

지능검사로서의 Raven Progressive Matrices 검사의 타당도연구: Standard Progressive Matrices 검사를 중심으로

전 덕 임 강 연 옥

충북대학교 심리학과

본 연구는 Raven Progressive Matrices검사의 한국판 개발을 위한 기초연구로서 대학생들을 대상으로 Standard Progressive Matrices검사(SPM)를 실시하여 지능검사로서의 SPM의 신뢰도와 타당도를 탐색하기 위한 목적으로 수행되었다. 405명의 대학생(남자 202명, 여자 203명)에게 지필검사형식의 SPM을 실시하였고, 그들 중 100명에게 KWIS를 함께 실시하였다. 이들의 SPM점수의 평균은 52.5(S.D.= 4.0)였고, 남녀의 차이는 발견되지 않았으며, 신뢰도는 내적일치도가 .73이었고 검사-재검사 신뢰도는 .62였다. SPM의 구성타당도를 확인하기 위하여 5개의 소검사점수를 변인으로 하여 요인분석을 한 결과 SPM은 단일요인으로 구성되어 있음이 밝혀졌고 그 요인은 전체변량의 39.2%를 설명하였다. 지능검사로서의 SPM의 공인타당도를 검토하기 위하여 SPM과 KWIS와의 상관관계를 조사한 결과 SPM은 전체 IQ와 동작성 IQ 및 언어성 IQ와 유의한 상관관계가 있음이 밝혀짐으로써 지능검사로서의 SPM의 타당성이 확인되었다. 그러나 언어성 IQ와 가장 높은 상관관계를 지닌 어휘문제가 SPM과는 거의 상관관계가 없다는 사실은 SPM을 이용하여 전체지능수준을 추정하려 한다면 어휘능력이 SPM과 함께 반드시 고려되어야 한다는 점을 시사하고 있다. 문항반응 이론에 의거하여 SPM의 문항들을 분석한 결과, SPM의 다수의 문항들이 쉬운 문항범위에 속하였고 추측도도 다소 높게 나왔으나 문항변별력은 대체로 적절한 수준에 속하였다. 이상과 같은 결과들은 SPM이 지능의 한 척도로서 신뢰롭고 타당한 검사임을 입증하고 있으나, 다양한 연령층의 정상인들과 환자들에 대한 후속연구를 통해서 본 연구의 결과가 다시 검증되어야 할 것이다.

“지능”은 총체적이고 복합적인 능력으로서 지각과 주의집중과정을 포함하며, 학습, 추상적인 추론, 정신적 표상형성, 문제해결, 의사결정 등의 고등정신과정 및 문제상황에서 효과적으로 반응하여 환경의 요구에 적응하는 능력 등을 포함한다(Sternberg, 1994). 이러한 지능의 평가는 환자의 인지적 능력과 현실생활에서의 적응능력에 대한 귀중한 정보를 제공해 주므로 임상 평가에서 없어서는 안될 중요한 부분을 차지하고 있다.

현재 임상장면에서 가장 많이 사용되고 있으며, 다양하고 풍부한 정보를 제공함으로써 지능검사로서의 우수성이 인정되고 있는 지능검사는 Wechsler 지능검사이다(Kaufman, 1990). 그러나 임상심리학자들은 환자가 오랜 시간동안 주의집중을 할 수 없는 상태에 있거나 임상장면에서의 시간적 제약으로 말미암아 1시간 이상이 소요되는 이런 지능검사를 실시하는 것이 현실적으로 불가능한 상황에 종종 처하게 된다. 이런 경우 임상심리학자들은 Wechsler 지능검사의 우수성을 인정하면서도, 검사시간이 보다 짧게 걸려서 피검자의 더 많은 협조를 기대할 수 있고, 실시와 채점이 간편하여 경제적이며 집단검사도 가능한 지능검사에 대한 필요성을 간절히 느끼게 된다. 이런 목적이 충족될 수 있는 지능검사들 중에서 전 세계적으로 널리 사용되고 있으며(Moran, 1986), 최근 들어서는 신경심리학적 도구로서의 유용성이 인정되어 특히 주목받고 있는 검사가 Raven Progressive Matrices검사이다.

Raven Progressive Matrices검사는 학교에서 습득된 지식과 언어 능력에 의해서 영향을 받는 기존의 지능검사들과는 달리 문화적, 교육적 배경이 영향을 주지 않도록 만들어진 검사로서, 지각적 관계를 파악하고 유추에 의해서 추론하는 능력을 측정하도록 고안된 비언어적 검사이다. 이 검사는 특히 시간제한이 없으므로 신체적 장애를 지닌 환자들이나 뇌손상환자들 또는 정서적 문제를 지니고 있거나 검사에 대한 동기가 부족한 사람 등과 같이 빨리 반응하기 어려운 사람들에게 매우 유용한 검사로

밝혀지고 있다(Jensen, 1980).

1. Raven Progressive Matrices 검사의 구성

Raven Progressive Matrices검사(이하 RPM으로 약칭함)는 피검자의 나이와 지적 능력수준에 따라서 선택적으로 사용되는 Standard Progressive Matrices(1938, 이하 SPM으로 약칭함), Colored Progressive Matrices(1947)와 Advanced Progressive Matrices(1965)의 3 종류의 검사로 구성되어 있다. SPM은 그 중에서 가장 널리 사용되고 있는 것으로 교육, 국적 또는 신체적 조건에 관계없이 모든 연령의 사람에게 사용될 수 있도록 고안되었다. 총 60문항으로 구성되어 있으며 평균 20-40분이면 마칠 수 있으나 시간제한은 없다. 피검자의 과제는 한 부분이 미완성되어 있는 특정한 의미가 없는 도형이나 패턴으로 구성된 matrix를 보고 그 미완성 부분에 들어맞는 조각을 6개 또는 8개의 보기 중에서 골라 답하는 것이다. 전체 60문항은 과제의 종류에 따라서 각각 12문항씩으로 구성된 5개의 소검사로 나누어져 있는데, 소검사 A는 연속적인 패턴(continuous patterns), 소검사 B는 모양유추(analogies of figures), 소검사 C는 모양의 점진적인 전개(progressed development of figures), 소검사 D는 모양의 배열과 조합(arrangement and combination of figures), 소검사 E는 전체모양을 그 구성부분으로 해체하는(resolving figures into their constituent parts) 과제들로 구성되어 있다. 소검사들은 A에서 E로 갈수록 점차 어려워지며 각 소검사내의 문항들도 쉬운 문제에서 어려운 문제의 순서로 난이도별로 배열되어 있다.

Colored Progressive Matrices는 SPM과 유사하나 도형이 SPM에 비해서 단순하고, 총문항의 수도 36문항으로 SPM에 비해서 적다. 검사도구는 피험자의 주의를 끌 수 있도록 도형의 바탕에 색채가 들어가 있는, 각각 12문항씩으로 구성된 3개의 소검사(A, Ab, B)로 구성되어 있다. SPM이 모든 연령의 사람들에게 광범위하게 쓰여지고 있는데 반해서

CPM은 주로 어린 아동과 노인들, 정신 지체자들이나 뇌손상환자들에게 사용되어지고 있다. Raven, Court와 Raven(1977)은 유추에 의한 추론을 할 수 있는 지적 발달이 일단 이루어졌다면 CPM보다는 SPM을 사용하는 것이 더 신뢰롭다고 주장하였다.

