



Assessing the Risk of Violent Extremism in Prisons: A Comparison between VERA-2R and DRAVY-3

Roberto M. Lobato¹, Sandra Chiclana^{2,3}, Laura Blanco², José L. González-Álvarez^{4,5} and Ángel Gómez^{2,6}

¹Departamento de Ciencias de la Salud, Universidad de Burgos, Spain; ²Departamento de Psicología Social y de las Organizaciones, Universidad Nacional de Educación a Distancia, Madrid, Spain; ³Secretaría General de Instituciones Penitenciarias, Ministerio del Interior, Madrid Spain; ⁴Centro de Investigación en Ciencias Forenses y de la Seguridad, Universidad Autónoma de Madrid, Madrid, Spain; ⁵Servicio de Psicología de la Guardia Civil, Ministerio del Interior, Spain; ⁶ARTIS International, St. Michaels, Maryland, United States

ARTICLE INFO

Article history:

Received 15 October 2024
Accepted 30 August 2025

Keywords:

Violent extremism
Risk assessment
Prison
VERA-2R
DRAVY-3

ABSTRACT

Developing tools to assess the risk of violent extremism in prison is a common concern for penitentiary institutions. In order for first-line practitioners in the penitentiary context to become familiar with and better understand these tools, improve their evaluations, and even adapt or develop new instruments, the aim of this manuscript is to compare the tools and to delve into the strengths and weaknesses of two instruments with different characteristics but similar objectives: the VERA-2R and the DRAVY-3. A systematic comparison of their development, characteristics, structure, indicators, and validity is presented. The analyses show that these tools differ from their theoretical conception and their more specific objectives. However, they also show similarities at the indicator level with several dimensions overlapping. Regarding their validity, both have limitations that should be addressed in future studies. Finally, a review of the main differences and limitations is presented, providing avenues to improve these tools.

La evaluación del riesgo de extremismo violento en prisiones: una comparación entre VERA-2R y DRAVY-3

RESUMEN

El desarrollo de herramientas para evaluar el riesgo de extremismo violento en prisión es una preocupación común de las instituciones penitenciarias. El objetivo de este manuscrito es comparar las fortalezas y debilidades de dos instrumentos con características diferentes, pero objetivos similares y profundizar en ellas, el VERA-2R y el DRAVY-3, con el fin de que los profesionales de primera línea en el contexto penitenciario conozcan y comprendan mejor estas herramientas, mejoren su evaluación e incluso adapten o desarrollen nuevos instrumentos. Se presenta una comparación sistemática de su desarrollo, características, estructura, indicadores y validez. Los análisis muestran que estas herramientas difieren en su concepción teórica y sus objetivos más específicos. Sin embargo, también muestran semejanzas en sus indicadores, con varias dimensiones superpuestas. En cuanto a su validez, ambas tienen limitaciones que deberían abordarse en futuros estudios. Finalmente, se ofrece una revisión de las principales diferencias y limitaciones, señalando algunas formas de mejorar estas herramientas.

Palabras clave:

Extremismo violento
Evaluación de riesgos
Prisión
VERA-2R
DRAVY-3

Risk assessment, a process that involves the systematic collection and interpretation of information relating to an individual to predict the likelihood that he/she will perform the behavior of concern in the future (Herrington & Roberts, 2012), has become one of the most important practices in prison management of violent extremism. In the context of prisons, risk assessment may help provide early warning signs of radicalization among prisoners, classify prisoners according to their level of dangerousness, assign them to different types of disengagement interventions, and manage prison release and probation (Monahan, 2012). Given the

relevance of these tasks, different tools have been developed for risk assessment in an attempt to minimize bias (Cook, 2014; Lloyd & Dean, 2015; Meloy & Gill, 2016).

The first, and probably best-known, violent extremism risk assessment tool is the VERA—Violent Extremism Risk Assessment (Pressman et al., 2016). Since its appearance, new tools with different characteristics have emerged to overcome some of its potential limitations. A useful strategy to improve the tools already existing, but also to help in the development of new tools, is to make a comparison between them (Fernandez & de Lasala, 2021;

Cite this article as: Lobato, R. M., Chiclana, S., Blanco, L., González-Álvarez, J. L., & Gómez, A. (2026). Assessing the risk of violent extremism in prisons: A comparison between VERA-2R and DRAVY-3. *Anuario de Psicología Jurídica*, 36, Article e260475, 1-11. <https://doi.org/10.5093/apj2026a8>

Funding: This research was supported by the ERC Advanced Grant Agreement N° 101018172, A Multi-Theory Multi-Method Approach for Preventing and Reducing Radicalization leading to Violence – MULTIPREV, and the project PID2021-124617OB-I00 funded by the Spanish Ministry of Science, Innovation, and Universities (Spain).

Correspondence: agomez@psi.uned.es (Á. Gómez).

Lloyd, 2019; Lobato & García-Coll, 2022; Scarcella et al., 2016; van der Heide et al., 2019). This strategy provides insights into the advantages and limitations of each tool as compared to others, what allows to improve them and to understand in which situations it is more appropriated to use one tool over another. Thus, in order to raise awareness and facilitate the selection of such tools in the prison context for first-line practitioners, this research aims to compare the most recent version of the VERA, as it is the VERA-2R (Pressman et al., 2016), with the DRAVY-3 (González-Álvarez et al., 2022), a recent instrument to assess the risk of violent jihadist extremism in prisons that has undergone an exhaustive process of revision and updating in its short history.

Therefore, after introducing the main characteristics of the VERA-2R and the DRAVY-3, the objectives of this investigation are: 1) to compare their structure, the differences, and the possible overlap of their indicators; 2) to examine their reliability and validity; and 3) to point out their limitations and areas for improvement.

Development of VERA and DRAVY

The VERA was developed in 2009 as a specific assessment tool for violent extremists and terrorists under the format of a structured professional judgement (Pressman, 2009). Initially, it was originated to be used by forensic psychologists and clinical experts in a high-risk correctional terrorism unit in Australia (Pressman & Flockton, 2012). It was developed from a review of the literature and discussions with security and intelligence professionals with experience in violent criminal extremists. At the time of its creation, the VERA was the first instrument of its kind (Pressman, 2009). In 2010, the protocol was updated with the release of its second version, the VERA-2 (Pressman & Flockton, 2010). This update was made based on feedback from evaluators who tested the tool in a high-security prison in Australia. Feedback was received from several groups of experts, such as psychologists and psychiatrists in the justice system, intelligence and security analysts, professionals working with radicalized individuals in prison settings, users from local and national law enforcement agencies, and specialists in risk assessment and terrorism (Pressman, 2016). Finally, after an additional review of the literature, the latest new version, the VERA-2-Revised (VERA-2R), appeared in 2015 (RAN, 2021). In addition, in recent years, a shorter version has been created, the Violent Extremism Screening Analysis (VESA) approach (Pressman & Davis, 2022), although there is still no publicly available data on its performance.

