

발달선별검사의 인지영역 타당화 연구*

이준석**

나사렛대학교 재활학부

《요약》

이 연구는 표준화중인 발달선별검사의 인지영역 문항들의 타당성을 살펴보고자 하였다. 발달검사의 경우에는 연령별 타당도가 중요한 의미를 지니므로 개별 문항 타당성을 평가하기 위해 문항분석(문항난이도, 문항변별도), 문항 내적 일치도 및 연령별·성별 차이검증을 실시하였다. 분석결과에서 각 문항의 난이도는 점진적으로 증가하는 것으로 나타났으며, 문항변별도도 양호한 것으로 나타났다. Cronbach의 알파 값은 .94로 나타나 문항들이 내적인 일관성도 지닌 것으로 나타났다. 생후 12개월 이전 영아에게 해당하는 문항들은 생후 12개월 이후 반응에서 친정효과를 보여주었으며, 2~3세 유아에게 적절한 문항들은 12개월 이후 행동빈도가 증가하여 4세 경에 접근선에 이르는 형태를 보여주었다. 반면에 60개월 이상 유아들이 수행할 수 있는 문항들은 18개월 이전 영유아들의 반응이 바닥효과를 보여주는 것으로 나타났다. 인지영역 전체점수는 태어나서 35개월까지 지속적으로 증가하였으며, 36개월 이후 조금씩 증가폭이 줄어들어 54개월 이후에는 일정 수준을 유지하는 것으로 나타났다. 개별 문항 역시 연령별 변화를 적절히 반영하는 것으로 나타났다. 그리고 영역별·문항별 성별차이에 대해 논의하였다.

주제어 : 발달검사, 인지발달, 검사타당도, 문항타당도

1. 서론

영유아의 부모뿐만 아니라 영유아의 성장과 발달에 관심을 가진 유아교육, 유아특수교육, 아동발달, 소아과학, 간호학 등의 관련분야 전문가, 보육현장의 교사에 이르기까지 영유아의 발달 수준을 확인하는데 많은 관심을 기울이고 있다. 각 분야에서 발달 사정에 대한 연구가 활발히 진행되고 있다는 사실이 이를 잘 반영해 주고 있다.

소아과학과 간호학 분야에서는 오가실(1976)에 의해 시작된 덴버발달검사 표준화가 이근(1987, 1990, 1996)의 표준화연구를 거쳐 한국형 Denver II(신희선, 한경자, 오가실, 오진주, 하미나, 2002) 표준화를 완성하였고, 한국형 영유아발달검사(대한소아과학회,

* 이 논문은 2006년도 나사렛대학교 학술연구비지원에 의해 연구되었음.

** 교신저자(joonsuk@kornu.ac.kr)

2002), 서울영유아발달선별검사(이근, 2000), 이화영아발달선별검사(이근, 2002)를 만들어 내었다. 아동학과 유아교육 분야에서는 베일리 유아발달척도 표준화를 위한 제경숙(1984)의 기초연구이래로 정문자, 이은혜, 박경자(1993), 배윤희(2002), 박혜원, 조복희, 최호정(2003), 조복희, 박혜원(2004), 박혜원(2006), 박혜원과 신민선(2006) 등의 연구를 거쳐 개정판 Bayley 영유아발달검사(K-BSID-II)가 표준화되었고, 영아발달 평가도구(이영자, 이종숙, 신은수, 곽향림, 이정옥, 2001), 유아발달 수준검사(국립교육평가원, 1996), 3·4세용 유아발달검사(손원경, 2004) 등이 개발되었으며, 유아용 조기선별검사 ESI-P(황해익, 손원경, 2002), ESI-R(황해익, 손원경, 2003), 걸음마기 아동의 발달선별척도(이정화, 2006), 영유아용 발달검사(이은혜, 박주희, 최혜영, 2005) 등에 대한 기초연구가 진행되고 있다. 특수교육분야에서는 한국판 ASQ(허계형, Squires, Yovanoff, 이준석, 2006), 한국판 DIAL-3(전병운, 조광순, 이기현, 이은상, 임재택, 2004), KISE 발달척도(국립특수교육원, 1996), 발달장애아동을 위한 진단·평가 도구(KEDAS, 전병운, 조광순, 이준석, 강미라, 2005) 등이 표준화되었고, 유아발달검사(윤치연, 2005), 발달선별검사(이준석, 2005a)가 표준화 과정 중에 있다.

발달검사에 대해 이같이 많은 관심을 기울이는 이유는 영유아의 발달 초기의 특성 때문이다. 초기 영유아 발달기에는 장애의 증상과 징후가 분명하게 드러나지 않고 전반적인 심리적·신체적 발달 지연 증상을 보이거나 일부 증상만 나타나는 경우도 있어서 임상적으로 알아보기가 힘들며 진단이 쉽지 않다(김진경, 2001). 장애상태로 발전될 위험성을 가진 아동을 조기에 적절한 절차에 따라 진단·평가를 하고, 해당 아동의 요구에 적절한 중재나 교육 서비스를 제공하면 장애를 예방하거나 최소화할 수 있고 더 나아가 국가의 특수교육 및 사회복지 예산을 절감하는 효과를 가져올 수 있다(이미선, 조광순, 김주영, 강병호, 2001).

발달지체 영·유아 선별 및 진단평가에 대한 고찰에서 노진아(2005)는 선별 및 진단평가를 위한 표준참조형 검사와 교육과정중심 진단 및 평가도구 등 다양한 검사도구들의 개발이 필요하며, 진단과 연계된 교육과정 마련이 요구된다고 주장하고 있다. 이준석과 조광순(2004)의 연구에서는 국내 영유아 발달선별 및 진단 검사도구에 대한 문제점들로 영역적절성, 장애 영유아 포괄성, 심리측정적 특성 등을 들고 있다. 검사도구의 심리측정적 특성에 대해서 Nuttall, Romero, 및 Kalesnik(1999)는 검사도구의 표준화과정, 환산점수표의 연령분화, 신뢰도, 타당도, 하위영역, 문항변화도 등의 기준을 제시하고 있다. 대부분의 발달 검사들은 검사-재검사 신뢰도와 내적일치도 등을 보고하고 있으나 구체적인 내용타당도 확인과정, 구성관련타당도 검증 절차 등의 내용이 보고되지 않은 경우도 있었다. 타당도와 관련된 자료들을 수집하는 일은 관련 이론이나 관심 구인에 대한 충분한 경험, 타당도 연구의 설계, 자료 분석과 해석 등의 종합적인 지식이 필요하기 때문에(오수학, 김병준, 2002) 결코 쉽지 않은 과정이지만 심리적 개념의 이론화 단계에서 반드시 해결되어야 할 문제이다.

심리측정도구의 타당도에 관한 새로운 타당화 프로그램을 개념검토, 내적구조검토, 외적관계검토 등의 삼단계로 제시한 오수학과 김병준(2002)은 심리측정도구의 타당화에서 내적구조검토가 가장 핵심적인 단계라고 주장하였다. 내적 구조란 관찰된 변인들 간의 내적 관계를 의미하는 것으로 관찰된 변인들간 내적일관성이 가장 큰 관심거리이다 (Benson, 1998). 연구경비, 연구시간 등의 실제적인 문제들을 고려하여 심리측정도구의 타당화 목적에 따라 요인분석, 일반화가능도 이론, 중다특성-중다방법 행렬, 문항반응이론 등 여러 방법들을 선별적으로 사용할 수 있다.

