

아동성폭력 피해 가능성 평가기준 개발 및 타당화*

김 태 경 이 영 호†

가톨릭대학교 심리학과

본 연구는 아동성폭력 피해 가능성 평가기준을 개발하고 타당화하는 것을 목적으로 하였다. 연구 1은 아동성폭력 피해 가능성 평가기준의 개발 단계로, 만2세 이상~만13세 미만의 성폭력 의심사례 307건을 객관증거와 행동정황증거의 유무를 기준으로 증거존재집단(객관증거집단, 객관-행동정황증거집단, 행동정황증거집단)과 증거부재집단으로 분류한 후 관련 전문가들이 아동성폭력 피해 가능성 평가를 위해 사용하고 있는 기준들이 두 집단을 유의하게 변별해주지를 검증하였다. 그 결과, 진술 시에 아동이 나타낸 감정반응, 진술신빙성점수, 및 최초진술동기에서 증거존재집단이 증거부재집단에 비해 유의하게 높은 점수를 나타내었다. 임상증상 중에서는 급성스트레스증상에서만 행동정황증거가 존재하는 두 집단이 증거부재집단에 비해 높은 점수를 보였고, 객관증거집단은 이들의 중간 정도에 위치하였다. 연령에 부적절한 성행동 증가는 증거부재집단이 증거존재집단에 비해 유의하게 높은 점수를 보였다. 그 밖의 임상증상에서는 집단 간 차이가 유의하게 나타나지 않았다. 변별변인을 탐색하기 위한 로지스틱 회귀분석 및 ROC 분석을 토대로 학령전기의 변별모델을 [진술감정×2 + 진술신빙성요소 + 최초진술동기 + 자극노출 시 주관적 고통감×2], 그리고 이 연령대의 기준점수를 6점으로 결정하였고, 학령기의 변별모델을 [진술감정×4 + 진술신빙성요소], 그리고 기준점수를 8점으로 결정하였다. 연구 2는 연구 1에서 개발한 평가기준의 타당화단계로, 101명의 성폭력 피해 의심아동 사례를 증거존재집단과 증거부재집단으로 구분한 후 새로 개발된 기준을 사용한 분류와의 일치율을 검증하였다. 그 결과 학령전기의 분류정확률이 95.7%였으며, 학령기의 분류정확률은 90.1%이었고, 두 연령대 모두 카파 계수와 민감도 및 특정도가 만족스럽게 나타났다. 널리 활용되고 있는 진술분석기법인 CBCA 점수를 통한 집단 분류 정확률과 새로 개발된 평가기준을 통한 분류정확률을 비교한 결과, 새로 개발된 평가기준의 정확률이 더 높게 나타났다. 이러한 결과를 종합하여 연구의 시사점과 의의, 제한점, 및 후속연구를 위한 제언을 논의하였다.

주요어 : 아동성폭력, 피해가능성, 평가기준, 타당화, CBCA

* 본 연구는 2010년 가톨릭대학교 심리학과 박사학위 청구논문의 일부임.

† 교신저자: 이영호, 가톨릭대학교 심리학과, 경기도 부천시 원미구 역곡2동 산43-1.

Fax : 032-345-5189, E-mail : yhlee@catholic.ac.kr

아동성폭력 사건의 주요 특징은 구강이나 하문 및 신체 부분을 이용한 성폭력이 많기 때문에 산부인과 등의 신체검사 상 이상이 없는 경우가 많으며, 피해자인 아동이 유일한 목격자인 사례가 많고, 물리적 및 신체적 증거나 용의자 자백과 같은 증거가 부재한 경우가 매우 많다는 것이다(Bays & Chadwick, 1993; Faller, 2007). 또한, 아동이 가해자로부터 비밀을 지키도록 자주 강요받으며 비밀을 지키기 위해 아동이 피해 사실을 잘 폭로하지 않기 때문에 피해가 지속되는 경우가 잦고(Cederborg, Lamb & Laurell, 2007; Goodman-Brown, Edelstein, Goodman, Jones & Gordon, 2003), 이러한 폭로의 지연은 객관적인 증거확보의 어려움을 가중시킨다.

성폭력피해 아동에게서 자주 나타나는 임상적 징후(예를 들어, 우울, 불안, 자살시도, 자해 등)가 성폭력 피해의 증거로 고려되기도 하지만, 이러한 증상들이 성폭력 피해 아동에게서만 특징적으로 나타나는 것은 아니기 때문에 성폭력 피해아동과 비피해 임상군을 구분하는 데에는 문제가 있다는 지적이 있어 왔다(Kendall-Tackett, Williams & Finkelhor, 1993; London, Bruck, Ceci & Shuman, 2005). 따라서 대안으로 아동의 진술과 그 진술의 신빙성이 사건의 발생 여부를 판단하는 중요한 기준으로 고려되고 있는데, 아동의 발달적 특성 상 다양한 변인들에 의해 회상의 정확도와 진술의 양에 차이를 보이며(김태경, 이영호, 2010), 암시나 유도에 의해 기억이 왜곡될 가능성이 성인에 비해 높고(Greenhoot, 2000; Ornstein & Haden, 2002/2009), 어린 아동이 출처탐지 곤란을 보이기 때문에(Leichtman, Morse, Dixon & Wilch-Ross, 2000; Saywitz & Lyon, 2002), 아동의 진술만을 근거로 피해를 입증하는 것 역시 매

우 어렵다.

아직까지 성폭력 피해 아동과 비피해 아동을 구별해준다고 일반적으로 수용될만한 과학적 절차나 기준은 알려져 있지 않다(Poole & Lindsay, 1998). 이는 피해의 가능성 판단 시 고려해야 할 변인들이 매우 다양하고, 각 변인들이 복잡한 방식으로 상호작용하며, 아동의 특성에 따라서 그러한 변인들이 미치는 영향력이 달라지기 때문이다. 그럼에도 불구하고 다양한 목적으로 다양한 기법과 배경을 가진 여러 사람들이 아동성폭력 피해의 가능성을 평가하고 있는 것이 현실이다(Poole & Lindsay, 1998). 전문가들마다 강조하는 요소들에 약간의 차이가 있기는 하나 많은 전문가들이 아동의 진술에 나타난 요소와 임상증상요소 및 진술정보의 타당도와 관련된 요소들을 중요하게 간주하는 것으로 알려져 있다.

전문가들이 공통적으로 중요하게 고려하는 진술관련 요소에는 세부내용의 풍부함(Conte, Sorenson, Fogarty, & Rosa, 1991; Faller, 2007; Heiman, 1992), 사건이 발생하게 된 맥락정보(Heiman, 1992; Kuehnle & Connell, 2009; Wehrspann, Steinhauer, & Klajner-Diamond, 1987; Wharff, 1998), 사건과 관련된 구체적이거나 독특한 세부내용(Conte, Sorenson, Fogarty, & Rosa, 1991; Sink, 1988a, 1988b; Steller & Koehnken, 1989; Wharff, 1998), 진술에 부합하는 정서반응(Faller, 2007; Heiman, 1992), 협박내용 혹은 비밀주의적 요소(Heiman, 1992; Kuehnle & Connell, 2009; Sink, 1988a, 1988b; Wehrspann, Steinhauer, & Klajner-Diamond, 1987), 피해자와 가해자 간의 상호작용 기술(Steller & Koehnken, 1989), 진술의 구조적 성질과 진술의 자발성(Faller, 2007; Heiman, 1992), 대화의 인용(Conte, Sorenson, Fogarty, & Rosa, 1991; Kuehnle &

Connell, 2009) 등이 있다. 이러한 내용들을 종합한 진술분석 기법이 개발되어 사용되고 있는데 준거기반내용분석(Criterion-Based Content Analysis; CBCA)(Steller & Koehnken, 1989)과 사실성 평가(Reality Monitoring; RM)(Sporer, 1997)가 대표적이다.

CBCA는 실제로 경험한 사건에 대한 진술이 허위나 상상에 근거한 진술과는 내용 및 질적인 측면에서 상이할 것이라는 Undeutsch(1984)의 가정에 근거하여 Steller와 Koehnken(1989)이 체계화한 아동진술분석 기법이다. CBCA는 허위진술보다 진실한 진술에서 자주 발견된다고 지지받은 19개의 기준을 평가하며, 일반적 요소와 표현요소, 진술내용요소, 동기부여요소 그리고 범죄의 전형성 관련 요소로 구분된다. 일반적 요소 범주에는 논리적 일관성, 체계적이지 않은 표현, 세부내용의 풍부함이 포함된다. 표현상 요소에는 맥락상의 깊이, 상호작용의 묘사, 대화의 인용, 사건 도중 예상치 못한 일 발생이 포함된다. 세 번째 범주인 진술내용에는 독특한 세부묘사, 부가적인 세부묘사, 이해하진 못했지만 정확하게 진술한 세부묘사, 간접적인 관련 요소, 주관적인 심리상태, 가해자의 정신 상태에 대한 진술이 포함된다. 동기부여요소에는 자발적인 수정, 기억부족 시인, 자기 자신의 진술에 관한 의심제기, 자기 비난, 가해자용서가 포함되며, 마지막 다섯 번째 범주에는 범죄의 전형성에 대한 묘사가 존재하는지가 포함된다.

CBCA의 효용성에 대해서는 아직까지 논쟁이 진행 중으로, 성폭력 피해 아동을 타당하게 탐지해낸다고 보고하는 연구가 있는 반면(Craig, Scheibe, Raskin, Kircher, & Dodd, 1999) 신뢰도가 충분치 못하다는 지적도 있다(Anson, Golding, & Gully, 1993). 특히 어린 아동의 진

술에 대해 CBCA를 실시하는 것은 적절치 않다는 지적이 있으며(Wells & Loftus, 1991), 기존의 연구 결과들이 피해 가능성이 매우 높은 아동과 매우 낮은 아동들만을 비교함으로써 CBCA의 효용성을 과대평가하도록 만들었을 뿐 모호한 사례에서는 CBCA의 판별력이 만족스럽지 못하다는 지적도 있다(Lamb, Sternberg, Esplin, Hershkowitz, & Orbach, 1997; Vrij, 2005). 또한, CBCA의 하위 요소별로 증명력이 상이하므로 상이한 가중치를 부여할 필요성을 제안하는 전문가도 있다(예를 들어, Blandon-Gitlin, Pezdek, Lindsay, & Hagen, 2008). CBCA가 가진 또 다른 제한점은 명확한 채점기준이 부재할 뿐 아니라, 진술 신빙성에 대한 최종 판단을 위한 기준이 없으며, 결과적으로 최종 판단이 평가자의 주관적 판단에 맡겨진다는 것이다(Mazzoni & Ambrosio, 2003).

