

연구보고 2015-14

# 국제비교를 위한 유아발달평가 기초연구

문무경 박원순 임준범

**육아정책연구소**  
Korea Institute of Child Care and Education



## 머 리 말

아마도 유아발달 평가는 유아교육과 보육 분야에서 가장 연구가 미진한 주제 중의 하나일 것이다. 유아 발달의 개인차 및 조기 선별과 낙인 우려 등, 유아평가를 지양해 온 전통적 관점에도 불구하고, 최근 유아발달 평가를 측정, 진단하려는 관심이 급증하고 있다. 이는 취학 전 유아를 대상으로 하는 서비스의 보편화 달성과 더불어 서비스의 질을 입증하고 향상시키고자 유아 발달수준과 학습성과 측정, 평가의 필요성이 대두된 연유에서이다.

과거 일부 국가에서 유아의 학교준비도 측정을 위한 평가도구를 개발하기 시작하여 2000년대 중반부터 국가 간의 비교로 확산하게 되었다. 한편, 최근 2~3년부터 유네스코, OECD, IEA 등의 국제기구들은 유아의 균형적인 발달 모니터링, 유아교육과 보육 정책의 효과 검증, 거시적인 관점에서 유아교육과 보육 정책적 시사점 제공 등의 목적으로 유아 발달수준에 대한 국제비교의 가능성을 탐색하기 시작하였다. 그동안 주로 학습성취수준에 대한 국제비교연구인 PISA, TIMMS, PERLS를 수행해 온 것과는 달리, 유아를 대상으로 하는 국제비교는 평가도구의 타당성 이전에 평가 목적에 대한 국가 간의 논의와 합의가 필요하다.

이러한 배경에서 본 연구는 유아발달 국제비교의 기초를 마련하기 위하여 기존의 국내외 유아 발달과 학습 평가도구들을 검토하고 유아발달 평가 관련 국제비교의 주요 결과를 분석하고자 하였다. 또한 유아발달 평가에 대한 국내외 전문가들의 의견을 수렴하여 평가 목적에 따라 국제비교를 위하여 요구되는 유아발달 평가도구의 개발 방안을 제시하고 유아발달 국제비교 연구 수행시의 고려해야 할 사항들을 논의하고자 수행되었다.

본 연구의 수행을 위하여 협조하여 준 국내 연구협의진, 조사에 참여한 전문가 및 해외 전문가들에게 깊은 감사의 뜻을 전하며, 본 연구가 향후 유아발달 국제비교를 위한 평가도구 개발을 위한 후속 연구에 유용한 기초 자료로 활용되기를 기대한다.

2015년 11월  
육아정책연구소  
소장 **우 남 희**



# 차 례

요약 .....	1
I. 서론 .....	11
1. 연구의 필요성 및 목적 .....	11
2. 연구내용 .....	12
3. 연구방법 .....	13
II. 이론적 배경 .....	16
1. 국내 국제비교 연구 동향 .....	16
2. 유아발달 국제비교 연구결과 활용 목적 .....	20
3. 국가수준의 유아발달 평가도구 검토 .....	22
III. 유아발달 국제비교 평가도구 분석 .....	45
1. EDI(Early Development Instrument) .....	45
2. iPIPS(International Performance Indicators for Primary School) .....	51
3. EAP-ECDS(East Asia Pacific-Early Childhood Development Scales) .....	56
4. MELQO(Measuring Early Learning Quality and Outcomes) .....	66
IV. 유아발달평가 국제비교연구 결과 분석 .....	74
1. 특별관심 집단 .....	74
2. 유아발달 차이의 요인 .....	80
3. 발달영역별, 성별, 지역별 수준 차이 .....	85
V. 유아발달 국제비교에 대한 전문가 의견 심층분석 .....	89
1. 조사 개요 .....	89
2. 1차 조사 결과 분석 .....	91
3. 2차 조사 결과 분석 .....	98
VI. 정책적 제언 .....	105
1. 유아발달 평가 목적에 따른 국제비교 방안 .....	105
2. 유아발달 국제비교 연구 수행의 고려사항 .....	112

참고문헌 .....	117
Abstract .....	125
부록 .....	127
부록 1. 전문가 의견조사지(1차) .....	129
부록 2. 전문가 의견조사지(2차) .....	135

## 표 차례

〈표 I-3-1〉 국내 전문가 협의진 .....	14
〈표 I-3-2〉 국외 유아발달평가 전문가 협의진 .....	14
〈표 II-1-1〉 IEA와 OECD 주관의 국제평가 목록 .....	16
〈표 II-1-2〉 PIRLS와 prePIRLS의 평가틀 .....	17
〈표 II-1-3〉 TIMSS 2011 개요 .....	18
〈표 II-1-4〉 PIAAC의 평가틀 .....	20
〈표 II-3-1〉 영유아 기본학습능력발달 평가도구의 발달영역별 지표 .....	23
〈표 II-3-2〉 한국어동태널 연도별 조사 실시 계획 .....	25
〈표 II-3-3〉 한국어동태널 차수별 발달영역 조사표 .....	26
〈표 II-3-4〉 국내 학교준비도 검사도구 목록 .....	30
〈표 II-3-5〉 인지영역 검사도구 설명 .....	35
〈표 II-3-6〉 1차년도 검사 일정 .....	40
〈표 II-3-7〉 2차년도 검사 일정 .....	40
〈표 III-1-1〉 EDI 발달영역 및 세부영역과 예시문항 .....	46
〈표 III-2-1〉 PIPS 검사 내용과 예문 .....	52
〈표 III-2-2〉 iPIPS 4대 영역 및 세부항목 .....	54
〈표 III-3-1〉 EAP-ECDS의 참여국가별 발달영역 .....	58
〈표 III-3-2〉 EAP-CDCS 항목 예시 .....	58
〈표 III-3-3〉 EAP-CDCS 영역별 최종 문항 수 .....	61
〈표 III-3-4〉 EAP-CDCS 부모설문지 영역별 구성 .....	62
〈표 III-4-1〉 단계, 학교 교육 단계, 대략적인 연령 범주 .....	68
〈표 III-4-2〉 학습 영역의 범세계적인 구성틀 (Global Framework of Learning Domains) .....	69
〈표 III-4-3〉 범세계적 추적을 위한 학업 지표 .....	71
〈표 III-4-4〉 국가지원 사항 .....	72
〈표 III-4-5〉 국외평가도구 종합 분류표 .....	73
〈표 V-1-1〉 국내·외 전문가 조사 문항 목록(1차) .....	90
〈표 V-1-2〉 국내·외 전문가 조사 문항 목록(2차) .....	91
〈표 V-2-1〉 유아발달 평가영역 및 세부영역 .....	94

〈표 V-3-1〉 평가 영역 내 하위 영역별 중요도 평균 .....	101
〈표 V-3-2〉 유아발달 평가도구 개발 시 고려사항 중요도 .....	102
〈표 VI-2-1〉 국제비교 연구의 타당성 유형 .....	114

## 그림 차례

[그림 II-1-1] PISA 2012 평가틀 .....	19
[그림 II-3-1] EPPE표본 추출 설계도 .....	34
[그림 II-3-2] NEPS의 이론적 틀 .....	39
[그림 II-3-3] NEPS의 수학적 역량의 틀 .....	41
[그림 III-3-1] 문항특성곡선(Item Characteristic Curve) .....	60
[그림 III-3-2] 동일 문항에 대한 문화적 차이 .....	63
[그림 III-3-3] 연령에 따른 국가별 EAP-ECDS 점수 차 .....	64
[그림 III-3-4] 성별에 따른 국가별 EAP-ECDS 점수 차 .....	64
[그림 III-3-5] 도시·농촌에 따른 국가별 EAP-ECDS 점수 차 .....	65
[그림 III-4-1] 학습 영역의 범세계적인 구성틀 .....	69



# 요약

## 1. 서론

### 가. 연구의 필요성 및 목적

- 유아발달에 대한 국제비교의 목적은 국가간 경쟁을 유도하는 것이 아니라 정책입안자가 포괄적, 비판적 관점에서 유아교육과 보육 정책을 조망하고 수립할 수 있도록 지원하는데 있음(OECD, 2015).
- 본 연구에서는 평가 목적에 따른 유아발달에 대한 국제비교 및 평가도구의 개발 방안, 유아발달 국제비교 연구 수행시의 고려해야 할 사항들을 제시함으로써 유아발달 국제비교의 가능성을 탐색하고자 함.
  - 이를 위하여 기존의 관련 유아 발달과 학습 평가도구들을 검토하고 유아 발달 평가 관련 국제비교의 주요 결과를 분석하고자 함.

### 나. 연구내용

- 국제비교를 위한 국외 유아발달 관련 평가도구 검토함.
  - EDI, iPIPS, ARNEC의 ECDS, 유네스코의 Learning metrics의 평가 목적, 대상연령, 평가 영역 및 세부 문항, 검사 방법과 절차, 신뢰도, 타당도, 소요시간, 비용, 장단점 등을 분석함.
- 유아발달 평가도구를 사용한 국제비교 연구의 주요결과 분석함.
- 유아발달 국제비교의 필요성과 목적, 구체적인 평가영역 및 도구 타당화 등에 대한 국내 전문가 의견 파악함.
- 평가 목적에 따른 유아발달 국제비교 및 이를 위한 평가도구 개발과 활용방안을 제시함. 또한 향후 국제비교 연구 수행과 연구결과의 정책적 활용시 고려해야할 사항을 제시함.

## 다. 연구방법

- 본 연구는 문헌연구를 중심으로 수행됨.
  - 유아발달 평가의 국내·외 현황 및 기존 국제비교 연구를 검토함
  - 영국의 iPIPS(International Performance Indicators of Primary School), 캐나다의 EDI(Early Development Instrument), ARNEC의 EAP-ECDS (East Asia-Pacific Early Childhood Development Scale), 유네스코의 Learning Metrics의 개괄적 특징 및 관련 국제비교 연구의 주요결과 분석함.
- 국내외 발달평가 전문가 협의진 구성 및 운영함.
  - 총 6회의 전문가 협의진과의 협의회 및 2회의 집중워크숍을 개최함.

## 2. 이론적 배경

### 가. 국내 국제비교 연구동향

- 현재 국내에서 이루어지고 있는 국제비교 연구들은 IEA와 OECD가 주관하는 총 6종의 연구가 있으며, 연구 대상은 초등학교 4학년부터 성인까지로 다양하나 유아를 대상으로 한 연구는 전무함.
  - 6종의 연구 내용은 읽기, 수, 과학, 문제해결력 및 창의성 등을 주된 영역으로 포함하고 있음.

### 나. 유아발달 국제비교 연구결과 활용 목적

- 유아발달 국제비교는 개별 유아가 발달의 제반 영역에서 균형 있게 발달하는가를 모니터링할 수 있음
- 유아발달 국제비교 연구결과는 유아의 교육과 보육 정책의 효과 검증에 활용됨.
- 유아발달 국제비교 연구결과를 통하여 서로 다른 교육 및 보육 제도와 정책의 차이의 비교가 가능하므로 거시적 시각에서 정책적 시사점 제공.

## 다. 국가수준의 유아발달 평가도구 검토

- 국내의 영유아 기본학습능력발달 평가도구는 언어, 인지, 사회·정서 발달 부문에서 영유아의 전인적 발달을 평가하고 발달지연을 조기에 발견하여 적절한 교육지원을 체계화하는 것을 목적으로 함.
- 한국아동패널 검사도구는 아동이 출생 이후 어떠한 변화과정을 거쳐 가면서 어떠한 환경에 적응하고 성장하는지 파악하기 위한 종단자료를 생성하는 연구로 신체, 인지, 언어, 사회 및 정서 발달을 영역으로 함.
- 국내의 학교준비도 및 학습준비도에 대한 다수의 검사가 유아의 초등학교 및 초등교육과정에 대한 준비도를 측정함.
- 3세에서 7세의 유아를 대상으로 학령 전 아동들의 교육 효과를 조사한 영국의 EPPE(the Effective Provision of Pre-School Education) 프로젝트는 취학 전(pre-school) 교육이 발달에 긍정적인 효과가 있으며, 그 효과는 초등학교에서도 지속되나 효과는 점차 약해지며, 조기 ECEC 기관 이용아 및 취약계층 유아에게는 다양한 배경의 또래가 긍정적인 효과가 있음을 보고하였음.
- 유아기부터 늦은 성인기까지의 전 교육과정을 분석한 독일의 국가교육패널 연구 (NEPS: National Educational Panel Study)는 교육환경, 사회 불평등과 교육의 의사결정, 이민자들의 교육 적응도, 전 생애교육으로의 회귀를 주제로 함.
- 펜실베이니아주에서 실시하는 Head Start REDI (REsearch-based, Developmentally Informed) 프로그램은 학교준비도를 향상시키기 위해 유치원 교육과정을 수정하여 그 효과를 측정한 결과, 일부 긍정적인 효과가 있음을 밝힘.

## 3. 유아발달 국제비교 평가도구 분석

### 가. EDI(Early Development Instrument)

- EDI는 교사가 검사를 수행하는 취학 전 아동의 건강과 발달을 측정하는 학교준비도 검사임.

- 신체적 건강, 사회적 역량, 정서적 성숙도, 언어와 인지발달, 의사소통 능력, 일반적 지식의 5개 영역 총 104개 문항으로 구성되었음.
- 캐나다에서 개발된 후 미국, 호주, 영국, 뉴질랜드 등 영어권 국가뿐 아니라 칠레, 네덜란드, 이집트, 자메이카, 코소보, 몰도바, 멕시코 등 세계 12개국에서 참여하였음.

#### **나. iPIPS(International Performance Indicators of Primary Schools)**

- 1994년 영국에서 시작된 Performance Indicators in Primary Schools On-entry Baseline assessment (PIPS)를 기반으로 개발된 교사의 도움으로 컴퓨터를 사용해 3-5세 유아에게 사용하는 학습준비도 검사 도구임.
- 인지 발달, 인성 사회 정서 발달, 신체 발달의 3 가지 발달영역과 함께 배경 정보까지 총 4개 영역으로 구성되어 있음.
- 종단 추적연구를 통하여 이 검사의 결과는 취약한 유아에게 지속적인 중재가 이루어질 경우 효과가 크지만, 연속된 부정적 학교경험은 이후 회복하기 어려운 학업의 퇴보를 보이는 것을 밝힘.
- 학생의 학업 수행에 학교보다는 교사의 변인이 더 큰 영향을 끼치는 것으로 밝혀짐.

#### **다. EAP-ECDS (East Asia Pacific-Early Childhood Development Scales)**

- UNICEF와 Open Society Foundation의 후원으로 ARNEC(Asia-Pacific Regional Network for Early Childhood)이 주관하여 개발한 3-5세 유아를 위한 발달측정도구임.
- 캄보디아, 라오스, 몽골, 필리핀, 태국, 바누아투, 베트남 등 7개국을 대상으로 개발됨.
- 학습에 대한 접근방법, 인지 발달, 문화적 지식 및 참여, 언어 및 발현적 문해, 운동 발달, 건강, 위생 및 안전, 사회정서발달 등 7개 범주로 구성되어 있으며 인구사회학적 요인에 대한 정보는 부모설문을 통해 수집함
- 지역별, 성별, 연령별 발달 결과가 문화 간의 차이를 보임.

## 라. MELQO (Measuring Early Learning Quality and Outcomes)

- UNESCO Institute for Statistics(UIS)와 Center for Universal Education (CUE)은 2012년 7월에 학습성과 평가지표 전문팀(Learning Metrics Task Force: LMTF)을 구성하여 아동 및 청소년에게 보다 나은 교육 경험을 제공하기 위한 국제자문위원단의 7개 항목의 권고사항을 개발하였으며, 이중 학습역량, 범세계적 학습성과 추적을 위한 평가지표, 국가적 지원이 포함됨.
- MELQO는 신체적 안녕, 사회 및 정서, 문화 및 예술, 문해 및 의사소통, 학업 방식 및 인지, 산술 및 수학, 과학 및 기술의 7개 학습영역에서의 학습역량의 강화를 강조함.
- 범세계적 학습성과 추적을 위한 평가지표는 모두를 위한 학습, 학업을 위한 연령 및 교육 사항, 읽기, 산술, 학교준비도, 세계인으로서의 시민성, 학업기회의 폭과 같은 7개의 평가영역을 제시함.
- MELQO는 기술적, 제도적, 정치적 지원 등 3가지 국가지원이 필요함을 권고하고 있음.

## 4. 유아발달평가 국제비교 연구 결과 분석

### 가. 특별관심 집단

- 유아발달의 국제비교 연구들은 특수아, 장애아, 학습문제를 가진 아동, 원주민 아동 등을 그 외의 집단과 구분하여 집단의 차이를 규명함.
- 특수 및 장애아는 대부분의 발달영역에서 통제집단에 비하여 점수가 낮음.
- 청각장애아의 경우, 일부 수 영역에서 통제집단보다 높은 점수를 받은 경우가 있는데, 이는 서양 언어에 비해 수체계가 더 쉬운 수화의 영향임.
- 이중언어 아동의 경우, 전반적으로 통제집단보다 높은 언어 점수를 보였으나, 제2언어 학습자의 경우, 통제집단보다 언어 점수가 낮음.
- 아동의 언어발달은 이후 학교교육과정에 따라 제1언어의 종류와 큰 상관이 없음.

## 나. 유아발달 차이 요인

- 아동의 건강은 학교준비도와 정적인 상관관계를 보임.
- 아동의 성별이 학교준비도에 영향을 미치나, 이는 생물학적인 특성에 기인한 것이기도 하지만 아동에 대한 부모나 사회의 태도의 차이의 결과일 수 있음.
- 아동의 나이(월령)에 따른 차이, 또한 입학연령에 따른 차이가 있으나 그 효과의 크기나 지속되는 기간에 대해서는 이견이 있음.
- 부모의 학력은 아동의 발달에 별 영향이 없으나 가계소득은 긍정적인 영향이 높음.
- 안정적인 가정의 형태가 전반적인 발달의 점수와 정적인 상관관계에 있음.
- 부모가 아동과 함께 책을 읽는 것이 아동의 발달에 미치는 영향은 아동 스스로의 책에 대한 선호에 비해 영향력이 낮았음.
- 부모의 흡연 및 우울은 아동의 발달에 부정적인 영향이 큼.

## 다. 발달영역별, 성별, 지역별 수준 차이

- 일부 연구에서 인지발달에 있어서 도시의 아동이 촌지역보다 점수가 높았으나 성차는 크게 유의하지 않음.
- 사회·정서발달의 경우, 지역별 성차가 다른 경향성을 보일 수 있음.
- 대부분의 지역에서 농촌 아동이 도시 아동에 비하여 높은 신체발달을 보임.
- 언어발달 및 건강과 위생 및 안전의 차이, 문화지식과 참여 정도의 차이, 그리고 학습방식의 차이에 있어서 전반적으로 도시의 아동이, 또한 여아가 높은 발달 점수를 보임.

## 5. 국제비교를 위한 유아발달평가에 대한 전문가 의견 심층분석

### 가. 조사 개요

- 조사목적은 유아발달 국제비교 목적 및 향후 연구수행을 위한 전문가 견해와 수요를 파악하고자 함.

□ 구체적인 조사 내용은 다음과 같음

- 다른 나라와의 유아발달을 비교하는 목적, 평가에 적합한 연령대상 선정, 평가영역 및 필수 하위문항, 특히, 기존의 국제비교연구에서 우리나라 유아에 게 실시하기 부적합한 문항 및 우리나라 유아발달평가의 고유한 문항에 대한 규명, 평가방법과 절차, 평가시기, 평가결과 활용방안 등

□ 조사 참여자의 특성

- 국내 외 유아발달 평가 전문가 45명
  - 참여한 학계 전문가의 전공은 교육학과 및 유아교육학과 및 국가연구기관에서 학생평가 관련 국제비교연구 경험이 있는 연구책임자에 국한함.

## 나. 1차 조사 결과

□ 우선 유아발달 국제비교의 필요성 및 목적에 대한 인식은 해당 전공영역에 따라서 찬성과 반대의 입장으로 구분됨.

- 주로 찬성의 입장에서는 교육과정과 성과를 비교분석하여 자국 교육의 현 상황과 교육개선을 위한 유용한 자료로 활용될 수 있다는 점을 주장한 반면, 반대의 입장에서는 유아기의 발달은 개인차가 크고 발달 패턴 또한 상이하기 때문에 유아의 발달정도가 과연 국제비교를 할 만한 준거가 되는가에 대한 의문을 제기함.

□ 유아발달 수준에 대한 국제비교 연구 대상은 대상집단과 국가로 구분됨.

- 국제비교에 적합한 대상집단 연령에 대하여 만 5세를 선택한 이유는 대부분이 유아교육의 마지막 단계이면서, 초등학교 입학직전이므로 유아교육의 총체적인 측면을 고려할 수 있다는 점임.
- 또한 저소득층, 농어촌, 대도시 저소득층의 유아들을 관심 범위에 두어야 하며, 보육기관 미이용 유아 또한 포함되어야 함을 응답함. 이외에 성별차이는 국제비교연구의 주요 관심사이므로 포함되어야 함을 지적함.
- 대상 국가로는 발달에 대한 많은 연구와 정책이 이루어지고 선진적인 유아관련정책을 추진 중인 국가로 벤치마킹 할 수 있는 국가를 선호함.

□ 유아발달 평가영역과 세부영역에 대하여 참여국가 간 공통적으로 측정 및

보고가 가능한 세부영역을 선정하는 작업이 우선적으로 필요하며, 이러한 단계는 평가도구를 만들기 이전에 참여하는 국가들의 합의로 우선되어야 한다는 점을 조사대상자 대부분이 제시하고 있었음.

- 국제비교를 위한 유아발달 평가도구 개발 관련하여 정책적 활용도를 최우선 순위로 선택함. 그 다음 순으로 평가도구의 질(타당도와 신뢰도 등)이 중요한 것으로 나타남.

#### 다. 2차 조사 결과

- 유아발달 국제비교 목적에 있어서 1순위는 국제적 수준에서 우리나라 유아 발달수준의 현 위치 점검, 2순위는 국제적 수준에서 유아의 발달과 학습 성과 비교, 3순위는 적극적인 조기중재 방안 마련, 4순위는 개발도상국과의 교류협력을 위한 정보제공 및 지원으로 조사됨.
- 국제비교를 위한 평가도구 개발과 관련하여 적합한 연령대는 대부분의 전문가들은 만 5세 이상을 선택함. 비교 대상 국가는 주로 OECD 중심의 선진국과 사회문화가 비슷한 동북아시아권역 국가로 한정지어졌으며, 특정 집단 설정에 대해서는 유아교육 및 보육기관 미이용 집단과 지역적 분류 집단이 가장 많은 응답으로 확인되었음.
- 국제수준의 타당화 과정에서의 고려 사항에 대한 의견은 다음과 같음
  - 표본의 크기: 대체적으로 모집단의 대표성과 집단 간 비교 시 효과크기를 고려해 정해지는 것이 바람직함.
  - 국가별 문항 수정 허용수준: 전체 문항 중 20~30% 정도는 개별국가에서 재량으로 함.
  - 적정 문항 수: 평가영역과 방법에 따라서 다르며, 검사 대상이 유아이기 때문에 20분 이내에 수행 될 수 있는 길이의 검사로, 100문항 내외가 적정함.
  - 총체적 발달평가 안에 특정영역 평가도구를 포함하여 진행함. 집중, 추적해야 할 영역들이 있는 경우는 전체 영역 중 3개 이하로 제한.
  - 사회정서영역 평가의 타당도 확보 또는 향상 방안: 여러 차례 관찰 평가를 실시해야만 타당도 확보 및 향상이 이루어짐. 행동지표(예. 일주일의 평균

문제행동 횟수)를 활용을 권장함.

- 평가자: 기관의 교사, 부모, 전문 평가자로 구분되나, 가능한 부모를 주평가자로 참여시키고 교사나 전문 평가자는 부평가자로 평가체제를 갖추는 것을 선호함.
- 평가 시기 및 시행 횟수: 연초에 1회로 한정하는 것이 현실적이나, 평가영역에 따라 시기 및 횟수를 융통성 있게 조정할 필요가 있음.
- 평가(검사)환경의 표준화 확보: 행동수행을 평가하는 경우를 대단위 검사에 포함시키기는 어렵기 때문에 검사를 사용하며, 검사의 동등성 확보가 요구됨.

## 6. 정책적 제언

### □ 조기 개입을 위한 국제수준의 유아발달 평가

- 각 국가별로 유아 단계에 나타나는 계층, 성별, 인종, 지역 등 사회경제적 배경에 따른 학습 출발점의 차이 파악 및 유아교육 단계 교육불평등의 영향을 미치는 관련 요인 비교 분석
- 이후 학교 학습의 차이와 관련성이 높은 영역을 좀 더 집중적으로 평가할 필요가 있음. 일반적으로 학교 학습 또는 성취도는 아동의 지적능력(기억력, 인지능력, 수와 셈하기 등), 언어발달 수준 등과 밀접하게 관련이 있음.
- 다양한 집단 간 차이를 잘 드러내지 못한다면 조기 개입을 위한 평가도구로서 적절하지 않을 수 있음.

### □ 학교준비도 국제비교 가능성 및 유의점

- 초등교육 입문 단계에서 일반적인 유아들의 학교준비도를 국제적으로 비교하고 학교준비도에 영향을 미치는 다양한 가정, 교육, 제도적 요인 파악해야 함.
- 학령기 아동에 비해 유아를 대상으로 한 평가는 평가 시간, 평가 방법 등에서 많은 제약을 받으므로, 평가의 타당도, 신뢰도를 확보할 수 있는 연구 설계 및 평가 도구 선정, 평가 시행이 필요함. 이에 따른 추가적인 비용과 노력을 감수해야 함.

□ 유아발달 국제비교연구 수행 시 고려사항은 아래와 같음.

- 연령의 통제
- 언어와 문화 요인
- 평가도구의 국제수준 타당화 과정

# I. 서론

## 1. 연구의 필요성 및 목적

유아발달 평가는 전통적으로 유아교육과 보육 분야에 있어서 소극적인 논의와 관심의 대상이었고 상대적으로 발전이 더딘 주제이다. 1920년대 진보주의교육의 대두와 더불어 아동에 대한 연구가 활발히 이루어지면서 아동 대상의 표준화검사들이 개발되었고, 이는 정책입안자들에 의해 정책의 효과를 확인하고자 활용되기 시작하였다. 그러나 검사 도구를 사용한 유아발달 평가는 조기 선별 및 낙인에 대한 우려로 그 발전이 제한적이었다.

그럼에도 불구하고 최근 국제적으로 유아발달 평가에 대한 관심이 높아져 가고 있다. 보편적 유아교육과 보육 서비스 기회 제공의 목표를 일정 수준 이상으로 달성한 OECD 회원국들을 중심으로 한 대부분의 선진국들이 이제는 서비스의 질 향상, 양질의 학습경험 제공에 주목하게 되었다. 유아기 서비스의 질을 담보하기 위하여 상당수의 국가들(예: 캐나다, 영국, 호주, 독일 등)은 자국에서는 물론 국제적 수준에서 지속적으로 확인 가능한 유아 발달과 학습 지표를 규명하고 이를 활용하여 유아를 평가하는데 노력을 기울이기 시작하였는데 이는 유아교육과 보육 정책 및 제도 실행의 효과를 확인하고 생애초기부터 유아의 필요에 부합한 보다 구체적인 서비스를 제공하여 평등한 기회를 보장하기 위함이다.

유아발달 평가의 목적은 개별유아간의 비교보다는 유아간의 평등을 보장하고, 유아의 현재 경험이나 교육환경을 검토하여 궁극적으로 유아교육과 보육 서비스의 질을 향상하는데 있다(Janus & Offord, 2007; Tymms, et al., 2009) 아울러, 유아발달에 대한 국제비교 역시, 국가 간 경쟁을 유도하는 것이 아니라 정책입안자가 포괄적, 비판적 관점에서 유아교육과 보육 정책을 조망하고 수립할 수 있도록 지원하기 위한 것이다(OECD, 2015). 즉, 국가 내의 사회적, 문화적 특수성으로 인해 국내 비교만으로는 프로그램의 효과가 발달에 미친 직접적인 영향을 밝히기 어려운 점을 국제적인 비교를 통하여 취약계층 영유아의 발달에 대한 프로그램의 효과를 명확히 검증할 수 있다(Tymms, et al., 2009). 이러한 맥

락에서 UNESCO의 Learning Metrics Task Force(2013)는 국제비교를 위한 유아 발달평가는 ‘공익’(public good)으로 간주되어야 하며, 평가도구 및 수집된 자료는 무상으로 활용 가능하여야 하며, 비용문제로 인하여 국제비교에서 특정 국가를 배제하지 않아야 함을 강조한다.

요컨대, 유아교육과 보육 서비스 기회의 ‘보편적 제공(universal access)’을 넘어서 모든 유아가 적절한 교육적 지원을 받으면 일정수준의 ‘보편적 성취(universal achievement)’를 할 수 있다는 방향 전환이 필요하며, 이를 위하여 유아발달 국제비교는 하나의 강력한 수단이 될 수 있다.

이러한 취지에서 본 연구에서는 유아발달 국제비교의 기초를 마련하기 위하여 기존의 국내외 유아 발달과 학습 평가도구들을 검토하고 유아발달 평가 관련 국제비교의 주요 결과를 분석하고자 한다. 또한 취약소외계층 유아의 발달과 학습 보장을 비롯하여 유아발달 평가에 대한 국내 전문가들의 의견을 수렴하여 평가 목적에 따라 국제비교를 위하여 요구되는 유아발달 평가도구의 개발 방안을 제시하고자 한다. 아울러, 유아발달 국제비교 연구 수행시의 고려해야 할 사항들을 제시하고자 한다.

## 2. 연구내용

본 연구의 목적을 달성하기 위해 설정한 연구내용은 다음과 같다.

첫째, 유아발달 국제비교 관련 선행연구는 소수에 불과하므로 우리나라가 참여해 오고 있는 국제기구의 초중등학생 대상 학습성과 국제비교 연구의 동향과 시사점을 파악한다. 현재, 국제비교를 위하여 사용가능한 종합적인 유아발달 평가도구가 소수에 불과하므로 국가수준에서 활용된 대표적인 종단연구에서 사용된 평가도구와 방법, 주요 결과를 검토한다.

둘째, 국제비교를 위한 국외 유아발달 관련 평가도구를 검토한다. 우선적으로 국제적으로 사용되고 있는 종합적 유아발달 평가도구(여러 영역의 발달을 동시에 측정)에 국한하여, EDI, iPIPS, ARNEC의 ECDS, 유네스코의 Learning metrics 등을 검토한다. 평가 목적, 대상연령, 평가 영역 및 세부 문항, 검사 방법과 절차, 신뢰도, 타당도, 소요시간, 비용, 장단점 등을 분석한다. 국제비교를 위한 유아발달 평가영역 범위, 평가도구간의 우선순위(정책적 중요도, 평가도구

의 질[타당도, 신뢰도, 비교가능성], 소요비용 수준, 유아와 교사에 주는 부담 정도 등에 비추어), 취약소외계층 유아의 기본학습능력 관리를 위한 평가 등에 대하여 분석한다.

셋째, 유아발달 평가도구를 사용한 국제비교연구의 주요결과들을 분석한다. 특별관심 집단(특수아, 이중 언어 사용 유아 등), 총체적, 발달영역별, 연령별, 성별, 지역별에 따른 결과, 결과의 차이를 가져오는 요인 및 정책적 활용, 시사점을 검토한다.

넷째, 유아발달 국제비교의 필요성과 목적, 구체적인 평가영역 및 도구 타당화 등에 대한 국내 전문가 의견을 파악한다.

다섯째, 평가 목적에 따라 국제비교를 위한 유아발달 평가도구 개발 및 활용 방안을 제시하고 관련 국제비교 연구 수행 시 고려해야 할 사항들을 논의하고자 한다. 구체적으로 국제비교를 위한 유아발달 평가영역 및 하위 영역을 도출하고, 적합한 유아연령, 소요비용 수준 등 평가도구 타당화를 위한 실질적인 제반 사항 등을 논의, 제시한다. 이를 토대로 향후 유아발달 국제비교 연구 수행 및 연구결과의 정책적 활용 시 고려해야 할 사항들을 제시한다.

### 3. 연구방법

#### 가. 문헌연구

문헌연구를 통해 유아발달 평가의 국내·외 현황 및 기존 국제비교 연구를 검토한다. 이를 위하여 총체적인 발달영역에 걸쳐 국제비교에 활용되는 영국의 iPIPS, 캐나다의 EDI(Early Development Instrument), 유네스코의 Learning Metrics, ARNEC의 ECDS(Early Childhood Development Scale) 등 평가도구의 개괄적 특징, 대상연령, 평가방법, 타당도와 신뢰도, 평가소요시간, 비용, 평가도구의 장단점 등에 대하여 종합적으로 검토하였다. 또한 관련 국제비교 연구의 주요 결과를 분석하였다.

#### 나. 국내외 전문가 협의진 구성 및 운영

국내 유아발달평가 전문가 6인으로 협의진을 구성하여 정책 수립에 기여하고

자 국제적 수준에서 유아발달 평가도구 개발을 위한 방안을 모색하였다. 유아발달 평가도구 개발 및 국제비교 연구 수행 방안 등을 논의하기 위하여 심리학, 아동학, 교육심리, 유아교육 등 유관분야 전문가를 섭외하여 6차례의 협의회 및 워크숍을 개최하여 의견을 수렴하였다.

<표 1-3-1> 국내 전문가 협의회

(가나다 순)

성명	소속	전문 영역
김성식	서울교대 교수	교육평가
김창환	한국교육개발원 선임연구위원	국제비교
신종호	서울대 교수	학습심리
양명희	경희대 교수	교육평가
옥현진	이화여대 교수	학업성취 국제비교
조지민	한국교육과정평가원 선임연구위원	교육평가

국외 국제비교 유아발달평가 전문가 7인은 주로 기존의 국제비교 유아발달 평가도구 개발자 및 국제기구 전문가를 포함하며, 특히 아시아 지역 전문가를 활용하여 아시아 국가들의 유아발달 평가에 대한 가치와 구체적인 평가영역과 지표 등에 대한 관점을 파악하였다. 국외 협의진에게는 해당전문가가 개발하였거나 활용한 평가도구 및 수행한 국제비교 연구를 통한 주요 결과 및 정책수립에의 기여 및 영향력 등에 대하여 정보 및 자문을 구하였다.

<표 1-3-2> 국외 유아발달평가 전문가 협의회

성명	소속	비고
Peter Tymms	영국 Durham대 교수	iPIPS 개발자
Nirmal Rao	홍콩대 교수	ARNEC ECDS 개발자
Maki Hayashikawa	유네스코 파리본부	Learning Metrics
Juliane Hencke	IEA	ECES Study 코디네이터
Kiyomi Akita	동경대 교수	OECD ECEC Network Vice-Chair
Minyi Li	북경사범대 교수	아동가족패널 연구
Rebecca Bull	싱가포르 NIE 교수	싱가포르 유치원교육 영향 평가(유아 평가)

## 다. 전문가 의견 조사 실시

국내외 유아발달 평가도구 검토 및 유아발달 국제비교연구의 주요결과 분석을 토대로 향후 국제비교를 위한 유아발달평가에 대한 국내 전문가의 의견을 파악하고자 조사를 실시하였다. 학계 전문가 총 45명을 대상으로 2차례의 의견 조사를 통하여 의견을 수렴하였다.

다른 국가와 유아발달을 비교하는 목적 및 목적에 따른 적합한 연령대상, 평가영역 및 필수 하위영역, 평가방법과 절차, 평가 시기, 평가도구 타당화, 평가 결과 활용방안 등에 대한 견해를 파악하였다.

## 라. 자료수집 및 결과 공유를 위한 국외출장 실시

본 연구의 일환으로 싱가포르 교육부 유아교육과를 방문하여 본 연구의 결과를 공유하고 최근 싱가포르 교육부에서 추진 중인 중인 국가수준의 유아발달 성과 평가 사례(Singapore Kindergarten Impact Project: SKIP)를 파악하였다. 또한 싱가포르에 소재하는 ARNEC (Asia Regional Network of Early Childhood) 사무국을 방문하여 홍콩대학팀에 위탁하여 개발한 Asia-Pacific Early Childhood Development Scale이 도구 타당화 과정에 참여한 개발도상국의 유아교육과 보육 정책에 미친 영향 및 평가도구 활용 사례, 향후 계획 등을 파악하였다.

## II. 이론적 배경

현재 유아발달 국제비교 연구는 매우 소수이며 국제기구에서도 최근 2~3년 전에서야 비로소 유아 대상 국제비교의 가능성을 탐색하기 시작하였다. 국내에서는 유아를 대상으로 발달 및 학습 수준에 대한 국제비교는 전무한 실정이므로, 본 장에서는 초중등학생을 대상으로 우리나라가 참여해 오고 있는 국제기구의 학습성취수준 국제비교연구인 PISA, TIMSS, PERLS 등의 평가도구 및 활용 방법을 살펴봄으로써 국제비교 연구의 동향 및 시사점을 파악하고자 한다. 또한, 국제비교 연구 내 검사도구들의 특성을 파악하여 국제비교 평가 시 중요한 사항들을 정리하고자 한다. 아울러, 유아를 대상으로 수행된 매우 제한적이므로 국가 수준에서 수행된 국내외의 유아발달 평가 및 종단연구를 평가도구와 방법을 중심으로 검토하고자 한다.

### 1. 국내 국제비교 연구 동향

현재 국내에서 이루어지고 있는 국제비교 연구들은 아래의 <표 II-1-1>과 같다. 총 6종의 국제비교 연구가 진행되고 있으며, 대상은 초등학교 4학년부터 성인까지 다양하게 구성되어 있다.

<표 II-1-1> IEA와 OECD 주관의 국제평가 목록

주관	평가	대상(학년)	영역	국내시행 담당	주기	첫시행
IEA	PIRLS	만 9세(초4)	읽기	-	5년	2001년
	prePIRLS	만 9세(초4)	읽기	-		2011년
	TIMSS	만 9세(초4)	수학, 과학	한국교육 과정평가원	4년	1995년
	ICILS	만 13세(중2)	수학, 과학			2013년
OECD	PISA	만 15세(고1)	읽기, 수학, 과학, 문제해결력	한국직업 능력개발원	3년	2000년
	PIAAC	만 16~65세	읽기, 수학, 문제해결력		-	2012년

자료: 육현진(2014). 국제 문식성 평가 분석을 통한 문식성 교육 시사점 탐색: PIRLS, PISA, PIAAC을 중심으로. 청람어문교육, 49, p. 40

PIRLS(Progress in International Reading Literacy Study)와 prePIRLS는 국제 읽기 발달연구로 국내에서 첫 시행이 된 것은 각각 2001년도와 2011년도에 시행되었다. 이 연구들은 만 9세를 대상으로 읽기영역에 대한 검사를 실시하며, 시행 주기는 5년이다. prePIRLS의 경우는 PIRLS보다 난이도가 낮은 수준의 검사라고 할 수 있다. 아래의 <표 II-1-2>는 PIRLS와 prePIRLS의 평가틀을 제시하였다.

<표 II-1-2> PIRLS와 prePIRLS의 평가틀

평가 도구	문항 유형	읽기 목적			인지 유형				
		문학적 경험	정보의 습득·활용	배점 비율	정보 확인	추론	해석·통합	점검·평가	배점 비율
PIRLS	선다형	40개 (40점)	34개 (34점)	43%	21개 (21점)	33개 (21점)	10개 (10점)	10개 (10점)	43%
	구성형	32개 (50점)	29개 (50점)	57%	12개 (17점)	13개 (16점)	28개 (55점)	8개 (12점)	57%
	배점 비율	52%	48%	-	22%	28%	37%	13%	-
pre PIRLS	선다형	31개 (31점)	26개 (26점)	43%	23개 (23점)	22개 (22점)	12개 (12점)	-	43%
	구성형	32개 (36점)	34개 (41점)	57%	34개 (37점)	13개 (14점)	19개 (26점)	-	57%
	배점 비율	50%	50%	-	45%	27%	28%	-	-

자료: 옥현진(2014). 국제 문식성 평가 분석을 통한 문식성 교육 시사점 탐색: PIRLS, PISA, PIACC을 중심으로. 청람어문교육, 49,

TIMSS(Trends in International Mathematics and Science Study)는 수학과 과학영역에 특화된 성취도 추이변화 국제연구로써, 국제수준에서 각 평가자들의 합의된 교육과정 및 성취기준을 토대로 한 종단연구이다(김수진, 2013, 재인용). 실제로 가장 최근에 결과 보고를 한 TIMSS 2011보고서에서 밝힌 기본적 도구 설명은 아래의 <표 II-1-3>과 같다.