The DRAVY (Detection of Violent Radicalization of Jihadist Etiology) was originated in 2018 in Spain (see Fernández, 2019; Nistal-Burón, 2019; Secretaría General de Instituciones Penitenciarias, 2018). Spain is distinguished by a large number of Islamist terrorist arrests, being one of the European countries with the highest levels of Islamist radicalization, and the site of some of Europe's worst terrorist attacks (e.g., the 2004 Madrid train bombings and the 2017 van and knife attacks in and around Barcelona). By Order PCI/179/2019, on February 22, 2019, the National Strategy against Terrorism (Ministerio del Interior, 2019), approved by the National Security Council, was published. In this plan, special references were made to Penitentiary Centers, urging the Penitentiary Authorities to carry out a follow-up and assessment of convicted and detained persons linked to jihadist oriented terrorist acts, as well as of those individuals involved during their stay in prison in violent extremist recruitment or indoctrination. The first version of the DRAVY was developed from a review of existing instruments, such as the TRAP-18 (Meloy & Gill, 2016), the VERA-2 (Pressman and Flockton, 2010), the ERG 22+ (Lloyd & Dean, 2015), and the Multi-Level Guidelines (MLG; Cook, 2014), and it was tested three times over a period of six months between each application (Loínaz, 2019). To overcome some of its limitations (e.g., poor predictive capacity and problems linked

to the definition of risk factors), and based on previous analyses and a review of the literature, the indicators were reformulated and some others were included (González-Álvarez et al., 2021). In this second version, with the aim of reducing subjectivity, the DRAVY-2 became a more actuarial instrument (addition of weighted indicators and establishment of cut-off points) than a clinical judgement instrument. Finally, in 2022, the third version, the DRAVY-3, appeared based on the inclusion of new indicators from research carried out in prisons (Gómez et al., 2021; Gómez, Atran, et al., 2022), the suggestions of former radicals who now help combat violent Islamic radicalization, and a detailed analysis of the results of previous applications (González-Álvarez et al., 2022).

Characteristics of the VERA-2R and the DRAVY-3

Goals

The VERA-2R focuses on the conscious use of violence for ideological purposes. Pressman and Flockton (2012, 2014) made a detailed revision of the VERA-2R, based on the assumption that terrorists are "normal" individuals who possess conscious control over their actions. The authors indicate that the VERA-2R is intended to assess the risk presented by ideologically motivated violent offenders. This commitment to ideology would allow them to perform violent actions (consciously) to achieve a greater good. This is a characteristic distinction from other common violent offenders (e.g., rapists, murderers, thieves, or aggressors). In this way, the VERA-2R protocol shares general objectives with other tools aimed at assessing the risk of violence, such as the Historical/Clinical/Risk Management-20 (HCR-20; Webster et al., 1997), focusing on establishing the risk of future violent acts and making use of the information obtained with the instrument to design interventions. While its first objective focuses on "predicting," the authors also emphasize that it is not a predictive tool and cannot determine who will reoffend. The result would consist of "a logical inductive estimate of the risk of violent extremism" (Pressman, 2016, p. 259). Similarly, the tool can also be used to provide early warning signs and risk trajectories for individuals who are suspected and under surveillance (Pressman, 2016), although the authors recommend caution in these cases and always keep legal and ethical issues in mind. Briefly, the VERA-2R allows for tracking the risk of violent extremism and monitoring it through repeated assessments over time to determine changes in risk and protection indicators (Lloyd, 2019). This also ties with their recommendation to use it to evaluate programs for disengagement (Pressman, 2016).

The DRAVY-3 aims to evaluate inmates (for terrorism-related and unrelated offenses) based on their level of general violence, extremist violence, ability to proselytize, and level of radicalization (Nistal-Burón, 2019). Additionally, it aims to establish the level of dangerousness (González-Álvarez et al., 2021). These purposes are framed within specific objectives related to aiding decision-making in terms of detection, follow-up, and intervention. The DRAVY-3 can be used to provide early warning signs of radicalization among prisoners. It is used for the classification of prisoners according to their level of risk of radicalization and use of extremist violence, and it is being used for the evaluation of disengagement programs (González-Álvarez et al., 2022; González-Álvarez et al., 2021). The tool has been applied in Spanish prisons at six-month intervals since its creation (González-Álvarez et al., 2021).

Target Population

The VERA-2R is specific to offenders of violent acts carried out in persistence of ideological objectives (Pressman & Flockton, 2014). Therefore, it is intended to assess the risk for "already radicalized

individuals" who have engaged in terrorist activities. The VERA-2R was developed to assess classic terrorists rather than lower-profile extremists (Herzog-Evans, 2018). Specifically, the VERA-2R finds its utility in assessing individuals incarcerated for acts of violent extremism or terrorism, but not for individuals incarcerated for common crimes but susceptible to be radicalized. In addition, it can be used with the entire spectrum of violent extremists (regardless of the ideology and severity of the crime): right-wing violent extremists, left-wing violent extremists, animal rights advocates, violent environmentalists, violent anarchists, violent anti-abortionists, and all other violent offenders motivated by social, religious, or political ideology (Pressman & Flockton, 2014), including lone actors (Lloyd, 2019). Having a single assessment tool for a broad spectrum of terrorists restricts the possibility of bias or focusing on a single ideology (Pressman & Flockton, 2012). It can also be used with women and men equally and with youth (Lloyd, 2019; Pressman, 2016). Finally, it should be noted that the target population also differs depending on when the assessment is applied, with some countries using it to assess risk during parole, while others use it for pre-trial risk assessment (van der Heide et al., 2019).

In contrast, the DRAVY-3, unlike the VERA-2R, is not designed to be used with the entire spectrum of violent extremists, but only with those who adhere to a jihadist ideology, regardless of the severity of the crime (González-Álvarez et al., 2022; González-Álvarez et al., 2021). The DRAVY-3 is applied to inmates over 18 years of age at different stages of radicalization. These target populations are in line with the classification of the Spanish prison system into groups A (inmates convicted for jihadist terrorism), B (inmates convicted for reasons unrelated to jihadist terrorism but suspected of radicalizing others in prisons), and C (inmates vulnerable to recruitment for further radicalization) (Santos-Hermoso et al., 2021). It has also been used with women, although they were excluded from the analyses in the tool's evaluations due to their low number (González-Álvarez et al., 2022).

Structure of Dimensions and Indicators

Because both tools have several versions, to evaluate and compare the dimensions and indicators we focus on the most recent versions: VERA-2R (Pressman et al., 2016) and DRAVY-3 (González-Álvarez et al., 2022). On the one hand, the dimensions are the theoretical constructs into which the indicators are grouped, which are the descriptions of the characteristics to be evaluated. On the other hand, the indicators are the specific factors that increase or decrease the risk and whose presence must be evaluated to determine the overall risk.

The VERA-2R has 34 indicators organized in five dimensions (Pressman et al., 2016; Pressman & Flockton, 2014). The first dimension, with seven indicators, refers to beliefs, attitudes, and ideology; perceived grievances and injustices; identification of causes or persons responsible for grievances; moral emotions; alienation; the individual's relationship with the laws and norms of the state; and affiliation with an extremist group (Pressman et al., 2016). The second dimension is focused on social context and intentions and presents seven indicators related to the conscious objective for using violence to support an ideology, and to the cultural and social contexts, such as preferences, personal contacts, family, and friends, which may serve to encourage the actual use of violence to achieve ideological objectives (Pressman et al., 2016). The third dimension is dedicated to history, action, and capacity, and contains six indicators. It is focused on individual's ability to plan and carry out a violent extremist attack. These include a criminal or violent past, the training that an individual has received, access to the people, resources, and materials needed to commit an

attack (Pressman et al., 2016). The fourth dimension corresponds to different commitment and motivations at the individual level and includes eight indicators (Pressman et al., 2016). The fifth dimension includes six protective and risk-mitigating indicators. Furthermore, the VERA-2R includes 31 additional indicators organized in five dimensions: criminal history, personal history, radicalization, personality traits, and psychiatric characteristics (Pressman et al., 2016).