일반적으로 요인분석과정을 통해 내적구조를 검토하지만, 요인분석결과는 개별 문항에 대한 검증이라기보다는 검사도구 전반에 걸친 정보의 요약에 해당하므로 발달검사를 구성하고 있는 개별 문항의 연령별 타당성을 보장해주지는 않는다. 발달 검사의 경우에는 연령에 따른 문항 반응 차별성이 발달의 규준적 의미를 함축하므로 개별 문항의 검토가 필수적이다. 연령에 따라 해당 문항은 천정효과나 바닥효과를 나타내야 할 것이다. 이같이 구체적으로 연령에 따른 문항의 타당성이 검증되어야 비로소 검사 전반에 걸친 내적 구조 검토도 의미를 갖게 될 것이다.

이 연구에서는 이준석(2005a)에 의해 개발되고 표준화과정 중에 있는 발달선별검사중 인지영역을 중심으로 개별 문항들의 타당성을 연령에 따른 변화를 통해 검토해 보고자 하였다.

II. 연구 방법

1. 연구대상

전국의 0~71개월의 영유아를 대상으로 표집 하였으며, 통계청의 2000년 인구통계자료를 참고하여 표집대상을 6개월 단위로 150명씩을 표집기준으로 설정하였다. 대도시와 중소도시를 중심으로 놀이방, 어린이집, 유치원, 개인 의원, 영아 전담 어린이집, 교회(영아부) 등에서 자료를 수집하였으며, 인구 비례에 의거하여 읍면 단위 지역에 있는 영유아들도 포함시켰다. 수집된 전체 2,340명의 표본자료에서 허위반응이 의심되거나 정상적인 응답이라고 보기 어려운 응답을 제외하는 자료 충실성 검토 작업을 거쳐 2,050명의 자료만을 추출하여 분석에 포함하였으며, 발달장애나 언어장애 및 발달지체가 있는 것으로 진단이 된 영유아들은 표집대상에 포함하지 않았다. 2,050명 영유아의 성별, 연령별 분포는 표 1과 같다.

<표 1> 연구 대상의 평균 연령 및 사례수

연령		0-5 개월	6-11 개월	12-17 개월	18-23 개월	24-29 개월	30-35 개월	36-41 개월	42-47 개월	48-53 개월	54-59 개월	60-65 개월	66-71 개월	전체
사례수	남자	65	61	69	70	88	93	111	122	114	82	82	89	1046
	여자	59	58	64	60	70	85	111	96	120	87	96	98	1004
	전체	124	119	133	130	158	178	222	218	234	169	178	187	2050
평균연령(개월)		2.76	8.63	14.51	20.43	26.68	32.58	38.80	44.27	50.35	56.25	62.48	68.60	38.76

2. 연구도구

발달선별 검사는 발달장애아동을 위한 진단검사(KEDAS, 전병운 외, 2005)를 토대로 하고 있다. KEDAS는 IDEA 규정에 따라 발달검사영역을 인지, 사회-정서, 운동, 적응, 의사소통 등 다섯 영역으로 분류하여 표준화하였으며, 정상시의 아동관찰과 부모나 교사와의 면담을 통해 알아 볼 수 있도록 체크리스트 형태로 개발되었다. 각 문항은 0점(전혀 할 수 없음)과 1점(부분적으로, 가끔 할 수 있음), 2점(항상 할 수 있음)의 3점 평정척도를 이용하여 평정하도록 구성되어 있다.

<표 2> 발달진단검사의 영역, 하위 영역 및 선별검사의 문항수

발달진단검사의 영역		발달진단검사 문항 수	발달선별검사 문항 수
인지 영역	기억	84	18
	추리 및 학문기술	27	
	개념발달	19	
		38	
사회·정서 영역	타인과의 상호작용	94	18
	감정·정서의 표현	32	
	자아개념	19	
	대처기능 및 사회적 역할	23	
		20	
운동 영역	큰 운동	76	18
	작은 운동	42	
		34	
적응 영역	식사	65	18
	착탈의	23	
	용변·위생	21	
		21	
의사소통 영역	이해	63	18
	표현	29	
		34	
전 체		382	90

발달 선별 검사는 난이도와 문항-영역점수간 상관점수를 토대로 다섯 영역에서 선택한 90개 문항(연령 수준당 2문항, 영역별 18문항)으로 구성되어 있다(표 2 참조). 의사소통의 문항은 각 연령 수준별로 이해와 표현 기술을 대표할 수 있도록 선택하였고, 운동

영역은 큰 운동(대근육)과 작은 운동(소근육)을 대표할 수 있도록 연령별로 2문항씩 채택하였다(발달선별검사에 대한 심리측정학적 정보는 이준석(2005a) 참조).

발달진단검사의 인지 영역은 개념적 기술과 능력을 평가하는 문항들로 구성되어 있으며 기억력, 추리 및 학문기술, 개념발달, 3개의 하부 영역으로 분류된다. 기억력은 관련 단서를 제시할 때 정보를 상기해 내는 능력을 평가하며, 추리 및 학문기술은 문제를 인지하여 해결할 수 있는 사고능력과 상황에 대한 분석과 평가, 학습 성취에 필요한 학문적 기능 등을 평가한다. 개념발달은 사물에 대한 개념을 이해하고 사물들 간의 관계를 도출해내는 능력을 평가한다.

발달선별검사의 인지영역은 진단검사의 인지행동 하위영역을 고려하지 않고 난이도와 문항-영역 총점간 상관점수를 토대로 연령당 2문항, 전체 18문항을 선별하였다.

3. 분석 방법

검사문항을 평가하기 위해 고전검사이론을 바탕으로 문항난이도와 문항변별도를 검토하였으며, 문항반응이론에 기반한 문항난이도 분석도 함께 실시하였다. 신뢰도는 Cronbach의 α 를 사용하여 문항의 내적일치도를 추정하였다. 그리고 연령별, 성별로 반응 차이가 있는 지를 검증하기 위해 χ^2 분석을 하였다. 영가설이 기각된 경우에는 사후분석(성태제, 2001)을 실시하였다.

III. 결과

1. 문항의 난이도와 변별도

고전적인 문항난이도 분석을 위해 ‘거의 할 수 없음’(0)을 오반응으로, ‘때때로 (절반정도) 그렇게 할 수 있음’(1)과 ‘항상 그렇게 할 수 있음’(2)을 정반응으로 각 반응에 0.5, 1.0 가중치를 부여하고, 전체반응 아동수로 나누어 정리하여 표 3에 제시하였다. 문항반응이론을 이용한 문항난이도 분석은 문항반응이론의 2모수 로지스틱 모형을 이용하여 난이도 모수와 변별도 모수를 추정하였고 난이도 모수를 기준으로 문항을 배열하였다. 문항번호가 증가할수록 문항난이도는 증가하는 것으로 나타났다.

문항변별도는 발달선별 검사의 전체점수를 기준으로 전체 대상자들을 상위 33%, 중간 33%, 하위 33% 세 집단으로 나누어 이 세 집단간 빈도차이를 χ^2 검증으로 확인하는 방법과 문항 점수와 총점 간의 양류상관계수로 추정하는 방법을 사용하였다. χ^2 값을 검증한 결과, 모든 문항이 통계적으로 유의한 것으로 나타나서 세 집단간 차이를 잘 변별하

는 것으로 나타났다. 양류상관계수의 유의성 검증결과도 모든 문항이 통계적으로 유의한 것으로 나타났다.

