TSAS

canadian network for research on terrorism, security and society

# Working Paper Series

---

No. 17-1

July 2017

## A Concurrent Evaluation of Threat Assessment Tools for the Individual Assessment of Terrorism

Stephen D. Hart, Alana N. Cook, D. Elaine Pressman, Steven Strang, & Yan L. Lim

Director: Lorne Dawson, University of Waterloo

The TSAS Working Paper Series can be found at http://tsas.ca/research/tsas-working-papers/

The Canadian Network for Research on Terrorism, Security, and Society TSAS is supported as a national strategic initiative funded by SSHRC and Public Safety Canada, along with the following departments of the federal government:

- Royal Canadian Mounted Police
- Canadian Security Intelligence Service
- Immigration, Refugees, and Citizenship Canada
- Correctional Services Canada
- Defence Research and Development Canada
- Global Affairs Canada
- Security Intelligence Review Committee
- Office of the Communications Security Establishment Commissioner

TSAS also receives financial support from several Canadian universities, including the University of Waterloo.


For more information, contact the Lorne Dawson, Department of Sociology and Legal Studies, University of Waterloo (ldawson@uwaterloo.ca).

Report Submitted to the Canadian Network for Research on Terrorism, Security and Society
Final Version, 3 June 2017

**Acknowledgements**

At the time this research was conducted, the primary affiliations of the authors were as follows: Stephen D. Hart, Alana N. Cook, and Yan L. Lim, Department of Psychology, Simon Fraser University; Elaine Pressman, Canadian Center for Security and Intelligence Studies, Norman Paterson School of International Affairs, Carleton University; and Steven Strang, Strategic Policy & External Relations, Federal Policing, Royal Canadian Mounted Police.

The consensus MLG and HCR-20 V3 ratings for the series of five terrorism cases in Study One are available from the first author on request.

The authors gratefully acknowledge support in the form of a research grant from the Canadian Network for Research on Terrorism, Security and Society (TSAS). We also thank the people who provided (anonymous) reviews of a previous draft of this report on behalf of TSAS, especially one reviewer who gave some helpful feedback that helped us to make more detailed recommendations for future research. The opinions expressed herein are those of the first author; although they reflect the consensus or majority opinions of all authors where possible, they do not necessarily reflect the opinions of any individual co-author; also, the opinions expressed here do not necessarily reflect the official views of TSAS or any agencies with which any of the authors are affiliated.

Address correspondence to Professor Stephen D. Hart, Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada, V5A 1S6 (hart@sfu.ca).

## Table of Contents

## 1.0 Introduction

Terrorism represents a significant concern for public safety in most countries around the world; depending on the precise definition used and the geopolitical region under investigation, it is a weekly or even daily occurrence (United Nations, 2015). In response, nation states and international bodies have developed counter-terrorism strategies. Canada is no exception. The federal government recognizes terrorism is a continuing threat to national security (Public Safety Canada, 2014) and has detailed a plan for dealing with the threat that comprises four basic elements: prevent, detect, deny, and respond (Public Safety Canada, 2013). Focusing specifically on the criminal justice system, the Canadian response to terrorism is guided by Part II.1 of the *Criminal Code* (1985), which specifies relevant offences and sets out procedures for investigative hearings, trial, sentencing, and application for recognizance with conditions to prevent terrorist activity.

### 1.1 Assessment and Management of Risk for Terrorism

Assessment and management of risk for terrorism is a cornerstone of effective counter-terrorism. A wide range of people are responsible for assessing and managing terrorism risk, including intelligence and security, law enforcement, institutional and community corrections, and forensic mental health professionals. In this report, for the sake of brevity, we will refer to them collectively as *threat assessment professionals;* and we will refer to the work done by them as *threat assessment*.

As we use the term here, threat assessment may be considered synonymous with what others call violence risk assessment (Meloy, Hoffmann, & Hart, 2013). Threat

assessment may be defined as comprising two types or phases of work that are conceptually distinct yet procedurally inseparable (see Hart, Douglas, & Guy, 2016). The first is to understand a person's potential for future terrorism, and the second is to develop plans for disrupting that potential. In the first phase, threat assessment professionals analyze a wide range of factors to consider what kinds of terrorist activity people might be involved in, what roles they might play in such activity, the time and location of the activity, the identity of potential victims, the motives or reasons for the activity, and any events or occurrences might exacerbate or mitigate the person's potential for terrorism. In the second phase, threat assessment professionals consider various strategies, tactics, and logistic to identify which steps reasonably could and should be taken to effectively mitigate the risks posed by people in light of any relevant legal, situational, and practical constraints.

Over the past 25 years, the field of threat assessment has matured considerably. There is now a well-developed evidence base concerning the use of decision support tools by threat assessment professionals (e.g., Meloy & Hoffmann, 2014; Otto & Douglas, 2010). The evidence base indicates that a number of tools can be used by a wide range of professionals in many different countries to make decisions about the assessment and management of risk for diverse forms of violence—including general violence, sexual violence, intimate partner violence, stalking, workplace violence, honour-based violence, and so forth—with good reliability and validity (Singh, Grann, & Fazel, 2011). A major advance in the field was the development of the Structured Professional Judgement approach to risk assessment, which may be defined as "an

analytical method used to understand and mitigate the risk for interpersonal violence posed by individual people that is discretionary in essence but relies on evidence-based guidelines to systematize the exercise of discretion" (Hart et al., 2016, p. 643). SPJ decision support tools, also referred to as SPJ guidelines, "are specific evaluative devices or procedures developed according to the SPJ approach that are intended to assess and manage risk for specific forms of violence or in specific contexts" (Hart et al., 2016, p. 643).

Until recently, there was a total absence of decision support tools for individual assessment of risk for terrorist activity. Intelligence analysts in the fields of military, policing, and security studies have developed and used decision support tools for terrorist groups for many years (e.g., Grabo, 2002; Heuer, 1999; Heuer & Pherson, 2011; Strang, 2005). But these tools were not developed to assess and manage the risk posed by specific individuals within terrorist groups, and they were not developed for use by the full range of threat assessment professionals. In a seminal call to action, Monahan (2012) discussed the problems of individual assessment of risk for terrorism, and drew four major conclusions. First, Monahan concluded, a prerequisite for development of adequate decision support tools is to specify the nature and scope of risk for terrorism. Second, the SPJ approach is most appropriate for developing such tools. Third, the priority for research is to identity robust individual risk factors for terrorism. Fourth, to identify empirically validated individual risk factors for terrorism, researchers require access to retrospective information on known groups of terrorists and non-terrorists.

### 1.2 SPJ Decision Support Tools for Individual Assessment of

## Risk for Terrorism: Description

Around the same time that Monahan's paper was written, three SPJ decision support tools for individual assessment of risk for terrorism were either developed or in the process of development. Below, we provide a brief overview of the three tools. Our review excludes several SPJ tools developed after the publication of Monahan's paper that were intended to serve as screening tools, such as the Extremism Risk Screen (see Lloyd & Dean, 2015), or to assess risk for specific forms of terrorism, such as the Cyber-VERA (Pressman & Ivan, 2016) and Terrorist Radicalization Assessment Protocol (TRAP-18; Meloy & Gill, 2016).

### 1.2.1 VERA/VERA 2

The first tool is the Violent Extremism Risk Assessment Protocol. The first consultation version of the VERA was developed in 2009 by Elaine Pressman (VERA; Pressman, 2009). As is obvious from the date of its publication, the VERA appeared prior to Monahan's (2012) paper and indeed was discussed in it. The second consultation version was developed in 2010 in collaboration with John Flockton, based on feedback from operational experts in terrorism and violent extremism as well as from risk assessment and terrorism subject matter experts from around the world (VERA 2; Pressman & Flockton, 2010, 2012).

Broadly speaking, as their names imply, the VERA and VERA 2 focuses on terrorism motivated by extremist ideology, that is, sociopolitical beliefs, attitudes, and views that justify the use of violence as a political act. Although the VERA had a strong but by no means exclusive emphasis on extremism associated with radical Islam, the

VERA 2 makes very clear that the tools were intended to be used with a wide spectrum of extremism, including, for example, that associated with nationalists, ecoterrorists, anarchists, and right- and left wing groups. The authors recommend that it is used as a complement to other relevant risk assessment tools to evaluate people who have "already been convicted of a violence extremist or terrorist-designated offence" (Pressman & Flockton, 2012, p. 244).

As part of the administration of the VERA (Pressman, 2009), evaluators consider the presence of 28 risk and protective factors, also referred to as indicators, in five conceptual domains: Attitude, Contextual, Historical, Protective, and Demographic factors. The factors reflect the characteristics of people that are considered important in radicalization and terrorism. Evaluators consider the presence of factors based on multiple sources of information. Presence ratings reflect the degree to which the factors are associated with enhanced risk (for the Attitude, Contextual, History, and Demographic domains) or mitigate risk (for the Protective domains) and are made on a 3-point scale (*Low*, *Medium*, or *High*). A final global risk judgement of risk level is made on a 3-point scale (*Low*, *Medium*, or *High*) after consideration of all risk factors. The domains and individual risk factors in the VERA are presented in Table 1.