&lt;표 II-1-3&gt; TIMSS 2011 개요

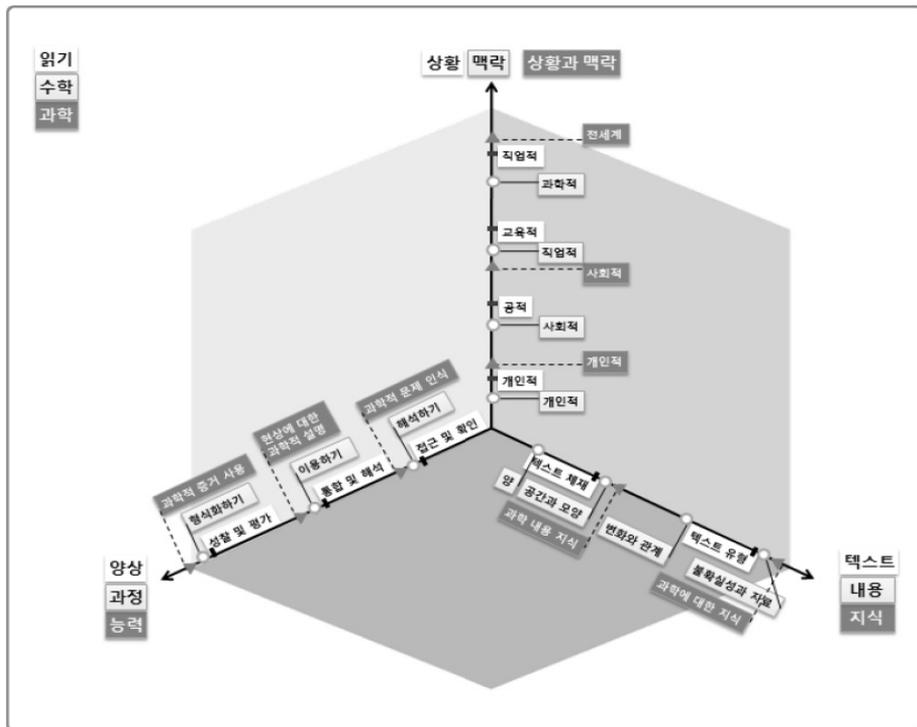
TIMSS 2011		the Trends in International Mathematics and Science Study 2011
연구특징	국제수준에서 합의된 교육과정을 토대로 한 종단적 연구	
연구목적	· 연구 참여국 학생들의 수학과 과학 성취도 수준 및 4년 주기의 변화추이 파악 · 성취도에 영향을 주는 교육 맥락적 정보들과 성취도와의 연계분석을 근거로 연구 참여국의 교육정책 수립의 기초 자료로 활용	
대상	초등학교 4학년, 중학교 2학년	
연구시기	2009년 ~2012년	
연구특징	국제수준에서 합의된 교육과정을 토대로 한 종단적 연구	
연구목적	· 연구 참여국 학생들의 수학과 과학 성취도 수준 및 4년 주기의 변화추이 파악 · 성취도에 영향을 주는 교육 맥락적 정보들과 성취도와의 연계분석을 근거로 연구 참여국의 교육정책 수립의 기초 자료로 활용	
대상	초등학교 4학년, 중학교 2학년	
연구시기	2009년 ~2012년	
평가 도구	성취도검사지	· 평가 문항을 총 14종의 검사지로 재구성함 · 검사지 1권당 2개의 수학 문항군과 2개의 과학 문항군이 포함되며, 이전 주기의 추이 문항이 포함됨
	설문지	초등학교 4학년(3종): 학교, 교사, 학생 중학교 2학년(4종): 학교, 교사(수학, 과학 교사 2종), 학생 · 시행시기: 2010년 12월 · 표집 및 참여인원
우리나라 시행	- 초등학교 4학년: 5329개 초등학교, 총 590547명의 대상 모집단 중 150개 학교, 4494명의 학생이 표집되어 실제로 4335명이 참여함 - 중학교 2학년: 2926개 중학교, 총 669919명의 대상 모집단 중 150개 학교, 5316명의 학생이 표집되어 실제로 5167명이 참여함	

자료: 김수진(2013). TIMSS 2011 우리나라 학생들의 수학·과학 성취 특성(PIM 2013-1). 한국교육과정평가원. p. 3

TIMSS는 4년 주기의 수학과 과학 성취도의 변화 추세를 파악하기 위한 종단 연구로써, 성취도에 영향을 주는 관련 정보들과 성취도와의 연계분석을 근거로 연구에 참여한 국가들의 정책 수립의 기초자료를 제공하는 목적을 가지고 있다(김수진, 2013, 재인용). 평가 대상은 초등학교 4학년과 중학교 2학년이며, 성취도 검사와 배경변인 설문으로 구성되어 있으며, 성취도 검사는 다시 수학과 과학

학영역으로 분류된다. 현재 TIMSS 2015가 계획 중이며, 한국교육과정평가원에서 주관하고 있다.

PISA(Programme for International Student Assessment)는 OECD에서 주관하는 학업성취도 국제비교 연구이다. 전 세계적으로 가장 많이 알려져 있는 국제 비교 연구 중 하나이며, 만 15세(국내 고등학교 1학년)를 대상으로 읽기, 수학, 과학, 문제해결력의 학업성취도를 평가하는 연구이다. 평가 주기는 3년으로, 다른 국제비교 연구들과는 달리 상대적으로 짧은 평가주기를 가지고 있다. PISA의 평가틀은 아래의 [그림 II-1-1]과 같다.



자료: 조지민 외(2012). OECD 국제 학업성취도 비교 연구: PISA 2012 본검사 시행보고서. 한국교육과정평가원 연구보고 RRE 2012-3-1. p. 20

[그림 II-1-1] PISA 2012 평가틀

PIAAC(Programme for the International Assessment of Adult Competencies)는 국제 성인역량 조사로 만 16세부터 65세까지 청소년기 이후 전 생애에 걸쳐

서 평가는 국제 비교 연구이다. 검사영역은 읽기, 수학, 문제해결력으로 구성되어 있으며, 실생활을 하면서 경험하게 되는 내용들이 문항 구성에 포함되어 있다. PIAAC의 평가들은 아래의 <표 II-1-4>와 같다.

<표 II-1-4> PIAAC의 평가들

	문식성	컴퓨터 기반 문제해결력
정의	<ul style="list-style-type: none"> <li>· 텍스트의 내용을 이해, 평가, 적용하며, 텍스트를 통해 사회에 참여하고 개인의 목적을 성취하며 지식과 잠재력을 발달시키는 데 필요한 능력</li> <li>· 단어와 문장을 이해하는 데에서 출발하여 복잡한 문자의 이해, 해석, 평가에 이르기까지의 다양한 기능을 포괄(단, 작문은 포함되지 않음)</li> </ul>	<ul style="list-style-type: none"> <li>· 정보 획득과 평가, 타인과의 의사소통, 그리고 실제적인 과업 수행을 위해 디지털 기술, 커뮤니케이션 도구와 네트워크를 활용하는 능력</li> <li>· 컴퓨터와 컴퓨터 네트워크를 통해 적절한 목표와 계획을 수립하고 정보에 접근 및 이를 활용함으로써 개인적인 업무 및 시민의 목적을 위해 문제를 해결하는 능력</li> </ul>
인지전략	<ul style="list-style-type: none"> <li>· 접근 및 확인</li> <li>· 통합 및 해석</li> <li>· 평가 및 반성</li> </ul>	<ul style="list-style-type: none"> <li>· 목표설정 및 진행과정 점검(모니터링)</li> <li>· 계획하기</li> <li>· 정보 수집 및 평가</li> <li>· 정보 활용</li> </ul>
맥락	<ul style="list-style-type: none"> <li>· 직장 관련</li> <li>· 개인</li> <li>· 사회 및 지역 사회</li> <li>· 교육과 연수</li> </ul>	<ul style="list-style-type: none"> <li>· 직장관련</li> <li>· 개인적</li> <li>· 사회와 지역</li> </ul>

자료: 옥현진(2014). 국제 문식성 평가 분석을 통한 문식성 교육 시사점 탐색: PIRLS, PISA, PIACC을 중심으로. 청람어문교육, 49, p.75

## 2. 유아발달 국제비교 연구결과 활용 목적

### 가. 유아의 균형적 발달 모니터링

유아의 발달에 대한 평가는 개별 유아가 발달의 제반 영역에서 고르게 발달하고 있는지 관찰할 수 있도록 한다. 발달 척도의 이와 같은 기능은 개인에게는 고른 발달을 추구할 수 있는 자료를 제공하게 한다. 특히 이 발달에 대한 평가가 국제적으로 비교될 수 있는 자료라고 한다면 한 국가의 맥락 속에서 유아들

간의 상대적인 발달의 정도에 대한 정보뿐만 아니라, 국가의 틀을 벗어나는 사회, 문화, 및 언어적 맥락을 모두 포함한 종합적인 발달 정보를 제공할 수 있다. 특히 이와 같이 유아의 발달에 대한 정확한 정보를 얻는 것은 개인적인 이점뿐만 아니라 정책 결정자들이 유아교육 및 보육의 목표를 설정하는 데 있어서 필수적인 정보이기도 하다.

## 나. 영유아 교육과 보육 정책의 효과 검증

호주의 태즈메니아주에서는 학령전기 (pre-school) 정책에 대한 효과성을 검증하기 위해 학교를 다니기 시작하며 의무적으로 iPIPS를 수행하도록 하고 있다 (Dixon, Humble, & Counihan, 2015). iPIPS의 결과는 Launching into Learning(LiL) 프로그램에 참여한 아동이 그렇지 않은 아동에 비해 10% 정도 높은 발달을 보였는데, 이것의 시사점은 영유아를 위한 교육 및 보육 정책을 시행할 때 발달에 대한 평가 척도를 사용하여 영유아를 위한 정책의 효과를 검증할 수 있었다는 점이다. 정책의 효과를 평가하기 위해서는 해당 정책에 따른 프로그램의 효과를 측정하는 고유한 척도를 개발하는 경우도 있다. 그러나 척도가 측정하는 바가 단지 해당 프로그램의 영향력의 유무를 검증하는 것을 넘어, 실제 아동의 일반적 발달이나 능력에 직접 영향을 미치는 지를 검증하는 것은 정책의 효과를 일반화하는 강력한 근거가 된다.

위의 호주의 사례에서는 iPIPS를 통한 유아의 발달에 대한 평가로 태즈메니아주의 LiL 프로그램의 효과를 검증할 수 있었다. LiL 프로그램에 참여한 유아가 그렇지 않은 유아에 비하여 더 높은 발달 지표 상의 결과를 보여 상대적으로 뒤쳐진 출발선상에서 학교생활을 시작하게 되는 취약계층의 영유아들에게 이 프로그램이 효과적일 수 있음을 밝힌 것이다. 특히 국내 비교만으로는 국가 내의 사회적, 문화적 특수성으로 인해 프로그램의 효과가 발달에 미친 직접적인 영향을 밝히기 어려운 점을 국제적인 비교를 통하여 취약계층의 영유아들의 발달에 대한 프로그램의 명확한 검증할 수 있었다.

한편 영국, 스코틀랜드, 뉴질랜드, 호주를 비교한 Merrell 과 Tymms (2007)의 연구는 학교 교육의 효과를 각 국가 내의 비교를 통하여 확인하였다. 비슷한 연령의 학교 교육을 경험한 아이들과 학교 교육을 경험하지 않은 아이들의 읽기 발달의 비교를 실시하였는데 그 발달상의 차이가 학교교육의 효과를 입증한 것이다. 이들 연구의 결과는 입학한 아동 중 가장 나이가 많은 아이들은 1년간 수

학한 아이들 중 가장 어린 아동에 비하여 나이 차이가 1개월에 불과하지만 이들의 읽기 수준의 차이는 약 2-3년의 차이를 보였다. 즉 초기 학교교육의 효과가 얼마나 강력한 지를 입증한 것으로 볼 수 있다. 특히, 이는 4개국에서 공통적으로 발견된 현상으로 특정한 정책이나 일국의 사회, 문화적 맥락과 관계없는 일반적인 결과임을 시사하는 것으로 볼 수 있다.

#### 다. 거시적 관점에서 유아교육과 보육에 대한 정책적 시사

Merrell 과 Tymms (2007)의 영국, 스코틀랜드, 호주, 뉴질랜드의 학령초기 아동에 대한 비교연구는 초등학교를 시작하는 연령이 서로 다른 국가들 간의 비교를 실시하였는데 연령과 학교교육의 경험의 효과를 국제간의 비교를 통해 보정함으로써 보다 정확하게 연령과 학교 교육의 효과를 통제할 수 있었다.

이중에서도 특히 눈여겨볼 수 있는 연구결과는 4개국의 비교에 있어서 호주 아동의 발달의 진보가 다른 국가의 아동에 비하여 낮은 점이다. 이 연구에서 그 차이에 대한 구체적인 원인을 규명하지는 못했기에 절대적인 기준이 되기엔 미흡하나 이를 통해 적어도 이러한 아동발달의 국제비교는 각 국가에서 학교교육의 유효성을 비교해 단점, 혹은 문제점을 인식하고 이에 대한 정책적 대안을 수립하는 계기를 제공하기에는 충분한 의미를 지니고 있다. 즉 전체적 발달의 비교를 통하여 한 국가의 교육 및 보육 정책이 더욱 관심을 기울여야 할 부분이 무엇인지를 알게 하고 이를 토대로 거시적인 관점에서 유아교육과 보육 정책을 수립하는 데 도움을 줄 수 있다.

### 3. 국가수준의 유아발달 평가도구 검토

유아발달을 평가하기 위한 국내의 도구 개발 및 표준화 연구는 다수 진행되어 왔다. 그러나 대부분의 도구들은 유아의 특정한 영역의 발달을 측정하는 것을 목적으로 한 것이며 유아의 총체적인 발달을 측정한 도구는 그리 많지 않다. 특히 문해능력이 제한적인 유아발달의 특성상 지필식 검사의 사용이 제한적이므로 도구의 타당도 및 신뢰도 확보가 용이하지 않다. 이에 본 연구에서는 우리나라의 대표적인 총체적 유아발달 평가도구와 학습준비도 검사도구를 살펴보고자 한다.

## 가. 국내 도구

### 1) 영유아 기본학습능력발달 평가도구

영유아 기본학습능력발달 평가도구는 언어, 인지, 사회·정서 발달 부문에서 영유아의 전인적 발달이라는 목표를 평가하기 위한 검사이며, 발달수준이 지연되는 영유아를 적시에 발견하여 지원을 체계화할 목적으로 개발된 도구이다(이종숙 외, 2008).

이 도구는 언어, 인지 사회·정서 발달영역의 연계성을 고려한 통합적인 평가를 통해 사회적으로 취약한 집단의 유아들이 기본 학습 능력에 도달하는데 기여할 수 있도록 개발되었다. 또한 영유아의 교실상황에서 관찰되는 3가지 영역의 행동과 언어를 중심으로 교사관찰 기반의 검사도구이다. 해당 연령에 따라 2세부터 5세까지 구성되어 있는데, 5세를 살펴보면 아래의 표와 같이 평가도구 지표 및 내용이 구성되어 있다(이종숙 외, 2008).

<표 II-3-1> 영유아 기본학습능력발달 평가도구의 발달영역별 지표

발달영역	지표
언어 (11문항)	단어 이해하기
	비유나 은유 이해하기
	사동사 이해하기
	농담 이해하기
	연결어미가 포함된 복문 사용하기
	범주어(분류어) 사용하기
	기본 말소리를 바르게 발음하기
	사물 설명하기
	글자 읽기
	자신의 이름쓰기
	끝음절 인식하여 말하기
인지 (8문항)	논리적 추론하기
	문제해결하기
	공간개념 갖기
	정보수집 및 조작하기
	분류하기

(표 II-3-1 계속)

발달영역	지표
인지 (8문항)	서열하기 수리적 책략 사용하기 패턴 만들기
	또래관계형성하기 사회적문제해결하기 도덕적 판단하기 사회적 놀이하기 자기정체성 형성하기 성인과 상호작용하기 자기존중감 갖기 가상놀이하기 환경에 대한 태도보이기 정서 조절하기 정서 인식하기 감정이입하기
사회·정서 (12문항)	

주: 이종숙 외(2008). 내용 참조, 표 재구성.

본 평가 도구는 유아교육기관에서 이루어지는 일상적인 일과활동 중에서 자연스럽게 관찰하는 것을 원칙으로 한다. 따라서 교사는 유아들의 일과활동을 주의 깊고 민감하게 관찰해야 하며, 관찰기관 동안 보이는 행동을 가능한 한 있는 그대로 평가하도록 하는 것이 특징이다.

## 2) 한국아동패널 검사도구(Panel Study on Korean Children)

‘한국아동패널 연구’는 아동이 출생 이후 어떠한 변화과정을 거쳐 가면서 어떠한 환경에 적응하고 성장하는지 파악하기 위한 종단자료를 생성하는 연구이다(육아정책연구소, 2014, 2015. 4. 3 인출). 한국아동패널은 인간발달의 생태이론(Bronfenbrenner, 1979)과 생애이론(Elder, 1998, Arredondo et al., 2006)을 이론적 틀로 반영하여, 아동을 둘러싼 환경을 체계적으로 유목화 하여 조사 영역을 구성하였다(장명림, 신나리, 박수연, 2006; 최윤경 외, 2014, 재인용). 조사영역에 따라 개발된 연구 모형은 아동, 부모, 가족 등의 인적 요소는 물론, 육아지원 서비스, 지역사회, 육아지원정책 등 아동 주변 환경의 제 분야를 포괄하여 시계열적, 다층적으로 분석하는 특징을 가지고 있다(장명림 외, 2006, 최윤경 외, 2014, 재인용).

<표 II-3-2> 한국아동패널 연도별 조사 실시 계획

	0세	1세	2세	3세	4세	5세	6세	7세 (초1)	9세 (초3)	12세 (초6)	심층조사	주요 연구 활동
1차년도 (2008년)												1차년도 조사 실시
2차년도 (2009년)											1차	2차년도 조사 실시
3차년도 (2010년)											부가조사 주)	3차년도 조사 실시 1차 국내 학술대회
4차년도 (2011년)												4차년도 조사 실시 2차 국제 학술대회
5차년도 (2012년)												5차년도 조사 실시 3차 국내 학술대회
6차년도 (2013년)											2차	6차년도 조사 실시 4차 국제 학술대회
7차년도 (2014년)												7차년도 조사 실시 5차 국내 학술대회
8차년도 (2015년)											3차	
9차년도 (2017년)												
10차년도 (2020년)												

자료: 최윤경 외(2014). 한국아동패널 2014. 육아정책연구소. p.15.

한국아동패널 2014에서 제시하는 차수별 아동 특성 변인 중 검사도구와 관련된 내용만 줄여서 살펴보면 아래의 <표 II-3-3>과 같다. <표 II-3-3>을 살펴보면 한국아동패널 연구에서는 아동의 발달 측면을 4가지 영역으로 나누어서 조사하고 있다(최윤경 외, 2014, 재인용). 신체발달, 인지발달, 언어발달, 사회·정서발달 측면 안에서 각각 하위영역 검사로 나누어지며, 신체발달 내에서는 기본적인 체중 신장 두위(1, 2차)는 모든 차수에서 조사되었으며, 운동기능 중 대, 소 근육 운동 발달은 아동이 나이가 먹어가면서 추가·보완되는 형태를 띠고 있었다. 인지발달의 경우는 5세 때부터 본격적으로 학업적 능력, 창의성, 지능, 학습준비도를 파악하고 있으며, 언어 발달의 경우 또한 표현 및 수용어휘력 파악과 조음, 발화, 읽기여부 등 포괄적인 언어검사를 시행하고 있다.

<표 II-3-3> 한국아동패널 차수별 발달영역 조사표

구분	하위변인	1차	2차	3차	4차	5차	6차	7차
신체 발달	체중, 신장, 두위(1,2차)	○	○	○	○	○	○	○
	운동발달, 미세운동/적응	○	○	○	-	-	-	-
	운동기능: 대/소근육 운동 발달	○	○	○	-	-	○	■
	우세 손	-	-	-	-	-	○	○
	기타(발달 전반)	-	○	-	-	-	-	-
인지 발달	문제해결	○	○	○	-	-	-	-
	인지발달: 학업적 능력	-	-	-	-	○	○	■
	창의성	-	-	-	-	○	-	-
	지능: 언어성, 동작성	-	-	-	-	-	⓪	-
	학습준비도	-	-	-	-	-	-	■
언어 발달	의사소통	○	○	○	-	-	-	-
	언어발달	○	○	○	-	○	○	○
	표현어휘력	-	-	-	○	-	⓪	■
	수용어휘력	-	-	-	○	-	○	○
	조음/발화/읽기여부	-	-	-	-	-	-	■
사회 · 정서 발달	기질	○	○	○	○	○	-	-
	애착	-	⓪	-	-	-	-	-
	사회성	○	○	○	-	-	-	-
	놀이/사회적 활동	-	○	○	○	○	-	-
	선호하는 놀이 영역	-	-	-	○	○	○	○
	또래 상호작용	-	-	-	○	○	○	○
	문제 행동	-	-	-	-	○	○	○
	자아개념	-	-	-	-	-	○	-
	자아존중감	-	-	-	-	-	-	■
	사회적 유능감	-	-	-	-	-	-	■
	주관적 행복감	-	-	-	-	-	-	■
	기관/교사 선호도	-	-	-	○	○	○	○

주: ⓪는 심층조사, ○는 일반조사, ■는 일반조사로 실시하되 2014년에 새로 추가되거나 보완된 경우를 의미함.

자료: 최윤경 외(2014). 한국아동패널 2014. 육아정책연구소. p.17 내용을 재구성함.

한편, 사회정서 발달의 경우, 연령별로 중요하게 파악되는 변인을 구체적으로

검사 또는 조사하였으며, 5세 이후로는 놀이영역, 또래상호작용, 문제행동, 자아 존중감, 사회적 유능감, 주관적 행복감 등, 현재 아동들의 발달에 이슈가 되는 하위영역별 검사를 진행하고 있다.

### 3) 학교준비도 검사

#### 가) 학교준비도의 개념

미국의 대표적인 유아의 패널 조사인 National Education Goals Panel(NEGP)에서는 학교준비도(School Readiness)와 관련된 5개 영역을 제시하였는데, 이는 신체적 건강과 근육발달(physical well-being and motor development), 사회·정서 발달(social & emotional development), 학습에 대한 동기와 태도(approaches toward learning(including motivation)), 언어발달(language development(including emergency literacy)), 그리고 수학과 과학이 포함된 인지 및 일반적 지식(cognition and general knowledge(including mathematics and science))이었다(Snow & Van Hemel, 2008; 방소영, 2013 재인용). 각각의 영역별로 살펴보면, 신체적 건강과 근육발달은 유아의 건강상태, 운동능력, 신체적 결함 또는 약물노출 여부에 관한 내용이 포함되며, 사회·정서적 발달은 긍정적인 자아인식, 사회적 상호작용능력, 정서표현 및 이해능력, 협동을 의미하며, 학습에 대한 동기와 태도는 과제수행에서의 열정, 호기심, 인내력 등을 포함하고 있다. 그리고 언어발달은 어휘능력과 글자 인식, 소리와 문자 조합을 이해하는 능력뿐만 아니라 이야기의 구성, 그림과 문자를 이용한 생각 표현 등 언어기반 능력의 다양한 능력을 포괄하며, 인지 및 일반적 지식은 형태인식, 공간지각 등의 인지적 요소와 사회적 관습, 사람들 간의 차이 인식 등의 능력을 포함한다(Enns, 2012; 방은정, 2015 재인용).

한편, 국내에서의 학교준비도는 학교생활의 모든 부분에 대한 적응능력과 수행능력을 위해서 필수적으로 요구되는 능력으로 정립하고 있으며, 이는 지적인 능력뿐만 아니라 신체적 능력, 사회·정서적인 능력의 준비정도까지 포괄하는 광의의 개념으로 볼 수 있다(방은정, 2015).

학교준비도와 비슷한 개념인 학습준비도(Learning Readiness)의 개념을 비교하면, 학습준비도는 개인이 특정 교재 또는 교육과정을 학습할 수 있는 능력을 갖추고 있는지 또는 학습을 할 수 있는 발달수준에 이르고 있는지를 말한다(Carlton & Winsler, 1999; 방소영, 2013; 방은정, 2015, 재인용). 주로 학습준비

도는 인지발달과 관련한 학습능력을 키우는데 초점을 두고 있는 반면, 학교준비도는 초등학교 입학 전 연령 아동의 학교생활에 대한 준비를 말한다(Kagan, 1990).

종합하면, 학습준비도는 학습과제의 성공적인 수행에 필요한 요인으로써 학습의 기능이 강조되는 개념이라면, 학교준비도는 학교생활의 전반적 적응과 학교생활의 성공적인 수행을 위하여 필수적으로 요구되는 것으로 인지적 준비 요구와 더불어 신체적 준비 정도, 사회·정서적 준비 정도까지 포함하는 보다 넓은 개념이라고 볼 수 있다(방은정, 2015).

#### 나) 국내 준비도 검사도구

국내의 준비도 검사는 해외의 검사 양상보다 상당히 제한되어 있다. 국내에서 표준화되고 타당도까지 확보한 검사도구는 임의도, 고종렬, 이혜선(1968)의 검사를 최초로 하여 다양한 준비도 검사가 있으며, 박연정(2007)의 연구에서 개발된 도구까지 포함하여 살펴보고자 한다.

##### (1) 학습준비도 검사(1968년도 개발)

임의도 외(1968)가 개발한 학습준비도 검사는 5개의 능력에 따른 9개의 검사로 구성되어 있으며, 5개의 능력은 언어능력, 공간지각, 상식, 수리력, 운동기능으로 나누어져 있고, 9개의 검사는 각각 5~12개의 문항으로 구성되어 있다. 문항은 모두 그림으로 구성되어 있으며, 초등학교 1학년을 대상으로 가, 나형의 두 종류로 구성되어 있다. 검사 기간은 54분 정도 소요되며, 전체 검사의 동형 검사 신뢰도는 0.89, 하위검사의 동형검사 신뢰도는 0.38~0.77이었다. 총점과 하위검사 간의 상관관계는 0.49~0.93이고, 하위검사 상호간의 상관계수는 0.15부터 0.67로 다양하게 나타났다.

##### (2) 학습준비도 검사(1978년도 개발)

김정권과 여광웅(1978)가 만든 검사는 First Grade Screening Test (1966년도 판)를 타당화 과정을 통하여 국내용으로 수정하여 전국 5,345명을 대상으로 표준화 한 것이다. 이 검사는 장애로 인하여 학업성취가 어려운 유아들을 가려내기 위한 선별검사이다. 신체 개념 발달, 시각-운동 협응, 정서적 성숙, 부모 인식, 놀이 인식, 지식, 지시 순종 영역 등이 포함되어 있으며, 총 27개의 문항으로 구성되어 있다. 서경희 외(2004)에 따르면 이 검사의 결과는 입학초기에 학습문제아동을 선별하는 자료로 활용할 수 있으나, 해당 검사만으로 유아의 선별

을 결정하는 일은 없어야 할 것이라고 지적하고 있다.

### (3) 학습준비도 검사(1979년도 개발)

이은혜(1979)가 개발한 검사는 미국의 Cooperative Preschool Inventory를 수정·보완한 개인용 검사이다. 검사 영역으로는 연상적 어휘, 수개념, 사회적 반응, 감각의 4가지 영역에 따라 13~19문항으로 전체 66개의 문항으로 구성되어 있다. 검사 소요시간은 15~25분이며, 검사를 시행하기 위해서 부수적인 도구(색종이 상자, 크레파스 등)들이 필요하다. 서울지역 아동 420명을 검사한 결과, 반분신뢰도는 0.94이고 KR-20 계수는 0.93으로 나타났다.

### (4) 학습준비도 검사 (1981년도 개발)

이연섭과 김성일(1981)에 의해 개발된 이 검사는, 인지 능력 평가(Cognitive Skills Assessment)와 발달수준 평가(Development Examination)를 통합하여 만든 검사이며, 대상 아동은 5~6세이다. 검사 시간은 20분 내외이며, 주변환경과의 관계, 협동 반응, 변별력, 기억력, 개념 형성 등의 능력을 측정하며, 이러한 영역은 다시 상식, 신체능력, 소근육 및 대근육 운동, 도형 구별, 수개념, 어휘 등 10개 하위영역으로 구분하여 학습준비도를 평가하고 있다. 본 학습준비도 검사와 학업 성취도와의 상관관계는 0.36~0.63이었으며, 타당도는 초등학교 교과목에 따라서 차이는 있었지만 높은 수준의 타당도를 보이는 것으로 나타났다(방소영, 2013 재인용).

### (5) 학습준비도 검사 (1987년도 개발)

박혜정, 강혜원, 장명림(1987)은 학습준비도 검사는 개념형성, 기억력, 구별력, 이해력, 운동 능력의 5개 영역으로 구성되어 있다. 기존의 학습준비도 관련 검사에서 개발되었던 문항을 참고로 50개의 문항을 선정하였으며, 유아의 월령단위까지 대상을 세분화하여 54개월~78개월까지의 미취학 유아 1,200여명을 대상으로 전국 단위 표준화를 시행하였다. 내적일치도 신뢰도 계수는 연령별로 0.80 이상이었고, 영역별로는 0.45~0.77로 나타났다. 검사-재검사 신뢰도 계수는 연령별로 0.87~0.91로 나타났으며, 영역별로는 0.54~0.86으로 나타났다.

### (6) 계몽학습준비도 검사 (1993년도 개발)

이영석 외(1993)가 개발한 검사는 언어생활, 수생활, 일상생활의 3개 영역으로 구성되어 있으며, 총 문항 수는 40문항이다. 집단검사와 개인검사의 혼합형으로 나누어져 있으며, 이에 따른 검사의 표준화는 2,000여명의 아동을 대상으로 실

시하였다. 전체 신뢰도는 0.79로 나타났으며, 영역별 신뢰도는 0.49~0.66으로 확인되었다고 보고하고 있다.

(7) 학습준비도 검사 (1995년도 개발)

윤명희(1995)는 국내외 학습준비도 검사와 유치원교육과정 6및 초등학교 6차 교육과정의 내용을 포함하여 검사를 개발하였다. 또한 검사 형태는 Metropolitan Readiness Test와 Early School Assessment의 형식을 영역별로 인용하였다. 이 검사는 언어, 수, 논리적 조작, 시·청각변별, 기억 영역의 총 32개의 문항으로 구성되어 있으며, 집단검사 형태로 유아 10~15명이 1개의 집단으로 실시할 수 있도록 설계되었다. 검사 진행방식은 검사를 진행하는 사람이 지시문을 읽고 유아가 검사지에 표시하는 형식이며, 유아의 발달 수준에 따라 총 2단계로 나누어져 있으며, 4세부터 5세는 1단계 검사를, 5세부터 6세는 2단계 검사를 이용하여 학습준비도를 측정하였다(방소영, 2013 재인용).

(8) 학교준비도 검사 (2007년도 개발)

박연정(2007)의 학교준비도 검사는 초등학교에 입학하기 전, 유아가 갖추어야 할 기본적 소양을 파악하기 위해 척도를 개발하였다. 이 검사는 담임교사가 평정하는 방식으로, 현재 유아의 능력이나 수준을 묻는 질문(리커트 척도)에 응답을 하게 되어 있으며, 단체생활, 정서, 상식, 학습능력으로 영역이 구성되어 있다.

<표 II-3-4> 국내 학교준비도 검사도구 목록

검사도구	제작자	측정영역 및 내용	특성
학습준비도검사 (1968)	임의도	<ul style="list-style-type: none"> <li>• 언어능력</li> <li>• 공간지각</li> <li>• 상식</li> <li>• 수리력</li> <li>• 운동기능</li> </ul>	<ul style="list-style-type: none"> <li>• 사지선다형</li> <li>• 54분소요</li> <li>• 초등 1학년 집단검사</li> </ul>
학습준비도검사 (1978)	김정권, 여광웅	<ul style="list-style-type: none"> <li>• 지식</li> <li>• 신체개념</li> <li>• 정서적지각</li> <li>• 부모상 지각</li> <li>• 놀이지각</li> <li>• 시간-운동협응</li> <li>• 지시순중</li> <li>• 기억</li> </ul>	<ul style="list-style-type: none"> <li>• FGST 원칙도 사용</li> <li>• 5개 지역별로 T점수 제시</li> <li>• 집단용 검사</li> </ul>
학습준비도검사 (1979)	이은혜	<ul style="list-style-type: none"> <li>• 사회적반응</li> <li>• 연상적어휘</li> <li>• 수개념</li> <li>• 감각개념</li> </ul>	<ul style="list-style-type: none"> <li>• 미국의 CPI 개정판 수정, 보완</li> <li>• 개인용 검사</li> <li>• 15-25분 소요</li> <li>• 연령에 따른 백분위 점수 제시</li> </ul>

(표 II-3-4 계속)

검사도구	제작자	측정영역 및 내용	특성
학습준비도검사 (1981)	이연섭, 김성일	<ul style="list-style-type: none"> <li>• 일반상식 (환경과의 관계)</li> <li>• 협응능력 (대소운동, 시각운동, 인물화)</li> <li>• 변별력(시각, 청각)</li> <li>• 기억력(청각기억, 시각기억)</li> <li>• 이해 및 개념 형성</li> </ul>	<ul style="list-style-type: none"> <li>• 5-6세 아동대상</li> <li>• 20분 정도 소요</li> </ul>
학습준비도검사 (1987)	박혜정, 강혜원, 장명림	<ul style="list-style-type: none"> <li>• 운동협응력</li> <li>• 변별력</li> <li>• 기억력</li> <li>• 개념형성</li> <li>• 이해력</li> </ul>	<ul style="list-style-type: none"> <li>• 4-6세까지의 취학 전 유아 대상</li> <li>• 전국단위 표준화</li> </ul>
계몽학습준비도 검사 (1993)	이영석, 구학봉, 노명환, 김승훈, 차미정, 고승자	<ul style="list-style-type: none"> <li>• 쓰기</li> <li>• 논리개념(분류, 서열, 도형, 시간, 공간)</li> <li>• 수리개념(수세기, 셈하기, 측정)</li> </ul>	<ul style="list-style-type: none"> <li>• 초등 1학년 대상</li> <li>• 집단 및 개인 혼합형</li> </ul>
학습준비도검사 (1995)	윤명희	<ul style="list-style-type: none"> <li>• 언어영역</li> <li>• 수개념</li> <li>• 논리적조작</li> <li>• 시각변별</li> <li>• 청각변별</li> <li>• 기억력</li> </ul>	<ul style="list-style-type: none"> <li>• 취학전 유아 대상</li> <li>• 2단계로 제작</li> </ul>
학습준비도 검사 (2007)	박연정	<ul style="list-style-type: none"> <li>• 단체생활태도</li> <li>• 정서교류능력</li> <li>• 일상생활지식</li> <li>• 학습활동능력</li> </ul>	<ul style="list-style-type: none"> <li>• 교사 평정</li> <li>• 리커트 척도</li> </ul>

주. 방소영(2013). 취학 전 유아의 학교준비도 검사 개발 및 타당화연구의 표(p.153)를 수정 보완함.

## 나. 국외도구

### 1) 영국 EPPE(the Effective Provision of Pre-School Education) Project

#### 가) 연구 개요

영국에서 실시된 EPPE(the Effective Provision of Pre-School Education) 프로젝트는 3세에서 7세의 아동들을 대상으로 한 아동발달에 관한 학령 전 아동들

의 교육 효과에 대한 내용을 조사한 것이다. 초기에 3000명의 3세 이상의 아동을 대상으로 6,7세까지(Key Stage 1)의 조사를 종단연구로 진행하였으며, 연속조사로 설계되었다. 부모, 가정환경, 학업 전 환경 등 다양한 환경 속에서 자란 아동들의 포함이 되었으며, 해당 아동들의 성장하면서 발달한 영역뿐만 아니라 이 프로젝트에 참여한 아동들과 다른 학령기의 아동들 간의 비교 연구까지 진행되어 있는 프로젝트이다(Sylva, Melhuish, Sammons, Siraj-Blatchford. 1999)

EPPE의 목표는 크게 다섯 가지의 질문을 가지고 있는데, 첫 번째로 학령기전 아동들에게 지적, 사회행동 발달에 영향을 주는 요인은 무엇인지, 두 번째로, 아동발달을 촉진시킬 수 있는 효과적인 방법이 있는지, 세 번째로, 효과적인 학령 전 시기 준비요소에는 어떤 것들이 있는지, 네 번째로, 아동발달에서 가정과 양육역사상 주요한 영향력은 무엇인지, 마지막으로 6,7세까지(Key Stage 1)의 단계까지 지속될 수 있는 영향요인이 있는가에 대해서 알아보는 목표를 가지고 있다(Sylva et al., 1999).

연구방법으로는 표준화된 아동 검사 시행, 학령 전 관계자와 학령 후 관계자가 모두 작성한 아동의 사회행동 프로파일, 부모면담, 학령 전 기관 면담, 질 관리 척도, 각종 사례 연구 및 관찰, 심층면담 등 다양한 방법으로 정보를 수집하였으며, 양적인 분석뿐만 아니라 부모, 교사의 광범위한 면담자료를 통한 분석 및 유아교육의 문헌분석까지 포함되어 있다.

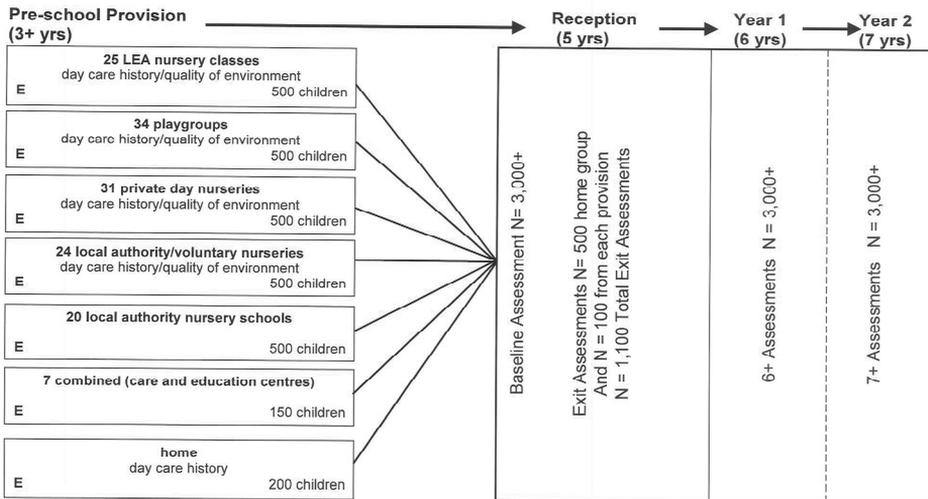
많은 자료를 기반으로 한 통계분석에서 사용한 방법은 부가가치 모형(value-added model)을 이용하여 학생 및 부모와 가정환경에 대한 것을 설명하였으며, EPPE 프로젝트 전반적으로 아동들 개개인의 수준에 맞게 층화된 다층모형을 사용하였다. EPPE는 영국의 3000명의 다양한 학생들을 표집하였고, 이러한 다양한 성장배경을 가진 아동들의 초기효과 및 개인차를 줄이기 위하여 다층모형을 사용하고 있다고 밝히고 있다(Sylva et al., 2004)

#### 나) 지역, 기관, 유아 표본 추출

개별적인 센터들의 효과와 여러 가지 형태의 기관 효과를 최대화시키기 위하여 EPPE에서의 표본 추출은 5개 지역의 6개의 지방정부를 선택하여 층화표집을 진행하였다. 이는 도시, 근교, 시골 지역을 포함하고 다양한 인종과 소외계층을 포함한 표집이었다. 이러한 표집은 실제로 다양한 지역적 맥락을 반영한 결

과 비교를 가능케 만든다. 6개의 주요한 기관의 형태는 놀이모임(playgroups), 공공보육시설(local authority day nurseries), 독지보육시설(voluntary day nurseries), 사립보육시설(private day nurseries), 사립 유아학교(nursery schools), and classes), 보육·교육 통합반 및 센터(classes, centres combining care and education)로 구분되어 각 기관형태 내에서 무선할당을 실시하여 센터를 선택하였다.