RM은 CBCA 점수가 사건에 대한 익숙함과 코칭에 의해 영향을 받기 때문에 이 점수를 근거로 피해 가능성을 평가할 경우 오긍정 오류의 가능성이 높고, CBCA에 이론적 근거가 부족하다는 비판이 일면서(Blandon-Gitlin, Pezdek, Rogers, & Brodie, 2005) CBCA의 대안으로 주목받게 되었다. RM이 바탕으로 하는 이론에 따르면, 경험된 사건의 기억특성들은 거짓으로 꾸며낸 기억의 특징과는 양적 및 질적으로 상이하며, 특히 지각적 정보(사건과 관련된 시각적 정보, 소리, 맛, 느낌 등), 맥락정보(사건이 발생한 시간과 장소 관련 정보), 그리고 정서적 정보(사건에 대한 느낌이나 통각정보)를 더 많이 담고 있다. 지어낸 사건 기억은 내적 작용에 의해 도출된 것이기 때문에 사고와 추론 같은 인지적 작동을 담고 있는 경향이 있으며, 모호하고 덜 구체적이다(Johnson, Hashtroudi, & Lindsay, 2009).

RM 기준은 연구자들마다 차이가 있다. 그 중에서 Sporer(1997)는 8개의 기준을 제시한 바 있으며 여기에는 진술의 명료함, 지각정보, 공간정보, 시간정보, 정서, 이야기의 재구성 가능성, 현실성, 인지적 작동이 포함된다. 인지적 작동은 거짓 진술에서 자주 나타나며 나머지는 진실한 진술에서 나타난다고 가정된다. RM 관련 연구들은 대부분 실험실에서 이루어진 것들인데, Vrij, Edward, Roberts와 Bull(2000)의 연구 결과 RM 점수가 진위를 71%~74% 정확하게 분류하였으며, Höfer, Akehurst와 Metzger(1996)의 연구에서는 70%, 그리고 Sporer (1997)의 연구에서는 75%~68% 정확하게 분류한 것으로 나타났다. 이러한 정확도는 CBCA와 비슷한 것이기는 하나, CBCA가 아동을 대상으로 한 결과인 것에 반해 RM 연구들은 대부분 성인을 대상으로 한 실험실 연구들이다. 또한, RM이 진짜 발생하였던 사건과 꾸며낸 사건을 효과적으로 구분해주는지도 아직까지 불분명하다(Otgaar, Candel, Merckelbach, & Wade, 2009).

임상증상요소에는 연령에 부적절한 성행동 증가(Conte et al., 1991; Corwin, 1988; Faller, 2007; Friedrich, 1997; Haskett, Wayland, Hutcheson, & Tavana, 1995; Heiman, 1992; Sgroi, Porter & Blick, 1982; Wehrspann et al., 1987; Wharff, 1998; Kuehnle & Connell, 2009)가 가장 많은 전문가들로부터 지지받고 있으며, 재경험과 공포, 회피, 악몽, 불안 등을 포함하는 급성스트레스 증상, 공격성, 가성숙, 학업관련 문제, 남성공포, 품행문제, 우울증, 배변문제, 수면문제, 자살충동, 해리, 수치심과 죄책감 등(Briere, 2005; Corwin, 1988; Elliott & Briere, 1994; Faller, 2007; Heiman, 1992; Sink, 1988a, 1988b)도 중요한 임상지표로 고려되고 있다.

그러나, 이러한 임상지표들은 비피해 임상집단에서도 자주 발견되는 것이기 때문에 특정도가 만족스럽지 못하다는 지적이 있다(예를 들어, Faller, 2007). 특히, 연령에 부적절한 성행동의 증가는 민감도가 높은 대신 특정도가 만족스럽지 못하기 때문에 이를 평가기준으로 사용할 때에는 반드시 아동의 연령과 성폭력 이외의 스트레스 사건 및 가족의 섹슈얼리티 등을 고려해야 한다(Friedrich, Grambsch, Damon, 1992).

정보의 타당도와 관련된 요소에는 진술의 일관성(Heiman, 1992; Kuehnle & Connell, 2009; Steller & Koehnken, 1989), 진술이 아동의 관점에서 이루어지는지(Faller, 2007; Jones & McGraw, 1987; Sink, 1988a), 반대 증거에도 불구하고 일관된 진술(Wehrspann et al., 1987), 허위진술특성(DeYoung, 1986; Faller, 2007), 코칭의 흔적(Heiman, 1992), 그리고 진술의 그럴듯함(Kuehnle & Connell, 2009) 등이 포함된다. 모든 전문가들은 피해의 가능성 평가 시에 반드시 이와 같은 타당성 관련 요소들을 충분히 검토하여 결론을 도출하도록 제안하고 있다. 무엇보다, 피해의 가능성이 제기된 후 공식적인 진술조사를 받기까지 아동이 누군가와 사건에 대해 나누었던 대화가 아동의 기억을 왜곡시킬 정도로 심각하게 유도적이거나 암시적이었었다는 전제조건이 충족되어야만 피해의 가능성 평가가 가능하다.

우리나라에서도 최근 십여 년 간 끔찍한 아동성폭력 사건이 연이어 발생하면서 아동이 형사사법 체계 안으로 들어오는 경우가 증가하였으며, 그 결과로 아동이 진술자로서 지니는 일반적인 특성, 특정 아동의 진술특성, 그리고 성폭력 피해의 가능성 평가 등과 관련한 아동심리전문가의 자문에 대한 요구가 급증하

게 되었다. 이러한 요구에 맞추어 2010년 4월 15일 제정된 「성폭력범죄의 처벌 등에 관한 특례법」에는 13세 미만 아동의 성폭력사건 수사 시 전문가 자문을 의무화하는 법 규정이 신설되었다. 그러나, 아직까지 우리나라에서는 아동의 진술특성이나 진술신빙성 관련 연구가 매우 적어서 몇몇 실험실 연구(예를 들어, 광금주, 김연수, 2003; 김재연, 이재연, 2000; 박자경, 1998; 이재연, 1998)와 문헌연구(예를 들어, 광금주, 이승진, 2006), 그리고 현장연구(예를 들어, 김태경, 이영호, 2010; 신의진, 김태경, 최경숙, 2008; 조은경, 2004)가 있을 뿐이다. 당연히 충분한 지식과 경험을 갖춘 전문가를 찾아보기도 매우 힘들다. 그럼에도 불구하고 현실적인 요구에 밀려 전문성을 갖추지 못한 사람들이 이러한 역할을 대신하고 있는데, 이러한 미봉책이 향후 심리학자에 대한 형사사법기관의 불신으로 이어질 가능성이 높아 염려스러운 실정이다. 이에, 본 연구는 우리 문화와 실정에 맞는 아동성폭력 피해 가능성 평가 기준을 개발하고 타당화하는 것을 목적으로 하였으며, 이러한 노력은 아동진술 관련 전문가의 육성에 이바지할 것이라 기대한다.

연구 1: 아동성폭력 피해 가능성 평가기준 개발

연구 1에서는 성폭력 피해가 의심되어 전문기관을 방문한 아동들의 자료를 분석하여 관련 전문가들이 사용하고 있는 기준들이 우리나라 성폭력 의심 아동의 피해 가능성 평가에 얼마나 도움이 되는지를 알아보고, 우리나라 실정에 맞는 평가기준을 개발하고자 하였다. 한편, 아동성폭력 피해 가능성 평가 시 정보

의 타당성과 관련된 요소들을 고려하는 것이 매우 중요한데, 이 요소들은 전문적인 지식과 경험을 토대로 한 질적 분석을 요하기 때문에 이를 수량화하는 것은 의미가 없으므로 분석에서 제외하였다.

방 법

연구대상

2007년 7월부터 2009년 11월까지 아동성폭력전담지원기관에 방문하여 사회사업가의 접수면담, 정신과 전문의의 정신의학적 평가, 및 임상심리전문의의 종합심리평가와 NICHD 조사 프로토콜(National Institute of Child Health and Development investigation protocol)에 의거하여 성폭력 사건에 대해 조사를 받은 만 2세 이상~만 13세 미만의 아동을 대상으로 하였다. 이 프로토콜은 Lamb, Sternberg, Esplin과 Orbach(1997)가 고안한 것으로, 수많은 아동조사 관련 연구결과들에 기초하여 마련된 조사 지침을 실제 조사에 적용 가능하도록 만들었으며 조사규칙에 대한 설명과 연습에서부터 실질적인 조사 주제로 이동하는 다양한 기법, 그리고 개방적인 탐색질문에서부터 유도질문에 이르는 일련의 질문방식들을 포함하고 있다.

316건 중 자료가 불충분한 아동을 제외하고 최종적으로 분석된 사례는 모두 307건이었다. 여아가 267명(87%)이었으며, 남아가 40명(13%)이었고, 평균 연령은 7.3세(SD=3.0, 범위 2~12세)였다. 학령전이 129명(42%)이었고, 학령기가 178명(58%)이었다. 연구기관에서는 개인신상정보를 제외한 자료를 성폭력 피해예방과 대책

마련을 위한 연구자료로 사용하는 것에 대한 동의서를 받고 있으며, 모든 사례가 이에 동의하였다.

측정도구

인구통계학적 정보 및 사건관련 변인 기록지

아동의 인구통계학적 특성과 사건 특성을 파악하기 위한 기록지로, 조은경(2004)의 연구에서 사용되었던 아동성폭력사건 기록조사표를 참고하여 고안하였다. 이 기록지를 통해 얻을 수 있는 정보로는 아동의 현재 연령, 성별, 최초피해로부터 피해가 인지된 시점까지의 소요시간, 피해자와 가해자의 관계, 사건의 심각도가 있다.

객관적인 증거 기록지

의학적 증거, 목격자 진술, 피의자의 자백, 물리·신체적 증거를 객관적 증거로 간주하였으며, 이를 파악하기 위한 기록지이다. 의학적 증거는 산부인과 전문의 등의 진단서 혹은 소견서를 근거로 평가하였고, 목격자 진술과 가해자 자백은 그 내용이 사건과 얼마나 관련성이 있으며 신뢰로운가를 고려하여 평가하였다. 물리·신체적 증거는 수사과정에서 경찰에 의해 수집된 성폭력 관련 증거들을 토대로 평가하였다. 최소 하나 이상의 증거범주에서 명백한 증거가 존재하는 경우에만 객관증거가 존재하는 것으로 간주하였으며, 모호한 증거만 여러 개 존재하는 경우에는 증거부재로 간주하였다. 객관적인 증거가 존재하는 아동은 170명(55.4%)이었으며, 이들의 평균 연령은 8.1세(SD=2.6)였고, 학령전기 49명(28.8%), 학령기 121명(71.2%)이었다.

행동정황증거 평가 기록지

성폭력 피해 가능성을 평가하기 위한 기준 개발을 위해서는 피해가 명백한 사례와 피해가 명백히 아닌 사례를 비교 검토하여 두 집단을 유의하게 변별해주는 변인을 찾아 기준을 마련하는 것이 가장 합당하다. 그러나, 성폭력 피해를 입은 아동 중에서도 객관적인 증거가 부재한 사례가 많기 때문에, 객관적 증거만을 기준으로 피해 여부를 평가하여 분류하는 것은 문제가 된다. 따라서, 본 연구에서는 객관적인 증거는 없으나 아동의 행동과 진술 및 맥락 정보를 감안해볼 때 성폭력 피해가 명백하다고 평가할 수 있는 사례를 따로 구분하였으며, 이와 같은 증거를 행동정황증거로 명명하고 성폭력 피해의 또 다른 증거로 고려하였다. 행동정황증거가 존재하는 아동은 모두 171명(55.7%)이었으며, 이 중 104건(60.8%)은 행동정황증거와 객관적인 증거가 함께 존재하였다.

