

Testing the application of violent extremism risk assessment to individuals who have radicalised in Australia: The case of the VERA-2R

Author

Belton, Emma, Cherney, Adrian

Published

2024

Journal Title

Trends & issues in crime and criminal justice

Version

Version of Record (VoR)

DOI

[10.52922/ti77338](https://doi.org/10.52922/ti77338)

Rights statement

This work is covered by copyright. You must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a specified licence, refer to the licence for details of permitted re-use. If you believe that this work infringes copyright please make a copyright takedown request using the form at <https://www.griffith.edu.au/copyright-matters>.

Downloaded from

<https://hdl.handle.net/10072/431843>

Griffith Research Online

<https://research-repository.griffith.edu.au>



Australian Government

Australian Institute of Criminology

Trends & issues in crime and criminal justice

No. 690

Abstract | Violent extremism risk assessment tools have become an important way of understanding and managing the threat posed by radicalised offenders in custodial and community settings. The aim of this study was to examine the applicability, interrater reliability and predictive validity of the Violent Extremism Risk Assessment—Version 2 Revised (VERA-2R) against data on individuals who have radicalised in Australia. Two trained assessors—the authors—completed VERA-2R risk assessments for a sample of 50 extremist offenders. Results showed differences in risk profiles between individuals who were violent and non-violent. It was found that the VERA-2R had good interrater reliability but low predictive validity.

Testing the application of violent extremism risk assessment to individuals who have radicalised in Australia: The case of the VERA-2R

Emma Belton and Adrian Cherney

Introduction

A key aim of violence risk assessment is to evaluate the propensity that a person will act in a dangerous way and then to develop strategies to mitigate or manage those behaviours (Borum 2015; Hart & Logan 2011). One approach to violence risk assessment is the use of the structured professional judgement (SPJ) framework. SPJ is defined as an analytical method used to understand and mitigate the risk of interpersonal violence posed by individual people that is discretionary in essence but relies on evidence-based guidelines (Hart & Logan 2011). SPJ has been widely adopted for the assessment of terrorist offenders and extremist violence risk (Logan & Lloyd 2019). One violent extremism risk assessment tool adopting the SPJ approach is the Violent Extremism Risk Assessment—Version 2 Revised (VERA-2R; Pressman et al. 2018).

Celebrating
50 years

The VERA-2R is used in Australia across correctional and community settings. It is used to inform assessments of individuals subject to Commonwealth terrorist offender preventative legislation, such as continuing detention orders (ie High Risk Terrorist Offender scheme, or Division 105A of the Criminal Code). It is in this context that the use of the VERA-2R has come under intense scrutiny and debate (Corner & Taylor 2023b; Cubitt & Wolbers 2023; Independent National Security Legislation Monitor 2023). The general consensus is that more research is needed on the VERA-2R and its application to violent extremists in Australia (Corner & Taylor 2023b; Cubitt & Wolbers 2022).

Given the influence of violent extremism risk assessment on decisions relating to pre- and post-detention, the management of extremist offenders and the forms of assistance provided to them, it is important that any tool provide valid and reliable assessments of risk. Briefly, validity as it pertains to violence risk assessment relates to whether an instrument (ie a specific tool) measures what it purports to measure. This includes, for example, whether a risk assessment tool can predict the risk of offending. Reliability, on the other hand, refers to the consistency of the tool in reaching risk judgements. This includes if there is consistency between assessors when it comes to their assessments of risk—termed interrater reliability (for more detail on specific validity and reliability criteria, see Groth-Marnat & Wright 2016: 117; Singh 2013).

There are debates and disagreements about approaches and methods to measure the validity and reliability of risk assessment tools (Singh 2013). This research does not set out to conclude whether the VERA-2R is more or less valid or reliable than other available violent extremism risk assessment tools. The aims of this project are:

- to examine whether the VERA-2R can help understand the types of factors that characterise violent extremists in Australia;
- to assess if the VERA-2R can discern between different risk levels in relation to violent and non-violent individuals; and
- to undertake a series of tests examining the interrater reliability and predictive validity of the VERA-2R.

In this paper we have provided a summary of the main results. Readers should refer to Cherney and Belton (2024) for more detail.

Violent Extremism Risk Assessment—Version 2 Revised

The VERA-2R is designed to capture contextual factors and risk indicators relevant to violent extremism as shown in Table 1 below (Pressman et al. 2018). The VERA-2R framework comprises 34 factors across five domains. Four of these domains reflect indicators of risk: *Beliefs, attitudes and ideology* (BA); *Social context and intention* (SCI); *History, action and capacity* (HAC); and *Commitment and motivation* (CM). A fifth domain, *Protective and risk mitigating* (P), considers protective factors. Eleven *Additional indicators* that capture background characteristics are also included (Pressman et al. 2018).

Table 1: VERA-2R domains and indicators
Beliefs, attitudes and ideology (BA)
BA 1 Commitment to ideology that justifies violence
BA 2 Perceived grievance and/or injustice
BA 3 Dehumanisation of designated targets associated with injustice
BA 4 Rejection of democratic society and values
BA 5 Expressed emotions in response to perceived injustice
BA 6 Hostility to national identity
BA 7 Lack of empathy and understanding for those outside one's own group
Social context and intention (SCI)
SCI 1 Seeker, user or developer of violent extremist materials
SCI 2 Target for attack identified (person, group, location)
SCI 3 Personal contact with violent extremists (informal or social context)
SCI 4 Expressed intention to commit acts of violent extremism
SCI 5 Expressed willingness and/or preparation to die for a cause or belief
SCI 6 Planning, preparation of acts of violent extremism
SCI 7 Susceptibility to influence, control or indoctrination
History, action and capacity (HAC)
HAC 1 Early exposure to violence-promoting, militant ideology
HAC 2 Network of family and friends involved in violent extremism
HAC 3 Violent criminal history
HAC 4 Strategic, paramilitary and/or explosives training
HAC 5 Training in extremist ideology in own country or abroad
HAC 6 Organisational skills and access to funding and sources of help
Commitment and motivation (CM)
CM 1 Motivated by perceived religious obligation and/or glorification
CM 2 Motivated by criminal opportunism
CM 3 Motivated by camaraderie, group belonging
CM 4 Motivated by moral obligation, moral superiority
CM 5 Motivated by excitement and adventure
CM 6 Forced participation in violent extremism
CM 7 Motivated by acquisition of status
CM 8 Motivated by a search for meaning and significance in life
Protective and risk mitigating (P)
P 1 Reinterpretation of the ideology
P 2 Rejection of violence as a means to achieve goals
P 3 Change in concept of the enemy
P 4 Participant in programmes against violent extremism
P 5 Support from the community for non-violence
P 6 Support from family members, other important persons for non-violence

Table 1: VERA-2R domains and indicators (cont.)

