

## 어린이집 프로그램 관찰척도의 개발과 타당화\*

Development and Validation of an Observation Scale for Day Care Programs

이 은 해\*\*

Rhee, Unhai

최 혜 영\*\*\*

Choi, Hye Yeong

송 혜 린\*\*\*

Song, Hye Rin

신 혜 영\*\*\*

Shin, Hye Young

### Abstract

This study aimed to develop an observation scale for assessing the quality of day care programs and to examine reliability and validity of the scale. Instruments included 3 self-assessment scales, the Evaluation Scale for Early Childhood Programs and an observation scale completed by 8 trained observers in 89 day care sites. Item analysis indicated that most items were acceptable in terms of item discrimination and item response distribution. Cronbach  $\alpha$ 's and inter-observer reliability were relatively high. Validity of the scale was evidenced by correlations between the observation scale and the criterion measure and inter-correlations among sub-areas of the observation scale. Scores on the observation scale were validated by quality ratings from self-assessment scores.

**Key Words** : 관찰척도(observation scale), 보육 프로그램 평가(assessment for day care programs), 척도의 개발(development of a scale), 척도의 타당화(validation of a scale)

\* 접수 2003년 2월 28일, 채택 2003년 4월 18일

\* 이 논문은 2002년도 삼성복지재단의 연구비 지원에 의해 이루어짐.

\*\* 교신저자: 연세대학교 생활과학대학 아동·가족 전공 교수, E-mail : uhrhee@yonsei. ac. kr

\*\*\* 연세대학교 대학원 아동·가족 전공 박사과정

## I. 서론

정부의 영유아 보육법 제정(1991)과 보육시설 확충 3개년 계획(1995)의 실시 후 우리나라의 보육시설과 보육 아동 수는 급격하게 증가하였다. 그러나 이러한 양적 팽창에 따라 질적인 수준도 함께 수반되고 있는가에 대해서는 의문이 제기되고 있다. 보육시설이 아직도 수요에 미치지 못하고 있어 계속 증가하는 추세에 있으므로 보육시설의 질을 평가하기 위한 다각적인 노력이 필수적으로 요구된다.

보육시설의 질에 관심을 두어야 하는 이유는 무엇보다도 영유아가 어린 연령부터 하루 중 깨어있는 대부분의 시간을 부모 아닌 대리 양육자와 함께 낯선 환경에서 보낸다는 것에 근거하고 있다. 보육시설의 경험은 영유아의 발달에 도움이 될 수도 있지만 때로는 발달을 저해하지 않는지 우려하는 시각도 있기 때문이다. 최근까지 이루어진 연구들을 종합해 보면, 질적으로 우수한 보육시설의 경험은 최소한 영유아의 발달에 지장을 초래하지는 않을 것이라는 조심스런 전망이다(Feisner-Feinberg & Burchinal, 1997; Phillips & Howes, 1987). 따라서 보육시설에서의 경험이 영유아의 사회·정서, 인지 발달에 도움이 될 수 있도록 보육시설의 질을 적합한 수준으로 유지하고 평가하는 것이 주요 과제이다.

보육시설의 평가는 최소 다음의 세 가지를 목표로 하고 있다(Friedman & Amadeo, 1999). 첫째, 보육시설의 인가 또는 허가를 위한 평가이다. 보육을 담당하는 행정기관은 시설 설립에 관한 인·허가 기준을 정하고 이에 따라 규제하고 관리할 책임이 있다. 일반적으로 시설의 공간(면적)과 설비, 영유아 집단크기, 교사 대 영유아 비율, 교사의 자격 등에 관한 기준이 이에 속한다. 이러한 요인은 보육의 질을 나타내는 구조적인

변인으로써 최소한의 질적 기준을 의미한다.

둘째, 전문가적 인준을 위한 평가이다. 대체로 전문가 집단에 의해 인준평가가 이루어지며, 질적 수준을 제고시키기 위해 기준을 높이 설정하고 있다. 전문가적 인준 평가의 종류로는 보육시설에 대한 것과 개인에 대한 인준으로 구분할 수 있다. 보육시설에 대한 인준평가는 NAEYC(1998)의 인준절차가 대표적이라 할 수 있으며, 개인에 대한 인준은 유아교육 전문가 심의회(Council for Early Childhood Professional Recognition)의 아동발달사 자격 부여에서 그 예를 찾아볼 수 있다.

셋째, 영유아 발달에 미치는 효과를 연구하기 위한 평가이다. 보육시설의 전반적인 질이 어떻게 영유아 발달에 영향을 미치는가 뿐만 아니라, 질을 보다 세분화한 구체적 변인들이 발달의 어느 영역에 더 많이 영향력을 갖는지, 그리고 가족 배경 변인과는 어떻게 상호작용을 하여 영유아에게 영향을 주는지에 이르기까지 다양하게 연구되고 있다. 이러한 효과를 단기 또는 장기적으로 추적함으로써 영유아 발달에 지속적으로 어떤 질적 요소가 더 영향력이 있는지를 추구하고 있다. 최근의 이 분야 연구는 시설의 인허가 기준인 구조적 변인보다 영유아가 직접 경험하는 측면인 과정적 변인에 초점이 모아지고 있다. 과정적 변인으로는 영유아와 교사간의 상호작용, 교사의 민감성과 반응성, 개별화된 배려, 교육자료의 발달적 적합성과 이용 가능성 등이 포함된다.

보육시설의 평가를 위해서는 이상의 세 가지 목표와 관련된 다양한 측정도구들을 필요로 한다. 첫 번째 범주에 속한 기준들은 객관적으로 측정이 가능하며, 행정적으로 규제되는 최소한

의 기준이므로 우리나라 같이 전국적인 기준이 적용되는 곳에서는 특별하게 전문적으로 고안된 측정도구를 필요로 하지 않는다. 그러나 인준을 위한 평가나 연구를 위한 평가에서는 보육의 질적 수준을 전반적으로 그리고 객관적으로 평가할 수 있는 양호한 측정도구가 필수적이다. 이 두 범주에 속하는 측정도구들은 대체로 시설의 원장이나 교사가 자체적으로 평가하기 위한 도구와 외부 관찰에 의한 도구로 나누어 볼 수 있다. 시설의 원장이나 교사가 자체적으로 평가할 때 프로그램의 질적 향상을 위해 내부 관점을 집약해 점검한다는 측면에서 이롭다. 그러나 자체 평가의 단점은 주관적 평가이므로 평가결과가 일반적으로 편파되기 쉽고, 차후 객관적 검증을 필요로 한다는 문제점이 있다. 실제 국내에서 보고된 바에 따르면, 담당 교사들의 자체평가 점수가 관찰자의 평가 점수보다 평균적으로 높은 경향을 보이고 있다(이은혜, 2002; 임재택, 1983). 반면 외부 관찰자에 의한 평가는 객관성의 증진 측면에서 우수하며, 관찰과 함께 문서를 점검하거나 질문을 함으로써 자료 수집의 폭을 넓힐 수 있어서 이롭다. 그러나 관찰자의 훈련과 현장 파견 등에 시간과 노력, 경비가 많이 들고 관찰 당시의 상황에 자료 수집이 제한되는 점을 유의해야 한다. 관찰에 의한 평가척도들은 이러한 실시상의 제한점에도 불구하고 인준평가의 최종단계에 적용되고 대부분의 평가 연구에서 광범위하게 사용되는 이유는 무엇보다도 결과의 객관성에 대한 신뢰라고 할 수 있다.

