

한국판 유아행동평가 척도 교사용(C-TRF)의 표준화 연구

김 영 아¹ 이 진¹ 김 유 진¹ 김 민 영¹ 오 경 자^{2*}
(주)휴노¹ 연세대학교 심리학과²

본 연구에서는 한국판 유아 행동평가척도 교사용(C-TRF)을 표준화하고 신뢰도와 타당도를 검증하기 위해 1,371명(남아 679명, 여아 692명)의 18개월~5세 영유아를 대상으로 자료를 수집하였다. 확인적 요인분석 결과, 미국판 원 검사의 요인구조를 한국에서도 적용할 수 있음을 확인하였고, 검사-재검사 신뢰도 값(.57~.82)이나 각 하위척도들의 내적 합치도 수준(Cronbach's $\alpha = .52 \sim .96$)은 일부 척도를 제외하고 전반적으로 양호한 것으로 나타났다. 한국판 C-TRF의 하위척도들은 상호 요인 간, 특히 내재화 및 외현화의 상위 요인 별로 높은 상관을 보였다. 문제행동총점을 비롯한 대부분의 척도에서 임상집단($n = 33$)과 비임상집단($n = 1,209$) 간에 중간 이상의 효과크기 수준 평균차이를 보였으며, 한국판 C-TRF 총점이 임상범위에 해당되는 경우, 그렇지 않은 경우 보다 어머니가 실시한 한국판 CBCL 1.5-5의 하위척도 점수가 모두 높은 것으로 나타나 수렴 및 변별 타당도를 뒷받침하였다. 성별, 연령대 별 척도 점수 비교 결과, 36개월 전후 남녀 네 집단으로 규준을 제시하는 것이 적합한 것으로 나타났다. 한국판 C-TRF의 임상 및 연구 장면에서의 활용에 대한 논의가 이루어졌다.

주요어: 한국판 C-TRF, 유아 문제행동, 교사 보고 행동평가척도, 표준화

급격한 신체적, 정신적 발달과 더불어 생애 초기의 사회적 관계 형성이 이루어지는 시기인 영유아기 아동의 정신건강에 대한 관심이 높아지고 있다. 문제 행동의 발현에 있어 조기 개입이 중요하다는 인식과 함께 영유아시기에 관찰되는 문제를

조기에 발견하기 위한 시도가 이어지고 있다. 어린 아동은 자신의 문제를 스스로 인식하여 표현하지 못하므로 영유아기 아동의 정서행동문제 진단에는 아동을 평소에 관찰하고 평가할 수 있는 성인의 보고가 중요하다. 유아와 많은 시간을 보내는 성인

* 교신저자: 오경자, E-mail: kjoh@yonsei.ac.kr

의 관찰은 많은 정보를 담고 있을 뿐 아니라, 유아를 바라보는 시각 자체가 유아를 대하는 태도에도 영향을 미칠 수 있다.

영유아의 문제행동을 성인 보고에 의거하여 선별하기 위한 검사들이 국내외에서 개발되어 사용되고 있다. Reynolds와 Kamphaus(1992)의 BASC (Behavior Assessment System for Children)는 2세 6개월부터 6세 유아의 행동에 대한 부모 평가 척도로 한국에서도 개발 연구가 진행되었다(이경숙, 신의진, 신동주, 전연진, 박진아, 2003; 조광순, 최지영, 2003). Goodman(1997)의 SDQ(Strengths and Difficulties Questionnaire)는 4~16세 아동을 대상으로 주의력, 또래관계, 친사회적 행동을 평가하고 있는 행동평가척도로 안정숙, 전성균, 한준규와 노경선(2003)에 의해 한국판으로 개발되었다. 18개월부터 만 5세 유아를 대상으로 하는 부모용 행동평가척도인 CBCL 1.5-5 (Achenbach & Rescorla, 2000)도 한국판으로 개발되어 전국적인 표준화를 통해 규준이 완성되고, 임상적 타당성이 검증되었다(김영아, 이진, 문수중, 김유진, 오경자, 2009; 오경자, 김영아, 2008; 이진, 김영아, 오경자, 2009). 이러한 척도들은 주로 부모의 평가에 의존한 것들이 많은데, 대개는 부모가 유아의 행동을 가장 많은 시간 관찰할 수 있고 양육 비중이 큰 점을 고려할 때 일차적으로 부모의 평가가 중시되는 것은 당연하다.

그러나 맞벌이 부부가 늘어나고 영유아 조기 교육에 대한 관심이 증가함에 따라 영유아가 보육기관이나 유아원 등의 교육기관에서 보내는 시간이 점차 길어지고 있어 교사평가의 중요성이 커지고 있다(신미숙, 2004). 또한 영유아와 일정한 시간을 보내는 교사의 평가를 활용할 때 여러 가지 이점이 있다. 첫째, 부모와의 직접적인 상호작용이 이루어지지 않는 가정 이외의 보육·교육기관에서의

영유아 행동은 해당 상황에서 영유아를 직접 지도하고 보호하는 교사의 관찰이 부모의 평가보다 신뢰로울 수 있다. 둘째, 한 명의 영유아에 대해 보고하는 부모의 보고와 달리, 여러 영유아들을 함께 관찰하는 교사는 보다 객관적인 시각에서 영유아의 행동을 평가할 수 있다. 셋째, 다른 영유아들이 함께 모여 단체생활을 하게 되는 유치원이나 유아원은 다양한 사회적 기술과 행동 양식을 배우는 생활의 장으로, 이후 발달에 큰 영향을 미치는 중요한 위험요인이 될 수 있는 적응과 관련된 문제행동을 파악할 수 있다. 넷째, 영유아의 행동에 대한 교사의 보고 내용이 부모와 일치하는 경우 자료의 신뢰성을 판단하는 근거가 된다. 다섯째, 교사와 부모의 평가가 일치하지 않은 경우에도 환경적 맥락에서 영유아의 문제행동을 유발하는 요인이 어떤 것인지를 탐색하거나 부모-교사 간 관점의 차이를 이해함으로써 영유아 행동에 대한 이해의 폭을 넓힐 수 있다. 동일한 대상이라 할지라도 서로 다른 관점을 지닌 평가자들과 서로 다른 환경 내에서 상호작용을 하면서 보이는 행동 양식은 큰 차이를 보일 수 있는 만큼, 평가 대상에게 의미가 있는 사회적 상호작용의 맥락을 반영할 수 있도록 한 명 이상의 평가자로부터 행동 평가의 결과를 얻는 것이 중요하다.

영유아 및 아동기의 문제를 조기 발견하고 대처하는 데 있어 교사 평가 도구가 중요한 축을 담당할 수 있다는 공감대를 기반으로 많은 학자와 전문가들이 교사에 의한 행동평가 척도를 개발해오고 있다(Achenbach, 1991; Cassidy & Asher, 1992; Gresham & Elliott, 1990; Quay & Peterson, 1987; Reynolds & Kamphaus, 1992).

최근 들어 우리나라에서도 교사의 평가에 기반한 행동평가 도구를 개발하려는 시도들이 있어왔다. 국내에서 개발된 교사 보고형 도구들의 상당수

는 외국 검사를 번안하여 표준화한 것(박은희 등, 2003; 소유경, 노주선, 김영신, 고선규, 고운주, 2002)으로 그나마도 초등학교 이후 취학 아동을 대상으로 하고 있는 것이 대부분이다. 신현숙, 박용재, 박주희와 류정희(2006)는 초등학생의 학교적응 수준을 평정하는 교사용 척도를 개발·타당화하여 학교 장면에서의 교사 평가를 통해 아동 행동에 대한 효율적 이해가 가능함을 밝힌 바 있고, 정현희(2002)는 미국에서 개발된 교사용-아동행동평가 척도인 T-CRS(Hightower, Work, Cowen, Spinell, Lotyczewski, Guare, & Rohrbeck, 1986)를 한국판으로 개발하여 신뢰도와 타당도를 검증하였다.

오경자와 이혜련(1989)은 Conners 부모 및 교사평가 척도 단축형(Abbreviated Conners' Parent Rating Scale: ACPRS, Abbreviated Conners' Teacher Rating Scale: ACTRS)의 신뢰도와 타당도를 검증하면서 아동의 과잉행동 문제에 대한 부모와 교사평가 자료를 비교하였다. 이 척도는 Conners(1969, 1970, 1973)에 의해 과잉행동 문제와 관련된 치료효과를 검증하기 위한 목적으로 개발된 93 문항의 부모용 척도와 39 문항의 교사용 척도를 Goyette, Conners와 Ulrich(1978)가 3세부터 17세 아동을 대상으로 하여 부모용-교사용 각각 10개의 문항으로 줄인 단축형 검사이다. 한국판 Conners 단축형 검사는 아동·청소년의 주의력 결핍 과잉행동 문제 평가에 쓰이는데, 부모와 교사 평가의 상관관계는 그다지 높지 않았으며, 상황에 따라 과잉활동적 행동이 다르게 나타남을 시사하였다.

유아를 대상으로 하는 국내 교사용 검사로 황혜정(2005)의 유아의 문제행동진단 검사, 이혜련과 양옥승(2006)의 교사용 유아의 사회적 유능감 평가 척도 등이 있기는 하나, 아직까지는 국내에서 표준화된 많은 유아 검사 도구들이 대체로 지능검사로 제한되어 있는 실정이다(윤진주, 안진경, 김지영,

2008 재인용). 특히 영유아의 발달 수준이나 교육 장면에서의 수행 수준과 관련된 행동의 평가 이외에 임상적으로 활용될 수 있는 교사용 문제행동 척도에 대한 연구가 충분히 이루어지지 않고 있다.

