

정상 영아 및 장애 위험 영아의 한국 Bayley 영유아발달검사(K-BSID-II) 수행 비교 : 3년 종단자료분석*

K-BSID-II Performance in Normal and High Risk Infants :
A Three Year Longitudinal Data Analysis*

박혜원(Hyewon Park Choi)¹⁾

ABSTRACT

Stability of the Korean Bayley Scale of Infant Development-II was tested in 305 infants by two measurements of varying intervals over a span of 5-36 months. Stability of K-BSID-II performance was $r=.62(p<.01)$, $r=.08(p<.18)$, $r=.69(p<.01)$, $r=.18(p<.01)$ for mental scale raw scores, mental scale index scores (MDI), psycho-motor scale raw scores and psycho-motor scale index scores(PDI), respectively. Stability was higher for the shorter test-retest interval group. Correlations between raw scores among infants with retest intervals of more than 2 years were stable on both mental and psycho-motor scales. MDI showed higher stability among high-risk infants than normal infants while PDI showed higher stability among normal infants. Testers and researchers should use both raw scores and index scores for better interpretations.

Key Words : 한국 Bayley 영유아 발달검사(K-BSID-II), 검사 안정성(Test-retest stability), 1-42개월 영유아(1-42 month olds), 3년 종단연구, 인지발달지수(MDI), 동작발달지수(PDI), 정상(normal), 장애위험(high risk).

I. 서론

의학과 복지의 향상으로 인간의 전 생애동안 진단(assessment)과 중재(mediation)의 역할은 점차 중요해 지고 있다. 국내에서는 특수교육법의

개정으로 학령기 아동의 진단과 중재가 활발해졌고 최근에는 학령 전 아동을 대상으로 하는 조기 진단과 교육이 사회의 역할이 되었다. 조기 교육과 진단에 대한 요구는 특수 아동뿐 아니라 일반 아동에게도 보편화되고 있다. 인적 자원과

* 이 논문은 2004년 한국학술진흥재단의 지원으로 연구되었음(KRF-2004-002-C00199).

¹⁾ 울산대학교 아동가정복지학과 교수

Corresponding Author : Hyewon Park Choi, Department of Child and Family Welfare, University of Ulsan, Ulsan, 680-749, Korea

E-mail : hyewonc@mail.ulsan.ac.kr

여성 인력의 중요성이 대두되면서 유아교육과 보육분야에서 대상 연령을 하향하는 추세이다. 교육부는 1992년부터 제 5차 유치원 교육과정에 3세를 포함하고 있고 보건복지부도 1991년 영유아보육법을 제정하여 출생 직후부터 전문적인 보육과 교육을 하기 위한 노력을 경주하고 있다. 이는 1960년대 이후 본격적으로 시작된 영아에 대한 연구 결과 영유아기는 인간이 살아가는데 필요한 기본적인 능력과 기술을 발달시키는 중요한 시기로 밝혀지고 있는 사실에서 기인한다. 적절한 연구 방법과 측정 도구를 통해 영아가 선천적인 지각 능력과 기억 등의 학습능력을 가지고 태어난다는 것이 밝혀지고 있는 것이다(Snow & McGaha, 2003).

장애에 대한 중재나 발달에 적합한 교육을 위해서는 올바른 진단이 선행되어야 한다. 이를 위해서는 영유아 발달에 대한 이해를 기초로 영유아의 발달을 진단할 수 있는 도구가 시급히 개발되어야 한다. 현재 국내에서는 영유아 발달을 진단하는 도구들이 몇 가지 사용되고 있으나 장애아 발달을 진단하는 검사이거나(강순구·조윤경, 1994; 김승국·김옥기, 1985), 간편용(김숙현·이경옥, 2001; 대한소아과학회, 2002)으로 개발되어 있을 뿐이어서 우리나라 영유아의 발달에 기초한 기준을 제시하는 발달검사가 필요한 실정이다.

영유아발달검사는 주로 미국에서 개발되어 왔는데 간편한 선별검사인 DDST, Brazelton 검사 등이 있고, 세계적으로 가장 많이 사용되며 발달지수까지 산출할 수 있는 것으로는 Bayley 영아발달검사(Bayley Infant Scale of Development-BSID)(Bayley, 1969)가 있다. 1930년대부터 개발(Bayley, 1932, 1933a, b; 1935; 1936)이 시작된 BSID는 1969년 미국에서 표준화된 이래 서구에서 뿐 아니라, 일본, 중국 등 동남아에서도 표준

화되거나 번안된 검사가 오래 전부터 사용되고 있다(Aggarwal, Chaudhari, Bhavé, Pandit, Barve, 1998; Brooks-Gunn, 1998). 미국에서는 BSID 출판 이후 20여 년간의 영아연구를 통해 밝혀진 사항을 고려하고 시대에 맞는 새 기준을 제시하기 위해 1993년에 2판인 Bayley Scale of Infant Development-II가 제작되었다(Bayley, 1993; Campbell, Siegel, Parr, & Ramey, 1986). 2판은 그 대상연령이 2-30개월용이던 1판과 달리 1-42개월용으로 확장되었으며 심리측정적인 특성에 대한 보다 자세한 검증이 이루어졌다(Baird, & Folsom, 1993; Black, & Matula, 1999; Goldstein, Fogle, Wieber, & O'Shea, 1995; Johnson, Cook, & Kullman, 1992; McChesne, 1995; Nellis, & Gridley, 1994).