Advanced Progressive Matrices는 우수한 지능을 가진 것으로 추정되는 피검자들에게 사용되며 지적 능력이라기 보다는 지적 효율성을 측정하는 검사이다. I과 II의 두 소검사로 구성되어 있으며 I은 II를 시작하기 전의 연습검사로 제시되고, II는 점차 매우 어렵고 복잡해지는 36개의 문항으로 구성되어 있다.

2. Raven Progressive Matrices 검사는 무엇을 측정하는 검사인가?

Spearman(1927)은 독립적으로 개발된 여러 지능 검사들간에 서로 높은 상관관계가 존재한다는 사실을 근거로 하여 지능검사들간에는 공통요인인 "g요인"이 존재하며, 지능은 이러한 일반요인(general factor: g요인)과 음악적 재능이나 기계적 능력과 같은 특수요인(special factor: s요인)으로 구성되어 있다는 지능의 2요인설을 주장하였다. 그는 또한 g요인은 과거에 경험하지 못했던 모호한 문제해결상황에 처하였을 때 그 상황에 대한 새로운 통찰을 얻고 문제를 해결할 수 있는 구성개념을 유추해내는 추론능력(eductive ability)과 이미 과거경험이나 학습을 통하여 습득한 지식이나 정보를 재생산하는 능력(reproductive ability)으로 구성되어 있다고 제안하였다.

Raven은 Spearman의 이러한 이론적인 입장에 따라서 RPM을 제작하였고, RPM으로 추론능력을 측정하고 어휘검사인 Mill Hill Vocabulary Scale을 사용하여 재생산능력을 측정한 후, 이 두 검사의 수행능력을 종합하였을 때에야 Spearman이 주장한 g 즉 일반적인 지적능력을 측정할 수 있다고 주장하였다. 그는(Raven et al., 1977) RPM은 관계를 추론하는 능력을 측정하기 위한 관찰력과 사고력 검사로서

그 사람의 과거 경험이나 현재의 언어적 의사소통 능력과는 무관하고 지능의 중요한 요소인 언어적 성분을 포함하고 있지 않으므로 RPM을 일반적 지능 검사로 여겨서는 안된다는 점을 강조하였다. 즉, 이미 습득된 지식과는 관계없는 지적 활동의 현재능력은 RPM을 통해서 측정하고 과거에 이미 습득된 어휘적 정보의 축적상태는 Mill Hill Vocabulary Scale을 통해서 측정함으로써만이 전체적인 지적 능력의 평가가 이루어질 수 있다는 것이다.

RPM이 무엇을 측정하는 검사인지 그 구성타당도에 관하여 이후 연구자들은 서로 다른 연구결과들을 보고하였다. Raven은 앞서 기술된 바와 같이 RPM이 일반지능(g)을 측정하는 검사로 사용되는 것에 대하여 주의를 주었지만, Spearman과 Jones(1950)는 RPM을 비언어적 검사들중에서 g를 측정하는 최고의 검사라고 하였고 McKenna(1984)도 RPM을 지능의 g요인을 측정할 수 있는 좋은 검사라고 주장하였다. 이들의 주장은 RPM에서 단일요인을 발견한 요인분석연구(Orme, 1966)에 의해서 뒷받침되었다. 그러나 다른 연구자들(Banks & Sinha, 1951; Dolke, 1976)은 요인분석결과 RPM에서 귀납적 추론, 지각속도, 공간능력, 지각적 정확성, 개념형성 등으로 이름붙일 수 있는 중다요인을 발견하였고, RPM의 점수를 일반적 지능(g)의 측정치로 사용하는 것에 반대하였다.

지능의 한 측정치로서 RPM의 효율성에 대한 연구는 기존 지능검사들과 RPM과의 상관관계를 연구하는 공인타당도(concurrent validity) 연구가 주종을 이루었다. 특히 WAIS와의 상관연구가 널리 행해졌는데 SPM의 점수가 WAIS-R의 전체 IQ점수와 매우 유의미한 정적상관관계(.75-.85)가 있음이 보고되었으며(Burke, 1985; Hall, 1957; Shaw, 1967) SPM점수가 WAIS-R의 전체 IQ점수의 추정치로 사용될 수 있음이 여러 연구에서 시사되었다(O'Leary, Rush, & Guastello, 1991; Shaw, 1967). Desia(1955)는 Wechsler-Bellevue 검사의 언어성 척도와 SPM과의 상관관계를 연구한 결과 두 검사

사이의 상관관계가 높지 않으며, 특히 상식문제와 어휘문제와의 상관관계가 가장 낮았다고 보고하였다. 그는 이 결과를 SPM이 현재의 지적 효율성을 재는 척도인데 반해서 Wechsler-Bellevue의 언어성 척도는 과거에 습득한 지식을 통하여 지능을 측정하기 때문인 것으로 해석하였다. 그러나 McLaurin와 Farrar(1973)는 SPM이 WAIS 전체 IQ와 동작성 IQ 뿐만 아니라 언어성 IQ와도 유의한 상관관계가 있다는 연구결과를 보고하였다. RPM의 언어적 요소의 포함여부는 추후연구가 필요한 과제로 남아있다.

3. 신경심리학적 평가도구로서의 Raven Progressive Matrices검사

신경심리학영역의 최근 성장과 함께 RPM은 지능 검사로서 뿐만 아니라 신경심리학적 평가도구로서도 그 유용성이 입증되고 있다(Milberg, Hebben, & Kaplan, 1986). 두뇌손상환자들은 일반적으로 병전(病前) 수준의 지적 능력을 유지하지 못하며 따라서 문제해결에 효과적으로 접근하지 못하고, 목표-지향적 행동이 체계적이지 못하며 손상의 부위와 정도에 따라서 다양한 인지적 능력들이 손상된다. 두뇌손상환자들의 지적 능력을 측정하는 다양한 지능검사들이 개발되었으나 적당한 수준의 시력을 지니고 간단한 지시만 이해하면 수행이 가능한 RPM은 언어장애나 운동장애를 지닌 뇌손상환자들의 지능측정을 위해서 기존의 지능검사들에 비해서 유용하게 쓰일 수 있다. 특히, RPM은 뇌기능변화에 매우 예민하기 때문에 유용한 두뇌손상의 선별도구로 사용될 수 있음이 시사되었다(Gainnotti, Miceli, & Caltagirine, 1979; Irving, Robinson, & McAdam, 1970).