<표 3> 발달선별검사의 인지영역 문항난이도와 문항변별도

문항 번호		1	2	3	4	5	6	7	8	9
난이도	고전	.94	.92	.92	.89	.88	.87	.85	.81	.77
	IRT	-2.95	-2.52	-2.06	-1.60	-1.40	-1.15	-1.13	-.98	-.54
변별도	χ^2 값	71.92	153.85	158.89	215.52	353.31	426.07	547.39	543.38	1157.56
	상관	.33	.37	.44	.50	.59	.60	.65	.64	.75
문항 번호		10	11	12	13	14	15	16	17	18
난이도	고전	.76	.61	.60	.58	.55	.49	.49	.37	.30
	IRT	-.48	-.17	.06	.13	.26	.57	.70	1.08	1.15
변별도	χ^2 값	1219.94	1691.35	1521.38	1792.33	1602.17	1616.76	1508.19	1275.77	974.51
	상관	.77	.77	.79	.77	.77	.71	.65	.57	.56

$p < .001$

2. 문항의 내적 일치도

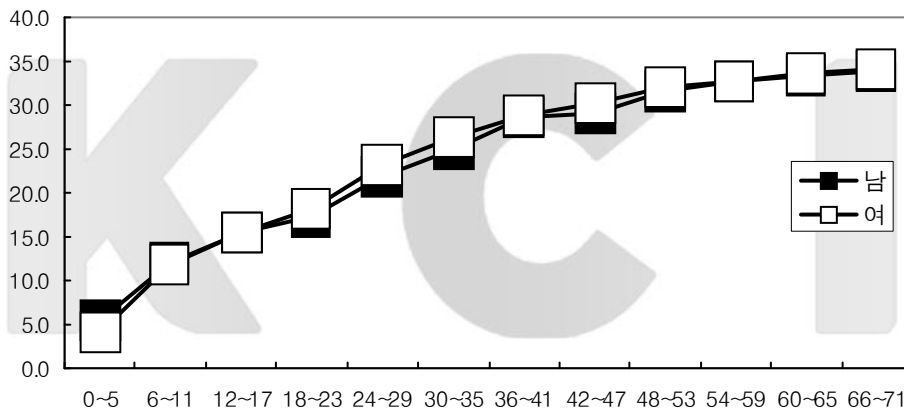
표 4는 전체 18개 문항의 내적 일치도 분석 결과를 보여주고 있다.

<표 4> 발달선별검사의 인지영역 문항 내적 일치도 분석결과

문항 번호	문항 평균	문항 분산	항목이 삭제된 경우 척도 분산	항목이 삭제된 경우 척도 분산	수정된 항목-전체 상관 관계	항목이 삭제된 경우 Cronbach의 α
인지 1	1.88	.140	83.742	83.742	.274	.942
인지 2	1.82	.219	83.272	83.272	.266	.942
인지 3	1.84	.218	82.112	82.112	.405	.940
인지 4	1.77	.316	80.643	80.643	.475	.939
인지 5	1.80	.297	79.651	79.651	.598	.938
인지 6	1.66	.432	78.351	78.351	.599	.937
인지 7	1.75	.386	78.097	78.097	.662	.936
인지 8	1.76	.351	78.561	78.561	.652	.937
인지 9	1.54	.642	74.158	74.158	.794	.933
인지10	1.52	.643	74.033	74.033	.803	.933
인지11	1.22	.835	72.228	72.228	.817	.933
인지12	1.35	.782	72.350	72.350	.839	.932
인지13	1.17	.786	72.471	72.471	.828	.932
인지14	1.25	.753	72.625	72.625	.837	.932
인지15	.98	.783	73.324	73.324	.769	.934
인지16	.98	.713	74.017	74.017	.759	.934
인지17	.75	.675	75.402	75.402	.679	.936
인지18	.60	.681	76.779	76.779	.574	.938

전체 문항의 Cronbach의 α 값은 .940 으로 적절한 수준의 내적 일치도를 보이고 있다. '수정된 항목-전체 상관 관계'의 경우, 문항2와 나머지 문항들의 합산 점수와의 상관계수가 .266으로 가장 낮게 나타났고, 문항12와 나머지 문항들의 합산 점수와의 상관계수가 .839로 가장 높게 나타났다. '항목이 삭제된 경우 Cronbach의 α '를 살펴보면 문항1과 문항2가 삭제될 때 Cronbach의 α 값이 증가하는 것으로 나타났으며, 문항3의 경우에는 동일하게, 그리고 다른 문항들은 감소하는 것으로 나타났다. 문항1, 2, 3은 출생 직후부터 나타나며, 1세 이전에 완결되는 인지행동을 측정하는 문항으로 난이도가 매우 낮은 문항이기 때문에 1세 이상의 유아와 아동의 문항반응 분산이 작아서 생긴 결과이다. 이는 해당 문항들의 분산값이나 '항목이 삭제된 경우 척도 분산'이 감소하는 정도가 다른 문항에 비해서 작은 것으로도 확인할 수 있다.

3. 연령에 따른 인지 점수의 변화



<그림 1> 연령대별 반응(인지 전체점수)

인지 영역에서 연령별 차이가 있는 지 검증하기 위해 분산분석을 실시한 결과, $F(11, 2038) = 919.87, p = .000$ 으로 연령별로 인지점수가 다른 것으로 나타났다. 연령별로 차이가 있는 지를 검증하기 위해 6개월 단위로 집단을 구분하여 도식화한 것이 그림 1이다. 구체적인 차이를 검토하기 위해 Scheffe 검증을 실시하였더니, 0-5개월, 6-11개월, 12-17개월, 18-23개월, 24-29개월, 30-35개월 집단은 각각 다른 연령집단과 유의한 차이가 있는 것으로 나타났으며, 36-41개월과 42-47개월 집단간 차이는 통계적으로 유의하지 않았고, 48-53개월 집단은 54-59개월 집단과 유의한 차이를 보이지 않았다. 그리고 54-59개월, 60-65개월, 66-71개월 집단은 유의한 차이를 보이지 않았다.

남녀간 성별차이가 있는지를 알아보기 위해 t검증을 실시한 결과, $t(2048) = 2.410, p$

= .016 으로 통계적으로 유의한 차이가 있는 것으로 나타났다. 즉 여아(M=26.13)가 남아(M=25.14)보다 인지점수가 더 높은 것으로 나타났다.

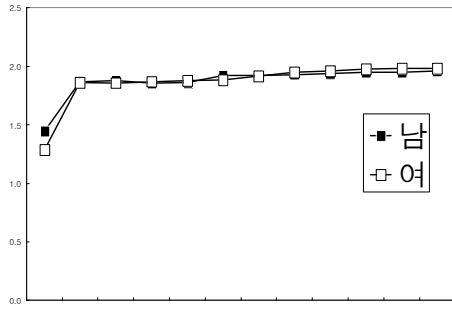
4. 연령에 따른 문항별 반응 변화

문항별로 연령별·성별 차이를 살펴보았다. 문항1(낮선 장소에 가면 고개를 돌려 주위를 살핀다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2(22, N = 2050) = 358.103, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 그림 2는 연령에 따른 문항1의 반응 변화를 (측정척도를 기준으로 하면 막대그래프를 사용하는 것이 더 적절하지만) 시각적으로 알아보기 쉽도록 선 그래프로 나타낸 것이다. 어느 연령집단간 응답비율에서 차이가 있는지를 검증하기 위해 '2'점(항상 할 수 있음) 반응을 기준으로 연령집단간 짝비교(pairwise comparison)를 실시한 결과에서 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며 다른 연령집단들간 차이는 유의하지 않았다.

