

자기보고식 성격측정도구의 재검사 신뢰도에서의 문화적 차이: 한국과 독일의 비교*

김 혜 은

김 수 정†

목포대학교 교육학과

본 연구의 목적은 문화적 특성에 따른 자기보고식 성격측정도구의 재검사 신뢰도의 차이를 검증하는 것이다. 일반적으로 동아시아인들이 관계적 맥락에 의해 영향을 많이 받기 때문에 심리적 특성의 일관성이 갖는 의미가 상대적으로 작다는 선행연구 결과들에 근거하여 한국과 독일의 대학생을 대상으로 상호의존성의 차이에 기초하여 자기보고식 성격측정도구의 재검사 신뢰도를 비교하였다. 조사에는 한국 대학생 257명과 독일 대학생 196명이 참여하였는데, 이들을 상호의존적 경향을 기준으로 각각 다시 두 집단으로 구별하였다. 측정도구는 NEO-PI-R 중 외향성요인의 온정과 사교성, 친화성요인의 이타성과 겸손, 성실성요인의 충실성과 자기규제 척도이며, 1개월 간격으로 2회 실시하여 재검사 신뢰도를 측정하였고, 이를 Fisher's Z로 환산한 후 일방향 z검증을 하였다. 그 결과 이타성 척도에서 한국과 독일의 상호의존성이 높은 집단이 낮은 집단보다 재검사 신뢰도가 유의하게 낮게 나타났으며, 겸손, 충실성, 자기규제 척도에서는 상호의존성이 높은 경우에 한국집단에서 독일집단보다 재검사 신뢰도가 유의하게 낮게 나타났다. 본 연구결과를 기초로 재검사 신뢰도의 문제를 상호의존성에 관련지어 논의하였다.

주요어 : 재검사 신뢰도, 자기보고식 심리측정, NEO-PI-R, 상호의존성

* 본 논문은 2013년도 목포대학교 교내연구비 지원에 의하여 연구되었음.

† 교신저자: 김수정, 목포대학교 사범대학 교육학과, (534-729) 전남 무안군 청계면 영산로 1666

E-mail: lmu30576@hanmail.net

성격특질이론은 개개인의 상대적으로 안정된 성격적 특성을 기반으로 다양한 상황에서 비교적 일관된 인간의 행동을 예측할 수 있을 것이라는 것을 전제한다(McCrae & Costa, 1994; Johnson, 1997). 이를 기초로 심리학의 제 분야에서는 성격측정도구를 활용하여 개인의 심리적 성향, 습관, 행동유형의 체계를 파악하려고 시도한다. 또한 심리측정을 통해 제반 심리적 특성 요인들의 관계를 규명함으로써 이론적 검증 뿐 아니라 진단과 심리치료의 효과를 분석하는 것을 가능하게 한다는 점에서 임상현장에서의 의미도 매우 크다(윤성민, 신희철, 2013; Chmielewski & Watson, 2009).

이와 같이 성격특질이론에서는 안정적인 심리적 특성에 기초한 개인의 성향을 탐색한다. 예를 들어 Schmitt, Allik, McCrae와 Benet-Martinez(2007)은 Costa와 McCrae(1992)의 NEO-PI-R을 활용하여 56개국에서 외향성, 친화성, 성실성, 신경증, 경험에 대한 개방성의 5요인 구조를 확인하였고 이를 기초로 NEO-PI-R 성격검사를 다양한 문화권에서 보편적으로 사용하는 것이 가능하다고 주장하였다.

한편 사회심리학자들은 개인의 성격적 특질 즉 성향 외에도 상황에 관심을 가지며 상황적 조건이 개인에게 어떤 영향을 미치는가에 대해서 탐색을 한다. 더 나아가 문화비교 심리학자들은 인간 성격의 일반적 범칙성보다는 각 문화권에서의 특성에 관심을 갖는다. 또한 성격이 사회적으로 구성된다는 측면에서 모든 문화에서의 보편성과 행동의 예측가능성에 대한 회의를 제기한다(Markus & Kitayama, 1998; Schweder, 1991).

예를 들어 문화비교 심리학자들은 집단주의 문화권에서 성격검사의 예언타당도가 더 낮을 수 있다고 주장한다. 이에 대한 근거로 개인주의 문화권에 비해 집단주의 문화권에서는 상황의 영향력이 크기 때문에 성격특질이 갖는 의미가 상대적으로 작고(Triandis, 1995), 특히 동아시아 사람들의 자아개념에는 일관성의 기능적인

가치가 중요하지 않다는 주장이 제기되었다(Heine, 2001). Kanagawa, Cross와 Markus(2001)는 일본인들이 다양한 관계적 맥락에서 자신의 능력, 역할 태도 등을 서술하는 내용에 있어 미국인들보다 관계적 맥락에 따라 더 차이를 보였다고 하였다. 마찬가지로 English와 Chen(2007)은 아시아계 미국인들이 유럽계 미국인들보다 인간관계와 관련된 상황에 따라 성격검사 결과에서의 일관성이 낮았다고 보고하였다. Church, Katigbak, Reyes, Salanga, Miramontes와 Adams(2008)의 연구에서도 유럽계 미국 대학생들보다 필리핀 대학생들에게서 관계적인 맥락에 따라 행동이 더욱 차이가 있는 것으로 나타났다. 이와 같은 연구 결과들에 대해서 몇몇 학자들이 아시아 문화권에서는 상황에 관계없이 일관적으로 행동하는 사람들을 사회적 맥락에 제대로 반응을 하지 못하는 것으로 해석하고, 이러한 경향을 오히려 융통성이 낮고 아직 성숙하지 못한 태도로 생각한다고 설명한 것과 연관 지을 수 있다(Kanagawa, Cross, & Markus, 2001; Markus & Kitayama, 1994).

한편 성격검사에서의 일관성이 낮다는 것이 심리측정 차원에서는 검사도구의 신뢰도가 낮은 것을 의미할 수 있다는 점에서 또 다른 차원에서의 논쟁점을 제기한다. 즉 심리측정 이론에서 신뢰도는 반복적인 측정과 다양한 상황에서의 일관성을 의미하는데(Cattell, 1986; Murphy & Davidshofer, 2001), 일관성이 낮다는 것을 실제적인 변화와 측정오류 모두로 해석할 수 있다는 것이다. 따라서 문화비교 심리학자들이 확인한 것과 같이 아시아 문화권에서 성격의 일관성이 낮게 나타나는 것은 검사도구의 낮은 신뢰도의 원인이 될 수도 있다. 그럼에도 이런 문제가 심각하게 논의되고 있지 않은 이유는 심리학 연구들에서 신뢰도지수로 흔히 Cronbach 알파 지수만 보고되고 있는 것과 무관하지 않을 것이다(Cortina, 1993). 심리측정의 일반화 이론에 의하면 신뢰도는 문항들 간의 일관성, 상황의 일관성, 검사자들 간의 일관성을 포함한다(John &

Soto, 2007). 이들은 서로 다른 측정오류를 분석하기 때문에 신뢰도의 내용이 상이하다. 예를 들어 Cronbach 알파는 자기보고식 검사도구의 어느 한 측정상황에서의 문항들 간의 일치도를 측정한다(Green, 2003; John & Soto, 2007). 따라서 Cronbach 알파는 이질적인 내용의 검사나 다차원 검사에서의 신뢰도는 낮게 평가하는 반면에 체계적인 측정오류가 있는 경우에는 신뢰도를 높게 평가하는 경향을 보인다(Cronbach, 1951). 즉 내적 일치도를 나타내는 알파계수는 단 한 번의 조사로 산출된다는 점에서 일시적 오류 즉 조사상황의 차이, 조사대상의 변화가 반영되지 않아 신뢰도가 과대평가될 수 있다는 지적이 있다. 특히 이런 오류가 조사를 하려고 하는 심리적 구성과 긍정적인 상관관계를 가질 때에는 더욱 문제가 될 수 있다(Osburn, 2000).