우리나라에서 개발된 교육 또는 보육 프로그램의 평가도구 중에서 관찰척도로 사용할 수 있는 종류는 극히 제한되어 있으며, 이들 중 대부분은 외국의 도구를 번역하여 적용한 것이다. 대표적인 관찰도구로 Harms와 Clifford(1980)의

ECERS(Early Childhood Environment Rating Scale)에 기초한 유아교육환경 평정척도(임재택, 1984)를 비롯하여, 최근 ECERS의 개정판(Harms, Clifford, & Cryer, 1998)을 번역하여 국내 적용 가능성을 조사한 예비연구(김윤지, 2003)가 있으며 ITERS(Infant/Toddler Environment Rating Scale)를 번안하여 영아/걸음마기 아동환경 평정척도 타당화 연구(현은강·이완정, 1997; 이완정·현은강, 1998)가 이루어진 바 있다. 관찰에 문서 점검 및 면접을 병행하여 자료 수집을 할 수 있는 도구로는 임재택·조희숙·황해익(1996)에 의한 영유아 보육 프로그램의 진단·평가척도가 개발된 바 있으며, 유아교육 프로그램 평가척도(강숙현, 1994)는 Abbott-Shim과 Sibley(1987)의 APECP (Assessment Profile for Early Childhood Programs)에 기초하고 있다. 이처럼 보육프로그램의 질을 평가할 수 있도록 국내에서 개발된 관찰척도는 소수에 불과하며 개발된 평가도구들도 대부분 외국의 도구에 기초하거나 유치원을 중심으로 한 유아교육에 초점이 맞추어져 있다. 따라서 우리나라 보육시설에서 운영되는 보육프로그램의 질을 전반적으로 측정하여 프로그램의 효과 또는 인준평가에 활용할 수 있는 양호한 관찰척도의 개발이 무엇보다도 필요하다고 하겠다.

본 연구는 우리나라 보육시설의 프로그램을 평가하는데 기여할 수 있는 다각적인 관점의 평가도구 세트 개발에 관련된 연구(이은혜·송혜린·신혜영, 2001)의 일환으로, 「어린이집 프로그램 관찰척도」의 개발과 척도의 타당화를 목적으로 하였다. 일차적으로 평가도구 세트 중에서 원장과 교사가 자체평가에 활용할 수 있는 3종류의 평가척도가 개발되었고(이은혜·송혜린·신혜영·최혜영, 2002), 본 관찰척도는 자체평가 결과를 보완 또는 확인하고 보다 객

관적인 현장평가를 위해서 고안된 것이다. 특히 본 척도는 우리나라 보육시설에 종사하는 전문가와 관련 학계 전문가의 견해를 반영하고, 우리나라 보건복지부의 시설 설비 기준에 기초하였으며, 유아반과 함께 영아반 프로그램 내용이 포함되었다는 점에서 기존의 다른 도구들과 구별될 수 있을 것이다.

이상에서 서술된 바와 같이 본 연구는 어린이집 프로그램 관찰척도의 개발과 타당화를 목

적으로 하며, 다음의 구체적인 연구문제가 제기되었다.

첫째, 어린이집 프로그램 관찰척도의 문항 양호도는 어떠한가?

둘째, 어린이집 프로그램 관찰척도의 신뢰도는 어떠한가?

셋째, 어린이집 프로그램 관찰척도의 타당도는 어떠한가?

## II. 연구방법

### 1. 연구대상

연구의 대상기관은 전국의 삼성 어린이집 30개 기관과 각 어린이집이 그 지역에서 추천한 국·공립 어린이집 34개 기관 그리고 해당 지역에서 연구 참여에 동의한 25개 기관 등 총 89개 어린이집을 대상으로 하였다.

연구대상 어린이집의 지역별 분포를 보면, 서울특별시 34개 기관(38.2%), 인천 및 경기도가 19개 기관(21.4%)으로 전체의 60%를 이루고 있으며, 부산, 대구, 울산과 경상도 20개 기관(22.5%), 광주와 전라도 10개 기관(11.3%), 대전과 충청도 6개 기관(6.8%)이 포함되었다.

대상 기관의 특성과 관련하여 학급수와 원아수 분포를 <표 1>에서 살펴보면, 5학급에서 8학급의 기관이 전체의 62%를 이루고 있었으며, 영아반은 2학급 이하가 58%, 유아반은 3~4학급의 기관이 65%였다. 혼합연령의 학급을 운영하지 않는 기관이 81%로 대부분이었고, 혼합연령을 운영하더라도 1~2학급에 그치는 것으로 나타났다.

<표 1> 연구 대상 기관의 학급수와 원아수 분포

구 분	학급수 분포		원아수 분포	
	학 급	기관수(%)	원 아	기관수(%)
전 체	4학급 이하	23(25.8)	100이하	53(59.6)
	5~8학급	55(61.8)	101이상	36(40.4)
	9학급 이상	11(12.4)		
영아반	2학급 이하	52(58.4)	20이하	42(47.2)
	3~4학급	27(30.3)	21~40	39(43.8)
	5학급 이상	10(11.2)	41이상	8(9.0)
유아반	2학급 이하	20(22.5)	70이하	49(55.1)
	3~4학급	58(65.2)	71~100	26(29.2)
	5학급 이상	11(12.4)	101이상	14(15.7)
혼 합 연령반	없음	72(80.9)	없음	72(80.9)
	1~2학급	16(18.0)	14~70	17(19.1)
	3학급	1(1.1)		

### 2. 연구도구

본 연구의 도구는 예비조사를 거쳐 개발된 「어린이집 프로그램 관찰척도」와 함께, 타당도 검증을 위해서 본 척도의 모체라 할 수 있는

자체평가용 어린이집 프로그램 평가척도 3종류와 외적 준거인 ‘유아교육 프로그램 평가척도’(이은혜·이기숙, 1996)가 선정되었다. 연구에 사용된 도구를 각각 서술하면 다음과 같다.