**Plan of Study**



E Entry Assessment

자료: Sylva, K., Sammons, P., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (1999). The Effective Provision of Pre-School Education (EPPE) Project: Intensive study of selected centres, p. 20.

[그림 II-3-1] EPPE표본 추출 설계도

각각의 유형별로 20-25개씩의 기관에서 20명씩의 유아가 모두 500명이 참여하여 6개의 유형 총 140 기관의 3,000명의 유아를 대상으로 하였다. 일부 지방정부에서는 특정 유형이 다른 지역에 비해 많지 않았고 일부 유형은 더 많은 경우도 있었다. 각 유형별 대표 기관은 지방정부별로 층화표집 하였다. 표집 과정에서 소규모의 기관들로 인해 최종 선택된 기관은 141개, 표집 유아는 3,000명이 넘었다. 유아교육을 경험하지 않은 효과를 검증하기 위해 200명 이상의 유아를 표집하였는데 그 수가 많지 않아 지역별로 그 수에 편차가 있다.

### 다) 유아 평가

EPPE에서 유아의 평가는 만 3세의 생일이 지난 후의 아동들을 대상으로 진행된다. 이 유아들은 4가지 인지적 영역을 평가받게 되는데, 언어 이해력, 작명 어휘, 비언어적 이해력, 공간지각능력을 평가받는다. 검사와 관련된 내용은 아래의 아래 표와 같다. 각 유아들의 사회정서 적응 능력은 아이들과 친근한 각 기관의 교사가 평가를 한다.

<표 II-3-5> 인지영역 검사도구 설명

검사명	검사내용	평가자
British Ability Scales Second Edition(BASII) (Elliot et al. 1996)	인지 발달 검사 패키지	
- 블록 쌓기	- 공간 지각	- EPPE 연구자
- 언어 이해	- 언어 능력	- EPPE 연구자
- 유사 그림	- 그림 추론 능력	- EPPE 연구자
- 작명 어휘	- 언어 능력	- EPPE 연구자
Adaptive Social Behavioural Inventory (ASBI)(Hogan et al. 1992)	사회적 행동과 감정 적응 능력	보육교직원
영어가 유창하지 못한 아동은 BASII 검사 중 비언어적 이해력과 공간지각 능력을 평가하는 검사와 사회정서 검사만 보게 된다.		

자료: Sylva, K., Sammons, P., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (1999). The Effective Provision of Pre-School Education (EPPE) Project: Intensive study of selected centres, p. 7.

이 검사들은 이후 유아발달 평가를 위한 기초 기준선으로 사용하기 위하여 선정되었다. 연구 검사관들은 지역별로 검사를 위한 훈련을 받고 신뢰도를 평가 받는다. 검사관은 유아에 대해 일대일로 평가한다. 만일 유아의 가정 언어가 영어가 아니고 영어가 능숙하지 않을 때는 통역을 고용할 수 있다.

한편, EPPE 프로젝트 내의 기관에 대한 조사는 지역의 현장 담당관(regional field officer)들은 각 기관을 정기적으로 방문하며 교사들을 관찰하고 각 기관에 대해 기록한다. 기관 특징에 대한 정보 역시 원장 면담을 통하여 얻는다. 이러한 기관 특징의 정보들은 집단의 크기, 교사 대 아동비율, 교사 훈련, 목표, 정책, 교육과정, 부모 참여 등이 있다. 지역 담당관들은 유아를 담당하고 있는 선임 담당관인 지역 책임자와 연락을 취하며 각 기관이 협력할 수 있도록 돕는다.

절차 질의 특징은 유아-교사 상호작용, 유아-유아 상호작용, 유아 활동의 구조화와 같은 일상적인 기능을 포함한다. 절차 질의 특징에 관한 정보는 Early Childhood Environment Rating Scale (ECERS)(Harms, Clifford & Cryer, 1998) 과 Caregiver Interaction Scale(CIS)(Arnett, 1989)을 사용하여 얻는다. 모든 현장 담당관들은 이들 도구를 사용하기 위한 훈련을 받았으며 신뢰도 검증을 마쳤다. ECERS의 하위 척도는 공간과 가구, 정기적 개인용무, 언어 및 추론, 활동, 상호작용, 프로그램 구성, 부모와 교사로 분류되어 있다. 또한 CIS는 보육교직원의 긍정적인 관계, 관대함, 무심함과 가혹함으로 하위영역이 구성되어 있다. 이후 유아들이 초등학교를 가게 되면 ECERS의 확장판은 좀 더 교육적인 내용들로 하위 영역이 구성되어 있는데, 언어, 수학, 과학, 환경과 다양성에 대한 영역으로 구성되어 있다.

## 라) 주요 연구결과

EPPE 프로젝트는 1997년부터 2003년까지 총 7년간 지속되면서 수많은 연구결과를 발간하였다. 2003년도에 발간한 최종보고서에서 제시한 효과적인 pre-school 교육에 관한 내용에서 제시한 총괄적인 결과는 크게 6가지로 구분하여 살펴볼 수 있다. 우선 첫 번째로, 학령 전 시기의 중요한 결과에는 pre-school의 경험은 모든 발달영역에서 긍정적인 효과를 가지는 것으로 확인되었다. 입학시기와 관련하여 빠른 입학은 지적영역 발달에 긍정적인 영향을 미친다고 한다. 또한 가정환경이 어려운 아동은 양질의 pre-school 경험이 매우 중요한 발달요인이며, 특히 다양한 사회적 배경을 지닌 아동들과의 혼합반 구성은 가정환경이 어려운 아동들에게 모든 영역이 발달하는데 더 큰 영향을 미치는 것으로 나타났다.

한편 아동 개개인의 환경과 pre-school의 환경 간의 큰 차이가 존재하는데, 몇몇의 환경적 요소들은 아동들에게 긍정적인 효과를 주는 것으로 확인되었고, 대부분의 조기교육 환경에서 양질의 교육환경이 제시되었지만, 유아원에서 좀 더 높은 수준의 교육환경을 유지하고 있는 것으로 나타났다고 보고하고 있다.

pre-school의 질과 pre-school에서의 특별한 실습의 효과성에 대한 영향에서는 질 높은 pre-schooling은 지적, 사회 행동발달에 긍정적인 영향을 주는 것으로 나타났고, 우수한 교사자원은 아동들로 하여금 높은 질 점수와 아동발달의 향상을 보장한다고 연구결과는 보고하고 있다. 질 평가에 포함되어 있는 아동과의

따뜻한 상호작용은 훈련받은 교사가 많을 때 영향력을 가지게 되고, 교육적인 관점과 사회발달적인 관점에서 상호보완적으로 동일하게 강조할 때, 아동들이 모든 발달영역에서 긍정적인 효과를 보인다고 주장한다. 또한 효과적인 유아교육은 가르치는 상호작용, 학습 환경에 대한 교육, 지속적인 생각의 공유가 존재해야하며, 모든 아동들은 가정교육의 질이 부모의 직업, 교육정도, 소득보다도 더 중요한 요소로 확인되었다. 결국 부모가 누구냐 보다 부모가 어떻게 하느냐가 중요하다는 결론을 EPPE 프로젝트에서는 나타내고 있는 것이다.

두 번째로, Key stage 1시기(6~7세)의 중요한 결과에서는 5가지로 나누어서 결과를 종합하고 있는데, 발달의 지속성에 대해서는 어느 정도 긍정적인 효과가 나타나고 있다고는 하지만 학령기로 접어들면서는 그 효과가 미미한 것으로 나타났다으며, Key stage 1 시기에 사회 행동적인 발달보다는 학습능력에 더 강한 영향력을 보이는 것으로 확인되었다. Pre-school의 질은 6세의 아동들이 읽기/수학 시험에서 큰 연관성이 있는 것으로 나타났으며, 7세에서도 연관성이 있지만 그 강도는 약해지는 것으로 확인되었다. 하지만 사회 행동발달에서는 유의미한 효과는 확인되지 않았다. Pre-school의 질과 사회 행동 발달과의 관계는 학령기 후 단계보다 학령기 전 단계에서 더 강한 관계가 있는 것으로 확인되었다. 또한 가정교육의 영향력은 학령기 전 시기와 마찬가지로 Key Stage 1에서도 아동 발달의 유의미한 영향력을 가지는 것으로 나타났다.

세 번째로, 어떠한 pre-school 경험이 아동발달에 차이를 주는지에 대해서는 pre-school 입학 시기와 다니는 기간, 서로 다른 집단의 아동들, pre-school 질이 가지고 있는 효과, 교사의 질, 교육과정의 균형, pre-school의 종류 등이 제시되었다. 입학시기와 관련하여 빠른 입학은 학령기 전 시기와 동일하게 지적영역 발달에 긍정적인 영향을 미친다고 보고하고 있으며, 서로 다른 집단의 아동들이 모여서 교육을 받는 환경 또한 취약아동들에게 상대적으로 긍정적인 효과를 보여주는 것으로 확인되었다. 한편 pre-school 질은 아동들이 6세에서 7세에 이르렀을 때, 학령기 전보다 상대적으로 약한 영향력을 가지지만 여전히 유의미한 효과를 나타내는 것으로 확인되었다.

학령기에서도 인지적인 영역과 함께 사회-정서적인 영역이 동시에 강조되는 교육과정에서 아동들의 균형적인 발달에 긍정적인 영향을 주었으며, 양육과 교육이 병행되는 유아원(nursery school)이 인지적 영역과 사회 발달에 더 큰 효

과성을 가지는 것으로 나타났다.

네 번째로, 가정환경이 아동 발달에 주는 영향에 대해서 면담을 통해 구체적으로 살펴보았다. 그 결과, 가정환경 중 인구학적인 배경, 학습 환경, EPPE를 참여하기 전 양육패턴이 영향을 미치는 요소로 나타났으며, 이러한 내용을 바탕으로 EPPE 프로젝트에서 Home Learning Environment(HLE) 측정도구를 만들어 냈다.

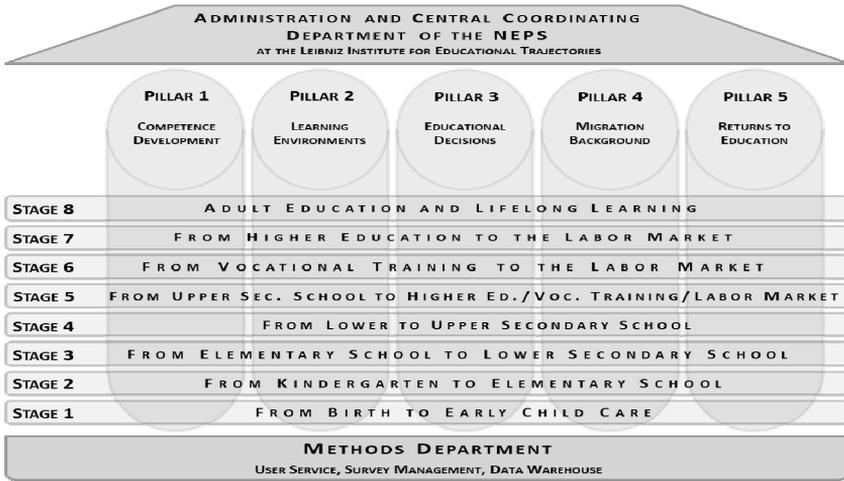
다섯 번째로, EPPE 프로젝트에 투입되기 전 조기교육에 대한 영향력을 분석한 결과는 3세 이전의 기관 경험이 반사회성 행동을 야기시키는 원인으로 나타났지만, 이는 시간이 지나 3~5세로 단계로 오면서 그 영향력은 많이 약해진다고 보고하고 있다.

여섯 번째로, 12가지 사례연구에 대한 결과는 다시 6가지의 3~5세 아동들의 주요한 특성으로 구분되는데, 성인-아동 간 언어적 상호작용의 질, 활동의 실행 능력, 지식과 교육과정의 이해, 어린 아동들이 학습하는 방법에 대한 지식, 아동을 대하는 성인들의 자세, 아동들의 높은 학습향상을 위한 부모 및 교사들의 역할, 아동들의 갈등요소에 대한 훈육 방법들을 제시하였다.

## 2) 독일 NEPS (National Educational Panel Study)

독일의 NEPS는 국가교육 패널 연구를 의미한다. 이 프로젝트는 유아시기부터 늦은 성인기까지 독일의 전 교육과정에 관련된 조사뿐만 아니라 이러한 교육과정으로 인한 성인기에 나타날 수 있는 가능성까지 분석하는 프로젝트이다.

NEPS는 개인에 미치는 영향 및 주요 교육과정이 사회에 미치는 영향을 중단적으로 분석하는 방법을 제시하고 있으며, 프로젝트의 주요 관심사인 개인의 역량 강화는 어떻게 이루어지고 있으며, 교육과정 내에서 여러 중요한 의사결정은 추후에 어떤 영향력을 미치는지, 가족이나 또래집단에서 어떠한 역량 및 학습을 하는지와 더불어 유치원, 학교, 대학, 직업교육, 심화교육 등 교육과정 구조에 따른 개인의 성취 역량에 초점을 맞추어 구성되어 있다(National Educational Panel Study, 2010)



자료: National Educational Panel Study. (2010). The National Educational Panel Study (NEPS) IEA - DPC contributions. In IEA General Assembly. Gaborone, Botswana, p. 15.

[그림 II-3-2] NEPS의 이론적 틀

[그림 II-3-2]를 살펴보면, 크게 5가지 측면(pillar)이 각 단계별로 포함이 되는데, 이는 전 생애에 걸친 역량 발달, 전 생애의 교육과정 속의 교육환경, 사회 불평등과 교육의 의사결정, 이민자들의 교육 적응도, 전 생애의 교육으로의 회귀의 주제가 있으며, 이러한 주제는 개인의 전 생애에 걸친 8단계에 각각의 요소로 연구주제화 된다. 이러한 8단계 중 첫 번째 단계와 두 번째 단계가 영유아의 특징을 나타내는 단계라 할 수 있는데, 첫 번째 단계에서는 만 4세 이하의 유아들이 어린 시절 교육과 역량 강화에는 어떤 것들이 있으며, 교육과정을 어떻게 이어가야하며, 가정 안팎으로 양육이나 교육적으로 준비를 해야 되는 요소가 어떤 것들이 있는지 등 유아가 앞으로 받게 될 교육에 대한 준비 및 유아기의 교육 및 양육 경험에 대한 내용들이 포괄적으로 조사하게 되며, 두 번째 단계에서는 거의 모든 아동들(전체의 96%)이 4세 이후가 되면서 유치원에 입학하게 되는 독일의 특성 상, 교육 체제에 진입하게 되는 유아들의 기관 적응에 관한 내용과 각 아동의 개별화된 교육과정을 추적하며 그에 따른 발달 변화를 조사하게 된다. 특히 2단계에서는 이론기반의 평가가 실시되며, 이를 통하여 교육경험에 따른 역량 발달을 확인할 수 있게 된다(National Educational Panel Study, 2010).

측정도구는 NEPS에서 두 번째 단계인 유치원 단계에서 실시되는 역량 검사를 살펴보고자 한다. 유치원 단계에서는 2년에 한번씩, 총 두 차례 검사가 실시되는데, 한 번의 검사는 2일간에 걸쳐서 검사가 진행된다. 구체적인 검사 일정은 아래의 표와 같다.

<표 II-3-6> 1차년도 검사 일정

1차년도 검사 1일차					
구인	문항수	소요시간	검사형태		
과학 역량	26	30분	그림문항		
자연과학영역의 질차적 메타인지	1	1분	그림문항		
1차년도 검사 2일차					
구인	문항수	소요시간	검사형태		
듣기평가	문단 수준 듣기평가(문법적 이해)	48	10분	그림문항	
	단어 수준 듣기평가(어휘적 이해)	77	20분	그림문항	
메타인지 영역	문법적 이해 역량	1	1분	그림문항	
	어휘영역 이해	1	1분	그림문항	

자료: National Educational Panel Study. (2010). The National Educational Panel Study (NEPS) IEA - DPC contributions. In IEA General Assembly. Gaborone, Botswana

<표 II-3-7> 2차년도 검사 일정

2차년도 검사 1일차					
구인	문항수	소요시간	검사형태		
수학 역량	26	30분	그림문항		
기본 인지능력	인식 속도	42	90초	지필식	
	추론	12	6분	지필식	
2차년도 검사 2일차					
구인	문항수	소요시간	검사형태		
단계적 특성 측정	초기 단계 철자	26	3분	언어, 도구 문항	
	음운 정보 처리	58	30분	언어, 도구 문항	
	보상지연	1	1분	언어, 도구 문항	
메타인지영역	수리적 역량	1	1분	그림문항	

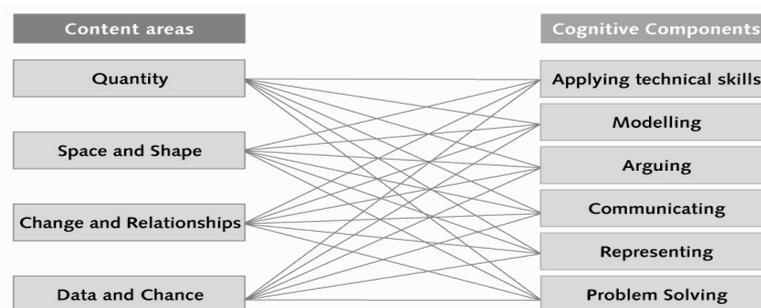
자료: National Educational Panel Study. (2010). The National Educational Panel Study (NEPS) IEA - DPC contributions. In IEA General Assembly. Gaborone, Botswana

과학역량은 NEPS 조사에서 환경, 기술, 건강의 맥락의 과학적 지식의 사용으로 이해되는 영역이다. 이러한 개념은 독일 교육부의 기준(Standard of the Conference of Ministers of Education for the medium-level school-leaving qualification)과 미국의 American Association for the Advancement of Science에서 활용된 가이드라인을 이용하여 문항을 추출하였으며, 선택된 문항은 물질의 속성, 시스템, 발달, 상호작용의 자연과학의 영역이 포함되어 있다.

메타인지(절차적 메타인지)영역에는 학습을 하는데 있어서 계획, 모니터링, 통제 활동들에 대한 적용능력이 포함되어 있는 영역이다. NEPS 안에서의 메타인지는 계획, 모니터링, 통제뿐만 아니라 각 영역에 따른 메타인지적 판단까지 포함되어 있는 문항으로 구성되어 있다.

듣기평가는 학교준비 뿐만 아니라 사회적 교류에서도 중요한 역량이 되는 언어적인 능력을 판단하는 목적으로 구성되어 있다. 특히 NEPS 안에서의 언어적 역량은 1차 조사에서는 단어와 문단수준으로 나누어서 검사를 진행하고 있으며, 2차 조사에서는 좀 더 구체적인 영역으로 나누어서 검사를 진행한다. 검사도구는 국제적으로 가장 많이 사용되고 있는 Peabody Picture Vocabulary Test를 독일버전으로 사용하지만, 1차년도에는 검사를 축소하여 사용하고 있다. 2차 검사에서는 TROG-D(Tests for Reviewing Grammatical Understanding)의 간편검사로 진행하고 있다.

수학역량에서는 PISA에서 제시한 수학적 문해력에 기반한 수학적 역량을 강조하고 있다(OECD, 2003, 24). 이러한 수학적 역량을 나이에 맞게 다시 문항을 재편하여 검사를 구성하고 있으며 각 영역은 아래와 같이 구성된다.



자료: National Educational Panel Study. (2010). The National Educational Panel Study (NEPS) IEA - DPC contributions. In IEA General Assembly. Gaborone, Botswana

[그림 11-3-3] NEPS의 수학적 역량의 틀

NEPS에서의 수학적 역량은 크게 양(Quantity), 공간과 모양(Space and Shape), 변화와 관계(Change and Relationship), 자료와 우연(Data and Chance)의 영역으로 나누어지고, 이러한 영역에 포함된 인지요소에는 기술적 능력 적용(Applying technical skills), 모형화(Modelling), 토론화(Arguing), 의사소통(Communicating), 대표성(Representing), 문제해결(Problem Solving)로 구성된다. 위의 구성은 PISA연구와 독일 국립 수학 교육 기준에서 도출한 내용이다.

기본 인지적 능력을 판단하는 것은 두 가지 검사로 나누어지는데, 그림 상징 검사(NEPS-BZT)와 추론검사(NEPS-MAT)이다. 그림 상징검사는 Digit-Symbol Test(Lang, Weiss, Stocker and von Rosenblatt, 2007)를 개량한 검사이며, 추론 검사는 기초적인 추론 검사를 그대로 사용하였다(Baltes & Lindenberger, 1999).

철자 및 음운 정보처리는 문자게임을 통하여 검사를 받게 되는데 크게 4가지 게임으로 구성되어 있다. 수 게임(The number game)은 “Kaufman Assessment Battery for Children”(K-ABC)검사에서 실행된 형태를 그대로 가져왔으며, 문장 완성 게임(The game “Build the right word”)은 독일의 “Test für Phonologische Bewusstheitsfähigkeiten(TPB)와 “Onset-Reim-Synthesieren”에서 공동작업으로 게임을 개발하였다. 라임게임(The rhyme game)은 Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten(BISC)에서 고안한 게임이며 단어의 짝을 만들어 구성하여 라임이 맞는지 틀리는지 구별해내는 게임이다. 또한 턴 게임(The turn game)은 Hamburg-Wechsler-Intelligenztest für Kinder III(HAWIK-III)에서 시행한 것을 가져온 게임이며, 이 게임으로 아동의 작업기억을 판별할 수 있다. 마지막으로 찾기 게임(The search game)은 원 검사에서 NEPS에 설계에 맞게 고안한 게임으로 아동들의 초기 운율과 음운학적인 인식을 하고 있는지 여부에 초점을 맞춘 게임이다.

보상지연은 자기통제 능력을 알아보기 위한 척도로 NEPS에서는 첫 검사날 마지막에 옷가방을 주면서 동시에 그 안의 내용물은 볼 수 없게 하여 당일 열어보지 않으면 다음 날 두 개의 옷가방을 주겠다고 하는 검사이다. 이러한 검사는 아동들이 더 나은 행동을 하는 것과 스트레스를 견디는 정도, 자아존중감 등을 판단할 수 있는 중요한 검사이다.

### 3) 미국의 헤드스타트 REDI

Head Start REDI (REsearch-based, Developmentally Informed) 프로그램은

펜실베이니아주에서 실시하는 프로그램으로 학교 준비도를 높이기 위해 유치원 교육과정을 변경하여 그 중재의 효과를 측정하는 프로그램이다. 비록 국제 발달 비교 연구는 아니지만 저소득층 자녀들의 학업성취도의 향상을 위한 프로그램과 평가로 실시한 프로그램의 효과를 측정하기 위한 다양한 도구의 사용이 특기할만하다. Welsh, Nix, Blair, Bierman와 Nelson (2010)에 의해서 보고된 연구에서는 학습준비도 검사는 저소득 가정을 돕기 위해 필요함을 역설하고 있는데, 이연구는 헤드스타트 교육을 받은 164명의 펜실베이니아주의 유아들을 대상으로 pre-K 시작시, pre-K 마칠 시, K 마칠 시에 각각 검사를 시행했다.

검사는 문해와 읽기, 수학, 인지능력과 언어를 측정하였다.

먼저 문해와 읽기에 대해서는 유아 초기 문해에 대한 활자 지식, 혼합, 엘리슨 척도 (Lonigan, Wagner, Torgesen, & Rashotte, 2007; Welsh et al., 2010 재인용)를 사용하였다. 활자지식은 유아가 글자나 단어 그림을 말하고, 글자를 읽는 것으로 구성되어 있으며, 혼합에서는 유아에게 단어의 부분을 조합하게 하거나 바른 그림을 가리키거나 전체 단어를 말하게 하는 음성 처리, 엘리슨 테스트에서는 복합어에서 일부를 생략한 부분을 재구성하도록 하거나 바른 그림을 가리키게 하는 것으로 구성되어 있다.

수학 검사는 우드콕-존슨 III 성취도 검사의 응용문제 척도 (Woodcock et al., 2001; Welsh et al., 2010 재인용)가 사용되었는데 이 검사는 손가락 두 개를 보이거나 사물을 세는 것과 같은 유아의 수 및 수량 이해를 검사한다.

인지검사는 세 가지가 실시되었는데 단어를 제시된 순서를 뒤집어 기억하는 검사, 썩기 두드리기 검사는 검사자가 썩기를 한번 두드리면 피험자는 두 번, 검사자가 두 번 두드리면 피험자는 한번 두드리는 검사인데 연습을 한 후 유아는 16가지로 혼합된 두드리기 과제를 수행한다. 마지막으로 차원변화 카드 분류하기는 파랗고 빨간 토끼와 배가 그려진 카드를 사용하는데, 유아는 먼저 색이나 모양으로 카드를 분류한 후에 나머지는 다른 차원으로 분류하게 하는 검사이다.

언어검사는 단어를 제시하고 이를 가장 잘 설명한 그림을 찾도록 하는 어휘 검사와 문장을 듣고 4가지 그림 중 의미가 가장 잘 맞는 그림을 고르는 문법 이해 검사와 종합 언어검사가 사용되었다.

연구 결과 헤드스타트 REDI 프로그램에 참여했던 유아들은 문해와 수학 검사 등에서 효과가 있음을 밝혔고, 장기 효과(유치원 졸업시)도 검출되었다고 주장하였다. 그러나 앞서 iPIPS 검토에서 살펴본 바와 같이 초기 1년 혹은 2년의

중재 효과로는 충분하지 않을 수 있다는 제언에 따라 초등학교 진학 후에까지 추적을 통해 그 효과의 지속여부를 확인할 필요가 있었을 것이다. 그러함에도 불구하고 이 연구를 통해 저소득층의 중재효과가 드러남에 따라 구체적인 프로그램의 효과의 원인을 규명하는 노력이 필요할 것이다.

#### 4) 싱가포르 SKIP<sup>1)</sup>

Singapore Kindergarten Impact Project(이하 SKIP)은 싱가포르 유치원 교육의 영향력을 평가하는 최초의 대단위 평가 프로젝트이다. 2012년 Lien재단의 Starting Well 보고서에서 싱가포르의 유아교육이 45개국 중 29위로 평가되자, 자국내 유아교육에 대한 전반적인 평가가 필요하다는 싱가포르 수상의 연설은 SKIP 프로젝트를 수행하는 계기가 되었다. 더불어, 2013년 ECDA (Early Childhood Development Agency, 유보 통합기관)의 발족과 함께, 정책적으로 유아발달 평가 프로젝트를 강조하는 분위기가 조성되어 SKIP 프로젝트를 추진하게 되었다.

SKIP의 목적은 싱가포르 유치원 교육의 과정적 요인(교사-아동 상호작용)과 구조적 요인(교실규모, 교사비율, 물리적 환경, 교사 질 등)이 아동의 학습경험에 어떻게 영향을 미치는가를 파악하는데 있다. SKIP에는 사회정서 역량, 수 개념, 언어 및 문해, 자기조절 능력, 대근육 운동능력을 포함하고, 가정환경요소를 조사하여 유아들의 개인차를 조사결과에 반영하고자 하였다.

SKIP은 각 지역별로 비례배분 할당하여 1,200명의 아동을 표집하였으며, 다음과 같은 4가지 종류의 자료를 수집하였다.

첫째, 유아의 환경(구조적 요인)과 교사-아동간의 상호작용 질(과정 요인)

- 프로그램의 질을 광범위하게 평가할 수 있는 ECERS-S도구를 활용하여 구조적요인과 과정 요인 조사
- 실제 운영 공간 및 가구배치, 육아 일상, 언어추론, 활동, 상호작용, 운영 프로그램 구조, 부모와 지원인력으로 구성되어 있음.
- 또한 ECERS-E 검사로 유아들의 문해능력, 수학, 과학 및 환경, 다양성 영역을 측정함. 교실의 환경을 평가하는 CLASS검사로 교실을 측정하는데, 이 검사 안에는 정서적 지원, 교실 조직, 교육지원의 3가지 영역으로 교실의 환경을 측정함.

1) 싱가포르 출장시 SKIP 책임자인 Dr. Rebecca Bull과의 면담내용을 정리함.

둘째, 유아 성과를 측정한다. 유아의 성과 측정은 크게 4가지 영역으로 사회 정서 역량, 수 개념, 언어 및 문해, 자기조절 능력, 대근육·소근육 운동능력으로 구분하고 있다. 각 영역에 있어서 다음과 같은 구체적인 검사 도구를 사용하였다. 자기조절 능력(머리-발끝-무릎-어깨 짚기 등), NEPSY-II 표상작업 도구, 작업 기억 및 실행능력 과업 측정, 수개념 및 언어, 비언어적 지적능력 측정, 사회정서 역량에 관하여 NEPSY-II 검사, 대·소근육은 IED-III 검사도구.

셋째, 교사 질문을 조사 차수에 따라 달리 구성하고 있다. 1차 질문지에서는 주로 교사 본인에 대한 내용(교육정도, 자격, 교사 경험, 외적 동기, 인성, 자기 효능감 등)을 조사하며, 2차부터 4차에서는 유아와 관련한 SDQ(사회정서 기능 평가 도구), CBRS와 SISS(학습과 관련한 행동 및 성과 채점 도구), BRIEF(자기 조절 능력 평가도구)를 사용한다.

마지막으로 부모에 대한 질문을 교사와 유사한 방식으로 구성하여 사용한다. 1차 질문지에서는 기본적인 인구학적 특성 및 부모 자신에 관한 내용을 조사, 2차부터 4차에서는 모국어, 집안에서의 학습활동, 잠자는 습관, 부모 건강, 유아 행동 특성에 관한 질문으로 자료를 수집한다.

이러한 SKIP 프로젝트는 종단적 연구로 2017년 말까지 추진될 예정이다. 2015년 1월부터 5월까지 1차 검사를 시행하였으며, 7월부터 8월까지 교실 관찰, 8월부터 11월까지 2차 검사를 진행하였다. 향후 2차례의 검사가 추가적으로 실시될 예정이다.

### III. 유아발달 국제비교 평가도구 분석

유아의 발달을 평가하기 위한 국외의 유아발달 평가도구의 개발 역시 국내의 유아발달 평가도구와 마찬가지로 다양한 영역별 도구가 개발되어 있으나 총체적 평가 도구, 특히 국제적으로 타당도를 인정받은 도구는 매우 소수에 불과하다. 본 연구에서는 유아의 발달을 총체적으로 평가하는 도구들 중에서 국제적 타당화의 과정에 힘을 기울이고 있는 도구들과, 국제기구들이 개발 및 확산과정에 있는 도구들을 중심으로 살펴보도록 한다.

#### 1. EDI (Early Development Instrument)

##### 가. 개요

EDI(Early Development Instrument; 이하 EDI)는 취학 직전 아동의 건강과 발달 상황에 대하여 측정하며, 전반적인 학교 준비도(School readiness)를 측정하는 도구로 교사가 검사자로 역할 한다(Janus & Offord, 2007). EDI는 캐나다 British Columbia주의 초등 및 유치원 교사들, 유아를 위한 행동단체(Early Years Action Group)와 부모와 가정 문해 센터(Parenting and Family Literacy Centres)의 자문을 받아 개발되었으며, 1998년도에 예비조사 이후, 문항의 적합도와 타당도 및 신뢰도 확인 작업을 거쳐서 2000년도에 검사의 최종버전이 출시되었다. 이후, 검사 대상자를 캐나다 전역으로 확대하여 2007년도에는 캐나다 내에서만 520,000명의 취학 전 아동들이 검사를 받았으며, 2011년도에는 캐나다, 호주, 미국, 자메이카 4개국의 자료로 검사도구의 타당도와 일반화 가능성에 대해서 연구를 진행하였다(Janus, et. al., 2011).

##### 나. 평가영역 및 문항 구성

EDI는 5개의 발달영역, 총 104개 문항으로 구성되어 있다. 5개의 영역은 각각 신체적 건강(Physical Health and Well-being), 사회적 역량(Social Competence),

정서적 성숙도(Emotional Maturity), 언어와 인지발달(Language and Cognitive Development), 의사소통 능력(Communication Skills), 일반적 지식(General Knowledge)으로 구성되어 있으며 이는 다시 16개의 세부 발달영역으로 구분되며 아래 <표 III-1-1>과 같다.

<표 III-1-1> EDI 발달영역 및 세부영역과 예시문항

EDI 발달영역 (EDI Domain)	세부영역 (Sub-domains)	예시문항 (Example item)
신체적 건강 (Physical Health and Well-being)	학교활동을 위한 신체적 준비도 (Physical readiness for school day)	배고픈 상태로 학교 도착한다. (arrives at school hungry)
	신체적 독립 (Physical independence)	동작을 잘 조정한다. (has well-coordinated movements)
	대근육·소근육 능력 (Gross and fine motor skills)	사물을 조작할 수 있다 (is able to manipulate objects)
사회적 역량 (Social Competence)	전반적 사회적 역량 (Overall social competence)	다른 아동과 잘 어울릴 수 있다 (Is able to get along with other children)
	책임감과 존경심 (Responsibility and respect)	행동에 대한 책임을 진다 (accepts responsibility for actions)
	학습에 대한 접근/학습방법 (Approaches to learning) 새로운 사물 탐색을 위한 준비도 (Readiness to explore new things)	독립적으로 공부한다. (works independently) 새로운 사물을 탐색하고 싶어한다 (Is eager to explore new items)
감정적 성숙도 (Emotional Maturity)	친사회적 행동과 타인에게 도움을 주는 행동 (Prosocial and helping behaviour)	곤경에 처한 다른 아동을 돕는다 (help other children in distress)
	불안하고 두려운 행동 (Anxious and fearful behaviour)	불행하거나 슬퍼보인다. (appears unhappy or sad)
	불행 또는 슬픔 (appear unhappy or sad)	물리적으로 싸운다. (gets into physical fights)
	활동과잉과 무관심 (Hyperactivity and inattention)	가만히 있지 못한다. (is restless)

(표 III-1-1 계속)

EDI 발달영역 (EDI Domain)	세부영역 (Sub-domains)	예시문항 (Example item)
	기본적인 읽기·쓰기 능력 (Basic literacy)	자신의 이름을 쓸 수 있다 (is able to write own name)
언어와 인지발달 (Language and Cognitive Development)	읽기·쓰기 능력과 셈하기와 기억 사용 (Interest in literacy/numeracy, and use memory)	수를 이용한 게임에 흥미가 있다 (Is interested in games involving numbers)
	발달된 읽기·쓰기 능력 (Advanced literacy)	문장을 읽을 수 있다 (Is able to read sentences)
	기본적인 셈하기 (Basic numeracy)	20까지 셀 수 있다 (Is able to count 20)
의사소통능력 (Communication Skills)	-	자신의 요구에 대해 명확히 소통하며, 다른 사람을 이해할 수 있다 (Is able to clearly communicate one's own needs and understand others)
일반적 지식 (General Knowledge)	-	세상에 관한 일반적 지식에 대하여 관심을 보인다. (Shows interest in general knowledge about the world)

자료: Janus et al(2007). Table 3.1 Domains, Sub-domains, and Sample Questions on the EDI (p. 5)

EDI 프로젝트에는 현재 전 세계적으로 12개국이 참여하고 있는데 EDI를 개발한 캐나다를 비롯하여, 미국, 호주, 칠레, 이집트, 영국, 네덜란드, 뉴질랜드가 참여하고 있다. 또한 자메이카, 코소보, 몰도바, 멕시코 등은 World Bank의 후원을 받아 참여하고 있다. 이들 참여하는 국가들 전체가 각 국가 내 취학전 아동의 전체가 참여하는 것은 아니다. 대부분의 국가들의 일부 지역에서 이 프로젝트에 참여하고 있지만, 호주에서는 60개 이상의 커뮤니티에서 EDI를 사용하여 취학 전 아동의 발달 실태를 분석하고 있어 적극적으로 EDI를 활용하고 있다(www.Offord.com, 2015. 3. 6. 인출).

EDI 문항의 특징은 특정 상황에서의 목표가 아닌 전반적 발달영역의 기초에 해당하는 내용들로 구성되어 있다는 점이다. 전반적인 발달은 문화적인 요인에

기반한 것이므로 참여하는 각 나라의 사정에 따라 문항을 바꿀 수 있는 것이 특징이다. 때문에 하나의 개발된 도구를 단순히 다른 나라에 번안하여 적용한 것이 아닌 참여국의 전문가들과 함께 문항 수정작업을 거친 대단위 검사이다.

요약하면, EDI는 유아의 주요 발달영역과 관련된 학교준비도를 평가하기 위해 고안된 교사평정용 검사도구이다(Janus, Brinkman, & Duku, 2011; Janus & Offord, 2007). EDI는 전인적 발달평가를 목적으로 하는 도구로 유치원 학급에 속한 아동들에게 개별적으로 적용할 수 있도록 구성되었지만, 이를 통해 수집된 자료는 학교나 지역, 국가단위로 합산하여서 결과의 해석에 있어서는 모든 유아의 점수 분포를 기준으로 집단수준을 파악하는 데만 사용하고 개인적 발달평가를 목적하고 있지는 않는다(Janus et al. 2011).

#### 다. 국제수준의 타당화 과정

캐나다에서 EDI가 개발된 이후 캐나다와 호주의 유아를 대상으로 신뢰도 및 타당도에 대한 검토가 이루어졌다(예: Janus & Offord, 2007; Forget-Dubois, Lemelin, Boivin, & Dionne, 2007). 캐나다에서 조사한 EDI의 구성타당도 및 동시타당도는 Janus와 Offord(2007)에 의해 보고되었다. Forget-Dubois 외(2007)는 아동의 유아기 EDI 점수와 초등학교 1학년 인지 검사에서의 성취도를 비교하여 EDI의 예측타당도를 조사하였다. Guhn과 Zumbo(2007)는 캐나다의 아동을 대상으로 조사한 EDI 자료를 통해 아동의 성별, 영어를 모국어로 하지 않는 경우, 원주민인지의 여부에 따라 문항 수준의 편향이 없음을 밝혔다.

호주에서는 EDI의 내용타당도(Hart, Brinkman, Blackmore 2003), 동시타당도와 구성타당도(Janus et al. 2011), 예측타당도(Brinkman et al., 2007)에 대한 연구가 이루어졌으며, 호주 원주민 적용에 대한 연구도 수행되었다(Goldfeld et al. 2009). EDI 영역은 미국 아동 표본을 대상으로 했을 때에도 타당한 심리측정도구이며, 자메이카 아동의 평가와도 연관되어 있음이 확인되었다(Janus et al. 2011; Janus & Offord, 2007).

영어권 국가들인 캐나다, 호주, 미국, 자메이카의 유아를 대상으로 EDI의 결과를 살펴보고 비교한 Janus et al.(2011)의 연구는 EDI가 각 국가에서 유아의 학습준비도를 측정하기에 적절한 도구인지 그 타당도를 검증하였다. 본 연구에 참여한 국가의 경제적 발달 수준은 1인당 국민총소득(GNI)을 기준으로 세계은행이 분류한 기준에 따라, 캐나다, 호주, 미국은 선진국으로, 자메이카는 개발도

상국으로 구분되었다.

