

심리측정학적 관점에서의 MMPI의 문제점 연구

김 중 슌

서울대학병원

박 동 건

고려대학교

본 연구의 주 목적은 심리측정학적인 관점에서 MMPI의 문제점을 고찰해 보는 것이었다. 대표준화 자료를 재분석하여 이론적인 신뢰도계수를 보다 정확히 추정해 보고자 했다. 그 결과 대표준화 당시에 제시된 대부분의 척도의 α 값과 그리 다르지 않았지만 *Mf*척도의 신뢰도 추정값은 매우 낮게 산출되었다. 그러나 신뢰도 추정치가 .7 이하인 척도는 MMPI처럼 연구의 목적이 아닌 개인의 진단을 목적으로 하는 검사의 경우에는 다소 낮기 때문에 척도점수가 불안정하여 진단에 세심한 주의를 기울여야 함을 지적하였다. 문항분석을 한 결과 52개 문항이 부적인 문항총점간 상관관계를 나타내었다. 이 부적의 상관관계가 부적의 요인간 상관에 기인하는지를 검증하기 위해서 탐색적 요인분석을 한 결과 대부분의 문항이 요인간 상관에 기인하여 부적인 문항총점간 상관을 나타낸 것이 아님을 밝혔다. 또한 준거 집단과 정상집단의 응답빈도율 자료로 또 다른 교차검증을 한 결과 여전히 다수의 문항이 문제가 있는 문항임이 확인되었다. 이러한 결과는 외국검사를 번역하여 사용할때 표준화 이전에 문항분석의 필요성을 시사한다는 사실과 이 연구의 제한점 및 추수연구의 방향을 논의하였다.

미국에서 발간된 수천종의 심리검사중에서 가장 많이 사용되는 검사는 웨슬러지능검사로 알려져 있다. 한편 그동안 성격검사중에서 가장 많이 연구되고 사용되는 것으로는 MMPI 였다. 최근에 활발히 논의되고 있는 성격의 5요인 이론에 바탕을 두고 있고 어느정도 심리측정학적인 고려도 있었다고 보고되고 있는 NEO-PI가 요즘 가장 많이 연구되는 성격검사의

하나이지만 MMPI는 여전히 가장 많이 쓰이는 검사이다. MMPI-2가 MMPI를 대체해 가고 있지만 MMPI가 아직도 많이 사용되고 있다고 한다(Adler, 1990). 사실 MMPI가 성격검사인지에 대해서 논란을 제기할 수도 있다. 검사를 소개하는 책에는 간혹 MMPI가 성격검사가 아닌 진단용검사라는 표제하에 별도로 분류되어 있어 MMPI가 정통적인 의미에서의 성격검사가

아님을 암시하기도 한다. 원래의 MMPI는 잘 정의된 성격특성을 측정하기 위해서 만들어진 것이 아니라 정신과 환자들의 진단과 분류의 목적으로 제작되었다. 근래에는 MMPI의 각 하위척도를 요인별로 분리하여 순수한 성격변인을 재는 도구로 이용하고 이해하려는 노력들을 하고 있다(Anastasi, 1988).

1940년대까지의 성격검사는 합리적방법(rational approaches to test construction)에 의해서 제작되었다. 합리적방법을 지지하는 검사개발자는 인간의 언어적 행위와 성격의 핵심적 요소와의 관계에 대한 충분한 사전 지식이 있어서 어떤 유형의 사람이 자신에 관해서 어떻게 말하리라는 것을 미리 알 수 있다고 가정하고 이를 바탕으로 문항을 작성하고 이 문항들을 척도별로 분류한다. 다시 말해서 측정하고자 하는 것과 내용상 관련된 것처럼 보이는 문항이 검사에 포함되었다. 순수한 합리적 입장에서 검사가 제작되는 경우에는 문항이 총점과 상관이 있음을 확인하는 절차를 거치지 않는다. 이런 식으로 만들어진 검사의 문항은 총점과 정적인 상관을 갖지 않을 가능성도 있고 때로는 부적인 상관을 가질 수도 있다. 우울증 환자에게 슬픈 느낌을 갖고 있는지를 물어보면 자신이 갖고 있는 느낌은 슬픔 이상의 것이라고 생각하고 아니므로 답변할 수도 있다. 이런 이유 때문에 합리적인 제작방식에 의해서 제안된 문항은 빈번히 그릇된 가정에 바탕을 두고 있다는 비판을 받는다(Nunnally & Bernstein, 1994).

1930년대 이론적인 실체를 측정의 절차와 동일시하는 조작주의(operationism)의 전통을 이어받은 경험지상주의(dust bowl empiricism)의 산물인 MMPI는 직업집단을 흥미에 의해서 변별할 목적으로 제작된 스트롱 흥미검사와 마찬가지로 준거에 의한 문항선정과 정답도출방식(criterion-oriented approaches)으로 만들어졌다(Rogers, 1995). 예를들면 우울증인 사람에 비해서 정상인이 사과를 더 잘 먹는다고 하면 사과를 좋아하는지 여부를 묻는 질문은 우울척도에 적합하다고 보게된다(김중술, 1988). 여기서 문항의 내용은 측정하려고 하는 심리적 특질과 무관하게 보여도 상관이 없다. 오히려 측정하고자 하는 구성개념과

문항이 관련이 없어 보인다는 사실이 장점으로 여겨진다.

이러한 방식으로 만들어진 MMPI가 임상장면에서 성공적으로 사용되자 그 후에 준거지향적인 방식을 채택하여 적지 않은 검사도구가 제작되게 되었다. 그러나 이러한 성공에도 불구하고 준거지향적 검사개발 방식은 심리측정학자에 의해서 비판받아 왔다. 그 비판은 세 가지로 요약될 수 있다. 첫째, 경험적 방식에 의해서 개발되었다고는 하지만 MMPI의 성공은 적절한 내용표집에 있었다. 즉 내용적으로 적합한 문항을 선정한 후에 경험적인 검증을 거쳤다. 사실 MMPI의 각 척도의 문항은 검사개발자의 머리에 되는대로 떠오른 생각을 아무렇게나 문항화 한 것이 아니다. 내담자 면담을 표준화 할 목적으로 임상면접에서 사용되고 있고 사용될 수 있다고 생각되는 문항들을 실제의 면담과 문헌을 바탕으로 작성한 많은 예비문항들 중에서 선정한 것이다. 즉, 준거와 상관이 있는 문항이 일차적으로 선정된 다음 내용적으로 당해 척도에 적합하지 않다고 판단되는 문항은 제외시키는 절차를 거쳐서 최종척도가 만들어졌다(Nunnally & Bernstein, 1994). 그러므로 어느 정도 내용타당화는 이루어졌다고 볼 수 있다. 그러나 이 과정을 척도의 일차원성의 확보의 단계로 확대 해석해서는 안될 것 같다. 검사제작자의 의도는 일차원성을 이룩하는 것이 아니라 진단집단을 내용적으로 명확히 나타내고 효과적으로 변별할 수 있는 문항을 선정하는 것이기 때문에 최종척도의 차원이 일차원이든 다차원이든 관심이 없었기 때문이다. 결국, 내용동질성 확립을 위한 문항분석을 실시하지 않았기 때문에 Nunnally (1978)의 주장과는 달리 내용동질성에 의한 검사제작은 아니었다. 따라서 MMPI는 전적으로 준거에 의해서만 검사를 제작한 것이 아니라 합리적인 제작방법도 병행하여 만들어진 검사라고 볼 수 있다. 그러나 제작당시나 그 이후에 척도총점과 문항과의 상관을 구하는 등의 문항분석을 하지 않았기 때문에 심리측정적인 문제점은 항상 내재해 있었다.