C-TRF(Caregiver-Teacher Report Form)는 18개월부터 만 5세 사이 영유아의 행동 문제를 영유아가 다니는 보육시설 및 유아교육기관의 교사가 해당 영유아의 행동을 평가하도록 하는 척도로, 전 연령대 행동평가 시스템인 ASEBA(Achenbach System of Empirically Based Assessment) 검사군 중 하나이다(Achenbach & Rescorla, 2000). 기본적으로 CBCL 1.5-5와 C-TRF가 포함된 ASEBA는 한 대상에 대한 여러 명의 평가 자료를 비교할 수 있도록 같은 연령대 검사들의 연계 사용을 고려하여 개발되었다. 따라서 C-TRF는 부모의 평가를 토대로 하는 CBCL 1.5-5와 유사한 구조와 문항으로 구성되어 있다. 다만 교사가 관찰하기 어려울 수 있는 수면 문제와 같은 문항들은 유치원에서 나타나기 쉬운 행동 문항으로 대체되어 있다.

본 연구에서는 CBCL 1.5-5의 한국판 개발에 이어서 CBCL 1.5-5와의 연계활용을 통해 평가의 객관성을 높이고, 임상적 유용성을 증진시킬 수 있을 것으로 기대되는 교사용 유아행동 평가 척도인 C-TRF를 한국판으로 개발하고자 하였다. 80개 이상의 문화권에서 폭넓게 사용되고 있는 ASEBA 검사군의 교사용 평가 도구인 C-TRF를 한국판으로 개발하여 한국 유아들의 문제행동 특성에 대한 자료를 누적하고 향후 한국 유아에게 적합한 도구로 조율되어나갈 수 있는 토대를 마련하고자 한다. 이를 위해 전국 표집을 통해 자료를 수집하고 신뢰도와 타당도를 평가하여 검사의 구성을 확정하고 규준을 마련하였다.

연구문제는 다음과 같다.

<연구문제 1> 미국판 C-TRF 원검사의 요인구

조를 한국의 영유아 집단 대상에 그대로 적용하는 것이 타당한가?

<연구문제 2> 한국판 C-TRF의 신뢰도(검사-재검사 신뢰도, 내적 합치도)와 타당도(수렴, 변별 타당도)는 적합한가?

<연구문제 3> 한국판 C-TRF의 점수를 통해 나타난 교사가 보고한 한국의 1.5~5세 영유아의 문제행동 수준은 어떠한가?

<연구문제 4> 한국판 C-TRF의 기준집단은 연령과 성별에 따라 어떻게 구분하는 것이 가장 통계적, 임상적으로 바람직한가?

방 법

연구대상

전국 각 지역에 분포된 유아교육 기관 22곳을 통해 수집된 18개월~83개월 사이의 유아 1,371명(남아: 679명, 여아: 692명)에 대한 평가가 기준 산출을 위한 자료로 활용되었다. 설문은 영유아가 속

해 있는 기관의 반 또는 그룹의 담임을 맡고 있는 교사, 혹은 보육자가 실시하도록 하였다. 총 1,404명의 영유아 평가 자료가 수집되었는데 분석을 위해 필요한 인구통계학적 변인(성별, 연령)이 포함되지 않은 사례와 행동평가 문항 100개 중 응답 결측치가 8개 이상인 사례, 그리고 해당 영유아가 임상기관에 내원한 적이 있다고 보고된 33개 사례를 제외한 1,371명의 자료가 기준 집단 분석에 포함되었다. 이는 미국판 C-TRF 원검사의 기준제작 방식(Achenbach & Rescorla, 2000)에서 사용한 기준집단 선별 기준과 동일하다.

임상기관에 방문한 적이 있는 사례를 중심으로 한 변별 타당도 분석에는 기준집단에서 제외된 33개 사례를 포함하였고, 영유아의 부모 평가와 교사 평가 간 관련성을 살펴보기 위한 분석에서는 부모(주 양육자) 평가와 교사 평가 자료가 모두 있는 1,185명(남아: 589명, 여아: 596명)의 자료를 분석 대상으로 하였다.

기준집단에 포함된 연령, 성별, 지역별 표집 사례수가 표 1에 제시되어있다. 주 대상자인 만 3~5

표 1. 기준집단의 연령, 성별, 지역별 표집 사례수 (N = 1,371)

연령	서울	경기도	충청/강원/제주	경상남북도	전라남북도	전체지역
18~35개월 (만 1.5/2세)	18 / 15	38 / 31	49 / 40	31 / 34	1 / 0	137/120 (10.0/8.8)
36~47개월 (만 3세)	27 / 29	35 / 40	42 / 32	37 / 56	22 / 27	163/184 (11.9/13.4)
48~59개월 (만 4세)	19 / 34	38 / 34	31 / 38	23 / 28	51 / 67	162/201 (11.8/14.7)
60~71개월 (만 5세)	18 / 20	17 / 16	22 / 21	13 / 8	87 / 84	157/149 (11.5/10.9)
72~83개월 (만 6세)	5 / 8	4 / 3	8 / 4	5 / 6	38 / 17	60/38 (4.4/2.8)
전체연령	87/106 (6.3/7.7)	132/124 (9.6/9.0)	152/135 (11.1/9.8)	109/132 (8.0/9.6)	199/195 (14.5/14.2)	679/692 (49.5/50.5)

남아/ 여아. ()안은 %

표 2. 지역별 인구분포: 기준집단과 전국평균(2005년 인구주택총조사 자료기반)

	서울	경기	충청/강원/제주	경상남북도	전라남북도
기준집단	14.0 (3)	18.6 (7)	20.9 (5)	17.6 (3)	28.7 (4)
전국평균 (2005년 인구주택총조사 0~5세 집단)	18.5	31.1	14.7	25.1	10.6

단위 %, ()안의 수는 표준화 과정에 참여한 유치원 및 보육원 등의 기관 수

표 3. 학력분포: 기준집단 부모와 전국평균(2005년 인구주택총조사 자료기반)

	기준집단 (부학력)	전국평균 (25~44세 남)	기준집단 (모학력)	전국평균 (25~44세 여)
대학원이상	10.1	5.9	3.8	3.6
대졸	50.5	51.1	53.8	42.4
고졸	20.4	38.5	26.8	47.3
중졸이하	1.4	4.6	.8	6.6
미기재	17.7	-	14.7	-
전체	100.0	100.0	100.0	100.0

단위 %

세의 연령별 인구비례는 통계청(2006)의 대한민국 연령별 인구전수조사 결과에 의한 인구비례를 반영하고 있다. C-TRF가 주로 만 5세까지의 유아를 대상으로 하지만 유치원에 다니는 기간 중에 일부 유아들이 만 6세가 된다는 점을 고려하여 만 6세라도 유치원에 다니고 있다면 C-TRF의 실시대상으로 포함시켰다. 표 1에 나타난 만 6세의 상대적 으로 낮은 비율은 유치원과 초등학교 1학년에 걸쳐 분포하고 있는 만 6세 중 유치원에 다니는 인구의 비율이 그대로 반영된 결과라고 볼 수 있다. 각 기관에 연령대별, 같은 반 단위로 C-TRF를 실시하도록 하였지만 18개월에서 36개월 미만(3세 미만)까지의 연령대도 상대적으로 적는데, 이는 이 시기 영유아들이 만 3세 이상에 비해 유아교육 기관에 다니는 비율이 낮고 반별 인원도 적어 수집 과정에 다소 제한을 받는 부분이 있었기 때문이다.

표 2는 기준집단 영유아의 지역적 분포 특성과

우리나라 전체 영유아 지역분포를 비교한 자료이며, 표 3은 이번에 수집된 영유아의 부모 학력 비율을 통계청의 우리나라 평균(통계청, 2006)과 비교한 결과이다. 기준집단 부모의 학력은 다소 높은 수준인 것으로 나타났다. 가정경제 수준에 대한 평가를 상, 중상, 중, 중하, 하의 5단계로 나누어 부모에게 보고하게 한 결과, “중”으로 평가한 비율이 43.3%로 가장 높았으며, 평균은 중보다 약간 낮은 수준이었다. 가구당 월평균 가계소득은 373만원(중앙값 316만원)으로 2007년 전국 가구당월평균 소득 금액보다 높고, 전국 도시근로자가구 평균과 거의 유사한 수준이었다.

평가에 참여한 교사는 모두 137명이었고, 교사 1인당 평균 9.63부의 검사를 실시한 것으로 나타났다. 평균 교사 경력은 4.82년이었고 교사와 영유아가 알고 지낸 기간은 평균 5.31개월이었다. 유치원이나 보육원 등의 기관 내에서 담임을 맡고 있는

그룹의 유아 수는 평균 20.37명이었으며, 교사의 학력 수준은 고등학교 졸업자 1.3%, 전문대 졸업자 53.9%, 대졸자 28.2%, 그리고 대학원 졸업자가 1.3% (무응답 15.3%)인 것으로 나타났다.

측정도구

1) 한국판 유아행동평가 척도-교사용(Caregiver-Teacher Report Form for Ages 1.5-5; C-TRF 1.5-5)

Achenbach와 Rescorla(2000)가 개발한 C-TRF를 한국어로 번역하여 사용하였다. 번역 과정은 임상심리 전문가가 1차 번역을 하고, 이를 1차 번역자를 포함, 대학원에서 임상심리학을 전공한 4인이 각 문항 번역의 정확성 및 번역 문항간의 일관성을 검토하였다. 1차 번역본의 검토 결과 의견이 일치하지 않은 문항에 대해서는 대안적 번역문을 정리한 후, 이를 C-TRF의 원저자와 영어/한국어를 사용하는 이중 언어 사용자에게 자문을 구하여 원래의 척도가 전달하려는 의미에 가장 가까운 번역을 택하여 2차 번역본을 구성하였다. 이를 다시 제3의 이중 언어 사용자의 역번역 과정과 그 결과에 대한 2차 검토를 통해 최종 번역본을 확정하였다. C-TRF나 CBCL 1.5-5가 속한 ASEBA 검사군의 경우, 연령대별 검사들의 연계 사용을 염두에 두고 개발되어, 상당수 문항이 동일한 행동을 부모와 교사의 시각으로 보고하도록 구성된 문항임을 고려하여 두 척도의 번역의 통일성을 유지하고자 하였다.

