

한국판 아동·청소년 행동평가 척도 교사용(TRF) 표준화 연구*

김 민 영

김 영 아

오 경 자[†]

(주)휴노

연세대학교 심리학과

본 연구에서는 한국판 아동·청소년 행동평가척도 교사용(TRF)을 표준화하고 신뢰도 및 타당도를 검증하였다. 한국판 TRF의 규준집단(남학생 476명, 여학생 462명)은 2005년 대한민국 인구 및 주택 총조사 자료를 참조하여 수집하였다. 확인적 요인분석 결과는 미국판 원검사의 요인구조를 한국판에도 동일하게 적용할 수 있음을 시사하였고, 각 하위척도들의 내적 합치도 수준(Cronbach's $\alpha=.59-.97$, 기타문제: .39)과 검사-재검사 신뢰도 값(.53-.96)이 모두 양호하였다. 내재화 및 외현화 척도에 속하는 하위 척도들과 상위 척도가 높은 상관을 보였고, 한국판 CBCL 6-18을 준거타당도 척도로 활용하여 TRF총점기준 비임상/임상 집단 간의 평균차이를 확인한 결과, 대부분의 척도에서 TRF총점기준 임상 집단이 비임상 집단보다 높은 점수를 나타내는 경향을 보였다. 단축형 Conners 척도와 관련이 있는 TRF의 척도 간에도 .77-.92 사이의 높은 상관관계를 보이는 것으로 나타나 수렴타당도 및 변별타당도를 지지하였다. 각 하위척도별 평균 점수를 학년급간과 성별에 따라 비교한 결과 초등/중·고등 학년급간과 남/여성별에 따라 규준을 제시하는 것이 적합한 것으로 나타났다. 한국판 TRF의 임상 및 연구 장면에서의 활용에 대한 논의가 이루어졌다.

주요어 : 한국판 TRF, 아동·청소년 문제행동, 행동평가척도, 표준화, 신뢰도, 타당도

* 본 논문에는 "ASEBA 아동·청소년 행동평가척도 매뉴얼(오경자, 김영아, 2011)"에 제시된 내용이 일부 포함되어 있음.

[†] 교신저자(Corresponding Author) : 오경자, 연세대학교 심리학과, 서울시 서대문구 신촌동 134
E-mail : kjoh@yonsei.ac.kr

교육과학기술부의 ‘학생 정신건강 검진 시범운영사업 연구보고서’에 따르면, 전국 245개 초·중·고등학교 학생 7만여 명을 대상으로 정신건강 선별검사를 실시한 결과 13%의 학생에게 정밀검진이 필요한 것으로 보고되었다(연합뉴스, 2009. 08. 23). 한국청소년상담원의 ‘2010 상담경향 분석 보고서’에서는 개인상담 호소문제 중 우울/위축, 반항성 및 품행성 장애 등의 정신건강 문제와 관련된 상담 비율이 2007년 8%에서 2010년에는 19%로 높은 증가율을 보인 것으로 나타났다(연합뉴스, 2011. 08. 22). 또한 최근 연이어 일어나고 있는 왕따 문제로 인한 자살 사건들은 학교에서의 아동·청소년의 적응 상태 및 정신 건강 수준 실태 파악과 문제 예방 및 해결책 마련이 시급함을 보여준다.

아동·청소년의 적응 상태 및 문제행동에 대한 평가는 부모, 교사, 본인, 임상/상담가 등 다양한 평가자의 보고를 통해 이루어질 수 있으며, 평가 대상이 동일한 경우라도 평가자가 누구인지가 중요한 요인이 될 수 있다. 이는 평가자 개개인이 가지고 있는 문제 행동 수준에 대한 일반적인 내적(자의적) 기준이 모두 다르고, 주로 어떤 상황에서 평가 대상자와 함께 있는가에 따라 어떤 평가자는 다른 평가자보다 평가 대상자의 상황 특이적(situation specific)인 행동에 민감할 수 있기 때문이다.

이와 같은 평가자 요인에 대한 고려가 적극적으로 반영된 ASEBA(Achenbach System of Empirically Based Assessment; Achenbach & Rescorla, 2001)는 평가 대상 및 평정자의 관계를 구조화하여 다양한 장면에서의 문제를 측정할 수 있도록 구성된 시스템으로, 여러 문화권에서 대표적인 행동평가 척도로 활용되어 온 CBCL(Child Behavior Checklist)을 비롯하여

그 동안 개발되어 온 개별 검사들을 독립적으로 포함하고 있다. 상당수 문항이 동일한 행동에 대해 평가자의 특성에 맞추어 평정할 수 있도록 구성된 것이 특징적으로, 생후 18개월에서 성인에 이르기까지 개인의 적응 상태 및 문제행동을 추적 비교할 수 있을 뿐 아니라 동일한 평가 대상에 대한 여러 평가자의 보고를 횡적으로 비교하는 것 또한 가능하다는 장점이 있다. 즉 전문가가 아닌 부모, 교사, 본인 등 다양한 평가자들이 평가 대상자의 문제행동에 대해 동일한 양식으로 보고하게 함으로써 우울이나 위축 등의 내재화 문제 및 비행이나 공격행동 같은 외현화 문제뿐만 아니라 주의집중문제 등 다양한 문제의 수준을 신뢰롭게 측정, 비교할 수 있다는 것이다. ASEBA 학령기용(School-Age Forms)에는 아동·청소년의 문제행동을 부모나 주변 성인들의 보고를 통해 평가하는 CBCL 6-18(Child Behavior Checklist for ages 6-18), 청소년이 자신의 문제행동에 대해 스스로 평가하는 YSR(Youth Self Report)과 함께 학생을 잘 알고 있는 교사나 학교 상담원 등이 학생의 적응 상태 및 문제행동을 평가하는 TRF(Teacher's Report Form)가 포함되어 있다. 전 세계적으로 ASEBA 검사군을 활용한 다양한 연구 결과들이 발표되고 있으며, 개별검사 뿐 아니라 ASEBA 검사군 안의 다양한 검사들을 함께 사용하여 비교 분석하는 연구들도 활발히 이루어지고 있다(예, Dedrick, Tan, & Marfo, 2008; Delaney-Black et al., 2000; Mychailysz, Mendez, & Kendall, 2010).

국내에서는 1991년 미국판 CBCL 4-18과 YSR을 표준화한 K-CBCL(아동·청소년 행동평가척도; 오경자, 이해련, 홍강의, 하은혜, 1997)과 K-YSR(청소년 자기행동 평가척도; 오경자, 하은혜, 이해련, 홍강의, 2001)이 사용되어왔으

며, 최근에는 2001년에 전반적으로 개편된 ASEBA 시스템에 포함된 아동·청소년 행동평가척도 부모용(CBCL 6-18) 및 청소년 행동평가척도 자기보고용(YSR)이 국내 규준에 맞추어 새롭게 표준화된 바 있다(오경자, 김영아, 2011).

그러나 아동·청소년기는 학교에서 오랜 시간을 보내며 선생님, 친구들과의 본격적인 상호작용이 시작되므로(이경남, 2002; 장휘숙, 1997), 아동·청소년 스스로의 보고나 가정에서 주양육자의 관찰뿐만 아니라 학교 장면에서의 문제 행동을 객관적으로 관찰, 평가하는 것이 중요해진다. 특히 학교는 여러 또래 아동·청소년이 함께 모여 단체생활을 하며 다양한 사회적 기술과 행동 양식을 배우는 생활의 장으로, 이후 발달에 큰 영향을 미치는 중요한 요인이 될 수 있는 적응과 관련된 문제를 확인해 볼 수 있다는 점에서 중요한 의미를 지닌다. 부모와의 직접적인 상호작용이 이루어지지 않는 학교에서 아동·청소년의 행동은 해당 상황에서 이들을 직접 지도하고 보호하는 교사가 관찰, 보고하는 것이 보다 신뢰로울 수 있다. 또한 한 명의 아동·청소년에 대해 평가하는 부모와는 달리, 여러 또래 아동·청소년들을 함께 관찰하는 교사의 경우보다 객관적인 시각에서 문제행동의 상대적 심각성을 평가할 수 있을 것이다. 사람들은 맥락적 상황에 따라 각기 다른 행동 양식을 취하기도 한다. 즉, 동일한 대상이라 할지라도 서로 다른 관점을 지닌 평가자들과 서로 다른 환경 내에서 상호작용을 하면서 보이는 행동 양식은 큰 차이를 보일 수 있다. 따라서 평가 대상에게 의미가 있는 사회적 상호작용의 맥락을 반영할 수 있도록 한 명 이상의 평가자로부터 행동 평가의 결과를 얻는 것이 중요하

다. 교사 평가를 부모 평가와 비교할 수 있다면, 아동·청소년 행동에 대한 이해의 폭을 넓히고 환경적 맥락에서 아동·청소년의 문제 행동을 유발하는 요인이 어떤 것인지를 탐색하거나 부모-교사 간 관점의 차이를 이해하는데 도움이 될 수 있을 것이다.

아동·청소년기의 문제를 조기 발견하고 대처하는데 있어 교사 평가가 중요한 축을 담당할 수 있다는 공감대를 기반으로 많은 학자와 전문가들이 교사에 의한 행동평가 척도를 개발해왔으며(Achenbach, 1991; Cassidy & Asher, 1992; Gresham & Elliott, 1990; Quay & Peterson, 1987; Reynolds & Kamphaus, 1992), 국내에서도 교사의 평가에 기반한 행동평가 도구를 개발하려는 시도들이 있어왔다(소유경, 노주선, 김영신, 고선규, 고윤주, 2002; 신현숙, 박용재, 박주희, 류경희, 2006; 정현희, 2002). 다만 그 대상이 아동·청소년을 포괄하기보다는 초등학생 등 일부 집단에 한정되거나 평가 영역이 ADHD 등 특정 문제에 국한되는 경우가 많았고, 교사의 평가 이외에 평가 대상자에 대한 다른 정보를 통합하여 객관적으로 비교하기에는 미흡한 점이 있었다. 한 개인에 대한 평가는 여러 장면에서, 여러 평가자에 의해, 여러 가지 방법으로 이루어질 경우 보다 높은 타당성을 획득할 수 있다는 전제(Hops, 1983) 하에, 동일한 문제행동에 대한 여러 평정자의 시각이 반영됨과 동시에 개인이 나타낼 수 있는 다양한 영역에서의 문제행동을 포괄적으로 측정하는 평가 도구의 개발 및 표준화가 중요하다 할 수 있을 것이다.

