Name of Tool	Violence Risk Scale Second Version (VRS-2)
Category	Violence Risk (Validated)
Author / Publisher	Wong and Gordon
Year	2001

Description

• The VRS is a 26-item actuarial risk assessment tool designed to assess the risk of violent re-offending for incarcerated individuals and forensic psychiatric patients being considered for community access.
• The tool consists of six static and twenty dynamic variables. It can be used to monitor variations in risk and motivation to change. The second edition includes an item 'criminal personality' intended to capture the characteristics of psychopathic individuals (Dolan et al., 2008).

Age Appropriateness

18+

Assessor Qualifications

Can be used by workers within the criminal justice system. No professional qualifications required.

Assessors are required to undertake an intensive training course.

Strengths

• Assess risk of violence using a combination of static and dynamic (changeable) risk factors, the latter can be used to identify treatment targets.
• Assesses risk changes as a function of treatment or variations over time.
• Assesses treatment readiness/motivation which can inform approaches to treatment.
• A discretionary clinical over-ride is available for situations that are not captured by the risk factors found in the tool.
• Can be used with females, aboriginal, psychopathic and mentally disordered individuals.

Empirical Grounding

The VRS can be considered as a 'conceptual actuarial scale,' since the risk predictors are primarily derived from the *Psychology of Criminal Conduct*, a text by Andrews and Bonta (2010) that utilises personality, cognitive-behavioural and social learning perspectives to conceptualize the psychology of criminal behaviour. Its static and dynamic risk factors are empirically and/or theoretically related to violent recidivism (Wong and Gordon, 2006).

Inter-Rater Reliability									

a) UK Research	• Doyle et al. (2012) - the VRS total score attained an ICC value of .96 in a sample of male and female patients discharged from acute mental health units. • Dolan et al. (2008) reported high correlation coefficients for inter-rater reliability of the VRS composite score, static subscale and dynamic subscale (ICCs = .89, .96 and .85 respectively).
b) International Research	• Wong and Parhar (2011) found an ICC value of .93 for the VRS total score in a sample of Canadian males on parole or other forms of conditional release in the community. • Lewis, Olver and Wong (2013) reported ICC values ranging between .82 to .84 for the VRS total score in a sample of high risk male Canadian individuals who offended with significant psychopathic traits. • Zhang et al. (2012) reported an ICC of .80 for the VRS total score in a sample of male and female Chinese forensic inpatients in the province of Sichuan, all of whom were suffering from significant mental disorders.

Validation History

General Predictive Accuracy									

The VRS has been developed for use in criminal justice and forensic psychiatric settings (see section 'Applicability: Mentally Disorders).

Validation History

Applicability: Females									

a) UK Research	• Dolan et al. (2008) reanalysed their results removing the data of female participants from the sample; however, this did not significantly alter the previous findings. It should be taken into consideration that the female sample size was very small (n=11).
b) International Research	• In an unpublished PhD thesis, Stewart (2011) looked at the VRS ratings of 101 federally sentence women in Canada were followed up for approximately 7 years in the community. ICC was .98, AUCs for violent recidivism and institutional misconduct were .84 and .78 respectively.

Validation History

Applicability: Ethnic Minorities									

The normative sample consists of approximately 45% aboriginal males (Wong & Gordon, 2006).

Validation History

Applicability: Mental Disorders									

a) UK Research	• Dolan et al. (2008) reported moderate to high predictive accuracy of the VRS-2 with the occurrence of an aggressive incident in relation to the composite score, the static subscale score and the dynamic subscale score (AUCs = .69, .60 and .70 respectively). The authors tentatively recommend the use of VRS-2 to predict inpatient violence. • Dolan and Fullam (2007) - the VRS was able to discriminate violent and non-violent patients. As an effect size test used to indicate the standardised difference between two means, Cohen's d was equal to .72. Patients who had engaged in institutional violence in the following 12 months post-assessment had higher mean VRS composite and subscale scores than the non-violent group.
b) International Research	• Lewis, Olver and Wong (2013) reported positive results using the instrument with high risk individuals with psychopathic traits. In a fixed 3-year follow-up period (n=110), the VRS post-treatment total score was predictive of violent reconvictions (AUC =.65); however, the pre-treatment total score was not significant (AUC=.60); with a variable follow-up period (n=150), both pre- and post-treatment total scores were significant (AUC=.60, .64 respectively). • Wong and Parhar (2011) reported AUC values of .83 and .72 in predicting violent and any re-offence respectively after 7 years of prospective follow up in the community • Wong and Gordon (2006) – the VRS had attained high AUC values in predicting recidivism in the following domains: 'all convictions' (AUC = .74), 'violent convictions' (AUC = .75), and 'non-violent convictions' (AUC = .72).

Contribution to Risk Practice

- The VRS consists of 20 dynamic factors that can be used to assess risk and identify treatment targets, inform the formulation of risk management plans and in release decision-making
- The VRS incorporates the Stages of Change model within the dynamic risk factors to assess treatment readiness and risk change. Using the combination of dynamic risk factors and the assessment of treatment readiness and change, a VRS assessment can also inform the levels of monitoring and rehabilitation efforts and risk change over time or with treatment (Wong and Gordon, 2006; Lewis et al., 2012; Olver et al., 2013).
- A screening version of the tool was developed to highlight which individuals may require more in-depth assessments or to be used for brief intake evaluations (contact authors for more information, see below).

Other Considerations

- The second edition (VRS-2) was an experimental version so named when it was under development. The content of the VRS and VRS-2 are essentially the same with only minor changes. Currently, the VRS is the appropriate name for the tool.
- For training to use the tool clinically, and for additional research on the tool, please contact the authors (s.wong@sasktel.net or audrey.gordon@outlook.com).

Name of Tool	Classification of Violence Risk (COVR)
Category	Violence Risk (Validated)
Author / Publisher	Monahan and Colleagues
Year	2005

Description

- The COVR is a self-report interactive software programme that aims to estimate the level of violence risk posed by individuals diagnosed with a mental disorder over a period of several months, post-discharge into the community.
- The tool assesses patients on 44 risk factors in estimating violence risk (Monahan, 2010).
- A 'classification tree' methodology is used in the COVR, with questions being asked in a sequence until the individual is assigned a risk category (Monahan, 2010).
- After the assessor has completed an interview with the participant, the software generates a report that consists of statistical estimates of the likelihood of future violence, including the confidence interval for that estimation of violence (Monahan, 2010).
- In 2007, McClusker expressed uncertainties about the instrument. Counteracting that, Meadows (2014) reported a study that "confirmed that COVR scores were predictive of re-hospitalization or violent recidivism."

Age Appropriateness

18-60

Assessor Qualifications

Assessors must possess the relevant qualifications (i.e. in the administration and interpretation of psychological tests) and training in the use of this tool.