둘째, 경험적 접근방법의 지지자들은 문항이 어떤 척도에 속하는지를 합리적으로 결정하는 것이 대단히

어렵다고 주장하고 있다. 그러나 문항분석을 해보면 척도내의 많은 문항들이 타당하지 않은 것으로 잘 못 판정되는데 이는 순전히 표집오차 때문이며 바로 이 점이 경험적 방법의 맹점이라는 견해도 있다.

셋째, 경험에 과도하게 의존하여 과학의 궁극적 목표인 현상의 체계적인 원인을 밝히는 작업을 소홀히 하게 만들었다는 것이다. 이러한 비판에도 불구하고 MMPI가 아직도 애용되고 있는 주된 이유 중의 하나는 많은 자료의 축적으로 인하여 검사상의 결함을 경험적으로 수정하여 임상적으로 비교적 올바른 판단을 할 수 있다는 믿음인 것 같다. 마치 자동차의 핸들의 방향과 자동차가 나아가는 방향이 일치하지 않을 때 그 비뚤어진 핸들에 운전자가 적응해 가는 것과 같다. 자동차에 사람을 맞추는 행위가 불합리한 것과 마찬가지로 임상자가 결함이 있는 검사를 다년간의 경험으로 보완하는 것 또한 비효율적이고 비합리적인 개선방식이라고 볼 수 있다.

MMPI의 문제점으로 가장 많이 거론되는 것이 저조한 신뢰도이다. 저조한 신뢰도가 준거지향적 검사 제작 방식의 단점으로서 지적되기도 하지만 사실 신뢰도는 문제가 되지 않을 수 있다. 왜냐하면 대부분의 연구에서 신뢰도 산출방식으로 내적합치도 추정방법을 사용하기 때문이다. 이질적인 문항들로 구성된 검사의 내적합치도 신뢰도 추정치는 정밀도가 낮기 때문에 신뢰할 수 없는 상황이 흔히 있기 때문이다. 이 경우 내적합치도 신뢰도 추정치는 이론적인 신뢰도계수의 하한값일 뿐이다. 그러므로 내적합치도 신뢰도 추정값이 낮게 산출되었다고 해도 이론적인 신뢰도계수는 작지 않은 값일 가능성에 희망을 가질 수 있다. 이때에는 전이효과만 주의한다면 검사재검사 신뢰도나 평행형 검사 신뢰도 추정치가 옳은 추정방법일 수 있다. 그러나 이것은 하나의 가능성일뿐 이론적인 신뢰도계수값이 실제로 낮을 가능성도 얼마든지 있기 때문에 검사를 제작할 때 내적합치도 신뢰도를 높이는 방식으로 검사를 제작하는 것이 바람직한 방법일 것이다.

MMPI에는 준거집단과 정상집단의 변별여부를 기준으로 하여 문항타당도가 높은 문항이 선정되었지만

검사의 예언타당도는 문항타당도의 합수만이 아니다. 검사의 예언타당도는 문항타당도의 평균이 크면 클수록 그리고 문항신뢰도의 평균이 작으면 작을수록 향상된다고 한다(Crocker & Algina, 1986). 이러한 공식에 의한 검사의 예측타당도는 문항신뢰도를 내적합치도 개념하에서 추정했을 때에만 유효하다. 이 공식은 중다회귀의 원리를 바탕으로 도출된 것인데 개별 문항을 독립변인으로 간주하고 있다. 그러나 예측타당도에 관한 위의 주장은 중다회귀에서의 한가지 중요한 가정을 간과하고 있다. 중다회귀에서는 각 독립변인의 신뢰도가 완벽하다는 가정을 하고 있다. 그러나 개별문항의 신뢰도는 낮을 수 밖에 없기 때문에 개별문항을 독립변인으로 간주하여 내리는 위와 같은 결론은 타당하지 않다. 그러나 문항을 의도적으로 동질적으로 묶지 않았지만 문항총점을 구한다는 것은 척도내에 동질의 문항을 군집으로 묶고 각 군집에 가중치를 동일하게 1로 주는 것이나 다름없을 가능성이 있고 그러므로 군집들의 신뢰도는 높을 것이기 때문에 신뢰도의 문제는 심각하지 않다고 주장할 수 있다. 그러나 이러한 주장은 우연적인 가능성에 의존할 뿐 그러한 의도를 체계적으로 구체화하지 않는 이상 설득력이 크지 않다. 또한 성격문항으로서 외적준거와 타당도가 높으면서 내적준거와는 상관이 낮은 문항은 실제로는 확보하기 어렵다고 한다.

MMPI 제작 당시의 진단집단처럼 예측하려는 외적준거가 다차원적이라면 그것을 예측하는 검사도 당연히 다차원적인 이질적 문항으로 구성되어야 한다. 그러나 외적준거가 다차원적이라는 심증이 있지만 어떤 유형의 차원이고 준거에 어떠한 비중이 있는지가 연구되지 않은 상태에서 경험적으로 구성된 문항들의 점수를 합산한다는 것은 문제가 아닐 수 없다. 왜냐하면 만일 척도를 구성하고 있는 문항들이 동일한 차원을 측정하지 않는다면 척도값의 해석이 모호해진다. 즉, 동일한 수준의 척도값은 상이한 성격특성을 의미할 수 있고 상이한 수준의 척도값은 동일한 성격특성을 의미할 수도 있기 때문이다 (김영환, 1982).

MMPI제작의 특이성 때문에 한 척도내에 서로간

상관이 없는 문항이 많이 포함되어 있다고 하더라도 단일의 척도점수를 산출하는 현재의 방식을 전적으로 부정하지 않음을 전제로 한다면 단일의 점수를 구성하는 척도내 문항은 부적인 표현의 문항의 경우에는 채점을 역으로 한 다음에는 척도점수와는 정적인 상관관을 가져야 할 것이다. 단일차원인 척도에서 척도값과 부적 상관관을 갖거나 상관관을 갖지 않는 문항의 처리는 비교적 간단하다. 그러나 다차원의 척도에서 0 이나 음수의 문항총점간 상관값을 갖는 문항의 경우에 채점방향(keying)이 거꾸로 되었다거나 의도한 구성개념을 실제로는 측정하고 있지 않다는 결론을 간단하게 내릴 수는 없다. 외적준거인 진단집단을 구성하는 환자가 다양한 유형의 특질을 지니고 있고 그러한 특성이 서로 부적인 상관관계를 갖는 특질이라면 그러한 특질들을 나타내는 증상 혹은 지표변인들간의 상관관계 역시 부적일 것이기 때문이다. 그러므로, 다차원 척도에서는 그 척도에 속하는 문항이고 채점방향도 문제가 없는 경우에도 문항-총점간 상관관계수가 음의 값이 될 가능성도 배제할 수 없다. 이러한 이론적 가능성에도 불구하고 만일 MMPI의 모든 척도에서 음의 상관값이 출현한다면 모든 척도에서 측정하는 심리장애적 특성내에 서로 음의 상관관을 갖는 차원들이 모두 포함되어 있다고 보아야 하는 어려움이 생긴다. 그러므로 음의 상관의 문항은 위에 언급한 가능성외에 우리나라 사람들의 반응 성향이 가져오는 특수성 때문에 채점방향을 역으로 해야 하지 않는지, 말의 표현이 잘못 됐는지 혹은 그 척도에 속하지 않은 문항은 아닌지를 다각도로 살펴볼 필요가 있다.