진술관련 요소 평가 기록지

진술관련 요소를 파악하기 위해 고안된 기록지로, 이를 통해 얻을 수 있는 정보에는 진술 중에 드러낸 감정반응(진술감정), 진술신빙성요소, 그리고 성폭력 피해를 처음으로 보고하게 된 때의 정황정보와 최초폭로유형이 있다. 진술감정은 진술조사 중에 아동이 어떠한 감정반응을 드러내는지(예를 들어, 불안, 공포, 회피, 분노, 성적 불쾌감, 죄책감, 수치심 등), 성적 각성을 드러내는지, 그리고 그러한 감정반응이 진술내용과 부합하는지를 관찰하여 그 빈도와 강도에 따라 없음(0점), 약간 관찰됨(1점), 강하게 관찰됨(2점)으로 평정하였다. 감정의 강도가 강하게 자주 드러날수록 높은 점수를 주었다. 30사례를 연구자 이외의 다른 임

상심리전문가가 재채점한 후 평정자간 일치도를 분석한 결과 평정자간 상관계수는 .82였다.

진술신빙성요소로는 많은 전문가들이 공통적으로 사용하고 있는 진술관련 요소인 세부 묘사의 양, 사건이 발생한 맥락정보, 성폭력 당시의 피해자와 가해자 간 상호작용묘사, 그리고 사건과 관련된 독특한(혹은 명료하거나 구체적인) 세부묘사를 고려하였다. 이러한 요소들은 해당 연구 결과들을 종합하여 개발된 CBCA에 이미 포함되어 있으며, 많은 CBCA 효용성 연구(예를 들어, Vrij, 2004; 조은경, 2004)에서 일관되게 피해 가능성이 높은 집단과 낮은 집단을 변별해준다고 보고된 요소들이기도 하다. 그러나, 18개 혹은 19개 요소들을 측정하는 CBCA는 개별 요소들의 상대적 변별력에 큰 차이가 있고 채점 기준이 명확하지 않다. 따라서, 본 연구에서는 CBCA 대신 선행연구들에서 공통적으로 높은 증명력을 보여주었던 네가지 요소만을 측정하였다. 각각의 요소는 없음(0점), 약간 나타남(1점), 뚜렷이 나타남(2점)으로 평정하였으며, 분석에는 각 점수를 합산한 총점이 사용되었다. 30사례에 대해서, 임상심리전문가의 1차 평가 결과와 또 다른 임상심리전문가에 의한 2차 평가 결과 간 일치도를 분석한 결과 평정자간 상관계수(interrater correlation coefficient; ICC)는 .84였다.

최초폭로유형은 진술일관성 및 최초진술 과정에서 있었을 수 있는 어른의 유도나 암시 가능성을 평가할 수 있도록 해주며, 진술에 대한 아동의 동기수준을 파악할 수 있는 중요한 정보이다. 따라서, 본 연구에서는 최초로 아동이 어떠한 맥락에서 어떠한 형태로 피해를 진술하였는지를 측정하였다. 최초진술유형을 토대로 아동의 최초 진술동기 수준을 평정하였으며, 비폭로 0점, 유도적인 탐색에서 폭

로 1점, 개방적인 탐색에서 폭로 2점, 그리고 의도적인 자발적 폭로를 3점으로 평정하였다. 연구가 이루어진 기관에서는 상담기록지에 최초 폭로유형을 기록하도록 되어 있었으며, 피해의 인지과정에 대해 상세하게 기록하도록 되어 있다. 연구자가 이 자료를 검토하여 재평가하였으며, 30사례에 대한 평정자 간 일치도를 분석한 결과 평정자간 상관계수는 .93이었다.

임상증상 평가 기록지

많은 전문가들이 성폭력 피해 아동에게서 자주 발견되는 임상증상들을 성폭력피해의 중요한 지표로 간주하고 있다(예를 들어, Adams, 2004; Heiman, 1992). 여기에는 재경험과 관련 자극 회피 등을 포함하는 급성스트레스 증상, 과도한 자위행위, 우울, 불안, 자살시도, 공격성 증가, 악몽 등이 포함된다. 본 연구에서는 이러한 임상증상들의 증거 가치를 검토하기 위해 6~18세의 평균 지능을 가진 아동에 대한 진단평가도구인 Korean Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version (K-SADS-PL-K)(Kim et al., 2004)의 선별 면담지 중 일부를 참고하여 임상증상 기록지를 구성하였다. 여기에는 우울장애, 조증, 정신증, 공황장애, 분리불안장애, 회피성장애(사회공포증), 광장공포증 및 특정공포증, 과불안(범불안장애), 강박장애, 유노증, 유분증, 신경성 식욕부진증, 신경성 폭식증, 주의력결핍 및 과잉행동장애, 반항성 도전장애, 품행장애, 틱장애, 및 외상후스트레스장애가 포함된다.

임상증상에 대한 평가는 임상심리전문가가 보호자 및 아동과의 면담을 통해 수집된 정보와 정신과 전문의의 진료 시 수집된 정보를

통합하여 이루어졌으며, 증상이 전혀 없으면 0 점, 약간 존재하나 심각한 정도는 아닌 경우 1 점, 그리고 매우 심각한 경우는 2점을 주었다. 아동과 보호자의 보고내용이 상이한 경우 상이한 이유를 검토하고 여러 정보들을 종합하여 최종적으로 평가하였다. 30사례를 무선표집하여 재평가한 후 일치도를 분석한 결과 평정자간 상관계수는 .84였다.

자료의 분석

우선, 증거의 종류에 따른 차이를 알아보기 위해 증거 유무 및 증거의 종류에 따라 객관 증거집단, 객관증거와 행동정황증거가 모두 존재하는 객관-행동정황증거집단, 행동정황증거집단, 및 두 종류의 증거가 모두 존재하지 않는 집단으로 분류하였다. 그런 다음 인구통계학적 및 사건관련 변인과 진술관련 요소 및 임상증상요소의 집단 간 차이를 검증하였다. 이를 위해 상관분석, 교차분석, 변량분석, 및 공변량분석이 이루어졌다. 분석 과정에서 연령의 영향력이 크게 나타났으므로 연령대별 집단 간 차이분석이 추가로 이루어졌다. 이러한 결과들을 토대로 증거의 종류와 무관하게 피해의 증거가 존재하는 집단과 부재하는 집단을 유의하게 변별해주는 변인을 탐색하여 판별기준점수를 도출하기 위한 이항 로지스틱 회귀분석이 이루어졌다. 분석은 학령전후로 구분하여 이루어졌다. 그리고, 실무자들이 피해 가능성 판단 시 사용할 수 있는 기준을 도출하기 위하여 로지스틱 회귀모형에 진입한 예측변인에 가중치를 부여한 모델을 만들고 ROC 분석을 통해 각 모델의 적합성을 검증, 가중모형을 선정하였다. 그런 후 가중모형의 점수대별 사례분포와 ROC 분석을 통한 민감

도 및 특정도 검토결과를 종합하여 종적으로 아동성폭력 피해가능성평가 기준을 도출하였다.

결 과

집단 간 차이검증

인구통계학적 및 사건관련 변인의 집단 간 차이 검증

인구통계학적 및 사건관련 변인의 기술 통계치와 집단 간 차이 검증 결과가 표 1에 제시되었다. 표에서 볼 수 있듯이, 증거부재집단이 나머지 세 집단에 비해 연령이 유의하게 낮았다. 또한, 사건관련 변인의 집단 간 차이 검증 결과 행동정황증거가 존재하는 두 집단이 객관증거만 존재하는 집단 및 증거부재집단에 비해 남아의 비율이 더 높았고, 1회 피해를 및 피해 직후 폭로율이 높았으며, 위협율도 훨씬 더 높았다. 삽입율은 객관증거와 행동정황증거가 함께 존재하는 집단이 증거부재집단에 비해 세배 이상 높았다.

연령이 거의 모든 변인과 높은 상관관계를 나타내었으므로, 연령대에 따른 인구통계학적 및 사건관련 변인의 차이를 추가적으로 검증하였다. 그 결과, 학령전기에 해당하는 6세 이하에서는 학령기인 7세 이상에 비해 당일 폭로율이 낮고, $\chi^2(1, N=307)=12.22, p<.05$, 가해자가 친밀한 타인인 사례가 많으며, $\chi^2(1, N=307)=35.11, p<.001$, 학령기에 비해 삽입율이 유의하게 낮고, $\chi^2(1, N=307)=17.72, p<.001$, 가해자의 위협율, $\chi^2(1, N=307)=15.88, p<.01$ 도 유의하게 낮았다.

표 1. 인구통계학적변인 및 사건관련 변인의 집단 간 차이검증

		객관증거 집단(n=66)		객관-행동정황 증거집단(n=104)		행동정황증거 집단(n=67)		증거부재 집단(n=70)		χ^2
		빈도	%	빈도	%	빈도	%	빈도	%	
		성별	여	49	74.2	96	92.3	63	94.0	
	남	17	25.8	8	7.7	4	6.0	11	15.7	
피해-인지 소요시간	당일	14	23.7	56	54.9	26	43.3	7	15.6	37.16***
	3개월 이내	24	40.7	17	16.7	10	16.7	18	40.0	
	1년 이내	11	18.6	14	13.7	11	18.3	13	28.9	
	2년 이내	6	10.2	11	10.8	7	11.7	5	11.1	
	2년 이상	4	6.8	4	3.9	6	10.0	2	4.4	
	모름	7		2		7		25		
피-가해자 관계	낯선사람	19	28.8	57	54.8	34	50.7	12	19.0	33.45***
	친밀한타인	24	36.4	14	13.5	19	28.4	24	38.1	
	친인척·가족	23	34.8	33	31.8	14	20.9	27	42.9	
	모름	0		0		0		7		
피해빈도	1회	21	34.4	61	59.8	39	62.9	10	34.4	17.91**
	2~9회	21	34.4	19	18.6	9	14.5	8	27.6	
	10회 이상	19	31.1	22	21.6	14	22.6	11	37.9	
	모름	5		2		5		41		
삽입여부	없음	50	76.9	56	53.8	48	72.7	49	86.0	21.33***
	있음	15	23.7	48	46.2	18	27.3	8	14.0	
	모름	1		0		1		13		
위협정도	없음	47	74.6	57	55.3	39	58.2	49	87.5	27.03***
	언어위협	10	15.8	26	25.2	22	32.8	6	10.7	
	무기위협	4	6.3	13	12.6	5	7.5	1	1.8	
	실제상해	2	3.2	7	6.8	1	1.5	0		
	모름	3		1		0		14		
연령		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i>
		7.4	2.6	8.6	2.6	7.6	3.0	5.1	2.8	

진술관련 요소의 집단 간 차이검증
진술관련 요소의 집단 간 차이검증 결과는 표 2에 제시되었다. 진술감정과 진술신빙성에

서 증거부재집단이 나머지 세 집단에 비해 유의하게 낮은 점수를 보였다. 최초진술유형을 토대로 측정 한 진술동기점수는 행동정황증거

표 2. 진술관련 요소의 집단 간 차이검증

	객관증거 n=66		객관-행동정황증거 n=104		행동정황증거 n=67		증거부재 n=70		F	Scheffé ^a
	M	SD	M	SD	M	SD	M	SD		
진술감정	1.18	.89	1.27	0.90	1.00	0.91	0.19	0.43	27.75***	1,2,3>4
진술신빙성	3.61	2.66	4.88	2.82	4.85	2.90	0.84	1.53	40.71***	1,2,3>4
최초진술동기	1.97	0.82	2.52	0.64	2.42	0.72	1.39	0.80	37.85***	2,3>1>4

^a 1=객관증거집단, 2=객관-행동정황증거집단, 3=행동정황증거집단, 4=증거부재집단

*** $p < .001$

가 존재하는 두 집단이 나머지 두 집단에 비해 유의하게 높았고, 객관증거집단이 증거부재집단에 비해 유의하게 높았다. 한편, 연령대별 진술관련요소의 차이를 검증한 결과, 세 변인 모두에서 학령기 아동이 학령전기 아동에 비해 유의하게 높았다, 순서대로 $F(1, 305)=39.49, p < .001$; $F(1, 305)=102.41, p < .001$; $F(1, 305)=51.68, p < .001$.