C-TRF는 영유아의 정서문제행동을 기술한 총 99개의 문항과 그 이외의 문제를 기술하도록 되어 있는 문항 1개를 포함해서 총 100개의 문항으로 구성되어 있다. 각 문항에 대해서는 유아를 담당하는 담임교사가 지난 2개월 내에 유아가 그 행동을 보였는지를 판단하여, “전혀 아니다: 0, 가끔 그렇

다: 1, 자주 그렇다: 2”의 3점 척도로 평가하도록 되어 있다. 따라서 점수의 범위는 0~200점이 된다.

미국판 C-TRF 원검사는 요인분석을 통하여 도출된 6개의 증후군 척도와 특정 요인으로 분류되지 않는 문제행동을 측정하는 1개의 기타문제 척도, 그리고 증후군 척도 및 기타문제 척도의 합으로 이루어진 내재화, 외현화, 문제행동총점 척도로 구성되어 있다. 6개 증후군 척도는 정서적 반응성, 불안/우울, 신체증상, 위축, 주의집중문제, 그리고 공격행동 척도이며, 이러한 척도들은 다시 내재화, 외현화, 문제행동총점의 3개 상위척도를 구성한다. 내재화 척도에는 정서적 반응성, 불안/우울, 신체증상, 위축 4개의 하위척도가 포함되고, 외현화 척도에는 주의집중문제와 공격행동 2개의 하위척도가 포함되며, 문제행동총점 척도는 문제행동 문항 전체의 합으로 점수화된다. CBCL 1.5-5에 포함되어 있던 수면문제 하위척도는 C-TRF에서는 제외되었다.

C-TRF에는 CBCL 1.5-5와 마찬가지로 경험적으로 도출된 증후군 척도를 보완하는 DSM진단척도(DSM-oriented scales)가 포함된다. DSM진단척도는 정신과 의사 및 심리학자로 구성된 9개 문화권의 전문가 16명으로 하여금 C-TRF의 각 문항을 유아에게서 자주 보고되는 DSM척도와 부합하는지를 평정하도록 하여 구성하였다(Achenbach & Rescorla, 2000). DSM(Diagnostic and Statistical Manual of Mental Disorders)이란 미국 정신의학회에서 발간해오는 정신장애의 진단 및 통계편람으로 정신장애 관련 분야에서 가장 광범위하게 사용되는 전통적인 진단기준이다(이춘재 등, 2002 참고). C-TRF를 비롯한 ASEBA 검사군의 증후군 척도는 경험적 자료를 기반으로 통계적 기법을 통해 함께 발생하는 문제들의 군집을 파악하여 구성한 것으로 범주적 진단 접근의 DSM방식과 보완

적인 장단점을 갖고 있는 것으로 비교되기도 하기 때문에 원저자는 임상장면에서의 활용을 고려하여 DSM방식의 척도를 추가하였다. DSM진단척도에는 정서문제, 불안문제, 전반적 발달문제, ADHD, 그리고 반항행동문제 등 5개의 척도가 포함된다.

2) 한국판 유아행동평가 척도-부모용(Korean version of the Child Behavior Checklist for ages 1.5-5; 한국판 CBCL 1.5-5)

한국판 C-TRF의 타당도 분석을 위해 Achenbach와 Rescorla(2000)가 개발한 CBCL 1.5-5를 오경자와 김영아(2009)가 한국어판으로 표준화한 한국판 CBCL 1.5-5를 함께 사용하였다. 앞서 언급한 바와 같이 C-TRF와 유사한 문항 및 요인 구조로 이루어진 한국판 CBCL 1.5-5는 0, 1, 2 척도의 100문항으로 7개의 증후군 척도(정서적 반응성, 불안/우울, 신체증상, 위축, 주의집중문제, 공격행동, 수면문제) 및 1개의 기타문제 척도와 5개의 DSM진단 척도(정서문제, 불안문제, 전반적 발달문제, ADHD, 반항행동문제)로 구성되어 있다. 또한 7개의 증후군 척도들 간의 합을 통해 구성된 상위 요인인 내재화, 외현화, 그리고 모든 증후군 척도와 기타문제 척도의 합으로 구성된 문제행동총점 척도가 있다. 전반적인 하위척도의 내적 합치도는 .56~.94 수준이었다(김영아 등, 2009).

자료 수집 절차 및 분석 방법

2006년부터 2007년에 걸쳐 자료 수집이 이루어졌다. 이는 한국판 CBCL 1.5-5를 표준화하기 위해 이루어진 자료 수집의 일환으로 이루어진 것이며 전반적인 절차는 한국판 CBCL 1.5-5의 자료 수집 과정과 동일하다. 전국을 5개 권역으로 나누어 대한민국 인구 주택 총조사(통계청, 2006)의 전국 지

역별 인구분포를 반영할 수 있는 유치원 및 보육기관을 선정하여 협조공문을 발송하고, 자료의 수집에 동의한 기관 22곳을 통해 자료를 수집하였다. 유치원에 보내진 설문지는 교사가 자신이 담당하고 있는 영유아에 대해 한국판 C-TRF를 실시하도록 하였다. 교사 1인당 너무 많은 영유아에 대한 평가를 실시할 경우 응답의 성실성에 문제가 생길 수 있으므로 교사 1인당 응답할 수 있는 C-TRF는 10개로 제한하였다. 더불어 교사가 특정 영유아를 선택하지 않도록 출석부 등과 같은 기관에서 보유한 영유아명단 번호로 임의로 선정된 영유아에 대해 평가하도록 하였다. 회수율은 약 96%였다. 이때 해당 기관 영유아의 부모에게 한국판 CBCL 1.5-5를 전달하여 실시하도록 하여 자료를 수집하였다. 검사-재검사 신뢰도를 산출하기 위해 6명의 교사로 하여금 최초로 한국판 C-TRF를 실시한지 3주 후에 다시 검사를 실시하도록 하여 61명의 재검사 자료를 얻을 수 있었다. 부모들에게는 검사 결과지와 도서를 제공하고, 교사들에게는 소정의 사례금을 제공하였다.

문제행동 문항의 결측값이 8개 이상 있는 자료는 제외하였으며, SPSS 16.0 for windows와 AMOS 16.0 평가판을 사용하여 신뢰도 및 타당도와 관련된 지표들을 분석하였다.

결 과

척도 구성

확인적 요인분석 C-TRF를 포함한 ASEBA 검사들은 미국에서 개발된 원 검사의 요인구조를 거의 유사하게 유지하면서 여러 문화권에서 널리 사용되고 있으며 그 임상적 유용성이 검증되어 왔다(Achenbach & Rescorla, 2000, 2001). 이처럼 동

일한 문항과 요인구조를 유지하여 검사가 표준화 될 경우, 문화권 간의 비교 연구가 용이하다는 장점이 있다. 따라서 C-TRF의 하위척도 구성이 우리나라의 영유아 집단을 평가하기 위해 필요한 요건들을 충족한다면 척도 구성을 동일하게 유지하는 것이 바람직할 것으로 보고, 이를 확인하기 위해 확인적 요인분석(CFA)을 실시하였다.

C-TRF와 같은 문제행동 평가척도를 임상군이 아닌 일반 정상집단에서 실시했을 경우, 표집의 대부분이 낮은 점수를 보고하는 정적 편포가 흔히 나타날 수 있다. 이러한 정적 편포는 요인 부하량에 영향을 미칠 수 있기 때문에, 척도의 원저자인 Achenbach와 Rescorla(2000)는 기준 집단 중앙값 이상의 자료를 확인적 요인분석에 포함시키고, 낮은 응답빈도에 의한 자료 해석의 왜곡을 방지하기 위해 0, 1, 2의 3점 척도를 0과 1의 이분 척도로 환산한 후 확인적 요인분석을 시행한 바 있다. 이번 한국판 C-TRF의 확인적 요인분석에서도 이와 동일한 방식을 적용하였다. 즉, 우리나라의 영유아를 대상으로 수집된 C-TRF의 표준화 자료에서 총점이 중앙값(12점) 이상인 703개의 사례를 포함하고, 기존의 3점 척도(0-1-2)를 0과 1의 이분 척도로 환산한 후 AMOS 16.0을 이용, 확인적 요인 분석을 실시하였다. 이 때 원자료를 통한 분석을 실시하였고 요인추출 방식은 Maximum Likelihood를 사용하였다. SRMR지수는 평균 .07로 한 개 요인 내 측정변수들의 단일 요인성 가정을 충족하였다. 모형의 효율성을 고려한 지수인 RMSEA는 .058, RMR은 .016로 적합한 수준을 나타내었다. 그 밖의 적합도 지수들(GFI=.716, CFI=.606, TLI=.590)은 기준에 충족되지 못했는데, C-TRF처럼 측정 항목이 많은 구조에서의 문항단위 분석에서 적합도가 저하되는 경향은 많은 연구에서 관찰되는 현상이다(Liu, Cheng, & Leung, 2011). 이러한 경우

측정 변수를 줄이기 위한 방안으로 소요인 총점을 측정 변수로 평가하는 항목묶음(parceling) 방식이 제시되기도 한다(Cai, Kaiser, & Hancock, 2004; Little, Cunningham, Shahar, & Widaman, 2002). 본 연구 자료에 이 방식을 적용하면 GFI=.99, CFI=.99, TLI=.97로 지표들이 상승되는 경향이 있었고, C-TRF 척도를 사용한 다른 연구들에서도 모형의 효율성을 고려한 지수가 적합한 수준일 경우 이를 근거로 요인구조의 적합성을 판단할 수 있다는 전제로 RMSEA값만을 기준으로 제시한 점등을 고려할 때(Achenbach & Rescorla, 2000), 한국판 C-TRF의 문항과 요인 간 구성 타당도는 일부 제약이 있기는 하나 원검사에서 검증된 수준의 적합도를 충족하고 있다고 판단된다.