As part of the administration of the VERA 2 (Pressman & Flockton, 2010), evaluators consider the presence of 31 risk and protective factors, also referred to as indicators, in five conceptual domains: Beliefs and Attitudes, Context and Intent, History and Capability, Commitment and Motivation, and Protective factors. The factors reflect the narratives and networks of people which are considered important in

radicalization and terrorism. Evaluators consider the presence of factors based on multiple sources of information. Presence ratings reflect the degree to which the factors enhance risk (for the Beliefs and Attitudes, Context and Intent, History and Capability, and Commitment and Motivation factors) or mitigate risk (for the Protective factors) are made on a 3-point scale (*Low*, *Moderate*, or *High*). A final global risk judgement of risk level is made on a 3-point scale (*Low*, *Moderate*, or *High*) after consideration of all risk factors. The domains and individual risk factors in the VERA 2 are presented in Table 2.

The VERA is an open access tool. It is available to the general public, and evaluators are not required to complete a specific training program prior to purchase or use of the tool. It is intended for use by threat assessment professionals from diverse backgrounds. It may be administered by individual professionals or by multi-disciplinary teams of professionals. In contrast, the VERA 2 is a restricted access tool. It is not available for purchase by the general public. Threat assessment professionals are required to undergo specific training prior to purchase and use of the tool. It is intended for use by threat assessment professionals from diverse backgrounds. It may be administered by individual professionals or by multi-disciplinary teams of professionals.

### 1.2.2 MLG

The second tool is the Multi-Level Guidelines (MLG; Cook, Hart, & Kropp, 2014), which grew out of Stephen Hart's work on the assessment and management of group-based violence (e.g., Hart, 2010; Hart & Dormond, 2009). The MLG defines group-based violence as the actual, attempted, or threatened physical injury of others that is

deliberate and unauthorized, perpetrated by one or more people whose decisions and behavior are influenced by a group to which they belong or with which they are affiliated. Thus, the concept of group-based violence includes the majority of terrorism, with the exception of some lone actor terrorism; but it also includes many forms of violence that is not terrorism (e.g., violence committed by criminal organizations, street gangs, new religious movements, clans, and ideologically-focused groups). The first version of the MLG was developed on the basis of a comprehensive, Campbell Collaboration-style systematic literature review conducted by Alana Cook as part of her doctoral dissertation, in which she evaluated the tool via surveys of subject matter experts and examining its use by threat assessment professionals (Cook, 2014; see also Cook, Hart, & Kropp, 2013). Based on the results of Cook's research, work on the final version of the MLG began in 2014 (Cook et al., 2014). The revision is intended to enhance the usability of the MLG by simplifying and clarifying the assessment procedure. We will focus on the most recent version of the MLG below.

Broadly speaking, as its name implies, the focus of the MLG is on individual-within-group dynamics relevant to violence—that is, individual and group dynamics that factors with and influence each to enhance or mitigate violence risk, based on a nested ecological model of violence. The authors recommend the MLG for use in conjunction with other relevant risk assessment tools to evaluate people who are known or suspected to have committed terrorist group-based violence, as well as those who may be at risk for terrorist group-based violence. The MLG may also be used to analyze the extent to which the terrorist violence perpetrated by a person should be considered

group-based versus individual (lone actor).

The structure of the MLG is modeled directly on that of commonly used SPJ guidelines, and in particular the third version of the Historical-Clinical-Risk Management—20 (HCR-20 V3; Douglas, Hart, Webster, & Belfrage, 2013). The MLG administration procedure comprises seven steps: evaluators gather relevant case information (Step 1); consider the presence and relevance of 16 basic risk factors, as well as any case-specific risk factors (Steps 2 and 3); develop an integrative formulation of terrorism risk based on risk factors that are present and relevant (Step 4); develop scenarios of future terrorism based on the formulation, as well as management plans based on those scenarios (Steps 5 and 6); and communicate various conclusory opinions about the nature of risks posed by the person (Step 7).

The 16 basic risk factors in the second version of the MLG (there were 20 risk factors in the first version) reflect four conceptual levels or domains of dynamics: Individual, Individual-in-Group, Group, and Group-in-Society. The Individual domain comprises factors relevant to people as individuals, irrespective of any groups to which they belong or are affiliated (e.g., mental health problems). These factors were modeled directly after those in the HCR-20 V3; indeed, the MLG manual permits evaluators to use the HCR-20 V3 risk factors as a substitute for the Individual domain factors in the MLG. The Individual-Group domain comprises factors relevant to people's identities, attitudes, and roles vis-à-vis groups (e.g., group-based identity). The Group domain comprises factors related to group processes and structures, irrespective of the individual person (e.g., group norms). Finally, the Group-Societal domain comprises

factors related to the broader social context in which the group exists and operates (e.g., intergroup conflict). The MLG risk factors are presented in Table 3. Ratings of presence are made for each risk factor on a 3-point scale (N = *no evidence the risk factor is present*, P = poss*ible or partial evidence the risk factor is present*, Y = *evidence the risk factor is definitely present*). Ratings of relevance (i.e., functional relevance with respect to the perpetration of violence) are also made on a 3-point scale (*Low*, *Moderate*, or *High*). Ratings of relevance (i.e., functional relevance with respect to the perpetration of violence) are also made on a 3-point scale (*Low*, *Moderate*, or *High*). The conclusory opinions made include: Future Violence, also known as Case Prioritization, reflecting overall likelihood that the person will commit group-based violence in the future; Serious Physical Harm, reflecting the risk that any group-based violence committed by the person in the future will result in life-threatening or lethal injury; and Imminent Violence, reflecting the risk that the person will commit group-based violence in the near future.

The MLG is an open access tool. It is available for purchase by the general public, and evaluators are not required to complete a specific training program prior to purchase or use of the tool. The tool is intended for use by threat assessment professionals from diverse backgrounds. It may be administered by individual professionals, although administration by multi-disciplinary teams of professionals is strongly recommended.

**1.2.3 SRG/ERG 22+**

The third tool is the Structured Risk Guidance for extremist offending, which first appeared in 2009 and in 2011 was subsequently revised and renamed the

Extremism Risk Guidelines (SRG and ERG 22+, respectively; Lloyd & Dean, 2015). The SRG and ERG 22+ were developed by the National Offender Management Service of England and Wales, with an important role played by Christopher Dean. The SRG and ERG 22+ were developed on the basis of systematic review of existing tools (including the VERA 2 and MLG) and consultation with subject matter experts (including Stephen Hart). We will focus below on the ERG 22+.

Broadly speaking, the ERG 22+ focuses on "pathway influences" (Lloyd & Dean, 2015, p. 48) that drove people to engage in terrorism-related offences and that may be targeted by intervention to facilitate disengagement or desistance. Like the original VERA, the ERG 22+ has a strong but not exclusive focus on extremism associated with radical Islam. The authors recommend the ERG 22+ to evaluate people who have been convicted of terrorism-related to assist in their correctional management and rehabilitation.

The administration procedure for the ERG 22+ is less structured than that of the VERA 2 or MLG. According to Lloyd and Dean (2015), when gathering and reviewing case information, evaluators focus on contextual circumstances, personal attributes, and actual or perceived benefits that contributed to past offending, as well as the contextual circumstances or personal attributes that might promote desistence from future offending. There is a list of risk factors that evaluators may review but do not need to code. Instead, the goal is for evaluators to develop a integrative case formulation based on all relevant factors. There are 22 basic risk factors from three domains (called *dimensions* in the tool): Engagement, Intent, and Capability. The risk factors in the ERG

22+ are presented in Table 4.

The ERG 22+ is a restricted access tool. It is not available for purchase by the general public. It is intended for use only by NOMS psychologists and probation officers. It is administered by individual professionals. Evaluators undergo specific training prior to use of the tool.

### 1.3 SPJ Decision Support Tools for Individual Assessment of Risk for Terrorism: Evaluation

As discussed previously, Monahan (2012) identified four prerequisites for the development of adequate decision support tools is to specify the nature and scope of risk for terrorism. It appears that the first three of those prerequisites have been met: multiple sets of SPJ guidelines have been developed that specify the nature and scope of risk for terrorism and include individual risk factors. The last prerequisite Monahan identified—empirical evaluation, and, in particular, retrospective studies comparing known groups of terrorist and non-terrorists—remains unmet.

Empirical evaluations of the VERA/VERA 2, MLG, and ERG 22+ to date is very limited and has been conducted, for the most part, by the people or agencies who developed them. There are two major reasons for this. First, it is difficult to systematically sample terrorists for research purposes. It is unlikely that researchers will be able to locate large numbers of terrorists who have not (yet) been apprehended and are willing to participate voluntarily in research; and terrorists who have been apprehended and detained in custody could be studied involuntarily, but they comprise a small subgroup that is unlikely to be representative of larger groups or

movements to which they belong. Second, even if one could identify a sample of terrorists for research purposes, case history information needed to conduct the research may be inaccessible for security reasons or simply absent due to incomplete investigation.

In the face of these challenges, researchers have pursued two basic research strategies. First, researchers have examined the extent to which the tools could be used and judged to be useful by threat assessment professionals in case studies or case series of known terrorists or in the actual day-to-day work of threat assessment professionals. Second, researchers have evaluated the interrater reliability of decisions made by threat assessment professionals who use the tools to evaluate case studies or a case series of known terrorists based on open source or sanitized information.

One example of the type of research done to date is a study by Beardsley and Beech (2013). They had two evaluators independently administer the VERA in a case series of 5 known terrorists based on open source information. They examined the extent to which the evaluators were able to make judgments concerning the presence of the VERA risk factors, as well as the interrater reliability of those judgments. They found that the evaluators rated most of the risk factors for each of the five cases; only 12 of 140 ratings (28 risk factors x 5 cases)—or about 9% of all ratings—were omitted. Also, there was variability across the 5 cases in the ratings for 24 of the 28 risk factors. Finally, the interrater reliability of the ratings was good: The overall agreement for the 140 ratings was about 86%, and the chance-corrected agreement as indexed by Cohen's $\kappa$ for each of the 28 risk factors was at least .76.