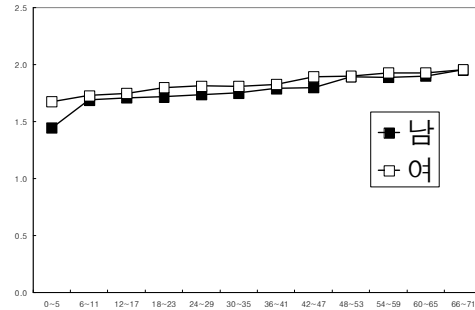
남녀간 성별차이가 있는지를 알아보기 위해 χ^2 검증을 실시한 결과, $\chi^2(2, N = 2050) = .876, p = .645$ 로 통계적으로 유의하지 않은 것으로 나타났다.

문항2(엄마가 안고 젖을 먹이려 하면 젖꼭지를 찾아 문다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2(22, N = 2050) = 130.752, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 유의한 경향성을 보이는 것으로 나타났다. 그림 3은 문항2의 연령별 반응변화를 보여준다.

남녀간 성별차이가 있는지를 알아보기 위해 χ^2 검증을 실시한 결과, $\chi^2(2, N = 2050) = 10.013, p = .007$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '1'점 반응에서는 남녀 차이가 나타나지 않았으며, '2'점 반응에서는 여아들이 남아들보다 응답비율이 조금 더 높게 나타나서 통계적 경향성을 보여주었다. '0'점 반응에서는 남아들의 4.8%(50명), 여아들의 2.3%(23명)가 응답하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 젖꼭지를 찾아 무는 행동의 발달이 조금 더 빠른 것으로 나타났다.



<그림 2> 연령대별 반응(문항1)



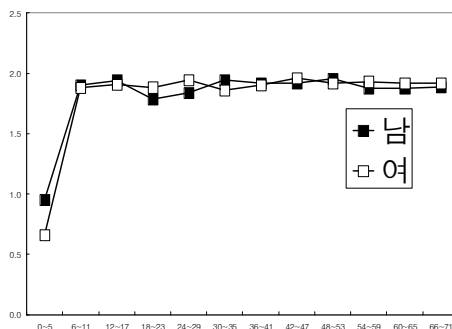
<그림 3> 연령대별 반응(문항2)

문항3(모빌에 달린 장난감을 잡기 위해 손을 뻗는다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 841.207, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며 다른 연령집단들간 차이는 유의하지 않았다. 그림 4는 문항3의 연령별 반응변화를 그래프로 나타낸 것이다.

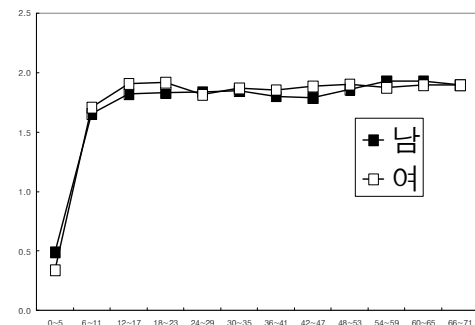
남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 1.766, p = .413$ 으로 통계적으로 유의하지 않은 것으로 나타났다.

문항4(손이 닿지 않는 곳에서 구르는 물체를 쳐다보고 잡으려 한다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 942.623, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며, 다른 연령집단들간 차이는 유의하지 않았다. 그림 5는 문항4의 연령별 반응변화를 보여준다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 2.055, p = .358$ 로 통계적으로 유의하지 않은 것으로 나타났다.



<그림 4> 연령대별 반응(문항3)



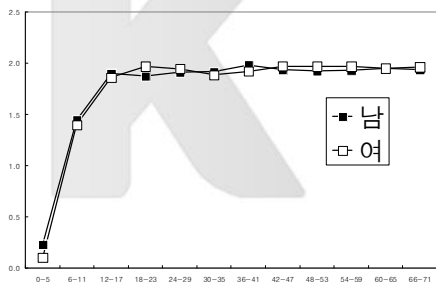
<그림 5> 연령대별 반응(문항4)

문항5(과자를 컵 안에 숨기면 그 과자를 찾는다.)에서 연령집단별 차이를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1497.306, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며, 6-11개월과 12개월 이상 연령집단간 차이도 모두 유의한 것으로 나타났다. 그림 6은 문항5의 연령별 반응변화를 그래프로 나타낸 것이다.

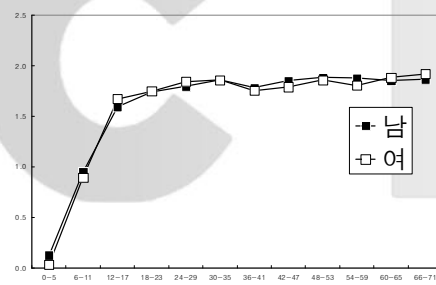
남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = .640, p = .726$ 으로 통계적으로 유의하지 않은 것으로 나타났다.

문항6(새로운 장난감에 호기심을 보인다.)의 연령집단별 차이를 χ^2 검증 실시 결과, $\chi^2 (22, N = 2050) = 1222.964, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며, 6-11개월과 12개월 이상 연령집단간 차이도 모두 유의한 것으로 나타났다. 그림 7은 문항6의 연령별 반응변화를 보여준다.

성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 2.997, p = .223$ 으로 통계적으로 유의하지 않은 것으로 나타났다.



<그림 6> 연령대별 반응(문항5)



<그림 7> 연령대별 반응(문항6)

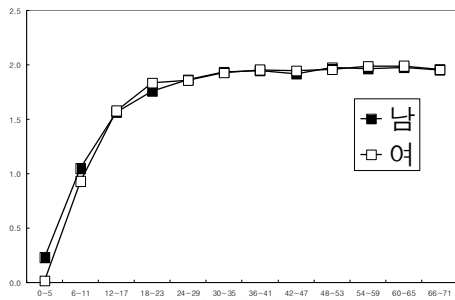
문항7(공이 시야 밖으로 굴러가면 공을 따라 주시한다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1339.840, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 집단간 차이가 모두 유의한 것으로 나타났으며, 6-11개월과 12-17개월간 연령 차이는 유의하지 않지만 6-11개월과 18개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 8은 문항7의 연령별 반응변화를 그래프로 나타낸 것이다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = .065, p = .968$ 로 통계적으로 유의하지 않은 것으로 나타났다.

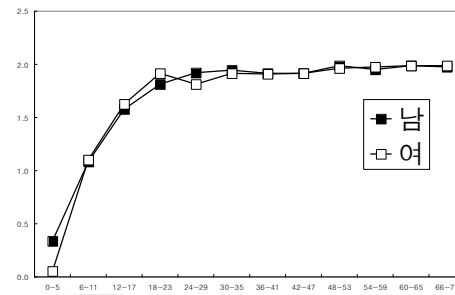
문항8(병 안에 있는 물건[예: 사탕]을 꺼내려고 병을 뒤집는다.)에서 연령집단별 차이를 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1359.863, p = .000$ 으로 통계적으로 유

의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월과 다른 연령 집단간 차이가 모두 유의하였으며, 6-11개월과 12-17개월간 연령 차이는 유의하지 않지만 6-11개월과 18개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 9는 문항8의 연령별 반응변화를 보여준다.

성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 1.593, p = .451$ 로 통계적으로 유의하지 않은 것으로 나타났다.



