



What works in threat and risk assessment, for whom, and in what circumstances?

The State of the Evidence

This work is supported by funding by the United States Department of Homeland Security, Science and Technology Directorate - Public Safety and Violence Prevention Evaluations
70RSAT21G00000002/70RSAT21FR0000084



NCITE NATIONAL COUNTERTERRORISM
INNOVATION, TECHNOLOGY,
AND EDUCATION CENTER
A U.S. DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

Developed in partnership with NCITE by:
Paul Gill, Zoe Marchment, Amber Seaward

University of College, London

Introduction

This output is from a much broader project that seeks to review current prevention programs; domestic and international, to research best practices in threat assessment, interventions, re-integrations, and counter-recidivism. In doing so, it aims to determine their effectiveness, review their efficacy and identify best practices undertaken by allies abroad in an effort to understand the similarities and differences in the known process of radicalization and mobilization for terrorism as compared to targeted violence.

A previous output by the teams at UCL and NCITE synthesized what was known about risk and threat assessment instruments, their psychometric properties, and their predictive capabilities. It was noted that the human element of threat and risk assessment processes remained unsynthesized. This 'state of the science' briefing note acts as a first step to such a synthesis and combines both basic and applied literature, peer reviewed scientific outputs, and so-called 'grey' literature. We argue that such human factors are far more consequential for real life harm reduction than an instrument's predictive powers, and especially so when low base rate events are the subject of enquiry.

To synthesize this evidence, we conducted another systematic review.¹ Five researchers sifted through 7,259 studies and judged their title and abstract based on whether the study: (a) concerned risk/threat assessment, (b) was an evaluation, (c) was a *process* evaluation, and (d) had some empiricism. 48 full papers made it through this initial sift. For the purpose of this

¹ We conducted a keyword search of titles and abstracts in PsycNet and the National Criminal Justice Reference Service for papers published from database inception until 11th November 2021, restricted by English language. Key words searched for issues related to problems of interest (insider*" OR "violen*" OR "terroris*" OR "radicali*" OR "crim*" OR "recidiv*" OR "offen*" OR "extremis*" OR "aggressi*" OR "threat*" OR "arrest*" OR "reoffen*" OR "re-offen*" OR "assault" OR "femicide" OR "counterproductive workplace behav*" OR "stalk*" OR "sex*" OR "homicide*" OR "killing*" OR "attack*" OR "murder*" OR "harass*" OR "shoot*" OR "fixat*"), threat/risk assessment ("risk assess*" OR "threat assess*" OR "risk manag*" OR "threat manag*" OR "case manag*" OR "lethality assess*" OR "danger assess*" OR "assess* risk" OR "assessment of risk" OR "manag* risk" OR "management of risk" OR "risk instrument*" OR "risk classif*" OR "risk predict*" OR "actuarial" OR "structured professional judgement" OR "SPJ"), and evaluations ("evaluat*" OR "effect*" OR "outcome*" OR "program*").

briefing note, 21 were excluded because they were not peer reviewed papers but will be incorporated into future updates. Only 27 studies therefore met all of our initial criteria (see reference list for full details).

To organize our results to date, we utilize the EMMIE framework, developed for the UK ‘What Works Centre for Crime Reduction’ at the Department of Security and Crime Science at University College London.² EMMIE is an acronym denoting five categories of evidence relevant to policing and crime prevention (see Box 1 below). It was inspired by the ‘realist’ approach to evaluation³, which directs evaluators to ask not only whether an intervention ‘worked’ – the dominant question in the crime prevention literature - but *how* an intervention worked (or not), *why*, for whom, and under what conditions is it more or less effective. Just as importantly, EMMIE calls attention to ways in which some interventions may inadvertently backfire under particular conditions. EMMIE was designed to help disentangle the many components of a complex intervention in order to generate insight into the features that support its success (or otherwise).