Bayley 영아발달검사에 대한 국내 연구를 살펴보면 1980년대 이후 꾸준히 척도화를 위한 탐색 연구와 타당도에 대한 연구가 수행되었다(민동욱·박혜원, 2003; 배윤희, 2001; 이일섭, 1992; 정문자, 1983; 정문자·이은혜·박경자, 1993; 정영미, 1992; 제경숙, 1984). 이러한 연구들은 Bayley 검사의 유용성에 대해 보고하고 있으나 한국 영유아를 위한 기준을 제작하지는 못하였다. 조복희·박혜원(2004)은 예비연구(박혜원·조복희·최호정, 2003)를 통해 개발된 한국 Bayley 영유아발달검사(Korean Bayley Scale of Infant Development-II, K-BSID-II)를 사용하여 정상분만한 1,700명의 영유아를 대상으로 표준화 연구를 실시하였다. 이 연구에서는 한국 기준을 제작하기에 앞서 미국 기준을 사용하여 한국 영유아의 수행을 보고하였는데 한국 영유아는 인지발달 척도에서는 미국 영유아의 수행과 유사하였으나 동작발달 척도에서는 미국 영유아에 비해 유의하게 높은 수행을 나타냄으로써 한국 기준 마련의 필요성을 보고하였다.

그런데 Bayley 영아발달검사 등의 영아발달검사가 보다 유용하게 활용되기 위해서는 검사 결과의 안정성에 대한 점검이 필요하다. 일반적으로 한 시점에 측정된 지능검사의 결과가 이후의 검사 수행과 얼마나 일치하는가 즉 안정성에 관한 관심은 매우 높으며 특히 어린 영아의 수행 안정성은 논쟁의 대상이 되어왔다(Bayley, 1970; Rubin, 1996; Sabourin, 1996). 영아기 발달에 관한 체계적인 종단 연구를 수행한 Bayley와 그녀의 동료들은 1929년에 시작하여 30년 이상 이루어진 종단적 연구(Bayley, 1970)에서 영아기 수행의 안정성이 만족스럽지 못하다고 밝혔다.

이후의 안정성 연구에서도 유사한 몇 가지 결론이 도출되었다(Jencks, 1972). 첫째, 극단적으로 낮은 점수를 제외하고는, 영아검사에서의 수행은 성장후의 IQ를 예언하지 못한다. 즉 영아기 수행 점수와 그 이후의 수행 점수간의 상관은 “0”에 가깝다. 둘째, 영아기 이후의 수행점수들은 서로 유의한 상관을 보인다. 셋째, 일반적으로 IQ의 안정성은 비교되는 수행 간 간격이 짧을수록 높고 또한 처음 검사 시의 월령이 높을수록 높다. 따라서 안정성은 연령과 함께 증가한다고 볼 수 있다.

그런데 이렇듯 영아기에 측정된 지능수행이 이후의 발달과 유의한 상관을 보이지 않았던 이유의 하나는 실제로 영아기의 지적 능력이 안정적이지 않기 때문일 수 있다. 그러나 다른 한편으로는 영아기에 측정하는 행동과 영아기 이후에 측정하는 행동 특성 간에 차이가 있기 때문일 수 있다. 즉 검사의 내용자체가 상당히 달라지기 때문에 검사 수행 간 상관이 낮을 수 있다(Miller, 1987). 이는 실제로 영아기의 지적 능력은 이후의 능력과 상관이 있음에도 불구하고 이를 측정하는 도구 또는 방법의 문제 때문에 안정성이 낮게 산출될 수 있음을 시사한다.

1993년에 표준화된 BSID-II는 1960년 이후 활

발히 연구되어 온 영아기 인지능력 즉 재인기억, 신기성 선호(Brody, 1992; McChesne, 1995) 등을 측정하여 안정성을 증가시키기 위해 노력하였다. Banks, & Ginsburg(1985), Brody(1992), Fantz & Fagan(1975) 등의 인지발달 연구자들은 아동이 복잡한 시각자극을 선호하는 정도는 신경학적인 성숙도와 비례하며 이러한 경향은 후의 수행을 예언할 수 있음을 보고하였다. 또한 사고, 추론, 단기 기억과 같은 정보처리 및 문제해결 능력이 어린 영아에게도 찾아볼 수 있음이 밝혀졌다. Gelman & Gallistel(1986)은 수세기의 발달은 세부적인 하위단계로 나누어지는데 초기 단계는 영아기에 획득된다는 것을 보고하였다. 따라서 이러한 성장 후 지적 수행과 밀접한 관계가 있는 측면을 측정하는 문항이 개발되어 BSID-II에 첨가되었다(Black, & Matula, 1999).

본 연구의 목적은 BSID-II를 번안하고 한국 영유아에 맞도록 수정한 K-BSID-II의 안정성을 살펴보고 이러한 안정성이 검사 연령이나 검사 간격, 그리고 1차 검사에서의 수행수준(장애위험 여부)에 따라 어떻게 달라지는지 분석하였다. 이는 K-BSID-II의 유용성을 밝히는데 중요한 역할을 할 수 있을 것이다. 연구문제는 다음과 같다.
<연구문제 1> 3년에 걸친 K-BSID-II 수행의 안정성은 어떠한가?

<연구문제 2> K-BSID-II 수행의 안정성은 검사 연령, 검사 간 간격, 장애위험에 따라 차이가 있는가?

II. 연구방법

1. 연구대상

본 종단 연구는 한국 Bayley 영유아발달검사

표준화(조복희·박혜원, 2004)연구 과정에 참여한 기준 집단과 타당도연구 집단(정상 기준 집단 아동 1700명, 장애 및 미숙아집단 등 타당도 연구참여 아동 124명)의 1,824명의 영아 중 2차 검사안내에 따라 재검사를 받은 아동 305명을 대상으로 이루어졌다. 이중 남아는 158명(51.5%), 여아는 147(48.5%)명이었다. 본 종단연구 대상 표집시 정상 기준집단 영유아 외에 뇌성마비 등 장애집단이나 미숙아등 장애위험집단(124명)을 대상으로 2차 검사안내를 하였으나 참여율이 극히 낮았고(총 16명), 이 중에는 1차 수행 시 원점수가 0인 경우가 있어 상관분석이 불가능하여 분석에서 제외하였다.

