

## 인물화 지능검사의 타당도: 아동용 웨슬러 지능검사와의 비교

황순택<sup>†</sup> 전영순 노은정 조윤진 여상우

충북대학교 심리학과

본 연구에서는 인물화에 의한 간편지능검사의 타당도를 검증하기 위해 DAP에서 산출된 IQ와 K-WISC-III에서 산출된 IQ 간 상관을 분석하였다. 6세~12세 아동 182명(남아 113명, 여아 69명; 정상아동 63명, 임상아동 119명)이 자료수집에 참여하였다. 자료의 분석결과 DAP와 K-WISC-III로 산출된 두 지능지수 간에 전체적으로 유의미한 정적 상관이 확인되었다. 지능수준과 연령을 기준으로 집단을 세분화하여 상관분석을 한 결과 IQ 120 이상인 집단과 취학전 집단을 제외한 모든 하위 집단에서 유의미한 정적 상관이 확인되었다. 또한, 제작연대가 다른 두 검사에서 산출된 IQ 간 현저한 차이가 있는 점을 감안하여 DAP IQ로 K-WISC-III IQ를 추정할 때의 회귀방적식을 산출하였다.

주요어 : 인물화 지능검사, IQ, 타당도, 한국 웨슬러 아동용 지능검사

---

<sup>†</sup> 교신저자(Corresponding Author) : 황순택 / 충북대학교 사회과학대학 심리학과 / 충북 청주시 개신동 12  
Fax : 043-271-1713 / E-mail : hstpsy@chungbuk.ac.kr

인물화를 비롯한 그림검사는 심리평가 장면에서 두 가지 용도로 활용되어 왔다. 그 중 하나는 자신의 언어로 표현하기 어려운 느낌과 생각, 스스로 인식하지 못하는 성격과 정서의 영역을 평가할 때 이러한 언어와 인식의 한계를 보완해 주는 투사적 성격검사로서의 용도이다. 다른 하나는 인지기능 및 발달수준을 평가하는 용도이다.

그림을 통해 아동의 인지발달을 평가하고 인지능력을 측정하려는 노력은 오래 전부터 있었다. 1920년대에 Burt, Goodenough 등의 연구자는 아동의 그림을 읽기, 쓰기, 산수 등 특정한 학습기능 및 일반적인 지능수준과 비교한 연구, 인물화 그림의 특정한 측면들과 아동의 정신연령 간의 상관을 본 연구 등을 수행하면서 그림검사를 지능검사로 사용할 수 있는 가능성을 탐색하였다.

그림검사를 지능검사의 한 대안으로 생각하게 된 데에는 그림이 후천적으로 습득된 지식의 영향을 받지 않는다는 점에서 상당한 강점이 있고, 또한 무엇보다도 실시가 용이한 데서 그 원인을 찾아볼 수 있다. Goodenough (1926)는 아동의 그림이 성격이나 흥미와 같은 비 지적 요소보다는 지적인 요소와 더 밀접한 관계가 있다고 가정하고 지능 측정의 수단으로 인물화 검사(Draw-A-Man Test; DAM)를 개발한 바 있다. DAM은 그림의 상세함, 신체 부분들 간의 비율, 운동협응의 정도 등을 고려하여 아동의 인지적 및 발달적 측면을 검사할 수 있도록 고안된 것으로, 일종의 지능검사로 분류할 수 있다. 그 후 Goodenough와 Harris (1963)는 Goodenough의 DAM을 바탕으로 방법을 재평가하고, 연령의 범위를 확대하고, 동형 검사를 개발하는 등 검사의 체계를 확장시켜 Goodenough-Harris Drawing Test(G-H; Goodenough

& Harris, 1963)를 완성하였다.

Goodenough와 Harris 외에도 몇몇 학자들이 그림검사를 지능의 측정에 활용할 수 있는 방안을 고안하였다. Buck은 인물화 검사를 성인에게 적용할 수 있도록 확장시켰으며, 인물뿐만 아니라 집, 나무까지도 그림검사의 한 부분으로 포함하여 House-Tree-Person( H-T-P)의 채점체계를 개발하여(Buck, 1948, 1966) 지능측정 도구로 발전시켰다. Koppitz(1968)는 아동의 인물화에 대한 발달점수(Koppitz Developmental Items; KDI)를 개발하였다. KDI는 G-H를 보다 간편하고 쉽게 하기 위해 고안된 것이다.