한편 Cronbach 알파와 재검사 신뢰도의 상관계수는 평균 $r=.25$ 로 높지 않다고 보고되고 있어 Cronbach 알파와 재검사 신뢰도는 개념적으로나 경험적으로 상이한 신뢰도를 측정하고 있음을 알 수 있다(Chmielewski & Watson, 2009). 신뢰도에 영향을 줄 수 있는 요인으로 Cronbach 알파에는 문항내용의 상호관련성, 다차원성, 모호성 외에도 피검사자 오류, 표집집단 분산이 있다. 반면에, 재검사 신뢰도에는 문항내용의 모호성, 피검사자 오류와 표집집단 분산 외에도 상태의 변화가 영향을 미치는 것으로 보고되었다(MaCrae, Kurtz, Yamagata, & Terracciano, 2011). 따라서 Schmidt, Lee와 Ilies(2003)는 Cronbach 알파와 재검사 신뢰도를 모두 보고할 것을 제안하고 있다.

측정오류를 제거하지 못할 경우 심리적 특성의 측정이나 이와 다른 변수들의 관계를 분석하는데 오류를 제공한다는 점에서 신뢰도지수의 정확한 활용은 중요한 의미를 갖는다(Watson, 2004). 신뢰도 지수로 Cronbach 알파가 가장 많이 보고되는 이유는 편리성 때문일 것이다(John & Soto, 2007). 그럼에도 MaCrae 등(2011)은 재검사 신뢰도가 Cronbach 알파보다 타당도를 더 유의하

게 예측한다고 하였는데 따라서 재검사 신뢰도에 대한 관심을 더욱 가질 필요가 있다. 한편 재검사 신뢰도는 측정하는 것이 번거롭고 검사내용이 기억될 수 있다는 문제점이 흔히 지적되고 있다. 그런데 앞서 언급한 바와 같이 사회적 맥락의 영향과 같은 문화적인 요인에 기인하여 심리측정에서의 일관성이 낮고 따라서 재검사 신뢰도가 더 낮게 나오는 문화권이 있을 수 있다는 점에서 주의를 기울일 필요가 있다.

이미 서술한 바와 같이 아시아권에서 서구문화권보다 자아에 대한 기술이나 인간관계와 관련된 성격검사에서의 일관성이 떨어진다는 연구결과를 기초로 신뢰도 지수에 영향을 미칠 수 있는 문화적 맥락의 탐색을 위해 집단주의를 설명하는 이론적 틀을 적용할 수 있을 것이다. 이는 본 연구에서 탐색하려고 하는 심리측정에서의 안정성과 관련하여 다음과 같은 문화적 특성을 정리할 수 있기 때문이다. 즉 집단주의 문화권에서는 개인주의 문화권보다 내집단과 외집단의 차이(Triandis, 1995)나 사적영역과 공적영역의 구별이 크며(Newman, 1993), 위계질서나 사회적 영향에 민감하며 체면을 중요시한다. 또한, 상황에 적합하게 행동을 하려는 경향이 강하며 의사소통에 있어 맥락을 강조한다(Noguchi, 2007). 따라서 관계적 맥락에 따라 행동의 차이를 보이는 경향이 클 수 있다.

같은 맥락에서 자아의 특성을 다른 사람들과의 관계적인 맥락에서 보려고 하는 상호의존적 자아상에 대한 이해를 통해 성격측정 상황에서의 일관성에 미치는 영향을 탐색할 수 있다. 예를 들어 Markus와 Kitayama(1991)는 자아표상으로 문화적 차이를 설명하는데, 서구 문화권의 사람들에게서는 독립성과 내적 특성, 개성을 중요시하는 독립적 자아상이 상대적으로 더 많이 강조된다고 하였다. 반면에 비서구 문화권의 사람들에게서는 관계, 사회적 맥락을 중요시하는 상호의존적 자아상이 강조된다고 설명하였다. 이와 같이 자아표상의 차이로 접근하는 시도는 각 문화권에서의 개인적 수준의 차이를 좀 더 쉽게

설명할 수 있다는 장점을 갖고 있다. 그럼에도 독립적 자아와 상호의존적 자아의 두 개의 요인으로 구별하는 것에 대해 심리측정 차원에서 비판적인 연구결과들이 있어(Fiske, 2002; Levine 등, 2003), 본 연구에서는 한국과 서구의 문화차이를 탐색하는데 있어 상호의존적 자아상에 한정하여 접근하려고 한다. 또한 한국에서 독립적 자아상과 상호의존적 자아상이 모두 높게 나타났다고 보고되고 있는 것도 고려하였다(한규석, 2009).

Singelis(1994)에 의하면 정도의 차이는 있으나 상호의존적 자아가 강한 사람들은 유동적이고 가변적인 자아 특성을 가지며 사회적 지위, 역할, 관계와 같이 외적이고 공적인 특성을 중요하게 여긴다. 또한 조화로운 대인관계와 다양한 상황에서의 적응능력을 중요하게 생각하기 때문에 다른 사람과의 관계를 고려하며 상황에 맞게 행동하려는 경향이 강하다. 따라서 주어진 상황에 적절하게 반응하지 못하는 것은 미성숙한 것이며 자신의 역할이나 위치에 맞게 행동하는 것이 개성을 잃어버리는 것이 아니라고 생각한다(Kanagawa, Cross, & Markus, 2001). 이와 같이 상호의존적 자아는 다수의 미결정적인 유동적 자아로 이루어지는 경향을 가지며, 상황과 타인과의 관계가 상호의존적 자아상에 능동적이고 지속적으로 통합되며 발달한다(Kondo, 1990; Lebra, 1994).

Kitayama, Park, Sevincer, Karasawa와 Uskul(2009)에 의하면 상호의존적인 자아가 강한 사람은 상황에 민감한 판단과 결정을 하기 때문에 상황귀인을 하는 경향이 강하다. Markus, Kitayama, Mullanly, Masuda, & Fryberg(1997)의 연구에서도 일본 학생들이 자신을 묘사할 때 상황을 언급하는 경우가 두 배 이상이 되는 것으로 보고되었다. 상호의존적인 자아가 강한 사람들에게는 성격검사에서도 같이 구체적인 상황이 제시되지 않고 일반적이고 추상적인 차원에서 자신의 성격을 판단해야 하는 경우 검사문항에 응답하는 것이 수월하지 않을 수 있다. 유럽계 미국 학생

들이 아시아계 학생들에 비해 성격검사에 응답을 하는 시간이 더 빨랐다는 Markus 등(1997)의 연구결과도 같은 맥락에서 이해할 수 있다.

이상의 선행 연구결과들을 종합하면 상호의존적 자아가 강한 사람들의 경우 상황이나 사회적 맥락에 따라 심리측정 결과의 일관성이 낮게 나타날 것으로 예상되고 신뢰도 특히 재검사 신뢰도가 더 낮게 나타날 수 있다는 문제를 제기할 수 있다. Markus & Kitayama(1998)도 상호의존적 자아가 강한 사람들의 행동이 상황에 따라 다르다면 이를 측정하는 도구에서의 신뢰도와 예측타당도의 문제가 있을 수 있다는 지적을 하였다. 그럼에도 이에 대한 경험적 연구는 거의 없는 실정이다. 물론 본 연구의 목적과 유사한 연구로 Oishi, Diener, Scollon과 Biswas-Diener(2004)가 미국과 일본에서 다양한 사회적 맥락에서의 정서경험의 일관성을 분석한 연구와 Church 등(2008)이 성격적 특성과 관련된 행동의 일관성을 검증한 연구가 있다. 그러나 자기보고식 성격검사 도구를 활용하여 문화적 맥락에 따른 재검사 신뢰도의 문제점을 직접적으로 다룬 연구는 아직 알려진 것이 없다.

본 연구에서는 재검사 신뢰도의 문제를 상호의존적 자아상과 연결하여 한국과 독일집단을 대상으로 탐색하려고 한다. 이처럼 한국과 독일집단을 비교하는 이유는 문화적 가치를 기준으로 73개 국가의 문화를 7개의 문화군으로 분류하고 비교를 한 Schwartz(2006)의 연구에서 한국을 포함한 유교문화권과 독일을 포함한 서유럽 문화권에서 문화적 가치에 있어 가장 대비를 이루고 있기 때문이다. 또한 본 연구에서와 같은 접근은 Matsumoto(2003)와 Owe 등(2013)이 주장한 것처럼 문화비교를 할 때 개인적 수준(상호의존성의 차이)과 거시적 수준(국가 간 비교)에서 모두 분석이 이루어질 필요성이 있다는 점을 고려한 것이다.