A second example is a study by Pressman and Flockton (2014; see also Pressman, 2014). The VERA 2 was administered to a group of 11 convicted terrorists and a comparison group of 11 convicted violent (non-terrorist) offenders, matched on the basis of age, sex, background and religion. Several other general violence or general criminality risk assessment tools were administered to the same groups, including Version 2 of the HCR-20 (Webster, Douglas, Eaves, & Hart, 1997), the Screening Version of the Hare Psychopathy Checklist—Revised (Hart, Cox, & Hare, 1995), and the Level of Service Inventory—Revised (Andrews & Bonta, 1995). The results indicated there were statistically significant differences between the terrorists and non-terrorist violent criminals on all risk tools. Specifically, the terrorist offenders were significantly *lower* risk than the non-terrorist violence offenders according to the general violence and general criminality risk assessment tools, but significantly *higher* risk according to the VERA 2.

A third example is Cook's doctoral research on the first version of the MLG (Cook, 2014). She recruited 46 threat assessment professionals to attended one of two workshops on the use of the MLG. The first day of each workshop was devoted to completion of a pre-training questionnaire and instruction in administration of the MLG. Professionals were then randomly assigned to small groups, and each professional in each group completed the MLG for 5 of 10 practice cases (7 of which involved terrorism, 3 of which involved organized crime or honour-based violence) based on sanitized case materials or open source information over the course of the following two days. The assignment of cases to groups was counter-balanced to control

for order effects. At the end of the three days, professionals completed a post-training questionnaire. Comparison of pre- and post-training questionnaire data indicated that professionals reported improvement in their level of confidence, knowledge, and competence related to violence risk assessment generally that was moderate in magnitude and statistically significant; and improvement in their level of confidence, knowledge, and competence related to risk assessment for group-based violence more specifically that was very large in magnitude and statistically significant. With respect to utility of the MLG, they found that professionals rated all 20 risk factors for each of the 10 cases. Also, there was variability across the 10 cases for each of the 20 risk factors. Finally, interrater reliability for the ratings was good: the median level of chance-corrected agreement for (as indexed by intraclass correlation coefficients for single ratings, absolute agreement method) for the 20 individual risk factors was .53, and for conclusory opinions related to overall risk was .71.

In addition to these, other evaluations of the assessment tools have been conducted, including a study on consumer satisfaction with the SRG/ERG 22+ in the National Offender Management Service (Webster, Kerr, & Tompkins, in press) and papers presented at professional meetings discussing the application of the MLG (e.g., Burton & Amat, 2013) or MLG and VERA 2 (e.g., van Kuijk & Voerman, 2016) in field settings by threat assessment professionals. But most of the evidence supporting the tools is anecdotal in nature, reflecting the opinions of law enforcement, intelligence, national security, and corrections agencies around the world that have used the tools.

### 1.4 Current Study

We decided to undertake further evaluation of SPJ decision support tools for individual assessment of risk for terrorism in a law enforcement setting, focusing on the VERA 2 and MLG (as the ERG 22+ was intended for use in correctional settings). Our original plan was to have a group of analysts, trained by the authors in the use of the VERA 2 and MLG, use both tools to assess a large case series of known or suspected offenders based on sanitized case materials extracted from actual police records. Unfortunately, although the plan had support in principle from stakeholders at the outset, we were unable to get the necessary approvals, as our police partner agency, the Royal Canadian Mounted Police, was unable to commit the resources for its staff to collate and sanitize the case materials. We considered training a group of independent researchers to use the VERA 2 and MLG and then complete a smaller case series of known or suspected offenders based on sanitized case materials or open source information, but we did not have the resources to pay for one of the authors, Elaine Pressman or John Flockton, to conduct official VERA 2 training and, as a consequence, we were not eligible to purchase the tool.

We therefore decided to evaluate the content of the MLG by conducting two different studies. In Study One, we examined risk ratings made using the MLG, HCR-20 V3 (the most widely used and best validated tool for individual assessment of risk for general violence), and the VERA (the first tool for individual assessment of risk for terrorism) in a series of five open-source cases, the same ones used by Beardsley and Beech (2013). First, we evaluated the interrater reliability of MLG risk ratings made using the MLG, expecting they would be *good*. Second, we examined the concurrent

validity of MLG and HCR-20 V3 risk ratings. We expected the association between summary risk ratings on the MLG and HCR-20 V3 would be *large and positive*. We also expected that ratings on the MLG Individual domain would have a *large and positive* association with HCR-20 V3 domain ratings, but the association with the ratings on the MLG Individual-in-Group, Group, and Group-in-Society domains would be *near-zero or negative*. Third, we examined the association between MLG and VERA ratings. We expected that ratings on the MLG Individual domain would have *near-zero or negative* associations with ratings for the VERA domains, but the association with the ratings on the MLG Individual-in-Group, Group, and Group-in-Society domains would be *large and positive*.

In Study Two, we conducted a conceptual analysis of the content overlap of the MLG and VERA 2. We expected overlap between risk factors in the Individual and Individual-in-Group domains of the MLG and the Beliefs and Attitudes domain of the VERA 2 would be *moderate to large*, but overlap between the Group and Group-in-Society domains of the MLG and all the VERA 2 domains would be *near-zero*.

## 2.0 Study One

### 2.1 Method

#### 2.1.1 Case Series

Beardsley and Beech (2013) presented VERA ratings for a series of five well-known cases of terrorism (Andreas Baader, Ikuo Hayashi, Theodore Kaczynski, Patrick Magee, and Timothy McVeigh) based on open-source information. They were selected to be diverse in terms of the nationality and extremist attitudes of the terrorists, the

extent to which they operated alone or as part of a group, and the role they played in any group to which they belonged. A detailed description of the cases and source material, as well as their ratings for VERA risk factors for each case, can be found in their paper.

We coded the MLG and HCR-20 V3 for the same five cases studied by Beardsley and Beach (2013). We attempted to locate the case materials to code the MLG and HCR-20 V3 that were identified by Beardsley and Beach (2013), and were successful in locating about 90% of those materials via the World Wide Web; 5 of the 51 URL links (10%) cited by Beardsley and Beach (2013) were broken. Despite this, the case materials available were sufficient in quantity and quality to permit administration of the MLG and HCR-20 V3.

**2.1.2 Procedure**

The five cases were assessed using the MLG and HCR-20 V3 by four evaluators based on the available open source information. All evaluators were graduate students in clinical-forensic psychology at Simon Fraser University who had extensive education and supervised practice in violence risk assessment.

**2.1.3 VERA**

As noted, the VERA risk factors for each case were taken from the paper by Beardsley and Beech (2013). Those ratings reflected the consensus judgment of two raters, following their independent administration of the VERA for the five cases. No details concerning the interrater reliability of the VERA ratings were reported in the original paper, save that individual presence ratings were all high (Cohen's $\boldsymbol{\kappa} \geq .76$).

### 2.1.4 HCR-20 V3

As noted previously, the HCR-20 V3 (Douglas et al., 2013) is a set of SPJ guidelines for assessing risk for general violence. The HCR-20 V3 was not developed to assess risk for terrorism and is clearly not sufficient for that purpose, yet its content is necessary for the assessment of risk for terrorism and the assessment of risk for general (i.e., non-terrorist) violence by terrorists (Dernevik, Beck, Grann, Hogue, & McGuire, 2009; Gudjonsson, 2009; Hart, 2010; Hart & Dormond, 2009). Administration of the HCR-20 V3 comprises seven steps. Evaluators gather relevant case information (Step 1); consider the presence and relevance of 20 basic risk factors, as well as any case-specific risk factors (Steps 2 and 3); develop an integrative formulation of violence risk based on risk factors that are present and relevant (Step 4); develop scenarios of future violence based on the formulation, as well as management plans based on those scenarios (Steps 5 and 6); and communicate various conclusory opinions about the nature of risks posed by the person (Step 7).

The 20 basic risk factors in the HCR-20 V3 are divided into three temporal domains; see Table 5. The 10 Historical factors reflect adjustment problems at any time in the past. The 5 Clinical factors reflect adjustment problems in the recent past. The 5 Risk Management factors reflect likely adjustment problems in the near future. Ratings of presence are made for each risk factor on a 3-point scale (N = *no evidence the risk factor is present*, P = *possible or partial evidence the risk factor is present*, Y = *evidence the risk factor is definitely present*). Ratings of relevance (i.e., functional relevance with respect to the perpetration of violence) are also made on a 3-point scale (*Low*, *Moderate*, or *High*). The

summary risk ratings made include: Future Violence, also known as Case Prioritization, reflecting overall likelihood that the person will commit violence in the future; Serious Physical Harm, reflecting the risk that any violence committed by the person in the future will result in life-threatening or lethal injury; and Imminent Violence, reflecting the risk that the person will commit violence in the near future. Summary risk ratings are also rated on a 3-point scale (*Low*, *Moderate*, or *High*).

In the current study, two evaluators administered the HCR-20 V3 for the 5 cases based on open source information concerning the psychosocial adjustment of terrorists at the time their criminal careers ended (i.e., due to apprehension or death). Summary risk ratings reflected judgments concerning terrorists' overall risk of future violence, assuming they were alive and released into the community shortly after their criminal careers ended. To minimize sequencing effects, one evaluator assessed the cases in alphabetical order and the other assessed them in reverse alphabetical order, blind to each other's ratings and to the MLG ratings. After finishing their assessments, the evaluators reviewed their ratings for each case one at a time and made a final set of joint consensus ratings.

