

한국 Bayley 영유아 발달검사 (K-BSID-II) 표준화연구: 예비연구

박혜원 조복희 최호정
울산대학교 아동가정복지전공 경희대학교 아동가족전공

1-42개월 사이의 영유아의 발달을 평가할 수 있는 BSID-II를 한국 영유아용으로 표준화하기 위해 1차로 도구와 절차를 개발하고자 예비연구를 수행하였다. 서울/경인/충청, 영남, 호남에서 각각 59명, 98명 그리고 29명의 영유아를 대상으로 검사를 실시하였다. 예비연구에서는 미국 번역, 수정판에 더해 10개의 추가 문항이 사용되었다. 한국의 영유아는 인지발달지수에서 미국의 영유아보다 낮은 경향을 보였고($M=96.9$, $SD=10.9$) 동작발달지수($M=100.4$, $SD=13.4$)에서는 차이가 없었다. 행동평정척도는 높은 내적신뢰도를 보였다(.97-.98). 인지척도와 동작척도 수행의 원점수간에는 매우 높은 상관(0.98)을 보였으나(0.98) 표준화된 지수간에는 중간정도의 상관(0.52)이 나타났다. 문항의 난이도 분석결과 연령집단에 따라 기저선과 천정선 문항이 부족한 경우가 발견되어 12개 연령집단에서는 다른 연령 집단문항을 추가적으로 사용되어야 함을 시사하였다. 본 예비연구를 통해 표준화연구를 위한 도구와 절차를 완성하였다.

주요어: 한국 Bayley 영유아발달검사(K-BSID-II), 표준화, 1-42개월 영유아, 인지척도, 동작척도, 행동평정척도

1990년 이후 발달장애에 대한 관심이 높아 지고 특히 특수 교육법의 개정으로 특수교육이 보다 보편화됨에 따라 조기 진단에 대한 요구는 매우 높다. 또한 현대 사회에서 취업모가 증가함에 따라 가정의 기관에서 영유아를 보호해야 할 필요성이 대두되고, 영유아보육과 유아교육 전반에서 교육대상 연령이 하향하고 있다. 따라서 어린 영유아들이 집단교육상황에 접하게 되므로 월령별 영유아의 발달 특징을 이해하고, 영유아의 잠재 가능성을 최대한으로 계발시킬 교육에 도움을 줄 수 있는 영유아 발달에 관련된 기초 자료가 시

이 논문은 2001년도 한국학술진흥재단의 지원에 의하여 연구되었음(KRF-2001-04-D00103).

교신저자 : 박혜원, E-mail: hyewonc@mail.ulsan.ac.kr

급히 요구되고 있다(이형신, 변상형, 1999).

국내에서는 어린 영유아를 위한 발달진단이나 검사도구의 개발을 위하여 1980년대부터 연구들이 수행되었으나 전국적 표집을 통해 규준제작을 완성하지 못한 것이 대부분이다. 국내에서 사용되었던 영유아용 발달검사를 살펴보면, 오가실(1976), 전민철 등(1997)이 덴버 발달선별검사를 번안하여 사용하였고, 이근(1987)은 이를 서울지역 아동만을 대상으로 표준화하였다. 또 KIDS(Kansas Infant Development Screen)를 사용하여 한국 영아발달 선별검사표(Korean Infant Development Chart)를 개발하는 연구가 수행되었다(이종린, 1987). 무엇보다도 제경숙(1983), 정문자, 이은혜, 박경자(1993)는 베일리 영유아발달 척도(BSID)를 번안하여 한국 표준화를 시도한 바 있으나 표준화작업을 마치지 못하였다. 박혜원(1993)은 탈습관화 능력을 통해 6-12개월 영아를 측정하는 FTII(Fagan Test of Infant Intelligence)를 연구용으로 사용하였으나, 국내 규준을 마련하지는 못하였다. 그 외에도 포테이지(강순구, 조운경, 1994)와 사회성숙도 검사(김승국, 김옥기, 1985) 등이 장애아의 발달수준을 가늠하기 위해 번안되어 사용되고 있으나 국내아동을 대상으로 한 규준이 없거나, 오래된 규준을 사용하여 실용성이 떨어진다. 최근 김숙현, 이경옥(2001)은 1-2세용 발달체크리스트를 개발하였고, 대한소아과학회(2002)에서 한국형 영유아발달검사를 개발하여 규준을 제시하고 있으나, 이들은 실시의 용이성을 강조하여 제작된 간편도구로서, 제한된 정보만을 제공할 수 있을 뿐이다.

본 연구에서는 1개월부터 42개월의 어린 영아를 대상으로 세계적으로 가장 널리 사용

되고 있는 Bayley 영유아발달검사를 한국 아동에 적합하게 표준화하기 위해 수행한 예비 연구를 보고하고자 한다. Bayley 영유아 발달검사는 1933년부터 사용되었던 캘리포니아 인지척도(1세용, 유아용), 그리고 캘리포니아 동작척도(영아용)를 기초로 1969년에 Bayley Infant Scale of Development(BSID)로 개발되었다(Bayley, 1932; 1933; 1935; 1936; 1965; 1969). 그 후 오랜 수정보완과정을 거쳐 1993년에 2판(Bayley, 1993)이 제작되었다. 2판은 그 대상연령이 2-30개월용이던 1판에서 1-42개월용으로 확장되었다. 또한 1960년 이후 활발히 전개된 영아 인지 연구의 결과를 참조하여 성장후의 지능을 보다 잘 예언할 수 있는 문항을 첨가하였고 또한 고위험집단이나 발달지연 영아의 평가를 위해 새로운 문항도 첨가한 것이 특징이다(Black, & Matula, 1999; Brook-Gunn, & Weinraub, 1983; Campbell, Siegel, Parr, & Ramey, 1986).

