

성폭력위험성평가의 진화: 근거기반 성폭력위험성프로토콜(RSVP)을 중심으로

서종한*·김희송**·김경일***

국 | 문 | 요 | 약

성폭력위험성프로토콜(Risk for Sexual Violence Protocol: 이하, RSVP)은 근거기반(Evidence-based) 구조화된전문가판단(Structured Professional Judgement: 이하, SPJ) 모델을 기초로 하여 만들어진 위험성평가도구이다. 북미와 유럽에서 활용되는 대표적 도구로 국내에서는 소개된 바가 없다. 기존 계리[혹은 통계]적 위험성평가(actuarial risk assessment: 이하, 'ARIAs') 도구와는 달리 임상가의 전문적인 의사결정 판단을 돕고 향후 범죄 예방을 위해 치료와 관리계획에 주요 초점을 두고 있다. 국내·외의 문헌과 학자들은 통상 3세대 혹은 4세대 모델로 평가 하고 있으며 범죄 예방의 효율성과 효과성을 높일 수 있는 접근 혹은 대안적 모델로 고려하고 있다. 본 연구의 목적은 국내에서는 생소한 근거기반 SPJ 개념을 소개하고 대표적인 도구인 RSVP를 국내에 타당화 하기 위해 전국 6개 교도소에 수용중인 재소자 50명을 대상으로 직접조사 방식으로 채점자간 신뢰도와 공인타당도를 확인하는 데 있다. 채점자간 신뢰도를 확인 하기 위해 평균 4명 이상의 훈련받은 채점자가 각 사례에 동시에 참여하였고 이 후 논의과정을 거친 후 독립적으로 개별 채점이 이루어졌다. 연구결과, 전체적으로 모든 영역과 단계에서 높은 신뢰도를 보였지만 사례관리(단계5)와 전문가판단요약(단계6)은 다른 단계에 비해 상대적으로 낮았다. 국내에서 대표적으로 사용하고 있는 Static-99R과 한국성폭력위험성평가(HAGURSOR) 간 공인타당도를 살펴 본 결과, 전체적으로 위험요인들 간에 높은 상관을 보였고 특히 국내에서 제작된 도구와는 단계 5와 6에서도 부분적으로 적정수준의 상관을 보였다. 전체적으로 본 연구를 통해 국내에서의 활용 가능성을 살펴볼 수 있었고 향후 예측 타당도 확인이 필요할 것으로 보인다.

- ❖ 주제어 : 근거기반, 구조화된전문가판단모델, 성폭력위험성평가, 성범죄자, 공인타당도, 채점자간 신뢰도

* 사이먼프레이저대학교 심리학과 법정신건강정책연구소 수석연구원

** 국립과학수사연구원 법심리과 과장

*** 교신저자: 아주대학교 심리학과 교수

I. 서론

이수정 (2010, p. 123¹⁾)의 정의에 따르면 위험성은 “범죄를 저지를 가능성 또는 재범 가능성으로 범죄자가 미래에 사회로 복귀했을 때 다시 범죄를 저지를지 평가하여 궁극적으로 다른 사회 구성원을 보호하고 범죄자를 효율적으로 통제하기 위한 것”으로 기술하였다. 한국에서 범죄 예측을 위해 사용되고 있는 대표적인 계리적 위험성평가 도구(actuarial instrument: 이하, ‘ARIAs’)는 한국성범죄재범위험성평가(KSORAS)²⁾, 한국판 사이코패스진단검사(the Korean version of PCL-R)³⁾ 등이다. 이런 도구의 최종 목적은 재범률 예측(prediction)에 기반을 두었고 형사사법체계에 오랜 기간 그 기능을 하고 있다. 하지만 예측이라는 패러다임은 국·내외적으로 범죄자의 범죄를 예방하고 사회 구성원을 적극적으로 보호하는 차원에서 효용성, 유용성, 효과성에 한계점을 드러내며 새로운 형태의 위험성평가 패러다임을 요구 받아오고 있다. 왜냐하면 계리적 평가도구는 향후 범죄자에게 임상가가 어떤 관리용이성과 계획을 갖고 다루어야 할지 어떤 부분에 초점을 맞춰 치료와 개입을 선택적으로 해야 할지에 대한 정보를 주지 못하고 있기 때문이다. Hart (2009)⁴⁾는 이런 이유로 인해 현재를 위험성 평가의 새로운 전환점이 대두되는 시대라고 규정한 바 있다

Desmarais와 Singh (2013)⁵⁾에 따르면 4세대 평가 방식을 제시하며 단순히 예측이라는 목적에서 벗어나 이제는 관리계획이라는 최종 지향점을 추구하는 것이라고 주장한 바 있다. 또한 최근 Monahan과 Skeem (2011, 2014)⁶⁾은 평가의 목적에 따

1) 이수정 (2010, p. 123). 최신 범죄 심리학. 서울: 북카페.

2) 이수정·고려진, 한국 성범죄자 위험성 평가척도(KSORAS) 실시요강. 서울: 맥스메디카 (2011).

3) 조은경·이수정 (2008). 한국판 표준화. PCL-R 전문가 지침서, (R. D. Hare 저, 조은경, 이수정 역, 2008). 서울: 학지사 심리검사 연구소.

4) Hart, S. D. (2009). Evidence-based assessment of risk for sexual violence. *Chapman Journal of Criminal Justice*, 1, 143-165.

5) Desmarais, S. L., & Singh, J. P. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States*. Washington, DC, Council of State Governments Justice Center.

6) Monahan, J., & Skeem, J. L. (2014). The evolution of violence risk assessment. *CNS Spectrums*, 19, 419-424; Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, 20(1), 38-42.

라 선별적으로 계리적 위험성 평가 도구를 사용해야 할지 임상적 평가 도구를 사용해야 할지를 판단해야 된다고 보았고 특정 목적에 상관없는 계리적 도구만의 사용을 지적하기도 했다. 특히 Hart, Michie와 Cooke⁷⁾ (2007, 2010, 2013)등은 지속적으로 규범정립적 통계가 갖는 오차 부분이 실제 법적 맥락에서 의사결정자로 하여금 잘 못 된 판단을 이끌거나 남용의 가능성을 제기해 오고 있다. 뿐만 Haag (2016)등과 Hart(2016)⁸⁾은 계리적 도구가 인종 등에 대한 문화적 차이 변인을 적절히 반영하지 못해 특정 지역 혹은 인종과 성별에 대한 재범 위험률을 과도하게 추정하고 있다고 보고 하였다. Kroner (2016)에 따르면, 캐나다 BC주 법정에서 원주민 혹은 선주민족(first people, Ewert)이 특정 계리적 도구(예, VRAG, Static-99)에 대한 불만 혹은 회피를 제기하여 소송 중이며 캐나다 대법원에서 이 점을 지적하며 배척 사유가 적절할 경우 계리적 도구 사용에 대한 기피를 할 수 있다고 규정하였다(판례, Ewert v. Canada)⁹⁾. 이런 대안적 모색에서 Douglas (2013)¹⁰⁾ 등은 폭력 위험성 개념화로 이루어진 근거기반 구조화된 전문가 판단 모델을 소개하였다.

최근까지 한국 뿐 아니라 북미 등에서 폭력 위험성 사례 개념화는 법정, 정신건강 등 분야에서 큰 관심을 받지 못했다. 아마도 개념화에 대해 충분히 생각할 수 있는 기반을 제공하지 않는 숫자 및 통계 알고리즘 기반의 통계적 위험성 평가 절차가 1990년대와 2000년대 광범위하게 사용되었기 때문인 것으로 보인다. 이런 흐름은

-
- 7) Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments. *The British Journal of Psychiatry*, 190(49), s60-s65; Cooke, D. J., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and human behavior*, 34(4), 259-274; Hart, S. D., & Cooke, D. J. (2013). Another look at the (Im-) precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences & the Law*, 31(1), 81-102.
- 8) Haag, A. M., Boyes, A., Cheng, J., MacNeil, A., & Wirove, R. (2016). An introduction to the issues of cross-cultural assessment inspired by Ewert v. Canada. *Journal of Threat Assessment and Management*, 3(2), 65; Hart, S. D. (2016). Culture and violence risk assessment: The case of Ewert v. Canada. *Journal of Threat Assessment and Management*, 3(2), 76.
- 9) Kroner, D. G. (2016). The Ewert v. Canada judgment: Moving forward. *Journal of Threat Assessment and Management*, 3(2), 122.
- 10) Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 version 3: assessing risk for violence*. Burnaby, BC, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.

최근 국내 문헌과 연구의 방향을 살펴보면 크게 차이가 없어 보인다.

하지만, 다행스럽게도 최근 폭력위험성평가는 숫자기반의 통계적 혹은 계리적 사고에서 점차 근거기반(evidence-based)의 법-임상적 사고로 이동하려는 움직임을 보이고 있으며, Hart(2009)¹¹⁾는 이러한 변화를 과거의 정형화된 공식(formula)에서 사례 개념화(case formulation)로의 이동이라고 표현한 바 있다. 보다 많은 학술연구와 결과가 발표되면서 근거기반 위험성 평가 절차 개발을 위한 탄탄한 발판이 마련되고 있다. 이러한 폭력 위험성 평가 절차는 법정신건강 환경뿐만 아니라 정신건강, 교정, 사법기관 등 여러 환경에서 적절히 활용할 수 있도록 고안되었다. 계리적 위험성 평가는 과거의 많은 사례에 대한 실제적 경험에 의존하기 때문에 객관성과 타당성을 확보¹²⁾할 수는 있으나 사후 예방을 위한 관리지침과 계획을 세우는데 정보를 줄 수 없어 한계점이 있음을 부인할 수 없다. 또한 이수정(2010), 정지숙(2012) 등과 최금화(2009)의 연구를 보면 국내 학자조차도 HCR-20¹³⁾ 등과 같은 도구를 일률적으로 예측을 위한 재범위험성평가 도구라 지칭하며 그 목적을 제대로 파악하지 못하는 듯 보인다¹⁴⁾. 엄밀히 말하자면 이런 도구는 임상적 평가 도구로 구조화된 판단 모델에 근거하여 제작 된 것이기 때문에 미래에 발생할 수 있는 재범 가능성을 예측할 수 있는 알고리즘, 계리적 산술과 가중치는 전혀 존재하지 않는다. 다시 말해 임상 전문가 즉 평가자의 판단이 사례정보에 근거하고 있고 이를 근거로 전문가가 시나리오와 미래 범죄를 관리하고 예방할 수 있는 소견을 제시할 수 있게 만들어진 근거 기반 평가도구이다. 본 논문의 일차적 목적은 한국 임상법심리학(clinical forensic psychology) 분야에서 전무한 폭력 위험성 법 사례 개념화(forensic case management)에 관한 논의를 촉진하고 실질적인 응용을 증진하는데 도움을 주어 효율적인 범죄 예방(prevention)을 하는 데 그 목적을 두었다.

Sackett과 Rosenberg (1995, p.71)에 따르면 근거기반(evidence-based¹⁵⁾)이란

11) Hart, S. D. 앞의 논문 참조.

12) 이수정·윤옥경. (2003). 범죄위험성의 평가와 활용방안. 한국심리학회지: 일반, 22(2), 99-126.

13) Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. 앞의 매뉴얼 책 참조.

14) 한국에서 소개된 논문 중에 구조화된전문가판단을 계리적 도구와 혼용하여 사용하고 있음. see e.g., 정지숙·신정·이장규. (2012). 재범위험성 평가도구 타당도 검증. 한국심리학회지: 일반, 31(3), 643-665; 최금화 (2009년). 폭력 위험성 측정 도구 (Violence Risk Assessment: HCR-20)의 타당도 연구. 경기대학교 대학원 석사 학위 논문.)

“관찰이나 연구, 통계 또는 잘 타당화 된 이론 형태의 관련 정보를 지침 혹은 바탕으로 하거나, 검토해 도출한 행동이나 결정”을 의미한다. 현재 여러 학문에서 폭넓게 사용하고 있는 이 개념은 의약분야에서는 보편화되어 있으며 “개별 환자 치료에 관한 의사결정 시 최근의 증거를 면밀하고 명시적으로 적절히 사용하는 것”이라는 고전적 정의 형태로 이를 사용하고 있다. 국내에서는 최기홍 (2013) 등¹⁶⁾이 ‘근거기반실천과 심리치료’ 연구에서 처음으로 소개한 바 있는 데 APA(2006, p.278)에서 제시한 “내담자 각각에게 행할 심리 치료를 정하는 의사 결정 과정에서 연구를 통해 근거한 확보된 심리치료를 임상적으로 숙련 된 치료자가 환자의 필요 가치와 선호 등의 맥락을 고려하여 내담자에게 적용하는 것”이라는 정의를 따랐다. 하지만 이 정의는 심리치료와 관련된 광의의 개념이라 볼 수 있기¹⁷⁾때문에 범법심리학적 개념과는 응용적 접근에서 의미하는 바가 다소 차이가 있을 수 있다.