For the purposes of the present study, we focused our analyses on consensus lifetime presence ("ever present") and relevance ratings for the 20 basic risk factors and on summary risk ratings. We did not evaluate the interrater reliability of the HCR-20 ratings, as those are clearly established in the literature.

**2.1.5 MLG**

Details concerning the content and administration procedure of the MLG have

already been presented and will not be repeated here. We made two modifications to the standard MLG administration procedure. First, for individual risk factors, evaluators were asked to rate presence based on information concerning the psychosocial adjustment of the terrorists at the time their criminal careers ended (i.e., at the time they were apprehended or died). Similarly, for conclusory opinions, evaluators were asked to rate the terrorists' overall risk of future violence, assuming they had lived and were released into the community shortly after their criminal careers ended.

Two evaluators, both of whom completed training in the use of the MLG, assessed all 5 cases using the second version of the MLG (Cook et al., 2015). To minimize sequencing effects, one evaluator assessed the cases in alphabetical order and the other assessed them in reverse alphabetical order, blind to each other's ratings and to the MLG ratings. After finishing their assessments, the evaluators reviewed their ratings for each case one at a time and made a final set of joint consensus ratings.

In the present study, we focused our analyses on lifetime presence ("ever present") and relevance ratings for the 16 individual risk factors and on ratings for summary risk ratings. Analyses of interrater reliability were based on the independent ratings and analyses of concurrent validity were based on the consensus ratings.

**2.1.6 Statistical Analyses**

For analytic purposes, we converted the VERA, HCR-20 V3, and MLG item ratings into numbers (0 = *Low*, 1 = *Moderate*, 2 = *High*) and then summed them to create numerical total and domain scores. We also converted the HCR-20 V3 and MLG summary risk ratings into numerical scores 0 = *Low*, 1 = *Moderate*, 2 = *High*). (Summary

risk ratings for the VERA were not reported in the study by Beardsley & Beech, 2013.)

We indexed interrater reliability of MLG risk ratings using a chance-corrected measure of agreement, the intraclass correlation coefficient for single ratings ($ICC_1$), 2-way random effects model, absolute agreement method. Following Fleiss (1981), ICCs were interpreted as follows: < .39 = *poor*, .40 to .59 = *fair*, .60 to .74 = *good*, and ≥ .75 = *excellent*. We indexed concurrent validity using Person product-moment correlations between risk ratings made using the MLG and those made using the VERA and HCR-20 V3.

## 2.2 Results

### 2.2.1 Interrater Reliability of the MLG

For summary risk ratings, the interrater reliability of the future violence rating fell in the *good* range, but the interrater reliability of the serious physical harm and imminent violence ratings fell in the *fair* range. These are summarized in Table 6.

The interrater reliability of presence ratings for MLG risk factors was in the *excellent* range for 14 of 16 risk factors, in the *good* range for one risk factor (GS3, Operating in an unstable context/environment), and in the *poor* range for only one risk factor (GS4, Threatened by or in conflict with other groups). The average was in the *excellent* range, *Mdn* $ICC_1$ = .95. Looking at numerical scores, the interrater reliability of total and domain scores all fell in the *excellent* range, even for the Group-in-Society domain, $ICC_1$ = .87; there are summarized in Table 7.

For relevance ratings, the interrater reliability of risk factors was in the *excellent* range for 9 of 16 risk factors, in the *good* range for 4 risk factors, and in the *poor* range for

3 risk factors. The average was in the *excellent* range, *Mdn* $ICC_1$ = .80. Looking at numerical scores, the interrater reliability of total and domain scores fell in the *excellent* range with the exception of the Group-in-Society domain, which fell in the *fair* range, $ICC_1$ = .57; these are also summarized in Table 7.

**2.2.2 Concurrent Validity: MLG versus HCR-20 V3**

The correlations between summary risk ratings on the MLG and HCR-20 V3 are presented in Table 8. As expected, 7 of the 9 correlations were *large* and statistically significant. Looking at the MLG, Future Violence ratings had larger correlations with the HCR-20 V3 ratings than did the MLG Serious Physical Harm or Imminent Violence ratings; in contrast, the HCR-20 V3 Serious Physical Harm ratings had larger correlations with the MLG ratings than did the HCR-20 V3 Future Violence or Imminent Violence ratings.

Next, we examined the association between various risk ratings made using the MLG and HCR-20 V3. The correlations between total and domain presence scores on the two tools are presented in Table 9, and the correlations between total and domain relevance scores are presented in Table 10. The overall pattern of findings was generally consistent with expectations. For both presence and relevance, MLG Individual domain scores had *positive* correlations with HCR-20 V3 total and domain scores, but ranged in magnitude from *small to large*; and the other MLG domain all had *near-zero or negative* correlations with the HCR-20 V3 total and domain scores. None of the correlations, however, was statistically significant.

**2.2.3 Concurrent Validity: MLG versus VERA**

Next, we examined the association between total and domain presence scores on the MLG and VERA. (Recall that there are no relevance ratings on the VERA.) These are presented in Table 11. They were consistent with expectations only in limited respects. First, the MLG Individual scores had *near-zero* or *negative* correlations with the VERA Contextual, Historical, and Protective domain scores; however, the correlations with the Attitude and Demographic domains were *positive* and *large*. None of these correlations was statistically significant. Second, only the VERA Contextual domain scores had *large*, *positive*, and statistically significant correlations with all the MLG domain scores (with the exception of scores on the Individual domain). The other VERA domain scores had correlations with the MLG domain scores that varied in direction and magnitude, and none of them was statistically significant.

### 2.3 Discussion

Several tentative conclusions can be drawn from the findings of Study One. First, there was no evidence of problems with the interrater reliability of risk ratings made using Version 2 of the MLG. The finding is consistent with past research on Version 1 of the MLG (Cook, 2014), as well as in past research using the HCR-20 V3 (Douglas et al., 2013), VERA (Beardlsey & Beech, 2013), and other SPJ guidelines (e.g., Otto & Douglas, 2011).

Second, the association between the MLG and HCR-20 V3 summary risk ratings was consistent with expectations. Specifically, the correlations among the summary risk ratings made using the two tools were large and positive. This may sound strange, as the tools are designed to assess risk for different things. But recall that the HCR-20 V3 is

designed to assess risk for general (i.e., any) violence, whereas the MLG is designed to assess risk for group-based violence, which is a specific form of violence that includes many (if not most) acts of terrorism. This means that anyone who is rated as a high risk for future group-based violence on the MLG should also be rated as a high risk for future violence on the HCR-20 V3; however, there may be many people who are rated as a high risk for general violence on the HCR-20 V3 who are not rated as a high risk for future group-based violence on the MLG. This may be true even for some terrorists, as not all terrorism is group-based violence—for example, some lone actor terrorists whose violence is linked to idiosyncratic beliefs or who operate in isolation, such as Theodore Kaczynski and Anders Bering Breivik.

Third, the association between MLG and HCR-20 V3 domain scores was consistent with expectations in some respects but not in others. On one hand, the presence and relevance ratings for the Individual domain of the MLG had positive correlations with the HCR-20 V3 total and domain ratings, whereas the other MLG domain and total ratings had negative or near-zero correlations with the with the HCR-20 V3 total and domain ratings. This was consistent with our expectations based on the fact that the risk factors in the Individual domain of the MLG were modelled after those in the HCR-20 V3. But contrary to expectations, none of the correlations was statistically significant (i.e., significantly different from 0). This may have been due to the very restricted variability in Individual domain ratings: all five cases received ratings of Y for I1, I2, and I3, and so the only difference between the cases in the Individual domain was for I4. Regardless, the pattern of findings suggests that the risk factors in the MLG

Individual-in-Group, Group, and Group-in-Society domains are generally unrelated to those in the HCR-20 V3.

Fourth, the association between the MLG and VERA domain scores was consistent with expectations only in limited respects. As expected, the MLG Individual domain scores were not significantly correlated with any of the VERA domain scores, although the correlations varied in direction and magnitude. But only the VERA Contextual domain scores had large, positive, and statistically significant correlations with the MLG Individual-in-Group, Group, and Group-in-Society domain scores. These findings suggest that the risk factors in the MLG Individual domain are generally unrelated to those in the VERA. Surprisingly, however, they also suggest that the overlap between the remaining MLG and VERA risk factors may be limited in scope, or perhaps that the content of the MLG and VERA domains is diverse—possibly too diverse to support the formation of composite domain scores.

## 3.0 Study Two

### 3.1 Method

Three researchers, all familiar with the content the VERA 2 and trained in the use of the second edition of the MLG, rated the content overlap risk factors on the two tools using a multi-step procedure. In the first step, Researcher A considered the VERA 2 risk factors one at a time and rated whether it overlapped in content with each of the MLG risk factors, while at the same time Researcher B considered the MLG risk factors one at a time and rated whether it overlapped in content with each of the VERA 2 risk factors. Researchers A and B worked independently in this step, that is, blind to each other's

ratings. They made ratings on a simple dichotomous scale (*no* versus *yes*) based on surface similarity of the risk factors.

In the second step, after completing their individual ratings, Researchers A and B broke the blind, discussed their ratings, and made a set of final consensus ratings of overlap among the risk factors using the same dichotomous scale.

