

한국 아동·청소년의 문제행동 변화 양상: 한국판 CBCL, YSR의 1, 2차 표준화 자료 비교

김민영

(주)휴노

이진

Harvard School
of Public Health

김영아

(주)휴노

하은혜

숙명여자대학교
아동복지학부오경자[†]

연세대학교

본 연구에서는 시기에 따른 문제행동 수준의 변화 양상을 보다 정확하게 확인하기 위한 심리측정학적 선행 조건으로 한국판 아동·청소년 행동평가척도 CBCL과 YSR의 1차(1991년 미국판 원검사 기반), 2차(2001년 미국판 원검사 기반) 전국 표준화 자료에 대한 측정 동등성 분석을 실시하였다. 분석 결과, CBCL과 YSR 모두 1, 2차 표준화 시기에 걸쳐 비교적 강한 수준의 측정 동등성이 지지되어 문제행동에 대한 전반적 평가 태도, 경향, 그리고 문항 및 척도가 두 시기의 평정자들에게 유사한 의미로 이해되고 있는 것으로 나타났다. 이는 두 시기의 문제행동 측정 체계가 안정적이라는 의미로, CBCL과 YSR의 1, 2차 표준화 시기에 따른 점수 비교를 통해 아동·청소년 문제행동 수준의 변화, 즉 문제행동 발현율의 변화를 살펴보는 것이 타당함을 알려준다. 각 표준화 시기에 따른 검사별로 600명씩 성별 및 연령(12-14세, 15-17세) 분포를 동일하게 맞춘 대응 표본 비교 집단을 구성하여 점수를 비교한 결과, 1, 2차 표준화 시기에 걸쳐 한국 아동·청소년 문제행동 수준이 부모보고(CBCL)에서는 감소했으나, 아동·청소년 본인보고(YSR)에서는 유의미한 변화가 없었다. 또한 두 시기의 문항별 점수의 상관계수가 CBCL, YSR 모두 매우 높아($r > .90$) 해당 기간에 걸친 문제행동 항목들 간의 상대적 발현 빈도가 유사한 수준인 것으로 나타났다. 본 연구에서 확인된 한국판 CBCL, YSR의 임상적 유용성과 한계점에 대한 논의가 이루어졌다.

주요어 : CBCL, YSR, 행동평가척도, 아동·청소년 문제행동, 문제행동 수준 변화 양상

[†] 교신저자: 오경자, 연세대학교, 서울특별시 서대문구 연세로 50

E-mail: kjoh@yonsei.ac.kr

기성세대들이 다음 세대를 보며 “우리 때는 안 그랬는데”라고 말하는 현상은 어느 시대에나 있어왔다. 실제로 특정 연령대의 행동 양상이 시기에 따라 변화하는 것인지를 확인하기 위해서는 문제행동의 발현율(prevalence)이 수년간에 걸쳐 어떻게 변화하는지(secular change)를 살펴볼 필요가 있다. 그런데 이와 같이 서로 다른 시기의 자료를 비교하여 문제행동의 변화 양상을 살펴보는 경우, 거시적인 관점에서 문제행동 발현율의 변화를 살펴보고자 하기 때문에 발현율 변화의 구체적인 예측변인을 가정하기가 쉽지 않다. 그럼에도 불구하고 시기에 따른 문제행동의 변화 양상을 확인하는 것이 중요한 이유는 다음과 같다. 첫째, 특정 문제행동의 발현율이 급격히 증가 또는 감소하는지를 확인함으로써 구체적인 역학조사를 통해 발현율 변화에 영향을 미치는 원인을 규명하고 대책을 마련할 필요가 있는 문제행동 영역을 발견하는 데 도움이 될 수 있다(Collishaw, Maughan, Goodman, & Pickles, 2004). 특히 아동·청소년의 적응과 발달에 밀접한 영향을 미치는 경제, 가족, 교육, 미디어 등의 사회적 요인들이 시간에 따라 변화하기 때문에 이러한 요인들을 고려하여 문제행동 발현율이 어떻게 변화하는지를 이해하는 것이 필요하다(Loeber, Burke, Lahey, Winters, & Zera, 2000). 둘째, 연령대 및 성별에 따라 발현율의 변화 양상이 어떻게 나타나는지를 확인함으로써 어떤 집단에 체계적인 개입이 시급한지를 파악할 수 있다(Tick, Van der Ende, & Verhulst, 2007). 셋째, 문제행동 측정 도구의 활용에 있어 새로운 기준 점수를 적용할 시기를 판단하는 데 도움이 될 수 있다(Au et al., 2004). DSM(Diagnostic and Statistical Manual of Mental Disorders; 정신 장애 진단 및 통계 편람)과 같

은 진단 기준도 꾸준히 개정 작업(American Psychiatric Association, 1952, 1968, 1980, 1987, 1994, 2000, 2013)이 이루어지고 있는 만큼 문제행동의 기준 점수 역시 시간의 흐름에 따라 개정이 필요할 수 있다.

아동·청소년 문제행동의 시간에 따른 변화 양상(time trend)에 대한 연구들은 비교 시기, 문제행동 영역, 연령, 성별, 평정자 등에 따라 다양한 결과를 보고하고 있다. 예를 들어, Achenbach와 동료들은 7-16세 미국 아동·청소년의 문제행동이 1976년에서 1989년 사이에 증가했지만(Achenbach & Howell, 1993), 1989년에서 1999년 사이에는 약간 감소했다고 보고했다(Achenbach, Dumenci, & Rescorla, 2003). 또한 네덜란드의 4-16세 아동·청소년을 대상으로 실시한 Verhulst, Ende와 Rietbergen(1997)의 연구에서는 1983-1993년의 10년간 아동·청소년의 문제행동에 큰 변화가 없었다. 이러한 결과들은 서로 다른 시기에 모집단으로부터 무선적으로 표집된 동일 연령대 아동·청소년 집단의 문제행동 수준이 다를 수 있으며, 비교 시기 및 문화권에 따라 변화 양상이 다를 수 있음을 보여준다. 즉, 국내에서도 아동·청소년의 문제행동 수준이 시기에 따라 변화하는지를 독자적으로 살펴볼 필요가 있음을 시사한다.

한편, Collishaw 등(2004)의 연구에서는 영국 15-16세 청소년의 품행문제가 1974년에서 1999년까지 유의미하게 증가한 반면, 정서문제는 1986년에서 1999년 사이에만 증가한 것으로 나타났다. 또한 Eimecke, Pauschardt, Remschmidt, Walter와 Mattejar(2011)의 연구에서 8-11세 독일 아동은 1987-2008년의 21년에 걸쳐 신체증상 영역에서만 문제행동 수준이 유의미하게 변화한 것으로 나타났다. 이와 같은 결과는

문제행동 영역에 따라 문제행동 수준의 시간에 따른 변화 양상이 다를 수 있기 때문에 다양한 문제행동 영역을 포괄적으로 살펴볼 필요가 있음을 의미한다.

문제행동 수준의 변화 양상을 살펴볼 때 표집의 연령대를 함께 고려할 필요성을 시사하는 연구 결과도 있다. 예를 들어, 네덜란드 6-16세 아동·청소년의 20년간 문제행동 변화를 조사한 Tick 등(2007)의 연구에서 1983-1993년에 6-11세 집단의 문제행동은 감소하고 12-16세 집단의 문제행동은 증가하였으나, 1993-2003년에는 모든 연령대에서 문제행동이 증가했다고 보고했다.

시간에 따른 문제행동 변화 양상이 아동·청소년의 성별에 따라 다를 수 있음을 보여주는 연구 결과들도 있다. 예를 들어, Sourander 등(2004)의 연구에서 8-9세 핀란드 남아의 문제행동이 1989-1999년의 10년 사이에 감소한 것에 비해 여아는 변화가 없었다. 또한 8-11세 독일 아동을 대상으로 한 Eimecke 등(2011)의 연구에서 1987년에는 남아가 여아보다 높은 수준의 신체증상 문제를 보였으나, 2008년에는 여아가 더 높은 수준을 보일 정도로 여아의 문제 수준 변화폭이 큰 것으로 나타났다.

또한 1983-2003년에 걸쳐 6-16세 네덜란드 아동·청소년을 대상으로 살펴본 Tick 등(2007)의 연구에서 부모는 자녀의 문제행동 수준 및 임상 비율이 약간 증가했다고 보고했으나, 교사는 학생들의 문제행동 수준에 변화가 거의 없는 것으로 보고했다. 이러한 결과는 문제행동 수준을 누가 평가하는지에 따라 아동·청소년 문제행동 발현율의 변화 양상에 차이가 있을 수 있음을 보여준다. 따라서 관찰된 발현율의 변화가 실재(實在)하는지 아니면 문제행동에 대한 평정자의 지각이 바뀐 것인지를

확인해 보아야 할 것이다. 게다가 기존 연구들에서 아동·청소년 행동 및 정서문제 평가에 있어 부모, 교사, 아동·청소년 본인 등 다양한 평정자 간에 전반적으로 낮은 일치도가 보고되고 있다(Achenbach, McConaughy, & Howell, 1987; Phares, 1997; Sawyer, Baghurst, & Clark, 1992; Stanger & Lewis, 1993). 이와 같은 평정자 간 차이는 어느 한 평정자 집단의 보고가 더 중요하거나 옳고 그르다기보다는 평정자에 따라 문제행동을 평가하는 우선적 기준이 다를 수 있으며 평정자와 평정 대상 간 관계의 맥락이 중요함을 의미한다(Youngstrom, Loeber, & Stouthamer-Loeber, 2000). 즉, 어느 한 집단의 평정자로부터 얻어진 정보에만 의존하기보다는 다양한 평정자들의 보고를 통합적으로 활용하여 문제행동을 이해하는 과정이 중요하기 때문에 평정자에 따른 변화 양상도 함께 살펴볼 필요가 있다.

그런데 이와 같이 아동·청소년 문제행동의 시간에 따른 변화 양상을 살펴본 기존 연구들에서는 관찰된 문제행동 점수의 변화가 실제로 시기에 따른 문제행동 수준의 변화를 반영하는 것인지, 아니면 측정 방법과 관련된 다른 요인들의 변화에 수반되는 현상인 것인지에 대해 구체적으로 확인하고 있지 않다. 만약 이들 연구에서 서로 다른 시기에 걸쳐 문제행동 평가 도구의 측정 동등성(measurement invariance/equivalence)이 확보되지 않았다면 관찰된 결과들은 재해석될 필요가 있다. 왜냐하면 심리계량적 측면에서 동일한 의미를 지니지 않는 평가 도구로 측정된 문제행동은 그 의미가 전혀 다를 수 있기 때문이다. Vandenberg와 Lance(2000)는 이와 같은 비교가 “발음만 비슷할 뿐 본질적으로 전혀 관련이 없는 두 단어인 샌드위치(sandwich)와 샌드웨이

(sand wedge)의 비교와 같다(p. 40)”고 표현하고 있다.

