

Raven Colored Progressive Matrices 검사의 표준화

강 연 옥 김 정 기

삼성서울병원 포항공과대학교

본 연구는 Raven Colored Progressive Matrices(CPM) 검사의 표준화연구로서 성별과 연령 및 학력의 인구통계학적 분포에 맞추어 선정된 20세에서 79세까지의 도시와 농촌에 거주하는 357명(남자 181명, 여자 176명)의 피검자에게 CPM을 실시하여 그 규준을 구하였다. 거주지역과 성별은 CPM의 수행에 영향을 미치지 않았으나 연령과 학력은 유의미한 효과를 나타내었고, CPM의 신뢰도는 내적일치도가 .90이었다. CPM의 구성타당도를 검토하기 위하여 전체 36개의 문항을 변인으로 하여 요인분석을 한 결과 CPM이 원래의 제작의도와 일치하는 3요인으로 구성되어 있음이 확인되었고, 지능검사로서의 CPM의 공인타당도를 검증하기 위하여 K-WAIS의 어휘문제와 CPM의 상관관계를 조사한 결과 CPM은 어휘력과 .65($p < .001$)의 유의미한 상관관계를 지니고 있음이 밝혀졌다.

Raven Progressive Matrices Test (RPM; Raven, Court, & Raven, 1977)은 지각적 능력과 시공간적 추론능력에 입각한 비언어성 지능검사로써 인정되어 전 세계적으로 널리 사용되어 왔다(Moran, 1986). RPM은 특히 학교에서 습득된 지식과 언어능력에 의해서 영향을 받는 기존의 지능검사들과는 달리 문화적, 교육적 배경이 영향을 주지 않도록 만들어진 검사로서, 지각적 관계를 파악하고 유추에 의해서 추론하는 능력을 측정하도록 고안된 검사이다. 이 검사는 특히 시간제한이 없으므로 신체적 장애를 지닌 환자들이나

뇌손상환자들 또는 정서적 장애를 지니고 있거나 검사에 대한 동기가 부족한 사람 등과 같이 검사장면에서 빨리 반응하기 어려운 사람들에게 매우 유용한 검사로 밝혀지고 있다(Jensen, 1980).

RPM은 피검자의 나이와 지적 능력에 따라서 선택적으로 사용되는 Standard Progressive Matrices Test (SPM), Colored Progressive Matrices Test(CPM)와 Advanced Progressive Matrices Test(APM)의 3종류의 검사로 구성되어 있다. SPM은 모든 연령의 사람들에게 사용될 수 있도록 고안된 검사로서 각각 12문항씩

을 포함한 5개의 소검사로 구성되어 있고, CPM은 어린이나 노인 그리고 지적 장애가 있는 사람들에게 적합하도록 고안되었으며, APM은 우수한 지능을 가진 것으로 추정되는 피검자들에게 사용될 목적으로 개발된 검사이다. 일단 유추에 의한 추론이 가능할 정도로 지적 능력이 발달하였으면 SPM을 사용하는 것이 더 적절하지만 지적 발달이 이런 단계에 도달하지 못하였거나 이미 발달되었던 지적인 능력이 손상되었을 때는 그 지적 능력의 쇠퇴정도를 측정하기 위하여 CPM을 사용하는 것이 더 타당하다는 주장이 저자들(Raven et al., 1977)에 의해서 제기되었고, CPM은 현재 어린 아동과 노인들, 정신지체자들이나 뇌손상환자들에게 임상장면에서 널리 사용되고 있다.

CPM은 SPM과는 달리 그림의 배경에 밝은 색깔을 넣어서 피검자의 주의를 끌고 검사에 대해서 피검자가 자발적으로 흥미를 갖도록 유도하였으며, 검사를 실시할 때 필요이상의 언어적 지시를 피할 수 있도록 고안되었다. CPM은 먼저 개발된 SPM의 소검사에서 가장 쉬운 두개의 소검사(A와 B)와 소검사 A와 B의 중간정도의 난이도를 유지하도록 새로이 제작된 소검사 Ab로 구성되어 있다. 소검사 A는 “단순하고 연속적이거나 점진적인 변화를 보이는 패턴을 완성하는 과제들”로 구성되어 있고, 소검사 Ab는 “비연속적인 패턴을 완성하는 과제들”로, 소검사 B는 “비연속적인 패턴을 완성하는 과제”와 “공간적인 유추에 의한 구체적 추론이나 논리적 유추에 의한 추상적 추론을 요구하는 과제”들로 구성되어 있다(Raven, Court, & Raven, 1990b). 각 소검사에 12문항씩, CPM은 총 36문항으로 이루어져 있으며 각 문항은 12세 아동 수준의 지적 발달에 도달한 사람이면 모두 답할 수 있는 수준으로 만들어졌다(Raven et al., 1977).