본 연구에서 설정한 연구문제는 '상호의존적 자아상에 따라 자기보고식 심리측정도구의 재검사 신뢰도의 차이는 어떠한가' 이러한 차이는 한

국과 독일에서 어떠한 양상을 보이는가?’이다.

방 법

연구대상

본 연구에는 한국의 대학생 321명과 독일 대학생 220명이 참가하였고 한국과 독일 모두 교직과정 수업시간에 조사가 이루어졌다. 이들은 모두 3-6학기 학생들로 1학년과 4학년 학생들을 대상으로 조사하지 않은 이유는 입학과 졸업 등으로 인한 주거환경의 변화, 취업, 교우관계의 변화 등과 같은 생활사 경험이 재검사 신뢰도를 측정하는데 불리할 수 있다는 사실을 고려하였기 때문이다(Chmielewski & Watson, 2009). 또한 재검사 간격이 실제의 변화가 발생할 수 있는 기간보다 짧고 문항과 응답내용을 기억하지 않도록 해야 한다는 점을 고려하여 1개월 간격으로(Schmidt 등, 2003; Watson, 2004), 두 차례에 걸쳐 조사가 이루어졌다. 전체 문항들의 응답 시간은 25분 정도 소요되었다.

한편 응답자의 익명성을 보호하면서 1차 조사와 2차 조사에서 동일 인물임을 확인할 수 있도록 하기 위해 설문지에 출생지, 어머니의 이름과 자신의 생일을 조합하여 표시하도록 하였다. 1차 조사에는 총 541명이 그리고 2차 조사에는 538명이 참가하였다. 두 차례에 걸쳐 설문지를 수거한 후 동일한 응답자의 설문지로 확인되지 못한 88부를 제외하였으며 본 연구의 최종적인 자료처리 대상은 한국에서 257명(여자: 145명, 남자: 112명), 독일에서 196명(여자: 109명, 남자: 87명)이었다.

측정도구

자기보고식 성격측정 도구로 Costa와 McCrae (1992)가 개발하고 Ostendorf와 Angleitner(2004)가 독일어로 번역하여 표준화 작업을 한 NEO-PI-R

을 활용하였다. 이를 연구자가 한국어로 번역하고 독문학 전공자가 역 번역을 한 후 차이가 있는 경우 독일의 심리학 전공 교수와 논의를 하여 수정하였다. NEO-PI-R은 5요인설에 기초하여 신경증(Neuroticism), 외향성(Extraversion), 경험에 대한 개방성(Openness to Experience), 친화성(Agreeableness), 성실성(Conscientiousness)을 측정하는 다면적 성격측정 도구이다. 한편 이들 5개 요인은 각각 6개의 척도로 구성되었고, 각 척도는 8개 문항으로 구성되었다.

일반적으로 5요인 중 신경성, 외향성, 친화성은 대인관계와 관련된 성격특질로, 그리고 성실성과 경험에 대한 개방성은 지적인 활동을 요구하는 과제와 관련된 성격특질로 분류된다(Oishi 등, 2004). 본 연구에서는 상호의존성에 따른 재검사 신뢰도의 문화적 차이가 어떠한 양상을 보이는가를 탐색하기 위하여 대인관계와 관련된 요인 중 외향성과 친화성을 선택하고 지적인 과제와 관련된 성격특질인 성실성을 선택하여 이들 중 각 2개의 척도로 재검사 신뢰도를 비교하려고 한다. 이로써 본 연구에 투입된 척도는 외향성요인의 온정과 사교성, 친화성요인의 이타성과 겸손, 그리고 성실성요인의 충실성과 자기규제로 모두 6개 척도의 자기보고식 문항 48개이다. 이들 척도는 Yoon, Schmidt와 Ilies(2002)가 한국인들을 대상으로 한 NEO-PI의 타당성 연구에서 내적 일치도나 요인부하량이 비교적 높은 것으로 보고되어서 본 연구에서 재검사 신뢰도를 분석하는데 적합한 것으로 판단되었다. 한편 신뢰도가 문항 수에 의해 영향을 받을 수 있다는 점을 고려할 때 NEO-PI-R의 모든 하위영역이 8개 문항으로 동일하다는 사실은 신뢰도 비교를 하는데 장점이라고 할 수 있다(McCrae 등, 2011). 이들은 모두 5점 척도로 제시되었다.

상호의존성을 측정하기 위해서는 Singelis(1994)의 Self-Construct Scale중 상호의존적 자아와 관련된 12개 문항을 7점 척도로 제시하였다. 본 연구의 측정도구의 Cronbach 알파지수는 다음과 같다.

표 1. 측정시기에 따른 조사도구의 Cronbach 알파지수

척도	한국		독일		전체	
	1차	2차	1차	2차	1차	2차
온정	.71	.74	.70	.76	.70	.74
사교성	.83	.84	.82	.84	.82	.84
이타성	.78	.75	.84	.84	.82	.80
겸손	.69	.70	.78	.79	.75	.74
충실성	.68	.69	.72	.76	.71	.74
자기규제	.81	.80	.90	.90	.87	.87
상호의존	.76	.75	.68	.68	.74	.73

본 연구에서 확인한 Cronbach 알파지수를 Nunnally & Bernstein(1994)에 의해 제시된 기준 .70과 비교했을 때 표1에서 보는 바와 같이 전반적으로 양호한 값을 보이고 있다. 특히 본 연구에서 활용된 각 척도의 Cronbach 알파지수 값은 McCrae 등(2011)과 안창호, 채준호(1997)의 연구에서 제시된 대학생을 대상으로 측정된 값에 비해 대체로 더 높은 것으로 확인되었다.

가설

본 연구에서는 한국과 독일의 대학생을 대상으로 상호의존적 경향의 차이에 따라 성격측정도구의 재검사 신뢰도의 차이가 있는가를 분석하려고 한다. 동아시아인들이 상황적 요인에 의해 영향을 많이 받기 때문에 심리적 특성의 일관성이 갖는 의미가 상대적으로 작다는 선행연구 결과들에 근거하여 다음과 같은 가설이 설정되었다.

1. 상호의존적 경향이 강한 집단에서는 그렇지 않은 집단에 비해 재검사 신뢰도가 더 낮게 나타날 것이다.
2. 독일집단에 비해 한국집단에서 재검사 신뢰도가 더 낮게 나타날 것이다.

자료분석

집단 간 재검사 신뢰도를 비교하기 위해 우선 각 척도의 1차와 2차 조사에서의 점수 간의 상관계수를 산출하였고, 이들을 Fisher's Z값으로 전환한 후 가설에 근거하여 일방향 z검증을 하였다. 상호의존적 경향에 따른 집단의 구별은 상호의존적 자아상의 중앙값을 기준으로 이루어졌는데 이를 위해 1차 조사와 2차 조사에서 동일하게 나타난 중앙값 4.75을 기준으로 두 집단을 나누어 상위집단과 하위집단으로 구별하였다. 이로써 한국집단에서는 상위집단이 149명, 하위집단이 108명이고 독일집단에서는 상위집단이 73명 하위집단이 123명이었다.

결 과

측정도구인 NEO-PI-R의 6개 척도의 재검사 신뢰도를 상호의존성과 국가에 따른 차이로 검증한 결과는 다음과 같다. 먼저 외향성요인의 온정과 사교성 척도와 관련된 결과를 살펴보면, 표 2에서 보는 바와 같이 전체집단의 온정과 사교성의 재검사 신뢰도는 각각 .80과 .81로 나타났다.