The DRAVY-3 is made up of 63 indicators organized in three dimensions (González-Álvarez et al., 2022). The first dimension refers to violence and includes 20 indicators for assessing the tendency toward violence and ease with which the move to action occurs at a given time. This dimension can be subdivided into two subdimensions: general violence, including 13 indicators about the use of violence, insults, threats, self-harm, and non-compliance with rules; and ideological violence, which encompasses 6 indicators that reflect organizational capacity, connections to violent extremism, and manifestations of the use of violence for ideological purposes. These indicators are directly related to violent behavior and underlying intentionality (Nistal-Burón, 2019). The second dimension is focused on personal radicalization and proselytization and includes 34 indicators that reflect the risk of recruiting or coercive behaviors and the risk of personal radicalization. It includes indicators such as lack of tolerance towards non-believers, the tendency to organize collective religious acts, the need for greater personal status, the tendency to isolate, or feelings of indifference towards victims. These indicators are related to recruitment behavior and/or individual radicalization processes (Nistal-Burón, 2019). The third dimension includes 9 indicators of changes in daily routines that constitute dynamic factors to monitor changes in behavior, the latter being absent in the VERA-2R.

Overlapping of Indicators

After describing and comparing the main characteristics of both tools, we will proceed to compare the overlap of the indicators in order to detect the aspects evaluated by both tools and the differential aspects not evaluated by both tools. We followed the strategy of Hart et al. (2017) to compare the indicators. It was decided to exclude from the comparison the protective factors from VERA-2R as well as the additional indicators, and the daily routine change indicators from DRAVY-3 since they were specific characteristics for each of the tool. In the first step, three experts (two social psychologists and a professional in evaluation within the prison system) individually considered the risk indicators of the VERA-2R, one at a time, and assessed whether they coincided with each of the indicators of the DRAVY-3. These experts rated on a simple dichotomous scale (yes/no) based on the similarity of risk indicators. Krippendorff's alpha coefficient was .48 showing low inter-rater reliability. In the second step, after completing their individual ratings, the two social psychologists broke the blind, discussed their ratings, and made a set of final consensus ratings of the overlap between the risk indicators using the same dichotomous scale. In the third step, the professional in the evaluation within the prison system judged the degree of overlap between the pairs of indicators obtained in the previous phase using a four-point scale (0 = none, 1 = low, 2 = moderate, and 3 = high).

The results are shown in Figure 1, which shows the overlap among the VERA-2R (columns) and the DRAVY-3 (rows) indicators, organized by dimensions. Individual cells are shaded to reflect assessments of the degree of overlap. Overall, there was an overlap of 31 pairs of indicators. There were 8 pairs with low overlap, 10 with moderate overlap, and 13 with high overlap. In

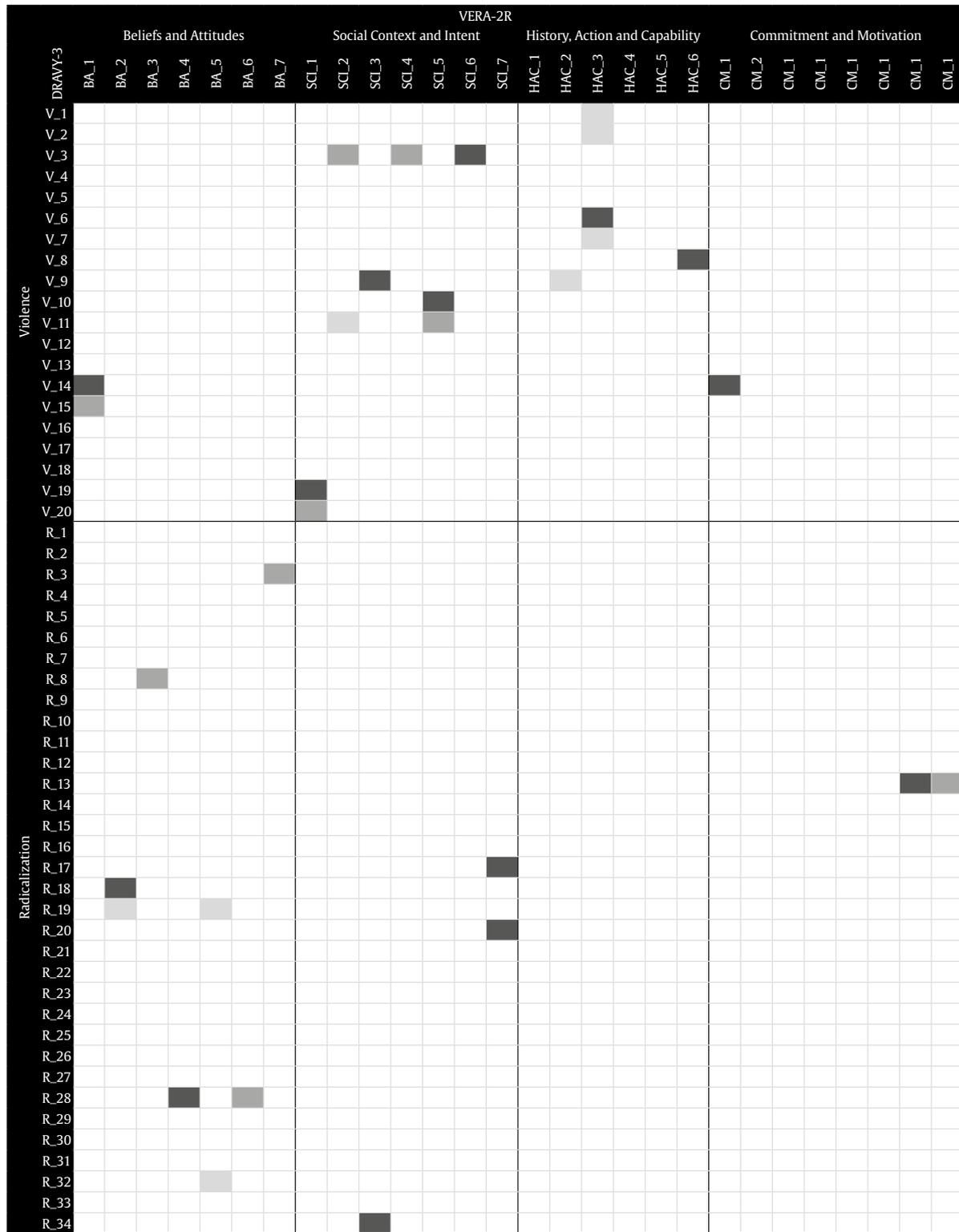


Figure 1. Degree of Overlap of VERA-2R and DRAVY-3 Indicators. Columns are VERA-2R indicators; Rows are DRAVY-3 indicators; Shading of cells reflects ratings of the degree of overlap: white = 0 or none; light gray = 1 or low; dark gray = 2 or moderate, and black = 3 or high.

sum, 19 indicators of the VERA-2R (out of 34) were represented in the DRAVY-3, which means that 55.88% of the indicators overlap with the DRAVY-3; and 23 indicators of DRAVY-3 (out of 54) were represented in VERA-2R, which accounts for 42.59% overlap.

Regarding the overlap on each of the dimensions (see [Figure 2](#)), all the VERA-2R indicators of beliefs and attitudes overlapped with at least one of the DRAVY-3 indicators of the radicalization and proselytization dimension, except for the first indicator

		VERA-2R			
		Beliefs and attitudes	Context and intentions	History, action and capacity	Commitment and motivations
DRAVY-3	Violence	10-15%	30-86%	30-50%	5-13%
	Radicalization and proselytization	18-86%	9-29%	0%	3-25%

Figure 2. Degree of Overlap of VERA-2R and DRAVY-3 Dimensions. The first percentage in each cell refers to the indicators within the corresponding DRAVY-3 dimension that overlap with those in the VERA-2R, while the second percentage refers to the indicators within the corresponding VERA-2R dimension that overlap with those in the DRAVY-3. Shading of cells reflects ratings of the degree of overlap: white = none; light gray = low; dark gray = moderate, and black = high.

which overlapped with indicators of the violence dimension. Similarly, all the VERA-2R social context and intention indicators overlapped with at least one of the DRAVY-3 indicators; while most overlapped with indicators from the violence dimension, the last indicator only overlapped with indicators from the radicalization and proselytization dimension. Regarding the history, action, and capacity indicators, only half (three of six) overlapped with DRAVY-3 indicators; the indicators that overlapped were in the violence dimension, and those that did not overlap referred to early exposure to violence and training. Finally, five of the motivational indicators that refer to motivations related to opportunism, group membership, moral obligation, adventure-seeking, and forced participation were not represented in DRAVY-3; the other three motivational indicators overlapped with two DRAVY-3 indicators from different dimensions.