신뢰도

한국판 C-TRF 하위 척도들의 재검사 신뢰도 및 내적 합치도(Cronbach's α)가 표 4에 제시되어 있다. 각 하위척도의 내적 합치도는 .52~.96의 범위를 보이고 있는데, 함께 제시된 미국 기준집단의 내적합치도 범위인 .52~.97과 거의 유사한 수준을 보였으며, 전반적인 신뢰도 수준은 양호한 편이었다. 3주 간격을 두고 실시한 재검사 상관은 .57~.82 범위로 나타났다. 미국판의 경우 8일 간격의 재검사 신뢰도를 조사한 바 있는데, 기간이 달라 수평적인 비교는 어려우나 .57~.91의 수준으로 대부분의 척도에서 한국판 C-TRF의 재검사 신뢰도보다 약간 높은 수준을 유지하였다. 미국판의 경우 신체증상 척도의 재검사 신뢰도가 .91로 가장 높았으나, 한국판 신체증상 척도 재검사 신뢰도는 .60으로 두 국가 간 자료의 차이가 가장 크게 나타났다.

척도의 신뢰도를 평가하기 위한 또 다른 방법으로 동일한 대상에 대하여 서로 다른 평가자의 평

표 4. 한국판 C-TRF의 검사-재검사 신뢰도 및 내적 합치도

요인	재검사 신뢰도		내적합치도		평정사간 일치도 CBCL1.5-5 / C-TRF	
	한국: 3주 간격 (n = 61)	미국: 8일 간격 (n = 59)	한국 (N = 1,371)	미국	한국	미국
정서적 반응성	.62**	.72**	.72	.71	.14**	.28**
불안/우울	.63**	.68**	.75	.76	.18**	.28**
신체증상	.60**	.91**	.52	.52	.05	.30**
위축	.59**	.77**	.80	.83	.17**	.29**
주의집중문제	.57**	.84**	.86	.89	.16**	.51**
공격행동	.82**	.89**	.93	.96	.19**	.55**
기타문제	.65**	-	.83	-	.09**	-
내재화	.69**	.77**	.89	.89	.15**	.30**
외현화	.79**	.89**	.94	.96	.19**	.58**
문제행동총점	.72**	.88**	.96	.97	.13**	.50**
DSM정서문제	.68**	.76**	.70	.78	.09**	.21**
DSM불안문제	.60**	.57**	.71	.68	.22**	.26**
DSM전반적발달문제	.64**	.83**	.79	.80	.23**	.42**
DSM ADHD	.65**	.79**	.88	.92	.17**	.52**
DSM반항행동문제	.78**	.87**	.81	.93	.17**	.42**

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

가 결과가 얼마나 일치하는지 알아보고자 하였다. 이를 위해 한 영유아에 대해 교사는 한국판 C-TRF, 주 양육자는 한국판 CBCL 1.5-5를 작성한 총 1,185개 사례의 점수 간 상관관계를 알아보았다. 상관분석은 한국판 C-TRF와 한국판 CBCL 1.5-5의 공통문항 82개에 대해 실시되었다. 그 결과 문제행동총점 척도에서는 $r=.13$, $p<.01$, 내재화, 외현화 척도는 각각 $r=.15$, $p<.01$ 과 $r=.19$, $p<.01$ 의 상관관을 보였다. 그리고 신체증상($r=.05$, $p=.08$)을 제외한 나머지 소척도들도 .09~.23 범위로 모두 통계적으로 유의미한 고른 상관관을 보이긴 하였으나, 그 크기는 미국판 C-TRF에 비해 전반적으로 낮은 수준이었다(표 4). 미국의 경우 가정을 방문하여 CBCL 1.5-5를 실시하고 부모를 통해 자녀의

교사에게 C-TRF를 실시하게 한 자료를 수집하였으나, 우리나라의 경우 유아교육 기관을 중심으로 교사 및 부모들에게 직접 자료를 수집하여 한 명의 교사가 여러 명의 유아를 평가하게 된 상황이 평정사간 일치도 수준에 영향을 미쳤을 가능성을 고려할 필요가 있다.

타당도

수렴타당도

척도의 수렴타당도를 알아보기 위해 Pearson의 적률상관계수를 산출한 결과, 모든 척도들은 서로 $p < .01$ 수준에서 유의한 상관관계를 보였으며, 문제행동총점 척도는 나머지 증후군 척도 및 DSM

표 5. 한국판 C-TRF의 각 하위척도 간 상관(N = 1,371)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.00														
2	.76	1.00													
3	.37	.34	1.00												
4	.53	.58	.20	1.00											
5	.46	.40	.16	.57	1.00										
6	.59	.45	.24	.41	.67	1.00									
7	.71	.67	.36	.65	.73	.77	1.00								
8	.86	.89	.47	.82	.55	.55	.78	1.00							
9	.59	.47	.23	.50	.83	.97	.82	.60	1.00						
10	.78	.72	.37	.70	.80	.87	.95	.85	.92	1.00					
11	.65	.71	.28	.76	.49	.45	.68	.81	.50	.71	1.00				
12	.73	.87	.31	.53	.40	.40	.65	.80	.43	.66	.60	1.00			
13	.68	.67	.28	.88	.51	.42	.68	.86	.48	.72	.68	.63	1.00		
14	.52	.43	.22	.50	.92	.82	.81	.56	.92	.87	.45	.42	.46	1.00	
15	.59	.43	.25	.42	.63	.92	.72	.55	.90	.83	.47	.36	.42	.74	1.00

1. 정서적 반응성, 2. 불안/우울, 3. 신체증상, 4. 위축, 5. 주의집중문제, 6. 공격행동, 7. 기타문제, 8. 내재화, 9. 외현화, 10. 문제 행동총점, 11. DSM 정서문제, 12. DSM 불안문제, 13. DSM 전반적 발달문제, 14. DSM ADHD, 15. DSM 반항행동문제.
표에 제시된 모든 상관관계는 $p < .01$ 수준에서 유의미하였음.

진단 척도들과 각기 .37~.95 사이의 비교적 높은 상관을 보였다. 한국판 C-TRF의 증후군 척도들은 크게 내재화 척도와 외현화 척도로 나누어지는데, 내재화 척도로 분류되는 증후군 척도들(정서적 반응성, 불안/우울, 신체증상, 위축)은 상대적으로 그 크기가 낮았던 신체증상($r=.47$)을 제외하고는 내재화 척도 점수와 .82~.89사이의 상관을 나타내었다. 외현화 척도의 하위 척도인 주의집중문제와 공격 행동 척도의 상관은 .67로 다른 내재화 척도 요인(.24~.59)과의 상관에 비해 높은 상관을 나타내고 있다(표 5).

변별타당도

유아기에 임상기관에 내원한 경험과 전문가의 도움을 받은 경험이 있는지 여부가 현재의 문제 행동 수준을 말해주다고 보기는 어려우나, 적응상의 문제를 보이고 있을 가능성을 고려하여 원저자는 임상기관에 내원한 적이 있거나 전문가의 도움을

받았던 적이 있는 것으로 보고된 임상집단(referred sample)을 정상 기준집단 구성에서 제외하였다. 한국판 C-TRF의 하위척도 점수에서 임상기관을 방문 경험이 있는 임상집단과 비임상집단(non-referred sample) 간의 점수 비교에서 임상집단이 더 높은 것으로 확인된다면, 임상집단 영유아의 문제 행동 수준을 제대로 측정하고 있음을 알려주는 하나의 기준이 될 수 있을 것이다. 물론 이러한 분류는 임상가의 진단이나 구조화된 면담도구에 의해 진단이 내려진 일반적인 임상 연구에서의 임상군과는 차이가 있을 것이다. 임상기관 방문 여부 문항에 '모른다'라고 응답하거나 응답을 하지 않은 경우는 제외하고, 임상기관에 방문한 적이 없다고 보고된 집단 1,209명의 자료와, 임상기관에 방문한 적이 있다고 보고된 집단 33명의 자료 간의 한국판 C-TRF 점수 차이를 비교해보았다(표 6). 그 결과, 임상집단의 문제행동총점 평균은 36.39점으로 비임상집단의 17.66점에 비해 높은 것으로 나타났다.

