

## K-WAIS 단축형의 타당도 연구

이 용 승

김 중 술

서울대학교병원  
신경정신과

서울대학교 의과대학  
정신과학 교실

본 연구에서는 김중술, 이용승 및 이민식(1994)이 제시한 K-WAIS의 몇가지 소검사를 통한 추정 IQ가 정신과 환자 집단에도 타당하게 적용될 수 있는지를 검토하고자 하였다. K-WAIS의 11가지 소검사 중에서 두가지 소검사(어휘문제, 토막짜기)의 환산점수의 합을 구하거나, 네가지 소검사(어휘문제, 산수문제, 토막짜기, 차례맞추기)의 환산 점수의 합을 구하여 전체 IQ를 추정하게 되는데, 본 연구에 포함된 정신과 환자 집단의 경우, 각각의 추정치는 FSIQ와 .92, .97의 높은 상관을 보였다. 두번째로 두가지 소검사를 통한 추정 IQ는 FSIQ와 유의미한 차이를 보인 반면에, 네가지 소검사를 통한 추정 IQ는 FSIQ와 유의미한 차이를 보이지 않았다. 세번째로 WAIS-R 단축형 연구들에서와 유사하게, 각 추정 IQ를 7개의 지능 하위 범주(최우수, 우수, 평균상, 평균, 평균하, 경계선, 정신지체)로 나누어 FSIQ의 범주화와 비교했을 때, 두가지 소검사를 통한 추정치는 58%, 네가지 소검사를 통한 추정치는 72%의 일치율을 보였다. 이는 두가지 혹은 네가지 소검사의 추정치를 사용하여 하위 범주로 분류하거나, 정확한 IQ 수치를 산출할 때에는 신중할 필요가 있다는 것을 시사한다. 결론적으로 두가지 혹은 네가지 소검사를 통해 전체 IQ를 매우 신뢰롭게 추정할 수 있지만, 정확한 IQ가 요구될 때에는 단축형을 통한 추정 IQ를 신중하게 사용해야 할 것이다.

김중술, 이용승 및 이민식(1994)은 K-WAIS의 11가지 소검사 중에서 각각 두가지 소검사와 네가지 소검사를 선택하여, 이들 환산 점수의 합을 통해 전체 IQ를 추정하는 방법을 제시하였다. 전체 지능 검사를 실시하는데에는 시간이 많이 걸리기 때문에, 몇가지 소검사의 조합으로 이루어진 단축형 검사로 한 개인의 전체 IQ를 타당하게 추정할 수 있다면 매우 유용

한 시도라고 볼 수 있다(Silverstein, 1982).

단축형은 두가지 소검사를 통해 전체 IQ를 추정하는 방법과 네가지 소검사를 통해 전체 IQ를 추정하는 방법 두가지가 있는데, 두가지 소검사 단축형은 어휘문제와 토막짜기의 조합으로, 그리고 네가지 소검사 단축형은 어휘문제, 산수문제, 토막짜기, 차례맞추기의 조합으로 구성되었다. 이들 소검사들의 환산점

수의 합으로 전체 IQ를 추정하게 되는데, 김중술 등(1994)의 연구에서는 소검사들의 환산점수의 합을 통한 전체 IQ 추정치와 FSIQ 간에 두가지 소검사의 경우는 .91, 네가지 소검사의 경우는 .95의 매우 높은 상관을 보였다. 본 연구에서는 이러한 두가지 혹은 네가지 소검사를 통한 IQ 추정치가 정신과 환자 집단에도 타당하게 적용될 수 있는 지를 평가해 보려고 하였다.

단축형 지능 검사의 임상적인 타당도의 기준은 Resnick과 Entin(1971)이 제시하였는데, 그들은 첫째, 단축형으로 추정한 IQ와 FSIQ 간의 상관이 유의미하게 높아야 하고, 둘째, 단축형 IQ의 평균과 FSIQ의 평균 간에는 유의미한 차이가 없어야 하며, 셋째, 피험자들을 7개의 지능 수준으로 범주화(최우수, 우수, 평균상, 평균, 평균하, 경계선, 정신지체 수준)했을 때, 단축형을 통한 범주화와 FSIQ를 통한 범주화 간에 큰 차이가 없어야 한다고 주장하였다. 그러므로 본 연구에서는 정신과 환자 집단의 경우 단축형을 통한 추정 IQ가 이러한 세가지 기준에 얼마나 부합하는 지를 평가하려고 한다.

