

한국 웨슬러 아동 지능검사(K-WISC-III) 단축형의 타당도*

전 영 순 황 순 택* 이 속 희

충북대학교 심리학과

본 연구는 한국 웨슬러 아동 지능검사(K-WISC-III)의 Doppelt 방식 단축형의 타당도를 검증하고자 하였다. Campbell(1998)이 WISC-III 단축형으로 제안한 7가지의 구체적인 단축형과 K-WISC-III 표준화 자료에서 요인부하량이 가장 높은 소검사들로 구성된 단축형을 검증 대상으로 하였다. 만 6세부터 16세 사이의 정상 및 임상 집단 총 294명을 무선적으로 두 집단으로 나누고 그 중 한 집단의 점수로 8 가지 단축형 각각에 대한 IQ 회귀공식을 산출하고 이 공식을 다른 한 집단에게 적용하여 단축형 추정 IQ를 산출하고 이 IQ와 이 집단의 실제 전체 IQ 간 상관 및 차이검증을 실시하였다. 그 결과 단축형 추정 IQ와 전체 IQ 간 높은 상관이 확인되었으며, 두 IQ의 평균차이는 '일반지식+빠진곳찾기'로 구성된 단축형을 제외한 나머지 7 가지 단축형에서는 통계적으로 유의미한 차이가 관찰되지 않았다. 또한, 단축형 추정 IQ와 전체 IQ를 지능범주로 분류하였을 때 범주분류의 상관이 높았다. 이상과 같은 결과는 K-WISC-III 단축형의 타당성을 말해주는 것으로 해석되었다.

주요어 : 지능검사, 단축형, 타당도, K-WISC-III

* 본 연구는 2007년도 충북대학교 학술연구지원사업의 연구비지원에 의해 연구되었음(This work was supported by the research grant of the Chungbuk National University in 2007).

† 교신저자(Corresponding Author) : 황순택 / 충북대학교 심리학과 / 충북 청주시 흥덕구 개신동
Fax : 043) 261-2187 / E-mail : hstpsy@chungbuk.ac.kr

웍슬러 지능검사는 여러 개의 소검사들로 구성된 전체척도 모두를 실시하여 개인의 지적 능력을 산출한다. 그러나 이 검사를 임상에서 사용하기 시작한 초기부터 전체 소검사를 다 실시하는데 드는 시간적, 인적 효율성에 대해 종종 문제가 제기되어 왔다. WAIS-III (Wechsler Adult Intelligence Scale-III; Wechsler, 1997)는 전체 소검사를 실시하는데 평균 75분 정도의 시간이 소요된다. Ryan, Lopez와 Werth (1998)의 연구에서 임상집단을 대상으로 WAIS-III를 실시했을 때 평균 91분(범위 54~136분) 소요되었다. K-WISC-III의 경우 정해진 10개의 소검사를 실시하는 데는 대략 50-70분 정도가 소요되며, 세 개의 보충 소검사는 추가로 10~15분이 더 소요된다(곽금주, 박혜원, 김청택, 2001a, b).

이처럼 긴 실시시간이 요구되는 검사를 임상장면에서 사용하는 데는 경우에 따라 상당한 어려움이 있을 수 있다. 예컨대, 만성 정신질환자들의 경우 주의집중력의 저하, 피로감, 동기 부족 등으로 인해 2시간가량 소요되는 전체 소검사를 한 번에 모두 실시하기가 어려울 때가 많다(Ryan, Georgemiler, & Mckinney, 1984). 두부 외상이나 질병으로 인한 대뇌 신경학적 기능의 장애가 있는 환자의 경우에도 주의력의 저하, 충동조절의 어려움, 기타 다양한 이유로 인해 전체 지능검사를 다 실시하기 어려운 경우가 흔히 있다. 성인용 지능검사는 물론 아동용 지능검사의 경우 이러한 어려움은 더 크다.

이러한 어려움 때문에 검사 전체를 실시했을 때 얻을 수 있는 정보의 일부를 포기하면서 검사의 일부만 실시하는 방안이 제안되었다. Silverstein은 웍슬러 지능검사를 실시하는데 많은 시간이 소요되는 만큼 일부 소검사만을

선택하여 구성된 단축형 검사로 피검자의 전체척도 지능을 타당하게 추정할 수 있다면 매우 유용한 시도라고 제안하고, 풍부한 임상적인 함의를 지닌 몇몇 WAIS-R 단축형을 제시한 바 있다(Silverstein, 1982, 1984, 1985). 그러나 지능검사 단축형은 피검자의 지적 능력에 대한 세밀한 정보를 상실할 수 있고, 또, 지능지수의 분류(IQ classification)에서 정확성이 손상될 수 있어 연구 또는 특수한 임상장면에 국한하여 제한적으로 사용되어 왔다(Silverstein, 1990).

우리나라에서도 몇몇 연구자들에 의해 지능검사 단축형의 타당도가 연구되었다. 성인용 지능검사 단축형에 대한 연구로 KWIS의 Doppelt 방식 단축형이 김중술, 김영환(1974)에 의해 처음 소개됐다. 김중술, 이용승, 이민식(1994)은 K-WAIS의 원자료를 분석하여 두 가지 소검사를 이용한 추정지능과 네 가지 소검사를 이용한 추정지능을 제시하면서 대략적인 지능을 알고자 할 때에는 이러한 단축형 검사를 이용한 추정이 매우 타당하다고 보고하였다. 또한 임상장면에서 K-WAIS를 이용한 이용승과 김중술(1995)의 단축형 연구도 있다. 우리나라에서 현재까지 가장 널리 사용되어온 Doppelt 방식 단축형은 어휘, 산수, 토막짜기, 차례맞추기 등 4가지 소검사로 이루어져 있는데, 지능 검사의 3개 요인 구조에 따르면 이 4가지 소검사는 언어적 이해(어휘 소검사)요인, 지각적 조직화(토막짜기, 차례맞추기 소검사)요인, 주의력(산수 소검사)요인을 측정한다고 볼 수 있다(임영란, 이우경, 이원혜, 박종원, 2000).