<그림 8> 연령대별 반응(문항7)



<그림 9> 연령대별 반응(문항8)

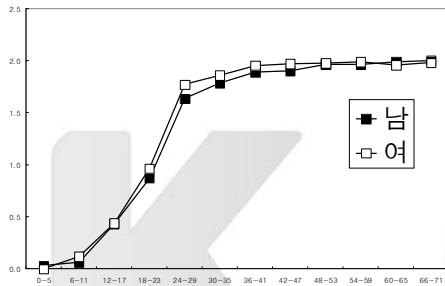
문항9(한 자리 숫자 두 개[예: 2, 4]를 정확하게 따라서 말한다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1706.988, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월, 6-11개월, 12-17개월 집단간 차이가 유의하지 않았으나, 0-5개월, 6-11개월 집단과 18개월 이상 연령집단간 차이가 모두 유의하였으며, 12-17개월 집단과 24개월 이상 연령집단간 차이 및 18-23개월 집단과 30개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 10은 문항9의 연령별 반응변화를 그래프로 나타낸 것이다.

남녀간 성별차이 χ^2 검증을 실시한 결과, $\chi^2 (2, N = 2050) = 7.857, p = .020$ 으로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '0'점 반응에서는 남녀 차이가 나타나지 않았으나, '1'점 반응에서는 남아들의 8.5%(89명), 여아들의 5.7%(57명)가 응답하여 통계적으로 유의한 것으로 나타났으며, '2'점 반응에서는 여아들(75.5%)이 남아들(71.0%)보다 조금 더 많이 응답한 것으로 나타나서 통계적 경향성을 보여주었다. 즉 여아들이 남아들보다 한 자리 숫자 두 개를 정확하게 따라하는 행동의 발달이 조금 더 빠른 것으로 보인다.

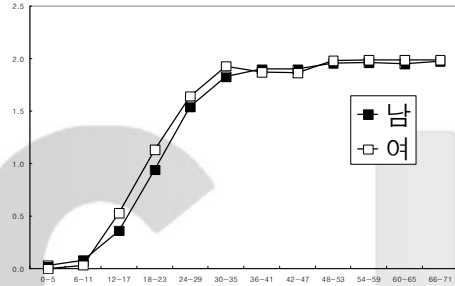
문항10("너[철수]는 누구야?"라고 물으면 "철수야[나야]"하고 말하거나 자신을 가리킨다.)에 대한 연령집단별 차이를 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1671.815, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한

결과, 0-5개월, 6-11개월, 12-17개월 집단간 차이가 유의하지 않았으나, 0-5개월, 6-11개월 집단과 18개월 이상 연령집단간 차이가 모두 유의하였으며, 12-17개월 집단과 24개월 이상 연령집단간 차이 및 18-23개월 집단과 30개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 11은 문항10의 연령별 반응변화를 보여준다.

남녀간 성별차이 χ^2 검증을 실시한 결과, $\chi^2 (2, N = 2050) = 9.401, p = .008$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '0'점 반응에서는 유의한 남녀 차이가 나타나지 않았으나, '1'점 반응에서는 남아들의 10.0%(105명), 여아들의 6.9%(69명)가 응답하였고, '2'점 반응에서는 여아들의 74.6%(749명), 남아들의 69.2%(724명)가 응답하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 자신에 대한 질문에 더 정확하게 반응하는 것으로 나타났다.



<그림 10> 연령대별 반응(문항9)



<그림 11> 연령대별 반응(문항10)

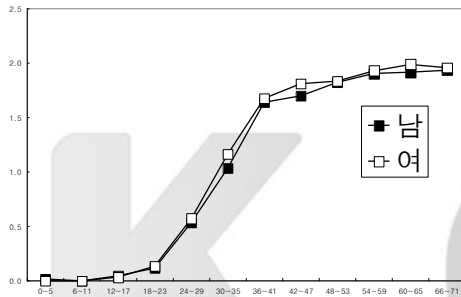
문항11(하나, 둘, 셋을 순서대로 센다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1699.247, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월, 6-11개월, 12-17개월, 18-23개월 집단간 차이가 유의하지 않았으나, 이 연령 집단들과 30개월 이상 연령집단간 차이가 모두 유의하였으며, 24-29개월 집단과 36개월 이상 연령집단간 차이, 그리고 30-35개월 집단과 36개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 12는 문항11의 연령별 반응변화를 그래프로 나타낸 것이다.

남녀간 성별차이를 χ^2 검정한 결과, $\chi^2 (2, N = 2050) = 7.171, p = .028$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '0'점과 '1'점 반응에서는 유의한 남녀 차이가 나타나지 않았으나, '2'점 반응에서는 여아들의 58.4%(586명), 남아들의 52.5%(549명)가 응답한 것으로 나타나서 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 순서대로 숫자를 헤아리는 행동을 더 빨리 하는 것으로 나타났다.

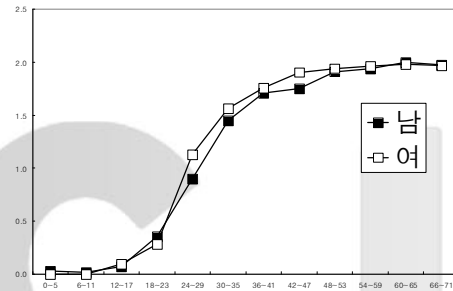
문항12(여러 모양의 도형을 제시하면, 같은 모양끼리 짝짓는다.)의 연령집단별 차이에 대해 χ^2 검증을 실시한 결과, $\chi^2 (22, N = 2050) = 1742.631, p = .000$ 으로 통계적으로

유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월, 6-11개월, 12-17개월 집단과 24개월 이상 연령집단간 차이가 모두 유의하였으며, 18-23개월 집단과 36개월 이상 연령집단간 차이, 그리고 24-29개월 집단과 36개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났다. 그림 13은 문항12의 연령별 반응변화를 보여준다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2(2, N = 2050) = 6.611, p = .037$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '0'점과 '1'점 반응에서는 유의한 남녀 차이가 나타나지 않았으나, '2'점 반응에서는 여아들의 65.2%(655명), 남아들의 59.8%(625명)가 반응하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 도형 짝짓기를 더 잘하는 것으로 나타났다.



<그림 12> 연령대별 반응(문항11)



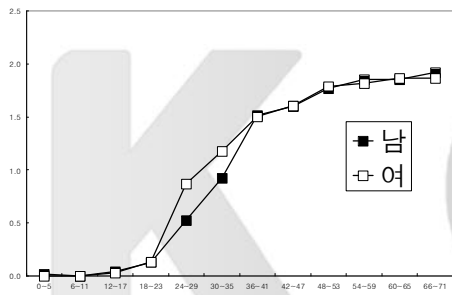
<그림 13> 연령대별 반응(문항12)

문항13(물건 세 개[예: 컵, 인형, 사탕, 열쇠]를 보여주고 모두 숨기면 숨겨진 물건의 이름을 모두 말한다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증을 실시한 결과, $\chi^2(22, N = 2050) = 1632.791, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월, 6-11개월, 12-17개월, 18-23개월 집단간 차이는 유의하지 않았으며, 이 연령집단들과 30개월 이상 연령집단간 차이는 모두 유의하였다. 그리고 24-29개월 집단과 36개월 이상 연령집단간 차이도 모두 유의한 것으로 나타났다. 그림 14는 문항13의 연령별 반응변화를 그래프로 나타낸 것이다.