본 연구는 2001년 미국판 ASEBA에 포함되어 있는 교사용 질문지인 TRF를 국내 아동·청소년을 대상으로 표준화하고, 그 신뢰도와 타당도를 평가하기 위한 목적으로 실시되었다.

이를 통해 학교 장면에서 교사의 보고를 통해 아동·청소년의 문제행동 수준을 파악할 수 있는 객관적인 측정 도구를 마련하고, CBCL 6-18, YSR과 함께 다중 평가자에 의한 평가 자료의 횡적인 평가를 가능케 함으로써 평가 대상을 보다 정확하게 이해하고 문제 해결적 개입에 대한 다양한 통찰을 얻게 될 수 있을 것으로 기대된다. 또한 ASEBA 검사군에 속하는 척도 중 최근에 한국판으로 표준화된 유아 행동문제평가척도 교사용(C-TRF; Caregiver-Teacher Report Form, 김영아, 이진, 김유진, 김민영, 오경자, 2011)과 더불어 유아기와 아동·청소년기에 이르는 연속적인 발달 과정에 대한 평가를 촉진할 수 있는 계기를 마련할 수 있을 것이다.

그러나 신뢰도와 타당도가 확보된 보편성이 있는 검사라 하더라도 한 문화권에서 최적화되어 개발된 검사가 다른 지역에서도 항상 최적화된다고 보기는 어렵다. 따라서 국내에서 미국판 척도를 사용하기 위해서는 문항별 요인구조 탐색, 기준 작성 등의 개발 단계에서 몇 차례 판단과 선택이 이루어져야 하고 이에 따라 개발 방향이 달라질 수 있다. ASEBA는 여러 문화권에서 널리 사용되며 횡문화적 비교가 가능한 보편성이 강한 검사이므로 한국판 TRF의 표준화에서는 이러한 기능을 최대한 유지할 수 있도록 하였다. 즉, 문항 번역에서부터 자료 수집, 통계 분석 방식은 원검사 개발 방식을 적용하는 것을 원칙으로 하되, 국내에서 적용하기에 적합하지 않은 부분이 있는지 검토하며 개발을 진행하였다.

본 연구의 연구문제는 다음과 같다.

1. 미국판 TRF 원검사의 요인구조를 한국 아동·청소년 집단에 그대로 적용하는 것이 타당한가?

2. 한국판 TRF의 신뢰도와 타당도는 적합한가?

3. 한국판 TRF 기준집단은 연령과 성별에 따라 어떻게 구분하는 것이 바람직한가?

4. 한국판 TRF 점수를 통해 나타난 교사가 보고한 한국 아동·청소년의 문제행동 수준은 미국 아동·청소년의 기준 집단과 비교했을 때 어떠한가?

방 법

연구 대상

2005년 대한민국 인구 및 주택 총 조사(통계청, 2006)의 전국 지역별 인구분포를 반영하여 각 지역의 초, 중, 고등학교 27곳을 선별한 후 학교로 TRF를 발송하여 해당 기관의 담임 교사가 아동·청소년의 행동에 대해 평가하도록 하였다. 그 결과, 총 957명의 자료가 수집되었는데 이 중 성별, 학년급간, 지역분포가 확인되지 않거나, 미국판 원검사의 표준화 과정에서와 마찬가지로 TRF의 문제행동척도 문항 중 결측치가 8개 이상인 사례를 제외한 938명(남: 476명; 여: 462명)의 자료를 한국판 TRF 표준화 기준집단으로 규정하고 분석하였다. 평가에 참여한 교사는 총 98명(남자: 17명, 여자: 54명, 미기입: 27명; 평균 연령 36.8세, 연령 범위: 22-58세)으로 평균교사경력 11.7년이었고, 교사 1인당 평균 9.53부의 검사를 실시한 것으로 나타났다.

기준집단에 포함된 학년급간, 성별, 지역별 표집 사례수가 표 1에 제시되어 있고, 표 2는 5-19세의 인구분포 자료를 4대 권역(서울/경기·강원·충청·제주/경상남북도/전라남북도)

표 1. 표준집단의 학년급간, 지역, 성별 표집 사례수(N=938)

학년	서울	경기·강원· 충청·제주	경상남북도	전라남북도	전체
초등 1	3 / 5	10 / 10	5 / 5	4 / 1	22 / 21 (4.6 / 4.5)
초등 2	5 / 4	11 / 9	5 / 5	4 / 6	25 / 24 (5.3 / 5.2)
초등 3	5 / 5	12 / 8	10 / 0	6 / 3	33 / 16 (6.9 / 3.5)
초등 4	9 / 9	10 / 10	5 / 5	5 / 5	29 / 29 (6.1 / 6.3)
초등 5	11 / 9	9 / 10	6 / 3	5 / 5	31 / 27 (6.5 / 5.8)
초등 6	10 / 10	9 / 10	5 / 5	2 / 2	26 / 27 (5.5 / 5.8)
중등 1	5 / 5	20 / 12	0 / 0	0 / 0	25 / 17 (5.3 / 3.7)
중등 2	34 / 32	20 / 20	29 / 32	0 / 0	83 / 84 (17.4 / 18.2)
중등 3	10 / 10	20 / 51	8 / 10	0 / 10	38 / 81 (8.0 / 17.5)
고등 1	20 / 10	31 / 48	30 / 10	0 / 2	81 / 70 (17.0 / 15.2)
고등 2	19 / 30	20 / 20	16 / 9	0 / 0	55 / 59 (11.6 / 12.8)
고등 3	18 / 0	0 / 0	10 / 7	0 / 0	28 / 7 (5.9 / 1.5)
전체	149 / 129 (31.3 / 27.9)	172 / 208 (36.1 / 45.0)	129 / 91 (27.1 / 19.7)	26 / 34 (5.5 / 7.4)	476 / 462 (50.7 / 49.3)

* 남/여. ()안은 %.

표 2. 지역 분포: 표준집단과 전국평균(2005년 인구주택총조사 자료기반)

	서울		경기·강원· 충청·제주		경상남북도		전라남북도	
	남	여	남	여	남	여	남	여
전국평균(5-19세)	18.4	18.6	43.4	44.2	27.4	26.1	10.8	11.0
전국평균(전체연령)	20.6	20.9	42.0	41.5	26.9	26.9	10.5	10.8
표준집단(초1-고3)	31.3	27.9	36.1	45.0	27.1	19.7	5.5	7.4

* 단위 %

으로 추출하여 비교한 결과이다. 이를 통해 표준집단이 국내 아동·청소년의 지역 및 권역 별 분포를 어느 정도 반영했음을 확인해볼 수 있었다. 단, 고등학교 집단의 경우 우리나라

입시 제도와 관련하여 상대적으로 섭외가 어려운 고 3 학생의 비율에 비해 1, 2학년의 비율이 좀 더 높았다.

연구 도구

한국판 아동·청소년 행동평가척도 교사용(Teacher's Report Form for Ages 6-18; TRF)

Achenbach와 Rescorla(2001)가 개발한 TRF를 한국어로 번역하여 사용하였다. 문항 번역은 TRF가 속한 ASEBA 검사군의 번역을 실시할 때 통합적으로 이루어진 것으로 임상심리 전문가가 1차 번역을 하고, 1차 번역자를 포함하여 대학원에서 임상심리학을 전공한 4인이 각 문항 번역의 정확성 및 번역 문항간의 일관성을 검토하였다. 1차 번역본의 검토 결과 의견이 일치하지 않는 문항에 대해서는 대안적 번역문을 정리한 후, TRF의 원저자(Achenbach)와 영어/한국어를 사용하는 이중 언어 사용자에게 자문을 구하여 원래의 척도가 전달하려는 의미에 가장 가까운 번역을 택하여 2차 번역본을 구성하였다. 이를 다시 제 3의 이중 언어 사용자의 역번역 과정과 그 결과에 대한 2차 검토를 통해 최종 번역본으로 확정하였다.

TRF는 '자기가 시작한 일을 끝내지 못한다', '학교에서 규율을 어긴다' 등과 같이 아동·청소년의 문제행동을 기술한 총 119개의 문항과, 제시된 문항 이외의 문제를 직접 기술하도록 되어 있는 개방형 문항 1개로 이루어져 있다. 각 문항은 교사가 지난 2개월 내에 해당 학생이 그 행동을 보였는가를 판단하여 "전혀 아니다: 0, 가끔 그렇다: 1, 자주 그렇다: 2"의 3점 척도로 평가하도록 되어 있다. 따라서 총 점의 범위는 0-240점이다.

미국판 TRF 원검사는 요인분석을 통하여 도출된 증후군 척도들과 특정 요인으로 분류되지 않는 문제행동을 측정하는 기타문제 척도로 이루어져 있다. 증후군 척도는 불안/우울,

위축/우울, 신체증상, 사회적 미성숙, 사고문제, 주의집중문제, 규칙위반, 공격행동의 8개 요인으로 이루어져 있다. 또한 증후군 척도 및 기타문제 척도들은 세 가지의 상위척도(내재화, 외현화, 문제행동총점)를 구성한다. 내재화 척도는 불안/우울, 위축/우울, 신체 증상의 세 가지 하위척도, 외현화 척도는 규칙위반, 공격행동의 두 가지 하위척도, 문제행동총점은 문제행동 문항 전체의 합으로 점수화된다.

또한 정신과의사 및 심리학자로 구성된 16개 문화권의 전문가 22명이 TRF의 각 문항을 아동·청소년에게서 자주 보고되는 DSM (Diagnostic and Statistical Manual of Mental Disorders) 척도와 부합하는지 평정하도록 하여 구성된 DSM 진단 척도(DSM-oriented scales)를 추가적으로 포함하여 경험적으로 도출된 증후군 척도를 보완하고자 하였다(Achenbach & Rescorla, 2001). DSM 정서문제, DSM 불안문제, DSM 신체화문제, DSM ADHD, DSM 반항행동문제, DSM 품행문제의 6개 요인으로 구성되어 있다.