따라서 시간에 따른 문제행동의 변화가 관찰되었다고 하더라도 실제로 문제행동의 수준이 변화했다는 결론을 내리기 전에 다음과 같은 가능성을 고려하여 해석할 필요가 있다. 첫째, 문제행동 수준에 대한 평정자들의 평가 태도 및 경향이 변화했을 수 있다. 다시 말해, 특정 문제행동에 대한 인식의 재고가 이루어졌거나(Padmanabhan, Sugunan, Brahmaphuran, Nandini, & Pavichran, 2003), 특정 문제행동에 대한 사회적 태도에 변화가 있었을 가능성이 있다(Bhatia, Khan, Mediratta, & Sharma, 1987). 예를 들어, 이전에는 문제로 여겨지지 않던 행동이 새롭게 문제행동으로 여겨지거나, 보고하는 것이 금기시되던 문제행동들을 보다 개방적인 태도로 보고하게 되었을 수 있다. 둘째, 문제행동 측정 도구의 의미가 변화했을 가능성이 있다(Collishaw et al., 2004; de Beurs, Fokkema, de Groot, de Keijsjer, & Kerkhof, 2014; Vandenberg & Lance, 2000). 사회·문화적 맥락의 변화에 따라 측정 도구의 안면·구성 타당도가 체계적으로 변화할 수 있고, 문항 및 척도가 측정하고자 하는 문제행동 요인 구조의 적합성이 측정 시기에 따라 다를 수 있다. 즉, 시간에 따른 문제행동 수준의 차이가 관찰될 경우 그것이 실제 문제행동 수준의 변화에 따른 결과라기보다는 문제행동 진단 기준, 평가 도구, 그리고 문제행동을 보고하는 방식이 시간에 따라 달라진 결과일 수 있다는 것이다(Collishaw et al., 2004).

사회·문화적 요인이 문제행동의 수준뿐 아니라 문제행동의 평가 방식에도 영향을 미칠 수 있기 때문에 문제행동 양상의 변화를 읽어내는 것은 단일 차원으로 접근하기 어려운 일

일 것이다. 문제행동 평균 점수의 차이가 두 시기의 변화를 보여주는 일차적인 정보일 수 있으나, 이에 앞서 문제행동에 대한 시각 및 평가의 틀이 동일하게 유지되고 있는지를 파악하는 것도 중요하다. 따라서 본 연구에서는 서로 다른 시점의 아동·청소년 문제행동 발현율을 비교하기 위한 심리측정학적 선행 조건으로 평가 도구의 시기에 따른 측정 동등성 분석 방법 및 결과 해석의 예시를 제시한 후, 한국 아동·청소년의 시간에 따른 문제행동 발현율의 변화 양상을 살펴보고자 한다.

특히 우리나라는 사회·경제적으로 매우 빠른 변화를 겪어왔으며, 이러한 급격한 변화가 아동·청소년 정신 건강에 영향을 미칠지도 모른다는 우려가 높아지고 있다. 최근 아동·청소년의 높은 자살률, 학교 폭력 증가, 흡연 및 음주 등 물질 남용의 급증과 같은 사회적 통계(강은정, 2007; 김용석, 모선영, 최현정, 이자영, 2008; 질병관리본부, 2014)가 이와 같은 우려를 뒷받침하고 있지만, 실질적으로 이러한 현상을 확인할 경험적 자료는 거의 없는 실정이다. 동일한 표집을 4-5년에 걸쳐 추적함으로써 문제행동 수준의 변화를 파악하려는 종단 연구가 이루어지기는 했으나 공격성 및 우울과 같은 특정 문제행동 영역에 제한되어 있는 경향이 있었고(김재철, 최지영, 2010; 서미정, 2009, 조정아, 2009; 좌현숙, 오승환, 2011), 전국적인 표집을 통해 특정 시기의 아동·청소년을 대표하는 서로 다른 코호트(cohort) 집단에서 다양한 문제행동의 발현율이 어떻게 다른지를 확인하는 연구는 아직 보고된 바 없다.

이와 같은 연구문제를 살펴보기 위해 본 연구에서는 CBCL(Child Behavior Checklist; 아동·청소년 행동평가척도 부모용)과 YSR(Youth

Self-Report; 청소년 행동평가척도 자기보고용)을 활용하고자 한다. CBCL과 YSR은 아동·청소년 문제행동을 평가하는 대표적인 경험적(empirical) 접근 방식의 행동평가척도로, 임상 현장에서 관찰되는 주요 문제행동들을 취합한 후 통계적 분석을 통해 증후군(syndrome)을 도출하는 방식으로 구성되었다. 출시된 이후 오랜 시간에 걸쳐 다양한 연구를 바탕으로 임상적 활용도를 증진시키기 위한 개정 작업이 진행되어 왔는데(Achenbach, 1991a; Achenbach & Edelbrock, 1983), 2000년대에 들어서는 그동안 개발되어 온 개별검사들의 평가 대상 및 평정자 관계를 구조화하여 전 연령대에 걸쳐 문제행동을 평가하는 ASEBA(Achenbach System of Empirically Based Assessment) 시스템으로 구축하면서 전반적인 개편이 이루어졌다(Achenbach & Rescorla, 2001). 이 과정에서 CBCL과 YSR에는 기존에 포함되었던 문항들 중 아동·청소년기에 주로 보이는 문제행동을 측정하는 데 다소 적합하지 않은 것으로 평가된 문항들이 제외되고 이를 대체하는 새로운 문항들이 포함되었다. 이로 인해 증후군 소척도를 구성하는 문항들의 분류가 바뀌는 등의 변화가 있었다.

국내에서는 CBCL과 YSR 모두 두 차례에 걸쳐 전국 규모의 표준화가 이루어졌다. 한국판 1차 표준화는 1991년 미국판 CBCL 4-18과 YSR을 기반으로 이루어졌으며(K-CBCL 아동·청소년 행동평가척도: 오경자, 이해련, 홍강의, 하은혜, 1997; K-YSR 청소년 자기행동 평가척도: 오경자, 하은혜, 이해련, 홍강의, 2001), 이후 ASEBA 시스템 도입 후 개편된 2001년 미국판 CBCL 6-18과 YSR을 기반으로 한국판 2차 표준화 작업이 이루어졌다(CBCL 6-18 아동·청소년 행동평가척도 부모용; YSR 청소년 행동평가척도 자기보고용; 오경자, 김영아,

2011). 앞서 언급한 바처럼 ASEBA 시스템 도입 전후로 일부 문항 구성에 변화가 있기는 했으나 대부분의 문항이 그대로 유지되고 있어 두 시기의 문제행동 평가에 대한 다양한 비교 분석이 가능하다.

따라서 본 연구에서는 한국판 CBCL과 YSR의 1, 2차 표준화 자료에 대한 측정 동등성 분석을 실시하여 두 시기의 문제행동 평가 도구 측정치를 직접적으로 비교할 수 있는지 확인한 후, 한국 아동·청소년의 문제행동 점수 변화 양상을 확인해 보고자 하였다. 이때, 부모보고(CBCL)와 청소년 자기보고(YSR) 자료를 함께 확인함으로써 문제 영역, 연령, 성별뿐만 아니라 평정자에 따라서 변화 양상이 어떻게 나타나는지를 탐색적으로 확인해 볼 수 있을 것이다.

방 법

연구 대상

본 연구에서는 한국판 CBCL과 YSR의 1, 2차 표준화 자료를 사용하였다. 두 시기의 평가 태도 및 경향의 유사성과 문항 및 척도가 평정자들에게 동일한 의미로 이해되고 있는지를 확인하는 측정 동등성 분석은 척도 전반에 대한 검토이므로 표준화 집단 전체를 대상으로 하였다. 실질적인 점수 비교 분석에서는 두 시기의 표본 표집이 다를 수 있음을 고려하여 성별 및 연령 분포를 동일하게 맞춘 대응 표본(matched sample)을 사용했다.

표준화 집단

한국판 CBCL¹⁾, YSR의 1차 표준화 자료는

표 1. 한국판 CBCL과 YSR의 1, 2차 표준화 집단 구성 (단위: 명)

검사	표준화 시기	연령 범위	남	여	전체
CBCL	1차 표준화	4-17세	1,964	1,944	3,908
	2차 표준화	6-18세	1,353	1,263	2,616
YSR	1차 표준화	12-17세	2,549	4,021	6,570
	2차 표준화	12-18세	861	812	1,673

주로 교육부가 발행한 1994년 교육연감에 제시된 각급학교 수, 학급 수 등의 통계자료에 의거해 유치원, 초, 중, 고등학교를 대상으로 전국 각지에서 학급 단위로 유층화 표집 방법을 사용하여 1995-1997년에 걸쳐 수집하였다. 한국판 CBCL, YSR의 2차 표준화 자료는 2005년 대한민국 인구 및 주택 총조사(통계청, 2006)의 전국 지역별 인구 분포를 반영하여 2007-2008년에 걸쳐 전국 각 지역의 초, 중, 고등학교로부터 수집하였다. 표준화 집단 구성은 표 1에 제시되어 있다.

1) CBCL 6-18의 전신이라고 볼 수 있는 CBCL 4-18(Achenbach & Edelbrock, 1983)이 출시된 이후, 국내에서 CBCL을 도입하기 위해 1988년부터 여러 차례의 예비 조사가 이루어졌다. 이 시기의 CBCL은 연령 및 성별 등에 따라 각기 다른 종류의 증후군들로 구성되어 있었는데 이처럼 단일화되지 않은 척도를 사용할 경우 실제 연구 및 임상 장면에서 여러 가지 어려움이 있는 것으로 나타났다. 따라서 원저자는 연령 및 성별 집단에 대해 동일한 문제행동 증후군 척도로 재구성한 검사를 출시하였다(Achenbach, 1991a). 한국판 CBCL 1차 표준화 작업은 원검사의 개편에 따라 1991년 미국판 척도 구조를 기반으로 이루어졌으며, 예비 조사에서 사용한 자료들도 포함되었다. 미국 원검사에서도 1991년판 요인 구조 도출 시 1983년판 연구에 활용되었던 자료를 일부 활용하였다(Achenbach & Rescorla, 2007, p.159)

대응 표본

대응 표본은 CBCL과 YSR 실시 대상 연령에 모두 포함되는 12-17세 자료에서 추출하였다. 표본을 표준화 시기(1차, 2차), 연령(12-14세, 15-17세), 성별(남, 여)에 따라 분류한 후, 각 집단에서 150명씩 무선 표집(random sampling)²⁾을 실시하여 각 검사별로 표준화 시기에 따라 (예: CBCL 1차 표준화) 600명씩 대응 표본 비교 집단을 구성하였다(표 2).