두뇌손상환자들이 보이는 RPM 반응양상에 대한 연구들(e.g., Gainnotti et al., 1979; Urmer, Morris, & Wendland, 1960)은 두뇌손상집단이 양적으로 뿐만 아니라 질적으로 통제집단과 다른 수행을 보였음을 보고하였다. 양적인 측면에서 두뇌손상환자들은 WAIS검사결과에 드러난 어떤 지적 감퇴보다도

더 저하된 RPM의 수행수준을 보였으며, 특히, 정신지체 아동과 노인성 치매환자에게서 가장 두드러지게 저하된 점수를 보였다(Raven, Court, & Raven, 1991). 알콜중독의 경우는 만성화될수록 수행점수가 저조하였고 금주를 하면 점수가 증진되었다(Fabian & Parson, 1983). 실행증 또는 약물중독환자의 경우도 점수가 낮았고(Kertesz & Hooper, 1982), 실어증환자도 낮은 점수를 보였는데 언어치료를 받은 후 검사점수가 증진되는 경향이 있었다(David & Skilbeck, 1984). 이는 RPM이 동작성 요인뿐만 아니라 언어성 요인도 포함하고 있다는 McLaurin와 Farrar(1973)의 주장과 일치하는 결과이며 RPM이 환자의 치료과정중 그 경과에 대한 가치있는 정보를 제공할 수 있음도 또한 시사한다(Raven, Court, & Raven, 1990). 질적인 차이로는 두뇌손상집단이 그림-배경의 반전오류를 많이 범하였고, 상대적으로 어려운 소검사들은 잘 맞추면서 가장 단순한 소검사인 소검사 A에서는 잘못된 등 문항난이도와 불일치하는 반응양상을 보였다(Gainnotti et al., 1979).

4. 연구목적

RPM의 임상적 유용성에도 불구하고 한국에서는 RPM이 거의 사용되지 않았고 따라서 그에 대한 기초연구도 전무한 실정이다. 단지 최근에 대학병원 신경과와 정신과를 중심으로 새롭게 사용되기 시작한 신경심리검사배터리인 "비엔나 검사총집(Vienna Test System)"에 SPM의 computerized version이 포함되어 있어서 환자들을 대상으로 SPM이 사용되기 시작하였으나, computer채점으로 얻어지는 SPM점수와 추정IQ점수가 한국에서의 표준화과정을 거친 것이 아니어서 임상적인 자료로 쓰여지지 못하고 있다. 본 연구는 전체 인구를 대상으로 한 RPM의 한국판 표준화를 위한 기초연구로서, 임상장면에서 더 손쉽게 사용할 수 있는 지필검사 방법으로 SPM을 대학생들에게 실시하여 SPM의 신뢰도와 타당도를 검토하고, 이들 결과와 문항

분석결과를 바탕으로 지능검사로서의 SPM의 유용성을 검토해 보고자 수행되었다.

방 법

1. 피검자

심리학 과목을 수강하는 405명의 대학생(남자 202명, 여자 203명)을 대상으로 하였다. 피검자의 평균연령은 남자가 21.7세(S.D.= 3.5) 여자가 20.1세(S.D.= 1.4)로 전체 평균연령은 20.9세(S.D.= 2.8)였다.

2. 검사도구 및 절차

검사도구로 SPM Booklet(Raven et al., 1991)와 KWIS(Korean Wechsler Adult Intelligence Test: 전용신, 서봉연, 이창우, 1963)가 사용되었다. SPM은 30-100명의 단위로 전체 405명에게 수업시간동안 집단검사로 실시되었다. SPM을 실시한 피검자들중 남녀 각각 50명씩 100명의 지원자를 모집하여 SPM 실시 후 2주일 이내에 KWIS검사를 연구자들이 실시하였다. SPM에 소요된 시간은 남자가 평균 22분(S.D.= 6.6), 여자가 20.3분(S.D.= 6.8)으로 전체 평균은 21.2분(S.D.= 6.8)이었고, KWIS검사에는 피검자당 약 1시간 30분 정도의 시간이 소요되었다.

3. 자료분석

요인분석(factor analysis)과 중다회귀분석(multiple regression analysis)을 위하여 SPSS-PC를 사용하였고, SPM 문항의 적절성을 검토하기 위해서 PC-BILOG(Binary LOGistics Models; 상태제, 1991)를 이용하여 문항반응이론에 따른 문항분석을 실시하였다.

결 과

1. 신뢰도

SPM의 내적일치도 계수 Cronbach's α 는 .73이었

고, 본 연구의 피검자들 중 42명에 대해서 2주일의 간격을 두고 재검사를 실시한 결과, 검사-재검사 신뢰도는 .62였다.

2. 점수의 분포

SPM점수의 백분위점수분포와 A에서 E까지 5개 소검사 각각의 평균점수(표준편차)는 표 1과 같다. 전체 점수의 평균은 52.5(S.D.= 4.0)로 이는 중, 고등학생집단을 대상으로 SPM의 computerized version을 실시하여 이철과 연구자들(이철, 유희정, 김창윤, 한오수, 박인호, 1993)이 보고한 만 13세에서의 41.9(S.D.= 9.4), 15세에서의 45.9(S.D.= 7.8)보다 유의하게 높은 점수이다. 대학생집단의 이와 같은 높은 수행수준은 SPM점수가 어려서부터 20세까지는 계속 상승하여 20세를 전후해서 절정을 이루고, 20-25세 이후에 나이와 함께 점차 감소된다는 것을 보여준 Raven 등(Raven, Court, & Raven, 1988)과 Persaud(1987)의 연구결과와 일치하는 결과이다.

한편 소검사 A에서 E로 갈수록 검사의 평균점수가 점점 낮아지고(A-B: $t(808)=7.48, p<.001$: B-C: $t(808)=6.12, p<.001$: C-D: $t(808)=9.70, p<.001$: D-E: $t(808)=12.09, p<.001$) 표준편차가 점차 커지는 것을 통해서 소검사 A에서 E로 갈수록 문항이 점차 어려워짐이 확인되었다.

표 1. SPM점수의 백분위점수와 SPM의 소검사별 평균점수(표준편차)

1) SPM점수의 백분위점수(percentile score)

백분위점수	SPM 점수
1	39
5	45
10	48
25	50
50	53
75	55
90	57
95	58
99	60

2) SPM의 소검사별 평균점수(표준편차)

소검사	남자(n=202)	여자(n=203)	전체(N=405)
A	11.64(.61)	11.68(.56)	11.66(.59)
B	11.15(1.27)	11.25(.98)	11.20(1.13)
C	10.96(1.15)	10.67(1.22)	10.77(1.19)
D	9.96(1.33)	10.07(1.45)	10.01(1.39)
E	8.91(1.94)	8.73(1.88)	8.82(1.91)
전체	52.51(4.12)	52.39(3.94)	52.45(4.03)

동양사회에서의 사회적 요인과 관련하여 SPM의 수행에서 성차이가 존재할 수도 있을 것이라는 Court(1983)의 예상과는 달리 본 연구결과에서는 전체검사점수 평균이 남자 52.5(S.D.= 4.1), 여자 52.4(S.D.= 3.9)로 남녀의 차이가 발견되지 않았다($t(403) = .3, ns$). 따라서 본 연구의 모든 자료분석은 남녀를 구분하지 않고 전체집단에 대해서 수행되었다.

3. 요인분석

SPM의 구성타당도를 검토하기 위한 요인분석에 앞서서 5개 소검사 각각의 내용타당도를 검토하였다. A에서 E까지의 소검사들이 Raven이 밝혔듯이 각각 연속적 패턴, 모양유추, 모양의 점진적 전개, 모양의 배열과 조합, 전체모양을 그 구성부분으로 해체하는 과제로 구성되어 있는지 소검사별로 각 문항들을 검토한 결과 모든 소검사들이 그 제작의 도대로 내적일관성을 지닌 문항들로 구성되어있음이 확인되었다.