1) 어린이집 프로그램 관찰척도

어린이집 프로그램 관찰척도는 보육시설의 환경요소와 영유아가 갖는 경험요소를 중심으로 5개 평가 영역으로 구성되어 총 73개 문항으로 이루어져 있다. 각 문항의 평가는 어린이집 하루의 주요 일과를 관찰하여 3, 2, 1점 중의 하나로 서술된 기준에 따라 평정하며, 현장에서 관찰이 불가능한 경우 ‘관찰 못함(N/O)’으로 판단한다. 평가 영역의 내용과 영역별 문항수는 다음과 같으며, <표 2>에 정리되어 있다.

(1) 물리적 환경

실내 공간의 크기, 공유 공간 내 시설·설비의 적합성과 안전성, 실외공간 구성과 설비 등 기관 전체의 물리적 환경을 내용으로 11문항으로 구성된다.

(2) 건강·안전·영양

실내 시설의 청결과 위생, 급간식 관리, 그리고 영유아의 건강과 안전 관리 등을 내용으로 11문항으로 구성된다.

(3) 학습환경

영아 보육실과 유아 보육실의 발달적으로 적합한 영역 구성, 그리고 각 영역의 설비와 다양한 교재·교구의 구비를 내용으로 영아반 8문항, 유아반 11문항, 모두 19문항으로 구성된다.

(4) 교육경험 및 활동

발달적으로 적합한 놀이와 학습활동의 진행 정도, 활동간의 균형과 선택 가능성 등을 내용으로 영아반 3문항, 유아반 11문항, 모두 14문항으로 구성된다.

(5) 교사-영유아 상호작용

교사와 영유아간 언어적 상호작용의 적절성, 교사의 수용적이고 반응적인 태도, 그리고 긍정적인 행동지도 등을 내용으로 영아반 8문항, 유아반 10문항, 모두 18문항으로 구성된다.

<표 2> 어린이집 프로그램 관찰척도의 영역 및 문항수

영역 및 하위영역	하위영역영역 문항수	영역 문항수
<b>1. 물리적 환경</b>		<b>11</b>
실내 공간과 설비	5	
실내외 시설과 설비의 안전	3	
실외 공간과 설비	3	
<b>2. 건강·안전·영양</b>		<b>11</b>
실내 시설의 청결과 위생	4	
급간식의 관리	2	
유아의 건강과 안전	5	
<b>3. 학습환경</b>		<b>19</b>
영아반: 보육실의 구성	1	
놀이감의 구비정도	7	
유아반: 보육실의 구성	4	
놀이감과 교재의 구비정도	7	
<b>4. 교육경험 및 활동</b>		<b>14</b>
영아반: 놀이 경험과 활동	3	
유아반: 활동의 균형과 선택가능성	4	
다양한 놀이와 활동	7	
<b>5. 교사-영유아 상호작용</b>		<b>18</b>
영아반: 언어적 상호작용	3	
교사의 수용적, 반응적 태도	2	
긍정적 행동지도	3	
유아반: 교사의 수용적, 반응적 태도	3	
언어적 상호작용	2	
긍정적 행동지도	3	
놀이참여와 확장 관련 상호작용	2	
<b>전 체</b>		<b>73</b>

## 2) 어린이집 운영관리 평가척도

운영관리 평가척도는 어린이집의 원장이나 원감이 어린이집 운영의 여러 측면에 대해 자체적으로 평가할 수 있는 척도이다. 평가영역과 영역별 문항수(괄호 안의 숫자)는 물리적 환경(13), 건강·안전·영양(12), 교직원(10), 가정 및 지역사회 연계(7), 전반적인 교육과정 운영(7)의 5개 영역 49문항으로 구성되어 있다.

## 3) 어린이집 영아반 평가척도

영아반 평가척도는 1세~3세 미만 영아를 담당하는 교사가 학급의 실제에 대해 자체적으로 평가할 수 있는 도구이다. 평가영역은 놀이 환경(11), 놀이활동(8), 일상적 양육(8), 건강과 안전(11), 교사-영아 상호작용(9)의 5개 영역 47문항으로 구성되어 있다.

## 4) 어린이집 유아반 평가척도

유아반 평가척도는 3세~5세 유아를 담당하는 교사가 학급의 실제에 대해 자체적으로 평가할 수 있는 도구이다. 평가영역은 학습환경(17), 교육과정 계획(7), 교육과정 운영(17), 건강과 안전(12), 교사-유아 상호작용(11)의 5개 영역 64문항으로 이루어져 있다.

## 5) 유아교육 프로그램 평가척도

외적 준거로 사용된 유아교육 프로그램 평가척도는 간편형이 적용되었으며, 이 척도는 3세~5세 유아를 위한 유아교육기관, 주로 유치원을 대상으로 한 자기평가도구이다. 평가영역은 일과계획 및 상호작용(12), 교육과정(20), 물리적 환경(16), 건강·안전·영양(13), 운영관리(19)의 5개 영역 80문항으로 이루어져 있다.

이상에서 관찰척도를 제외한 4개 척도들은

모두 3점 척도로써 3, 2, 1점에 해당되는 평가기준이 서술문으로 제시되고 자신의 실제 현장에 가장 가까운 서술문을 택해 응답하도록 되어 있다.

## 3. 연구의 절차

### 1) 관찰척도의 영역 및 문항 선정

어린이집 프로그램 관찰척도는 1차적으로 제작된 자체 평가용 운영관리 평가척도, 영아반 평가척도, 유아반 평가척도(이은혜 외, 2002)를 모체로 하여 개발되었다. 관찰척도의 문항은 자체 평가척도의 예비실시와 문항분석 결과에 기초하여, 첫째 관찰 가능성이 있고, 둘째 양호한 문항으로 선정하되 보육 시설의 환경 요소와 영유아가 갖는 경험 요소를 중심으로 5개 영역을 구성하였다.

이에 따라 예비 실시용 ‘어린이집 프로그램 관찰척도’는 프로그램의 전반적 질을 평가하는 5개 영역을 선정하고, 각 영역별로 11~20개 문항을 이루어 총 80문항으로 제작되었다. 평가영역은 ‘물리적 환경(11문항)’, ‘건강·안전·영양(15문항)’, ‘학습환경(20문항)’, ‘교육 경험 및 활동(19문항)’, 그리고 ‘교사-영유아 상호작용(15문항)’이었다. 예비척도 문항에 대한 관찰 평정은 3단계 기술평정으로써 질적 수준에 따라 3점, 2점, 1점으로 표기하도록 하였다. 관찰척도는 관찰자의 현장기록을 용이하게 돕기 위해 관찰척도 기록지를 별도로 제작하였으며, 관찰척도 기록지는 각 문항과 기준의 내용을 핵심만 요약하여 표로 만든 것이다. 또한 제한된 시간 내에 관찰이 안되는 문항을 위해 각 문항마다 관찰 안됨(N/O) 항목을 추가시키고, 기준 외의 관찰사항을 자유롭게 기록할 수 있는 여백을 포함시켰다.