Strengths

- Assessment is fairly quick to administer
- Software based assessment that can eliminate sources of error and has the practicality of screening large samples (Snowden, et al. 2009).
- The COVR provides an estimate of risk that the practitioner can consider in relation to any other information they hold about the patient (Monahan, 2010).

Empirical Grounding

- The tool was developed from the MacArthur Violence Risk Study (Monahan et al., 2001). Variables from the study that predicted future violence were then used to classify participants into risk categories using the iterative Classification Tree System. This is an interactive model of violence,

considering combinations of risk factors in order to classify an individual into a risk level (Monahan, 2010).

•The COVR implements actuarial methods to a long-established modelling approach used in the medical field to inform professional judgements (Monahan et al., 2001).

Inter-Rater Reliability	■ ■ ■ ☐ ☐ ☐ ☐ ☐
a) UK Research	None available at present.
b) International Research	•Monahan (2010) - Using a dataset of 385 interviews, high kappa coefficients were found.

Validation History

General Predictive Accuracy	■ ■ ■ ■ ■ ■ ■ ■ ■
a) UK Research	•Meadows (2014) indicated: "The current study assessed success rates in predicting violent recidivism among forensic and civilly committed inpatients released from a state psychiatric hospital utilizing the violence risk categories generated from the COVR. This study confirmed that COVR scores were predictive of re-hospitalization or violent recidivism." •Snowden, Gray and Taylor (2010) found that applying the instrument in a UK sample indicated "COVR was a good predictor of both verbal and physical aggression. Its predictive ability was similar to that of the VRAG, although the VRAG was a better predictor of violence to property."
b) International Research	None available at present

Validation History

Applicability: Females	■ ■ ■ ☐ ☐ ☐ ☐ ☐

This tool can be applied to female adults who have offended; however, there is limited research validation relating to this population.

Validation History

Applicability: Ethnic Minorities	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

No Empirical Evidence Available.

Validation History

Applicability: Mental Disorders	
a) UK Research	• Doyle et al. (2010) - no significant findings were found in relation to the tool's predictive accuracy and inpatient violence in a 20-week post-discharge follow-up. • Snowden et al. (2009) - in a six-month follow-up the COVR presented the ability to predict physical aggression (AUC = .73); however, it was unable to predict property offences and verbal aggression.
b) International Research	• Sturup, Kristiansson and Lindqvist (2011) - in a 20-week follow-up period post-discharge, the COVR attained a high AUC value of .77 in predicting violent offences within a sample of 331 forensic mental health patients. • McDermott et al. (2011) - COVR has modest predictive accuracy in relation to physical aggression by psychiatric patients (AUC = .73) in the 20-week follow-up period subsequent to administration of the tool. • Monahan (2010) – the COVR has the ability to discriminate between low and high-risk groups. The estimated rate of recidivism was 1.2% for low risk and 63.6% for the high-risk group. The observed rates of recidivism in the prospective sample were 9% and 49% for the low and high-risk groups respectively. • Monahan et al. (2005) – the COVR obtained an AUC value of .70 (sensitivity = .75 and specificity = .77) in their follow up investigation of patients from the MacArthur study. • Persson et al. (2017) assessed 200 individuals (193 of which were followed-up) who were undergoing forensic psychiatric investigation in Stockholm. The predictive validity for the COVR tool was found to be modest with an AUC of .61.

Contribution to Risk Practice

• The tool is useful in aggregating the individual's self-reports in relation to the formulation of the risk of reoffending.
• Issues regarding the reliability of self-report (Snowden et al., 2009).
• The tool provides a statistical probability of short-term risk of recidivism (i.e. within a one-year period following discharge from a secure facility). No other supplementary information is included

that would provide a justification for the statistical information generated by the COVR (Snowden et al., 2009).
• Information taken from patient's files may be inaccurate or incomplete.
• Mixed findings regarding its predictive utility.
• The COVR states the risk category of an individual (e.g. 'low', 'high'); however, it does not clarify why an individual is placed at a certain level of risk (Snowden et al., 2009).
• Assessors should note that the COVR has been normed on mentally disordered populations; however, its predictive accuracy lessens depending on the type of recidivism being investigated.

Other Considerations

• Violent incidents are measured by patient self-report, official police records, hospital records, and collateral informants. Total scores are given in a probability format (a percent range for likely violence being committed within the next several months), a frequency format (e.g., for every 100 persons similar to the patient being assessed, between 20 and 32 will commit a violent act over the next several months), and a categorical format (classes of risk, including very low, low, average, high, and very high) (Kennedy et al., 2007).
• The COVR has been validated for clinical use with acute psychiatric patients who are being considered for release into the community and should be administered by practitioners in mental health disciplines. If the COVR was to be used with other populations, caution should be exercised (Monahan et al., 2005).
• Monahan and colleagues (2005) recommend that practitioners adopt a risk assessment procedure that is two-fold in nature: administer the COVR instrument; thereafter review the risk estimate generated by the tool. This will allow for additional considerations of risk or protective factors not covered in the COVR assessment. There is also the possibility for conflicting information in the patient's records. If this cannot be verified, it is advised the practitioner marks the answer as 'missing' (Monahan, 2010).

Name of Tool	Historical Clinical Risk-20 (HCR-20) – developed into HCR-20V3
Category	Violence Risk (Validated)
Author / Publisher	Webster and colleagues
Year	2013

Description

- The HCR-20 is a 20-item structured clinical guide for the assessment of violence risk intended for use with civil psychiatric, community, forensic, and criminal justice populations.
- The instrument has a tripartite temporal focus, comprising the following: ten historical variables ('H' Scale), looking at a history of problems with violent behaviours and attitudes, employment, relationships, mental and personality disorders and antisocial behaviours; five clinical variables ('C' Scale), highlighting recent or current problems with psychosocial, mental health and behavioural functioning; five risk management factors ('R' Scale), encompassing relevant past, present, and future considerations with regards to living conditions, services, personal support and stress. All of these scales should be reviewed regularly (Douglas et al., 2014).
- The third version of HCR-20 (HCR-20V3) was published in 2013 and the encompassing factor on personality now considers all disorder symptoms. The 'relevance rating' allows for the rating of the presence and relevance of each risk factor to be evaluated, allowing for assessments to be individualised (Logan, 2014).
- The HCR-20 prioritises cases as low/routine, moderate/elevated or high/urgent. A low/routine rating suggests the person is not in need or any special interventions or monitoring. Moderate/elevated risk indicates special management and increased monitoring is needed. The high/urgent prioritisation requires immediate action, which could include hospitalisation or suspending a conditional release (Brunt, 2013).

Age Appropriateness

18-65

Assessor Qualifications

Assessors must possess a degree, certificate or licence to practice within health care settings.