표 6. 임상(referred) 집단과 비임상(non-referred) 집단의 한국판 C-TRF 척도 평균

	비임상집단 (n = 1,209)	임상집단 (n = 33)	F (차이검증)	Cohen's D (효과크기)
정서적반응성	1.43 (1.95)	2.67 (2.19)	12.83**	.63
불안/우울	1.80 (2.19)	3.00 (2.30)	9.63**	.55
신체증상	.39 (.89)	.76 (1.58)	5.18*	.40
위축	1.77 (2.52)	5.18 (3.72)	57.08**	1.30
주의집중문제	2.12 (3.01)	5.55 (5.12)	39.65**	1.09
공격행동	4.83 (6.63)	8.09 (8.18)	7.67**	.49
기타문제	5.33 (5.26)	11.15 (7.87)	38.15**	1.07
내재화	5.39 (6.08)	11.61 (6.66)	33.44**	1.01
외현화	6.95 (8.90)	13.64 (12.24)	17.75**	.74
문제행동총점	17.66 (18.23)	36.39 (24.56)	33.24**	1.00
DSM정서문제	1.04 (1.64)	2.52 (2.39)	25.34**	.88
DSM불안문제	1.06 (1.66)	2.03 (1.74)	10.99**	.58
DSM전반적발달문제	2.31 (2.94)	7.00 (4.42)	79.28**	1.52
DSM ADHD	3.51 (4.34)	7.24 (6.41)	23.08**	.84
DSM반항행동문제	1.22 (2.01)	2.33 (2.64)	9.71**	.55

() 안은 표준편차, * $p < .05$, ** $p < .01$

표 7. 한국판 C-TRF 총점 기준 임상-비임상 집단 간의 한국판 CBCL 1.5-5 하위척도별 점수평균 및 차이 비교

	C-TRF 총점기준 비임상 (n = 980)	C-TRF 총점기준 임상 (n = 205)	전체집단 (N = 1,185)	F
CBCL 정서적반응성	2.33 (2.27)	2.92 (2.44)	2.43 (2.31)	11.00**
CBCL 불안/우울	2.63 (2.20)	3.10 (2.48)	2.71 (2.25)	7.19**
CBCL 신체증상	1.58 (1.72)	1.86 (1.95)	1.63 (1.76)	4.10*
CBCL 위축	1.53 (1.70)	2.01 (2.08)	1.61 (1.78)	12.31**
CBCL 수면문제	2.47 (2.19)	2.79 (2.30)	2.52 (2.21)	3.65
CBCL 주의집중문제	1.41 (1.50)	1.77 (1.55)	1.48 (1.52)	9.16**
CBCL 공격행동	7.10 (5.29)	8.96 (5.83)	7.42 (5.43)	20.27**
CBCL 기타문제	7.60 (4.89)	8.92 (5.28)	7.83 (4.98)	12.09**
CBCL 내재화	8.08 (6.18)	9.88 (7.47)	8.39 (6.45)	13.35**
CBCL 외현화	8.51 (6.25)	10.73 (6.82)	8.90 (6.40)	20.63**
CBCL 문제행동총점	26.66 (17.02)	32.32 (18.79)	27.64 (17.47)	18.09**
CBCL DSM정서문제	2.37 (2.01)	2.94 (2.16)	2.47 (2.04)	13.14**
CBCL DSM불안문제	3.50 (2.60)	4.11 (2.98)	3.60 (2.68)	8.88**
CBCL DSM전반적발달문제	2.78 (2.58)	3.32 (3.15)	2.88 (2.69)	6.83**
CBCL DSM ADHD	2.86 (2.22)	3.54 (2.33)	2.98 (2.25)	15.33**
CBCL DSM반항행동문제	2.11 (1.93)	2.67 (1.99)	2.20 (1.96)	14.23**

() 안은 표준편차, * $p < .05$, ** $p < .01$

또한 모든 하위 척도에서 임상집단은 유의하게 높은 점수를 보였으며, Cohen(1988)의 효과크기로 평가한 결과 모두 .40 이상의 효과크기를 나타내어, 중간 이상의 수준을 보였다. 임상집단 33명 중 남아는 21명으로 비임상 집단에 비해 남아 비율이 더 높았는데, 일반적으로 남아의 문제 행동 점수가 높기는 하지만 본 연구의 임상집단에서는 남아의 문제행동총점 평균이 31.10, 여아는 45.64로 단순히 남아가 더 많음으로 나타난 점수 차이는 아닌 것으로 보인다.

또한 한국판 C-TRF의 총점을 기준으로, C-TRF의 원저자가 준임상 기준으로 제시하고 있는 60T (84%ile) 이상의 점수를 보고한 집단과 60T 미만의 점수를 보고한 집단 간의 한국판 CBCL 1.5-5의 점수 차이를 비교해 보았다. 즉, 한국판 CBCL 1.5-5를 준거타당도 척도로 활용하여 한국판 C-TRF의 변별력을 확인해보고자 하였다. 이를 위해 한국판 C-TRF 기준집단 1,371명의 자료 중에서, 한국판 CBCL 1.5-5 자료가 함께 있는 1,185명의 자료에 대해서 분석을 실시하였다. 그 결과, 한국판 C-TRF의 총점에서 60T 이상을 받은 집단이 그렇지 않은 집단보다 한국판 CBCL 1.5-5의 수면문제 하위척도를 제외한 모든 하위척도들에 대해서 통계적으로 유의미하게 높은 점수를 보고하였다(표 7). 한국판 CBCL 1.5-5의 수면문제 척도에 대해서 유의미한 차이가 나타나지 않은 것은 C-TRF의 경우 미국판 원검사나 한국판 검사에서는 영유아 교사가 직접적으로 관찰하기 어려운 영유아의 수면 행동과 관련된 ‘수면문제’ 하위 척도가 없어 총점에 반영되지 않은 점이 영향을 미친 것으로 추정된다.

마지막으로 한국판 CBCL 1.5-5를 준거 검사로 사용하여 한국판 C-TRF의 변별력을 검증하기 위해 판별분석을 실시하였다. 이 분석에서도 한국판

CBCL 1.5-5의 임상집단 분류 기준으로 60T 점수를 사용하였다. 그 결과, 판별 함수의 정준 상관 계수(Canonical Correlation)는 .11로 통계적으로 유의미하였으며(Wilks $\lambda = .99, p < .05$), 비임상집단과 임상집단을 정확하게 분류해내는 비율은 65.1%인 것으로 나타났다(표 8). 이러한 결과는 한국판 C-TRF를 이용하여 임상집단과 비임상집단을 판별하는 것이 가능하며, 임상집단 선별 도구로서의 유용성이 지지되고 있음을 의미한다.

하위 척도의 평균

한국판 C-TRF 척도의 특성을 파악하고 기준을 작성하기 위해 수집된 기준집단의 전체 평균 및 성별, 연령별 평균 점수를 확인하고 미국판 원검사의 기준집단이 보고한 평균 점수 수준과 비교해 보았다. 한국판 C-TRF 기준 집단 남녀별 하위 척도 평균 점수 및 미국판 원검사 기준 집단의 남녀별 C-TRF 하위 척도 평균 점수와 그 차이 수준이 표 9에 제시되어 있다. 먼저 한국 기준집단에서 남녀 간의 차이를 검증해본 결과, 신체증상($F(1, 1369)=4.63, p<.05$), 위축($F(1, 1369)=23.53, p<.01$), 주의집중문제($F(1, 1369)=41.28, p<.01$), 공격행동($F(1, 1369)=13.48, p<.01$), 기타문제($F(1, 1369)=5.05, p<.05$), 외현화($F(1, 1369)=23.98, p<.01$), 문제행동총점($F(1, 1369)=11.00, p<.01$), DSM 전반적발달문제($F(1, 1369)=11.12, p<.01$), DSM ADHD($F(1,$

표 8. 한국판 C-TRF의 모든 하위척도를 이용하여 CBCL 1.5-5로 분류된 임상·비임상 집단 여부를 분류한 결과

실제집단	사례수	예측된 집단	
		비임상	임상
비임상	992	683 (68.85)	309 (31.15)
임상	193	105 (54.40)	88 (45.60)

() 안의 값은 % . 집단을 정확하게 분류한 비율은 65.10%임.

표 9. 한국판 C-TRF의 남녀별 평균 및 미국판 C-TRF의 규준집단과의 척도별 평균 비교

	한국전체 (N = 1,371)			미국자료 (N = 1,192)		한국-미국 차이 (t 값)	
	남	여	남녀 차 F	남	여	남	여
	(n = 679)	(n = 692)		(n = 588)	(n = 604)		
정서적 반응성	1.35 (1.85)	1.53 (2.06)	2.69	1.50 (1.90)	1.30 (1.90)	1.42	-2.09*
불안/우울	1.76 (2.20)	1.85 (2.22)	.60	2.10 (2.20)	2.20 (2.40)	2.74**	2.71**
신체증상	.35 (.84)	.45 (.98)	4.63*	.50 (1.00)	.70 (1.20)	2.86**	4.07**
위축	2.16 (2.85)	1.49 (2.30)	23.53**	2.80 (3.20)	2.30 (2.90)	3.73**	5.52**
주의집중문제	2.68 (3.39)	1.62 (2.63)	41.28**	3.60 (3.70)	2.60 (3.40)	4.58**	5.74**
공격행동	5.62 (7.11)	4.29 (6.29)	13.48**	6.90 (8.50)	5.30 (7.60)	2.88**	2.58**
기타문제	5.77 (5.75)	5.11 (5.01)	5.05*	-	-	-	-
내재화	5.63 (6.29)	5.32 (6.18)	.82	6.80 (6.60)	6.40 (6.90)	3.21**	2.95**
외현화	8.30 (9.68)	5.92 (8.30)	23.98**	10.50 (11.30)	8.00 (10.10)	3.69**	4.02**
문제행동총점	19.69 (19.75)	16.35 (17.52)	11.00**	23.10 (20.90)	19.60 (20.90)	2.97**	3.01**
DSM정서문제	1.12 (1.81)	1.00 (1.55)	1.85	1.20 (2.00)	1.20 (1.90)	.74	2.06*
DSM불안문제	1.06 (1.66)	1.10 (1.73)	.19	1.10 (1.60)	1.30 (1.90)	.44	1.97*
DSM전반적발달문제	2.62 (3.14)	2.08 (2.82)	11.12**	3.20 (3.20)	2.60 (3.00)	3.24**	3.20**
DSMADHD	4.18 (4.71)	2.92 (3.99)	28.70**	5.50 (5.50)	4.10 (4.90)	4.54**	4.71**
DSM반항행동문제	1.46 (2.16)	1.07 (1.91)	12.10**	2.10 (3.00)	1.70 (2.70)	4.29**	4.78**

() 안은 표준편차, * $p < .05$, ** $p < .01$

1369)=28.70, $p < .01$), 그리고 DSM 반항행동문제 ($F(1, 1369)=12.10$, $p < .01$) 하위척도에서 유의미한 차이가 나타났다. 대체로 여아는 남아에 비해 외현화 문제로 분류되거나 외현화 문제와 관련된 하위 척도에서 문제행동 수준이 더 낮게 보고되는 경향이 있었다.