이처럼 각 소검사의 내용타당도가 검증되었으므로 각 소검사 점수를 변인으로하여 공통요인분석(common factor analysis)의 주축분해법(principal axis factoring)을 사용하여 요인분석하였다. 고유치(eigenvalue)와 scree도표를 검토한 결과 1개의 요인을 추출하는 것이 타당하다는 결론에 이르렀는데 이 요인의 고유치는 1.96이었고 총변량의 39.2%를 설명하였다. 각 소검사의 요인부하량은 표 2와 같다. 요인부하량이 0.4이상인 변인만을 유의미한 변인으로

고려하는 Gorsuch(1974)의 기준에 의해서 살펴보면 5개 소검사들 중 소검사 A는 유의미하지 않은 변인으로 밝혀졌다.

표 2. SPM검사의 요인구조

SPM 소검사	요인부하량
A	.29
B	.40
C	.62
D	.51
E	.62
설명변량	39.2%

4. SPM과 KWIS와의 상관관계 및 지능지수(IQ)의 추정

본 연구의 요인분석에서 밝혀진 것처럼 SPM이 하나의 단일요인으로 구성되어 있는 검사라면 SPM이 측정하는 것은 무엇인가? 그 단일요인의 성질을 밝히고 공인타당도(concurrent validity)를 검토하기 위해서 SPM과 함께 100명의 피검자들에게 실시된 KWIS검사자료를 이용하여 SPM의 총점과 KWIS의 11개 소검사의 환산점수, 언어성 IQ, 동작성 IQ 및 전체 IQ와의 상관계수(Pearson product-moment correlation coefficient: r)를 구하였다(표 3).

SPM은 전체 IQ($r=.451$) 및 동작성 IQ($r=.443$)와 유의미한 상관관계가 있음이 밝혀졌고, 동작성 IQ 보다는 낮은 수준이나 언어성 IQ도 역시 SPM과 유의미한 상관관계($r=.290$)를 지니고 있었다. 각 소검사별로 살펴보면 SPM은 토막짜기($r=.391$)와 가장 높은 상관관계를 나타냈고, 바퀴쓰기를 제외한 모든 동작성검사들에 대하여 차례맞추기($r=.291$), 빠진곳찾기($r=.283$), 모양맞추기($r=.262$)의 순으로 유의미한 상관관계를 나타내었다. 언어성검사중에서는 산수문제($r=.347$)가 토막짜기와 비슷한 수준의 높은 상관관계를 나타내었고, 공통성문제($r=.$

SPM	상식 문제	이해 문제	산수 문제	공동성 문제	숫자 따라하기	어휘 문제	바꿔 쓰기	빠진 곳 찾기	토막 짜기	차레 맞추기	모양 맞추기	언어성 IQ	동작성 IQ	전체 IQ	
SPM	1.000														
상식문제	.116	1.000													
이해문제	-.033	.163	1.000												
산수문제	.347***	.058	-.027	1.000											
공동성문제	.198*	.043	.114	.359***	1.000										
숫자따라하기	.195	-.037	-.020	.215*	.000	1.000									
어휘문제	.066	.197	.223*	.174	.206*	.167	1.000								
바꿔쓰기	.111	-.142	-.204*	.184	.243*	-.069	-.011	1.000							
빠진곳찾기	.283**	.230*	.114	.221*	.115	.127	.088	.036	1.000						
토막짜기	.391***	.015	-.024	.303**	.353***	.236*	.015	.054	.450***	1.000					
차레맞추기	.291**	.168	-.004	.099	.238*	.036	-.029	-.047	.339**	.428***	1.000				
모양맞추기	.262**	.266**	-.032	.129	.191	.155	.180	.122	.189	.206*	.324**	1.000			
언어성 IQ	.290**	.348***	.438***	.538***	.526***	.526***	.625***	.011	.253*	.240*	.108	.252*	1.000		
동작성 IQ	.443***	.189	-.027	.290**	.368***	.158	.094	.361***	.641***	.669***	.713***	.634***	.304**	1.000	
전체 IQ	.451***	.343***	.260**	.515***	.562***	.428***	.451***	.221*	.551***	.552***	.497***	.540***	.819***	.792***	1.000

* $p < .05$ ** $p < .01$ *** $p < .001$ (2-tailed significance)

표 4. 회귀분석결과

예언변인	R ²	B	SE B	Beta	t	p
SPM	.203	.437	.083	.423	5.284	.000
어휘문제	.178	1.531	.290	.423	5.288	.000
(상수)		76.754	5.887		13.038	.000

198)가 약한 상관관계를 나타내었을 뿐, 상식문제, 이해문제, 어휘문제 등과 SPM은 거의 상관관계가 없는 것으로 나타났다. 이와 같은 SPM과 KWIS의 소검사들과의 관계에서 추론하여 볼 때 본 연구에서 SPM과 언어성 IQ가 유의미한 상관관계를 나타낸 것은 SPM과 산수문제, 공통성문제 및 숫자따라하기와의 상관관계에서 비롯된 것임을 알 수 있다.

한편, 어휘문제가 언어성검사들중에서 언어성 IQ와 상관관계가 가장 높으나($r=.625$) SPM과 어휘문제는 서로 상관관계가 거의 없으며($r=.066$), SPM과 전체 IQ와의 상관관계의 크기($r=.451$)가 어휘문제와 전체 IQ와의 상관관계의 크기와 같다는 사실은 Raven이 SPM을 제작하였을 때 의도한 바대로 SPM과 어휘검사의 결과가 종합되어야 지능이 정확하게 측정될 수 있다는 점을 시사한다. 따라서 회귀분석을 통해서(표 4 참조) SPM점수와 어휘문제점수를 가지고 KWIS의 전체 IQ점수를 추정할 수 있는 회귀방정식을 다음과 같이 구할 수 있다.

$$\text{KWIS 전체 IQ} = .437(\text{SPM점수}) + 1.531(\text{어휘문제점수}) + 76.754$$

표 4에서 볼 수 있듯이 SPM과 어휘문제 각각의 Beta계수의 값이 서로 같음은 SPM점수와 어휘문제 점수가 전체IQ점수에 기여하는 정도가 같음을 시사한다.

5. 문항분석

SPM 문항들의 난이도와 변별력을 검토하기 위해서 문항분석을 실시하였다. 문항의 난이도를 “정확 통과율”로 표현하고, 문항의 변별력을 각 문항이 전체검사점수에 비추어 개인차를 어느 정도나 잘 반영하는지 “문항·총점 상관계수” 등으로 표시하는 고전검사이론에 의거한 문항분석은 쉬운 검사도구를 사용하느냐 어려운 검사도구를 사용하느냐에 따라 피검자의 수준이 다르게 측정되고, 피검자집단이 높은 능력의 집단이나 또는 낮은 능력의 집단이나에 따라 검사문항에 대한 통계치들이 영향을 받는다는 단점을 지니고 있다(성태제, 1991). SPM이 본 연구의 피검자인 대학생들에게 쉬운 검사임이 점수분포에서 이미 밝혀진 바 있으므로 고전검사이론에 의거한 문항분석을 한다면 그로부터 얻어진 난이도나 변별도 등의 수치들이 대학생집단이라는 피검자집단의 특성에 의해서 영향을 받을 가능성이 시사된다. 따라서 본 연구에서는 문항반응이론¹⁾에 의거하여 SPM 문항들을 분석하였다. 문항반응이론은 각 문항들이 고유한 특성을 지니고 있기 때문에 피험자 집단의 특성에 의해서 문항의 특성이 달리 추정되지 않으며(문항특성의 불변성), 피험자는 개인이 지니고 있는 고유한 능력 수준이 있기 때문에 어떤 검사도구를 택하여도 달리 추정되지 않는다는 입장(피험자 능력의 불변성)에서 문항을 분석하는 이론이다.