역사적으로 BSID는 일반 영아의 발달을 평가하는 심리측정의 전통에서 출발하였다(Bayley, 1969). 따라서 BSID는 위험을 지닌 영아들의 진단에 가장 많이 사용되었으나, 임상가들은 일반 영유아들을 같은 연령의 또래와 비교할 수 있는 다양한 정보를 제공하거나 발달을 추적하는데 유용하다고 보고하고 있다(Gauthier, Bauer, Messinger, & Closius, 1999; Goldstein, Fogle, Wiever, & O'Shea, 1995; Johnson, Cook, & Kullman, 1992). 그러므로 Bayley 검사는 다음과 같은 용도로 사용되고 있다. 첫째, 아동의 발달 진단 세트검사의 한 부분으로 유용하게 사용된다. 둘째, 중재 프로그램을 실시한 이후 아동의 발달과정을 추적하는데 유용하게 사용된다. 아동이 영재아동이나 발달지

체로 진단되고 나면 그에 맞는 교육이 실시 되는데 이때 제공되는 교육이 아동의 수준에 적합하고 효과가 있는지 판단할 수 있게 해 준다. 셋째, 영아 발달에 대한 부모 교육 도구로 사용된다. 영유아의 발달에 대해 부모에게 설명하고 아이의 강점과 약점에 대해 보다 많은 이해를 할 수 있도록 도와준다. 그 외에도 Bayley 검사의 넷째 용도는 아동의 발달을 측정하는 각종 연구도구로서의 기능이다. BSID는 이러한 다양한 용도로 세계의 많은 국가에서 사용되어 왔기 때문에 본 도구를 국내에 표준화 하는 것은 많은 기존 외국 연구들(Black, & Matula, 1999; Olson, Bates, & Bales, 1984; Papousek, & Papousek, 1992; Rosengren, Behrend, & Perlmutter, 1993)과 비교할 수 있는 장점을 지닌다.

본 예비 연구에서는 BSID-II를 번안한 K-BSID-II 표준화연구용 도구가 한국아동에게 실시하기에 적합한지 그 내용과 절차를 탐색하는 것을 목적으로 한다. 보다 구체적으로는 시행과 채점에 대한 검사자의 반응을 조사하며, 무엇보다도 월령별 문항의 난이도, 문항에 대한 영아의 흥미도, 그리고 K-BSID-II 3가지 척도의 심리측정적인 특성을 분석하여 표준화 연구를 위한 기틀을 마련하고자 한다.

BSID-II의 이해

BSID-II의 구성 BSID-II는 인지척도(Mental Scale), 동작척도(Motor Scale), 행동평정척도(BRS: Behavior Rating Scale) 세 가지 척도로 구성되어 있다. BSID-II 개발에서 중요시된 기본 전제의 하나는 생애 초기의 성장이 빠르게 진행되어 발달면에서 다른 시기와는 구

별되므로 이 단계에 적합한 특수한 절차와 방법이 고안하는 것이었다. 즉 서로 다른 연령의 영아가 흥미를 갖고 참여할 수 있는 자극을 사용함으로써 관찰하려는 능력을 발달에 적합한 상황에서 수집하려고 하였다.

표 1. BSID-II의 구성

구 성	세부내용
인 지 척 도	178문항으로 구성되었으며 BSID의 기존문항에서 29개를 삭제하고 53개를 새롭게 추가하였다. 인지, 언어, 사회/정서 발달에 역점을 둔 척도이다. 인지발달 : 영아의 재인기억과 습관화, 시각적 선호/시각적 정확성, 문제 해결, 수개념과 수세기 언어발달 : 아동의 문법규칙 이해의 발달과 어휘의 측정, 표현/수용 언어의 측정 사회/정서 발달 : 검사기간 동안 검사자에 대한 아동의 반응
동 작 척 도	111문항으로 구성되었으며 BSID의 기존문항에서 8개 삭제하고 44개를 새롭게 추가하였다. 기본적인 움직임 패턴과 반응, 소근육 운동조절, 역동적 움직임, 순응적 움직임, 시각-운동 협응 발달을 측정한다. 동작 수준 : 6개월 이상 영아의 소근육 운동과 대근육 운동수준. 다양한 자세로 움직이는 능력 측정. 감각통합 : 23개월에서 37개월의 연령에게 적절한 문항으로 촉각과 시각각이 포함되었다. 지각-운동통합 : 37개월 이상 영아에게는 동작 계획측정, 13개월에서 42개월 영아에게는 소근육 운동 조절을 측정.
행 동 평 정 척 도	30개 문항으로 구성되었으며 아동의 미묘한 질적 차이와 이후 미래의 기능을 잘 예측해 주는 척도이다. 아동의 주의/각성, 과제 지향성, 정서 조절과 동작수준 및 기타의 하위 요인으로 구성되어 있다.

BSID-II 문항구성은 BSID에 사용되었던 문

항의 대부분이 그대로 유지되고 있으나, 2~30개월의 연령범위에서 1~42개월로 범위가 확장되면서 새로운 문항들을 첨가되었고, 기존 문항들 중 일부분이 수정되었으며, 적절치 못한 문항은 삭제되었다. 특히 정보처리, 기억, 학습을 측정하는 새로운 문항들이 포함되었으며, 언어 기능을 보다 상세히 평가할 수 있게 되었다. 결과 BSID-II는 인지척도 178문항, 동작척도 111문항, 행동평정척도 30문항으로 구성되어 있다. 검사의 구성 및 내용은 표 1과 같다. 따라서 이 척도는 영아의 대근육 운동, 소근육 운동, 사회/정서적 능력, 언어 능력, 인지/적응 능력 등을 평가하는 데 유용하다.