전체 표준화 집단과 대응 표본의 문제행동 총점(총 문제행동; 최종 분석에 포함된 99개 문항의 응답 총점)에 대한 Simple t test (Uitenbroek, 1997)를 실시한 결과, 그 차이가 유의미하지 않은 것으로 나타나($|t|s < 1.0$) 대응 표본을 전체 표준화 집단의 대표 표본으로 사용할 수 있을 것으로 보인다(표 3).

연구 도구

CBCL과 YSR은 아동·청소년기에 주로 나타날 수 있는 정서행동문제가 기술된 각 문항에 대해 지난 6개월 내에 아동·청소년이 그 행동을 보였는지를 판단하여, '전혀 해당되지 않

2) 무선 표집은 SPSS 15.0 for windows 평가판에서 사례수의 조건부 선택(select case)의 무작위 표본 추출(random samples of cases) 방식을 통해 이루어졌다.

표 2. 대응 표본 비교 집단 구성 (단위: 명)

검사	표준화 시기	연령 범위	남	여	전체
CBCL	1차 표준화	12-14세	150	150	300
		15-17세	150	150	300
		전체	300	300	600
	2차 표준화	12-14세	150	150	300
		15-17세	150	150	300
		전체	300	300	600
YSR	1차 표준화	12-14세	150	150	300
		15-17세	150	150	300
		전체	300	300	600
	2차 표준화	12-14세	150	150	300
		15-17세	150	150	300
		전체	300	300	600

표 3. 전체 표준화 집단과 대응 표본의 문제행동총점(총 문제행동) 평균 비교

	CBCL			YSR		
	전체 표준화 집단	대응 표본	<i>t</i>	전체 표준화 집단	대응 표본	<i>t</i>
1차 표준화	23.47 (18.18)	22.81 (17.67)	.85	38.71 (21.39)	38.72 (21.00)	-.01
2차 표준화	13.79 (13.71)	14.36 (13.56)	-.93	37.77 (19.26)	37.85 (19.86)	-.09

주. () 안은 표준편차, 문제행동총점은 본 연구의 최종 분석에 포함된 99개 문항의 응답 총점임.

는다: 0, 가끔 그렇거나 그런 편이다: 1, 자주 그런 일이 있거나 많이 그렇다: 2'의 3점 척도로 평가하도록 되어 있다. CBCL은 부모(주양육자), YSR은 아동·청소년 본인이 실시한다.

CBCL(Child Behavior Checklist; 아동·청소년 행동평가척도 부모용)

1차 표준화를 통해 출시된 K-CBCL 아동·청소년 행동평가척도(오경자 등, 1997)는 1991년 미국판 CBCL 4-18(Achenbach, 1991a)을 표준

화한 것이며, 2차 표준화를 통해 출시되어 현재 사용 중인 한국판 CBCL 6-18 아동·청소년 행동평가척도 부모용(오경자, 김영아, 2011)은 2001년 미국판 CBCL 6-18(Achenbach & Rescorla, 2001)을 표준화한 것이다. 1991년과 2001년 미국판 원검사는 대상 연령대의 범위 및 일부 문항, 그리고 각 요인을 구성하는 해당 문항들에 다소 차이가 있지만, 대부분의 문항은 동일하게 유지되었고, 한국판 검사들은 각각 해당 시기 원검사의 형식을 거의 그대로 유

지하고 있다. 1, 2차 표준화에 따른 한국판 CBCL 검사들, 즉 K-CBCL(오경자 등, 1997)과 한국판 CBCL 6-18(오경자, 김영아, 2011)은 모두 1-112번까지는 아동·청소년의 행동에 대해 기술되어 있는 항목에 응답하는 객관식 문항으로 이루어져 있으며, 추가적으로 보이는 문제행동이 있을 경우 직접 기술하고 그 수준을 평정하는 개방형 문항인 113번 문항이 있다. 신체적 문제를 측정하는 56번 문항이 8개의 하위 문항으로 구성되어 있어 실질적인 문항 수는 총 120개이다. 이들 문항은 불안/우울(우울/불안), 위축/우울(위축), 신체증상, 사회적 미성숙, 사고문제(사고의 문제), 주의집중문제, 규칙위반(비행), 공격행동(공격성)의 8개 증후군 척도와 기타문제 척도를 구성한다.³⁾ 이중 불안/우울, 위축/우울, 신체증상 척도는 내재화 상위척도를 구성하고, 규칙위반과 공격행동은 외현화 상위척도를 구성하며, 8개 증후군 척도와 기타문제는 문제행동총점(총 문제행동)을 이룬다. 한국판 1차 표준화 K-CBCL의 신뢰도(Cronbach's α)는 .62-.86, 한국판 2차 표준화 CBCL 6-18의 신뢰도는 .62-.95이었다.

YSR(Youth Self-Report; 청소년 행동평가척도 자기보고용)

1차 표준화를 통해 출시된 K-YSR 청소년 자기행동 평가척도(오경자 등, 2001)는 1991년 미국판 YSR(Achenbach, 1991b)을 표준화한 것이며, 2차 표준화를 통해 출시되어 현재 사용 중인 YSR 청소년 행동평가척도 자기보고용(오경자, 김영아, 2011)은 2001년 미국판 YSR

(Achenbach & Rescorla, 2001)을 표준화한 것이다. 미국 원검사는 1991년판에서 2001년판으로 개편되는 과정에서 문항 내용 및 각 요인을 구성하는 문항들의 분류가 일부 바뀌었으나 대부분의 문항은 동일하게 유지되었고, 한국판 검사들은 각각 해당 시기 원검사의 형식을 거의 그대로 유지하고 있다. 1, 2차 표준화에 따른 한국판 YSR 검사들, 즉 K-YSR(오경자 등, 2001)과 한국판 YSR(오경자, 김영아, 2011)은 모두 1-112번 문항으로 이루어져 있는데, 부모가 보고하는 CBCL과 달리 아동·청소년 본인이 보고하는 YSR에서는 자신의 행동에 대한 부정적인 점만을 떠올려 보고하게 하는 것이 응답에 부정적인 영향을 줄 것을 고려하여 보완적으로 “나는 꽤 정직하다”, “새로운 것을 해보길 좋아한다”와 같이 긍정적인 행동 문항 14개가 포함되어 있으며, 이 문항들은 문제행동 척도 점수 계산에서 제외된다. 신체적 문제를 측정하는 56번 문항이 8개의 하위 문항으로 구성되어 있어 실질적인 문제행동 문항 수는 총 105개이며, CBCL과 동일한 8개의 증후군 척도와 기타문제 척도, 상위척도인 내재화, 외현화, 문제행동총점(총 문제행동) 척도로 구성되어 있다. 한국판 1차 표준화 K-YSR의 신뢰도(Cronbach's α)는 .63-.85, 한국판 2차 표준화 YSR의 신뢰도는 .59-.93이었다.

분석 방법

본 연구에서는 부모와 아동·청소년 본인의 보고에 따른 문제행동 변화 양상을 함께 검토해보기 위해 CBCL과 YSR의 공통 문항인 105개 문제행동 문항 중에서 1, 2차 표준화 사이에 내용이 바뀐 6개 문항을 제외한 99개 문항을 최종 분석에 포함하였다. 한국판 CBCL 및

3) 본 논문에서 증후군 척도명은 2차 표준화 이후 현재 사용 중인 명칭을 기준으로 서술하고 있으며, 괄호 안은 2차 표준화 이전에 사용되었던 명칭이다.

YSR 2차 표준화의 기반이 되고 현재까지 널리 사용되고 있는 2001년도 미국판 원검사의 요인 구조 중 ASEBA 검사군의 핵심 척도인 문제행동 증후군 척도를 중심으로 분석이 이루어졌다.

먼저 CBCL과 YSR 검사별로 표준화 집단 전체(표 1 참고)에 대해 일련의 확인적 요인분석을 통한 측정 동등성 분석(Vandenberg & Lance, 2000)을 실시하여 1, 2차 표준화 시기 각각의 표준화 집단에서 측정 모형이 동등한 적합성을 지니는지를 확인해 보고자 하였다. 이 분석은 서로 다른 평정자, 평정 시기 등과 같은 다양한 상황에 걸쳐 기본적인 측정 모형에 동등성 제약을 점차 더해가면서 확인적 요인분석의 적합도가 어떻게 변해 가는지를 확인함으로써 측정 동등성이 얼마나 강력하게 지지되는지를 보여준다.

그 다음 1, 2차 표준화 시기에 따른 문제행동 점수 비교에 앞서 여러 연구들에서 사회경제수준 및 부모의 학력 등과 같은 사회·인구학적 변인과 아동·청소년 문제행동의 연관성에 대해 보고하고 있는 것(남재봉, 2009; 장영애, 이영자, 2011; 하은혜, 오경자, 송동호, 2003; Zukauskienė, Ignataviciene, & Daukantaite, 2003)을 고려하여 관련 변수들의 통제가 필요한지를 확인해 보았다. 이에 대한 분석 결과를 토대로 부학력과 모학력을 공변인으로 하고, 성별 및 연령 분포를 동일하게 맞춘 대응 표본(표 2 참고)에 대해 CBCL과 YSR 검사별로 1, 2차 표준화 시기에 따른 문제행동 점수에 대한 차이 검증을 실시하였다. 표준화 시기, 연령, 성별에 따른 아동·청소년 문제행동 점수를 비교해볼 수 있도록 여러 증후군 하위척도들로 구성되는 문제행동총점, 내재화, 외현화 상위척도에 대해서는 각각 별도의

ANCOVA를 실시하고, 8개 증후군 하위척도에 대해서는 다변량 분석의 관점에서 한 번에 MANCOVA를 실시하였다. 마지막으로 CBCL, YSR 각각의 두 표준화 시기에 따른 문항별 점수에 대한 Pearson 상관 분석을 실시하였다.

분석에는 AMOS 7.0 및 SPSS 15.0 for windows 평가판을 사용했다.