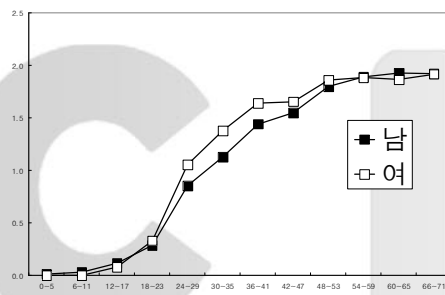
성별차이 χ^2 검증 결과는 $\chi^2(2, N = 2050) = 8.824, p = .012$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '1'점 반응에서는 유의한 남녀 차이가 나타나지 않았으나, '2'점 반응에서는 여아들의 51.5%(517명), 남아들의 46.5%(486명)가 반응하여 통계적 경향성을 보이고 있고, '0'점 반응에서는 남아들의 35.4%(370명), 여아들의 29.3%(294명)가 반응하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 숨겨진 대상을 찾는 데 필요한 대상 영속성 개념이 더 빨리 나타나는 것으로 보인다.

문항14(“왜 사용하지?”라는 질문에 정확히 대답한다[예: “왜 난로를 켜지?”, “왜 우산을 쓰지?”].)에서 연령집단별 차이를 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1678.867, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 0-5개월, 6-11개월, 12-17개월, 18-23개월 집단간 차이는 유의하지 않았으며, 이 연령집단들과 24개월 이상 연령집단간 차이는 모두 유의하였다. 그리고 24-29개월 집단과 36개월 이상 연령집단간 차이도 모두 유의한 것으로 나타났다. 그림 15는 문항 14의 연령별 반응변화를 보여준다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 10.265, p = .006$ 으로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과에서 '0'점과 '1'점 반응은 유의한 남녀 차이가 나타나지 않았으나, '2'점 반응에서는 여아들의 56.8%(570명), 남아들의 49.7%(520명)가 반응하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 이유를 더 명확하게 표현하는 것으로 나타났다.



<그림 14> 연령대별 반응(문항13)



<그림 15> 연령대별 반응(문항14)

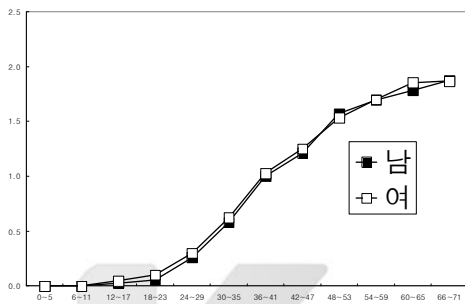
문항15(여러 개의 물건 중에서 함께 사용되는 물건을 짝짓는다[예: 망치와 못, 신발과 양말, 실과 바늘].)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1554.285, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 29개월 이하 집단과 36개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났으며 42-47개월 집단과 54개월 이상 집단간 차이도 유의한 것으로 나타났다. 그림 16은 문항15의 연령별 반응변화를 그래프로 나타낸 것이다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 3.119, p = .210$ 으로 통계적으로 유의하지 않은 것으로 나타났다.

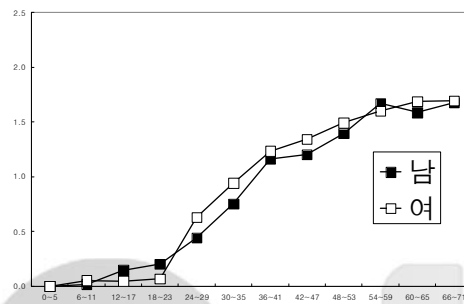
문항16(악기의 소리를 듣고 그 악기의 이름을 말한다.)의 연령집단별 차이를 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1365.394, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 23개월 이하 집단과 36개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났으며, 30-35개월 집단과 48개월 이상 집

단간 차이도 유의한 것으로 나타났다. 그림 17은 문항16의 연령별 반응변화를 보여준다.

성별차이 χ^2 검증 결과, $\chi^2 (2, N = 2050) = 9.159, p = .010$ 으로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과, '1'점 반응에서 유의한 남녀 차이가 나타나지 않았으나, '0'점 반응에서는 남아들의 39.1%(409명), 여아들의 33.7%(338명)가 반응하였으며, '2'점 반응에서는 여아들의 37.8%(380명), 남아들의 32.0%(335명)가 반응하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 악기와 악기소리에 대한 개념 짚짓기를 더 잘하는 것으로 나타났다.



<그림 16> 연령대별 반응(문항15)



<그림 17> 연령대별 반응(문항16)

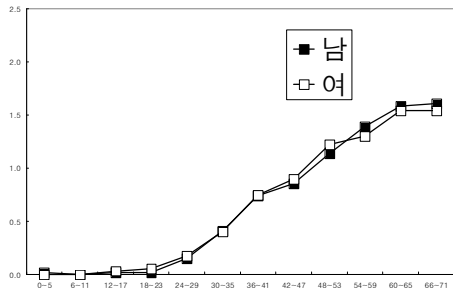
문항17(자신의 이름과 주소를 말한다.)에서 연령집단별로 차이가 있는지를 검증하기 위해 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1210.808, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 23개월 이하 집단과 36개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났으며, 24-29개월 집단과 42개월 이상 집단, 그리고 30-35개월 집단과 48개월 이상 집단간 차이는 모두 유의한 것으로 나타났다. 그림 18은 문항17의 연령별 반응변화를 그래프로 나타낸 것이다.

남녀간 성별차이를 χ^2 검증한 결과, $\chi^2 (2, N = 2050) = 3.308, p = .191$ 로 통계적으로 유의하지 않은 것으로 나타났다.

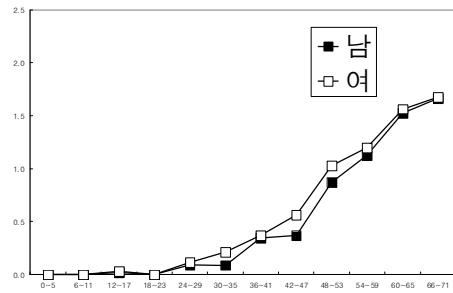
문항18(요일 이름을 순서대로 말한다.)에서 연령집단별 차이를 χ^2 검증한 결과, $\chi^2 (22, N = 2050) = 1214.271, p = .000$ 으로 통계적으로 유의한 차이를 보여주었다. 구체적 차이 검증을 위해 짝비교를 실시한 결과, 41개월 이하 집단과 48개월 이상 연령집단간 차이는 모두 유의한 것으로 나타났으며, 42-47개월 집단과 54개월 이상 집단, 그리고 48-53개월 집단과 54개월 이상 집단간 차이는 모두 유의한 것으로 나타났다. 그림 19는 문항18의 연령별 반응변화를 보여준다.

남녀간 성별차이 χ^2 검증 결과, $\chi^2 (2, N = 2050) = 11.242, p = .004$ 로 통계적으로 유의한 것으로 나타났다. 구체적인 차이검증을 실시한 결과, '1'점 반응에서 유의한 남녀 차이가 나타나지 않았으나, '0'점 반응에서는 남아들의 65.4%(684명), 여아들의

58.7%(589명)가 반응하였으며, '2'점 반응에서는 여아들의 24.8%(249명), 남아들의 19.3%(202명)가 반응하여 통계적으로 유의한 차이를 보여주었다. 즉 여아들이 남아들보다 요일과 요일 순서 개념에 대해 더 많이 알고 있는 것으로 나타났다.