인물화에 의한 지능평가의 타당도를 검증하기 위해 비네 지능검사 또는 웨슬러 지능검사에서 산출된 지능지수와 비교하는 연구가 다양하게 이루어졌다. Koppitz(1968)는 KDI와 WISC 간에 높은 상관을 얻어 지능평가 도구로서의 KDI의 타당도를 확인한 바 있다. Abell, Von Briesen, 그리고 Watz(1996)의 연구에서도 아동들이 그런 인물화에 대해 G-H, KDI 두 채점체계로 점수를 산출하고 그들의 WISC-R 및 비네 지능검사에서 산출된 지능지수와 비교하였다. 그 결과 두 체계 모두 WISC-R, 비네 지능검사 IQ 간에 높은 상관을 확인할 수 있었다. 이러한 연구결과들은 G-H, KDI와 같은 인물화를 이용한 지능평가 체계가 상당한 정도의 타당성이 있음을 말해주는 것이다.

인물화에 의한 지능평가는 간편성과 실용성의 장점을 가지고 있어 여러 가지 임상적 제한으로 비네, 웨슬러처럼 정교하고 종합적이지만 상대적으로 많은 시간이 소요되는 지능검사를 실시하기 어려운 상황에서 사용하기에 용이하다. G-H, KDI 등은 실시 시간이 매우 짧고 과제가 간단하여 신체적 또는 정신적 기

능이 매우 제한되어 있는 사람들이나 어린 아동들에게도 어렵지 않게 실시할 수 있는 장점이 있다. 이 때문에 일반집단 뿐 아니라 임상진단, 연령, 그리고 지능수준 등에 의해 구분되는 다양한 하위집단들을 대상으로 인물화 지능검사의 타당도 연구가 수행되었다.

임상집단 연구의 한 예로, Tramill, Edwards, 그리고 Tramill(1980)는 학습곤란을 경험하고 있는 학생들에게 DAP와 WISC-R을 실시하여 두 검사에서 산출된 지능지수를 비교한 결과 두 점수간 유의한 상관을 확인하였다. 이 결과는 학습곤란 집단에서도 G-H, KDI와 같은 인물화 지능검사가 타당하게 사용될 수 있음을 말해주는 것이다. Aikman, Belter, 그리고 Finch(1992)는 지능과 학업성취 간의 관계를 연구하는데 이 검사를 사용하기도 하였다.

Goodenough가 발전시킨 인물화 검사는 지능을 평가함에 있어 비네 그리고 웨슬러 지능검사에 견줄 정도로 염밀하지는 않지만, 지능검사로서의 가치를 인정받아 미국 뿐만 아니라 많은 나라에서 표준화되어 지능 측정 도구로 사용되어 왔다. 이 검사는 특히 나이가 어린 아동이나 신체적 또는 정신적 장해를 보이는 사람들에게 널리 사용되어 왔고, 지능, 성격, 적응문제, 사회병질, 정신지체 등의 다양한 진단영역에서 사용되어 왔다.

우리나라에서는 Goodenough-Harris Drawing Test(Goodenough, & Harris, 1963)를 한국 아동들에게 표준화시킨 ‘인물화에 의한 간편 지능검사’(김재은, 김동극, 여광웅, 1973)가 제작된 바 있다. 국내에서 이 검사는 일부 임상장면에서 사용되어 오기는 했지만 검사의 타당도에 관한 연구가 충분히 이루어지지 못하였다. 인물화 검사가 가지는 간편성과 실용성 등의 장점을 감안할 때 검사의 타당도를 확인하는

것은 이 검사의 활용을 촉진하는데 도움이 될 것이다.