## 2) 관찰척도의 예비조사 및 결과

관찰척도의 예비조사는 2001년 8월 21일부터 9월 7일 사이에 19개 어린이집을 임의 표집하여 이루어졌다. 이들 기관은 자체 평가용 척도의 자료 수집 대상인 87개 기관 중에서 삼성 어린이집 16곳, 국공립 어린이집 3곳을 표집하였고, 지역별로는 서울·경인 지역이 13곳, 중부지역 2곳, 영호남 지역 4곳이었다. 현장관찰은 유아교육 현장경험이 있는 아동학 석사 4명에 의해 훈련과정을 거쳐서 실시되었다. 관찰자는 각 기관을 오전 9시경부터 오후 1시경까지 방문하여 기관 전체의 물리적 환경과 건강·안전·영양에 관해 관찰하고, 기관의 영아반 1학급과 유아반 1학급을 무선 표집하여 보육실 내의 학습환경, 교육활동, 교사와의 상호작용을 관찰하였다. 관찰한 내용은 관찰척도 기록지에 3점 척도로 표기하고 관찰이 안되는 문항은 N/O에 의문이나 특기 사항은 여백에 기록하도록 하였다.

예비 조사를 통해 수집된 자료는, 먼저 기술적 자료를 위해 영역별 평균과 표준편차를 산출하고, 관찰척도의 신뢰도 검토를 위해 영역별 Cronbach  $\alpha$ 계수와 관찰자간 신뢰도를 산출하였으며, 문항 검토를 위해 문항별 반응 분포를 살펴보았다.

분석 결과로 첫째, 관찰척도의 영역별 문항 점수의 평균은 3점 척도에서 2.03부터 2.39로 나타났다. 영역별로 보면, 물리적 환경이 2.39, 학습환경과 교사-영유아 상호작용이 2.36으로 높은 편이었고, 그 다음 건강·안전·영양은 2.14, 그리고 교육경험 및 활동이 2.03으로 가장 낮았다.

둘째, 척도의 내적 합치도 지수인 Cronbach  $\alpha$ 계수와 관찰자간 신뢰도를 산출한 결과, 전체 척도에서  $\alpha$ 계수는 .85로 나타났고 영역별로는

.80에서 .91의 범위로 나타나 본 척도가 영역별로 일관된 내용을 평가하고 있다고 해석할 수 있다. 관찰자간 신뢰도는 2명이 독립적으로 관찰한 결과의 상관계수가 전체 척도에서 .79, 영역별로는 .70에서 .85로 나타나서 신뢰도 면에서 수용할만한 수준을 보여주었다.

셋째, 척도를 이루고 있는 전체 문항 각각에 대해 문항 평정 점수별 빈도, 문항 평균 및 표준편차, 그리고 문항 삭제에 따른  $\alpha$ 계수의 변화를 살펴보고, 문항 내용을 고려하여 삭제하거나 수정하였다. 결과적으로 영아반의 교육경험 및 활동에 관련된 문항들은 통합하였고 건강·영양·안전에서  $\alpha$ 를 낮추는 2개 문항이 삭제되었다. 반면 교사-영유아 상호작용에서는 영아반과 유아반 문항의 균형 및 내용의 중요도에 의해 3문항이 추가 포함되었다.

이상의 예비 조사 결과에 기초하여 어린이집 프로그램 관찰척도는 5개 영역의 73문항으로 완성되었다.

## 3) 관찰척도의 본조사

자료수집은 대상 어린이집 89개소에서 2002년 1월 10일에서 2월 8일 사이의 약 4주간에 걸쳐 이루어졌으며, 자체 평가척도를 우송함과 동시에 기관별로 협조 의뢰 공문을 발송하였다. 운영관리 평가척도는 해당 기관의 원장이나 원감에 의해 평가하도록 했으며, 영아반 평가척도와 유아반 평가척도는 원장이 해당 연령을 한 학급씩 임의로 선정하여 해당 학급의 교사가 스스로 평가하도록 한 후 반송용 우편으로 모두 수거하였다.

현장방문 관찰자는 예비조사에 참여한 관찰자 4명을 포함하여 모두 8명으로 현장경험이 있고 아동학 석·박사 과정생 또는 석사 후 연구원으로 구성되었다. 관찰자 훈련과 예비 관찰

을 거쳐 관찰자간 신뢰도 산출을 한 후, 필요하면 추가 관찰 훈련을 하였다. 본 관찰은 각 기관을 1명의 관찰자가 오전 9시부터 오후 1시경까지 방문하여 실시하였다. 현장 방문 시 준거척도인 유아교육 프로그램 평가척도 간편형을 배부하고 당일 관찰이 끝난 후 수거하도록 하였다.

본 조사에서 수집된 자료로는 기술적인 자료와 함께 문항 양호도와 척도의 신뢰도를 다시 검증하고 척도의 타당도 분석을 하였다. 문항 양호도는 문항 반응 분포와 문항 변별도로 검토하였으며, 척도의 신뢰도 검증을 위해서는

총점 및 하위 영역별 Cronbach  $\alpha$ 와 관찰자간 신뢰도를 산출하였다. 타당도 검증을 위해 구인 타당도와 공인 타당도에 관련된 자료가 분석되었다. 구인타당도 검증을 위해 첫째로 관찰척도의 결과가 원장 및 교사에 의한 자체평가를 얼마나 잘 반영하는지를 알아보고자 세 종류의 자체 평가척도 총점에 의해 상, 중, 하 집단을 나누고, 이 세 집단의 관찰척도 영역별 평균 점수 차이를 F검증하였다. 둘째로 어린이집 관찰척도의 영역간 상호상관을 분석하였다. 공인 타당도는 본 척도와 외적 준거간의 상관계수를 내어 검토하였다.

### Ⅲ. 연구 결과

어린이집 관찰척도의 문항 양호도, 신뢰도, 타당도 자료를 제시하기에 앞서 관찰척도 점수의 일반적 경향을 파악하기 위해 영역별로 문항점수를 합산한 영역 점수분포와 문항점수의 분포를 살펴보면 다음의 <표 3>과 같다.

<표 3> 어린이집 관찰척도의 영역별 점수 및 문항점수 분포

어린이집 관찰척도	영역 점수		문항 점수
	범위	평균(표준편차)	평균(표준편차)
① 물리적 환경	11-33	25.01 ( 4.87)	2.27 (.44)
② 건강·안전·영양	11-33	24.65 ( 3.70)	2.24 (.34)
③ 학습환경	19-57	39.82 ( 9.32)	2.10 (.49)
④ 교육경험 및 활동	14-42	27.37 ( 5.36)	1.96 (.38)
⑤ 교사·영유아 상호작용	18-54	40.93 ( 7.51)	2.27 (.42)
총 점	73-219	157.79 (26.91)	2.16 (.37)

<표 3>에서 보는 바와 같이 영역별 문항 점수의 평균은 3점 척도에서 1.96~2.27로 나타났다. 하위 영역 중에서는 ‘교육경험 및 활동’의 평균이 1.96으로 가장 낮고 ‘물리적 환경’(2.27)과 ‘교사·영유아 상호작용’(2.27)은 다른 영역에 비해 높게 나타났다. 영역 점수의 평균은 개별 기관의 관찰점수를 산출하여 해석할 때 하나의 잠정적인 기준으로 해석할 수 있을 것이다.