Additional indicators
CH Criminal history
CH 1 Client of the juvenile system/convicted for non-violent offence(s)
CH 2 Non-compliance with conditions or supervision
PH Personal history
PH 1 Violence in family
PH 2 Problematic upbringing and/or placed in juvenile care
PH 3 Problems with school and work
MD Mental disorder
MD 1 Personality disorder
MD 2 Depressive disorder and/or suicide attempts
MD 3 Psychotic and schizophrenic disorder
MD 4 Autism spectrum disorder
MD 5 Post-traumatic stress disorder
MD 6 Substance use disorder

When applying the VERA-2R, users are required to score each indicator listed in the four main risk domains (BA, SCI, HAC and CM) according to the ratings of ‘high’, ‘moderate’, and ‘low’. For the indicators listed in the *Protective and risk mitigating* domain (P), users are required to score them in reverse, in that low scores indicate the absence of a protective factor. When it comes to the *Additional indicators*, they are rated as either ‘present’ or ‘absent’. The authors of the VERA-2R state that a range of information sources should be used when applying the VERA-2R. If no information relating to a particular indicator is present, then it should not be rated. Drawing on this information and the use of SPJ, VERA-2R users are advised to interpret the weighting of the indicators to arrive at an overall risk assessment and case formulation. The authors of the VERA-2R state that this does not comprise a numerical score relating to a risk level. In the VERA-2R manual this overall risk score is stated as either low, moderate or high, and it is also suggested that an alternative approach can include low, low-moderate, moderate, moderate-high or high risk ratings. For more detail on the VERA-2R scoring method, refer to Cherney and Belton (2024); de Bruin et al. (2022); and Pressman et al. (2018).

Project methodology

Methods

The research sample for this project was selected from the Profiles of Individual Radicalisation in Australia (PIRA) database. The PIRA captures data on individuals who have radicalised in Australia from 1985 to 2022. PIRA draws on a range of open sources and includes individuals who adhere to Islamist, far-right, far-left and single-issue ideologies. Based on assessments of the reliability of collected source materials and according to a codebook (see Belton & Cherney 2023), individuals in PIRA are coded across 122 variables relating to background, demographics, group affiliation and contextual information. Data are first qualitatively captured and then transferred into quantitative indicators. It is this mix of data that was used to conduct VERA-2R assessments on the selected sample. There are limitations in using open-source materials to compile profiles of individuals who have radicalised to extremism, relating to impartiality, accuracy and completeness. There can be potentially missing information, for instance, compared with closed sources or psychological evaluations that may include information related to risk assessment. However, open sources can be just as detailed as information derived from closed sources—for example, police files (Gill et al. 2019). For more detail on the PIRA database including inclusion criteria, coding framework and interrater reliability, see Belton, Cherney and Zahnow (2023).

Sample and data

VERA-2R risk assessments were completed on a total of 50 cases captured in the PIRA sample, consisting of individuals who had participated in a range of terrorism-related acts. This sample size was chosen to limit assessor fatigue. We chose cases with the largest number of data sources and the most reliable information to enhance the accuracy of the assessments by minimising the possibility of missing data. To identify PIRA cases, they were first ranked according to the number of sources used, then by those with available court data (eg sentencing transcripts), which is considered to be the most reliable form of open-source data. On average, cases were compiled using 29 separate data sources ($SD=13.04$), with 88 percent ($n=44$) having available court data.

The above selection process resulted in a sample comprising 37 violent and 13 non-violent individuals ($N=50$). Violent extremists in PIRA are defined as individuals who attempted or actually engaged in ideologically motivated behaviour intending to cause harm, injury or death. To be considered violent, there had to be clear evidence of operational plans that were aimed at engaging in violent acts. Those considered to be non-violent were individuals who came to hold radical views, who did not engage in violent action but who did undertake non-violent activities motivated by an extremist ideology (eg financing a terrorist organisation or viewing, compiling and disseminating extremist material). Hence the categorisation of 'violent' includes those conducting violent acts themselves and individuals who engaged in preparatory acts of violence. Existing studies have made similar definitional distinctions between violent and non-violent extremists when analysing comparable datasets (see Becker 2019; Schuurman & Carthy 2023).

The inclusion of non-violent extremists ensured the sample represented individuals from a range of backgrounds who participated in ideologically motivated behaviours that are not always characterised by violence. This sample was selected to ensure the data as closely as possible reflected the variation and complexity of risk profiles and hence varying levels of risk based on behavioural outcomes. Islamist and far-right aligned individuals were included to further ensure the data captured a range of ideological motivations. Most cases, however, were aligned with Islamist or jihadist ideologies (88%), with only six far-right cases included (12%).

Coding VERA-2R indicators

Qualitative information in the PIRA database was used to measure the VERA-2R risk indicators as being either low, moderate or high. PIRA data were manually assessed against the VERA-2R risk domains. Identical to what is stated in the VERA-2R manual, specific indicators for the domains *Beliefs, attitudes and ideology* (BA), *Social context and intention* (SCI), *History, action and capacity* (HAC), *Commitment and motivation* (CM) and *Protective and risk mitigating factors* (P) were scored—ie coded—on a rating scale as 0=low, 1=moderate, or 2=high. The remaining 11 *Additional indicators* were scored on a dichotomous scale as 1=present or 0=absent.

Both authors attended a three-day official training course conducted by the Commonwealth Department of Home Affairs on the practical use and application of the VERA-2R risk assessment tool and were given copies of the manual authored by Pressman and colleagues, and assessor forms that are used in the field. Scoring practices were informed by the VERA-2R manual and this training, and guidelines detailing specific instructions, parameters and explanations for scoring each of the VERA-2R indicators. To ensure that the assessments closely reflected how the VERA-2R is applied and used in practice, the authors liaised with VERA-2R trainers and other users to clarify the correct scoring of indicators. Both the first and second author assessed the sample of 50 extremists separately and blind. Once the ratings for all indicators for each profile were completed, a final risk judgement rating was given. Both authors gave either a final rating of *low, moderate or high*.

Results

Reliability testing of the VERA-2R

The level of consistency among ratings of the VERA-2R was examined through a range of analysis to test for interrater reliability. This was analysed using four measures: percentage agreement, Cohen's kappa, Krippendorff's alpha and intraclass correlation coefficient (ICC). We only report results from the ICC here because it is considered a more robust measure of interrater reliability (for a full explanation of all the measures of reliability and results from each measure, see Cherney & Belton 2024). The ICC was examined using a two-way random effects model and absolute agreement type. The interpretation of ICCs was based on critical values for single measures. Interpretations of the ICC value follow Koo and Li (2016) based on 95 percent confidence intervals of ICC values, where <0.50 is considered poor, 0.50 to 0.75 moderate, and 0.75 to 0.90 good, with values >0.90 rated as excellent. It has been pointed out that an ICC of 0.61 is desirable and commonly reported for established tools (Douglas & Reeves 2010), but ICC values above 0.75 are preferred (Vincent et al. 2012).