이들 재검사 신뢰도에 대한 집단 간 비교를 한 결과 온정척도에서 상호의존성에 따른 차이를 검증한 결과는 $z=.95$ 로 일방향 검증에서의 $z_{95\%}=-1.65$ 를 기준으로 판단했을 때 상위집단이 하위집단보다 재검사 신뢰도가 낮게 나타나지 않았다. 국가 간의 차이를 검증한 결과도 $z=-.84$ 로 한국집단의 재검사 신뢰도는 독일집단의 사교성 재검사 신뢰도보다 유의하게 낮게 나타나지 않았다.

사교성 척도에서는 상호의존성에 따른 차이를 검증한 결과 $z=1.23$ 으로 일방향 검증에서의 $z_{95\%}=-1.65$ 를 기준으로 판단했을 때 상위집단이 하위집단보다 재검사 신뢰도가 낮게 나타나지 않았다. 국가 간의 차이를 검증한 결과도 $z=-.31$

표 2. 각 척도의 재검사 신뢰도 값

		온정	사교성	이타성	겸손	충실성	자기규제
한국	상	.80	.83	.69	.67	.65	.79
	하	.74	.79	.79	.76	.76	.83
	전체	.79	.81	.77	.70	.69	.82
독일	상	.86	.85	.69	.82	.82	.87
	하	.80	.80	.82	.80	.80	.86
	전체	.82	.82	.79	.81	.81	.87
전체	상	.81	.83	.70	.77	.73	.83
	하	.78	.79	.81	.81	.79	.85
	전체	.80	.81	.78	.79	.78	.85

로 한국집단의 재검사 신뢰도는 독일집단의 사교성 재검사 신뢰도보다 유의하게 낮게 나타나지 않았다.

친화성요인의 경우 전체 집단의 이타성과 겸손 척도의 재검사 신뢰도는 각각 .78과 .79였다. 이들 재검사 신뢰도에 대한 상호의존성과 국가에 따른 집단 간 비교를 한 결과 이타성의 경우 국가 간의 비교에서는 $z=-.54$ 로 한국집단의 이타성 재검사 신뢰도는 독일집단의 재검사 신뢰도보다 유의하게 낮게 나타나지 않았다. 반면에 상호의존성에 따른 차이를 검증한 결과는 $z=-2.75$ 로 일방향 검증에서의 $z_{99\%}=-2.33$ 을 기준으로 판단했을 때 $p<.01$ 수준에서 상위집단이 하위집단보다 재검사 신뢰도가 유의하게 낮게 나타났다.

겸손 척도에서는 국가 간의 차이를 검증한 결과는 $z=-.262$ 로 $z_{99\%}=-2.33$ 을 기준으로 판단했을 때 $p<.01$ 수준에서 한국집단의 겸손 척도의 재검사 신뢰도는 독일집단의 재검사 신뢰도보다 낮게 나타났다. 반면에 상호의존성에 따른 차이를 검증한 결과는 $z=-1.06$ 으로 상위집단이 하위집단보다 재검사 신뢰도가 유의하게 낮게 나타나지 않았다.

성실성요인의 경우 전체집단의 충실성과 자기규제 척도의 재검사 신뢰도는 각각 .78과 .85

이었다. 이들 재검사 신뢰도에 대한 상호의존성과 국가에 따른 집단 간 비교를 한 결과 충실성의 경우 상호의존성에 따른 차이를 검증한 결과는 $z=-1.37$ 로 상위집단이 하위집단보다 재검사 신뢰도가 유의하게 낮게 나타나지 않았다. 반면에 국가 간의 차이를 검증한 결과는 $z=-2.83$ 으로 일방향 검증에서의 $z_{99\%}=-2.33$ 을 기준으로 판단했을 때 $p<.01$ 수준에서 한국집단의 충실성 재검사 신뢰도는 독일집단의 재검사 신뢰도보다 유의하게 낮게 나타났다.

자기규제 척도에서 상호의존성에 따른 차이를 검증한 결과는 $z=-.72$ 로 상위집단이 하위집단보다 재검사 신뢰도가 낮게 나타나지 않았다. 반면에 국가 간의 차이를 검증한 결과 $z=-1.85$ 로 일방향 검증에서의 $z_{95\%}=-1.65$ 를 기준으로 판단했을 때 $p<.05$ 수준에서 한국집단의 자기규제 재검사 신뢰도는 독일집단의 재검사 신뢰도보다 유의하게 낮게 나타났다. 이상의 결과에 기초하여 국가와 상호의존성의 상호작용효과를 알아보기 위해 이들 변인들을 조합하여 집단 간 비교를 한 결과는 표 3과 같다.

표 3에서 보는 바와 같이 온정과 사교성 척도의 경우 한국과 독일 집단 모두에서 상호의존성에 따른 재검사 신뢰도에 있어 $z_{95\%}=-1.65$ 와 비교했을 때 상위집단이 하위집단보다 재검사 신

표 3. 각 척도 재검사 신뢰도의 국가와 상호의존성에 따른 z값 비교

	온정	사교성	이타성	겸손	충실성	자기규제
한국집단의 상/하 비교	1.17	.91	-1.75*	-1.41	-1.72*	-.91
독일집단의 상/하 비교	1.40	1.05	-2.06*	.47	.40	.26
상위집단의 한/독 비교	-1.38	-.47	-.00	-2.34**	-2.61**	-1.80*
하위집단의 한/독 비교	-1.05	-.20	-.64	-.67	-.75	-.79

*p<.05, **p<.01

뢰도가 유의하게 낮게 나타나지 않았다. 또한 상호의존성의 상·하위집단 모두에서도 마찬가지로 국가에 의한 재검사 신뢰도에 있어 한국집단의 재검사 신뢰도가 독일집단보다 낮게 나타나지 않았다.

이타성의 경우 상호의존성의 상·하위집단 모두에서 국가에 의한 재검사 신뢰도에 있어 한국집단의 재검사 신뢰도가 독일집단보다 유의하게 낮게 나타나지 않았다. 반면에 한국과 독일 집단 모두에서 상호의존성에 따른 재검사 신뢰도에 있어 한국 집단은 $z=-1.75$ 이고 독일 집단은 $z=-2.06$ 으로 일방향 검증에서의 $z_{.95}=-1.65$ 를 기준으로 판단했을 때 $p<.05$ 수준에서 상위집단의 재검사 신뢰도가 하위집단보다 유의하게 낮게 나타났다. 즉 상호의존성에 의한 재검사 신뢰도의 차이는 한국과 독일 모두에서 나타났음을 알 수 있다.

겸손 척도에서는 한국과 독일 집단 모두에서 상호의존성에 따른 재검사 신뢰도에 있어 상위집단이 하위집단보다 재검사 신뢰도가 낮게 나타나지 않았다. 상호의존성의 하위집단에서도 국가에 의한 재검사 신뢰도에 있어 한국집단의 재검사 신뢰도가 독일집단보다 낮게 나타나지 않았다. 반면에 상호의존성의 상위집단에서는 $z=-2.34$ 로 일방향 검증에서의 $z_{.95}=-2.33$ 을 기준으로 판단했을 때 $p<.01$ 수준에서 한국집단이 독일집단보다 재검사 신뢰도가 유의하게 낮게 나타났다(그림 1 참조).

충실성 척도의 경우 독일 집단에서 상호의존

성에 따른 재검사 신뢰도에 있어 상위집단이 하위집단보다 재검사 신뢰도가 유의하게 낮게 나타나지 않았다. 반면에 한국 집단에서는 $z=-1.72$ 로 일방향 검증에서의 $z_{.95}=-1.65$ 를 기준으로 판단했을 때 $p<.05$ 수준에서 상위집단에서 하위집단보다 재검사 신뢰도가 유의하게 낮게 나타났다.