BSID-II의 실시. BSID-II는 연령집단별로 해당연령의 문항을 제시하도록 되어 있는데 1-12개월까지 매 1개월, 13, 14-16, 17-19, 20-22, 23-25, 26-28, 29-31, 32-34, 35-37, 38-42개월용으로 모두 22개 문항세트로 나뉜다. 실시문항과 별도로 발달지수를 산출할 수 있는 근거가 되는 기준은 38개로 나뉘어 있다: 3세이하의 아동에게는 1개월 단위로 기준이 제시되어 있으며, 그외에 37-39개월, 40-42개월용 기준이 있어 총 38개이다. 이러한 복잡한 실시 절차를 감안하여 지침서에서는 검사자들에게 실시문항목록(Cue sheet)을 활용하도록 하고 있다. 이는 영아의 월령에 맞추어 검사에 사용되는 준비물과 개념적 관련성에 따라 검사 문항을 재배치시킨 것이다. 그러나 실시문항의 순서는 영아의 반응에 따라 바꿀 수 있어 융통성을 지니고 있다.

검사 실시시 부모는 아동과 함께 있을 수 있으며, 검사자는 필요한 경우 부모가 아동의 반응을 유도해 내거나 아동을 안정시키도록

요청할 수 있다. 검사자가 아동과 라포 형성을 잘 하지 못하거나 흥미를 유지시키지 못할 경우 검사 결과의 타당도를 저해할 수 있다.

검사에 필요한 도구는 판매되는 검사도구 안에 대부분 포함되어 있으나, 기타 보조도구(초시계, 표준크기의 계단)나 일회용으로 사용하여야 하는 백지 및 휴지(영유아의 침을 닦거나 도구에 묻은 침을 닦을 때 사용)등은 검사자가 별도로 준비하여야 한다. 또한 넓은 마루 공간이 필요하다. 검사에 소요되는 시간은 25분(1개월이하)에서 60분(15개월 이상)사이로 평균 45분정도이나 10%정도의 경우는 75분 이상이 소요되는 것으로 알려져 있다. 기록용지에는 쉬운 문항부터 어려운 것으로 난이도에 따라 배열되어 있고 평균 월령과 정상 범위(95%)가 문항마다 제시되어 있다.

BSID-II의 표준화과정과 심리측정적 특성. BSID-II는 미국 전가정을 대표하는(성, 인종, 교육, 지역에 따라 층화표집) 1,700명의 아동을 대상으로 표준화되었다. 표집 대상은 1, 2, 3, 4, 5, 6, 8, 10, 12, 15, 18, 21, 24, 27, 30, 36, 42개월의 17개 연령집단별로 100명(남여 동수)의 영유아이며, 1-5개월 사이의 영아는 각 개월에서 \pm 1주내에 출생한 경우, 6-24개월 사이의 영아는 \pm 2주, 그 이후의 유아는 \pm 3주내에 출생한 정상적인 아동으로 하였다. BSID-II는 BSID, DDST, DAS, McCarthy Scale 등과의 공인 타당도가 증명되었다(Bayley, 1993).

BSID-II의 해석은 인지발달지수(MDI)와 동작발달지수(PDI)의 점수에 따라 범주를 정하고 발달영역 프로파일에서 인지, 언어, 사회성, 동작 기능에 대한 발달연령을 계산해 내는 것으로 가능하다. 발달지수들의 진단 범주는 표 2에 제시되어 있다.

표 2. 발달지수의 진단범주

점 수	범 주
115 이상	가속 수행 (Accelerated Performance)
85 - 114	정상 범주 (Within Normal Limits)
70 - 84	약간 지연 (Mildly Delayed Performance)
69 이하	심각한 지연 (Significantly Delayed Performance)

BSID-II의 행동평정척도(BRS)는 임상적으로 BSID의 행동평정척도(IBR)보다 더 유용하다. BSID-II의 행동평정척도는 5개요인(각성/주의집중, 과제지향성/몰입, 정서조절, 동작수준, 기타)으로 구성된 것으로 밝혀졌다.

Bayley의 보고(1993)에 의하면 BSID-II와 BSID의 인지척도 사이에는 중간정도의 상관($r=.62$)이 있었으며, BSID-II의 인지척도가 BSID 척도보다 평균 12점 정도가 낮았다. 동작척도 간에도 역시 중간 정도의 상관이 있었으며($r=.63$), BSID-II의 동작척도가 BSID 척도보다 평균점수가 7점 정도 낮았다. BSID-II에서 인지척도와 동작척도간의 상관은 어린 연령에서 더 높았다. 따라서 어릴 경우 운동 기능 장애가 있는 아동은 인지척도와 동작척도 둘 다 점수가 낮을 수 있다. 성장후 IQ와의 상관은 일반적으로 어린 영아의 경우에는 낮지만 24개월부터 점차 증가한다. 그러나 BSID-II는 조기 발달 지체를 발견하는 데 유의미한 예측력이 있다. 임상적으로 볼 때 운동 기능 장애나 비정상적인 운동 기능은 항상되기도 하지만 정상 범위를 이탈한 인지기능은 거의 유사한 상태로 유지되거나 더 악화된다고 보고되고 있다(Bayley, 1993). 인지

척도와 동작척도 각각에서 평균은 100이며 최하점수는 50점이다.