Across all the VERA-2R domains (including risk and protective domains and additional indicators) the mean ICC value was 0.86, indicating good interrater reliability. Overall, the average amount of agreement across these categories can be considered good (Koo & Li 2016). Tables 2 to 5 below present the mean ICC values for the four major VERA-2R risk domains as well as the ICC values for each risk indicator. The ICC values for risk judgement ratings between assessors indicated good reliability (0.83, 95% CI [0.72, 0.90], $p < 0.001$). Out of the four major risk domains, *Beliefs, attitudes and ideology* (BA) received the highest level of overall interrater reliability (ICC=0.87). All indicators showed good to excellent interrater agreement (see Table 2).

Beliefs, attitudes and ideology (BA)	ICC	95% CI
Mean domain BA	0.87	
BA 1 Commitment to ideology that justifies violence	0.83***	0.48, 0.83
BA 2 Perceived grievance and/or injustice	0.98***	0.96, 0.98
BA 3 Dehumanisation of designated targets associated with injustice	0.88***	0.79, 0.93
BA 4 Rejection of democratic society and values	0.87***	0.78, 0.92
BA 5 Expressed emotions in response to perceived injustice	0.90***	0.82, 0.94
BA 6 Hostility to national identity	0.85***	0.74, 0.90
BA 7 Lack of empathy and understanding for those outside one's own group	0.78***	0.64, 0.87

***statistically significant at $p < 0.001$

Note: $N=50$. 95% CI=95% confidence interval. ICC=intraclass correlation coefficient

Three of the seven indicators in the *Social context and intention* (SCI) domain demonstrated excellent interrater reliability and two had good interrater reliability (Table 3). However, this domain also demonstrated the poorest levels of interrater agreement for a single indicator. The indicator *Susceptibility to influence, control and indoctrination* (SCI 7) showed poor agreement (ICC=0.45). ICC values for the variable *Seeker, user or developer of violent extremist materials* (SCI 1) presented as having moderate interrater agreement. The average amount of agreement across these indicators can be classified as good (ICC=0.81; Table 3).

Indicators relating to *History, action and capacity* received the second highest level of overall interrater agreement (Table 4). The indicator with the lowest agreement between raters, *Training in extremist ideology in own country or abroad* (HAC 5), still demonstrated moderate interrater reliability (ICC=0.71).

Table 3: ICC values for social context and intention VERA-2R indicators

Social context and intention (SCI)	ICC	95% CI
Mean domain SCI	0.81	
SCI 1 Seeker, user or developer of violent extremist materials	0.63***	0.43, 0.78
SCI 2 Target for attack identified (person, group, location)	0.94***	0.90, 0.97
SCI 3 Personal contact with violent extremists (informal or social context)	0.89***	0.81, 0.94
SCI 4 Expressed intention to commit acts of violent extremism	0.94***	0.89, 0.96
SCI 5 Expressed willingness and/or preparation to die for a cause or belief	0.83***	0.72, 0.90
SCI 6 Planning, preparation of acts of violent extremism	0.96***	0.92, 0.97
SCI 7 Susceptibility to influence, control, indoctrination	0.45***	0.08, 0.68

***statistically significant at $p < 0.001$

Note: $N=50$. 95% CI=95% confidence interval. ICC=intraclass correlation coefficient

Table 4: ICC values for history, action and capacity VERA-2R indicators

History, action and capacity (HAC)	ICC	95% CI
Mean domain HAC	0.82	
HAC 1 Early exposure to pro-violence, militant ideology	0.80***	0.66, 0.88
HAC 2 Network of family and friends involved in violent extremism	0.92***	0.86, 0.95
HAC 3 Previous criminal violence	0.84***	0.72, 0.90
HAC 4 Strategic, paramilitary and/or explosives training	0.82***	0.67, 0.89
HAC 5 Training in extremist ideology in own country or abroad	0.71***	0.53, 0.82
HAC 6 Organisational skills, access to funding and sources of help	0.80***	0.67, 0.88

***statistically significant at $p < 0.001$

Note: $N=50$. 95% CI=95% confidence interval. ICC=intraclass correlation coefficient

The *Commitment and motivation* (CM) domain had the lowest average interrater reliability (ICC=0.80). The indicator *Motivated by acquisition of status* (CM 7) incurred the lowest ICC value, indicating moderate agreement, and this was bordering on poor reliability (ICC=0.55). One possible reason for this result relates to gaps in the data sources used to conduct assessments and that it is potentially a highly subjective domain given its focus on motivations. Hence it leads to greater variation in the judged relevance and intensity of certain indicators. Further commentary on this indicator is provided in the main report (see Cherney & Belton 2024).

Table 5: ICC values for commitment and motivation VERA-2R indicators

Commitment and motivation (CM)	ICC	95% CI
Mean domain CM	0.80	
CM 1 Motivated by perceived religious obligation and/or glorification	0.92***	0.86, 0.95
CM 2 Motivated by criminal opportunism	0.87***	0.78, 0.92
CM 3 Motivated by camaraderie, group belonging	0.74***	0.57, 0.84
CM 4 Motivated by moral obligation, moral superiority	0.76***	0.60, 0.85
CM 5 Motivated by excitement and adventure	0.85***	0.75, 0.91
CM 6 Forced participation in violent extremism	1.00***	1.00, 1.00
CM 7 Motivated by acquisition of status	0.55***	0.32, 0.71
CM 8 Motivated by a search for meaning and significance in life	0.68***	0.49, 0.80

***statistically significant at $p < 0.001$

Note: $N=50$. 95% CI=95% confidence interval. ICC=intraclass correlation coefficient

Patterns across VERA-2R risk domains

A combined dataset was created to assess risk patterns across the sample. That is, the two VERA-2R datasets of the same 50 cases that had been rated by both authors were combined into one dataset to allow for various forms of statistical analysis. This final dataset only included the four major risk domains (BA, SCI, HAC and CM) and the additional indicators. A defaulting system was created to minimise bias in the data (see Cherney & Belton 2024 for more detail).

In summary, the sample contained 50 extremists who were mostly male (94%) and between 15 and 50 years old ($M=25.96$, $SD=8.86$). A large portion were first or second generation immigrants (78%, $n=39$). At the time of their offending, just under half were married ($n=22$), 42 percent had not completed their high school education, and 10 percent had enrolled in tertiary education. Individuals in the sample were suspected of, or convicted for, a variety of ideologically motivated activities. This included espousing extremist rhetoric online, compiling and disseminating extremist materials, financing a terrorist organisation, facilitating overseas travel, logistical support such as procuring weapons, and conspiring or planning a domestic terrorist attack. Seventy percent ($n=35$) were formally charged and convicted of a terrorism-related offence.