결 과

확인적 요인분석을 통한 측정 동등성 (Measurement Equivalence) 분석

측정변수인 문항의 수가 많고 요인 구조가 복잡한 CBCL과 YSR의 경우, 개별 문항들을 측정변인들로 가정한 모형에 대해 확인적 요인분석을 실시하면 모형의 적합도 수준이 좋지 않은 것으로 보고되었다(Achenbach & Rescorla, 2001). 이런 경우, 측정변수의 수를 줄여서 상위 요인들의 구조를 통해 적합도를 확인하는 항목묶음(parceling) 방식을 적용할 수 있다(Little, Cunningham, Shahar, & Widaman, 2002). 따라서 본 연구에서는 현재 사용하고 있는 CBCL과 YSR 측정 구조 모형을 기반으로 그림 1에 제시된 모형과 같이 항목묶음 방식의 요인분석을 실시하였다.

모형 적합도 수준이 강력한 동등성 제약 하에서도 유사한 수준으로 나타날 경우, 다양한 상황에 걸쳐 측정 구조의 적합성이 유지되고 있음을 의미한다. 비교의 대상이 되는 모형 간에 χ^2 의 변화량이 유의미하면 모형 적합도 수준이 현저히 낮아졌다고 판단할 수 있으나 (Vandenberg & Lance, 2000), 표집의 수가 큰 경우 χ^2 값은 작은 자유도 차이의 모형들에 대해

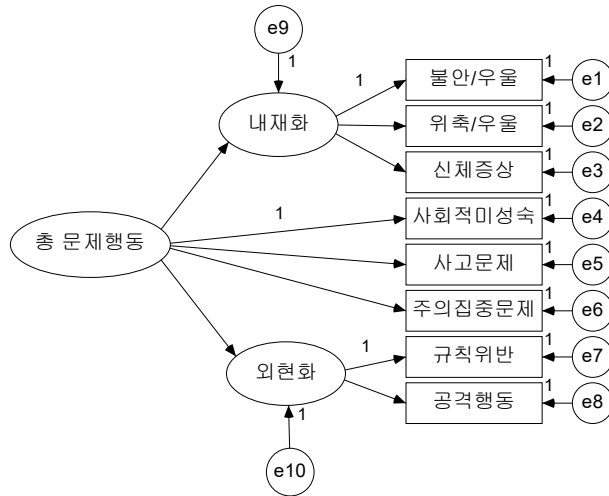


그림 1. 측정 동등성 분석에 사용된 CBCL 및 YSR 구조 모형

서 모형의 실제 적합도 수준과 관계없이 유의미한 차이를 나타낼 가능성이 매우 높다 (Meade, Johnson, & Braddy, 2008). 따라서 본 연구에서는 동등성 제약 전-후에 걸쳐 CFI (Comparative Fit Index)의 변화량(ΔCFI)이 .02보다 작은지(Vandenberg & Lance, 2000), 그리고 RMSEA(Root Mean Square Error of Approximation)의 90% 신뢰구간에 겹치는 부분이 있는지(Wang & Russell, 2005)를 통해 측정 동등성 지지 여부를 평가하였다. 즉, 동등성 제약 이후 CFI가 .02 이상 감소함과 동시에 RMSEA가 동등성 제약 이전 모델에 비해 90% 신뢰구간을 초과하여 증가하는 경우 측정 동등성이 지지되지 않는 것으로 판단하였다(Lee et al., 2014).

총 7개의 모형에 대한 확인적 요인분석을 실시하였다(표 4). 모형 1은 기저 모형으로 어떠한 동등성 제약도 가해지지 않은 상태에서 1, 2차 표준화 자료에 대해 형태 동일성(configural equivalence)을 검증한다. 검사의 원저자가 제시하는 기준인 RMSEA가 CBCL은 .061, YSR은 .066으로 우수한 적합도 수준을 보이는

것으로 나타났으며, CFI도 두 검사 모두 .96 이상으로 우수한 수준을 보였다(양호한 적합도 수준: $RMSEA < .08$, $CFI > .90$; Byrne, 2001; Holve, 1995). 따라서 CBCL, YSR 모두 1, 2차 표준화 자료에서 요인 구조가 동일한 의미를 지니고 있다고 볼 수 있다.

모형 2, 3은 측정 동일성(metric equivalence) 제약 모형으로, 기저 모형에서 확인된 요인 구조 외에 추가적으로 문항 합산 값들로 이루어진 증후군 하위척도(예: 불안/우울)가 1차 요인(예: 내재화)을 통해, 1차 요인이 2차 요인(총 문제행동)을 통해 설명될 수 있는 변량이 두 시기의 표준화 집단에서 동등한지를 검증한다. 즉, 모형 2는 증후군 하위척도와 1차 요인 간의 요인 부하량에 동일성 제약을 추가한 것이고, 모형 3은 모형 2의 제약에 1차 요인과 위계적인 2차 요인 간 요인 부하량에 동일성 제약을 추가한 것이다. 분석 결과, CBCL, YSR 모두 모형 2, 3에서 $\Delta CFI < .005$ 로 모형 적합도 수준에 유의미한 변화가 없었고, RMSEA 90% 신뢰구간에 겹치는 부분이 존재

표 4. CBCL, YSR의 1, 2차 표준화 집단에 대한 측정 동등성 분석 결과

	CBCL				YSR			
	CFI	△CFI	RMSEA	RMSEA 90% 신뢰구간	CFI	△CFI	RMSEA	RMSEA 90% 신뢰구간
모형 1	.967	-	.061	.058-.065	.963	-	.066	.063-.069
모형 2	.963	.004	.061	.057-.064	.960	.003	.064	.061-.067
모형 3	.963	.000	.059	.056-.063	.960	.000	.063	.060-.066
모형 4	.939	.024	.070	.068-.073	.942	.018	.069	.067-.072
모형 5	.936	.003	.070	.068-.073	.942	.000	.068	.065-.070
모형 6	.931	.005	.073	.070-.075	.942	.000	.068	.065-.070
모형 7	.877	.054	.090	.088-.093	.937	.005	.066	.063-.068

주. △CFI: 이전 모형과 해당 모형에서의 CFI 변화량

- 모형 1. 형태 동일성 기저 모형
- 모형 2. 측정 동일성 제약 모형-1차 요인
- 모형 3. 측정 동일성 제약 모형-2차 요인
- 모형 4. 절편 동일성 제약 모형
- 모형 5. 요인 분산 동일성 제약 모형-1차 요인
- 모형 6. 요인 분산 동일성 제약 모형-2차 요인
- 모형 7. 잔차 동일성 제약 모형

하는 것으로 나타나 측정 동등성이 지지되었다. 따라서 CBCL, YSR 모두 증후군 하위척도와 1차 요인, 1차 요인과 2차 요인 간 관련성이 1, 2차 표준화 집단에서 동일한 의미로 해석되고 있다고 볼 수 있다.

모형 4는 측정변인(문항 합)들의 초기치(절편; intercept)에 동일성 제약을 추가함으로써 절편 동일성(scalar equivalence)을 검증한다. CBCL에서는 RMSEA 90% 신뢰구간에 겹치는 부분이 사라졌으며, CFI 변화량(△CFI=.024)을 기준으로 보았을 때에도 모형 적합도 지수가 유의미하게 낮아져 절편 동일성 가정이 기각되었다. 이는 CBCL 1, 2차 표준화 집단에서 하위척도 측정의 초기치에 차이가 있다는 것으로, 두 표준화 시기에 걸쳐 응답 편향이나

응답 역치에 있어 주목할 만한 차이가 관찰되었음을 의미한다. 구체적으로 2차 표준화 집단(측정변인 추정 초기치 $M=1.53$)의 측정 초기치가 1차 표준화 집단(측정변인 추정 초기치 $M=2.61$)에 비해 더 낮았다. 이처럼 RMSEA 신뢰구간 및 CFI 변화량을 기준으로 보았을 때 CBCL의 하위척도 측정 초기치에 차이가 있는 것으로 나타났음에도 불구하고, CBCL의 원저자가 제시하는 모형 적합도 지수인 RMSEA의 절대적인 적합도는 양호한 수준($RMSEA < .08$)이었음을 고려할 필요가 있다. 양호한 수준의 RMSEA는 측정 동등성 제약이 존재하는 상태에서의 한국판 CBCL 측정 모형이 1, 2차 표준화 시기에 걸쳐 적합함을 의미하므로, 큰 틀에서 모형의 적절성을 해치지 않

고 있음을 알려준다. 이를 고려하면 1, 2차 표준화 시기에서 초기치의 차이가 전반적인 CBCL의 측정 모형에 미치는 영향이 우려할만한 수준으로 큰 것은 아닐 수 있으며 두 표준화 시기에 걸쳐 여전히 일정 수준의 측정 동등성이 지지되고 있는 것으로 볼 수 있다. YSR의 경우, RMSEA 90% 신뢰구간의 중복 구간이 사라지기는 했으나, CFI 변화량($\Delta CFI = .018$)을 기준으로 보았을 때에는 1, 2차 표준화 시기에 걸쳐 측정 동등성이 지지되고 있는 것으로 나타나 절편 동일성 가정이 기각되지 않았다.

모형 5, 6은 추가적으로 두 시기의 표준화 집단에서 1차 요인 및 2차 요인의 변량에 차이가 있는지를 확인하는 요인 분산 동일성(equivalence of factor variance) 제약 모형이다. CBCL과 YSR 모두 모형 5, 6에서 측정 동등성이 유지되고 있는 결과를 보였다. 즉, 1차 요인과 2차 요인이 갖고 있는 자체 변량(요인 수준의 정보)이 1, 2차 표준화 집단에서 일관적이라고 볼 수 있다.

마지막 모형 7은 증후군 하위척도들의 오차항이 두 시기의 표준화 집단에서 동일한지를 확인하는 잔차 동일성(equivalence of residual variance) 제약 모형이다. CBCL의 경우, 여전히 CFI와 RMSEA 모두 양호한 수준에 근접한 적합도를 나타내기는 하였으나, CFI의 변화량은 .054였고, RMSEA 90% 신뢰구간에 겹치는 부분이 없어 요인 분산의 잔차에 대한 측정 동등성이 지지되지 않았다. 이는 CBCL에서 상위척도(내재화, 외현화, 총 문제행동)를 통해 설명하고 있는 증후군 하위척도(예: 불안/우울)들의 변량이 1, 2차 표준화 집단에 따라 다를 수 있음을 의미한다. 구체적으로 하위척도의 잔차(상위척도를 통해 설명되지 않는 변량)가 1차 표준화

집단(오차항 추정치 $M=2.77$)보다 2차 표준화 집단(오차항 추정치 $M=1.59$)에서 더 작은 것으로 나타났다. 그러나 YSR은 모형 7에서도 강한 측정 동등성이 지지되어 척도에서 가정한 상위척도를 통해 설명하고 있는 증후군 하위척도들의 변량, 즉 잔차를 제외한 나머지 변량이 1, 2차 표준화 집단에 대해 신뢰로운 것으로 나타났다.