The EMMIE framework is especially relevant to threat and risk assessment due to its multi-method focus which combines process and outcome evaluation traditions. It offers a dynamic, holistic means of understanding how processes work. This helps to elicit working theory, and tease out contextual variation, which is likely given the vast array of contexts in which threat and risk assessments play out.

Box 1 – The EMMIE framework

Effect – has *it* worked?

The first ‘E’ of EMMIE refers to ‘effect’ size. Typically, this focuses on the ‘effect’ of an intervention but can also be extended to other non-traditional effects (e.g. the formulation itself, the management plan).

Mechanism – how did *it* work?

The first ‘M’ refers to the ‘mechanism’ through which an intervention brings about its effect – the ‘active ingredient’ so to speak. This is important in determining what needs to be done to produce (or avoid) a given outcome.

Moderator – what conditions are needed for *it* to work?

The second ‘M’ refers to ‘moderators’ (or ‘contexts’) – the conditions that are instrumental for an intervention to activate the mechanisms.

Implementation – what was found to be needed to put *it* in place?

The ‘I’ refers to ‘implementation’ conditions that support or obstruct delivery of the intervention (this would include reliability testing).

² See <http://whatworks.college.police.uk/toolkit/About-the-Crime-Reduction-Toolkit/Pages/About.aspx>

³ Pawson, R. and Tilley, N. (1997). *Realistic Evaluation*. London, UK: Sage.

Economics – is *it* cost effective?

Finally, the second ‘E’ refers to ‘economics’ – what the intervention will cost in relation to outputs, outcomes or benefits.⁴

The following pages briefly summarize the state of the science emanating from published peer-reviewed process evaluations of threat and risk assessment procedures:

Effects (*has it worked?*)

Key Finding	Evaluation Method	Instrument / Setting	Citation
Users report risk assessment instruments as useful and valuable to everyday practice.	Police officer survey (n=213)	Range of intimate partner violence risk assessment instruments	Campbell et al. (2018)
	Practitioner survey (n=42)	Practitioners dealing with child sexual exploitation	Franklin et al. (2018)
	Practitioner survey (n=28)	Short-Term Assessment of Risk and Treatability: Adolescent Version (START:AV)	Sher & Gralton (2014)
Assessment outcomes strongly impact on perceptions of offending likelihood but are rarely the sole criteria that factors into an opinion.	User survey (n=109)	Static-99R	Chevalier et al. (2015)
	Practitioner survey (n=42)	Practitioners dealing with child sexual exploitation	Franklin et al. (2018)

Mechanisms (*how did it work?*)

Key Finding	Evaluation Method	Instrument / Setting	Citation
Assessments help prioritize key risk and protective factors that require treatment/intervention/boosting.	Case file analysis (n=107)	Case Management Strategies	Dhaliwal et al. (1994)

⁴ Manning, M., Johnson, S.D., Tilley, N., Wong, G.T. and Vorsina, M. (2016). *Economic Analysis and Efficiency in Policing, Criminal Justice and Crime Reduction: What Works?*. Palgrave Macmillan.

	Case file analysis (n=216)	Offender Assessment System (OASys)	Kewley et al. (2015)
	Practitioner survey (n=28)	START:AV	Sher & Gralton (2014)
Users report risk assessments help focus minds on problems/concerns, map risk and level of risk, guide decision-making and client interviews, help with team building, provide a consistency of language and are useful to potential victims to show concerns being highlighted.	Practitioner survey (n=42)	Practitioners dealing with child sexual exploitation	Franklin et al. (2018)
	Practitioner survey (n=88)	Wales Applied Risk Research Network (WARRN), Child and adolescent mental health services	Gray et al. (2019)
	Practitioner interviews (n=10)	Multi-Agency Risk Assessment Conferences	Robinson (2006)
	Practitioner survey (n=28)	START:AV	Sher & Gralton (2014)
Risk assessments open avenues for specific types of treatment and interventions.	Referrer interviews (n=31)	National Health Service (NHS) Lothian Sex Offender Liaison Service	Judge et al. (2014)
Risk assessment can negate potential biases about certain types of victims.	Case file analysis (n=867)	Brief Spousal Assault Form for the Evaluation of Risk (B-SAFER)	Storey & Strand (2017)
Risk assessments (a) by partner agencies for police and (b) by police for partner agencies, help improve and inform decision-making, objective analysis, and case management.	Police officer survey (n=213)	Range of intimate partner violence risk assessment instruments	Campbell et al. (2018)
	Referrer survey (n=14)	The Integrated Threat and Risk Assessment Centre	Ennis et al. (2015)