#### 1. 관찰척도의 문항 양호도

어린이집 관찰척도의 문항 양호도를 분석하기 위해 총 73개 문항에 대한 문항 반응분포, 문항 변별도, 그리고 문항 삭제 시 Cronbach  $\alpha$  등을 산출하였다. 문항 반응분포는 각 문항의 기준점수별 반응빈도(%), 평균과 표준편차를 산출하여 지나치게 편포된 문항과 문항에 90% 이상이 3점에 반응하는 천정효과(ceiling effect)



여부를 검토하였다.

문항 변별도는 각 문항이 프로그램 질의 차이를 바르게 변별하는지 알아보기 위해서 척도의 총점에 의해 대상기관을 상, 하 집단으로 이등분하고 상, 하 집단 간 문항반응의  $x^2$  와 Cramer V를 산출하였다. 관찰척도의 73개 문항에 대한 분석 결과는 <표 4>에 제시되어 있다.

<표 4> 어린이집 관찰척도의 문항분석 결과

어린이집 관찰척도	문항반응 점수범위	문항수	천정효과 문항변별도	
			90%	$x^2 = N.S.$
① 물리적 환경	1.80~2.58	11	0	0
② 건강·안전·영양	1.47~2.79	11	0	2
③ 학습환경	1.66~2.79	19	0	0
④ 교육경험 및 활동	1.31~2.54	14	0	2
⑤ 교사·영유아 상호작용	1.68~2.76	18	0	0
계	1.31~2.79	73	0	4

<표 4>에 나타난 바와 같이 문항 반응점수의 평균은 3점 척도에서 1.31~2.79사이에 분포하며, 천정효과를 나타내는 문항은 없는 것으로 나타났다. 문항 변별도 분석에서 대부분의 문항이 양호하였으나  $x^2$ 값이 유의하지 않은, 즉 문항이 프로그램의 질을 구별하지 못하는 경우가 모두 4개 문항으로 나타났다. 이를 영역별로 살펴보면 '건강·안전·영양' 영역의 '칫솔과 컵의 관리'와 '전기 콘센트의 안전덮개'와 관련된 2개 문항, '교육경험 및 활동' 영역의 '영아 언어활동, 미술활동 또는 음률활동'과 '유아 물놀이와 모래놀이'와 관련된 문항 2개였다. 이 4개 문항은 문항 내용의 중요성과 문항 삭제시의  $\alpha$ 등을 고려한 결과 삭제하기보다 포함시키기로 결정하였다.

## 2. 관찰척도의 신뢰도

관찰척도의 신뢰도 검증은 내적 합치도인 Cronbach  $\alpha$ 와 관찰자간 신뢰도를 분석하였다. 척도의 하위영역 및 전체에 대한  $\alpha$ 계수를 제시하면 다음의 <표 5>와 같다.

<표 5> 어린이집 관찰척도의 Cronbach  $\alpha$ 계수

어린이집 관찰척도	문항수	Cronbach $\alpha$
① 물리적 환경	11	.85
② 건강·안전·영양	11	.83
③ 학습환경	19	.93
④ 교육경험 및 활동	14	.85
⑤ 교사·영유아 상호작용	18	.92
전 체	73	.95

<표 5>에 의하면, 관찰척도의 내적 합치도는 전체 척도에서 .95, 영역별로는 .83에서 .93에 이르고 있다. 즉, 관찰척도의 전체 문항 및 각 영역별 문항간에 내적 일관성이 어느 정도 높다고 할 수 있다.

관찰자간 신뢰도는 훈련된 8명의 관찰자가 2명씩 짝지어 3개 기관을 독립적으로 관찰한

<표 6> 관찰자간 점수의 상관계수 및 평정 일치도

어린이집 관찰척도	상관계수	평정 일치도(%)
① 물리적 환경	.74	84
② 건강·안전·영양	.95	84
③ 학습환경	.76	82
④ 교육경험 및 활동	.96	85
⑤ 교사·영유아 상호작용	.96	83
전 체	.92	84

후, 영역 및 전체에 대해 상관계수( $r$ )와 문항별 평정의 일치도(%)를 산출하였다. 관찰자간 신뢰도 자료는 <표 6>에 제시되어 있다.

<표 6>에서 보면 관찰자간 영역별 상관계수는 .74에서 .96으로, 평정의 일치도는 82%에서 85%로 나타났다. 전체 척도에서 상관계수는 .92를, 일치도는 84%를 보이고 있다. 따라서 척도의 내적 합치도와 관찰자간 신뢰도 수준은 비교적 양호하다고 할 수 있다.

### 3. 관찰척도의 타당도

관찰척도의 구인을 탐색하기 위한 첫 번째 분석으로 관찰척도의 결과가 원장 및 교사에 의한 자체평가를 얼마나 잘 반영하는지 분석하기 위해, 세 종류의 자체 평가척도 총점에 따라 89개 기관을 상, 중, 하 세 집단으로 나누고, 이 세 집단의 관찰척도 영역별 평균의 차이를 검증하였다. 먼저 운영관리 평가척도에 의한 상, 중, 하 집단의 관찰척도 영역별 평균점수와 차이

<표 7> 운영관리 평가척도 상·중·하 집단에 따른 관찰척도 영역별 점수의 차이

관찰척도	운영관리			F
	상(n=30) M(SD)	중(n=29) M(SD)	하(n=30) M(SD)	
① 물리적 환경	2.62( .31) <sup>a</sup>	2.22( .45) <sup>b</sup>	1.98( .28) <sup>c</sup>	24.75***
② 건강·안전·영양	2.57( .26) <sup>a</sup>	2.32( .33) <sup>b</sup>	2.16( .32) <sup>b</sup>	13.54***
③ 학습환경	2.47( .39) <sup>a</sup>	2.07( .49) <sup>b</sup>	1.76( .26) <sup>c</sup>	25.12***
④ 교육경험 및 활동	2.32( .42) <sup>a</sup>	1.99( .43) <sup>b</sup>	1.86( .31) <sup>b</sup>	11.17***
⑤ 교사·영유아 상호작용	2.60( .40) <sup>a</sup>	2.25( .43) <sup>b</sup>	2.18( .36) <sup>b</sup>	9.68***

abc 다른 문자는 Sheffè검증결과 유의한 차이를 나타낸다.

\*\*\*p<.001

에 대한 F검증 결과는 다음의 <표 7>과 같다.