In general, the VERA-2R indicators are represented in the latest version of the DRAVY-3, with the exception of half of the history, action, and capacity indicators, and most of the commitment and motivation indicators. In contrast, seven indicators from the violence dimension of DRAVY-3 were not represented in the VERA-2R. These indicators focus on threats and other incidents inside prisons, and the intervention of prohibited materials. As for the radicalization and proselytization dimension, twenty-four did not overlap with any of the VERA-2R indicators. Broadly speaking, these indicators were related to behaviors linked to the precepts of jihadist ideology, such as performing purification practices or organizing and attending unauthorized religious acts, behaviors related to proselytizing, such as changing religion or performing acts of proselytizing, physical signs such as a military-style haircut or a cardinal on the left foot, and the latest indicators included after research in prison, such as admiration, identity fusion, or surveillance.

Reliability and Validity of the Tools

After conducting a systematic review of the tools aimed at assessing extremism and radicalization, Scarcella et al. (2016) concluded that the psychometric properties of the instruments were weak and had room for improvement. They also highlighted the lack of validity assessment of the instruments themselves. Since the evaluation made by Scarcella et al. (2016), new instruments have been developed, and more studies have been conducted to validate these tools. However, some of the problems identified are still present, in part due to the limited number of terrorists assessed, the difficulty in accessing to this population, and for obtaining information about the validity of the instruments (Hart et al., 2017). This section summarizes and compares the evidence of

reliability and validity that previous studies have found in the different versions of the VERA and the DRAVY.

Reliability

Internal consistency or reliability measures the degree to which each item of an instrument measures the same characteristic (Higgins & Straub, 2006). In the case of VERA-2R, an independent study by Thijssen et al. (2022) reported Cronbach's alphas between .42 and .88 for the different dimensions. Beardsley and Beech (2013) computed inter-rater reliability or equivalence. Two independent evaluators achieved 85.7% agreement by analyzing five cases of terrorists using data retrieved from the internet. Cohen's kappa was equal to or greater than .76 in all cases. Furthermore, de Bruin et al. (2022; see also Duits & Kempes, 2023) conducted an instrumental study where two assessors rated a sample of 30 convicted of terrorist offences in the Netherlands using the VERA-2R. They found that the inter-rater reliability, represented by the inter-class correlation, ranged from .73 to .85 in the five dimensions. The authors also evaluated the intra-rater reliability using one assessor who rated a sample of 33 cases twice with an interval of six months. The inter-class correlation ranged from .80 to .96 in the different dimensions. Last, Cherney and Belton (2024), evaluating 50 cases from the Profiles of Individual Radicalisation in Australia (PIRA; Belton et al., 2023) database, found an inter-class correlation of .86 across all the dimensions.

In the case of DRAVY, two major public validations have been performed with the second and third versions conducted by prison system personnel, together with academic personnel (González-Álvarez et al., 2022; González-Álvarez et al., 2021). However, no independent external evaluations have been conducted to date. For reliability, the Kuder-Richardson (KR-20) formula reported values of .78 for general violence, .53 for ideological violence, and .89 for radicalization and proselytism (González-Álvarez et al., 2022). To date, there are no open evaluations that evaluate the inter-rater or the intra-rater reliability. Taking together, it seems that the reliability of the VERA and the DRAVY are respectable and similar, although more external evaluations are needed, mainly in the case of DRAVY.

Validity

Validity indicates that what is to be measured is accurate. Different types of validity can be distinguished (Higgins & Straub, 2006). Among these, we highlight content validity, construct validity, discriminant validity, and predictive validity.

Content Validity. Content validity refers to the adequacy of the indicators of an instrument to assess a concept or domain of interest.

It encompasses aspects such as the semantic clarity of the indicators, their consistency, and their appropriateness for the assessed dimensions. The content validity of the VERA-2 was assessed based on expert judgment (Pressman, 2016). More than 60 professional security and intelligence analysts working in the counterterrorism field have evaluated the importance of these indicators. These evaluators reported that most indicators used were either very important or highly important. They were also asked to indicate the factors that were not present among the indicators, according to their experience. The analysts recommended the inclusion of indicators related to cyber behavior, coercion, and the search for significance or status.

Content validity was also tested for the DRAVY-2. During its reformulation, the prison professionals involved in the task tried to ensure that the indicators were representative of what they were intended to measure (violence and radicalization), and that they were written in language understandable to the team of professionals who were going to apply it (González-Álvarez et al., 2021). Moreover, prison professionals evaluated the format of the tool, the comprehension of the indicators, and the relevance of the indicators within its dimensions (González-Álvarez et al., 2021). In general, both tools have adequate content validity, although it would be desirable to know more about the experience and knowledge of the professionals who evaluated them.

Construct Validity. Construct validity refers to whether the tool measures the target construct and whether it is well defined. In the case of the VERA, Beardsley and Beech (2013) conducted an evaluation using the VERA-2 with a sample of five terrorists with different ideologies, providing construct validity. Information was collected using open sources obtained from the internet. The researchers found that attitudinal indicators were present in all terrorists, that these contextual indicators varied, being more prevalent in terrorists who were part of an organization, that historical indicators were inconsistent, which seemed to suggest that early experiences were not particularly relevant in determining who would become a terrorist, and that protective factor indicators identified extremists who were less likely to commit terrorist acts in the future. The authors pointed out that the limitations of this analyses were the sample size and that all the cases chosen were male domestic terrorists who had carried out extreme terrorist crimes with a large number of fatalities. In addition, Thijssen et al. (2022) conducted another study testing the VERA-2R with a sample of jihadi detainees in the Netherlands. They included different groups (i.e., terrorism convicts, released terrorist suspects, violent criminals, criminals who prepared for a terrorist attack, returned foreign fighters, criminals who attempted to travel to Syria or Iraq, and juvenile offenders) and found significant differences between those convicted and those who had been not convicted in the dimensions of beliefs, attitudes, and ideology, social context and intention, and history, action, and capacity. Convicted participants had higher scores in these three dimensions. Moreover, violent extremists were significantly more willing and trained to commit a violent extremist act; violent extremists who planned a terrorist attack were less likely than other violent extremists to be motivated by moral duty; detainees who attempted to travel to Syria or Iraq were significantly more likely to adhere to an ideology that justified violence than other violent extremists; returned foreign fighters had more frequently attended training, although they also exhibited more protective factors; and younger violent extremists scored significantly higher on specific risk factors.

In the case of the DRAVY-3, multivariate statistical techniques have been used to test construct validity (González-Álvarez et al., 2022). The sample consisted of three groups of inmates related to terrorism (groups A, inmates convicted for jihadist terrorism; B, inmates convicted for reasons unrelated to terrorism but suspected of radicalizing others in prisons; and C, inmates vulnerable to recruitment) and two control groups, consisting of inmates under surveillance who showed incipient signs of radicalization and Muslim inmates who did not demonstrate any sign of radicalization. Factor

analyses were performed separately for the violence dimension and the radicalization and proselytization dimension, indicating that the violence indicators were divided into two sub-dimensions (general and ideological violence), and that the radicalization and proselytization indicators constituted a unidimensional scale. However, some indicators were discarded because of their low factor loading. In addition, the three groups related to terrorism had higher scores than the control groups in almost all indicators (only ten indicators did not differentiate between the groups) and, at the same time, differences were found between these three groups related to terrorism. In particular, group A presented the largest number of significant indicators, seven decreasing indicators of general violence, three of extremist ideological violence, and many of the radicalization indicators with an upward trend; group B was characterized by a larger set of indicators, which also showed an upward trend; and group C presented the least significant indicators, although all of them ascended, some of violence, and others of radicalization, especially vulnerability. In brief, both tools appear to have adequate construct validity, demonstrating that they are suitable for assessing the proposed dimensions.