1989년 MMPI 한국판 재표준화가 실시될 때 미국 원판을 기초로 하여 원문에 맞게 번역하고 각 문항들이 측정하려는 의도를 검토하고 문항내용중 문화적 차이가 있는 문항은 한국적 문화적 상황에 맞도록 번역하였다는 보고가 있었다(김영환 등, 1994). 이때 전문가의 주관적 판단에 의한 일차적인 검토가 있은후 전국적인 표본을 대상으로 한 규준화는 이루어졌으나 규준화 이전에 문항분석을 통한 객관적 검증과정이 생략되었다. 이러한 이유 때문에 문화적인 차이로 인

하여 우리나라에는 적합하지 않은 문항이 포함되었을 가능성이 있다.

최근에 필자중의 한 사람은 심리측정학적인 고려를 상당히 기울여서 제작되었다고 알려진 Personality Research Form을 번역하여 문항분석을 한 적이 있다. 이 검사의 각 척도는 하나의 구성개념을 설정하고 이를 나타내는 문항을 내용적으로 선정하고 문항분석을 통하여 거의 동질적으로 제작되었다. 그러나 분석결과 원래의 척도에 따라 묶이지 않는 문항이 적지 않음을 발견하였다. 원본의 개별문항에 대한 심리측정적인 고려가 상당하였고 번역에 나름대로 충실하고자 했음에도 불구하고 아마도 문화적인 차이 때문에 이 같은 현상이 일어났다고 생각된다.

성격검사는 시대와 문화를 반영하는 것이기 때문에 있는 그대로 수입하여 번역하여 사용할 때 상당히 주의 기울여야 한다. 더욱이 외국의 검사를 번역하여 사용하고 기존의 외국의 해석집을 많이 참고하는 경우에는 심리측정학적인 고려는 더욱 절실하다. 동일 국가내에서도 하위문화집단에 따라 문항의 적합성에 차이가 있는데 문화적인 배경이 동서로 나뉘어 판이하게 다른 경우의 성격검사의 번역은 실제로 제 2의 검사제작과 같은 단순한 번역이상의 노력이 요구된다. 문항분석을 통해서 적절하지 않은 문항 혹은 채점방향이 역으로 되어야 하는 문항들을 밝힐 필요가 있다. 재표준화 이전의 구판을 대상으로 한 연구이긴 하지만 김영환의 논문에서는 준거타당적 집단비교에 의한 연구에서 기존의 점수부여체계에 이상이 있을 가능성이 있는 문항이 다수 있음이 보고되고 있다.

MMPI-2가 제작된 배경은 시간의 흐름에 따라 문화의 변천으로 인하여 기존의 문항중에는 타당도가 소멸된 것이 있다고 판단하고 또한 적절한 표본에 의한 표준화가 주된 목적이었다(Greene, 1991). 우리나라 MMPI판도 동일한 이유로 개정할 필요가 있으며 시간적 흐름에 따른 변화이외에도 문화적인 차이에 의한 각 문항의 적합성여부에 대한 심리측정적 검토가 이루어져야 한다. 따라서 본 논문은 MMPI의 척도별로 문항분석을 실시하여 각 척도에 포함되어 있는 문항들의 적합성을 검토하는 것이 목적이다. 또한

개별척도를 해석하건 코드유형으로 해석하건 측정오차의 정도는 해석의 신뢰성에 지대한 영향을 주는 만큼 각 척도의 신뢰도를 재검토하고자 한다.

방법 및 절차

분석자료는 1990년 재표준화 자료를 사용하였다. 재표준화자료중에서 무선반응을 했다고 의심되는 사례를 제외한 다음 분석하였다. 각 문항에 대한 문항총점간 상관계수값을 산출하였다. 김영환(1982)의 논문의 응답빈도비율 자료를 사용하여 문항총점간 상관계수값에서 문제가 된다고 판단된 문항에 대해서는 비율의 차이에 관한 z 검증을 실시하였다. 또한 척도내 여타문항들과 상이한 차원이면서 음의 상관을 갖는 차원에 속하는 문항인지 아니면 채점방향을 변환해야 되는 문항인지의 증거를 찾는 하나의 방안으로서 요인분석을 실시하였다. 탐색적분석의 입장이기 때문에 공변량보다는 문항의 변량을 설명하는 요인을 추출하기 위해서 주성분분석을 실시하였다. 이론적인 신뢰도계수값을 알아보기 위해서 내적합치도 방식이외에 평행형검사 방식에 의하여 신뢰도를 추정하였다.

결과

신뢰도

각 척도의 문항을 문항곤란도의 순으로 배열한 다음 홀수번의 문항점수를 합한 총점과 짝수번의 문항점수를 합한 총점을 각각 산출하여 두 총점간 상관을 내고 Spearman-Brown 공식에 의해서 교정하여 검사의 신뢰도를 추정하였다. 이렇게 하여 두 하위검사의 평균점수를 같게하여 평행형검사를 만들어 이론적인 신뢰도 계수를 얻고자 하였다. 이는 신뢰도이론중 평행검사모형의 개념을 바탕으로 한 것이다. 대부분의 척도가 다차원적이기 때문에 영역표집모형(domain

sampling model)의 개념을 바탕으로한 내적합치도 신뢰도 추정방식에 의한 신뢰도 추정값의 정밀도가 떨어지기 때문에 평행검사모형이 더 적합하다고 할 수 있다. 각 척도에 대한 재표준화 당시 발표된 알파값, 논문에서 재표준화자료로 다시 산출한 알파값, 평행형검사 신뢰도 추정값이 표 1에 제시되어 있다.

표 1. 척도별 신뢰도계수 추정치

척도	반분신뢰도	교정신뢰도	본연구 α	재표준화 α
F	.57	.73	.81	.83
K	.52	.68	.66	.66
HS	.69	.82	.80	.80
D	.48	.65	.61	.56
HY	.46	.63	.63	.55
PD	.34	.50	.52	.68
MF(남)	.24	.38	.29	.61
MF(여)	.11	.20	.24	.52
PA	.39	.56	.62	.66
PT	.74	.85	.83	.78
SC	.74	.85	.86	.84
MA	.38	.56	.57	.60
SI	.59	.74	.74	.77

표 1에서 보듯이 재표준화 당시와 본 논문에서의 알파값에서 다소 차이가 있다. 특히 mf척도의 알파값은 크게 차이가 나서 재표준화 당시의 추정치의 절반이하였다. 무작위응답의 가능성이 있는 사람을 배제한 것일뿐 근본적으로 동일한 자료를 분석한 것인데 큰 차이가 난 이유는 분명치 않다. 평행형검사 신뢰도 추정치는 내적합치도 신뢰도계수 추정치와 그리 다르지 않았다. D, Hs, K, Mf(남), Pt의 평행형 추정치가 약간 향상된 반면 F, Ma, Sc, Pa, Pd, Mf(여)척도의 평행형 추정값은 오히려 약간 감소되었고 나머지 척도의 신뢰도 추정치는 변함이 없었다. 따라서 α 값이 이론적인 신뢰도계수를 과소 추정하지 않다고 잠정적인 결론을 내릴 수 있다.