먼저 각 문항이 문항반응이론의 3모수 모형에 적합한지를 보여주는 chi-square검증을 실시하여 본

1) 문항반응이론에 대한 자세한 내용은 성태제(1991)를 참조하기 바랍니다.

논 의

연구의 SPM 자료가 문항반응이론 모형에 적합한가를 살펴보았다. 그 결과 4문항(C-8, D-10, E-10, E-11)을 제외하고는 본 자료가 문항반응이론에 의한 분석에 타당한 자료인 것으로 나타났다. 문항반응이론에 의거한 "문항 난이도"를 산출한 결과 SPM의 다수의 문항들이(42문항) 쉬운 문항범위(-.5 이하)에 속하였고, C-12와 E-9는 중간범위(-0.5 - +0.5), D-11, E-11, E-12는 어려운 범위(+0.5 - +2.0)에 속하였으며, D-12는 매우 어려운 범위(+2.0 이상)에 속한 문항인 것으로 나타났다(부록 1 참조). 문항이 피험자를 능력에 따라 변별하는 정도를 나타내는 "문항 변별력"은 총 41문항이 적절한 범위(.39-.79)에 속하였고, 4문항(E-2, E-3, E-4, E-10)은 높은 범위 (.80-.99)에 속하였으며, E-5, E-11, E-12는 변별력이 매우 높은 범위(1.0 이상)에 속하는 것으로 밝혀졌다. 전체 60문항의 평균 변별도는 .553(S.D. = .193)으로 SPM이 전체적으로 적절한 수준의 변별력을 지니고 있음을 나타내었다. 능력이 전혀없는 사람이 추측해서 문항의 답을 맞출 확률인 "문항 추측도"의 평균은 .146(S.D.=.012)이었다. SPM은 보기가 6개인 문항 24개와 8개인 문항 36개로 구성되어 있으므로 이 추측도의 수준은 다소 높은 수준에 속하는 것으로 사료된다.

문항반응이론에 근거하여 얻은 검사정보곡선은 검사가 전체능력범위에서 얼마나 정확하게 능력을 측정하는가를 보여주고, 검사가 목적에 맞게 제작되었는가에 관한 정보를 준다. 그림 1의 검사정보곡선은 본 연구에 참여한 전체 피검자의 능력 수준을 0을 중간점, 1을 측정단위로 정하고 능력범위가 $-\infty$ 에서 $+\infty$ 까지 존재한다고 가정할 때, SPM은 그 능력 수준이 -2.5에 해당하는 사람들에게 대해서 측정의 오차없이 검사에서 측정하고자 하는 능력을 가장 정확하게 측정할 수 있는 검사임을 보여준다.

RPM의 가장 큰 잇점중의 하나로 자주 언급되는 것은 RPM이 문화적으로 공평한, 즉 문화적 배경이 큰 영향을 미치지 않는 검사라는 점이다. 이 점은 여러 국가들의 표준화된 기준들이 문화권간에 차이가 없이 상당히 안정되어 있다는 사실과(Raven, Court, & Raven, 1988) SPM과 다른 지능검사들과의 상관계수가 문화권에 관계없이 비슷하다는 사실(Carlson & Jensen, 1981; Powers, Barkan, & Jones, 1986)에서 입증되었다. 최근에 교도소 수감자를 대상으로 행해진 연구에서 Templer(1992)는 SPM의 수행수준은 정규교육의 정도와 무관하며 인지능력을 측정하는 다른 검사들에 비해서 SPM에서 민족간의 차이가 덜 나타났다고 보고하였다. 그러나 Owen(1992)은 남아프리카의 4개 인종을 대상으로 연구한 결과 신뢰도, 문항난이도, 문항변별력 등의 심리측정적인 측면에서는 SPM이 문화적으로 공정한 검사이지만, 수행수준에 있어서는 인종간에 큰 차이가 있음을 발견하였고 인종별로 각기 다른 기준이 필요하다고 주장하였다. 한편 Freeman(1984, 재인용)은 8세와 12세의 시골과 도시의 흑인 국민학교 아이들을 대상으로 SPM을 실시한 연구에서 시골아이들보다 도시의 아이들의 SPM점수가 더 높은 것을 발견하였고, SPM을 수행하는데 필요한 인지양식이 시골집단보다 도시집단내에서 더 발달된다는 결론을 내렸다. 즉 다양한 외적 자극과 문화적 영향이 SPM의 수행과 상관되어 있다는 것이다. Burg와 Belmont(1990)은 또한 같은 점수를 받더라도 다른 문화권의 사람들은 다른 해결 전략을 사용할 수 있으므로 문항의 세심한 분석이 이루어져야 한다는 점을 지적하였다. 그러나 Jensen(1980)은 문화적이거나 인종적인 차이가 있는 것처럼 보이는 이런 결과들이 실제로는 사회경제적 지위(Bartlett, Newbrough, & Tulkin, 1972)나 교육수준(Vincent & Cox, 1974)이 검사결과에 영향을 미친 것이라고 주장하였다. 본 연구에서 얻은 대학생들