BSID 및 BSID-II의 국내 연구 베일리 척도에 대한 국내연구를 살펴보면 1983년 제경숙은 Bayley 영아발달검사(BSID)의 척도화 가능성과 문항의 양호도를 분석한 결과 우리나라 영유아의 발달 경향을 효과적으로 변별해 줄 수 있다고 보고하였다. 정문자(1983)는 BSID를 사용하여 가정방문 프로그램효과를 살펴 보았다. 그 후 정문자, 이은혜, 박경자(1993)는 BSID를 표준화하기 위한 예비적 탐색목적으로 서울지역의 2-30개월사이의 아동을 14개 연령집단별로 285명을 표집하여 검사를 실시하였다. 그 결과 이 척도의 문항은 대부분의 경우 국내 영유아의 발달 경향을 효과적으로 변별해주며 총점 분포를 통해서도 전반적으로 연령에 따라 발달하는 경향이 있었다고 보고하였다. 또한 인지척도와 동작척도 문항의 연령 배치는 국내 유아에게 적용함에 있어 크게 무리가 없어 가까운 월령 내에서 부분적인 수정만으로도 활용가능하다고 보고하였다. 끝으로 신뢰도에 있어서 문항내적 일치도 면에서 만족할 만하며, 검사자간 신뢰도는 대체로 양호하였다. 그리고 이 시기의 영아들에게는 대체로 인지기능과 동작기능간에 정적 관계가 있었으며, 우리나라 영아들은 미국 영아에 비해 인지기능 중 그림책과 관련된 행동, 대상 영속성 개념등에서 다소 빠른 발달을 보였으며, 조작 행동과 언어발달에서 다소 늦은 경향이 나타났다고 보고하였다. 그러나 이 연구에 후속하여 표준화는 이루어지지 못하였다.

그 외에도 정영미(1992)는 연세대학교 의과 대학 연동세브란스병원에서 출생한 6,7,8개월

사이의 영아 40명을 대상으로 BSID를 실시한 결과 대상 유아의 평균 인지발달지수는 101.4 ±7.1(6개월), 102.2±9.0(7개월), 97.3±3.4(8개월)으로 감소하였으나 전반적으로 유용한 도구임을 보고하였다. 이일섭(1992)은 연세대학교 의과대학 영동세브란스병원 신생아 집중치료실에서 치료를 받은 환아중 고위험 신생아로 분류되어 추적관리를 필요로 하였던 고위험 저출생체중아 42명을 대상으로, 환아의 교정연령 9개월에 인지발달 및 동작발달지수를 검사하였으며, 발달이상과 관계가 있다고 알려진 여러 주산기 위험요인과 Bayley 발달지수와의 관계를 분석하였다. 분석 결과 고위험 저출생체중아 42명에서 Bayley 인지발달지수 및 동작발달지수의 평균 및 표준편차는 99.4±12.8과 101.9±16.6으로 동작발달지수가 인지발달지수보다 높았다고 보고하였다. 배윤희(2001)는 국내 최초로 BSID-II를 사용하여 울산광역시 소재 어린이집 5곳을 선정하여 의학적 질병이 없는 정상아이를 67명(남자-37명, 여자-30명)에게 검사를 실시하였다. 결과 BSID-II의 수행은 연령과 성별에 따라 차이가 있으며, 인지척도에서 .88, 동작척도에서 .74-.79의 높은 신뢰도를 보였다고 보고하였다. 또한 인지척도와 동작척도간 높은 정적 상관을 보였으며 K-WPPSI(박혜원, 박금주, 박광배, 1996)와도 높은 공준타당도를 보인다는 것을 밝히고 본 도구의 표준화가능성을 열어주었다.

방법

연구 대상

BSID-II의 도구 및 절차의 특성을 살펴보기

위해 17개 월령집단에서 186명을 표집하였다. 각 지역별, 월령별, 성별로 표집된 영유아의 분포는 표 3에 제시하였다. 본 예비연구에서는 규준을 작성하기 위한 연구가 아니므로 표준화연구에서 계획하고 있는 인구통계학적 변인(지역, 월령별, 성별, 및 부모학력)에 따라 정확한 층화표집을 실시하지 않았다. 서울/경인지역, 충청, 영남, 호남지역의 전국에서 훈련받는 검사자가 주거하는 지역에서 10명 내외를 검사하는 것을 목표로 하되 검사자별로 월령과 성(gender)을 지정해 주었다.

표 3. 예비연구에 참여한 영유아의 지역, 연령 및 성별분포

성/ 월령	1,2,3, 8,10,12, 24,27,30, 4,5,6 15,18,21 36,42 개월 개월 개월				계	총계
	서울/경인/충청	남 6	6	11		
영남	남 12	18	15	45	98	
	여 17	23	13	53		
호남	남 5	4	4	13	29	
	여 3	3	10	16		
계	남 23	28	30	81	186	
	여 22	34	49	105		
총계	45	63	79	186	186	

표집 대상은 심각한 의학적 합병증 없이 태어난 만기(full-term) 출산아(出産兒)이며 최근에 정신적·신체적·행동적 문제가 없다고 판명되었거나 치료를 받은 기록이 없다고 보고된 정상 유아만을 대상으로 하였다. 이들은 개별 접촉, 지역 의료기관, 어린이집 등 영유아 보육기관과 사회복지시설, 그리고 한국에서 영유아를 위한 조기 교육프로그램을 시행

하는 사교육기관을 통해 모집하였다. 최종적으로 표집된 영아의 지역분포는 서울/경인/충청 59명, 영남 98명, 호남 29명이었다.