한국판 아동·청소년 행동평가척도 부모용(Child Behavior Checklist for Ages 6-18; CBCL 6-18), 청소년 행동평가척도 자기보고용(Youth Self-Report; YSR)

한국판 TRF의 타당도 분석을 위해 Achenbach와 Rescorla(2001)가 개발한 CBCL 6-18, YSR을 한국판으로 표준화한 아동·청소년 행동평가척도 부모용과 청소년 행동평가척도 자기보고용(오경자, 김영아, 2011)을 함께 사용하였다. 두 검사 모두 0-1-2 척도로 구성되어 있으며, 8개의 증후군 척도(불안/우울, 위축/우울, 신체증상, 사회적 미성숙, 사고문제, 주의집중문제, 규칙위반, 공격행동) 및 기타문제 척도

들과 이들 증후군 척도들 간의 합으로 구성된 상위 요인인 내재화 척도, 외현화 척도, 문제 행동총점이 있다. 또한 6개의 DSM 진단 척도(DSM 정서문제, DSM 불안문제, DSM 신체화 문제, DSM ADHD, DSM 반항행동문제, DSM 품행문제)도 동일하게 구성되어 있다. CBCL 6-18은 총 120문항, YSR은 총 119문항으로 이루어져 있다. 한국판 CBCL6-18의 내적 합치도는 .58-.95, YSR은 .63-.85 수준이었다(오경자, 김영아, 2011).

단축형 Conners 척도(Abbreviated Conners' Teacher Rating Scale: ACTRS)

한국판 TRF의 타당도 분석을 위해 단축형 Conners 척도를 함께 사용하였다. 단축형 Conners 평정 척도는 본래 Conners(1973)가 교사용 척도로 개발한 39개 문항을 Goyette, Conners와 Ulrich(1978)가 10개의 문항으로 축소한 검사도구로, 국내에서는 오경자와 이혜련(1989)이 번안하였다. 한국판 Conners 단축형 검사는 주의력 결핍 과잉 행동 문제 평가에 주로 쓰이며, 0-3점의 4점 척도로 구성되어 있고 점수가 높을수록 문제 행동이 심각한 것으로 볼 수 있다. 한국판 단축형 Conners 척도의 내적 합치도(Cronbach's α)는 .92였다.

자료 수집 절차 및 분석 방법

2007-2008년에 걸쳐 대한민국 인구 및 주택 총조사(통계청, 2006)의 전국 지역별 인구분포를 반영하여 각 지역의 초, 중, 고등학교 27곳으로부터 규준집단을 표집하였다. 한국판 TRF를 각 학교로 보내어 교사들이 담당하고 있는 학생에 대해 평가를 실시하게끔 하였다. 한 교사가 해당 학급 내 전체 학생의 검사를 실

시할 경우 그 부담이 크고 응답의 성실성에 문제가 생길 수 있으며, 학급 내 한두 명의 학생에 대해서만 평가할 경우 선정의 편향이 불가피 할 수 있다는 점을 고려하여 교사 1인당 응답할 수 있는 검사지를 10개 내외 범위로 정해 전달하였다. 우리나라 아동·청소년들의 평균적인 특성 자료를 얻어야 하므로, 어떤 특성이 있는 학생들만 선택해 평가하지 않도록 하기 위해 해당 반 학생 10명을 무작위로 뽑도록 했다. 즉, 학생 선택 시 성적이나 문제 행동과 같은 기준으로 선정되지 않아야 함을 강조하여 무작위 표집이 될 수 있도록 하였다. 단, 성비는 차이가 나지 않도록 반반씩 뽑도록 하였다. 또한 아동·청소년의 문제 행동 수준에 대한 부모, 학생 본인, 교사의 평가를 비교하기 위해 일부 학생의 부모에게는 한국판 CBCL 6-18을, 학생들에게는 YSR을 실시하도록 하였다. 이 자료들은 각각 한국판 CBCL 6-18과 YSR 검사 개발을 위한 자료로도 활용이 되었다. TRF와 함께 실시한 CBCL 6-18과 YSR은 각각 607부, 395부였다(표 3).

표 3. 평가자 간 점수 비교를 위해 한국판 TRF와 함께 실시한 검사의 자료 수

	TRF + CBCL	TRF + YSR
부 수	607	395

추가적으로 검사-재검사 신뢰도를 알아보기 위해 서울 지역의 중학교 8개 학급과 Y대학교의 교육대학원 심리학 관련과목을 수강하는 교사 7명을 대상으로 교사 1인당 학생 5명에 대해 TRF를 작성하고, 1주일 후 다시 동일한 학생에 대해 재평가하도록 하여 총 75명의 검사-재검사 자료를 수집하였다. 이 자료는 규준

집단에 포함되지 않았다.

본 표준화 연구에서는 표준집단($N=938$)에 대한 문제행동증후군 척도와 DSM 진단 척도를 중심으로 결과를 제시하고자 한다. 신뢰도 및 타당도와 관련된 지표 분석을 위해서 SPSS 15.0 for windows와 AMOS 7.0 평가판을 사용하였다.

결 과

확인적 요인분석(CFA: Confirmatory Factor Analysis)

TRF를 포함한 ASEBA 검사들은 여러 문화권에서 원 검사의 요인구조를 거의 그대로 유지하며 사용되고 있어 축적된 자료의 활용이나 문화권 간의 비교 연구 측면에서 큰 이점이 되고 있다(Achenbach & Rescorla, 2000, 2001). 따라서 미국판 TRF의 척도 구성이 우리나라 아동·청소년 집단을 평가하기 위해 필요한 요건들을 충족한다면 동일한 구성을 유지하는 것이 바람직한 것이다.

TRF는 아동·청소년의 문제행동과 관련된 증상을 측정하므로 가정 및 학교에서 잘 적응하고 있는 일반 집단에서 실시할 경우 문제행동의 보고 비율이 낮아 정적 편포 경향이 나타날 수 있으며, 이러한 정적 편포는 요인 부하량에 영향을 미칠 수 있다. 따라서 척도의 원저자인 Achenbach와 Rescorla(2001)는 표준 집단 문제행동총점의 중앙값 이상인 자료만 확인적 요인분석 자료에 포함시키고, 낮은 응답 빈도에 의한 자료 해석의 왜곡을 방지하기 위해 3점 척도(0-1-2)를 이분 척도(0-1)로 환산한 후 확인적 요인분석을 시행한 바 있다. 한국

판 TRF의 확인적 요인 분석에서도 이와 동일한 방식을 적용하여 문제행동총점이 중앙값(10점) 이상인 479개의 사례만을 포함하고, 0-1-2 척도를 0(기존의 0점)과 1(기존의 1, 2점)의 이분척도로 재코딩한 후 확인적 요인분석을 실시하였다. 먼저 한국판 TRF의 모든 문항을 측정변인으로, 문제행동증후군 척도의 하위요인을 잠재변인으로 한 문항수준의 확인적 요인분석을 실시한 결과, RMSEA가 .056으로 미국판 검사의 .070과 마찬가지로 우수한 수준으로 나타났으나 그 밖의 적합도 지수들($CFI=.502$, $TLI=.490$, $GFI=.589$)은 적합도 수준이 좋지 못했다(양호한 적합도 수준: $RMSEA < .08$, CFI , TLI , $GFI > .90$; Byrne, 2001; Holsie, 1995). 이와 같은 결과는 측정 변인인 문항과 그 상위 요인의 수가 많을 경우 발생하는 현상으로, TRF의 원저자 역시 RMSEA 수준만 제시하고 있다(Achenbach & Rescorla, 2001). 추정해야 하는 경로가 늘어남에 따라 모델이 복잡해지고 상대적인 설명 변량이 약화되어 모델의 적합도 수준이 저하되는 경우는 이전 연구에서도 보고된 바 있고(Cai, Kaiser, & Hancock, 2004; Little, Cunningham, Shahar, & Widaman, 2002), 한국판 CBCL 1.5-5 (김영아, 이진, 문수중, 김유진, 오경자, 2009), C-TRF(김영아, 이진, 김유진, 김민영, 오경자, 2011)의 경우에서도 유사한 양상이 나타났다.

이와 같은 경우, 각 요인별 합산 값을 하나의 측정변인으로 가정하고 하위척도의 상위 요인인 내재화, 외현화, 문제행동총점을 상위 요인으로 가정하여 항목묶음(parceling) 확인적 요인분석을 사용하도록 제안된 바 있다(Little, Cunningham, Shahar, & Widaman, 2002). 항목묶음 확인적 요인분석을 위해서는 먼저 특정 문항들의 합산 값으로 제시되는 잠재변수의 일

표 4. 한국판 TRF 표집집단에 대한 원 TRF 요인 구조의 확인적 요인분석($n=479$)

	χ^2	df	RMSEA	CFI	TLI	GFI
문항수준	15083.411	6076	.056	.502	.490	.589
항목묶음	263.269	17	.174	.835	.728	.884

차원성이 검증되어야 한다(Segars & Grover, 1993; Somers, Nelson, & Karimi, 2003). 따라서 개별 문항을 측정변인으로 두고 해당 문항들의 합으로 가정될 수 있는 요인과의 구조에서 SRMR(Standardized Root Mean Square Residual)을 확인해 본 결과, .043-.094의 범위로 모두 .10보다 작아 일차원성을 만족시키는 것으로 나타났다. 항목묶음 확인적 요인분석 결과, CFI=.835, TLI=.728, GFI=.884로 지표들이 상승되는 경향이 있었다(표 4). 이와 같은 요인분석 결과들을 바탕으로, 한국판 TRF의 문항과 요인 간 구성 타당도는 일부 제약이 있기는 하나 원저자가 제시하는 기준을 충족한 점과 척도 활용의 효율성을 고려할 때 원검사의 구조를 적용할 수 있을 것으로 판단하였다.

신뢰도

한국판 TRF 척도들의 내적 합치도(Cronbach's α), 평가자간 일치도 및 재검사 신뢰도가 표 5에 제시되어 있다.

각 척도의 내적 합치도는 기타문제(.39)를 제외하면 .59-.97의 범위를 보이고 있는데, 함께 제시된 미국 규준집단의 내적합치도 범위 .72-.97에 비해 유사하거나 일부 다소 낮은 경향은 있었으나 전반적인 신뢰도 수준은 모두 양호한 편이었다.