The most common risk factors among the sample were associated with beliefs, attitudes and ideology; specifically, indicators that measured grievances (BA 1, BA 2 and BA 5) all received high ratings in over 66 percent of cases. Risk indicators relating to social networks, which in the VERA-2R are captured across two domains, were also reported as high (HAC 2: 64%, $n=32$; SCI 3: 68%, $n=34$). Measures of intent such as expressing the intention to commit, or planning, acts of violent extremism were also high (see Figure 2). This suggests that the *Beliefs, attitudes and ideology* (BA) and *Social context and intention* (SCI) domains appear to dominate risk factors associated with a sample of Australian extremists.

Results show some factors were less common and, in some cases, not present. Early exposure to militant ideologies and motivations relating to criminal opportunism and excitement and adventure were rare across the sample. There were no instances where individuals were judged as being forced to participate in extremism (CM 6; see Figure 4). Low prevalence of indicators seemed to cluster in the *Commitment and motivation* (CM) domain (see Figures 1 to 5 below). The most common additional indicators were related to problematic upbringings or being placed in juvenile care and issues with school or work. Substance abuse and mental disorders were evident. This indicates unstable personal backgrounds, with mental disorders present in the sample (see Figure 5).

Figure 1: Distribution of indicators across low, moderate and high risk ratings for beliefs, attitudes and ideology VERA-2R domain (N=50)

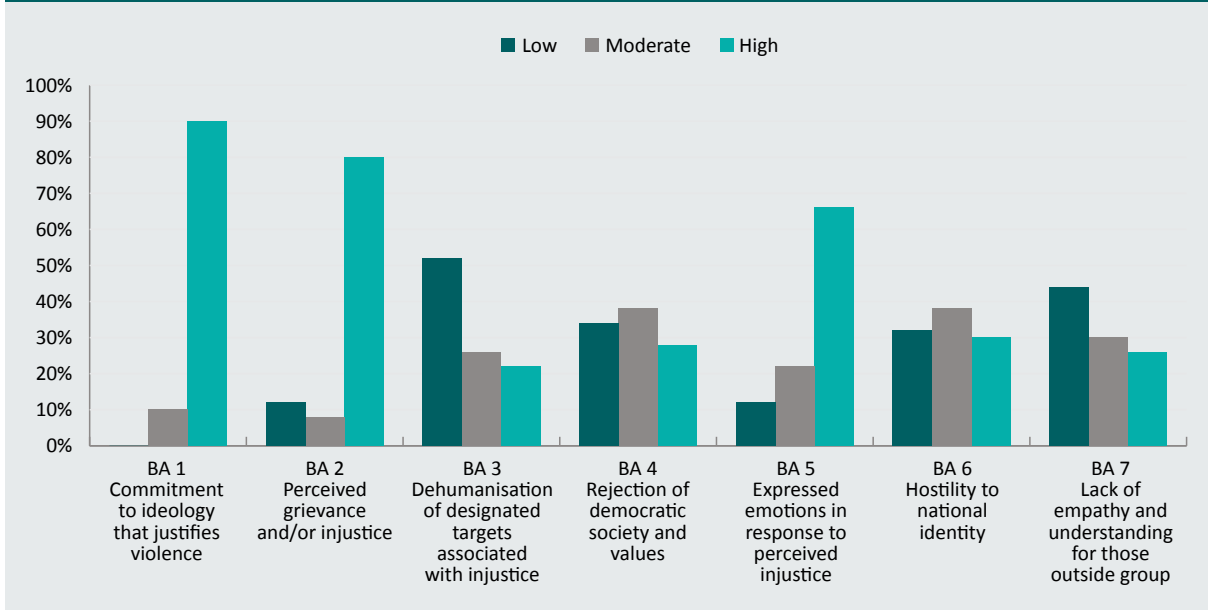


Figure 2: Distribution of indicators across low, moderate and high risk ratings for social context and intention VERA-2R domain (N=50)

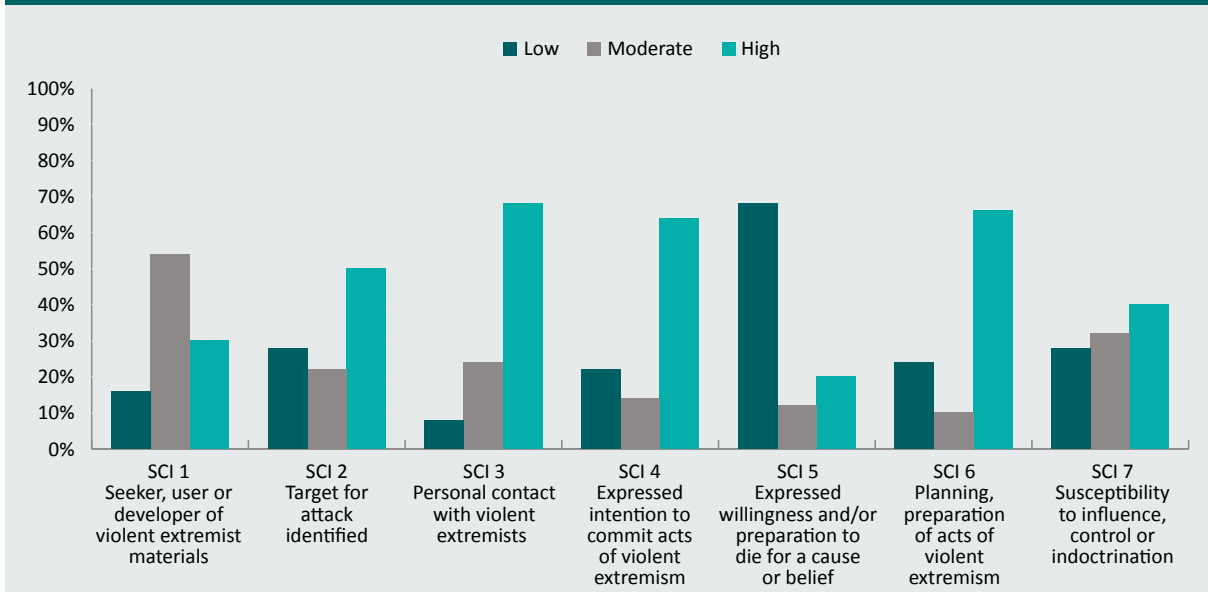


Figure 3: Distribution of indicators across low, moderate and high risk ratings for history, action and capacity VERA-2R domain (N=50)