Hart (2009, pp.273-4)에 따르면 근거기반 폭력 위험성 평가는 다음과 같이 정의될 수 있다. “a)향후 대인 폭력 행위 가능성을 파악하거나 (b)폭력 방지에 필요한 조치에 대한 의사결정을 위해 가장 과학적이고 전문적인 지식에 근거하여 사람들에게 대한 정보를 수집하는 과정이다.”¹⁸⁾. 하지만, 원칙적으로 이 정의를 수용하는 이들조차도 실제로 어떻게 이 정의를 적용할지 상당한 이견을 보이고 있다. 일반적 측면에서 근거기반 의사결정 개념은 여러 분야에서 폭넓게 인정을 받으며 수용되고 있는 추세이다. 하지만 세부적으로는 법 정신건강 전문가들 간 상당한 이견을 보이고 있다. 근거기반 의사결정의 기본 정의조차도 의견이 매우 분분하다¹⁹⁾. 뿐만 아니라 근거기

15) Sackett과 Rosenberg (1995)가 *British Medical Journal*에 정의했던 근거기반 의미는 다음과 같다: *an action or decision was guided by, based on, or made after reviewing relevant information in the form of observation, research, statistics, or well-validated theory*; Sackett, D. L., Rosenberg, W. M., Muir Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. it's about integrating individual clinical expertise and the best external evidence. *British Medical Journal*, 312, 71-72.

16) 최기홍, 임민경, 이지혜, 이한나, 김태동. (2013년). 근거기반실천과 심리치료. 한국심리학회지: 일반, 32(1), 251-270.

17) APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-Based Practice in Psychology. *American Psychologist*, 61(4), 271- 285.

18) Hart, S. D. 앞의 논문 참조.

19) Justice, L. (2008). Evidence-based terminology. *American Journal of Speech-Language Pathology*, 17, 324-325; Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three

반 의사결정 원리와 응용에 대해서도 뚜렷하게 의견 상 불일치가 나타나고 있다²⁰⁾.

구조화된 전문가 판단 과정은 광의 및 협의 정의 두 측면에서 모두 근거 기반의 의미를 갖는다. 첫째, 안내지침서는 과학·전문적인 문헌들에 의해 직접적으로 잘 알려진 형식으로 안내되거나 구조화되어 있다. 둘째, 구조화된 전문가 판단 모델을 사용하여 만들어진 폭력 위험성 평가가 ‘신뢰성과 타당성’이 있다는 견해를 뒷받침할 만한 상당한 과학적 증거가 있다²¹⁾. 국내에서는 정지숙 (2012) 등의 연구에서 ‘구조화된 전문적 판단’으로 소개문헌에서 처음으로 언급하였고 3세대 도구로 소개하였으나 실제 외국 문헌에서는²²⁾ 4세대 모델에 가깝다고 보고하고 있다. 사례관리계획(치료, 관리감독, 피해자안전보호조치 등)과 전문요약판단(적극적 개입, 향후 위험성 수준, 재검토 등) 등을 포함하고 있어 향후 범죄자 개입과 처우에 대한 정보를 제공하고 있기 때문이다. 해당 전문가들과의 논의과정을 통해 용어적 정의가 ‘전문적’ 속성의 맥락보다는 전문가 즉 행위 주체자의 속성이 맞다는 의견을 들어 본 연구에서는 ‘구조화된 전문가 판단’이라고 명명하도록 하겠다²³⁾.

구조화된 전문가 판단(*SPJ; structured professional judgement*) 안내지침을 적용하여 만들어진 모델 중 대표적인 검사도구가 성폭력위험성프로토콜(The Risk for Sexual Violence Protocol²⁴⁾: 이하, ‘RSVP’)라고 할 수 있다. 이 접근법은 크게 두 가지 형태로 검사 공식을 개념화하였다. 첫째, 과거 발생한 폭력을 ‘의사결정이론

controversies and a caveat. *Health Affairs*, 24, 163-173; Timmermans, S., & Mauck, A. (2005). The promises and pitfalls of evidence-based medicine. *Health Affairs*, 24, 18-28.

20) Miles, A., Polychronis, A., & Grey, J. E. (2006). The evidence-based health care debate. 2006. Where are we now? *Journal of Evaluation in Clinical Practice*, 12, 239-247.

21) Otto, R. K., & Douglas, K. S. (Eds.). (2011). *Handbook of violence risk assessment*. Routledge.

22) Monahan, J., Steadman, H. J., Appelbaum, P. S., Grisso, T., Mulvey, E. P., Roth, L. H., ... & Silver, E. (2006). The classification of violence risk. *Behavioral Sciences & the Law*, 24(6), 721-730; Otto, R. K., & Douglas, K. S. (Eds.). (2011). *Handbook of violence risk assessment*. Routledge.

23) 본 연구자는 해당 용어의 정확한 개념 정의를 위해 밴쿠버 사이먼프레이저대학교 심리학과에서 Kevin S. Douglas, Stephen D. Hart, Laura Guy, and David Cooke 등과 자세히 논하였음.

24) 국내에서는 아직 소개되지 않았음. see e.g., D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, British Columbia: Mental Health, Law, and Policy Institute, Simon Fraser University.

프레임워크(*decision theory framework*)²⁵⁾을 통해 분석한다. 그리고 미래 폭력에 대해 예측하고 시나리오 계획에 근거하여 사례관리 계획을 수립한다. 국내에서는 구조화된 판단 모델을 접목한 위험성평가도구는 현재까지 타당화 되었거나 이를 목적으로 만들어지지 않고 있다.

의사결정이론 프레임워크에 따르면, 위험성 평가의 중요한 과제는 사람이 폭력을 하기로 결정한 동기와 방식 그리고 이런 의사결정에 영향을 촉진시키거나 방해한 다양한 위험 요인을 파악하는 것이다. 위험요인들은 의사결정에 영향을 주며 몇 가지 인과적 역할을 할 수 있기 때문이다. 이런 위험 요인들은 크게 개인의 의사결정을 ‘동기화(*motivate*)’ 하고 ‘탈억제화(*disinhibit*)’ 시키며 ‘불안정화(*destabilize*)’ 시킨다고 볼 수 있다. 구체적으로 ‘동기 유발 요인’은 범죄자로 하여금 폭력을 통해 얻을 수 있는 ‘보상과 자각된 이윤’을 증가시켜 주는 역할을 한다. ‘탈억제 유발 요인’은 폭력을 통해 치를 수 있는 ‘지각된 대가와 부정적 결과의 심각성을 감소’시키는 기능을 한다. ‘불안정화 유발 요인’은 일반적으로 ‘개인의 인지, 기억, 사고 능력 등을 방해해서 의사결정을 모니터링하거나 통제하지 못하게끔’ 한다. 이런 맥락에서, 의사결정 이론에 따르면 위험성을 관리 할 때 과제는 ‘효과적인 전략’을 수립해서 개인의 의사결정이 ‘친-사회적’으로 이루어지게끔 격려하거나 혹은 반대로 폭력과 관련된 의사결정을 어렵게 만들어 비폭력적인 행동을 하게끔 이끄는 것이다. 결국, RSVP와 같은 SPJ 가이드라인은 과학적이며 전문적인 문헌의 체계적인 검토를 거쳐 일반적으로 중요하게 여겨질 수 있는 위험 요인에 평가자가 주의를 두게끔 이끌어 주거나 직접 의사결정이론을 활용하여 실제 사례에서 위험요인 관련성을 적용하고 이해하는 데 도움을 주도록 하였다.

국내에서 윤정숙과 이수정 (2012)이 기술한 바와 같이, 범죄자가 범죄를 저지를 가능성, 범죄를 발생시키는 범죄자의 욕구, 그리고 치료진행시 개인의 특수성을 고려해야 한다는²⁶⁾ 위험-욕구-반응(*Risk-Need-Responsivity; RNR*²⁷⁾) 접근처럼, SPJ

25) Hart, S. D. 앞의 논문 참조 (pp. 143-165.에서 의사결정프레임워크에 대하여 구체적으로 기술).

26) 윤정숙, 이수정 (2012). 성범죄자 치료 프로그램의 주요 원칙과 방향 정립을 위한 이론적 고찰. 한국심리학회지: 일반, 31(3), 847-875.

27) Andrews, D. A. (2012). *The risk-need-responsivity (RNR) model of correctional assessment and treatment*. Using social science to reduce violent offending, 127-156.

접근법에 있어 개념화는 잘 알려지고 정립된 이론에 근거한 장점이 있다. 적절한 치료가 범죄 가능성에 따라 강도나 기간이 달라지고 범죄 욕구에 치료 내용이 선별적으로 적용되어야 하고 개인의 성향과 능력에 따른 적절한 치료방식을 선택해야 하는 측면은 곧 개별기술적 접근을 지향하고 있고 SPJ 접근도 요인 간 인과적 역할을 개념화하거나 개별 사례의 미래 폭력을 예측, 사례 관리 계획을 개발하는데 그리고 위험요인을 식별하는 과정을 구조화하여 구체적인 지침을 줄 수 있다는 점에서 이와 비슷하다고 볼 수 있다. 하지만 RNR과는 달리 평가자가 초점을 맞추는 위험 요인들은 일반 범죄 행동보다는 폭력과 관련이 있다. 또한 SPJ 접근법이 구조화되었긴 하지만 RNR 접근보다 더 융통성이 있어서 새롭거나 특이한 맥락에서 잘 응용될 수 있다고 본다.

RSVP는 민·형사상 장면에서 다용하게 활용하기 위해 제작된 것으로 위험성 평가(risk assessment)와 관리계획(management plan) 결정에 관련된 전반적인 지침을 제공하여 각 사례별로 미래 계획을 구체화하는 데 적극적인 도움을 줄 수 있다. 첫 번째, 매뉴얼에 제시된 지침사항에 따라 면담을 포함하여 필요한 사례 정보를 수집해야 한다. 두 번째, 평가자는 5개 요인과 22개의 개별 위험 요인 및 추가적으로 확인된 사례별 위험 요인(그 이외의 위험요인)의 존재유무 여부를 코딩한다. 존재 여부 채점은 시기별로는 과거(통상 평가 일을 기준으로 12개월 이전)와 최근(통상 평가 일을 기준으로 당해 12개월 이내)으로 이루어져 있다. 각 기간 별 존재 여부 채점은 3점 척도(없음, 가능성 혹은 부분적, 있음)이고 정보가 완전하지 못할 경우에는 생략 할 수도 있다. 3단계에서 평가자는 개별 위험 요인들의 관련성을 결정하는데, 관련성(relevance)은 평가자가 기능적으로 미래 발생 가능성이 있을 것 같은 성폭력 위험요인이나 이를 방지하기 위해 만든 관리 전략효과를 현저히 손상시킬 수 있는 요인인 경우 해당될 수 있다. 관련성 채점도 3점 순서 척도를 사용하고 있다(관련 없음, 아마도 혹은 부분적으로 관련 있음, 관련 있음). 네 번째 단계, 평가자는 미래 발생 가능성이 있는 성폭력 시나리오(scenarios)를 확인하고 이를 기술한다. 1단계에서 수집한 성폭력 이력에 대한 정보와 2와 3단계에서 존재유무 및 관련성이 있다고 한 위험요인, 그리고 미래 가능 주거환경 등을 종합적으로 고려하여 미래에 발생할 수 있는 일들을 추측한다.

다음으로 가능한 미래(possible futures) 혹은 우려되는 결과(feared outcome)의 기술을 ‘시나리오’라고 하는데, 계획을 촉진시키는 방식으로 복잡한 예측을 단순화시킨 간단한 묘사²⁸⁾라고도 할 수 있다. 평가자는 다양한 시나리오를 적극적으로 만들 필요가 있고 성폭력에 대한 일반 지식과 가까운 사례 정보를 근거로 실현 가능성이 낮은 시나리오는 제거 할 필요가 있다. 성폭력위험성프로토콜에 사용 된 절차는 시나리오 계획 세우기 방법론(scenario planning methodology)²⁹⁾에서 도출되었고 이 방법론을 통해 상당히 불확실한 상황에서 계획을 수립해야 하는 분야에서도 수십년 동안 성공적으로 적용되어 오고 있다.