확인적 요인분석을 통한 측정 동등성 분석 결과, CBCL, YSR 모두 1, 2차 표준화 집단에 대해 비교적 강한 수준의 측정 동등성이 지지되었다. 즉, 현재 사용하고 있는 CBCL과 YSR 측정 구조 모형을 두 표준화 시기의 측정 자료에 적용했을 때 주목할 만한 차이가 있다고 보기 어려우며, 문제행동 측정 체계가 비교적 안정적이라고 할 수 있다. 다만, CBCL의 경우 두 표준화 시기에 걸쳐 전반적인 문제행동 평정의 기준치(reference point)가 달라졌고, 상위척도를 바탕으로 한 하위척도 설명 변량 측면에서의 신뢰도에 변화가 있었다는 점이 관찰되었다.

1, 2차 표준화 시기에 따른 문제행동 점수 변화

1, 2차 표준화 시기에 따른 문제행동 점수 비교에 앞서 사회·인구학적 변인들을 통제할 필요가 있는지를 확인해 보았다. 두 표본 모두 경제수준과 부모 학력 정보를 포함하고 있으나, 경제수준 관련 정보는 구체적인 기준에 의해 측정된 것이 아니라 부모가 주관적으로 지각한 경제수준이므로 객관적인 통제를 하는데 제한이 있었다. 대응 표본의 부모 학력 분포를 비교해본 결과에서는 CBCL, YSR 모두 표준화 시기 사이에 부학력 및 모학력의

중졸 이하 비율이 감소하고, 대졸 이상 고학력이 증가한 것으로 나타났다($\chi^2_6 > 129.6$, $df=2$, $p < 0.001$). CBCL의 경우, 1, 2차 표준화 시기 대응 표본 전체에 대해 부학력 및 모학력이 높을수록 문제행동 수준이 전반적으로 낮은 것으로 나타났다. 부학력이 대졸 이상인 경우 문제행동총점은 $M=15.99(SD=15.29)$, 중졸 이하인 경우 $M=23.71(SD=19.00)$ 이었고, 모학력이 대졸 이상인 경우 $M=16.25(SD=15.10)$, 중졸 이하인 경우 $M=23.37(SD=17.70)$ 이었다. 이처럼 부모의 학력이 낮을수록 보고된 문제행

동 점수가 높게 나타나는 경향은 2005년 수집된 영유아 집단에 대한 한국판 CBCL 1.5-5의 자료에서도 관찰되었다(김혜진, 김영아, 이진, 오경자, 2009). YSR의 경우, CBCL만큼 뚜렷한 경향성이나 현격한 차이를 보이지는 않았으나, 전반적으로 부학력 및 모학력이 높을수록 문제행동 수준이 낮은 경향은 동일했다. YSR의 1, 2차 표준화 시기 대응 표본 전체에 대해 부학력이 대졸 이상인 경우 문제행동총점은 $M=38.41(SD=20.80)$, 중졸 이하인 경우 $M=39.81(SD=21.69)$ 이었고, 모학력이 대졸 이상인

표 5. CBCL 척도 점수의 추정 평균치와 설명 변량

	평균 (표준오차)		설명 변량 (% of explained variance)				
	1차 표준화	2차 표준화	표준화 시기	연령	성별	표준화 시기 × 연령	표준화 시기 × 성별
<u>ANCOVA</u>							
문제행동총점	22.28 (.75)	15.74 (.81)	3.09**	.01	.26	.01	.08
내재화	7.68 (.28)	4.94 (.30)	3.93**	.10	.08	.00	.00
외현화	5.50 (.23)	3.88 (.25)	1.99**	.00	.72**	.01	.21
<u>MANCOVA</u>							
불안/우울	3.22 (.14)	2.45 (.15)	1.32**	.05	.11	.05	.02
위축/우울	1.94 (.09)	1.18 (.10)	3.00**	.09	.42*	.11	.01
신체증상	2.53 (.11)	1.30 (.12)	5.16**	.08	.69**	.00	.00
사회적 미성숙	2.04 (.10)	1.61 (.11)	.80**	.63**	.02	.07	.15
사고문제	1.63 (.09)	1.17 (.09)	1.17**	.11	.98**	.00	.25
주의집중문제	2.89 (.11)	2.40 (.12)	.78**	.12	1.96**	.08	.00
규칙위반	1.23 (.07)	.85 (.08)	1.11**	.02	2.37**	.08	.43*
공격행동	4.26 (.18)	3.03 (.19)	1.96**	.01	.21	.00	.11

주. 설명 변량(% of explained variance): "(partial η^2) × 100"의 값으로, 해당 변수(예: 표준화 시기)에 의해 각 척도가 설명되는 전체 분산의 비율을 의미함(Dattalo, 2008). 표준화 시기의 경우, '1차 표준화'=1, '2차 표준화'=2로 코딩하였으며, 연령은 '12-14세'=1, '15-17세'=2, 성별은 '남'=0, '여'=1로 코딩하여 분석하였음. * $p < .05$, ** $p < .01$.

경우 $M=37.06(SD=20.85)$, 중졸 이하인 경우 $M=39.55(SD=21.53)$ 이었다. 따라서 부학력 및 모학력이 아동·청소년의 문제행동 점수에 미쳤을 가능성을 통제 한 후, 1, 2차 표준화 시기에 따른 문제행동 점수 변화를 비교하였다.

분석 결과, CBCL은 모든 척도에서 1차 표준화 시기보다 2차 표준화 시기에 점수가 유의미하게 감소한 것으로 나타났다(표 5). 규칙위반 척도 점수는 1, 2차 표준화 시기에 걸쳐 남자의 점수 변화폭이 여자에 비해 큰 것으로 나타나 표준화 시기와 성별의 유의미한 상호

작용이 나타나기는 했지만, 두 시기 모두 남자(1차/2차 $M=1.58/.99$)가 여자(1차/2차 $M=.89/.70$)보다 높은 점수를 보이고, 남녀 모두 1차보다 2차에 점수가 낮아진 경향은 동일하였다. 이를 제외하면 나머지 모든 척도에서 연령과 성별은 표준화 시기와 유의미한 상호작용을 보이지 않았다. 한편, 1, 2차 표준화 시기에 걸쳐 연령 및 성별에서 차이를 보이는 척도들도 있었다. 예를 들어, 사회적 미성숙은 연령이 증가함에 따라 점수가 유의미하게 감소했다(12-14세: $M=1.99(SE=.10)$, 15-17세: $M=$

표 6. YSR 척도 점수의 추정 평균치와 설명 변량

	평균 (표준오차)		설명 변량 (% of explained variance)				
	1차 표준화	2차 표준화	표준화 시기	연령	성별	표준화 시기 × 연령	표준화 시기 × 성별
<u>ANCOVA</u>							
문제행동총점	38.46 (.98)	37.96 (.99)	.01	.34*	.19	.02	.00
내재화	12.79 (.38)	11.80 (.38)	.32	.34*	2.03**	.01	.07
외현화	9.97 (.29)	10.51 (.30)	.15	.01	.74**	.00	.01
<u>MANCOVA</u>							
불안/우울	5.79 (.20)	5.95 (.20)	.03	.32	1.23**	.04	.01
위축/우울	3.41 (.12)	2.88 (.12)	.92**	.24	.66**	.06	.11
신체증상	3.60 (.13)	2.97 (.13)	1.03**	.14	2.80**	.03	.12
사회적 미성숙	3.83 (.14)	3.85 (.14)	.00	.21	.04	.02	.13
사고문제	3.37 (.14)	3.05 (.14)	.24	.39*	.01	.02	.18
주의집중문제	4.51 (.12)	4.70 (.12)	.12	1.16**	.04	.17	.07
규칙위반	2.74 (.10)	2.47 (.10)	.33	.21	3.86**	.04	.22
공격행동	7.23 (.22)	8.05 (.22)	.63**	.01	.05	.03	.01

주. 설명 변량(% of explained variance): "(partial η^2) × 100"의 값으로, 해당 변수(예: 표준화 시기)에 의해 각 척도가 설명되는 전체 분산의 비율을 의미함(Dattalo, 2008). 표준화 시기의 경우, '1차 표준화'=1, '2차 표준화'=2로 코딩하였으며, 연령은 '12-14세'=1, '15-17세'=2, 성별은 '남'=0, '여'=1로 코딩하여 분석하였음. * $p<.05$, ** $p<.01$.

1.66($SE=.10$)). 성차가 나타난 척도에서는 신체 증상(남자: $M=1.73(SE=0.11)$, 여자: $M=2.10(SE=.10)$)을 제외하면, 남자가 여자보다 높은 점수를 보였다: 외현화(남자: $M=5.10(SE=.23)$, 여자: $M=4.27(SE=.22)$), 위축/우울(남자: $M=1.68(SE=.09)$, 여자: $M=1.44(SE=.08)$), 사고문제(남자: $M=1.58(SE=.09)$, 여자: $M=1.22(SE=.08)$), 주의집중문제(남자: $M=2.98(SE=.11)$, 여자: $M=2.31(SE=.11)$), 규칙위반(남자: $M=1.28(SE=.07)$, 여자: $M=.79(SE=.07)$).