Convergent Validity. Convergent validity refers to the ability to detect a relationship between the concept of interest and a concept that has similar significance (Higgins & Straub, 2006). Using conceptual analysis, Hart et al. (2017) compared Multi-Level Guidelines (MLG; Cook, 2014) and the VERA. In a first comparison using the first version of the VERA, the authors found that the overlap between the risk factors of both tools was limited. In a second comparison, this time using the VERA-2, the authors found that each of the VERA-2 risk factors had substantial overlap with one or more MLG risk factors. However, this overlap was asymmetric. Not all MLG factors were present in the VERA-2, whereas most of the VERA-2 content could be explained by only three MLG risk factors. The authors concluded that many of the risk factors included as specific in the VERA-2 reflect what the MLG considers more general problems (individual and individual-group dimensions).

In the case of the DRAVY-3, to the best of our knowledge the convergent validity has not yet been tested. Therefore, regarding convergent validity, more research needs to be done for both, the VERA-2R and, especially, the DRAVY-3.

Discriminant Validity. Discriminant validity refers to the ability to detect the absence of a significant relationship between the concept of interest and another that has the opposite significance (Higgins & Straub, 2006). In this regard, the indicators of the first version of the VERA were compared with those of other tools aimed at assessing the risk of other forms of violence (Pressman, 2009). Specifically, the Historical/Clinical/Risk Management-20 (HCR-20; Webster et al., 1997) was used to assess general violence risk, and the Structured Assessment of Violence Risk in Youth (SAVRY; Borum et al., 2006) was used to assess general violence in youth. The results showed that only 12% of the HCR-20 indicators (3 out of 25) and 28% of the SAVRY indicators (7 out of 25) were relevant for assessing violent extremism and overlapped with VERA. Consistently, most of the indicators used to assess violent extremism in the VERA were not included in these tools. Similarly, in another study, the use of the VERA-2 was compared with other tools aimed at assessing general violence and psychopathy (Pressman, 2016). Specifically, the HCR-20 (Webster et al., 1997), Violence Risk Scale-Screening Version (VRS-SV; Wong & Gordon, 2006), Psychopathy Check List-Screening Version (PCL:SV; Hart et al., 1995), and Level of Service Inventory-Revised (LSI-R; Andrews & Bonta, 1995), with a group of high-security inmates convicted of violent extremism and/or terrorism-related offenses, and another group of common criminals known for their violence and non-ideological motivation (control group). The results showed that those convicted of terrorism scored higher on the VERA-2 indicators, whereas the control group scored higher on the other tools.

Discriminant validity has not been tested for DRAVY-3. Hence, further comparisons with the DRAVY-3 and other tools aimed at assessing other types of violence would be desirable.

Predictive Validity. In the case of the VERA, predictive ability is not considered an adequate or realistic goal because of the dynamic nature of the radicalization process (Pressman, 2016). In line with this assumption, the authors did not conduct studies that considered the predictive validity of the tool. However, Thijssen et al. (2022) conducted a study in which they tested the predictive validity of the VERA-2R dimensions for predicting terrorism conviction (vs. no terrorism conviction but suspected terrorist acts). The results showed that none of the dimensions were related to terrorism convictions. Furthermore, Cherney and Belton (2024) evaluated 50 cases (37 violent and 13 non-violent individuals) from the PIRA database and found that 57% of violent extremists were judged as high risk (sensitivity) and 69% of non-violent extremists were correctly identified as low risk (specificity). The ROC curve showed an area under the curve (AUC) of .63 indicating poor predictive validity (Horcajo-Gil et al., 2019).

Since its second version, the DRAVY has become a more actuarial tool. Consequently, its predictive power was tested using the second and third version. Focusing on the validation of the last version (González-Álvarez et al., 2022), the predictive capacity of the tool was tested to classify the most dangerous inmates (determined by experts from penitentiary institutions) as opposed to the least dangerous. Dangerousness was used as a proxy for recidivism, as actual recidivism was limited. Twenty-three indicators (three from the violence scale and twenty from the radicalism and proselytism scale) formed a subscale predicting dangerousness. Specifically, the model correctly classified 83.3 percent of the most dangerous inmates (sensitivity) and 66.9 percent of those who were not most dangerous (specificity). The AUC was of .82 indicating an excellent predictive validity (Horcajo-Gil et al., 2019). In this case, more

evidence is needed in the case of VERA-2R to clarify whether it really has predictive capacity and to determine what it predicts, while in the case of DRAVY-3 it is necessary to clarify the utility of the indicators that do not have predictive capacity, in addition to other issues with the dangerousness variable, which are specified in the following section.

Limitations and Areas for Improvement

In this section, we analyze the main limitations and potential areas of improvement identified when comparing the two tools. These limitations are stated in terms of the type of tools they constitute, their generalizability, their starting points in their development, the nature of their indicators and their overlap, and their validity (for a summary see Table 1).

First, the VERA-2R is a structured professional judgment tool and it has some limitations inherent to this type of instruments, such as the amount of time and resources required for its implementation, the need for adequate training, or the subjective part of the final decision (Logan, 2017; Muñoz-Vicente & López-Ossorio, 2016). For instance, probation officers in the Netherlands stopped using the VERA-2R because of capacity problems and lack of information (Sumpter, 2020). The DRAVY-3 is constituted as an actuarial tool, which also has limitations, such as those related to individualization and flexibility, and the sample dependency with which the algorithms are constructed—the so-called “shrinkage” (Hart et al., 2016). In fact, no evidence has been reported regarding the use of the algorithm to predict the risk posed by new users, the weightings of the various indicators are not specified, and it remains unclear how the final score is constructed. As a result, final decisions fall to the evaluators or decision-makers, thereby incorporating the inherent biases of structured professional

Table 1. Overview of the Comparison between VERA and DRAVY

	VERA-2R	DRAVY-3
Development	Adopts the form of a structured professional judgement; developed through literature review and experts feedback; implemented in several countries	Started as a structured professional judgement tool but in the next versions is considered an actuarial tool; developed through literature review, statistical procedures, research, and experts feedback; implemented in Spain
Characteristics		
Goals	Assess risk of violence; predict future violence; and design interventions	Assess risk of violence, ability to proselytize, and vulnerability to radicalization; predict future violence; and design interventions
Target population	Individuals who have engaged in terrorist activities; entire spectrum of violent extremists	Individuals who have engaged in terrorist activities, who proselytize, and who are vulnerable to recruitment; focus on jihadist extremism
Structure		
Dimensions	Five dimensions: beliefs, attitudes, and ideology; social context and intention; history, action, and capacity; commitment and motivation; and protective and risk-mitigating indicators; plus additional indicators	Two dimensions: violence (general violence and ideological violence), and personal radicalization and proselytization; plus changes in daily routines indicators
Indicators	34 indicators group in 5 dimensions plus 31 additional indicators group in 5 dimensions	54 indicators group in two dimensions plus 9 changes in daily routines indicators
Overlap	19 indicators of VERA-2R are represented in DRAVY-3 (55.88%)	23 indicators of DRAVY-3 are represented in VERA-2R (42.59%)
Reliability and validity		
Reliability	Cronbach's alphas between 0.42 and 0.88; inter-rater ICCs between 0.73 and 0.85; intra-rater ICCs between 0.80 and 0.96	Kuder-Richardson (KR-20) values between 0.53 and 0.89
Content validity	Expert judgement with professional security and intelligence analysts working in the counterterrorism	Expert judgement with prison professionals
Construct validity	Assess terrorists with different ideologies and distinguish between convicted and non-convicted terrorist	Discriminates between inmates related to terrorism and two control groups of inmates non-radicalized
Convergent validity	Converges with the individual and individual-group dimensions of the MLG	Untested
Discriminant validity	Discriminates in comparison with HCR-20, SAVRY, VRS-SV, PCL:SV, and LSI-R	Untested
Predictive validity	Classified 57% of the violent extremists and 69% of the non-violent extremists	Classified 83.3% of the most dangerous inmates and 66.9% of the non-dangerous

Note. ICC = inter-class correlation.

judgement. An area of improvement relies on an option that seems to be followed by the DRAVY-3, although this is not made explicit. This would consist of using the tool in an actuarial manner to obtain a single score, but relying on expert judgment alongside that score when making a final decision.