도구

번역된 문항과 검사도구는 전공 교수와 기존 BSID 검사사용자에 의해 검토되고, 문화적인 차이(예, 아동의 얼굴은 한국아동모습으로)는 수정되었다. BSID-II 원본에 기초하여 연구용 검사도구 30세트를 제작하였다. 검사 문항과 도구 검토과정중 문화적인 차이가 나타날 수 있는 문항에 대비하여 표 4에 제시된 것과 같은 대처가능한 문항을 10개(동작척도 3개, 인지척도 7개) 첨가하였다. 각 문항은 4세 미만의 영유아에게 실시할 수 있는 인지, 동작 발달지표로 흔히 사용되는 것들로 아동학의 문헌이나 유사한 검사(예:K-WPPSI)들에서 응용한 것이다. 각 첨가문항은 추정되는 월령에 해당되는 문항뒤에 배치하였고, 원도구의 기록용지의 형식을 최대한 활용하기 위해 해당 문항번호는 전체 문항의 번호를 수정하는 대신 바로 전 문항번호뒤에 하위번호를 첨가하는 형식(63-1)을 취하였다.

검사자훈련

예비연구에서 검사자는 서울/경인 9명, 충청 2명, 영남 7명, 호남 2명으로 구성되었는데 아동학 관련분야 석사과정 이상의 학력을 가진 사람으로 기존에 사용되고 있는 다양한 지능검사에 대한 경험이 있었다. 이들은 2회에 걸친 교육과 훈련 검사를 받았다. 먼저 영유아를 대상으로 BSID 검사 및 BSID-II 검사

표 4. 예비연구에 추가된 문항

문항번호	문항	검사실시 월령	
63-1	낮을 가린다.	5-9	
130-1	간단한 지시를 수행한다.	20-28	
140-1	반대어를 안다.	23-37	
인지	145-1	가장 긴 것에 대해 안다.	23-37
	145-2	가장 많은 양에 대해 안다.	23-37
	145-3	똑같은 양에 대해 안다.	23-37
175-1	문장 따라하기	38-42	
77-1	공을 앞으로 찬다.	14-25	
동작	107-1	발끝을 붙여 걷는다.	35-42
	111-1	공을 받는다.	35-42

경험을 쌓은 전문 검사자가 이 척도에 대한 전반적인 오리엔테이션과 검사 도구 사용에 대한 구체적 시범을 통해 지도하였다. 1차 훈련을 마친 후 검사자는 1-2명의 영유아를 대상으로 실습한 연습 사례를 제출하여 기록용지상에 나타난 시행 또는 채점 과정의 모든 오류 또는 문제에 대해 서면과 구두로 보고받았다.

검사 절차

각 아동은 가정과 교육기관 그리고 주거지역에 가까운 대학실험실(경희대학교, 울산대학교, 부산대학교) 등에서 개별적으로 검사를 받았다. 검사평균시간은 영유아의 연령에 따라 비례하며 증가하여 30분에서 1시간정도가 소요되었다.

결 과

분석에 앞서 검사자간 신뢰도를 구하기 위해 10개월에서 36개월 사이의 20명의 영유아(1세 미만 6명, 1-2세 6명 2세 이상 8명)를 대상으로 두명의 검사자가 채점한 내용을 분석하였다. 한 검사자는 검사실시와 채점을 하였고, 동시에 다른 검사자는 아동을 관찰하며 채점한 결과 채점의 일치도를 살펴보았다. 결과 인지척도는 .99, 동작척도는 .99 그리고 행동평정척도는 5개 요인에서 .93에서 .83사이의 상관을 보였다.

본 예비연구에서는 지역 인구비례 표집을 하지 못하였기에 지역에 따른 분석은 생략하였다. 표 5에서 한국 영유아의 K-BSID-II 수행을 보면 인지발달지수는 96.89(10.87), 동작발달지수는 100.43으로 미국 영유아와 유사하였으나 인지발달지수는 낮은 편이다. 성차는 유의하지 않았다.

표 5. 한국 영유아의 BSID-II 수행

성	N	인지지수	동작지수
		M(SD)	M(SD)
남	81	96.28(12.22)	99.23(13.25)
여	105	97.36(9.74)	101.35(13.42)
전체	186	96.89(10.87)	100.43(13.35)
t		-.670	-1.073

비록 BSID를 사용하여 검사내용에 차이가 있지만 인지지수보다 동작지수가 높다는 선행연구(정영미, 1992)와 일치하며 또한 BSID-II를 사용한 배윤희의 연구결과와도 유사하였다.

표 6에는 행동평정척도(BRS)의 내적 타당도를 분석하였다. 미국 검사의 원저자인 Bayley

는 행동평정척도에 대하여 많은 관심을 가지고 그 유용성에 대해 강조하였다(Black, & Manula, 1999). 특히 BRS가 미국과 문화적인 차이등으로 인해 차이가 있을 수 있었으나 한국 영유아의 경우에도 내적 타당도는 .97에서 .98사이로 매우 높았다.