이에 비해 YSR 상위척도들은 모두 표준화 시기에 따른 차이가 유의미하지 않았다(표 6). 증후군 하위척도 중에서는 내재화 영역에 포함되는 위축/우울과 신체증상 척도 점수가 1, 2차 표준화 사이에 유의미한 감소를 보인 반면, 외현화 영역에 포함되는 공격행동 척도 점수는 유의미하게 증가한 것으로 나타났다. 모든 척도에서 연령과 성별은 표준화 시기와 유의미한 상호작용을 보이지 않았으나, 1, 2차 표준화 시기에 걸쳐 연령 차이가 나타난 척도들에서는 연령이 증가할수록 높은 점수를 보고하는 것으로 나타났다: 문제행동총점(12-14세: $M=37.02(SE=.96)$, 15-17세: $M=39.40(SE=.92)$), 내재화(12-14세: $M=11.84(SE=.37)$, 15-17세: $M=12.76(SE=.36)$), 사고문제(12-14세: $M=3.03(SE=.13)$, 15-17세: $M=3.39(SE=.13)$), 주의집중문제(12-14세: $M=4.33(SE=.12)$, 15-17세: $M=4.88(SE=.11)$). 또한 1, 2차 표준화 시기에 걸쳐 내재화 관련 영역에서는 여자가 남자보다 높은 점수를 보였다: 내재화(남자: $M=11.17(SE=.36)$, 여자: $M=13.43(SE=.36)$), 불안/우울(남자: $M=5.41(SE=.19)$, 여자: $M=6.33(SE=.19)$), 위축/우울(남자: $M=2.95(SE=.11)$, 여자: $M=3.34(SE=.11)$), 신체증상(남자: $M=2.81(SE=.13)$, 여자: $M=3.75(SE=.13)$). 반면, 1, 2차 표준

화 시기에 걸쳐 외현화 관련 영역에서는 남자가 여자보다 높은 점수를 보였다: 외현화(남자: $M=10.77(SE=.28)$, 여자: $M=9.71(SE=.29)$), 규칙위반(남자: $M=3.04(SE=.10)$, 여자: $M=2.17(SE=.10)$).

1, 2차 표준화 시기에 걸친 문항별 점수의 상관관계(stability of item rank)

Achenbach 등(2003)이 제시한 방식을 따라 각 표준화 시기별로 대응 표본 전체 자료에 대해 각 문항의 평균을 계산한 후, 두 시기의 문항 평균들의 상관계수를 산출하였다. 분석 결과, CBCL에서 $r=.90(p<.01)$, YSR에서 $r=.92(p<.01)$ 로 두 시기의 상관계수가 매우 높은 것으로 나타났는데, 이는 해당 기간에 걸친 문제행동 항목들 간의 상대적 발현 빈도가 매우 유사한 수준임을 의미한다.

논 의

본 연구에서는 대표적인 아동·청소년 행동 평가척도인 CBCL과 YSR을 통해 한국 아동·청소년의 문제행동 변화 양상을 확인해 보았다. 서로 다른 시기의 자료를 비교하기 위해서는 문제행동에 대한 평가 태도 및 경향, 그리고 문항 및 척도의 의미 변화가 있었는지를 확인할 필요가 있다. 특히 전국적 규모로 이루어진 두 번의 표준화 시기 사이에 CBCL, YSR이 속한 ASEBA 검사군의 개편이 이루어졌기 때문에 먼저 측정 동등성 분석을 통해 두 시기의 검사를 비교할 수 있는지 확인한 후, 1, 2차 표준화 시기에 따른 점수 변화를 살펴 보았다. 본 연구의 결과 및 논의점은 다음과

같다.

첫째, 확인적 요인분석을 통한 측정 동등성 분석 결과, CBCL, YSR 모두 1, 2차 표준화 집단에 대해 비교적 강한 수준의 측정 동등성이 지지되었다. 이는 심리측정적 관점에서 두 번의 검사 표준화가 이루어진 10년 이상의 기간에 걸쳐 CBCL과 YSR의 문제행동 측정 체계가 안정적임을 의미하며, 1, 2차 표준화 시기에 CBCL 및 YSR을 통해 측정된 문제행동 점수를 비교할 수 있음을 시사한다. 총 7개 단계의 측정 동등성 검증에서 YSR은 모든 동등성 제약 모형에 걸쳐 강한 측정 동등성이 지지되었으나, CBCL은 하위척도 측정의 초기치에서 1, 2차 표준화 집단에 차이가 있는 것으로 나타났다. 즉, 부모보고의 경우 두 시기 중 한 집단에서 점수를 높거나 낮게 보고하는 경향이 있을 수 있는데 실제로 CBCL의 점수를 비교해 본 결과, 1차 표준화 때보다 2차 표준화에서 부모들이 자녀의 문제행동 평가 시 동일한 문제행동에 보다 낮은 점수를 부여하는 엄격한(strict) 평가 기준을 사용한 것으로 나타났다. 본 연구 결과만으로 이와 같은 결과가 나타난 이유를 명확히 알 수는 없으나, 교육 현장에서 아동·청소년의 문제행동을 정신건강 관리의 관점에서 체계적으로 접근하려는 시도들이 이루어진 것과 관련이 있을 수 있다(이진석, 2011). 이러한 사회적 변화가 부모의 자녀 문제행동에 대한 관심을 증대시켰고, 이에 따라 아동·청소년의 다양한 문제행동 양상 및 수준에 대한 이해도가 전반적으로 높아지면서 자녀의 문제행동을 평가할 때 명확히 해당 사항이 있는 항목에 대해서만 실제로 증상이 있다고 보고하는 경향이 있었을 가능성이 있다(Loeber, Green, & Lahey, 1990). 더불어 과거에 비해 교육열이 높아지고(홍봉선, 2004) 생활기

록부와 같은 새로운 평가 방식의 도입에 따라(강정진, 1997) 자녀의 학교생활 및 검사/시험 등에 민감해진 부모들이 보다 방어적으로 반응했기 때문일 수도 있을 것이다.

또한 CBCL에서는 1, 2차 표준화 시기에 걸쳐 상위척도를 통해 설명될 수 있는 하위척도의 변량에 차이가 있는 것으로 나타났는데, 2차 표준화 자료에서의 잔차(상위척도를 통해 설명되지 않은 변량)가 1차 표준화 자료에 비해 더 작았다. 이는 2차 표준화에서 1차 표준화 때보다 CBCL의 상위척도가 하위척도들을 효과적으로 설명하고 있음을 의미한다.

둘째, 1, 2차 표준화 시기에 걸쳐 한국 아동·청소년 문제행동 수준이 부모보고(CBCL)에서는 감소한 것으로 나타났으나, 아동·청소년 본인보고(YSR)에서는 큰 변화가 없었다. 이는 측정 동등성 분석에서 CBCL의 경우 초기치에 차이가 있었고, YSR에서는 차이가 없었던 것과 일맥상통하는 결과이다.

부모보고와 자기보고에서 1, 2차 표준화 시기에 걸쳐 문제행동 수준의 변화 양상이 다르게 나타난 것은 사람들마다 문제행동 수준에 대한 내적 기준이 다르며, 그 기준은 평정자 집단 및 평가 시기에 따라 다른 양상으로 변화할 수 있음을 의미한다. 본 연구 결과에서 두 번의 표준화 시기에 걸친 문제행동 판단 기준의 변화는 부모에게서만 관찰된 독특한 현상으로 보인다. 따라서 서로 다른 시기에 CBCL을 통해 측정된 점수의 수준, 즉 평균값 자체를 비교하여 문제행동 수준의 변화를 살펴볼 때에는 실제로 아동·청소년 문제행동 수준이 달라졌을 가능성뿐만 아니라 CBCL을 작성하는 부모의 평가 태도 및 방식에 변화가 있었는지를 함께 고려하여 해석해야 할 것이다. 현재 사용 중인 한국판 CBCL, YSR에는 2

차 표준화 당시 수집한 전국 표집을 바탕으로 한 새로운 기준 점수가 적용되고 있는데, CBCL의 경우 이번 연구에서 관찰된 것과 같은 부모의 문제행동 평가 경향의 변화를 적절히 반영할 수 있는 바람직한 개정 절차였던 것으로 판단된다. YSR에서는 이와 같은 변화가 관찰되지는 않았지만, 사회·문화적 맥락에 따른 평가 경향성에 변화가 있을 가능성을 고려하여 정기적인 개정 절차를 거칠 필요성이 시사된다.

또한 부모보고에서만 점수가 변화한 것은 앞서 서술한 바와 같이 부모가 방어적으로 반응했을 가능성을 포함하여, 아동·청소년 문제행동에 대한 부모의 반응 양식이 아동·청소년 본인의 반응 양식보다 사회·문화적 변화의 영향을 더 많이 받았을 가능성이 있다. 예를 들어, 성격의 보고 방식에 대한 Connelly와 Ones(2010)의 메타 분석 연구에서 제시한 모델에 따르면, 타인보고의 경우 자기보고와는 달리 평가하고자 하는 특성이 발현되는 상황 단서를 활용해야하는 추가적인 인지 과정을 거치게 된다. 따라서 부모보고와 같은 타인보고의 경우 사회·문화적 영향에 더 민감할 가능성이 있다.

한편, 두 시기에 걸쳐 YSR에서는 점수가 유사한 수준이었던 반면 CBCL에서는 점수가 전반적으로 낮아진 경향에 따라 1차 표준화 시기에 비해 2차 표준화 시기에서 CBCL과 YSR의 점수, 즉 부모와 아동·청소년이 보고한 문제행동 발현율의 차이가 더 벌어진 것으로 볼 수 있기도 하다.

본 연구는 문제행동 수준의 변화 양상에 초점을 맞추었기 때문에 각 표준화 시기에서의 CBCL과 YSR의 점수 차이에 대한 구체적인 분석을 실시하지는 않았으나, 1, 2차 표준화 시

기 모두에서 YSR 점수가 CBCL 점수보다 높은 것을 확인할 수 있었다(표 5, 6 참고; 문제행동 총점 M : 1차 표준화 $CBCL/YSR=22.28/38.46$, 2차 표준화 $CBCL/YSR=15.74/37.96$). 많은 연구들에서 아동·청소년 행동 및 정서문제 평가에 있어 부모-자녀 간 일치도가 높지 않음을 보고하고 있는데(강현정, 국승희, 2010; Goodman, 2001; Kemper, Gerhardstein, Repper, & Kistner, 2003), 본 연구 결과는 한 시점에서의 평정자 간 점수 차이뿐만 아니라 사회·문화적 맥락에 따른 각 평정자 집단의 평가 경향성의 변화를 함께 살펴볼 필요가 있음을 시사한다는 점에서 의의가 있다. 현재 사용 중인 한국판 CBCL과 YSR을 표준화할 때 교사가 평가하는 TRF(Teacher's Report Form; 아동·청소년 행동평가척도 교사용; 오경자, 김영아, 2011)도 함께 표준화되었으므로, 추후 연구에서는 교사의 평가 결과를 함께 비교 분석해봄으로써 보다 더 다각적인 평가가 가능할 것으로 기대된다.