	Referrer interviews (n=31)	NHS Lothian Sex Offender Liaison Service	Judge et al. (2014)
	Embedded observatory design	Multi-Agency Risk Assessment Conferences	Robinson (2006)
Training significantly improves identification of (critical) risk factors, risk levels, and the quantity and quality of suggested management strategies which correspond to identified risk factors. Training also improves a participants' rationale for their decisions. Trainees reported greater levels of confidence, and perceived competence in risk assessment.	Practitioner survey (n=96)	Spousal Assault Risk Assessment (SARA), Stalking Assessment and Management (SAM), Historical Clinical and Risk Management 20 (HCR20)	Storey et al. (2011)
Training on a specific instrument significantly improved self-reported skills in formulation, risk management/safety planning, and risk communication. Training also improved self-reported confidence, service user safety and general public safety, and ability to save lives. Training also improves information sharing and communication across agencies.	Practitioner survey (n=88)	WARRN, Child and adolescent mental health services	Gray et al. (2019)
	Practitioner survey (n=28)	START:AV	Sher & Gralton (2014)
	Practitioner survey (n=85)	Clinical Assessment of Risk Decision Support (CARDS), Clinical sites	Watts et al. (2004)
Threat assessment training improves understanding of basic concepts and guidelines for conducting a threat assessment, the understanding of specific forms of violence, helps responses to threats of violence and motivates participants to use threat assessment principles in their practice.	Education practitioners (n=4,666)	Comprehensive Student Threat Assessment Guidelines	Stohlman et al. (2020)
Training in risk formulation significantly improves user perception of instrument effectiveness, anticipated impact upon future work, anticipated impact upon managing risk, perceptions of how easy it is to complete, and perceived relevance to clinical practice.	Clinician survey (n=131)	HCR20-v3	Covernton et al. (2019)

Risk assessment training increases the likelihood of a formulation and risk management plan being articulated even in the absence of a specific instrument being used.	Case file analysis (n=100)	Hospital setting	Sundrum & Browne (2004)
Limited or inadequate resourcing negatively impacts the ability for risk management plans to be actioned.	Referrer survey (n=14)	The Integrated Threat and Risk Assessment Centre	Ennis et al. (2015)

Moderators (*what conditions are needed for it to work?*)

Key Finding	Evaluation Method	Instrument / Setting	Citation
Thorough assessments require information from a range of sources. Failure to do so risks an incomplete picture and harm being realized. Poor information sharing practices can negatively impact the success of threat assessment. Issues regarding confidentiality are a consistent concern.	Single case study	Threat assessment in schools	Goodrum et al. (2018)
	Practitioner interviews (n=10)	Multi-Agency Risk Assessment Conferences	Robinson (2006)
The inability to consider alternative scenarios in changing and complex environments can negatively impact threat assessments.	Single case study	Government cybersecurity department	Ceric & Holland (2019)
Older and more experienced respondents report higher perceived usefulness for risk assessment instruments.	Police officers (n=213)	Range of intimate partner violence risk assessment instruments	Campbell et al. (2018)
High staff turnover and poor knowledge of existing resources negatively impacts multi-agency working.	Referrer survey (n=14)	The Integrated Threat and Risk Assessment Centre	Ennis et al. (2015)
Different professional backgrounds benefit equally from training.	Education practitioners (n=4,666)	Comprehensive Student Threat Assessment Guidelines	Stohlman et al. (2020)
Refresher workshops significantly increase adherence to guidance protocols.	5 users coding 60+ cases each	START:AV	De Beuf et al. (2020)