<그림 18> 연령대별 반응(문항17)



<그림 19> 연령대별 반응(문항18)

IV. 논의

이 연구는 짧은 시간 안에 간편하게 영유아의 발달 상태를 살펴볼 수 있도록 표준화 중인 발달선별검사의 인지영역 타당도를 문항별로 살펴보고자 하였다. 다른 검사들에 비해 발달 검사에서는 연령별 타당도가 중요한 함의를 지니므로 개별 문항을 평가하기 위해 문항분석(문항난이도, 문항변별도), 문항 내적 일치도 및 연령별, 성별 차이검증을 실시하였다.

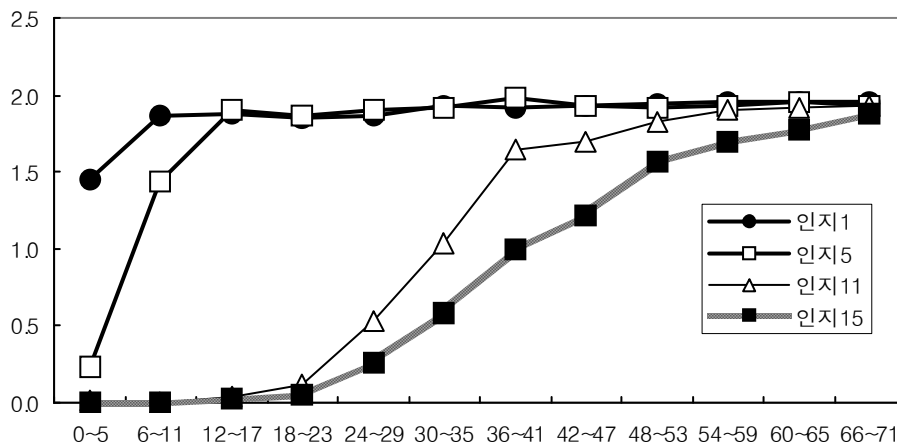
고전적인 문항난이도 분석과 문항반응이론을 이용한 문항난이도 분석을 통해 살펴본 문항의 난이도는 문항이 배열된 순서에 따라 점진적으로 증가하는 것으로 나타났다. 또한 분류된 집단 차이를 잘 변별하는 것으로 나타났으며, 양류상관계수도 모두 유의한 것으로 나타나서 양호한 문항변별도를 보여주는 것으로 나타났다.

신뢰도를 알아보기 위해 실시한 문항의 내적 일치도 분석에서 전체 문항의 Cronbach의 α 값은 .940으로 높게 나타났고, 출생 직후의 인지행동을 측정하는 문항1, 2, 3을 제외하고는 문항이 삭제될 때 Cronbach의 α 값이 감소하는 것으로 나타나 문항들이 내적 일관성을 가지고 있는 것으로 나타났다.

연령에 따른 인지능력 변화를 살펴보기 위해 6개월 단위로 집단을 구분하여 연령별 차이를 검증한 결과, 태어나서 35개월까지 영아들의 인지 능력은 지속적으로 증가하였으며, 36개월 이후 조금씩 증가폭이 줄어들어 54개월 이후에는 인지능력이 일정 수준을 유지하는 것으로 나타났다. 이는 유아의 지능검사에서 나타나는 하위검사점수의 완만한 증가(예: 문수백, 변창진, 1997)와도 수렴되는 결과이다.

인지영역 전체점수를 종속변인으로 남녀차이를 검증한 결과에서는 여아의 인지점수가 남아보다 더 높은 것으로 나타났다. 이준석(2005b)의 운동영역 타당화 연구에서는 전체 점수에서 남녀 성별차이가 나타나지 않았으나 소근육 운동점수에서는 여아의 점수가 남아점수보다 높았으며, 적응행동검사의 실제적 기술에서 남학생보다 여학생의 인지점수가 높은 것으로 나타난 결과와 일치한다(이준석, 유재연, 신현기, 전병운, 고등영, 2005). 연령에 따른 문항별 반응변화를 성별로 구분하여 살펴보았더니 숫자개념, 대상영속성 개념, 원인 찾기, 요일개념 등 대부분의 영역에서 여아가 남아보다 조금 더 높은 점수를 보여주었다.

0~71개월 아동을 대상으로 한 인지영역의 문항별 연령 변화를 요약해 보면 그림 20에서 볼 수 있는 바와 같이 크게 네 종류로 나눌 수 있다. 첫 번째 유형의 문항은 인지에서 기본이 되는 개념을 측정하는 문항으로 장애아동을 진단하기 위해 검사를 실시하는 경우를 대비하여 선택한 문항이다. 문항 1번과 문항 2번이 여기에 해당하는 문항이며, 출생 직후 대부분의 영아들이 반응을 보이는 문항 유형이다. 두 번째 유형의 문항은 18개월 이전 영아들에게 해당하는 문항으로 생후 18개월 이내 행동이 완결되는 경우이다. 이는 18개월 이전 영아들이 반응할 수 있는 문항이기 때문에 18개월 이후 영유아들의 반응은 천정효과를 보여주고 있다. 문항 3에서 문항 8번까지 여섯 문항이 여기에 해당된다. 세 번째는 문항 9번부터 문항 12번까지 12개월 이후 행동이 나타나기 시작해서 약 4세 경에 이르러 행동이 완성되는 유형이다. 일반적으로 점진적인 변화를 보여주는 발달의 대표적인 형태라고 할 수 있다. 마지막으로 문항 13번부터 문항 18번까지 해당하는 유형으로 24개월 이후 66개월 이상까지 행동빈도가 증가하는 경우이다. 3세 이상의 유아들이 수행할 수 있는 문항이기 때문에 18개월 이전 영유아들은 전혀 수행하지 못하는 바닥효과를 보여준다.



<그림 20> 문항별 연령변화 유형

이 같은 문항 유형분류를 통해 확인된 발달선별검사의 인지영역 개별 문항들의 타당성을 토대로 탐색적 요인분석이나 확인적 요인분석, 일반화가능도 이론을 이용한 D-연구 분석 등을 통해 검사 전반에 걸친 내적 구조 검토를 거쳐야 좀 더 의미 있는 결과를 얻을 수 있을 것이다. 현재까지 개발된 검사들은 일반적으로 개별 문항에 대한 분석보다는 요인분석을 활용한 구성관련타당도 검토에 치중하는 편이며(예, 이정화, 2006), 검사 개발과정상의 편의를 위해 연령별로 일부 문항만을 실시하여 전반적인 타당성을 하위영역간 상호상관계수로 확인하는 경우도 보고되고 있다(예, 이은혜 외, 2005). 그렇지만 검사의 타당도를 한 가지 방법만으로 검증하는 것은 제한적일 수밖에 없다. 요인분석은 검사문항 전반에 걸친 정보를 요약해 주지만 개별 문항의 연령별 타당성을 보장해 주지는 않는다. 개별 문항의 난이도가 적절하다고 해서 전반적인 검사 난이도가 적절하거나 구성개념이 제대로 측정된다고 주장할 수도 없다. 발달 검사의 경우에는 연령에 따른 문항 반응 차별성이 발달의 규준적 의미를 함축하므로 개별 문항의 검토후 전체적인 내적 구조 검토를 거쳐야 할 것이다.