또한 전반적으로 문제행동 측정 시기(표준화 시기)는 연령 및 성별과 유의미한 상호작용을 보이지 않았다. 즉, 1, 2차 표준화 시기에 걸쳐 부모보고(CBCL)에서 점수가 감소한 경향을 보인 것과 자기보고(YSR)에서 유사한 점수를 보인 경향이 연령과 성별에 따라 다르지 않았다. 그러나 본 연구에서는 부모의 학력 수준을 통제된 상태에서 비교하였으며, 사회·문화적 맥락의 변화에 따라 다른 효과를 보일 가능성을 배제할 수 없다. 따라서 추후 연구에서는 아동·청소년의 문제행동 수준이 다양한 사회·문화적 요인에 따라 어떻게 변화하는지를 함께 확인해 보아야 할 것이다.

서구 문화권에서 이루어진 시간에 따른 문제행동의 변화 양상에 대한 연구들과는 시기

및 기간, 대상 연령이 달라서 직접적인 비교가 어렵기는 하나, 아동·청소년의 문제 행동이 시간에 따라 전반적으로 감소하거나 (Achenbach, Dumenci, & Rescorla, 2002), 비슷한 수준을 유지하는(Sourander et al., 2004; Verhulst et al., 1997) 것은 본 연구의 결과와 유사한 결과이기도 하다. 그러나 한국 아동·청소년의 자기보고(YSR)에서 내재화 관련 하위척도인 위축/우울, 신체증상 영역은 유의미한 감소를, 외현화 관련 하위척도인 공격행동은 유의미한 증가를 보인 것은 내재화 영역에서 문제행동이 증가하고(Eimecke et al., 2011; Tick et al., 2007), 외현화 영역은 큰 변화가 없었던(Eimecke et al., 2011) 기존 연구들과 차이가 있다.

한편, 아동·청소년기에는 신체 생리적으로 급격한 발달이 이루어지고 자아의식이 높아지게 되는데, 발달심리학에서는 이 시기를 전환기 또는 과도기라고 지칭한다. Shapiro, Siegel, Scovill과 Hays(1998)는 이 시기에 새로운 자아와 환경, 사회적 기대에 대한 적응 과정에서 수반되는 정서적 갈등과 스트레스가 다양한 문제행동을 일으킬 수 있다고 보았으며, Allen, Moore와 Kupermine(1997)은 12-17세에 걸쳐 문제행동이 급격히 증가하는 경향이 있음을 밝혔다. 본 연구에서도 CBCL 사회적 미성숙에서 연령이 증가함에 따라 점수가 유의미하게 감소한 것을 제외하면, CBCL, YSR 모두 전반적으로 연령이 증가함에 따라 문제행동이 증가하는 경향을 보였다.

또한 성별에 따라 정서적 민감성 및 대처 방식에 차이가 있고(함진선, 이상한, 2010), 사회 및 부모로부터 기대되는 역할이 다르며, 남자와 여자가 서로 다른 사회화 과정을 거치게 됨에 따라(김지현, 권연희, 2010) 구체적으

로 드러나는 문제의 양상에 남녀 간 차이가 있을 수 있다(구미향, 이양희, 2003; 김신애, 이형실, 임수경, 2008; 이현지, 하은혜, 오경자, 2005; Rescorla et al., 2007). 본 연구 결과, 내재화 관련 영역에서 CBCL은 성차가 유의미하지 않거나 일관적이지 않은 양상을 보였으나, YSR에서는 전반적으로 여자가 남자보다 높은 점수를 보였다. 한편, 외현화 관련 영역은 CBCL과 YSR 모두 남자가 여자보다 높은 점수를 보이는 경향이 있었다. 불안, 우울과 같은 내재화 문제는 아동·청소년기에 본격적으로 발달하게 되는데, 규칙위반, 공격행동 등과 같은 외현화 문제에 비해 겉으로 드러나는 행동으로 타인이 관찰하기가 상대적으로 어렵다. 즉, 내재화 문제는 본인이 직접 보고하지 않는 한 타인이 파악하기가 쉽지 않다. 게다가 아동·청소년기는 또래와의 상호작용이 중요해지는 시기이기 때문에 본인이 겪는 어려움을 호소하는 대상이 반드시 부모는 아닐 수 있다. 이처럼 문제 영역 및 청소년기의 발달적 특성으로 인해 문제 영역에 따른 성차가 평정자에 따라 다르게 나타났을 가능성이 있다. Achenbach와 Rescorla(2007)는 이와 같이 문제 영역에 따라 평정자 간 보고가 일치하지 않을 경우, 내재화 문제에 대해서는 자기보고, 외현화 문제에 대해서는 부모보고에 비중을 두고 이해하는 것이 효과적이라고 제안하였다 (p. 255).

셋째, 1, 2차 표준화 시기에 걸친 문항별 점수의 상관계수는 CBCL, YSR 모두 매우 높은 수준으로 나타났다($r > .90$). 이는 다른 문화권에서 이루어진 문제행동의 변화 양상에 대한 연구 결과와 유사한 수준(Achenbach et al., 2003: $r = .94-.98$; Eimecke et al., 2011: $r = .90$)으로, 두 번의 표준화가 이루어진 10년 이상의

기간에 걸쳐 한국 아동·청소년의 문제행동 항목들 간의 상대적 발현 빈도에 큰 변화가 없음을 의미한다. 특히, 부모보고에서 문제행동 수준이 크게 감소한 것으로 보고되었음에도 불구하고, 한국 아동·청소년이 어려움을 겪고 있는 문제행동 영역은 크게 달라지지 않았음을 보여준다. 추후 연구에서 개별 문항 각각에 대한 검토를 통해 한국 아동·청소년이 특히 어려움을 겪는 부분을 파악한다면, 아동·청소년의 정신건강 증진을 위한 대책 마련에 있어 보다 구체적이고 실질적인 도움이 될 수 있을 것이다.

본 연구는 문제행동이 시간에 따라 어떻게 변화해 가는지를 확인해 보는 과정에서 평가 시기에 따른 평가 도구 및 자료 수집 방식의 차이, 진단 기준의 변화 등과 같은 방법론적 한계를 최소화하기 위해 서로 다른 시기에 실시한 평가 자료의 측정 동등성을 확인하고, 아동·청소년의 문제행동 점수 변화 양상을 부모와 아동·청소년 본인 평가를 통해 다각적으로 확인해 보았다는 점에서 의의가 있다.

연구 결과, 한국 아동·청소년의 문제행동이 증가하고 있다는 일반적인 우려와는 달리, 전반적인 문제행동 수준이 증가한 경향은 나타나지 않았다. 실제로 한국 아동·청소년의 문제행동 수준이 악화되지 않고 비슷한 수준을 유지하거나 개선되었을 수 있을 것이다. 그러나 CBCL과 YSR이 우울, 주의집중문제, 공격행동 등 문제행동 영역 전반을 평가하고 있기는 하지만, CBCL과 YSR에서 평가되지 않는 영역의 문제가 증가했을 가능성도 배제할 수는 없다. 따라서 일반 대중이 느끼는 아동·청소년의 구체적인 문제행동과 CBCL, YSR 문항 및 요인들과의 관계를 체계적으로 검증하는 연구도 필요할 것이다.

또한 본 연구 결과에서 부모와 아동·청소년 본인, 즉 평정자에 따라 문제행동 변화 양상이 다르게 나타난 것은 사회·문화적 변화에 따른 부모의 축소보고로 인한 것일 수도 있고, 아동·청소년 본인이 지각하고 보고하는 문제행동 수준이 일관적으로 과장되어 있을 가능성도 완전히 배제하기는 어렵다. 게다가 문제 영역의 특성에 따라 어떤 평정자는 다른 평정자보다 평가 대상자의 상황 특이적 (situation specific)인 행동에 민감할 수 있고(Tett & Guterman, 2000), 개개인마다 문제행동을 인식하고 기억하는 것과 문제행동의 수준에 대한 역치(내적 기준)가 다르기 때문에 평정자 간 차이가 나타날 수도 있다(Achenbach & Rescorla, 2007, p. 29). 따라서 정신병리학 연구에서 한 명 이상의 다양한 평정자의 평가 자료를 함께 통합, 해석하는 것이 중요함을 보여준다. 그리고 평정자 간 일치도를 살펴보는 연구에서 단순히 한 시점에서의 점수 차이만 비교하는 방식에서 벗어나 여러 시점에서의 점수를 함께 살펴보고, 각 시점에서의 맥락과 이에 영향을 줄 수 있는 사회·문화적 요인들의 차이를 고려할 필요가 있음을 알려준다. 향후 보다 다양한 평정자 집단의 평가 자료를 활용하여, 보다 넓은 연령대를 포함한 장기적이고 지속적인 연구 분석이 이루어진다면 아동·청소년의 문제행동 변화 양상을 체계적으로 이해하는 데 도움이 될 수 있을 것이다. 나아가 사회·문화적 환경에 적합한 치료 계획 및 정책 등을 수립하고, 새로운 기준 점수를 적용할 시점을 파악하는 데 유용한 정보를 제공할 수 있을 것으로 보인다. 추후 표준화 자료와 같은 일반 집단뿐만 아니라, 임상 집단 자료에 대해서도 시간에 따른 문제행동 수준의 변화 양상에 대한 연구가 이루어진다면,

척도의 임상적 효용성을 높이고 다양한 영역에서 어려움을 겪고 있는 아동·청소년에게 실질적인 도움을 줄 수 있을 것으로 기대된다.