다섯 번째 단계에서, 평가자는 위험 요인과 시나리오를 고려하여 성폭력 위험성 관리(management)에 필요한 세부적인 전략(strategies)을 짜게 된다. 전략을 도출하는 과정에서는 각 시나리오에 따라 발생할 수 있는 성폭력 가능 여부 뿐 아니라 개별 위험 요인 간 미래 관련성 측면을 동시에 고려해야 한다. 포괄적인 위험 관리 전략을 보장하기 위해서 평가자가 모니터링(monitring), 치료(treatment), 감독(supervision), 그리고 피해자 안전 계획(victim safety planning) 등 네 가지 측면을 적극적으로 고려하도록 하였다. 이 네 가지 측면에서 평가자는 특정 전략을 확인한 후 가까운 시일에 해당 사례에 어떤 방식으로 적용해야할지 구체적이면서도 분명하게 숙고하여 결정해야 한다. 5단계에서 훌륭한 위험성 관리 계획 작성은 4단계에서 만들어진 시나리오 수준에 상당부분 직·간접적으로 의존하게 된다.

6단계, 평가자는 사례에 따라 전체 위험요인에 대한 전문가 판단을 문서화해야 할 필요가 있다. 이 단계에서 전문가들끼리 토의 과정을 거쳐 만들어진 개별 전문가 판단은 향후 책임관리(liability management) 측면에서 매우 중요한 부분이 될 수도 있다. 평가자는 사례 우선순위 혹은 전체 위험요인, 심각한 위해에 대한 위험, 범죄자가 암시한 다른 위험요인, 즉각적으로 취해져야 할 행동, 사례 (재)검토에 대한 유발요인 혹은 중요한 날짜 등을 고려하여 판단을 내리도록 한다. 중요한 점은, 성폭력위험성프로토콜은 다양한 위험요인에 근거하여 위험성을 산출할 수 있는 공식

28) Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. 앞의 책자(시나리오 부분) 참조.

29) Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. 앞의 책자 (시나리오 부분) 참조.

이나 알고리즘을 제시하지는 않지만, 평가자 자신의 자율재량으로 제기한 위험요인과 위험요인 관리 측면에서 다른 요인들의 의미나 관련성을 고려할 수 있도록 하였다. 본 연구의 목적은 앞서 서론 전반에 걸쳐 논의 한 바와 같이 현재까지 국내에 소개 된 바 없는 RSVP를 소개하고 이 도구가 국내에서 타당한 지 확인하고자 하였다. 또한 개별 위험요인이 평가자간에 신뢰성 있게 반응을 하는지 다른 도구들과의 관련성이 어떤지를 확인함으로써 국내 활용가능성을 평가하고자 하였다.

II. 연구 방법

1. 채점자 훈련 및 슈퍼비전

RSVP을 사용하기 위해서는 이론과 슈퍼비전으로 이루어진 적용 및 실습, 그리고 강의 혹은 워크숍 참여 등을 통한 다양한 방법으로 전문가로부터 적절한 훈련과 슈퍼비전을 받도록 규정하고 있다. 본 연구에서 RSVP 저자 Stephen D. Hart와 Kevin Douglas 교수를 초청하여 5일간 집중 교육을 실시하였다. 이 후 한 달간 원저자와 서종한 박사가 개별 평가자에게 3케이스 식 슈퍼비전을 실시했다. 슈퍼비전 이후 2일간 평가와 채점에 대해 전문보수교육을 진행했다. 보수교육은 RSVP 정보와 시행 주제에 강조점을 두며 구체적인 지침을 다시 검토하여 세부 사항에 대해 연구자간 논의하는 방식으로 진행했다. 두 번째, 성폭력 위험성 평가와 관련해 최근 문헌에 있어 중요한 발전이었는지 검토하고 논의했다. 전체적으로 저자의 권고 사항에 따르면, 교육 시간은 18시간의 학습하도록 하였고 슈퍼비전은 최소 한 달간의 기간을 두어 3케이스 이상 전문가의 피드백을 받도록 하였다. 전체 채점자 구성은 교정본부와 각 해당 청 심리치료센터에 소속된 교정심리학자 12명과 일반대학원 심리학과 박사과정 중에 있던 연구원 12명이었다. 3명이 한 조가 되어 총 8개의 팀을 이루었다.

2. 조사 대상

본 연구에서는 재소자 50명의 성폭력 범죄자를 대상으로 조사를 실시했다. 이 중 3명은 참여 거부와 응답 불충분으로 인해 평가에서 제외했다. 조사 대상자들의 평균 나이는 35.90세($SD=10.23$)이었고 모두 남성 성범죄자들로 구성되었다. 참가자 모두 전국에 분포한 교도소와 구치소(안양, 수원, 군산, 의정부, 포항, 청주)에 수용 중인 기결범이었다. 이들의 과거 평균 범법 건수는 3.98건($SD=5.42$)이었다.

3. 연구 진행과정

각 교도소에 연구자들이 직접 방문하여 대상자들의 개인 공식 기록 문서를 열람 후 1시간가량 검토하였고 이 후 면대면(FTF) 반구조화된 임상면담 방식(SSCI)으로 연구를 진행했다. 우선 개인공식 자료를 검토하는 데 동의서를 받고 동의과정을 구하는 과정에서 이탈자는 3명 있었다. 연구자에게도 개인 및 민감 정보 비밀 준수 서약을 작성하도록 했다. 연구 참여자 모두 IRB 정기 교육을 이수했다. 검토 정보로는 일반적으로, (교정)보라미시스템에 입력된 수용자 행동관찰 정보(징벌, 이송, 출력 등), 판결문, 수용자 심리치료 및 평가 내용, 범죄사실 기록 등을 참고했다.

RSVP 면담 과정은 Hart et al. (2003)이 매뉴얼에서 권장하는 방식에 따라 진행했다. 매뉴얼에서 권장하는 시행 순서는 문서기록과 부차적 정보를 먼저 검토하여 피면담자와 관련된 정보를 파악하고 면담을 실시하도록 권고하고 있다. 이 과정에서 RSVP 면담 가이드(interview guideline, 본 연구 목적으로 한국어로 번역)와 워크시트(worksheet; Hart et al., 2003; 본 연구 목적으로 한국어로 번역)를 활용하도록 하였다. 면담 과정에서 피검사자의 진술 왜곡과 부정, 축소 여부, 작화 및 거짓말 등에 대해 평가하도록 하였다. 모든 면담 과정이 끝나면 피면담자에게 소정의 보상이 주어졌다. 각 사례 당 3명의 채점자가 동시에 면담에 참여했다. 이들 중 한명은 면담을 진행했고 나머지 두 명은 면담에 참관했으며 부가적인 질문을 하도록 했다. 이 후 채점자는 면담자료, 부가적 자료, 범죄기록, 판결문, 교정심리평가 등을 통합적으로 참고하여 개별적으로 평가했다.

4. 측정도구

가. STATIC-99R³⁰⁾

이 도구는 오직 정적(static) 범죄관련 변인들만으로 이루어진 성범죄 재범 위험성 평가도구이다. Static-99r은 연령(출소에정나이), 친밀한 관계경험(2년 이상 성적 친밀감으로 동거), 본건 범죄 중 비성적 폭력범죄(본건만 해당), 과거범죄 중 비성적 폭력범죄(본건 제외), 과거범죄 중 성범죄 횟수(본건 제외), 과거 형선고 받은 횟수(본건 제외), 비접촉 성범죄(본건 포함), 비친족 피해자(본건 포함), 비면식 피해자(본건 포함), 남성 피해자(본건 포함) 10개 항목으로 구성되어 있다. 총 점수의 범위는 0점에서 10점이다. Helmus, Hanson et al. (2012)은 Static-99r와 성범죄 재범률 사이의 상관관계는 높고($r=.56$)이고 ROC의AUC는 .69에서 .70로 보고하였다.

나. 성폭력사범 재범위험성 평가³¹⁾

HAGSOR(한림성범죄자위험성평가도구; 조은경, 2010) 최근에 한림대학교에서 개발한 도구로 현재 교정시설에서 성범죄자의 재범위험성을 평가하기 위한 도구로 활용되고 있다. HAGSOR는 10개의 정적위험요인과 13개의 동적위험요인을 평가하며, 두 요인의 점수를 합산하여 최종적으로 위험성을 산출한다. 출소 예정 나이, 성범죄기록(첫 성범죄 연령, 이전 성범죄 유죄선고 횟수), 일반범죄기록(전체 유죄선고 횟수, 비성적 폭력범죄 유죄선고, 보호관찰/집행유예 위반여부), 성적일탈 및 피해자와의 관계(친족 피해자, 친인척 피해자, 남성 피해자, 13세 미만 피해자) 등으로 구성되어 있고 점수 범위는 0점에서 10점 사이이다. 타당화 검사에서 Static-2002³²⁾, RM2000/S³³⁾, KSORAS와 같은 다른 재범위험성 평가도구들과 .50 이상의 높은 상

30) Phenix, A., Fernandez, Y., Harris, A. J. R., Helmus, M., Hanson, K., & Thornton, D. (2016). *Static-99R Coding Rules Revised - 2016*.

31) 조은경 (2010). 성폭력사범재범위험성평가척도 개발(HAGSOR). 과천: 법무부[법무부용역연구보고서].

32) Canada. Solicitor General Canada. (2003). *Notes on the Development of Static-2002* [electronic Resource]. Solicitor General Canada.

33) Thornton, D. (2007). *Scoring guide for the Risk Matrix 2000.9/SVC*: February 2007 version.

관관계를 보였으며, 재범 예측도에 대한 연구가 추후 보완될 예정이다.

다. 성폭력위험성프로토콜

평가자에게 <표1>과 <표2>에 제시된 바와 같이 한국어로 번역된 RSVP 매뉴얼을 제공하였고 데이터 수집 평가지와 면담 가이드 내 질문지를 제공했다.

<표1> RSVP 단계 2(위험항목 존재유무), 3(관련성) 및 6(요약판단)에 대한 항목 요약기술

단계 2 & 3: 위험항목 존재 그리고 관련성 확인	
※ 채점자에게 각 위험 항목에 대한 존재 (과거 및 최근)와 관련성에 대해 평정해 줄 것을 요구(있음, 부분적, 없음)	
A. 성폭력 전력	내용 및 세부 항목
성폭력 만성도	개인이 저지른 성폭력 지속성과 빈도
성폭력 다양성	개인이 저지른 다양한 유형의 성폭력 범죄
성폭력 심화	개인이 저지른 성범죄가 시간이 지남에 따라 점차적으로 더 빈번해지고, 심각해지거나, 혹은 다양해지는 행동 패턴
성폭력시 신체적 강제	성범죄 도중의 행동 또는 성범죄를 위해 피해자에게 상해를 입히거나 신체적 상해에 대한 공포감을 불러일으키는 행동
성폭력시 심리적 강제	성범죄를 저지르는 과정 또는 좀 더 저지를 때 일어나는 행동으로 피해자에게 손실에 대한 위협 혹은 지위, 특권, 호의, 혹은 애정을 약속하는 것
B. 심리적 적응	
성폭력에 대한 극단적 축소 혹은 부인	성폭력 행동, 행동에 대한 책임, 그리고 행동의 결과를 인정하지 않거나 책임을 받아들이지 않는 것
성폭력을 지지 혹은 용인하는 태도	성폭력을 지지하거나 이를 용인하는 태도, 믿음, 가치, 혹은 인지상태
자기 인식 문제	개인을 성폭력 위험에 처하게 하는 요인·과정에 대한 인식부족
스트레스 혹은 대처능력 문제	외부 사건에 대해 불안정한 심리사회적 적응을 보이거나 민감한 정도를 나타냄
아동학대의 결과로 인해 나타난 문제	아동기, 청소년기에 학대당한 경험으로 인해 심리사회적 적응 문제이며, 성폭력 위험성 증가와 관련됨
C. 정신 장애	
성적 일탈성	일탈된 성적 각성의 안정된 패턴
사이코패스 성격장애	PCL-R (Hare, 1991, 2003) 사이코패스 체크리스트를 통한 평가로 정의
주요 정신질환	개인의 인지, 정서, 또는 행동에 있어 상당한 손상
약물 남용 문제	심리 사회적 기능을 손상시키는 합법/불법적인 약물을 얻으려는 개인의 시도