일반적으로 BSID-II의 경우 각 척도별 수행은 <표 1>과 같이 85-115점 사이는 정상범주, 115이상은 우수 그리고 85-70사이는 경도지체, 70이하는 중도지체로 구분하고 있다(Bayley, 1993). <표 1>에는 K-BSID-II의 표준화연구에서 실제 밝혀진 분포와 분류방식을 제시하였다. 따라서 본 연구에서는 1차 수행에서 인지지수나 동작지수가 85점미만의 수행을 나타낸 아동을 장애위험집단으로 정의하여 이 집단아동들의 수행을 정상 집단의 자료와 비교하였다. <표 2>에는 1, 2차 검사 시 아동의 월령별 분포를 제시하였다.

<표 1> K-BSID-II의 수행범주 분류 (박혜원·조복희, 2005)

지수범위	범주	이론적 정상 분포(%)	실제 분포(%)	
			인지척도	동작척도
115 이상	우수	16.0	17.1	18.1
85-114	정상	68.0	69.2	70.7
70-84	경도지체	13.5	11.1	8.9
69 이하	중도지체	2.5	2.6	2.3

<표 2> 1차 검사 및 2차 검사시 월령별 참여아동수

검사시 월령	1차 검사 참여아동수(%)	2차 검사 참여아동수(%)
6개월 이하	166(54.4)	3(1.0)
6개월~12개월	58(19.0)	55(18.0)
13개월~24개월	63(20.7)	104(34.1)
25개월~42개월	18(5.7)	144(49.9)
총합계	305(100.0)	305(100.0)

2. 연구도구

K-BSID-II는 인지척도(Mental Scale), 동작척도(Motor Scale), 행동평정척도(BRS : Behavior Rating Scale) 세 가지 척도로 구성되어 있는데 인지척도는 179문항, 동작척도는 112문항, 행동평정척도는 30문항으로 구성되어 있다. 행동평정척도의 경우 월령집단에 따라 측정하는 하위척도가 달라져 종단적 분석에는 적절하지 않아 본 종단분석에서는 제외하였다. 인지 척도와 동작척도는 22개 연령집단별로 해당연령의 문항세트를 제시하도록 되어 있어 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14-16, 17-19, 20-22, 23-25, 26-28, 29-31, 32-34, 35-37, 38-42개월용으로 나뉜다. 예비 연구(박혜원 등, 2003)결과 인지척도의 채점자간 일치도는 $r=.99$, 동작척도의 경우는 $r=.99$ 였으며 표준화연구(조복희·박혜원, 2004)에서 보고된 것은 각각 $r=.91$, $r=.99$ 에 해당하였다. 또한 본 도구의 지침서에 따르면(박혜원·조복희, 2005), K-BSID-II와 한국 웨슬러 유아지능검사(K-WPPSI)와의 상관은 $r=.65$ (인지척도), $r=.52$ (동작척도) 등으로 공준 타당도 등도 검증되었다. 그 외에 도구에 대한 세부적인 설명은 박혜원·조복희·최호정(2003) 및 조복희·박혜원(2004)을 참조하라.

3. 연구절차 및 자료분석

종단연구를 위해 전국의 한국 베일리 영유아 발달검사의 표준화연구에 참여하였던 아동전체를 대상으로 우편으로 2차 검사에 대한 홍보를 실시하고 검사희망 접수를 받았다. 모든 대상 아동에게 1회의 안내문을 발송하였고 특히 3년에 걸친 종단 자료를 수집하기 위해 만 1세 이하의 참여아동에게는 추가로 1회의 안내문을 발송하여 참여를 독려했다. 희망자에 한해 실제 검사 장소 및 시간을 정하기 위해 전화로 다시 한 번 접촉하였다. 2차 검사를 실시하기 위해서 참여한 검사자는 총 9명이며 아동학 석사과정중의 학생으로 검사에 대한 이론적 지식, 실시방법 및 채점방법에 있어서 숙련된 검사자이다. 1차 검사 기간은 2002년 3월~2003년 2월까지이며 2차 검사기간은 2003년 3월~2005년 6월까지였다. 현재 K-BSID-II의 규준은 출판되지 않아 본 연구의 분석에 사용된 수행점수는 표준화연구와 마찬가지로 미국의 규준을 사용하였다. 수집된 자료는 SPSS/PC+ 12.K 프로그램을 활용하여 분석하였다.

III. 결 과

<표 3>은 성별에 따른 K-BSID-II 척도별 수행점수이다. 1차 검사에서 인지지수를 보면 남아는 98.49(8.07), 여아는 97.26(9.04)으로 남아가 높은 경향을 보이지만 그 차이는 유의미하지 않았다. 동작지수 분석에서는 남아가 98.66(12.67), 여아가 100.09(12.58)로 여아가 높은 경향이었지만 통계적으로 유의미하지는 않았다. 2차 검사에서는 인지지수와 동작지수 모두에서 여아가 96.97(12.03), 104.33(13.82)으로 남아의

<표 3> 성별에 따른 K-BSID-II 척도별 1, 2차 수행(평균(SD))

시기	척도	성별		t
		남(n=158)	여(n=147)	
1차	인지지수	98.49(8.07)	97.26(9.04)	1.34
	동작지수	98.66(12.67)	100.09(12.58)	-0.985
2차	인지지수	94.51(12.53)	96.97(12.03)	-1.76
	동작지수	101.68(12.61)	104.33(13.82)	-1.75

평균인 94.51(12.53)과 101.68(12.61)보다 높게 나타났으나 역시 성에 따른 유의미한 차이는 없었다. 따라서 추후 안정성 분석에서 성별에 따른 분석은 하지 않았다.