미국과 한국 영유아의 C-TRF 평균 자료를 Simple t test로 비교한 결과(Uitenbroek, 1997), 전반적으로 미국의 자료에 비해 문제행동 수준이 낮은 것으로 나타났다. 특히 주의집중문제, 공격행동

과 같은 외현화 문제와 관련된 하위척도의 경우 그 차이가 더욱 크게 나타나 남녀 모두 한국 영유아들이 미국의 영유아에 비해 낮은 수준의 문제를 보이고 있었다. 이는 한국 영유아가 외현화 문제에 있어서 미국의 영유아보다 낮은 수준의 문제 행동을 보고하는 것으로 나타났던 한국판 CBCL 1.5-5의 연구에서도 관찰되었던 결과(김영아 등, 2009)이며, 주로 공격행동과 같은 외현화 문제에서는 미국에 비해 동양권 국가에서 점수가 낮게 관찰된다는 현상과도 일치하는 결과이다(Achenbach &

표 10. 한국판 C-TRF의 연령대별 하위척도 점수 비교

	1세 (n = 48)	2세 (n = 209)	3세 (n = 347)	4세 (n = 363)	5세 (n = 306)	6세 (n = 98)	F
정서적 반응성	2.35 (2.02)	2.05 (2.27)	1.46 (1.92)	1.34 (2.01)	1.16 (1.75)	.88 (1.36)	9.45**
불안/우울	2.50 (2.12)	2.50 (2.64)	1.79 (2.16)	1.70 (2.22)	1.53 (1.99)	1.34 (1.63)	7.21**
신체증상	.10 (.37)	.46 (.91)	.41 (.88)	.47 (1.09)	.34 (.81)	.31 (.79)	2.04*
위축	2.79 (3.33)	2.79 (3.23)	1.96 (2.64)	1.68 (2.58)	1.27 (1.90)	1.03 (1.78)	12.53**
주의집중문제	2.58 (2.75)	3.47 (3.54)	2.12 (2.93)	2.19 (3.27)	1.57 (2.63)	.83 (1.99)	14.42**
공격행동	7.67 (4.44)	8.41 (8.32)	4.93 (6.17)	4.54 (6.92)	3.41 (5.75)	2.69 (4.73)	19.46**
기타문제	7.13 (3.88)	7.91 (6.27)	5.65 (5.11)	5.26 (5.65)	4.11 (4.60)	3.36 (4.01)	17.48**
내재화	7.75 (6.08)	7.80 (7.56)	5.61 (6.12)	5.20 (6.32)	4.31 (5.20)	3.55 (4.27)	11.75**
외현화	10.25 (6.01)	11.88 (10.92)	7.04 (8.30)	6.73 (9.48)	4.98 (7.80)	3.52 (6.16)	20.62**
문제행동총점	25.13 (13.44)	27.59 (22.51)	18.31 (17.38)	17.19 (19.45)	13.40 (15.76)	10.43 (12.54)	20.83**
DSM정서문제	1.63 (1.94)	1.62 (2.04)	1.21 (1.85)	.96 (1.54)	.66 (1.26)	.67 (1.29)	11.33**
DSM불안문제	2.02 (1.63)	1.75 (2.17)	1.04 (1.60)	.95 (1.73)	.80 (1.36)	.61 (1.07)	13.72**
DSM전반적발달문제	3.63 (3.30)	3.55 (3.66)	2.50 (2.93)	2.11 (3.00)	1.70 (2.35)	1.46 (2.21)	14.33**
DSM ADHD	4.85 (3.40)	5.50 (5.13)	3.56 (4.16)	3.54 (4.66)	2.58 (3.77)	1.72 (3.05)	16.23**
DSM반향행동문제	2.13 (1.76)	2.32 (2.57)	1.27 (1.95)	1.07 (1.94)	.84 (1.72)	.59 (1.60)	19.38**

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

Rescorla, 2000).

규준의 작성에 앞서 연령대에 따라 다른 규준을 적용하는 것이 필요할지를 결정하기 위해 연령대별로 한국판 C-TRF의 점수 수준을 비교해본 결과가 표 10에 제시되어 있다. 분석 결과, 전반적인 한국판 C-TRF의 하위척도에서 연령대별로 유의미한 차이가 나타났다. 연령이 어릴수록 문제행동 점수가 높은 경향이 있는데, 특히 만 2세(24개월

이상~36개월 미만)와 3세(36개월 이상~48개월 미만) 사이, 즉 36개월을 전후로 해서 전반적인 하위 척도들에 대해 가장 급격한 점수변화를 보이고 있다. 이와 같은 결과는 한국판 CBCL 1.5-5의 개발 연구(김영아 등, 2009)를 비롯하여 손원경과 황해익(2004), 송진숙(2000), Shelov(1998) 등의 연구에서 영유아 발달상의 중요한 차이가 나타나는 시기가 2~3세 사이였다는 점과 일치한다.

표 11. 36개월 전·후, 그리고 남·녀별 집단의 한국판 C-TRF 하위척도 평균 비교

	남아 (n = 679)		여아 (n = 692)	
	36개월 전후 집단간 차이		36개월 전후 집단간 차이	
	점수차	F	점수차	F
정서적 반응성	-0.51	12.93**	-0.54	11.88**
불안/우울	-0.56	11.00**	-0.45	6.99*
신체증상	-0.12	3.45	0.11	2.10
위축	-1.14	27.79**	-0.62	12.37**
주의집중문제	-0.69	6.94**	-1.02	26.46**
공격행동	-1.78	10.63**	-3.19	46.86**
기타문제	-1.91	18.89**	-2.07	30.19**
내재화	-2.33	23.70**	-1.50	10.12**
외현화	-2.47	11.05**	-4.21	46.78**
문제행동총점	-6.70	19.83**	-7.77	35.20**
DSM정서문제	-0.70	25.70**	-0.46	15.51**
DSM불안문제	-0.54	18.17**	-0.50	14.41**
DSM전반적발달문제	-1.30	30.08**	-0.87	16.57**
DSM ADHD	-1.05	8.37**	-1.77	35.01**
DSM반항행동문제	-0.65	15.65**	-0.91	41.00**

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

이번에 표준화한 C-TRF의 경우, 한국판 CBCL 1.5-5의 경우에서보다 남녀 간 차이가 보다 두드러지게 나타나고 있으므로, 영유아의 성별에 따른 기준 집단을 산출함과 동시에, 한국판 CBCL 1.5-5와 마찬가지로 36개월을 기준으로 하여 기준 집단을 구별하는 것이 타당할 것으로 보인다. 성별 및 연령별로 한국판 C-TRF 하위척도 점수간의 차이가 유의미한지를 검증하기 위해 집단 간 ANOVA를 실시한 결과는 표 11에 제시되어 있다. 그 결과, 신체증상을 제외한 모든 하위척도에서 남녀 집단 각각에 대해 36개월 전후로 뚜렷한 차이를 보이고 있음을 확인할 수 있으며, 성별과 연령에 따른 개별적인 기준을 마련하는 것이 바람직함을 지지해 준다. 미국판 원 검사의 경우에도, CBCL 1.5-5의 경우에는 남녀 집단 간 기준을 따로 산출하지 않은 반면, C-TRF 검사에 대해서는 남·녀 집단 간

기준을 개별적으로 산출하여 제시하고 있다.

논 의

본 연구에서는 Achenbach와 Rescorla(2000)가 개발한 C-TRF를 우리말로 번역하여 한국판으로 개발하고, 전국의 영유아를 대상으로 표준화 작업을 진행하였으며, 측정 도구로서의 신뢰도와 타당도를 살펴보았다. 이번 표준화 연구를 통해 확인한 내용은 다음과 같다.

우선 본 논문의 목적이 미국에서 개발된 유아 교사용 척도를 한국판으로 개발하는 것이므로, 원 척도의 구조가 한국의 영유아를 대상으로 한 경우에도 유용성이 있는지 확인하기 위해 확인적 요인 분석을 실시하였다. 그 결과, 적합도 지수들이 모두 기준에 부합하지는 못했다. 그러나 복잡한 요인

구조와 많은 문항이 특정 지수에 대한 적합도 수준을 저하시킬 수 있다는 점(Little, Cunningham, Shahar, & Widaman, 2002)과 관련하여 원저자가 모델 검증 방식에서 사용한 적합도 지수인 RMSEA는 .058로 적합한 수준으로 나타났다. 또한 문항별 요인부하량 수준 등의 기타 통계 지표들이 양호하고, 이미 개발된 한국판 CBCL 1.5-5와의 연계 활용 및 비교문화 연구의 가능성 등을 총체적으로 고려할 때 미국판 원 척도와 동일한 문항과 요인 구조를 유지하는 형태로 개발하는 것이 바람직할 것으로 판단하였다. 향후 한국판 C-TRF가 유아교육 현장에서 활용되며 누적되는 자료들을 통해 보다 유용성이 있는 구조가 발견되면 이를 반영할 수 있을 것이다.