<표 7>에 의하면, F검증 결과 집단간 차이가 모든 관찰척도의 영역에서 유의한 것으로 나타났다. F검증이 유의하므로 어떤 집단간의 차이가 유의도에 기여했는가를 알기 위해 사후검사로 Sheffè검증을 실시하였다. 이 결과에 따르면, 상집단은 중집단 또는 하집단과 모든 영역에서 유의한 차이를 보이고 있다. 중집단과 하집단간의 차이가 유의한 영역은 ‘물리적 환경’과 ‘학습환경’이며, 다른 세 영역에서는 중집단과 하집단간 차이가 유의하지 않았다. 따라서 관찰척도의 점수는 원장이 운영관리 측면에 대해 자체적으로 평가한 결과를 잘 반영해 준다고 해석할 수 있다. 특히 물리적 환경과 학습환경 등 환경의 물리적 요소가 많이 포함된 내용에서는 상, 중, 하집단의 구분이 더 분명하였다.

그 다음 영아반 평가척도의 총점에 의해 나눈 세 집단의 관찰척도 영역별 평균, 표준편차, 그리고 차이검증 결과를 보면 <표 8>과 같다.

영아반 교사에 의한 자체평가 결과는 관찰의

<표 8> 영아반 평가척도 상·중·하 집단에 따른 관찰척도 하위영역별 점수의 차이

관찰척도	영아반			F
	상(n=30) M(SD)	중(n=29) M(SD)	하(n=30) M(SD)	
1. 물리적 환경	2.61( .31) <sup>a</sup>	2.21( .40) <sup>b</sup>	2.01( .37) <sup>b</sup>	21.54***
2. 건강·안전·영양	2.54( .28) <sup>a</sup>	2.34( .35) <sup>b</sup>	2.16( .29) <sup>b</sup>	11.53***
3. 학습환경	2.41( .44) <sup>a</sup>	1.98( .45) <sup>b</sup>	1.91( .42) <sup>b</sup>	11.18***
4. 교육경험 및 활동	2.39( .30) <sup>a</sup>	1.89( .38) <sup>b</sup>	1.88( .39) <sup>b</sup>	19.79***
5. 교사·영유아 상호작용	2.64( .31) <sup>a</sup>	2.24( .40) <sup>b</sup>	2.14( .41) <sup>b</sup>	14.99***

abc 다른 문자는 Sheffè 검증결과 유의한 차이를 나타낸다.

\*\*\*p<.001

하위영역에서 모두 세 집단간 유의한 차이가 나타났다. 사후 검증결과, 상집단은 중집단 또는 하집단과 모든 관찰영역에서 유의한 차이를 보였다. 그러나 중집단과 하집단간의 차이는 Sheffé에서 유의하지 않았다. 따라서 관찰척도의 점수는 영아반 교사가 자체적으로 평가한 결과를 반영하며, 특히 자체평가의 상집단이 중 또는 하집단보다 높게 관찰된다고 할 수 있다.

마지막으로 유아반 평가척도의 총점에 따른 세 집단간의 관찰척도 영역별 차이검증은 <표 9>에 제시되어 있다.

이 자료에서도 F검증 결과가 모두 유의한 것으로 나타났으며, Sheffé검증 결과는 상집단이 하집단과는 유의한 차이를 모든 영역에서 보이고 있다. 특히 ‘물리적 환경’ 영역에 대한 관찰점수는 유아반의 자체평가 결과를 상, 중, 하 세 집단간 분명히 구분하였다. 그러나 다른 영역에서 중집단은 상 또는 하집단과 유의한 차이가 없었다. 이상에서 제시한 바와 같이, 관찰척도 점수는 자체적으로 평가한 운영관리, 영아반 실제, 유아반 실재를 상당 부분 반영한다고 할 수 있다.

관찰척도의 구인을 탐색하기 위한 두번째 분석으로 척도의 5개 하위영역간 상호 상관계수를 산출하였고, 이를 <표 10>에 제시하였다.

<표 9> 유아반 평가척도 상·중·하 집단에 따른 관찰척도 하위영역별 점수의 차이

관찰척도	유아반			F
	상(n=30) M(SD)	중(n=29) M(SD)	하(n=30) M(SD)	
1. 물리적 환경	2.57(.37) <sup>a</sup>	2.28(.39) <sup>b</sup>	1.99(.36) <sup>c</sup>	18.12***
2. 건강·안전·영양	2.54(.28) <sup>a</sup>	2.34(.38) <sup>ab</sup>	2.17(.28) <sup>b</sup>	10.57***
3. 학습환경	2.33(.49) <sup>a</sup>	2.19(.45) <sup>a</sup>	1.79(.34) <sup>b</sup>	12.66***
4. 교육경험 및 활동	2.30(.40) <sup>a</sup>	2.12(.40) <sup>a</sup>	1.74(.29) <sup>b</sup>	18.15***
5. 교사-영유아 상호작용	2.50(.45) <sup>a</sup>	2.43(.41) <sup>a</sup>	2.09(.32) <sup>b</sup>	9.05***

abc 다른 문자는 Sheffé 검증결과 유의한 차이를 나타낸다.  
\*\*\*p<.001

<표 10>에 의하면, 하위영역간의 상호상관이 .52~.79로 모두 유의한 정적 상관을 나타내고 있어 관찰척도 총점으로 어린이집의 질적 수준을 요약할 수 있다고 하겠다. 하위영역별로 보면, 기관의 전반적인 ‘물리적 환경’ 영역이 다른 네 영역과 .60~.79의 높은 정적 상관 관계를 나타내고 있다. 또한 ‘학습환경’은 ‘교육경험 및 활동’과 .77의 상관을, ‘교사-영유아 상호작용’과 .73의 높은 상관을 보여주고 있다. 한편 ‘교사-영

<표 10> 어린이집 관찰척도 하위영역간의 상호상관(N=89)

어린이집 관찰척도	① 물리적 환경	② 건강·안전·영양	③ 학습환경	④ 교육경험 및 활동	⑤ 교사-영유아 상호작용
① 물리적 환경	-				
② 건강·안전·영양	.71***	-			
③ 학습환경	.79***	.63***	-		
④ 교육경험 및 활동	.71***	.56***	.77***	-	
⑤ 교사-영유아 상호작용	.60***	.52***	.73***	.78***	-

\*\*\*p<.001

<표 11> 어린이집 관찰척도와 유아교육 프로그램 평가척도 영역간의 상관(N=89)

관찰척도 유아교육 프로그램	① 물리적 환경	② 건강·안전·영양	③ 학습환경	④ 교육경험 및 활동	⑤ 교사·영유아 상호작용
① 일과계획 및 상호작용	.42***	.45***	.28**	.32**	.25*
② 교육과정	.55***	.46***	.46***	.44***	.40***
③ 물리적 환경	.53***	.38***	.49***	.40***	.30**
④ 건강·안전·영양	.42***	.51***	.27**	.36***	.31**
⑤ 운영관리	.40***	.31**	.41***	.38***	.27**
유아교육 프로그램 평가척도 총점	.61***	.51***	.52***	.49***	.40***

\*\*\*p<.001 \*\*p<.01 \*p<.05

유아 상호작용'은 '물리적 환경'(.60)이나 '건강·안전·영양'(.52)보다 '학습환경'(.73)이나 '교육경험 및 활동'(.78)과 더 높은 상관을 나타낸다. 이러한 상관 패턴은 본 척도가 보육시설의 환경 요소와 영유아에게 제공되는 경험요소를 중심으로 평가하고자 한 본래의 목적을 어느 정도 반영하는 것으로 해석할 수 있다.