척도의 신뢰도 평가의 방법으로 동일한 대상에 대하여 서로 다른 평가자의 평가 결과가

얼마나 일치하는지 알아보기 위해 교사가 평가한 TRF와 함께 주 양육자가 평가한 CBCL 6-18, 학생 본인이 평가한 YSR 점수 간 상관관계를 알아보았다. 상관분석은 비교 검사 간 공통문항에 대해서만 실시되었는데, TRF와 CBCL 6-18의 공통문항은 92개, TRF와 YSR의 공통문항은 90개였다. 그 결과 문제행동총점에서는 TRF와 CBCL 6-18이 $r=.16(p<.01)$, TRF와 YSR이 $r=.17(p<.01)$ 의 상관을 보였다. 내재화 척도에서는 각각 $.16(p<.01)$ 와 $.27(p<.01)$, 외현화 척도에서는 $.27(p<.01)$ 와 $.34(p<.01)$ 의 상관을 보였다. TRF와 CBCL 6-18의 불안/우울($r=.06, p=.14$), DSM 불안문제($r=.06, p=.13$)를 제외한 나머지 소척도들은 모두 통계적으로 유의미한 고른 상관을 보이긴 하였으나, 그 크기는 .50 미만으로 전반적으로 낮은 수준이었다(표 5). 미국판에서도 부모와 교사 평가는 .12-.44, 학생 본인과 교사 평가는 .05-.32 수준의 낮은 일치도를 보였다(Achenbach & Rescorla, 2001).

서울 지역의 초·중·고등학교 15개 학급 담임교사가 해당 학급 학생 75명에 대해 평가하고, 동일 학생들에 대해 1주일 후 재평가한 검사-재검사 상관은 .53-.96 범위로 나타났다. 16일 간격으로 재검사 신뢰도를 조사한 미국판 원검사와 기간이 달라 수평적인 비교가 어렵기는 하나, 미국판 검사-재검사 상관은 .60-.95 수준으로 한국판 TRF의 검사-재검사 신뢰도와 유사하였다.

표 5. 한국판 TRF의 내적 합치도, 평가자간 일치도 및 검사-재검사 신뢰도

	내적합치도		평정자간일치도		검사-재검사 신뢰도	
	한국판	미국판	교사-주양육자	교사-본인	한국판(7일)	미국판(16일)
	(N=938)	(N=3,086)	(n=607)	(n=395)	(N=75)	(N=44)
불안/우울	.80	.86	.06	.13**	.87**	.89
위축/우울	.81	.81	.28**	.22**	.91**	.60
신체증상	.71	.72	.18**	.16**	.68**	.83
사회적미성숙	.75	.82	.29**	.21**	.87**	.95
사고문제	.62	.72	.10*	.15**	.95**	.72
주의집중문제	.94	.95	.35**	.30**	.96**	.95
규칙위반	.83	.95	.34**	.42**	.90**	.83
공격행동	.92	.95	.22**	.28**	.93**	.88
기타문제	.39	-	.21**	.18**	.85**	-
내재화	.87	.90	.16**	.17**	.91**	.86
외현화	.94	.95	.27**	.34**	.94**	.89
문제행동총점	.97	.97	.16**	.17**	.95**	.95
DSM 정서문제	.72	.76	.15**	.14**	.87**	.62
DSM 불안문제	.59	.73	.06	.11*	.78**	.73
DSM 신체화문제	.66	.80	.21**	.13**	.53**	.73
DSM ADHD	.92	.94	.32**	.32**	.94**	.95
DSM 반항행동문제	.78	.90	.19**	.21**	.92**	.91
DSM 품행문제	.89	.90	.39**	.34**	.91**	.71

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

타당도

한국판 TRF 척도의 수렴타당도를 알아보기 위해 Pearson의 적률상관계수를 산출한 결과, 모든 요인들은 $p < .01$ 수준에서 유의한 상관관계를 보였으며 문제행동총점 척도는 하위 증후군 척도들과 .49-.92 사이의 비교적 높은 상관을 보였다. 특히 내재화 문제로 분류되는 불안/우울, 위축/우울, 신체증상 하위 척도들은

내재화 척도 점수와 .63-.89 사이의 비교적 높은 상관을 나타내었고, 외현화 문제로 분류되는 규칙위반, 공격행동 하위 척도는 외현화 척도와 각각 .89, .98의 상대적으로 높은 상관을 보이고 있다(표 6).

미국 TRF 원검사에서 준임상집단의 기준으로 제시하고 있는 60T(84%ile) 기준을 한국판 TRF에 적용하는 것이 타당한지 확인하기 위해 한국판 CBCL 6-18을 준거타당도 척도로 활용

표 6. 한국판 TRF의 각 척도 간 상관($N=938$)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.00																	
2	.51	1.00																
3	.46	.34	1.00															
4	.56	.54	.38	1.00														
5	.55	.41	.46	.54	1.00													
6	.47	.40	.34	.60	.51	1.00												
7	.33	.22	.32	.38	.42	.72	1.00											
8	.47	.24	.30	.55	.54	.78	.78	1.00										
9	.49	.42	.42	.46	.45	.57	.52	.47	1.00									
10	.89	.81	.63	.64	.60	.51	.36	.44	.56	1.00								
11	.45	.25	.33	.53	.53	.80	.89	.98	.51	.44	1.00							
12	.68	.54	.49	.74	.67	.92	.78	.87	.67	.73	.88	1.00						
13	.63	.81	.53	.64	.53	.62	.40	.44	.56	.83	.45	.73	1.00					
14	.82	.54	.45	.61	.52	.45	.28	.38	.46	.80	.37	.62	.61	1.00				
15	.36	.25	.93	.27	.36	.26	.28	.25	.34	.53	.27	.40	.39	.35	1.00			
16	.38	.24	.27	.53	.46	.95	.74	.82	.51	.38	.84	.87	.48	.36	.21	1.00		
17	.40	.21	.26	.47	.43	.73	.76	.92	.41	.38	.91	.79	.38	.33	.21	.76	1.00	
18	.33	.23	.28	.46	.48	.75	.92	.88	.49	.35	.94	.81	.40	.28	.23	.78	.82	1.00

1. 불안/우울, 2. 위축/우울, 3. 신체증상, 4. 사회적미성숙, 5. 사고문제, 6. 주의집중문제, 7. 규칙위반, 8. 공격행동, 9. 기타문제, 10. 내재화, 11. 외현화, 12. 문제행동총점, 13. DSM 정서문제, 14. DSM 불안문제, 15. DSM 신체화문제, 16. DSM ADHD, 17. DSM 반항행동문제, 18. DSM 품행문제. 표에 제시된 모든 상관관계는 $p < .01$ 수준에서 유의미하였음.

하여 한국판 TRF의 변별력을 확인해보고자 하였다. 즉, 한국판 TRF 표준집단 938명 중 한국판 CBCL 6-18을 함께 실시한 607명을 60T (84%ile) 이상의 TRF 문제행동총점을 보고한 집단과 60T 미만의 점수를 보고한 집단으로 분류한 뒤, 두 집단의 한국판 CBCL 6-18 점수 차이를 비교해보았다. 분석 결과 신체증상, 기타문제, DSM 불안문제, DSM 신체화문제를 제외한 하위척도들에 대해서 한국판 TRF의 총점에서 60T 이상을 받은 집단이 그렇지 않은 집단보다 CBCL 6-18에서도 높은 점수를 나타내는 경향을 보였다(표 7).

단축형 Conners 척도와 TRF 척도가 어느 정도 상관을 보이는지 알아보고자 표준집단에

포함된 144명에 대해서 추가적으로 단축형 Conners 척도 10문항을 함께 평가하도록 하였다. 한국판 Conners 단축형 검사는 아동·청소년의 주의력 결핍 과잉 행동 문제 평가에 쓰이므로, 이와 관련이 있는 TRF의 척도(주의집중문제, DSM ADHD)에서 높은 상관성이 나타난다면 한국판 TRF의 타당도를 지지해주는 결과가 될 수 있을 것이다. 한국판 TRF의 주의집중문제는 집중곤란과 과잉행동충동성의 합으로 구성되고, DSM ADHD는 DSM 집중곤란과 DSM 과잉행동충동성의 합으로 구성된다. 따라서 좀 더 세부적으로 살펴보기 위해 본 분석에서는 4개의 하위 요인을 추가하였다. 분석 결과, Conners 척도는 앞서 언급한 TRF의

표 7. 한국판 TRF 총점 기준 임상-정상 집단 간의 한국판 CBCL 6-18 하위척도별 점수평균 및 차이 비교

	TRF문제행동총점 기준 - 비임상 (n=506)	TRF문제행동총점 기준 - 임상 (n=101)	전체 (n=607)	F
CBCL 불안/우울	2.08 (2.44)	2.67 (2.79)	2.18 (2.51)	4.76*
CBCL 위축/우울	1.25 (1.72)	1.79 (2.18)	1.34 (1.81)	7.57**
CBCL 신체증상	1.33 (1.94)	1.45 (1.86)	1.35 (1.93)	.31
CBCL 사회적미성숙	1.54 (1.93)	1.98 (2.59)	1.61 (2.06)	3.88*
CBCL 사고문제	.90 (1.44)	1.50 (1.82)	1.00 (1.53)	13.37**
CBCL 주의집중문제	2.75 (2.88)	4.64 (3.81)	3.07 (3.13)	32.32**
CBCL 규칙위반	.81 (1.29)	1.86 (2.63)	.99 (1.64)	36.57**
CBCL 공격행동	2.89 (3.37)	4.42 (4.59)	3.15 (3.64)	15.06**
CBCL 기타문제	1.63 (1.92)	1.99 (1.90)	1.69 (1.92)	3.01
CBCL 내재화	4.66 (4.91)	5.91 (5.68)	4.87 (5.06)	5.19*
CBCL 외현화	3.71 (4.33)	6.28 (6.64)	4.13 (4.88)	24.28**
CBCL 문제행동총점	15.19 (14.29)	22.31 (18.24)	16.37 (15.23)	18.94**
CBCL DSM 정서문제	1.67 (1.95)	2.20 (2.15)	1.75 (2.00)	6.03*
CBCL DSM 불안문제	1.02 (1.36)	1.30 (1.43)	1.06 (1.38)	3.53
CBCL DSM 신체화문제	.72 (1.25)	.73 (1.19)	.72 (1.24)	.00
CBCL DSM ADHD	1.98 (2.19)	3.18 (2.71)	2.18 (2.32)	23.23**
CBCL DSM 반항행동문제	1.26 (1.46)	1.93 (1.78)	1.37 (1.54)	16.27**
CBCL DSM 품행문제	.78 (1.43)	1.88 (2.85)	.96 (1.79)	33.55**

*: $p < .05$, **: $p < .01$, ()안은 표준편차

관련 척도에서 .77-.92의 높은 상관을 보였고, 그 외에 규칙위반, 공격행동, 외현화, 문제행동총점 등과도 높은 상관을 보이는 것으로 나타났다(표 8).