문항분석

타당도 척도인 L, F, K, 척도는 분석에서 제외했고 Mf척도는 신뢰도가 낮고 포함된 문항들의 문항총점간 상관계수값이 대체로 낮기 때문에 분석을 하지 않았다. 김영환의 논문에는 Hy, Ma, Si의 척도에 상응하는 응답을 자료가 제시되어 있지 않기 때문에 비율에 차이에 대한 교차검증을 하지 않았다.

Hs 척도 이 척도의 경우 문항총점간 상관계수값(r_{ii})이 음수인 문항은 없다. 그러나 r_{ii} 값이 0에 가까운 문항들이 몇 개 있다. 만일 다차원적 척도가 아니라 단일 차원의 구성개념을 측정하는 경우라면 0에 가까운 값을 갖는 문항은 무조건 제외되어야 하지만 척도내 차원간에 상관이 없는 경우에는 각 차원을 나타내는 문항이 얼마나 포함되어 있는지에 따라서 0에 가까운 값이 나올 수도 있다. 즉 어떤 차원의 문항의 수가 상대적으로 적게 포함되어 있는 경우 그 차원에 속하는 문항들은 상대적으로 낮은 문항총점간 상관을 가질 수 있기 때문이다. 130번 문항의 상관값은 통계적으로는 유의미하지만 척도내 다른 문항에 비해서 낮은 편이다. 요인분석한 결과에 의하면 이 문항은 요인 2에 속해 있어서 낮은 값이 나올 가능성은 없는 것 같다. 그렇다면 이 문항은 척도에 포함시키지 않는 것이 좋을 것도 같다.

D 척도 D척도의 경우에는 11개의 문항의 r_{ii} 값이 음수이다. 그런데 문항번호 30, 39, 95, 155, 191, 241, 263은 아니오가, 문항번호 104, 130, 145, 193은 그렇다가 현행 이상방향으로의 채점방식이다. 앞서도 언급한대로 척도가 만일 단일차원으로 구성되어 있다면 음수인 문항들은 문장표현이 잘못 되었든 혹은 다른 차원에 속하는 문항이라고 단정할 수 있기 때문에 수정되거나 제외되어야 한다. 그러나 문항내 차원이 다원적이고 차원간 상관이 음수인 경우에는 때에 따라서는 문항총점간 상관이 음수일 수 있기 때문에 차원간 상관이 부적인 지를 알아보기 위해서 요인분석을 하였다. 이 연구의 모든 요인분석은 탐색적인 입

장에서 공통변량을 1로 하는 주성분분석을 하였다. 요인추출은 스크리검사를 기준으로 하여 2개 혹은 3개를 추출하였다. 이는 과소요인추출이 될 가능성이 있지만 본 연구에서는 척도의 차원을 검토하는 것이 목적이 아니라 문항총점간 상관을 교차검증하기 위해서 요인분석을 이용하는 것이기 때문에 적은 요인이라도 크게 문제될 것이 없다고 판단하였다. 오히려 요인을 과다추출하여 다른 문항과 상대적으로 부적인 요인부하값을 보이는 것을 막을 수도 있으므로 적은 수의 요인을 추출하였다. 요인회전은 δ 값을 0으로 하여 가장 큰 사각을 구성하게끔 회전을 하였다. D 척도의 문항총점간 상관계수값, 준거집단과 정상집단의 응답비율, 비율차이에 대한 검통통계량 z 값 과 요인부하값이 표 2에 제시되어있다.

척도의 요인간 상관은 -0.187로 부적 상관을 나타내었다. 이것은 문항총점간 상관이 척도내 이질적인 차원간 상관 때문에 음수가 나올 가능성도 있음을 시사해 주는 것이다. 그러나 요인간 상관이 작기 때문에 직교회전도 시도해 보았다. 결과적으로 얻은 요인 패턴과 요인구조행렬상에서 r_{ii} 값이 음수인 문항의 요인부하와 부호의 방향이 사각회전의 결과와 직교회전의 결과가 크게 다르지 않았다. 따라서 요인간 상관도 그리 큰 편도 아니고 요인구조행렬상의 문항의 요인부하도 크게 다르지 않기 때문에 부적상관을 갖는 문항들을 반드시 정당화 할 수는 없다. 특히 15번 문항과 16번 문항은 r_{ii} 의 부적인 값이 상당히 크기 때문에 더욱 요인간 상관이 음수이기 때문일 가능성은 희박한 것 같다. 따라서 문항총점간 상관계수값을 근거로 한다면 음수값을 갖는 모든 문항들은 채점방향을 바꿔야 하거나 제거되는 것이 좋을 것 같다. 김영환의 논문에 제시된 진단집단과 통제집단의 응답비율을 비교하는 검증을 한 결과 241번 문항을 제외한 문항들이 $\alpha=0.05$ 수준에서 통계적으로 유의미하였다. 이는 MMPI를 처음 제작할 때 문항선정시 사용한 방법과 일치한다. 30번 문항의 z 값은 -4.36으로 통계적으로 유의미하고 김영환의 논문에서 준거집단에서 그렇다라고 답한 사람 비율은 43.8%로서 정상집단의 23.5% 보다 크므로 현행의 채점방식과는 달리 그렇

표 2. D 척도 문항총점간 상관, 응답빈도율 및 요인분석결과

문항	r_{ii}	정상집단	준거집단	Z	1요인	2요인	표준화집단
30	-.0249	23.5/76.5	43.8/56.2	-3.41	-.058	.514	73.4/26.5
39	-.0316	20.2/79.8	45.8/54.2	-4.36	-.034	.555	65.4/34.6
95	-.0001	99.4/ 0.6	86.5/13.5	4.44	-.185	-.115	37.6/62.3
155	-.0042	20.9/79.1	65.6/34.4	-7.18	-.007	-.082	47.1/52.8
191	-.0993	18.4/81.6	60.4/39.6	-6.89	-.255	.290	51.4/48.5
241	-.0567	4.3/95.7	8.4/91.6	-1.36	-.244	.438	40.1/59.7
263	-.0319	3.7/96.3	21.1/78.9	-4.48	-.187	.277	20.7/79.2
104	-.1459	4.3/95.7	13.5/86.5	-2.68	-.199	.183	28.5/71.4
130	-.0706	0.0/100.0	3.1/96.9	-2.26	-.182	-.006	90.5/ 9.4
145	-.1118	11.8/88.9	32.6/67.4	-4.08	-.019	-.540	50.8/49.0
193	-.0807	1.8/98.2	11.5/88.5	-3.34	-.231	.075	80.4/19.4