임상증상의 집단 간 차이검증

표 3에 임상증상의 집단 간 차이검증 결과가 제시되어 있다. 표에서 볼 수 있듯이, 재경험, 성폭력 사건과 관련된 자극 노출 시 고통감, 그리고 사건 관련 상황이나 활동에 대한 회피 증상에서 행동정황증거가 존재하는 두 집단이 나머지 두 집단에 비해 유의하게 높았다. 불안과 공포반응에서도 행동정황증거가 존재하는 집단이 증거부재집단에 비해 유의하게 높았으며, 객관-행동정황증거집단이 객관증거집단에 비해 유의하게 높았다. 과다경계나 놀람반응은 행동정황증거가 존재하는 두 집단이 증거부재집단에 비해 유의하게 높았으며, 해리/착각/환각 증상에서는 객관-행동정황증거집단이 증거부재집단에 비해 유의하게 높았고,

악몽과 수면문제는 행동정황증거집단이 객관증거집단에 비해 유의하게 높았다.

한편, 연령대에 따른 임상증상의 차이를 추가로 분석한 결과, 재경험, $F(1, 305)=39.77, p < .001$, 자극노출 시 고통감, $F(1, 305)=20.83, p < .001$, 관련 상황이나 활동의 회피, $F(1, 305)=16.20, p < .001$, 과다 경계 및 놀람, $F(1, 305)=9.13, p < .01$, 해리/착각/환각, $F(1, 305)=6.01, p < .05$, 불안/공포, $F(1, 305)=8.76, p < .01$, 우울/무기력, $F(1, 305)=13.81, p < .01$ 에서 학령기가 학령 전기에 비해 유의하게 높은 점수를 보였다. 성행동 증가, $F(1, 305)=24.19, p < .001$, 성과 관련된 독특한 진술의 반복, $F(1, 305)=4.21, p < .05$, 급성적인 일상생활 변화, $F(1, 305)=5.41, p < .05$, 및 배설장애/퇴행, $F(1, 305)=13.55, p < .001$ 에서는 학령전기가 학령기보다 유의하게 높은 점수를 나타내었다.

피해 가능성 판단기준변인 탐색을 위한 로지스틱 회귀분석

객관증거와 행동정황증거 중 하나 이상이 존재하는 집단을 ‘증거존재집단’으로, 두 증거가 모두 부재한 집단을 ‘증거부재집단’으로 재

표 3. 임상증상의 집단 간 차이검증

	객관증거 n=66		객관-행동정서증거 n=104		행동정서증거 n=67		증거부재 n=70		F	Scheffé ^a
	M	SD	M	SD	M	SD	M	SD		
재경험	0.23	0.63	1.58	.82	1.60	0.80	0		122.64***	1,4<2,3
자극노출시고통	0.12	0.41	0.89	.99	1.07	1.01	0.01	0.12	137.73***	1,4<2,3
관련상황활동회피	0.14	0.49	0.96	1.00	1.13	1.00	0.01	0.12	38.43***	1,4<2,3
불안/공포반응	0.50	0.85	0.99	1.00	0.85	0.99	0.26	0.63	11.14***	4<2,3 1<2
과다경계/놀람	0.09	0.38	0.31	.73	0.37	0.78	0		6.64***	4<2,3
해리/착각/환각	0.09	0.38	0.09	.40	0.21	0.62	0.04	0.27	3.70*	4<3
악몽/수면문제	0.27	0.67	0.63	.93	0.75	0.96	0.39	0.75	4.60**	1<3
흥미감소	0.03	0.25	0.30	.71	0.09	0.38	0		1.48	
집중문제	0.06	0.35	0.12	.47	0.09	0.42	0.09	0.41	0.24	
성행동증가	0.86	0.99	0.26	.65	0.34	0.73	0.87	0.93	12.63***	1,4>2,3
우울/무기력	0.32	0.68	0.34	.75	0.18	0.58	0.17	0.54	1.42	
자살관련문제	0.03	0.25	0.06	.34	0.12	0.48	0.03	0.24	1.06	
짜증분노	0.38	0.78	0.45	.84	0.68	0.95	0.46	0.83	1.60	
성특이진술반복	0.03	0.25	0.02	.20	0.03	0.24	0		0.35	
성적불쾌감	0.06	0.35	0.06	.34	0.09	0.42	0		0.96	
신체화증상	0		0.06	.34	0.03	0.24	0.03	0.24	0.71	
품행문제	0		0		0		0.01	0.12	1.13	
급성일상변화	0.05	0.27	0.04	.28	0.09	0.42	0.14	0.49	1.32	
배설장애/퇴행	0.23	0.63	0.13	.50	0.21	0.62	0.20	0.60	0.43	
죄책감/수치심	0.09	0.42	0.08	.39	0.18	0.58	0		2.30	
폭식	0.03	0.25	0.00		0		0.03	0.24	0.84	

^a 1=객관증거집단, 2=객관-행동정황증거집단, 3=행동정황증거집단, 4=증거부재집단

p*<.05, *p*<.01, ****p*<.001

분류한 후, 두 집단을 유의하게 변별해주는 변인을 탐색하기 위한 이항 로지스틱 회귀분석을 실시하였다. 집단 간 차이검증 과정에서 연령대별로 상이한 결과가 나타났으므로, 만6세 이하(학령전기)와 만7세 이상(학령기)으로 나누어 분석을 실시하였다. 분석 시에는 앞서

이루어진 변량분석에서 유의한 집단 간 차이를 나타낸 변인들을 선정하여 투입하였으며, 변인 간 공선성을 최소화하기 위해 우도비 통계량에 근거한 전진단계 선택방식을 채택하였다.

분석 결과는 표 4와 5에 제시되어 있다. 학

표 4. 증거존재 대 증거부재 집단의 로지스틱 회귀분석 결과 : 학령전기

	B	SE	Wald	p	Exp(B)
상수	3.11	.69	20.12	.000	22.48
진술감정	1.66	.39	23.38	.000	.191
진술신빙성	.70	.18	14.80	.000	.495
최초진술동기	.79	.36	4.85	.026	.452
자극노출시고통감	1.80	.82	4.85	.028	.166

주. 증거존재집단의 n= 74, 증거부재집단의 n= 55

표 5. 증거존재 대 증거부재 집단의 로지스틱 회귀분석 결과 : 학령기

	B	SE	Wald	p	Exp(B)
상수	-.30	.49	.37	.541	1.351
진술감정	1.97	.64	9.63	.002	.139
진술신빙성	.45	.14	10.77	.001	.641

주. 증거존재집단의 n= 163, 증거부재집단의 n= 15

령전기에서는 진술감정점수, 진술신빙성점수, 최초진술동기점수, 그리고 성폭력 관련 자극노출 시 고통감점수가 두 집단을 유의하게 변별해주는 변인으로 도출되었으며, 이 모형을 사용한 집단 분류결과 증거존재집단의 분류정확률이 85.1%, 증거부재집단의 분류정확률 83.6%, 그리고 전체 분류정확률이 84.5%로 나타났다. 학령기에서는 진술감정점수와 진술신빙성점수가 두 집단을 유의하게 변별해주는 변인으로 도출되었으며, 이 모형을 사용한 집단분류결과 증거존재집단의 분류정확률 98.2%, 증거부재집단의 분류정확률 40.0%, 전체 분류정확률 93.3%로 나타났다.

아동성폭력 피해 가능성 평가기준 도출

아동성폭력 피해 가능성 평가기준을 도출하

기 위해 로지스틱 회귀분석의 최종모형에 진입한 예측변인의 설명력 또는 기여 정도로 해석될 수 있는 회귀계수(B)의 상대적 비율을 기준으로 하는 가중배점 모델을 만들었으며, 각 모델에 대한 ROC 분석을 통해 이 모델들의 적합성을 검증하였다. 가중배점 모델을 선정한 다음에는 모델의 변별기준점을 임의로 선정하여 각 점수별 곡선아래 면적 값(AUC)과 각 점수대별 피험자의 분포를 고려하여 변별기준점을 결정하였다.

한편, 변별기준점의 결정 시에는 개발된 도구가 활용될 장면에서 허용할 수 있는 오긍정율과 오부정율의 허용 범위를 고려해야만 한다. 일반적으로 발견적인 가치가 중요한 임상적인 관점에서는 오부정으로 인해 아동이 위험한 환경으로 되돌아감으로써 재피해를 당할 가능성이 높기 때문에 오부정률을 낮추는 것

이 중요하며, 수사적인 관점에서는 피해의 입증 못지않게 비가해자를 가해자로 잘못 판단할 가능성을 줄이는 것이 중요하기 때문에 가능한 오긍정률이 낮아야 한다는 보수적인 입장을 취한다. 본 연구를 통해 개발될 평가기준은 수사적 목적이 좀 더 중요한 장면에서 사용될 것으로 예상되므로, 변별 기준점수의 결정 시 비피해 아동으로부터 피해 아동을 변별해내는 정확도를 말하는 특정도가 적어도 .95 이상이며 피해 아동을 피해 아동으로 분류해내는 정확도를 말하는 민감도가 .70 이상이 되는 점수를 우선적으로 고려하였고, 점수의 가감에 따른 민감도와 특정도의 변동 폭도 추가로 고려하였다.