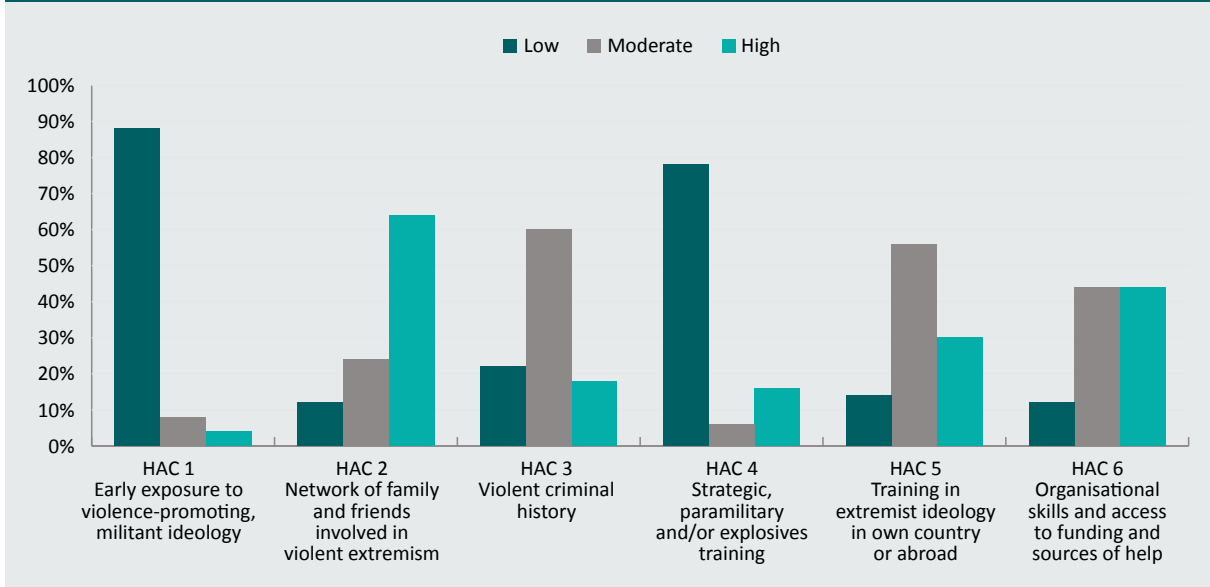


Figure 4: Distribution of indicators across low, moderate and high risk ratings for commitment and motivation VERA-2R domain (N=50)

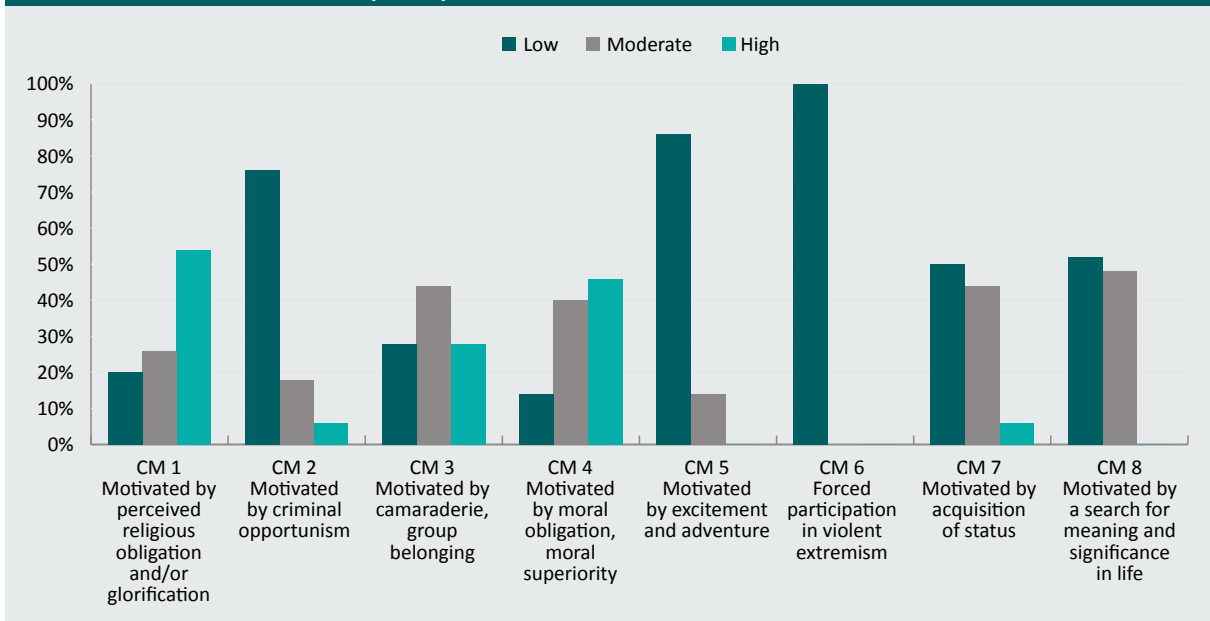
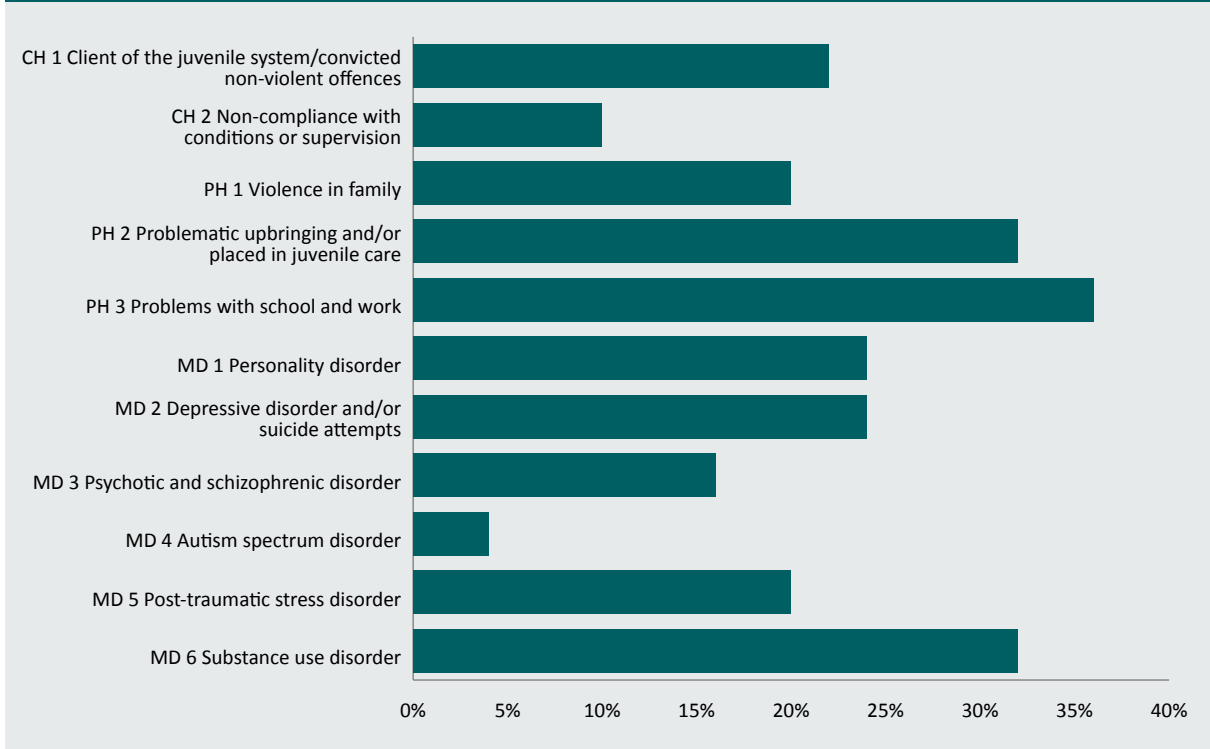