Slow assessments lead to negative perceptions from referrers.	Referrer survey (n=14)	The Integrated Threat and Risk Assessment Centre	Ennis et al. (2015)
Without standardized training, users of the same instrument who have different professional backgrounds are likely to (a) use the instrument differently and (b) interpret a single instrument's risk scores differently.	User survey (n=95)	Psychopathy Checklist Revised	Boccaccini et al. (2017)
	User survey (n=109)	Static-99R	Chevalier et al. (2015)
	Case file analysis (n=72)	Static-99R, Static-99R, Minnesota Sex Offender Screening Tool-Revised (MnSOST-R)	Murrie et al. (2009)
	Case file analysis (n=398)	Psychopathy Checklist Revised	Murrie et al. (2012)
Different trainers, providing the same training materials, can have significantly different impacts upon how much participants learn.	Education practitioners (n=4,666)	Comprehensive Student Threat Assessment Guidelines	Stohlman et al. (2020)

Implementation (what was needed to be put in place?)

Key Finding	Evaluation Method	Instrument / Setting	Citation
"Risk assessment matters, but only when implemented well."	Multi-site process and outcome evaluation	Youth Level of Service/Case Management Inventory; Structured Assessment of Violence Risk in Youth (SAVRY)	Vincent et al. (2016)
Users often do not adhere to the guidance in a variety of ways including reporting risk, coding, and document keeping. Quality assurance, refresher workshops and supervisory input is therefore needed.	User survey (n=95)	Psychopathy Checklist Revised	Boccaccini et al. (2017)
	User survey (n=109)	Static-99R	Chevalier et al. (2015)
	Field study	START:AV	De Beuf et al. (2020)
	Case file analysis (n=34)	HCR20v2	Gough et al. (2015)

	User survey (n=95)	Psychopathy Checklist Revised	Boccaccini et al. (2017)
Practitioners might push back against risk assessment instruments that do not provide for professional discretion in decision-making. It is important to provide users with some discretion and avoid a 'tick box' culture.	Police officer survey (n=213)	Range of intimate partner violence risk assessment instruments	Campbell et al. (2018)
	Practitioner survey (n=42)	Practitioners dealing with child sexual exploitation	Franklin et al. (2018)
Training parameters and content need to be signposted in advance.	Practitioner survey (n=27)	HCR-20v1	Garrett & Rowe (2004)
Understanding user needs is essential for continued upskilling.	Practitioner survey (n=88)	WARRN, Child and adolescent mental health services	Gray et al. (2019)
	Practitioner survey (n=28)	START:AV	Sher & Gralton (2014)
Users report documentation and processes can take too long and/or are resource intensive. Systems need to be put in place to reduce this burden.	Practitioner survey (n=88)	WARRN, Child and adolescent mental health services	Gray et al. (2019)
	Referrer interviews (n=31)	NHS Lothian Sex Offender Liaison Service	Judge et al. (2014)
	Practitioner interviews (n=10)	Multi-Agency Risk Assessment Conferences	Robinson (2006)
	Practitioner survey (n=28)	START:AV	Sher & Gralton (2014)
	Practitioner survey (n=85)	CARDS, Clinical sites	Watts et al. (2004)

Economics (is it cost effective?)