In the third step, Researcher C took each pair of overlapping items identified by Researchers A and B and made a judgment of the *degree* of overlap on a 3-point scale (0 = *none*, 1 = *low*, 2 = *moderate*, 3 = *high*) based on the extent to which the names, definitions, and descriptions of the risk factors were similar from intensional, extensional, and ostensional perspectives. The intensional perspective focuses on core principles or properties; the extensional perspective, on enumeration of specific features; and the ostensional perspective, on illustrative features or exemplars. In practical terms, we expect that the overlap rating for a given pair of VERA 2 and MLG risk factors will reflect the magnitude of the correlations among presence ratings for those two risk factors observed in field research.

## 3.2 Results

### 3.2.1 Overall Overlap

Table 12 illustrates of the overall degree of overlap among the VERA 2 and MLG risk factors. The figure is a cross-tabulation of the VERA 2 risk factors (rows) and MLG risk factors (columns), organized by domains. The individual cells are shaded to reflect ratings of the degree of overlap: white or empty = 0 or *none*, light gray = 1 or *low,* dark gray = 2 or *moderate*, and black = 3 or *high*. Overall, there was at least *low* overlap among

80 pairs of risk factors out of a total of 31 x 16 = 496 possible pairs, or 16% of the theoretical maximum.

Overlap as indicated by the numerical ratings is summarized in Table 13. The overall overlap (i.e., Total by Total) was 167 out a possible total of 496 x 3 = 1488, or 11% of the theoretical maximum.

**3.2.2 VERA 2 Risk Factors**

Looking at the rows of Table 12, all VERA 2 risk factors overlapped to some degree with at least one MLG risk factor. The degree of overlap for the VERA 2 risk factors was similar across the 5 domains. To clarify this, Table 13 presents the total of numerical ratings of overlap within each of the domains, expressed as a percentage of the theoretical maximum. The percentage of overlap for the 5 domains clustered between from 9% and 13%.

Turning back to Table 12, it is apparent that three VERA 2 risk factors accounted for most of the overlap with the MLG. P4 (Involvement with non-violent, de-radicalization, offence-related programs) had *high* overlap with 4 MLG risk factors and *moderate* overlap with one risk factor. BA5 (Feelings of hate, frustration, persecution, alienation) had *high* overlap with 2 MLG risk factors and *moderate* overlap with another 2 risk factors. CI6 (Expressed intent to plan, prepare violent action) had *high* overlap with one MLG risk factor and *moderate* overlap with 3 risk factors. In terms of total numerical ratings for these VERA 2 risk factors, P4 had a score of 14 (29% of the theoretical maximum of the theoretical maximum of 16 x 3 = 48); BA5 had a score of 10 (21% of the theoretical maximum); and CI6 had a score of 9 (19% of the theoretical

maximum). Together, these three risk factors account for 20% of the total observed overlap between the VERA 2 and MLG.

In most other cases, the degree of overlap for the VERA 2 risk factors was *moderate* or *high* with at least one risk factor or, in the alternative, *low* with multiple MLG risk factors. There were three exceptions: HC1 (Early exposure to pro-violence militant ideology), CM5 (Driven by criminal opportunism), and P5 (Community support for non-violence) each had *low* overlap with one MLG risk factor. In numerical terms, their overlap ratings all were 1 (2% of the theoretical maximum).

**3.2.3 MLG Risk Factors**

Looking next at the columns of Table 12, only 12 of the 16 MLG risk factors overlapped to some degree with the VERA 2 risk factors. The pattern of overlap for MLG risk factors differed across domains: Only 8 of the 16 MLG risk factors, all from the Individual and Individual-in-Group domains, had moderate or high overlap with at least one VERA 2 risk factor. Second, the other 8 MLG risk factors, all from the Group and Group-in-Society domains, had much less overlap with the VERA 2 risk factors. Looking at Table 13, the percentage overlap for the I and I-G domains was 20% and 23%, respectively, whereas the percentage overlap for the G and G-S domains was 2% and 1%, respectively.

Three MLG risk factors accounted for most of the overlap with the VERA 2 risk factors. I2 (Attitude problems) had *high* overlap with 11 VERA 2 risk factors, *moderate* overlap with 7 risk factors, and *low* overlap with another 7 risk factors. IG4 (Negative attitude toward people outside the group) had *high* overlap with 7 VERA 2 risk factors,

*moderate* overlap with another 7 risk factors, and *low* overlap with 1 risk factor. IG3 (Strong commitment to group) had *high* overlap with 4 VERA 2 risk factors, *moderate* overlap with 6 risk factors, and *low* overlap with 4 risk factors. In terms of total numerical ratings for these MLG risk factors, I2 had a score of 54 (58% of the theoretical maximum of the theoretical maximum of 31 x 3 = 93); IG4 had a score of 36 (39% of the theoretical maximum); and IG3 had a score of 28 (30% of the theoretical maximum). Together, these three risk factors account for 71% of the total observed overlap between the MLG and VERA 2.

A total of four MLG risk factors had no overlap with VERA 2 risk factors: G4 (Strong leadership/power structure), GS2 (Socially isolated/isolative), GS3 (Operating in an unstable context/environment), and GS4 (Threatened by or in conflict with other groups). The numerical ratings for each was 0 (0% of the theoretical maximum).

### 3.3 Discussion

At first glance, the overall overlap findings in Table 12 suggest relatively low overall overlap between the VERA 2 and MLG risk factors. But a closer look indicates this is not the case. First, each of the VERA 2 risk factors had substantial overlap with one or more MLG risk factors, and the overlap was consistent across VERA 2 domains. Second, there is substantial overlap between the MLG and VERA 2, although this varies markedly as a function of MLG domain. Specifically, the I and IG domains overlap substantially with the VERA 2, with a clear majority of that overlap being attributable to three MLG risk factors; in contrast, the G and GS domains overlap very little, if at all, with the VERA 2.

This asymmetric association between the content of the VERA 2 and MLG suggests that there is relatively little unique or non-redundant content in the VERA 2. Indeed, most of the content of the VERA 2 could be accounted for by 3 MLG risk factors, suggesting that many of the VERA 2 risk factors reflect specific aspects of what the MLG considers to be more general problems (i.e., attitudes that support or condone criminality or violence, negative attitudes toward people outside the group, group commitment). The G and GS domains of the MLG, however, have considerable unique or non-redundant content. Put simply, most of what is measured by the VERA 2 risk factors is also measured by the MLG risk factors in the I and IG domains, but what is measured by the MLG risk factors in the G and GS domains is not measured by the VERA.

## 4.0 Conclusion

### 4.1 Summary

Despite the obvious limitations of the studies described herein—primarily, forced reliance on a series of 5 open-source cases for Study One, and inability to access the VERA 2 for Study Two—our findings, and indeed the research process itself, provided valuable insight into the nature of the HCR-20 V3, VERA and VERA 2, and MLG and how these tools could be used in a complementary manner for terrorism risk assessment. Let us discuss what we learned about each of the tools in turn.

With respect to the HCR-20 V3, this tool provides a comprehensive framework for evaluation of individual-level risk factors for violence (i.e., those reflecting a person's social and psychological adjustment, both past and recent). Consideration of individual-level risk factors is essential in all terrorism risk assessments, for three reasons. First, terrorism is a specific form of violence, which in turn is a specific form of antisocial behavior. Terrorism is differentiated from other forms of violence by the presence of motives and goals related to furthering social or political change. Second, although all terrorism is characterized by the presence of motives and goals related to furthering social or political change, other motives and goals may also be present, some of which may reflect personal or idiosyncratic (as opposed to shared) desires, beliefs, attitudes, pathology, and activity. Third, many or even most people who commit violence are not specialists, engaging in only a single type of violence and for the same motive or goal, but rather engage in violence that is diverse in nature. This means it would be foolish to assume that a person who is known or suspected of terrorism can

be at risk only for terrorism and not for other forms of violence—that is, as foolish as assuming that all people known or suspected of terrorism are at risk for other forms of violence. The point here is that terrorists are heterogeneous, which is why comprehensive risk assessment is essential for good case management and why consideration of individual-level risk factors is essential in comprehensive risk assessment. Our findings indicate that the MLG provides a rather general or crude assessment of individual-level risk factors (those in the Individual domain), and the VERA/VERA 2 provide a very limited assessment of them, suggesting that it may be important to include the HCR-20 V3 (or similar tool) in all comprehensive terrorism risk assessments. Consistent with the conclusion, Version 2 of the MLG explicitly permits and even encourages evaluators to use the HCR-20 V3 risk factors instead of the MLG's Individual domain risk factors.

With respect to the VERA/VERA 2, these tools provide a detailed analysis of "extremist" desires, belief, and attitudes—those that support or condone terrorism. The VERA/VERA 2 overlap very little with HCR-20 V3, except for a single risk factor (H9, Violent attitudes). The VERA/VERA 2 overlap more with the MLG, although the overlap is limited almost exclusively to two MLG domains (Individual and Individual-in-Group) and primarily to two MLG risk factors (I2, Attitude problems; IG4, Negative attitude toward people outside the group). On the surface, these findings suggest that the VERA/VERA 2 has content that is non-redundant with the HCR-20 V3 and MLG and therefore has clear "added value" in terrorism risk assessment. But the results of Study Two indicated that every VERA 2 risk factor overlapped to some extent with one

or more MLG risk factors, whereas a number of MLG risk factors did not overlap with VERA 2 risk factors. It appears that the VERA 2 risk factors are framed at a different level of specificity than those in the MLG: the former are more specific, whereas the latter are more general. Put differently, many of the VERA 2 risk factors reflect different facets or aspects of extremist desires, belief, and attitudes, rather than distinct risk factors. This suggests to us that the VERA 2 may be useful, or even essential, for the detailed assessment of extremist desires, belief, and attitudes that are assessed only in general terms by the HCR-20 V3 (risk factor H9) or MLG (risk factors I2 and IG4).