표 6. BSID-II 행동평정척도 내적타당도(Cronbach α)

월령	명	항목수	Cronbach α
1-5	33	20	.97
6-12	45	27	.98
13-42	108	28	.98

표 7에는 K-BSID-II의 소척도별 상관을 분석하였다. 인지원점수의 경우 인지지수와는 .264(p<.01)의 상관을 보였으며 동작 원점수는 동작지수와 .454(p<.01)의 상관을 보였다. 인지척도와 동작척도간 상관을 보면 원점수 간에는 .99의 높은 상관을 보였으나 지수간에는 .52의 상관을 보였다. K-BSID-II의 원점수와 지수점수간의 상관은 미국 원 BSID-II의 경우(.98, .44)와 비교할 때 유사하거나 더 높았다. 끝으로 검사문항별 난이도를 살펴본 결과 표 8과 같이 90%이상의 수준을 나타내는 천정선이하의 문항이나 15%이하를 나타내는 기저선 이하의 문항이 없는 경우가 발생함을 알 수 있다. 따라서 표준화연구에서는 근접연령의 문항세트로부터 일부 문항을 첨가하여 검사를 실시하여야 할 것을 시사하였다. 기저선이하, 천장선이상의 수행을 나타낸 문항의 수를 표 8에 제시하였다.

표 7. BSID-II 척도별 수행상관

	인지척도 원점수	인지 발달지수	동작척도 원점수	동작 발달지수	행동평정척도 (과제지향성)	행동평정척도 (정서조절)	행동평정 척도 (동작수준)	행동평정 척도 (기타)
인지척도 원점수	1.000							
인지 발달지수	.264**	1.000						
동작척도 원점수	.986**	.203**	1.000					
동작 발달지수	.412**	.524**	.454**	1.000				
행동평정척도 (과제지향성)	.154	.361**	.177	.328*	1.000			
행동평정척도 (정서조절)	.403**	.487**	.414**	.444**	.836**	1.000		
행동평정척도 (동작수준)	.425**	.396**	.401**	.331*	.708**	.728**	1.000	
행동평정척도 (기타)	.036	.180	.045	.170	.609**	.402**	.473**	1.000

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

표 8. 천정선이상, 기저선이하 수행의 문항수

월령	인지척도 문항수				동작 척도 문항수			
	95% +	90% +	15% -	10% -	95% +	90% +	15% -	10% -
1	2	4	2	2	4	6	2	2
2	6	6	5	5	4	4	6	6
3	1	1	1	2	5	5	1	1
4	0	0	2	9	1	1	2	5
5	5	5	5	5	3	3	0	0
6	5	5	1	1	3	3	2	2
8	1	1	2	3	0	2	0	0
10	3	3	1	2	4	5	2	2
12	0	0	4	2	2	2	2	1
15	1	1	2	2	0	0	4	3
18	2	2	5	5	2	2	0	0
21	1	1	6	3	0	0	2	1
24	1	2	1	4	0	2	0	0
27	0	1	2	2	0	2	1	1
30	1	1	0	2	1	1	2	2
36	2	4	0	0	2	2	1	1
42	3	4	0	0	0	2	0	0

결론 및 논의

발달 장애아동을 위한 진단도구나 발달에 적합한 교육을 위해서 유아의 잠재적 문제나 가능성을 조기 발견할 수 있는 도구에 대한 요구는 매우 높다(박혜원, 2000). 국내에서는 영유아에 대한 진단검사는 외국의 것이 그대로 번안되어 사용되거나, 시기적으로 오래되고 또는 간편형으로 제작되어(김숙현, 이경옥, 2001; 김승국, 김옥기, 1985; 대한소아과학회, 2002; 오가실, 1976; 이근, 1987) 영유아발달에 대한 자세한 정보를 제공하지 못하고 있다.

Bayley 검사는 국내에서 지난 20년가까이 표준화를 위한 연구가 간헐적으로 수행되었으나 국내의 다른 영아검사 연구(송창화, 1964; 우영근, 1962; 이영석, 1995; 이은해, 고윤주, 1988; 이영자 등, 2001; 이종린 등, 1987; 정문자 등, 1993; 정혜영, 1995)와 마찬가지로 검사도구의 제작, 검사실시상의 어려움 그리고

무엇보다도 영아의 표집에 많은 어려움이 있어 표준화가 이뤄지지 못했다. 그러나 최근 신생아들조차 세상을 이해하는 틀을 가지고 태어나 어린 영아기부터 학습과 발달이 매우 중요하다는 것이 강조되고 있다(EBS, 2002). 따라서 영아기부터 아이에 맞는 자극과 환경 조성이 필요하고 따라서 아동을 올바르게 진단하거나 발달평가를 하는 것이 중요하다는 인식이 확산되고 있다. 한편 국내에서는 조산아동의 출생이 10%에 육박하는 등 발달장애위험을 가진 아동이 많아지면서 조기 진단에 대한 연구와 도구의 개발은 더 이상 늦출 수 없는 과제이다.

본 연구에서는 오랜 개발 역사를 지니고 전세계적으로 보편화되어 있는 Bayley 영유아 발달검사를 한국 아동의 수준에 적합하게 개발하고 규준을 개발하기 위해 예비연구를 수행하였다. 특히 Bayley 영유아발달검사는 양적인 점수를 산출하거나 명명하는데 그치는 것이 아니라, 인지와 동작능력의 발달을 표 1에 제시된 것과 같은 다양한 측면에서 측정함으로써 개별 아동의 발달상의 강점과 약점을 자세히 알 수 있다(Aggarwal, Chaudhari, Bhave, Pandit, & Barve, 1998; King, & Seegmiller, 1973; Pelco, & Sarpolis, 1999; Sabourin, 1996; Wasserman, & Matula, 1996). 특히 아이만이 가지는 독특한 학습방식은 어떤 것인지 등에 대해 알게 해주는 이점을 가진 발달 검사로 보다 효율적인 교육을 위한 중요한 자료가 된다고 할 수 있다.