두 번째로, 내적 일치도(Cronbach's α) 및 재검사 신뢰도는 하위척도에 따라 차이는 있었지만 전반적으로 적절한 수준인 것으로 나타나고 있다. C-TRF는 개별 문항들도 임상적으로 각각 의미가 있게 평가하고 확인하도록 되어 있기 때문에 도구로서의 신뢰도 수준이 지지되었다.

반면 영유아에 대한 부모 평가와 교사 평가의 상관관계를 통해서 본 일치도는 통계적으로 유의미하긴 하였지만 .05~.23의 수준으로 낮게 나타났다. 평정자간 시차가 다를 수 있다는 전제 자체가 교사를 통한 평가척도의 개발 목적이기도 한 점을 고려할 때, 한국판 C-TRF 및 CBCL 1.5-5를 통해 나타난 교사와 부모 평가의 낮은 상관 결과는 유아 행동 자료를 수집할 때 부모 뿐 아니라 교사의 보고도 함께 참고해야할 필요성을 시사한다고 볼 수 있다. 실제로 영유아 문제 행동에 대한 부모-교사 평정 간 일치도와 관련된 기존 연구들은 부모와 교사 평가 간에 상당한 불일치가 나타남을 밝히고 있다(Hay, Pawlby, Sharp, Schmücker, Mills, Allen, & Kumar, 1999; Kumpulainen, Rasanen,

Henttonen, Moilanen, Piha, Puura, Tamminen, & Almqvist, 1999; Phares, 1997; Stanger & Lewis, 1993; Winsler & Wallace, 2002; Youngstrom, Loeber, & Stouthamer-Loeber, 2000). 이는 임상 장면에서 주로 부모의 보고에 의한 행동평가 자료를 활용하지만 이것이 제한적일 수 있으며 교사의 평가 자료를 함께 확인할 수 있다면 유아의 행동에 대한 포괄적인 정보를 파악할 수 있음을 말해준다. 이러한 현상이 부모와 교사의 영유아 문제 행동 평가에 대한 내적 기준의 차이에서 기인한 결과인지, 영유아가 가정과 교육기관에서 전혀 다른 행동을 보이기 때문인지, 그 구체적인 원인을 파악할 수 있다면 다중 평가자의 평가 결과를 통합하여 영유아의 행동을 정확하게 파악하고 이해하는데 도움이 될 수 있을 것이다.

셋째, 구성 타당도를 평가하기 위해 실시한 한국판 C-TRF 하위척도들 간의 상관관계는 모두 통계적으로 유의한 것으로 나타났고, 특히 내재화 척도와 이를 구성하는 대부분의 하위척도들 간의 상관관계 및 외현화 척도를 구성하는 주의집중문제, 공격행동 소척도 간의 상관관계가 상대적으로 높아 한국판 C-TRF의 수렴 타당도가 지지되었다. 다만, 예외적으로 신체증상 척도만이 내재화 척도와 높지 않은 상관을 보여, 교사들이 유아를 정서적 반응성, 불안/우울, 위축 차원에서 평가하는 것과 신체증상 문제에 대해 평가하는 것 사이에는 상관관계가 약한 것으로 보인다.

신체증상 척도에서 신뢰도나 구성타당도가 낮게 나타나며 원척도와 다른 특성을 보이는 현상은 한국판 CBCL 1.5-5 연구에서도 보고된 바 있다(김영아 등, 2009). 따라서 신체증상 척도 점수에 대해서는 제한적으로 해석해야 하며 소척도 점수 뿐 아니라 점수 상승을 가져온 개별문항들을 함께 검토하는 것이 필요할 것이다. 또한 우리나라 성인

들이 유아의 신체증상을 평가하는 기준에 있어 어떠한 문화적인 차이가 있는지 탐색하는 연구가 필요할 것으로 보인다.

넷째, 한국판 C-TRF를 이용한 교사 보고를 통해 임상적으로 의미 있는 문제행동을 보이는 영유아들을 선별할 수 있는지 평가하고자 영유아의 임상기관 방문 경험 여부가 있는 집단과 나머지 기준집단에서 교사평가 점수 차이 검증을 실시한 결과 한국판 C-TRF의 임상적 변별력이 지지되었다. 또한 C-TRF의 문제행동총점 임상범위 기준인 60T 이상·미만 집단에 대한 CBCL 1.5-5 하위척도의 평균 점수를 비교하고 판별분석을 실시해본 결과에서도 C-TRF의 변별력이 지지되었다. 대상 집단 전체로 볼 때는 부모와 교사 보고 간 상관성이 높지 않은 것으로 나타났음에도 불구하고, 교사에 의해 문제행동이 많이 보고되는 상위 점수대 영유아들은 부모도 문제행동을 많이 보고할 가능성이 높은 것으로 나타나, 영유아 교육 기관에서 C-TRF를 활용한 교사 평가를 통해 문제행동을 보이는 영유아들을 선별하는 것이 유용함을 말해준다. 다만 이는 개발 단계에서 최소한의 변별력을 평가해 보기 위해 원검사에서 가이드라인으로 제시한 절단점에 대해 문제가능성이 높은 집단과 그렇지 않은 집단이 차이를 보이는지 평가해 본 결과라는 것을 염두에 두어야 한다. 절단점은 일반적으로 낮춰 잡으면 문제가 있는 집단이 빠짐없이 선별될 가능성이 높아지는 대신 정상집단을 포함할 가능성이 높아지고, 높은 절단점을 사용하면 반대의 현상이 일어나므로 절대적 기준이 있다기보다 사용 목적에 따라 적합한 절단점을 선택할 수 있다. 향후 정확한 진단으로 구분된 다양한 임상집단에 대한 연구가 뒤따르며 여러 절단점에 대해 민감도와 특이도를 반영한 변별력 정보가 제공될 때 척도의 임상적 활용도가 더욱 높아질 것이다.

마지막으로 척도의 평균을 성별과 연령 집단에 대하여 각각 비교한 결과, 주로 외현화 문제에서 성별 차이가 나타났고, 36개월 연령대 집단을 전후로 가장 많은 하위척도에서 유의미한 차이가 나타났다. 부모가 평가하는 CBCL 1.5-5에 비해 교사평가인 C-TRF에서 상대적으로 더 영유아의 성별에 따른 문제행동 점수 차이가 더 큰 것으로 나타났는데 이는 한국과 미국 기준 집단 모두에서 나타나는 현상으로 보인다(김영아 등, 2009; Achenbach & Rescorla, 2000). 가정에서 주로 자신의 자녀만 상대하며 자녀와 같은 성의 또래와 비교하여 평가할 가능성이 많은 부모에 비해, 여러 영유아들을 평가하는 교사는 영유아들을 상대적으로 평가하면서 성별에 따른 유아의 문제행동 비율을 달리 보고하게 된 것일 수 있다. 또한 남아와 여아가 가정과 유아교육기관에서 달리 행동할 가능성도 배제할 수는 없어, 향후 원인에 대한 탐색이 필요한 것으로 보인다.

한국판 C-TRF는 1.5~5세 연령대 영유아를 대상으로 한 검사이긴 하지만, 영유아 시기일수록 짧은 시간 내에 큰 발달적 변화가 이루어질 수 있고, 행동 수준에 대한 기대치가 그만큼 달라질 수 있기 때문에 해당 성별, 연령대를 가장 잘 반영할 수 있는 기준이 마련되어야 검사의 유용한 활용에 도움이 될 수 있을 것으로 판단하였다. 이에 따라 한국의 기준집단을 대상으로 한 검사 결과에서 중요한 차이가 있는 것으로 나타난 성별(남·여), 연령별(36개월 미만·이상) 집단에 대해 개별적인 기준 점수를 마련하는 것이 바람직할 것으로 판단된다.

이번 연구는 우리나라의 영유아 교육 및 연구 현장에서 널리 사용될 수 있는 교사 평가용 유아 문제행동 평가도구를 번안하고 전국의 영유아를 대상으로 표준화한 후, 측정 도구로서의 신뢰도와 타당도를 살펴본 데 의의가 있다. 영유아 발달 선

별 및 진단 검사도구가 충분치 않은 상황에서(이준석, 조광순, 2004) 한국의 전체 영유아 인구분포를 반영한 폭넓은 기준집단을 바탕으로 개발된 한국판 C-TRF는 먼저 개발된 한국판 CBCL 1.5-5와 함께 활용됨으로써 영유아 문제 행동에 대한 객관적이고 다각적인 차원에서의 이해를 증진시키는데 도움이 될 수 있을 것으로 기대된다. 또한 영유아기 이후, 아동·청소년기 연령대와 연속성을 가질 수 있는 유사한 체계의 ASEBA 검사들을 발달적 연속선상에서 함께 활용할 수 있다는 점은 추후 장기 종단적 평생 발달 연구를 가능케 한다는 측면에서 의미가 있으며, 미국판 검사와의 동일성을 유지하고 있기 때문에 비교 문화 연구에서도 활용될 수 있을 것이다.