## 연구방법

1994년 1월에서 8월까지 서울대학교병원 신경정신과 임상심리실에 심리검사가 의뢰된 피험자 72명을 대상으로 하여 자료가 분석되었다. 연령별로는 16세에서 63세까지의 환자가 포함되었으며, 평균 연령은 30.79( $SD=10.15$ )세였다. 이 중 남자는 54명, 여자는 18명이 포함되었으며, 대학 재학 이상이 35명, 고졸 19명, 중졸 9명, 국졸 8명, 국퇴이하 1명의 교육 수준을 보였다. 진단별로 보면, 정신분열증 10명, 조울증 1명, 성격장애 11명, 신경증 26명, 뇌손상 환자를 포함하여 교통사고 감정으로 의뢰된 환자 17명, 기타 7명이었다. 이들은 이미 전체 지능 검사가 실시된 상태였으며, 각 단축형으로 추정한 전체 IQ는 김중술 등(1994)이 제시한 표로 산출하였다.

## 결 과

전체 지능 검사를 통한 FSIQ, 두가지 소검사로 추정한 전체 IQ, 네가지 소검사로 추정한 전체 IQ는 각각 100.06( $SD=16.59$ ), 102.57( $SD=16.08$ ), 99.81( $SD=16.64$ )의 평균과 표준편차를 보였으며, 두가지 소검사 추정 IQ와 FSIQ의 상관은 .92, 네가지 소검사 추정 IQ와 FSIQ의 상관은 .97로 매우 높은 상관을 보였다. 이는 Resnick 등(1971)이 제시한 첫번째 기준을 만족한다.

다음으로 두가지 소검사 추정 IQ와 FSIQ, 네가지 소검사 추정 IQ와 FSIQ간의 t-검증을 각각 해보았는데, 두가지 소검사를 통한 추정치는 FSIQ와 유의미한 차이를 보인 반면에( $t(72)=3.29, p<.05$ ), 네가지 소검사를 통한 추정치의 경우는 유의미한 차이를 보이지 않았다( $t(72)=-.49$ ). 이는 네가지 소검사 단축형은 FSIQ와 평균에서 차이가 나지 않았지만, 두가지 소검사 단축형의 평균은 FSIQ보다 다소 높음을 의미한다. 그러므로 두가지 소검사 단축형의 경우는 Resnick 등(1971)이 제시한 두번째 기준을 만족시키지 못했는데, 이는 Ryan, Larsen 및 Pfriftera(1983)의 WAIS-R 단축형 연구와도 같은 결과이다. Silverstein(1985)이 개관한 소검사를 통한 단축형 지능 연구들에서도 8개 연구 중에서 4개 연구가 이 두번째 기준을 만족시키지 못하고 추정 IQ와 FSIQ 간에 유의미한 차이를 보였다.

세번째로 7개의 지능 하위 수준으로 범주화했을 때 두가지 소검사를 통한 추정 IQ와 FSIQ, 네가지 소검사를 통한 추정 IQ와 FSIQ 간의 일치 정도를 보면, 두가지 소검사 추정치의 경우는 58%였고, 네가지 소검사 추정치의 경우는 72%의 일치율을 보였다. 네가지 소검사를 통한 추정치에 비해 특히 두가지 소검사를 통한 추정치의 일치율이 비교적 낮다고 볼 수 있는데, WAIS-R 단축형에 관한 연구들에서도 일치율은 50-77%로 비교적 낮은 편이었다(Silverstein, 1985). 그러므로 추정 IQ를 통해 지능 수준을 범주화하거나, 정확한 IQ 측정이 요구되는 상황에서는 매우 조심스럽게 단축형을 통한 추정 IQ를 사용해야 할 것

이다.

다음으로는 진단 하위집단별로 단축형 IQ와 FSIQ 간의 상관을 보았는데, 첫번째로 정신병집단(정신분열증 10명, 조울증 1명)의 경우는, 네가지 소검사를 통한 추정 IQ와 FSIQ 간의 상관은 .92였고, 두가지 소검사를 통한 추정 IQ와 FSIQ는 .91의 상관을 보였다. 신경증집단(26명)의 경우는 각각 .97, .94의 상관을 보였으며, 뇌손상 환자를 포함한 교통사고감정 환자집단(17명)은 .96, .88, 성격장애 환자집단(11명)은 .96, .89의 상관을 보였다.

이러한 결과를 보전대, 네가지 소검사를 사용한 단축형은 .92에서 .97까지 높은 상관을 보여, 임상집단의 하위 유형중에서 어떤 집단에서도 매우 유용하게 전체 IQ를 추정한다고 볼 수 있다. 상대적으로 정신병 집단의 경우 네가지 소검사를 통한 추정 IQ와 FSIQ는 .92로 다소 낮은 상관을 보이는데, 이는 현재 정신병으로 입원한 환자들이 주의집중의 곤란을 호소하고, 논리적인 추리력 등에 장애가 있는 경우가 많기 때문에 산수문제와 차례맞추기의 수행이 다른 소검사들에 비해 다소 저조할 수 있다.