공인타당도는 본 관찰척도와 외적 준거인 유아교육 프로그램 평가척도 점수간의 상관계수를 산출하여 검토하였다. 이 결과가 <표 11>에 제시되어 있다.

<표 11>에서 보면 어린이집 프로그램 관찰척도와 자체 평가한 유아교육 프로그램 평가척도의 영역간 상관계수의 범위가 .25~.55로 모두 유의한 정적 상관이 나타났다. 이처럼 어린이집

관찰척도와 유아교육 프로그램의 질을 평가하는 외적 준거간 관련성이 높다는 것은 공인타당도의 증거라 할 수 있다.

두 척도에서 개념적으로 관련된 영역간의 상관계수를 살펴보면, 프로그램 관찰척도의 '물리적 환경'은 유아교육 프로그램 평가척도의 '물리적 환경'과 .53(p<.001), 두 척도의 '건강·안전·영양' 간에는 .51(p<.001), '학습환경'과 준거척도의 '교육과정', '물리적 환경'과는 각각 .46(p<.001)과 .49(p<.001), '교육경험 및 활동'은 '교육과정'과 .44(p<.001)의 상관을 보여 관련이 적은 다른 영역들간의 상관계수보다 대체로 높은 것을 알 수 있다. 즉, 개념적으로 관련성이 높은 영역간의 상관계수가 더 높은 것은 수렴 타당도의 증거로 해석할 수 있다.

#### IV. 논의 및 결론

본 연구에서는 「어린이집 프로그램 관찰척도」를 개발하고 타당도를 검증하고자 하였으며, 주요 결과를 다음에 논의하고 결론을 제시하고자

한다.

첫째, 어린이집 관찰척도의 문항 평균은 자체 평가척도의 관련영역 평균(이은혜 외, 2002)과

비교할 때 전체적으로 낮은 경향이 있었다. 이러한 경향은 예비조사와 타당화 절차 모두에서 발견되었다. 이는 외부 관찰자의 관찰이 내부의 자체평가보다 객관적이기 때문에 현장 관찰에서 더 실제에 가까운 자료를 제공한다고 해석될 수 있으며, 이러한 이유 때문에 자체평가만으로는 평가결과를 활용하는데 한계가 있다고 하겠다. 또 다른 해석으로는 관찰시간의 제한 때문에 어린이집에서 때때로 실시되고 있는 활동일지라도 관찰당시에 관찰되지 못할 수도 있다. 어느 해석이 더 설득력이 있는지는 추후 연구를 통해서 계속 탐구되어야 할 것이다.

둘째, 관찰척도의 문항 양호도 분석 결과에 의하면, 73개 문항은 4개 문항을 제외하고 문항반응 분포와 문항 변별도에서 양호한 것으로 나타났다. 이는 무엇보다도 관찰척도를 구성할 때 자체평가척도 문항 중에서 양호한 문항으로 선정한 것에 기인하는 것이다. 문항 변별도가 약한 4개 문항 중 2개는 영유아의 안전과 건강 관리 측면에서 반드시 포함되어야 할 문항이었고, '교육경험 및 활동' 영역의 2개 문항은 영아의 학습경험을 측정하는 소수 문항 중의 하나라는 것과 물·모래 놀이에 관련된 것이었다. 이러한 문항을 포함시키고도 Cronbach  $\alpha$  계수가 해당영역에서 각각 .83(건강·안전·영양), .85(교육경험 및 활동)였기 때문에 문항을 삭제하지 않고 포함시키게 되었다.

셋째, 관찰척도의 신뢰도 검증 결과에 따르면, 척도의 전체 또는 각 영역별  $\alpha$ 의 범위가 .83~.93으로 내적 일관성이 높다고 할 수 있으며, 관찰자간 신뢰도도 상관계수 .74~.96으로 비교적 양호하다고 볼 수 있다. 다만 관찰자간 신뢰도 검증에서 대상 기관수가 적으며, 예비조사와 본 조사에 걸쳐 관찰 훈련이 충분히 이루어지고 현장 경험이 많은 관찰자들이 다수

관찰을 했다는 점에 유의할 필요가 있다. 관찰의 신뢰도는 관찰자의 자격과 훈련정도에 따라 차이가 클 수 있기 때문이다.

넷째, 척도의 구인 타당도와 공인 타당도 검증결과에 의하면, 관찰척도의 타당성을 보여주는 증거들이 제시되었다. 먼저 세 가지 자체평가 결과의 타당성을 관찰척도로 점검한 결과에 의하면, 자체평가척도에 의한 질적 수준의 상, 중, 하를 관찰척도의 영역별 평균점수가 의미 있게 구별해주는 것으로 나타났다. 이는 관찰척도만으로도 보육 프로그램의 질을 구별할 수 있다는 점에서 관찰척도의 타당성을 제시하는 것이며, 나아가 자체평가 결과를 인준하는 과정에서 유용하게 활용될 수 있음을 시사하는 것이다. 그 다음 관찰척도의 하위 영역간의 상관은 .52~.79의 유의한 정적 상관을 모두 나타냄으로써 관찰척도의 총점이 어린이집 프로그램의 질을 요약해주는 지표로써 사용될 수 있다고 해석할 수 있다.

관찰척도의 공인 타당도는 외적 준거인 유아교육 프로그램 평가척도와 상관계수를 통해 검증되었으며, 대체로 유의한 정적 상관을 나타냈다. 이는 교육프로그램의 질을 전반적으로 측정하는 측면에서 본 관찰척도의 공인타당도를 확인하는 증거라 할 수 있다. 특히 하위영역별 상관계수에서 개념적으로 관련된 영역간의 상관이 그렇지 않은 영역간의 상관보다 높은 경향을 나타냄으로써 수렴타당도를 시사하였다. 국내에 유아교육 또는 보육 프로그램을 평가하는 양호한 도구들이 제한되어 있어서 외적 준거척도를 선택하는데 어려움이 있으나, 본 척도의 타당도에 관련된 증거는 지속적으로 탐구되어야 할 과제이다.