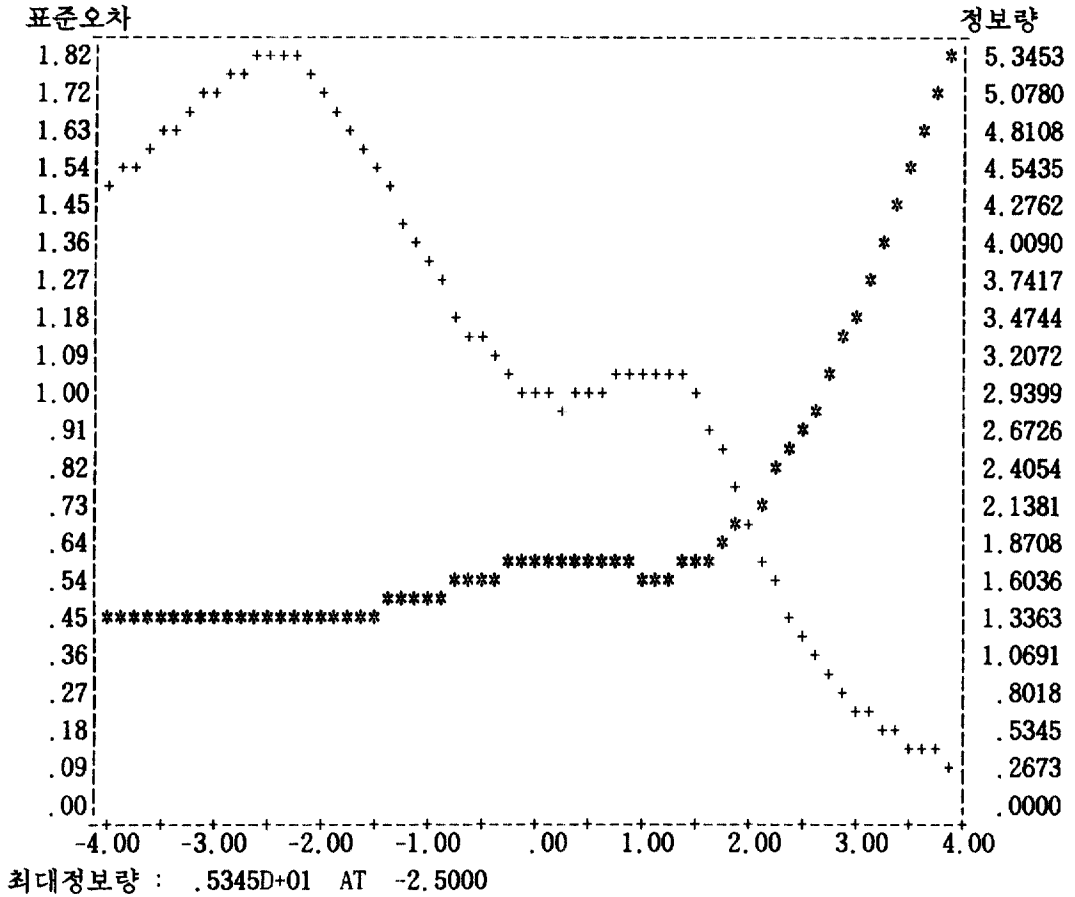


그림 1. SPM의 검사정보곡선

의 표준은 50 percentile점수가 52.5점으로 영국 (44점)이나 미국(45점)의 표준보다 상당히 높은 점수를 나타내고 있다. 이런 차이가 미국과 영국의 표준이 완성된 1979년보다 다양한 문화적 혜택 및 현대문명의 발달로 전반적인 지적 수준이 향상된 때문인지 또는 미국과 영국의 표준은 다양한 교육수준의 피검자들로부터 얻어진 것인데 반해서 본 연구의 피검자들은 대학생들로만 구성되었다는 차이에 기인한 것인지, 또는 다른 어떤 문화적이거나 인종적인 차이를 반영하는 것인지에 관해서는 추후연구가 필요로 된다.

요인분석결과 SPM의 5개 소검사로부터 단일요인

이 추출되었으므로 SPM이 무엇을 측정하는 검사인지를 확인하기 위해서 SPM의 총점과 KWIS검사와의 상관계수를 산출하였다. 그 결과 SPM은 KWIS의 전체 IQ와 .451의 유의한 상관관계를 지니고 있고, 동작성 IQ 뿐만 아니라 언어성 IQ와도 유의한 상관관계가 있음이 발견되었다. SPM과 전체 IQ와의 유의한 상관관계는 SPM이 지능의 한 척도임을 입증해 주며, 동작성 IQ 및 동작성 소검사들과 SPM과의 유의한 상관관계는 "지각적 관계를 파악하고 유추에 의해서 논리적으로 추론하는 능력을 측정하는 비언어적 검사"라는 SPM의 특성으로 미루어 예견되었던 결과이다. 그러나 SPM과 언어성 IQ와의 상관관계는

어떻게 설명될 수 있을까? SPM과 KWIS의 언어성 소검사들과의 상관관계를 분석하여 보면, SPM과 언어성 IQ와의 상관관계는 언어성검사들중에서 어휘문제나 상식,이해문제 등과 같이 상대적으로 언어적 요소를 많이 지닌 소검사들과 SPM과의 상관관계에서 비롯된 것이 아니라, 산수문제와 공통성문제, 숫자따라하기와 같이 언어적 요소가 상대적으로 적은 소검사들과 SPM과의 상관관계를 반영한다는 것을 알 수 있다. 그렇다면 이 검사들은 SPM과 어떤 공통요소를 지니고 있는가? 소검사들중에서 토막짜기 다음으로 SPM과 높은 상관관계를 보인 산수문제의 경우에는 산수문제가 측정하는 능력중의 하나로 알려져 있는 "일반적인 추론능력(general reasoning ability)" (Glasser & Zimmerman, 1967)과 SPM의 문제해결에 요구되는 추론능력과의 유사성에서 그 상관관계를 유추할 수 있으나, 공통성문제나 숫자따라하기와 SPM의 공통요소에 대하여서는 각 검사과제들에 대한 보다 자세한 분석이 행해진 후에야 밝혀질 수 있을 것이다. 상관관계의 연구에서 또 한 가지 주목되어지는 사실은 KWIS의 언어성검사들중에서 언어성 IQ와 가장 상관관계가 높은 어휘문제가 SPM과는 상관관계가 거의 없으나, 전체 IQ와는 SPM과 동일한 수준의 상관관계를 지니고 있다는 사실이다. 현시점에서 SPM자체가 언어적 요소를 전혀 포함하고 있지 않다고 말할 수는 없으나, SPM과 어휘문제와의 위의 관계를 고려한다면, SPM점수로 전체지능수준을 추정할 때에는 반드시 어휘검사를 실시하여 그 점수를 SPM 점수와 함께 고려하여야 할 것으로 사료된다. 이는 SPM이 어휘검사와 함께 사용되어야 정확한 "g"를 측정할 수 있다는 Raven 등(1977)의 주장과 일치된 결론이며, SPM점수만을 가지고 IQ를 추정하는 Vienna Test System의 방법에 문제가 있음을 시사한다.

SPM 문항의 난이도와 변별력에 대하여서는 많은 비판이 제기되어 왔다(Banks & Sinha, 1951). 즉 SPM에는 쉬운 문항과 중간 정도의 난이도를 지닌 문항이 너무 많아서 변별력이 떨어지고(Dolke, 1976; Slater, 1948), 문항배열에 있어서도 문제점이

있음(Keir, 1949)이 지적되었다. 특히 소검사 A는 선행연구들에서(e.g., Eisenburg, Levy, & Kirsch, 1970) 너무 쉬운 문항이 많고 변별력도 떨어진다는 비난을 받아 왔는데, 요인분석결과 추출된 단일요인에 대해서 소검사 A만이 유의하지 않은 변인으로 밝혀진 본 연구의 결과도 역시 소검사 A가 수정되거나 재구성될 필요가 있음을 시사하고 있다. 또한 본 연구에서 행해진 문항반응이론에 의한 문항분석 결과에 의하면 선행연구에서 밝혀진 것처럼 SPM 문항들중 문항난이도에 있어서 쉬운 범위에 속하는 문항이 매우 많고 중간범위나 어려운 범위에 속하는 문항의 수가 상대적으로 적었으며, 문항추측도는 다소 높은 것으로 나타났으나 문항변별력이 있어서는 60문항중에서 48문항이 적절한 수준 또는 우수한 수준에 속하는 것으로 밝혀졌다. 본 연구의 이와 같은 결과는 각 문항의 난이도, 추측도, 변별도를 고려하여 문항을 제외시키거나 문항의 순서를 바꾸는 등 검사를 재구성하여야 할 필요성이 있음을 시사한다.