Assessors must also possess the necessary training and experience in the administration, scoring and interpretation of clinical behavioural assessment instruments and be familiar with professional and research literature in the study of violence. It is possible for a team of professionals to complete the tool: a psychiatrist could complete the items relating to mental illness; a psychologist could look at the personality disorder and psychopathy items; a social worker may complete items pertaining to social history and future plans (Douglas and Reeves, 2010).

Strengths

- Large research base.
- The HCR-20 has the capacity to guide clinical judgement about intervention and risk management (Gray et al., 2008).
- The inclusion of a clinical formulation in the HCR-20 exploring the motivating factors for violence and potential future risk scenarios affords the evaluator the opportunity to think about violence in real-world scenarios (Brunt, 2013).

Empirical Grounding									

- Research has shown the HCR-20 includes static and dynamic factors that have sound empirical grounding (Douglas et al., 2005).

- The HCR-20 has been subject to more than 200 empirical validations (Douglas et al., 2014).

Inter-Rater Reliability									
a) UK Research	• Doyle et al. (2014) found that "the HCR-20V3 demonstrated very good inter-rater reliability and significantly predicted community violence at six and twelve months post-discharge, with ROC AUCs of .73 and .70 respectively." • Gray et al. (2008) - ICC of .80 found for the HCR-20^{V2}. • Doyle and Dolan (2006) found ICC values of .85 and .83 for the clinical and risk management items of the HCR-20.
b) International Research	• Mills et al. (2007) - the original HCR-20 achieved an ICC value of .85 in a Canadian sample of incarcerated males. • Douglas et al.'s (2002-2008) review of previous research containing showed ICC value of .73 and above for the HCR-20 across different sample populations. • Douglas and Belfrage (2014) found inter-rater reliability was evident for the version 3 of HCR-20. • Green et al. (2016): "Results indicated higher inter-rater reliability on scoring risk factors among males as compared to females, calling for future research into the role of item indicators across genders and possible differences in interpretations of scoring guidelines." • Cawood (2017) found the inter-rater reliability of the HCR-20 V3 was significant with an ICC of .72.

Validation History

General Predictive Accuracy	
• The HCR-20 was developed from consideration of the empirical literature concerning factors that relate to violence. • There are 16 new sub-items in the Historical scale in version 3, which prompt the rater to look in more detail at a wider range of historical information (Doyle et al., 2014).	
a) UK Research	None available at present.
b) International Research	• Abbiati and colleagues (2014) applied risk assessment instruments to 52 violent offenders in a Swiss prison to evaluate predictions for physical, any and other misconduct. Total scores were good for physically violent misconduct (AUC=0.80), fair for any misconduct (AUC=0.72) and poor for other misconduct (AUC=0.67).

Validation History

Applicability: Females	
a) UK Research	• Coid et al. (2009) - the 'H' scale generated AUC values of .70 to .73 for female offenders.
b) International Research	• Garcia-Mansilla, Rosenfeld and Cruise (2011) - the total score for the 'H' and 'C' scales had moderate predictive accuracy for community violence (AUC= .60); although when separating the AUC value for the 'C' scale alone did not have significant predictive accuracy. • Schapp et al. (2009) - the HCR-20 score did not predict general and violent recidivism in female psychiatric patients. • Strub, Douglas and Nicholls (2016) study used a sample of 52 men and 48 women – "Results indicated that the HCR-20 as well as its components predicted both the occurrence and imminence of violent outcomes and gender did not moderate those relationships." • The HCR-20^{V3} was coded alongside other risk assessment tools to check predictive accuracy for 78 female forensic psychiatric patients over a period of 11.8 years. Findings suggest that the HCR-20^{V3} showed significant predictive accuracy. The clinical scale of the tool was significant for predicting violent recidivism (de Vogel, Bruggeman and Lancel, 2019).

Validation History

Applicability: Ethnic Minorities	■	■	■						

a) UK Research	• Snowden, Gray and Taylor (2010) - the HCR-20 generated moderate to high AUCs of .72 and .66 for White and Black mentally disordered offenders respectively.
b) International Research	• Fujii et al. (2005) - composite HCR-20 score achieved moderate to high AUC values for native Hawaiian and Euro-American groups (.73 and .64 respectively); although for Asian Americans the value was lower (.58). There were no significant differences between AUC values for these ethnic groups.

Validation History

Applicability: Mental Disorders	■	■	■	■	■	■	■	■	■

a) UK Research	• O'Shea et al. (2015) maintained that their study demonstrated that "after controlling for a range of potential covariates, the HCR-20 is a significant predictor of inpatient aggression in people with an ID (intellectual disability) and performs as well as for a comparison group of mentally disordered individuals without ID. The potency of HCR-20 subscales and items varied between the ID and comparison groups suggesting important target areas for improved prediction and risk management interventions in those with ID." • Coid et al. (2009) – the HCR-20 obtained moderate AUC values for violent recidivism and acquisitive reconviction in male offenders (.67 and .69 respectively). The HCR-20 also generated moderate to high predictive accuracy for female offenders. • Ho et al. (2009) - ROC analyses revealed that the 'H' scale had moderate to high predictive accuracy for predicting minor violence (AUC = .619), serious violence (AUC = .74), and any violent incidents (AUC = .61) in a psychiatric sample. • Lindsay et al. (2008) - the HCR-20 obtained a relatively high AUC of .72 in a sample of offenders with learning disabilities. • A survey of 43 mental health clinicians in a secure hospital found the historical and clinical subscales of the

	HCR-20 were perceived to be the most relevant to violence prediction (Dickens and O'Shea, 2017).
b) International Research	• Campbell, French and Gendreau (2009) - meta-analysis highlighted the predictive reliability of the HCR-20 in regard to institutional violent recidivism (K = 11, (n = 758) Z+ = .28). • Mills et al. (2007) found an AUC value of .73 in their pseudo-prospective study of 83 incarcerated males. • A study by Arai et al. (2016) examined the records of forensic psychiatric patients from 2008-2015 to test the predictive accuracy of the HCR-20. Results from ROC analyses indicate that the clinical and risk subscales of the HCR-20 showed good predictive accuracy, although the historical one failed to do so. • Sada and colleagues (2016) utilised the HCR-20 on 225 patients within a Mexican psychiatric facility. It was found that violent behaviour was more severe in the patients within the high-risk category, thus suggesting the HCR-20 is a suitable instrument to predict risk of violence. • Vitacco et al. (2016) assessed data from 116 forensic inpatients and found that higher scores in the risk scale of the HCR-20 had a link to a greater likelihood of not being released from or having to return to a forensic facility after release. The authors conclude that clinicians should perhaps consider community-based variables when evaluating forensic patients due to be released back into the community. • The predictive validity of the HCR-20 was examined in a sample of 136 forensic psychiatric patients in Australia. Findings showed that the total score, historical and risk management scales all had moderate to large positive correlations with reconvictions (Shepherd, Campbell and Ogloff, 2018). • Jeandarme et al. (2017) carried out a study in 3 forensic medium security units in Belgium. The results indicated that the HCR-20 only shown predictive accuracy for low risk individuals, whilst it was not accurate for high-risk patients.