그러나 Bayley 영유아발달검사(BSID)는 실시상의 어려움(Baird, & Folsom, 1993; Dickerson, 1997; Gauthier, Bauer, Messinger, & Closius, 1999; Mayes, 1997; Ross, & Lawson, 1997),

검사결과의 안정성(Rubin, 1996; Sabourin, 1996) 등에서 많은 논의가 있어 왔다(Bayley, 1969, Black, & Matula, 1999, Nellis, & Gridley, 1994). 따라서 BSID-II는 안정적인 인지능력으로 간주되는 습관화/탈습관화 능력이나 기타 인지 능력에 관한 항목을 첨가하여 이러한 문제를 보완한 것으로 알려 있다(Goldstein, Fogle, Wieber, & O'Shea, 1995; McChesne, 1995).

가장 먼저 BSID-II를 한국 영유아에게 실시한 결과 비록 일부 문항의 경우 실시방법과 채점방법에 대한 어려움이 있었으나 검사자간 신뢰도에서 입증된 바와 같이 검사자들은 신뢰롭게 검사를 실시할 수 있었다는 것을 보고하였다. 또한 영유아들의 모집이 어려웠으나 실제 검사에 참여한 경우 검사를 중단해야 하는 등의 어려움은 거의 없었다. 따라서 검사자의 입장에서나 피검사자(영유아)의 입장 모두에서 본 검사의 실시 및 절차상의 문제는 없는 것으로 사료된다.

K-BSID-II 수행 결과를 살펴본 표 1에서 한국 아동은 미국 아동에 비해 동작발달 능력은 유사하나 언어와 인지 능력은 상대적으로 낮은 것으로 나타났다. 이는 앞으로 표준화연구를 통하여 보다 많은 아동을 대상으로 확인할 필요가 있다. 성차분석결과 유의하지는 않으나 여아의 수행이 높은 경향을 보였다. 많은 연구에서 밝힌 바와 같이 생의 초기에는 여아의 발달이 빠름을 다시 한번 확인하였다(Bayley, 1935; Mussen, 1973).

행동평정척도와 관련하여 미국문화와의 차이 등으로 인해 내용타당도분석결과에 큰 관심을 두고 분석하였다. 그러나 표 6에 제시된 바와 같이 행동평정척도(BRS)의 내적 타당도 분석결과에서도 매우 높은 신뢰도가 나타나

이 척도의 한국화에도 문제가 없는 것을 사료된다. 끝으로 K-BSID-II의 소척도별 상관분석결과에서도 원 척도와 유사한 경향을 보이며 안정적인 상관유형이 나타났다.

본 연구를 통하여 BSID-II의 실시절차, 도구등이 문화적인 차이가 있는 것만 변형되었을 때 한국 영유아에게 적용하는데 문제가 없음을 알 수 있었다. 그러나 표준화연구에서는 표 8에서 밝혀진 것과 같이 천장선과 기저선이 부족한 연령의 경우 앞 뒤 연령집단의 문항을 실시하도록 연령별 문항세트를 조정할 필요성이 밝혀졌다. 앞으로 지역인구분포, 부모교육수준 등의 가정환경까지 고려한 층화표집을 통하여 전국 영유아를 대상으로 한 표준화연구가 수행되어야 할 것이다. 또한 이러한 연구를 통하여 한국 영유아발달을 측정하고 다른 국가의 영유아와 비교할 수 있는 진단도구가 개발되어, 임상 장면과 연구분야에서 많이 활용될 수 있어야 할 것이다.