1. K-BSID-II 수행의 안정성

<표 4>는 1차 검사 시 아동의 월령에 따라 K-BSID-II 척도별 1, 2차시 수행을 나타낸 것이다. 전체 연령집단에서 1차 검사 시 인지지수는 97.93(SD=8.51)이며 동작지수는 99.41(SD=11.26)로 나타났다. 2차 검사의 경우도 인지지수 95.98(SD=12.07), 동작지수 103.25(SD=13.13)로 인지지수보다 동작지수에서 높은 양상을 보이고 있다. 이는 조복희·박혜원(2004)의 연구에서 밝힌 바와 같이 한국 영유아의 동작성 척도 수행이 미국의 영유아에 비해 높은 것과 일치하는 결과이다. 인지지수의 경우 2차 검사에서 1차 검사보다 수행이 낮아지는 경향을 보이는 반면 동작지수는 높아져서 유의한 차이가 있었다. 연령집단 별로 살펴볼 때 1차 검사의 인지지수를 제외하고는 연령별로 유의한 수행차이를 보이고 있으나 연령에 따른 일관적인 경향은 나타나지 않았다.

<표 5>에 제시된 바와 같이 1차 수행과 2차 수행간의 상관을 보면 인지척도 원점수간에는 $r=.62(p<.01)$, 동작척도 원점수간에는 $r=.69$

〈표 4〉 1차 검사 월령에 따른 K-BSID-II 척도별 1, 2차 수행

시기	척도	1-6개월	7~12개월	12~24개월	24-42개월	Total	F
		(n=166)	(n=58)	(n=63)	(n=18)	(N=305)	
		M(SD)	M(SD)	M(SD)	M(SD)	M(SD)	
1차	인지지수	98.17(8.05)	98.69(8.13)	97.43(9.59)	95.06(8.51)	97.93(8.51)	.957
	동작지수	96.60(9.43)	103.31(9.41)	102.11(14.05)	103.28(14.59)	99.41(11.26)	5.723**
2차	인지지수	96.84(11.78)	91.69(12.32)	97.19(12.85)	97.56(8.45)	95.98(12.07)	3.100*
	동작지수	101.73(12.36)	102.59(14.79)	106.19(12.71)	109.11(13.58)	103.25(13.13)	3.580*

* $p < .05$ ** $p < .01$

($p < .01$), 그리고 동작척도 지수점수 간에 $r = .18$ ($p < .01$)의 유의한 상관을 보였다. 그러나 인지척도의 지수 간에는 유의한 상관이 도출되지 못하였다($r = .08, p > .10$).

2. 검사연령, 검사간격 및 장애위험에 따른 K-BSID-II 수행의 안정성

Miller(1987)는 일반적으로 지적 수행의 안정성에 영향을 미치는 시간적 요인으로 검사 연령, 검사간 간격을 언급하였다. 일반적으로 검사 안정성(검사-재검사간 상관)은 처음 검사 시의 연령이 높을수록, 그리고 검사 간 간격이 짧을수록 커진다. 본 연구에서는 출생 직후부터 초기 3년

반에 걸쳐 검사 시의 연령과 1차, 2차 검사간의 간격에 따라 수행의 안정성이 어떻게 달라지는지 살펴보았다.

그 외에도 일반적으로 발달검사들은 정상수준의 아동보다 정신지체 등 발달장애가 있는 경우 예언력이 큰 것으로 보고되고 있다(Bayley, 1970). 본 연구에서는 1차 수행에서 인지나 동작 척도에서 85점미만의 수행을 나타낸 아동을 장애위험집단으로 정의하여 이들의 수행 안정성을 정상 아동의 수행안정성과 비교하였다.

1) 검사 연령에 따른 수행의 안정성

검사연령에 따라 안정성이 어떻게 달라지는지 살펴보기 위해 1차 검사 월령의 중앙값(6개월)을

〈표 5〉 K-BSID II 척도별 1, 2차 수행상관

(N=305)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.04						
1차 검사 동작원점수	.94**	-.02					
1차 검사 동작지수	.25**	.37**	.32**				
2차 검사 인지원점수	.62**	-.00	.66**	.15**			
2차 검사 인지지수	.05	.08	.02	.09	.22**		
2차 검사 동작원점수	.65**	-.02	.69**	.17**	.97**	.16**	
2차 검사 동작지수	.18**	.04	.16**	.18**	.21**	.50**	.34**

** $p < .01$

〈표 6〉 1차 검사시 월령이 6개월이하인 집단의 K-BSIDII 척도별 1, 2차 수행상관 (N=166)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.22**						
1차 검사 동작원점수	.73**	.11					
1차 검사 동작지수	.08	.37**	.25**				
2차 검사 인지원점수	.18**	.02	.27**	-.06			
2차 검사 인지지수	-.01	-.05	.00	.03	.08		
2차 검사 동작원점수	.25**	.02	.31**	-.06	.96**	-.01	
2차 검사 동작지수	.03	.02	-.06	.10	-.03	.45**	.12

**p<.01

〈표 7〉 1차 검사시 월령이 7개월이상인 아동의 K-BSIDII 척도별 1, 2차 수행상관 (N=139)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.16*						
1차 검사 동작원점수	.90**	-.02					
1차 검사 동작지수	.15*	.36**	.33**				
2차 검사 인지원점수	.53**	.04	.52**	.10			
2차 검사 인지지수	.33**	.20*	.30**	.14	.20**		
2차 검사 동작원점수	.51**	.01	.51**	.11	.96**	.60**	
2차 검사 동작지수	.20**	.07	.21**	.15	.44**	.58**	.60**

*p<.05 **p<.01

중심으로 2집단으로 구분하여 2회에 걸친 검사수행간의 상관을 분석한 결과를 <표 6, 7>에 제시하였다. 선행연구들(Jencks, 1972; McCall, 1983)과

마찬가지로 전반적으로 1차 검사 시 월령이 6개월 이하인 집단(n=166)의 수행안정성보다 1차 검사 시 월령이 7개월 이상인 집단(n=139)의 안