Contribution to Risk Practice

- The HCR-20 has been translated into sixteen languages and is used across various continents: North and South America, Asia, Europe and Australia (Douglas and Reeves, 2010).
- The HCR-20 can identify a number of risk and responsivity factors relevant to the individual's risk of violent recidivism.
- Many of the factors identified by the tool can act as targets for treatment/change (e.g. insight, relationship factors) and the instrument can aid decisions regarding the level of monitoring and supervisory strategies, in relation to individuals who pose minimal to high levels of risk for recidivism.
- The HCR-20 can aid assessors in developing risk formulations and risk management strategies.
- Doyle et al. (2014) reports in a study of the third version: "Findings support the hypotheses that (1) the HCR-20 V3 and sub-scales can be coded with satisfactory agreement across different raters, and (2) patients with high scores at discharge on HCR-20 V3 were significantly more likely to be violent than service users with low baseline scores at six and 12 months post discharge in the community."
- The definition of violence provided with the HCR-20 extends to threatened and attempted violence. This means it could be useful to assess risk in cases of violence that do not involve physical harm such as stalking or causing psychological damage (Douglas and Reeves, 2010).

Other Considerations

- The time period for which an assessment is produced needs to be considered. Snowden and colleagues (2007) state that the 'C' scale of HCR-20 is found to be a good predictor of institutional violence over the next 3 months but a poor predictor of reconviction over a period of several years.
- The authors advise that the dynamic items (i.e. the clinical and risk management) are capable of indexing change. In addition, some of the Historical items may not necessarily be 'static' (e.g. changes in the offender's relationship or employment status) (Douglas et al., 2001).
- The HCR-20 does not provide numerical estimates of risk for violence. It is advised that assessors keep abreast of research about the impact of social factors on violence risk and to consider this when applying HCR-20 assessments across various social groups (Douglas and Reeves, 2010).
- Dr. Vogel has developed the Female Additional Manual (FAM) which forms an additional supplement to the HCR-20 in relation to assessing violence in women (Vogel et al., 2012; see the 'Responsivity Section').
- Few studies have used the categorical risk ratings to determine the predictive utility of the HCR-20 (de Vogel and de Ruiter, 2005).
- The HCR-20 should be completed using information obtained from interviews with the individual and other collateral information.
- The focus on mental health and the requirement that the assessor is well-versed in mental health interviews is a limitation of the HCR-20 instrument, making it best suited for use with those being managed or moving out of inpatient treatment facilities (Brunt, 2013).
- For more information on HCR-20 (Version 3) please visit: http://kdouglas.wordpress.com/hcr-20/hcr-20/

Name of Tool	Short Term Assessment of Risk and Treatability (START)
Category	Violence Risk (Validated)
Author / Publisher	Webster and colleagues
Year	2009

Description

- The START is a 20-item structured professional judgement tool designed to structure regular clinical assessments within inpatient and community contexts.
- The tool is intended to assess, document, communicate and manage risk across diverse settings.
- The 20 items included in the START are drawn from research that have shown these variables to be associated with seven risk estimates/adverse outcomes to individuals with mental health problems and personality dysfunctions, as well as persons who come into conflict with the law.
- Assessors code the items according to two scales presented in the tool: (1) Strength and (2) Vulnerability.
- The items are rated on a 3-point Likert scale from 0 to 2 and can be coded as both a strength and a vulnerability.
- The START includes seven risk estimates which include violence, suicide and self-harm. The risk estimates are derived from the consideration of the ratings from the strength and vulnerability scales.
- The tool was initially designed to capture dynamic vulnerabilities and strengths while generating a framework for periodic assessment of risk to inform clinical progress reviews. It should inform treatment, daily management and decision-making.
- The START is intended for use with adults diagnosed with mental, personality and substance-related disorders. It is relevant to inpatient and community psychiatric, forensic and correctional populations.

Age Appropriateness

16+

Assessor Qualifications

Experienced clinicians from a mental health background. Assessors are required to have participated in relevant training for this tool. It can be completed either by an individual practitioner or jointly by a clinical team via group discussion and reaching a consensus.

Strengths

- The tool considers strengths rather than being purely risk orientated (Nicholls et al., 2006).
- Collins et al. (2008) found that clinicians deemed START as appropriate, easy to use and clinically useful.
- START is intended for use in both inpatient units and outpatient services.

Empirical Grounding	

• The manual claims that the tool is grounded in the HCR-20 and relevant studies of acute violence (Webster et al., 2004). The authors drew upon research from civil psychiatry, forensic psychiatry and corrections reflecting studies from both institutional and community settings.

• The START is a concise clinical guide for the dynamic assessment of short-term (i.e. weeks to months) risk for violence (to self and others) and treatability.

Inter-Rater Reliability	
a) UK Research	• Timmins, Evans and Tully (2018) assessed the inter-rater reliability of START across disciplines, recruiting psychiatrists, mental health nurses, psychologists and occupational therapists to rate 20 case items and 7 risk estimates for two test cases. Good to excellent IRR was found for START items; whilst moderate-to-poor IRR was found for risk estimates amongst raters. There were clear differences between disciplines at item levels, highlighting the importance of collaborating as a team when completing risk assessments.
b) International Research	• Desmarais et al. (2012a) found ICCs of .93 for strength scores, .95 for vulnerability scores and .85 for risk estimates respectively. • Nicholls et al. (2006) – the START attained excellent inter-rater reliability (ICC) in various settings; Psychiatry (.80), Nursing (.88) and Social work (.92). • Dickens and O'Shea (2015) reported "Inter-rater reliability for coding the SOS (Start Outcome Scale) from progress notes was in the excellent range: Cohen's Kappa ranged from .83 to 1.00, the lowest being for self-neglect and the highest for self-harm and physical aggression." • Viljoen et al. (2011) - the START strength and vulnerability scale total scores attained good ICC values of .62, and .68. • Wilson et al. (2010) - found ICCs of .85, .90 and .81 for the strength, vulnerability and risk estimates respectively. • Crocker et al. (2011) carried out START assessments on 42 individuals at a civil psychiatric unit in Canada. An inter-rater reliability check on six patients six months later found that there was low IRR for total risk score of .38, whereas total strength score was strong at .81.

	•A study by O'Shea, Picchioni and Dickens (2016) of 22 adults in a secure mental hospital found that the inter-rater reliability for START items was in the excellent range.