신경심리학영역의 최근 성장과 함께 CPM은 지능 검사로서 뿐만 아니라 신경심리학적 평가도구로서도 그 유용성이 입증되고 있다(Milberg, Hebben, & Kaplan, 1986). 두뇌손상환자들은 일반적으로 병전(病前) 수준의 지적 능력을 유지하지 못하므로 문제해결에 효과적으로 접근하지 못하고 목표-지향적 행동이 체계적이지 못하며 손상의 부위와 정도에 따라서 다

양한 인지적 능력들이 손상된다. 두뇌손상환자들의 이러한 지적 능력을 측정하는 다양한 지능검사들이 개발되었으나 적당한 수준의 시력을 지니고 간단한 지시만 이해하면 수행이 가능한 CPM은 언어장애나 운동장애를 지닌 뇌손상환자들의 지능측정을 위해서 기존의 지능검사들보다 유용하게 쓰이고 있다. 더우기, CPM은 두뇌기능의 변화에 매우 예민하기 때문에 두뇌손상의 여부를 변별하는 선별도구(screening device)로 유용하게 사용될 수 있음이 선행연구에서 보고된 바 있다(Gainnotti, Miceli, & Caltagirine, 1979; Irving, Robinson, & McAdam, 1970).

두뇌손상환자들이 보이는 CPM의 반응양상에 대한 연구들(e.g., Gainnotti et al., 1979; Urmer, Morris, & Wendland, 1960)은 두뇌손상집단이 CPM수행시에 양적으로 뿐만 아니라 질적으로도 통제집단과 다른 수행을 보였음을 보고하였다. 양적인 측면에서 두뇌손상환자들은 WAIS의 소검사에서 드러난 어떤 지적인 감퇴보다도 더 저하된 CPM의 수행수준을 보였으며, 특히 정신지체 아동과 노인성 치매환자에게서 가장 두드러지게 저하된 점수가 관찰되었다(Raven, Court, & Raven, 1991). 알콜중독의 경우는 만성화될수록 CPM의 수행점수가 저조하였고 금주를 하면 그 점수가 증진되었다(Fabian & Parson, 1983). 실행증환자(Kertesz & Hooper, 1982) 뿐만 아니라 실어증환자도 CPM에서 낮은 점수를 얻었고 언어치료를 받은 후 그 점수가 증진되는 경향이 있었다(David & Skilbeck, 1984). 이는 CPM이 동작성 요인뿐만 아니라 언어성 요인도 포함하고 있다는 McLaurin과 Farrar (1973)의 주장을 뒷받침하는 결과이며 CPM이 이런 환자들의 치료과정중 그 경과에 대한 가치있는 정보를 제공할 수 있음을 시사한다(Raven, Court, & Raven, 1990a). 두뇌손상환자들이 CPM 검사중에 나타내는 질적인 반응특성으로는 전경-배경(figure-ground)의 반전오류와 상대적으로 어려운 소검사들은 잘 맞추면서 가장 단순한 소검사인 소검사 A에서는 오류를 보이는 것과 같은 문항난이도와 불일치하는 반응양상이 관찰되었다(Gainnotti et al., 1979).