〈표 8〉 검사 간격이 12개월 미만인 아동의 K-BSIDII 척도별 1, 2차 수행상관 (N=181)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.05						
1차 검사 동작원점수	.94**	.00					
1차 검사 동작지수	.27**	.45**	.29**				
2차 검사 인지원점수	.90**	-.03	.95**	.20**			
2차 검사 인지지수	-.04	.19**	-.08	.13	.10		
2차 검사 동작원점수	.90**	-.06	.93**	.19**	.98**	.05	
2차 검사 동작지수	.15*	.08	.11	.19**	.18*	.46**	.30**

*p<.05 **p<.01

〈표 9〉 검사 간격이 12개월 이상인 집단의 K-BSID-II 척도별 1, 2차 수행상관 (N=124)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.07						
1차 검사 동작원점수	.93**	-.03					
1차 검사 동작지수	.23*	.23*	.38**				
2차 검사 인지원점수	.58**	-.08	.57**	.10			
2차 검사 인지지수	.17	-.04	.14	.05	.70**		
2차 검사 동작원점수	.65**	-.06	.65**	.16	.92**	.51**	
2차 검사 동작지수	.25**	-.00	.25**	.17	.47**	.53**	.66**

* $p < .05$ ** $p < .01$

정성이 높은 경향을 보였다. 월령이 낮은 집단은 인지원점수의 안정성이 $r = .18(p < .05)$ 이었으나 월령이 높은 집단은 $r = .53(p < .01)$ 으로 증가했으며 동작원점수의 경우도 $r = .31$ 에서 $r = .51$ 로 높아짐을 알 수 있다. 인지 지수의 경우는 월령이 낮은 집단은 유의한 상관이 도출되지 못했지만, 월령이 높은 집단에서는 $r = .20(p < .05)$ 로 낮지만 유의한 차이를 보이고 있다. 반면 동작지수의 경우 월령이 낮은 집단과 높은 집단 모두 유의한 상관을 보이지는 못했다.

2) 검사 간격에 따른 수행의 안정성

1차 검사와 2차 검사 간격은 최소 5개월에서

최대 36개월이었다. 전체 아동 중 검사간격이 12개월 미만인 집단이 181명(53.9%), 검사간격이 12개월 이상인 집단은 124명(46.1%)이었다. 12개월을 중심으로 두 집단으로 구분 후 각 집단 내에서 1차 검사와 2차 검사수행 간 상관관계를 표 8, 9에서 살펴보았다. 검사간격이 12개월 미만인 집단에서는 원점수와 지수점수 모두 1, 2차 수행 간에 유의한 상관이 도출되었다. 그러나 표 9에 제시된 바와 같이 검사간격이 큰 집단의 경우 원점수에서는 수행 안정성이 유의하였으나(인지척도 $r = .58$, 동작척도 $r = .65$), 지수에서는 유의한 상관을 보이지 않았다.

〈표 10〉에는 검사간격이 12개월 이상인 아동

〈표 10〉 검사 간격이 24개월 이상인 집단의 K-BSID-II 척도별 1, 2차 수행상관 (N=32)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.53**						
1차 검사 동작원점수	.67**	.25					
1차 검사 동작지수	.15	.31	.61**				
2차 검사 인지원점수	.63**	.10	.71*	.32			
2차 검사 인지지수	.37**	-.04	.33	.19	.79**		
2차 검사 동작원점수	.57**	.08	.66**	.31	.89**	.60**	
2차 검사 동작지수	.34**	-.02	.32	.22	.68**	.71**	.85**

* $p < .05$ ** $p < .01$

〈표 11〉 장애위험 및 성별에 따른 K-BSID-II 척도별 1, 2차 수행(평균(SD))

시기	척도	정상집단(n=264)		장애집단(n=41)	
		남(n=139)	여(n=125)	남(n=19)	여(n=22)
1차	인지지수	99.92(6.95)	99.50(6.83)	88.00(8.11)	84.55(9.67)
	동작지수	101.41(10.36)	102.34(11.74)	78.58(9.73)	87.32(9.02)
2차	인지지수	94.42(12.88)	97.56(12.13)	95.16(9.89)	93.59(11.06)
	동작지수	101.29(12.58)	105.28(14.07)	104.53(12.90)	98.91(11.10)

124명중 간격이 24개월-36개월 사이인 아동 32명에 대한 상관을 분석한 결과를 제시하였는데 이 경우에도 지수 간에는 상관이 유의하지 않았지만 인지 원점수는 $r=.63(p<.01)$, 동작 원점수는 $r=.66(p<.01)$ 의 유의한 상관을 나타내고 있다.

4) 장애위험에 따른 수행의 안정성

본 종단 연구는 K-BSID-II 검사가 가능한 생후 42개월동안 3년에 걸친 종단 자료수집을 위해 생후 1개월부터 1세 이전의 영유아를 주대상으로 하여 자료를 수집하였다(전체 대상의 67%). 그런데 어린 영아의 경우 진단이 가능한 뇌성마비 등 심한 장애집단과 미숙아를 대상으로 장애 집단을 선정할 수 있었으나 이들 대부분의 경우 수행이 매우 낮아 기본 점수 외에는 수행에 차이를 보이지 못한 경우가 많아 상관분석을 위주로 하는 종단분석에 적합하지 않았다. 따라서 본 연구

에서 검사를 수행할 수 있는 영유아중 인지척도나 동작척도에서 85미만의 지체를 보이는 영유아를 장애위험집단으로 정의하여 장애 위험집단 영유아(n=41)와 정상집단 영유아(n=264)의 수행(표 11)과 수행의 안정성(표 12 및 표 13)을 비교하였다.