폭력적 혹은 자살사고	자신이나 타인에게 위해를 가하려는 사고, 충동, 환상 경험
D. 사회적 적응	
성적으로 친밀한 관계상의 문제	개인이 욕구, 능력 또는 기회가 있든 없든 적절 연령의 상대와 안정적인 성적 관계를 만들거나 유지하는데 실패하는 것
비성적 대인관계상의 문제	개인의 욕구, 능력, 혹은 기회 때문에 긍정적인 사회지지망을 형성하고 유지하는 데 실패를 겪는 것
고용관련 문제	안정적인 고용을 확립하고 유지하지 못하는 문제
비(非)성적 범죄	본질적으로 성적이지 않은 심각한 범죄행위

E. 관리용이성

계획 문제	친사회적으로 생애계획을 세우고 이를 이루려는 데 발생하는 문제
치료 문제	개인의 심리사회적 적응 상 결함을 다룰 목적으로 고안된 재활서비스(교정, 사회 복귀) 혜택에 대한 실패
관리감독 문제	성폭력 재범 가능성을 (더) 어렵게 만들고자 고안한 서비스를 수혜 받지 못함

단계 6: 전문가판단 요약

※채점자에게 아래와 같이 요약 판단을 내려 줄 것을 요청

1. 사례 우선순위	성폭력 범죄를 막기 위해 필요한 노력이나 개입 정도: 고/임박한, 보통/심화되는, 저/일상적인
2. 심각한 신체적 위해 위험성	개인이 저지를 수 있는 성폭력의 심각성과 임박성: 고, 보통, 저
3. 긴급조치의 필요성	즉각적 조치에 대한 필요성: 그렇다, 아마도, 아니다
4. 제기된 다른 위험성	성폭력 이외에 타인을 위해할 정도의 중대한 위험을 낳을 수 있는 모든 우려점: 그렇다, 아마도, 아니다

각 사례에 대해 9페이지 가량의 자료 수집 평가지(강제 선택)은 각 반응 변인에 대해 수집된 자료를 바탕으로 강제적으로 선택할 수 있게끔 항목으로 이루어져 있다. <표1>과 <표2>를 보면 사용된 자료수집 항목에 대한 간단한 기술과 세부 항목이 제시되어 있다. 본 연구를 위해 연구자가 임의적으로 단계 4, 5, 그리고 6에서는 몇 가지 항목으로 세분화하여 구분하였다. 평가자에게 단계 2와 3 그리고 단계 6에 대해 기술된 세부 항목 중에 하나를 선택해 줄 것을 요청했다. 그리고 단계 4와 5에서도 채점자에게 본 연구를 위해 만든 항목 중 하나를 선택해 줄 것을 요청했다. 각 사례의 마지막 부분에서는 채점자가 전체 성폭력 위험성에 대한 전체적인 추정

〈표2〉 RSVP 단계 4 (시나리오 계획), 5 (사례관리) 관련 추가 연구 항목 기술

단계 4: 시나리오 계획	
※ 채점자에게 가능한 ‘반복’ 그리고 ‘심화’ 범행 시나리오를 그려줄 것을 요청. 이 후 시나리오 특징을 확인하기 위해 아래와 같이 제시된 강제 선택 항목에 반응해 주도록 요청	
1. 성폭력 유형	성폭력 유형: 성적침해(예, 성희롱 등, 성기노출죄, 성추행, 강간(심각한 폭력 없음), 강간(심각한 폭력 동반), 강간살인
2. 피해자	시나리오에서 가능성 있는 피해자: 취학 전 남자, 취학 전 여자, 남자 청소년, 여자 청소년, 성인 남성, 성인 여성
3. 심리적 위해	심리적 위해 수준: 없음, 경미한(단기/가벼운 정서 고통, 보통(중기/보통의 정서 고통), 심각한(장기/주요한 정서 고통 및 심리장애, PTSD)
3. 신체적 위해	신체적 위해 수준: 없음, 경미한(가벼운 찰과상), 보통(자상과 타박상), 심각한(심각한 자상과 타박상, (거의) 치명적인
4. 임박성	성폭력 임박성: 1~4주, 6개월, 12개월, 몇 년
5. 빈도/지속기간	성폭력 빈도: 없음, 한두 번, 여러 번, 반복적인/자주 지속기간: 위험성 만성 또는 급성 여부: 만성, 급성
6. 가능성	성폭력 가능성: 매우 낮은, 낮은, 보통, 높은, 매우 높은
단계 5: 사례관리	
※채점자에게 가장 적절한 모니터링에 대해 강제선택 추천을 하게끔 요구하였음	
추천 모니터링 수준	모니터링 수준: 정기적인 전문가 접촉, 관련 전문가와의 중간 수준의 약속 전화통화, 관련 전문가와 높은 수준의 대면접촉과 전화통화

5. 자료 분석

급내상관계수, 개별 RSVP 항목에 대한 채점자간 신뢰도는 ICC,(개별 점수가 상관관계)와 ICC_c(평가자 평균가 상관관계)를 통해 평가하였다. 단계 2, 3 그리고 6 단계에서 표준화된 RSVP 위험성 요인에 대해 개별적으로 분석이 이루어졌고 추가적으로 단계 4와 5에서 부분적인 항목도 함께 분석했다. ICC 모델은 평균 변동성 추정치에 근거하고 있고 집단 간 분산과 전체 분산 간 비율로 개념화 할 수 있다. ICC는 모든 판단과 항목에 걸쳐 나타나는 전체 분산에 똑같은 테스트 항목에 대한 다른 판단의 변산도를 비교하여 정정기회 동의 척도(chance-corrected agreement)를 제공해 준다. 특히 2명 이상의 채점자가 있거나 데이터가 범주화 속성을 띠는 경우 신뢰도를 평가하는데 추천되는 통계적 방식이다³⁴⁾.

34) Uebersax, J. (2011). Statistical Methods for Rater and Diagnostic Agreement.(2009).

ICC가 가중 Kappa와 산술적으로 동일한 값을 산출³⁵⁾할 뿐 아니라 법-임상심리 연구 분야에서 더 일반적으로 사용하고 있기에 선택되었고 이후 다른 연구 결과와도 비교할 수 있을 것으로 보인다. ICC 통계치에 대한 몇 가지 형태가 존재하고³⁶⁾ 각각이 다른 연구 설계에 적합할 수 있다. 모든 채점자가 각 사례를 평가했고 채점자와 사례 모두 더 큰 전집에서 표집 될 수 있는 것으로 보이기 때문에 ICC₂(이원 무선효과: two-way random effect)가 여기서는 더 적합할 것으로 보인다. 각 사례와 채점자에 대한 무선평가가 순수하게 이루어지지 않았지만, 본 연구가 유목내 평균효과 혹은 고정효과(population average effect or fixed effect)와 반대로 대상자-특정효과(subject-specific effect)와 관련 있기 때문에 무선평가를 가정했다. 또한 가정과 해석이 다르기는 하지만 다른 이원 ICC 모델에 대한 수치가 동일하다(Nichols, 1998). ICC는 절대 동의(absolute agreement)와 단일 측정(single measure)으로 나누어져 계산된다.

RSVP 채점자간 신뢰도를 평가했던 이전 연구를 따라 ICC 해석 가이드라인은 Fleiss (1981)³⁷⁾을 따랐다. 기준은 다음과 같다. ICC<.39인 경우 나쁨, .40 < ICC < .59인 경우 보통, .60 < ICC < .74인 경우 좋음, .75 < ICC인 경우 아주 좋음으로 평가했다. 추가적으로, Landis와 Koch (1977)³⁸⁾는 0이하 이면 나쁨, 0-.20은 약간, .21-.40은 적절한, 4.1-.60은 보통, .61-.80 좋은, 그리고 .81-1.00은 거의 완벽 한으로 평가할 수 있다.

ICC 표본크기, Walter, Eliasziw 그리고 Donner (1998)³⁹⁾는 ICCs에 사용되어 하

35) Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency*; Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

36) McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, 1(1), 30.

37) Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), 105-112.

38) Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

39) Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine*, 17(1), 101-110.

는 적정 수의 대상자와 채점자 수를 계산할 수 있는 방정식을 제공한 바 있다. PASS(Power and Sample Size for Windows⁴⁰⁾)가 이 방정식을 잘 나타내는 데 본 연구를 위해 필요한 채점자와 대상자 수를 계산하는 데 사용했지만 본 연구에서는 PASS에서 제시하는 적정 수의 대상자를 채우지는 못했다. 이 점에 대한 한계점은 논의에서 좀 더 자세히 기술하겠다.

백분율 일치성, 백분율 일치성도 실용적 수준에서 있는 대략적인 일치성에 대한 정보를 제공해 준다 (Uebersax, 2009)⁴¹⁾. 본 연구에서는 백분율 즉 평균과(가까운 정수로 반올림 한 값)에 대한 일치성을 살펴보았다. 각 일치성 측정은 각각 평균과 일치하는 관찰비율을 나타낸다고 볼 수 있다. 백분율 일치성은 각 항목과 전체 사례 및 채점자를 대상으로 계산되어졌다. 본 연구에서 평균을 사용하여 백분율 일치성을 알아 본 것은 실제 RSVP 전문가 채점에서 논의과정과 의견 조율과정이 중요한 단면을 이루고 있기 때문이다.

III. 연구 결과

1. 연구문제: RSVP를 사용하여 위험성 판단을 했을 경우 채점자 간 신뢰도

〈표3〉 RSVP 존재 채점, 관련성 채점, 전문가 요약판단 채점에서의 ICC

	(최근) 존재		(미래) 관련성	
	ICC ₁	ICC ₂	ICC ₁	ICC ₂
단계 2 & 3: A 성폭력전력	.91(.86-.95)	.98(.96-.99)	.91(.85-.95)	.98(.96-.99)
단계 2 & 3: B 심리적 적응	.93(.89-.96)	.98(.97-.99)	.91(.86-.95)	.98(.96-.99)
단계 2 & 3: C 주요 정신질환	.88(.81-.93)	.97(.94-.98)	.90(.84-.94)	.97(.95-.98)
단계 2 & 3: D 사회적 적응	.89(.83-.94)	.97(.95-.98)	.87(.80-.93)	.97(.94-.98)
단계 2 & 3: E 관리용이성	.87(.80-.93)	.97(.94-.98)	.84(.76-.91)	.96(.93-.98)

40) Hintze, J. L. (2008). *Power analysis and sample size system (PASS) for windows User's Guide I*. NCSS. USA: Utah, Kaysville.

41) Uebersax, J. S. (2009). *User guide for the MH program (vers. 1.2)*. Statistical Methods for Rater Agreement website. 2006.

	(최근) 존재		(미래) 관련성	
	ICC ₁	ICC ₂	ICC ₁	ICC ₂
단계 2 & 3: 전체	.95(.93-.97)	.99(.98-.99)	.95(.92-.97)	.99(.98-.99)
단계 6: 전문가요약판단	.65(.50-.78)	.88(.80-.93)	N/A	N/A

<표3>과 <표4>는 RSVP 단계 2, 3, 6단계에 대한 전체 ICC를 나타내고 있다. Fleiss (1981) 기준에 따르면 전문가 요약 판단을 제외한 모든 영역에서 아주 좋은 신뢰도(>.75)를 보이고 있고 단계6은 좋은 신뢰도(.60<ICC<.74)를 보이고 있다.