Figure 5: Prevalence of additional indicators (N=50)



VERA-2R indicators across violent compared with non-violent extremists

A chi-square test was used to determine the association between risk indicators for the four major VERA-2R domains and the two groups of extremists. For cell frequencies fewer than five, Fisher's exact test was reported instead of chi-square. Phi coefficients were calculated as a measure of effect size and were interpreted as 0.10=small, 0.30=medium and 0.50=large. Results from analysis (Table 6) show four indicators to be associated with violent extremists, all of which were part of the *Beliefs, attitudes and ideology* domain. Violent extremists were more likely to have a perceived grievance ($p=0.033$, $\phi=0.342$) and express emotions in response to these grievances ($p=0.033$, $\phi=0.342$), compared with those who were classified as non-violent. Violent extremists were also more likely to dehumanise designated targets associated with perceived injustice compared with non-violent individuals ($p=0.050$, $\phi=0.296$). Compared with non-violent extremists, individuals who were categorised as violent demonstrated hostility towards a national identity ($p=0.050$, $\phi=0.278$).

Non-violent extremists were significantly more likely to have received paramilitary or explosives training ($p=0.023$, $\phi= -0.346$) than those who were classified as violent. Upon closer inspection of the data, it was found that those who received training were involved in facilitating overseas travel, and producing and disseminating terrorist manuals, and were leaders of terrorist organisations.

Table 6: Major risk VERA-2R indicators association with violent and non-violent extremists					
VERA-2R domain risk indicators	Violent (n=37)		Non-violent (n=13)		Sig (p)
	n	%	n	%	
Beliefs, attitudes and ideology indicators (BA)					
BA 1 Commitment to ideology that justifies violence	37	100.0	13	100.0	
BA 2 Perceived grievance and/or injustice	35	94.6	9	69.2	*
BA 3 Dehumanisation of designated targets associated with injustice	21	56.8	3	23.1	*
BA 4 Rejection of democratic society and values	25	67.6	8	61.5	
BA 5 Expressed emotions in response to perceived injustice	35	94.6	9	69.2	*
BA 6 Hostility to national identity	28	75.7	6	46.2	*
BA 7 Lack of empathy and understanding for those outside one's own group	22	59.5	6	46.2	
Social context and intention indicators (SCI)					
SCI 1 Seeker, user or developer of violent extremist materials	30	81.1	12	92.3	
SCI 2 Target for attack identified	28	75.7	8	61.5	
SCI 3 Personal contact with violent extremists	34	91.9	12	92.3	
SCI 4 Expressed intention to commit acts of violent extremism	30	81.1	9	69.2	
SCI 5 Expressed willingness and/or preparation to die for a cause or belief	12	32.4	4	30.8	
SCI 6 Planning, preparation of acts of violent extremism	29	78.4	9	69.2	
SCI 7 Susceptibility to influence, control or indoctrination	29	78.4	7	53.8	
History, action and capacity indicators (HAC)					
HAC 1 Early exposure to violence-promoting, militant ideology	5	13.5	1	7.7	
HAC 2 Network of family and friends involved in violent extremism	33	89.2	11	84.6	
HAC 3 Violent criminal history	29	78.4	10	76.9	
HAC 4 Strategic, paramilitary and/or explosives training	5	13.5	6	46.2	*
HAC 5 Training in extremist ideology in own country or abroad	32	86.5	11	84.6	
HAC 6 Organisational skills and access to funding and sources of help	32	86.5	12	92.3	

Table 6: Major risk VERA-2R indicators association with violent and non-violent extremists (cont.)

VERA-2R domain risk indicators	Violent (n=37)		Non-violent (n=13)		Sig (p)
	n	%	n	%	
Commitment and motivation indicators (CM)					
CM 1 Motivated by perceived religious obligation and/or glorification	30	81.1	10	76.9	
CM 2 Motivated by criminal opportunism	8	21.6	4	30.8	
CM 3 Motivated by camaraderie, group belonging	28	75.7	8	61.5	
CM 4 Motivated by moral obligation, moral superiority	32	86.5	11	84.6	
CM 5 Motivated by excitement and adventure	7	18.9	0	0.0	
CM 6 Forced participation in violent extremism	0	0.0	0	0.0	
CM 7 Motivated by acquisition of status	20	54.1	5	38.5	
CM 8 Motivated by a search for meaning and significance	20	54.1	4	30.8	

*statistically significant at $p < 0.05$

Note: Due to the small sample size the original 3-point scores for risk indicators (0=low, 1=moderate, 2=high) were converted into a single dichotomous variable measuring the presence or absence of a risk indicator variable (0=absence, 1=presence), where presence equates to moderate (1) and high ratings (2), and absence equals low (0) ratings

Predictive validity of the VERA-2R

As recommended in the literature, the validity of the VERA-2R was tested using a range of performance indicators: sensitivity and specificity, positive predictive values (PPV), negative predictive values (NPV) and receiver operating characteristic area under the curve (AUC) (Singh 2013).

To perform the four validity tests—sensitivity and specificity, positive predictive values, and negative predictive values—the final risk judgement ratings (low, moderate and high) were converted into a dichotomous outcome variable (high risk vs low risk). This was because these types of tests require a dichotomous indicator (Singh 2013). To meet this requirement we converted what was a 3-point judgement into a dichotomous measure, where 0=low risk and 1=high risk. We decided to combine low and moderate ratings into a low-risk category, and high ratings would represent the high-risk category. This allowed a base threshold for predictive validity to be produced. While categorising individuals who had moderate risk ratings as low (rather than moderate and high being combined), this possibly underestimates the level of risk for some offenders. However, it was decided this was more appropriate than combining moderate and high-risk categories to ensure our approach was conservative. There was an equal distribution between the categories of high risk ($n=25$, 50%) and low risk ($n=25$, 50%).