그리고 한국 웍슬러 아동용 지능검사(K-WISC)에 대한 단축형 연구도 진행되었다. 우선 Doppelt(1956) 방식으로 이창우와 서봉연

(1974)은 정상아동들을 대상으로 하여 몇 개의 소검사를 실시한 후 각 소검사간의 상관계수들의 합을 이용하여 전체지능지수를 추정할 수 있도록 하였다. 또한 Satz와 Mogel(1962)방식으로 각 소검사에서 몇 개의 문항을 채택하여 전체지능지수의 추정치를 산출하는 방법으로 김명권과 김중술(1985)이 임상집단 아동을 대상으로 실시하여, 단축형의 타당성을 지지해 주는 결과를 보고하였다. 박혜원(2001)은 한국 웨슬러 유아지능검사(Korean Wechsler Preschool and Primary Scale of intelligence; K-WPPSI)의 단축형에 Doppelt 방식을 이용하여 연구하였는데, 높은 신뢰도와 임상적 타당도 기준은 충족시켰으나, 표준편차의 범위가 다소 넓게 나타나 단축형을 사용하기 위해서는 새로운 기준을 마련해야 함을 시사하였다.

단축형을 구성할 때 고려해야 할 사항으로는 1) 단축형을 어떻게 구성할 것인지, 2) 단축형 검사에서 얻어진 원점수로 IQ를 어떻게 추정할 것인지, 3) 단축형의 타당도를 어떻게 판단할 것인지 등이 있다. 웨슬러 지능검사의 단축형을 구성하는 방법으로는 몇몇 소검사를 선택하는 Doppelt 방식(Doppelt, 1956)과 모든 소검사의 일부 문항만을 선택하는 Satz-Mogel 방식(Satz & Mogel, 1962)이 대표적이다. 일부 소검사만을 선택하는 방식은 다양한 정신과정을 평가하는데 제한이 있으며, 그에 비해 전체 소검사들의 일부 문항만을 선택하는 것은 원척도가 가지고 있는 다양한 기능을 반영할 수 있다. Wolfson과 Bachelis(1960), Yudin(1966) 등은 Satz-Mogel 방식 단축형의 타당도를 검증하였으며, Kaufman(1972,1975)과 Silverstein(1967, 1982)은 Doppelt 방식 단축형의 타당도를 검증하였다.

Doppelt 방식 단축형에서 소검사를 선택하

는 구체적인 방법으로 Levy(1968)는 1) 전체척도 점수와 가장 높은 상관을 가지고 있는 소검사들을 표집(Scale sampling)하는 방법, 2) 전체척도에서 언어성과 동작성을 각각 가장 잘 대표하는 소검사들을 표집하는 유층적(Stratified) 방법, 3) 요인분석에서 산출되는 요인부하량을 바탕으로 소검사들을 선택하는 요인표집(Factor sampling) 방법, 4) 실시시간이나 피검자의 장애 여부, 정서 상태 및 기타 검사자의 개인적인 경험에 근거하여 소검사를 다르게 표집하는 방법 등을 제안하였다.

단축형에 의거하여 추정지능을 산출하는 방법으로는 Satz-Mogel 방식의 경우 각 소검사의 실시문항 대 전체문항 비율을 역으로 곱해서 해당 소검사의 추정 원점수를 산출할 수 있다. Doppelt 방식의 경우 Satz-Mogel 방식에서처럼 비율공식을 이용하는 방법과 회귀(Regression estimates)를 이용하여 추정하는 방식이 있다(예를 들어, Doppelt, 1956; Silverstein, 1974). 여기서, 전체 소검사 수 대비 사용되는 소검사의 비율을 이용하여 전체 지능지수를 추정하는 방식은 개념이 단순하고 실제 사용상 간편한 점이 장점이 될 수 있으나 각 소검사가 웨슬러 지능검사의 전체척도를 대표하는 정도에 있어 개념적으로 그리고 실제적으로 차이가 있는 만큼 조악한 측면이 있다. 회귀를 이용하여 전체 지능을 추정하는 방식은 추정을 위한 회귀 방정식을 산출해야 하는 번거로움이 있으나 심리측정적인 측면에서 볼 때 보다 정교하다는 장점이 있다.

검사 전체를 실시하여 얻어지는 지능지수를 대신하여 단축형 검사에서 추정된 지능지수를 사용하기 위해서는 추정된 지능지수가 원래의 지능지수와 일치도가 높아야 할 것이다. 이 때문에 단축형에 의해 추정된 지능지수의 임

상적 타당도를 검증하기 위한 연구들이 다양하게 시도되었다. 대부분의 단축형 지능검사의 타당도 연구는 Rabin(1943) 과 Resnick-Entin(1971)이 제시한 임상적 타당도 기준을 사용하였다. Rabin과 Resnick-Entin이 제시한 기준은 다음과 같다: 1) 전체척도 지능지수와 단축형에 의한 추정 지능지수 간에 높은 상관관계가 있어야 한다; 2) 전체척도 지능지수 평균과 단축형 추정된 지능지수 평균 간의 차이가 유의미하지 않아야 한다; 3) 전체척도 지능지수와 단축형 추정 지능지수의 범주적 진단분류(경계선, 보통상, 최우수 등)에서 높은 일치율을 보여야 한다.