다라고 응답한 사람을 이상방향으로 채점해야 한다. 39번 문항의 카이제곱값은 9.2로 통계적으로 유의미하고 준거집단의 그렇다의 반응율이 45.8%, 정상집단이 13.5%로 그렇다가 아닌 아니오를 이상방향으로 채점해야 한다. 신판과 구판의 문항 95의 내용이 전혀 다르기 때문에 김영환의 논문결과와의 비교는 가능하지 않다. 95번 문항의 문항총점간 상관이 -.0001로 0에 가깝기 때문에 D척도에 속하지 않는 문항일 가능성도 있다. 155번 문항의 경우는 김영환 논문에서 사용한 구판의 질문 내용이 신판의 동일 문항의 내용과 반대로 되어 있다는 점을 감안하여 그렇다와 아니다의 응답비율을 거꾸로 해석해야 한다. 두 집단간 차이가 통계적으로 유의미하고 준거집단의 아니다의 응답비율이 65.6%가 정상인 집단의 20.9%보다 크고 문항총점간 상관은 -.0042이기 때문에 기존의 채점방향과는 달리 아니다에 점수를 주어야 할 것이다. 191번 문항의 r_{ii} 은 -.0993이고 준거집단이 그렇다고 답한 비율이 60.4%로서 정상집단의 18.4%보다 크고 통계적으로 유의미하기 때문에 기존의 채점방향을 바꾸는 것이 낫다고 생각된다. 241번 문항의 r_{ii} 값이 -.0567인 반면 정상인과 준거집단의 응답율의 차이는 $\alpha = 0.05$ 수준에서 통계적으로 유의미하지는 않았다. 다만 그렇다로 답한 비율이 준거집단 8.4%, 정상집단 4.3%로 기존채점방향이 그렇다 대신 아니다일 수 있

다. 263번 문항은 기준집단의 그렇다는 응답율이 21.1%로서 정상집단의 3.7%보다 크며 통계적으로 유의미하므로 기존의 채점방향인 아니오 보다는 그렇다가 채점방향으로 적합하며 r_{ii} 값 -.0319가 이를 뒷받침하고 있다. 104번 문항의 r_{ii} 이 -.1459로서 표집오차에 의한 음수값으로 보기에는 상당히 큰 편이기 때문에 현행 채점방향이 전환되어야 한다는 지지자료를 김영환의 논문에서 기대했으나 정상집단에서는 그렇다로 답한 사람이 4.3%, 준거집단에서는 13.5%로서 기존의 채점방향인 그렇다가 적합한 것으로 나타났다. 요인분석결과에서는 51번 문항(나는 내 친구들 못지 않게 건강하다)이 가장 큰 요인부하값을 갖는 제 1 요인에서의 요인부하값이 -.17로서 r_{ii} 의 결과와 일치한다. 130번 문항에 정상집단의 사람은 모두다 그렇다고 답변했지만 준거집단 사람의 96.9%가 그렇다고 답변해서 원래의 채점방향인 그렇다 보다는 아니오가 더 적합한 채점방향인 것으로 보인다. 문항총점간 상관이 -.1118로서 비교적 큰 값을 갖고 있는 145번 문항에 그렇다고 답변한 사람의 비율이 기준집단은 32.6%, 정상집단은 11.1%로서 기존의 채점방향인 아니오가 부적합한 것으로 판정할 수 있다. 193번 문항은 구판과 신판의 문항내용이 반대로 표현되어 있기 때문에 해석에 주의를 기울여야 하지만 응답비율로만 보면 기존의 채점방향이 바뀌어져야 한다. 구판의 채

점방향은 아니오 인데 기준집단의 그렇다의 응답율이 11.5%, 정상집단의 응답율이 1.8% 로서 문항총점간 상관값이 의미하는 바와 같이 현행의 채점방향이 역전되어야 할 것으로 나타났다.

Hy 척도 요인분석결과 스크리검사를 통해서 2요인을 추출하기로 하고 $\delta=0$ 의 사각회전을 하였다. 그 결과 표 3에 제시된 바와같이 요인간의 상관이 -.15이었다.

표 3. Hy 척도 문항총점간 상관 및 요인분석결과

문항	r_{ii}	1요인	2요인
30	-.0491	-.123	.390
129	-.0504	-.159	.428
136	-.0103	-.101	.328
147	-.0441	-.170	.474
213	-.0063	-.062	.215
267	-.0389	-.147	.388

그러므로 r_{ii} 이 부적인 문항들이 반드시 좋지 않은 문항이라고 단정할 수는 없다. r_{ii} 이 음수인 문항인 문항 30, 129, 136, 147, 213, 267의 요인부하를 살펴 보았다. 147번 문항과 129번 문항은 요인부하의 절대값이 제 2 요인에 크게 걸려 있지만 제 1 요인의 영향력을 더 많이 받고 있음을 알 수 있다. 오히려 238번 처럼 요인 2에 대한 부하량의 크기가 부적으로 큰 문항은 정적인 r_{ii} 을 보이고 있다. 따라서 요인 1에 속하는 문항들의 변산이 요인 2에 속하는 문항들의 변산보다 상대적으로 크고 이것이 척도점수의 결정에 영향을 많이 끼치기 때문에 요인 1에 부적인 부하량을 어느 정도 이상 갖는 문항은 문항총점간 상관의 음수로 나타난 것임을 짐작할 수 있다. 요인 1의 설명변량이라고 할 수 있는 고유근(eigenvalue)이 크기 때문에 이 요인이 총점에서 차지하는 비율이 높음을 알 수 있다. 직교회전에서 요인부하의 절대값에서 다소 차이가 날 뿐 사각회전의 결과와 크게 다르지 않았다. 따라서 부적인 문항총점상관을 갖는 문항이 반드시 차원간 상관의 부적이기 때문이라고 단정할

수 없다. 그러나 문항총점간 상관의 개념과 요인분석에서의 부하량의 개념이 반드시 일치하는 경우가 아니기 때문에 이러한 추론이 반드시 옳다고 단정할 수는 없다.

30번 문항에 그렇다로 대답한 준거집단 사람의 비율은 43.8%, 정상집단은 23.5%로 그렇다는 답변이 이상방향으로 채점되어야 한다. 그렇다로 129번 문항에 답한 사람의 비율은 기준집단에서는 39.6%, 정상집단에서는 9.3%로 r_{ii} -.0504가 의미하는 바 대로 채점방향의 역전가능성을 확인해준다. 136번 문항의 r_{ii} 은 -.0103이고 기준집단에서 아니오로 답한 사람의 비율이 47.9%, 정상집단에서는 81.5%이어서 원래의 아니오는 이상방향으로의 채점응답이 될 수 없다. 147번 문항도 채점의 역전가능성이 확실시되는 문항이다. 원래의 채점방향은 아니오인 반면 r_{ii} 이 -.0441이고 기준집단에 아니오라고 답한 사람이 32.6%이고 정상집단에는 66.7%이기 때문에 그렇다는 반응을 채점방향으로 설정하는 것이 바람직하다. 267번 문항은 아니오라고 답변한 사람을 이상방향으로 채점하기로 되어 있으나 우울신경증집단의 아니오의 응답비율은 37.5%, 정상집단은 72.2%이므로 그렇다는 응답에 점수를 부여하는 것이 적합한 것 같다.