이처럼 외국에서는 CPM이 일반적인 지적 기능의

평가뿐만 아니라 국소적인 두뇌손상으로 인한 인지기능의 평가를 포함한 여러 연구에서 사용되고 있고 임상적으로도 널리 활용되고 있는데도 불구하고 아직 한국에는 CPM을 포함한 RPM 전반에 대한 기초연구가 매우 부족한 실정이다. SPM의 경우에는 서울의 중고생과 대학생 및 20-30대 성인들을 대상으로 규준을 제시한 연구(이철, 유희정, 김창윤, 한오수와 박인호, 1993, 1996)와 대학생들을 대상으로 지능검사로서의 SPM의 신뢰도와 타당도를 검토한 연구(전덕임과 강연옥, 1994) 등 몇 편의 연구가 행해진 바 있으나 CPM의 경우에는 유치원아동부터 초등학교 6학년 아동까지의 규준을 제시한 연구(윤미라, 1996)가 한편 있을 뿐이다. 임상에서 두뇌손상환자나 노인들에게 RPM을 사용하고자 할 때 임상가들은 검사의 난이도와 검사에 소요되는 시간 등을 고려하여 SPM보다는 CPM을 선택하게 되지만 표준화된 규준이 없어서 환자의 수행수준에 대한 정확한 평가를 행하지 못하고 있다. 본 연구는 이런 필요에서 CPM의 신뢰도와 타당도를 검증하고 청년부터 노년까지의 전체 연령층에 대한 표준화된 CPM의 규준을 제공할 목적으로 수행되었다.

방법

1. 피검자

서울·경기지방과 청주를 중심으로 한 충청북도 및 포함을 중심으로 한 경상북도 지방에서 357명(남자 181명, 여자 176명; 서울 138명, 충북 104명, 경북 115명)의 피검자가 검사에 참여하였다. 표 1에 제시된 바와 같이 20세부터 79세까지를 10년씩 6범주로 나누고, 통계청(1993, 1994)의 사회지표와 인구주택총조사자료에 근거하여 각 연령범주에서의 남녀별 학력분포에 맞추어 남녀 피검자들을 선정하였다. 피검자들은 모두 신경과나 정신과적 질환의 병력이 없는 사람들이었고 이들의 평균연령은 남자가 49.9세(S.D.=17.6) 여자가 49.1세(S.D.=17.2)로 전체 평균연령은 49.5세(S.D.=17.4)였다.

2. 검사도구 및 절차

표준화연구중인 신경심리학검사총집의 일부로서 다른 신경심리학검사와 함께 booklet form의 CPM

표 1. 연령과 학력에 따른 전체 피검자의 분포

연령	성별	무학	국졸	중졸	고졸	대졸	합계
20 - 29	남			3	19	8	30
	녀			3	21	6	30
30 - 39	남		2	5	15	8	30
	녀		5	6	14	4	29
40 - 49	남	1	5	4	11	7	27
	녀		9	6	9	4	29
50 - 59	남	1	11	7	8	3	30
	녀	8	14	4	5	1	32
60 - 69	남	7	13	7	5	2	34
	녀	16	9	3	2		30
70 - 79	남	11	12	3	3	1	30
	녀	17	7	2			26
합계		61	87	53	112	44	357

검사(Raven et al., 1990b)와 K-WAIS(Korean-Wechsler Adult Intelligence Scale: 임상심리학회, 1992)의 어휘 문제를 개인검사로 실시하였다. CPM은 K-CVLT(Korean-California Verbal Learning Test; 김정기와 강연옥, 인쇄중)의 단기지연 단서회상단계와 장기지연 자유회상단계 사이의 20분 지연기간동안 실시되었고 어휘문제는 K-CVLT가 끝나고 난 직후에 실시되었다.

3. 자료분석

기본적인 기술통계와 요인분석 등의 자료분석을 위하여 SPSS-PC+를 사용하였고, 결과 분석에서 서울과 지방의 지역차이($t_{1355} = 1.54, ns$)나 남녀의 차이($t_{1355} = .84, ns$)가 관찰되지 않았으므로 이들 변인을 통합한 후 모든 분석을 행하였다.

결과

1. 신뢰도

CPM의 내적 일치도 계수 Cronbach's α 는 피검자 전체에 대하여 .93이었고 연령범주별로 살펴보면 .80-.93의 범위에 걸쳐 있었다. 이는 외국의 피검자들을 대상으로 한 Carlson과 Jensen(1981; .65-.85) 등의 선행연구결과와 일치하는 결과이다.

2. 점수의 분포

전체 피검자에 대한 CPM점수의 문항별 정답율(%)과 평균(표준편차) 및 소검사 A, Ab, B 각각의 평균점수(표준편차)는 표 2와 같다. 이를 피검자의 연령범주에 따라서 다시 나누어 제시한 것이 표 3이다. 소검사의 난이도를 비교해 본 결과 Raven 등(Raven et al., 1990b)이 CPM의 제작시에 의도한 대로 소검사 A, Ab, B의 순서로 점차 어려워진다는 것이 밝혀졌다(A vs. Ab; $t_{1355} = 5.29, p < .001$; Ab vs. B; $t_{1355} = 9.29, p < .001$).