많은 연구에서 Rabin과 Resnick-Entin의 첫 번째 기준인 전체척도 지능지수와 단축형 추정 지능지수 간에 높은 상관관계가 확인되었지만 두 번째 기준인 두 지능지수 간 차이의 동등성은 충족되지 않았다. 예를 들어, GoH(1980), Ryan, Georgemiller과 MacKinney(1984), Ryan, Larsen과 Prifitera(1983), Silverstein(1985) 등의 연구에서 일부 소검사만 사용하여 단축형 추정 지능을 산출했을 때 전체척도 지능지수와 높은 상관은 있었지만 평균차이 또한 유의미한 것으로 나타났다. 일반적으로 두 점수분포 간의 차이가 임상적으로 볼 때 의미를 둘 정도로 크지 않더라도 표본의 크기가 커짐에 따라 차이가 통계적으로 유의미한 것으로 판단될 수 있다. 이런 점에서 이 두 번째 기준을 일괄적으로 적용하는 데 대해서는 다소간 논란이 있다. 평균차이의 통계적 유의도 검증 뿐만 아니라 함께 연구에 사용된 표본의 크기, 지능지수 차이의 크기 등도 고려해야 한다.

Rabin과 Resnick-Entin이 제시한 타당도 기준 중 마지막 기준, 즉, 두 지능지수 간의 범주적 진단분류의 일치율 또한 많은 논란이 되어왔

다. 전체척도 지능지수와 단축형 추정 지능지수가 완벽하게 일치하여 상관이 1.00이 되지 않는 한 불일치는 존재할 수밖에 없다.

Silverstein(1985)의 연구에서는 두 지능지수간 상관이 .95일 때조차도 불일치율이 25%를 초과하였다. GoH(1980)의 연구에서는 표집된 아동들 중 45%가 다른 지능 분류 범주에 해당되었다. Ryan, Georgemiller과 MacKinney(1984)의 연구에서 2개 소검사 단축형을 통한 지능분류의 불일치율은 37%였으며, 4개 소검사를 사용한 단축형에서도 지능분류의 불일치율은 25%나 되었다. Rabin과 Resnick-Entin의 세 번째 타당도 기준은 임상장면에서 중요한 의미를 지닐 수 있지만 이에 관한 이러한 연구결과들이 기준이 심리측정적인 측면에서는 충족시키기가 대단히 어려운 기준임을 말해주는 것이다. 즉, 전체척도 지능지수와 단축형 추정 지능지수 간의 지능지수 1점 차이가 지능분류의 불일치로 집계될 수도 있는 반면, 지능지수 10점 차이는 정확하게 일치하는 분류로 집계될 수도 있어 심리측정적인 측면에서는 그다지 합리적이지 못한 기준이기도 하다.

개인용 지능검사를 실시하는 데는 상당한 시간이 소요되기 때문에 특별한 상황과 조건에서는 검사 전체를 실시하는 것이 불가능한 경우가 있다. 이 때문에 검사 전체를 실시했을 때 얻을 수 있는 정보의 일부를 포기하면서 검사의 일부만 실시하는 단축형이 Silverstein을 비롯한 여러 연구자들에 의해 다양하게 제안되고 검토되어 왔다. 본 연구는 한국 웨슬러 아동지능검사(K-WISC-III) 단축형의 타당도를 검증하기 위한 것이다. 본 연구에서 검증하고자 하는 구체적인 단축형은 Doppelt 방식에 의거하여 Campbell(1998)이 제안한 단축형과 K-WISC-III 규준 자료에서 추출된 요인을

중심으로 한 단축형이다. 타당도의 판단은 Rabin(1943)과 Resnick-Entin(1971)이 제시한 기준을 사용하였다.

방 법

연구 참여자

서울과 청주 지역에 거주하는 만 6세 0개월부터 16세 4개월 사이의 아동 총 294명(남아 181명, 여아 113명; 정상집단 130명, 임상집단 146명)이 본 연구에 참여하였다. 정상집단 아동의 경우 청주에서 유치원, 초등학교, 중학교에 재학 중인 학생들로 이루어져 있으며, 임상집단은 소아정신과 외래를 방문하여 심리평가와 함께 소아정신과 전문의로부터 정신과적 진단을 받은 아동들로 이루어져 있다.

연구 도구

본 연구에서 사용한 한국 웨슬러 아동지능검사(Korean-Wechsler Intelligence Scale for Children-III: K-WISC-III)는 미국의 WISC 개정 3판(WISC-III; Wechsler, 1992)의 한국판이다. K-WISC-III는 광금주, 박혜원, 김청택(2001a,b)에 의해 표준화되었다.

연구절차

임상아동 자료는 소아 신경정신과 병원에 내원한 아동의 심리검사 자료 중 K-WISC-III 자료를 수집하였다. 이들 검사자료는 병원에서 임상심리 전문가가 또는 임상심리수련생에 의해 실시된 것들이다. 정상아동 자료는 심리

학과 대학원에서 심리평가 수업을 수강하고 실습을 한 대학원생 12명이 실시하였으며, K-WISC-III 지침서에 따라 전체척도를 실시하여 얻었다. 실시시간은 라포형성 시간을 포함해 평균 1시간 50분 정도가 소요되었다.