<표 11>에서 집단과 성별에 따른 K-BSID-II 척도별 1, 2차 수행을 살펴보면 정상집단에서 남아의 1차 인지지수는 99.92(6.95), 여아의 1차 인지지수는 99.50(6.83)으로 남아가 높게 나타났지만 그 차이는 유의미하지 않았다. 동작지수 분석에서는 남아가 101.41(10.36), 여아가 102.34(11.74)로 여아가 높게 나타났지만 역시 유의미한 차이는 없었다. 2차 검사에서는 인지지수와 동작지수 각각 여아가 99.56(12.13), 105.28(14.07)로 남아의 평균인 94.42(12.88)과 101.29(12.58)보다 높은 경향을 보였다. 장애위험 집단

〈표 12〉 정상 집단의 K-BSIDII 척도별 1, 2차 수행상관 (N=264)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.08						
1차 검사 동작원점수	.97**	.05					
1차 검사 동작지수	.41**	.21**	.44**				
2차 검사 인지원점수	.62**	.04	.65**	.21**			
2차 검사 인지지수	.04	.01	.03	.07	.20**		
2차 검사 동작원점수	.65**	.03	.68**	.24**	.97**	.14*	
2차 검사 동작지수	.16**	-.03	.16*	.20**	.20*	.50**	.33**

* $p<.05$ ** $p<.01$

〈표 13〉 장애위험 집단의 K-BSID-II 척도별 1, 2차 수행상관

(N=41)

	1차 검사 인지원점수	1차 검사 인지지수	1차 검사 동작원점수	1차 검사 동작지수	2차 검사 인지원점수	2차 검사 인지지수	2차 검사 동작원점수
1차 검사 인지지수	.30						
1차 검사 동작원점수	.79**	-.19					
1차 검사 동작지수	.01	-.30**	.13				
2차 검사 인지원점수	.63**	-.03	.70**	.17			
2차 검사 인지지수	.17	.43**	.03	.16	.35*		
2차 검사 동작원점수	.68**	-.02	.73**	.15	.98**	.29	
2차 검사 동작지수	.39**	.30	.18	.04	.35*	.46**	.46**

* $p < .05$ ** $p < .01$

에서도 남녀 간에는 일관적인 차이를 보이지 않았으나 1차 수행에 비해 2차 수행 시 유의한 증가를 보이고 있다.

<표 12>에서 정상 집단에서 1, 2차 수행상관을 보면 인지원점수간은 $r = .62(p < .01)$, 인지지수간은 $r = .01$, 동작 원점수간은 $r = .68(p < .01)$, 동작 지수간은 $r = .20(p < .01)$ 으로 나타났다. <표 13>의 장애위험집단의 각 상관은 $r = .63(p < .01)$, $r = .43(p < .01)$, $r = .73(p < .01)$, 그리고 $r = .04(p > .10)$ 로 동작 지수를 제외하고는 사례수가 적음에도 불구하고 정상 집단에 비해 더 높은 상관유형을 보여 주고 있다. 또한 정상 집단에서는 인지지수의 안정성이 낮은 반면 장애집단에서는 동작 지수의 안정성이 낮은 것으로 나타났다.

IV. 논 의

본 연구에서는 심리측정적인 측면에서 매우 정교하고 신뢰로운 영아발달검사의 하나로 알려진 1개월~42개월 영유아용 BSID-II를 한국형으로 개발한 한국 Bayley 영유아 발달검사(K-BSID-II) 수행의 안정성을 알아보고 또한 검사연령, 검사간격 그리고 장애위험에 따른 안정

성의 변화를 분석하였다.

일찍이 Bayley 영유아검사의 기초가 되는 영아 발달 검사들을 대상으로 안정성을 연구한 Bayley (Bayley, 1965; 1970)는 BSID를 포함한 영아 발달 검사의 결과는 아동의 현재 발달수준을 설명하는데 유용하지만 성장후의 지능에 대한 예언력 즉 안정성은 낮은 것으로 보고하였다. 그러나 Bayley는 발달 지연이 뚜렷한 영아의 경우 BSID의 수행수준은 예언력(안정성)을 지니는 것으로 보고하고 있고 다른 연구자들(Fried, & Watkinson, 1988)도 장애아동의 진단에는 영아기의 측정이 유용함을 보고하고 있다.

아직 미국의 경우도 BSID-II에 대한 종단적인 자료에 대한 연구보고는 찾아보기 힘들지만 본 연구에서는 2002-2003년의 한국 표준화 및 타당도연구의 표본을 대상으로 5개월에서 36개월의 간격으로 2회 수행을 측정하여 영아기동안의 수행안정성을 분석하였다. <표 5>에 제시된 바와 같이 인지원점수간 상관은 인지척도 $r = .91$, 동작척도 $r = .89$ 에 이르는 매우 높은 정적 상관을 보였으나 지수 간 상관은 유의하지만 낮은 경향(인지척도 $r = .18$, 동작척도 $r = .24$)을 보이고 있다. Bayley 검사는 편차척도로서 평균 100, 표준편차 15인 분포에서 도출되는 지수를 산정한다.

이는 월령에 따라 증가하는 원점수를 월령별 표준점수 내에서 평균과 편차를 고려하여 환산된 표준점수로서 일반적으로 원점수에 비해 안정성이 낮다(McCall, 1983). 그럼에도 불구하고 본 연구에서 5개월에서 36개월을 간격으로 검사를 실시했을 때 인지지수를 제외한 모든 수행에서 상관이 도출됨으로써 K-BSID-II의 안정성은 확보된 것으로 해석할 수 있다. 또한 본 연구결과 원점수의 안정성은 매우 높아 영아기 수행의 안정성을 살펴보기 위해서는 원점수를 활용하는 것이 유용함을 시사하고 있다.