참고문헌

- 강은정 (2007). 한국 아동 정신건강 현황과 정책과제. 보건복지포럼, 60-72.
- 강정진 (1997). 포트폴리오 평가의 실과 적용 방안. 한국실과교육학회 1997년도 학술발표대회, 91-113.
- 강현정, 국승희 (2010). 내재화 장애 청소년과 외현화 장애 청소년의 문제행동에 대한 부모 보고와 청소년 보고의 양상. 한국심리학회지: 임상, 29(4), 907-925.
- 교육부 (1994). 교육연감.
- 구미향, 이양희 (2003). 아동기 두려움과 행동 문제의 특성 및 관련성 탐색. 초등교육연구, 16(1), 341-363.
- 김신애, 이형실, 임수경 (2008). 남녀 청소년의 가족체계유형과 학교생활적응 및 행동문제의 관련성. 한국가정과교육학회지, 20(4), 1-17.
- 김용석, 모선영, 최현정, 이자영 (2008). 아동, 청소년의 정서행동문제 현황과 지원방안에 관한 연구. 아동복지 연구소 보고서, 1 (단일호), 1-128.
- 김재철, 최지영 (2010). 청소년의 외현적 공격성 및 내면화된 문제행동의 개인별 변화에 대한 친한 친구 관련 변수의 영향. 한국청소년연구, 21(1), 123-148.
- 김지현, 권연희 (2010). 어머니 양육행동, 아동의 정서조절 및 스트레스 대처행동이 남아와 여아의 행동문제에 미치는 영향. 대한가정학회지, 48(3), 1-13.
- 김혜진, 김영아, 이진, 오경자 (2009). 사회적 구학적 변인에 따른 한국 유아의 문제행동 특성. 한국심리학회지: 발달, 22(3), 125-144.
- 남재봉 (2009). 사회경제적 지위, 학업성적과 청소년 비행. 사회과학연구, 26(2), 3-33.
- 서미정 (2009). 초기 청소년의 외현적 공격성 변화와 비행, 우울/불안 및 학업성취감: 잠재성장분석. 한국청소년연구, 20(2), 141-167.
- 오경자, 김영아 (2011). ASEBA 아동·청소년 행동평가척도 매뉴얼. 서울:(주)휴노.
- 오경자, 이혜련, 홍강의, 하은혜 (1997). K-CBCL 아동·청소년 행동평가검사. 서울: 중앙적성 출판사.
- 오경자, 하은혜, 이혜련, 홍강의 (2001). K-YSR 청소년 자기행동 평가척도. 서울: 중앙적성출판사.
- 이진석 (2011). 학생정신건강관리 체계운영실태분석을 통한 문제점 확인 및 개선방안연구. 교육부 연구 보고서.
- 이현지, 하은혜, 오경자 (2005). 청소년기 정서 및 행동문제의 성차. 소아청소년정신의학, 16(1), 117-123.
- 장영애, 이영자 (2011). 아동이 지각한 부부갈등 및 부모-자녀간 의사소통이 아동의 스트레스와 문제행동에 미치는 영향. 한국가족치료학회지, 19(3), 183-205.
- 조정아 (2009). 선행모형을 적용한 청소년의 우울 변화에 관한 종단연구: 변화경향과 개인차에 대한 성별·부모·또래·교사 요인 검증. 한국청소년연구, 20(3), 167-192.
- 좌현숙, 오승환 (2011) 내재화문제의 종단적 변화와 예측요인에 대한 성별차이 분석.

- 청소년복지연구, 13(4), 119-147.
- 질병관리본부 (2014). 제10차(2014년) 청소년건강행태온라인조사 통계. 질병관리본부.
- 통계청 (2006). 2005 대한민국 인구 및 주택총조사.
- 하은혜, 오경자, 송동호 (2003). 청소년기 우울 및 불안증상 공존집단의 심리사회적 위험 요인. *한국심리학회지: 일반*, 22(2), 127-144.
- 함진선, 이장한 (2010). 성별에 따라 성격차원과 사랑유형이 성적 부정의 질투에 미치는 영향. *한국심리학회지: 사회 및 성격*, 24(3), 109-123.
- 홍봉선 (2004). 우리나라 교육복지의 방향과 과제. *한국사회복지학*, 56(1), 253-282.
- Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, Vt: University of Vermont.
- Achenbach, T. M. (1991b). *Manual for the Youth Self-Report and 1991 Profile*. Burlington, Vt: University of Vermont.
- Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002). Ten-year comparisons of problems and competencies for national samples of youth self, parent, and teacher reports. *Journal of Emotional and Behavioral Disorders*, 10(4), 194-203.
- Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2003). Are American children's problems still getting worse? A 23-year comparison. *Journal of Abnormal Child Psychology*, 31(1), 1-11.
- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist/4-18 and Revised Child Behavior Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., & Howell, C. T. (1993). Are American children's problems getting worse? A 13-year comparison. *Journal of the American Academy of Child & Adolescent Psychiatry*, 32(6), 1145-1154.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implication of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.
- 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.
- Achenbach, T. M., & Rescorla, L. A. (2007). *Multicultural Understanding of Child and Adolescent Psychopathology*. New York: The Guilford Press.
- Allen, J. P., Moore, C. M., & Kupermine, G. P. (1997). Developmental approaches to understanding adolescent deviance. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. Weisz (Eds.), *Developmental Psychopathology: Perspectives on Risk and Disorder* (pp. 548-567). Cambridge, England: Cambridge University Press.
- American Psychiatric Association (1952, 1968, 1980, 1987, 1994, 2000, 2013). *Diagnostic and Statistical Manual of Mental Disorders* (1st ed., 2nd ed., 3rd ed., 3rd ed., rev., 4th ed.; 4th ed., text rev., 5th ed.). Washington, DC: Author.
- Au, R., Seshadri, S., Wolf, P. A., Elias, M. F.,

- Elias, P. K., Sullivan, L., Beiser, A., & D'Agostino, R. B. (2004). New norms for a new generation: cognitive performance in the Framingham offspring cohort. *Experimental Aging Research, 30*(4), 333-358.
- Bhatia, S. C., Khan, M. H., Mediratta, R. P., & Sharma, A. (1987). High risk suicide factors across cultures. *International Journal of Social Psychiatry, 33*(3), 226-236.
- Byrne, B. M. (2001). *Structural equation modelling with AMOS: Basic Concepts, Applications and Programming*. Mahwah, NJ: Erlbaum.
- Collishaw, S., Maughan, B., Goodman, R., & Pickles, A. (2004). Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry, 45*(8), 1350-1362.
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*(6), 1092-1122.
- Dattalo, P. (2008). *Determining Sample Size: Balancing Power, Precision, and Practicality*. New York: Oxford University Press.
- de Beurs, D. P., Fokkema, M., de Groot, M. H., de Keijser, J., & Kerkhof, A. J. (2014). Longitudinal measurement invariance of the Beck Scale for Suicide Ideation. *Psychiatry Research, 225*(3), 368-373.
- Eimecke, S., Pauschardt, J., Remschmidt, H., Walter, R., & Mattejat, F. (2011). Time Trends in Psychopathology. *Zeitschrift für Kinder-und Jugendpsychiatrie und Psychotherapie, 39*(3), 187-195.
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1337-1345.
- 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.
- 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.
- Lee, J., Huang, Y. H., Robertson, M. M., Murphy, L. A., Garabet, A., & Chang, W. R. (2014). External validity of a generic safety climate scale for lone workers across different industries and companies. *Accident Analysis & Prevention, 63*, 138-145.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151-173.
- Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: a review of the past 10 years, part I. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(12), 1468-1484.
- Loeber, R., Green, S. M., & Lahey, B. B. (1990). Mental health professionals' perception of the utility of children, mothers, and teachers as

- informants on childhood psychopathology. *Journal of Clinical Child Psychology, 19*(2), 136-143.
- Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568-592.
- Padmanabhan, V. T., Sugunan, A. P., Brahmaphuran, C. K., Nandini, K., & Pavithran, K. (2003). Heritable anomalies among the inhabitants of regions of normal and high background radiation in Kerala: results of a cohort study, 1988-1994. *International Journal of Health Services, 34*(3), 483-515.
- Phares, V. (1997). Accuracy of informants: Do parents think that mother knows best? *Journal of Abnormal Child Psychology, 25*(2), 165-171.
- Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., ... & Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders, 15*(3), 130-142.
- Sawyer, M. G., Baghurst, P., & Clark, J. (1992). Differences between reports from children, parents and teachers: Implications for epidemicological studies. *Australian and New Zealand Journal of Psychiatry, 26*, 652-660.
- Shapiro, R., Siegel, A. W., Scovill, L. C., & Hays, J. (1998). Risk-taking patterns of female adolescents: What they do and why. *Journal of Adolescence, 21*(2), 143-159.
- Sourander, A., Santalahti, P., Haavisto, A., Piha, J., Ikäheimo, K., & Helenius, H. (2004). Have there been changes in children's psychiatric symptoms and mental health service use? A 10-year comparison from Finland. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*(9), 1134-1145.
- 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.
- Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality, 34*(4), 397-423.
- Tick, N. T., Van der Ende, J., & Verhulst, F. C. (2007). Twenty year trends in emotional and behavioral problems in Dutch children in a changing society. *Acta Psychiatrica Scandinavica, 116*(6), 473-482.
- Uitenbroek, D. G. (1997). "SISA Binomial," <http://www.quantitativeskills.com/sisa/statistics/t-test.htm>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70.
- Verhulst, F. C., Ende, J. V. D., & Rietbergen, A. (1997). Ten year time trends of psychopathology in Dutch children and adolescents: no evidence for strong trends. *Acta Psychiatrica Scandinavica, 96*(1), 7-13.
- Wang, M., & Russell, S. S. (2005). Measurement

- equivalence of the job descriptive index across Chinese and American workers: Results from confirmatory factor analysis and item response theory. *Educational and Psychological Measurement*, 65, 709-732.
- 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(6), 1038-1050.
- Zukauskiene, R., Ignataviciene, K., & Daukantaite, D. (2003). Subscales scores of the Lithuanian version of CBCL: Preliminary data on the emotional and behavioural problems in childhood and adolescence. *European Child & Adolescent Psychiatry*, 12(3), 136-143.
- 1차원고접수 : 2015. 02. 27.
수정원고접수 : 2015. 12. 03.
최종게재결정 : 2015. 12. 07.

Secular Changes in Behavior Problems of Korean Youths Based on the Korean-CBCL and -YSR

Min Young Kim	Jin Lee	Young Ah Kim	Eun Hye Ha	Kyung Ja Oh
Huno Inc.	Harvard School of Public Health	Huno Inc.	Sookmyung Women's University	Yonsei University

The aim of this study was to see whether the prevalence of the behavioral and emotional problems of Korean youths has changed across the first and the second phases of the nation-wide standardization for the Korean-CBCL and -YSR. Instruments for the first- and second-phase were based on the original US 1991 and 2001 versions, respectively. Independent samples recruited respectively for the two phases of the standardization of the Korean-CBCL and -YSR were utilized. Measurement Equivalence tests supported that both CBCL and YSR measurement constructs have consistent meanings across the standardization phases. To compare CBCL and YSR scores of the two cohort groups, 600 cases (300 boys and 300 girls) between the ages of 12 and 17 years were randomly selected from each of the standardization samples. There was a significant decline in CBCL (parents report) while there was little change in YSR (self report) across the first and the second standardization phases. Pearson r correlation showed very high stability ($r > .90$) in the rank ordering of item scores across the two phases. Implications of using CBCL and YSR as a clinical and research tool were discussed.

Key words : Child Behavior Checklist, Youth Self-Report, youth behavior problems, time trend in psychopathology