본 연구에서 채택한 단축형 방식은 Doppelt 방식으로, Campbell(1998)의 연구에서 사용된 WISC-III 단축형 7 가지와 함께 한국 웨슬러 아동지능검사(K-WISC-III) 표준화 자료에서 제시한 4요인에서 요인부하량이 가장 높은 소검사를 중심으로 구성된 단축형(어휘-토막짜기-산수-기호쓰기 소검사)을 대상으로 하였다. Campbell(1998)이 사용한 단축형 구성은 그 이전까지의 다양한 연구에서 발표된 단축형을 바탕으로 한 것이다. Campbell(1998) 7 가지 단축형의 구성은 다음과 같다:

상식-빠진 곳 찾기(I-PC)

이 두 척도는 전체척도(Full Scale), 언어성 척도(Verbal Scale)와 동작성 척도(Performance Scale)와 높은 상관을 나타내는 소검사로, 피험자의 동기나 피로와 관련된 문제를 피할 수 있는 실시시간으로 구성되었다(Kaufman, Ishikuma, & Kaufman-Packer, 1991).

어휘-토막짜기(V-BD)

어휘-토막짜기 소검사는 전체척도와 높은 상관($r = .77, .76$)을 보여주며, 두 소검사는 언어성 척도와 동작성 척도에서 일반능력 g 요인을 가장 잘 설명하는 소검사이다(Herrera-Graf, Dipert, & Hinton, 1996).

공통성-어휘-차레맞추기-토막짜기(S-V-PA-BD)

이 단축형은 심리측정적인 요소와 임상적인

상황을 고려하여 선택된 소검사들이다(Kaufman, Kaufman, Balgopal, & McLean, 1996).

상식-산수-빠진곳찾기-기호쓰기(I-A-PC-Cd

이 단축형은 실용적인 측면 특히 채점의 용이성과 간결성을 우선적으로 고려하여 선택한 소검사들이다(Kaufman, Kaufman, Balgopal, & McLean, 1996).

공통성-산수-빠진곳찾기-토막짜기(S-A-PC-BD)

이 구성은 소척도 점수-전체 지능지수간 높은 상관, 실시 및 채점 시간의 단축, 라포 형성의 용이성, 임상적 해석의 풍부성 등 심리

측정적인 측면과 임상적 측면 그리고 실질적인 측면 모두를 다 고려하여 구성된 것이다(Kaufman, Kaufman, Balgopal, & McLean, 1996).

빠진곳찾기-상식-기호쓰기-토막짜기-어휘(PC-I-Cd-BD-V)

단축형 구성에 있어서 언어성 척도와 동작성 척도가 적어도 각각 2개씩은 포함돼야 하며, 그러한 소검사들은 전체척도와 높은 상관을 보여야 하고, 다양한 정신과정, 실시자의 효율성, 임상적인 측면 등을 반영해야 한다는 Kaufman(1990)의 제안에 의거하여 구성된 것이다(Dumont & Faro, 1993).

표 1. K-WISC-III 단축형 구성 및 추정 지능지수 계산 공식

| 단축형 | 소 검사 구성 | 회귀 공식 |
|-----|------------------------------------|---|
| ① | I-PC (Kaufman et al., 1991) | FIQ = 45.634 + 3.307 I+ 2.199 PC |
| ② | V-BD (Herrera-Graf et al., 1996) | FIQ = 38.098 + 3.211 V + 2.8157 BD |
| ③ | S-V-PA-BD (Kaufman et al., 1996) | FIQ = 30.502 + 1.090 S + 2.270 V+ 1.638 PA + 1.796 BD |
| ④ | I-A-PC-Cd (Kaufman et al., 1996) | FIQ = 37.3074 + 1.775 I + 1.802 A + 1.598 PC + 1.282 Cd |
| ⑤ | S-A-PC-BD (Kaufman et al., 1996) | FIQ = 31.779 + 1.913 S + 1.546 A + 1.690 PC + 1.478 BD |
| ⑥ | PC-I-Cd-BD-V (Dumont & Faro, 1993) | FIQ = 33.016 + 1.180 PC +0.829 I + 1.162 Cd + 1.475 BD +2.158 V |
| ⑦ | V-S-PC-BD-A-Cd (Donders, 1997) | FIQ = 30.991 + 1.796 V + 0.780 S+ 1.342 PC + 1.176 BD +0.866 A + 0.908 Cd |
| ⑧ | V-BD-A-Cd (곽금주 등, 2001b) | FIQ = 35.171 + 2.426 V+ 1.876 BD + 1.013 A + 1.188 Cd |

주: K-WISC-III = Korean-Wechsler Intelligence Scale for Children-III; 회귀공식은 본 연구에 사용된 아동 집단을 집단1과 집단2로 무선허당 하여, 그중 집단 1에서 산출한 회귀식이다. FIQ(Full IQ); I (Informantion: 상식); PC (Picture Completion: 빠진곳찾기); V (Vocabulary: 어휘); BD (Block Design: 토막짜기); S (Similarities: 공통성); PA (Picture Arrangement: 차례맞추기); A (Arithmetic: 산수); Cd (Coding: 기호쓰기).

어휘-공통성-빠진곳찾기-토막짜기-산수-기호쓰기(V-S-PC-BD-A-Cd)

이 구성은 WISC-III 요인분석을 바탕으로 소검사를 선정한 것이다(Donders, 1997). 언어 이해 요인 중 상식과 공통성 소검사, 지각조직 요인에서 토막짜기 소검사와 빠진곳찾기 소검사, 주의집중 요인에서 산수 소검사, 처리 속도 요인에서 기호쓰기 소검사가 각각 선택되었다.