검사정보곡선은 SPM이 특히 지적능력수준이 낮은 집단에 대하여 정확하고 신뢰로운 정보를 제공할 수 있음을 보여주었다. 문항반응이론에 의한 문항분석은 그 결과가 피검사집단의 능력수준에 의해서 영향을 받지 않는다는 장점을 지니고 있으나, 문항반응이론을 적용할 때에 반드시 기억하여야 하는 것은 "동일한 연령집단내에서"라는 단서조건이다. 그러므로 본 연구의 문항분석결과는 본 연구에 피검자로서 참여한 대학생들과 같은 나이 또래인 모든 "20대 초반의 성인들"에게 한정된 결과로서, 그 나이의 사람들에게 SPM은 변별력은 있으나 매우 쉬운 검사이며, 그들 중 지적 능력수준이 매우 낮은 사람들(예를 들면, 정신지체나 두뇌손상환자들)에 대해서 가장 측정의 오차없이 SPM이 측정하고자 하는 능력을 정확하게 측정할 수 있다는 사실을 보여주고 있다. 따라서, SPM을 노인집단이나 아동집단과 같은 다른 연령집단에 실시하였을 때 문항의 난이도, 변별도, 추측도가 어떻게 달라지고, 그 나이

집단에서는 SPM이 어떤 수준의 능력을 지닌 사람들에게 적절한 검사가 될 지는 아직 미지수이다. 앞으로 다양한 연령층의 정상인 및 환자들에 대한 연구가 수행되고 그 자료가 종합된다면 SPM의 특성이 보다 분명히 밝혀질 것이며, 그를 바탕으로 하여 검사 문항을 재구성함으로써 보다 우수한 "한국판 SPM"을 만들 수 있을 것이다. 끝으로, 본 연구는 지필 검사로 실시되었으나 병원장면에서는 Vienna Test System을 이용하여 전산화된 검사로 실시되는 경우가 많으므로 지필검사로 실시할 때와 전산화된 검사를 실시할 때의 피검자의 수행차이에 관한 비교 연구가 수행되어야 할 것이다.

참고문헌

- 성태제(1991, 편역). *문항반응이론입문*. 서울: 양서원.
- 이철, 유희정, 김창윤, 한오수, 박인호 (1993). 비엔나 검사총집의 임상적용을 위한 한국 표준화 연구(I): 서울 학생군에 대한 Standard Progressive Matrices 검사. *신경정신의학*, 32(2), 252-258.
- 전용신, 서봉연, 이창우 (1963). KWIS 실시요강. 서울: 중앙교육연구원.
- Banks, C., & Sinha, U. (1951). An item analysis of Progressive Matrices test and Binet. *British Journal of Psychology, Statistical Section*, 4, 91-94.
- Bartlett, D. P., Newbrough, J. R., & Tulkin, S. R. (1972). Raven Progressive Matrices: An item and set analysis of subjects grouped by race, sex, and social class. *Journal of Consulting and Clinical Psychology*, 38(1), 154.
- Burg, B., & Belmont, I. (1990). Mental abilities of children from different cultural backgrounds in Israel. *Journal of Cross-Cultural Psychology*, 21, 90-108.
- Burke, H. R. (1985). Raven's Progressive Matrices(1938): More on norms, reliability and validity. *Journal of Clinical Psychology*, 41(2), 231-235.
- Carlson, J. S. & Jensen, C. M. (1981). Reliability of the Raven Colored Progressive Matrices Test: Age and ethnic group comparisons. *Journal of Consulting Psychology*, 49(3), 320-322.
- Court, J. H. (1983). Sex differences in performance on Raven's Progressive Matrices: A review. *The Alberta Journal of Educational Research*, 29(1), 54-74.
- David, R. & Skilbeck, C. (1984). Raven IQ and language recovery following stroke. *Journal of Clinical Neuropsychology*, 6(3), 302-308.
- Desia, M. (1955). The relationship of the Wechsler-Bellvue verbal scale and the Progressive Matrices. *Journal of Consulting Psychology*, 19, 60.
- Dolke, A. M. (1976). Investigation into certain psychometric properties of Raven's Standard Progressive Matrices Test. *Indian Journal of Psychology*, 51(3), 225-236.
- Eisenberg, B. S., Levy, B. I., & Kirsch, A. A. (1970). Some psychometric properties of the Raven Progressive Matrices test. *Experimental Publication System*, 7, 230-232.
- Fabian, M., & Parsons, O. (1983). Differential improvement of cognitive function in recovering alcoholic women. *Journal of Abnormal Psychology* 82(1), 87-95.
- Gainotti, G., Miceli, G., & Caltagirone, C. (1979). The meaning of poor performance obtained on Raven's Coloured Matrices by brain-damaged patients. *Italian Journal of Psychology*, 6(2), 135-145.
- Glasser, A. J., & Zimmerman, I. L. (1967). *Clinical interpretations of the Wechsler*

- Intelligence Scale for Children*. New York: Grune & Stratton.
- Gorsuch, R. L. (1974). *Factor analysis*. Philadelphia, PA: Saunders.
- Hall, J. C. (1957). Correlation of a modified form of Raven's Progressive Matrices(1938) with the Wechsler Adult Intelligence Scale. *Journal of Consulting Psychology*, 21(1), 23-26.
- Irving, G., Robinson, R., & Mcadam, A. (1970). The validity of some cognitive tests in the diagnosis of dementia. *British Journal of Psychiatry*, 117, 149-156.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn and Bacon, Inc.
- Keir, G. (1949). The Progressive Matrices as to school children. *British Journal of Psychology, Statistical Section 2*, 140-150.
- Kertesz, A., & Hooper, P. (1982). Praxis and language: The extent and variety of aphasia. *Neuropsychologia*, 20(3), 275-286.
- McKenna, F. P. (1984). Measures of field dependency: Cognitive style or cognitive ability? *Journal of Personality and Social Psychology*, 47, 593-603.
- McLaurin, W. A., & Farrar, W. E. (1973). Validities of the Progressive Matrices Tests against IQ and grade point average. *Psychological Reports*, 32, 803-806.
- Milberg, W. P., Hebben, N., & Kaplan, E. (1986). The Boston process approach to neuropsychological assessment. In I. Grant & K. Adams(Eds.), *Neuropsychological assessment of neuropsychiatric disorders*. New York: Oxford University Press.
- Moran, A. P. (1986). The reliability and validity of Raven's Standard Progressive Matrices for Irish apprentices. *International Review of Applied Psychology*, 35, 533-538.
- O'Leary, U. M., Rusch, K. M., & Guastello, S. J. (1991). Estimating age stratified WAIS-R IQs from scores on the Raven's Standard Progressive Matrices. *Journal of Clinical Psychology*, 47(2), 277-284.
- Orme, J. E. (1966). Hypothetically true norms for the Progressive Matrices Tests. *Journal of Human Development*, 9, 223-230.
- Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Difference*, 13(2), 149-159.
- Persaud, G. (1987). Sex and age differences on the Raven's Matrices. *Perceptual and Motor Skills*, 67, 45-46.
- Powers, S., Barkan, J. H., & Jones, R. B. (1986). Reliability of the Standard Progressive Matrices Test for Hispanic and Anglo-American children. *Perceptual and Motor Skills*, 62, 348-350.
- Raven, J. C., Court, J. H., & Raven, J. (1977). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. London: H.K. Lewis.
- Raven, J. C., Court, J. H., & Raven, J. (1988). *Raven manual section 3. Standard Progressive Matrices: 1988 edition*. LA: Oxford Psychologists Press.
- Raven J., C. Court, J. H., & Raven, J. (1990) : *Raven manual research supplement 3. American and international norms: Neuropsychological*. LA: Oxford Psychologists Press.
- Raven, J. C., Court, J. H., & Raven, J. (1991). *Raven manual section 1. General overview:*