Sensitivity and specificity

Sensitivity and specificity tests are regarded as the simplest indicator to measure the predictive validity of an instrument (Singh et al. 2013). Scores calculate the proportion of individuals who engaged in violent acts and who were judged to be at high risk (sensitivity) and the proportion of individuals who engaged in non-violent acts who were judged to be at low risk (specificity). The results of tests of sensitivity showed that the VERA-2R had a sensitivity value of 57 percent. That is, 57 percent of violent extremists were judged as high risk. The tests for specificity found a value of 69 percent, indicating that 69 percent of non-violent extremists were correctly identified as low risk during assessments.

Positive and negative predictive values

Two additional performance indicators are positive predictive values (PPV) and negative predictive values (NPV), which measure high and low risk calibration (Singh 2013). These estimate the likelihood of the VERA-2R tool correctly predicting an outcome—in this instance, that someone who was assessed as high risk was violent (PPV), and that someone who was assessed as low risk did not offend violently (NPV). In this study, the former includes preparatory acts of violence (see above distinction between violent and non-violent extremists as captured in the sample). A PPV of 84 percent showed that, of the extremists classified as high risk using the VERA-2R, 84 percent were violent. The NPV showed that, among those assessed as low risk, 36 percent of cases were non-violent. That is, of the individuals who were categorised as low risk, 36 percent did not commit an act of violent extremism. It should be noted that these values are influenced by base rates. PPV and NPV rely heavily on the distribution of the outcome variable, which in this study was whether extremists were violent or non-violent. In our study, violent offending base rates were higher than non-violent offending base rates.

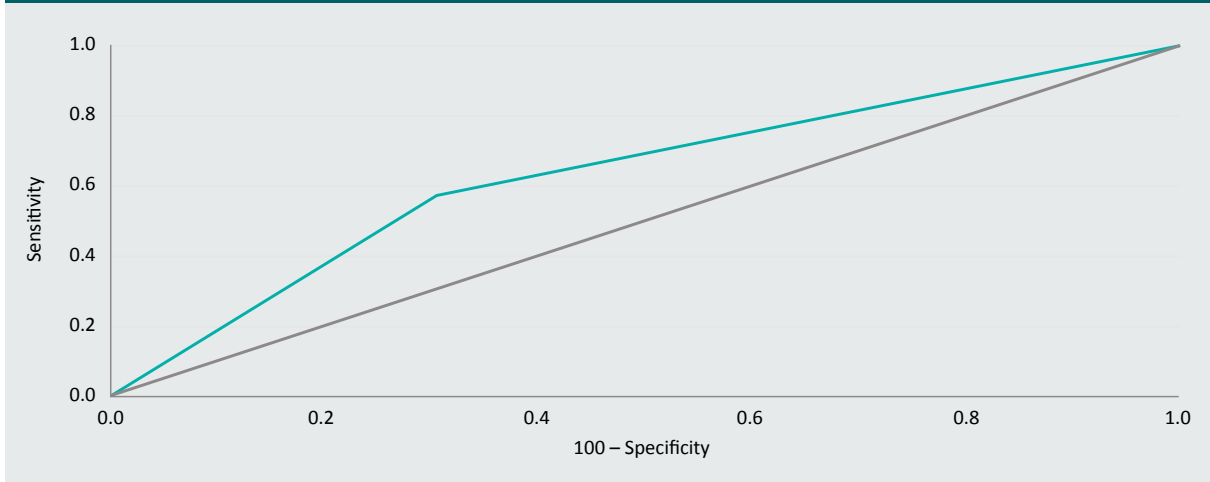
Area under the curve

The receiver operating characteristic analysis is another method of testing for predictive validity (Singh et al. 2013). It assesses the predictive accuracy of a tool by providing what is termed an area under the curve (AUC) statistic. It requires a dichotomous outcome (here violent vs non-violent extremist behaviour) and categorical risk judgements (Singh et al. 2013). Simply put, the AUC estimates the probability that a randomly selected violent individual received a higher risk classification on the VERA-2R than a randomly selected non-violent individual.

Consistent with the previous analysis, we converted risk judgements into a dichotomous value where VERA-2R judgements of low and moderate = 0 and high-risk judgements = 1. AUC values of >0.80 are considered 'excellent', 0.70 to 0.79 are regarded as 'acceptable' and 0.60 to 0.69 rated as 'weak'. For violence risk assessments, AUC values over 0.70 are considered 'moderate' while values over 0.75 are 'good' (Geraghty & Woodhams 2015). Risk classifications were measured using a dichotomous low vs high risk rating (see Figure 6). The AUC value here was 0.630 (95% CI [0.455, 0.805]), and the outcome was not significant ($p>0.1$), indicating the VERA-2R had poor predictive validity.

Caution should be taken when interpreting the AUC as it does not perfectly determine when a risk assessment tool is accurate enough to make practically useful distinctions between violent and non-violent subgroups (Mossman 2013). It should be noted that small sample sizes (less than 200) can impact the accuracy of the result (Singh 2013). Hence, when interpreting the validity test, readers should keep in mind the low power of small sample sizes such as the one employed in this study.

Figure 6: Area under the curve analysis for VERA-2R risk ratings (low vs high)



Conclusion

The results from this study show that the VERA-2R had good interrater reliability. The ICC values for interrater reliability indicated consistency in risk judgements and in the assessment of available information. This provides evidence that the VERA-2R can lead to consistent judgements, particularly when undertaken by people trained in administering the tool. However, a different result may have been achieved with a larger and more varied number of assessors. It should be pointed out that a minority of specific indicators demonstrated moderate and low interrater reliability. One reason for this may relate to the ambiguity and overlap surrounding some indicators (see Cherney & Belton 2024).

When using the VERA-2R domains to map trends across the sample, the results showed that some indicators do not seem to be associated with an Australian extremist cohort. This included, for example, early exposure to militant ideologies (HAC 1), motivation related to criminal opportunism (CM 2) or forced participation in extremism (CM 6). Among the sample, certain indicators were prevalent pertaining to the presence of grievances and the influence of social networks. Both factors have been identified in the international and Australian literature as playing a role in radicalisation to violent extremism (Bright, Whelan & Harris-Hogan 2020; Corner & Taylor 2023a; LaFree et al. 2018). Based on whether someone is violent or non-violent, particular risk factors can vary according to the VERA-2R, with violent individuals in the sample having a clustering of factors in the *Beliefs, attitudes and ideology* (BA) domain. This related to grievances and a sense of injustice, dehumanisation, emotional expressions in response to perceived injustice, and hostility to a national identity. It is perhaps the presence of these risk factors that help these individuals overcome inhibitions relating to the use of violence and provide justifications for violent action and are hence more pronounced.

This is relevant to scenario planning in relation to the case management of extremist offenders and the need to weight particular factors as being more important to violent outcomes and for them to be the focus of intervention work.