본 연구는 한국판 웨슬러 아동용 지능검사를 타당도 준거로 하여 Goodenough와 Harris(1963)가 제작하고 김재은 등(1973)이 한국판으로 표준화한 인물화 지능검사의 타당도를 검증하기 위한 것이다. 일반적으로, 다른 지능검사를 외적 준거로 사용하여 특정한 지능검사의 타당도를 검증하기 위해서는 적어도 두 검사에서 산출된 지능지수 간의 상관이 통계적으로 유의미해야 하고, 두 지능지수의 차이가 통계적으로 유의미하지 않아야 한다(Resnick & Entin, 1971). Resnick과 Entin은 또한 두 검사에서 산출된 지능지수를 “경계선”, “보통상”, “최우수”와 같은 지능수준 범주로 분류했을 때도 분류 일치율이 높아야 한다고 제안하였으나 이는 임상적인 측면에서 의미가 있는 하지만 심리측정적 측면에서는 큰 의미를 두기 어려운 기준이다. 또한, 두 지능지수 간 상관이 유의미하고 차이가 유의미하지 않아야 된다는 두 가지 판단기준 중 두 번째 기준은 현재 시점에서 DAP의 타당도를 평가하는데 적용하기가 어렵다. 인물화 지능검사와 본 연구에서 타당도 준거로 사용할 웨슬러 아동용 지능검사(K-WISC-III)는 규준 제작 시점이 현저하게 다르기 때문이다.

검사가 제작된 후 많은 시간이 경과한 경우 그 검사에서 산출되는 IQ는 실제에 비해 과대 추정되는, 이를 바 Flynn 효과(Flynn, 1984, 1987)가 나타나기 때문에 제작시기가 다른 두 검사에서 각각 산출된 IQ의 차이에 대한 유의도 검증은 무의미하다. 이 때문에 본 연구에서는 첫 번째 기준만 적용하여 인물화 지능검사의 타당성을 살펴보고자 하였다. 실제로 지능측정 도구로서의 인물화 검사의 타

당도를 검증한 많은 연구에서 비네 또는 웨슬러 지능검사와의 상관만으로 검증하였으며, 두 지능지수 간의 차이검증을 한 경우는 찾아 보기 어렵다.

본 연구에서는 또한 두 검사 간 유의미한 상관이 확인되어 인물화 지능검사의 타당성이 시사되는 경우 인물화 지능검사에서의 수행으로 K-WISC-III IQ를 추정할 수 있는 회귀방식을 산출하고자 한다. 이는 인물화 지능검사가 제작된 후 30년 이상 지난 현재 시점에서 그 규준을 지금 그대로 사용하기는 어려운 만큼 인물화 지능검사 규준 상의 IQ를 현 시점에서 가장 최근에 제작된 아동용 지능검사인 K-WISC-III IQ로 환산하기 위한 것이다.

## 방 법

### 연구참여자

충청지역에 거주하는 만 6세~12세 아동 182명(남아 113명, 여아 69명)이 연구에 참여하였다(표 1). 이들은 청주와 대전 및 그 인근 지역에서 소아정신과 외래 진료를 받고 있는 119명의 임상 아동과 청주지역에서 유치원, 초등학교, 중학교에 다니는 63명의 정상 아동으로 구성되어 있다. 자료수집은 임상아동의 경우 전체 수집기간 동안 각 병원의 임상심리 전문가가 실시한 심리평가 자료에서 K-WISC-III와 인물화 검사 둘 다 실시된 사례를 수집

하였다. 정상아동의 경우 심리학과 대학원에서 심리평가와 임상실습 과목을 이수한 대학원생들이 실시하였다.

### 연구 도구

#### 인물화에 의한 간편지능검사

인물화에 의한 간편지능검사(Draw-a-Person; DAP)는 김재은, 김동극, 여광웅(1973)[Goodenough-Harris Drawing Test(Goodenough & Harris, 1963)]를 원본으로 하여 우리나라에 맞도록 개정한 것으로, 만 2세부터 12세 까지의 취학전 및 초등학교 아동, 정신지체아, 특수학교, 특수학급 아동에게 사용할 수 있도록 만들어진 지능검사이다. 이 검사는 남자상과 여자상의 두 척도로 이루어져 있으며, 아동이 그린 남자상과 여자상을 각각 60여개의 채점 기준에 따라 채점한 후 규준에 비추어 지능지수를 산출하게 된다. 김재은 등이 보고한 이 검사의 채점자간 신뢰도는  $r = .93\sim .97$ 로 매우 높다. 또한, 타당도를 확인하기 위해 지능성숙 검사(김난수, 1963)와 비교한 결과 두 검사에서 산출된 점수간 상관이  $r = .37\sim .42$  정도로 나타나 매우 높지는 않지만 외국의 선행 연구들과 비슷한 수준이었다. 본 연구에서 일부 자료(40명)를 대상으로 채점자간 일치도를 산출한 결과  $r = .91\sim .97$ 의 높은 신뢰도가 확인되었다.