이를 위해 연구자들은 EDI의 발달 영역이 4개 국가에서 유사한 심리측정의 특성을 나타내는지를 검증하고, 발달영역 중 특히 언어 및 인지발달 영역과 수용언어 검사 간의 관련성에 있어서 국가 간의 차이가 있는지를 확인하고자 하였다. 수용언어 검사로는 개정판 Peabody 그림 어휘 검사(Peabody Picture Vocabulary Test-Revised; PPVT-R, Dunn, Dunn & Price, 1989)가 활용되었다. PPVT-R 자료는 캐나다와 호주, 자메이카의 유아로부터 수집되었다.

캐나다에서 사용되는 EDI의 원본을 중심으로 하되, 국가에 따라 약간의 수정이 이루어졌다. 원래 EDI는 총 5개의 발달영역, 104개의 문항으로 구성되었으나, 미국에서 시행된 조사에서는 하나의 문항을 제거하였고, 호주의 경우에는 척도의 정확성을 저해하지 않는 것으로 확인된 9개의 문항을 조사에서 제외하였다. 각 국가별로 문항의 단어를 수정하여 조사에 활용하였다.

## 라. 타당화 주요결과

### 1) EDI의 심리측정 특성에 대한 국가 간 비교 분석

Janus et al.(2011)에서는 각 국가에서 수집된 EDI 자료를 비교해 본 결과를 통해 대부분의 발달영역에서 측정된 심리 특성이 유사하다고 보고하였다. 다만 신뢰도를 산출한 결과, 신체적 건강 및 웰빙(Physical Health and Well-being) 영역의 경우에는 가장 낮은 수준의 내적일관성이 나타났다. 특히 신체적 독립(Physical Independence)의 세부영역에서는 Cronbach의  $\alpha$ 가 .09에서 .40에 머무르면서, 현저히 낮은 신뢰도를 나타냈다. 연구자들은 이처럼 낮은 일관성이 나타난 이유에 대해, 신체적 건강 및 웰빙의 세부영역과 관련된 문항의 응답에 편차가 거의 없기 때문으로 설명했다. 그러나 일부 항목에서 낮은 요인 부하량으로 인한 한계가 있음에도, 신체적 건강 및 웰빙 영역은 그 내용에 있어서 여전히 중요성을 가지고 있다고 강조하였다.

이후 타당도 검증을 위한 분석에서는 다른 세 나라에 비해 표본 크기가 작은 자메이카가 분석 대상에서 제외되었다. 타당도 검증을 위해 먼저 선행연구와 동일한 요인 수를 지정하여 탐색적 요인분석을 시행함으로써 요인 부하량을 검토하고, 이후 구조방정식 모형을 이용하여 확인적 요인분석을 실시하였다. 자료 분석의 결과, EDI의 세부영역에서는 항목에 따라 근소한 차이가 있었으나, 전체

적인 발달영역에서는 각 국가 간에 유사한 요인 부하량을 가지는 것으로 나타났다. 이를 통해 세 나라에서 모두 EDI 요인 구조의 일관성과 유사성이 유지됨이 확인되었다.

모형의 적합도 분석을 위해 카이제곱( $\chi^2$ ) 값 외에도 CFI(comparative fit index), TLI(Tucker-Kewis Index), RMSEA(Root Mean Square Error of Approximation)의 세 지수를 산출하여 비교하였다. 모든 모형에 대한  $\chi^2$ 값은 통계적으로 유의했으나, 표본의 크기를 고려할 때 이를 통해 얻을 수 있는 정보는 제한적이다. 이를 보완하기 위해 함께 살펴본 CFI 및 TLI 지수는 모형이 통계적으로 적합함을 보여주었다. 그러나 RMSEA 지수(0.063-0.228)를 고려할 때 EDI 발달영역의 항목들은 개념적으로 조합이 이루어지기는 하였으나 적합도가 높은 편은 아니라는 사실이 확인되었다. 한편, 각 국가별 적합도 분석 결과를 비교했을 때 모두 유사한 양상을 나타냈다. 즉 5개의 발달영역 가운데 언어 및 인지발달에서 가장 좋은 수준의 적합도가 확인된 반면, 정서적 성숙은 적합도가 가장 낮은 영역으로 밝혀졌다.

## 2) EDI의 언어 및 인지발달 영역과 PPVT-R의 관련성

Janus 외(2011)는 EDI의 발달영역 중 언어와 관련된 영역의 점수를 어휘 검사 결과와 비교함으로써 관련 양상을 밝히고자 하였다. 이를 위해 캐나다, 호주, 자메이카에서 수집된 EDI와 PPVT-R 결과를 분석하였다. 그 결과, 두 검사 간에 유의한 상관관계가 있으며 이러한 양상은 세 나라에서 공통적으로 나타났다( $r=.19-.38$ ,  $p<.001$ ).

이와 더불어 선별 검사도구로서의 특성을 조사하기 위해서 두 검사 결과를 비교하여 민감성(sensitivity), 특수성(specificity), 정적 예측도(positive predictive value), 부적 예측도(negative predictive value)를 분석하였다. 그 결과 캐나다, 호주, 자메이카에서 모두 민감성은 낮고(37.1%-50%), 특이성은 높게 나타났다(85.5%-88.1%). 또한 이러한 국가들에서는 공통적으로 정적 예측도가 낮고(0.22-0.30) 부적 예측도가 높은 것(0.92-0.94)으로 밝혀졌다. 이는 낮은 EDI 점수가 수용언어 능력이 결여되었음을 의미한다고 보기는 어려우나, EDI에서 낮은 점수를 받지 않은 아동은 수용언어에 문제가 없을 가능성이 높다는 사실을 의미한다.

EDI팀은 PPVT-R이 국제적으로 통용되는 도구라는 점에서 EDI와 비교 평가

할 수 있는 척도로 선정하였다. 그러나 EDI는 아동 발달에 대한 총체적인 측정(holistic measure)을 위해 고안된 도구라는 점에서 EDI팀은 PPVT-R과 같이 특정 영역의 능력을 평가하는 척도와는 근본적으로 성격이 다르다는 점을 한계로 지적하였다.

## 마. 시사점

4개의 영어권 국가들을 대상으로 자료를 수집하여 분석하는 국제비교를 시도한 Janus et al.(2011)의 연구는 캐나다와 호주, 미국에서는 모집단을 대표할만한 수준에서 연구가 이루어졌다는 사실에 의의가 있다. 이 연구의 대상이 된 4개국에서 EDI는 취학하는 데 필요한 아동의 발달 관련 정보를 제공하였다.

실질적으로 EDI가 훨씬 더 많은 국가에서 사용되고 있다는 사실을 감안할 때 더 많은 지역의 아동을 대상으로 한 타당도 및 신뢰도의 추가적인 조사가 이루어질 필요가 있다. 특히 비영어권 국가와 개발도상국을 대상으로 한 자료가 구축되어야 할 것이다.

이와 더불어 후속연구에서는 대상 아동에 대해 보다 구체적인 배경 정보를 포함시킴으로써 EDI의 영역별 점수가 가지는 잠재적 예측 요인들에 대해 탐색해 보아야 할 것이다. 또한 EDI를 통한 아동발달 측정과 함께 대상 아동의 맥락적 정보를 함께 수집한다면 서로 다른 나라에 거주하는 유아가 학교준비도를 갖추는 데 필요한 내용을 파악할 수 있다는 점에서, 해당 지역 및 국가의 관련 상황을 지원할 수 있을 것으로 본다. 특히 모집단 측정을 한 EDI 자료를 수집하여 교육 및 보건 행정 데이터베이스와 연계할 경우, 학교준비도의 예측하는 핵심 요소들을 밝혀내는 데 도움이 될 수 있을 것이다. 더 나아가 학교 시스템과 연계함으로써 아동의 교육적 성취 결과를 예측하는 EDI의 예측타당도를 높일 수 있을 것이다.

## 2. iPIPS (International Performance Indicators of Primary Schools)

### 가. iPIPS의 기원

iPIPS (International Performance Indicators of Primary Schools)검사는 영국 Durham 대학 Centre for Evaluation and Monitoring (CEM)에서 개발된 아동

발달 평가도구이다. iPIPS 검사는 1994년 영국 학교에서 학습의 출발점을 파악하기 위하여 학생들의 읽기와 수학 능력을 측정하고자 시작된 Performance Indicators in Primary Schools On-entry Baseline assessment (PIPS)를 기원으로 하고 있다(Tymms, Merrell, & Jones, 2004).

PIPS는 1994년에 개발된 이후 3-5세 유아에서부터 초등교육을 마친 후까지 추적한 자료를 축적하고 있다. PIPS는 원래 지필식 검사로도 개발되었지만 컴퓨터를 활용할 수 있도록 발전하였으며, 이는 추후 PIPS가 국제적으로 평가도구로 거듭날 수 있는 계기가 되었다. 컴퓨터를 활용하는 PIPS의 내용과 예시문항은 다음과 같다.

<표 III-2-1> PIPS 검사 내용과 예문

내용	예문
손글씨(Handwriting)	자신의 이름 쓰기
어휘(Vocabulary)	연속된 그림에 등장하는 사물 말하기
읽기에 대한 사고 (Ideas about reading)	활자에 대한 개념 파악
음운 인식 (Phonological awareness)	운율과 반복
글자 식별(Letter identification)	대소문자가 섞인 것을 순서대로 나열
단어 인식과 (문장) 읽기 (와 이해) (Word recognition and reading)	
수학적 사고(Ideas about mathematics)	수학적 개념의 이해
수세기(Counting and numerosity)	
합계(Sums)	연산기호 없는 덧셈과 뺄셈
모양 식별(Shape identification)	
자릿수 식별(Digit identification)	
수학문제(Mathematics problems)	연산기호를 사용한 합계 포함

자료: Tymms, P., Merrell, C., & Jones, P. I. (2004). Using baseline assessment data to make international comparisons., 30(5), p. 675 내용 정리.

iPIPS의 특징 중 하나는 교사가 검사자로 유아를 평가한다는 점이다. 즉 유아에 대한 사전 정보가 충분한 검사자만이 가능한 평가를 교사가 수행할 수 있다는 장점을 지니고 있다. iPIPS의 인지발달 영역의 문항들에 대해 교사는 컴퓨터 화면의 진행에 따라 유아에게 문제를 읽어주며 답을 고르게하고, 유아의 응

답을 기록한다. 나머지 영역은 교사가 직접 응답하는 문항들로 구성되어 있다. iPIPS의 또다른 특징은 컴퓨터로 검사를 수행한다는 점이다. 컴퓨터로 검사를 치르는 덕분에 검사 직후 검사 결과를 바로 얻을 수 있으며 또한 그 결과는 온라인에 연결된 컴퓨터를 통해 서버로 전송된다. 또한 컴퓨터를 검사에 활용하여 응답자의 수준에 맞는 문제를 제시하여 응답자에게 최소한의 문제를 제시하고도 지필식 검사에서 묻는 몇 배의 문항으로 얻을 수 있는 결과를 대신할 수 있다는 장점을 가지고 있다.

## 나. 국제수준의 타당화

Tymms, Merrell과 Jones(2004)는 PIPS의 국제비교를 위한 사전연구를 통해서 수학, 읽기, 어휘에 관한 뉴질랜드, 스코틀랜드, 호주의 일반인과 원주민, 영국 일반인 및 청각장애, 영어가 모국어가 아닌 학생 등 다양한 집단을 대상으로 검사를 실시하였다. 각각의 집단의 문화적인 차이에 대한 일반적인 우려와는 달리 대부분 유사한 난이도의 유형을 구성함으로써 국가 또는 문화 간의 비교에 무리가 없음을 보여주었다. 물론 예외적으로 난이도의 차이를 드러낸 경우도 있었는데 이는 국가 간의 차이보다는 청각장애아 집단과 같이 언어사용의 차이에 따른 결과를 반영한 것이어서 이들이 보여준 난이도의 차이는 오히려 이질적인 집단의 검사 및 비교가 불가하다는 단점이라기보다는 집단의 특성을 구별해주는 특징을 잘 드러낸 장점으로 이해할 수도 있다. 이러한 PIPS의 국제비교 결과는 Third International Mathematics and Science Study (TIMSS)나, Programme for International Student Assessment (PISA)와 같은 유명한 국제비교 연구들과 마찬가지로 유아의 학습준비도에 대한 국가 간 비교의 가능성을 확인하게 되었다.

이처럼 국제비교는 단지 다른 나라의 집단의 비교뿐 아니라 장애아 집단, 모국어가 주류언어가 아닌 집단이나 원주민 집단과 같은 다양한 소수 집단에 대한 비교도 가능함을 시사하였다. 즉 발달에 대한 국제비교 척도를 국가 내의 다양한 집단의 유아들을 주류 유아들과 비교를 통하여 소외계층과 같은 관심을 기울여야 할 유아들을 위한 조기 중재를 위한 매우 중요한 교육정책의 의사결정을 돕는 도구가 될 수 있다는 것이다.

이와 같은 PIPS의 국제비교를 위해 탄생한 iPIPS는 영국을 비롯하여 남아프리카 공화국, 러시아, 중국 등 점차 다양한 지역 및 국가에서 실시되고 있으며, 아동의 발달수준을 자국 내에서의 비교뿐만 아니라 국제적인 비교정책을 입안

하기 위한 기초자료를 생성하는 것을 목표로 하고 있다. 즉 유아기 발달 평가의 성과가 국가 정책 입안에 반영될 수 있도록 하는 객관적인 검사도구도구로 발전해가고 있는 것이다.

## 다. iPIPS의 평가영역

iPIPS는 인지 발달(Cognitive development), 인성·사회·정서 발달(Personal, social and emotional development), 신체 발달(Physical development)의 3 가지 발달영역과 배경정보, 총 4개 영역으로 구성되어 있다. 아래의 <표 II-1-2>는 각 영역별 세부영역 또는 문항 내용이다.

<표 III-2-2> iPIPS 4대 영역 및 세부항목

영역	세부 내용
인지발달 (Cognitive Development)	필기(Handwriting)
	어휘(Vocabulary)
	읽기에 대한 사고(Ideas about reading)
	음운 인식(Phonological awareness)
	글자 식별(Letter identification)
	단어 인식과 읽기(Word recognition and reading)
	수학적 사고(Ideas about mathematics)
	셈하기와 수 사용 능력(Counting and ability to use numbers)
	합산(Sums)
	모양 식별(Shape identification)
인성·사회·정서 발달 (Personal, social and emotional Development)	숫자 자릿수 식별(Digit identification)
	수학 문제(Mathematic problems)
	단기기억(Short term memory)
신체 발달 (Physical Development)	학교환경 적응능력(Adjusting to the school environment)
	인성 발달(Personal development)
	사회적, 정서적 발달(Social and emotional development)
배경 정보 (Contextual Information)	키와 몸무게(Height and weight)
	소근육발달(Fine motor development)
	대근육발달(Gross motor development)
	아동과 학교에 대한 배경 정보(Background information on the children and schools)

자료: www.iPIPS.org 홈페이지 참조(2014년 업데이트), 2015.03.06. 인출

위에 표에 제시되어 있는 바와 같이, 인지발달 영역은 13개의 세부 항목으로 구성되어 있다. 필기는 아동에게 자신의 이름을 쓰게 하며, 어휘는 일련의 사진에 담긴 사물을 묻고, 읽기에 대한 사고는 아동의 활자에 대한 개념을 평가한다. 또 음운 인식은 운율에 대한 것을, 글자 식별은 알파벳 대소문자를 구별해서 인식할 수 있는가를, 단어 인식과 읽기는 단어와 문장에 대한 이해를 평가한다. 수학적 사고는 아동의 수학적 개념에 대한 이해를 평가하며, 셈하기와 수 사용능력에 대한 평가, 합산에서 연산 기호를 사용하지 않고 덧셈과 뺄셈 문제를 제공하며, 모양 식별, 숫자 자릿수 식별, 그리고 수학문제에서는 연산 기호를 사용한 합산, 강조된 단추들의 순서를 연상하는 단기기억 평가로 구성되어 있다.

본 도구의 인지발달 부분은 컴퓨터나 태블릿 PC 등의 휴대용 기기를 이용해 아동이 검사를 받을 수 있도록 훈련된 검사자가 실시하며 약 20분 정도가 소요된다. 컴퓨터 프로그램은 아동에게 말로 문제를 제시하고 아동은 문제에 따라 화면을 짚거나 대답을 한다. 검사자는 아동의 반응을 화면에 적용하고 진행을 돕는다. 프로그램의 각 부분은 아동이 몇 개의 문제를 틀릴 때까지 문항의 난이도를 높여가며 제시하고, 아동이 일부 문제를 틀리게 되면 다음 부분으로 넘어가게 된다.

인성·사회·정서 발달의 평가는 정규교육의 첫째 동안의 아동의 인성, 사회, 및 정서 발달 평가를 위한 것이다. 이 평가는 인지 발달의 평가와는 달리 교실에서 매일의 상호작용과 관찰을 통해 얻은 아동에 대한 정보를 바탕으로 담임 교사가 답하게 되어 있다. 교사는 총 3개 하위 영역, 11개 문항으로 구성된 질문에 답을 하는데 통상 교사가 한 아이를 위해 답하는데 5분에서 10분 정도가 소요된다. 하위 영역은 학교환경에의 적응, 인성 발달, 사회·정서발달의 세 가지로 구성되어 있다. 이 문항들에 대한 답은 해당 아동이 5단계로 나뉜 루브릭에 기술된 내용 중 가장 잘 설명된 내용을 고르는 것으로 구성되어 있다. 평가는 학기가 시작되고 몇 주 지난 뒤와 학년을 마치는 때 다시 실시하게 된다.

신체발달의 평가는 세부 영역으로 키와 체중, 소근육 발달, 그리고 대근육 발달로 구성되어 있다.

배경정보 영역은 교사를 위한 설문과 부모나 양육자를 위한 설문으로 구성되어 있는데 아동과 학교에 대한 배경정보(예를 들면 나이, 장애, 사회·경제적 지위 등)를 얻기 위한 것이다.

## 라. 시사점

iPIPS의 국제비교에 대한 Tymms, Jones, Albone과 Henderson (2009)의 연구 및 Merrell과 Tymms (2007) 연구의 시사점 가운데 가장 주목할 만한 것은 첫째, 유아기의 조기 중재는 중요하다는 점이다. 특히 수학, 읽기에 있어서 초기 중재 여부의 차이는 어휘보다 더 컸다는 것을 밝혔다.

둘째, 그러나 일회성 중재는 이후 효과가 미미하였으며 지속적인 중재는 이후 그 효과가 지속되는 것을 관찰할 수 있었다. 반면 연속된 부정적인 학교경험은 이후에 회복하기 어려운 학업의 퇴보로 이어지는 치명적인 결과를 보였다. 이러한 결과가 유아발달평가의 국제비교에 주는 시사점은 영유아의 초기 평가 결과에 대한 처방은 직후에 개선되었다고 할지라도 추가적인 관심을 가지고 꾸준히 그 얻은 성취를 유지하게끔 노력할 필요가 있다는 점이다.

셋째, 이들의 국가 간 발달 비교의 결과는 일관되게 학교의 효과보다 교사의 효과가 더 컸음을 나타냈다. 즉 학교변인, 학교의 효과는 상대적으로 교사 변인, 교사의 효과의 크기에 비해 작다는 결과가 주는 정책적 함의는 매우 중요하다. 즉 학교 단위의 정책에 대한 관심은 교사의 발전에 대해 더 많이 관심을 기울이는 것으로 변화해야할 것으로 보인다. 마지막으로 iPIPS 도구를 통한 국제비교 연구의 데이터는 수학 및 읽기에 대한 조기 중재가 매우 이른 시기, 심지어 4세에서도 효과가 있었음을 입증하였다. 많은 나라들의 의무교육 및 무상교육의 시기가 점점 더 앞당겨지고 있으며, 이러한 연구들을 통해 그 가능성이 검증된 것이기에 시사하는 바가 크다고 할 수 있다.

## 3. EAP-ECDS (East Asia Pacific-Early Childhood Development Scales)

### 가. 개발 배경

EAP-ECDS(East Asia Pacific-Early Childhood Development Scales)는 만 3세~만 5세를 대상으로 하는 발달측정도구로, UNICEF와 Open Society Foundation의 후원을 받은 ARNEC(Asia-Pacific Regional Network for Early Childhood)이 주관하고 홍콩대학교 교육학부가 실제로 수행하는 프로젝트

를 통하여 개발되었다(Sun & Becher, 2014).

2002년에 UNICEF가 미국 콜롬비아대학교와 파트너십을 맺으며 ELDS(Early Learning and Developmental Standards)을 개발한 것이 본 도구의 시초라 할 수 있다. EAP-ECDS는 2009년까지 49개 국가가 ELDS를 각각 국가의 서로 다른 문화에 따라 적용하였을 뿐 아니라, 라오스의 '학교준비도 역량', 캄보디아의 '발달평가시험', 베트남 '영유아발달 및 학습 체크리스트'의 개발에 직접적인 영향을 미쳤다(Rao, et. al, 2015).

본 도구개발은 7개국(캄보디아, 라오스, 몽골, 필리핀, 태국, 바누아투, 베트남)을 대상으로 하였으며, 홍콩대학교 교육학부는 기존에 개발한 HKECDS(Hong Kong Early Child Development Scale)를 토대로 EAP-ECDS을 개발하였다(Rao, et. al, 2015).

## 나. 개발 및 타당화 과정

EAP-ECDS의 개발은 총 3단계로 추진되었다(Rao, et. al, 2015). 1단계는 2010년 후반에 시작되어 문헌검토로 진행되었다. 각국의 항목들을 합쳐서 재코딩한 다음, 전문가에게 자문을 구해 검사지를 완성하였다. 2단계에서는 예비조사 연구를 수행되었는데 다양한 문화적 특징을 포괄하기 위해 몽골, 피지, 중국을 선정하였다. 이후, 수정보완 사항 반영 및 타당화 검토 후 2013년 초에 최종검사지를 개발하였다. 3단계에서는 2013년 6월부터 12월까지 6개 국가의 만3-5세 아동을 대상으로 표본을 수집하였다.

### 1) 1단계

검사지를 작성하기 위해 7개 국가에서 각각 정한 영유아발달 및 학습 기준을 파악하였다.

&lt;표 III-3-1&gt; EAP-ECDS의 참여국가별 발달영역

캄보디아	라오스	몽골	필리핀	태국	바누아투	베트남
신체 및 건강 도덕 및 문화 사회 및 정서 인지 및 추론 언어	신체, 위생 및 안전 언어 인지 및 기본 개념 정서, 사회, 정신(Mental)	신체 사회정 서 창의력 수학	신체 건강, 웰빙 및 운동, 언어, pre-readi ng pre-Math 인지 사회정서 인성 및 가치	운동 및 신체 웰빙 사회 정서 인지 언어 도덕 창의	신체, 운동 건강 및 웰빙 공동체 생활 및 학습 Spiritual and Character 언어 인지	신체 정서 및 사회 언어 및 의사소통 인지

자료: Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2015). Report on Technical Support for the Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS), UNICEF.

각국의 발달영역을 파악한 후, 다음과 같은 과정으로 검사가 개발되었다. 먼저 7개 국가의 ELDS를 합친 총 1,738개의 지표를 주제에 맞추어 분류하였다. 각 카테고리에 있는 지표 수의 비율에 맞추어 각 영역에 몇 문항씩 포함할 것인가를 결정하였다. 이를 토대로 각 영역의 비율에 맞추어 총 100개의 문항을 선정하였다.

평가영역은 총 7개 범주로 분류되었다. 7개 평가영역은 학습에 대한 접근방법(Approaches to Learning), 인지 발달(Cognitive Development), 문화적 지식 및 참여(Cultural Knowledge and Participation), 언어 및 발현적 문해(Language and Emergent Literacy), 운동 발달(Motor Development), 건강, 위생 및 안전(Health, Hygiene and Safety), 사회정서발달(Socio-Emotional Development)로 구성되었다. 구체적인 문항 예시는 다음 표에 제시되어 있다.

&lt;표 III-3-2&gt; EAP-CDCS 항목 예시

항 목	내 용
준비물	1. 종이 한 장, 연필 한 자루
진행과정	평가자는 아동에게 아동 자신의 그림을 그리도록 요청한다.
지시 사항	여기 펜과 종이 가 있어. 네가 서 있는 모습을 여기에 그려볼 수 있겠나?

(표 III-3-2 계속)

국 가	내 용
점수 체계	머리 그리고/혹은 몸통에 추가로 한 가지 특징이 있다. <input type="checkbox"/> 0 <input type="checkbox"/> 1
	팔 그리고/혹은 다리를 알아볼 수 있다. <input type="checkbox"/> 0 <input type="checkbox"/> 1
	손 그리고/혹은 발을 알아볼 수 있다. <input type="checkbox"/> 0 <input type="checkbox"/> 1
	얼굴에 한 가지 알아볼 수 있는 특징이 있다 (눈, 입, 코, 귀). <input type="checkbox"/> 0 <input type="checkbox"/> 1
	얼굴에 알아볼 수 있는 특징이 하나 이상 추가적으로 있다 (눈, 입, 코, 귀). <input type="checkbox"/> 0 <input type="checkbox"/> 1

자료: Sun, J. & Becher, Y. (2014). Early Childhood Development in the East Asia Pacific: Development & Findings from the East Asia Pacific - Early Child Development Scales (EAP-ECDS), p. 12.

이러한 과정을 거쳐, 홍콩대학팀은 참여국 전문가로 구성된 워크숍에서 받은 피드백을 토대로 문항, 점수 체계 및 매뉴얼을 수정 및 개발하였다. 마지막으로 전문가들이 방법의 적절성을 검토하고 일부 수정대안을 제시하며 검사지 초안을 완성하였다.

## 2) 2단계

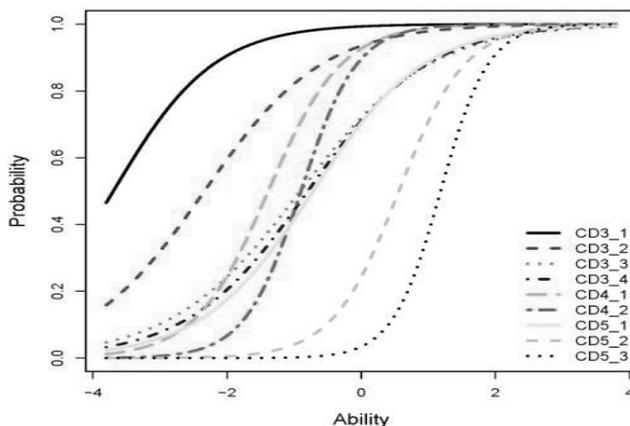
검사지 완성 후, 예비조사 연구를 통해 범국가적으로 사용하기 위한 타당성을 검증하였다. 대상 국가에 검사지 초안을 발송하여 번역 및 검토를 하여 현장 조사를 한 다음, 수정을 거듭했다. 이후, 대상 국가 중 다양한 문화권을 가지고 있다고 판단되는 세 국가, 몽골, 피지, 중국을 선정하여 시범연구를 시행하였다. 현장 조사 후 몇 가지의 수정 제안사항은 다음과 같다.

- 아동에게 더 친숙한 자료를 이용할 것
- 아동이 이해할 수 있는 언어를 사용할 것
- 정해진 표준 지침대로 할 것
- 아동들이 응답하는 방법을 여러 가지로 인식할 것: 꼭 말로 해야 할 필요는 없으며 행동을 인식하는 방법 역시 고려할 것
- 아동과 검사자의 래포 형성을 위해 노력할 것

예비조사 연구는 같이 진행되었다. 대상 국가에 검사지 패키지 전달 -> 몽골, 피지에 검사 방법 및 매뉴얼 전달 -> 몽골, 피지에서 검사 실시 -> 중국에서의 자료 수집 준비 및 몽골, 피지 자료 입력 -> 중국에서의 자료 수집 및 현장 지원 -> 몽골, 피지 자료 분석 및 보고서 작성 순으로 이루어졌다.

각 국의 표본 수는 각각 120명이었으며 성비를 약 5대 5로 맞추어 진행하였으며, 검사 소요시간은 몽골이 약 45분, 피지는 약 40분, 그리고 중국은 약 80분으로 나타났다. 본 검사지의 신뢰도는 Cronbach's alpha가 .70 이상인 것으로 확인되었고, 타당도는 내용타당도인 전문가 검토를 거쳤다. 또한, 연령별 타당도는 연령이 높은 아동이 낮은 아동에 비해 더 유의한 것으로 나타났다. 문항 적합도는 각 문항의 차이를 잘 반영하는가, 문항 변별도, 문항 유용도를 검증하였다.

먼저 Model Selection(모델 선정) 과정에서 문항반응이론(Item Response Theory)을 측정하였다. 문항반응이론을 통해 피검사자의 능력이나 심리적 특성과 같은 '잠재적 특성'을 측정하고 문항이 이원, 다원화되는지 확인하고자 하였다. 1-모수 로지스틱 모형(Rasch model), 2-모수 로지스틱 잠재 특성 모형(two-parameter logistic latent trait model), 그리고 Birnbaum's three-parameter model을 검토한 결과, 2-모수 로지스틱 잠재 특성 모형이 가장 적합한 것으로 나타났다(AIC=2689.19, LRT=119.45,  $p < 0.001$ ) (Rao, 2012).



자료: Rao, N. (2012). The development and validation of contextually relevant scales to assess and monitor child development and learning in the East Asia Pacific Region. p. 49.

[그림 III-3-1] 문항특성곡선(Item Characteristic Curve)

또한 문항곡선의 특성을 분석하여 각 검사 문항에 대한 정답 확률과 능력 척도 간의 관계를 살펴보았는데, 그래프와 같이 기울기가 높을수록 그 문항은 능력의 상하집단을 더 잘 변별하는 것으로 나타났다. 이 곡선의 문항 난이도와 문항변별도는 난이도를 의미하는 가로축, 추측도를 의미하는 세로축으로 파악할 수 있다. 예비조사 연구 결과, 본 검사지는 총 100문항에서 85문항으로 선별되었으며, 자세한 항목은 다음과 같다.

<표 III-3-3> EAP-CDCS 영역별 최종 문항 수

영역	문항 수
1. 학습접근방법	7문항
2. 인지 발달	21문항
3. Cultural Knowledge and Participation	10문항
4. Language and Emergent Literacy	16문항
5. 운동 발달	7문항
6. 건강, 위생 및 안전	9문항
7. 사회-정서 발달	15문항
총 계	85문항

자료: Sun, J., & Becher, Y. (2014). Early Childhood Development in the East Asia Pacific: Development & Findings from the East Asia Pacific - Early Child Development Scales (EAP-ECDS). (20 p.) "The Revised Scale after Phase II"

이러한 결과를 바탕으로 예비조사 연구의 몇 가지 제한점이 발견되었다. 먼저 심리측정의 문제로, 수집할 수 있는 자료의 한계, 요인 분석의 어려움과 신뢰도, 소요시간 관리, 그리고 문화적으로 더 적절한 문항을 개발해야 한다는 지적이 있었다. 또한, 아동들이 검사과정 도중 힘들어한다는 이에 아동이 피곤해지는 경우 검사 중간에 나갈 수 있게 하자는 제안이 있었다. 더불어 검사 시 검사자와 아동이 매뉴얼 이상의 내용을 하는 경우 전체 검사시간에 지장을 준다는 점이 지적되었는데 아동이 검사에 보다 집중할 수 있도록 검사자가 되도록 필요한 말만 하도록 권장되었다.

마지막으로 검사자가 의도치 않게 기록지를 아동에게 노출하여 아동을 불안하게 하는 상황, 아동이 의욕적으로 검사에 참여하도록 하는 방법 역시 논의되었다. 이에 아동에게 칭찬하는 표현을 많이 활용하되, 물질적 제공은 금지하는 것으로 논의되었다.

### 3) 3단계

#### 가) 3단계 연구과정

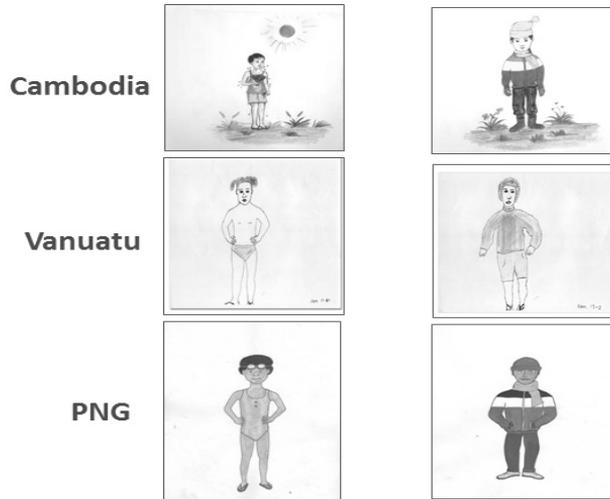
2단계 예비조사 연구 결과를 바탕으로 3단계에서는 1,2단계 참여국가 중 일부가 중도포기 혹은 새로 참여하면서 최종 6개 국가(캄보디아, 중국, 몽골, 파푸아 뉴기니, 동티모르, 바누아투)에서 검사도구의 타당성을 검증하였다. 2013년 6월부터 12월까지 표본 수집이 진행되었고 표본은 거주지역(도시, 도서·산간 지역), 성별, 영유아기관 이용 유무를 고려하여 각국마다 900~1,803명을 선별하였다. EAP-ECDS 지침 매뉴얼, 채점표, 부모설문지는 각국의 언어로 번역되었다. 부모 설문지는 사회인구학적 요인을 조사하기 위해 배부되었으며, 가족구성원, 가족 자산, 가족 내 학습 지원자 등을 조사하였다. 또한, 아동의 초기학습 및 발달, 건강 및 습관 등을 파악하고자 하였다. 부모설문지의 아동에 대한 조사내용은 <표 III-3-4>와 같다.

<표 III-3-4> AP-CDCS 부모설문지 영역별 구성

영역	능력 수준
인지 발달	새로운 내용을 학습하고 문제를 해결할 수 있는가
사회정서 발달	타인에 대한 배려심 및 감정조절과 같은 사회적 능력을 보여주는가
운동 발달	달리기, 제자리뛰기가 가능한가 젓가락, 숟가락/연필/펜을 쥌 수 있는가
언어 및 발현적 문해	언어 능력
건강, 위생 및 안전	위생 습관이 잡혀있는가(예-혼자서 손을 씻는 것) 안전 규칙을 준수하는가(예-뜨겁거나 위험한 것은 만지지 않는 것)
문화적 지식 및 참여	지역사회의 중요한 행사에 참여하는가(축제 포함)
학습접근방법	새로운 일에 집중할 수 있는가(TV 시청 제외)

자료: Sun, J., & Becher, Y. (2014). Early Childhood Development in the East Asia Pacific: Development & Findings from the East Asia Pacific - Early Child Development Scales (EAP-ECDS). (41 p.) "Items for parents' rating of children's competence"

아동용 설문지는 각국의 문화 차이를 고려하여 만들어진 문항이 제공되었고 아래 예시와 같다. 아래는 17번 문항으로, 그림에 맞는 계절을 맞추는 내용이다. 중국, 바누아투, 파푸아뉴기니가 여름, 겨울을 나타내는 그림을 보여주지만 문화권 특성에 맞춘 의복, 소품을 제공함으로써 아동들이 쉽게 이해할 수 있도록 구성되었다.



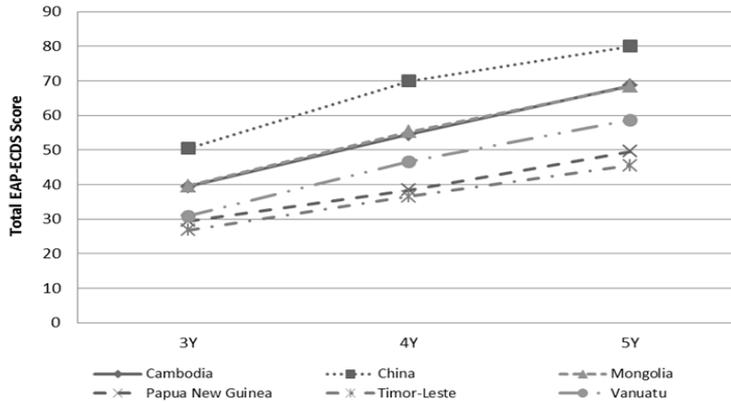
자료: Sun, J., & Becher, Y. (2014).. Early Childhood Development in the East Asia Pacific: Development & Findings from the East Asia Pacific - Early Child Development Scales (EAP-ECDS). (25 p.) “Country adaptations: Item 17- Predicts season or weather”

[그림 III-3-2] 동일 문항에 대한 문화적 차이

나) 3차 주요결과

본 척도는 내적합치도 Cronbach’s alpha(0.74~0.95)와 내용타당도를 확인하여 타당도를 검증하였고, 각 영역과 그룹을 비교하기 위해 MANOVA로 3(연령)×2(성별)×2(지역)로 분석하였다. 아동검사 결과와 부모 평가는 일관성이 있는 것으로 나타났으며 문항 분석은 예비조사 연구와 마찬가지로 적합도, 문항변별 기능 분석, 문항정보 곡선, 영역별 검사정보 곡선(test information function curves)을 확인하였다. 상관관계는 0.14~0.47(p<.001) 수준으로 나타났으며, 영유아기관 이용 유무, 가구소득, 모 학력, 성별, 그리고 연령에 따라 차이를 보였다(P<.001). 특히, 영유아기관 등원유무(6.52)와 연령(12..28)은 다른 영역에 비해 설명력이 높은 것으로 나타났다(Rao, et al, 2015).

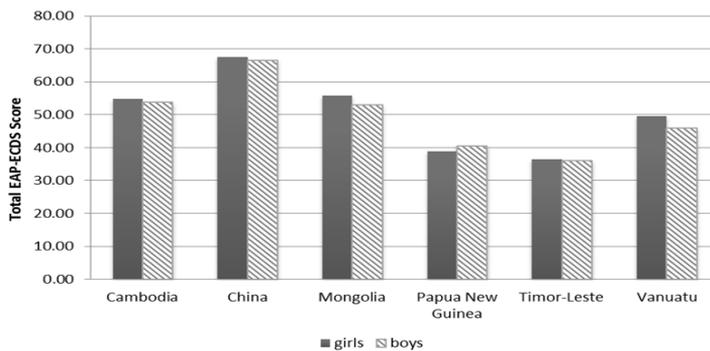
여기서 주목할 부분은 연령, 성별, 지역별 차이이다. 아래는 국가별 EAP-ECDS 점수를 연령별로 나타낸 그래프이다. 6개 국가 모두 연령에 비례하여 높은 점수를 보여주고 있으며 만 3세와 5세의 차이는 평균 20점 이상으로 나타났다.



자료: Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2015). Report on Technical Support for the Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS). UNICEF. p. 259.

[그림 III-3-3] 연령에 따른 국가별 EAP-ECDS 점수 차

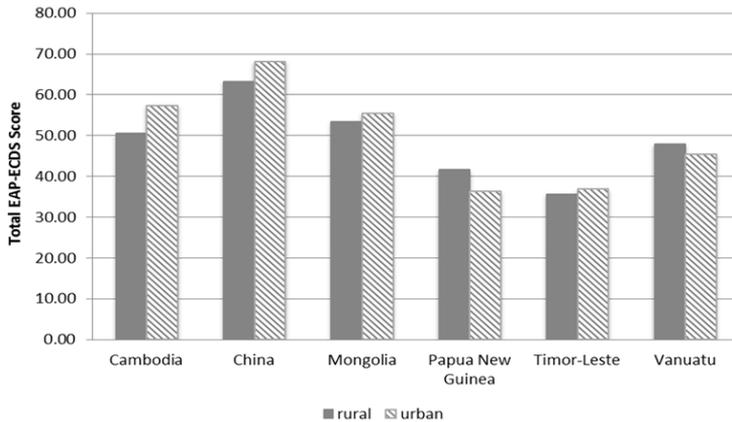
성별에서도 차이를 볼 수 있었는데, 파푸아뉴기니, 동티모르를 제외한 캄보디아, 중국, 몽골, 바누아투는 모두 남아가 여아보다 더 높은 점수를 받았으며, 이는 성별 간 불균형을 의미한다.