학령전기의 평가기준 도출

학령전기에서는 진술감정, 진술신빙성, 최초 진술동기, 그리고 성폭력 관련 자극 노출 시의 고통감을 예측변인으로 한 회귀모형이 도출되었다. 따라서, 진술신빙성의 회귀계수를 기준점수(1.00)로 하고 나머지 변인의 상대적 비율을 고려하여 가중비율을 결정하였다. 이렇게 해서 도출된 가중배점 모델이 표 6에 제시되었다. 각 모델에 대한 ROC 분석 결과, 모델 3의 AUC 값이 .918로 가장 높았다. 대체로 .71이상인 경우 예측 정확도가 높은 것

으로, .64 미만인 경우 낮은 것으로, 그리고 그 사이의 값을 가지는 경우 중간 정도의 예측 정확도가 있는 것으로 평가되므로(Eher, Rettenberger, Schiling, & Pfafflin, 2008, 노일석, 2009에서 재인용), 이 AUC 값은 예측 정확도가 상당히 높은 것으로 평가될 수 있다. 따라서, 3번 모델을 최종 가중배점 모델로 선택하였다.

가중배점모델을 선정한 후, 변별기준점수를 결정하기 위해 각 점수대별 피험자의 분포와 변별 기준점수 별 AUC 값을 살펴보았다. 점수대별 피험자의 분포를 검토한 결과, 증거부재 집단은 1점(45.5%)이 최빈치였고 5점의 누적비율이 92.7%였으며, 증거존재집단은 5점이 최빈치였고 10점 이상의 누적비율이 75.7%였다. 기준점수별 AUC 값과 민감도 및 특정도를 살펴본 결과는 표 7에 제시되어 있다. 4점에서 민감도가 .878, 특정도 .818이었으며, 5점으로 상향할 경우 민감도는 .026 감소하고 특정도는 .055 증가하였고, 6점으로 상향할 경우 민감도가 .148 감소하고 특정도는 .054 증가하였다. 7점으로 상향할 경우에는 특정도가 .95 이상으로 높아졌으나 민감도가 .70 미만으로 낮아졌다. 이러한 점들을 모두 고려하여 학령전기의 기준점수를 6점으로 결정하였다.

표 6. 로지스틱 회귀모형에 진입한 예측변인들에 대한 가중배점모델 : 학령전기

	회귀계수(B)	회귀계수 가중배점				
		모델1	모델2	모델3	모델4	모델5
진술감정	1.66(2.37)	3	3	2	2	1
진술신빙성	.70(1.00)	1	1	1	1	1
최초진술동기	.79(1.13)	2	1	1	1	1
자극노출시감정	1.80(2.57)	3	3	2	1	1

도출된 가중배점 모델이 표 8에 제시되었다. 각 모델에 대한 ROC 분석 결과, 모델 2의 AUC 값이 .92로 가장 높은 편에 속하였으므로 이 모델을 선정하였다. 변별기준점수를 결정하기 위해 각 점수대별 피험자의 분포를 살펴본 결과, 증거부재사례는 0점(40.0%)이 최빈치였으며, 6점의 누적비율이 86.7%였다. 증거존재집단의 최빈치는 16점(23.3%)이었으며, 13점 이상인 비율이 66.9%였다.

기준점수별 AUC 값과 민감도 및 특정도는 표 9에 제시되어 있다. 변별 기준점수가 5점일 때에 민감도 .871, 특정도 .800으로 나타났으며, 6점으로 상향하면 민감도가 .024 낮아지고 특정도에는 변화가 없었다. 7점으로 상향하면 민감도는 .050 낮아지고 특정도는 .067 높아졌고, 8점으로 상향하면 민감도는 .056 낮아지고 특정도는 .13 높아졌으며, 이때의 민감도는 .80, 특정도는 1이었다. 따라서 기준점수를 8점으로 결정하였다.

연구 2. 아동성폭력 피해 가능성 평가기준의 타당화

연구 2에서는 연구 1에서 개발한 아동성폭력 피해 가능성 평가기준을 실제 사례에 적용하여 판별적중률 분석을 통해 이 기준의 타당도를 검증하였다. 이와 함께 가장 널리 활용되고 연구가 활발하게 진행되고 있는 아동진술분석기법인 CBCA와의 분류정확도 비교도 이루어졌다.

방 법

연구대상

2009년 12월부터 2010년 5월까지 전국에 있는 아동성폭력 전담 지원 기관에 방문한 만 2세 이상~13세 미만의 아동을 대상으로 하였다. 총 113건의 사례가 수집되었으며, 정보가 불충분하거나 절차적으로 문제가 있는 사례를 제외하고 최종적으로 분석에 사용된 사례는 모두 101건이었다. 여아가 93명(92.1%), 남아는 8명(7.9%)이었으며, 평균연령은 7.3세(SD= 3.0)였고, 학령전기 48명(47.5%), 그리고 학령기가 53명(52.5%)이었다. 객관적인 증거만 존재하는 아동이 38명(37.6%), 객관증거와 행동정황증거가 함께 존재하는 아동 18명(17.8%), 행동정황 증거만 존재하는 아동 25명(24.8%), 그리고 두 종류의 증거가 모두 존재하지 않는 아동 20명(19.8%)이었다.

측정도구

성폭력 피해가 의심되어 아동성폭력 전담지원 기관에 방문한 아동을 대상으로 연구 1을 통해 마련된 아동성폭력 평가 기록지가 작성되었다. 이 기록지는 아동 및 보호자와의 면담과 아동과의 사건면담 과정에서 기록하도록 되어 있으며, 이를 통해 얻을 수 있는 정보에는 피해아동 및 가해자의 기본적인 인적사항, 사건관련 정보, 피해를 지지하는 의료적 증거, 목격자 증언, 가해자 자백, 물리·신체적 증거, 행동정황증거, 급성스트레스 증상을 포함한 임상증상, 최초로 피해를 보고 혹은 인지하게 된 경위와 최초 폭로유형 및 최초 진술내용, 진술조사 시의 아동 진술내용, 진술 시의 감정반응, 진술신빙성요소가 있다.

연구절차

평가자 교육

전국에 있는 아동성폭력전담지원기관에 근무하는 임상심리전문가들을 대상으로 아동성폭력평가와 관련한 사전교육을 시행하였다. 사전 교육은 하루 8시간에 걸쳐 진술조사 지침서(김태경, 이영호, 2010)를 활용한 진술조사실습(진술조사사례 시청 및 조사연습)과 아동성폭력평가 기록지 작성으로 구성되었다.

자료수집 및 채점

평가자가 아동성폭력평가 기록지와 진술조사 CD 및 녹취록을 우편 혹은 인편으로 연구자에게 보냈으며, 수집된 자료를 토대로 연구자가 피해 가능성을 평가하였다. 분석 과정에서 정보가 모호하거나 누락된 경우 평가자에게 전화로 추가 정보를 수집하였으며, 그럼에도 불구하고 정보가 부족한 경우에는 분석에서 제외하였다. 교육에 참석한 임상심리전문가 11명 중 5명이 자료를 보냈으며, 이들의 아동성폭력 피해자를 대상으로 심리학적 평가를 수행한 경험은 평균 2.6년이었다.

자료의 분석

객관증거나 행동정황증거가 존재하는 아동을 증거존재집단으로, 그리고 두 종류의 증거가 모두 부재한 아동을 증거부재집단으로 분류한 후 연구 1에서 도출된 성폭력 피해 가능성 판단 기준변인에서의 집단 간 차이를 검증하였다. 그런 다음 연구 1에서 도출된 기준점수로 집단을 분류하여 ROC 분석을 통해 민감도와 특정도를 검토하였고, 증거유무로 판단한 결과와 기준점수로 분류한 결과의 일치율(판단적중률)을 *Kappa* 분석을 통해 검토하였다. 이와 함께 CBCA의 분류정확도를 검증하고, 연구 1에서 개발된 평가기준을 통한 분류정확도와 비교하였다.

결 과

성폭력 피해 가능성 관련 변인의 집단 간 차이 검증

연령대 별로 성폭력 피해 가능성 관련 변인의 증거존재 및 증거부재집단의 집단 간 차이를 검증한 결과가 표 10과 11에 제시되어 있

표 10. 성폭력 피해 가능성 관련 변인의 집단 간 차이 검증 : 학령전기

	증거존재집단 <i>n</i> =33		증거부재집단 <i>n</i> =14		<i>F</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
진술감정	1.39	0.86	0.36	0.75	15.29***
진술신빙성	3.39	2.22	0.36	0.75	37.70***
최초진술동기	2.48	0.67	1.14	0.66	39.93***
자극노출시고통감	0.85	0.91	0.21	0.58	5.81*

p*<.05, *p*<.01, ****p*<.001

표 11. 성폭력 피해 가능성 관련 변인의 집단 간 차이 검증 : 학령기

	증거존재집단 n=48		증거부재집단 n=6		F
	M	SD	M	SD	
진술감정	1.67	0.56	0.67	0.52	17.22***
진술신빙성	5.59	2.22	3.67	2.16	4.10*
최초진술동기	2.56	0.71	1.83	0.98	5.15*
자극노출시고통감	1.31	0.88	0.17	0.41	9.80**

* $p < .05$, ** $p < .01$, *** $p < .001$

다. 표에서 볼 수 있듯이, 두 연령대 모두 증거존재집단이 증거부재집단에 비해 진술감정, 진술신빙성요소, 최초진술동기, 및 자극 노출 시의 고통감이 유의하게 높았다.

판별적중률 검증

연구 1에서 도출된 평가기준이 증거존재집단과 증거부재집단을 얼마나 정확하게 예측하는지를 알아보기 위해 Kappa 분석을 실시하였다. 분석 결과는 표 12에 제시되었다. 표에서 볼 수 있듯이, 연구 1에 비해 특정도가 낮아

졌기는 하나 학령전기의 경우 연구 1에서 결정한 기준점수인 6점에서 전체 정확분류율(95.7%)이 가장 높았고 카파계수(.84)도 가장 높았으며, 학령기에서도 연구 1에서 결정한 기준점수인 8점에서 높은 수준의 민감도(.92)와 특정도(.83), 및 분류정확도(90.6%)를 나타내었고, 카파계수도 .61로 비교적 양호하였다.