With respect to the MLG, our findings suggest that this tool provides a satisfactory, although admittedly rather general, assessment of individual-level risk factors and extremist desires, belief, and attitudes (i.e., through the risk factor in the Individual and Individual-in-Group domains), but also uniquely assesses higher-level risk factors (i.e., those in the Group and Group-in-Society domains) that are not captured by the HCR-20 V3 or VERA/VERA 2. We found the MLG particularly helpful in evaluating cases in which the terrorism was truly group-based violence (i.e., the person was acting in concert with and supported by others), such as in the case of Patrick Magee, versus individual violence (i.e., the person was motivated by idiosyncratic desires, belief, and attitudes), such as in the case of Theodore Kaczynski. We note, however, that the MLG is intended to assess risk for the former, not the latter. For example, our analysis of Theodore Kaczynski was that, although he presented a high risk for (individual) violence as assessed by the HCR-20 V3, he was a low risk for group-based violence as assessed by the MLG.

### 4.2 Recommendations

Based on our experiences conducting the research described herein, we conclude with recommendations for research and practice.

First, with respect to research, there is a clear need for multiple studies that directly compare tools that may be useful for terrorism risk assessment, including the HCR-20 V3, VERA 2, and MLG, as well as the ERG 22+ and TRAP-18. One possible approach is large-scale empirical research to examine the concurrent validity of the tools in a series of routine case files from national security, law enforcement, or corrections settings. Each series should include at least 25 cases, and ideally 50 or more cases, to permit appropriate statistical analyses of interrater reliability and concurrent validity. Of particular interest would be a series comprising multiple members of the same terrorist group, with those members having played different roles in the planning or perpetration of violence; similarly, it would be useful to have a series of lone actor terrorists. Based on our experience, we strongly recommend that the evaluators who administer the tools should be adequately trained (i.e., by the authors of the tools) and experienced. The tools are quite complex and sophisticated, and their optimal use requires more than a cursory review of the manuals or a day of training and practice. It would also be ideal for different evaluators to administer each tool (to prevent contamination of ratings) and to have multiple independent evaluators for each tool (to permit evaluation of interrater reliability and subsequent generation of consensus ratings for evaluation of concurrent validity). But other approaches to research are both possible and potentially valuable. For example, there is a need for more detailed

analysis of the tools using qualitative methods to identify similarities and differences with respect to key terms and concepts (such as "terrorism," "ideology," "attitudes," "group," or "violence") and theoretical assumptions (such as the nature and extent of the association between "attitudes" and "violence"). Similarly, detailed case studies may help to clarify similarities and differences in administration procedures, such as the way in which they handle information that is missing, incomplete, or uncertain. (Of course, there are many other research topics that could be discussed aside from the concurrent validity of the tools—such as consumer satisfaction with them, their legal admissibility, and the extent to which they guide case management—but they are beyond the scope of the current paper.)

Second, with respect to practice, in light of the incomplete overlap among the tools observed in this study, we believe it is premature to conclude that one or more tools is not useful or should be avoided. Instead, we recommend that comprehensive terrorism risk assessments routinely incorporate multiple tools. Evaluators should consider using tools such as the HCR-20 V3 to assess risk for general violence. These tools can provide good information concerning the presence and relevance individual-level risk factors. This information can be incorporated in any terrorism risk assessment, but also can be used to assess a person's risk for non-terrorist violence—a hazard that may be relevant in a substantial proportion of terrorism cases. Evaluators should also consider using tools such as the VERA 2 (or ERG 22+) that focus on the presence and relevance of risk factors related to extremist desires, beliefs, and attitudes. Finally, evaluators should consider using tools that assess the presence and relevance of risk

factors related to group dynamics, such as the MLG. Note that using multiple tools does not increase the time necessary to conduct a risk assessment, as the clear majority of assessment time is spent gathering information; once that information has been gathered, analyzing it within multiple frameworks (i.e., different tools) requires minimal cost in terms of time and effort. Although the process of reconciling the findings of different assessment tools can be difficult, we believe it has the potential to generate important insights into and deeper understanding of cases.

## 5.0 References

Andrews, D. A., & Bonta, J. L. (1995). *LSI–R: The Level of Service Inventory – Revised*. Toronto: Multi-Health Systems.

Beardsley, N. L., & Beech, A. R. (2013). Applying the violent extremist risk assessment (VERA) to a sample of terrorist case studies. *Journal of Aggression, Conflict and Peace Research*, *5*, 4-15. doi: 10.1108/17596591311290713

Burton, S., & Amat, G. (2013, April). *Application of police authored threat assessment: Witness protection cases involving organized crime*. Paper presented to the annual meeting of the Association of European Threat Assessment Professionals, Vienna, Austria.

Cook, A. N. (2014). *Risk assessment and management of group-based violence*. Unpublished doctoral dissertation, Department of Psychology, Simon Fraser University, Burnaby, Canada. Available at: http://summit.sfu.ca/system/files/iritems1/14289/etd8437_ACook.pdf

Cook, A. N., Hart, S. D., & Kropp, P. R. (2013). *Multi-Level Guidelines for the assessment and management of group-based violence*, Version 1. Burnaby, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.

Cook, A. N., Hart, S. D., & Kropp, P. R. (2014). *Multi-Level Guidelines for the assessment and management of group-based violence*, Version 2. Burnaby, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.

*Criminal Code,* RSC 1985, c C-46 s 745.

Dernevik, M., Beck, A., Grann, M., Hogue, T., & McGuire, J. (2009). The use of

psychiatric and psychological evidence in the assessment of terrorist offenders. *Journal of Forensic Psychiatry & Psychology*, 20, 508-515. doi: 10.1080/13501760902771217

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V 3: Assessing risk of violence – User guide*. Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*, 2nd ed. New York: John Wiley & Sons.

Grabo, C. M. (2002). *Anticipating surprise: Analysis for strategic warning*. Available at: http://www.ni-u.edu/ni_press/pdf/Anticipating_Surprise_Analysis.pdf

Gudjonsson, G. H. (2009). The assessment of terrorist offenders: A commentary on the Dernevik et al. article and suggestions for future directions. *Journal of Forensic Psychiatry and Psychology, 20*, 516-519. doi: 10.1080/13501760902771233

Hart, S. D. (2010, March). Risk for terrorism: A multi-level, multi-disciplinary approach to threat assessment. In M. Scalora (Chair), *The psychology of terrorism and extremist violence*. Invited address, annual meeting of the American Psychology-Law Society (Div. 41 of the American Psychological Association), Vancouver, Canada.

Hart, S. D., & Dormond, K. (2009, August). *Multi-level threat assessment of extremist violence*. Paper presented at the annual meeting of the Association of Threat Assessment Professionals, Anaheim, California.

Hart, S. D., Douglas, K. S., & Guy, L. S. (2016). The structured professional judgment approach to violence risk assessment: Origins, nature, and advances. In L. Craig &

M. Rettenberger (Volume Eds.) and D. Boer (Series Ed.), *The Wiley handbook on the theories, assessment, treatment of sexual offending: Volume II. Assessment* (pp. 643-666). Oxford, UK: Wiley.

Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *Manual for the Screening Version of the Hare Psychopathy Checklist–Revised (PCL:SV)*. Toronto: Multi-Health Systems.

Hart, S. D., Douglas, K. S., & Guy, L. S. (2016). The structured professional judgment approach to violence risk assessment: Origins, nature, and advances. In L. Craig & M. Rettenberger (Volume Eds.) and D. Boer (Series Ed.), *The Wiley handbook on the theories, assessment, treatment of sexual offending: Volume II. Assessment* (pp. 643-666). Oxford, UK: Wiley.

Heuer, R. J., Jr. (1999). *Psychology of intelligence analysis*. Langley, VA: Centre for the Study of Intelligence, Central Intelligence Agency. Available at: https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/PsychofIntelNew.pdf

Heuer, R. J., Jr., & Pherson, R. H. (2011). *Structured analytic techniques for intelligence analysis*. Washington, DC: CQ Press.

van Kuijk, P., & Voerman, B. (2016, April). *Developing a risk assessment and management approach in extremist cases*. Paper presented to the annual meeting of the Association of European Threat Assessment Professionals, Ghent, Belgium.

Lloyd, M., & Dean, C. (2016). The development of structured guidelines for assessing risk in extremist offenders. *Journal of Threat Assessment and Management*, *2*, 40-52.

http://dx.doi.org/10.1037/tam0000035

Meloy, J. R., & Gill, P. (2016). The lone-actor terrorist and the TRAP-18. *Journal of Threat Assessment and Management*, *3*, 37-52. doi:10.1037/tam0000061

Meloy, J. R., Hart, S. D., & Hoffmann, J. (2014). Threat assessment and threat management. In J. R. Meloy & J. Hoffmann (Eds.), *International handbook of threat assessm*ent (pp. 3-17). New York: Oxford.

Meloy, J. R., & Hoffmann, J. (Eds.). (2014). *International handbook of threat assess*ment. New York: Oxford.