이러한 수행점수에 따르면 정신병 집단의 경우, 어휘문제와 차례맞추기 두가지 소검사에 더하여 산수문제와 차례맞추기를 조합한 네가지 단축형이 두가지 단축형에 비해 FSIQ와 더 높은 상관을 보인다고 기대하기 어려울 것이다. 실제로 산수문제와 차례맞추기를 제외한 두가지 소검사 단축형의 상관은 .91로, 다른 하위 집단과는 달리 추정 IQ와 FSIQ 간의 상관에 있어 네가지 소검사 단축형과 거의 차이가 나지 않았다. 하지만 진단 하위집단에 상관없이 두가지 소검사를 통한 추정 IQ는 네가지 소검사 단축형을 통한 추정 IQ에 비해 FSIQ와 낮은 상관을 보였다(.88에서 .94).

Ryan(1983)은 정확한 IQ 추정을 하는데 단축형을 사용하기는 어렵지만, 보통 지능 집단(90이상)과 보통이하 지능 집단(89이하)을 구분하는 선별 장치로서는 유용하다는 결론을 내렸다. 본 연구에 포함된 임상 집단에서는, 90이상의 FSIQ를 가진 54명 중에서 두가지와 네가지 소검사 단축형 모두 52명(96%)을 90

이상이라고 분류하였으며, 단지 2명(4%) 만을 90이하라고 분류하였다. 또한 89이하의 FSIQ를 가진 18명 중에서 두가지 소검사 단축형은 12명(67%), 네가지 소검사 단축형은 15명(83%)을 89이하라고 올바르게 분류하였으며, 각각 6명(33%), 3명(17%)을 잘못 분류하였다. 그러므로 두가지와 네가지 소검사를 통한 추정 IQ 모두, 보통 집단과 보통이하 집단을 매우 정확하게 분류한다고 볼 수 있다.

Ryan(1983)은 두가지 소검사를 통한 단축형 IQ로 보통 지능 집단과 보통이하 지능 집단을 구분하였는데, 보통 지능 집단에서는 89%, 보통이하 지능 집단에서는 83%의 정확한 분류 일치율을 보였다. 보통 지능 집단의 경우는 본 연구가 더 높은 분류 일치율을 보이고 있는 반면에(96% 대 89%), 보통이하 지능 집단의 경우는 Ryan(1983)의 연구가 더 높은 분류 일치율을 보이고 있다(67% 대 83%).

## 논 의

두가지 소검사를 통한 추정 IQ는 FSIQ와 높은 상관을 보여 Resnick 등(1971)이 제시한 첫번째 기준에는 부합하였으나, 두번째 기준과 세번째 기준을 만족시키지 못했다. 즉, 두가지 소검사를 통한 추정 IQ와 FSIQ는 유의미한 차이를 보였으며, 7개의 범주화를 했을 때 일치율은 58%로 저조하였다. 네가지 소검사를 통한 추정 IQ와 FSIQ 역시 높은 상관을 보였으며, 네가지 소검사를 통한 추정 IQ와 FSIQ 간의  $t$ 검증에서도 유의미한 차이를 보이지 않았다.

따라서 네가지 소검사를 통한 추정 IQ는 Resnick 등(1971)이 제시한 첫번째 기준과 두번째 기준에 부합하는 결과를 보였다. 하지만 범주화를 했을 때의 일치율은 72%로 두가지 소검사 단축형에 비해서는 다소 높은 편이지만 Resnick 등(1971)이 제시한 세번째 기준을 만족한다고 보기는 어렵다. 그러므로 단축형 IQ를 통해 정확한 IQ 추정을 하는 것은 다소 어렵다고 볼 수 있으며, 이러한 용도로 지능 평가가 이루어질 때에는 신중하게 단축형을 사용할 필요가 있을 것

이다.

Resnick 등(1971)의 기준들 중에서 이 세번째 기준에 대해서는 Silverstein(1985)이 비판을 하고 있다. 그에 의하면 단축형을 통한 추정 IQ와 FSIQ가 거의 완벽한 상관을 보이는 경우에도 추정 IQ와 FSIQ 간의 일치율이 높아지지 않으며, 또한 지능 하위 수준으로 범주화할 때 IQ 점수 상에서 1점 차이가 나더라도 두 IQ간에 불일치를 보이는 경우가 있는 반면에, 19점 차이가 나더라도 일치하는 경우가 있어 임의적인 범주화 기준을 적용하기에는 무리가 있다고 주장하였다.