규준 집단의 구분

한국판 TRF의 특성을 파악하고 규준을 작성하기 위해 수집된 자료의 전체 평균 및 성별,

학년급간별 평균 점수를 확인해본 결과가 표 9, 10에 제시되어 있다.

먼저 성별에 따른 점수 차이를 살펴보기 위해 수집된 자료를 성별(남, 여) 집단으로 나눈 후 각 문제행동 증후군 척도, 내재화, 외현화, 문제행동총점 및 DSM 진단 척도 별 점수 차이 여부를 살펴보았다(표 8). 그 결과, 불안/우울, 주의집중문제, 규칙위반, 내재화, DSM ADHD, DSM 품행문제 척도에 대해 유의미한

표 8. Conners 척도와 TRF 하위척도 상관관계 (n=144)

TRF 하위척도	Conners 척도와의 상관관계수
불안/우울	.34**
위축/우울	.24**
신체증상	.21*
사회적미성숙	.51**
사고문제	.55**
집중곤란	.77**
과잉행동충동성	.87**
주의집중문제	.87**
규칙위반	.83**
공격행동	.85**
기타문제	.49**
내재화	.34**
외현화	.90**
문제행동총점	.86**
DSM 정서문제	.42**
DSM 불안문제	.31**
DSM 신체화문제	.15
DSM 집중곤란	.83**
DSM 과잉행동충동성	.87**
DSM ADHD	.92**
DSM 반항행동문제	.81**
DSM 품행문제	.84**

** $p < .01$

차이가 나타남을 확인할 수 있었다. 전체적으로 살펴보면 불안/우울, 신체증상 등 내재화 척도에서는 여아가 남아보다 더 높은 문제행동 수준을 보이는 것으로 나타났으나, 규칙위반, 공격행동 등 외현화 척도에서는 남아가

여아보다 더 높은 문제행동 수준을 보이는 것으로 나타났다. 이와 같은 결과들은 성별에 따라 주로 보고되는 문제 행동의 종류나 정도가 다를 수 있으므로 성별에 따라 달리 적용될 수 있는 기준이 필요함을 시사한다.

미국판 원검사인 TRF는 학년급간 기준 대신 연령을 기준으로 11세 전후의 연령 집단에 대해 하위척도 별 점수 차이를 제시한 바 있다. 그러나 연령으로 집단을 구분하게 될 경우 생년월일의 차이로 인해 만 11-12세에 해당하는 아동이 초등학교에 다닐지, 중학교에 다닐지 불분명한 경우가 발생하게 된다. 또한 학령기 아동의 경우, 수치상의 나이보다는 학생이 속한 환경(초등학교인지 중·고등학교인지)이 사회심리적 적응에 보다 중요한 영향을 미칠 수 있다. 따라서 한국판 TRF에서는 연령 대신 학교 및 학년 급간을 기준으로, 초등학교 1학년-6학년, 중학교 1학년-고등학교 3학년 두 개의 학년 급간에 대하여 한국판 TRF의 하위척도별 점수 차이를 살펴보았다(표 10). 분석 결과, 기타문제를 제외한 모든 척도에서 초등학생이 중·고등학생 보다 높은 점수를 보였으며 불안/우울, 신체증상, 사회적 미성숙, 사고문제, 공격행동, 내재화, 외현화, 문제행동총점, DSM 불안문제, DSM 신체화문제, DSM 반항행동문제 척도에서는 그 차이가 유의미한 것으로 나타났다.

이러한 자료의 통계적 특성과 더불어 우리나라에서 일반적으로 청소년기와 아동기를 구분하는 기준이 초등학교의 졸업 여부로 받아들여지고 있다는 점을 고려할 때 성별과 더불어 초등학교/중·고등학교 집단에 대해 각각 고유한 기준을 산출하는 것이 척도의 임상적 활용 측면에서 유용할 것으로 판단된다. 따라서 한국판 TRF의 문제행동 점수의 기준은 2

표 9. 한국판 TRF 척도의 성별 평균 점수 및 차이 검증

	남 ($n=476$)	여 ($n=462$)	전체 ($N=938$)	차이 검증 F
불안/우울	1.64 (2.62)	2.32 (3.19)	1.98 (2.93)	12.91**
위축/우울	1.49 (2.44)	1.39 (2.20)	1.44 (2.33)	.43
신체증상	.42 (1.10)	.51 (1.31)	.46 (1.21)	1.38
사회적 미성숙	1.37 (2.42)	1.34 (2.03)	1.36 (2.24)	.02
사고문제	.47 (1.13)	.45 (1.13)	.46 (1.13)	.09
주의집중문제	6.69 (8.59)	5.44 (7.77)	6.08 (8.22)	5.43*
규칙위반	1.49 (2.71)	1.13 (2.44)	1.31 (2.59)	4.66*
공격행동	3.30 (5.79)	2.82 (4.89)	3.07 (5.37)	1.89
기타문제	1.03 (1.37)	1.05 (1.38)	1.04 (1.38)	.04
내재화	3.54 (5.00)	4.22 (5.49)	3.88 (5.26)	3.89*
외현화	4.79 (8.02)	3.95 (7.03)	4.38 (7.55)	2.95
문제행동총점	17.89 (22.02)	16.45 (20.75)	17.18 (21.41)	1.07
DSM 정서문제	1.32 (2.02)	1.30 (2.02)	1.31 (2.02)	.02
DSM 불안문제	.55 (1.13)	.69 (1.18)	.62 (1.16)	3.43
DSM 신체화문제	.26 (.76)	.32 (.99)	.29 (.88)	1.09
DSM ADHD	3.60 (5.09)	2.72 (4.43)	3.17 (4.79)	8.08**
DSM 반항행동문제	1.01 (1.73)	.89 (1.65)	.95 (1.69)	1.12
DSM 품행문제	1.90 (3.64)	1.25 (2.89)	1.58 (3.30)	9.28**

*: $p < .05$, **: $p < .01$, ()안은 표준편차

표 10. 한국판 TRF 척도의 학년급간 별 평균 점수 및 차이 검증

	초등학생 ($n=310$)	중·고등학생 ($n=628$)	전체 ($N=938$)	차이검증 F
불안/우울	2.57 (3.27)	1.69 (2.71)	1.98 (2.93)	19.10**
위축/우울	1.56 (2.44)	1.37 (2.27)	1.44 (2.33)	1.41
신체증상	.61 (1.38)	.39 (1.11)	.46 (1.21)	6.99**
사회적 미성숙	1.77 (2.60)	1.15 (2.00)	1.36 (2.24)	16.29**
사고문제	.62 (1.40)	.38 (.96)	.46 (1.13)	9.19**
주의집중문제	6.43 (8.27)	5.90 (8.19)	6.08 (8.22)	.84
규칙위반	1.33 (2.26)	1.30 (2.74)	1.31 (2.59)	.02
공격행동	3.80 (5.85)	2.70 (5.08)	3.07 (5.37)	8.82**
기타문제	.95 (1.42)	1.08 (1.35)	1.04 (1.38)	1.98
내재화	4.74 (5.86)	3.45 (4.89)	3.88 (5.26)	12.73**
외현화	5.13 (7.80)	4.00 (7.41)	4.38 (7.55)	4.63*
문제행동총점	19.63 (22.86)	15.97 (20.56)	17.18 (21.41)	6.12*
DSM 정서문제	1.42 (2.15)	1.26 (1.95)	1.31 (2.02)	1.30
DSM 불안문제	.76 (1.25)	.56 (1.11)	.62 (1.16)	6.25*
DSM 신체화문제	.39 (1.01)	.24 (.80)	.29 (.88)	5.96*
DSM ADHD	3.44 (4.75)	3.03 (4.81)	3.17 (4.79)	1.46
DSM 반항행동문제	1.20 (1.84)	.83 (1.60)	.95 (1.69)	9.90**
DSM 품행문제	1.88 (3.48)	1.43 (3.20)	1.58 (3.30)	3.79

*: $p < .05$, **: $p < .01$, ()안은 표준편차

(성별: 남, 여)*2(연령: 초등 1-6학년, 중등 1-고
등 3학년)의 총 네 집단으로 나누어 작성되었
다.

미국과 한국 기준집단의 척도 별 평균 비교

한국어로 번안된 TRF를 통해 측정된 기준집
단의 성별, 연령 집단 별 문제행동 평균과 표
준편차가 표 11-1, 11-2에 제시되어 있다. 2008
년 수집한 한국 기준집단의 문제행동 점수 수
준에 대한 상대적 이해를 돕기 위해 미국판

원검사 제작시 표집 된 기준 집단의 문제행동
영역별 평균 점수를 함께 제시하였다. Simple *t*
test를 적용해서 한국의 기준집단과 미국의 규
준집단 간 TRF 하위척도의 점수를 비교한 결
과(Uitenbroek, 1997) 초등학생 집단의 경우, 한
국 기준집단은 주의집중문제와 DSM ADHD
척도에서 미국 기준집단보다 유의미하게 낮은
문제행동 수준을 보이고 있는 것으로 나타났
다. 반면, 한국 여자 기준집단은 사회문제, 한
국 남자 기준집단은 공격행동 영역에서 미국
기준집단보다 높은 문제행동 수준을 보이고