Key Finding	Evaluation Method	Instrument / Setting	Citation
Electronic training programs for risk assessment are equally as effective as face-to-face programs as measured by performance in a post-training skill acquisition test. However, the estimated per-trainee	Practitioner survey (n=87)	Ontario Domestic Assault Risk Assessment (ODARA)	Hilton & Ham (2015)

cost of electronic training was one-third of the cost.			
Risk assessments lead to more appropriate allocation of intervention resources that are matched to risk level.	Case file analysis (n=464)	Structured Assessment of Violence Risk in Youth (SAVRY)	Vincent et al. (2012)

Conclusion

Our systematic sift for evidence regarding the user element of risk assessment and management found just 27 studies. To put this into context, our previously submitted literature review of the psychometric properties of risk assessment instruments covered 2,108 empirical studies. This is a huge disparity and reflects a real lack of focus on the human aspect to threat and risk assessment. Only 9 of the 27 studies have been published in the past five years which also suggests that this is not an emerging field of research which will inevitably be much greater in size in the coming years. Real concerted support for such designs and inquiries are required from funding agencies, practitioners, and academics alike to truly unpack the human element of risk assessment and management. The 27 studies we identified tentatively suggest that even if a risk assessment instrument had perfect psychometric properties, it might struggle in the field if the human element is overlooked. Likewise, a highly calibrated and well implemented human element can negate the raw edges of a badly designed risk assessment instrument. Our state of the science summary, based on the EMMIE framework, tentatively suggests the following conclusion:

- Risk assessment instruments **positively effect assessments** but we need to further understand and identify the additional cues (outside of the instrument) that practitioners use to make assessments as well as their scientific merits.
- These **positive effects occur via multiple mechanisms** and further research should compare these mechanisms' effect sizes in order to identify which should be prioritized for successful implementation.
- Likewise, there are **multiple site-specific and individual-user moderators** highlighting the need for evaluations to be context-specific and disentangle the effects from multiple levels of analysis.
- Successful risk assessment **requires concerted focus on implementation challenges** and highlights the need for greater inputs from research fields such as organizational psychology which have traditionally not traversed in this space.
- There is an **almost absent understanding of the cost-effectiveness** of risk assessment instruments which requires immediate focus and attention.

References

- Boccaccini, M. T., Chevalier, C. S., Murrie, D. C., & Varela, J. G. (2017). Psychopathy Checklist–Revised use and reporting practices in sexually violent predator evaluations. *Sexual Abuse, 29*(6), 592-614.
- Campbell, M. A., Gill, C., & Ballucci, D. (2018). Informing police response to intimate partner violence: Predictors of perceived usefulness of risk assessment screening. *Journal of police and criminal psychology, 33*(2), 175-187.
- Ceric, A., & Holland, P. (2019). The role of cognitive biases in anticipating and responding to cyberattacks. *Information Technology & People.*
- Chevalier, C. S., Boccaccini, M. T., Murrie, D. C., & Varela, J. G. (2015). Static-99R reporting practices in sexually violent predator cases: Does norm selection reflect adversarial allegiance?. *Law and Human Behavior, 39*(3), 209.
- Covernton, E. E., Moores, A., & Lowenstein, J. A. (2019). Changing clinicians' perceptions of the role that risk formulation and the HCR-20v3 play in the assessment and management of violence. *Journal of Forensic Practice.*
- De Beuf, T. L., de Vogel, V., & de Ruiter, C. (2020). Adherence to structured risk assessment guidelines: Development and preliminary evaluation of an adherence scale for the START: AV. *Journal of Forensic Psychology Research and Practice, 20*(5), 413-435.
- Dhaliwal, G. K., Porporino, F., & Ross, R. R. (1994). Assessment of criminogenic factors, program assignment, and recidivism. *Criminal Justice and Behavior, 21*(4), 454-467.
- Ennis, L., Hargreaves, T., & Gulayets, M. (2015). The integrated threat and risk assessment centre: A program evaluation investigating the implementation of threat management recommendations. *Journal of Threat Assessment and Management, 2*(2), 114.
- Franklin, A., Brown, S., & Brady, G. (2018). The use of tools and checklists to assess the risk of child sexual exploitation: lessons from UK practice. *Journal of child sexual abuse, 27*(8), 978-997.
- Garrett, T., & Rowe, R. (2004). Violence risk assessment training for local mental health services. *The British Journal of Forensic Practice.*
- Goodrum, S., Thompson, A. J., Ward, K. C., & Woodward, W. (2018). A case study on threat assessment: Learning critical lessons to prevent school violence. *Journal of Threat Assessment and Management, 5*(3), 121.
- Gough, K., Richardson, C., & Weeks, H. (2015). An audit of service-user involvement and quality of HCR-20 version 2 risk assessments on rehabilitation and low secure wards. *Journal of Psychiatric Intensive Care, 11*(S1).
- Gray, N. S., Tiller, J., & Snowden, R. J. (2019). WARRN—a formulation-based risk assessment procedure for Child and Adolescent Mental Health Services (CAMHS): the view of clinicians. *Journal of Forensic Practice.*
- Hilton, N. Z., & Ham, E. (2015). Cost-effectiveness of electronic training in domestic violence risk assessment: ODARA 101. *Journal of Interpersonal Violence, 30*(6), 1065-1073.
- Judge, J., Quayle, E., O'Rourke, S., Russell, K., & Darjee, R. (2014). Referrers' views of structured professional judgement risk assessment of sexual offenders: A qualitative study. *Journal of sexual aggression, 20*(1), 94-109.