한편 상호의존성의 하위집단에서는 국가에 의한 재검사 신뢰도에 있어 한국집단의 재검사

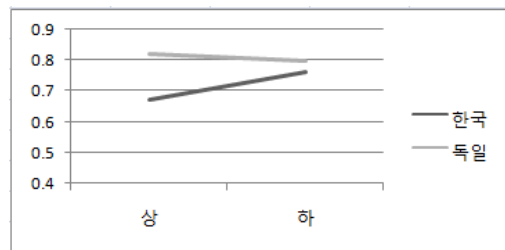


그림 1. 겸손 척도에서의 국가와 상호의존성의 상호작용 효과

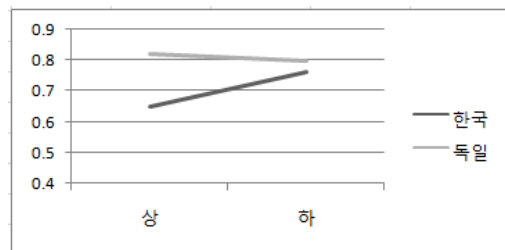


그림 2. 충실성 척도에서의 국가와 상호의존성의 상호작용 효과

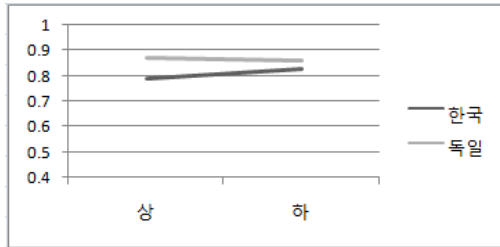


그림 3. 자기규제 척도에서의 국가와 상호의존성의 상호작용 효과

신뢰도가 독일집단보다 낮게 나타나지 않았다. 반면에 상호의존성의 상위집단에서는 $z=-2.61$ 로 일방향 검증에서의 $z_{99\%}=-2.33$ 을 기준으로 판단했을 때 $p<.01$ 수준에서 한국집단의 충실성 재검사 신뢰도는 독일집단의 재검사 신뢰도보다 유의하게 낮게 나타났다(그림 2 참조).

자기규제 척도에서는 한국과 독일 집단 모두에서 상호의존성에 따른 재검사 신뢰도에 있어 상위집단이 하위집단보다 재검사 신뢰도가 낮게 나타나지 않았다. 또한 상호의존성의 하위집단에서도 국가에 의한 재검사 신뢰도에 있어 한국집단의 재검사 신뢰도가 독일집단보다 낮게 나타나지 않았다. 반면에 상호의존성의 상위집단에서 국가 간의 차이를 검증한 결과 $z=-1.80$ 로 일방향 검증에서의 $z_{95\%}=-1.65$ 를 기준으로 판단했을 때 $p<.05$ 수준에서 한국집단이 독일집단보다 재검사 신뢰도가 낮게 나타났다(그림 3 참조).

논 의

본 연구의 결과 이타성 척도에서 상호의존성이 높은 집단이 낮은 집단보다 재검사 신뢰도가 유의하게 낮게 나타났으며, 겸손, 충실성, 자기규제 척도에서는 한국집단에서 독일집단보다 재검사 신뢰도가 유의하게 낮게 나타났다. 따라서 상호의존적 경향이 강한 집단에서는 그렇지 않은 집단에 비해 재검사 신뢰도가 더 낮게 나타

날 것이라는 가설 1과 독일집단에 비해 한국집단에서 재검사 신뢰도가 더 낮게 나타날 것이라는 가설 2는 부분적으로 확인되었다. 이에 대해 자세한 논의를 하기 전에 본 연구에서 확인된 재검사 신뢰도가 어느 정도 수용할 만하지에 대한 검토를 하려고 한다. 우선 온정과 사교성의 재검사 신뢰도가 각각 .80과 .81이고, 이타성과 겸손은 각각 .78과 .79이며, 충실성과 자기규제는 각각 .78과 .85이었다. 자기규제의 재검사 신뢰도가 가장 높고 이타성과 충실성의 재검사 신뢰도가 상대적으로 낮게 나타났다. McCrae 등 (2011)의 연구에서도 자기규제는 재검사 신뢰도가 높은 척도이고 이타성과 충실성은 낮은 척도에 속하는 것으로 보고되어서 본 연구의 결과와 유사한 경향을 보였다.

한편 Caruso(2000)는 이타성과 겸손이 포함된 친화성 영역의 재검사 신뢰도가 낮다는 사실에 근거하여 친화성을 성격특질이 아니라 상태(state)로 보아야 할 것이라는 견해를 피력하였다. 이와 같이 친화성을 기본 성격특질로 보는 것에 대해 Eysenck(1992)도 회의적인 입장을 보였다. 그러나 재검사 신뢰도가 낮다고 하여 성격특질이 아니라 상태라고 규정하는 것보다는 신뢰도가 낮은 이유를 찾아 분석하는 것이 더 우선되어야 할 것이고 이를 탐색하는 것이 본 연구의 목적과도 부합한다.

본 연구에서의 척도별 재검사 신뢰도를 다시 국가별로 살펴보면, 한국에서의 재검사 신뢰도 지수는(온정:.79, 사교성:.81, 이타성:.77, 겸손:.70, 충실성:.69, 자기규제:.82) Piemont과 Chae(1997)가 1주일 간격으로 측정한 재검사 신뢰도 지수(온정:.84 사교성:.85, 이타성:.74, 겸손:.74, 충실성:.55, 자기규제:.75, N=58)와 유의한 차이를 보이지 않았다($z=1.04, .65, -.50, .63, -1.56, -1.24$ ns). 독일집단의 경우에도(온정:.82, 사교성:.82, 이타성:.79, 겸손:.81, 충실성:.81, 자기규제:.87) Ostendorf와 Angleitner(2004)가 대학생들을 대상으로 1개월 간격으로 측정한 재검사 신뢰도 지수(온정성:.83 사교성:.87, 이타성:.78, 겸손:.79 충실

성: .80, 자기규제: .88, N=70)와 비교했을 때 유의한 차이를 보이지 않았다($z=.28, 1.24, -.18, -.31, -.25, .23, ns$). 따라서 본 연구에서 산출된 재검사 신뢰도 지수는 수용할 만하다고 평가할 수 있다.

한편 Caruso(2000)는 표집집단에 따라 신뢰도 값이 다를 수밖에 없다고 하였는데, 이러한 이유로 Pedhazur와 Schmelkin(1991)은 검사의 신뢰도 지수를 언급하는 것은 오해의 소지가 있음을 지적하였다. 같은 맥락에서 Thompson(1994)도 검사의 신뢰도가 아니라 검사 측정치(scores)의 신뢰도라고 표현하는 것이 더 정확하다고 주장하였다. 따라서 본 연구결과 확인된 재검사 신뢰도 값을 NEO-PI-R의 신뢰도로 일반화하는 것은 타당하지 않을 것이다.

각 척도별 결과에 대한 논의는 다음과 같다. 이타성 척도의 경우 독일과 한국 모두에서 상호의존적 자아상에 따라 집단 간 차이를 보였는데, 상호의존적 경향이 강한 집단의 재검사 신뢰도가 더 낮게 나타났다. 이러한 결과에 대한 근거를 척도의 내용과 관련지어 분석할 필요가 있다. 이타성 척도의 내용을 살펴보면 이 척도는 '다른 사람들'과의 관계, 배려에 관한 내용들을 포함하고 있다. 첫 번째 응답을 한 것과 두 번째 응답을 한 사이에 1개월의 시간 간격이 있었는데 이른바 '다른 사람들'이 누구인지 두 시점에서의 표상 내용이 다를 수 있고 이에 따라 응답내용이 달라질 수 있는데, 이러한 경향은 상호의존적 경향이 강할수록 더욱 클 수 있다는 것이다. 즉 자아의 특성을 다른 사람들과의 관계적인 맥락에서 보는 경향이 강한 경우 이타성과 같은 척도에서 재검사 신뢰도가 낮게 나타날 수 있다는 것이다.

이와는 달리 온정과 사교성에서는 상호의존성의 효과와 국가 간의 차이가 나타나지 않았다. 이와 같은 결과에 대한 이유를 다시 척도의 내용들을 기초로 분석해보면, 사교성의 경우에는 다른 사람과의 교제나 사귀는 것을 원하는 정도를 측정하고 있는데(Ostendorf & Angleleitner,

2004), 온정 척도에서처럼 '일반 대중', '처음 만나는 사람들'과 같이 내집단이 아닌 외집단과 관련된 내용이 많아서 상호의존적인 경향에 의한 영향을 받지 않았을 것이라고 추측할 수 있다(Singelis, 1994). 한편 이타성 척도에서는 다른 사람들에 대해 배려를 하거나 돕는 것과 같이 (Ostendorf & Angleleitner, 2004) 온정이나 사교성에 비해 다른 사람들과 더 밀도 있는 상호작용을 하는 내용으로 구성되어 있어 상대방이 누구인가에 따라 응답내용이 달라질 수 있는 가능성이 더 크다고 할 수 있다.