Our analysis indicates the VERA-2R tool has some capacity to delineate between levels of risk but was not necessarily predictive. That is, our analysis showed differences in VERA-2R risk indicators associated with violent and non-violent individuals, with violent individuals demonstrating specific behavioural and motivational traits. However, the predictive validity of the VERA-2R was found to be low. This result must be interpreted with caution because of our sample size. Also, any debate and conclusion made about the predictive validity of the VERA-2R needs to be qualified by the challenges faced in predicting extremist offending and reoffending given its low base rate in Australia and the absence of data on recidivism. It also needs to be prefaced with a consideration of factors that influence engagement in, and disengagement from, violent extremism. This requires tools and approaches that focus not only on risks but also on individual needs and that aim to promote behavioural change relating to disengagement from violent extremism.

Acknowledgements

The project was funded through a Criminology Research Grant: CRG 40/21–22.

References

URLs correct as at December 2023

Becker MH 2019. When extremists become violent: examining the association between social control, social learning, and engagement in violent extremism. *Studies in Conflict and Terrorism* 44(12): 1104–1124. <https://doi.org/10.1080/1057610X.2019.1626093>

Belton E & Cherney A 2023. *Profiles of Individual Radicalisation in Australia (PIRA) Codebook* (draft). University of Queensland

Belton E, Cherney A & Zahnow R 2023. Profiles of Individual Radicalisation in Australia (PIRA): Introducing an Australian Open-Source Extremist Database. *Perspectives on Terrorism* 17(1): 18–35

Borum R 2015. Assessing risk for terrorism involvement. *Journal of Threat Assessment and Management* 2(2): 63–87

Bright D, Whelan C & Harris-Hogan S 2020. On the durability of terrorist networks: Revealing the hidden connections between jihadist cells. *Studies in Conflict & Terrorism* 43(7): 638–656. <https://doi.org/10.1080/1057610X.2018.1494411>

Cherney A & Belton E 2024. *Testing the reliability and validity of the VERA-2R on individuals who have radicalised in Australia*. Report to the Criminology Research Advisory Council. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/crg77345>

Corner E & Taylor H 2023a. *Grievance-fuelled violence: Modelling the process of grievance development*. Research Report no. 27. Canberra: Australian Institute of Criminology. <https://doi.org/10.52922/rr78917>

- Corner E & Taylor H 2023b. *Testing the reliability, validity and equity of terrorism risk assessment instruments*. Centre for Social Research and Methods, Australian National University (FOI released document version). <https://www.homeaffairs.gov.au/foi/files/2023/fa-230400097-document-released-part-1.pdf>
- Cubitt T & Wolbers H 2023. *Review of violent extremism risk assessment tools in Division 104 control orders and Division 105A post-sentence orders*. Canberra: Australian Institute of Criminology. <https://www.aic.gov.au/publications/special/special-14>
- de Bruin A, Duits N, Kempes M & Prinsen M 2022. *Interrater and intrarater reliability of the Violent Extremism Risk Assessment tool*. Netherlands: Custodial Institutions Agency, Ministry of Justice, Science and Education
- Douglas KS & Reeves KA 2010. Historical-Clinical-Risk Management-20 (HCR-20) Violence Risk Assessment Scheme: Rationale, application, and empirical overview. In RK Otto & KS Douglas (eds), *Handbook of violence risk assessment*. Routledge/Taylor & Francis Group: 147–185
- Geraghty K & Woodhams J 2015. The predictive validity of risk assessment tools for female offenders: A systematic review. *Aggression and Violent Behavior* 21. <https://doi.org/10.1016/j.avb.2015.01.002>
- Gill P, Corner E, McKee A, Hitchen P & Betley P 2019. What do closed source data tell us about lone actor terrorist behavior? A research note. *Terrorism and Political Violence*. <https://doi.org/10.1080/09546553.2019.1668781>
- Groth-Marnat G & Wright JA 2016 *Handbook of psychological assessment* (6th ed). Hoboken, NJ: John Wiley & Sons
- Hart SD & Logan C 2011. Formulation of violence risk using evidence-based assessments: The structured professional judgment approach. In P Sturmey & M McMurrin (eds), *Forensic case formulation*. Chichester, UK: Wiley-Blackwell: 83–106
- Independent National Security Legislation Monitor 2023. *Review into Division 105A (and related provisions) of the Criminal Code Act 1995* (Cth). Commonwealth of Australia
- Koo TK & Li MY 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine* 15(2): 155–163
- LaFree G, Jensen MA, James PA & Safer-Lichtenstein A 2018. Correlates of violent political extremism in the United States. *Criminology* 56(2): 233–268
- Logan C & Lloyd M 2019. Violent extremism: A comparison of approaches to assessing and managing risk. *Legal and Criminological Psychology* 24: 141–161
- Mossman D 2013. Evaluating risk assessments using receiver operating characteristic analysis: Rationale, advantages, insights, and limitations. *Behavioral Sciences & the Law* 31(1): 23–39. <https://doi.org/10.1002/bsl.2050>
- Pressman E, Duits N, Rinne T & Flockton J 2018. *Violent Extremism Risk Assessment—Version 2 Revised*. Utrecht, Netherlands: Netherlands Institute of Forensic Psychiatry and Psychology

Schuurman B & Carthy SL 2023. Understanding (non) involvement in terrorist violence: What sets extremists who use terrorist violence apart from those who do not? *Criminology & Public Policy*. <https://doi.org/10.1111/1745-9133.12626>

Singh JP 2013. Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & the Law* 31(1): 8–22

Singh JP, Desmarais SL & Van Dorn RA 2013. Measurement of predictive validity in violence risk assessment studies: A second-order systematic review. *Behavioral Sciences & the Law* 31(1): 55–73

Vincent GM, Guy LS, Fusco SL & Gershenson BG 2012. Field reliability of the SAVRY with juvenile probation officers: Implications for training. *Law and Human Behavior* 36(3): 225–236

Emma Belton is a Research Fellow in the Griffith Criminology Institute at Griffith University.

Adrian Cherney is a Professor in the School of Social Science at the University of Queensland.

General editor, *Trends & issues in crime and criminal justice* series: Dr Rick Brown, Deputy Director, Australian Institute of Criminology. Note: *Trends & issues in crime and criminal justice* papers are peer reviewed. For a complete list and the full text of the papers in the *Trends & issues in crime and criminal justice* series, visit the AIC website: www.aic.gov.au

ISSN 1836-2206 (Online) ISBN 978 1 922877 33 8 (Online)

<https://doi.org/10.52922/ti77338>

©Australian Institute of Criminology 2024

GPO Box 1936
Canberra ACT 2601, Australia

Tel: 02 6268 7166

Disclaimer: This research paper does not necessarily reflect the policy position of the Australian Government

www.aic.gov.au