자료: Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2015). Report on Technical Support for the Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS). UNICEF. p. 259.

[그림 III-3-4] 성별에 따른 국가별 EAP-ECDS 점수 차

지역별에서도 역시 도시-농촌 간의 차이를 확인할 수 있다. 파푸아뉴기니와 바누아투를 제외한 캄보디아, 중국, 몽골 그리고 동티모르는 농촌보다 도시에 사는 아동들의 전 영역에서 발달점수가 더 높은 것으로 나타났다. 마찬가지로 도시-농촌 간의 불균형을 나타내었다.



자료: Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2015). Report on Technical Support for the Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS), UNICEF. p. 260.

[그림 III-3-5] 도시·농촌에 따른 국가별 EAP-ECDS 점수 차

## 다. 시사점

본 연구는 첫 번째로 지역을 다루는 아동발달도구라는 점에서 의의가 있으며, 상향식 접근으로 도구평가를 발달시켰으며, 각 문화마다 맥락에 맞추어 제공했다는 점에서 의미있다고 볼 수 있다. 본 연구결과는 도·농, 성별에 따른 차이, 부모참여 증진, 영유아보육 및 교육 커리큘럼 발달 및 검토에 유용하게 사용될 수 있을 것이다. 후속 연구발달을 위해서는 접근성이 낮은 지역에서의 자료 수집, 언어 선택, 번역 절차를 더욱 철저히 해야 하며, 문화적으로 적절한지 철저히 검사하며, 직·간접 평가 중 어느 것이 적합한지 고려해야 하며, 도구사용자 및 운영자를 훈련시킬 수 있는 방법이 고안되어야 할 것이다(Sun & Becher, 2014: 54-56).

## 4. MELQO (Measuring Early Learning Quality and Outcomes)

### 가. 학습성과 평가지표 개발 배경

국가와 지역에 따라 교육에 대한 장벽과 이를 극복하기 위한 해결책은 다르지만, 교육이 인간의 기본적인 권리이며 개인에게 기회를 제공하는 관문이라는 점에는 이견이 없다. 이러한 상황에서 중요한 것은 단지 학교에서 보내는 시간이 아니라, 그 안에서 습득하는 학습 내용과 학습의 질적 수준이다.

그러나 현실적으로 초중등 교육과정을 이수한 다수의 아동·청소년은 이후의 삶을 영위하는 데 필요한 최소한의 지식이나 기술, 능력을 갖추지 못한 상태에 머물러 있다. UNESCO(2012)의 보고에 따르면, 취학하지 못하거나 초등학교 입학 후 4년 이내에 교육과정에서 탈락하는 아동이 약 1억2천만 명에 이르며, 적어도 2억5천만 명의 초등학교생이 읽기, 쓰기, 셈하기에서 발달수준에 적합한 학업 기술을 제대로 수행해 내지 못하는 것으로 보고하고 있다.

또한 열악한 교육수준은 아동의 미래를 위태롭게 하는 요인이 될 수 있다. 이는 소득이 낮은 국가에 사는 아동에게만 국한된 이야기가 아니라, 부유한 나라에 사는 아동에게도 해당되는 문제이다. 따라서 자신이 속한 국가나 처한 상황에 관계없이, 모든 아동이 기본적인 교육을 받을 수 있도록 할 필요가 있다.

지난 십여 년 간 국제적인 차원에서 모든 아동에 대해 초등교육의 기회를 확대하기 위해 노력해 왔으나, 내용적인 면에서는 여전히 매우 낮은 수준에 머물러 있는 실정이다. 지금까지는 교육적 혜택을 받지 못하는 지역의 아동들이 학교교육을 받을 수 있도록 접근성을 높이는 데 초점을 맞추어 왔다면, 앞으로는 교육적으로 열악한 환경에 있는 지역의 아동도 일정 수준에 도달할 수 있도록 초점을 맞추어야 하는 것이다.

이러한 맥락에서 학습의 질적 수준과 학업성취 결과를 평가할 수 있는 범세계적 기준이 필요하다. 기존에 여러 국가에서 학습성과 수준을 평가하는 척도들을 개발하고 사용해 왔으나, 이를 이용한 국가 간 학습의 질적 수준 및 학업성취 결과를 비교·평가하고 문제점을 진단하는 데에는 한계가 있었다. 경험적인 증거에 기반하여 학습에서의 위기를 진단하고 이에 부합하는 정책을 강구하기 위해서는 먼저 세계의 각 지역에 거주하는 아동의 학습성과 수준을 체계적으로 평정할 수 있는 도구가 필요하게 되었다.

이를 위해 UIS와 CUE는 2012년 7월에 학습성과 평가지표 전문팀(Learning Metrics Task Force: LMTF)를 구성하였다. 매우 포괄적인 분야에서 다양한 관계자들이 모인 이 자리에서는 전 세계의 모든 아동과 청소년에게 더 나은 교육의 경험을 만들어 주어야 한다는 궁극적인 목적에 따라 국제적인 수준에서의 평가 기준을 개발해야 한다는 데 의견을 모았다. 여기에는 30개의 회원 기관과 186명의 기술적 전문가들로 구성된 working group, 118개 국가에서 온 1,700명 이상의 자문위원들이 참여하여, 7개 항목으로 구성된 권고사항을 개발하였다. 이 권고사항에는 글로벌 패러다임의 전환, 학습 역량, 범세계적 학습성과 추적을 위한 평가지표, 국가지원, 평등, 공익으로써의 평가 강조, 실행으로 구성되어 있다.

## 나. 학습성과 평가지표 내용

앞서 제시한 7개의 권고사항 중 학습성과 평가지표와 직접적인 관련이 있는 사항은 학습 역량, 범세계적 학습성과 추적을 위한 평가지표, 국가지원이다. 이 3가지 권고사항에서 평가지표에 구체성을 확보하고자 세부내용을 살펴보도록 한다.

### 1) 학습 역량

모든 아동과 청소년은 7개 학습 영역에 걸쳐 역량을 개발한다. 학습성과 평가지표 전문팀(LMTF)는 교육체계가 아동과 청소년들에게 그들의 미래의 삶과 생계를 준비하는 데 꼭 필요한 7개 영역에서 학업을 완수할 수 있는 기회를 제공해야 한다고 권고한다. 세계 전역의 교육과정은 유아기부터 초기 중등교육(lower secondary school)에 이르기까지 이 7개의 영역에 초점을 맞추어야 할 것을 강조한다. 이는 아동이 이후에 학업을 지속하던 직업을 가지게 되는 간에 상관없이 모든 아동이 기본적으로 갖추어야 할 지식과 기술이라는 점에서 중등교육까지로 제한하였다. 학습 영역의 글로벌 체계는 정규교육과정에 한정하여 적용하기보다, 학업이 이루어지는 곳이라면 어느 곳이라든지(공동체 교육 시스템, 비정규 교육 프로그램 등) 폭넓게 적용할 수 있어야 한다.

#### 가) 아동의 학습시기

유아기, 초등교육기, 중등교육기 등과 같은 학습 단계 또는 학교교육 단계에 대한 구분은 지역이나 국가, 때로는 아동 개인에 따라 서로 다르게 나누어질 수 있다. 학업은 연속적으로 이루어진다는 개념에 따라 학교 교육의 단계를 구분하

면 아래의 표와 같이 제시할 수 있다.

<표 III-4-1> 단계, 학교 교육 단계, 대략적인 연령 범주  
(Stage, Schooling Levels and Approximate Age Spans for Developing a Framework)

단계	학교 교육 단계	대략적인 연령 범주(세)
유아기 (Early childhood)	출생 후부터 취학 전까지, 유아교육 포함 (ISCED 0)	0-8
초등교육기 (Primary)	초등교육 (ISCED 1)	5-15
초등교육 이후 시기 (Postprimary)	초기 중등교육기 (ISCED 2)	10-19

주. 여기서의 학교교육 단계는 1997년 개정된 교육에 대한 국제표준분류(International Standard Classification of Education: ISCED, UNESCO, 1997)를 근거로 함. 연령 범주는 아동발달의 단계를 포괄적으로 반영한 것으로, 학교교육 단계에 직접적으로 상응하지 않음.

자료: Learning Metrics Task Force. (2013). Toward Universal Learning: Recommendations from the Learning Metrics Task Force. (p. 19) table 1: Stages, Schooling Levels and Approximate Age Spans for Developing a Framework.

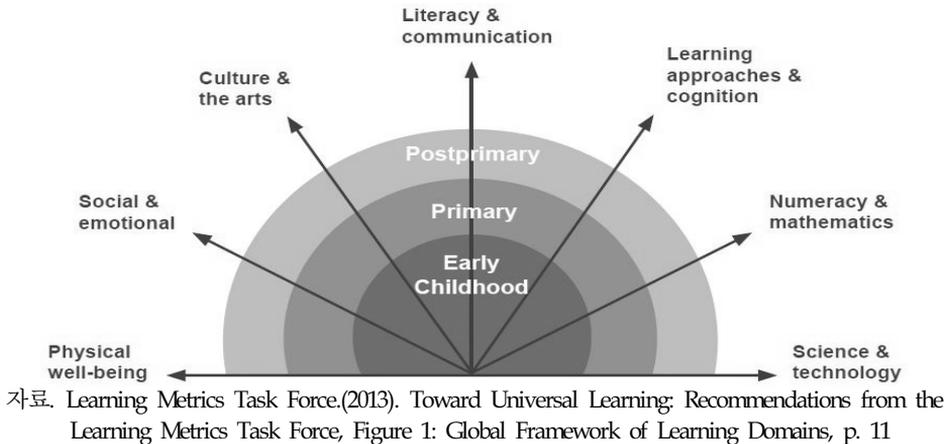
#### 나) 아동의 학습 공간

앞서 제시한 것처럼 학습성과 평가지표 전문팀(LMTF)는 대상 아동의 범주를 세 단계(유아기, 초등교육기, 초등교육 이후 시기)로 구분하면서 이에 맞추어 권고사항을 제시하였는데, 이는 학업이 학교에서만 이루어지지 않는다는 사실을 전제로 한다. 소외계층을 포함하여 많은 아동과 청소년이 비공식적인 교육 프로그램을 통해 학습과정을 거치고 있음을 반영한 것이다. 소득 수준이 높은 국가에서는 많은 유아들이 정규 유아교육 과정에 참여하지만, 세계의 여러 국가에 속한 다수의 유아들은 가정이나 지역사회에서 부모나 형제 혹은 다른 구성원에 의해 체계화되지 않은 비공식적인 교육을 받고 있는 상황이다. 또한 중등교육과정이나 이를 대신할 만한 다른 학습 프로그램을 이수하지 않은 아동들의 경우에는 일터나 가족, 지역 등과 같은 데에서 비공식적이고 구조화되지 않은 방식으로 초등학교 과정 이후의 교육을 받게 된다는 사실을 염두에 두어야 한다 (Lüftenegger et al., 2012).

#### 다) 학습 영역의 범세계적인 구성틀(Global Framework of Learning Domains)

국가나 지역에 따라 학업이 이루어지는 시기와 장소가 다르고, 상이한 교육

구조를 갖추고 있는 상황에서 어떤 결과가 범세계적인 차원에서 모두에게 좋을 것인지를 정의하는 것은 하나의 도전이다. 그러나 연구 결과를 바탕으로 학습성과 평가지표 전문팀(LMTF)은 모든 아동과 청소년에게 초기 중등교육 수준에 이를 때까지 요구되는 7개의 학습 영역을 제안하였으며, 이를 통해 포괄적이고 전인적인 틀을 마련하고자 하였다[그림 III-4-1].



[그림 III-4-1] 학습 영역의 범세계적인 구성틀  
(Global Framework of Learning Domains)

여기에 제시된 모든 영역은 유아기부터 초등교육 이후의 시기에 이르기까지 폭넓게 적용 가능하지만, 일부 영역은 특정 시기에 더 적합할 수 있다. 다음의 표에 각 영역에 대한 설명과 해당 하위영역의 예를 제시하였다<표 III-4-2>.

<표 III-4-2> 학습 영역의 범세계적인 구성틀 (Global Framework of Learning Domains)

영역	설명	하위영역의 예*
신체적 웰빙	아동·청소년의 신체 활용, 운동 조절 능력의 발달, 적절한 영양·운동·보건 위생·안전 실천의 이해 및 습취에 관한 방법	<ul style="list-style-type: none"> <li>신체적 건강 및 위생</li> <li>음식과 영양</li> <li>신체적 활동</li> </ul>
사회 및 정서	아동·청소년이 성인 및 또래와 관계를 형성하고 유지하는 방법, 다른 사람과의 관계 안에서 자기 자신에 대해 인식하는 방법 포함	<ul style="list-style-type: none"> <li>사회 및 공동체의 가치</li> <li>시민적 가치</li> <li>정신적 건강 및 웰빙</li> </ul>

(표 III-4-2 계속)

영역	설명	하위영역의 예*
문화 및 예술	미술, 공연, 춤, 창조적 동작의 영역과 관련된 활동 및 시각 예술, 미디어 아트, 문학 등을 포괄하는 창의적 표현. 가족·학교·지역사회·국가 안에서 이루어지는 문화적 경험	<ul style="list-style-type: none"> <li>• 창조적 예술</li> <li>• 문화적 지식</li> <li>• 자신에 대한 정체성 및 공동체의 정체감</li> <li>• 다양성에 대한 지각 및 존중</li> </ul>
문해 및 의사소통	아동·청소년이 속한 사회에서 사용하는 일차적 언어를 통한 의사소통. 말하기, 듣기, 쓰기, 다양한 매체를 통해 듣거나 읽은 내용에 대해 이해하기를 모두 포함	<ul style="list-style-type: none"> <li>• 말하기와 듣기</li> <li>• 어휘</li> <li>• 쓰기</li> <li>• 읽기</li> </ul>
학업 방식 및 인지	학업 방식은 학업에 대한 학습자의 관여(engagement), 동기, 참여를 의미. 인지는 이러한 여러 접근 방식을 통해 학습 내용을 습득하는 정신적 과정을 뜻함	<ul style="list-style-type: none"> <li>• 지속과 주의</li> <li>• 협력</li> <li>• 문제해결</li> <li>• 자기지시(self-direction)</li> <li>• 비판적 사고</li> </ul>
산술 및 수학	계량적 언어 및 수를 통한 과학은 주변 환경에서 관찰되는 현상을 재현하는 데 보편적으로 활용됨	<ul style="list-style-type: none"> <li>• 수 개념 및 조작</li> <li>• 기하학 및 패턴</li> <li>• 수학의 응용</li> <li>• 자료와 통계</li> </ul>
과학 및 기술	과학은 물리적 법칙과 보편적 사실을 아우르는 특정한 지식 또는 지식의 본체나 체계. 기술은 문제해결을 위한 도구의 창작 및 활용과 관련	<ul style="list-style-type: none"> <li>• 과학적 연구</li> <li>• 생명과학</li> <li>• 물리학</li> <li>• 지구과학</li> <li>• 디지털 기술의 인식과 활용</li> </ul>

주: \* 여기에 제시된 예는 하위영역의 이해를 돕기 위한 것으로, TF는 아동의 교육 단계(유아기, 초등교육기, 초등교육 이후 시기)에 따라 구체적인 내용을 제시하였음.

자료: Learning Metrics Task Force (2013). Toward Universal Learning: Recommendations from the Learning Metrics Task Force, Table 2: Global Framework of Learning Domains, p. 21.

## 2) 범세계적 학습성과 추적을 위한 평가지표

TF는 모든 국가에 적용하는 것을 목적으로 하는 학습지표를 마련할 것을 권고한다. 이들 지표는 아동이 어떤 교육을 받아왔는지를 살펴보는 것에서 더 나아가, 보다 근본적인 학습기회를 평가하게 된다. 각 지표는 기존에 제작된 평가에 대한 광범위한 검토 및 모든 국가에 적합한 학습의 전 영역을 다루고자 하는 원의를 바탕으로 선정되었다.

<표 III-4-3> 범세계적 추적을 위한 학업 지표

평가 영역	지표 설명
모두를 위한 학습	수행 평가와 학업에 대한 평가를 하나의 지표로 결합 (초등학교 과정 후반부에 읽기 능력의 유창성을 평가)
학업을 위한 연령 및 교육 사항	제때에 입학했는지를 비롯하여, 학업 수행 과정 및 과정 완수에 대해 평가. 취학하지 못하거나 중도에 탈락한 아동을 파악하기 위한 모집단 기반의 지표
읽기	3학년까지의 읽기와 관련된 기본적 능력 및 초등학교 과정 종료 시점에서의 유창성 측정
산술	초등학교 과정 수료 시점에서의 기본적인 산술 능력 및 초기 중등과정에서의 숙련도를 측정
학교준비도	유아기 학습의 허용 수준 및 초등학교 입학 시점에서의 영역별 발달 수준 측정
세계인으로서의 시민성(Citizen of the World)	청소년들 사이에서 자신이 속한 지역사회, 국가, 그리고 세계의 일원으로서 성공하기 위해 갖추어야 하는 가치 및 기술에 대해 평가
학업 기회의 폭	학습에 대한 7개의 전 영역에 걸쳐 학업 기회에 어느 정도 노출되어 있는지를 추적

자료: Learning Metrics Task Force. (2013). Toward Universal Learning: Recommendations from the Learning Metrics Task Force, Figure 6: Learning Indicators for Global Tracking, p. 23.

이제까지 사용되어 오던 지표들은 교육 기회에 대한 접근성, 성취 결과, 문해 및 산술과 관련된 기술을 측정하고 평가하는 데 초점이 맞추어졌다. 그러나 여러 국가에서 유아기 학습 및 발달, 문제해결기술, 비판적 사고력과 같이, 이전에 다루었던 것과는 다른 영역에서 학업의 진행상황을 추적할 필요가 있다는 요구를 가지게 되었다. 이에 대한 응답으로 학습성과 평가지표 전문팀(LMTF)은 지표를 구성하는 내용을 다각적으로 선정하고자 했다. 각 국가들의 입장에서 처한 상황에 맞는 목표 및 요구에 맞추어 무엇을 측정할 것인지를 분명하게 결정하기 위해서는 이에 관해 논의를 거쳐야 하고 우선순위를 정해야 하며, 실제적인 행동을 취해야 할 것이다.

### 3) 국가 지원

TF는 각 국가들의 평가 체계의 강화와, 궁극적으로는 학습성과 수준의 향상을 지원한다. 평가 및 학업의 질적 수준을 향상시키기 위해 필요한 대응책은 각국이 처한 상황과 맥락에 따라 서로 다르지만, 어떤 식으로든 교육에 대한 평가

체계를 구축하기 위해 모든 국가들이 노력하고 있다. 국가 정부 및 비정부 관계자들은 현재 평가가 이루어지는 부분과 아동·청소년이 배워야 한다고 여기는 국가 비전 사이의 간극을 최소화하면서 학업성취 결과를 향상시키고자 한다. TF는 학업에 대한 평가 및 성취 결과를 향상시키는 데 필요한 평가 체계의 질적 수준을 진단하고, 관련자들을 소집하며, 기술적·재정적 지원을 평가하는 절차가 필요하다는 사실에 동의한다. 이를 위해서는 기술적·제도적·정치적 영역에서 이해당사자들의 기술과 지원이 뒷받침되어야 한다. TF는 새로운 독립 기구를 창설하기보다, 각 분야의 관계자들 간에 협력할 것을 제안한다. 즉 국가적인 차원에서는 교사 및 시민사회와 더불어 교육부와 다른 정부 기관이, 지역적인 차원에서는 지역의 경제적·교육적 기구가 여기에 해당한다. 그리고 국제적인 차원에서는 민간 분야의 관련자들과 재단, UN 산하 기관, 시민사회, 교원기관, 연구소 및 기타 기관들이 이 권고사항에 따르기 위해 협력해야 한다.

<표 III-4-4> 국가지원 사항

기술적	<ul style="list-style-type: none"> <li>• 범세계적으로 추적하기 위해 합의된 학습 평가 추진</li> <li>• 포괄적인 협의 절차를 통해 국가 내 및 국가 간 추적을 위한 새로운 평가 개발</li> <li>• 평가도구의 질 확인 절차 마련</li> </ul>
	<ul style="list-style-type: none"> <li>• 기존의 평가 체계가 가진 강점 및 약점에 대해 자기진단할 수 있는 도구 개발</li> </ul>
제도적	<ul style="list-style-type: none"> <li>• 실천 지역 지원</li> <li>• 전문성, 교훈, 효과적인 실천 사례에 대한 공유 활성화</li> </ul>
	<ul style="list-style-type: none"> <li>• 평가의 문화 조성 및 평가 자료의 효과적인 활용 촉진</li> </ul>
정치적	<ul style="list-style-type: none"> <li>• 학습에 대한 접근에서 접근 이상의 학업으로 초점을 전환하도록 장려</li> <li>• 상향식으로 시민사회의 학습 측정 및 향상과 연계</li> <li>• 평가를 위한 자원 동원</li> </ul>

자료: Learning Metrics Task Force. (2013). Toward Universal Learning: Recommendations from the Learning Metrics Task Force, Figure 3: Supporting Countries, p. 12.

## 다. 타당화 추진 계획

2015년에 일련의 주요 문항들이 구축될 것이며, 측정도구와 관련 매뉴얼 및 가이드라인에 대한 예비조사가 실시될 예정이다. 도구 타당화 절차는 지역, 언어, 예산 규모의 다양성을 수용하기 위해 15개국에서 실시되고 있다.

예비조사 과정에서는 측정 도구에 대한 신뢰도와 타당성을 파악할 뿐만 아니라, 모든 사람에게 적용 가능한 모니터링 시스템에 도구를 통합할 수 있는 방안을 고안할 것이다. 2016년 초반까지 도구와 매뉴얼들은 어느 기관이나 단체 및

국가에서든 자유롭게 오픈 소스로 이용가능하게 될 것이다.

궁극적 목적은 모든 아동의 초기 학습과 발달의 향상을 위해 각 국가가 측정 도구들을 자국의 문화와 맥락에 맞게 적용할 수 있도록 하는 것이며, 현 모니터링 및 평가 프로그램의 전체 시스템에 측정 과정을 포함시키는데 있다.

<표 III-4-5> 국외평가도구 종합 분류표

EDI		iPIPS		EAP-ECDS		Learning Metrics	
영역	하위영역	영역	하위영역	영역	하위영역	영역	하위영역
신체적 건강	학교활동의 물리적 준비도 물리적 독립 대근육 소근육	신체 발달	키와 몸무게 소근육발달 대근육발달	운동발달	달리기, 체자리 뛰기	신체적 웰빙	신체적 건강 및 위생 음식영양 신체적 활동
				건강, 위생 및 안전	위생 습관		
사회적 역량	전반적 사회적 역량 책임과 존중심 배움에 대한 접근 새로운 사람 탐험을 위한 준비도	인성·사회·정서 발달	학업환경 적응능력 인성 발달 사회적, 정서적 발달	사회-정서 발달	타인에 대한 배려심 및 감정조절 등의 사회적 능력	사회 및 정서	사회 및 공동체 가치 사민적 가치 장신적 건강 및 웰빙
언어와 인지 발달	기본적인 읽기·쓰기 능력	인지 발달	필기 어휘 알파벳대안사모 음운 인식 글자 식별 단어인식과 읽기 수학적 사고 셈하기와 수 합산 모양 식별 숫자재질수식별	언어 및 발현적 문해	언어 능력	학업 방식 및 인지	자수와 주위 협력 문제해결 자기자시 (self-direction) 비판적 사고 수개념 및 조작 기하학 패턴 수학의 응용 자료와 통계 과학 연구
				인지 발달		언어 발달	새로운 내용 학습 및 문제 해결
	의사 소통 능력		수학 문제 단기 기억		학습 접근 방법		
일반 상식		배경 정보	아동과 학교에 대한 배경 정보	문화적 지식과 참여	지역사회 행사에 참여	문화 및 예술	창조적 예술 문화적 지식 자신/공동체 정체감 다양성

## IV. 유아발달평가 국제비교 연구결과 분석

앞서 밝힌 바와 같이 유아발달을 총체적으로 평가한 도구는 그 수가 많지 않고 국제적으로 타당화의 과정을 거치고 있는 도구는 더욱 드물다. 본 장에서는 유아발달평가의 결과를 국제적으로 비교한 연구결과를 분석함으로써 특별관심 집단(주로 취약소외계층 유아) 선정, 유아발달 차이의 요인, 발달영역별, 성별, 지역별 수준에 따른 차이를 검토하였다.

### 1. 특별관심 집단

Janus, Brinkman과 Duku(2011)는 캐나다 전역에서 두 차례에 걸쳐 수집된 EDI 기준자료(normative data)를 이용하여 아동의 특성에 따른 발달의 차이를 살펴보았다. 이들이 사용한 자료는 캐나다 전역에서 두 차례에 걸쳐 광범위하게 수집된 것으로, 두 번째 자료는 심리측정분석(psychometric analysis) 결과에 기초하여 일부 문항에 대한 코딩을 수정하여 사용하였다. 1999/2000년부터 2003/2004년까지 수집된 1차 자료수집에는 캐나다의 10개 주가 모두 참여했고, 2004/2005년부터 2006/2007년에 수집된 2차 자료수집에는 3개 연해주(Maritime Provinces; Nova Scotia, New Brunswick, Prince Edward Island)를 제외한 7개 주를 대상으로 하였다. 이 연구에서는 기본적으로 두 자료가 모두 활용되었으나, 분석대상을 세분화하여 보다 구체적인 내용을 살펴보고자 할 때에는 보다 최근의 자료인 2차 조사의 자료를 이용하였다. 정해진 기준에 따라 아동의 연령이 확인되지 않거나, EDI의 두 영역 이상에 대해 정보가 누락된 표본과 특수아는 기준표본에서 제외되었다.

Janus 외(2011)은 EDI의 특정 문항에 응답한 결과를 토대로 다음과 같은 하위 집단을 설정하였다. 1) 특수아, 2) (영어를 사용하는 학교에서) 영어 또는 (불어를 사용하는 학교에서) 불어를 모국어로 사용하지 않는 아동; 제 2언어 학습자, 3) 특수한 학습문제를 가진 아동(학급 내에서 학업을 수행하는 데 영향을 미치는 문제를 가진 아동으로, 특수아와 별도로 구분함), 4) 특별한 기술을 가진 아동(children with special skills), 5) 추가적인 평가가 필요한 아동; 특수한 도

움을 필요로 하는 아동과 구분, 6) 원주민 신분의 아동. 이상의 6가지 하위집단 중 이들 연구에서는 앞의 두 집단, 즉 특수한 도움이 필요한 아동과 영어나 불어를 모국어로 사용하지 않는 아동에 초점을 맞추었다. 이를 통해 EDI로 측정된 학교준비도를 하위집단별로 살펴보고, 앞의 6개 하위집단에 속하지 않는 아동으로 구성된 통제집단 아동의 학교준비도와 비교함으로써 아동의 특성에 따른 학교준비도의 차이를 확인하였다.

PIPS를 이용한 Tymms 외(2004)의 연구에서는 2002년 학교를 다니기 시작한 뉴질랜드, 영국, 스코틀랜드, 네델란드, 호주의 유아들을<sup>2)</sup> 대상으로 연구를 하였다. 자료의 수집은 각국에서 수집된 자료들 중에서 집단별로 1000명씩을 무선 표집한 표본을 활용하였는데 가용한 자료가 1000개 미만인 3개의 집단은 표본의 수가 1000보다 작았다. 연구의 집단 비교를 위한 흥미로운 점은 집단의 구분이 국가에 머물지 않고, 영국의 경우 청각장애아 집단, 영어가 제2 언어인 집단을 모집단에서 별도로 구별하여 집단을 구성하였고, 호주의 경우 유치원에 가지 시작한 유아와 일찍 초등학교를 시작한 유아, 그리고 선주민 집단으로 구분하였다.

### 1) 특수아 및 장애아

먼저 특수아 집단과 통제집단의 발달영역별 EDI 평균 점수를 비교한 결과, 전 영역에 걸쳐 통제집단 아동의 점수가 높게 나타났다. 구체적인 발달영역별로 비교해 보면, 의사소통 기술 및 일반 지식 영역에서 평균 차이가 가장 크게 나타났다. 사회적 능력, 언어 및 인지발달, 신체적 건강 및 웰빙, 정서적 성숙 순으로 이어졌다. 가장 점수 차이가 크게 나타난 의사소통 기술 및 일반지식 영역은 상대와 서로의 생각 및 요구에 대해 효과적으로 주고받거나, 다른 사람들이 의사소통하는 내용을 이해하며, 분명하게 말하는 것과 같은 능력과 관련되어 있다(Janus & Offord, 2007). 이러한 특성에 비추어 본다면 이 영역은 특히 아동의 말하기나 언어 문제의 영향을 받는 것을 알 수 있다.

어느 한 발달영역에 대한 아동의 어려움은 다른 발달영역에도 영향을 미친다는 점에 대해서도 생각해 볼 필요가 있다. 가령 아동의 신체적 손상(physical impairment)은 또래와 어울리는 데 어려움을 겪게 할 수도 있고, 학습에 필요한

2) 학교에 다니고 있지만 모두 평균연령이 6세 미만으로 우리의 유치원 연령에 해당하는 유아를 대상으로 함

운동기술을 획득하는 데 장애 요인으로 작용할 수도 있다. 아동이 행동문제를 가지고 있는 경우에도 그 문제가 사회적 관계를 형성하는 데 걸림돌로 작용할 수 있으며, 충동조절이나 주의집중과 같은 영역에서 자기조절과 관련된 문제를 야기할 수 있다.

2차 조사에서 획득한 자료를 토대로 아동이 가진 문제유형에 따라 분류하면 신체장애, 시각 손상, 청각 손상, 말하기 문제, 학습장애, 정서문제, 행동문제의 7개 범주로 구분할 수 있다. 이때 각각의 범주는 배타적이지 않은 것으로, 한 아동이 두 가지 이상의 문제를 가지는 경우도 포함되어 있으며, 아동이 특수교육을 받는지의 여부와는 연관하지 않았다.

한 명의 아동이 동시에 두 가지 이상의 문제를 가질 수 있다는 점을 감안하여 집단 간 발달영역별 EDI 점수 차이를 비교해 본 결과, 전반적으로 의사소통 및 일반 지식 영역이 아동의 문제에 대한 영향을 가장 많이 받는 경향이 관찰되었다. 특징적으로 정서 및 행동문제를 가진 아동의 집단에서는 사회적 능력에서 낮은 점수가 나타났다. 각 문제 유형 집단과 통제집단 간에 발달영역별 EDI 점수의 차이를 비교해 본 결과, 시각 및 청각의 손상과 같은 감각의 문제를 가진 아동 집단보다는 신체적 장애나 말하기 문제, 학습 장애, 정서 및 행동 문제 등의 비감각적인 문제를 가진 아동의 집단에서 더 큰 점수 차이가 확인되었다.

주목해야 할 점은 특수한 도움을 필요로 한다고 분류된 아동이 반드시 장애 진단을 받았음을 의미하지는 않는다는 것이다. 따라서 이 집단에 해당하는 많은 아동들이 여러 발달영역에서 어려움을 겪고 있음에도 불구하고 학교위원회로부터 적절한 혜택을 받지 못하는 상황에 놓여 있다.

Tymms 외(2004)는 수발달의 측정 중 세 자릿 수 에서 청각장애아 집단이 다른 집단보다 더 높은 점수를 얻은 것에 주의를 기울였다. 그 원인에 대하여 이들은 검사에 사용된 수화(British Sign Language, BSL)가 영어나, 화란어에 비하여 더 쉽기 때문으로 추정하였다. 그러나 청각장애아 집단은 읽기에서는 다른 집단에 비하여 매우 어려움을 겪은 것으로 나타났다. 또한 전반적인 검사 결과에 있어서도 전체 집단들 중에서 영어가 제2언어인 집단과 더불어 가장 낮은 점수를 기록하여 청각장애가 이들의 총체적인 학교준비도에 부정적인 영향을 미치고 있음을 밝혔다.

결론적으로, EDI나 PIPS를 활용한 비교 모두에서 학교에서 학습을 하는 데에 있어서 특수한 교육적 도움을 필요로 하거나 장애를 가진 아동들은 전형적으로

발달하는 또래와 비교할 때 전체 발달영역에서 낮게 평가되고 있음이 나타났다. 그러나 일련의 분석 결과는 특수아일지라도 자신이 가진 문제의 유형에 따라 학교준비도의 측면에서 다른 발달적 특성을 나타낸다는 사실을 확인시켜 주기도 하였다. EDI를 활용한 Janus 외(2011)의 연구와 Tymms 외(2004)의 연구는 이처럼 특수한 상태에 있는 아동의 빈도가 낮은 상황에서 모집단 수준의 자료를 다루었다는 측면에서 의의가 있다.

## 2) 이중/다언어 아동

Janus 외(2011)은 캐나다 아동들을 대상으로 모국어에 따라 하위집단을 구분하고 각 집단의 학교준비도 관련 특성을 살펴보았다. 역사적으로 볼 때 캐나다는 이민자들로 이루어진 국가라는 특수성을 가지고 있다. 초기 이민자들은 대부분 유럽 출신으로, 1970년대 초반에 새롭게 유입된 유럽인 이민자의 비율은 60%를 넘어서기도 했다. 그러나 1980년대에 들어서면서부터는 아시아와 중동 지역의 이민자가 증가하기 시작하면서, 현재는 전세계 각 대륙에서 이민 온 사람들이 캐나다 국민의 상당수를 이루고 있다. 이때 출신 국가에 따른 이민자의 구성 비율이나 이들이 캐나다에 유입된 시기와 같은 특성들은 이후 세대 자녀의 언어 환경과도 관련되어 있다.

또한 캐나다는 불어와 영어를 통용하는 나라로, 지역에 따라 모국어(제1언어)가 다를 수 있다. 여기서는 아동이 속한 학교(기관)에서 제1언어로 사용하는 언어를 아동의 모국어로 간주하였다. 연구자들은 EDI 자료 분석을 위해 아동의 언어적 특성에 따라 다음과 같은 3개의 집단을 구분하였다. 1) 영어를 사용하는 학교에서 모국어로 영어만 사용하는 아동과 불어를 사용하는 학교에서 불어만을 사용하는 아동은 '언어 통제집단(Language-control group)'으로 묶어서 분류하였다. 이는 모국어가 무엇인지를 묻는 질문에 대해 '다른 언어'라고 응답한 아동은 '다양한 언어(Diverse-language)'라고 응답한 아동과 별도로 구분되었다. 2) 다언어 범주에 속하면서 학교에서 쓰는 교수언어(language of instruction)를 제2외국어로 하는 아동은 '제2언어 학습자(Second Language Learners)'로 분류하였다. 3) 다양한 언어 범주에 속하지만 학교에서 사용하는 언어를 제2외국어로 하지 않는 아동은 '비 제2언어 학습자(Non-Second Language Learners)' 또는 '이중 언어 사용자(bilingual)'로 간주하였다. 특수아는 이러한 표본에서 제외되었다. 아동의 모국어와 관련된 분석에서 언어 통제집단이 차지하는 비율은 1차 자료

에서 74.3%, 2차 자료에서 75.2%였으며, 다양한 언어 범주에 속하는 아동의 비율은 1차 조사에서 20.9%, 2차 조사에서 22.3%에 해당했다.

다언어 집단과 언어 통제집단의 EDI 발달영역별 평균 점수를 비교해 본 결과, 의사소통기술 및 일반 지식 영역에서만 언어 통제집단의 점수가 다소 높았다. 다언어 집단 중 제2언어 학습자의 경우에는 모든 영역에서 언어 통제집단에 비해 낮은 점수를 보였는데, 그 차이는 의사소통기술 및 일반 지식 영역에서 가장 크게 나타났고, 언어 및 인지발달에서는 중간 수준의 차이가 확인되었다. 상대적으로 이중 언어를 사용하는 아동 집단은 의사소통기술 및 일반 지식을 제외한 모든 영역에서 통제집단보다 높은 점수를 나타냈으나, 그 차이는 통계적으로 근소한 수준에 머물렀다.

이 분석의 결과에 따르면 아동의 언어적 특성이 다르더라도 언어 및 인지발달 상태는 크게 차이가 나타나지 않았다. 즉 아동이 언어를 충분히 익히지 못한 채로 학교에 입학했다고 하더라도 학급에서 이루어지는 읽기 활동을 통해 나아질 수 있고, 글자나 숫자를 배울 수 있으며, 암기 기술이나 시간 개념과 관련된 능력을 발휘할 수 있다. 따라서 이런 아동의 경우에는 언어구사력을 다 갖춘 아동에 비해 스스로를 표현하는 데 있어서 유창성이 떨어질 수도 있지만, 이후 정상적으로 발달이 지속되면 앞서 있던 아동을 따라잡을 수 있는 잠재력을 가지고 있다.

반면, 의사소통 및 일반 지식 영역은 언어적으로 표현된 지시나 말의 내용을 온전히 이해하여 즉각적으로 반응하는 방식으로 언어를 사용해서 행동하는 아동의 능력을 반영한다. 더 나아가 이 영역은 아동이 입학 전에 학교에서 사용하는 언어와 문화를 경험한 양과도 관련이 있다. 이런 맥락에서 본다면 제2언어 학습자가 나타내는 발달에서의 격차가 의사소통 기술 및 일반 지식에서 더 두드러지는 이유를 이해할 수 있다.

추가적으로, 제2언어 학습자 및 이중 언어 사용자의 비율에 있어서 성차는 나타나지 않았다. 성별과 연령을 통제 했을 때에도 언어 통제집단과 제2언어 학습자, 이중 언어 사용자 간 EDI 점수 차이는 유의한 것으로 확인되었다.

Tymms 외(2004)의 연구에서는 앞서 언급한 청각장애아 집단과 함께 영국의 영어가 제2언어인 EAL(English as an Additional Language) 집단의 PIPS 점수가 가장 낮았음을 확인하였다. 그러나 연구자들은 이러한 결과에 대하여 이들의 의사소통의 문제에 기인한 것으로 이해하였다. 즉 이들의 검사가 영어로 진행되

었으며 특히 언어와 관련된 문항들 역시 이제 막 배우기 시작한 영어로 진행된 검사에서 이들이 더 낮은 점수를 얻은 것은 어쩌면 당연한 결과일 수 있다. 유사한 이유로 호주의 선주민 아동들 역시 전체적으로 호주의 다른 두 집단과 비교해서 전체 영역에서 낮은 점수를 얻었다.

Janus 외(2011)은 학교준비도의 측면에서 모국어로 사용하는 언어에 따른 아동발달 특성의 차이를 살펴보기 위해 EDI 조사 결과를 토대로 '비공식 모국어'로 가장 사용 빈도가 높은 상위 12개 언어를 선정하였다. 여기에는 편자브어, 스페인어, 광둥어, 만다린어(표준 중국어), 아랍어, 필리핀어(타갈로그어), 우르두어, 베트남어, 타밀어, 포르투갈어, 이탈리아어, 폴란드어가 해당되었다. 이러한 언어를 사용하는 아동의 비율은 1차 조사에서 7.7%, 2차 조사에서 10.3%에 이르렀다. 각 언어별 비교에는 2차 조사 자료만을 활용했기 때문에 여기에서 해당 언어 사용자의 비율이 0.4% 이상이 되는 경우에만 분석이 이루어졌다. 분석 대상 아동 중에는 어려서부터 두 개 이상의 언어(예: 표준 중국어와 광둥어, 포르투갈어와 스페인어 등)를 같이 배우는 경우 등이 포함되어 있기 때문에, 각 집단이 상호 배타적이지 않은 상태에서 제 2언어 학습자와 이중 언어 사용자를 구분하였다.