〈표4〉 성폭력 전력(단계 3)과 심리적 적응 영역(단계 3)에서의 ICC

	ICC			
	사례 수	평정자수	ICC ₁	ICC ₂
영역 A: 성폭력 전력	47	4		
1. 만성도				
존재 (과거)	47	5	.88(.82-.93)	.97(.95-.98)
존재 (최근)	47	4	.87(.81-.93)	.97(.94-.98)
관련성 (미래)	47	4	.82(.72-.89)	.94(.91-.97)
2. 다양성				
존재 (과거)	47	4	.89(.82-.93)	.97(.95-.98)
존재 (최근)	47	4	.87(.80-.92)	.96(.94-.98)
관련성 (미래)	47	4	.87(.80-.92)	.96(.94-.98)
3. 심화				
존재 (과거)	47	4	.86(.79-.92)	.96(.93-.98)
존재 (현재)	47	4	.89(.82-.93)	.97(.95-.98)
관련성 (미래)	47	4	.84(.75-.90)	.95(.92-.97)
4. 신체적 강제				
존재 (과거)	47	4	.84(.75-.96)	.95(.92-.97)
존재 (최근)	47	4	.72(.59-.83)	.91(.85-.95)
관련성 (미래)	47	4	.72(.59-.83)	.91(.85-.95)
5. 심리적 강제				
존재 (과거)	47	4	.81(.71-.88)	.94(.90-.97)
존재 (최근)	47	4	.87(.80-.92)	.96(.94-.98)
관련성 (미래)	47	4	.80(.70-.88)	.94(.90-.96)
영역평균			.74(.62-.84)	.92(.87-.95)
영역 B: 심리적 적응				

	ICC			
	사례 수	평정자수	ICC ₁	ICC ₂
<i>6. 극단적 축소 혹은 부인</i>				
존재 (과거)	47	4	.83(.70-.88)	.94(.90-.96)
존재 (최근)	47	4	.85(.77-.91)	.96(.93-.97)
관련성 (미래)	47	4	.81(.71-.88)	.94(.90-.96)
<i>7. 지지·용인하는 태도</i>				
존재 (과거)	47	4	.70(.56-.81)	.90(.83-.94)
존재 (최근)	47	4	.76(.64-.85)	.92(.88-.96)
관련성 (미래)	47	4	.69(.55-.81)	.90(.83-.94)
<i>8. 자기 인식 문제</i>				
존재 (과거)	47	4	.77(.65-.86)	.93(.88-.96)
존재 (최근)	47	4	.88(.81-.93)	.96(.94-.98)
관련성 (미래)	47	4	.84(.75-.90)	.95(.92-.97)
<i>9. 스트레스/대처능력 문제</i>				
존재 (과거)	47	4	.91(.86-.95)	.97(.96-.98)
존재 (최근)	47	4	.96(.93-.97)	.99(.98-.99)
관련성 (미래)	47	4	.96(.93-.97)	.99(.98-.99)
<i>10. 이동학대 문제</i>				
존재 (과거)	47	4	.90(.85-.94)	.97(.95-.98)
존재 (최근)	47	4	.88(.81-.93)	.96(.94-.98)
관련성 (미래)	47	4	.88(.81-.93)	.96(.94-.98)
영역평균			.84(.76-.91)	.95(.92-.97)

RSVP 영역 내 각 항목에 대한 채점자간 신뢰도를 구했다. Landis와 Koch (1977)의 기준에 따르면 성폭력 전력 영역에서는 12개 항목에서 거의 ‘완벽할 정도의’ 일치성을 보여주었고 나머지 3개 항목인 신체적 강제(최근, 미래)와 심리적 강제(미래)도 ‘좋은’ 일치도를 보였다. 심리적 적응에서는 4개 항목을 제외한 모두에서 ‘완벽한’ 일치성을 보였고 ‘지지·용인하는 태도’가 상대적으로 다른 항목에 비해 낮은 일치도(과거: $ICC_1=.70$, $ICC_2=.90$; 최근: $ICC_1=.76$, $ICC_2=.92$; 미래: $ICC_1=.69$, $ICC_2=.90$)를 보였지만 여전히 좋은 신뢰도를 나타냈다.

〈표5〉 주요 정신장애(단계 3) 및 사회적 적응 영역(단계 3)에서의 ICC

	ICC			
	사례 수	평정자수	ICC ₁	ICC ₂
영역 C: 정신 장애				
11. 성적일탈성				
존재 (과거)	47	4	.81(.72-.89)	.94(.91-.97)
존재 (최근)	47	4	.78(.68-.87)	.93(.89-.96)
관련성 (미래)	47	4	.77(.65-.86)	.93(.88-.96)
12. 사이코패스성격장애				
존재 (과거)	47	4	.75(.63-.85)	.92(.87-.95)
존재 (최근)	47	4	.73(.61-.84)	.91(.86-.95)
관련성 (미래)	47	4	.75(.63-.85)	.92(.87-.95)
13. 주요 정신질환				
존재 (과거)	47	4	.84(.76-.91)	.95(.92-.97)
존재 (최근)	47	4	.86(.78-.92)	.96(.93-.97)
관련성 (미래)	47	4	.85(.77-.91)	.96(.93-.97)
14. 약물 남용 문제				
존재 (과거)	47	4	.86(.79-.92)	.96(.93-.97)
존재 (최근)	47	4	.82(.72-.89)	.94(.91-.97)
관련성 (미래)	47	4	.84(.76-.90)	.95(.92-.97)
15. 폭력적·자살사고				
존재 (과거)	47	4	.78(.68-.87)	.93(.89-.96)
존재 (최근)	47	4	.72(.59-.83)	.91(.85-.95)
관련성 (미래)	47	4	.74(.61-.84)	.92(.86-.95)
영역평균			.72(.59-.83)	.91(.85-.95)
영역 D: 사회적 적응				
16. 성적 친밀한 관계상 문제				
존재 (과거)	47	4	.84(.76-.90)	.95(.92-.97)
존재 (최근)	47	4	.70(.57-.81)	.90(.84-.94)
관련성 (미래)	47	4	.73(.60-.83)	.91(.85-.95)
17. 비성적 대인관계상 문제				
존재 (과거)	47	4	.83(.74-.90)	.95(.92-.97)
존재 (최근)	47	4	.82(.73-.89)	.95(.91-.97)
관련성 (미래)	47	4	.80(.69-.88)	.94(.90-.96)

	ICC			
	사례 수	평정자수	ICC ₁	ICC ₂
18. 고용관련 문제				
존재 (과거)	47	4	.82(.73-.89)	.95(.91-.97)
존재 (최근)	47	4	.88(.82-.93)	.96(.94-.98)
관련성 (미래)	47	4	.85(.77-.91)	.95(.93-.97)
19. 비(非)성적 범죄				
존재 (과거)	47	4	.86(.79-.92)	.96(.94-.98)
존재 (최근)	47	4	.88(.82-.93)	.96(.94-.98)
관련성 (미래)	47	4	.80(.70-.88)	.94(.90-.96)
영역평균			.70(.57-.82)	.90(.84-.94)

<표5>는 정신장애 영역과 사회적 적응 영역 세부 항목 15개에 대한 ICC를 보여 주고 있다. Fleiss (1981)의 기준으로 보면, ‘아주 좋은’ 신뢰도 (ICC>.75)를 보이는 항목은 11개이며 그 이외 4개 항목도 ‘좋은’ (.61<ICC<.80) 일치도를 보여 주었다. 사회적 적응영역에서는 10개 항목에서 ‘아주 좋은’ 신뢰도를 보였고 나머지 2개 항목도 ‘좋은’ 신뢰도를 보였다.

3.1.1 ‘RSVP 단계 6’의 채점자간 신뢰도

<표6> 관리용이성(단계 3) 및 전문가 판단 요약(단계 6)에서의 ICC

	ICC			
	사례수	평정자수	ICC ₁	ICC ₂
영역 E: 관리용이성				
20. 계획 문제				
존재 (과거)	47	4	.85(.77-.91)	.95(.93-.97)
존재 (최근)	47	4	.80(.70-.88)	.94(.90-.96)
관련성 (미래)	47	4	.78(.68-.87)	.93(.89-.96)
21. 치료 문제				
존재 (과거)	47	4	.88(.81-.93)	.96(.94-.98)
존재 (최근)	47	4	.90(.84-.94)	.97(.95-.98)
관련성 (미래)	47	4	.87(.80-.92)	.96(.94-.98)
22. 관리감독 문제				

	ICC			
	사례수	평정자수	ICC ₁	ICC ₂
존재 (과거)	47	4	.71(.58-.82)	.90(.84-.94)
존재 (최근)	47	4	.72(.59-.85)	.91(.85-.95)
관련성 (미래)	47	4	.65(.51-.78)	.88(.80-.93)
영역평균			.85(.77-.91)	.95(.93-.97)
단계 6: 전문가 판단 요약				
사례 우선순위	47	4	.71(.58-.82)	.91(.85-.95)
심각한 신체 위해 위험성	47	4	.67(.53-.79)	.89(.81-.93)
긴급조치의 필요성	47	4	.51(.35-.68)	.81(.68-.89)
제기된 다른 위험성	47	4	.46(.30-.64)	.77(.63-.87)
사례검토	47	4	.44(.27-.62)	.76(.60-.86)

<표6>은 관리용이성 총 9개 항목에 따른 신뢰도와 단계 6의 4가지 전문가 요약 판단 항목에 대한 신뢰도를 보여주고 있다. 관리용이성 측면에서는 Landis와 Koch (1977)의 기준을 적용해 보면, 5개 항목에서 완벽한 일치성을 보이며 높은 신뢰도를 나타냈고 그 이외에 .61 이상의 좋은 신뢰도를 보였다. 다만 상대적으로 관리감독 항목이 ‘존재’와 ‘관련성’부분에서 낮은 신뢰도 (관련성: ICC₁=.65, ICC₂=.88)를 보였다. 단계 6의 요약판단에서는 사례우선 순위 (ICC₁=.71, ICC₂=.91)와 심각한 신체 위해 위험성 (ICC₁=.67, ICC₂=.89)이 좋은 신뢰도를 보인 반면, 긴급조치 필요성 (ICC₁=.51, ICC₂=.81), 제기된 다른 위험성 (ICC₁=.46, ICC₂=.77) 그리고 사례검토 (ICC₁=.44, ICC₂=.76)는 보통 정도의 신뢰도 수준을 나타냈다.

〈표7〉 전문가 요약 판단(단계 6)에서의 평균 백분율 일치도

	낮음	보통	높음
단계 6: 요약판단 항목 채점을 모두 페어링 (두개씩 짝을 지음) 한 이후 이들 간 평균 백분율 일치도			
사례우선순위			
낮음	14.3%	10.6%	4.7%
보통	2.4%	26.4%	9.0%
높음	0.0%	1.4%	31.4%
심각한 신체적 위해 위험성			
낮음	20.7%	10.9%	1.7%
보통	2.9%	37.1%	5.4%

	낮음	보통	높음
높음	0,0%	6,3%	15,0%
긴급조치 필요성			
낮음	13,0%	9,2%	2,2%
보통	10,5%	27,5%	12,0%
높음	0,0%	10,2%	14,6%
제기된 다른 위험성			
낮음	2,4%	5,3%	1,8%
보통	4,7%	23,4%	8,1%
높음	0,0%	12,4%	34,9%

note. 백분율은 6개의 페어링 간 평균을 말함 (즉, A-B 채점; A-C 채점; A-D 채점; B-C 채점; B-D 채점; C-D 채점)

신뢰도 분석과 함께 전문가 요약 판단에 속해 있는 4개 항목에 대한 *평균 백분율 일치도(average percentage agreement)*를 살펴보았다. 총 47케이스에 대해 각각 6가지 형태의 페어링(두 개씩 짝짓기)이 존재하며, 이를테면 A-B, A-C, A-D, B-C, B-D, 그리고 C-D 등이고 이들 6가지 페어링 간 평균을 산출하여 일치도를 알아보았다.

사례우선순위의 경우 페어링 채점 중 72.1%가 완전한 일치도를 보였다(낮음-낮음, 보통-보통, 높음-높음). 하나의 범주차이(예, 낮음-보통, 보통-높음)가 나는 백분율은 23.4%이며 두개의 범주차이(예, 낮음-높음)가 나는 경우는 4.7%였다. 심각한 신체적 위해 위험성은 72.8%가 완전한 일치도를 보였다(즉, 낮음-낮음, 보통-보통, 높음-높음). 하나의 범주 차이를 보이는 경우(예, 낮음-보통, 보통-높음 등), 19.4%이며 두개 범주 차이가 나는 경우(낮음-높음), 1.7%였다.

긴급조치 필요성은 55.1%가 완전한 일치도를 보였고, 하나 범주 차이를 보이는 경우는 19.4%, 두 개의 범주 차이를 보이는 경우는 2.2%였다. 제기된 다른 위험성에서는 60.7%가 완전한 일치도를 보였고, 하나의 범주 차이를 보이는 경우는 17.7%, 두 개의 범주 차이를 보이는 경우는 1.8%였다. 전체적으로 사례우선 순위와 심각한 신체적 위해 위험성 부분은 70%를 상회하는 높은 일치성을 보인 반면, 긴급조치 필요성은 50% 조금 넘는 수준으로 상대적으로 낮은 일치성을 보였다.