표 11-1. 한국판 TRF와 미국판 TRF 기준집단의 척도 별 평균 점수: 초등학생 집단

	한국 (2008년 자료)		미국 (2001년 자료)		<i>t</i> 값 (한국-미국 평균차)	
	남 (<i>n</i> =166)	여 (<i>n</i> =144)	남 (<i>n</i> =539)	여 (<i>n</i> =588)	남	여
불안/우울	2.18 (2.91)	3.01 (3.60)	2.50 (3.20)	2.70 (3.60)	-1.21	.93
위축/우울	1.56 (2.56)	1.57 (2.31)	1.40 (2.40)	1.60 (2.50)	.71	-.14
신체증상	.55 (1.33)	.68 (1.43)	.50 (1.40)	.60 (1.50)	.42	.60
사회적 미성숙	1.72 (2.63)	1.83 (2.56)	1.40 (2.20)	1.20 (2.10)	1.42	2.74**
사고문제	.70 (1.52)	.52 (1.25)	.60 (1.40)	.40 (1.00)	.76	1.07
주의집중문제	8.02 (8.90)	4.58 (7.08)	11.10 (11.00)	6.60 (8.60)	-3.68**	-2.93**
규칙위반	1.62 (2.56)	.99 (1.80)	1.60 (2.40)	.90 (1.90)	.09	.53
공격행동	4.56 (6.48)	2.93 (4.91)	3.40 (5.70)	2.10 (4.30)	2.07*	1.86
기타문제	.87 (1.39)	1.03 (1.45)	-	-	-	-
내재화	4.29 (5.56)	5.26 (6.15)	4.40 (5.30)	4.80 (5.90)	-.23	.81
외현화	6.18 (8.72)	3.92 (6.39)	5.00 (7.80)	3.00 (5.90)	1.56	1.57
문제행동총점	21.79 (24.20)	17.15 (21.03)	23.20 (23.10)	16.70 (19.80)	-.66	.23
DSM 정서문제	1.46 (2.13)	1.36 (2.19)	1.20 (2.00)	.90 (1.80)	1.40	2.34*
DSM 불안문제	.67 (1.17)	.85 (1.34)	.70 (1.20)	.70 (1.40)	-.29	1.19
DSM 신체화문제	.36 (.93)	.42 (1.09)	.30 (1.20)	.50 (1.40)	.68	-.74
DSM ADHD	4.49 (5.32)	2.22 (3.64)	6.00 (6.30)	3.50 (4.90)	-3.06**	-3.51**
DSM 반항행동문제	1.41 (1.94)	.95 (1.69)	1.30 (2.10)	.70 (1.60)	.63	1.61
DSM 품행문제	2.51 (4.15)	1.15 (2.31)	2.20 (3.60)	1.00 (2.50)	.87	.69

*: $p < .05$, **: $p < .01$, ()안은 표준편차

표 11-2. 한국판 TRF와 미국판 TRF 표준집단의 척도 별 평균 점수: 중·고등학생 집단

	한국 (2008년 자료)		미국 (2001년 자료)		<i>t</i> 값 (한국-미국 평균차)	
	남 (<i>n</i> =310)	여 (<i>n</i> =318)	남 (<i>n</i> =580)	여 (<i>n</i> =612)	남	여
불안/우울	1.35 (2.40)	2.01 (2.94)	2.00 (2.90)	2.20 (3.30)	-3.57**	-.90
위축/우울	1.45 (2.38)	1.30 (2.15)	1.80 (2.90)	1.60 (2.50)	-1.93	-1.91
신체증상	.35 (.95)	.43 (1.25)	.50 (1.30)	.40 (1.20)	-1.97*	.35
사회문제	1.17 (2.28)	1.13 (1.70)	1.10 (2.30)	.90 (2.00)	.44	1.84
사교적미성숙	.35 (.84)	.41 (1.07)	2.30 (1.30)	.30 (1.00)	-27.07**	1.52
주의집중문제	5.98 (8.36)	5.83 (8.04)	10.60 (11.20)	5.50 (8.00)	-6.95**	.60
규칙위반	1.42 (2.79)	1.19 (2.69)	1.90 (3.20)	1.20 (2.40)	-2.32*	-.06
공격행동	2.63 (5.28)	2.77 (4.89)	3.00 (5.70)	2.10 (4.90)	-.97	1.98*
기타문제	1.11 (1.36)	1.05 (1.35)	-	-	-	-
내재화	3.14 (4.64)	3.75 (5.11)	4.20 (5.60)	4.10 (5.80)	-3.02**	-.95
외현화	4.05 (7.52)	3.96 (7.31)	4.80 (8.40)	3.30 (6.90)	-1.36	1.33
문제행동총점	15.80 (20.50)	16.13 (20.65)	22.40 (25.30)	14.90 (21.20)	-4.21**	.85
DSM 정서적문제	1.24 (1.97)	1.27 (1.94)	1.60 (2.40)	1.10 (2.10)	-2.40*	1.23
DSM 불안문제	.49 (1.11)	.62 (1.10)	.60 (1.20)	.70 (1.30)	-1.37	-.99
DSM 신체화문제	.21 (.64)	.27 (.93)	.30 (1.00)	.30 (1.00)	-1.63	-.46
DSM ADHD	3.13 (4.90)	2.94 (4.73)	5.40 (6.20)	2.80 (4.50)	-5.99**	.44
DSM 반항행동문제	.79 (1.57)	.86 (1.63)	1.20 (2.10)	.80 (1.70)	-3.29**	.53
DSM 품행문제	1.57 (3.29)	1.29 (3.11)	1.90 (3.70)	1.10 (2.80)	-1.36	.91

*: $p < .05$, **: $p < .01$, ()안은 표준편차

있는 것으로 나타났다.

중·고등학생 집단의 경우, 한국 여자 표준집단이 공격행동에서 미국 표준집단보다 높은 문제행동 수준을 나타냈으나 대부분의 척도에서 한국 남자 표준집단이 미국 남자 표준집단보다 문제행동 수준이 낮은 것으로 나타났다. 특히 주의집중문제 척도에서는 성별에 상관없이 미국 표준집단보다 한국 표준집단의 문제행동 수준이 전반적으로 낮은 것으로 나타났다.

논 의

본 연구에서는 Achenbach와 Rescorla(2001)가 개발한 TRF를 국내 아동·청소년을 대상으로 표준화하고 측정 도구로서의 신뢰도와 타당도를 살펴보았다.

우선 미국 원척도의 구조가 한국의 아동·청소년을 대상으로 한 경우에도 적합한지를 알아보기 위해 확인적 요인 분석을 실시하였다. 그 결과 모든 적합도 지수들이 기준에 부합하지는 못했으나, 원저자가 모델 검증 방식

에서 사용한 적합도 지수인 RMSEA(.056)는 적합한 수준으로 나타났다. 이는 복잡한 요인 구조와 많은 문항이 특정 지수에 대한 적합도 수준을 저하시켰기 때문일 수 있다(Little, Cunningham, Shahar, & Widaman, 2002). 항목목록 확인적 요인분석을 실시해 본 결과에서는 적합도 지수들이 대체로 양호하게 나타났고, 2011년 새롭게 표준화된 한국판 CBCL 6-18, YSR과의 연계 활용 및 비교문화 연구의 가능성 등을 총체적으로 고려할 때 미국 원척도와 동일한 문항과 요인구조를 유지하는 형태로 개발하는 것이 바람직한 것으로 판단하였다. 그러나 문화권마다 중요시하는 가치(value)가 다른 것처럼 문제행동 수준의 유형 중에서도 어떤 문제행동을 더 심각하게 받아들이는 가에는 차이가 있을 수 있다. 따라서 향후 한국판 TRF가 교육 현장에서 활용되며 한국 아동·청소년 고유의 특성 및 유용성이 있는 구조가 발견된다면 이를 적극 반영해야 할 것이다.

두 번째로, 한국판 TRF의 신뢰도와 타당도를 살펴보았다. 먼저, 신뢰도를 평가하기 위해 내적 일치도(Cronbach's α) 및 재검사 신뢰도를 살펴본 결과 하위척도에 따라 차이는 있었지만 전반적으로 적절한 수준인 것으로 나타났다. 반면 아동·청소년에 대한 교사, 부모, 학생 평가의 상관관계를 통해서 본 일치도는 대부분 통계적으로 유의미하기는 했지만 그 수준은 대체로 낮은 편이었다. 미국판의 보고자간 평가 일치도 수준도 한국판에서와 마찬가지로 높지 않은 편이었다(Achenbach & Rescorla, 2001). 한국판과 미국판 척도에 따른 소척도별 내적 합치도 수준은 측정 오차에 따른 변량을 교정하지 않은 값이므로 직접적인 비교는 어렵지만, 전반적으로 한국판 척도의 내적 합치도 수준은 미국판에 비해 낮은 경향성이 나타

났다. 미국의 경우 각 가정을 방문하여 CBCL 6-18, YSR을 실시하고 부모를 통해 자녀의 교사에게 TRF를 실시하게 한 자료를 수집한 반면 우리나라에서는 교육 기관을 통해 교사 및 부모들로부터 자료를 수집함에 따라 한 명의 교사가 여러 명의 아동·청소년을 평가하게 되었는데, 이러한 자료 수집 방식의 차이가 평가자의 평정 태도 및 평가자 간 일치도에 영향을 미쳤을 가능성을 고려해야 할 것이다.

그러나 자료 수집 방식의 특성을 배제하더라도 아동·청소년 문제행동 평가에 있어 평정자 간에 상당한 불일치가 존재함을 보여주는 연구들이 상당수 존재한다(Kemper, Gerhardstein, Repper, & Kistner, 2003; Kumpulainen et al., 1999; Phares, 1997; Stanger & Lewis, 1993; Winsler & Wallace, 2002; Youngstrom, Loeber, & Stouthamer-Loeber, 2000). 평정자간 시각이 다를 수 있다는 전제 자체가 교사를 통한 평가척도의 개발 목적이기도 한 점을 생각해볼 때 이는 상식적으로 타당한 결과이며, 이러한 불일치성은 다양한 평가자들을 통한 정보 수집의 필요성을 역설적으로 보여준다. 즉 개인이 나타내는 문제행동의 영역이나 수준 등은 평가자에 따라 다르게 보고될 수 있기 때문에 자기보고나 부모 보고, 교사 보고의 어느 한 측면에 의존할 경우 평가 대상자의 적응 상태 및 문제행동이 과대 또는 과소 추정될 우려가 있다는 것이다. 실제 여러 평정자의 보고를 활용했을 때에 좀 더 정확한 진단을 내릴 수 있다는 연구 결과가 이를 뒷받침한다(Cantwell, Lewinsohn, Rohde, & Seeley, 1997). 다중평가자의 보고 결과를 수집, 통합하여 참조함으로써 문제행동의 수준에 대한 다차원적인 이해를 얻는 한편 평가자 간 차이가 내적 기준의 차이에서 기인한 결과인

지, 혹은 아동·청소년이 각 장면에서 다른 행동을 보이기 때문인지를 파악할 수 있다면 아동·청소년의 행동을 보다 정확하게 파악하고 이해하는데 도움이 될 수 있을 것이다.