표 2. CPM문항별 정답율(%)과 평균(표준편차)

문항	정답율 (%)	평균(표준편차)
A1	96.4	.964 (.188)
A2	98.0	.980 (.139)
A3	98.3	.983 (.129)
A4	96.9	.969 (.173)
A5	97.5	.975 (.157)
A6	94.7	.947 (.225)
A7	81.0	.810 (.393)
A8	81.8	.818 (.386)
A9	80.4	.804 (.398)
A10	73.9	.739 (.440)
A11	62.2	.622 (.486)
A12	42.6	.426 (.495)
Ab1	98.0	.980 (.139)
Ab2	94.1	.941 (.236)
Ab3	91.6	.916 (.278)
Ab4	82.9	.829 (.377)
Ab5	84.9	.851 (.356)
Ab6	72.8	.728 (.445)
Ab7	85.2	.852 (.356)
Ab8	70.9	.709 (.455)
Ab9	80.1	.801 (.400)
Ab10	76.5	.765 (.425)
Ab11	76.2	.762 (.427)
Ab12	41.7	.417 (.494)
B1	99.2	.992 (.091)
B2	93.0	.930 (.256)
B3	93.8	.938 (.241)
B4	91.3	.913 (.282)
B5	86.3	.863 (.345)
B6	82.1	.821 (.384)
B7	70.9	.709 (.455)
B8	46.5	.465 (.499)
B9	55.2	.552 (.498)
B10	61.3	.613 (.488)
B11	45.4	.454 (.499)
B12	37.3	.381 (.503)
전체		28.062 (7.198)

또한 CPM의 점수는 피검자의 연령에 따라서 유의미한 차이가 있음이 관찰되었고($F_{15, 331} = 9.23, p < .001$) 학력 또한 CPM점수에 유의미한 영향을 미치는 것으로

표 3. CPM 소검사의 연령별 평균점수(표준편차)

연령	CPM 소검사			전체
	A	Ab	B	
20 - 29	11.417 (.869)	11.267 (1.247)	10.733 (1.656)	33.350 (3.483)
30 - 39	11.034 (1.217)	11.017 (1.345)	9.983 (2.013)	32.034 (3.926)
40 - 49	10.768 (1.537)	10.473 (1.687)	10.000 (2.018)	30.875 (5.441)
50 - 59	10.081 (1.876)	9.677 (2.455)	8.710 (2.633)	27.952 (6.723)
60 - 69	8.672 (2.047)	7.875 (3.283)	6.688 (2.938)	23.250 (7.392)
70 - 79	8.286 (1.856)	7.018 (2.832)	5.714 (1.846)	21.018 (5.551)
전체	10.036 (1.999)	9.548 (2.772)	8.630 (2.889)	28.062 (7.198)

로 밝혀졌으므로($F_{(8, 330)} = 19.41, p < .001$) 각 연령과 학력별 CPM의 수행수준을 따로 표 4에 제시하였다. 각 연령별 피검자의 전체평균점수는 비록 각 연령범주에 속하는 피검자의 수가 충분히 크지는 않다 하더라도(27-38명) 그 연령범주에 해당하는 성별과 학력별 분포에 맞추어 수집한 피검자들의 평균점수이므로 대표성이 있다고 사료된다. 그러나 연령/학력별로 각 cell에 제시된 평균점수들중에는 그 cell에 해당하는 피검자의 수가 매우 적은 사례가 있으므로 후속 연구로 피검자의 수가 충원될 때까지 그러한 cell의 평균점수는 임상적 적용시에 매우 조심스럽게 사용되어야 할 것이다.