Pd 척도 82번 문항의 r_{ii} 값은 표 4에서 볼 수 있듯이 -.0515이며 정해진 채점방향은 아니오이다. 비록 통계적으로는 유의미하지 않지만 정상집단에 대한 아니오의 응답빈도율이 81.4%, 준거집단에 대한 응답빈도율이 79.3% 로서 아니오보다는 그렇다는 반응이 성격장애를 나타내는 것으로 판정하는 것이 나올 것 같다. 혹은 김영환의 연구결과를 근거로 한다면 이 문항은 Pd척도에 적합하지 않다고 보여진다. 134번 문항은 r_{ii} 값과 경험적 비율자료에 의한 채점방향의 결정이 일치하지 않는 경우이다. 문항총점간 상관값은 -.0935로서 채점방향이 바뀌어야 하는 경우이거나 적합한 차원의 문항이 아닌 경우라고 판단할 수 있는 반면 성격장애집단의 아니오의 응답빈도율이 46.6%이고 정상집단에 대한 응답율은 28.4%로서 현행 채점방향인 아니오가 적합한 것으로 문항총점간 상관과

표 4. Pd 척도 문항총점간 상관, 응답빈도율 및 요인분석결과

문항	r_{ii}	정상집단	준거집단	Z	1요인	2요인	표준화집단
82	-.0515	18.6/81.4	20.7/79.3	-.35	.003	-.323	40.0/59.4
134	-.0935	71.6/28.4	53.4/46.6	2.53	-.377	-.030	63.6/36.1
141	-.0625	36.2/63.8	34.5/65.5	.23	-.069	-.286	63.9/36.0
171	-.1053	39.5/60.5	41.4/58.6	-.25	.066	-.470	68.9/31.0
180	-.0581	11.1/69.9	46.6/53.4	-5.76	.086	-.413	60.3/39.4
183	-.0110	85.8/14.2	61.4/38.6	3.95	-.103	-.001	32.3/67.6
231	-.0602	26.7/73.3	28.1/71.9	-.21	-.329	.010	40.5/57.7
235	-.0125	40.6/60.0	59.6/40.4	-2.50	-.205	.081	55.6/44.2
248	-.0976	27.2/72.8	17.9/82.1	1.41	-.451	-.001	41.0/58.8
267	-.0712	27.8/72.2	26.3/73.7	.22	-.037	-.380	46.3/53.5
287	-.0071	49.1/50.9	36.2/63.8	1.69	-.194	.268	47.0/52.7
296	-.0408	60.5/39.5	39.3/60.7	2.79	-.360	.018	65.4/34.2

상충된 결과를 보인다. 주성분분석결과 두 요인간의 상관인 .11의 양수값이므로 차원간 상관인 음수이기 때문에 Pd 문항들의 r_{ii} 값이 음수가 되었을 가능성은 없다. 요인구조행렬에서의 부하값이 두 요인에서 모두 음수이므로 채점방향에 문제가 있다고 판단할 수 있다. 혹은 이 문항이 미국 원판에서는 타당한 문항이었지만 문화적인 차이 때문에 우리나라에서는 Pd 척도에 합당한 문항이 아닐 수도 있다. 그러나 이 문항의 부하값이 -.37로 비교적 큰 음수 값을 가지므로 척도에는 소속되나 채점방향을 바꾸는 것이 더 좋을 것 같다. 141번 문항의 r_{ii} 값은 -.0625이고 현행 채점 방향은 아니오이다. 이 문항은 0.05의 유의수준에서 통계적으로 유의미하지는 않지만 준거집단의 아니오의 응답빈도율은 65.5%, 정상집단은 63.8%로 채점의 역전은 불필요한 것처럼 보인다. 이 문항도 r_{ii} 와 반응빈도의 분석결과가 일치하지 않았다. 171번 문항의 r_{ii} 값은 -.1053이고 통계적으로 유의미하지 않지만 준거집단의 아니오의 반응율이 정상집단의 그것에 비해서 크기 때문에 정해진 방향으로의 채점이 적합하지 않은 것 같다. 180번 문항의 채점방향은 아니오인데 준거집단 사람들의 아니오의 응답빈도비율은 53.4%이고 정상집단은 69.9%이므로 원래의 채점방향이 역전되어야 할 것 같다. 183번 문항의 r_{ii} 는 음수이지만 응답빈도율 자료는 원래의 채점방향이 옳음을 암

시하고 있다. 그러나 구판의 문항진술은 “게으름뱅이에 돈을 주는 것을 반대한다”이나 신평에서는 “거지에게 돈 주는 것을 나는 반대한다”로 되어 있어서 신평은 구판과는 달리 이타적인 행동을 측정하는 것 같다. 표준화집단에서의 그렇다로 응답한 사람은 32.3% 아니오라는 반응은 67.6%로 구판의 문항반응빈도율의 그렇다 85.8%, 아니다 14.2% 와 대조를 이룬다. 따라서 신평을 이용한 r_{ii} 의 결과와 구판의 반응빈도율의 결과의 차이는 문항내용의 차이에 기인한 것일 가능성이 높다. 재표준자료의 응답빈도율을 정상인자료로 간주한다면 이는 r_{ii} 의 결과를 지지하고 있다고 볼 수 있다.

Pa 척도. 이 척도에서는 표 5에 나타나 있듯이 9개의 문항의 r_{ii} 가 음수값을 갖고 있다. 문항총점간 상관의 절대값이 가장 큰 327번 문항에 대한 편집성 정신분열증집단의 아니오의 응답비율이 43.5%, 정상집단은 77.3%이므로 아니오로 응답하는 사람을 이상집단으로 분류한다면 부적절할 것이다. 그 다음으로 큰 r_{ii} 의 음수값을 갖는 문항은 348번으로 정상집단의 아니오라고 응답한 비율이 78.5%로 준거집단의 58.4%보다 크므로 채점방향은 원 채점방향과는 다른 그렇다가 되어야 할 것 같다. 117번 문항의 경우 준거집단 중에 아니오라고 응답한 사람의 빈도율은 58.5%,

표 5. Pa 척도 문항총점간 상관, 응답빈도율 및 요인분석결과

문항	r_{ii}	정상집단	준거집단	Z	1요인	2요인	표준화집단
93	-.0438	86.3/13.7	80.4/19.6	1.40	-.033	.535	84.2/15.7
109	-.0346	49.4/50.6	58.7/41.3	-1.65	-.012	.360	79.5/20.4
117	-.1376	10.4/89.6	41.5/58.5	-6.32	-.298	.242	30.4/69.3
124	-.0047	72.0/28.0	85.1/14.9	-2.81	.050	.456	86.5/13.3
313	-.0031	76.1/23.9	70.9/29.1	1.04	-.026	.278	69.2/30.6
316	-.0109	32.7/67.3	60.7/39.3	-4.96	-.075	.543	78.7/21.1
319	-.0852	19.9/80.1	61.3/38.7	-7.48	-.134	.479	68.0/31.7
327	-.2033	22.7/77.3	56.5/43.5	-6.13	-.278	.217	40.0/59.7
348	-.1866	21.5/78.5	41.6/57.4	-3.84	-.317	.108	25.7/74.2