새로 개발된 평가기준과 CBCA의 분류정확률 비교

이상적으로는 다른 평가기준과의 교차타당

표 12. 기준점수별 판별적중률

연령대	기준점수	증거존재집단분류		증거부재집단분류		전체분류 정확률	Kappa
		정확분류(%)	민감도	정확분류(%)	특정도		
학령전 n=47	≥4	32(97.0)	.97	11(78.6)	.79	91.5	.79
	≥5	32(97.0)	.97	11(78.6)	.79	91.5	.79
	≥6	32(97.0)	.97	12(85.7)	.86	95.7	.84
	≥7	29(87.9)	.88	13(92.9)	.93	89.4	.76
학령기 n=54	≥6	45(93.8)	.94	2(33.0)	.33	87.0	.29
	≥7	45(93.8)	.94	3(50.0)	.50	88.9	.44
	≥8	44(91.7)	.92	5(55.6)	.83	90.1	.62
	≥9	41(85.4)	.85	5(83.3)	.83	85.2	.48

표 13. 연구 1에서 개발된 평가기준과 CBCA 점수의 분류정확률 비교

연령대	평가기준	증거존재집단분류		증거부재집단분류		전체분류 정확률
		빈도	%	빈도	%	
학령전 n=47	신개발기준	33	100	13	92.9	90.6
	CBCA	29	87.9	11	78.6	85.1
학령기 n=54	신개발기준	47	97.9	2	33.3	90.7
	CBCA	48	100	0	0	88.9

도 검증이 필요하나, 본 연구에서는 현실적인 제약으로 인해 판별적중률 분석을 통해 타당도를 검증하였고, 추가적으로 전문가들이 아동의 진술신빙성 평가 도구로 사용하고 있는 CBCA 점수를 통한 집단 분류정확률과 본 연구에서 개발한 평가기준을 통한 분류정확률을 비교·검토하였다. 그 결과는 표 13에 제시되어 있다. 표에서 볼 수 있듯이, 학령전기에서 새로 개발된 평가기준의 분류정확률은 증거존재집단 100%, 증거부재집단 92.9%, 그리고 전체 90.6%였으며, CBCA의 분류정확률은 각각 87.9%와 78.6%, 그리고 85.1%로 나타났다. 학령기에서는 새로 개발된 평가기준의 분류정확률이 증거존재집단 97.9%, 증거부재집단 33.3%, 그리고 전체 90.7%였으며, 증거부재집단은 각각 100%, 0%, 그리고 88.9%로 나타났다.

논 의

본 연구에서는 우리나라에서 아동성폭력평가와 관련한 전문가 자문에 대한 요구가 급증하고 있으나 아직까지 관련 전문지식과 경험을 갖춘 전문가가 매우 부족하며, 이를 빌미로 타당하지 못한 절차와 기준을 적용하여 자

격 조건을 갖추지 않은 사람들이 아동성폭력 피해 가능성을 판단하고 있는 현실에 문제를 제기하고, 선행연구자와 관련 전문가들이 공통적으로 사용하고 있는 아동성폭력 피해 가능성 평가기준을 성폭력 피해 의심 사례에 적용하여 자료를 분석함으로써 우리문화와 실정에 맞는 성폭력 피해 가능성 평가기준을 개발하고 타당화하는 것을 목적으로 하였다.

연구 1은 개발단계로, 성폭력 피해가 의심되어 전문기관에 내방한 307건의 사례를 의료적 증거, 목격자 증거, 가해자 자백, 및 물리·신체적 증거를 포함하는 객관적 증거나 아동의 피해 진술과 행동증상 및 정황증거를 포함하는 행동정황증거가 존재하는 집단 및 이러한 증거가 존재하지 않는 증거부재집단으로 구분하여 관련 전문가들이 성폭력 피해의 가능성 평가 시에 중요하게 고려하는 변인들에서의 집단 간 차이를 검증하고, 그 결과를 토대로 증거가 존재하는 피해 명백 집단과 증거부재집단을 변별해주는 예측변인을 탐색하고 변별변인 및 변별 기준점수를 도출하여 아동성폭력 피해가능성 평가기준을 개발하였다.

집단 간 차이 검증 결과, 증거가 존재하는 집단이 증거부재집단에 비해 진술감정과 진술신빙성요소 및 최초진술동기를 포함한 진술관련 요소에서 유의하게 높은 점수를 나타내었

다. 급성스트레스 증상에서는 행동정황증거가 존재하는 두 집단이 증거부재집단에 비해 유의하게 높은 점수를 나타내었으며, 객관증거만 존재하는 집단은 이들의 중간에 위치하였다. 그 밖에, 관련 전문가나 선행연구자들이 사용하고 있는 행동증거에 포함되는 우울이나 자살사고 등을 포함하는 임상증상들은 성폭력 피해 가능성 판단을 위한 독립적인 지표가 되기 어려운 것으로 나타났다.

연령에 부적절한 성행동 증가에서 행동정황증거가 존재하는 두 집단이 나머지 두 집단에 비해 유의하게 ‘낮은’ 점수를 나타내어 주목할 만하였다. 사건관련 변인의 집단 간 차이를 함께 감안해볼 때, 이러한 결과는 연령에 부적절한 성행동의 증가는 아는 사람으로 부터의 지속 피해의 후유증일 수 있으며, 특히 학령기에서는 성행동 증가가 의미 있는 성폭력 피해의 지표일 가능성을 시사하는 것일 수 있다. 다만, Friedrich 등(1992)이 지적한 바 있듯이 이 지표를 활용할 때는 반드시 가족의 습관과 섹슈얼리티, 그리고 최근에 경험한 스트레스 사건의 종류와 강도를 충분히 고려할 필요가 있어 보이며, 사례의 특성을 충분히 고려하는 것 역시 중요해 보인다.

연령은 거의 모든 변인과 유의한 상관관계를 드러내었으며, 학령전기와 학령기로 구분하였을 때 사뭇 다른 양상을 드러내었다. 따라서, 학령전후로 구분하여 집단 변별변인을 탐색한 결과 학령전기의 변별변인은 진술감정, 진술신빙성요소, 최초진술동기, 그리고 성폭력과 관련된 자극 노출 시의 주관적 고통감을 포함한 회귀모델이 도출되었다. 이 연령대의 가중배점 모델은 [진술감정×2 + 진술신빙성요소 + 최초진술동기점수 + 자극노출 시 주관적 고통감×2] 이었고, 변별 기준점수는 6점

으로 결정되었다. 학령기에서는 진술감정과 진술신빙성요소만이 유의한 변별변인으로 도출되었으며, 이 연령대의 가중배점 모델은 [진술감정×4 + 진술신빙성요소] 였고, 변별 기준점수는 8점이었다.

이 처럼 연령대에 따라서 상이한 변별변인이 도출되었으며 공통적으로 포함된 변인들의 가중치도 연령에 따라 달라졌다는 것은 연령을 고려하지 않고 아동성폭력 피해가능성을 평가하는 것이 지닌 위험성을 반영한다. 학령전기에서만 최초진술동기가 유의한 변별요인으로 도출되었는데, 이는 어린 아동이 지닌 유도나 암시의 취약성과 관련하여 이해될 필요가 있다. 선행연구(예를 들어, Cassel & Bjorkland, 1995; Ceci, Toglia, & Ross, 1990; Fivush, Peterson, & Schwarzmuller, 2002/2009)에 따르면, 성인의 암시나 유도, 특히 그것이 반복될 경우 나이 어린 아동은 성인의 질문에 담긴 암시정보를 자신의 실제기억으로 받아들이며 그 결과 허위 기억을 발달시킬 수 있다.

물론, 어린 아동의 성폭력 사건의 특성 상 나이 든 아동에 비해 객관적 증거가 더 드물게 나타난다는 점에서 이 연령대에서 피해아동이 비피해 아동으로 잘못 분류되었을 가능성이 있기는 하지만(Hewitt, 1999), 증거가 부재한 나이 어린 아동 중에 최초로 보호자의 유도질문에 대한 반응으로 성폭력 피해를 보고한 아동의 비율이 무려 74.5%로 매우 높게 나타났던 점을 감안해볼 때, 나이 어린 아동의 경우 성폭력 피해를 입지 않았음에도 불구하고 보호자의 유도질문과정에서 피해 가능성이 부당하게 제기되는 사례의 비율 *false allegation rate*이 나이 든 아동에 비해 훨씬 높을 가능성은 분명 존재한다고 볼 수 있다. 따라서, 나이 어린 아동의 성폭력 피해 가능성 평가 시에는 최초진

술 시의 정황 및 아동과 청취자 간에 나눈 대화의 내용과 형식 등을 면밀히 고려할 필요가 있다. 이 같은 맥락에서 볼 때 본 연구에서 학령전기에서 최초진술동기가 유의한 변별요인으로 도출된 것은 당연한 결과로 보인다. 하지만, 학령기에서도 최초진술동기의 집단 간 차이가 유의하였으므로 이들의 피해 가능성 평가 시에도 최초진술동기를 중요하게 고려할 가치가 있다.

성폭력 사건과 관련된 자극 노출 시의 주관적 고통감이 학령전기에서만 유의한 변별요인으로 나타났다. 두 연령군에서 공통되게 자극 노출 시의 고통감을 포함한 급성스트레스증상에서 유의한 집단 간 차이를 보였으나, 학령기에서는 진술감정과 진술신빙성요소의 설명력이 워낙 크고 자극노출 시의 고통감이 진술감정과 더 높은 공선성을 지니고 있기 때문에 변별변인에서 제외된 것으로 보인다. 따라서, 아동성폭력 피해 가능성 평가의 실제에서는 자극 노출 시의 주관적 고통감 이외의 급성스트레스 증상들을 면밀히 탐색하여 최종 판단 시 고려할 필요가 있다.

두 연령대에서 진술감정과 진술신빙성요소가 공통적인 변별요인으로 도출되기는 하였으나, 상대적 기여도에 차이가 있었다. 이러한 결과는 어쩌면 당연한 것으로, 어린 아동의 경우 나이든 아동에 비해 언어적 유창성이 부족하기 때문에 진술신빙성요소에서 상대적으로 낮은 점수를 얻을 수밖에 없다. 그렇기 때문에 나이든 아동과 같은 비중으로 진술신빙성요소를 고려하는 경우 오분류, 특히 오부정 오류의 가능성이 높아질 수 있다. 실제로 본 연구에서 학령전기에 비해 학령기에서 진술신빙성요소가 유의하게 높았다. 또한, 나이 어린 아동들은 성폭력을 놀이로 인식하여 고통감을

경험하지 않는 경우가 많으며, 이를 반영하듯 나이든 아동에 비해 진술 중에 드러내는 감정 반응이 유의하게 적었다. 이러한 이유로 어린 아동에서 나이든 아동들에 비해 진술감정의 상대적 변별가가 낮아졌을 수 있다.

비록 빈도는 낮으나 성적 불쾌감과 독특한 성적인 내용의 자발적이고 반복적인 언급은 증거존재집단에서만 나타났다. 따라서, 아동성폭력 피해의 가능성 평가 시에 이러한 증상을 면밀하게 평가하여 최종판단 시 고려할 필요가 있다. 하지만, 몇몇 전문가나 연구자들(예를 들어, Adams, 2004; 조은경, 2004)과 같이 임상증상을 독립적인 피해의 증거로 사용하는 것은 오긍정 오류의 위험이 높을 수 있음을 유념해야하며, 증상의 발현이나 악화에 영향력을 행사하였을 수 있는 성폭력 이외의 변인들에 대한 탐색이 반드시 이루어질 필요가 있다.