Schmddt 등(2003)은 이른바 특수요인 오류 즉 피검사자와 문항의 상호작용이 검사결과의 오류의 원인이 될 수 있다고 했는데 이를 본 연구결과와 관련지어 본다면, 문항내용에 따라 상호의존적인 경향이 높은 사람들의 경우에는 대인관계 상황에서 모호한 참조집단을 포함하는 검사문항에 응답할 때 즉 누구와의 관계를 의미하는지 분명하지 않는 경우에는 검사문항에 대해 달리 반응을 할 수 있을 것이라는 가정을 할 수 있다.

또한 본 연구의 결과를 English와 Chen(2007)의 연구결과와도 관련지어 논의를 할 수 있는데, 그들 연구에서 아시아계 미국인들이 유럽계 미국인들보다 사회적 환경의 변화에 따라 각 맥락에 맞는 복수의 자아개념을 갖고 있고 따라서 자아개념의 일관성이 낮지만, 관계적 맥락이 변화하지 않으면 시간의 변화에도 불구하고 안정성을 보이는 것으로 나타났다. 즉 본 연구에서의 온정과 사교성 척도에서와 같이 내집단이 아닌 외집단과 관련된 내용에 대하여 설문 하거나 문항내용에서 거론되는 사람들이 어떤 집단에 속하는지 비교적 명확하게 구별되는 경우에는 상호의존적 경향의 영향력이 줄어들고 비교적 안정적인 응답을 함으로써 재검사 신뢰도에 영향을 미치지 않는 반면에 이타성 척도에서처럼 누구와의 상호작용 상황을 염두에 두고 설문 문항에 응답을 하느냐는 것이 분명하지 않은 경우에는 상호의존적 경향에 의한 차이를 보이는 것

으로 판단할 수 있다. 따라서 본 연구 결과 특히 상호의존적인 경향이 강한 집단에서 이타성 척도의 재검사 신뢰도가 낮게 나타난 점이 앞에서 언급한 McCrae 등(2011)이나 Caruso(2000)가 지적한 이타성 척도를 포함한 친화성 요인의 신뢰도 문제에 대한 부분적인 설명을 제공한다고 할 수 있다.

한편 겸손, 충실성과 자기규제 척도에서는 국가와 상호의존성의 상호작용 효과가 나타나 상호의존성이 높은 집단에서만 한국집단의 재검사 신뢰도가 독일집단에서보다 더 낮게 나타났다. 이러한 결과에 대해 척도의 내용을 분석하면서 살펴보면, 자기규제성의 경우에는 ‘일의 시작과 완성’ 혹은 ‘원칙의 수행’을 위한 자기관리를 하는 내용들을 포함하고 있는데, 한국 집단의 경우 어떤 일이냐에 따라 혹은 어떤 원칙이냐에 따라 즉 상황 혹은 맥락에 따라서 달라질 수 있다는 태도가 상호의존성이 높은 집단에서 영향을 미친 것으로 볼 수 있다. 이러한 경향은 충실성에서도 유사하게 나타났다. 한편 겸손척도에서는 능력이나 성과를 드러내거나 자신의 우월성을 평가하는 내용을 포함하고 있는데 (Ostendorf & Angleleitner, 2004) 한국의 상호의존성이 높은 사람들이 상황에 맞추어 행동을 하는 경향이 높아 검사 문항에 응답을 하면서 어떤 상황을 염두에 두고 답을 하는가에 따라 1, 2차 검사 시에 일관성이 낮아진 것으로 해석할 수 있다. 흔히 시차를 두고 응답내용이 달라질 경우 이를 실제적인 변화로 해석을 하는 경향이 있는데 본 연구에서 살펴본 바와 같이 이에선 실제적인 변화 뿐 아니라 한국집단의 경우에는 상호의존적 경향에 의한 측정오류가 포함되었을 수 있다. 따라서 이러한 문항들을 포함하는 심리측정의 결과를 해석할 때 특히 상호의존적 경향이 강한 사람들에게서 더욱 유의할 필요가 있음을 알 수 있다.

겸손, 충실성, 자기규제 척도에서 상호의존성이 높은 집단에서만 한국과 독일에서의 재검사 신뢰도에서 차이를 보인 것은 이타성 척도에서

상호의존성의 영향이 한국과 독일 모두에서 나타난 것과는 차이를 보인다. 즉 다른 사람들을 배려하고 친절하게 행동하는 것과 같은 내용을 포함하는 이타성 척도에서는 한국과 독일 모두에서 상호의존성에 의해 재검사 신뢰도에 의해 차이를 보인 반면에, ‘일이나 원칙의 수행’과 관련된 자기규제나 충실성 척도에 있어서는 독일에 비해 한국에서 상호의존성의 영향이 더 강한 것으로 나타났다. 충실성의 경우 의무에 대한 이해에 있어 독일에서 한국에 비해 준법행위가 더 강조되는 것과도 관련지어 해석할 수 있다 (Hoppe-Graff & Kim, 2004). 즉 맡겨진 과업이나 책임을 이행하는 것을 법이나 규칙의 준수와 관련지어 생각하는 경향이 더 강한 독일의 경우 상황에 따라 달리 행동하는 경향이 상대적으로 더 적어 재검사 신뢰도가 더 높게 나타났다고 할 수 있다.

또한 한국의 상호의존성이 높은 집단에서 일, 혹은 성과와 관련된 상황에서 일관성이 상대적으로 낮게 나타난 것에 대해 공과 사의 구별이 상대적으로 약한 한국사회의 인정주의적 경향이 (한규석, 2009) 반영되었을 수 있다는 해석도 가능하다. 그러나 추후 연구를 통해 한국의 인정주의적 경향과 상호의존성의 관계가 규명되어야 이와 같은 해석에 대한 정당성이 부여될 것이다.

Kitayama 등(2009)은 서구 문화권에서 개성의 표현, 개인적 목표의 수행, 자기발전, 높은 자존감 유지와 같은 문화적 과제를 수행하기 위해서 초점을 맞춘(focused) 주의(Masuda & Nisbett, 2001), 성향 귀인(Morris & Peng, 1994)을 보이는 반면에 동아시아 문화권에서는 자기비판과 적응, 다른 사람들로부터 인정을 받기 위해서 주변을 살피고 사회적 상황에 주의를 기울이는 것 외에도, 맥락과 상황에 민감한 판단과 결정을 하고 전체적 주의와 상황귀인을 한다고 하였는데, 본 연구의 결과에서 한국의 상호의존성이 높은 사람들이 일을 수행하거나 자신의 능력 혹은 성과를 드러내는 경우에도 상황이나 맥락에

민감하게 반응을 보이고 이러한 경향이 재검사 신뢰도에 영향을 미친 것으로 이해할 수 있다.

Peng과 Nisbett(1999)은 (유학적) 변증법적 사고를 비일관적이고 모순적인 내용의 수용, 복잡성과 변화에 대한 기대, 전체적 사고와 관련된 사고체계로 설명하였다. 이를 응용하여 이들 연구자들은 아시아계 미국인들이 유럽계보다 자아개념의 일관성이 낮은 것과(Choi & Choi, 2002; English & Chen, 2007) 상반되는 감정을 동시에 보고하는 경향(Kitayama, Markus, & Kurokawa, 2000; Schimmack, Oishi, & Diener, 2005; Scollon, Diener, Oishi, & Biswas-Diener, 2005)을 설명할 수 있었다. 즉 아시아계 사람들이 상황에 맞추어 행동과 사고를 조절하는 유연성이 강하고 상황의 의미를 크게 생각하며 상황적 요인을 고려하는 경향이 강해 모순적인 내용을 포함하는 것을 자연스럽게 수용하고 이를 줄이려는 노력을 더 적게 하는 사고경향이 있다는 것이다. 반면에 서양에서는 합리적이고 분석적인 사고를 강조하는 선형적 사고를 하는 경향이 있어 내적인 일관성을 강조하고 따라서 자아개념도 비교적 일관적이고 탈맥락적인 경향을 보인다(Spencer-Rodgers, Peng, Wang, & Hou, 2004). 요약컨대 서양인들은 다양한 맥락에 걸쳐 비교적 일관성을 갖는 반면에 동아시아인들은 맥락에 따라 다양한 모습을 보이는 경향을 갖는다는 것이다. 본 연구의 결과에서는 이러한 문화적 차이가 한국의 상호의존적 경향이 높은 집단에서 더욱 강조되어 겸손, 충실성, 자기규제 척도의 재검사 신뢰도의 차이로 나타났을 가능성이 있다고 해석할 수 있다.