그 다음 한국판 TRF의 타당도를 살펴보기 위해 TRF 척도 간 상관 분석, CBCL 6-18을 준거 검사로 활용한 점수 비교, 단축형 Conners 척도와 상관 분석을 실시한 결과, 한국판 TRF의 측정 도구로서의 타당도가 지지되었다. 한국판 TRF 척도들 간의 상관관계는 모두 통계적으로 유의한 것으로 나타났고, 특히 내재화, 외현화 척도와 이들을 구성하는 하위척도들 간의 상관관계가 상대적으로 높아 한국판 TRF의 수렴 타당도가 지지되었다. TRF 문제행동총점 임상범위 기준인 60T 이상/미만 집단에 대한 한국판 CBCL 6-18 척도의 평균 점수를 비교한 결과에서 TRF의 변별력이 지지되었다. 앞서 대상 집단 전체에 대한 평정자간 일치도에서 부모와 교사 보고 간 상관이 높지 않은 것으로 나타났음에도 불구하고, 교사에 의해 문제행동이 많은 것으로 보고된 상위 점수대 아동·청소년들은 부모 평가에서도 문제행동이 높게 보고될 가능성이 높은 것으로 나타난 것이다. 이는 교육 기관에서 TRF를 활용한 교사 평가를 통해 문제행동을 보이는 아동·청소년들을 선별하는 것이 유용함을 말해준다고 할 수 있다. 향후 실제 임상집단으로 분류된 집단과의 비교 분석을 통해 TRF의 임상적 변별력을 확인해야하겠지만, 본 분석 결과를 통해 임상집단 구분 기준으로 미국판 원 검사에서 사용하고 있는 60T 기준을 활용하는데에 무리가 없음을 일차적인 수준에서 확인해 볼 수 있었다. 또한 TRF 중 주의력과 관련된 척도들은 단축형 Conners 척도와 높은 상관관계를 보여 TRF의 검사 문항이 측정하고자

하는 내용 영역을 잘 반영하고 있음을 부분적으로나마 확인해볼 수 있었다.

셋째, 인구통계학적 집단 별 척도의 평균을 살펴본 결과 중요한 차이가 있는 것으로 나타난 성별(남/여)과 학년급간(초등학생/중·고등학생)을 기준으로 나누어진 네 개의 집단에 대해 각각 고유한 기준을 산출하는 것이 척도의 임상적 활용 측면에서 유용할 것으로 판단되었다. 성별에 따른 문제행동 점수를 비교해본 결과, 주로 내재화 척도에서는 여학생이, 외현화 척도에서는 남학생이 더 높은 문제행동 수준을 보이는 것으로 나타났다. 이러한 성차는 여러 선행 연구들(구미향, 이양희, 2003; 김연, 황혜정, 2005; McGee, Silva, & Williams, 1984)의 결과와 일치하며, 서울시 소아청소년 광역정신보건센터에서 발표한 2005년도 역학 연구 결과에서 남아가 여아보다 적대적 반항장애, 품행장애 등의 행동장애를 보이는 비율이 높았던 것과도 일맥상통한다. 한국판 CBCL 6-18, YSR의 표준화 연구(오경자, 김영아, 2011)에서도 성차가 나타난 하위 척도의 내용을 살펴보면 남아는 외현화, 여아는 내재화와 관련된 문제행동을 더 높게 보고하는 경향이 있었으며, 이러한 양상은 제 2차 성징이 나타나는 아동·청소년기에 접어들면서 성별에 따라 구체적으로 드러나는 문제의 양상에 있어서 중요한 차이를 나타내고 있을 가능성을 지지한다고 볼 수 있을 것이다.

학년급간 별 문제행동 점수를 비교해본 결과에서는 기타문제 척도를 제외한 모든 척도에서 초등학생이 중·고등학생 보다 높은 점수를 보였다. 한국판 TRF에서 나타난 학년급간 별 점수 차이와 더불어 우리나라에서 일반적으로 청소년기와 아동기를 구분하는 기준이 초등학교의 졸업 여부로 받아들여지고 있다는

점을 고려할 때 초등학교/중·고등학교 집단에 대해 각각 고유한 기준을 산출하는 것이 척도의 임상적 활용 측면에서 유용할 것으로 판단되었다. 한편 초등학생이 대부분의 척도에서 높은 점수를 나타낸 본 연구 결과는 위축/우울, 신체증상, 주의집중문제 등의 척도에서 중·고등학생이 유의미하게 높은 점수를 보고하였던 한국판 CBCL 6-18의 결과나 저학년일수록 행동장애의 비율이, 고학년일수록 주요우울장애 비율이 높았던 서울시 역학 연구 결과(2006)와 일치하지 않는 부분이다. 추후 연구를 통해 이러한 차이가 부모와 교사 평정자의 차이 때문인지, 혹은 다른 이유가 있는 것인지에 대한 원인 파악이 이루어져야 할 것이다. 또한 국내 교육 제도의 특성 상, 교사와 학생이 함께 보내는 물리적인 시간은 중·고등학교에서 훨씬 긴 것에 비해 교사-학생 간 상호작용 및 관찰의 기회는 담임교사가 전 과목을 가르치는 초등학교에서 더욱 많을 수 있다. 이처럼 학년급간에 따라 교사가 각 학생에 대해 파악하거나 인지하고 있는 정도의 차이가 있을 수 있으므로, 교사와 학생이 실제 상호작용하는 시간이나 방식, 학생에 대해 알고 있는 정보의 양이나 알고 지낸 기간 등 다양한 요인이 교사의 학생 평가에 어떠한 영향을 미치는가에 대한 연구도 함께 이루어져야 할 것이다. 또한 TRF는 초등학교 1학년 부터 고등학교 3학년까지 동일한 문항을 적용하여 문제행동을 평가하고 있는데, TRF 문항들이 국내에서는 특정 연령 집단에서의 문제행동을 보다 효과적으로 대표하거나 대표하지 못하고 있을 가능성이 있을 수 있음에 유의하고, 향후 이와 관련한 체계적인 연구가 이루어져야 할 것이다.

마지막으로 본 연구 자료를 미국 아동·청

소년의 문제행동 수준을 비교해 본 결과, 전반적으로 우리나라 아동·청소년들이 미국 아동·청소년 보다 낮은 수준의 문제행동을 보였다. 이와 같은 문화에 따른 문제행동 점수 차이는 기존 연구들에서도 찾아볼 수 있다. 예를 들어, 태국과 미국 아동·청소년에 대한 교사 평가 결과를 비교한 연구에서는 태국 아동·청소년의 점수가 거의 모든 부분에서 높게 나타났다(Weisz et al., 1989). 또한 ASEBA 검사군을 활용한 네덜란드, 터키, 모로코 비교 연구에서는 모로코 아동·청소년의 외현화 문제행동 수준이 네덜란드, 터키에 비해 높은 것으로 나타났으며(Stevens et al., 2003), 미국(mainland)과 푸에르토리코의 교사 보고 비교 연구에서는 푸에르토리코 아동·청소년의 점수가 더 높게 보고되었다(Achenbach et al., 1990). 푸에르토리코 부모 보고에서도 미국 부모에 비해 아동·청소년의 문제행동 점수가 더 높게 보고되었는데, 아동·청소년 본인 보고에서는 푸에르토리코 아동·청소년의 점수가 유의미하게 낮게 보고된 점을 주목할 만하다(Achenbach et al., 1990). 이와 반대로 일본 부모는 미국 부모에 비해 자녀의 문제행동총점을 낮게 보고했으나, 일본 청소년 스스로는 미국 청소년에 비해 약간 더 높은 점수를 보고한 것을 보인 연구도 있다(Rescorla et al., 2007). 즉, 비교 문화 연구 상에서 나타나는 차이가 실제로 문화권에 따른 아동·청소년의 문제행동 수준이나 영역에 차이가 있기 때문인지, 혹은 문제행동을 평가하는 평정자의 내적 기준에 차이가 있기 때문인지에 대해서는 보다 체계적인 연구가 필요할 것이다.

본 연구는 우리나라의 아동·청소년 교육 및 연구 현장에서 널리 사용될 수 있는 교사용 평가도구를 전국의 아동·청소년을 대상으

로 표준화 한 후, 측정 도구로서의 신뢰도와 타당도를 살펴본데 의의가 있다. 학교는 여러 아동·청소년들이 함께 모여 단체 생활을 통해 다양한 사회적 기술과 행동 양식을 배우는 곳으로, 이후 발달에 큰 영향을 미치는 중요한 요인이 될 수 있는 적응 관련 문제행동을 파악할 수 있다는 점에서 교사의 평가가 중요한 역할을 할 수 있다. 아동·청소년 발달 선별 및 진단 검사도구가 충분치 않은 상황에서 (이준석, 조광순, 2004) 한국의 전체 아동·청소년 인구분포를 반영한 폭넓은 기준집단을 바탕으로 개발된 한국판 TRF는 한국판 CBCL 6-18, YSR과 함께 활용됨으로써 아동·청소년 문제 행동에 대한 객관적이고 다각적 차원의 이해를 증진시키는데 도움이 될 수 있을 것으로 기대된다. 또한 영유아기 이후, 아동·청소년기 연령대와 연속성을 가질 수 있는 ASEBA 검사들을 발달적 연속선상에서 함께 활용할 수 있다는 점은 추후 장기 종단적 평생 발달 연구를 가능하게 한다는 측면에서 의미가 있으며, 미국판 검사와의 동일성을 유지하고 있기 때문에 비교 문화연구에서도 활용될 수 있을 것이다.

본 연구를 통해 한국판 TRF의 측정 도구로서 기본적인 요건들은 확인되었으나, 한국판 TRF를 표준화하는 것에 초점을 맞추었기 때문에 교사의 특성 자체와 관련한 분석에 있어서는 한계가 있었다. 본 연구에 참여한 교사는 주로 담임교사로, 대체로 2학기(10-12월)에 작성했다. 그러나 교사가 대상 학생을 얼마나 잘 알고 있는지, 대상 학생과 알고 지낸 시간은 어느 정도인지, 대상 학생과의 관계는 어떤지 등의 요인도 교사 평가에서는 중요한 요소가 될 수 있다. 따라서 추후 연구에서는 이러한 교사의 특성에 따른 구체적인 분석이 이

루어져야 할 것이다. 또한 향후 추가적인 연구를 통해 보완되어야 할 부분들도 남아있다. 첫 번째는 평정자간 일치도에 대한 보다 상세한 연구가 필요하다는 점이다. 평정자 간 일치도가 낮은 수준에 머물렀던 이유가 무엇에 기인하는 것인지 파악할 필요가 있다. 평정자간의 차이를 설명할 수 있는 기제가 밝혀지면 다중 평가 결과를 통합하는 ASEBA 행동 평가 척도의 활용도가 더욱 높아질 것이다. 둘째로 임상기관에 내원한 아동·청소년들을 대상으로 세분화된 임상 진단 집단에 대한 변별력 검토가 이루어져야 할 필요성이 있다. 구체적인 문제 행동 및 임상 집단을 변별하는 데에 한국판 TRF의 하위척도들이 유의미한 역할을 할 수 있는지 검증 될 때 한국판 TRF의 임상적, 학문적 활용이 보다 활발하게 이루어질 수 있을 것이다. 셋째, 우리나라 문화에 특정한 교사-학생간의 상호관계 특성, 교육 환경 등의 요인을 고려하여 지속적으로 관찰되는 요인 구조나 임상적으로 유용한 부가적인 하위척도가 있는지 탐색해가는 연구가 필요하다. 넷째, TRF 검사 내부에는 평정의 신뢰도나 타당도를 점검할 수 있는 장치가 포함되어 있지 않는데 이를 보완할 수 있는 방법에 대한 추가적인 연구도 필요할 것이다.