한편, 변별변인을 도출하기 위한 로지스틱 회귀분석 과정에서 학령기의 분류정확률이 40%에 불과하였다. 증거부재사례 중 절반 이상이 낮은 사람으로 부터의 성폭력 피해를 ‘자발적’으로 폭로하였으며, 피해에 대해 상당히 자세하고 구체적으로 진술하였던 점에서, 이러한 결과는 본 연구에서 성폭력 피해를 입은 다수의 아동을 피해 가능성이 낮은 증거부재집단으로 잘못 분류됨으로써 초래된 것일 가능성이 있다. 따라서 관련 전문가들(예를 들어, Steller & Koehnken, 1989; Craig et al., 1999)이 제안한 바와 같이, 아동-특히 나이든 아동의 경우 다른 지표들은 상승하지 않으나 아동이 피해에 대해 구체적이고 명료하게 진술한다면 피해의 가능성을 면밀히 고려할 필요가 있어 보인다. 물론, 학령기 아동 중 증거부재 집단으로 분류된 사례가 15명으로 매우 적기

때문에 이러한 해석을 위해서는 좀 더 많은 사례를 대상으로 한 진술내용분석 연구가 필요하겠다.

진술내용요소뿐 아니라 진술 과정에서 관찰된 감정반응과 임상요소, 그 중에서도 사건과 직접적으로 관련이 있다고 볼 수 있는 급성스트레스 관련 증상, 그리고 최초폭로유형과 같은 변인들이 변별요인으로 도출되었다는 점은 아동성폭력 피해 가능성 평가 영역에서 시사하는 바가 크다. 최근 들어 우리나라에서도 진술내용분석에 상당히 많은 관심이 쏠리고 있는데, 본 연구 결과를 토대로 해서 볼 때 피해가능성 평가를 위해서는 진술내용에 대한 기계적인 분석을 넘어서 아동과 아동이 속한 환경에 대한 이해가 필수적으로 요구되며, 특히 임상평가가 진술내용분석 못지않게 중요한 것으로 보인다. 이와 관련하여 Meyer 등(2001)은 진술내용에만 전적으로 의존하여 피해 가능성을 판단하는 평가자들은 자신들의 의뢰인을 충분히 이해하지 못하는 경향이 있다고 지적한 바 있다.

한편, 피해가 명백한 사례에서 아동이 피해를 부인하거나 이전 진술을 취소하거나 혹은 매우 유도적인 질문이 있는 후에야 비로소 피해를 보고하는 경우가 분명 존재하며, 반대로 피해를 입증할 만한 객관적 혹은 행동정황증거는 부재하나 피해를 구체적이고 명료하게 진술하여 피해의 가능성을 배제하기 곤란한 사례도 존재한다. 증거존재 사례 중 16.9%가 누군가의 유도질문이 있는 후에야 비로소 피해를 보고하였으며, 그 비율이 학령전기에서 두 배 이상 높았다(학령전기 27.0%, 학령기 12.3%). 이러한 사례들 중 87.5%가 지목된 피해자가 친밀한 대상이었으며 피해가 2회 이상이었음을 감안해볼 때, 친밀한 대상으로부터

성폭력 피해를 입은 경우에 피해의 폭로가 더욱 빈번하게 지연되거나 폭로되지 못한 채 비밀로 남겨지는 경향이 강한 것으로 보인다. 이러한 결과는 선행연구(예를 들어, Cederborg et al., 2007; Goodman-Brown et al., 2003)와 일관된 것으로, 특히 어린 아동에게서 그러한 경향이 더욱 강한 것 같으며, 이는 어린 아동의 성폭력 피해입증 곤란을 더욱 가중시키는 요인으로 작용할 수 있다.

또한, 일반적이지는 않으나 Summit(1983, 1992)의 주장과 일치하게 친족성폭력 사례에서 종종 증거가 명백함에도 불구하고 아동이 피해를 부인하는 사례를 만날 수 있으며, 가족 이외라고 하더라도 평소 애정 어린 관계를 유지하였던 대상으로부터의 성폭력 피해 사례에서도 진술취소나 폭로의 지연 등이 자주 관찰된다(김태경, 이영호, 2010; Lippert, Cross, Jones, & Walsh, 2009). 그렇기 때문에, 특정 아동의 성폭력피해 가능성을 평가할 때에는 시점에 따른 아동의 진술 변화 추이와 진술특성, 아동과 보호자 요인, 조사 요인 등을 면밀하게 검토하여 최종 판단을 내려야 할 것이다.

타당화 단계인 연구 2에서는 성폭력 피해가 의심되어 전문기관에 방문한 아동들을 성폭력 피해가 명백한 집단과 지지증거가 전무한 증거부재집단으로 구분한 후 연구 1에서 개발한 평가기준이 이 두 집단을 얼마나 정확하게 변별해주는지 분석하였다. 분석 결과, 두 연령대에서 모두 연구 1을 통해 개발된 평가기준이 성폭력 피해 의심아동을 90% 이상 정확하게 분류하였으며, 연구 1에서 제안한 기준점수에서 가장 높은 분류정확률과 민감도와 특정도 및 카파 계수를 나타내었다. 또한, 새로 개발된 평가기준의 분류정확률이 CBCA 점수를 통한 집단 분류 정확률에 비해 더 높았다. 이러

한 결과는 연구 1에서 개발된 평가기준이 타당한 아동성폭력 피해 가능성 평가기준이며, 널리 활용되고 있는 CBCA 보다 더 분류정확도가 높음을 반영한다. 이와 함께, 새로 개발된 평가기준이 채점이 용이하고 간편하며 (Mazzoni & Ambrosio, 2003 참조), 아동임상경험을 가진 사람이라면 일정 기간의 교육과 지도 감독 후에 사용이 가능할 수 있음을 감안할 때, 현장에서의 활용도가 CBCA에 비해 더 높을 것으로 기대된다.

끝으로, 본 연구의 제한점 및 추후 연구를 위한 제언을 하자면 다음과 같다. 첫째, 거의 모든 현장연구가 그러하듯이 본 연구에서 증거부재집단으로 분류된 아동이 정말로 피해를 입지 않은 아동이라고 보는 것은 불가능하다. 이 때문에 증거부재집단에 실제 피해아동이 포함되었을 가능성이 있으므로 실제 사례에 본 연구결과에서 도출된 기준을 적용할 때는 아동이 속한 집단의 기저율과 가용한 정보들을 충분히 검토할 필요가 있다. 둘째, 사건관련 변인에서 집단 간 차이가 나타났음에도 불구하고 이러한 변인들을 모두 통제하기에는 표집수가 충분치 못하여 가장 영향력이 컸던 연령만 통제하였다. 따라서 이 평가기준을 적용할 때는 사건관련 변인을 충분히 고려해야만 할 것이다. 셋째, 학령기의 증거부재 사례가 15건으로 적었기 때문에 결과가 편향되었을 가능성이 있으므로 향후 학령기 증거부재 사례를 충분히 표집하여 재분석해볼 필요가 있다. 마지막으로, 본 연구에서는 질적인 분석을 요하는 정보의 타당도와 관련된 요소들이 포함되지 않았다. 정보의 타당도와 관련된 요소에는 보호자와 아동의 허위진술동기, 이차적 이득의 존재유무, 아동의 피암시성 수준, 애착유형, 쟁점이 되고 있는 성폭력 사건 이

외의 스트레스 사건 유무와 그 영향력, 사건에 대한 보호자의 태도 등을 포함한 다양한 변인들이 포함되며, 이러한 정보에 대한 평가는 상당한 경험적 지식이 필요하다. 본 연구를 통해 도출된 기준점수를 실제 사례에 적용하여 성폭력 피해의 가능성을 판단하고자 할 때에는 반드시 수집된 정보의 타당도를 면밀하게 검토해야 한다.

국내외에서 변별변인 및 기준점수를 찾고자 한 선행연구가 없어 비교가 사실상 불가능하지만, 위와 같은 제한점에도 불구하고 아동성폭력 피해 가능성 판단을 위한 기준을 개발하기 위한 국내 최초의 시도라는 점에서 본 연구가 지니는 의의는 상당하다. 본 연구를 통해 개발된 피해 가능성 평가기준은 관련 연구나 전문가 양성이 거의 이루어지지 않아 아동성폭력 평가 영역에서 황무지나 마차가지인 우리나라에서 유용하게 사용될 것이라 기대된다.

참고문헌

- 곽금주, 김연수 (2003). 취학전 아동증언에서 참여여부, 질문 및 질문자의 특성에 따른 기억의 정확성. *한국심리학회지: 발달*, 16, 1-19.
- 곽금주, 이승진 (2006). 아동증언에 영향을 주는 요인들. *한국심리학회지: 일반*, 25, 13-40.
- 김재연, 이재연 (2000). 유아증언의 신뢰성연구. *아동학회지*, 21(3), 54-68.
- 김태경, 이영호 (2010). 아동진술조사지침서: 성폭력피해아동의 조사적 면담을 중심으로. 서울: 두감람나무.

- 김태경, 이영호 (2010). 성폭력 피해 아동의 진술양상. *한국심리학회지: 임상*, 29, 301-325
- 김태경, 최지영, 임자영, 윤소미, 신의진 (2006). 성폭력피해 아동의 로샤반응 특성: 로샤검사가 성폭력 피해의 증거를 제공해주는가? *한국심리학회지: 임상*, 25, 765-779.
- 노일석 (2009). 청소년 절도사범 재범예측요인: 절도 소년보호 관찰대상자 재범위험성 평가도구(LJP-RRAR) 개발연구. *한국심리학회지: 일반*, 28, 449-470.
- 박자경 (1998). 유도질문이 아동진술에 미치는 영향. *충북대학교 석사학위 청구논문*.
- 신의진, 김태경, 최경숙 (2008). 성폭력피해 아동의 진술에 영향을 미치는 요인조사. *국회 법제사법위원회*.
- 이재연 (1998). 아동증언의 신뢰성판단에 대한 발달적 접근. *아동권리연구*, 2(2), 75-90.
- 조은경 (2004). 성폭력피해아동 진술에 대한 증거기반 내용분석의 타당화를 위한 연구. *한국형사정책연구원*.
- Adams, J. A. (2004). In training, medical evaluation of suspected child sexual abuse. *Journal Pediatric Adolescent Gynecology*, 17, 191-197.
- Anson, D. A., Golding, S. L., Gully, K. J. (1993) Child sexual abuse allegations: Reliability of criteria-based content analysis. *Law and Human Behavior*, 17, 331 - 341.
- Bays, J., & Chadwick, D. (1993). Medical diagnosis of the sexually abused child. *Child Abuse & Neglect*, 17, 91-110.
- Blandon-Gitlin, I., Pezdek, K., Rogers, M. & Brodie, L. (2005). Detecting deception in children: An experimental study of the effect of event familiarity on CBCA ratings, *Law & Human Behavior*.
- Blandon-Gitlin, I., Pezdek, K., Lindsay, D. S., & Hagen, L. (2008). *Criteria-based content analysis of true and suggested accounts of event*. Manuscript submitted for publication.
- Briere, J. (2005). *Trauma Symptom Checklist for Young Children: Professional Manual*. Odessa, FL: Psychological Assessment Resources, Inc.
- Cassel, W. S., & Bjorkland, D. F. (1995). Developmental patterns of eyewitness memory and suggestibility: An ecologically based short-term longitudinal study. *Law and Human Behavior*, 19, 507-532.
- Ceci, S. J., Toglia, M. P., & Ross, D. F. (1990). The suggestibility of preschooler's recollections: Historical perspectives on current problems. In R. Fivush & J. A. Hudon (Eds.), *Knowing and remembering in young children* (pp. 285-300). New York: Cambridge University Press.
- Cederborg, A. C., Lamb, M., & Laurell, O. (2007). Delay of disclosure, minimization and denial when the evidence is unambiguous. A multi victim case. In M. Pipe, M. Lamb, Y. Orbach, and A. C. Cederborg (Eds.). *Child sexual abuse: Disclosure, Delay and Denial* (pp. 159-174). Mahwah, NJ: Erlbaum.
- Conte, J. R., Sorenson, E., Fogarty, L., Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *American Journal of Orthopsychiatry*, 61, 428-437.
- Corwin, D. L. (1988). Early diagnosis of child sexual abuse: Diminishing the lasting effects. In G. Wyatt & G. Powell(Eds). *Lasting effects*