남자상, 여자상 모두 각각 60점이 만점이며, 한 채점기준 당 1점씩 해당되는 채점기준의

표 1. 연구참여자의 집단별 연령과 성 분포

	전체집단(n=182)	정상집단(n=63)	임상집단(n=119)
성(남/여)	113 / 69	32 / 31	81 / 38
연령	8.92±1.81 세	9.11±1.78 세	8.82±1.82 세

개수를 원점수로 한다. 이렇게 얻어진 원점수를 규준과 대조하여 지능지수와 백분위 점수로 환산한다. 산출된 지능지수는 원점수를 연령별 평균과 표준편차로 환산한 편차지능지수이며, 규준은 그림을 그린 아동의 성별과 그려진 그림의 성별에 따라 4가지로 만들어져 있다. 또한 아동의 지능수준을 해석하는데 지역적인 특성을 고려하기 위하여 전국규준과 지역규준이 따로 제시되어 있다. 본 연구에서는 전국규준에 의거하여 지능지수를 산출하였으며, 남자그림과 여자그림에서 각각 산출되는 두 개의 지능지수를 평균하여 개인의 최종적인 지능지수로 사용하였다.

### 한국 웨슬러 아동용 지능검사

한국 웨슬러 아동용 지능검사(Korean-Wechsler Intelligence Scale for Children-III; K-WISC-III)는 미국의 WISC 개정 3판(WISC-III; Wechsler, 1991)의 한국판이다. 곽금주, 박혜원, 김청택(2001)이 표준화하였으며, 만 6세부터 16세까지의 아동의 지능을 측정하는 대표적인 개인용 검사도구로 사용되고 있다. 6개 언어성 소검사와 7개 동작성 소검사로 구성되어 있으며, 검사결과로 언어성 지능지수, 동작성 지능지수, 전체 지능지수 등 세 개의 지능지수가 산출된다. 산출된 지능지수는 원점수를 평균 100, 표준편차 15로 환산한 편차지능지수이다.

표 2. DAP와 K-WISC-III에서 산출된 지능지수 평균

	DAP IQ 평균	K-WISC-III IQ 평균
전체집단(n=182)	111.99±15.50	97.10±21.14
정상집단(n=63)	121.02±12.87	108.83±14.84
임상집단(n=119)	107.22±14.67	90.79±21.37

주. DAP = Draw-A-Person Test,

K-WISC-III = Korean-Wechsler Intelligence Scale for Children-III

### 결과

#### DAP IQ와 K-WISC-III IQ 간 상관

전체 연구참여자들의 DAP IQ 평균은 111.99, K-WISC-III IQ 평균은 97.10으로, DAP가 있는 K-WISC-III보다 15점정도 높게 IQ를 산출하였다(표 2). DAP IQ와 K-WISC-III IQ 간 상관분석 결과 유의미한 정적 상관이 있었다, 전체집단에서  $r$  (180)=.47,  $p<.01$ ; 정상집단에서  $r$  (63)=.30,  $p<.05$ ; 임상집단에서  $r$ (117)=.38,  $p<.01$ . 이 결과는 정상집단과 임상집단 모두에서 K-WISC-III에서 높은/낮은 지능지수를 얻는 아동은 DAP에서도 높은/낮은 지능지수를 얻는다는 것을 의미한다.