Validation History
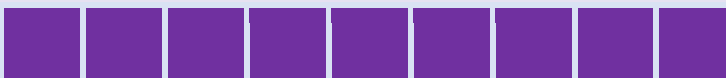
General Predictive Accuracy	
a) UK Research	•Braithwaite et al. (2010) suggested there was partial support for the predictive validity of the instrument. Both the strength and vulnerability scales significantly predicted aggression against others and suicidality (AUC= .65 for each scale and behaviour). AUCs of .67 and .63 were generated for substance abuse in the strength and vulnerability scales respectively. Neither scale, however, significantly predicted the occurrence of self-harm, suicidality, self-neglect or victimisation (AUCs ranging from .52 - .58). •Gray et al. (2011) tested the START in a limited population study of 51 mentally disordered patients. The SPJ scores were able to predict violence to others, verbal aggression, self-harm and victimisation (AUCs of .65, .70, ,86 and .67 respectively). The strength and risk scores varied in their ability to predict certain behaviours. The strength scores were poor predictors for all behaviours bar self-harm (AUC= .61), with an AUC range of .21-.47. The risk scores were better predictors with an AUC range of .60-.74 for all behaviours; the only exception to this is for self-harm which generated an AUC of .48.
b) International Research	•Crocker et al. (2011) found that whilst START total risk scores showed good predictive accuracy in relation to physical aggression for periods of 1 and 3 months (AUC ranging from .65-.77), they were not as accurate for the long-term of 6 to 12 months. Individuals displaying physical and property aggression had higher risk and lower strength scores on the START. •O'Shea, Picchioni and Dickens (2016) found that the inclusion of strengths improved the predictive accuracy of the START tool. The percentage of cases correctly classified increased from 0.6% to 4.4%. The specific risk estimates scale showed increased predictive accuracy over both the vulnerability and strength scales, showing moderate to large predictive accuracy for all behaviours (AUCs range from .640-.783), bar self-neglect (AUC of .546).

- In a study examining aggression data retrieved from institutional records, START strength and vulnerability total scores predicted all forms of aggression, bar physical aggression towards objects, Moderate to large effect sizes were generated for any aggression, verbal aggression and physical behaviours (others) with AUCs ranging from .65-.90. For physical aggression against objects, an AUC of .62 was generated in the strength total score (Cartwright et al., 2018).

- de Vogel, Bruggeman and Lancel (2019) coded file information for 78 female forensic psychiatric patients using a number of structured professional judgement tools. The START Vulnerability scores showed moderate and large predictive accuracy for all recidivism in medium and long term follow-ups (AUCs of 0.748 and 0.698 respectively), as well as for violent recidivism (AUCs of .697 and .704 for medium and long term respectively).

Validation History

Applicability: Females	██	██	██	██	██	██			

a) UK Research	• O'Shea and Dickens (2015) found START was a stronger predictor of aggression and self-harm in women than men. • Quinn et al. (2013) found significant predictive validity for adverse incidents at the one month time point and this then diminished over time. Females were rated as having significantly less strengths and more risks than males.
b) International Research	• Viljoen et al. (2011) - in a 3-year follow-up in a sample of female forensic patients, the START strength and vulnerability scores showed moderate to large AUCs at .70 and .80 respectively The results show the START scales were predictive of successful reintegration into the community (defined as the absence of readmission to hospital and the presence of an absolute discharge decision) in a sample of female forensic patients.

Validation History

Applicability: Ethnic Minorities									

No Empirical Evidence Available

Validation History

Applicability: Mental Disorders	
a) UK Research	• Gray et al. (2011) - the vulnerability scale was moderately predictive of violence to others (AUC =.68). The strength scale had a significant negative correlation with violence to others (r=.-.42) and a corresponding AUC value of .21. Low scores on the strength scale were, thus, predictive of violence. • Predictive validity of START was evident when Marriott et al. (2017) administered the tool to 527 inpatients within a secure mental health facility in the United Kingdom. • Quinn et al. (2013) discovered that START scores were capable of distinguishing between those with mental disorders at the various stages of their care pathways. • Alderman, Major and Brooks (2016) used the START to examine 4559 aggression recordings related to 76 patients with an acquired brain injury. The START risk of violence was classed as low and high for 50% and 13.7% of the sample respectively; suggesting the need for specific tools to be developed for use in patients with ABI.
b) International Research	• Crocker et al. (2011) carried out a longitudinal study, which indicated that START was well integrated into a Canadian unit's administrative activities. • Wilson et al. (2010) - in a 12-month follow up, the strength and vulnerability total scores and the final risk estimates significantly predicted any aggressive acts with AUCs ranging from .82 to .89. • Chu et al. (2011) - in a 1-month follow period, the START vulnerability total scores attained high AUC values in predicting inpatient aggression (.76), interpersonal violence (.78) and verbal threat (.77). Similarly the strength total scores predicted inpatient aggression (.71) and interpersonal violence (.75) but not verbal threat. • Braithwaite et al. (2010) - the vulnerability scale significantly predicted physical aggression against others (AUC = .66) in a 2-year follow-up period. • Nicholls et al. (2006) - START generated moderate to high AUC values for a broad range of aggressive behaviours in a psychiatric hospital: verbal aggression against others (.72), physical aggression against objects

	(.67), physical aggression against others (.70) and sexual inappropriateness.

Contribution to Risk Practice

- The START has the ability to create awareness of risk factors and strengths presented by the individual. Findings from previous research also suggest that it may be useful for distinguishing between types of patients (Nicholls et al., 2006; Quinn et al., 2013).
- The START includes dynamic factors and the strengths of individuals, which could inform offence analyses and risk formulations.
- The use of the START can help identify factors that are important targets for treatment, intervention and management planning.
- Repeated assessments using the START can aid assessors in monitoring changes in risk level and identify necessary changes in risk management strategies.
- The START can aid the assessor in examining potential improvement/deterioration in identified risk, responsivity and protective factors which, in turn, can also inform risk management strategies. Further, the tool allows for other harmful scenarios to be considered for individuals, e.g. suicide, substance abuse, self-harm and self-neglect (O'Shea, Picchioni and Dickens, 2016).
- O'Shea and Dickens (2015) reported: "The study provides limited support for the START by demonstrating the predictive validity of its specific risk estimates for substance abuse and unauthorised leave. High negative predictive values suggest the tool may be of most utility in screening out low risk individuals from unnecessary restrictive interventions; very low positive predictive values suggest caution before implementing restrictive interventions in those rated at elevated risk."
- START is routinely used within forensic mental health populations in the United Kingdom and is recommended by the *Department of Health.*
- Staff members at a forensic high secure unit in Norway were surveyed about the START. It was felt by 68% of respondents that the existing and potential needs of patients were covered by the tool. Moreover, 73% agreed that using the START tool contributed to a more systematic risk assessment and management process (Kroppan et al., 2011).