일반적으로 검사 수행의 안정성은 초기 검사 연령이 증가함에 따라 높아져 처음 검사 시기가 늦을수록 그리고 두 검사 간 간격이 짧을수록 높아진다(McCall, 1983; Miller, 1987). <표 6, 7>을 비교해 볼 때 본 연구에서도 선행연구와 마찬가지로 검사 시 월령이 높았던 집단(월령 7개월 이상)이 낮았던 집단(월령 6개월 이하)에 비해 안정성이 높은 것을 보여주었다. 검사간격(12개월 미만 vs. 12개월 이상)에 따른 안정성 분석결과도 표 8, 9에서 비교해 볼 때 선행 연구와 마찬가지로 1차 검사와 2차 검사의 간격이 작은 집단에서 인지지수와 동작지수의 안정성이 더 큰 경향을 보여주고 있다. 무엇보다도 표 10에 제시된 것과 같이 만 42개월 미만의 영유아기동안 만 2년 이상의 간격으로 검사를 2회 실시한 경우에도 원점수에서는 6할대의 높은 수행안정성을 보이고 있다. 지수의 경우 앞에서 밝힌 바와 같이 원점수에 비해 편차가 줄어들고 본 연구에서 검사 간격이 2년 이상인 집단의 경우 32명이라는 비교적 소수였음에도 불구하고 안정성이 유의한 것은 본 검사의 안정성에 대한 매우 고무적인 결과로 사료된다.

의학의 발달로 영아기 사망률이 낮아지고 조산아의 생존률이 높아지면서 발달장애에 대한

관심이 고조되고 있고 정상 영유아의 경우보다 장애집단이나 발달 지연 등 비정상적인 발달양상을 보이는 아동에 대한 조기 진단의 요구는 더 높다(민동옥·박혜원, 2003; Bayley, 1932; Fried, & Watkinson, 1988). Bayley 영유아발달검사의 일반적인 진단 범주에 따라 1차 한국 Bayley 영유아 발달검사수행에서 지수가 85미만의 점수로 정상범주(85-115)에서 벗어난 아동과 정상이상의 수행을 보인 아동의 수행을 비교하였을 때 표 11에 제시된 바와 같이 일반 아동의 수행이 1차로부터 2차 검사사이에 변화되는 것보다 장애위험 집단의 수행에 큰 폭의 향상이 있었음을 알 수 있다. 이는 발달 초기에 영아들 특히 발달이 지연된 영아들이 정상 궤도를 향한 발달(catch-up)추세를 반영한다고 볼 수 있다. 또한 본 종단연구에 참여하여 1차 검사결과와 피드백을 통해 영아의 발달지체가능성에 대해 알게 됨으로써 적극적인 중재가 이루어 졌을 가능성을 간과할 수 없다. 끝으로 본 연구의 장애 위험집단의 경우 Miller(1987)가 언급하는 통계적 회귀(statistical regression)의 영향으로 2차 검사 시 평균으로의 상향현상이 더욱 두드러지게 나타났을 가능성도 배제할 수 없다.

본 연구결과 안정성에 있어 정상 집단과 장애 위험집단의 경우 대조적인 양상을 보인 것이 흥미롭다. 정상 아동의 경우 동작 척도가 인지 척도에 비해 안정성이 높은 반면($r=.20, p<.01$ vs. $r=.01, p>.10$), 장애위험집단의 경우 인지척도의 안정성이 동작 척도에 비해 높았다($r=.43, p<.01$ vs. $r=.04, p>.10$). 이는 장애 아동의 경우 1, 2차 검사 간에 수행향상이 동작 척도에서 더욱 두드러지게 나타나고 있음에 따른 현상이다. 이는 인지발달의 장애 유무를 판단하는 데에는 동작 척도보다 인지척도가 더 유용할 수 있다는 것을 시사하고 있다. 본 검사의 인지척도에 채인기억, 신기성 선호 및 문제해결과 같은 기초적인 인지

능력을 반영하는 문항이 다수 포함되어 있기 때문일 수 있다.

본 연구에서는 출생 후 42개월(3년 6개월)동안 만 3년간의 종단연구를 수행하기 위해 월령이 6개월 이하인 매우 어린 연령의 아동을 중심으로 자료가 수집되었고 따라서 이렇게 어린 시기에 장애진단이 어려워 장애진단 집단대신 장애 위험집단을 중심으로 자료분석이 이루어졌다는 한계가 있으므로 앞으로 장애 진단이 가능한 영아기 후반의 아동을 대상으로 비교적 단기간에 걸쳐서라도 장애 아동집단의 수행안정성을 비교하는 연구를 수행해야 할 것이다.

그 외에도 3년이라는 오랜 기간에 걸쳐 어린 영아를 섭외하는 어려움과 검사 비용 등의 현실적인 어려움 때문에 인구통계학적 변인에 따른 층화표집을 실시하지 못하고 재검사에 응한 아동만을 대상으로 하였다. 따라서 각 비교 집단별로 인구통계학적 변인에 있어 동질성 확보가 어려웠다. 후속연구에서는 인구 통계학적 변인을 고려한 표집을 실시함으로써 집단별로 보다 정확한 안정성 평가와 이에 미치는 영향에 대한 분석이 이루어져야 할 것이다. 무엇보다도 K-BSID-II의 유용성을 높이기 위하여 유아기이후까지 지속적으로 추적하여 보다 장기적인 안정성에 대한 연구가 필요하다.

참 고 문 헌

강순구·조윤경(1994). *포테이지 아동발달 지침서*. 서울 : 특수교육.
 김숙현·이경옥(2001). 영아발달체크리스트 표준화연구. 한국유아교육학회 연차학술대회.
 김승국·김옥기(1985). *사회성숙도 검사*. 서울 : 중앙적성.
 대한소아과학회(2002). *한국형 영유아발달검사*. 서