RSVP 단계 4와 5 채점자간 신뢰도

〈표8〉 시나리오 계획(단계 4) 및 사례관리(단계 5)에서 ICC

	ICC			
	사례수	평정자수	ICC ₁	ICC ₂
단계 4: 시나리오 계획				
<i>반복시나리오</i>				
시나리오 유형	47	4	.85(.77-.91)	.95(.93-.97)
시나리오 피해자	47	4	.83(.74-.90)	.95(.92-.97)
심리적 위해	47	4	.49(.33-.66)	.79(.66-.88)
신체적 위해	47	4	.49(.33-.66)	.79(.66-.88)
생명위협 가능성	47	4	.64(.49-.77)	.87(.79-.93)
임박성	47	4	.65(.50-.78)	.88(.80-.93)
빈도	47	4	.68(.54-.80)	.89(.82-.94)
지속기간(만성/급성)	47	4	.56(.40-.71)	.84(.73-.91)
시나리오 가능성	47	4	.65(.51-.78)	.88(.80-.93)
<i>심화시나리오</i>				
시나리오 유형	47	4	.78(.67-.87)	.93(.89-.96)
시나리오 피해자	47	4	.58(.43-.73)	.85(.75-.91)
심리적 위해	47	4	.30(.14-.49)	.63(.39-.79)
신체적 위해	47	4	.61(.46-.75)	.86(.77-.92)
생명위협 가능성	47	4	.46(.29-.63)	.77(.62-.87)
임박성	47	4	.51(.35-.68)	.81(.68-.89)
빈도	47	4	.20(.05-.39)	.50(.17-.72)
지속기간(만성/급성)	47	4	.50(.34-.67)	.80(.67-.89)
시나리오 가능성	47	4	.39(.22-.57)	.72(.53-.84)
단계 5: 사례관리				
추천된 모니터링 수준(반복)	47	4	.55(.38-.70)	.83(.71-.90)
추천된 모니터링 수준(심화)	47	4	.60(.44-.74)	.85(.76-.92)

<표8>은 시나리오 항목을 반복 및 심화로 나누어 살펴 본 결과를 보여주고 있다. Landis와 Koch (1977)의 분류 기준에 따라, 반복시나리오에서는 시나리오 유형과 피해자에서 ‘아주 좋은’ 신뢰도를 보였고 생명위협 가능성, 임박성 그리고 시나리오 가능성은 ‘좋은’ 신뢰도를 보였다. 심리적 위해와 신체적 위해 그리고 지속기간은 ‘보통’ 정도의 신뢰도를 나타냈다. 심화시나리오는 상대적으로 반복시나리오보다

떨어지는 신뢰도를 보였다. ‘아주 좋은’ 신뢰도를 보이는 항목은 존재하지 않았고 유일하게 신체적 위해만 ‘좋은’ 수준의 신뢰도를 보였다. 그 이외에 시나리오 피해자, 생명위험 가능성, 지속기간은 ‘보통’ 수준이었고 심리적 위해와 시나리오 가능성은 ‘적정’ 수준의 일치도를 보였다. 단계 5인 사례관리는 반복 및 심화 시나리오 관련 모니터링 수준에서 모두 ‘보통’ 수준의 신뢰도를 보였다.

연구문제2: RSVP 공인타당도

RSVP ‘존재’ 항목 공인타당도

〈표9〉 RSVP 22 ‘존재 (최근)’ 항목과 STATIC-99R 및 HAGSOR 상관관계 (r)

RSVP 존재 (최근) 항목	STATIC-99R TOTAL	STATIC-99R RISK CAT.	HAGSOR TOTAL	HAGSOR RISK CAT.
A. 성폭력 전력				
만성도	.108	.043	.394**	.376**
다양성	.324*	.189	.417***	.408***
심화	.241	.116	.190	.219*
신체적 강제	-.006	.009	-.052	.071
심리적 강제	.038	.000	-.143	-.118
A 영역 점수	.406***	.243**	.430***	.414***
B. 심리적 적응				
극단적 축소 혹은 부인	.019	.055	.132	.004
지지 혹은 용인하는 태도	.493***	.352**	.139	.042
자기 인식 문제	.174	.122	.471***	.480***
스트레스/대처능력 문제	-.007	-.103	.422***	.524***
아동학대 문제	.443***	.341**	.605***	.541***
B 영역점수	.449***	.303**	.659***	.571***
C. 정신 장애				
성적일탈성	.263*	.112	.321**	.260*
사이코패스성격장애	.042	.074	.011	.162
주요 정신 질환	.394**	.283*	.494***	.460***
약물 남용 문제	.184	.230*	.350**	.382**
폭력적·자살사고	.261*	.208*	.348**	.452**

<i>RSVP 존재 (최근) 항목</i>	<i>STATIC-99R TOTAL</i>	<i>STATIC-99R RISK CAT.</i>	<i>HAGSOR TOTAL</i>	<i>HAGSOR RISK CAT.</i>
C 영역점수	.426***	.351**	.537***	.575***
D. 사회적 적응				
성적으로 친밀한 관계상의 문제	.175	.128	.214*	.117
비성적 대인관계상의 문제	.385**	.417***	.463***	.404***
고용관련 문제	.276**	.257*	.379**	.411***
비(非)성적 범죄	-.015	.013	.153	.261*
D 영역점수	.402***	.369**	.547***	.543***
E. 관리용이성				
계획 문제	.490***	.501***	.634***	.657***
치료 문제	.337**	.272*	.363**	.323**
관리감독 문제	.275*	.221*	.586**	.494***
E 영역점수	.579***	.531***	.682***	.620***
전체점수				

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

<표9>는 RSVP 22개 항목에 대한 존재 관련 공인타당도를 나타내고 있다. 성폭력 전력 영역 중 ‘다양성’이 상대적으로 두 준거 척도와 높은 관련성을 보였다. 성폭력 전력 전체 영역 점수와는 높은 상관관계를 나타냈다. 심리적 적응 중 ‘아동학대로 결과로 인한 문제’와 두 준거 척도 간에 유의미한 상관이 존재했다. 심리적 적응 영역 점수도 높은 상관관계를 보였다.

정신장애 중 ‘주요 정신질환’과 ‘폭력적/자살 사고’는 두 준거 척도 간 유의미한 상관을 보였고 영역 점수와도 높은 상관을 보였다. 사회적 적응 영역 중 ‘비성적 대인관계상의 문제’와 ‘고용관련문제’가 유의미한 상관을 보였고 영역 점수 또한 높은 상관을 보였다. 관리 용이성은 모두 항목에서 관련성을 보였고 HAGSOR가 STATIC-99R보다 대체로 더 높은 상관을 보였다.

RSVP ‘관련성’ 항목 공인타당도

〈표10〉 RSVP 22개 ‘관련성’ 항목과 STATIC-99R 및 HAGSOR간 상관관계 (r)

RSVP 22개 (관련성) 항목	STATIC-99R TOTAL	STATIC-99R RISK CAT.	HAGSOR TOTAL	HAGSOR RISK CAT.
A. 성폭력 전력				
만성도	.020	-.097	.167	.119
다양성	.317**	.139	.282*	.228*
심화	.169	.073	.084	.148
신체적 강제	.103	.110	.200*	.362**
심리적 강제	.078	.032	-.102	-.024
영역점수	.406***	.243*	.430***	.414***
B. 심리적 적응				
극단적 축소 혹은 부인	.029	.008	.272*	.130
지지 혹은 용인하는 태도	.506***	.371**	.174	.037
자기 인식 문제	.139	.085	.471***	.480***
스트레스/대처능력 문제	-.055	-.152	.287*	.369**
아동학대 문제	.428***	.320**	.577***	.504***
영역점수	.449***	.303**	.659***	.571***
C. 정신 장애				
성적 일탈성	.359**	.221*	.410***	.384**
사이코패스성격장애	.042	.074	.011	.162
주요 정신 질환	.362**	.254**	.431***	.377**
약물 남용 문제	.239*	.286**	.367**	.389***
폭력적·자살사고	.212*	.175	.366**	.455***
영역점수	.426***	.351**	.537***	.575***
D. 사회적 적응				
성적으로 친밀한 관계상의 문제	.200*	.140	.214*	.117
비성적 대인관계상의 문제	.302**	.304**	.382**	.316**
고용관련 문제	.316**	.269**	.289*	.329**
비(非)성적 범죄	.094	.086	.275*	.412***
영역점수	.402***	.369**	.547***	.543***
E. 관리용이성				
계획 문제	.396**	.395**	.490***	.453***
치료 문제	.337**	.272*	.363**	.323**
관리감독 문제	.329**	.236*	.562***	.469**
영역점수	.579***	.531***	.682***	.620***
전체점수				

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

<표10>은 RSVP 22개 항목에 대한 ‘관련성’ 관련 공인타당도를 보여주고 있다. 앞서 살펴본 ‘존재’ 항목과도 전체적으로 비슷한 결과를 보였다. 성폭력 전력 영역 중 ‘다양성’이 상대적으로 두 준거 척도와 높은 관련성을 보였고 신체적 강제는 HAGSOR와 상대적으로 관련성이 있었다. 성폭력 전력 전체 영역 점수와는 높은 상관관계를 보였다. 심리적 적응 중 ‘아동학대로 결과로 인한 문제’와 두 준거 척도 간에 유의미한 상관이 존재했다. 특히 성폭력을 지지하거나 용인하는 태도는 STATIC-99R과, 반면 성폭력에 대한 자기 인식문제는 HAGSOR와 높은 상관성을 나타냈다.

정신장애 중 ‘성적일탈성’, ‘주요 정신질환’ 그리고 ‘약물 남용문제’는 두 준거 척도와 모두 상관관계를 보였다. 정신장애 영역 점수와도 높은 상관을 보였다. 비슷하게 사회적 적응 영역 중 ‘비성적 대인관계상의 문제’와 ‘고용관련문제’가 유의미한 상관을 보였고 영역 점수 또한 높은 상관을 보였다. 관리 용이성은 모두 항목에서 일관성 있게 높은 관련성을 보였으며, HAGSOR가 STATIC-99R보다 대체로 더 높은 상관 지수를 보였다.

RSVP ‘시나리오 계획’ 항목 공인타당도

<표11> RSVP 시나리오 계획 항목과 STATIC-99R 및 HAGSOR 간 상관관계 (r)

RSVP ‘시나리오 계획’ 항목	STATIC-99R TOTAL	STATIC-99R RISK CAT.	HAGSOR TOTAL	HAGSOR RISK CAT.
단계 4: 시나리오 계획				
반복시나리오				
시나리오 유형	.139	.085	.328**	.373**
시나리오 피해자 유형	-.091	-.075	.164*	.302**
심리적 위해	-.064	-.113	-.031	-.086
생명위협 가능성	.198	.109	.379**	.450***
신체적 위해	.165	.096	.278**	.260**
임박성	.055	.143	.061	.041
빈도(급성/만성)	-.051	-.122	.031	-.004
지속기간	-.131	-.062	-.213	-.129
시나리오 가능성	-.016	-.029	.055	.094

<i>RSVP '시나리오 계획' 항목</i>	<i>STATIC-99R TOTAL</i>	<i>STATIC-99R RISK CAT.</i>	<i>HAGSOR TOTAL</i>	<i>HAGSOR RISK CAT.</i>
심화시나리오				
시나리오 유형	.044	-.027	.253*	.362**
시나리오 피해자 유형	-.215	-.193	-.087	-.054
심리적 위해	-.008	-.081	-.039	-.026
신체적 위해	.044	-.034	.198	.338**
생명위협가능성	.041	-.069	.258*	.283*
임박성	.109	.135	-.040	-.031
빈도(급성/만성)	.144	.077	.388**	.224*
지속기간	-.262*	-.199*	-.379**	-.270*
시나리오 가능성	.015	-.004	.253*	.241*
영역 평균				

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

<표11>은 RSVP 시나리오 계획 9개 항목과 두 준거 척도 간에 상관 정도를 보여주고 있다. 전체적으로 STATIC-99R과는 상관성을 보이고 있지는 않지만, HAGSOR와는 어느 정도의 상관관계를 보여주고 있다.

구체적으로 반복 시나리오에서는 ‘시나리오 유형’, ‘피해자 유형’, ‘생명위협 가능성’, ‘신체적 위해’ 등 5개 항목과 관련성을 보여주었다. 비슷하게도 심화 시나리오에서는 ‘시나리오 유형’, ‘신체적 위해’, ‘생명위협 가능성’, ‘빈도’ 등 6개 항목에서 유의미한 상관을 보였다.