Monahan, J. (2012). The individual risk assessment of terrorism. *Psychology, Public Policy, and Law, 18*, 167-205. doi: 10.1037/a0025792

Otto, R. K., & Douglas, K. S. (Eds.). (2010). *Handbook of violence risk assessment*. New York: Routledge.

Pressman, D. E. (2009). *Risk assessment decision for violent political extremism*. Ottawa: Public Safety Canada. Available at: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/2009-02-rdv/2009-02-rdv-eng.pdf

Pressman, D. E. (2014). *Violent extremist risk assessment in intelligence analysis: Benefits and validity of the VERA 2 protocol*. Proceedings of the XIX International Conference, "Mihai Viteazul," Bucharest, Romania.

Pressman, D. E., & Flockton, J. (2010). *VERA 2. Violent Extremism Risk Assessment, Version 2 Manual*.

Pressman, D. E., & Flockton, J. (2012). Calibrating risk for violent political extremist and

terrorists: The VERA 2 structured assessment. *British Journal of Forensic Practice, 14*, 237-251.

Pressman, D. E., & Flockton, J. (2014). Violent extremist risk assessment: Issues and applications of the VERA 2 to the high security prison setting. In A. Silke (Ed.), *Prisons, terrorism and extremism: Critical Issues in management, radicalisation and reform* (pp. 122-143). New York: Routledge.

Pressman, D. E., & Ivan, C. (2016). Internet use and violent extremism: A cyber-VERA risk assessment protocol. In M. Khader, L. S. Neo, G. Ong, E. Tan, & J. Chin (Eds.), *Countering violent extremism and radicalisation in the digital era* (pp. 391-409). Hershey, PA: Information Science Reference.

Public Safety Canada. (2013). *Canada's counter-terrorism strategy*. Ottawa: Government of Canada.

Public Safety Canada. (2014). *Public report on the terrorist threat*. Ottawa: Government of Canada.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and meta-regression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*, 499-513. doi: 10.1016/j.cpr.2010.11.00

Strang, S. (2005). *Project SLEIPNIR: An analytic technique for operational priority setting*. https://analysis.mitre.org/proceedings/Final_Papers_Files/135_Camera_Ready_Paper.pdf

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for*

*violence*, Version 2. Burnaby, Canada: Mental Health, Law, & Policy Institute, Simon Fraser University.

Webster, S., Kerr, J., & Tompkins, C. (in press). *A process evaluation of the Structured Risk Guidance for extremist offenders: Final report*. London, UK: National Centre for Social Research.

United Nations. (2015). *United Nations action to counter terrorism*. Available at: http://www.un.org/en/terrorism/

Table 1

*VERA: Domains and Risk Factors*

| Domain | | Risk Factor |
|---|---|---|
| Attitude | A1 | Attachment to ideology justifying violence |
| | A2 | Perception of injustice and grievances |
| | A3 | Identification of target of injustice |
| | A4 | Dehumanisation of identified target |
| | A5 | Internalized martyrdom to die for cause |
| | A6 | Alienation from society and rejection of values |
| | A7 | Hate, frustration, persecution |
| | A8 | Need for group bonding and belonging |
| | A9 | Identity problems |
| | A10 | Empathy for those outside own group |
| Contextual | C1 | User of extremist websites |
| | C2 | Community support for violent action |
| | C3 | Direct contact with violent extremists |
| | C4 | Anger at political decisions, actions of country |
| Historical | H1 | Early exposure to violence |
| | H2 | Family/friends involvement in violent action |
| | H3 | Prior criminal violence |
| | H4 | State-sponsored military, paramilitary training |
| | H5 | Travel for non-state sponsored training/fighting |
| | H6 | Glorification of violent action |
| Protective | P1 | Shift in ideology |
| | P2 | Rejection of violence to obtain goals |
| | P3 | Change of vision of enemy |
| | P4 | Constructive political involvement |
| | P5 | Significant other/community support |

*Table continues…*

| Domain | | Risk Factor |
|---|---|---|
| Demographic | D1 | Sex (male) |
| | D2 | Married (< 1 year) |
| | D3 | Age (young) |

*Note.* VERA = Consultative Version 1 of the Violent Extremism Risk Assessment Protocol (Pressman, 2009).

Table 2

*VERA 2: Domains and Risk Factors*

| Domain | | Risk Factor |
|---|---|---|
| Beliefs and Attitudes | BA1 | Commitment to ideology justifying violence |
| | BA2 | Victim of injustice and grievances |
| | BA3 | Dehumanization/demonization of identified targets of injustice |
| | BA4 | Rejection of democratic society and values |
| | BA5 | Feelings of hate, frustration, persecution, alienation |
| | BA6 | Hostility to national collective identity |
| | BA7 | Lack of empathy, understanding outside own group |
| Context and Intent | CI1 | Seeker, consumer, developer of violent extremist materials |
| | CI2 | Identification of target (person, place, group) in response to perceived injustice |
| | CI3 | Personal contact with violent extremists |
| | CI4 | Anger and expressed intent to act violently |
| | CI5 | Expressed desire to die for cause or martyrdom |
| | CI6 | Expressed intent to plan, prepare violent action |
| | CI7 | Susceptible to influence, authority, indoctrination |
| History and Capability | HC1 | Early exposure to pro-violence militant ideology |
| | HC2 | Network (family, friends) involved in violent action |
| | HC3 | Prior criminal history of violence |
| | HC4 | Tactical, paramilitary, explosives training |
| | HC5 | Extremist ideological training |
| | HC6 | Access to funds, resources, organizational skills |
| Commitment and Motivation | CM1 | Glorification of violent action |
| | CM2 | Driven by criminal opportunism |
| | CM3 | Commitment to group, group ideology |
| | CM4 | Driven by moral imperative, moral superiority |
| | CM5 | Driven by excitement, adventure |

*Table continues…*

| Domain | | Risk Factor |
|---|---|---|
| Protective | P1 | Re-interpretation of ideology less rigid, absolute |
| | P2 | Rejection of violence to obtain goals |
| | P3 | Change of vision of enemy |
| | P4 | Involvement with non-violent, de-radicalization, offence-related programs |
| | P5 | Community support for non-violence |
| | P6 | Family support for non-violence |

*Note.* VERA 2 = Consultative Version 2 of the Violent Extremism Risk Assessment Protocol (Pressman & Flockton, 2012).

Table 3

*MLG Version 2: Domains and Risk Factors*

| Domain | | Risk Factor |
|---|---|---|
| Individual | I1. | Conduct problems |
| | I2. | Attitude problems |
| | I3. | Social adjustment problems |
| | I4. | Mental health problems |
| Individual-Group | IG1. | Strong group-based identity |
| | IG2. | Violent role or status in group |
| | IG3. | Strong commitment to group |
| | IG4. | Negative attitude toward people outside the group |
| Group | G1. | History of violence |
| | G2. | Violent norms or goals |
| | G3. | Strong cohesion |
| | G4. | Strong leadership/power structure |
| Group-Societal | GS1. | Large in size/scope |
| | GS2. | Socially isolated/isolative |
| | GS3. | Operating in an unstable context/environment |
| | GS4. | Threatened by or in conflict with other groups |

*Note*. MLG = Version 2 of the Multi-Level Guidelines (Cook et al., 2014).

Table 4

*ERG 22+: Domains and Risk Factors*

| Domain | | Risk Factor |
|---|---|---|
| Engagement | E1 | Need to redress injustice and express grievance |
| | E2 | Need to defend against threat |
| | E3 | Need for identity, meaning, belonging |
| | E4 | Need for status |
| | E5 | Need for excitement, comradeship or adventure |
| | E6 | Need for dominance |
| | E7 | Susceptibility to indoctrination |
| | E8 | Political/moral motivation |
| | E9 | Opportunistic involvement |
| | E10 | Family or friends support extremist offending |
| | E11 | Transitional periods |
| | E12 | Group influence and control |
| | E13 | Mental health |
| Intent | I1 | Over-identification with a group or cause |
| | I2 | Us and Them thinking |
| | I3 | Dehumanisation of the enemy |
| | I4 | Attitudes that justify offending |
| | I5 | Harmful means to an end |
| | I6 | Harmful end objectives |
| Capability | C1 | Individual knowledge, skills and competencies |
| | C2 | Access to networks, funding and equipment |
| | C3 | Criminal history |
| + | Any other factor | |

*Note*. ERG 22+ = Extremism Risk Guidelines (Lloyd & Dean, 2015).

Table 5

*HCR-20 V3: Domains and Risk Factors*

| Domain | | Risk Factor |
|---|---|---|
| Historical | H1 | Violence |
| | H2 | Other antisocial behavior |
| | H3 | Relationships |
| | H4 | Employment |
| | H5 | Substance use |
| | H6 | Major mental disorder |
| | H7 | Personality disorder |
| | H8 | Traumatic experiences |
| | H9 | Violent attitudes |
| | H10 | Treatment or supervision response |
| Clinical | C1 | Insight |
| | C2 | Violent ideation or intent |
| | C3 | Symptoms of a major mental disorder |
| | C4 | Instability |
| | C5 | Treatment or supervision response |
| Risk Management | R1 | Professional services and plans |
| | R2 | Living situation |
| | R3 | Personal support |
| | R4 | Treatment and supervision response |
| | R5 | Stress or coping |

*Note*. HCR-20 V3 = Version 3 of the Historical-Clinical-Risk Management—20 (Douglas et al., 2013).