3. 요인분석

CPM의 구성타당도를 검토하기 위한 요인분석에 앞서서 3개 소검사 각각의 내용타당도를 검토하였다. 소검사 A, Ab, B의 문항들이 Raven(Raven et al., 1990b)이 밝혀듯이 연속적인 패턴의 완성, 비연속적인 패턴의 완성, 점진적인 변화를 보이는 패턴의 완성, 유추에 의한 구체적 추론, 유추에 의한 추상적 추론 과제 등으로 구성되어 있는지 소검사별로 각 문항들을 본 연구자들이 검토한 결과 각 소검사의 문항들이 그 제작의도대로 구성되어 있음을 확인하였다. 그러나 CPM은 SPM과는 달리 각 소검사들이 하나의 일관적인 내용들로 구성된 것이 아니라 한 소

검사내에 여러 종류의 과제들이 복합적으로 들어 있고 특히 소검사 B의 처음 5문항은 소검사 Ab의 과제와 같은 종류의 과제들로 구성되어 있었다. 따라서 각 소검사의 총점이 아니라 각 소검사의 문항들 하나하나의 점수를 변인으로 하고 공통요인분석(common factor analysis)의 주축분해법(principal axis factoring)을 사용하여 요인분석을 실시하였다. 고유치(eigenvalue)와 scree도표를 검토한 결과 고유치가 2.0 이상인 3개의 요인을 추출하는 것이 타당하다는 결론에 이르렀는데 이 3개의 요인은 총분산의 42.2%를 설명하였다. 사각회전후의 3개 요인의 최종고유치값은 각각 9.86, 2.09, 1.44였고 각 요인은 총분산의 27.4%, 5.8%, 4.0%를 설명하였으며 따라서 총분산의 37.2%가 이 세 요인에 의해서 설명된다는 결과를 얻었다. 사각회전후 각 요인에 속하는 변인과 요인부하량은 표 5와 같다. 요인부하량이 0.4이상인 변인만을 유의미한 변인으로 고려하는 Gorsuch(1974)의 기준을 적용하여 살펴보면 Factor 1에 속한 문항들은 대부분 소검사 Ab의 문항들로서(9문항) “비연속적 패턴의 완성(discrete pattern-completion)”이 요구되는 문항들이다. Factor 2는 소검사 B의 문항들과 소검사 A와 Ab의 가장 어려운 문항(A12, Ab12)을 포함하며 “유추에 의한 구체적이거나 추상적인 추론(concrete or abstract reasoning by analogy)”을 필요로 하는 문항들로 구성되어 있고, Factor 3은 소검사 A의 문항들과 소검사 Ab와 B의 가장 쉬운 첫 번째 문항(Ab1,

표 4. 연령과 학력별 CPM검사의 평균점수(표준편차)

연령	무학	국졸	중졸	고졸	대졸	전체
20 - 29			25.667 (6.088) n=6	34.200 (1.539) n=40	34.214 (1.847) n=14	33.350 (3.483) n=60
30 - 39		28.571 (5.028) n=7	30.818 (3.250) n=11	32.379 (4.039) n=29	34.333 (1.073) n=12	32.034 (3.926) n=59
40 - 49	21.000 (0.000) n=1	27.467 (5.705) n=15	31.300 (3.802) n=10	31.450 (6.126) n=20	34.091 (1.814) n=11	30.875 (5.441) n=56
50 - 59	24.375 (7.927) n=8	26.423 (5.873) n=26	28.364 (7.852) n=11	31.154 (5.742) n=13	33.500 (2.646) n=4	27.952 (6.723) n=62
60 - 69	19.591 (7.042) n=22	23.870 (7.783) n=23	23.400 (4.789) n=10	30.571 (2.637) n=7	30.000 (7.071) n=2	23.250 (7.392) n=64
70 - 79	18.667 (5.505) n=27	23.200 (4.652) n=20	20.400 (3.715) n=5	26.333 (5.508) n=3	28.000 (0.000) n=1	21.018 (5.551) n=56
전체	19.967 (6.488) n=61	25.563 (6.326) n=87	27.434 (6.222) n=53	32.446 (4.310) n=112	33.818 (2.315) n=44	28.062 (7.198) N=357

B1)을 포함하며 "연속적인 패턴의 완성(continuous pattern-completion)"을 요구하는 문항들로 구성되어 있다.

4. CPM과 어휘검사의 상관관계

지능의 한 측정치로서의 CPM의 타당도를 검토하기 위하여 인지기능검사들중에서 전통적으로 전체 IQ와 가장 높은 상관관계를 지니고 있는 것으로 보고되었으며 언어성 IQ를 대표한다고 인정되고 있는 어휘능력과 CPM점수와의 상관관계를 산출하였다. 어휘검사로서는 K-WAIS의 어휘문제 소검사를 사용하였고, 64세 이상의 노인들의 경우에는 표준화된 환산점수가 K-WAIS의 규준에 제시되어

있지 않으므로 CPM점수와 어휘능력간의 상관관계(Pearson product moment correlation: r)는 어휘문제의 원점수를 사용하여 구하였다. 그 결과, CPM과 어휘문제간에는 매우 유의미한 상관관계가 있음이 밝혀졌다($r_{(357)} = .652$ $p < .001$).