RSVP ‘사례관리 및 요약판단’ 항목 공인타당도

<표12> RSVP ‘사례관리 및 요약판단’ 항목과 STATIC-99R 및 HAGSOR 간 상관관계 (r)

<i>RSVP (사례관리 및 요약판단) 항목</i>	<i>STATIC-99R TOTAL</i>	<i>STATIC-99R RISK CAT.</i>	<i>HAGSOR TOTAL</i>	<i>HAGSOR RISK CAT.</i>
단계 5: 사례관리				
모니터링 수준 (반복시나리오)	.024	.026	-.067	-.038
모니터링 수준 (심화시나리오)	.152	.133	.281*	.216*
단계 6: 전문가 요약 판단				
사례 우선순위	.107	-.012	.149	.211*

<i>RSVP (사례관리 및 요약판단) 항목</i>	<i>STATIC-99R TOTAL</i>	<i>STATIC-99R RISK CAT.</i>	<i>HAGSOR TOTAL</i>	<i>HAGSOR RISK CAT.</i>
심각한 신체적 위해 위험성	.251*	.190*	.518***	.628***
긴급조치의 필요성	-.162	-.202*	-.140	-.094
제기된 다른 위험성	.066	.054	-.082	.019
사례검토	-.143	.019	-.240*	-.189*

전체점수

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

<표12>는 RSVP 단계 5 시나리오에 따른 ‘모니터링’ 수준과 단계 6 전문가 ‘요약 판단’에 대한 공인타당도이다. STATIC-99R과는 ‘심각한 신체적 위해 위험성’과 상관을 보였고, HAGSOR와는 ‘모니터링’ 수준, ‘심각한 신체적 위해 위험성’ 그리고 ‘사례검토’에서 유의미한 상관을 보였다. 전체적으로 한국에서 타당화 된 위험성 평가 척도 HAGSOR과 높은 관련성을 보여주었다.

IV. 논의

전체적으로 선행연구⁴²⁾와 일치하게도, 본 연구 결과 높은 채점자간 신뢰도를 보였다⁴³⁾. 흥미로운 점은 참가자들로부터 심화 시나리오보다 반복 시나리오에서 좀 더 안정적인 값을 얻을 수 있었다. 이는 보통 전문가들이 의사결정을 할 때 이전에

42) Hart, S. D. (2003, April). *Assessing risk for sexual violence: The risk for sexual violence protocol (RSVP)*. In annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria; Watt, K. A., & Jackson, K. (2008). *Interrater and structural reliabilities of the Risk for Sexual Violence Protocol (RSVP)*. In annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria; Jackson, K. J. (2016). *Validation of the Risk for Sexual Violence Protocol in Adult Sexual Offenders* (Doctoral dissertation, Arts & Social Sciences: Department of Psychology); Ryon & Hart. (2016). *Validation the Risk for Sexual Violence Protocol in Sexual Offenders* (Master thesis, Arts & Social Sciences: Department of Psychology).

43) see e.g., Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health*, 11(2), 119-133.

일어나지 않은 시나리오와는 반대로 실제 일어났던 과거 행동을 근거로 이루어진다는 점을 이해하면 납득할 만한 점이다.

전문가 요약판단은 전체적으로 단계 2 & 3보다 낮은 수준의 신뢰도를 보였다. 긴급조치의 필요성, 제기된 다른 위험성, 사례검토 등의 항목이 특히 낮았는데 이는 Sutherland et al. (2016)과 Jackson (2016)⁴⁴⁾의 논문과 일치하는 부분이다.

전문가 요약 판단(단계6)에 평균 백분율을 살펴본 결과 전체적으로 50% 이상의 높은 일치율을 보였고 그 중에서도 사례우선순위 매기기와 심각한 신체적 위해 위험성은 상대적으로 높은 일치율을 보였다. 3명 이상의 전문가가 실무에서 평가와 관련된 결과를 바탕으로 논의 과정을 거칠 때 이루어질 수 있는 긴급조치 필요성에 대해서는 다양한 태도로 보였다. 긴급조치는 당장 필요한 개입할 뿐 아니라 부수적으로 동반되는 전략에 대해 추가적인 의사결정이 필요하기 때문에 해당 전문가의 위치와 경력에 따라 다르게 고려하는 것으로 보인다.

공인 타당도 측면에서는 RSVP 5개 영역과 STATIC-99R 및 HAGSOR 간 상관관계는(존재와 관련성 측면에서) 유의한 수준으로 비교적 높게 나타났다(.40-.68). 외국에 제작되어 국내에 사용 중인 STATIC-99R 보다는 한국형 성폭력 재범위험성 평가를 위해 개발된 도구인 HAGSOR가 상대적으로 높은 상관관계를 보였다. 본 연구를 위해 인위적으로 범주화로 묶은 시나리오에서는 STATIC-99R과는 관련성이 거의 나타나지 않았다. 다만 HAGSOR와는 부분적으로 몇 가지 항목(예, 성폭력 유형, 피해자, 생명위협가능성, 신체적 위해)에서 유의한 상관을 보였다. 심화 시나리오는 반복 시나리오보다 상대적으로 많은 상관관계를 보였는데 빈도 지속기간 시나리오 가능성에서 유의미한 수준에서 높은 상관을 보였다. 전문가 요약 판단 중에 심각한 신체적 위험 위험성과 두 척도 간에 유의미한 상관을 보였다.

본 연구는 기존의 다른 연구⁴⁵⁾와는 달리, 실제 채점자들이 직접 참가자들을 대상

44) see e.g., Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., & Stocks, R. 앞의 논문 참조; Jackson, K. J. 앞의 (박사학위)논문 참조.

45) Ryon & Hart. 앞의 논문 참조; Hart, S. D. 앞의 발표논문 참조; Watt, K. A., & Jackson, K. 앞의 발표논문 참조; Watt, K. A., Klaver, J. A., Eaves, D., & Hart, S. D. (2002, March). *Clinical assessment of risk for sexual violence: a normality study of risk communication practices*. Paper presented at the Bennial Meeting of the America Psychology-Law Society (Div. 41 of the American Psychological Associations), Austin, Texas.

으로 인터뷰 과정이 이루어졌고, 이후 사건 파일을 검토한 다음 개별 채점이 이루어졌다. 단순히 파일 검토 과정만 거쳤거나 후향적으로 이전 파일 혹은 실제 사건에 근거한 설명문(예, vignette)을 검토한 후에 평정이 이루어지지 않았다. 무엇보다도 최소 3명 이상의 채점자가 참여하였고 이 후 논의 과정을 거침으로써 본 도구의 목적에 맞게끔 연구가 진행되었다. 전체 및 각 영역 별 점수 이외에도 5개 영역에 대한 세부 항목 22개에 대한 신뢰를 살펴보았는데 이전에 연구되지 않은 부분으로 그 의의가 크다고 볼 수 있다.

이 연구의 가장 큰 장점은 채점자 모두 연구 참가자와는 비면식 관계에 있는 상태였다. 즉, 비면식 관계에서 채점자들은 참가자들의 특징을 세세하게 잡아 낼 수 있었을 뿐만 아니라, 방법론 측면에서도 이러한 비면식 관계가 평정에 영향을 미치지 않았음을 의미하기 때문에 채점자간 오류가 최소화 될 수 있었다.

공인타당도 연구에서, RSVP 평정 자체가 다른 두 준거 척도 평정하는 과정에 영향을 주었을 수 도 있다. 이런 이유로 인해 다소 과도하게 상관관계가 도출 될 가능성도 배제할 수 없다. 본 연구에서 4명의 채점자가 동참을 했지만 47명의 사례에서 탄탄한 연구 설계를 제공하기 위해서는 최소한 9-10명의 채점자가 동참하는 것이 가장 바람직하다(Power and Sample Size for Windows⁴⁶⁾). 본 연구는 채점자가 신뢰도를 알아보기 국내에서 실시한 탐색적 수준의 연구로 향후 연구에서는 앞서 언급했던 것처럼 표본사례수와 채점자간 적정 수를 고려하여 이루어져야 한다. 마지막으로 채점자간 신뢰도를 구할 때 실제 임상가와 연구자 간 차이를 알아 볼 필요가 있다. Sutherland et al (2016)⁴⁷⁾이 제시한 것처럼 전문가들의 전문성과 감독이 실제 채점에 어떤 영향을 주었는지 면밀하게 검토할 필요가 있다.

RSVP는 광범위한 영역의 성폭력 위험성 평가에 적용 될 수 있다. 다른 실무 지침서들처럼 연구, 교육 및 훈련, 그리고 중요한 사건 검토에 활용될 수 있다⁴⁸⁾. 첫

46) Hintze, J. L. (2008). *Power analysis and sample size system (PASS) for windows User's Guide I*. NCSS. USA: Utah, Kaysville.

47) Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., & Stocks, R. 앞의 논문 참조.

48) Reed, G. M., McLaughlin, C. J., & Newman, R. (2002). American Psychological Association policy in context. *American Psychologist*, 57(12), 1041-1047.

번째, 성폭력과 관련된 범죄로 체포되었을 때, 범죄자들의 범죄혐의 유형 혹은 범죄 이력을 보면 공판 전 석방을 불허할지 또는 예측된 위험을 관리하기 위해 어떤 조건 하에 공판 전 석방을 해야 할 지에 대해 의문을 제기할 수 있다. 이수정과 윤옥경 (2003)이 제시한 바와 같이, 판결 전 시행된 성폭력 위험성 평가는 판사가 피고인에게 다이버전, 조건부 혹은 조건 없는 석방을 고려할 때 도움이 될 수 있다⁴⁹⁾. 판결이 난 후 성폭력 위험성 평가는 판사가 두 가지 대안 형 가운데 결정하거나(예: 보호관찰 또는 수감) 지역사회 내 관리감독 조건을 만들거나 권고 할 때(예: 법원명령치료(감호), 거짓말탐지기를 사용한 관리(감독)) 도움이 될 수 있다⁵⁰⁾.

두 번째, 처분 판결 이후, RSVP는 기관 혹은 지역사회 배정을 위해 수용 평가를 진행하는 교정 직원에게 도움이 된다. 평가 결과는 치료 계획 수립뿐만 아니라 적절성, 그리고 가족 방문, 일시석방 등 조건을 구성할 때 사용될 수 있다. 구류 판결을 받았을 때, 석방 이전에 실시한 성폭력 위험성 평가는 기관의 교정 직원 혹은 가석방 심의 위원회가 석방의 적합성 혹은 조건부 석방을 결정하는데 도움이 되고, 또한 석방 후 치료 혹은 관리 계획을 수립하는데 도움이 된다. 지역사회내 거주하며 보호관찰 기간이 끝나가는 범죄자에게 교정직원은 위험성 평가 결과를 통해 공식적인 경고를 고지해야 한다. 이런 노력을 통해 교정직원은 해당사건이 종결되기 전 윤리적 그리고 법적의무를 면할 수 있게 된다.

세 번째, 많은 국외 형사사법체계에서는 “성폭력 포식자”라고 알려진, “위험한” 혹은 “고위험” 성범죄자에게 부정기적 형·민사적 수감을 허용하는 법규를 제정하였는데(예: Fitch & Hammen, 2003⁵¹⁾). 성폭력 위험성 평가는 이런 법적 절차에서 중요한 역할을 할 수 있다. 네 번째, 과거 성폭력을 저질렀다고 알려진 혹은 성폭력 혐의가 있다고 추정되는 사람이 (양)부모, 보호자, 혹은 아동 서비스 제공자 및 다른 취약한 집단의 보호자로서 적합한지 판단하기 위해 성폭력 위험성 평가에 위탁될 수 있다. 또한 민사제도 안에서 가정폭력에 관심이 많아지고 있는 데 양육/출입 평

49) 이춘화. (2001). 소년에 대한 조사제도의 문제점과 개선방안. *소년보호연구*, (3), 26-85. 한국의 판결 전 조사서에 대한 부분을 참조.

50) 이수정 & 윤옥경. 위 논문 참조.

51) Fitch, W. L., & Hammen, D. A. (2003). The new generation of sex offender commitment laws: Which states have them and how do they work?