Second, the VERA-2R is a standard tool aimed at assessing all cases of radicalization, regardless of ideology, while DRAVY-3 focuses explicitly on the radicalization of jihadist ideology. As a standard tool for all ideologies and contexts, the VERA-2R does not seem to work in some contexts because of their particularities. In this regard, the VERA-2 was attempted to be implemented in Indonesia, but over time, the project began to disintegrate. The instrument, supposedly adapted, was not considered suitable for the context, given its complexity (Sumpter, 2020). Therefore, the DRAVY-3 has the advantage of being created for a specific context with its legal particularities. However, it remains to be seen whether it could be exported to other contexts. First, many of the indicators are specific to jihadist terrorism, making their use for assessing cases involving other ideologies related to terrorism unfeasible. Second, the tool is designed to assess the risk posed by individuals convicted of terrorism-related offenses, potential proselytizers, and vulnerable individuals, based on the needs of the Spanish prison system. Consequently, its use in other countries would only be feasible in prison systems with similar management structures—specifically, those that manage terrorist inmates using a segregation-dispersion model in which inmates are concentrated in a few facilities but are not mixed with the general prison population due to the level of isolation (Lobato & García-Coll, 2022)—and in which both proselytism and vulnerable individuals are considered relevant targets for intervention (Marrero & Berdún-Carrión, 2021). Moreover, the DRAVY-3 not only seeks to assess the risk of individuals involved in terrorism, such as the VERA-2R, but also to assess cases of proselytizing and individuals vulnerable to radicalization. This breadth of vision generates a problem related to the definition of the hazard (Herrington & Roberts, 2012; Roberts & Horgan, 2008). It is not clear which hazard is to be assessed, whether it is recidivism (using dangerousness as a proxy), the beginning of a radicalization process, or vulnerability to recruitment. Moreover, these dimensions do not conform to these hazard classifications.

Third, although it does not conform to a specific theory, the VERA-2R's dimensions are based on a more theoretical conception (Monahan, 2012). Herzog-Evans (2018) emphasized that the VERA's dimensions resemble engagement, intent, and capability, which fits the theory of reasoned action explaining the relationship between attitudes and behaviors (Ajzen & Fishbein, 2005). Therefore, we can say that its dimensions are theory-driven. In contrast, the DRAVY-3's dimensions appear to arise from the legal and prison system that distinguishes between prisoners of terrorist offenses, proselytizers, and individuals vulnerable to radicalization. Although some indicators are supported by research and align with certain theories (e.g., the identity fusion theory; Gómez, Atran, et al., 2022), the dimensions do not conform to any single theoretical framework a priori. Instead, they appear to combine indicators from different theories. Therefore, we can say that the dimensions are legally-driven. This distinction ties with what was previously mentioned regarding the difficulty of using a standard or theory-driven tool in a specific context, and the difficulty of exporting a contextual or legally-driven tool.

Fourth, a shared limitation lies in the combination of dynamic and static risk indicators. Although static indicators can be useful in assessing the vulnerability of individuals to radicalization, they will not undergo variations, so their usefulness in detecting changes or the effect of an intervention will be limited (Douglas & Skeem, 2005). Both tools include both types of risk indicators. In the case of the VERA-2R, static indicators are grouped mainly in the dimension

of history, action, and capacity. In the case of the DRAVY-3, they are found in the dimension of violence, although they are mixed with dynamic indicators and not restricted to a single subdimension. While static indicators are easier to assess and require fewer inferences, dynamic indicators are more useful for prediction. However, they typically require ongoing assessments over time and demand greater inferential judgment from evaluators, thereby increasing the risk of error (Muñoz-Vicente & López-Ossorio, 2016). A useful approach for both tools is to explicitly distinguish between these static and dynamic indicators, establish different dimensions for them, and evaluate their usefulness in terms of classification and prediction. In this regard, the DRAVY-3 incorporates daily routine indicators that could be added to the VERA-2R, although no evidence is provided on their validity, and their relationship with the other dimensions is not clear.

Fifth, the analyses revealed that, broadly, half of the indicators of both tools overlapped. A priori, such overlap may appear desirable as it provides convergent validity and could suggest that these indicators are essential for assessing risk regardless of the context. However, the overlap could also be seen as problematic given that the DRAVY was developed with reference to the VERA-2, among other tools, and adopted some of its indicators (Loínaz, 2019). In this regard, the predictive validity analysis of the DRAVY-3 showed that many of the overlapping indicators with the VERA-2R were of limited utility, particularly those located within the tendency toward violence dimension (González-Álvarez et al., 2022). Therefore, while referencing other tools during the development of an assessment tool can be useful, it is advisable—over time and based on empirical evidence—to revisit the initial assumptions made about certain indicators and eliminating those that are not useful. From an opposing perspective, considering the indicators that do not overlap, they offer the possibility of including new indicators in both tools to improve them. The VERA-2R would benefit from the inclusion of indicators related to specific behaviors within the prison. Likewise, it would be advisable to integrate some dimensions with specific indicators of behavior associated with the ideology related to terrorism; for this, some of the DRAVY-3 indicators would be useful in the case of jihadism. An update with indicators from the latest research could also improve this tool (e.g., Gómez, Atran, et al., 2022). In turn, the DRAVY-3 could include more indicators regarding past use of violence and weapons training, although this would have an impact on the increase in static risk factors. Further-more, the greatest contribution would be in the motivational aspect, that is, the inclusion of new indicators to assess different motivations related to radicalization. Finally, a major shortcoming is the lack of protection indicators. Although the VERA-2R includes some, it is desirable to create better evidence-based indicators that are not just reverse indicators (e.g., Wolfowicz et al., 2021). This shortcoming may lead to an overestimation of risk which, combined with the predominance of static indicators, could hinder the detection of progress in users.