논의

RPM의 가장 큰 잇점중의 하나로 자주 언급되는 것은 RPM이 문화적으로 공평한, 즉 문화적 배경이 검수수행에 큰 영향을 미치지 않는 검사라는 점이다. 이 점은 여러 국가들에서 표준화된 RPM 규준들이 문화권간에 차이가 없이 상당히 안정되어 있다는 사실

표 5. CPM검사의 요인구조

CPM 문항	Factor 1	Factor 2	Factor 3
Ab7	.78541	.05009	-.03382
B6	.68538	-.03200	-.22149
Ab9	.66236	-.08488	-.08247
Ab4	.66227	.04521	.05684
Ab6	.64422	-.22021	-.06436
B5	.62970	-.02834	-.00950
Ab3	.61942	.07619	-.00784
Ab5	.61195	.00438	-.05672
Ab8	.60508	-.23040	-.09618
A7	.56409	-.15395	.00977
A9	.56050	-.18988	.06550
B2	.55925	.14056	.14332
A10	.50683	-.27955	-.04689
Ab2	.48465	.11222	.31478
B3	.47382	.04284	.03609
A11	.43268	-.30516	.00807
A6	.42775	.00363	.14898
Ab10	.40956	-.19754	.05594
B4	.38926	-.06352	.12559
B7	.36876	-.29798	.00447
Ab11	.31465	-.23174	.09804
B9	.14067	-.71431	.06310
B8	.11406	-.69499	-.01201
B10	.21577	-.66317	.04816
B11	.15345	-.62848	.02204
Ab12	.09929	-.60112	-.00729
B12	.06890	-.57149	.00519
A12	.12697	-.45096	-.02667
A8	.19958	-.22911	.14184
A1	-.04437	-.12328	.00500
B1	.06816	.10800	.62148
A5	-.02620	-.09824	.61267
A4	.02573	-.02030	.52670
Ab1	.14137	.17486	.47873
A3	-.03918	-.04050	.43247
A2	-.07803	-.09415	.40515
설명변량(%)	27.4	5.8	4.0

(Raven, Court, & Raven, 1988)과 RPM과 다른 지능 검사들과의 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 규준이 필요하다고 주장하였다. 한편 Freeman(1984, 재인용)은 8세와 12세의 시골과 도시의 흑인 국민학교 아이들을 대상으로 CPM을 실시한 연구에서 시골아이들 보다 도시의 아이들의 CPM점수가 더 높은 것을 발견하였고 CPM을 수행하는데 필요한 인지양식이 시골집단보다 도시집단내에서 더 발달 된다는 결론을 내렸다. 즉 다양한 외적 자극과 문화적 영향이 CPM의 수행과 상관성이 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 연구에서는 한국인들의 경우 대도시와 지방(시골)간에 CPM수행수준에 있어 큰 차이가 없으며 남녀의 수행수준도 역시 차이가 없으나 연령과 학력이 CPM의 수행수준에 큰 영향을 미치는 중요한 변인으로 밝혀졌다.

CPM이 무엇을 측정하는 검사인지 그 구성타당도에 관하여 연구자들은 서로 다른 연구결과를 보고하였다. Raven은 RPM을 어휘능력과 종합하지 않고 단독으로 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 주장하

였다. 이들의 주장은 RPM에서 단일요인을 발견한 요인분석연구(Orme, 1966)에 의해서 뒷받침되었다. 그러나 다른 연구자들(Banks & Shiha, 1951; Dolke, 1976)은 요인분석결과 RPM에서 귀납적 추론, 지각 속도, 공간능력, 지각적 정확성, 개념형성 등으로 이 름붙일 수 있는 중다요인을 발견하였고, RPM점수를 일반적 지능(g)의 측정치로 사용하는 것에 반대하였다. 본 연구에서는 요인분석결과 CPM에서 3요인을 발견하였고 Factor 1은 비연속적인 패턴의 완성, Factor 2는 유추에 의한 구체적이거나 추상적인 추론, Factor 3은 연속적인 패턴의 완성 등을 요구하는 문항들로 구성되었음을 확인하였다. 요인분석결과 밝혀진 이와 같은 CPM의 구조는 원래의 검사제작 의도와 일치하고 있으므로 CPM의 구성타당도를 입증하는 결과로 사료되나 CPM이 지능의 한 측정치로서 얼마나 타당하고 효율적인가의 문제는 이 요인분석의 결과로는 판단할 수 없는 문제이다.