- 1991 edition. LA: Oxford Psychologists Press.
- Shaw, D. J. (1967). Estimating WAIS IQ from Progressive Matrices scores. *Journal of Clinical Psychology, 23*, 184- 185.
- Slater, P. (1948). Comments on "The comparative assessment of intellectual ability". *British Journal of Psychology, 39*, 20-21.
- Spearman, C. (1927). *The abilities of man*. New York: MacMillan.
- Spearman, C., & Jones, L. L. W. (1950). *Human ability*. London: MacMillan.
- Sternberg, R. J. (1994). Intelligence. In V. S. Ramachandran(Ed.), *Encyclopedia of human behavior* (pp.662-673). Vol.23. New York: Academic Press.
- Templer, D. I. (1992). Prison norms for Raven's Standard Progressive Matrices. *Perceptual and Motor Skills, 74*, 1193-1194.
- Urmer, A. H., Morris, A. B., & Wenaland, L. V. (1960). The effect of brain damage on Raven's Progressive Matrices. *Journal of Clinical Psychology, 16*, 182-185.
- Vincent, K. R., & Cox, J. A. (1974). A re-evaluation of Raven's Standard Progressive Matrices. *Journal of Psychology, 88*, 299-3

**A validity study
on the Raven Progressive Matrices Test
as an intelligence test**

Deug-Im Jeun and Yeonwook Kang

Chungbuk National University

The present study was conducted to explore the reliability and the validity of the Raven Progressive Matrices Test as an intelligence test in Koreans. The booklet form of the Standard Progressive Matrices (SPM) was administered to 405 college students (202 males, 203 females). Among them, 100 volunteers were also tested with the KWIS. The mean of the SPM scores was 52.5(s.d.=4.0). The sex difference was not found. The estimate of the test-retest reliability was .62 and the internal consistency (Cronbach's alpha) .73. The construct validity of the SPM was investigated by factor-analyzing the 5 set scores. A single factor that explained 39.2% of the variance was extracted from the 5 sets, indicating the high internal consistency of the SPM. The fact that there were significant correlations between the SPM and KWIS's Full, Verbal, and Performance Scale IQs indicates that the SPM is a valid measure of intelligence. That the Vocabulary subtest having the highest correlation with Verbal Scale IQ was barely correlated with the SPM, however, suggests that both the SPM and the vocabulary scale should be considered together for assessing general intelligence(g). According to the item analysis based on the item response theory, it was found that the SPM consists of too many easy items and the average "guessing" was in the slightly high range, while the average "item discrimination" was in the moderate to high range. In sum, the present results indicate that the SPM is a reliable and valid measure of intelligence, although these results should be validated by the future studies that will be conducted to other aged normal and patient groups.

부록 1. 문항반응이론에 의한 SPM의 문항분석결과

문항	난이도	변별도	추측도	문항	난이도	변별도	추측도
A-1	-6.758 2.851*	.434 .186*	.149 .080*	B-10	-4.096 .955*	.632 .175*	.147 .079*
A-3	-5.238 2.199*	.640 .352*	.149 .080*	B-11	-2.649 .612*	.445 .107*	.148 .079*
A-4	-6.796 2.696*	.403 .160*	.149 .080*	B-12	-1.504 .327*	.557 .112*	.141 .075*
A-5	-7.413 3.561*	.490 .236*	.149 .080*	C-1	-6.514 2.402*	.453 .170*	.149 .080*
A-7	-8.325 3.900*	.377 .161*	.149 .080*	C-2	-6.849 2.928*	.378 .162*	.149 .080*
A-8	-4.543 1.344*	.495 .165*	.149 .080*	C-3	-7.110 3.225*	.448 .203*	.149 .080*
A-9	-4.846 1.423*	.708 .276*	.148 .079*	C-4	-4.641 1.349*	.333 .099*	.150 .080*
A-10	-4.681 1.574*	.512 .198*	.149 .080*	C-5	-4.807 1.699*	.517 .212*	.149 .080*
A-11	-4.374 1.223*	.391 .116*	.150 .080*	C-6	-2.625 .496*	.670 .160*	.147 .078*
A-12	-2.100 .578*	.335 .078*	.152 .081*	C-7	-4.380 1.299*	.579 .209*	.149 .080*
B-1	-4.965 1.742*	.792 .511*	.149 .080*	C-8	-1.650 .396*	.463 .097*	.148 .079*
B-2	-3.995 1.077*	.624 .211*	.149 .079*	C-9	-3.515 .917*	.366 .099*	.151 .080*
B-3	-5.790 1.887*	.567 .205*	.149 .080*	C-10	-1.584 .345*	.562 .117*	.156 .082*
B-4	-5.220 1.766*	.469 .173*	.149 .080*	C-11	-1.980 .468*	.442 .099*	.149 .079*
B-5	-3.236 .671*	.775 .223*	.146 .078*	C-12	-.181 .182*	.729 .149*	.125 .063*
B-6	-4.049 1.007*	.382 .097*	.148 .079*	D-1	-7.888 3.554*	.367 .156*	.149 .080*
B-7	-3.059 .773*	.403 .108*	.152 .081*	D-2	-4.645 1.317*	.429 .131*	.150 .080*

문항	난이도	변별도	추측도	문항	난이도	변별도	추측도
B-8	-3.000 .751*	.361 .089*	.149 .079*	D-3	-3.708 1.125*	.578 .223*	.151 .080*
B-9	-3.565 .918*	.356 .092*	.149 .079*	D-4	-4.269 1.275*	.473 .159*	.149 .080*
D-5	-4.771 1.421*	.521 .177*	.149 .080*	E-3	-2.473 .407*	.833 .191*	.143 .077*
D-6	-4.414 1.276*	.413 .129*	.150 .080*	E-4	-2.087 .330*	.840 .185*	.156 .082*
D-7	-2.906 .730*	.368 .092*	.151 .081*	E-5	-2.303 .351*	.018 .275*	.156 .082*
D-8	-1.934 .462*	.424 .091*	.143 .077*	E-6	-1.547 .333*	.596 .128*	.151 .080*
D-9	-1.816 .360*	.583 .120*	.148 .079*	E-7	-.940 .254*	.686 .142*	.160 .082*
D-10	-3.153 .712*	.422 .100*	.151 .081*	E-8	-1.356 .293*	.626 .128*	.145 .077*
D-11	.707 .257*	.526 .125*	.131 .064*	E-9	-1.153 .377*	.407 .083*	.148 .078*
D-12	2.092 .452*	.511 .164*	.112 .050*	E-10	.461 .166*	.797 .179*	.114 .056*
E-1	-4.584 1.244*	.435 .127*	.149 .080*	E-11	1.310 .156*	.263 .447*	.088 .029*
E-2	-2.909 .508*	.844 .220*	.149 .079*	E-12	1.269 .173*	.051 .341*	.106 .036*