Dimensions and Counterproductive Behaviors. *International Journal of Selection and Assessment*, 10(1/2), 117-125.
- Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for Personnel Selection: An Overview of Constructs, Methods and Techniques. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *International Handbook of Work and Organizational Psychology, Vol. 1.*, London: Sage.
- Sandal, G. M., Musson D., Helmreich, R. L., & Gravdal, L. (2005). Social desirability bias in personality testing: Implications for astronaut selection. *Acta astronautica*, 57, 634-41.
- Schmit, M. J., & Ryan, A. M. (1993). The big five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, 78, 966-974.
- Smith, D. B. (1996). *The Big Five in personnel selection: Reexamining frame of reference effects.* Unpublished master thesis, University of Maryland, College Park.
- Terracciano, A. (2003). The Italian version of the NEO-PI-R: Conceptual and empirical support for the use of targeted rotation. *Personality and Individual Differences*, 35, 1859 - 1872.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703 - 742.
- Topping, G. D., & O'Gorman, J. G. (1997). Effects of faking set on validity of the NEO-FFI. *Personality and Individual Differences*, 23(1), 117-124.
- Tsaousis, I., & Nikolaou, I. (2001). The Stability of the Five-Factor Model of Personality in Personnel Selection and Assessment. *International Journal of Selection and Assessment*, 9, 290-301.
- Viswesvaran, C., & Ones, D. S. (1999). Meta-analysis of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59, 197-210.
- Wiggins, J. S., & Pincus, A. L. (1989). Conceptions of personality disorder and dimensions of personality. *Psychological Assessment*, 1, 305 - 316.
- Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review*, 12, 250-264.
- 1차 원고접수 : 2010. 7. 19
2차 원고접수 : 2010. 10. 25
3차 원고접수 : 2011. 1. 23
4차 원고접수 : 2011. 2. 14
최종계재결정 : 2011. 2. 18

Comparison of factor scores and factor structures of NEO-PI R between job applicants and normative group

Sung-Cheol Jung

The Catholic University of Korea, Department of Psychology

The present study aimed to analyse the extent to which Positive Presentation Management influences on factor scores and factor structure of NEO-PI R personality inventory in the context of personnel selection. Subjects of present study are 21,349 applicants who applied for a French public transport organization and filled out NEO-PI R personality inventory during 5 years. Comparisons with French norm showed significant differences in scores of all of five factors: scores of Neuroticism and Openness are lower than French norm (deflation of factor scores) and those of Extraversion, Agreeableness, and Conscientiousness are higher than French norm (inflation of factor scores). Confirmatory factor analysis and Procrustes rotation demonstrated that variations of factor scores due to Positive Presentation Management did not affect 5-dimensional factor structure. From these results, we can conclude that applicants' Positive Presentation Management affect factor scores but does not affect structural validity of NEO-PI R. We further proposed a norm of NEO-PI R scores for individuals in contexts of high-stakes testing.

Key words : personality, FFM(Five Factor Model), positive presentation management, social desirability, personnel selection, applicant, structural validity, faking