마지막으로 이 연구에서 얻어진 문항들의 연령별 발달경향은 단기간의 횡단 조사를 통하여 구한 자료에 근거하고 있다는 제한점을 염두에 둘 필요가 있을 것이다. 횡단조사 연구는 현재 시점에서 발달 현상만을 제시해 줄 뿐, 시간의 경과에 따른 영유아들의 발달 변화 양상을 밝혀주지는 못하기 때문이다. 따라서 실질적인 발달경향을 살펴보기 위해서는 30년 이상 종단적 연구를 실시한 Bayley(1970; 박혜원, 2006에서 재인용)이나 3년 종단자료를 분석한 박혜원(2006)의 연구처럼 종단 연구를 통해 이 연구에서 살펴본 발달 경향, 그리고 문항의 타당성에 대한 수렴적 증거를 확보할 필요가 있을 것이다.

참고문헌

- 국립교육평가원 (1996). **유치원 교육평가연구-유아발달수준검사**. 서울: 국립교육평가원.
- 국립특수교육원 (1996). **KISE 발달척도**. 경기: 국립특수교육원.
- 국립특수교육원 (1998). **한국의 특수교육 지표**. 경기: 국립특수교육원.
- 김진경 (2001). 현재 보건소에서 사용 중인 영유아 발달선별검사의 비교. **한국모자보건학회 학술대회 연례집 (2001년도 춘계)**, 14-24.
- 노진아 (2005). 발달지체 영·유아 선별 및 진단평가의 최근 연구동향에 관한 고찰. **유아특수교육연구**, 5(1), 5-24.
- 대한소아과학회 (2002). **한국형 영유아 발달 검사**. 대한소아과학회.
- 문수백, 변창진 (1997). **K-ABC 해석요강**. 서울: 학지사.
- 박혜원 (2006). 정상 영아 및 장애 위험 영아의 한국 Bayley 영유아발달검사(K-BSID-2) 수행 비교: 3년 종단자료분석. **아동학회지**, 27(1), 153-166.
- 박혜원, 신민선 (2006). 한국 Bayley 영유아 발달검사수행의 안정성: 종단 연구. **대한가정학회지**, 44(2), 103-112.
- 박혜원, 조복희, 최호정 (2003). 한국 Bayley 영유아 발달검사(K-BSID-II) 표준화 연구: 예비연구. **한**

- 국심리학회지:발달, 16(4), 121-134.
- 배윤희 (2002). 베일리 영유아발달 검사-II의 한국 표준화를 위한 예비연구. 석사학위 논문, 울산대학교 대학원.
- 성태제 (2001). **현대 기초통계학의 이해와 적용**(개정판). 서울: 교육과학사.
- 손원경 (2004). 3·4세용 유아발달검사 개발연구. 박사학위 논문, 부산대학교 대학원.
- 신희선, 한경자, 오가실, 오진주, 하미나 (2002). **한국형 Denver 2 검사지침서**. 서울: 현문사.
- 오가실 (1976). Denver Developmental Screening Test의 한국표준화를 위한 기초 연구. **간호학논집**, 1, 93-116
- 오수학, 김병준 (2002). 심리측정도구의 타당도에 관한 새로운 개념과 검증방법. **한국스포츠심리학회지**, 13(3), 111-126.
- 윤치연 (2005). 유아 발달검사 표준화 예비연구. **정서·행동장애연구**, 21(2), 179-192.
- 이근 (1987). 서울 아동의 덴버 발육 선별검사 및 한국판 DDST. **소아과**, 30(9), 958-971.
- 이근 (1990). **한국판 덴버발달선별검사**. 서울: 이화여자대학교 부속병원.
- 이근 (1996). 덴버II발육선별검사와 서울 아동의 발달에 관한 비교 연구. **소아과**, 39(9), 1210-1215.
- 이근 (2000). 서울 영유아 발달선별검사. **소아과**, 43(3), 335-343.
- 이근 (2002). 이화 영아 발달 선별 검사에 대한 연구, **소아과**, 45(6), 700-711.
- 이미선, 조광순, 김주영, 강병호 (2001). **장애 영·유아 조기발견 및 진단·평가방안 연구**. 안산: 국립 특수교육원.
- 이영자, 이종숙, 신은수, 곽향림, 이정옥 (2001). 교사평정용 유아발달 평가도구. **아동학회지**, 22(2), 255-275.
- 이은혜, 박주희, 최혜영 (2005). 영유아용(만 2세~5세) 발달검사 개발을 위한 예비연구. **아동학회지**, 26(6), 73-88.
- 이정화 (2006). 걸음마기 아동의 발달선별척도의 개발에 관한 연구. **한국심리학회지:발달**, 19(1), 105-125.
- 이준석 (2005a). 발달선별검사 개발을 위한 연구. **정서·행동장애연구**, 21(3), 25-43.
- 이준석 (2005b). 발달선별검사의 운동영역 타당화 연구. **특수교육저널:이론과 실천**, 6(3), 93-111.
- 이준석, 유재연, 신현기, 전병운, 고등영 (2005). 적응행동검사의 실제적 기술 영역 타당화 연구. 한국 특수교육학회 2005년 추계학술대회 발표논문집, 95-108.
- 이준석, 조광순 (2004). 국내 영·유아 발달 선별 및 진단 검사도구에 대한 고찰: 표준화중인 검사도구들을 중심으로. **특수교육저널:이론과 실천**, 5(1), 1-26.
- 전병운, 조광순, 이기현, 이은상, 임재택 (2004). **K-DIAL3 지침서**. 서울: 도서출판 특수교육.
- 전병운, 조광순, 이준석, 강미라 (2005). 발달장애 아동을 위한 진단·평가 도구의 개발 및 웹기반 시스템 구축. **특수교육학연구**, 40(1), 113-151.
- 정문자, 이은혜, 박경자 (1993). 베일리 유아발달 척도의 한국 표준화를 위한 예비연구. **아동학회지**, 14(1), 5-21.
- 제경숙 (1984). 베일리의 유아발달검사의 척도화를 위한 연구. **아동학회지**, 5, 53-64.
- 조복희, 박혜원 (2004). 한국 Bayley 영유아발달검사(K-BSID-II) 표준화연구(1): 지역, 성별 및 모의 교육수준에 따른 수행분석. **한국심리학회지:발달**, 17(1), 191-206.
- 허계형, Squires, J., Yovanoff, P., 이준석 (2006). 한국 영유아를 대상으로 한 발달 선별도구의 심리측정학적 연구. **정신지체연구**, 8(3), 261-278.
- 황혜익, 손원경 (2002). ESI-P의 타당화를 위한 기초연구. **열린유아교육연구**, 7(3), 291-306.
- 황혜익, 손원경 (2003). 유아용 발달선별검사도구 개발연구: ESI-R에 나타난 발달경향과 기준점 탐색을 중심으로. **열린유아교육연구**, 8(2), 241-259.
- Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice*, 18(1), 10-22.
- Nuttall, E. V., Romero, I., & Kalesnik, J. (1999). *Assessing and screening preschoolers: Psychological and educational dimensions*. Needham Heights, MA: Allyn & Bacon.

A Validation study of Cognitive Area of Developmental Screening Inventory

Joonsuk Yi

Division of Rehabilitation, Korea Nazarene University

<Abstract>

The purpose of this research was to conduct item validity analysis of a cognitive area of developmental screening inventory assessing developmental state of preschoolers. Explorations of item analysis (item difficulty and discrimination) and item internal consistency yielded evidence of content validity. The cognitive skills of the infants until 35 months increased gradually, and 54 month after it appeared with the fact that the motor ability maintains a stable level. Each item also reflected an age change appropriately. Against the result regarding gender difference was discussed.

Key Words : developmental test, cognitive development, test validity, item validity