참 고 문 헌

- 강순구·조윤경(1994). 포테이지 아동발달 지침서. 서울: 특수교육
- 김남희(1996). 영유아 발달선별검사 개발을 위한 기초연구. 부산대학교 교육대학원 석사학위논문.
- 김숙현, 이경옥(2001). 영아발달체크리스트 표준화 연구. 2001. 한국유아교육학회 연차학술대회.
- 김승국·김육기(1985). 사회성숙도 검사요강. 서울: 중앙적성출판사.
- 대한소아과학회(2002). 한국형 영유아발달검사.
- 박혜원, 곽금주, 박광배(1996). 한국 웨슬러 유아지능검사의 표준화: 표준화연구(1). 한국심리학회지:발달, 8(2), 33-41.
- 박혜원(2000). 아동지능검사의 개요. 한국 웨슬러 유아지능검사 워크샵, 초급, 7-19.
- 배윤희(2001). 베일리 영유아발달 검사-II의 한국 표준화를 위한 예비연구. 울산대학교 석사학위논문.
- 송창화(1964). 한국 영유아의 운동, 행동발달에 관한 연구. 경북의대잡지, 5(1), 195-233.
- 오가실(1976). Denver Developmental Screening Test의 표준화를 위한 연구. 간호학 논집, 1.
- 우영근(1962). 우리나라 유아의 정신발육에 관한 연구. 전남대학교논문집, 7, 233-335.
- 이 근(1987). 서울 아동의 덴버 발육 선별 검사 및 한국판 DDST. 소아과, 30(9), 958-971.
- 이영석(1995). 한국 유아의 연령별 발달수준 확인 연구. 미래유아교육학회지, 1권, 1-66.
- 이영자, 이종숙, 신은수, 곽향림, 이정옥(2001). 교사 평정용 영아발달 평가도구. 아동학회지, 2(2), 23-47.
- 이은혜·고윤주(1988). 아동의 행동발달 평정척도 개발에 관한 연구. 아동학회지, 9(2), 1-28.
- 이일섭(1992). Bayley 발달검사를 이용한 고위험 저출생체중아의 발달평가. 연세대학교 대학원 석사학위논문.
- 이종린·최증환·문형로·조세진(1987). 한국 영아 발달 스크리닝 검사에 관한 연구. 소아과, 30(8), 852-866.
- 이형신, 변상형(1999). 신생아 가사환자에서 주산기 위험 인자 및 신생아합병증과 Bayley검사 결과와의 관계에 관한 연구. 대한신생아학회지, 6(1).
- 전민철, 김영훈, 정승연, 이인구, 김종완, 황경태(1997). 발달지체아에서 Denver II의 유용성에 대한 연구. 대한소아신경학회, 5(1), 111-118.
- 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 석사학위논문.
- 정혜영(1995). Early Screening Inventory의 표준화를 위한 예비연구. 부산대학교 교육대학원 석사학위논문.
- 제경숙(1983). 베일리 유아발달검사의 표준화를 위한 예비연구. 연세대학교 대학원 석사학위논문.
- 한국교육방송(EBS)(2002). 아기성장정보고서.
- Aggarwal, P., Chaudhari, S., Bhave, 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.
- Bayley, N.(1932). A study of the crying of infants during mental and physical tests. *Journal of Genetic Psychology*.
- Bayley, N. (1993). *Bayley Scales of Infant Development. 2nd Edition: Manual*. San Antonio: Harcourt Brace & Company.
- Bayley, N.(1933). *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.
- Black, M. M. & Matula, K.(1999). *Essentials of Bayley Scales of Infant Development II Assessment*. New York: John Wiley & Sons.
- Brook-Gunn, J., & Weinraub, M. (1983). Origins of infant intelligence testing. In M. Lewis (Ed.) *Origins of intelligence: Infancy and early childhood(2nd ed, 25-66)*. 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.
- Dickerson, M.(1997). Potential scoring problems using the Bayley Scales of Infant Development-II Mental Scale. *Journal of Early Intervention*, 21(1), 36-44.
- Gauthier, S. M., Bauer, C. R., Messinger, D. S., & Closius, J. M. (1999). The Bayley Scale of Infant Development II: Where to start. *Journal of Developmental and Behavioral Pediatrics*, 20, 75-79.
- Goldstein, D.J., Fogle, E.E., Wiebeer, 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 Psychodeucational Assessment*, 13, 391-396.
- 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 Intervention*, 16(4), 353-359.
- King, W., & Seegmiller, B. (1973). Performance of 14- to 22-month-old black first-born infants on two tests of cognitive development: The Bayley scales and the Infant Psychological Development Scale. *Developmental Psychology*, 8, 317-326.

- Mayes, S. D. (1997). Potential scoring problems using the Bayley Scale of Infant Development-II Mental Scale. *Journal of Early Intervention, 21*, 36-44.
- 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.
- Mussen, P.(1973). *The psychological development of the child*. N.Y.:Prentice-Hall, Inc.
- Nellis, L., & Gridley, B. E. (1994). Review of the Bayley Scale of Infant Development-Second Edition. *Journal of School Psychology, 32*, 201-209.
- Olson, S. L., Bates, J. E., & Bayles, K. (1984). Mother-infant interaction and the development of individual differences in children's cognitive competence. *Developmental Psychology, 20*, 166-179.
- Papousek, H., & Papousek, M. (1992). Beyond emotional bonding: The role of preverbal communication in mental growth and health, *Infant Mental Health Journal, 13*, 43-53.
- Pelco, L. E. & Sarpolis, D. L.(1999). Concurrent validity of the Mayes Motor Free Compilation (MMFC) for assessing mental ability in children with physical impairments. *International Journal of Disability, Development and education, 46*(4), 487-492.
- Rosengren, K. S., Behrend, D. A., & Perlmutter, M. (1993). Parental influences on children's cognition. In R. Pasnak, & M. L. Howe (Eds.) *Emerging Themes in Cognitive Development: Volume II: Competencies*. New York: Springer-Verlag Inc.
- Ross, G., & Lawson, K. (1997). Using the Bayley-II: Unresolved issues in assessing the development of prematurely born children. *Journal of Development and Behavioral Pediatrics, 18*, 109-111.
- 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).
- Wasserman, J. D. & Matula, K. (1996). The factor structure of the behavior rating scale of the Bayley scales of Infant development-II. *Education and Psychological Measurement, 56*(3), 460-474.

The standardization of Korean Bayley Scales of Infant Development (K-BSID-II): A Pilot Study

Hyewon Park Choi
Dept. of Child & Family Welfare,
University of Ulsan

Bokhee Cho Hojung Choi
Dept. of Child & Family Studies
Kyunghee University

As the preliminary study for the development of the test material for Korean Bayley Scales of Infant Development-II, 186 infants from Seoul and its vicinities, middle-, south regions of Korea were tested with the pilot version of K-BSID-II. The pilot test included ten extra items in addition to translated BSID-II items. Korean infants showed a slightly lower Mental Developmental Index ($M=96.9$, $SD=10.9$) on BSID-II than American counterparts but the Psychomotor Developmental Index ($M=100.4$, $SD=13.4$) was very similar to that of the US norm sample. There was high reliability of BSID-II Behavior Rating Scale in all age groups (.97-.98). The correlation between the MDI and PDI was moderate (.52). There were high inter-scoring reliabilities for MDI, PDI and BRS. Item difficulty analyses suggested rearrangements of items and the addition of test items from other age groups for some age groups for the standardization study. The results provided basic materials for the standardization study of Korean BSID-II.

Keywords: Korean Bayley Scale of Infant Development, standardization, 1-42 month olds, mental scale, motor scale, behavior rating scale