- of child sexual abuse(pp. 251-269). Newbury Park. CA:Sage.
- Craig, R. A., Scheibe, R., Raskin, D. C., Kircher, J. C., & Dodd, D. (1999). Interviewer questions and content analysis of children's statements of sexual abuse. *Applied Developmental Science, 3*, 77-85.
- DeYoung, M. (1986). A conceptual model for judging the truthfulness of young child's allegation of sexual abuse. *American Journal of Orthopsychiatry, 56*, 550-559
- Eher, R., Rettenberger, M., Schilling, F., & Pfafflin, F. (2008). Failure of static-99 and SORAG to predict relevant reoffense categories in relevant sexual offender subtype: A prospective study. *Sexual Offender Treatment, 3*, 132-145.
- Elliott, D. M., & Briere, J. (1994). Forensic sexual abuse evaluations of older children: Disclosures and symptomatology. *Behavioral Sciences and the Law, 12*, 261-277.
- Faller, K. C. (2007). *Interviewing children about sexual abuse: controversies and best practice*. NY: Oxford University Press.
- Fivush, R., Peterson, C., & Schwarzmuller, A. (2009). 아동진술의 신빙성: 질문기법의 맥락에서 (김태경과 윤소미 역). 진술조사의 맥락에서 본 기억과 피암시성(pp. 373-400). 서울: 시그마프레스(원전은 2002에 출판).
- Friedrich, W. N. (1997). *Child Sexual Behavior Inventory: Professional Manual*. Odessa, FL:Psychological Assessment Resources.
- Friedrich, W. N., Grambsch, P., Damon, L. (1992) Child Sexual Behavior Inventory. Normative and clinical comparisons. *Psychological Assessment, 4*, 303-311.
- Goodman-Brown, T. B., Edelstein, R. S., Goodman, G. S., Jones, D. P. H., & Gordon, D. S. (2003). Why children tell: A model of children's disclosure of sexual abuse. *Child Abuse & Neglect, 27*, 525-540.
- Greenhoot, A. F. (2000). Remembering and understanding: The effects of changes in underlying knowledge on children's recollections. *Child Development, 71*, 1309-1328.
- Haskett, M. E., Wayland, K., Hutcheson, J. S., & Tavana, T. (1995). Substantiation of sexual abuse allegations: Factors involved in the decision-making process. *Journal of Child Sexual Abuse, 4*, 19-47.
- Heiman, M. (1992). Annotation. Putting the puzzle together: validating allegations of child sexual abuse. *Journal of Child Psychology and Psychiatry, 33*, 311 - 329.
- Hewitt, S. K. (1999). Protecting our children from sexual abuse : Why our current system is not working childhood sexual behavior. In C. T. Hendrix & W. Westgate (Eds.). *Assessing allegations of sexual abuse In preschool children : Understanding small voices*. Thousand Oaks:Sage.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (2009). Source monitoring. *Psychol Bull, 135*, 638-677.
- Jones, D. P., & McGraw, J. M. (1987). Reliable and fictitious accounts of sexual abuse to children. *Journal of Interpersonal Violence, 2*, 27-45.
- Kim, Y. S., Cheon, K. A., Kim, B. N., Chang, S.

- A., Yoo, H. J., Kim, J. W., Cho, S. C., Seo, D. H., Bae, M. O., So, Y. K., Noh, J. S., Koh, Y. J., McBurnett, K., Leventhal, B. (2004). The reliability and validity of Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version-Korean Version(K-SADS-PL-K). *Yonsei Med*, 45, 81-89.
- Kuehnle, K., & Connell, M. (2009). *The Evaluation of Child Sexual Abuse Allegations: A Comprehensive Guide to Assessment and Testimony*. NJ:Wiley & Sons, Inc.
- Lamb, M. E., Orbach, Y., Hershkowitz, I., Explin, P. W., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD Investigative Interview Protocol. *Child Abuse & Neglect*, 31, 1201-1231.
- Leitchman, M., Morse, M. B., Dixon, A., & Wilch-Ross, M. (2000). Source monitoring and suggestibility: An individual differences approach. In K. P. Roberts & M. Blades (Eds.), *Chodren's source monitoring* (pp.257-288). Mahcah, NJ: Lawrence Erlbaum Associates.
- Lippert, T., Cross, T. P., Jones, L., & Walsh, W. (2009). Telling interviewers about sexual abuse: Predictors of child disclosure at forensic interviews. *Child Maltreatment*, 14, 100-113.
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, 11, 194-226.
- Mazzoni, G. K., & Ambrosio, L. (2003). analisi del resoconto testimoniale in bambini: impiego del metodo di analisi del contenuto C.B.C.A. in bambini di 7 anni. *Psicologia e Giustizia*. (Analyzing child witness reports: using the C.B.C.A. in 7-years-old children). *Archive*, 1.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., Reed, G. M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *The American psychologist*, 56, 128-65.
- Ornstein, P. A., & Haden, C. A. (2009). 기억의 발달: 아동증언의 이해 (김태경과 윤소미 역). *진술조사의 맥락에서 본 기억과 피암시성* (pp. 33-66). 서울: 시그마프레스. (원전은 2002에 출판)
- Otgaar, H., Candel, I., Merckelbach, H., & Wade, K.A. (2009). Abducted by a UFO: Prevalence information affects young children's false memories for an implausible event(pp. 115-125). *Lire le compte rendu sur PsychoTémoins*.
- Poole, D. A., & Lindsay, D. S. (1998). Assessing the accuracy of young children's reports: Lessons from the investigation of child sexual abuse. *Applied & Preventive Psychology*, 9, 1-26.
- Sgroi, S., Blick, L., & Porter, F. (1982). A conceptual framework for child sexual abuse. In S. Sgroi (Ed.), *Handbook of clinical intervention in child sexual abuse* (pp. 9 - 38). Lexington, Massachusetts: D. C. Heath and Company.

- Sink, E. (1988b). Studies of true and false allegations: A critical review. In E. B. Nicholson (Ed.), *Sexual abuse allegation custody and visitation cases* (pp. 37-47). Washington, DC: American Bar Association.
- Sink, F. (1988a). A hierarchical model for evaluation of child sexual abuse. *American Journal of Orthopsychiatry*, 58, 129-135.
- Sporer, S. L. (1997). The less traveled road to truth: Verbal cues in detection deception in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology*, 11, 373-397.
- Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217-245). New York: Springer.
- Summit, R. C. (1983). The child sexual abuse accommodation syndrome. *Child Abuse & Neglect*, 7, 177-193.
- Summit, R. C. (1992). Abuse of the Child Sexual Abuse Accommodation Syndrome. *Journal of Child Sexual Abuse*, 1, 153-163.
- Undeutsch, U. (1984). Courtroom evaluation of eyewitness testimony. *International Review of Applied Psychology*, 13, 51-67.
- Vrij, A. (2005). Criteria-based content analysis: A qualitative review of the first 37 studies. *Psychology, Public Policy, and Law*, 11, 3-41.
- Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2004). Let me inform you how to tell a convincing story: CBCA and reality monitoring scores as a function of age, coaching, and deception. *Canadian Journal of Behavioral Science*, 24, 239-263.
- Vrij, A., Edward, K., Roberts, K., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behaviour. *Journal of Nonverbal Behavior*, 24, 239-263.
- Wehrspann, W. H., Steinhauer, P., & Klajner-Diamond, H. (1987). Criteria and methodology for assessing credibility of sexual abuse allegation. *Canadian Journal of Psychiatry*, 32, 615-623.
- Wells, G. L., & Loftus, E. F. (1991). Is this child fabricating? Reactions to a new assessment technique. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 168-171). Washington, DC: American Psychological Association.
- Wharff, E. A. (1998). A study of decision-making criteria in child sexual abuse evaluations. *Dissertation Abstracts International. Section A: Humanities & Social Sciences*, 59(1-A), 0326

원고접수일자 : 2010. 7. 10.

게재확정일자 : 2010. 9. 3.

Development and Validation of Assessment Standard for Determining Probability of Child Sexual Abuse

Tae-Kyoung Kim

Young-Ho Lee

Department of Psychology, the Catholic University of Korea

The aims of this study was to develop and validate assessment standard for determining probability of child sexual abuse. In study 1, assessment standard was developed. For this aims, 307 alleged sexual abuse cases which were composed of 2~12 years old were assigned to one of objective evidence group, objective evidence and behavior · context evidence group, behavior · context evidence group, and no-evidence group. Emotional response(congruous emotional response to statement content), statement credibility elements(quantity of details, contextual details, interactions between victim-perpetrator, unusual or specific or declare details), first disclosure type, and clinical symptom(ex. reexperience, hyper-arousal, avoidance, depression, anxiety, increased sexual behavior) were included in dependent variables. These variables were commonly used by experts to assessment the possibility of child sexual abuse. In result of group difference analysis, the groups that objective or behavior · context evidence existed were higher than no-evidence group in emotional response, statement credibility elements, and first disclosure type. Among clinical symptoms, only in acute stress reaction symptom both behavior · context evidence exist group were higher than no evidence group, and objective group was located between behavior · context group and no-evidence group. In age-inappropriate sexual behavior, objective group and no evidence group were higher than both behavior · context evidence exist group. In result of binary logistic regression for discriminant variable to evidence group versus no-evidence group, weighted model was [emotional response score×2 + statement credibility score + first disclosure score + subjective distress level during exposure to stimuli related to sexual abuse event×2] for preschool age, and cutoff point was determined score 6. Weighted model was [emotional response score×4 + statement credibility score] , and cutoff point was determined score 8 for school age. The aims of study 2 validated new assessment standard. For this aims, 101 alleged sexual abuse cases were classified evidence group versus no-evidence group. And the hit ratio were analysis between evidence based group classification and new assessment standard based group classification. In result, in both age group the hit ratios were more than 90%, and *Kappa* coefficients were relatively high. And, New assessment standard was revealed higher correct classify rate than CBCA. Consequently, new developed assessment standard for determining probability of child sexual abuse was validated. Implications of these findings were discussed.

Key words : child sexual abuse, assessment standard, determining probability, validation, CBCA