Other Considerations

- Doyle et al. (2008) reported uncertainty over time frame in which risk and strengths are applied from a survey conducted with users of the START.
- Dickens and O'Shea (2015) suggested for lower risk patients assessment at 3 month intervals was appropriate. For those with elevated risk rating more frequent assessments were warranted.
- The START can be completed by a single clinician or by the patient's multi-disciplinary team.
- Fewer validation studies have been conducted on samples that consist solely of female patients and patients of other ethnic backgrounds.
- An electronic START Integrated Treatment Plan (START ITP) has been developed and is being pilot tested in Canada (Leech, personal communication, January 2013).
- Research is ongoing for the START and its use in different settings (e.g. jail diversion programs, Desmarais et al., 2012a).
- An abbreviated manual is available for use with adolescents (Short Term Assessment of Risk and Treatability: Adolescent Version; START-AV) (Nicholls et al. 2010) and the full manual is in preparation by Dr. Viljoen and colleagues. Pilot investigations and other studies have been conducted on the START-AV (see Desmarais et al., 2012b; Viljoen et al., 2012).

- Those using the START tool are to consider any indicators that there are threats of harm that are real, enactable, acute and targeted. Assessors should be mindful of T.H.R.E.A.T in emergency situations where a comprehensive review of the evidence is not possible. (O'Shea, Picchioni and Dickens, 2016).
- Potential limitations of the START tool are it may be too general for certain patients or groups of patients (e.g. those with learning disabilities) (Kroppan et al., 2017).

Name of Tool	Violence Risk Appraisal Guide-Revised (VRAG-R)
Category	Violence Risk (Validated)
Author / Publisher	Quinsey, Harris, Rice and Cormier
Year	2013

Description

•The Violence Risk Appraisal Guide-Revised (VRAG-R) is a 12-item actuarial risk assessment instrument for the prediction of violent recidivism among male forensic psychiatric patients.
•The instrument was revised in 2013 in order to make it easier to score. Four of the original VRAG items were dropped for using outdated diagnostic criteria or because they have been shown not to be fully applicable to individuals who committed sexual offences (Hertz et al., 2019).
•The instrument utilises the clinical records as a basis for scoring rather than structured interviews or questionnaires (Harris et al., 2015).
•The VRAG-R provides a numerical estimate of the risk of violent recidivism. It is suitable for males aged 18 years and older who have committed serious, violent or sexual offences.

Age Appropriateness

18+

Assessor Qualifications

Professional expertise and training on instrument.

Strengths

•The VRAG-R has a large literature base.
•This tool can be used in combination with historical notes and criminal records (Thomson et al., 2008).

Empirical Grounding

The VRAG was developed from file reviews of 618 male criminal offenders and forensic patients who were initially being assessed for criminal responsibility, fitness to stand trial and/ or being treated in a secure setting; this sample was followed for 7 years. Subsequently, the tool was recalibrated with an extended sample of 800 individuals and followed over a period of 10 years (Quinsey et al., 2006). It was revised in 2013 to make it easier to score. The results of the developmental sample of the VRAG-R showed good predictive accuracy with an AUC of .76 (Harris, Rice and Quinsey, 2016).

Inter-Rater Reliability	
a) UK Research	• Gray et al. (2007) – the VRAG obtained a high ICC value of .95 • Doyle and Dolan (2006) found an inter-rater reliability value of .99 between three raters based on seven cases.
b) International Research	• Rossegger et al. (2011) - the VRAG obtained an ICC value of .95. • Endrass et al. (2008) utilised Krippendorff's alpha to determine the VRAG's inter-rater reliability. The VRAG attained an excellent inter-rater reliability coefficient of .89. • Mills et al. (2007) found an ICC value of .95 for the VRAG in a sample of incarcerated Canadian offenders. • Olver and Sewall (2018) found the VRAG-R displayed excellent inter-rater reliability across 35 randomly selected double-coded cases, with an ICC value of .97.

Validation History

General Predictive Accuracy	
• The VRAG has been devised for use in forensic psychiatric settings (see section 'IV. Mentally Disordered Offenders'). As previously mentioned, other studies have also tested its validity in offenders without psychiatric diagnoses (Langton et al., 2007; Loza and Dhaliwal, 1997).	
a) UK Research	None available at present.
b) International Research	• Rice, Harris and Lang (2013) developed a revised version of the VRAG (VRAG-R), making it easier-to-score. Both the revised version and the original VRAG yielded high predictive accuracy with an approximate ROC of .75. • The VRAG was administered to 52 violent offenders in a Swiss prison to test its ability to predict misconduct. The VRAG displayed good predictive validity for physically violent misconduct and any misconduct (AUCs of 0.83 and 0.81 respectively); fair predictive validity was shown for other misconduct (AUC=0.73) (Abbiati et al., 2014). • Using a sample of 296 sex offenders followed-up over 17.6 years, Olver and Sewall (2018) found the VRAG-R scores demonstrated moderate to large predictive

accuracy for sexual (AUC range=60-.67) and violent (AUC range=.70-.78) recidivism respectively.

• In the first European cross-validation study of the VRAG-R, 534 individuals convicted of a sexual offence were followed up for an average of 7.62 years. The VRAG-R showed moderate to large predictive accuracy for violent, general and sexual recidivism (AUCs of .75, .78 and .63 respectively). It was found that predictive accuracy for sexual recidivism was only significant for those convicted of child sexual abuse offences but not for that who committed them against adult victims (Hertz et al., 2019).

• When applied to a sample of 597 male juvenile sexual offenders, the VRAG-R showed potential strength in predicting non-sexual violent recidivism. It was found, however, that elevated offence severity and adverse childhood experiences encumbered the predictive accuracy of the tool, particularly in the cases of sexual recidivism (Barra et al., 2018).

Validation History

Applicability: Females	■	■	■						

a) UK Research	• Coid et al. (2009) – the VRAG generated moderate predictive accuracy of recidivism in a sample of female offenders with ROC values ranging between .65 to .66.
b) International Research	• Eisenbarth et al. (2012) - the VRAG demonstrated good accuracy in predicting general recidivism in a sample of 80 German female offenders (AUC = .72). • Hastings et al. (2011) - the VRAG was unable to predict institutional misconduct and post-release recidivism in female offenders.

Validation History

Applicability: Ethnic Minorities	■	■	■						

a) UK Research	• Snowden, Gray and Taylor (2010) - in a two-year follow-up, the VRAG obtained an AUC value of .74 in predicting violent reconvictions for offenders of Black ethnic origin.
b) International Research	None available at present.