Table 6

*Interrater Reliability: MLG Summary Risk Ratings*

| Summary Risk Rating | $ICC_1$ | 95%CI | *p* |
|---|---|---|---|
| Future Violence | .71 | [-.43, .97] | .074 |
| Serious Physical Harm | .40 | [-.23, .90] | .138 |
| Imminent Violence | .50 | [-.27, .93] | .126 |

*Note*. $N = 5$. $ICC_1$ = intraclass correlation coefficient, single rater, 2-way random effects model, absolute agreement.

Table 7

*Interrater Reliability: MLG Presence and Relevance Ratings, Total and Domain*

| | Presence | | | Relevance | | |
|---|---|---|---|---|---|---|
| **Total/Domain** | **$ICC_1$** | **95%CI** | ***p*** | **$ICC_1$** | **95%CI** | ***p*** |
| **Total** | **.99** | **[.90, 1.00]** | **< .001** | **.89** | **[.38, .99]** | **.010** |
| Individual | 1.00 | -- | -- | .80 | [.10, .98] | .028 |
| Individual-in-Group | .95 | [.65, .99] | .002 | .78 | [-.02, .97] | .039 |
| Group | .99 | [.93, 1.00] | < .001 | .92 | [.51, .99] | .003 |
| Group-in-Society | .87 | [.16, .99] | .017 | .57 | [-.55, .95] | .139 |

*Note*. $N = 5$. $ICC_1$ = intraclass correlation coefficient, single rater, 2-way random effects model, absolute agreement. -- = could not be calculated due to lack of variability (perfect agreement).

Table 8

*Concurrent Validity: Correlation Between MLG and HCR-20 V3 Summary Risk Ratings*

| | HCR-20 V3 | | |
|---|---|---|---|
| **MLG** | **Future Violence** | **Serious Physical Harm** | **Imminent Violence** |
| Future Violence | .71* | 1.00** | .71* |
| Serious Physical Harm | .50 | .83** | .60* |
| Imminent Violence | .54* | .71** | .33 |

*Note*. $N = 5$. * $p < .10$; ** $p < .05$.

Table 9

*Concurrent Validity: Correlation Between MLG and HCR-20 V3 Presence Ratings, Total and Domain*

| | HCR-20 V3 | | | |
|---|---|---|---|---|
| **MLG** | **Total** | **Historical** | **Clinical** | **Risk Management** |
| **Total** | **-.45** | **-.27** | **-.36** | **-.57** |
| Individual | .37 | .55 | .16 | .22 |
| Individual-in-Group | -.24 | -.10 | -.10 | -.40 |
| Group | -.39 | -.25 | -.25 | -.52 |
| Group-in-Society | -.67 | -.49 | -.68 | -.66 |

*Note*. $N = 5$. All *n.s.*

Table 10

*Concurrent Validity: Correlation Between MLG and HCR-20 V3 Relevance Ratings, Total and Domain*

| | HCR-20 V3 | | | |
|---|---|---|---|---|
| **MLG** | **Total** | **Historical** | **Clinical** | **Risk Management** |
| **Total** | **-.10** | **-.04** | **-.13** | **-.13** |
| Individual | .58 | .52 | .44 | .70 |
| Individual-in-Group | .06 | .15 | .02 | -.01 |
| Group | -.06 | -.03 | -.04 | -.11 |
| Group-in-Society | -.45 | -.40 | -.47 | -.45 |

*Note*. $N = 5$. All *n.s*.

Table 11

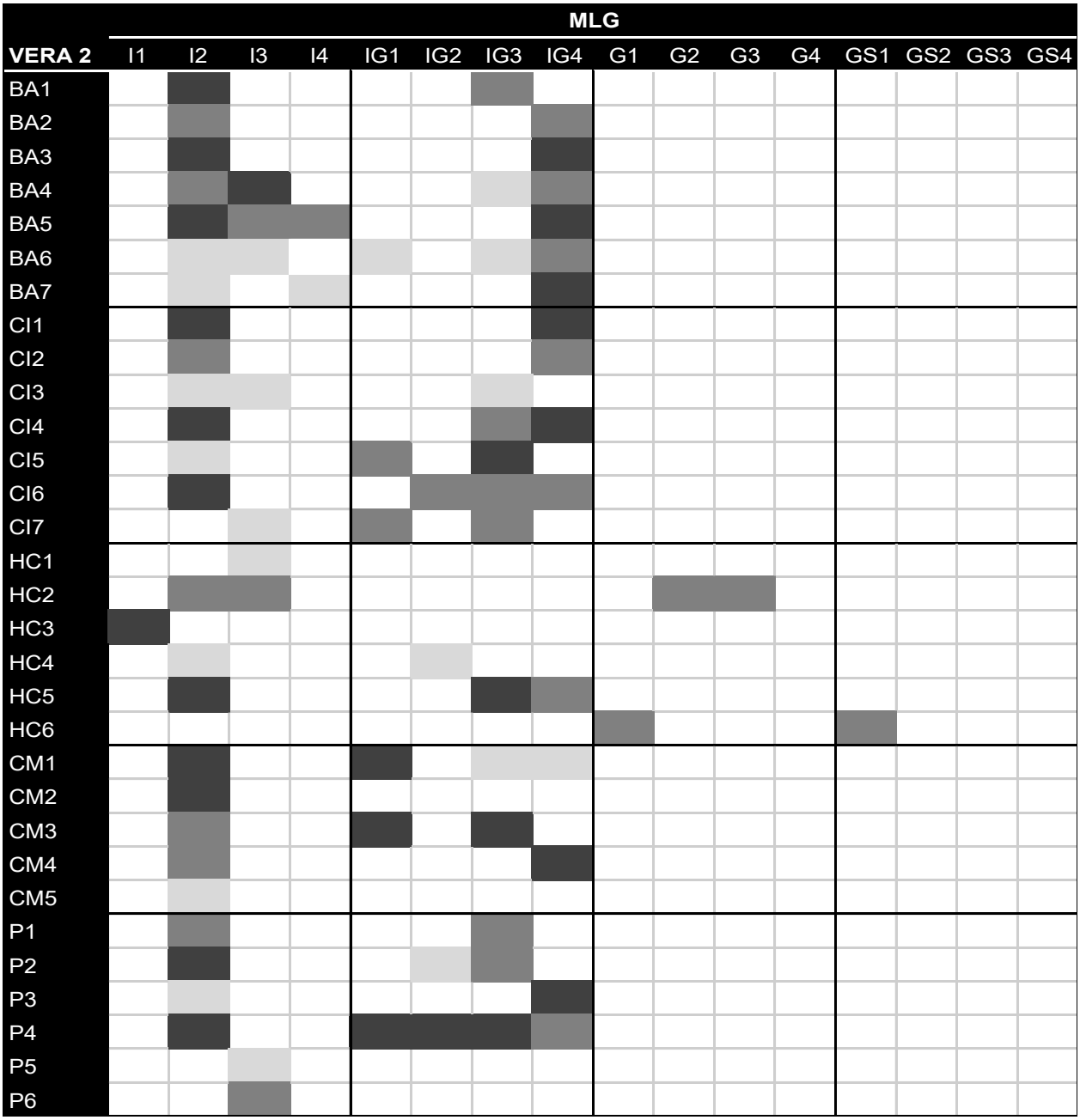
*Concurrent Validity: Correlation Between MLG and VERA Presence Ratings, Domain and Total*

| | VERA | | | | | |
|---|---|---|---|---|---|---|
| **MLG** | **Total** | **A** | **C** | **H** | **P** | **D** |
| **Total** | **.69** | **-.34** | **.96**** | **.57** | **.62** | **-.62** |
| Individual | -.17 | .56 | -.63 | .00 | -.65 | .65 |
| Individual-in-Group | .60 | -.38 | .90** | .56 | .47 | -.47 |
| Group | .56 | -.47 | .90** | .48 | .61 | -.61 |
| Group-in-Society | .73 | -.17 | .92** | .49 | .74 | -.74 |

*Note.* $N = 5$. For VERA domains, A = Attitude, C = Contextual, H = Historical, P = Protective, D = Demographic. * $p < .10$; ** $p < .05$.

Table 12

*Degree of Overlap of VERA 2 and MLG Risk Factors*




*Note.* Rows are VERA 2 risk factors; columns are MLG risk factors. VERA 2 domains: BA = Beliefs and Attitudes; CI = Context and Intent; HC = History and Capability; and CM = Commitment and Motivation. MLG domains: I = Individual; I-G = Individual-in-Group: G = Group; and G-S = Group-in-Society. Shading of cells reflects ratings of the degree of overlap: white or empty = 0 or *none*; light gray = 1 or *low*; dark gray = 2 or *moderate*, and black = 3 or *high*.

Table 13

*Degree of Overlap of VERA 2 and MLG: Domains and Total*

| | MLG | | | | |
|---|---|---|---|---|---|
| **VERA 2** | **I** | **I-G** | **G** | **G-S** | **Total** |
| BA | 29% | 24% | 0% | 0% | 13% |
| CI | 18% | 31% | 0% | 0% | 12% |
| HC | 17% | 8% | 8% | 3% | 9% |
| CM | 18% | 23% | 0% | 0% | 10% |
| P | 17% | 26% | 0% | 0% | 11% |
| Total | 20% | 23% | 2% | 1% | 11% |

*Note.* VERA 2 domains: BA = Beliefs and Attitudes; CI = Context and Intent; HC = History and Capability; and CM = Commitment and Motivation. MLG domains: I = Individual; I-G = Individual-in-Group: G = Group; and G-S = Group-in-Society.