먼저, 다언어 집단 언어 통제집단의 EDI 점수를 비교했을 때에는 대부분의 경우에서 다언어 집단에 비해 언어 통제집단의 점수가 더 높게 나타났다. 그러나 다언어 집단을 제2언어 학습자와 이중 언어 사용자 집단으로 나누었을 때에는 대체로 이중 언어 사용자 집단의 점수가 높았고, 언어 통제집단의 점수가 그 뒤를 잇는 경향이 관찰되었다. 특히 이중 언어 사용자 집단과 제2언어 학습자 집단을 비교했을 때에는 전반적으로 이중 언어 사용자 집단에서 더 높은 점수를 얻었으며, 언어 및 인지발달 영역과 의사소통 및 일반 지식 영역에 있어서는 모든 언어 집단에서 이중 언어 사용자 집단에 속한 아동이 높은 발달적 성취를 보였다.

이때 주지할 점은, 모국어의 종류에 따라 일부 언어 사용자 집단에서 다른 양상의 결과가 나타났다는 사실이다. 특히 과반수의 언어 집단에서 EDI 발달영역 중 의사소통 및 일반 지식의 영역에서는 이중 언어 집단에 비해 언어 통제집단의 점수가 더 높게 나타났다. 따라서 비공식적 언어 사용자의 발달적 결과에 대해 일반화하기 보다는 각 언어집단별 문화적 신념이나 실제 등에 따라 아동의 발달에 다른 영향을 미칠 수 있음에 유념해야 한다.

연구자들은 아동의 언어 발달 특성과 관련된 발달적 결과에 부모가 취학 전 자녀에 대해 가지는 기대나 신념, 학업을 수행하기 이전 단계에 이루어진 아동의 경험, 새로운 문화나 분위기, 일상에 대한 적응 과정 등이 매우 중요한 역할을 한다고 지적하였다.

## 2. 유아발달 차이의 요인

Janus와 Offord(2007)의 연구에서 가장 중심적으로 다룬 주제는 초등학교에 입학하는 시점에서 아동의 학교준비도에 차이를 발생시키는 위험요인들을 탐색하는 것이다. 이를 위해 앞에서 제시한 방식으로 대상 아동 및 부모와 관련된 특성들을 조사하여 EDI로 평가한 학교준비도와 관련된성을 분석하고, 이를 통해 산출된 결과가 가지는 의미를 밝히고자 하였다. 이러한 맥락에서 연구결과를 살펴보면, 아동의 학교준비도를 통해 확인된 취약성은 아동의 건강상태 및 성별과 가장 밀접하게 관련되어 있으며, 이러한 격차는 저소득 가정에서 더 두드러졌다. 가족구조와 아동의 연령, 부모 참여, 부모의 건강 등도 아동의 취약성과 관련이 있는 것으로 나타났으나, 앞에 제시한 요인들에 비해서는 영향력이 낮은 편이다.

### 가. 아동 변인

#### 1) 아동의 건강상태

먼저 아동의 건강상태에 대해 살펴보면, 건강 상태가 양호한 아동에 비해 그렇지 않은 아동이 학교준비도에서 더 취약한 경향이 나타났다. 특히 출생 시 저체중이었던지의 여부나 조사 시점에서 평가한 아동의 건강상태보다는 Health Utilities Index(HUI; Horseman, Furlong, Feeny, & Torrance, 2003)으로 측정된 기능적 측면에서의 건강상태가 학교준비도와 더 밀접한 관련이 있는 것으로 확인되었다. 특수한 도움이 필요하다고 판단되는 아동이 취학 시점에서 어려움을 겪을 수 있다는 것에 대해서는 일반적으로 잘 알고 있다. 그러나 건강상태의 하위요인(suboptimal health status)이 아동의 학교준비도에 어떤 영향을 미치는지는 충분히 고려되지 못하므로, 이러한 결과를 토대로 아동의 건강상태가 보다

신중하게 다루어질 필요가 있다.

## 2) 아동의 성별

캐나다 아동을 대상으로 한 EDI연구에서 여아에 비해 남아의 취약성이 2.3배 더 높게 나타났다. 성별은 환경적인 요인이 아니라 인구학적 특성이기 때문에, 성차를 유발하는 기제에 초점을 맞춘 중재가 필요하다. 뇌의 발달 과정에서부터 성차가 나타나기 시작하지만, 이렇게 생물학적 요인에서 기인한 차이 중의 일부에 대해서는 성인의 태도와 반응이 그 격차를 확대시키거나 유지하는 역할을 한다. 가령, 언어 영역의 좌뇌 편재화는 여아의 경우가 더 일찍 이루어지기도 하지만(Bornstein, Hahn, & Haynes, 2004), 어머니들이 아들보다는 딸과 더 많은 상호작용을 하는 경향(Clearfield & Nelson, 2006)이 초기 언어발달에서의 성차를 가중시킨다는 것이다. 이러한 양상은 사회정서적 측면의 발달에서도 볼 수 있다. 이처럼 생물학적 특성으로 인한 성차가 부모의 행동으로 인해 시간이 지나면서도 줄곧 이어지며(Huttenlocher, Bryk, Seltzer, & Lyons, 1991), 초등학교 입학 시기부터 그 이후까지의 아동 능력의 차이를 유발한다. 교육현장에서 학업이 이루어지는 기존의 방식이 구조화된 학습활동에 더 잘 따르는 여아에게 맞추어져 있다는 인식이 확산되면서, 남아에게 효과적인 교수 방법을 개발하고 촉진해야 할 필요성이 대두되고 있다(Chaplin, Cole, & Zahn-Waxler, 2005; Sink & Spencer, 2005; Zill & West, 2001). 중요한 점은 아동의 학습과 발달에 있어서 성차가 있다는 사실이 확인되는 만큼, 이를 줄이기 위한 노력이 이루어져야 할 것이다.

반면, 일부 연구에서는 유아의 성차가 일정한 경향을 보이지는 않는 것으로 보고하고 있기도 하며 학령기 아이들에 대한 연구에서도 성차는 영역에 따라 각기 다른 결과를 내기도 한다(Csapó, Molnár, & Nagy, 2014).

## 3) 아동의 월령

초등학교 입학 시 아동의 학교준비도와 관련이 있는 또 하나의 변인으로 아동의 월령이 확인되었다. 1998년에 미국에서 대규모 표본을 대상으로 이루어진 조사에서도 입학 시 아동의 나이가 6세에 가까울수록 언어, 수학, 사회적 능력, 추론 기술과 같은 모든 영역에서 더 좋은 능력과 기술을 발휘하는 것으로 보고된 바 있다(Zill & West 2001). 그러나 같은 학년 내에서 아동의 월령이 더 높

다는 것이 어떤 효과를 발생시키는지에 대해서는 분분한 결과가 있었다(예: Byrd, Weitzman, & Auiger, 1997; Meisels, 1992). 그리고 이 시기에 나타나는 차이가 고등학교 때나 대학 입학 시기에까지 이어질 수 있다는 결과가 있는가 하면(Bedard & Dhuey, 2006), 초등학교 입학 후 수 년 안에 차이를 좁힐 수 있다고 보고한 연구도 있었다(Stipek & Byler, 2001). Tymms 외(2004)는 학교 시작 연령을 중요한 변인으로 보고 있는데 이는 국가 간의 비교에 있어서 유아 발달의 차이의 원인으로 종종 지적되기도 한다. 한편, Bedard와 Dhuey(2006)에 의하면 미국 아동의 5%가 의도적으로 유치원에 들어가는 시기를 한 해 늦추는데, 이는 특히 사회경제적 지위가 높은 계층에 속한 어린 아동과 관련이 있는 것으로 나타났다. 캐나다의 아동에 대해서는 이 같은 양상이 확인되지 않았지만, 아동의 나이는 초등학교 입학 시 학교준비도의 차이와 이후 학교생활의 성패에 대해 지속적으로 영향을 미칠 수 있다는 점에 대해서는 추론이 가능하다.

## 나. 가정 변인

### 1) 가계소득

가정의 가계소득은 이 시기 아동의 취약성과 매우 밀접하게 관련된 요인으로 확인되었다. 반면 부모의 학력 수준과의 관련성은 이에 한참 못 미치는 것으로 나타났다. 이러한 결과와는 달리 Phillips와 Shonkoff(2000)는 생애 초기 아동의 성취에 대한 영향력을 비교했을 때 가계소득보다는 어머니의 학력이 더 큰 비중을 차지하는 것으로 보고하며, 이에 대한 세 가지 근거를 제시하였다. 첫 번째는 SES가 그 자체로써 영향을 미치기보다, 부모의 상호작용이라든가 보육의 질이나 영양과 같은 자원들을 통해 아동의 발달에 관여한다는 것이며, 두 번째는 학력 수준이 높은 여성들도 출산 직후 몇 년 간은 자신의 교육 수준에 상응하는 소득을 올리기 어렵다는 것이다. 그리고 세 번째는 어머니의 학력이 높을수록 소득수준과 관계없이 보다 적절한 양육 전략 및 양육 행동을 취하고, 더 나은 영양 공급을 이루는 경향이 있다는 것이다.

Janus 외(2007)의 연구는 캐나다 가정을 대상으로 이 같은 관련성을 광범위하게 조사한 최초의 연구이기 때문에, 다른 국가에서 이루어진 선행연구 결과와 단적으로 비교하기에는 조심스러운 부분이 있다. 캐나다에서의 교육과 소득 수준 간의 불균형이 충분히 크지 않을 수 있으며, 각 변인을 다루는 방식이나 아

동의 발달적 결과를 측정하는 방식에서 차이가 있기 때문일 수도 있다. 이를 보면, 어머니의 학력이 아동의 지식 및 기술에 강한 영향력을 발휘하는 것으로 확인된 West와 Rose(2000)의 연구에서는 소득을 별도의 변인으로 다루는 대신, 위험 요인을 구성하는 네 가지 요인들 중의 일부로 상정하였다. 또한 앞서 제시한 세 가지 근거 가운데 두 번째에 해당하는 내용은 캐나다의 상황에 적용되지 않을 가능성도 제시되었다. 이러한 내용들을 감안한다면 아동의 학교준비도에 대한 가계 소득의 영향력에 대해 재고할 필요가 있다.

## 2) 가족 구성 형태

아동의 학교준비도와 관련된 또 다른 가족 관련 요소로 가족 구성의 형태가 지적되었다. EDI의 발달영역 중 의사소통 기술 영역을 제외한 4개의 발달 영역에서 가족 구조는 가계소득 수준보다 더 중요한 역할을 하는 변인으로 확인되었다. 아동이 속한 가정이 한부모 가족의 형태를 띠고 있는지의 여부는 EDI의 모든 발달영역과 관련이 있었다. 그러나 아동의 취약성에 초점을 맞추었을 때 더 강한 연관성을 나타내는 것은 온전한 가족(intact family)과 함께 생활하는지의 여부였다. 즉 아동의 출생 시점부터 조사가 이루어진 시점까지 생물학적 부모와 함께 지냈다는 사실이 취약 시 아동의 취약성을 예방할 수 있는 강력한 보호요인으로 작용할 수 있음을 시사한다.

## 3) 부모 참여

부모 참여와 관련하여 Janus 외(2007)는 두 개의 차원에서 접근하였다. 하나는 부모가 자녀와 함께 책을 읽는 빈도(frequency of reading with child)이고, 다른 하나는 아동이 책이나 잡지를 보는 빈도(frequency of the child looking at book and magazines)였다. 여기서의 빈도는 매일 일상적으로 이루어지는 경우와 한 주 동안 몇 차례, 혹은 그 이하로 이루어지는 경우를 이분법적으로 구분한 것이었다. 분석 결과, 매일 부모와 함께 책을 읽지 않는 것보다는 아동 스스로 책이나 잡지를 자주 보지 않는 것이 취약성과 더 관련되어 있는 것으로 나타났다. 또한 아동이 책이나 잡지를 보는 빈도만을 가지고 비교했을 때에는 일상적으로 책을 보는 아동에 비해 그렇지 않은 아동의 취약성이 1.3배가량 더 높은 것으로 검증되었다. 아동의 발달 특성에 대한 조사와 책을 읽는 빈도에 대한 조사가 같은 시기를 기준으로(횡단적으로) 이루어졌기 때문에, 인과관계를 추론

할 수 있는 민감성은 떨어질 수 있다.

이와 관련하여 유아를 대상으로 이루어진 Raikes 외(2006)의 종단연구에 따르면, 성취 결과를 조사한 시점과 같은 시기의 독서량보다는 이전(2세 때)의 책읽기가 더욱 강한 예측요인으로 작용한다. 이 같은 사실을 감안하면, 단지 부모와 함께 책을 읽는다는 사실보다는 아동이 스스로 활자에 다가가기까지의 능력과 의지가 학교준비도를 갖추는 데 대해 훨씬 분명한 지표로 작용했을 가능성이 있다. 즉 5세 무렵의 아동들에게 향상되는 독립성이 다른 어떤 요소보다 아동이 책에 다가갈 것인지를 선택하는 데 영향을 미칠 수 있는데, 그것이 취학 시점의 학교준비도 차이로 이어지게 된다는 것이다. 만일 이러한 기제가 성립한다면, 어렸을 때부터 아동에게 문해 자료에 대한 접근성과 가용성이 확보되는 것이 무엇보다 중요하다고 볼 수 있다.

#### 4) 부모 건강 및 흡연

어머니의 우울은 유아의 가정에서의 행동문제와 깊은 관련이 있다. 이는 실제 어머니의 우울이 직접적으로 유아의 발달에 영향을 미쳤을 수도 있지만 또한 어머니가 자신의 유아의 발달을 직접 평가한 경우 자신의 자녀의 행동을 특별히 더 문제 행동으로 평가했을 가능성도 있다는 점을 인식할 필요가 있다 (McWayne, Cheung, Wright, & Hahs-Vaughn, 2012).

한편, 부모의 건강과 관련된 지표 중 흡연은 아동의 발달에서 취약성이 나타날 가능성을 1.3배까지 높이는 요인으로 작용한다는 사실이 확인되었다. 흡연이 낮은 학력이나 소득수준과 관련된 요인으로 다루어진 적은 있으나, 여기서는 소득과 별개의 요인으로 부모의 흡연이 자녀의 발달에 부정적인 영향을 초래할 수 있음을 추론할 수 있는 근거를 제시하였다. 이 연구에서는 부모의 흡연과 자녀의 발달을 같은 시기에 조사하여 이들 간의 관련성을 밝혔으나, 태내기의 아동이 부모의 흡연으로 인해 영향을 받게 되는 뇌신경행동학적결과일 가능성에 대해서도 고려할 필요가 있다(Napiorkowski, et al. 1996).

이 연구는 횡단적 자료를 사용했기 때문에 변인들 간의 인과관계를 밝히는 데에는 제약이 있었다. 취약성에 대해서 이분법적 변수를 사용했다는 점에서도 한계가 지적되었다. 그러나 캐나다 아동을 대상으로 EDI를 통해 측정된 아동의 학교준비도가 가정의 사회경제적 지위나 인구학적 특성, 가족 관련 요인에 민감하다는 사실을 확인시켜 주었다는 점에서 의의를 찾을 수 있다.

### 3. 발달영역별, 성별, 지역별 수준 차이

EDI로만 평정한 집단에 비해 보강된 표본에는 여아의 비중이 높았으며, 월령이 근소하지만 유의한 수준에서 낮은 것으로 나타났다. 성별과 나이를 통제한 상태에서 EDI 발달영역별 점수를 비교했을 때에는 정서적 성숙을 제외한 네 영역에서 보강된 표본의 아동이 더 높은 점수를 나타냈다. 정서적 성숙 영역에 대해서는 두 집단 간에 유의한 차이가 나타나지 않았다. 다른 국제비교 검사도구인 EAP- ECDS의 차이는 각 발달영역 별로 제시하였다<sup>3)</sup>.

#### 1) 인지 발달의 차이

캄보디아의 아동들은 도시의 아동들이 지방의 아동들보다 월등히 더 나은 인지발달 성과를 보였다. 중국의 아동들도 동일하게 도시의 아동들이 지방의 아동들보다 모든 연령에서 월등히 더 나은 성과를 보였으며, 그 차이는 4세 때 가장 크고, 5세 때 가장 작은 것으로 나타났다. 이 원인은 조기교육의 실시에 따라서 차이가 줄어든 것으로 추측된다.

몽골은 모든 아동들이 나이가 들어감에 따라 인지능력이 증가하는 것으로 나타났다. 또한 여자아동들의 성장이 남자아동들보다 더 높은 것으로 나타났지만, 통계적으로 유의한 차이를 나타내지는 않았다. 그리고 5세의 도시 아동들이 같은 나이의 지방 아동들보다 인지능력이 뛰어난 것으로 나타났지만, 3세나 4세에서의 도시와 지방의 차이는 유의미하지 않은 것으로 나타났다.

파푸아뉴기니의 아동들은 남자 아동들이 여자 아동들 보다 인지능력이 높은 것으로 나타났고, 지방의 아동들이 도시의 아동들보다 인지발달이 더 높은 것으로 나타났다. 이는 다른 국가들과의 큰 차이점이라고 할 수 있다.

동티모르 아동들은 도시 아동들이 지방의 아동들 보다 높은 것으로 나타났지만, 바누아투의 경우, 지방의 아동들이 도시의 아동들보다 인지발달이 더 높은 것으로 나타났다.

#### 2) 사회정서적 발달의 차이

캄보디아 아동들의 사회정서적인 발달 측면에서는 도시의 여자 아동들이 남

3) EAP-ECDS의 지역차이는 ARNEC에서 보고하는 국가별 연구보고서를 요약한 내용임 (www.arnecc.net에서 2015년 6. 15일자 인출)

자 아동들보다 더 큰 발달을 보였지만, 지방에서의 남녀 차이는 나타나지 않았다. 중국의 경우도 여자 아동들이 남자 아동들보다 더 큰 발달을 보였고, 도시의 아동들이 지방의 아동보다 더 큰 발달정도를 가지는 것으로 나타났다.

몽골의 경우, 여자 아동들이 남자 아동들보다 더 큰 발달을 보였고, 5세의 도시 아동들이 같은 나이의 지방 아동들보다 사회정서 발달이 뛰어난 것으로 나타났다지만, 3세나 4세에서의 도시와 지방의 차이는 유의미하지 않은 것으로 나타났다.

파푸아뉴기니의 아동들은 지방의 아동들이 도시의 아동들보다 더 높은 것으로 나타났다지만, 동티모르 아동들은 연령에 따른 차이가 나타나지 않았고, 바누아투의 경우, 여자 아동이 남자 아동보다 더 높은 것으로 나타났다.

### 3) 신체발달의 차이

캄보디아 아동들은 신체발달 측면에서 지방의 아이들 보다 도시의 아이들이 훨씬 발달정도가 크게 나타났으며, 5세의 남자 아동들이 5세의 여자 아동들보다 더 큰 발달 정도를 나타냈다. 그러나 3~4세에서는 성별에 따른 큰 차이를 보이지 않았다.

중국과 몽골, 파푸아뉴기니, 바누아투의 아동들은 지방의 아이들이 도시의 아이들보다 훨씬 발달정도가 크게 나타났으며, 동티모르 아동들은 도시 아동들이 지방의 아동들 보다 높은 것으로 나타났다.

### 4) 언어적 발달의 차이

캄보디아 아동들의 언어적 발달 측면에서는 여자 아동들이 남자 아동들보다 뛰어났으며, 도시의 아동들이 지방의 아동들보다 모든 연령대에서 더 나은 언어와 발생적 문해 능력을 가지고 있었고, 나이가 들어감에 따라 그 차이는 증가하는 것으로 확인되었다. 중국과 몽골의 경우도, 여자 아동들이 남자 아동들보다 뛰어났으며, 도시의 아동들이 지방의 아동들보다 더 나은 언어와 발생적 문해 능력을 가지고 있는 것으로 확인되었다.

파푸아뉴기니의 아동들의 연령에 따른 특별한 차이를 보이지 않았으며, 동티모르 아동들은 도시의 아동들이 지방의 아동들보다 월등히 높게 나타났으며, 바누아투의 경우, 지방의 여자 아동이 지방의 남자 아동보다 높은 수준을 가지고 있었지만 도시에서의 성별차이는 나타나지 않았다.

## 5) 건강과 위생 및 안전의 차이

캄보디아 아동들의 건강과 위생, 안전한 측면에서는 도시의 아동들이 지방의 아동들보다 모든 연령대에서 더 나은 건강과 위생, 안전 상태를 가지고 있는 것으로 나타났으며, 나이가 들어감에 따라 그 차이는 증가하는 것으로 확인되었다. 중국의 경우, 건강과 위생, 안전한 측면에서도 여자 아동들이 남자 아동들보다 뛰어났으며, 캄보디아와 마찬가지로 도시의 아동들이 지방의 아동들보다 더 나은 건강과 위생, 안전 상태를 가지고 있는 것으로 나타났다.

몽골의 아동들은 여자 아동들이 남자 아동들보다 뛰어났으며, 도시의 아동들이 지방의 아동들보다 더 나은 건강과 위생, 안전 상태를 가지고 있는 것으로 나타났지만, 파푸아뉴기니 아동들은 지방의 아동들이 도시의 아동들보다 더 나은 건강과 위생, 안전 상태를 가지고 있는 것으로 나타났다.

동티모르 아동들의 건강과 위생, 안전한 측면에서는 3~4세의 아동들에게서는 여자 아동이 남자 아동보다 높게, 5세의 아동에서는 반대의 패턴으로 나타났으며, 바누아투의 아동들은 지방의 남자 아동이 도시의 남자 아동보다 더 나은 건강과 위생, 안전 상태를 가지고 있는 것으로 나타났다.

## 6) 문화지식과 참여정도의 차이

캄보디아 아동들의 문화지식과 참여정도는 여자 아동들이 남자 아동들보다 높게 나타났고 도시 아동들이 지방아동들보다 더 나은 문화지식과 참여정도를 가지는 것으로 나타났다. 중국의 경우도 도시 아동들이 지방아동들보다 더 나은 문화지식과 참여정도를 가지는 것으로 나타났다.

몽골의 아동들은 여자 아동들이 남자 아동들보다 문화지식 참여정도가 높았으며, 5세 아동들에게서는 도시 아동들이 지방아동들보다 더 나은 문화지식과 참여정도를 가지는 것으로 나타났지만, 3~4세의 아동들에게는 차이가 없는 것으로 확인되었다.

파푸아뉴기니와 동티모르의 아동들은 문화지식과 참여정도에서 지방의 아동들이 도시의 아동들보다 훨씬 더 나은 문화지식과 참여정도를 가지는 것으로 나타났고, 비누아트는 도시-지방간 차이와 성별차이가 보이지 않았다.

## 7) 학습 방식의 차이

캄보디아 아동들의 학습 방식은 도시 아동들이 지방 아동들보다 더 많은 기회를 가지는 것으로 나타났다. 중국의 아동들은 여자 아동들이 남자 아동들보다 뛰어났으며, 캄보디아 아동들과 비슷하게 도시의 아동들이 지방의 아동들보다 더 나은 학습방식을 가지고 있는 것으로 나타났다.

몽골 아동의 학습방식은 3~4세에서는 성별에 따른 차이가 없는 것으로 나타났다지만, 5세에서는 여자 아동들이 남자 아동들보다 뛰어났으며, 도시의 아동들이 지방의 아동들보다 더 학습 방식이 더 나은 것으로 나타났다.

과푸아뉴기니와 동티모르 아동들은 학습에 대한 접근에서만 도시의 아동들이 지방의 아동들보다 더 접근방식이 우수했으나, 바누아투에서는 도시-지방간 차이와 성별차이가 보이지 않았다.

## V. 유아발달 국제비교에 대한 전문가 의견 심층분석

이 장에서는 유아발달에 대한 국제비교의 목적, 발달평가 영역, 핵심 지표 및 평가 방법과 절차, 평가도구 타당화 과정, 결과해석 및 정책적 활용 등에 대한 전문가의 의견을 파악하기 위한 조사 결과를 분석하였다. 또한 전문가 심층분석을 통하여 의견 조사 결과를 구체화시켰으며 평가 목적별로 시나리오 분석을 실시하였다.

### 1. 조사 개요

#### 가. 조사 목적

본 전문가 의견조사의 목적은 국가 간의 유아발달을 비교하는 목적과 더불어 평가에 적합한 연령대상 선정, 평가영역 및 필수 하위문항, 특히, 기존의 국제비교연구에서 우리나라 유아에게 실시하기 부적합한 문항 및 우리나라 유아발달 평가의 고유한 문항 규명, 평가 방법과 절차, 평가 시기, 평가결과 활용방안 등에 대한 견해와 수요를 파악하는데 있다.

#### 나. 조사 참여자 특성

설문조사의 참여자는 유아발달 평가 전문가 45명이며, 해당 전문가들은 교육학 및 유아교육학 전공자들로 국가연구기관에서 과거 국제비교 연구를 진행하였거나, 현재 국제비교연구를 진행하고 있는 연구책임자로 대상을 한정하였다.

#### 다. 조사 내용 및 방법

전문가 의견조사는 총 2회 실시되었으며, 1차 조사에서는 총 11개의 개방형 문항(open questions)으로 구성된 의견조사지를 바탕으로 진행되었으며 연구진은 전문가 3인과의 자문회의를 통하여 의견조사를 검토하였다. 문항 내용은 유

아발달 수준에 대한 국제비교의 필요성과 목적, 대상연령 및 선정 이유, 비교대상 국가, 평가 영역, 활용가능한 평가도구 여부, 평가도구 개발 시 우선적으로 고려해야 할 사항, 타당화 고려 사항, 국제 결과 활용방안 등으로 구성하였다. 아래의 문항 내용을 바탕으로 전문가 대상 온라인 조사 및 협의회 개최를 통하여 의견을 수렴하였다.

<표 V-1-1> 전문가 의견조사 문항 목록(1차)

번호	문항내용
1	유아발달 수준 국제 비교의 필요성 및 목적
2	유아발달 수준에 대한 국제비교 연구 수행 시 특별한 관심 대상 집단 (예: 저소득층 남아, 거주지역별, 유아교육과 보육기관 미이용 아동, 성별 등) 및 그 이유
3	유아발달 수준에 대한 국제비교 연구 수행시 적합한 연령 및 그 이유
4	유아발달 국제비교 적합한 대상국가
5	국제비교 유아발달 평가도구를 개발 시 포함되어야 할 평가영역(및 세부 영역)들과 그 이유(예: ① 신체, 언어, 인지, 사회정서 영역 등 ② 각 영역별로 반드시 포함되어야 할 핵심적인 세부 하위 영역 3~4가지)
6	기존에 개발된 유아발달 평가도구 중 국제비교를 위해 활용 가능한 도구 추천
7	국제비교를 위한 유아발달 평가도구를 개발함에 있어서 우선적으로 고려해야 할 사항.
8	유아발달 평가도구에 대한 국제수준의 타당화 과정에서 고려해야 할 사항
9	유아발달 수준에 대한 국제비교 결과를 국가정책의 성과평가 등에 활용하기 위하여 고려하여야 할 점(특정한 정책의 성과 모니터링, 학교준비도 점검 등에 활용시)
10	국제비교를 위한 유아발달 평가도구 개발에 대한 추가적인 제안 사항 (예: 아시아권 또는 우리나라 유아발달의 특성 반영, 유아 발달 국제비교를 위한 국가적 차원의 지원 등)
11	관련 참고문헌 및 웹사이트 링크

2차 조사는 1차 조사에서 수렴된 유아발달 평가를 국제비교하는 목적에 따라서 각기 다른 수준의 문항들을 제시하여 각 목적에 따른 유아발달평가 개발 시 나리오를 분석하였다. 총 12문항으로 구성되며 1 ~ 10번은 평가의 목적에 따라서 응답하게 되고, 11 ~ 12번 문항은 하위 평가영역의 중요성에 대한 인식과 평가도구 개발 시 중요하게 고려해야 할 사항에 대해서 전문가 의견을 조사하였다.

<표 V-1-2> 국내·외 전문가 조사 문항 목록(2차)

목적	국제적 유아의 발달과 학습 성과 비교	국제적 수준에서 우리나라 유아 발달 수준의 현위치 점검	조기중재 방안 마련	개발도상국과의 교류협력을 위한 정보 제공 및 지원
문항1	목적의 우선순위			
문항2	적합한 연령대			
문항3	비교 대상국가			
문항4	특정 관심 집단 설정			
문항5	필요한 표본 수			
문항6	국가별 문항 수정 가능 비율			
문항7	적정 문항 수(총검사 문항 수)			
문항8	적합한 평가자			
문항9	적절한 평가 시행 횟수 및 시기			
문항10	평가 영역			
문항11	공통: 포함되어야 할 평가 영역의 중요성			
문항12	공통: 국제비교 유아발달 평가도구를 개발 시 고려사항			

## 2. 1차 조사 결과 분석

### 가. 유아발달 국제비교의 필요성 및 목적에 대한 인식

본 연구의 목적과 연관성이 많은 필요성 및 목적에 대한 인식에서는 해당 전문가들의 전공영역에 따라서 유아발달 국제비교의 필요성 및 목적에 대한 인식이 상이하게 나타났다. 우선 국제비교에 대한 찬성과 반대의 입장으로 나누어졌다. 주로 찬성의 입장에서는 교육과정 성과를 비교분석함으로써 자국 교육의 현상 상황과 교육개선을 위한 유용한 자료로 활용될 수 있다. 유아기의 발달요소는 이후 개인이 성장하는데 중요한 영향을 미치기에 우리나라 유아들이 타 국가 유아들과 비교해 건강한 발달 패턴을 보이고 있는지를 확인하는 것이 중요한 의미를 가지고 있다고 주장하였다.

또한 21세기 글로벌 지식기반사회에서 교육 관련 주제 및 이슈는 더 이상 국가수준에 국한된 것이 아니라 국제사회의 의제이며, 지역 또는 국가차원의 연구와 학교 현장에 대한 이해뿐만 아니라 국제차원에서의 비교가 중요한 고려대상이 되고 있다는 점을 제시하였고, OECD가 진행하는 PISA, TALIS, PISAC 등

주요한 국제비교연구에 이미 참여하여 교육적 수요를 정책적으로 반영할 수 있는 계기를 마련했기 때문에 유아발달수준의 국제비교 또한 사회적 수요에 의한 유아교육정책이 아닌 교육적 수요에 대한 유아교육정책수립에 기여할 수 있다는 점에서 찬성한다는 입장을 표명하였다.

한편, 반대의 입장에서는 유아의 발달정도가 과연 국제비교를 할 만한 준거가 되는지에 대한 평가이다. 유아기의 발달은 개인차가 크고 발달 패턴 또한 상이하기 때문에 이를 하나의 기준으로 평가하기가 매우 모호하다는 점을 지적하였다. 또한 유아발달의 정도를 국제적으로 비교하는 것은 각국의 교육과 보육의 상황과 학교교육의 정도, 정책의 반영 정도가 다르기 때문에, 국제비교를 통하여 어떠한 교육적 의미를 도출시킬 것인가에 대한 의문점이 제기된다. 반대의 입장에서 또 하나의 우려는 결과의 활용이다. 국제비교를 통한 연구들에서 순위를 제공하는 것은 가장 지양해야할 점인데, 유아기의 발달정도에 대하여 순위를 제공한다면 이는 큰 사회적 위험요소가 될 수 있다는 점이 제기되었다.

## 나. 유아발달 국제비교연구 대상

유아발달 수준에 대한 국제비교 연구 대상에 대한 질문은 크게 대상 집단과 대상 국가로 나누어서 조사하였다. 대상 집단에는 대상 연령과 함께, 특별히 관심을 두어야 하는 대상을 질문하였으며, 대상 국가에서는 비교가 가능한 대상 국가가 어떠한 나라가 있는지를 조사하였다.

### 1) 대상 집단

대상 집단 중 먼저 연령에 대해서는 유아교육의 마지막 단계인 만 6세가 가장 많은 의견이 모아졌다. 이밖에도 상당수의 전문가들이 만3세도 대상 집단으로 지목하였다. 만 6세를 선택한 이유로는 대부분이 유아교육의 마지막 단계이면서, 초등학교 입학직전이므로 유아교육의 총체적인 측면을 고려할 수 있다는 점에서 중요한 연령대인 것으로 나타났으며, 만 3세의 경우는 인지/정서/사회성/행동 측면에서 유의미한 결과들을 관찰할 수 있으며, 누리과정이 시작되는 시점이라는 점에서도 적합하다는 의견이 있었다.

한편, 구체적인 연령을 제시하지 않은 이유에 대해서는 국가별로 학교교육 시스템이 매우 상이하기 때문에 공통적으로 초등학교 입학 이전 시기와 이후를

구분하여 비교대상 국가 모두에서 공통적으로 나타나는 입학이전 시기의 연령을 선택하는 것이 필요해서 구체적인 연령을 대상으로 삼기는 매우 어려울 것이라는 의견과 평가요소(언어, 사회성, 신체활동 등)와 평가방법(관찰, 자기보고, 부모, 교사 보고 등)에 따라서 연령 적합성 기준은 달라질 수밖에 없다는 의견이 제시되었다.

특별한 관심 집단을 묻는 질문에서 대해서는 전문가 각각의 의견들이 다양하게 나타났다. 국제비교의 일차적인 목적을 유아발달에 필요한 최소한의 지원, 그리고 이를 통한 최소한의 형평성 확보에 둔다면 저소득층, 농어촌, 대도시 저소득층의 유아들을 관심 범위에 두어야 하며, 보육·교육기관 미이용 유아 또한 포함되어야 한다고 조사에서 나타났다.

또한 현실적으로 누리과정의 정책적인 효과를 보기 위해서는 반드시 교육·보육기관 미이용자가 비교 대상이 되어야 하지만 현실적으로 대부분의 유아들이 기관을 이용하고 있기 때문에 비교하기 어려우며, 대상조차 많은 수를 차지하고 있지 않아 기본적으로 전체를 대표할 수 있는 표본을 구성한 뒤, 중요 소집단을 하위집단으로 구별하여 추적·비교 하는 것이 좋은 방법이 될 것이라는 제안이 제시되었다.

마지막으로 성별차이(gender difference)에 대한 집단을 살펴보는 것이 제안되었다. 성별차이는 OECD의 학령기 국제비교 연구에서 중요하게 다루는 주제 중 하나이며, 성별 차이 극복 방안이 세계의 주요 연구 흐름 중 하나이기 때문에 국내 유아발달 수준에서도 반드시 성별차이에 대한 대상 집단을 나누어 확인하는 것이 필요하다는 의견이 나왔다.

## 2) 대상 국가

대상 국가에 대한 질문에 대해서는 유아 발달에 대한 많은 연구와 정책이 이루어지고 선진적인 유아관련정책을 추진 중인 곳으로, 우리가 벤치마킹 할 수 있는 국가이면 좋겠다는 의견이 많았다. 예를 들어 영국과 미국과 같이 선진국을 중심으로 범위를 한정하거나, 프랑스나 북유럽 국가와 같이 선진적인 육아관련 정책을 추진 중인 국가이면 좋겠다는 의견이었다.

또 다른 의견으로는 가능한 우리와 유사한 사회적 특성을 가지거나 유사한 사회적 정책이 중시되는 국가가 포함되면 좋겠다는 의견이 있었다. 예를 들어 대만, 싱가포르, 그리스, 핀란드, 스페인을 추천하였다.

## 다. 유아발달 평가영역 및 세부영역

유아발달 평가영역과 세부영역에 대한 조사에서는 참여국가 간 공통적으로 측정 및 보고가 가능한 세부영역을 선정하는 작업이 우선적으로 필요하며, 이러한 단계는 평가도구를 만들기 이전에 참여하는 국가들의 합의로 우선되어야 한다는 점을 조사대상자 대부분이 제시하고 있었다. 응답자들의 의견을 종합하면 각 영역에는 다음과 같은 세부영역들을 제시하였다.

<표 V-2-1> 유아발달 평가영역 및 세부영역

평가영역	세부영역
(공통의견)	
신체	지각, 소근육, 대근육, 이동, 기본신체조건/발달, 기본운동능력, 기초 위생습관, 식사횟수 및 정도, 간식의 종류와 정도, 편식
인지	수감각, 분류, 기억, 추론, 창의성, 기억력, 셈하기, 공간인지능력, 기초문제해결력, 이해력, 사고력, 학습준비도
언어	듣기, 말하기, 쓰기, 의사소통, 음성언어듣기, 음성언어말하기, 음성언어상호작용, 문자 언어에 대한 기초인식, 어휘력
사회정서	개인의 이해, 집단활동 참여하기, 상호작용, 배려와 협력, 또래관계(소통), 부모관계(소통), 자기개념, 적응행동, 문화적 다양성/상호존중, 기초시민의식(질서의식, 규범준수 등), 또래인지(관계), 감정조절, 친구와 보내는 시간
(개별의견)	
표현	신체표현, 음악표현, 미술표현, 종합표현
학습정서	동기, 태도, 자아존중감
동기	과제에 대한 흥미, 관심, 취미, 여가활동
정의(심리사회)	친구관계, 책임감, 도덕, 질서, 규칙지키기 등 기본적인 예의범절

## 라. 국제비교를 위한 유아발달 평가도구 개발 관련

### 1) 평가도구 개발을 위한 고려사항

유아발달 평가도구를 개발함에 있어서 고려사항에 대한 응답내용을 종합해보면 가장 높은 순위로 나온 것은 정책적 활용도였다. 정책적 활용도를 최우선 순위로 선택한 이유에 대해서는 많은 예산과 노력이 들어가는 평가의 시행이

정책적인 시사점이 없다면 연구차원의 시행이 나올 것이라는 의견과 국제비교를 위한 자료 수집이 어떤 목적으로 사용되어야 하는가에 대한 분명한 목적의식이 가장 중요하다는 의견이 있었다. 또한 국제적인 비교뿐만 아니라 자국의 정책 활용 또한 목적으로 진행되어야 하며, 정책적 활용에서 평가의 결과가 단순히 유아 발달 정도를 비교하는 수준이 아니라 유아교육 전반(교육과정, 시설 및 환경, 교사 등)에 대해 정책적 시사점이 도출될 수 있도록 고려해야 한다는 점에서 정책적 활용도를 높은 순위로 선택한 것이다. 정책적 활용도 다음으로는 평가도구의 질에 대한 의견이 많았다. 평가도구를 개발함에 있어서 평가도구의 질은 우선적 고려대상이 아니라 반드시 갖추어야 할 요소로, 사용되는 평가도구가 국제비교를 위해서는 타당도, 신뢰도 측면에서 질적인 기준을 충족시킬 수 있어야 한다고 의견이 제시되었다. 국가 간의 동일한 조건에서 유아발달이 평가될 수 있도록 평가도구, 평가과정의 표준화가 필요하다는 측면에서 평가도구의 질이 담보되어야 한다고 주장하였다.

세 번째는 국제비교 가능성 여부이다. 평가의 개념, 평가 모형, 자료 분석 및 해석에 있어서 국제비교 가능성이 확보되지 않으면 연구의 의의를 상실할 수 있으며, 비교 연구이므로 목적에 맞는 자료 수집이 이루어지려면 다양한 국가가 참여해야하기 때문에 국제비교가 실제로 적용될 수 있는지의 여부가 중요한 고려사항이 된다고 제시하였다.

이 밖의 내용으로는 기존 평가도구와의 중복성, 실시의 용이성, 배경변인 설정, 비용 및 시간, 참가자의 심리적 부담, 표집의 적절성 등이 제시되었다. 기존 평가도구와의 중복성은 이미 ESP(Educational and Social Progress)가 초등학교 1학년부터 6학년까지의 종단연구로 시작이 되기에 기존에 개발된 유아발달 평가도구가 있다면 이들 간의 중복성에 대한 검토가 필요하다고 하였으며, 실시의 용이성은 평가 실시로 인한 부담이 크면, 자료 수집을 지속적으로 진행하기 어려움이 따르기 때문에 검사 길이와도 연관 지어 확인해볼 필요가 있다고 제시하였다.

## 2) 국제수준의 평가도구 타당화 과정을 위한 고려사항

국제수준의 타당화 과정에서의 고려 사항에 대하여 선행연구들을 참조한 결과를 바탕으로 8가지 문항을 선정하여 제시하였다. 표본의 크기, 국가별 문항 수정의 정도, 적정 총 문항 수, 총체적 발달평가 도구와 특정영역 평가도구의

차이, 사회정서영역 평가의 타당도 확보 또는 향상 방안, 평가자(교사, 부모, 전문검사자), 평가 시기 및 시행 횟수, 평가(검사)환경의 표준화 확보이며 기타의 견까지 포함 총 9개의 문항으로 구성하여 전문가들에게 질문하였다.