가에 있어, 과거 혹은 지속되는 성폭력 문제가 대두될 때마다 성폭력 위험성 평가를 포함해야 한다⁵²⁾. 마지막으로, 사실상 복미의 모든 사법체계에서는 법령적, 관습법적, 그리고 도덕적 의무로서 의료 및 사회 복지 서비스 제공자들이 특정 상황에서 위험에 처한 사람들에게 경고 혹은 보호하는 의무를 부과하고 있다.⁵³⁾ 미국에서는 일반적으로 “메간 법”으로 알려져 있는 신상공개법령에 기인한 것일 수 있다. 성폭력 위험성 평가는 전문가들이 범죄자가 가지고 있는 위험이 법에 따라 “합리적이고 가능성이 있는지” 혹은 “심각한지(예: 임박한, 심각한 신체적 혹은 심리적 상해를 초래할 수 있는) 판단 내릴 때 도움을 줄 수 있다.

52) Herman, S. P. (1997). Practice parameters for child custody evaluation. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(10), 57S-68S.

53) Ad Hoc Committee on Legal and Ethical Issues in the Treatment of Interpersonal Violence. (1997). *potential problems for psychologists working with the area of interpersonal violence* [Brochure]. Washington, DC: APA; Herbert, P. B. (2002). The duty to warn: A reconsideration and critique. *JOURNAL-AMERICAN ACADEMY OF PSYCHIATRY AND THE LAW*, 30, 417-424.

참고문헌

- 윤정숙, 이수정 (2012). 성범죄자 치료 프로그램의 주요 원칙과 방향 정립을 위한 이론적 고찰. 한국심리학회지: 일반, 31(3), 847-875.
- 이수정 (2010). 최신 범죄 심리학. 서울: 북카페.
- 이수정·고려진 (2011), 한국 성범죄자 위험성 평가척도(KSORAS) 실시요강. 서울: 맥스메디카 .
- 이수정·윤옥경. (2003). 범죄위험성의 평가와 활용방안. 한국심리학회지: 일반, 22(2), 99-126.
- 이춘화 (2001). 소년에 대한 조사제도의 문제점과 개선방안. 소년보호연구, (3), 26-85.
- 조은경 (2010). 성폭력사범재범위험성평가척도 개발(HAGSOR). 과천: 법무부[법무부용역연구보고서].
- 조은경·이수정 (2008). 한국판 표준화. PCL-R 전문가 지침서, (R. D. Hare 저, 조은경, 이수정 역, 2008). 서울: 학지사 심리검사 연구소.
- 정지숙·신정·이장규. (2012). 재범위험성 평가도구 타당도 검증. 한국심리학회지: 일반, 31(3), 643-665.
- 최금화 (2009). 폭력 위험성 측정 도구 (Violence Risk Assessment: HCR-20)의 타당도 연구. 경기대학교 대학원 석사 학위 논문).
- 최기홍, 임민경, 이지혜, 이한나, 김태동. (2013). 근거기반실천과 심리치료. 한국심리학회지: 일반, 32(1), 251-270
- Ad Hoc Committee on Legal and Ethical Issues in the Treatment of Interpersonal Violence. (1997). potential problems for psychologists working with the area of interpersonal violence [Brochure]. Washington, DC: APA.
- Andrews, D. A. (2012). The risk-need-responsivity (RNR) model of correctional assessment and treatment. *Using social science to reduce violent offending*, 127-156.
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-Based

- Practice in Psychology. *American Psychologist*, 61(4), 271- 285.
- Canada. Solicitor General Canada. (2003). *Notes on the Development of Static-2002* [electronic Resource]. Solicitor General Canada.
- Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American journal of mental deficiency*, 86(2), 127-137.
- Cooke, D. J., & Michie, C. (2010). Limitations of diagnostic precision and predictive utility in the individual case: A challenge for forensic practice. *Law and human behavior*, 34(4), 259-274.
- Desmarais, S. L., & Singh, J. P. (2013). *Risk assessment instruments validated and implemented in correctional settings in the United States*. Washington, DC, Council of State Governments Justice Center.
- Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20 version 3: assessing risk for violence*. Burnaby, BC, Canada: Mental Health, Law and Policy Institute, Simon Fraser University.
- Fitch, W. L., & Hammen, D. A. (2003). *The new generation of sex offender commitment laws: Which states have them and how do they work?*. In Winick, Bruce J. (Ed); La Fond, John Q. (Ed). (2003). *Protecting society from sexually dangerous offenders: Law, justice, and therapy.*, (pp. 27-39). Washington, DC, US: American Psychological Association.
- Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement*, 5(1), 105-112.
- Haag, A. M., Boyes, A., Cheng, J., MacNeil, A., & Wirove, R. (2016). An introduction to the issues of cross-cultural assessment inspired by Ewert v. Canada. *Journal of Threat Assessment and Management*, 3(2), 65.
- Hart, S. D. (2003, April). *Assessing risk for sexual violence: The risk for sexual violence protocol (RSVP)*. In annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria.

- Hart, S. D. (2009). Evidence-based assessment of risk for sexual violence. *Chapman Journal of Criminal Justice*, 1, 143-165.
- Hart, S. D. (2016). Culture and violence risk assessment: The case of Ewert v. Canada. *Journal of Threat Assessment and Management*, 3(2), 76.
- Hart, S. D., & Cooke, D. J. (2013). Another look at the (Im) precision of individual risk estimates made using actuarial risk assessment instruments. *Behavioral Sciences & the Law*, 31(1), 81-102.
- Hart, S. D., Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence. Burnaby, British Columbia: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments. *The British Journal of Psychiatry*, 190(49), 60-65.
- Herbert, P. B. (2002). The duty to warn: A reconsideration and critique. *Journal-American Academy of Psychiatry and The Law*, 30, 417-424.
- Herman, S. P. (1997). Practice parameters for child custody evaluation. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(10), 57-68.
- Hintze, J. L. (2008). *Power analysis and sample size system (PASS) for windows User's Guide I*. NCSS. USA: Utah, Kaysville.
- Jackson, K. J. (2016). Validation of the Risk for Sexual Violence Protocol in Adult Sexual Offenders (Doctoral dissertation, Arts & Social Sciences: Department of Psychology).
- Justice, L. (2008). Evidence-based terminology. *American Journal of Speech-Language Pathology*, 17, 324-325.
- Kroner, D. G. (2016). The Ewert v. Canada judgment: Moving forward. *Journal of Threat Assessment and Management*, 3(2), 122.

- Kropp, P. R., Hart, S. D., & Lyon, D. (2008). *Guidelines for Stalking Assessment and Management (SAM): User manual*. Vancouver, Canada: ProActive ReSolutions Inc.
- Kropp, P. R., Laws, D. R., Klaver, J., Logan, C., & Watt, K. A. (2003). *The Risk for Sexual Violence Protocol (RSVP): Structured professional guidelines for assessing risk of sexual violence*. Burnaby, British Columbia: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, 1(1), 30.
- Miles, A., Polychronis, A., & Grey, J. E. (2006). The evidence-based health care debate . 2006. Where are we now? *Journal of Evaluation in Clinical Practice*, 12, 239-247.
- Monahan, J., & Skeem, J. L. (2014). The evolution of violence risk assessment. *CNS Spectrums*, 19, 419-424.
- Monahan, J., Steadman, H. J., Appelbaum, P. S., Grisso, T., Mulvey, E. P., Roth, L. H., & Silver, E. (2006). The classification of violence risk. *Behavioral Sciences & the Law*, 24(6), 721-730.
- Otto, R. K., & Douglas, K. S. (Eds.). (2011). *Handbook of violence risk assessment*. Routledge.
- Phenix, A., Fernandez, Y., Harris, A. J. R., Helmus, M., Hanson, K., & Thornton, D. (2016). Static-99R Coding Rules Revised - 2016.
- Pomerol, J.-C. (2001). Scenario development and practical decision making under uncertainty. *Decision Support Systems*, 31, 197-204.
- Reed, G. M., McLaughlin, C. J., & Newman, R. (2002). American Psychological Association policy in context. *American Psychologist*, 57(12), 1041-1047.
- Ringland, G. (1998). *Scenario planning: Managing for the future*. Chichester,

- UK: Wiley.
- Ryon & Hart. (2016). Validation the Risk for Sexual Violence Protocol in Sexual Offenders (Master thesis, Arts & Social Sciences: Department of Psychology).
- Sackett, D. L., Rosenberg, W. M., Muir Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. it's about integrating individual clinical expertise and the best external evidence. *British Medical Journal*, 312, 71-72.
- Skeem, J. L., & Monahan, J. (2011). Current directions in violence risk assessment. *Current Directions in Psychological Science*, 20(1), 38-42.
- Sutherland, A. A., Johnstone, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., ... & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the risk for sexual violence protocol. *International Journal of Forensic Mental Health*, 11(2), 119-133.
- Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three controversies and a caveat. *Health Affairs*, 24, 163-173.
- Timmermans, S., & Mauck, A. (2005). The promises and pitfalls of evidence-based medicine. *Health Affairs*, 24, 18-28.
- Thornton, D. (2007). Scoring guide for the Risk Matrix 2000.9/SVC: February 2007 version. Retrieved from [http://www.cfcp.bham.ac.uk/Extras/SCORING%20GUIDE%20FOR%20RISK%20MATRIX%202000.9-%20SVC%20\(ver.%20Feb%202007\).pdf](http://www.cfcp.bham.ac.uk/Extras/SCORING%20GUIDE%20FOR%20RISK%20MATRIX%202000.9-%20SVC%20(ver.%20Feb%202007).pdf)
- Uebersax, J. (2011). Statistical Methods for Rater and Diagnostic Agreement.(2009).
- Uebersax, J. S. (2009). *User guide for the MH program (vers. 1.2)*. Statistical Methods for Rater Agreement website. 2006.
- van der Heijden, K. (1994). *Probabilistic planning and scenario planning*. In G. right & P; Ayton (Eds.), *Subjective probability* (pp. 549-572). Chichester,

UK: Wiley.

- Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine*, 17(1), 101-110.
- Watt, K. A., & Jackson, K. (2008). Interrater and structural reliabilities of the Risk for Sexual Violence Protocol (RSVP). In annual meeting of the International Association of Forensic Mental Health Services, Vienna, Austria.
- Watt, K. A., Klaver, J. A., Eaves, D., & Hart, S. D. (2002, March). Clinical assessment of risk for sexual violence: a normality study of risk communication practices. Paper presented at the Biennial Meeting of the American Psychology-Law Society (Div. 41 of the American Psychological Associations), Austin, Texas.

The evolution of Sexual Violence Risk Assessment: The Risk for Sexual Violence Protocol based on SPJ

Sea, Jong-han* · Kim, Hee-song** · Kim, Kyung-il***

The Risk for Sexual Violence Protocol is a representative risk assessment instrument which is based on Structured Professional Judgement approach, following evidence-base principals. Compared to actuarial instruments, the instrument was developed to help clinician's decision to make risk management and treatment plan based on case scenarios. On this account, a substantial number of clinicians have selected the RSVP across the world, mainly e.g., Europe and North America. Previous literatures and scholars have also evaluated the RSVP as one of fourth generation's assessments and approved the utility and efficacy in terms of crime prevention consisting of idiographic intervention and future risk management plan. Yet, the RSVP has not been validated in Korean settings. Our goal was to help to stimulate discussion about the critical role of formulation in risk assessment and to examine options for best practice in terms of RSVP. Then, we attempted to validate RSVP to check whether the usage possibility in Korea correctional settings is a proper option for forensic practitioners. In the study, fifty inmates were selected for interrater reliability and concurrent validity. In the process of study, raters were first trained through intensive five-day workshop and three-case supervision, and then participated in interviews with inmates. Static-99R and HAGSOR were adopted for the concurrent validity. As a result of the study, twenty risk factors of the RSVP were strongly appeared to be high in interrater reliability, while step 5 and 6 were

* Senior researcher, Mental Health, Law, and Policy Institute, Simon Fraser University.

** Chief, Dept. of Forensic Psychology, National Forensic Service.

*** Correspondence author: Professor, Dept. of Psychology, Ajou University.

relatively lowered compared to step 4. In terms of concurrent validity, on the whole, twenty RSVP risk factors were highly correlated with two measures, whereas step 5 and 6 were partially correlated or not correlated. Specifically, compared to Static-99R, HAGSOR (developed in Korea) was more correlated with step 4, 5, and 6. Through this empirical study, the RSVP seems to provide proper interrater validity and concurrent validity, which is potentially well applicable in Korea correctional setting, however still the caution on step 5 and 6 of the RSVP might be required.

❖ Keyword: Evidence-base, Sexual Violence, Risk Factors, Formulation, RSVP, SPJ