울 : 대한소아과학회.
 민동옥·박혜원(2003). 조산아에 대한 베일리 영유아 발달검사 실시 및 채점 방식에 관한 연구. 한국 아동학회 추계학술대회.
 박혜원·조복희·최호정(2003). 한국 Bayley 영유아 발달검사 (K-BSID-II) 표준화연구 : 예비연구. *한국심리학회지 : 발달*, 16(4), 12-134.
 박혜원·조복희(미출간원고). 한국 Bayley 영유아 발달검사 : 해석지침서.
 배윤희(2001). 베일리 영유아발달 검사-II의 한국 표준화를 위한 예비연구. 울산대학교 석사학위논문.
 이일섭(1992). Bayley 발달검사를 이용한 고위험 저체중출생아의 발달평가. 연세 대학교 석사학위논문.
 정문자(1983). 유아발달을 위한 가정방문 프로그램 연구. *대한 가정학회지*, 21(4), 129-137.
 정문자·이은혜·박경자(1993). 베일리 유아발달 척도의 한국 표준화를 위한 예비연구. *아동학회지*, 14(1), 5-21.
 정영미(1992). Bayley 발달검사를 이용한 유아의 발달평가. 연세대학교대학원 석사 학위논문.
 제경숙(1984). 베일리의 유아발달검사의 척도화를 위한 연구. 연세대학교 석사학위 논문.
 조복희·박혜원(2004). 한국 Bayley 영유아발달검사 (K-BSID-II) 표준화연구(1) : 지역, 성별 및 모의 교육 수준에 따른 수행분석. *한국심리학회지 : 발달*, 17(1), 191-206.
 Aggarwal, P., Chaudhari, S., Bhawe, S., Pandit, & A., Barve, S.(1998). Clinical predictors of outcome in hypoxic ischaemic encephalopathy in term infants. *Annals of Tropical Paediatrics*, 18, 117-121.
 Baird, M., & Folsom, J.(1993). Concurrent validity of the Kent Scoring Adaptation of the Bayley Scales of Infant Development. *Diagnostique*, 18(2), 135-143.
 Banks, M., & Ginsburg, A.(1985). Infant visual preferences : A review and new theoretical treatment. In H. Reese(Ed.), *Advances in child development and behavior*(19, 207-246). Orlando, FL : Academic

- Press.
- Bayley, N.(1932). A study of the crying of infants during mental and physical tests. *Journal of Genetic Psychology*, 40, 306-329.
- Bayley, N.(1933a). Mental growth during the first three years. *Genetic Psychology Monographs*, 14, 1-92.
- Bayley, N.(1933b). *The California First-Year Mental Scale*. Berkeley : University of California Press.
- Bayley, N.(1935). *The development of motor abilities during the first three years*. Monographs of the Society for Research in Child Development.
- Bayley, N.(1936). *The California Infant Scale of Motor Development*. Berkeley : University of California Press.
- Bayley, N.(1965). Comparisons of mental and motor test scores for age, 1-15 months by sex, birth order, race, geographical location, and education of parents. *Child Development*, 36, 379-411.
- Bayley, N.(1969). *Bayley scales of infant development*. New York : Psychological Corporation.
- Bayley, N.(1970). Development of mental abilities. In P. H. Mussen(Ed.), *Carmichael's manual of child psychology*(3rd ed., Vol. 1, pp. 1163-1209). New York : Wiley.
- Bayley, N.(1993). *Bayley scales of infant development-II*. San Antonio : Psychological Corporation.
- Black, M. M., & Matula, K.(1999). *Essentials of Bayley Scales of Infant Development II Assessment*. New York : Wiley & Sons.
- Brody, N.(1992). *Intelligence*(2nd ed.). New York : Academic Press.
- Brooks-Gunn, J., & Weinraub, M.(1983). *Origins of infant intelligence testing. Origins of intelligence : Infancy and early childhood*(2nd ed, 25-66). In M. Lewis (Ed.) New York : Plenum Press.
- Campbell, S. K., Siegel, E., Parr, C. A., & Ramey, C. T.(1986). Evidence for the need to renorm the Bayley Scale of Infant Development based on the performance of population-based sample of 12-month-old infants. *Topics in Early Childhood Special Education*, 6, 83-96.
- Fantz, R., & Fagan, J.(1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child Development*, 46, 3-18.
- Fried, P. A., & Watkinson, B.(1988). 12.-and 24-month neurobehavioural follow-up of children prenatally exposed to marihuana, cigarettes and alcohol. *Neurotoxicology and Teratology*, 10, 305-313.
- Gelman, R., & Gallistel, D. R.(1986). *The child's understanding of number*. Cambridge, MA : Harvard University Press.
- Goldstein, D. J., Fogle, E. E., Wieber, J. L., & O'Shea, T. M.(1995). Comparison of the Bayley Scales of Infant Development-Second Edition and the Bayley Scales of Infant Development with premature infants. *Journal of Psychoeducational Assessment*, 13, 391-396.
- Jencks, C.(1972). *Inequality*. New York : Basic Book.
- Johnson, L. J., Cook, M. J., & Kullman, A. J.(1992). An examination of the concurrent validity of the Battelle Developmental Inventory as compared with the Vineland Adaptive Scales and the Bayley Scales of Infant Development. *Journal of Early Invention*, 16(4), 353-359.
- McCall, R.(1983). A conceptual approach to early mental development. In M. Lewis(Ed.), *Origins of intelligence*(2nd Ed.). New York : Plenum.
- McChesne, S.(1995). Concurrent validity of the Fagan test of infant intelligence with the Bayley Scales of Infant Development for infants from low socioeconomic and middle to upper socioeconomic status. University of South Carolina.
- Miller, S.(1987). *Developmental research methods*. New York : Prentice Hall.
- Nellis, L., & Gridley, B. E.(1994). Review of the Bayley Scale of Infant Development-Second Edition. *Journal of School Psychology*, 32, 201-209.

Rubin, B. G.(1996). A Test-retest study of the mental scale of the Bayley Scales of infant development : Second-age children with significant developmental disabilities. Temple University.

Sabourin, J. A.(1996). The stability of IQ in a clinic-referred child sample as measured by the Bayley

Scales of Infant development, the Griffiths mental development scales and the Kaufman Assessment Battery for Children. Laurentian University of Sudbury(Canada).

Snow, C., & McGaha, C.(3rd. Ed.)(2003). *Infant development*. New York : Prentice Hall.

2005년 10월 31일 투고 : 2006년 1월 28일 채택