지능의 한 측정치로서 RPM의 효율성에 대한 연구는 기존 지능검사들과 RPM과의 상관관계를 연구하는 공인타당도(concurrent validity)의 연구가 주종을 이루었다. 특히 WAIS와의 상관연구가 널리 행해졌는데 SPM의 점수가 WAIS-R의 전체 IQ점수와 매우 유의미한 정적상관관계(.75-.85)가 있음이 선행연구들(Burke, 1985; Hall, 1957; Shaw, 1967)에서 보고되었으며 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 언어성척도와 SPM과의 상관관계를 연구한 결과 두 검사사이의 상관관계가 높지 않으며, 특히 상식문제와 어휘문제와의 상관관계가 가장 낮았다고 보고하였다. 그는 이 결과를 SPM이 현재의 지적 효율성을 재는 척도인데 반해서 Wechsler-Bellevue의 언어성척도는 과거에 습득한 지식을 통하여 지능을 측정하기 때문인 것으로 해석하였다. 그러나 McLaurin과 Farrar(1973)는 SPM이 WAIS 전체 IQ와 동작성 IQ 뿐만 아니라 언어성 IQ와도 유의한 상관관계가 있다는 연구 결과를 보고한 바 있다. 전덕임과 강연욱(1994)은

SPM과 KWIS를 한국의 대학생들에게 실시하여 그 상관관계를 조사한 결과 SPM이 전체 IQ, 동작성 IQ 및 언어성 IQ와 유의한 상관관계가 있으나, 언어성 IQ와 가장 상관관계가 높은 어휘문제가 SPM과는 거의 상관관계가 없음을 발견하였다. 이는 지능검사로서의 SPM의 타당도를 입증하는 결과이나 SPM을 이용하여 전체지능을 추정하려 한다면 어휘능력이 SPM과 반드시 함께 고려되어야 한다는 점을 시사하는 결과로 해석된다.

Moran(1972)은 Stanford Binet검사와 CPM사이에서 .78의 상관관계를 발견하였고, McLeod와 Rubin(1962)은 정신과환자들에게 WAIS와 CPM을 실시하여 전체 IQ와는 .72, 동작성 IQ와는 .70, 언어성 IQ와는 .58의 상관관계를 발견하였다. 한국의 어린이들을 대상으로한 윤미라(1996)의 연구에 의하면 유치원과 국민학교 어린이들에게 CPM과 KEDI-WISC를 실시하고 그 상관관계를 분석한 결과, CPM은 숫자따라하기를 제외한 모든 KEDI-WISC의 소검사들과 .47-.67의 유의미한 상관관계를 지니고 있고, 언어성 IQ, 동작성 IQ 및 전체 IQ와도 .43-.49의 유의미한 상관관계를 지니고 있음을 발견하였다. 특히 토막짜기 소검사와 CPM과의 가장 높은 상관관계수(.67)는 CPM검사의 심리측정적 특성이 공간지각적 능력과 비언어적이고 추상적인 사고능력과 관계가 깊다는 선행연구의 결과를 뒷받침하는 결과로 사료된다. 본 연구에서는 대학생을 대상으로 한 전덕임과 강연욱(1994)의 SPM 연구결과와는 달리 CPM점수와 K-WAIS의 어휘문제 점수간에 매우 유의미한 상관관계가 있음이 발견되었다. 이런 결과는 CPM의 문항들이 SPM의 문항들보다 너무 쉽고 단순하기 때문에 SPM에서 요구되는 추상적이고 공간적인 비언어적 추론과정이 CPM수행시에는 SPM 수행시만큼 많이 필요로 되지 않아서 비롯된 결과일 것으로 추측되나 이에 대한 확실한 규명을 위해서는 CPM과제에 대한 좀더 자세한 분석과 지능검사에 포함되어 있는 다른 소검사들과 CPM과의 상관관계에 대한 연구가 먼저 수행되어야 할 것으로 사료된다. 그러나 본 연구에서 밝혀진 것처럼 정상인에게 있어 CPM의 점수와 어휘력이 정적인 상관

관계에 있다면, 과거에 습득한 지식을 재생산해내는 어휘력은 뇌손상후에도 잘 손상되지 않으나 지적 활동의 현재상태를 측정하는 CPM의 수행수준은 뇌손상의 유무에 따라서 민감하게 영향을 받는다고 알려져 있으므로 두뇌손상이 의심되는 환자에게서 CPM 점수와 어휘력점수간에 큰 차이가 발견되었을 때 이를 두뇌손상을 의미하는 중요한 지표로 해석할 수 있을 것이다.