Sixth, although several investigations have evaluated the psychometric properties of both tools, there is room for improvement (Lloyd, 2019). The indicators need to be refined to improve internal reliability. Regarding construct and convergent validity, one of the main criticisms focuses on the fact that neither the credentials nor the number of experts involved in such a process is detailed (Herzog-Evans, 2018). Therefore, a more transparent validation procedure is required, at least in the case of the VERA-2R. Similarly, although both tools indicate that they can and have been applied to women, they are not adapted in either case (Loínaz, 2016). Neither tool has identified the specific underlying processes that explain female radicalization (Gómez, Chiclana, et al., 2022), so they probably lack content validity in this respect. In addition, in the case of the DRAVY-3, dimensionality has been analyzed with the dimensions separately, so it would be interesting to analyze

them as a whole. Moreover, although our investigation provides some evidence of convergent validity, more studies using other tools are needed. The same is true for divergent validity, which, in the case of the DRAVY-3, has no evidence. In terms of predictive validity, the VERA-2R needs new studies to explore this type of validity, the DRAVY-3 requires new studies using other samples to test whether the indicators intended for this purpose still maintain sensitivity and specificity (van der Heide et al., 2019), and both need to demonstrate that changes in the indicators are associated with changes in risk (Wilson et al., 2013). Additionally, the case of DRAVY presents other issues related to the dependent variable of dangerousness. On the one hand, dangerousness and recidivism are not the same, so it could be argued that the predictive validity of the tool concerning recidivism has not been tested, leaving it unclear whether this is an intended objective of the tool (Teijón-Alcalá, 2023). On the other hand, the fact that the same prison experts who created the dangerousness classification were also the ones using the tool to assess the same individuals increases the likelihood of good predictive validity, suggesting that a significant bias may underlie these results (Teijón-Alcalá, 2023). These issues with predictive validity are critical, as inaccurate risk classification may lead to false positives, resulting in the stigmatisation of inmates, and false negatives, which may underestimate the actual risk (see, for instance, the case of Usman Khan; Weeks, 2021).

Last, unlike other tools created from criminology (e.g., Andrews & Bonta, 2016), the VERA-2R and the DRAVY-3 are not linked to a treatment plan (Herzog-Evans, 2018; Logan & Sellers, 2021). In other words, there are no specific measures or programs recommended depending on the level of risk, which could affect penitentiary decisions regarding these inmates, as such decisions may be influenced by biases with limited grounding in empirical evidence. This would be a costly but critical step that would provide a quantum leap for these tools.

Discussion

The VERA-2R and the DRAVY-3 are tools for assessing violent extremism in the prison context. In the present study, we conducted a systematic evaluation comparing their development, characteristics, structure, overlap of their indicators, reliability, and validity. This comparison highlighted that both tools, despite being intended for use in the prison context, present severe differences. On the one hand, the VERA-2R is a structured professional judgment designed to assess the risk of violent acts by individuals involved in terrorist activities regardless of their ideology. On the other hand, the DRAVY-3 is an actuarial tool aimed at assessing and predicting the risk of violence, proselytizing, and vulnerability to being recruited, and therefore its target population are individuals who have engaged in terrorist activities, who proselytize, and who are vulnerable to recruitment inside the jihadist spectrum.

Regarding their structure and indicators, the VERA-2R maintains theory-driven dimensions and its indicators differ from the DRAVY-3 in that they include aspects of early exposure to violence and training, and different motivations. For its part, the DRAVY-3's dimensions are legally-driven (in compliance with prison regulations) and include indicators not present in the VERA-2R related to behavior inside the prison, acts of proselytization, and behaviors related to jihadist ideology. With respect to the evidence of reliability and validity, both present acceptable reliability and content and construct validity. However, there is no evidence of convergent and divergent validity for the DRAVY-3, while there is in the VERA-2R; and the DRAVY-3 does not present predictive validity (with some concerns) unlike the VERA-2R.

In addition, we outlined some of the main limitations of both tools. Broadly, the VERA-2R presents limitations when using it in a

specific context when trying to evaluate different ideologies, while the DRAVY-3 is a contextual tool that would present problems if it were used in another context. Likewise, both tools combine static and dynamic indicators, so they are limited in detecting changes and effects of possible interventions.

Among the contributions of this investigation, we highlight the keys provided that can be useful when updating these tools or developing new ones. First, it is necessary to establish clear objectives and, in the case of multiple objectives, to make it clear which dimensions or indicators evaluate each of the objectives and in which specific population. Second, in order to improve the assessment, it is necessary to adapt the tools to the context. The tools should integrate indicators of specific behaviors in the context, in this case the prison, and related to the ideology. Third, the differences between static and dynamic indicators should be made explicit and determine how each type of indicator influences the final evaluation or in which cases each is more useful. Fourth, the gender perspective must be taken into account. Both tools indicate that they are valid for use with women, but they do not include specific indicators or validity to prove their usefulness in this population. Fifth, the tools should be subjected to more validity tests, preferably with different populations, including control groups, and being conducted by personnel not involved in their development. Also, given that we are talking about a dynamic and contextual phenomenon, the tools should be updated based on these investigations. Sixth, one of the major limitations is the lack of management guidelines or treatments associated with risk. These tools should aim to associate possible interventions or ways of managing risk associated with the results of their application.

Conclusion

We conclude by emphasizing the value of this comparison. This research has shown the main limitations of the VERA-2R and the DRAVY-3, so it can assist first-line practitioners in selecting a tool by considering its benefits and limitations. The comparison brings out some of the strengths of the tools, as well as their objectives and target populations; consequently, this can provide ideas for improving or contextualizing tools with similar objectives. Finally, also when it comes to building new versions of these tools or taking them to other contexts, the comparison is relevant because it provides some ideas on the directions to take and the new research that should be performed to continue providing validity. In brief, it is expected that this review will be useful for first-line practitioners in the penitentiary context to become familiar with and better understand these instruments, improve their evaluations, and even adapt or develop new instruments to assess the risk of violent extremism.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Lawrence Erlbaum Associates.
- Andrews, D. A., & Bonta, J. (1995). *LS-R: The Level of Service Inventory-Revised*. Multi-Health Systems.
- Andrews, D. A., & Bonta, J. (2016). *The psychology of criminal conduct*. Routledge.
- 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(1), 4-15. <https://doi.org/10.1108/17596591311290713>
- Belton, E., Cherney, A., & Zahnow, R. (2023). Profiles of individual radicalisation in Australia (PIRA): Introducing an Australian open-source ex-