#### 지능수준별 집단에서의 DAP IQ와 K-WISC-III IQ 간 상관

지능수준에 따라 두 지능지수 간의 관계가 동일하게 유지되는지 확인하기 위해 수준별로 두 지능지수 간 상관을 산출하였다. 본 연구에 표집된 표본을 K-WISC-III IQ에 따라 1) 경계선과 정신지체 집단(IQ 79이하, 40명), 2) 평균하, 평균, 그리고 평균상 집단(IQ 80~119, 116명), 3) 우수, 최우수 집단(IQ 120이상, 26명)으로 구분하였다. 지능수준에 따라 분류한

세 집단에서 각각 DAP IQ와 K-WISC-III IQ 간 분석결과 IQ 79 이하인 집단과 IQ 80~119 인 집단의 경우 상관이 유의미하였으나 IQ 120 이상 집단에서는 유의미하지 않은 부적 상관을 보였다, IQ 79 이하 집단에서  $r(38)=.39, p<.05$ ; IQ 80~119 집단에서  $r(116)=.31, p<.01$ ; IQ 119 이상 집단에서  $r(26)=-.16, ns$ .

#### 연령별 집단에서의 DAP IQ와 K-WISC-III IQ 간 상관

연령수준에 따라 두 지능지수 간의 관계가 변하는지 확인하기 위해 각 연령집단 별로 두 지능지수 간 상관을 산출하였다. 전체 연구참여자들을 연령에 따라 미취학 아동(6세, 17명), 초등학교 저학년 아동(7~9세, 96명), 초등학교 고학년 아동(10~12세, 67명)으로 분류하였다. 세 집단에서 DAP IQ와 K-WISC-III IQ 간 상관은 취학전 아동 집단에서는 유의미하지 않았으나, 초등학교 취학 집단에서는 유의미한 정적 상관을 보였다, 취학전 아동 집단에서  $r(17)=.17, ns$ ; 초등학교 저학년 집단에서

$r(96)=.49, p<.01$ ; 초등학교 고학년 집단에서  $r(67)=.48, p<.01$ .

#### K-WISC-III 추정 IQ 회귀방정식

만약 DAP IQ와 K-WISC-III IQ 간 상관이 매우 높고, 또, 평균의 차이가 유의하지 않다면 두 지능검사가 유사한 수준의 지능지수를 산출해준다는 것을 의미한다. 본 연구에서 두 지능지수 간 높은 상관은 확인되었다. 별도로 차이검증은 하지 않았으나 두 지능지수 분포는 IQ  $111.99\pm 15.50$ 와  $97.10\pm 21.14$ 로, 큰 차이를 보이고 있다. 두 검사의 규준제작 시기의 차이를 감안할 때 어떤 지능검사를 사용하는 가에 따라 한 개인에 대해 다른 지능지수를 산출하게 될 것임은 분명하다. 30여년 전에 제작된 인물화 지능검사 규준을 새로 제작하지 않는 한 현재로는 DAP로 타당한 지능지수를 산출하는 것이 불가능하며, 그렇지 않은 현 시점에서 DAP로 지능수준을 판단하기 위해서는 DAP IQ와 K-WISC-III IQ 간 상관을 이용한 회귀방정식을 산출하여 일정한 교정을

표 3. DAP IQ로 K-WISC-III IQ를 추정할 때의 회귀방정식

집단 구분	회귀방정식
전체	WISC IQ = $24.70 + .65 \text{ DAP IQ}$
정상집단	WISC IQ = $67.34 + .34 \text{ DAP IQ}$
임상집단	WISC IQ = $30.70 + .57 \text{ DAP IQ}$
취학전(6세)	WISC IQ = $65.22 + .24 \text{ DAP IQ}, ns$
초등학교 저학년(7~9세)	WISC IQ = $23.00 + .67 \text{ DAP IQ}$
초등학교 고학년(10~12세)	WISC IQ = $16.41 + .72 \text{ DAP IQ}$
IQ 79 이하	WISC IQ = $28.73 + .36 \text{ DAP IQ}$
IQ 80~119	WISC IQ = $74.96 + .23 \text{ DAP IQ}$
IQ 120 이상	WISC IQ = $132.60 - .06 \text{ DAP IQ}, ns$

해주는 것이 도움이 될 수 있다.