Validation History

Applicability: Mental Disorders	
a) UK Research	• Doyle et al. (2012) - the VRAG moderately predicted post-discharge violence in a sample of patients discharged from acute mental health units (AUC = .65). • Coid et al. (2009) - the VRAG appeared to outperform the HCR-20 and the PCL:R, demonstrating a ROC area of .70 for violent recidivism; the VRAG scores also predicted acquisitive reconviction in males with a ROC of .71. • Ho et al. (2009) - AUC analysis revealed that the VRAG had moderate to high accuracy in predicting minor violence (.70), serious violence (.74) and any violent incidents (.68). • Snowden et al. (2009) – the VRAG obtained a ROC value of .77 in a sample of male psychiatric patients.
b) International Research	• A study within an Australian clinical forensic practice found that the revised version of the VRAG (VRAG-R) demonstrated predictive validity for recidivism (Brookstein, Daffern and Ogloff, 2016). • Glover and colleagues (2017) tested the VRAG-R on a sample of 120 male correctional individuals. Results indicated that the VRAG-R gave moderate levels of predictive validity for general and violent recidivism that was able to be sustained over time. • Camilleri and Quinsey (2011) report the VRAG has "good predictive accuracy with psychiatric patients of lower intelligence". • Verbrugge et al. (2011) - the VRAG total score attained AUC values of .79 and .92 for violent and general recidivism respectively in a sample of 59 community-based offenders with intellectual disabilities. • In their retrospective study, Rice et al. (2008) showed that the VRAG had the ability to discriminate risk between non-intellectually disabled individuals (control) and intellectually disabled sex offenders. • Kröner et al. (2007) - the tool demonstrated moderate accuracy (AUC) in predicting general (.70) and violent recidivism (.70) in a sample of German male offenders undergoing clinical evaluation for criminal responsibility.

	• Pouls and Jeandarme (2018) collected VRAG scores for 52 offenders with intellectual disabilities (OIDs). AUCs were non-significant; although a trend towards significance was evident for physical aggression (AUC=0.74). The results show that the VRAG overestimated risk of OIDs and was only accurate in identifying low-risk individuals.

Contribution to Risk Practice

- The VRAG-R has the ability to create awareness of static risk factors and can prompt further assessment of the risk of reoffending.
- The VRAG-R shows some consideration for responsivity issues (e.g. psychopathy).

Other Considerations

- Since the VRAG is composed solely of static factors, the tool does not have the capacity to inform treatment protocol or monitor offender progress or motivation for intervention (Daffern, 2007).
- In a review of cases decided in United States federal courts, it was found that the VRAG was mainly introduced by the prosecution as a measure of violence risk and was rarely challenged (Cox et al., 2018).
- The VRAG was found to have high concurrent validity with SAQ total scores (Andreau-Rodriguez, Peña-Fernández, and Loza, 2016).
- The tool also relies on PCL:R rating scores as part of the predictive measurement. A study by Doyle, Dolan and McGovern (2002) found that the PCL:SV was a significant contributor to the predictive validity of the VRAG.
- For more information on the VRAG-R, please visit the following website: http://www.vrag-r.org/

Name of Tool	Terrorist Radicalization Assessment Protocol-18 (TRAP-18)
Category	Violence Risk (Validated)
Author / Publisher	Meloy
Year	2016

Description

- The TRAP-18 is an SPJ tool for assessing individuals who potentially may engage in lone-actor terrorism. It is intended to be used by mental health, intelligence, law enforcement and security professionals to manage operational data on a person of concern and prioritise cases based upon the presence or absence of warning behaviours and characteristics. It is not an actuarial instrument designed to specifically predict acts of lone actor terrorism (Meloy, 2017).
- It is two-fold in nature consisting of eight proximal warning behaviours (pathway, fixation, identification, novel aggression, energy burst, leakage, directly communicated threat and last resort behaviour) and ten distal characteristics (personal grievance and moral outrage, framed by an ideology, failure to affiliate with extremist or other group, dependence on the virtual community, thwarting of occupational goals, changing in thinking and emotion, failure of sexual intimate pair-bonding, mental disorder, creativity and innovation, history of criminal violence) (Meloy et al., 2019).
- All the proximal warning behaviours are dynamic and based on patterns of behaviour, whilst several of the distal characteristics (e.g. history of mental disorder) are static risk factors. Although protective factors are not explicitly included, the absence of certain indicators (proximal warning behaviours and distal characteristics) are protective. Further, the narrative questions ask about the presence of protective factors in individual cases (Meloy, 2019).
- The focus of the TRAP-18 is on patterns of behaviour rather than distinctive variables, i.e. it is not intended to predict who will or will not commit an act of terrorism; rather, the tool can be used to help assign resources by informing on which individuals should receive priority attention. The results generated from using the tool indicate whether a case requires active management (where one or more warning behaviours are present), or monitoring (where only distal characteristics exist) (Meloy & Genzman, 2016; Meloy, 2018; Meloy, 2019).
- To have the most reliable assessment using the TRAP-18, three sources of data should be used: a direct interview (this may be clinical or non-clinical and may or may not involve psychometric testing); collateral interviews with those who are acquainted with the individual and are aware of their behaviour; and the individual's public records, including law enforcement and national security documents if available. It is recognised, however, that a direct interview may not be feasible, necessary, or wise in certain cases (Meloy, 2019).

Age Appropriateness

18+

Assessor Qualifications

Mental health, intelligence, law enforcement, and counter-terrorism professionals with caseloads or supervisory responsibilities. Assessors are required to attend the standardised training course lasting 6 – 7 hours in person or online.

Strengths

• An overview of the strengths of the TRAP-18 was compiled by experts in risk assessment. It was noted that basing some of the distal factors on psychoanalytic theory provides a clinical understanding that may inform risk assessment and intervention. As a tool, it also has the potential to contribute to the prioritisation of cases in a pre-crime scenario, as well as formulation, re-formulation and ongoing risk management in a post-crime situation (Lloyd, 2019).
• The tool can be used regardless of ideology (Meloy & Gill, 2016).

Empirical Grounding

• The TRAP-18 is underpinned by theoretical and empirical literature on lone-actor terrorism and extremism. Its theoretical underpinnings include theory and research on targeted violence, object relations and attachment theories, gestalt psychology (an attempt to understand meaningful perceptions in a chaotic world system) and psychobiological foundations for predatory violence (Meloy, 2019).

Inter-Rater Reliability

a) UK Research	No empirical evidence available at present.
b) International Research	• Meloy et al., (2015) investigated the TRAP-18 using a sample of 22 individuals who had committed terrorism in Europe over a period of thirty-five years. Three hundred and ninety six codings were undertaken by two raters who are experts in threat assessment and management. The mean inter-rater reliability was found to be 0.895. The IRR range for items was good to excellent ranging from 0.68 to 1.0 for the warning behaviours and 0.75 to 1.0 for the distal characteristics.