그러므로 이러한 주장을 고려한다면, 단축형을 통한 추정 IQ가 세 기준을 모두 만족시키기는 어렵다고 볼 수 있다. Ryan(1983)도 임상적인 관점에서 Resnick 등(1971)이 제시한 세번째 기준을 만족시키지 못하는 것이 커다란 결점이 될 수 없다고 주장하였는데, 우선은 단축형 지능이 일반적인 지능 평가에는 사용되지 않고 있으며, 보통 이상의 지능을 가진 사람과 보통 이하의 지능을 가진 사람을 변별하는데는 매우 유용하다는 것이다. 그에 의하면 90이상의 IQ를 보이는 환자는 지적인 기능이 손상되지 않았을 확률이 매우 높다고 결론내릴 수 있고, 89이하의 IQ를 보이는 경우에는 보다 적절한 진단을 내리기 위해 부가적인 인지 기능을 평가하는 검사를 할 필요가 있다는 것을 지적해준다는 것이다.

결론적으로, 두가지 소검사와 네가지 소검사 단축형은 FSIQ와 각각 .92, .97로 높은 상관을 보여 단축형을 통해 전체 IQ를 추정하는 것이 타당하다는 결론을 내릴 수 있겠다. 특히 네가지 소검사를 통한 추정 IQ는 정신병 환자집단이 .92로 다소 낮은 것을 제외하고는 FSIQ와의 상관이 .96에서 .97로 매우 높은 상관을 보이고 있다. 그러므로 두가지 소검사와 네가지 소검사를 통한 단축형 지능을 정신과 환자집단을 포함하는 다양한 임상 장면에서 매우 유용하게 사용할 수 있을 것이다. 다만 보다 신중하게 단축형 지능을 적용하려고 할 때에는 두가지 소검사 보다는 네가지 소검사를 통해 전체 IQ를 추정하는 것이 보다 안정적이라고 보겠다.

부가적으로 타당도 연구에 포함된 임상집단을 정신병 집단, 신경증집단, 뇌손상 환자를 포함한 교통사고 감정 환자집단, 성격장애 환자집단 등으로 나누어 분석해 보았는데, 하위 집단별로도 특별한 차이없이 단축형을 통한 추정 IQ가 FSIQ와 높은 상관을 보이고 있다(.88에서 .97). 그러므로 정신과 환자 하위 집단에 상관없이 전체 IQ를 추정하는데 단축형 지능을 유용하게 사용할 수 있을 것이다. 다만, 정신병 집단의 경우 네가지 소검사를 통한 단축형 지능의 추정치가 다른 집단에 비해 FSIQ가 다소 낮은 상관(.92)을 보여서, 다른 집단보다는 보다 신중하게 단축형을 통한 추정 IQ를 적용할 필요가 있다고 하겠다.

그밖에 고려할 점으로는 전체 지능 검사를 다 실시한 후에 몇가지 소검사를 선택하는 것과 전체 지능 검사를 다 실시하지 않고 몇가지 소검사를 실시하는 것이 차이가 있을 수 있다는 것이다. 본 연구는 이미 전체 지능 검사가 실시된 이후에 몇가지 소검사를 선택하여 전체 IQ를 추정하였는데, 정확한 상관을 보려면 단축형 지능에 포함된 소검사만을 실시하여 FSIQ와 상관을 보아야 할 것이다. 두가지나 네가지 소검사만을 실시하는 경우와 전체 소검사를 다 실시하는 경우 그 수행 상에 차이가 있을 수 있기 때문이다. 그러므로 앞으로는 이러한 점을 고려하여 축적한 자료를 통해 단축형 지능을 재검토 할 필요가 있다고 하겠다.

## 참고문헌

- 김중술 · 이용승 · 이민식. (1994). K-WAIS의 단축형에 관한 연구. *정신의학*, 19, 3, 121-126.
- 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. (1983). Clinical utility of a WAIS-R short form. *Journal of Clinical Psychology*, 39, 261-262.
- 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.
- Silverstein, A.B. (1982). Two- and four-subtest short forms of the wechsler adult intelligence scale-revised. *Journal of Consulting and Clinical Psychology*, 50, 415-418.
- Silverstein, A.B. (1985). An appraisal of three criteria for evaluating the usefulness of WAIS-R short forms. *Journal of Clinical Psychology*, 41, 5, 676-680.

## Validity of Short Forms of the Korean-Wechsler Adult Intelligence Scale

**Yong-Seung Lee**

Depart. of Neuropsychiatry  
Seoul National University Hospital

**Zoung-Soul Kim**

Depart. of Psychiatry, College of Medicine  
Seoul National University

The validity of the two- and four-subset short forms of the Korean-Wechsler Adult Intelligence Scale was investigated using a sample of 72 psychiatric patients. Both short forms correlated highly with the Full Scale IQ (.92 and .97). However, pairwise t-tests computed between FSIQ and each short form indicated that a two-subtest short form yielded a significantly higher IQ, whereas a four-subtest short form did not significantly differ from the FSIQ. Finally, the rate of agreement with the Full Scale IQ, in terms of Wechsler's seven-category classification, was relatively low (58% and 72%) for each short form. It is concluded that the two- and four-subtest short forms should be used with caution in situations where a precise IQ measurement is required.