DAP IQ로 K-WISC-III IQ를 추정하기 위해 DAP IQ를 독립변인, K-WISC-III IQ를 종속변인으로 하는 회귀방정식을 산출하였다. 전체 집단, 지능수준별, 연령별로 구분하여 9개의 회귀방정식을 산출하였다(표 3). 취학전 아동(N=17) 집단과 IQ 120 이상(N=26) 집단에서는 유의미한 회귀방정식이 산출되지 못하였으며, 이를 제외한 모든 집단에서 유의미한 회귀방정식이 산출되었다.

## 논 의

본 연구에서는 인물화에 의한 간편 지능검사(김재은 등, 1973)의 타당도를 검증하기 위해 DAP IQ와 K-WISC-III IQ 간 상관을 분석하였다. 충청지역에 거주하는 만 6세~12세 아동 182명(남아 113명, 여아 69명; 정상집단 119명, 임상집단 63명)이 연구에 참여하였다.

연구참여자들의 DAP IQ 평균은 112, K-WISC-III IQ의 평균은 97로, 전체적으로 볼 때 DAP가 K-WISC-III에 비해 약 15점정도 높은 IQ를 산출하였다. 상관분석 결과, DAP IQ 와 K-WISC-III IQ 간 상관이 전체 집단에서 .47, 정상집단에서 .30, 임상집단에서 .38로, 유의미한 정적 상관을 보였다. 이 상관계수는 Scott(1981)의 연구결과와 비슷한 수준이며, Abell, Heiberger, 그리고 Johnson(1994)의 연구에 비해서는 높은 값이다. 이 결과는 정상집단과 임상집단 모두에서 K-WISC-III에서 높은/낮은 지능지수를 얻는 아동은 DAP에서도 높은/낮은 지능지수를 얻게 된다는 것을 의미한다. 이는 K-WISC-III를 준거로 볼 때 인물화에 의한 간편형 지능검사에서 얻어진 지능지수가 타당함

을 말해주는 것이다. 여기서, 전체집단에 비해 정상집단과 임상집단으로 구분하여 따로 상관을 산출하였을 때 상관계수가 다소 낮게 산출된 것은 지능지수 값의 폭(range)이 축소되기 때문에 나타나는 인공적인(artifact) 현상(탁진국, 2007)으로 추측된다.

지능수준에 따라 세 집단으로 구분하여 검증한 결과 DAP IQ와 K-WISC-III IQ 간의 상관이 IQ 79 이하 집단에서  $r=.39$ , IQ 80~119 집단에서  $r=.31$ 로 유의미한 정적 상관을 보였으나 IQ 120 이상인 집단에서는 유의한 상관이 나타나지 않았다. 이 결과는 인물화 검사와 웨슬러 지능검사에서 산출된 IQ 간의 상관을 분석하여 IQ 80 이하인 집단에서 가장 높은 일치율을 보고하였던 Aikman 등(1992)의 연구와 유사한 결과이다. IQ 120 이상인 집단에서 유의한 상관이 나타나지 않은 것은 우선, DAP를 통한 지능평가가 고도의 지적 능력을 판별하기는 미흡한 검사임을 말해주는 것일 수 있다. 지능수준이 아주 높지 않은 경우나 발달장애를 가진 임상집단에서는 DAP를 통한 지능수준의 추정 또는 MR 여부의 판단이 시간과 비용 면에서 효율적일 수 있지만, 일정 수준 이상의 지적 성취를 이룬 아동들에서는 DAP가 지능수준 차이를 섬세하게 드러내지 못한다는 임상적 인상이 이러한 가능성을 뒷받침한다. 그밖에, IQ 수준을 기준으로 집단이 세분화되는 과정에서 이 집단의 크기가 가장 작았고, 또, 점수의 범위가 축소되었기 때문에 나타난 인공적인 현상일 가능성도 고려해 볼 수 있다.

IQ 120 이상인 집단에서 유의한 상관이 나타나지 않은 본 연구결과가 고능력 집단의 경우 인물화 검사가 지능수준을 잘 반영하지 못하기 때문인지, 자료의 축소 때문에 나타난

현상인지, 아니면 두 가능성의 모두 작용한 것인지 현재로는 판단하기 어렵다. 추후 연구에서는 고능력 집단에서 인물화 수행의 개인 차가 지능수준 외의 다른 변인으로 설명 가능 한지 체계적으로 검토해볼 필요가 있겠다. 또한, 지능수준이 이 범위에 속하는 사례수를 다른 집단과 유사한 정도로 표집하여 상관을 산출해볼 필요가 있겠다.