참고문헌

- 김정기, 강연옥 (인쇄중). 한국판 캘리포니아 언어학 습검사(K-CVLT) 검사요강.
- 윤미라 (1996). Colored Progressive Matrices 검사의 한국판개발을 위한 예비연구. 충남대학교 석사 학위 논문.
- 이 철, 유희정, 김창윤, 한오수, 박인호 (1993). 비엔나 검사총집의 임상적용을 위한 한국 표준화 연구(I): 서울학생군에 대한 Standard Progressive Matrices 검사. *신경정신의학*, 32(2), 252-258.
- 이 철, 유희정, 김창윤, 한오수, 박인호 (1996). 비엔나 검사총집의 임상적용을 위한 한국 표준화 연구(II): 서울 성인군에 대한 Standard Progressive Matrices 검사. *신경정신의학*, 35(1), 133-143.
- 임상심리학회(편) (1992). K-WAIS 실시요강. 서울: 한국가이던스.
- 전덕임, 강연옥 (1994). 지능검사로써의 Raven Progressive Matrices 검사의 타당도연구: Standard Progressive Matrices 검사를 중심으로. *한국심리학회지: 임상*, 13(1), 65-81.
- 통계청 (1993). 인구주택총조사보고서. 서울: 통계청.
- 통계청 (1994). 한국의 사회지표. 서울: 통계청.
- 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.
- 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.
- 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.
- 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., & McAadam, A. (1970). The validity of some cognitive tests in the diagnosis of demementia. *British Journal of Psychiatry*, 117, 149-156.
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- 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.
- MiLeod, H. N., & Rubin, J. (1962). Correlation between Raven Progressive Matrices and the WAIS. *Journal of Consulting Psychology*, 26 (2), 190-191.
- 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, R. E. (1972). Progressive Matrices and the educationally disadvantaged. *Mental Retardation*, 10, 9.
- 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.
- 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. (1990a) : *Raven manual research supplement 3. American and international norms: Neuropsychological*. LA: Oxford Psychologists Press.
- Raven, J. C., Court, J. H., & Raven, J. (1990b). *Raven manual: Section 2. Coloured Progressive Matrices: 1990 edition*. L.A.: Western Psychological Services.
- Raven, J. C., Court, J. H., & Raven, J. (1991). *Raven manual: Section 1. General overview: 1991 edition*. L.A.: Western Psychological Services.
- Shaw, D. J. (1967). Estimating WAIS IQ from Progressive Matrices scores. *Journal of Clinical Psychology*, 23, 184-185.
- Spearman, C., & Jones, L. L. W. (1950). *Human*

ability. London: Macmillan.

Templer, D. I. (1992). Prison norms for Raven's Standard Progressive Matrices. *Perceptual and Motor Skills*, 74, 1193-1194.

Urner, 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, 299-303.

A Normative Study of the Raven Colored Progressive Matrices Test

Yeonwook Kang

Jung K. Kim

Samsung Medical Center Pohang University of Science
and Technology

As a normative study of the Raven Colored Progressive Matrices (CPM), the booklet form of the CPM was administered to 357 healthy individuals (181 males and 176 females) who were selected by stratified sampling procedure reflecting the recent Korean census data in terms of major demographic variables, such as sex, age, education. Although there was no difference of the CPM scores between men and women and between urban and rural, it was found that age and education influence the CPM scores. Therefore, normative data were provided for age and education group separately. The internal consistency (Cronbach's alpha) of the CPM was .90. The construct validity of the CPM was investigated by factor-analyzing the 36 items, yielding 3 factors that were congruent with the original 3 subsets. A significant correlation ($r=.65$, $p<.001$) was found between the CPM and the Vocabulary score of the KWIS, indicating that the CPM is a valid measure of intelligence.