지수와 단축형 추정 지능지수 간의 상관계수를 구하고, 두 지능지수의 차이를 검증하기 위해 대응표본 *t* 검증(Paired *t*-test)을 하였다. 마지막으로 웨슬러 지능의 범주적 진단분류에 따라 전체척도 지능지수와 단축형 추정 지능지수의 범주분류를 결정한 후 Spearman의 rho 계수로 두 범주 간 상관계수를 구하였으며, 분류 일치도를 산출하였다.

분석 방법

자료 분석은 단축형 8가지 방식에 따른 각각의 추정 지능지수 계산은 회귀공식을 이용하였다. 먼저 전체 연구 참여자 294명을 무선적으로 두 집단으로 나누고 첫 번째 집단에서 각 단축형에 포함되는 소검사들로 전체 지능지수를 예측하는 회귀공식(표 1)을 산출한 다음 두 번째 집단에 이 회귀 공식을 적용하여 추정 지능지수를 산출하였다. 이 추정 지능지수의 타당도를 검증하기 위해 전체척도 지능

결 과

K-WISC-III 전체척도에서 전체 아동, 정상아동, 임상아동의 지능지수는 각각 97.33 ± 21.55 , 107.94 ± 14.49 , 88.99 ± 22.55 였다. 회귀식 산출에 사용된 1 집단($n=147$)의 전체 지능지수는 97.08 ± 22.17 , 회귀식을 적용하여 두 지능지수의 상관, 차이검증, 범주적 분류 일치율의 검증에 사용된 2 집단($n=147$)의 전체 지능지수는 97.71 ± 20.97 였으며, 두 집단 간 지능지수의 차이는 유의하지 않았다, $t=0.286$, $p>.10$.

표 2. K-WISC-III 전체척도 IQ 및 회귀식으로 추정된 단축형 IQ의 평균, 표준편차

| | | 평 균 | 표준편차 |
|-----------------------------|------------------|-------|-------|
| 전체척도 IQ | | 97.71 | 20.97 |
| 단 축 형 추 정 IQ | ① I-PC | 99.52 | 16.34 |
| | ② V-BD | 97.65 | 18.09 |
| | ③ S-V-PA-BD | 97.52 | 19.00 |
| | ④ I-A-PC-Cd | 98.39 | 17.64 |
| | ⑤ S-A-PC-BD | 97.72 | 18.51 |
| | ⑥ PC-I-Cd-BD-V | 97.68 | 18.44 |
| | ⑦ V-S-PC-BD-A-Cd | 97.32 | 18.50 |
| | ⑧ V-BD-A-Cd | 97.32 | 18.31 |

주. 단축형을 구성하는 소검사 약자는 표 1의 각주 참조.

1 집단에서 단축형에 포함되어 있는 소검사의 평가치를 이용하여 전체 지능지수를 추정할 때의 회귀식을 산출하였으며, 이 회귀식을 2 집단에 적용하여 추정된 단축형 지능지수를 산출하였다. 단축형에 의해 추정된 지능지수와 전체척도에서 유도된 지능지수는 표 2에 제시하였다.

K-WISC-III 전체척도 지능지수와 단축형 추정 지능지수 간의 상관

전체집단의 상관계수 r 의 범위는 .88 ~ .96였으며, 평균 .93 이었다(표 3). 8 개의 추정 방식(I-PC, V-BD, S-V-PA-BD, I-A-PC-Cd, S-A-PC-BD, PC-I-Cd-BD-V, V-S-PC-BD-A-Cd, V-BD-A-Cd)에서의 상관은 각각 .88, .89, .95, .94, .93, .95, .96, .93 이었으며, 모든 상관계수는 $p < .001$ 수준에서 유의하였다. 이 결과는 단축형에 의한 추정 지능지수와 전체척도 지능지수가 높은 상관관계를 가지고 있음을 말해준다.

K-WISC-III 전체척도 지능과 단축형 추정지능 간 차이검증

전체척도 지능과 단축형 추정지능들 간의 차이의 절대값 범위는 0.01 ~ 1.93 이었다. 전체척도 지능지수와 단축형 추정 지능지수의 동등성을 확인하기 위해 대응표본 t 검증(Paired t test)을 실시하였다(표 3). 그 결과 상식-빠진곳찾기(I-PC) 단축형, $t=2.27, p<.05$,을 제외한 모든 단축형에서 산출된 추정 지능지수와 전체척도 지능지수의 평균차가 유의하지 않다는 것을 의미한다.

K-WISC-III 전체척도 지능과 단축형 추정지능의 Wechsler 범주적 진단 분류간 상관 및 일치도

K-WISC-III 전체척도를 이용하여 산출한 전체지능과 일부 소검사만을 선택한 단축형으로부터 추정된 지능지수를 웨슬러의 범주적 분류에 따라 7개 범주(정신지체, 경계선, 평균하,

표 3. K-WISC-III 전체척도 지능과 회귀식으로 추정된 단축형 지능 간 상관 및 차이 검증

| 단축형 | 소검사 구성 | r | 평균 차 | 표준편차 | t | 유의수준 |
|-----|----------------|--------|-------|-------|-------|----------|
| ① | I-PC | .88*** | 1.93 | 10.30 | 2.27 | $p=.025$ |
| ② | V-BD | .89*** | -0.07 | 9.69 | -0.09 | $p=.932$ |
| ③ | S-V-PA-BD | .95*** | -0.20 | 6.45 | -0.37 | $p=.709$ |
| ④ | I-A-PC-Cd | .94*** | 0.80 | 7.66 | 1.26 | $p=.210$ |
| ⑤ | S-A-PC-BD | .93*** | 0.01 | 7.99 | 0.01 | $p=.991$ |
| ⑥ | PC-I-Cd-BD-V | .95*** | 0.09 | 6.55 | 0.17 | $p=.862$ |
| ⑦ | V-S-PC-BD-A-Cd | .96*** | -0.40 | 6.29 | -0.76 | $p=.447$ |
| ⑧ | V-BD-A-Cd | .93*** | -0.39 | 7.78 | -0.61 | $p=.543$ |

*** $p < .001$

주: 단축형을 구성하는 소검사 약자는 표 1의 각주 참조.