연령집단별 검증에서는 취학전 아동 집단에서는 유의미한 상관이 산출되지 않았으나 초등학교 저학년과 고학년 아동 집단에서는 .48~.49의 유의미한 정적 상관이 확인되었다. 초등학교 연령에서 가장 높은 상관이 확인된 본 연구결과는 Abell 등(1996) 연구에서도 관찰된 바 있다. 취학전 아동 집단에서 유의한 상관이 나타나지 않은 것 또한 본 연구의 집단 구분 과정에서 취학전(n=17) 아동이 초등학교 저학년(n=96)과 고학년(n=67)에 비해 사례수가 현저하게 작았는데, 이 때문에 낮은 상관이 산출된 것인지 아니면 이 연령층에서 인물화 검사의 수행이 지적 능력의 발달을 잘 반영하지 못함을 시사하는 것인지 판단하기 어렵다. 다만, Goodenough(1926)와 그 이후의 많은 연구자들이 특히 아동의 인물화 검사가 학습기능, 지능수준, 정신연령, 인지발달 등과 높은 상관이 있음을 밝혀온 점을 감안할 때 후자보다는 전자의 가능성성이 더 클 것으로 추측된다. 그러나 DAP가 정교한 시각-운동 협응을 필요로 하는 만큼 일정한 신체적 연령에 도달한 후에 라야 지적 능력이 이 검사에 반영될 수 있는지도 모른다. 추후 연구에서 이 연령대의 집단의 크기를 충분히 했을 때 의미 있는 상관이 나타나는지, 또, 취학전 연령 아동의 시각-운동 협응능력과 인물화 검사 수행 간 어떤 관계가 있는지 등을 알아볼 필요가 있겠다.

한편, DAP IQ와 K-WISC-III IQ는 전체집단을 기준으로 할 때 약 15점 정도의 차이가 있었는데, 이러한 차이 중 상당 부분은 두 검사의 규준제작 시점상 30년 가까운 차이에서 오는 Flynn 효과로 설명될 수 있다. 이 경우 두 IQ의 차이검증은 의미가 없다. 본 연구에서 지능수준이 높은(IQ 120 이상) 집단과 미취학 연령 집단을 제외한 대부분의 정상 및 임상 집단에서 유의한 상관이 관찰되어 이 검사에서 산출된 IQ의 타당성을 확인되었지만 이 검사에서 산출되는 IQ를 직접 사용하기는 어렵다. 이 문제를 극복하기 위해서는 근본적으로 인물화 지능검사의 규준이 새로 제작되어야 한다. 본 연구에서는 회귀방정식을 이용하여 DAP에서 산출된 IQ로 K-WISC-III IQ를 추정할 수 있는 방안을 제시하였다.

인지기능, 특히 지능 측정의 한 방안으로 인물화 검사를 활용하는 방안과 그 타당성에 대해 최근까지도 다양하게 연구가 수행되고 있다(예를 들어, Bensur & Eliot, 1993; Ericsson, Hilleras, Holman, Jorn, Forssell, Almk, Ronnberg, & Winblad, 1994; Levin & Bus, 2003; Plubrukarn, & Theeramanoparp, 2003). 우리나라에서는 임상 현장에서 여전에 따라 인물화 검사를 통해 지능을 측정하는 경우가 있지만 지능평가 도구로서의 인물화 검사에 대한 타당도를 검증한 연구는 찾아보기 어렵다. 본 연구결과는 시간과 비용, 피검자의 주의지속 능력 또는 흥미 등의 제한으로 인해 웨슬러 지능검사를 실시하기 어려운 경우 또는 대략적인 지능수준의 평가만 필요한 상황에서 인물화 검사가 지능 지수 산출을 위한 타당한 대안이 될 수 있음을 시사한다.