• In a study by Challacombe & Lucas (2018), two raters evaluated the whole sample (n=58) using the TRAP-18. Average Cohen's Kappa was good for proximal characteristics (k=.687) and excellent for distal characteristics (k=.812). The average inter-rater reliability for the entire TRAP-18 was found to be excellent (k=.757). |

Validation History

General Predictive Accuracy										
a) UK Research	No empirical evidence at present.									
b) International Research	•There have been a number of validation studies carried out on the TRAP-18. An examination of the postdictive validity of warning behaviours was first carried out by looking at school shooting case studies. Although not a form of terrorism because school shootings are usually not motivated by political or ideological reasons, these attacks are similarly unpredictable and have the potential to cause mass causalities (Meloy et al., 2014). •Another study showed that the TRAP-18 was generalisable across various types of terrorism: jihadists, right-wing extremists and single issue attacks (see Meloy & Gill, 2016 for further details). •After examining 111 lone-actor terrorist attacks in the United States and Europe, Meloy & Gill (2016) found that the TRAP-18 was able to discriminate between lone-actor terrorists who successfully carried out their attacks compared to those whose attacks were thwarted. The five variables found to be significantly different were fixation, creativity/innovation, failure in sexually intimate pair bonding, pathway and less likely to be dependent upon a virtual community. •Böckler et al., (2015) used the TRAP-18 to assess the case of the 2011 Frankfurt Airport Attack. Carrying out a qualitative analysis of investigation and court files found that the perpetrator showed six proximal warning behaviours and nine distal characteristics. Tracing the various stages of the individual's life highlighted several triggers towards him drawing upon jihadist ideologies. •Challacombe & Lucas (2018) applied the TRAP-18 to a series of violent and non-violent incidents involving Sovereign Citizens in the US. Chi-square and binary logistic regression analyses were used to test the ability of the TRAP-18 to predict violent outcomes. The full model was statistically significant (x^2=33.88), suggesting TRAP-18 was able to distinguish between individual cases that were violent and non-violent. •Erlandsson & Meloy (2018) assessed the 2015 Swedish school attack in Trollhättan using the TRAP-18. The perpetrator met 7 out of the 8 proximal warning behaviours and 8 out of 10 distal characteristics on the TRAP-18 instrument. Based on these results, the authors									

concluded there is an excellent goodness of fit between this incident and other cases of individual terrorism in Europe and North America.

• Goodwill & Meloy (2019) used a combined sample of North American terrorist attackers (n=33) and non-attackers (n=23) to plot the potential clustering (co-occurrence) of risk factors. Findings indicated that proximal warning behaviours are present in attackers and largely absent in non-attackers, whilst distal characteristics are evident in both groups. Three of the distal characteristics (*personal grievance and moral outrage*, *ideological framing,* and *changes in thinking and emotion*) cluster with both the proximal warning behaviours and the attackers. This suggests both that these distal factors co-occur more in attackers than non-attackers, and that there is an increased likelihood of finding proximal warning behaviours than any of the remaining seven distal characteristics

Validation History

Applicability: Females									
a) UK Research	No empirical evidence at present.								
b) International Research	No empirical evidence at present.								

Validation History

Applicability: Ethnic Minorities									
a) UK Research	No empirical evidence at present.								
b) International Research	No empirical evidence at present.								

Validation History

Applicability: Mental Disorders	■	■	■						
a) UK Research	No empirical evidence at present.								
b) International Research	• Fernández García-Andrade et al. (2019) applied the TRAP-18 to 44 patients with severe mental illness, who had a criminal history and were in situations of social exclusion. High predictive validity was demonstrated for the TRAP-18 (AUC=1.00), indicating it could be a useful tool for assessing the risk of terrorist radicalisation in								

	mentally ill individuals, particularly those with a history of being in prison and living in socially secluded situations.

Contribution to Risk Practice

• Tools like the HCR-20 V3 and WAVR-21 could function as a 'gateway' instrument to allow for a more individualised assessment of the behaviours and motivations associated with lone actor terrorism using the TRAP-18 (Meloy, 2018). Guldimann & Meloy (2020) suggest using other tools such as the HCR-20 V3 in conjunction with the TRAP-18.

• Meloy et al., (2019) coded 2 non-random samples of convenience: 33 cases of a lethal terrorist attack in the United States; 23 individuals who posed a national security concern but did not mount an attack: the latter group were either successfully risk managed for three years, or had no intent to mount an attack. Half of the TRAP-18 indicators were found to be significantly different between the samples with medium to large effect sizes (φ=.35-.70). The three warning behaviours that were not significantly different between the groups were fixation, novel aggression and leakage. Due to its retrospective design, no inferences were made about predictive validity on the basis of this study.

• Looking at a sample of 22 individuals who had committed terrorism in Europe, Meloy et al. (2016) found that 'content validity' (the extent to which a measure represents all the facets of a given construct) was evident in 72% of the variables of the TRAP-18.

• In a review of the literature, Guldimann & Meloy (2020) described the inter-rater reliability of the TRAP-18 as excellent, with research showing promise in terms of content, criterion, discriminative, and predictive validity. They found that several of the proximal warning behaviours - pathway, fixation, identification, leakage, energy burst, and last resort – were commonly found in the research, while "directly communicated threat" was not prominent. However, they caution that its absence should not be interpreted to mean that no threat exists.

Other Considerations

• A group of experts reviewed the strengths and limitations of the TRAP-18. Some limitations noted were that:
  - o the focus on lone actors potentially limits its utility;
  - o a full assessment involving a direct interview, psychometric testing, and complete information sources may not be entirely realistic in a pre-crime scenario; and
  - o some of the more psychoanalytic distal factors may be difficult to make sense of without clinical expertise.

• In terms of the strengths of the tool, they noted that:
  - o it can be used in risk management and prevention, potentially discriminating between empty and real threats;
  - o it can potentially assist with case prioritisation in a pre-crime scenario, in addition to formulation, re-formulation and ongoing risk management in a post-crime situation; and
  - o a clinical understanding of subjects relevant to terrorism underpin and inform the tool (Lloyd, 2019).

• Brugh et al., (2020) applied the TRAP-18 to a sample (n=77) of jihadism-inspired lone actor terrorists in Europe and the US, using only publically available information from the Western Jihadism Project database. Of 18 items, only four were rated Present more often than they were rated Absent or Unknown (Pathway, Identification, Personal Grievance, Framed by Ideology). In comparing the US and European samples, the items Fixation, Energy Burst, Leakage, and

Dependence on the Virtual Community were more common the US sample. The study produced three "false negatives:" three cases were not recommended for Active Risk Management, contrary to expectations for a sample of confirmed lone actor terrorists. The authors conclude that using TRAP-18 with only publically available information raises questions about the tool's feasibility in this setting. Consideration should also be given to potential difficulties using the TRAP-18 across different geopolitical contexts where there may be differences in how information is gathered and made publicly available.

•The TRAP-18 is owned, copyrighted and trademarked in the United States by Dr Meloy, with distribution and sales licensed to *Multihealth Systems, Inc.* (mhs.com) (Meloy, 2019).

•Training in the TRAP-18 is available from the *Global Institute of Forensic Research* through their online on demand resources (gifrinc.com). This company is owned by Multihealth Systems.

•Further information about the tool and its author may be found here:
http://drreidmeloy.com/training/trap-18/