본 연구의 결론에 앞서 몇 가지 제한점을 지적하면, 첫째 관찰척도의 본 조사 실시 대상기관이

89개로 전국적인 표집은 하였으나, 삼성어린이집과 국공립어린이집이 64개로 외부인의 관찰을 허용할 만큼 비교적 우수한 질의 보육시설일 가능성이 높다. 따라서 본 척도가 양질의 기관을 평가하는데는 무리가 없으나, 질이 낮은 기관에 적용할 때 제한적일 수 있다. 둘째, 문항별 1, 2, 3점의 평정기준을 관찰 가능하고 명료하게 서술하려 했으나, 관찰자에 따라서는 기준을 다르게 해석할 가능성이 완전히 배제되기는 어렵다. 따라서 관찰자 훈련을 위해 평정기준에 대한 상세한 해설을 포함한 「사용 지침서」가 앞으로 필요할 것이다. 그리고 엄격한 관찰자 훈련을 통해서

만 본 척도의 유용성이 보장될 것이다. 셋째, 척도의 타당도에 관련된 증거들이 제시되기는 하였으나, 타당도는 새로운 증거와 다른 접근방법으로 계속해서 탐구되어야 할 과제이다.

결론적으로 본 연구에서 개발된 「어린이집 프로그램 관찰척도」는 대체로 양호한 문항으로 구성되어 있으며, 척도의 신뢰도가 우수한 편이고 타당도의 증거도 제시되어 있다. 본 척도는 어린이집 프로그램의 질을 전반적으로 평가하기 위해서 외부 관찰자의 객관적인 평가도구로 그리고 자체평가 결과를 인준하는 과정에서 활용될 수 있을 것으로 기대된다.

## 참 고 문 헌

- 강숙현 (1994). *유아교육 프로그램 평가척도의 이해와 활용*. 서울 : 동문사.
- 권영례, 이영자, Harms(1999). 유치원과 초등학교 저학년 교육환경 평가척도 개발 연구. *교육학연구*, 37(3), 341-364.
- 김윤지(2003). 유아교육환경 평정척도 개정판(ECERS-R)의 국내적용을 위한 예비연구. 연세대학교 대학원. 석사학위 논문.
- 이기숙(1989). 한국 유아교육기관의 질적 분석 기준 개발에 관한 연구. *논총(한국문화연구원)*, 55, 295-320.
- 이대균, 이기우(1999). 유치원 기관평가 도구 개발을 위한 연구. *열린유아교육연구*, 4(1), 213-235.
- 이영석(1990). 한국 유아교육시설기관의 표준평가척도 개발에 관한 예비연구. *성균관대학교 사회과학*, 29(2), 169-221.
- 이영자, Harms, 권영례(1998). 유치원과 초등학교 저학년 교육환경 평가척도 개발 예비연구. *초등교육연구*, 12(1), 141-160.
- 이 옥(2000). 보육시설 인증제 도입에 따른 평가항목 개발 연구의 성과와 과제. *한국영유아 보육학*, 21, 11-37.
- 이완정, 현은강(1998). 영아/걸음마기 아동 환경평정척도 타당화 연구(II). *대한가정학회지*, 36(6), 95-108.
- 이은혜(2002). 보육시설 프로그램 평가도구의 개발과 타당화. *영유아 교육기관의 인정제*, 삼성복지재단과 이화여대 유아교육학과 BK21 핵심연구팀 공동 주최 국제학술대회 자료집, pp. 93-130.
- 이은혜, 송혜린, 신혜영(2001). *삼성 어린이집 운영 및 프로그램 평가도구 개발연구*. 삼성복지재단.
- 이은혜, 송혜린, 신혜영, 최혜영(2002). 어린이집 프로그램 평가척도의 개발을 위한 예비연구. *아동학회지*, 23(4), 199-213.
- 이은혜, 이기숙(1994). 유아교육 프로그램 평가척도 개발에 관한 예비 연구. *아동학회지*, 15(1), 5-21.
- 이은혜, 이기숙(1995). 유아교육 프로그램 평가척도의 타당화 연구. *유아교육연구*, 16(1), 157-178.
- 이은혜, 이기숙(1996). *유아교육 프로그램 평가척도 : 교사를 위한 활용지침서*. 서울 : 창지사.
- 임재택(1983). 유아교육환경 평정척도의 표준화를 위한 예비 연구. *교육논집(부산대학교 사범대학)*,

- 10, 107-143.
- 임재택, 조희숙, 황해익(1996). 영유아 보육프로그램의 진단·평가척도 개발 연구. *성곡논총*, 29(4), 401-500.
- 정기원, 오미영, 안현애(1995). 보육시설 평가기준 및 평가체계 개발. 한국보건사회 연구원.
- 최경애(2000). 보육시설의 영아기 프로그램 평가척도 개발을 위한 기초연구: 영아보육프로그램 평가 및 운영 실태 조사. *사회과학연구(덕성여자대학교)*, 6, 191-214.
- 현은강, 이완정(1997). 영아/결혼마기 아동 환경평정 척도 타당화 연구(I). *대한가정학회지*, 35(6), 235-248.
- Abbott-Shim, M., & Sibley, A.(1987). *Assessment Profile for Early Childhood Programs*. Atlanta, GA : Quality Assist.
- Bredenkamp, S., & Willer, B. A.(1996). *NAEYC accreditation : A decade of learning and the years ahead*. Washington, DC : NAEYC.
- Friedman, S. L., & Amadeo, J.(1999). The child-care environment : Conceptualizations, assessments, and issues. In S. L. Friedman and T. D. Wachs (Eds.), *Measuring environment across the life span* (pp. 127-165), Washington, DC : APA.
- Gullo, D. F.(1994). *Understanding assessment and evaluation in early childhood education*. New York : Teachers College Press.
- Harms, T., & Clifford, R. M.(1980). *Early Childhood Environment Rating Scale*. New York : Teachers College Press.
- Harms, T., Clifford, R. M., & Cryer, D.(1998). *Early Childhood Environment Rating Scale*(Rev. ed.). New York : Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M.(1990). *Infant/Toddler Environment Rating Scale*. New York : Teachers College Press.
- Katz, L. G.(1992). Early childhood programs : Multiple perspectives on quality. *Childhood Education*, 69(2), 66-71.
- NAEYC(1998). *Accreditation Criteria and Procedures of the National Association for the Education of Young Children, 1998 Edition*. Washington, D C : Author.
- Peisner-Feinberg, E., & Burchinal, M.(1997). Relations between preschool children's child care experiences and concurrent development : The Cost Quality and Outcome Study. *Merrill-Palmer Quarterly*, 43(3), 451-477.
- Phillips, D. A., & Howes, C.(1987). Indicators of quality in child care : Review of research. In D. A. Phillips(Ed.), *Quality in child care : What does research tell us?*(pp. 1-20). Washington, DC : NAEYC.
- Spodek, B., & Saracho, O. N.(Eds.) (1997). *Issues in early childhood educational assessment and evaluation*. New York : Teachers College Press.