정상집단에서는 89.6%이므로 아니오라고 응답한 사람에게 1 점을 주는 현재의 채점방식에 이의를 제기할 수 있다. 93번 문항의 현재의 채점방향은 아니오이고 정상집단의 아니오라는 응답빈도율이 준거집단보다 크므로 r_{ii} 의 결과와 일치하여 현행채점방식에 문제가 있는 것 같다. 109번 문항도 아니오가 현행 채점방향인데 r_{ii} 값과 응답빈도율의 결과는 그렇다의 반응을 이상방향으로 채점하는 것이 옳은 것임을 시사하고 있다. 124번 문항의 r_{ii} 값은 약한 음수값을 갖지만 편집성 정신분열증집단에서의 아니오의 응답빈도율이 14.9%, 정상집단에서는 28.0%로서 현행 채점방향인 아니오는 문제가 있다. 그렇다면 107번 문항처럼 r_{ii} 값이 양수인 경우도 분석의 타당성을 검증하기 위해서 검토할 필요가 있다. 즉 문항총점간 상관이 문제가 없는 문항의 응답빈도율의 양태를 살펴볼 필요가 있다. 편집성 정신분열증 집단이 이 문항에 아니오라고 응답한 사람의 비율이 78.8%이고, 정상집단에서는 65.4%이므로 원 채점 방향인 아니오가 맞는 채점방식이므로 r_{ii} 의 결과와 반응율의 결과가 일치하고 있다. 313번 문항에 대한 준거집단의 아니오 응답빈도율이 29.1%, 정상집단은 23.9%로 현행 채점방향인 아니오에 일단은 문제가 없는 듯 하지만 집단간 비율의 차이 통계량이 유의미 하지 않고 r_{ii} 값도 -.0031로 0에 대단히 가깝기 때문에 이 척도에 속하지 않는 문항일 가능성도 배제할 수 없다. 그러나 124번 문항의 r_{ii} 값도 -.0047로 매우 작다는 사실로 반론을

제기할 수도 있겠다. 316번 문항의 경우 준거집단의 그렇다의 반응빈도율은 60.7%, 아니오의 빈도율은 32.7%이므로 현행 아니오 보다는 그렇다의 응답을 이상방향으로 채점하는 것이 옳을 것 같다. 319번 문항에 대한 편집성 정신분열증 집단의 그렇다의 반응빈도율이 61.3%, 정상집단의 경우 19.9%로 문항총점간 상관의 결과와는 달리 채점방향이 문제가 없는 것 같다.

Pt 척도 이 척도에는 343번 문항만 문항총점간 상관값이 음수이다. 표 6에 제시되어 있듯이 불안신경증집단의 사람들이 이 문항에 그렇다고 한 비율이 71.4%이고 정상집단의 경우에는 69.3%로서 비율의 차이에 대한 검증통계값은 통계적으로 유의미하지 않지만 그렇다의 반응을 이상방향으로 채점해야 함을 암시하는 결과이고 그것이 원 채점 방향이다.

교차검증을 위하여 요인분석결과를 살펴보면 사각회전 후의 요인간 상관값이 .27로 양수이고 요인구조행렬에서의 부하값이 요인 1과 요인 2에서 각각 -.017과 -.16로 모두 음수이므로 채점방향이 역전되어야 하거나 본 척도에 속하지 않을 문항으로 판단할 수 있다.

Sc 척도 표 7에 제시된 바와 같이 이 척도에는 r_{ii} 이 음수인 문항이 모두 4개이다. 196번 문항의 경우 만성정신분열증집단에서는 40.4%, 정상집단은 20.2%

표 6. Pt 척도 문항총점간 상관, 응답빈도율 및 요인분석결과

문항	r_{ii}	정상집단	준거집단	Z	1요인	2요인	표준화집단
343	-.0767	69.3/30.7	71.4/28.6	-.30	-.017	-.164	58.3/41.5

표 7. Sc 척도 문항총점간 상관, 응답빈도율 및 요인분석결과

문항	r_{ii}	정상집단	준거집단	Z	1요인	2요인	표준화집단
196	-.0568	79.8/20.2	73.1/26.9	1.43	-.080	-.009	86.8/13.1
306	-.1751	33.1/66.9	46.4/53.6	-2.46	-.189	-.084	13.5/86.4
322	-.2429	15.3/84.7	50.0/50.0	-6.70	-.202	-.276	59.1/40.5
104	-.1678	4.3/95.7	24.8/75.2	-5.26	-.173	-.057	28.5/71.4

이 아니므로 응답하여 아니오를 현행대로 이상방향으로 채점하는 것이 좋다는 결과가 나왔다. 요인분석결과 $\delta=0$ 로 사각회전했을 때의 요인간 상관이 .27 이고 요인 1과 요인 2에서의 부하값이 -.08, -.009로 모두 음수이므로 채점방향에 문제가 있는 문항으로 여겨진다. 306번 문항의 r_{ii} 값은 비교적 큰 음수이지만 준거집단이 그렇다고 할 확률이 38.8%, 정상집단은 33.1%로 비록 비율의 차이는 작지만 현행의 채점방향을 지지해 준다. 그러나 이 문항의 요인분석결과 요인구조행렬에서의 요인들과 이 문항사이의 상관이 모두 음수이므로 이 문항도 문제가 있다고 볼 수 있다. 322번 문항의 r_{ii} 값은 -.24로서 채점방향이 역전되어야 함을 강하게 암시하고 있고 준거집단의 아니오 반응빈도가 정상집단에 비해서 통계적으로 유의미하게 작기 때문에 아니오라는 원 채점방향은 문제가 있다. 104번 문항의 원 채점방향이 그렇다고 준거집단의 그렇다의 반응빈도율이 정상집단보다 크기 때문에 반응빈도율의 결과만으로는 문제가 없는 듯하지만 문항총점간 상관과 요인분석결과는 이와는 상반된 결과이다. 전체적으로 요인간 상관이 .26으로 양수이고 논의되고 있는 문항들이 두 요인 모두에 부적인 부하값을 갖기 때문에 문제가 있는 문항으로 보인다.

Ma 척도 이 척도에서는 표 8에 제시되어 있듯이 총 5개의 문항이 부적인 r_{ii} 값을 갖고 있다.

표 8. Ma 척도 문항총점간 상관 및 요인분석결과

문항	r_{ii}	1요인	2요인
101	-.1301	-.232	-.091
105	-.0697	-.153	.120
120	-.1085	-.249	.141
148	-.1770	-.291	-.043
289	-.0830	-.148	-.035
240	-.0342	-.170	.276

101번 문항의 r_{ii} 값은 -.1301로서 표집오차에 의해서 나타난 값으로는 보기 어렵다. 더욱이 척도에 대한 주성분분석결과 요인간 상관계수값이 .06이므로 척도내 차원간의 상관이 음수에 기인하여 r_{ii} 가 음수가 되었을 가능성은 거의 없다고 볼 수 있다. 또한 두 요인 모두에 음의 요인부하값을 갖기 때문에 현재의 문항채점방향이 역전되어야 함이 확실히 한 것 같다. 148번 문항도 양쪽요인에 대한 요인부하가 모두 음의 값을 갖기 때문에 r_{ii} 값이 큰 음수값을 갖는다. 따라서 이 문항의 채점방향도 거꾸로 되어야 할 것 같다. 105번, 120번, 289번, 240번 문항은 제 1 요인에만 부적인 요인부하값을 갖지만 가장 많은 변량을 설명하는 요인이기 때문에 r_{ii} 이 음수값이 되는 원인으로 작용했다고 볼 수 있다.

Si 척도 이 척도를 주성분분석하여 2 요인을 추출

하고 사각회전한 결과 요인간 상관은 -.03으로서 두 요인은 사실상 서로 독립적이라고 볼 수 있다. 따라서 요인간 상관 때문에 r_{ii} 값이 음수가 나올 수 있는 가능성은 희박하다.