표 4. K-WISC-III 전체척도 지능과 회귀식으로 추정된 단축형 지능간의 Wechsler 지능의 범주적 진단분류의 상관(roh) 및 일치도

| 단축형 | 소검사 구성 | 전체척도 지능지수 | |
|-----|----------------|-----------|----------|
| | | roh | 범주분류 일치도 |
| ① | I-PC | .80*** | .50 |
| ② | V-BD | .86*** | .59 |
| ③ | S-V-PA-BD | .91*** | .70 |
| ④ | I-A-PC-Cd | .88*** | .57 |
| ⑤ | S-A-PC-BD | .85*** | .62 |
| ⑥ | PC-I-Cd-BD-V | .89*** | .68 |
| ⑦ | V-S-PC-BD-A-Cd | .92*** | .71 |
| ⑧ | V-BD-A-Cd | .85*** | .57 |

*** $p < .001$

주: 단축형을 구성하는 소검사 약자는 표 1의 각주 참조.

평균, 평균상, 우수, 최우수)로 각각 구분하고, 두 지능범주 간 일치도를 Spearman의 roh 계수로 산출하고 이에 대한 유의도 검증을 실시하였다(표 4). 그 결과 8 개의 단축형에서 roh 계수는 .804~.918로 모두 $p < .001$ 수준에서 유의한 상관이 산출되었다. 이 결과는 단축형에 의해 추정된 지능지수의 범주분류와 전체척도 지능지수의 범주분류가 상당히 일치한다는 것을 시사한다. 다음으로 정확분류 사례/전체 사례로 두 지능지수의 범주분류 일치도를 비교한 결과 유형에 따라 .50~.71의 일치도를 보였다(표 4).

논 의

본 연구는 한국 웨슬러 아동지능검사(K-WISC-III)의 Doppelt 방식 단축형의 타당도를 검증하기 위한 것이다. Campbell(1998)이 제안

한 7 가지 단축형과 K-WISC-III 규준자료에서 추출된 요인을 중심으로 한 단축형 등 총 8 가지 단축형 구성에 대해 타당도를 검증하였다. 만 6세~16세의 일반아동과 임상아동(294명)의 아동용 웨슬러 지능검사 자료를 무선적으로 두 집단으로 나누어 첫 번째 집단에서는 8 개 단축형 각각으로부터 지능지수를 추정하는 회귀식을 산출하였다. 이 회귀식을 두 번째 집단의 자료에 적용하여 지능지수를 추정하고 단축형에 의해 추정된 지능지수와 전체척도에서 산출된 지능지수 간의 상관검증, 차이검증, 일치도 검증을 시행하였다.

그 결과 첫째, 전체척도 지능지수와 단축형에서 추정된 지능지수 간 상관의 유의도 검증에서는 8 가지 단축형 모두에서 .88~.96의 비교적 높은 상관이 관찰되었다. 이 결과는 Rabin(1943)과 Resnick-Entin(1971)이 제시한 세 가지 기준 중 첫 번째 기준을 충족시키는 것이다.

둘째, 전체척도 지능지수와 추정된 지능지

수 간 대응표본 차이검증에서는 상식과 빠진 곳 찾기로 이루어진 단축형을 제외한 7 가지 단축형에서 통계적으로 의미 있는 차이가 관찰되지 않아 Rabin과 Resnick-Entin의 두 번째 기준을 충족시키는 것이다. 상식과 빠진 곳 찾기 척도는 전체척도(Full Scale), 언어성 척도(Verbal Scale)와 동작성 척도(Performance Scale)와 높은 상관을 나타내는 소검사이다. 이 조합은 실시에 소요되는 시간이 짧아 피험자의 동기나 피로와 관련된 문제를 피할 수 있는 장점이 있지만(Kaufman, Ishikuma, & Kaufman-Packer, 1991) 척도의 개수가 두 개에 불과해 점수의 안정성을 확보하는데 어려움이 있다. 이 조합으로 추정된 지능지수와 전체지능 간 의미 있는 차이가 나타난 본 연구결과는 이러한 점수의 불안정성에서 기인하는 것일 수 있다. 어휘-토막 짜기 조합의 경우에도 두 개의 척도로 구성되어 있지만 본 연구에서 두 점수분포 간 의미 있는 차이가 관찰되지 않았는데, 이 조합의 경우 두 척도가 전체 지능지수와 상관이 보다 높고 일반 능력 g 요인을 가장 잘 설명하는 척도라는 점에서(Herrera-Graf, Dipert, & Hinton, 1996) 차이가 있다. 물론, 이 경우에도 전체 지능지수와 상관이 상대적으로 다소 낮아 점수가 안정적이지 못함을 알 수 있다.