참고문헌

- 구미향, 이양희 (2003). 아동기 두려움과 행동 문제의 특성 및 관련성 탐색. *초등교육연구*, 16(1), 341-363.
- 김 연, 황혜정 (2005). 부모와 자녀가 지각한 촉진적 의사소통이 아동의 문제행동에 미치는 영향. *열린유아교육학회*, 10(4), 69-

- 88.
- 김영아, 이 진, 김유진, 김민영, 오경자 (2011). 한국판 유아행동평가 척도 교사용(C-TRF)의 표준화 연구. *한국심리학회지: 발달*, 24(2), 65-86.
- 김영아, 이 진, 문수중, 김유진, 오경자 (2009). 한국판 CBCL 1.5-5의 표준화 연구. *한국심리학회: 임상*, 28(1), 117-136.
- 서울시 소아청소년 광역 정신보건센터 (2006). 서울시 소아청소년 정신장애 유병률 조사: 2005년도 역학사업보고서.
- 소유경, 노주선, 김영신, 고선규, 고윤주 (2002). 한국어판 부모, 교사 ADHD 평가 척도의 신뢰도와 타당도 연구. *신경정신의학*, 41(2), 283-289.
- 신현숙, 박용재, 박주희, 류정희 (2006). 교사보고형 초등학생 학교적응행동 평정척도의 개발과 타당도 검증. *한국심리학회지: 학교*, 3(2), 1-26.
- 연합뉴스 (2009.08.23). 초중고생 13% 정신건강 적신호.
- 연합뉴스 (2011.08.22). 청소년, 정신건강 문제 늘었다.
- 오경자, 김영아 (2011). ASEBA 아동청소년 행동평가척도 매뉴얼. 서울: (주)휴노.
- 오경자, 이해련 (1989). 주의력 결핍 과잉 활동 증 평가도구로서의 단축형 Conners 평가척도의 활용. *한국심리학회지: 임상*, 8(1), 135-142.
- 오경자, 이해련, 홍강의, 하은혜 (1997). K-CBCL 아동 · 청소년 행동평가검사. 서울: 중앙적성 출판사.
- 오경자, 하은혜, 이해련, 홍강의 (2001). K-YSR 청소년 자기행동 평가척도. 서울: 중앙적성출판사.
- 이경남 (2002). 청소년의 애착, 자아존중감 및 충동통제가 비행에 미치는 영향. *대한가정학회지*, 40(8), 191-206.
- 이준석, 조광순 (2004). 국내 영 · 유아 발달 선별 및 진단 검사도구에 대한 고찰: 표준화중인 검사도구들을 중심으로. *특수교육저널: 이론과 실천*, 5(1), 1-26.
- 장휘숙 (1997). 청소년기의 애착과 자아존중감 및 자기효능감의 관계. *인간발달연구*, 4(1), 88-106.
- 정현희 (2002). 한국판 교사용-아동평정척도(T-CRS)의 신뢰도와 타당도. *한국심리학회지: 발달*, 15(2), 73-88.
- 통계청 (2006). 2005년 대한민국 인구 및 주택 총조사.
- Achenbach, T. M. (1991). *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF Profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., Bird, H. R., Canino, G., Phares, V., Bould, M. S., & Rubio-Stipec, M. (1990). Epidemiological Comparisons of Puerto Rican and U.S. Mainland Children: Parent, Teacher, and Self-Reports. *Journal of the American Academy of Child & Adolescent Psychiatry*, 29(1), 84-93.
- Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for 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 and Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth,

- and Families.
- Byrne, B. M. (2001). *Structural equation modelling with AMOS: Basic concepts, applications and programming*. Mahwah, NJ: Erlbaum.
- 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.
- Cantwell, D. P., Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1997). Correspondence Between Adolescent Report and Parent Report of Psychiatric Diagnostic Data. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(5), 610-619.
- Cassidy, J., & Asher, S. R. (1992). Loneliness and peer relations in young children. *Child Development*, 63, 350-365.
- Conners, C. K. (1973). Rating scales for use in drug studies with children. *Psychopharmacology Bulletin (Special Issue, Pharmacotherapy of Children)*, 24-29.
- Dedrick, Tan, & Marfo (2008). Factor structure of the Child Behavior Checklist/6-18 in a sample of girls adopted from China. *Psychological Assessment*, 20, 70-75.
- Delaney-Black, Covington, Templin, Ager, Nordstrom-Klee, Martier, Leddick, Czerwinski, & Sokol (2000). Teacher-assessed behavior of children prenatally exposed cocaine. *Pediatrics*, 106, 782-791.
- Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on the revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology*, 6, 221-236.
- Gresham F. M., & Elliott, S. N. (1990). *Manual for the Social Skills Rating System*. Circle Pines MN: American Guidance Service.
- Holye, R. H. (1995). *The structural equation modelling approach: Basic concepts and fundamental issues*. In R.H. Hoyle (Ed.), *Structural equation modelling, concepts, issues, and applications* (pp.1-15). Thousand Oaks, CA: Sage.
- Hops, H. (1983). Children's social competence and skill: Current research practices and future directions. *Behavior Therapy*, 14(1), 3-18.
- Kemper, T. S., Gerhardstein, R., Repper, K. K., & Kistner, J. A. (2003). Mother - Child Agreement on Reports of Internalizing Symptoms Among Children Referred for Evaluation of ADHD. *Journal of Psychopathology and Behavioral Assessment*, 25(4), 239-250.
- 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*, 9, 151-173.
- McGee, R., Silva, P. A., & Williams, S. (1984). Behaviour problems in a population of seven-year-old children: Prevalence, stability, and

- types of disorder-a research report. *Journal of Child Psychology and Psychiatry*, 25, 251-259.
- Mychailysz, Mendez, & Kendall (2010). School Functioning in Youth with and without Anxiety Disorders: Comparisons by Diagnosis and Comorbidity. *School Psychology Review*, 39(1), 106-121.
- 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*. Available from the senior author at P. O. Box 248074, University of Miami, Coral Gables, FL 33124.
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., Bird, H., Broberg, A., Dobrea, A., Döpfner, M., Erol, N., Forns, M., Hannesdottir, H., Kanbayashi, Y., Lambert, M. C., Leung, P., Minaei, A., Mulatu, M. S., Novik, T. S., Oh, K., Roussos, A., Sawyer, M., Simsek, Z., Steinhausen, H., Weintraub, S., Metzke, C. W., Wolanczyk, T., Zilber, N., Zukauskienė, R., Verhulst, F. (2007). Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology*, 75, 351-358.
- Reynolds, C. R., & Kamphaus, R. W. (1992). *BASC: Behavior assessment system for children*. Circle Pines, MN: American Guidance Service.
- Segars, A. H., & Grover, V. (1993). Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *MIS Quarterly*, 17, 517-525.
- Somers, T., Nelson, K., & Karimi, J. (2003). Confirmatory factor analysis of the end-user computing satisfaction instrument: Replication within an ERP domain. *Decision Sciences*, 34, 595-621.
- 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.
- Stevens, G. W. J. M., Pels, T., Bengi-Arslan, L., Verhulst, F. C., Vollebergh, W. A. M., & Crijnen, A. A. M. (2003). Parent, teacher and self-reported problem behavior in The Netherlands Comparing Moroccan immigrant with Dutch and with Turkish immigrant children and adolescents. *Social Psychiatry and Psychiatric Epidemiology*, 38(10), 576-585.
- Uitenbroek, D. G. (1997). "SISA Binomial," uthampton: D. G. Uitenbroek. Retrieved 1 January, 2004, from the World Wide Web: <http://home.clara.net/sisa/binomial.htm>.
- Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Achenbach, T. M., & Trevathan, D. (1989). Epidemiology of Behavioral and Emotional Problems among Thai and American Children: Teacher Reports for Ages 6-11. *Journal of Child Psychology and Psychiatry*, 30(3), 471-484.
- 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.

원 고 접 수 일 : 2012. 06. 08.

수정원고접수일 : 2012. 08. 09.

최종게재결정일 : 2012. 08. 13.

Standardization Study for the Korean Version of the Teacher's Report Form

Min-Young Kim

Young-Ah Kim

Kyung-Ja Oh

Huno Inc.

Department of Psychology,
Yonsei University

The purpose of this study was to develop the norms and test the reliability and validity of the Korean version of the Teacher's Report Form (TRF). The selection of a normative sample for the TRF (476 boys and 462 girls) was based on data from the 2005 Korean Population and Housing Census. A confirmatory factor analysis supported that the factor structure of the original TRF can be plausibly applied to the Korean version of the TRF. Reliability was assessed in terms of internal consistency (.59-.97; other problems: .39) and test-retest reliability (.53-.96). The subscales of the TRF showed high correlations with its higher-order-factors such as internalizing and externalizing problems, and this provided good evidence of convergent validity. When the TRF was administered to a clinical sample ($n=101$) and normative sample ($n=506$), discriminant validity was confirmed by the mean differences on the subscales between the two samples. When differences in mean subscale scores for different age and sex groups were explored, it would be reasonable to use independent norms for two age groups (the first year ~ the sixth year of elementary school / middle school freshman ~ high school senior) and gender (boys / girls) for the Korean version of the TRF. The implications of using the TRF in Korea both as a clinical tool and a research instrument are discussed.

Key words : Teacher's Report Form, child behavior problems, school-age, standardization