- Kewley, S., Beech, A., Harkins, L., & Bonsall, H. (2015). Effective risk management planning for those convicted of sexual offending. *Journal of Aggression, Conflict and Peace Research*.
- Murrie, D. C., Boccaccini, M. T., Turner, D. B., Meeks, M., Woods, C., & Tussey, C. (2009). Rater (dis) agreement on risk assessment measures in sexually violent predator proceedings: Evidence of adversarial allegiance in forensic evaluation?. *Psychology, Public Policy, and Law*, 15(1), 19.
- Murrie, D. C., Boccaccini, M. T., Caperton, J., & Rufino, K. (2012). Field validity of the Psychopathy Checklist–Revised in sex offender risk assessment. *Psychological Assessment*, 24(2), 524.
- Robinson, A. L. (2006). Reducing repeat victimization among high-risk victims of domestic violence: The benefits of a coordinated community response in Cardiff, Wales. *Violence against women*, 12(8), 761-788.
- Sher, M. A., & Gralton, E. (2014). Implementation of the START: AV in a secure adolescent service. *Journal of Forensic Practice*.
- Stohlman, S., Konold, T., & Cornell, D. (2020). Evaluation of threat assessment training for school personnel. *Journal of Threat Assessment and Management*.
- Storey, J. E., & Strand, S. (2017). The influence of victim vulnerability and gender on police officers' assessment of intimate partner violence risk. *Journal of family violence*, 32(1), 125-134.
- Storey, J. E., Gibas, A. L., Reeves, K. A., & Hart, S. D. (2011). Evaluation of a violence risk (threat) assessment training program for police and other criminal justice professionals. *Criminal Justice and Behavior*, 38(6), 554-564.
- Sundrum, F., & Browne, S. (2004). Risk factors for violence: an audit on the effect of a tutorial on risk assessment. *Irish journal of psychological medicine*, 21(1), 28-29.
- Vincent, G. M., Guy, L. S., Gershenson, B. G., & McCabe, P. (2012). Does risk assessment make a difference? Results of implementing the SAVRY in juvenile probation. *Behavioral sciences & the law*, 30(4), 384-405.
- Vincent, G. M., Guy, L. S., Perrault, R. T., & Gershenson, B. (2016). Risk assessment matters, but only when implemented well: A multisite study in juvenile probation. *Law and Human Behavior*, 40(6), 683.
- Watts, D., Bindman, J., Slade, M., Holloway, F., Rosen, A., & Thornicroft, G. (2004). Clinical assessment of risk decision support (CARDS): The development and evaluation of a feasible violence risk assessment for routine psychiatric practice. *Journal of Mental Health*, 13(6), 569-581.