끝으로, 두 지능지수 간 진단분류의 일치도를 검토한 결과 8개 단축형에서 roh 계수로는 .80~.92의 높은 일치도를 보였지만 정확분류 사례/전체 사례로 일치도를 계산했을 때는 .50~.71의 일치율을 기록하여 임상장면에서 실제 사용과정에서는 진단범주의 불일치가 상당한 정도로 나타날 수 있을 것으로 예상된다. 연속변인으로 가정되는 지능지수를 10점 간격의 불연속적 범주변인으로 변환하는 것은 임

상적 목적상 필요에 의해 채택한 것이기는 하지만 기본적으로 심리측정적 정밀성을 포기하는 것이다. 전체 적도 지능지수와 단축형 추정 지능지수 간의 상관이 완벽하고 두 점수분포의 차이가 전혀 없이 완전히 일치하는 경우가 아닌 한 무시할 수 없는 불일치가 존재할 수 있다. 따라서, 이 기준은 임상적으로 매우 중요하기는 하지만 심리측정적인 측면에서는 근원적으로 충족시키기 어려운 기준이다. 실제로 단축형의 타당도를 검증한 GoH(1980), Mumpower(1964), Sliverstein(1982, 1984, 1985) 등의 여러 선행 연구에서도 진단분류에서 불일치가 상당히 큰 것으로 나타나 본 연구결과와 일치하고 있다.

이러한 결과는 Rabin(1943)과 Resnick - Entin (1971)이 제시한 세 가지 기준에 따를 때 기준 1)은 모든 단축형에서 충족되고 있고, 기준 2)는 일반지식과 토막짜기로 이루어진 단축형을 제외한 7 가지 단축형에서 충족되고 있다. 기준 3)의 경우 두 지능지수의 진단분류의 상관은 매우 높지만 실제 사례들 중 약 30~50%에서 진단분류의 불일치가 있을 것으로 예상되어 기준을 충분히 충족시키지는 못하고 있다. 세 번째 기준은 별개로 하고 볼 때 이상과 같은 본 연구의 결과는 연구에서 타당도를 검증하고자 한 8 가지 단축형 중 상식-빠진 곳 찾기로 이루어진 조합을 제외한 7 가지 단축형이 타당함을 시사한다.

웍슬러 지능검사의 단축형은 전체 검사가 제공하는 다양한 정보의 일부를 포기하는 것인 만큼 원칙적으로는 권장할 만한 절차는 아니다. 그러나 임상장면에서 수검자의 상태나 특별한 상황 때문에 전체 검사를 실시하는 것이 어렵거나 불가능한 경우라면 일부 정보를 포기하더라도 사용해볼 가치는 있다. 본 연구

는 이러한 상황에서 단축형을 사용할 때의 타당성을 검토한 것이다. 본 연구결과에 따르면 Campbell(1998)이 제안한 단축형들 중 상식-빠진 곳 찾기 조합을 제외한 6 가지 조합과 K-WISC-III 기준자료의 요인분석에서 도출된 어휘-토막짜기-산수-기호쓰기 조합 단축형을 통해 추정된 전체지능이 상당한 정도의 타당성을 확보하고 있다.

본 연구에서 여러 가지 방식의 단축형에 대해 타당도를 검증하였으나 구체적인 임상장면에서 어떤 단축형을 사용하는 것이 가장 바람직할지의 문제는 다루지 못하였다. 각 단축형을 구성하고 있는 척도들의 속성과 각 단축형을 고안한 연구자들의 취지를 참조하여 임상가가 개인이 선택해야 할 문제이지만, 단축형에 포함되는 척도의 수가 많을수록 점수의 신뢰도가 증가하는 만큼 가능한 많은 척도들이 포함된 단축형을 사용하는 것이 바람직할 것으로 보인다.

본 연구에서 사례수가 충분하지 않아 정상 집단과 임상집단에 대해, 그리고 연령수준 별로 별도의 검증하지는 않았으나 추후 연구에서는 정교한 추정을 위해서는 보다 세분화된 집단을 대상으로 타당도 검증을 해볼 필요가 있다. 진단, 연령 등 하위 집단에 따라 최적의 회귀식이 달라질 수 있을 것이기 때문이다. 또한, 표준화집단의 자료의 회귀분석을 통해 공식적인 단축형 추정 IQ 산출공식을 제공할 필요가 있겠다.

참고문헌

곽금주, 박혜원, 김청택 (2001a). 한국 웨슬러 아동지능검사(K-WISC-III) 표준화를 위한

예비연구. 한국심리학회지: 발달, 14, 43-59.

곽금주, 박혜원, 김청택 (2001b). 한국 웨슬러 아동지능검사(K-WISC-III) 지침서. 서울: 특수교육.

김명권, 김중술 (1985). K-WISC의 단축형에 관한 연구. 한국심리학회, 85년도 학술발표 초록.

김중술, 김영환 (1974). KWIS의 단축형에 관한 연구-Doppelt형. 장병립 교수 회갑기념 논총, 1-11.

김중술, 이용승, 이민식 (1994). K-WAIS의 단축형에 관한 연구. 정신의학, 19(3), 121-126.

박혜원 (2001). 한국 웨슬러 유아지능검사의 간편형 개발. 아동학회지, 22(2), 1-13.

이용승, 김중술 (1995). K-WISC 단축형의 타당도 연구. 한국심리학회지: 임상, 14(1), 111-116.

이창우와 서봉연 (1974). 한국판 웨슬러 아동용 지능검사 실시요강. 서울: 배영사.

임영란, 이우경, 이원혜, 박종원 (2000). 한국 웨슬러 지능검사(KWIS) 단축형의 정확성 및 타당도에 관한 연구. 한국심리학회지: 임상, 19(3), 563-574.