본 연구를 통해 한국판 C-TRF의 측정 도구로서 기본적인 요건들은 확인되었으나 향후 추가적인 연구를 통해 보완되어야 할 부분들이 있다. 첫 번째는 평정자간 일치도에 대한 보다 상세한 연구가 필요하다는 점이다. 부모-교사 평가 간 상관관계가 낮은 수준에 머물렀던 이유가 무엇에 기인하는 것인지 파악할 필요가 있다. 성인의 보고를 통한 영유아행동 평가에 있어서 그 차이를 설명할 수 있는 기제가 밝혀지면 다중 평가 결과를 통합하는 ASEBA 행동 평가 척도의 활용도가 더욱 높아질 것이다. 둘째로 임상집단에 대한 변별력 검토가 임상기관에 내원한 영유아들을 대상으로 세분화된 진단 집단에 대해서 이루어져야 할 필요성이 있다. 구체적인 문제 행동을 변별하는데 한국판 C-TRF의 하위척도들이 의미가 있는지 검증될 때 한국판 C-TRF의 임상적, 학문적 활용이 보다 활발하게 이루어질 수 있을 것이다. 셋째, 우리나라 문화에 특수한 교사-영유아 상호관계 특성, 교육 환경 등의 요인을 고려하여 지속적으로 관찰되는 임상적으로 유용한 부가적인 하위척도가 있는

지 탐색해가며 우리나라 유아의 특성을 반영할 수 있는 연구가 필요하다.

참 고 문 헌

- 김영아, 이진, 문수종, 김유진, 오경자 (2009). 한국판 CBCL 1.5-5의 표준화 연구. *한국심리학회: 임상*, 28(1), 117-136.
- 박은희, 소유경, 김영신, 최낙경, 김세주, 노주선, 고윤주 (2003). 한국어판 Conners 부모 및 교사용 평가 척도의 신뢰도와 타당도에 대한 예비적 연구. *소아청소년 정신의학*, 14, 1-14.
- 소유경, 노주선, 김영신, 고선규, 고윤주 (2002). 한국어판 부모, 교사 ADHD 평가척도의 신뢰도와 타당도 연구. *신경정신의학*, 41(2), 283-289.
- 손원경, 황해익 (2004). 3, 4세용 유아발달검사 개발연구. *아동교육연구*, 24(5), 43-62.
- 송진숙 (2000). 영유아 발달 척도의 개발을 위한 기초 연구: 1-36개월 영유아를 중심으로. *한국생활과학회지*, 9(4), 405-419.
- 신미숙 (2004). 조기교육과 유치원교육에 대한 학부모의 견해 : 유아기 자녀를 둔 어머니들을 대상으로. *경남대학교 대학원 석사학위 청구 논문*.
- 신현숙, 박용재, 박주희, 류정희 (2006). 교사보고형 초등학생 학교적응행동 평정척도의 개발과 타당도 검증. *한국심리학회지: 학교*, 3(2), 1-26.
- 안정숙, 전성균, 한준규, 노경선 (2003). 한국어판 강점, 난점 설문지(Strengths and Difficulties Questionnaire) 개발. *신경정신의학*, 42(1), 141-147.
- 오경자, 김영아 (2009). CBCL 1.5-5 유아행동평가 척도-부모용. 서울: (주)휴노컨설팅.
- 오경자, 이혜련 (1989). 주의력 결핍 과잉 활동증 평가도구로서의 단축형 Conners 평가척도의

- 활용. 한국심리학회지: 임상, 8(1), 135-142.
- 윤진주, 안진경, 김지영 (2008). 유치원 교사의 교육과정 평가와 유아 평가에 대한 실태 및 개선방안에 관한 연구. 한국생활과학회지, 17(3), 413-423.
- 재인용 (이순형, 이성욱, 권혜진, 이소은, 황혜신, 이혜승, 한유진, 정윤주, 이영미, 이옥경, 성미영, 권기남, 김지현 (2007). 보육과정. 서울: 학지사.)
- 이경숙, 신의진, 신동주, 전연진, 박진아 (2003). 한국판 영유아 행동문제 평가척도(Korean Behavior Assessment System for Children; K-BASC) 표준화 예비연구. 한국심리학회지: 발달, 16(4), 175-191.
- 이준석, 조광순 (2004). 국내 영·유아 발달 선별 및 진단 검사도구에 대한 고찰 : 표준화중인 검사도구들을 중심으로. 특수교육저널: 이론과 실천, 5(1), 1-26.
- 이진, 김영아, 오경자 (2009). 한국판 CBCL1.5-5의 변별타당도 및 임상적 유용성. 한국심리학회: 임상, 28(1), 171-186.
- 이춘재, 성현란, 송길연, 혜경, 김혜리, 박혜원, 광금주, 장유경, 박선미, 황상민, 이은희, 이도현, 조정호, 김선아 공역 (2002). 발달정신병리학: 영아기부터 청소년기까지. (pp.88-93). 중앙적성출판사. (Wenar, C. (1994). *Developmental Psychopathology: from infancy through adolescence* (3rd ed.). McGraw-Hill, Inc.)
- 이혜원, 양옥승 (2006). 교사용 유아의 사회적 유능감 평가척도 개발 연구, 유아교육연구, 26(2), 69-92.
- 정현희 (2002). 한국판 교사용-아동평정척도(T-CRS)의 신뢰도와 타당도. 한국심리학회지: 발달, 15(2), 73-88.
- 조광순, 최지영 (2003). 정서·행동장애 진단도구 BASC 4-5세 유아용 부모평정척도 국내 표준화 1차 연구. 특수교육학연구, 38(1), 291-314.
- 통계청 (2006). 2005 대한민국 인구 주택 총조사.
- 황혜정 (2005). 유아의 문제행동 진단을 위한 부모교사용 검사 개발 연구. 아동교육연구, 25(2), 235-257.
- Achenbach, T. M. (1991). *Manual for Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Dept. of Psychiatry.
- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Cai, X., Kaiser, A. P., & Hancock, T. B. (2004). Parent and teacher agreement on Child Behavior Checklist items in a sample of preschoolers from low-income and predominantly African American families. *Journal of Clinical Adolescent Psychology*, 33, 303-312.
- Cassidy, J., & Asher, S. R. (1992). Loneliness and peer relations in young children. *Child Development*, 63, 350-365.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- Conners, C. K. (1969). A teacher rating scale for use in drug studies with children.

- American Journal of Psychiatry*, 126, 884-888.
- Conners, C. K. (1970). Symptom patterns in hyperactive, neurotic, and normal children. *Child Development*, 41, 667-682.
- Conners, C. K. (1973). Rating scales for use in drug studies with children. *Psychopharmacology Bulletin (Special Issue, Pharmacotherapy with Children)*, 24-29.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology*, 6, 221-236.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System*. Circle Pines MN: American Guidance Service.
- Hay, D. F., Pawlby, S., Sharp, D., Schmücker, G., Mills, A., Allen, H., & Kumar, R. (1999). Parents' judgments about young children's problems: Why mothers and fathers might disagree yet still predict later outcomes. *Journal of Child Psychology and Psychiatry*, 40, 1249 - 1258.
- Hightower, A. D., Work, W. C., Cowen, E. L., Spinell, A. P., Lotyczewski, B. S., Guare, J. C., & Rohrbeck, C. A. (1986). *The Child Rating Scale: The development and psychometric refinement of a socioemotional self-rating scale for young school children*. Manuscript submitted for publication.
- Kumpulainen, K., Rasanen, E., Henttonen, I., Moilanen, I., Piha, J., Puura, K., Tamminen, T., & Almqvist, F. (1999). Children's behavioral/ emotional problems: A comparison of parents' and teachers' reports for elementary school-aged children. *European Child and Adolescent Psychiatry*, 8, 41 - 47.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question and weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 151-173.
- Liu, J., Cheng, H., & Leung, P. W. L. (2011). The application of the preschool child behavior checklist and the Caregiver - Teacher Report Form to mainland Chinese children: Syndrome structure, gender differences, country effects, and inter-informant agreement. *Journal of Abnormal Child Psychology*, 39, 251 - 264.
- Phares, V. (1997). Accuracy of informants: Do parents think that mother knows best? *Journal of Abnormal Child Psychology*, 25, 165 - 171.
- Quay, H. C., & Peterson, D. R. (1987). *Manual for the Revised Behavior Problem Checklist*. Coral Gables, FL: University of Miami.
- Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.
- Shelov, S. P. (1998). *Caring for Your Baby and Young Child: Birth to Age 5*. New York: Bantam Books.
- Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers and children on internalizing and externalizing behavior

- problems. *Journal of Clinical Child & Adolescent Psychology*, 22, 107 - 116.
- Uitenbroek, D. G. (1997). "SISA Binomial". <http://www.quantitativeskills.com/sisa/distributions/binomial.htm>
- Winsler, A., & Wallace, G. L. (2002). Behavior problems and social skills in preschool children: Parent - teacher agreement and relations with classroom observations. *Early Education and Development*, 13, 41 - 58.
- Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology*, 68, 1038 - 1050.
-

1차 원고 접수: 2011. 04. 11

수정 원고 접수: 2011. 05. 16

최종 게재 결정: 2011. 05. 17

Standardization Study for the Korean Version of the Caregiver-Teacher Report Form

Young-Ah Kim¹ Jin Lee¹ Yu-Jin Kim¹ Min-Young Kim¹ Kyung-Ja Oh²
Huno Inc.¹ Department of Psychology, Yonsei University²

The purpose of this study was to standardize the Korean version of Caregiver-Teacher Report Form (C-TRF). 1,371 preschoolers (679 boys and 692 girls) age from 1.5-5 were selected as the normative sample. The confirmatory factor analysis showed that the original C-TRF's factor structure can be appropriately applied to the Korean version of C-TRF. The reliability was tested via internal consistency (.52~.96) and test-retest reliability (.57~.82). The mean differences and Cohen's effect sizes of the subscales between clinical and normative sample revealed the scale's discriminant validity. Differences of mean subscale scores for different age and sex groups showed that 4 independent norms for sex (male vs female) and 2 age groups (18~35month vs 36~72month) would be reasonable for the scale.

Keywords: Korean version of C-TRF, Caregiver-Teacher Report Form, preschooler, standardization