표본의 크기에 대해서는 대체적으로 모집단의 대표성과 집단 간 비교 시 효과크기를 고려해 정해지는 것이 바람직하다는 의견이 다수였다. 현재 OECD PISA연구에서 만 15세 학생 5200명, IEA ICILS연구는 중학교 2학년 학생 2900명, IEA TIMSS연구에서는 초등학교 4학년 4300명으로, 적어도 3000명 이상을 대상으로 하는 것이 대단위 평가도구를 타당화 시킬 수 있는 기준선이 된다고 한다.

국가별 문항 수정의 정도는 전체 문항 중 20~30% 정도는 개별국가에서 정할 수 있도록 하는 것에 의견이 모아졌다. 국제비교를 기본 목적으로 하는 평가라면 모든 문항은 동일한 조건이어야 하며, 이를 위해 기존 국제 비교에서는 국제기구(OECD, UNESCO, IEA 등)와 각국의 평가 시행 담당 기관 간에 상당 기간 긴밀한 상호작용을 하며, 이를 통해 모든 국가에서 통용 가능한 문항을 선제하는 과정을 거치게 된다. 또한 국제기구에서는 각국의 번안 내용을 몇 차례에 걸쳐 검수하여 국가별 유불리 문제를 엄격하게 관리하기 때문에 국가별 문항 수정의 정도는 점점 적어지는 추세라고 밝혔다.

적정 문항 수에 대해서는 평가영역과 방법에 따라서 고려가 필요하며, 검사 대상이 유아이기 때문에 대체로 20분 이내에 수행 될 수 있는 길이의 검사여야 할 것이며, 구체적인 문항 수는 100개에서 300문항 내외가 적정하다는 의견을 제시하였다. 또한 유아를 대상으로 한 번의 검사로 진행되기 어려운 점을 고려하여 여러 회로 나누어서 평가할 수 있도록 하는 것이 용이하다는 의견도 함께 제시하였다.

총체적 발달평가 도구와 특정영역 평가도구의 차이 항목에서는 대체적으로 총체적 발달평가 안에 특정영역 평가도구를 포함시키면서 진행해야 한다는 의견이 많았고, 집중, 추적해야 할 영역들이 있는 경우는 전체 영역 중 3개 이하로 제한하여 특정 영역 평가를 하는 것이 바람직하다는 의견으로 모아졌다.

사회정서영역 평가의 타당도 확보 또는 향상 방안은 여러 차례 관찰 평가를 실시해야만 타당도 확보 및 향상이 이루어지지만, 대단위 검사의 특성 상 제한되는 사항이므로, 그에 따른 대안으로 행동지표(예. 일주일의 평균 문제행동 횟수)를 가능한 한 많이 활용을 하는 방향으로 가는 것에 대한 의견이 나왔다.

평가자에 대해서는 의견이 여러 갈래로 나누어졌다. 기관의 교사가 참여해야 한다는 의견과 부모가 평가해야 한다는 의견, 전문 평가자가 관찰해야 한다는 의견으로 나누어졌다. 이러한 의견을 종합하여 전문가협의회에서 나온 의견은 가능한 부모를 주평가자로 참여시키고, 교사나 전문 평가자는 부평가자로 평가 체제를 만드는 것이 주요한 방법이 될 수 있다는 의견이 제시되었다.

평가 시기 및 시행 횟수와 관련한 의견에서는 연초에 1회로 한정시키는 것이 현실 가능한 평가 시기 및 시행 횟수라고 대다수의 전문가가 제시하였으나, 연초에 실시하는 것이 비교국가의 학제 시스템을 고려해 조정될 필요성에 대해서도 의견이 나왔다.

평가(검사)환경의 표준화 확보 문항에서는 행동수행을 평가하는 경우를 대단위 검사에 포함시키기는 어렵기 때문에 검사를 사용할 것인데, 검사를 사용하는 경우 검사의 동등성 보장이 중요한 이슈로 나타날 것이라는 의견이 제시되었다. 또한 검사자의 사전교육을 통한 평가 과정과 절차에 관한 표준화 의견도 함께 제시되었다.

## 마. 유아발달수준 국제비교 연구결과 활용 시 고려사항

유아 발달 수준에 대한 국제비교 결과를 국가정책의 성과 평가 등에 활용하기 위하여 고려해야 할 점을 묻는 질문에 대해서는 국제비교의 결과 자체를 국가정책의 성과를 모니터링하고 평가하는데 활용하기는 다소 무리가 있다는 의견이 지배적이었다. 평가의 내용이 국가정책과 연계되기 어렵기 때문에 오히려 평가 결과의 해석에 대한 오류가 있을 시 잘못된 성과 평가가 될 수 있다는 점을 고려해야 할 것이며, 특히 사회·문화적 특성에 상대적으로 많은 영향을 받는 유아발달 정도는 그 결과의 해석이 다양할 수 있기 때문에 각별히 주의해야 된다는 의견이 있었다.

이와 반대되는 의견으로, 우리나라 유아교육이 앞으로 나아가야 할 방향을 확인하기 위해서는 우리나라와 비슷한 문화권에 속하면서 보다 선진적인 국가들이라 할 수 있는 일본, 싱가포르, 대만 등에 대한 자료와 정보들을 집중적으로 비교분석하여 이들 국가와 우리나라의 차이가 무엇인지 확인할 수 있는 문항들을 포함함으로써 우리나라 유아교육의 현 상태와 수준, 이후 정책을 전개하기 위한 방향 설정 등에 필요한 정보를 얻을 수 있을 것이라는 희망적인 의견도 포함되어 있었다.

추가적인 제안 사항으로는 세계적인 국제비교보다는 지역 내 국가 간의 비교 분석을 먼저 실시할 것을 제안한 의견이 있었다. 아시아권 내에서의 비교분석 연구를 수행하고 그 과정과 결과에 기반하여 국제비교를 위한 활용점을 찾는 것이 프로세스 상으로 적절하며, 가장 사회·문화적인 영향을 받는 시기인 학교 교육 이전의 가정 및 사회의 다양한 맥락적 공통점이 지역수준에서 보다 비교 우위가 있을 수 있다는 이유에서라고 밝혔다.

한편, 국제비교의 수행 방식에 대한 사항도 고려할 필요가 있다는 의견이 제시되었다. 예산의 확보, 국가 참여 방식이 우리나라가 주도하고 다른 국가는 참여하는 방식인지, 아니면 콘소시움과 같이 공동 시행 및 분담이 이루어지는 것인지에 대한 연구수행방식에 따라서 내용과 과정이 달라질 수 있기 때문에 이를 확인해야 한다는 의견과 함께 정책의 활용도를 높이기 위해서는 우리나라 유아교육 단계에서의 쟁점을 연구문제로 다룰 필요가 있다는 의견이 제시되었다.

마지막으로 국제비교의 특성 상 참여국들의 결과가 비교되어 순위가 매겨지게 되면 유아발달 정도의 국제 순위에 대한 부작용이 매우 클 것으로 예상되기 때문에 이러한 역기능을 최소화 할 수 있는 방안이 평가 설계단계에서부터 고려되어야 한다는 의견이 제시되었다.

### 3. 2차 조사 결과 분석

2차 조사 결과는 1차 조사에서 수합된 국제비교 목적에 따른 평가도구 개발 절차에 대해서 제시하였다. 1차 조사에서 수합된 4가지의 목적에 따라서 평가 대상, 비교대상 국가, 특정한 집단 설정 여부, 표본 수, 문항 수정 비율, 총 검사 문항 수, 평가자, 평가 시행 횟수 및 시기, 평가영역 등을 응답하였으며, 이러한 내용분석은 4가지 목적에 필요한 평가절차를 구성할 수 있다. 응답결과, 국제비교 목적에 따라서 상이한 영역들이 발견되었는데, 우선 비교 목적에 대한 우선 순위를 확인한 뒤, 전반적인 내용을 제시하고, 각 목적별 특이점을 서술하였다.

#### 가. 유아발달 국제비교 목적의 우선순위

유아발달 국제비교 목적의 우선순위에서 1순위는 45명 중 20명이 응답한 국

제적 수준에서 우리나라 유아 발달수준의 현 위치 점검으로 확인되었다. 2순위는 12명이 응답한 국제적 수준에서 유아의 발달과 학습 성과 비교였으며, 3순위는 적극적인 조기중재 방안 마련, 4순위는 개발도상국과의 교류협력을 위한 정보제공 및 지원으로 나타났다.

대부분의 전문가들은 국제적인 수준에서 우리나라 유아 발달 수준의 현 위치 점검을 위해 국제적 수준의 유아발달평가가 필요하다고 응답하였으며, 정책적으로 활용할 시에도 본 목적은 중요한 가이드라인이 될 수 있다는 지적도 함께 나왔다.

2순위의 국제적 수준에서 유아의 발달과 학습 성과 비교는 정책적으로 누리 과정에 대한 평가가 이루어지는 것을 포함한 목적이라 할 수 있겠다. 국가 간의 비교를 통하여 정책적 우수성을 평가하기 위한 목적이 국제비교 평가도구의 큰 목적 중에 하나이기 때문에 전문가들은 본 목적을 선택했다고 응답하였다.

3순위의 적극적인 조기중재는 영국의 EPPE 프로젝트와 같이 유아들이 아주 어린 나이에서부터 개입을 하여 최대한 출발선을 맞추겠다는 목적과 비슷하다. 국제적으로 유아의 발달수준을 비교하여, 현저하게 낮은 발달 수준을 보이는 유아 집단에게 즉각적으로 개입하여 발달 수준을 맞춘다면 소수 집단과 다문화 가정 등에 큰 도움을 줄 수 있을 것이라고 응답하였다.

4순위의 개발도상국과의 교류협력을 위한 정보제공 및 지원은 우리나라를 롤 모델로 하는 여러 국가들에게 우리의 현재 유아 발달 수준 정보를 제공하고 정책 공유 및 정보 제공하여 많은 개발도상국과의 교류협력을 위함이라고 응답하였다.

## 나. 전반적인 평가 도구 설계 사항

국제비교를 위한 평가 도구 개발과 관련하여 적합한 연령대는 대부분의 전문가들은 만 5세 이상을 선택하였다. 이 연령대의 유아들은 실제로 검사를 받을 수 있다는 것이 가장 큰 이유였다. 비교 대상 국가는 주로 OECD 중심의 선진국과 사회문화가 비슷한 동북아시아권역 국가로 응답이 모아졌다. 특정 집단 설정에 대해서는 유아교육 및 보육기관 미이용 집단과 지역적 분류 집단이 가장 많은 응답으로 확인되었다. 현재 누리과정을 실시하고 있는 우리나라에서는 대부분의 유아들이 어린이집이나 유치원과 같은 기관을 이용하고 있어서 일부에 불과하지만 기관을 이용하지 않는 미이용 유아들이 존재하고 있다. 이들 미이용

유아들은 표본수가 작기는 하지만 꼭 포함되어야 할 집단이라고 전문가들은 강조하였다. 또한 지역적 분류는 지역에 따라서 상이한 정부정책이 존재하기 때문에 지역적 분류가 배경변인에 필요한 사항임을 제시하였다.

해당 목적에 필요한 표본 수로는 3,000~5,000명과 5,000~ 0,000명이 동일하게 가장 많은 응답수를 보였으며, 조사의 실효성을 생각해 봤을 때는 2,000~3,000명이 적절하다는 의견 또한 확인되었다. 일반적으로 대단위 검사에서 가장 많이 사용되는 목표 표본 수는 측정모형과도 관련이 있는데, 문항반응이론(Item Response Theory) 모형을 대단위 검사 모형으로 사용할 시, 모형의 안정성을 위하여 대략 3,000명~5,000명으로 표본 수를 설정하는 것으로 알려져 있다 (Hambleton, 1990).

한편, 총 검사의 문항 수에 대한 질문에 전문가들은 50~75문항이 적절하다는 의견이 가장 많았다. 적합한 평가자에 대한 응답에서는 전문평가자와 교사가 가장 많은 응답을 나타내고 있었다. 객관적 수치를 측정하는 경우가 아니라면 훈련된 전문평가자가 관찰평가를 실시하여 검사의 신뢰도 확보가 중요하기 때문이라는 의견을 제시하였다. 또한 국가별 문항 수정 가능 비율에 대해서는 전체 검사의 10% 이하로 설정하는 것이 적절하다는 응답이 가장 많았으며, 30% 이하, 20% 이하 순으로 나타났다.

적절한 평가 시행 횟수 및 시기에 대해서는 연간 1회가 가장 응답이 많았지만, 비용이나 시행 준비 절차 등을 고려한다면 매년 실시하는 것은 무리가 있다고 하여 2년 또는 4년에 1회 실시하는 응답도 비슷한 비중으로 응답이 나왔다.

포함되어야 할 평가영역에 대해서는 모든 전문가들이 인지발달, 언어발달, 사회정서발달, 신체발달을 포함하는 것에 대하여 이견이 없었다. 추가적인 의견으로는 평가영역에 문화발달 영역에 대한 내용이 언급되었다. 평가 영역 내 하위 영역별 중요도에 대한 응답은 아래의 표와 같다. 인지발달 영역에서는 수 감각 및 셈하기 하위영역이 가장 중요한 것으로 인식하고 있었으며, 분류 및 공간 인지와 학교준비도가 두 번째로 중요한 것으로 인식하고 있었다. 언어발달 영역에서는 말하기, 듣기, 읽기와 의사소통 능력이 가장 중요한 것으로 인식되고 있었으며, 다른 발달 영역보다 상대적으로 높은 평균점수를 나타내는 것으로 확인되었다. 사회정서발달영역에서는 타인의 감정 파악 및 감정조절과 또래인지(관계)가 가장 중요한 것으로 확인되었으며, 부모관계 또한 중요한 수준으로 인식하는 것으로 나타났다. 마지막으로 신체발달 영역은 지각 및 기본 운동능력이

가장 높았으며 나머지 영역은 타영역에 비해 상대적으로 낮은 수준의 중요도를 보였다.

<표 V-3-1> 평가 영역 내 하위 영역별 중요도 평균 (N=45)

	구분	평균	표준 편차
인지 발달	수 감각 및 셈하기	4.41	0.84
	분류 및 공간인지능력	4.23	0.91
	기억 및 모방	3.48	0.84
	기초문제해결력	3.47	1.07
	기초추론능력	3.36	1.05
	학교준비도	3.81	0.63
언어 발달	말하기	4.44	0.84
	듣기	4.32	0.82
	읽기	4.17	1.10
	쓰기	3.85	1.03
	의사소통능력	4.17	0.87
	어휘력	3.95	0.73
	문자언어 기초인식	3.61	1.07
사회 정서 발달	집단생활능력	3.60	1.00
	타인감정파악 및 감정조절	4.23	0.78
	또래인지(관계)	4.21	0.63
	부모관계	4.19	0.56
	기초시민의식	3.16	0.87
	문화다양성	3.07	0.81
	상호존중	3.50	0.70
신체 발달	찬예육구, 승인육구, 성취육구, 불안 우울	3.63	0.69
	대근육	3.59	0.97
	소근육	3.59	0.97
	지각 및 기본운동능력	4.23	0.78
	기초위생습관	3.83	0.63
	식사횟수와 정도	3.33	0.82
	키, 몸무게	3.91	1.01

국제비교를 위한 유아발달 평가도구를 개발함에 있어서 고려해야 할 사항에 대한 중요도는 아래의 표와 같다. 총 11가지 영역으로 1차 조사에서 개방형질문으로 물어본 내용을 5점 척도로 전환하여 다시 질문하였다.

전문가들이 선택한 가장 중요한 항목은 평가도구의 질(타당도 및 신뢰도)과 정책적 활용도가 높은 수준의 중요도를 나타냈으며, 대부분의 사항이 중요한 고

려사항이라는 점과 함께 상대적으로 낮은 수준의 중요도를 가진 항목은 기존 평가도구와의 중복성으로 조사되었다.

<표 V-3-2> 유아발달 평가도구 개발 시 고려사항 중요도(N=45)

고려사항	평균	표준편차
정책적 활용도	4.47	0.51
평가도구의 질(타당도 및 신뢰도)	4.52	0.70
국제비교 가능성	4.33	0.94
문화적 다양성 및 차이 반영 정도	3.99	0.87
현지화의 용이성	4.07	0.81
비용	4.02	0.66
검사길이	4.02	0.66
참여자의 심리적 부담	3.73	0.48
기존 평가도구와의 중복성	3.35	0.67
배경변인 설정	4.18	0.56
표집의 적절성	4.13	0.99

## 다. 목적별 특이사항

### 1) 국제적 수준에서 유아의 발달과 학습성과 비교

유아의 발달과 학습성과의 비교를 위해서는 기본적으로 교육과정 또는 교육 정책 및 각 나라의 정책적 환경이 중요한 투입변수(input)가 되어야하며, 성과지표(outcome)는 동일한 척도로 판단해야 한다. 세계시민 덕목과 기본적 학습 소양 등의 평가도구는 동일한 척도를 만들기 위한 이러한 목적을 가지고 개발된 것이다. 이런 평가도구들은 정책효과에 대한 국가별 특징을 비교하할 수 있다는 장점을 지니고 있다.

적합한 연령대로는 대부분의 목적과 동일하게 만 5세 이상의 응답이 가장 많았으며, 비교대상 국가로는 OECD 중심의 선진국들로 응답이 가장 많았다. 이러한 응답이 나타나게 된 이유는 학습성과의 비교는 선진국과 비교해야 정책적 시사점을 이끌어내기 쉽고, 다른 나라에서 영향력이 큰 정책들을 벤치마킹 할 수 있다는 것이기 때문이다. 특정한 집단 설정은 다른 목적들과 유사하게 나타났으며, 대부분의 항목에서 전반적 평가도구의 설계 사항과 동일한 결과를 나타내었다.

## 2) 국제적 수준에서 우리나라 유아 발달수준의 현 위치 점검

국제적 수준에서 우리나라 유아 발달수준의 현 위치 점검은 다른 목적들보다 우선순위가 가장 높게 나온 목적이다. 전문가들이 판단했을 때, 국제비교의 목적에 가장 가까운 목적이라고 판단한 것이기 때문에 이 목적에서는 대규모 검사의 특징들이 많이 담겨있는 것으로 확인되었다. 우선 적합한 연령대는 만 5세 이상이라고 응답했다. 만 5세 이상이어야 평가자에 관계없이 검사가 진행될 수 있다고 판단한 것이다. 비교대상 국가로는 OECD 중심의 선진국들로 응답하였으며, 특정한 집단 설정은 기관 미이용 집단과 저소득층 및 사회적 배려 대상, 다문화 가정, 장애유아, 지역적 분류 집단, 성별 등 다양한 집단들이 응답에서 제시되었다.

표본 수는 3000명에서 5000명 수준의 표본이 필요한 것으로 확인되었으며, 국가별 문항 수정 가능 비율은 10% 이하로 응답한 전문가가 많은 것으로 나타났다. 적정 문항 수는 50~75문항 정도로 선택하였는데, 다른 국제비교 검사보다 문항 수가 적은 이유에 대해서는 문항이 유아를 대상으로 한 내용이기 때문이라고 응답하였다. 적합한 평가자로는 전문평가자와 교사가 비슷한 비율로 응답을 하였으며, 적절한 평가 시행 횟수 및 시기는 연간 1회가 가장 높은 응답을 보였지만 2년에 1회와 4년에 1회라고 응답한 비율도 높은 것으로 확인되었다. 현재 PISA나 TIMSS의 경우도 3년 또는 4년에 한 번씩 검사를 시행하고 있기 때문에 2년 이상에 1회라고 응답한 비율이 높게 나타났다고 볼 수 있다. 하지만 유아의 경우, 변화 속도가 빠르며, 외부의 환경적인 영향을 더 많이 받기 때문에 연 1회라고 응답한 것이 가장 많은 수를 차지한 이유라고 볼 수 있다.

## 3) 국제비교를 통한 적극적인 조기 중재 방안 마련

다른 목적과 동일하게 적합한 연령대는 만 5세 이상이라고 응답했지만, 만 2세라고 응답한 수도 큰 비중을 차지하고 있었다. 이는 조기 중재를 실제로 시행했었던 영국의 EPPE 프로젝트의 내용이 반영된 것이라고 볼 수 있다. 비교대상 국가는 OECD 중심의 선진국들로 많은 전문가들이 응답하였으며, 특정한 집단 설정은 유아교육 및 보육기관 미이용 집단과 저소득층 및 사회적 배려 대상이 있는 집단에 많은 응답을 하였다. 조기 중재 방안은 저소득층 및 사회적 배려 대상을 초점으로 둘 것이기 때문에 표본으로서의 집단 반영이 반드시 필요한 것이다.

표본 수는 다른 목적들과 마찬가지로 3000명에서 5000명 수준의 표본이 필요한 것으로 확인되었으며, 국가별 문항 수정 가능 비율은 20% 이하로 응답한 것이 많은 수를 차지하고 있었다. 적정 문항 수는 50~75문항 정도로 선택하였으며, 적합한 평가자로는 전문평가자가 다수의 응답으로 나타났다. 적절한 평가 시행 횟수 및 시기에서는 다른 목적과 차별화되는 점이 발견되었는데, 아주 어린 나이에는 학기당 1회라고 응답한 비율이 상당한 수준을 보였다. 이는 어린나이에 조기중재를 한 이후, 나타나는 정책적 효과를 판단하기 위함이라고 전문가들이 의견을 제시한 것이다.

#### 4) 국제비교를 통한 적극적인 조기 중재 방안 마련

개발도상국과의 교류협력을 목적으로 한 정보제공 및 지원차원에서의 국제비교는 현재 아시아 태평양 지역의 유아발달에 대한 국제비교 연구나 개발이 전무하므로, 평가지표의 개발은 큰 의미가 있다는 필요성을 가지고 있다. 우선 적합한 연령대는 다른 목적들과 동일하게 만 5세 이상을 대상으로 설정하는 의견이 주를 이루었지만, 비교대상 국가에서는 개발도상국 중심의 아시아 권역으로 한정되는 특징이 있었다. 또한 특정한 집단 설정에는 다른 목적들과 비슷한 양상으로 유아교육 및 보육기관 미이용 집단과 지역적 분류 집단을 설정하는 의견이 제시되었으며, 표본 수는 다른 목적들과는 달리 상대적으로 1000명 이하의 응답이 많이 제시되었다.

한편, 문항 수정 가능 비율은 전체적으로 10% 이하 응답이 가장 많이 나타난 것이 특징이다. 총 검사의 문항 수는 50문항 이하로 간단하게 보는 것이 좋다는 응답이 많이 있었으며, 적합한 평가자는 전문평가자로 다른 목적에서 나타나는 응답과 동일하게 나타났다. 하지만, 적절한 시행 횟수 및 시기로는 특징적으로 연 1회가 가장 많이 응답했고, 상대적으로 다른 목적에서의 2년에 1번, 4년에 1번에 응답한 전문가 또한 연 1회로 응답하였다. 이렇게 응답한 이유는 개발도상국의 경우, 정책 사항이나 교육제도에 있어서 급변할 수 있는 여지가 있기 때문에 이러한 사항을 반영한 것이라고 의견을 제시하였다.

## VI. 정책적 제언

### 1. 유아발달 평가 목적에 따른 국제비교 방안

#### 가. 조기 개입을 위한 국제수준의 유아발달 평가

V장의 전문가 델파이 조사결과에서 제시된 바와 같이 유아발달 평가의 목적에 따라 전반적인 유아발달 수준 평가, 학습성과 평가, 조기개입을 위한 평가, 국제개발협력을 위한 평가로 구분할 수 있다. 본 연구에서는 연구 목적에 부합하여 취약소외계층 유아의 출발점 평등을 위한 조기개입 및 일반 유아 대상의 학교준비도 국제비교를 위한 평가도구 개발 방안을 중심으로 제시하고자 한다.

##### 1) 목적

이 평가는 각 국가별로 유아 단계에 나타나는 계층, 성별, 인종, 지역 등 사회경제적 배경에 따른 학습 출발점의 차이를 파악하는 것에 초점을 둔다. 이는 초·중등학교 학습의 기초가 되는 유아발달과 학습능력을 측정하여 각 국가별로 유아교육 단계 교육불평등에 영향을 미치는 관련 요인들을 비교하여 분석하기 위한 것이다.

궁극적으로는 교육불평등 해소를 위해 국가별로 시행하는 조기 개입 정책에 대한 정보를 함께 조사함으로써, 사회 수준에서 계층간의 교육불평등을 감소하는데 유아교육의 조기 개입이 어떤 효과를 갖고 있는가를 비교하고 유아교육 조기 개입 정책의 효과성을 높이기 위한 시사점을 얻고자 한다.

##### 2) 대상

조기 개입을 위한 유아발달 평가는 이후 학교 학습과 학생 성취 격차에 결정적 영향을 미치는 시점을 확인하는 것이 중요하다. 이런 점에서 평가 대상은 유아교육에 해당되는 만 3세부터 만 5세로 설정하는 것이 적절하다.

이러한 경우, 3개 연령대를 모두 조사하는 방법과 특정한 연령대(예: 만 5세)만을 조사하는 방법이 모두 가능하다. 어떤 방식을 취하는가는 예산 규모에 따

라 좌우된다. 예산 상 제약이 없다면 모든 연령대를 조사하는 것이 바람직하고, 예산 제약이 있어 하나의 연령대만 선택해야 한다면 아동들의 발달 상 차이가 어느 정도 드러날 수 있고 아동들이 검사에 어느 정도 참여할 수 있는 시기인 만 5세 시점이 보다 적절할 것이다. 만 5세는 유아교육의 마지막 단계에 해당된다는 점도 중요하게 고려될 수 있다.

이 평가는 조기 개입에 주된 초점을 두고 있으므로 국가 평균적 발달 수준보다는 한 국가 내 여러 사회 집단 간 발달 수준의 차이를 잘 보여줄 수 있는 것이 중요하다. 이런 점에서 관심 집단, 예컨대 성별, 지역, 사회계층, 인종(민족) 등에 따른 아동의 발달수준 차이를 잘 파악할 수 있도록 평가 대상에서 이들 집단들이 집중적으로 고려될 필요가 있다.

### 3) 참여국가의 범위

사회계층, 인종, 집단 간 교육 격차와 불평등이 사회문제로 부각될 정도로 사회계층 분화가 비교적 뚜렷하고 이를 해소하기 위한 유아 단계 조기 개입 정책이 시도되고 있는 국가가 주요 대상으로 포함되어야 할 것이다. 여기에는 주로 계층 분화가 심화되어 있는 선진국(예: 미국, 영국, 프랑스 등), 조기 개입과 교육 평등 수준이 높은 북유럽 복지국가(예: 핀란드) 등을 필수적으로 포함하는 것이 타당할 것이다.

이와 함께 문화적 동질성을 조건으로 비교 검토하기 위하여 아시아 지역(예, 일본)을 포함하는 것을 고려해 볼 수 있다.

### 4) 평가 영역

일반적으로 유아 발달 수준을 평가할 때 사용되는 영역이 모두 포함될 수 있으며, 다음과 같은 영역들이 있다.

- 인지발달: 수 감각 및 셈하기, 기억 및 모방, 분류 및 공간인지능력
- 언어발달: 말하기, 듣기, 읽기, 쓰기, 어휘력
- 사회정서발달: 타인감정파악, 감정조절, 상호존중, 친애욕구
- 신체발달: 키, 몸무게, 기초위생습관, 기본운동능력
- 학교준비도: 위에서 제시한 발달적 측면 이외에 학교 학습을 위해 필요한 사회화(기초규범, 규칙준수, 공동생활 등)

다만, 조기 개입을 위한 사회 계층 간 불평등 지점을 확인하는 목적을 지니므로, 이후 학교 학습의 차이와 관련성이 높은 영역을 좀 더 집중적으로 평가할 필요가 있다. 일반적으로 학교 학습 또는 성취도는 아동의 지적능력(기억력, 인지능력, 수와 셈하기 등), 언어발달 수준 등과 밀접하게 관련되어 있다.

## 5) 평가 타당화 과정

검사도구의 국제수준 타당화는 두 가지 방식으로 진행될 수 있다. 하나는 참여 국가가 공동으로 개발하여 타당화과정을 거치는 것이고, 다른 하나는 한국이 주도적으로 개발한 후 각 국가의 상황에 맞게 타당화 과정을 거치는 것이다. 전자는 한국에서 개발한 것을 각 국가의 문화에 맞게 변안하는 과정에서 발생할 수 있는 편향의 가능성을 줄여주는 장점이 있으며, 후자는 시간과 비용 측면에서 타당화 과정을 좀 더 효율적으로 수행할 수 있으며 우리나라의 쟁점을 국제적으로 비교하는 데 장점이 있다.

조기 개입에 대한 검사도구의 국제수준 타당화 과정에서 집중적으로 검토해야 할 사항은 한 국가 내에서 사회 계층 또는 다양한 집단 간 기초 학습의 차이를 잘 분별해 낼 수 있는가를 확인하는 데 있다. 개발된 검사도구가 아동들의 전반적인 발달 수준을 잘 측정·평가하지만 다양한 집단 간 차이를 잘 드러내지 못한다면 조기 개입을 위한 평가도구로서 적절하지 않을 수 있다. 국제수준 타당화 과정에서는 참여 국가들의 맥락 속에서 다양한 집단의 차이를 어느 정도 타당하게 측정할 수 있는가를 확인할 필요가 있다.

### 가) 표본수

타당화 과정을 위한 표본 수는 모집단의 1% 정도로 한다. 만 5세를 조사 연령으로 할 경우 우리나라의 경우 표본 수가 약 4,700명 수준이 될 것이다. 이와 같은 표본 규모는 유아 단계의 아동들을 조사하기에는 매우 많은 숫자이다.

본 조사가 아닌 평가도구의 국가별 타당화를 충실하게 수행하기 위해서는 충분한 사례수가 확보될 필요가 있다. 특히, 조기 개입을 위한 유아발달 평가의 경우는 여러 사회 집단 간 차이를 타당하게 추정해내야 하므로 표본 사례수의 충분성은 보다 더 요구된다.

그러나 평가 시행을 위한 예산 상 제약을 고려하지 않을 수 없다. 유아단계 발달평가는 초중등학교의 성취도평가와 달리 지필평가가 불가능하므로 비용의

증가는 더 뚜렷할 수밖에 없다. 이런 점을 고려하여 평가도구의 국가별 타당화가 완료되고 실제 평가가 시행될 때에는 국가별로 0.5% 정도의 표본이 적정할 것으로 보인다(우리나라의 경우 약 2,300명 수준).

#### 나) 문항 수정 가능 비율

국제 비교의 타당성을 높이기 위해서 문항 수정 비율을 최소화하는 것이 필요하다. 평가 문항이나 설문 문항 중 국가 간 문화적 차이로 조사 측정이 불가능한 경우에 한하여 문항 수정을 허용할 수 있다. 그러한 경우에도 평가 문항의 내용, 수준(난이도) 등의 동일성을 확보할 수 있도록 각 국가의 전문가들이 공동으로 참여하여 수정 사항을 검토하는 과정이 필요하다.

각 국가마다 조기 개입과 관련하여 문제의식이나 이슈가 다를 수 있다. 우리나라의 경우 조기 개입이 무상의 누리과정 시행으로 이루어진다면, 다른 국가의 경우 일부 취약 계층에 보상교육(compensatory education)이 추진될 수 있다. 조기 개입과 관련하여 어떤 방법이 효과적인지 알아보고자 하는 이슈가 제기될 수 있을 것이다. 이와 같은 점을 반영하기 위해서 특정 국가의 특수한 문제를 연구하기 위한 문항이나 정보 수집을 부가 조사(평가)의 형태로 추가할 수 있다.

#### 다) 적정 문항 수(검사 길이)

인지발달이나 언어발달을 평가하는 것은 아동을 직접 대상으로 해야 한다. 이와 같은 경우, 평가 문항이나 검사 길이는 평가대상 아동의 집중력을 고려하여 약 20분 정도로 구성될 필요가 있다. 만약 평가 영역이 많거나 평가 항목이 많아 시간이 더 소요될 경우는 평가(검사) 시행 횟수를 늘려서 한 번의 검사 시간을 적절히 통제할 필요가 있다.

교사나 훈련된 평가자를 통한 간접 평가의 방식의 경우도 20시간을 넘지 않는 것이 바람직하다. 교사나 성인을 통한 간접 평가는 한 평정자가 여러 아동들을 평가해야 하는 상황이기 때문에, 문항 수가 너무 많거나 시간이 너무 많이 소요될 경우 전체적인 검사 타당성이 저하될 가능성이 있다.

#### 라) 평가자

기억력, 어휘력 등 인지발달과 언어능력과 같이 아동을 직접 평가해야 하는 검사도구의 경우는 전문 훈련을 받은 평가자가 평가할 필요가 있다.

사회정서발달 영역, 학습준비도 등과 같이 아동이 교실에서 행동하는 것을

관찰해야 하는 경우는 교사가 응답할 수 있을 것이다. 이 경우에도 교사 간 편차를 줄이고 객관적인 정보 평가가 이루어지기 위해서는 문항의 내용과 평정 기준에 대한 훈련이 필요하다.

배경 정보(성별, 연령, 가정 배경, 거주지역 정보 등)나 조기 개입 경험에 대한 사항은 평가 또는 검사도구보다는 설문지 조사를 통하여 학부모가 응답해주는 방식으로 이루어질 수 있다.

#### 마) 평가 시행 횟수 및 시기

평가 시행 횟수는 동일한 표본 아동에게 여러 번 반복 시행될 수 있고(중단 연구), 한 시점에서 횡단적으로 시행될 수 있다. 전자는 평가 영역의 발달 궤적을 아동 개인별로 추적할 수 있어 조기 개입이나 정책의 효과를 좀 더 분명하게 분석할 수 있는 장점이 있다. 반면, 후자는 인과적 분석에 제한점이 있지만 연구 수행과 표본 관리에서 수월한 측면이 있어서 전반적인 유아발달 수준의 격차와 유발 요인에 대한 전반적인 실태파악에 유리한 측면이 있다.

다양한 평가 횟수와 주기가 시도될 수 있겠으나, 타당한 방식을 결정하기 위해서는 평가 목적을 일차적으로 고려할 필요가 있다. 조기 개입을 위한 유아발달 평가는 이후 학교 학습과 학생 성취 격차에 결정적 영향을 미치는 발달 영역을 확인하는 것이 중요하다. 이는 이 평가가 한 시점에서 나타나는 다양한 집단 간의 격차를 정확히 확인하는 데 초점을 둘 필요가 있다는 것을 의미한다.

이런 점에서 보면, 조기 개입을 위한 평가 시행 횟수는 1회만 시행하는 방법도 충분히 의미가 있다고 할 수 있다. 매년 만 5세 시점에서 반복하여 평가할 수 있고, 만 3세부터 만 5세까지 아동을 동시에 평가하되 3년 마다 주기적으로 평가할 수도 있다. 물론, 평가 목적으로 조기 개입 정책의 효과를 강조하게 된다면, 이를 위해서 중단 조사와 평가가 시도될 수 있다.

## 나. 학교준비도 국제비교를 위한 유아발달평가

### 1) 목적

초등교육 입문 단계에서 일반적인 유아들의 학교준비도를 국제적으로 비교하고 학교준비도에 영향을 미치는 다양한 가정, 교육, 제도적 요인을 파악한다. 이를 통해 일반 유아의 학교준비도 제고를 위한 정책적 시사점을 얻는다.

## 2) 참여 국가 범위

원칙적으로 세계 모든 국가를 대상으로 한다. 단, 우리나라가 유아교육 발전을 위해 벤치마킹하고자 하는 대상은 주로 선진국들이므로, 한국의 입장에서 비교 대상 국가로 관심을 갖는 국가는 OECD 중심의 선진국이 될 것이다.

## 3) 평가 대상, 표집크기 및 평가 주기

평가 대상 연령은 국가별 초등교육 입문 직전 연령으로 하며, 우리나라의 경우는 만 5세이다. 평가 대상 표집 및 결과 분석 시 고려할 하위 집단은 성별, 지역별(도시화 정도를 반영), 유아교육과 보육 기관 이용 여부에 따라 구성한다.

표집 크기는 약 2,000명이 되어야 하며, 유아 대상의 평가는 성인이나 학령기 아동에 비해 더 많은 평가 비용, 시간을 필요로 하므로 평가 대상 규모를 작게 설정하는 것이 현실적이다. 평가 주기는 매 3년에 1회 평가를 실시한다.

## 4) 평가 영역

인지발달과 사회정서발달 영역을 반드시 포함한다. 인지발달영역에는 언어발달(말하기, 듣기), 수 감각 및 셈하기, 기초추론능력을 포함하고, 사회정서발달영역에는 부모관계, 타인감정과악 및 감정 조절, 불안과 우울의 하위영역을 포함할 수 있다. 애착이 특정한 사람과의 친밀감을 뜻한다면 사회성은 보다 다양한 사람과의 긍정적인 관계 형성을 의미한다면 정서성은 자기이해, 공감, 충동통제, 자기조절 등 정서적 측면에 대한 이해라 할 수 있다. 따라서 사회정서 발달수준 평가는 다른 사람과의 상호작용을 돕는 의사소통 능력과 유아 자신의 행동을 조직화하고 조절하는 능력에 대한 평가라고 볼 수 있다.

우리나라 교육과학기술부(2012)의 인성교육비전보고서에서는 사회성과 정서성 모두를 인성교육의 중요차원으로 제시하고 있으며, 미국의 ‘학업 및 사회정서학습을 위한 협력단체’(Collaborative for Academic, Social and Emotional Learning)에서도 ‘자기 인식(Self-Awareness)’ ‘자기관리(Self-Management)’ ‘사회적 인식(Social Awareness)’ ‘대인관계의 기술(Relationship Skills)’ ‘책임 있는 의사결정(Responsible Decision Making)’을 사회정서적 요인으로 구성하고 있다.

아동기에 이러한 사회정서적 능력의 기초를 형성하고 이에 대한 교육을 진행하는 것은, 이후 발달에 매우 중요한 기반이 될 수 있는 바 지적 성취 못지않

게 중요한 부분이므로, 이러한 사회정서적 능력을 유아발달 평가지표로 포함하는 것은, 누리과정 정책에 대한 효과평가 뿐 아니라 앞으로 우리나라 유아교육이 나아가 할 방향에서도 핵심적으로 판단된다. 특히 요즘 들어 많이 나타나는 유아들의 분노, 공격성, 불안 등 중요 정서적 반응들을 점검하고 확인하는 구체적이고 실증적인 국가수준의 노력이 필요해 보인다. 유아발달평가 연구를 앞으로 수행하게 될 때 사회정서적 발달 영역의 문항들은 이미 국내외 유아교육 연구에서도 많이 활용되었던 문항들을 추출하여 이를 수정 보완하여 사용할 수 있을 것으로 판단된다.

## 5) 평가자

학교준비도 평가를 위해서는 다양한 평가자를 활용하는 것이 보다 정확하고 효과적인 평가가 이루어질 수 있을 것이다. 인지발달 영역의 경우, 평가 관련 훈련을 받은 교사(유치원과 어린이집을 이용하는 만5세 대상)와 전문평가자(유치원과 어린이집 미이용 만5세 대상)를 모두 활용한다. 사회정서발달 영역의 경우 타인감정 파악 및 감정 조절, 불안, 우울 등에 대하여 전문평가자가 검사하고, 부모관계에 대한 문항은 부모를 평가자로 활용한다.

## 6) 평가도구 타당화과정의 고려사항

먼저, 평가문항이나 설문지를 번안하는 과정에서 국가 간 문화적 차이를 고려해 문항 내용의 수정을 부분적으로 허용할 수 있다. 단, 이 경우 국제비교 연구를 총괄하는 기구의 검토와 승인을 받도록 한다. 또한 국가별 필요나 관심사에 따라 해당 국가에서 필요로 하는 정보를 수집하기 위한 설문 문항을 일부 추가할 수 있도록 한다.