McCrae 등(2011)은 심리검사에서 왜 사람들이 상황의 변화에 따라 상이한 응답을 선택하는가에 대한 질문을 했는데 본 연구결과를 통해 상호의존적 경향이 이에 대한 부분적인 답을 제공할 가능성이 있다고 할 수 있다. 따라서 맥락에 따라 다르게 응답할 가능성이 있는 문항들을 포함한 성격검사의 경우 상호의존성이 강한 사람들의 결과를 해석하는데 있어 주의를 할 필요가

있다는 점을 지적할 수 있다.

본 연구에서 살펴 본 6개의 척도 중 4개의 척도 즉 이타성, 겸손, 충실성, 자기규제 척도에서 상호의존성이 높은 경우에 특히 한국 집단에서 재검사 신뢰도가 유의하게 더 낮게 나타났다는 점에 주목할 필요가 있다. 그런데 이러한 척도들의 문항들은 다른 사람을 돕거나 배려하는 것과 같은 상호작용을 요구하는 대인관계에서 대상이 누구인지를 명확히 하지 않고 있거나 상황에 따라서 달리 행동할 수 있는 가능성이 있는 내용들을 포함하고 있다. 따라서 상호의존성의 영향을 축소하기 위해서는 대인관계와 관련된 문항에서의 참조집단과 상황에 대한 서술을 좀 더 분명히 하는 것을 고려할 수 있다. 그럼에도 상호의존성에 의해 상대적으로 영향을 덜 받는 성격특성은 무엇인지, 문항을 어떻게 구성하는 것이 좋을지에 대해 분석하는 후속 연구들이 이루어질 필요가 있다.

일반적으로 신뢰도에 영향을 주는 요인으로 문항 내용 외에도 지시문, 응답양식 등이 있다(Chmielewski & Watson, 2009). 과거에 대한 질문을 하는 문항이나 응답양식이 형용사 리스트인 경우에 문장인 경우보다 신뢰도가 더 높고 각 개인의 정체성과 관련된 주요특질과 관찰 가능한 특질을 측정하는 경우 신뢰도가 더 높다(John & Robins, 1993). 따라서 본 연구에서 사용한 조사도구의 문항들이 관찰 가능한 행동과 그렇지 않은 특성에 대한 질문들로 복합적으로 구성되었다는 점이 재검사 신뢰도에 영향을 주었을 가능성도 고려할 수 있다.

이 외에도 McCrae 등(2011)은 재검사 신뢰도에 영향을 미치는 피검사자 오류의 예로 지능이나 문자해독력(Allik, Laidra, Realo, & Pullmann, 2004) 외에도 비협조적인 태도, 무작위성 응답을 들었는데, 본 연구의 결과를 통해 상호의존성도 피검사자 오류의 원인이 될 수 있다는 가정을 할 수 있다. 문제는 상호의존성이 체계적 오류 혹은 무작위성 오류의 성격을 갖는지는 아직 분명하지는 않다. 물론 본 연구의 결과를 확대 해석

하는 것은 문제가 있을 수 있으나, 만일 상호의 존성이 체계적인 오류의 원인이 된다면, 심리측정 도구를 구성할 때 이러한 오류를 통제하기 위한 방안을 마련할 필요가 있을 것이다. 그러나 보다 근본적으로는 성격측정도구를 활용하여 인간의 심리적 특성을 일반적이고 추상적인 차원에서 측정할 수 있다는 장점에도 불구하고 구체적인 상황을 고려한 다양한 측정방법의 보완이 요구된다고 할 수 있다.

한편 한국과 서구문화권을 비교하기 위해 다양한 하위속성을 기준으로 다차원적인 분석을 할 필요가 있음을 고려할 때(Triandis, Bontempo, Villareal, Asai, & Lucca, 1988), 본 연구에서 비교의 기준으로 적용했던 상호의존성 즉 자아표상의 차이 외에도 추후연구를 통해 개인목표와 집단목표의 충돌, 교환 혹은 정의 관계의 차이, 사회적 행위의 결정인자로서의 사회적 규범과 개인의 태도의 차이와 같은(한규석, 신수진, 1999) 다른 기준들을 적용하여 재검사 신뢰도의 문제를 탐색할 수 있을 것이다.

Vacha-Haase(1998)는 신뢰도 일반화 연구의 필요성을 제기하였는데, 이는 다양한 표집집단에 따라 검사의 신뢰도 지수에 영향을 주는 요인을 분석하는 것이다. 본 연구에서도 이러한 점에서 신뢰도 지수에 영향을 주는 요인으로 상호의존성을 기준으로 문화적 차이를 분석했다는 의미가 있음에도 불구하고 다음과 같은 연구의 한계점을 지적할 수 있다.

본 연구에서는 NEO-PI-R의 하위척도 여섯 개만을 사용하여 분석을 하였고 한국과 독일의 대학생들만을 대상으로 조사를 하였기 때문에 이를 일반화하는데 제한점이 있다. 따라서 다른 연구대상 집단도 조사함으로써 신뢰도에 영향을 주는 요인들을 분석할 필요가 있다. 아울러 다른 심리측정도구에도 확대 적용하고 재검사 측정 기간과 횟수를 다양하게 적용하는 추후연구들을 통해 신뢰도에 영향을 미치는 변인들을 체계적으로 탐색함으로써 자기보고식 심리측정도구의 결과를 좀 더 명확하게 해석하는데 기여하

는 것이 요구된다.

참고문헌

- 안창규, 채준호 (1997). NEO-PI-R의 한국표준화를 위한 연구. *한국심리학회지: 상담과 심리치료*, 9(1), 443-473.
- 윤성민, 신희철 (2013). 행복증진을 위한 개입 활동들의 효과에 영향을 미치는 성격 5요인의 역할 탐색. *한국심리학회지: 상담 및 심리치료*, 25(2), 275-308.
- 한규석 (2009). *사회심리학이 이해*. 서울: 학지사.
- 한규석, 신수진 (1999). 한국인의 선호가치 변화 - 수직적 집단주의에서 수평적 개인주의로. *한국심리학회지: 사회 및 성격*, 13(2), 293-310.
- Allik, J., Laidra, K., Realo, A., & Pullmann, H. (2004). Personality development from 12 to 18 years of age: Changes in mean levels and structures of traits. *European Journal of Personality*, 18, 445-462.
- Caruso, J. C. (2000). Reliability generalization of the NEO personality scales. *Educational and Psychological Measurement*, 60(2), 236-254.
- Cattell, R. B. (1986). The psychometric properties of tests: Consistency, validity, and efficiency. In R. B. Cattell & R. C. Johnson (Eds.), *Functional psychological testing* (pp. 54-78). New York: Brunner/Mazel.
- Chmielewski, M., & Watson, D. (2009). What is being assessed and why it matters: The impact of transient error on trait research. *Journal of Personality and Social Psychology*, 97(1), 186-202.
- Choi, I., & Choi, Y. (2002). Culture and self-concept flexibility. *Personality and Social Psychology Bulletin*, 28, 1508-1517.
- Church, A. T., Katigbak, M. S., Reyes, J. A. S., Salanga, M. G. C., Miramontes, L. A., & Adams, N. B. (2008). Prediction and cross-