Campbell, J. M. (1998). Internal and external validity of seven Wechsler Intelligence Scale for Children-Third edition short forms in a sample of psychiatric inpatients. *Psychological Assessment*, 10, 431-434.

Donders, J. (1997). A short form of the WISC-III for clinical use. *Psychological Assessment*, 9, 15-20.

Doppelt, J. E. (1956). Estimating the full scale score on the Wechsler Adult intelligence

- Scale from scores on four subtests. *Journal of consulting psychology*, 20, 63-66.
- Dumont, R., & Faro, C. (1993). A WISC-III short form for learning-disabled students. *Psychology in the Schools*, 30, 212-219.
- GoH, D. S. (1980). Note on selection of WISC-R short forms for different uses. *Journal of clinical psychology*, 36, 319-321.
- Herrera-Graf, M., Dipert, Z. J., & Hinton, R. N. (1996). Exploring the effective use of the Vocabulary/Block Design short form with a special school population. *Educational and Psychological Measurement*, 56, 522-528.
- Kaufman, A. S. (1972). A short form of the Wechsler Preschool and Primary Scale of Intelligence. *Journal of Consulting and Clinical Psychology*, 39, 361-369.
- Kaufman, A. S. (1975). Factor analysis of the WISC-R at eleven age levels between 6 1/2 and 16 1/2 years. *Journal of Consulting and Clinical Psychology*, 43, 135-147.
- Kaufman, A. S. (1990). Evaluation of Kaufman's short forms of the WAIS-R with psychiatric inpatients. *Journal of Clinical Psychology*, 48(2), 239-245
- Kaufman, A. S., Ishikuma, T., & Kaufman-Packer, J. L. (1991). Amazingly short forms of the WAIS-R. *Journal of Psychoeducational Assessment*, 9, 4-15.
- Kaufman, A. S., & Kaufman, J. C., Balgopal, R., & McLean, J. E. (1996). Comparison of three WISC-III short forms: Weighing psychometric, clinical, and practical factors. *Journal of Clinical Child Psychology*, 25, 97-105.
- Levy, P. (1968). Short-form tests: A methodological review. *Psychological Bulletin*, 69, 410-416.
- Mumpower, D. L. (1964). The fallacy of the short form. *Journal of Clinical Psychology*, 20, 111-113.
- Rabin, A. I. (1943). A short form of the Wechsler-Bellevue test. *Journal of Applied Psychology*, 27, 320-324.
- 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.
- Ryan, J. J., Georgemiller, R., & McKinney, R. (1984). Application of the four-subtest WAIS-R short form with an older clinical sample. *Journal of clinical psychology*. 40, 1033-1036.
- Ryan, J. J., Larsen, J., & Prifitera, A. (1983). Validity of two-and four-subtest short forms of the WAIS-R in a psychiatric sample. *Journal of Consulting and Clinical Psychology*, 51, 460.
- Ryan, J. J., Lopez, S. J., & Werth, T. R. (1998). Administration time estimates for WAIS-III Subtests, scales, and short forms. *Journal of Psychoeducational Assessment*, 16, 315-323.
- Satz, P., & Mogel, S. (1962). Abbreviation of the WAIS for clinical use. *Journal of clinical psychology*, 18, 77-79.
- Silverstein, A. B. (1967). A short form of the WISC and WAIS for screening purposes. *psychological reports*, 20, 682.
- Silverstein, A. B. (1974). A short-short form of the WISC-R for screening purposes. *Psychological Reports*, 35, 817-818.

- Silverstein, A. B. (1982). Factor structure of the wechsler adult intelligence scale-revised. *Journal of consulting and clinical psychology, 50*, 661-664.
- Silverstein, A. B. (1984). Standard errors for short forms of wechsler's intelligence scales with deviant subjects. *Journal of consulting and clinical psychology, 52*, 913-914.
- Silverstein, A. B. (1985). An appraisal of three criteria for evaluating the usefulness of WAIS-R short forms. *Journal of clinical psychology, 41*, 676-680.
- Silverstein, A. B. (1990). Short forms of individual intelligence tests. *Psychological Assessment, 2*, 3-11.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wolfson, W., & Bachelis, L. (1960). An abbreviated form of the WAIS Verbal Scale. *Journal of Clinical Psychology, 16*, 21.
- Yudin, L. W. (1966). An abbreviated form of the WISC for use with emotionally disturbed children. *Journal of consulting psychology, 30*, 272-275.
- 원고접수일 : 2007. 12. 27.
게재결정일 : 2008. 1. 21.

Validity of the Korean Wechsler Intelligence Scale for Children-III Short forms

Young Soon Jun

Soon-Taek Hwang

Suk-Hi Lee

Department of Psychology, Chungbuk National University

This study was conducted to examine the validity of the short forms of the Korean Wechsler Intelligence Scale for Children-Third Edition(K-WISC-III; Kwak et al., 2001). K-WISC-III data from 294 normal and clinical samples aged 6 ~ 16 were used to examine eight short forms of Doppelt type short forms of the K-WISC-III: seven from Campbell(1998)'s study and one from K-WISC-III factor analysis. Data from 294 subjects were randomly divided into two groups. The first group was used to make regression equations of the eight short forms to estimate IQs. The second group was used to compare the estimated IQs of the eight short forms with IQs determined using the full scale. There were high levels of correlations between estimated IQ and full scale IQ. Mean differences between two IQs were not significant, except for the "Information+Picture Completion" short form. Spearman's Rho correlations between the two IQ categories were very high. The results demonstrated the validity of the short forms of the K-WISC-III.

Key words : Intelligence scale, Short form, Validity, K-WISC-III