학령기 아동에 비해 유아를 대상으로 한 평가는 평가 시간, 평가 방법 등에서 많은 제약을 받는다. 따라서 평가의 타당도, 신뢰도를 확보할 수 있는 방향으로 연구 설계 및 평가 도구 선정, 평가 시행이 이루어져야 하며, 이 과정에서 추가적인 비용과 노력이 요구되는 것을 감수해야 할 것이다.

배경변인을 설정함에 있어서 연구 결과의 정책적 활용도를 높이기 위해, 유아 발달 및 학습준비도와 관련된 다양한 가정, 교육, 제도적 변인들에 대한 정보를 수집해야 한다.

아울러 평가 결과가 아동, 학부모, 교사 등 관련자들에게 고부담 평가가 되지

않도록 유의해야 한다. 고부담 평가의 경우 높은 평가 결과를 얻기 위한 왜곡된 노력으로 평가 결과의 타당성이 결여될 수 있으며, 예상치 않은 부작용이 발생할 수 있다.

## 2. 유아발달 국제비교 연구 수행의 고려사항

국제 발달비교에 있어서 고려해야 할 대표적인 사항은 학교교육의 시작 연령과 언어 및 문화이다. iPIPS를 이용한 국제 비교연구에서 연구자들은 국제 비교를 위해 연령의 효과를 효과적으로 통제하기 위해 많은 방법을 동원하였고 이를 성공적으로 통제할 수 있었다. 또한 언어 및 문화는 국제 비교연구를 위한 도구의 개발 시에 가장 많은 관심을 기울여야 할 부분이다.

### 가. 연령의 통제

iPIPS 국제비교 시 러시아의 경우 정규학교교육의 시작 시기는 평균 7세 반이다. 국제 평균에 비해 늦은 학교 교육의 시작은 동일 학년에 대한 국제 비교에 있어서 러시아의 결과가 두드러지게 높게 인식되는 착시 현상을 경험하게 된다. Kardanova 외(2014)의 국제 발달비교 연구에서 iPIPS의 러시아어 최초 번안된 버전에서는 학생들 너무 높은 점수를 얻어 천장효과를 보였다. 즉 단순한 국제 발달비교는 이러한 국가 간의 학제의 차이를 무시한 채 실제 차이를 왜곡할 우려가 있는 것이다.

한편, 동일한 학년의 아동이 타국의 아동에 비하여 나이가 많을 경우 학제를 무시하고 연령을 기준으로 평가할 경우에는 국제 발달 비교가 교육의 효과를 무시하여 실제 발달 정보를 축소하여 이해하게 만드는 오류를 범할 수 있다.

### 나. 언어 및 문화 요인

서로 다른 문화적/언어적 배경을 지닌 집단 간의 비교는 도구의 타당성의 문제와 직결되는 문제이다. 특히 언어는 서로 다른 언어체계를 단순 비교로 국제 간 비교를 실시하기에 어려움이 클 수밖에 없다. Kardanova 외(2014)는 평가도구를 다른 언어로 번역하는 데 있어서 다음의 네 가지를 고려해 평가할 것을 제안하고 있다.

첫째, 단어나 문장의 의미를 어떻게 전달하고 있는지 확인해야 한다. 언어의 차이로 인해서 전달하고자 하는 단어와 문장이 완벽하게 번역되기 어려운 것은 당연하다. 이를 감안하여 때로는 단어와 문장의 의미가 얼마나 충실하게 전달되고 있는지를 판단할 필요가 있다.

둘째, 문항의 내용을 살펴보아야 한다. 문항에서 다루고 있는 내용이 단순히 번역의 범위를 벗어나 실제 담고자 하는 내용이 충실하게 담겨진 것인지 확인이 필요하다.

셋째, 문항으로 측정하고자 하는 능력을 확인해야 한다. 문항의 내적 타당도에 해당하는 부분으로, 실제 문항으로 측정하고자 하는 능력이 번역된 문항에서도 일치되는 지를 확인해야 한다.

넷째, 번역의 과정에서 변화되는 문항이 원 문항의 난이도, 문항이 측정하고자 하는 바를 얼마나 잘 반영하고 있는가를 유의하여 보아야 한다. 문항이 번역되는 과정에서 해당 언어를 사용하는 국가의 사회 및 문화적 맥락에 따라서 동일한 문항이 원래 문항에 비하여 난이도가 더 높아지거나 낮아지는 문제가 발생할 수 있다. 예를 들어 서구에서 1/4(quarter)의 개념이 사회 문화적으로 많이 사용되고 있으나 한국의 문화에서는 그렇지 못하다. 그러므로 해외에서 개발된 도구를 한국에 적용할 경우 동일한 문항이 상대적으로 더 난이도가 높은 문항으로 탈바꿈되는 문제가 발생할 수 있다.

연령과 사회, 문화, 언어의 문제를 해결하기 위해 Kardanova 외(2014)의 연구에서는 문제를 대폭 수정하거나 낮은 수준의 문제들을 제거해서 평가의 유의미성을 높이는 노력을 펼쳤다. 그러나 이러한 문제의 변경은 또한 국제비교를 어렵게 한다는 문제가 발생하게 한다. 때문에 모든 문항을 동일하게 구성할 수는 없으나, 국가 간의 비교가 이루어지게 하도록 충분한 정도의 공통문제를 확보할 필요가 있다. 한편 또다른 방안으로는 사회, 문화, 언어적 차이가 크게 반영되는 항목들은 국제비교에서 제외하는 방안이 있을 수 있다.

실제 Kardanova 외(2014)의 연구에서는 국제비교를 위해서는 수학문항 등만을 비교하였는데, 언어와 문화의 차이를 고려할 때 언어발달을 국가 간에 비교하기가 용이하지 않았기 때문이다. 이러한 제한적인 발달의 비교에 회의적일 수 있음에도 불구하고 이들의 연구는 수학 능력의 발달이 첫째 학교에서 꾸준히 증가하며 학교에 입학한 첫째의 중요성을 밝힌 중요한 시사점을 보였다.

## 다. 평가도구의 국제수준 타당화 과정

다양한 언어의 문제로 인해 Kardanova 외(2014)은 유아발달 평가도구를 국제

화하는 과정을 3단계로 나누어 살펴볼 것을 제안하고 있다. 첫째 단계에서는 일차 번역이다. 단순한 번역의 과정을 거치지만, 이는 직역을 의미하는 것은 아니다. 이에, 개별적인 번역자가 2개의 버전으로 각기 도구를 번역하여 이들을 비교하여 최적의 문항으로 재구성할 것을 제안하고 있다.

둘째 단계에서는 단순 번역된 도구의 문항을 원 도구의 언어로 재 번역하는 것이다. 이와 같은 번역-재번역의 과정을 통해 주요하게 평가해야할 점은 1) 원래 문항과 번역된 문항의 실제 의미에 어떠한 차이가 있는지, 2) 문항의 형식 간에 차이가 있는지, 3) 원래 문항과 번역 문항의 제시 방법의 차이가 있는지, 4) 번역된 문항이 문화적으로 적합한지, 5) 핵심 단어의 변경이 적정한지, 6) 번역으로 인한 문장의 길이나 복잡성에 차이가 있는지를 검토할 필요가 있다.

셋째 단계에서는 검사 환경의 동등성 확보하는 것이다. 각 국의 문화나 언어의 차이뿐 아니라 학제나 검사에 대한 환경도 국가 간에 다를 수밖에 없다. 이러한 환경적인 요인이 평가에 미치는 영향을 최소화하기 위해 노력하여야 한다.

Hambleton과 Zenisky(2011) 역시 국제비교연구들이 서양 국가 맥락을 중심으로 개발된 도구를 주로 사용하고 있으므로 각 나라별로 현지화 또는 개별문화에 대해서 타당성을 확보해야 함을 강조하였다. 일반적으로 타당성을 확보하기 위한 연구는 설계, 자료수집, 자료 분석의 3단계로 진행되며 타당성을 확보하기 위한 연구 설계 시, 종류에 따라서 영역 및 세부내용을 확인해야할 필요가 있다. 타당성의 유형은 크게 개념적, 문화적, 언어적, 측정학적 타당성의 4가지로 구분되며, 각각의 하위 요인들은 다음 표와 같다.

<표 VI-2-1> 국제비교 연구의 타당성 유형

영역	타당화 유형	서술과 예시
개념	개념기반	각기 다른 문화권에서의 차이를 수용하는 타당성 (예: 유명인사에 대한 친숙성은 각 국가별로 다르기 때문에 이러한 차이를 동등화시켜야 함)
	이론기반	이론에 기반한 타당성 (예: 단기기억 공간 검사는 짧은 자극을 주어야 하는데, 각 문화별로 다르게 자극을 주어야 함)

(표 VI-2-1 계속)

영역	타당화 유형	서술과 예시
문화	용어상 / 사실기반	문화적 또는 국가적 특징에 따라서 세부적으로 수용하는 타당성(예: 통화 전환(달러->원))
	규범기반	규범, 가치, 실제 등의 문화적 차이를 수용하는 적응성(예: 체면이 깎이는 것에 대한 회피)
언어	언어학 기반	언어의 구조적 차이를 수용하는 타당성 (예: 영어로 Friend는 남녀 모두에게 통용되나, 다수 국가의 언어들은 성별에 따른 명사에 따라서 남자, 여자 친구로 구별되어 있음)
	화용론 기반	대화로 언어사용의 차이를 수용 (예: 면담자의 요구에 의한 단순명료함의 정도)
측정	친숙함/식별력 기반	각 문화별로 각기 다른 익숙한 자극을 평가하는 절차를 설명하는 타당성 (예: 그림검사에서 사용되는 각기 다른 그림)
	형식 기반	문화 또는 반응의 형식에 따른 문화적 차이를 최소화하는 타당성 (예: 응답척도 극단적 응답을 줄이기 위한 응답척도 타당성)

자료: Hambleton, R., & Zenisky, A. (2011). Cross-cultural research methods in psychology, p.37

유아발달 국제비교는 평가의 목적에 따라 평가내용과 도구 및 제반 평가방법과 절차 등이 달라진다. 기존의 유아발달 국제비교 연구는 주로 자국의 취학전 유아를 대상으로 학교준비도를 평가하기 위하여 출발하여 여러 국가간의 비교로 확산된 경우와 최근 글로벌 맥락에서 국제기구들이 주도하여 개발하고 있는 유아발달 국제비교 연구의 2가지 형태로 추진되고 있다.

양자는 공통적으로 한 국가의 맥락을 벗어나는 사회, 문화, 및 언어적 맥락을 포함한 종합적인 유아 발달의 정보를 제공할 수 있다. 유아발달의 국제비교는 각 국가에서 유아교육과 보육의 유효성을 비교해 취약점을 인식하고 이에 대한 정책적 대안을 수립하는 데 기여할 수 있음을 인식하여 향후 평가도구 개발 및 활용에 적극적이고 구체적인 관심을 기울일 필요가 있을 것이다.

올해는 국제비교를 위한 유아발달 평가도구 개발 및 비교연구 수행시 고려해야 할 제반 사항들을 검토하는 기초연구를 수행하였다. 이러한 결과를 토대로 후속적으로 국가수준에서 우선적으로 요구되는 평가의 목적을 선정하고 이에

요구되는 국제비교를 위한 유아발달 평가도구 개발의 타당성 연구가 수행되고, 그 다음에 평가도구가 개발되고, 국제수준의 타당화 과정을 거치는 단계로 중장기적인 안목에서 계획, 추진되어야 할 것이다.

## 참고문헌

- 교육과학기술부(2012). 인성교육 비전 수립 및 실천 방안 연구. 정책연구 2012-41.
- 김수진(2013). TIMSS 2011 우리나라 학생들의 수학 · 과학 성취 특성. 한국교육과정평가원 연구보고 PIM 2013-1.
- 김정권, 여광웅(1978). 정신박약아동 판별연구: 1학년생 선별검사, 서울: 자행회.
- 박연정 (2007). 학령전 아동의 학교준비도 척도개발 및 학교준비도와 관련 변인 연구. 고려대학교 대학원 박사학위논문.
- 박혜정, 강혜원, 장명림(1987). 유아학습준비도 검사개발 연구. 한국교육개발원 연구보고. RR 87-39
- 방소영(2013). 취학전 유아의 학교준비도 검사 개발 및 타당화연구. 경기대학교 대학원 박사학위논문.
- 방은정(2015). 유치원 교사의 교수행동이 유아의 학교준비도에 미치는 영향. 고려대학교 대학원 박사학위논문.
- 서경희, 이상복, 이효신, 서은정, 권명옥, 신윤희, & 이미경. (2004). 행동 분석 시스템을 통한 발달장애 아동의 문제행동 패턴 양식. 한국정서·행동장애아교육학회 학술발표논문집, 101-126.
- 육현진(2014). 국제 문식성 평가 분석을 통한 문식성 교육 시사점 탐색: PIRLS, PISA, PIACC을 중심으로. 청람어문교육. 49.
- 윤명희(1995). 유아학습준비도 검사의 개발 및 표준화에 관한 연구. 교육학연구, 33(1), 117-142.
- 이연섭, 김성일(1981). 유치원 아동의 학습준비도 연구(III). 한국교육개발원 연구보고 제 134집.
- 이영석, 구학봉, 노명환, 김승훈, 차미정, 고승자(1993). 계몽학습준비도검사 개발 연구. 계몽 아동 연구소 연구보고, 6.
- 이은혜(1979). 학교준비도 검사 표준화를 위한 조사 연구. 교육학연구, 17(1), 49-60.

- 이은혜(2006). 아동발달의 평가와 측정. 서울: 교문사.
- 이종숙, 신은수, 박은혜, 김영태, 곽영숙, 유영의 외(2008). 영유아의 언어 인지 사회·정서 발달 평가 도구 지침서. 제주특별자치도교육청·교육과학기술부.
- 임의도, 고종렬, 이혜선(1968). 학습준비도 검사 실시 요강: 초등학교 1학년용. 서울: 중앙교육연구소.
- 장명립, 신나리, 박수연(2006). 아동패널조사 모형개발 연구. 육아정책연구소 기본연구보고서, 1-101.
- 조지민, 동효관, 옥현진, 임혜미, 정혜경, 손수경, 배제성 외(2012). OECD 국제 학업성취도 비교 연구: PISA 2012 본검사 시행보고서. 한국교육과정평가원 연구보고 RRE 2012-3-1.
- 최윤경, 배윤진, 송신영, 임준범, 이예진, 김소아, 김신경(2014). 한국아동패널 2014. 육아정책연구소 연구보고 2014-33
- Arnett, J. (1989). Caregivers in day-care centers: Does training matter?. *Journal of Applied Developmental Psychology*, 10(4), 541-552.
- Arredondo, E. M., Elder, J. P., Ayala, G. X., Campbell, N., Baquero, B., & Duerksen, S. (2006). Is parenting style related to children's healthy eating and physical activity in Latino families?. *Health education research*, 21(6), 862-871.
- Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. In *Annual review of psychology* (pp. 471-507). Annual Reviews.
- Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, 1437-1472.
- Bornstein, M. H., Hahn, C. S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language*, 24(3), 267-304.
- Brinkman, S. A., Silburn, S., Lawrence, D., Goldfeld, S., Sayers, M., & Oberklaid, F. (2007). Investigating the validity of the Australian early development index. *Early Education and Development*, 18(3),

427-451.

- Bronfenbrenner, U. (1979). The ecology of human development: Experiments by design and nature.
- Byrd, R. S., Weitzman, M., & Auinger, P. (1997). Increased behavior problems associated with delayed school entry and delayed school progress. *Pediatrics*, 100(4), 654-661.
- Carlton, M. P., & Winsler, A. (1999). School readiness: The need for a paradigm shift. *School Psychology Review*, (28), 338-352.
- Chaplin, T. M., Cole, P. M., & Zahn-Waxler, C. (2005). Parental socialization of emotion expression: gender differences and relations to child adjustment. *Emotion*, 5(1), 80.
- Clearfield, M. W., & Nelson, N. M. (2006). Sex differences in mothers' speech and play behavior with 6-, 9-, and 14-month-old infants. *Sex Roles*, 54(1-2), 127-137.
- Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-Based Assessment of School Readiness and Early Reasoning. *Journal of Educational Psychology*, 106(3), 639 - 650.
- Dixon, P., Humble, S., & Counihan, C. (2015). *Handbook of International Development and Education*, Edward Elgar Publishing. [www.iPIPS.org](http://www.iPIPS.org)
- Dunn, R. S., Dunn, K. J., & Price, G. E. (1989). Learning style inventory (LSI). Price Systems, Incorporated (PO Box 1818, Lawrence 66044).
- Elder, G. H. (1998). The life course as developmental theory. *Child development*, 69(1), 1-12.
- Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales second edition (BAS II): administration and scoring manual*. London: NFER-Nelson.
- Enns, C. (2012). School readiness: Helping communities get children ready for school and schools ready for children. *Child Trends Research Brief*. (ERIC Document Reproduction Service No, ED 352171)

- Forget-Dubois, N., Lemelin, J. P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development, 18*(3), 405-426.
- Goldfeld, S., Sayers, M., Brinkman, S., Silburn, S., & Oberklaid, F. (2009). The process and policy challenges of adapting and implementing the Early Development Instrument in Australia. *Early Education and Development, 20*(6), 978-991.
- Guhn, M., Gadermann, A., & Zumbo, B. D. (2007). Does the EDI measure school readiness in the same way across different groups of children?. *Early Education and Development, 18*(3), 453-472.
- Hambleton, R. K. (1990). Item response theory: introduction and bibliography. *Psicothema, 2*(1), 97-107.
- Hambleton, R. K., & Zenisky, A. L. (2011). Translating and adapting tests for cross-cultural assessments. *Cross-cultural research methods in psychology, 46-70*.
- Harms, T., Clifford, M., & Cryer, D. (1998). *Early Childhood Rating Scale, Revised Edition (ECERS-R)*.
- Hart, B., Brinkman, S., & Blackmore, S. (2003). *How Well Are We Raising Our Children in the North Metropolitan Area?: Results of the Early Development Instrument*. Perth,, Western Australia: Population Health Program, North Metropolitan Health Service.
- Hogan, A. E., Scott, K. G., & Bauer, C. R. (1992). The Adaptive Social Behavior Inventory (ASBI): A new assessment of social competence in high-risk three-year-olds. *Journal of Psychoeducational Assessment, 10*(3), 230-239.
- Horsman, J. R., Furlong, W. J., Feeny, D. H., & Torrance, G. (2013). Health Utilities Index (Hui®): Population Reference Statistics. *Value in Health, 7*(16), A595.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender.

- Developmental psychology, 27(2), 236.
- Janus, M & Offord, D(2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science*. 39(1) 1-22.
- Janus, M., Brinkman, S. a., & Duku, E. K. (2011). Validity and Psychometric Properties of the Early Development Instrument in Canada, Australia, United States, and Jamaica. *Social Indicators Research*, 103(2), 283 - 297. <http://doi.org/10.1007/s11205-011-9846-1>
- Janus, M., Brinkman, S., Duku, E., Hertzman, C., Santos, R., Sayers, M., ... Walsh, C. (2007). The Early Development Instrument : A Population-based Measure for Communities. Guide.
- Janus, M., Sally, B., Eric, D., Clyde, H., Robert, S., Mary, S., Joanne, S., Cindy, W.(2011). *The Early Development Instrument: A Population-based Measure for Communities*. Offord: centre for child studies.
- Kagan, S. L.(1990). Readiness 2000: Rethinking rhetoric and responsibility. *Phi Delta Kappan*, Dec, 272-279.
- Kardanova, Ivanova, Merrell, Hawker, & Tymms (2014). The Role of the iPIPS Assessment in Providing High Quality Value Added Information on School and System Effectiveness within and between Countries, National Research University Higher School of Economics.
- Lang, F., Weiss, D., Stocker, A., & von Rosenblatt, B. (2007). The returns to cognitive abilities and personality traits in Germany. *Schmollers Jahrbuch: Journal of Applied Social Science Studies/Zeitschrift für Wirtschafts-und Sozialwissenschaften*, 127(1), 183-192.
- Learning Metrics Task Force. (2013). *Toward Universal Learning: Recommendations from the Learning Metrics Task Fore*. Montreal and Washington, D. C.: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution.
- Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). Test of preschool early literacy. Austin, TX: Pro-Ed.

- Lüftenegger, M., Schober, B., Van de Schoot, R., Wagner, P., Finsterwald, M., & Spiel, C. (2012). Lifelong learning as a goal - Do autonomy and self-regulation in school result in well prepared pupils?. *Learning and Instruction, 22*(1), 27-36.
- McWayne, C. M., Cheung, K., Wright, L. E. G., & Hahs-Vaughn, D. L. (2012). Patterns of school readiness among head start children: Meaningful within-group variability during the transition to kindergarten. *Journal of Educational Psychology, 104*(3), 862 - 878.
- Meisel, J. M. (Ed.). (1992). *The acquisition of verb placement* (Vol. 16). Springer Science & Business Media.
- Merrell, C. & Tymms, P. (2007). What children know and can do when they start school and how this varies between countries, *Journal of Early Childhood Research 5*(2), 115-134.
- Napiorkowski, B., Lester, B. M., Freier, M. C., Brunner, S., Dietz, L., Nadra, A., & Oh, W. (1996). Effects of in utero substance exposure on infant neurobehavior. *Pediatrics, 98*(1), 71-75.
- National Educational Panel Study. (2010). *The National Educational Panel Study (NEPS) IEA - DPC contributions*. In IEA General Assembly. Gaborone, Botswana.
- OECD, Development, Economic Analysis, & Statistics Division. (2003). *OECD Science, Technology and Industry Scoreboard 2003*. Canongate US.
- OECD (2015). *Starting Strong IV: Monitoring quality in early childhood education and care*. Paris: Author.
- Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From Neurons to Neighborhoods:: The Science of Early Childhood Development*. National Academies Press.
- Raikes, H., Alexander Pan, B., Luze, G., Tamis LeMonda, C. S., Brooks Gunn, J., Constantine, J., ... & Rodriguez, E. T. (2006). Mother - child bookreading in low income families: Correlates and outcomes during the first three years of life. *Child development, 77*(4), 924-953.

- Rao, N. (2012). The development and validation of contextually relevant scales to assess and monitor child development and learning in the East Asia Pacific Region. In Pre-Conference Seminar on Early Child Development: Recent research on brain development and assessment tools. Hong Kong.
- Rao, N., Sun, J., Ng, M., Becher, Y., Lee, D., Ip, P., & Bacon-Shone, J. (2015). Report on Technical Support for the Validation, Finalization and Adoption of the East Asia-Pacific Early Child Development Scales (EAP-ECDS), UNICEF.
- Sink, C., & Spencer, L. (2005). My Class Inventory-Short Form as an accountability tool for elementary school counselors to measure classroom climate. *Professional School Counseling*, 9(1), 37-48.
- Snow, C. E., & Van Hemel, S. B. (2008). Early childhood assessment: Why, what, and how? National Research Council of the National Academies Report.
- Stipek, D., & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Journal of Applied Developmental Psychology*, 22(2), 175-189.
- Sun, J., & Becher, Y. (2014). Early Childhood Development in the East Asia Pacific: Development & Findings from the East Asia Pacific - Early Child Development Scales (EAP-ECDS). In ARNEC Conference 2014. Manila.
- Sylva, K., Sammons, P., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (1999). *The Effective Provision of Pre-School Education (EPPE) Project: Intensive study of selected centres*. Retrieved from <http://eprints.ioe.ac.uk/5302/>
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., Smees, R., ... & Sadler, S. (2004). The effective provision of pre-school education (EPPE) project.
- Tymms, P., Jones, P. I., Albone, S., & Henderson, B. (2009). The first seven years at school. *Educational Assessment, Evaluation and Accountability*, 21, 67 - 80.

- Tymms, P., Merrell, C., & Jones, P. I. (2004). Using baseline assessment data to make international comparisons., 30(5), 673 - 689.
- Welsh, J. a., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, 102(1), 43 - 53.
- West, C. M., & Rose, S. (2000). Dating aggression among low income African American youth an examination of gender differences and antagonistic beliefs. *Violence Against Women*, 6(5), 470-494.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Woodcock-Johnson® III NU Tests of Achievement.
- Zill, N., & West, J. (2001). *Entering Kindergarten: A Portrait of American Children When They Begin School. Findings from the Condition of Education, 2000.*

www.Offord.com(2011). 2015. 3. 6일자 인출

www.iPIPS.org(2014). 2015. 3. 6일자 인출

www.worldbank.org(2014). 2015. 3. 8일자 인출

www.uis.unesco.org(2014). 2015. 4. 1일자 인출

http://panel.kicce.re.kr(2014). 2015. 4. 3일자 인출

www.arnec.net(2014). 2015. 6. 15일자 인출

## Abstract

# Issues and Tools of the International Assessment on Young Children's Learning and Development

Mugyeong Moon, Wonsoon Park, Junbum Lim

Assessment of young children has traditionally been one of under-researched areas because of a strong advocacy on individual difference of development in the early years as well as of concerns on negative aspects of early screening and labeling. As most of advanced countries have achieved universal ECEC services, they come to focus on enhancing quality of services, which, in turn, directs their attention to the linkage between quality and child development and learning outcomes. Utilizing results of international assessment on young children became an important agenda across countries in monitoring holistic development of children, identifying effects of, and informing ECEC policies.

In this context, the study reviewed currently available assessment tools, such as EDI (Early Development Instrument), iPIPS (International Performance Indicators for Primary School), EAP-ECDS (Asia Pacific-Early Childhood Development Scales), MELQO (Measuring Early Learning Quality and Outcomes) and analyzed major findings of studies with these tools.

A group of domestic and international experts provided their views on developing international assessment tools and strategies for international studies on assessing young children. A total of 45 domestic experts also participated in two rounds of in-depth survey.

The study concluded that there can be four different purposes of international assessment on children including assessing child's overall developmental, child outcomes of certain policies, for early intervention, and for international cooperation with developing countries. In accordance of

purposes of international assessment, varied are target groups, assessment domains and sub-domains, methods and procedures, timing, validation of tools, and utilization of assessment results. Some issues are discussed in assessing young children across countries, such as cultural contexts and languages. Further studies need to be conducted to develop and validate assessment tools both in regional and global contexts, reflecting upon key values and abilities required for children in the present and the future.

## 부 록

---

부록 1. 전문가 의견조사지(1차)

부록 2. 전문가 의견조사지(2차)



## 【국제비교를 위한 유아발달평가 기초 연구】 전문가 의견 조사 (1차)

안녕하십니까?

육아정책연구소에서는 2015년도 기본과제로 “국제비교를 위한 유아발달평가 기초 연구”를 수행하고 있습니다. 본 연구 수행의 일환으로 유아발달 국제비교의 목적, 발달평가 영역, 핵심 지표 및 방법과 절차, 평가도구 타당화 과정, 결과해석 및 정책적 활용 등에 대한 국내 전문가의 의견을 파악하고자 델파이 조사를 실시합니다.

각 문항에 자유롭게 의견을 개진해 주시고, 작성하신 의견 관련 참고문헌 출처 및 웹사이트 링크 등을 적어주시기 바랍니다.

개별 응답내용은 본 연구의 목적이외에는 다른 용도로 활용되지 않으며, 비밀이 보장됨을 말씀드립니다. 조사에 응해 주셔서 감사합니다.

2015년 9월

육아정책연구소장 우남희

성명		소속	
전공			

※ 각 문항에 대하여 형식이나 분량에 관계없이 전문가 의견을 자유롭게 기술해 주시기 바랍니다.

1. 유아발달 수준의 국제 비교는 왜 필요하며 어떤 목적으로 수행되어야 한다고 생각하십니까?

2. 유아발달 수준에 대한 국제비교 연구를 수행한다면 어떤 연령이 적합하다고 생각하십니까? 그 이유를 함께 적어주십시오

3. 유아발달 수준에 대한 국제비교 연구를 수행한다면 특별히 관심을 기울여야 하는 대상 집단 (예: 저소득층 남아, 거주지역별, 유아교육과 보육기관 미이용 아동, 성별 등)이 있습니까? 대상 집단과 그 이유를 함께 적어 주십시오.

4. 유아발달 국제비교 대상으로 어떤 국가가 적합하다고 생각하십니까?

5. 국제비교를 위한 유아발달 평가도구를 개발한다면, 포함되어야 할 평가 영역(및 세부 영역)들과 그 이유를 함께 적어주십시오. (예: ① 신체, 언어, 인지, 사회정서 영역 등 및 ② 각 영역별로 반드시 포함되어야 할 핵심적인 세부/하위 영역 3~4가지)

평가영역	세부영역	포함되어야 하는 이유 (기존 평가 도구에서 누락된 영역 및 세부 영역을 기입시 이유를 간단히 적어주십시오)

\* 칸이 더 필요하신 경우 추가하시기 바랍니다.

6. 국제비교를 위하여 기존에 개발된 유아발달 평가도구 중 활용가능한 도구가 있다면, 그 이유와 함께 추천해 주십시오.

7. 국제비교를 위한 유아발달 평가도구를 개발함에 있어서 우선적으로 고려해야 할 사항은 무엇입니까? 중요한 순서대로 적어주시고 그 이유를 적어주십시오.

(예: 정책적 활용도, 평가도구의 질- 타당도, 신뢰도, 국제비교 가능성, 문화적 다양성/차이 반영, 현지화 용이성, 비용, 검사 시간, 유아와 교사에 주는 부담정도 등)

순 위	고려 사항	이 유
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

8. 유아발달 평가도구에 대한 국제수준의 타당화 과정에서 고려해야 할 사항들에 대해 적어주십시오.

6-1	표본 크기	
6-2	국가별 문항 수정의 정도 (전체 문항 중 % 또는 공통문항 비율 등)	
6-3	적정 총 문항 수(의 범위) (총체적 발달평가도구의 경우)	
6-4	총체적 발달평가 도구와 특정영역 평가도구의 차이	

6-5	사회정서영역 평가의 타당도 확보 또는 향상 방안	
6-6	평가자(교사, 부모, 전문검사자)	
6-7	평가 시기 및 시행 횟수	
6-8	평가(검사)환경의 표준화 확보	
6-9	기타 (적어주세요)	

9. 유아발달 수준에 대한 국제비교 결과를 국가정책의 성과 평가 등에 활용하기 위하여 고려하여야 할 점은 무엇이라고 생각하십니까? (특정한 정책성과 모니터링, 학교준비도 점검 등에 활용시)

10. 국제비교를 위한 유아발달 평가도구 개발에 대한 추가적인 제안 사항을 적어주십시오(예: 아시아권 또는 우리나라 유아발달의 특성을 반영하는 평가지표, 유아발달 국제비교를 위한 국가적 차원의 지원 등)

11. 작성하신 의견과 관련된 참고문헌 및 웹사이트 링크 등을 적어주십시오.

## 【국제비교를 위한 유아발달평가 기초 연구】

### 전문가 의견 조사 (2차)

안녕하십니까?

육아정책연구소에서는 2015년도 기본과제로 “국제비교를 위한 유아발달평가 기초 연구”를 수행하고 있습니다. 본 연구 수행의 일환으로 유아발달 국제비교의 목적, 발달평가 영역, 핵심 지표 및 방법과 절차, 평가도구 타당화 과정, 결과해석 및 정책적 활용 등에 대한 국내 전문가의 의견을 파악하고자 2차 델파이조사를 실시합니다.

작성하신 내용 중 관련 참고문헌이 있으시면, 제일 마지막 문항에 적어 주시기 바랍니다. 개별 응답내용은 본 연구의 목적이외에는 다른 용도로 활용되지 않으며, 비밀이 보장됨을 말씀드립니다. 조사에 응해 주셔서 감사합니다.

2015년 10월

육아정책연구소장 우남희

성명		소속	
전공			

※ 다음의 표의 문항에 해당 번호를 기입해 주시고, 자유의견을 작성해 주십시오.

- 1) 문항 1에 기술된 유아발달 국제비교 목적의 우선순위를 1, 2, 3, 4로 표시하시고, 각 평가 목적에 따라 문항 2[ 11번까지 ] 응답해 주십시오.
- 2) 보기에 제시되지 않은 다른 의견이 있으시거나 설명, 보완 등이 필요하신 경우 자유의견에 작성해 주십시오.
- 3) 문항 2, 3, 4, 8, 10은 중복응답 가능합니다.

1. 유아발달 국제비교 목적의 우선순위(1, 2, 3, 4)를 기입해 주십시오.	2. 적합한 연령대	3. 비교대상 국가	4. 특정한 집단 설정
① 만1세 미만 ② 만1세 ③ 만2세 ④ 만3세 ⑤ 만4세	⑥ 만5세 이상 ⑦ 기타	① 개발도상국 중심의 아시아권역 ② OECD 중심의 선진국들 ③ 사회문화가 비슷한 동북아시아권역 ④ 북유럽권역 ⑤ 미주권역	① 유아교육 및 보육기관 미이용 집단 ② 저소득층 집단 및 사회적 배려 대상 ③ 장애유아 집단 ④ 다문화 가정 유아 ⑤ 지역적 분류 집단 ⑥ 성별
국제적 수준에서 유아의 발달과 학습 성과 비교 (교육과정 등의 input에 대한 성과)	번호: <자유의견>	번호: <자유의견>	번호: <자유의견>
국제적 수준에서 우리나라 유아 발달수준의 현 위치 점검	번호: <자유의견>	번호: <자유의견>	번호: <자유의견>
국제적 수준에서의 유아 발달수준을 비교 분석하여 적극적인 조기 중재 방안 마련 (early intervention)	번호: <자유의견>	번호: <자유의견>	번호: <자유의견>
개발도상국과의 교류협력을 위한 정보제공 및 지원	번호: <자유의견>	번호: <자유의견>	번호: <자유의견>

<p>유아발달 국제비교의 목적</p>	<p>4. 해당목적에서의 필요한 표본 수</p> <p>① 1,000명 이하      ⑥ 10,000명 이상                  ② 1,000명~2,000명      ⑦ 전수조사                  ③ 2,000명~3,000명                  ④ 3,000명~5,000명                  ⑤ 5,000명~10,000명</p> <p>빈호:</p>	<p>5. 국가별 문항 수정 가능 비율은</p> <p>진제 문항 중:</p> <p>① 10% 이하                  ② 20% 이하                  ③ 30% 이하                  ④ 50% 이하</p> <p>빈호:</p>	<p>6. 적정 문항 수 (총 검사 문항 수)</p> <p>① 50문항 이하      ⑥ 150문항 이상                  ② 50~75문항 미만                  ③ 75~100문항 미만                  ④ 100~125문항 미만                  ⑤ 125~150문항 미만</p> <p>빈호:</p>
<p>국제적 수준에서 유아의 발달과 학습 성과 비교 (교육과정 등의 input에 대한 성과)</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>
<p>국제적 수준에서 우리나라 유아 발달수준의 현 위치 점검</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>
<p>국제적 수준에서의 유아 발달수준을 비교 분석하여 <b>적극적인 조기 중재 방안 마련</b>(early intervention)</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>
<p><b>개발도상국과의 교류협력</b>을 위한 정보제공 및 지원</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>	<p>&lt;자유의견&gt;</p> <p>빈호:</p>

유아 발달 국제 비교의 목적	8. 적절한 평가 시행 횟수 및 시기	9. 포함된 평가 시행 횟수 및 시기	10. 포함되어야 할 평가 영역 (대영역)을 모두 표시주세요.
<p>유아 발달 국제 비교의 목적</p> <p>국가적 수준에서 유아의 발달과 학습 성과 비교 (교육과정 등의 input에 대한 성과)</p>	<p>① 교사 ② 부모 ③ 유아(또래) ④ 전문평가자</p> <p>번호: &lt;자유의견&gt;</p>	<p>① 학기당 1회 ② 연 1회 ③ 2년에 1회 ④ 4년에 1회</p> <p>번호: &lt;자유의견&gt;</p>	<p>① 인지 발달 ② 언어 발달 ③ 사회 정서 발달 ④ 신체 발달 ⑤ 기타</p> <p>번호: &lt;자유의견&gt;</p>
<p>국가적 수준에서 우리나라 유아 발달 수준의 현 위치 점검</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>
<p>국가적 수준에서 유아 발달 수준을 비교 분석하여 <b>적극적인 조기 중재 방안 마련</b>(early intervention)</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>
<p>개발도상국과의 교류 협력을 위한 정보 제공 및 지원</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>	<p>번호: &lt;자유의견&gt;</p>

11. 국제비교를 위한 유아발달 평가도구 개발시에 포함되어야 할 영역별 하위영역의 중요도에 대하여 표시해 주십시오.

대영역	세부영역	중요도				
		전혀 중요하지 않음	중요하 지 않음	보통	중요	매우 중요
인지발달	수 감각 및 셈하기	①	②	③	④	⑤
	분류 및 공간인지능력	①	②	③	④	⑤
	기억 및 모방	①	②	③	④	⑤
	기초문제해결력	①	②	③	④	⑤
	기초추론능력	①	②	③	④	⑤
	학교준비도	①	②	③	④	⑤
<자유 의견> 인지발달 내 추가해야할 하위영역						
언어발달	말하기	①	②	③	④	⑤
	듣기	①	②	③	④	⑤
	읽기	①	②	③	④	⑤
	쓰기	①	②	③	④	⑤
	의사소통능력	①	②	③	④	⑤
	어휘력	①	②	③	④	⑤
	문자읽어 기초인식	①	②	③	④	⑤
<자유 의견> 언어발달 내 추가해야할 하위영역						
사회정서발달	집단생활능력	①	②	③	④	⑤
	타인감정 파악 및 감정조절	①	②	③	④	⑤
	또래인지(관계)	①	②	③	④	⑤
	부모관계	①	②	③	④	⑤
	기초시민의식	①	②	③	④	⑤
	문화다양성	①	②	③	④	⑤
	상호존중	①	②	③	④	⑤
	친애유구, 증인유구, 성취유구, 불안/우울	①	②	③	④	⑤
<자유 의견> 사회정서발달 내 추가해야할 하위영역						
신체발달	대근육	①	②	③	④	⑤
	소근육	①	②	③	④	⑤
	지각 및 기본운동능력	①	②	③	④	⑤
	기초위생습관	①	②	③	④	⑤
	식사횟수와 정도	①	②	③	④	⑤
키, 몸무게	①	②	③	④	⑤	
<자유 의견> 신체발달 내 추가해야할 하위영역						

[이외에 추가해야할 대영역과 하위영역에 대하여 적어주십시오]

12. 국제비교를 위한 유아발달 평가도구를 개발함에 있어서 고려해야 할 사항 중 각 항목의 중요도를 체크해주시시오.

구분	중요도				
	전혀 중요하지 않음	중요하지 않음	보통	중요	매우 중요
1. 정책적 활용도	①	②	③	④	⑤
2. 평가도구의 질(타당도 및 신뢰도)	①	②	③	④	⑤
3. 국제비교 가능성	①	②	③	④	⑤
4. 문화적 다양성 및 차이 반영 정도	①	②	③	④	⑤
5. 현지화의 용이성	①	②	③	④	⑤
6. 비용	①	②	③	④	⑤
7. 검사 길이(검사 시간)	①	②	③	④	⑤
8. 참여자의 심리적 부담	①	②	③	④	⑤
9. 기존 평가도구와의 중복성	①	②	③	④	⑤
10. 배경변인 설정	①	②	③	④	⑤
11. 표집의 적절성	①	②	③	④	⑤

<자유 의견>

13. 응답하신 내용과 관련된 참고 문헌이 있으시면 적어주시시오.

연구보고 2015-14

---

국제비교를 위한 유아발달평가 기초연구

---

**발행일** 2015년 11월

**발행인** 우남희

**발행처** 육아정책연구소

**주 소** 서울시 서초구 남부순환로 2558 외교센터 빌딩 3층, 4층

전화: 02) 398-7700

팩스: 02) 730-3313

<http://www.kicce.re.kr>

**인쇄처** 현대아트컴

---

보고서 내용의 무단 복제를 금함.

ISBN 979-11-85941-39-4 93330