표 9. Si 척도 문항총점간 상관 및 요인분석결과

문항	r_{ii}	1요인	2요인
359	-.2110	-.464	.134
469	-.0566	-.247	.180
481	-.0761	-.340	.220

표 9에서 보듯이 359번 문항의 r_{ii} 은 -.21으로서 대단히 크기 때문에 요인간 상관이 부적인 것에 기인한 것이라고 설명하기는 어렵다. 469번 문항은 점수의 분산에 대한 설명변량이 가장 큰 제 1 요인에서의 요인부하가 -.24 이므로 문항채점방향이 역전되어야 하는 경우인 것 같다. 481번 문항 역시 제 1 요인에 대한 요인부하가 -.34이고 제 2요인에 대한 요인부하가 .22로서 문항의 채점방향을 재고해야 할 필요가 있다고 생각된다.

논 의

본 연구에서는 심리측정학적인 면에서 MMPI에 적지않은 문제점이 있음을 밝혔다. 절대적인 기준은 없지만 .7 이하의 신뢰도값은 상업적으로 널리 사용되는 MMPI같은 심리검사의 경우에는 대체로 낮다고 판단되기 때문에 그만큼 측정오차가 많이 개재되어 척도점수를 신뢰롭게 해석하는데 어려움이 있을 것으로 예상할 수 있다. 개별척도의 해석도 문제가 되지만 측정오차로 인하여 코드나 프로파일의 불안정하기 때문에 코드타입이나 프로파일에 근거한 진단에는 세심한 주의를 기울여야 할 것 같다. 특히 이 논문에서 산출된 Mf 척도의 신뢰도 추정치가 정확하다면 이 척도를 근거로 어떠한 성격적 해석이나 임상적 진단도 삼가하는 것이 좋을 듯 싶다. 이 연구에서는 또한 상당수의 문항의 채점방향이 역전되어야 함을 보여주고

있다. 미국판 MMPI도 심리측정적으로 문제가 있음이 지적되고 있지만 한국판은 원판이 갖는 문제점 이외에 문화적 차이에 의한 부적절성까지 포함되어 이러한 결과를 가져온 것으로 짐작된다. 그러나 이 문항들이 왜 역전되어야 하는지를 내용적인 면에서 고찰하지 못했다. 사실 뚜렷한 문화적인 차이점이나 논리적인 문제점이 발견되지 않는 한 어떤 문항이 왜 문화권에 따라 달리 채점되어야 하는지에 대한 이유를 찾기는 용이하지 않을 것이다.

문항총점간 상관의 결과와 김영환의 논문의 결과가 일치 하지 않는 문항이 적지 않았다. 불일치의 가능한 이유로 김영환의 논문에서는 현행 채점방향과 관계없이 거의 모든 문항에서 준거집단의 응답비율이 정상집단의 응답비율보다 높았다. 이러한 환자집단의 목중반응의 경향의 이유는 분명하지 않지만 이 때문에 문항총점간 상관의 결과와 일치되지 않은 경우가 많았지 않나 생각된다. 이는 표준화집단의 응답비율을 김영환의 정상집단의 응답비율과 비교해 보면 그 결과의 특이성을 발견할 수 있다.

이 연구는 몇 가지 제한점을 갖는다. 첫째 대표준화자료를 사용한 것이 표본의 크기 면에서는 장점일 수도 있지만 표본에 포함된 대부분의 사람이 일상생활을 영위해 나가는데 임상적으로 큰 문제를 갖고 있지 않으며 정신과적 문제를 갖고 있는 사람이 적절히 의도적으로 표집되지 않았기 때문에 연구결과의 해석에 제한이 따른다. 둘째, 오직 비교검증을 위하여 주성분분석을 행하긴 했지만 요인 수의 결정등을 오직 통계적으로 처리하였기 때문에 그에 따른 결과의 해석이 자의적일 수 있었다. 셋째, 평행형 신뢰도 추정치를 산출할때 오직 반분검사의 평균을 유사하게 했고 평행검사가 되기위한 그밖의 다른 요건은 고려하지 않았기 때문에 만족스러운 평행형 추정치라고 보기 어려울 수도 있다.

이러한 제한점에도 불구하고 본 연구에서 문제가 있다고 판정된 문항에 대해서는 신중히 선정된 환자집단을 이용하여 그 문항이 기준집단과 정상집단을 구별하여 주는지를 검증하는 후속 조치가 이루어져야 한다. 그때까지는 문제가 되는 문항을 제외한 척도점

수와 원래의 모든 문항을 포함한 척도점수의 예측력을 비교하는 개별적 사례연구가 있어야 하겠다. 궁극적으로는 서로 상관이 없는 차원은 과감히 분리하여 세분된 척도를 제작하여 점수해석에 모호함을 없애야 한다. 코드형은 척도의 다차원성을 보완하기 위해서 나온 교육책일 수 있으므로 이 보다는 새로운 검사를 제작하는 것이 더 나을 수도 있다. 또한 외국에서 만든 널리 알려진 검사도 심리측정적으로는 문제점을 내포하고 있다는 사실도 염두에 둘 필요가 있다. 본 연구는 외국검사를 번역하여 사용하고자 할 때도 새로운 검사를 제작하는 마음가짐으로 문항내용의 타당성 검토, 문항분석, 규준화 등의 절차를 빠짐없이 거쳐야 함을 시사해 주고 있다.

참고문헌

- 김영환 (1982). MMPI의 진단변별기능분석. 고려대학교 대학원 박사학위 논문.
- 김영환, 김재환, 김중술, 노명래, 신동균, 염태호, 오상우(1994). 다면적 인성검사 실시요강. 서울: 한국가이던스
- 김중술(1988). 다면적 인성검사: MMPI의 임상적 해석. 서울: 서울대 출판부.
- Adler, T. (1990, April). Does the new MMPI beat the classic ? *APA Monitor*.
- Anastasi, A. (1988). *Psychological Testing (6th ed.)*. New York: Macmillan.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical & Modern Test Theory*. Orlando, Fl.: Holt, Rinehart and Winston.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement Theory for the Behavioral Sciences*. San Francisco, Ca: W. H. Freeman and Co.
- Greene, R. L. (1991). *The MMPI-2/MMPI: An Interpretive Manual*. Boston, Ma; Allyn and Bacon.
- Nunnally, J. C. (1978). *Psychometric Theory (2nd ed.)*. New York: McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory (3rd ed.)*. New York: McGraw-Hill.
- Rogers, T. B. (1995). *The Psychological Testing Enterprise*. Pacific Grove, Ca.: Brooks/Cole.

A Psychometrical examination of the Korean version of MMPI

Zoung-Sul Kim

Dong-Gun Park

Seoul National University Hospital

Korea University

An aim of the study was to examine the Korean version of MMPI in a psychometrical perspectives. The study attempted to more accurately reestimate the theoretical reliabilities of the MMPI subscales utilizing the cases in the restandardization. It was found that although the computed α reliabilities for most of the scales were similar to those in the restandardizaion, α for the Mf sclae was much lower than before. It was pointed out that because the α values less than .7 estimated for many of the MMPI scales, which were not high enough for the individual disgnoses, may cause the scale scores to be unstable, diagnoes based on the scores should be done very cautiously. The item analysis revealed that the 52 items were negatively correlated with the scale scores. To explore the possibilities that the negative item-total correlations were due to the negative correlations among the dimensions of each scale, an exploratory principal component analyses were performed. The present results suggested that most of the negative item-scale score correlations did not result from the negatively correlated scale dimensions. A cross validation using the response proportions of the criterion group and the normal group confirmed the results of the item-total correlations. The results indicated that an item analysis be indispensable prior to norming. Some considerions for future research study and the limitations of the present study were discussed.