- situational consistency of daily behavior across cultures: Testing trait and cultural psychology perspectives. *Journal of Research in Personality*, 42, 1199-1215.
- Cortina J. (1993). What is coefficient alpha: an examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory and NEO Five-Factor Inventory Professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Cronbach, L. J., (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- English, T., & Chen, S. (2007). Culture and self-concept stability: Consistency across and within context among Asian Americans and European Americans. *Journal of Personality and Social Psychology*, 93, 478-490.
- Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Differences*, 13, 667-673.
- Fiske, A. P. (2002). Using individualism and collectivism to compare cultures-A critique of the validity and measurement of the constructs: Comment on Oyserman et al. (2002). *Psychological Bulletin*, 128, 78-88.
- Green, S. B. (2003). A coefficient alpha for test-retest data. *Psychological Methods*, 8, 88-101.
- Heine, S. J. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of Personality*, 69, 881-906.
- Hoppe-Graff, S. & Kim, H. -O. (2004). Understanding rights and duties in different cultures and contexts: Observations from German and Korean adolescents. In N.J. Finkel & F.M. Moghaddam (Eds.), *Human rights and duties: Empirical contributions and normative commentaries* (pp. 49-73). Washington, DC: APA Press
- 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, 521-551.
- John, O. P., & Soto, C. H. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461-494). New York: Guilford Press.
- Johnson, J. A. (1997). Units of analysis for the description and explanation of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 73-93). San Diego: Academic Press.
- Kanagawa, C., Cross, S. E., & Markus, H. R. (2001). "Who am I?" The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, 27, 90-103.
- Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States. *Cognition and Emotion*, 14, 93-124.
- Kitayama, S., Park, H., Sevincer, A. T., Karasawa, M., & Uskul, A. (2009). A cultural task analysis of implicit independence: Comparing North America, Western Europe, and East Asia. *Journal of Personality and Social Psychology*, 97(2), 236-255.
- Kondo, D. (1990). *Crafting selves: Power, gender and discourses of identity in a Japanese workplace*. Chicago: University of Chicago Press.
- Lebra, T. S. (1994). Mother and child in Japanese socialization: A Japan-U.S. comparison. In P. M. Greenfield & R. R. Cocking (Eds.), *Cross-cultural roots of minority child development* (pp. 259-274). Hillsdale, NJ: Lawrence Erlbaum.

- Levine, T. R., Bresnahan, M. J., Park, H., Lapinski, M. K., Wittenbaum, G. M., Shearman, S. M., Lee, S., Chung, D., & Ohashi, R. (2003). Self-construal scales lack validity. *Human Communication Research, 29*(2), 210-252.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.
- Markus, H. R., & Kitayama, S. (1994). A collective fear of the collective: Implications for selves and theories of selves. *Personality and Social Psychology Bulletin, 20*, 568-579.
- Markus, H. R., & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross-Cultural Psychology, 29*, 63-87.
- Markus, H. R., & Kitayama, S., Mullally, P., Masuda, T., & Fryberg, S. (1997). *Of selves and selfways: Patterns of individuality and uniformity in identity*. Stanford, CA: Stanford University Press.
- Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81*, 922-934.
- Matsumoto, D. (2003). The discrepancy between consensual-level culture and individual-level culture. *Culture & Psychology, 9*(1), 89-95.
- McCrae, R. R., & Costa, P. T. (1994). The stability of personality: Observations and evaluations. *Current Directions in Psychological Science, 3*, 173-175.
- McCrae, R. R., Kurtz, J. W., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review, 15*(1), 28-50.
- Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology, 67*, 949-971.
- Murphy, K. R., & Davidshofer, C. O. (2001). *Psychological testing: Principles and applications* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Newman, L. S. (1993). How individualists interpret behavior: Idiocentrism and spontaneous trait inference. *Social Cognition, 11*, 243-269.
- Noguchi, K. (2007). Examination of the content of individualism/collectivism scales in cultural comparisons of the USA and Japan. *Asian Journal of Social Psychology, 10*, 131-144.
- Nunnally, J. C., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Oishi, S., Diener, E., Scollon, C. N., & Biswas-Diener, R. (2004). Cross-situational consistency of affective experiences across cultures. *Journal of Personality and Social Psychology, 86*, 460-472.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343-355.
- Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R Manual*. Göttingen: Hogrefe.
- Owe, E. 외 78인 (2013). Contextualism as an important facet of individualism-collectivism: Personhood beliefs across 37 national groups. *Journal of Cross-Cultural Psychology, 44*(1), 24-45.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist, 54*, 741-754.
- Pietmont, R. L., & Chae, J. (1997). Cross-cultural generalizability of the five-factor model of personality: Development and validation of the NEO-PI-R for Koreans. *Journal of Cross-Cultural Psychology, 28*(2), 131-155.
- Schimmack, U., Oishi, S., & Diener, E. (2005). Individualism: A valid and important dimension

- of cultural differences. *Personality and Social Psychology Review*, 9, 17-31.
- Schmidt, F. L., Lee, H., & Ilies, R. (2003). Beyond alpha: An empirical examination of the effects of different sources of measurement error on reliability estimates for measures of individual differences constructs. *Psychological Methods*, 8, 206-224.
- Schmitt, D. P., Allik, J., McCrae, R. R. & Benet-Martinez, V. (2007). The geographic distribution of big five personality traits: Patterns and Profiles of Human Self-Description Across 56 Nations. *Journal of Cross-Cultural Psychology*, 38, 173-212.
- Schwartz, S. H. (2006). A theory of cultural value orientations: explication and application. *Comparative Sociology*, 5(2), 137-182.
- Schweder, R. A. (1991). *Thinking through cultures: Expeditions in cultural psychology*. Cambridge, MA: Harvard University Press.
- Scollon, C. N., Diener, E., Oishi, S., & Biswas-Diener, R. (2005). An experience sampling and cross-cultural investigation of the relation between pleasant and unpleasant affect. *Cognition and Emotion*, 19, 27-52.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20(5), 580-591.
- Spencer-Rodgers, J., Peng, K., Wang, L., & Hou, Y. (2004). Dialectical self-esteem and East-West differences in psychological well-being. *Personality and Social Psychology Bulletin*, 30, 1416-1432.
- Thompson, B. (1994). Guidelines for authors. *Educational and Psychological Measurement*, 54, 837-847.
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.
- Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, 54(2), 323-338.
- Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, 58, 6-20.
- Watson, D. (2004). Stability versus change, dependability versus error: Issues in the assessment of personality over time. *Journal of Research in Personality*, 38, 319-350.
- Yoon, K., Schmidt, F., & Ilies, R. (2002). Cross-cultural construct validity of the five-factor models of personality among Korean employees. *Journal of Cross-Cultural Psychology*, 33(3), 217-235.
- 1 차원고접수 : 2015. 06. 09.
수정원고접수 : 2015. 07. 19.
최종게재결정 : 2015. 07. 23.

Cultural differences in the stability of self-report measures: A comparison of Korean and German students

Hye-On Kim

Sujeong Kim

Mokpo National University

This study aims to examine if the stability of self-report measures vary with culture. People from East Asian countries are generally assumed to behave more flexibly in different interpersonal contexts and to put less emphasis on the stability of one's self-image as compared to Westerners. It is hypothesized that this difference is related to the stability of self-report measures and would be explained by interdependence. A questionnaire study was conducted with 257 Korean and 196 German college students. Both samples were divided in two subsamples according to their level of interdependence. Data on the retest-reliability of self-report measures were collected through the repeated assessment of 6 NEO-PI-R scales: warmth, gregariousness, altruism, modesty, dutifulness and self-discipline. There was a one-month interval between the two measurements. Retest-reliability coefficients were transformed into the Fisher's Z scores, and data analysis was carried out for these scores by one-tailed z-tests. Our results partially support the assumption of cultural effects on the stability of self-report measures. For the altruism scale, retest-reliability was significantly lower in the subsample with a higher level of interdependence both in Korea and in Germany. For the modesty, dutifulness and self-discipline scale Retest-reliability was significantly lower in students from Korea who also showed higher level of interdependence. The results were discussed in terms of interdependence.

Keywords : retest-reliability, self-report measures, NEO-PI-R, interdependence