## 참고문헌

- 곽금주, 박혜원, 김청택 (2001). 한국 웨슬러 아동지능검사(K-WISC-III) 지침서. 서울: 특수교육.
- 김난수 (1963). 지능성숙검사 실시요강. 서울: 중앙교육연구소.
- 김재은, 김동극, 여광웅 (1973). 인물화에 의한 간편 지능검사. 서울: 교육과학사.
- 탁진국 (2007). 심리검사. 학지사.
- Abell, S. C., Von Briesen P. D., & Watz L. S. (1996). Intellectual evaluations of children using human figure drawings: An empirical investigation of two methods. *Journal of Clinical Psychology*, 52, 67-74.
- Abell, S. C., Heiberger, A. M., & Johnson, J. E. (1994). Cognitive evaluations of young adults using human figure drawings: An empirical investigation of two methods. *Journal of Clinical Psychology*, 50, 900-905.
- Aikman, K. G., Belter, R. W., & Finch, A. J., JR. (1992). Human figure drawings: Validity in assessing intellectual level and academic achievement. *Journal of Clinical Psychology*, 48, 114-120.
- Bensur, B., & Eliot, J. (1993). Case's Developmental Model and Children's Drawings. *Perceptual and Motor Skills*, 76, 371-375.
- Buck, J. (1948). *The House-Tree-Person technique*. Los Angeles: Western Psychological Services.
- Buck, J. (1966). *The House-Tree-Person technique: Revised manual*. Los Angeles: Western Psychological Services.
- Ericsson, K., Hilleras, P., Holman, K., Jorn, A., Forssell, L. G., Almk, O., Ronnberg, L., & Winblad, B. (1994). The Short Human Figure Drawing Scale for Evaluation of Suspect Cognitive Dysfunction in Old Age. *Archives of Gerontology and Geriatrics*, 19, 243-251.
- Flynn, J. R. (1984). The Mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51.
- Flynn, J. R. (1987). Massive Gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Goodenough, F. L. (1926). *Measurement of intelligence by drawing*. New York: Harcourt, Barce, & World.
- Goodenough, F. L., & Harris, D. B. (1963). *The Goodenough-Harris Drawing Test*. New York: Harcourt, Brace & World.
- Koppitz, E. (1968). *Psychological evaluation of Children's human figure drawings*. New York: Grune & Stratton.
- Levin, I., & Bus, A. G. (2003). How is Emergent Writing Based on Drawing? Analyses of Children's Products and their Sorting by Children and Mothers. *Developmental Psychology*, 39, 891-905.
- Plubrukarn, R., & Theeramanoparp, S. (2003). Human figure drawing test: validity in assessing intelligence in children aged 3-10 years. *Journal of Medical Association of Thailand*, 86 Suppl 3, 5610-5617.
- Resnick, R. J., & Entin, A. D. (1971). Is an abbreviated form of the WISC valid for Afro-Americans? *Journal of Consulting and Clinical Psychology*, 36, 97-99.

- Scott, L. (1981). Measuring intelligence with the Goodenough-Harris Drawing Test. *Psychological Bulletin, 89*, 483-505.
- Tramill, J. L., Edwards, R. P., & Tramill, J. K. (1980). Comparison of the Goodenough-Harris Drawing Test and the WISC-R for Children Experiencing Academic Difficulties. *Perceptual and Motor Skills, 50*, 543-546.
- Wechsler, D. (1991). *Manual of the Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX: The Psychological Corp.

원고접수일 : 2007. 6. 15.

제재결정일 : 2007. 7. 6.

## Validity of DAP as Intelligence Test: Comparison of IQs between DAP and K-WISC-III

Soon-Taeg Hwang Young-Soon Jun Eun-Jung Noh Yun-Jin Jo Yeo, Sang Woo

Department of Psychology, Chungbuk National University

This study was to test the validity of DAP as intelligence test. IQ scores of DAP and K-WISC-III of the 182 normal and clinical children aged 6-12 years were compared. There were significant positive correlations between two IQs of all participants. In comparisons of IQs of subgroups divided by IQ level and age, correlations of two IQs in the subgroups divided by IQ level were significant except in the subgroup of IQ 120 and above. Also Correlations of two IQs in the subgroups by age were significant except in the subgroup of the preschooler. The regression equations were proposed for transformation DAP IQ into K-WISC-III IQ.

*Key words : Drawing A Person Test, IQ, Validity, K-WISC-III*