- tremist database. *Perspectives on Terrorism*, 17(1), 18-35. <https://doi.org/10.19165/TFTQ8338>
- Borum, R., Bartel, P., & Forth, A. (2006). *SAVRY: Professional manual for structured assessment of violence risk in youth*. Psychological Assessment Resources Inc.
- Cherney, A., & Belton, E. (2024). *Testing the reliability and validity of the VERA-2R on individuals who have radicalised in Australia*. Australian Institute of Criminology. <https://doi.org/10.52922/crg77345>
- Cook, A. N. (2014). *Risk assessment and management of group-based violence* [Tesis doctoral, Simon Fraser University]. <https://doi.org/10.1002/9781119035848.ch5>
- De Bruin, A., Duits, N., Kempes, M., & Prinsen, M. (2022). *Interrater and intrarater reliability of the violent extremism risk assessment tool*. Netherlands Institute of Psychiatry and Psychology (NIPF). <https://www.nipf.nl/documenten/rapporten/2023/03/02/interrater-and-intrarater-reliability-of-the-violent-extremism-risk-assessment-tool>
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11(3), 347-383. <https://doi.org/10.1037/1076-8971.11.3.347>
- Duits, N., & Kempes, M. (2023). Interrater and intra-rater reliability of the VERA-2R tool. *Frontiers in Psychiatry*, 14, Article 1236295. <https://doi.org/10.3389/fpsy.2023.1236295>
- Fernández, C. (2019). Valoración del riesgo del radicalismo violento en el medio penitenciario [Risk assessment of violent radicalism in the prison environment]. In R. Bermejo & I. Bazaga (Eds.), *Radicalización violenta en España: detección, gestión y respuesta* (pp. 253-259). Tirant Lo Blanch.
- Fernandez, C., & de Lasala, F. (2021). *Risk assessment in prison*. Radicalisation Awareness Network (RAN). https://ec.europa.eu/home-affairs/document/download/54d9f0f6-2e79-467e-900a-c17802c44f51_en
- Gómez, Á., Atran, S., Chinchilla, J., Vázquez, A., López-Rodríguez, L., Paredes, B., Martínez, M., Blanco, L., Alba, B., Bautista, H., Fernández, S., Pozuelo-Rubio, F., González-Álvarez, J. L., Chiclana, S., Valladares-Narganes, H., Alonso, M., Ruiz-Alvarado, A., López-Novo, J. L., & Davis, R. (2022). Willingness to sacrifice among convicted Islamist terrorists versus violent gang members and other criminals. *Scientific Reports*, 12(1), Article 2596. <https://doi.org/10.1038/s41598-022-06590-0>
- Gómez, Á., Bélanger, J. J., Chinchilla, J., Vázquez, A., Schumpe, B. M., Nisa, C. F., & Chiclana, S. (2021). Admiration for Islamist groups encourages self-sacrifice through identity fusion. *Humanities and Social Sciences Communications*, 8, Article 54. <https://doi.org/10.1057/s41599-021-00734-9>
- Gómez, Á., Chiclana, S., Chinchilla, J., Blanco, L., Alba, B., Bautista, H., & Pozuelo-Rubio, F. (2022). The mirage of the jihad. Disenchantment as the pathway to disengagement of female jihadists. A case study about radicalization in Spanish prisons [El espejismo de la yihad. El desencanto como camino para la desvinculación de las mujeres yihadistas. Un estudio de caso sobre la radicalización en las cárceles españolas]. *International Journal of Social Psychology*, 37(3), 586-617. <https://doi.org/10.1080/02134748.2022.2096254>
- González-Álvarez, J. L., Santos-Hermoso, J., Gómez, Á., López-Novo, J. L., Buquerín-Pascual, S., Pozuelo-Rubio, F., Fernández-Gómez, C., & Chiclana, S. (2022). A theoretical, empirical, and methodologically based instrument to assess the risk of violent jihadist radicalization in prisons: The DRAVY-3. *Terrorism and Political Violence*, 36(3), 283-299. <https://doi.org/10.1080/09546553.2022.2145194>
- González-Álvarez, J. L., Santos-Hermoso, J., López-Novo, J. L., Chiclana, S., Pozuelo-Rubio, F., Fernández-Gómez, C., & Lanza-Comenau, C. (2021). *Construcción y validación de la herramienta de clasificación y de valoración del riesgo de radicalismo violento en el ámbito penitenciario* [Construction and validation of a tool for the classification and assessment of the risk of violent radicalism in the prison environment]. Secretaría General de Instituciones Penitenciarias. <https://cpage.mpr.gob.es/producto/construccion-y-validacion-de-una-herramienta-de-clasificacion-y-de-valoracion-del-riesgo-de-radicalismo-violento-en-el-ambito-penitenciario/>
- Hart, S. D., Cook, A. N., Pressman, E., Strang, S., & Lim, Y. L. (2017). *A concurrent evaluation of threat assessment tools for the individual assessment of terrorism*. Canadian Network for Research on Terrorism, Security, and Society (TSAS). <https://www.tsas.ca/publications/a-concurrent-evaluation-of-threat-assessment-tools-for-the-individual-assessment-of-terrorism/>
- Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *Manual for the screening-version of the Hare Psychopathy Checklist-Revised (PCL:SV)*. 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, & D. Boer (Eds.), *The Wiley handbook on the theories, assessment, treatment of sexual offending: Volume II. Assessment* (pp. 643-666). Wiley.
- Herrington, V., & Roberts, K. (2012). Risk assessment in counter terrorism. In U. Kumar & M. K. Mandal (Eds.), *Countering terrorism: Psychosocial strategies* (pp. 282-305). SAGE Publications.
- Herzog-Evans, M. (2018). A comparison of two structured professional judgment tools for violent extremism and their relevance in the French context. *European Journal of Probation*, 10(1), 3-27. <https://doi.org/10.1177/2066220317749140>
- Higgins, P. A., & Straub, A. J. (2006). Understanding the error of our ways: Mapping the concepts of validity and reliability. *Nursing Outlook*, 54(1), 23-29. <https://doi.org/10.1016/j.outlook.2004.12.004>
- Horcajo-Gil, P. J., Dujó-López, V., Andreu-Rodríguez, J. M., & Marín-Rullán, M. (2019). Valoración y gestión del riesgo de reincidencia delictiva en menores infractores: Una revisión de instrumentos [Assessment and management of the risk of criminal recidivism in juvenile offenders: A review of instruments]. *Anuario de Psicología Jurídica*, 29(1), 41-53. <https://doi.org/10.5093/apj2018a15>
- Lloyd, M. (2019). *Extremist risk assessment: A directory*. Centre for Research and Evidence on Security Threats (CREST). <https://crestresearch.ac.uk/resources/extremism-risk-assessment-directory/>
- Lloyd, M., & Dean, C. (2015). The development of structured guidelines for assessing risk in extremist offenders. *Journal of Threat Assessment and Management*, 2(1), 40-52. <https://doi.org/10.1037/tam0000035>
- Lobato, R. M., & García-Coll, J. (2022). *La encrucijada entre la radicalización y la desradicalización: teorías, herramientas y aspectos aplicados* [The crossroad between radicalization and deradicalization: Theories, tools, and applied aspects]. Los Libros de la Catarata.
- Logan, C. (2017). Reporting structured professional judgement. In S. Brown, E. Bowen, & D. Prescott (Eds.), *The forensic psychologist's report writing guide* (pp. 8293). Routledge.
- Logan, C., & Sellers, R. (2021). Risk assessment and management in violent extremism: A primer for mental health practitioners. *The Journal of Forensic Psychiatry & Psychology*, 32(3), 355-377. <https://doi.org/10.1080/14789949.2020.1859591>
- Loinaz, I. (2016). Cuando "el" delincuente es "ella": intervención con mujeres violentas [When "the" offender is "she": Intervention with violent women]. *Anuario de Psicología Jurídica*, 26(1), 41-50. <https://doi.org/10.1016/j.apj.2016.04.006>
- Loinaz, I. (2019). *Validación del protocolo para la valoración del riesgo de radicalismo violento* [Validation of the protocol for the assessment of the risk of violent radicalism] (unpublished manuscript). Secretaría General de Instituciones Penitenciarias.
- Marrero, I., & Berdún-Carrión, S. (2021). Jihadist prisoners in Spain and the application of the high security prison regime. *European Journal of Criminology*, 20(2), 586-603. <https://doi.org/10.1177/14773708211018946>
- Meloy, J. R., & Gill, P. (2016). The lone-actor terrorist and the TRAP-18. *Journal of Threat Assessment and Management*, 3(1), 37-52. <https://doi.org/10.1037/tam0000061>
- Ministerio del Interior. (2019). *Estrategia nacional contra el terrorismo* [National Strategy against Terrorism]. Gobierno de España. <https://www.dsn.gob.es/es/file/2883/download?token=goAB5q9t>
- Monahan, J. (2012). The individual risk assessment of terrorism. *Psychology, Public Policy, and Law*, 18(2), 167-205. <https://doi.org/10.1037/a0025792>
- Muñoz-Vicente, J. M., & López-Ossorio, J. J. (2016). Valoración psicológica del riesgo de violencia: alcance y limitaciones para su uso en el contexto forense [Psychological assessment of the risk of violence: Scope and limitations for its use in the forensic setting]. *Anuario de Psicología Jurídica*, 26(1), 130-140. <https://doi.org/10.1016/j.apj.2016.04.005>
- Nistal Burón, J. (2019). Radicalización en prisiones [Radicalization in prisons]. In R. Bermejo & I. Bazaga (Eds.), *Radicalización violenta en España: detección, gestión y respuesta* (pp. 121-131). Tirant lo Blanch.
- Pressman, E. (2009). *Risk assessment decisions for violent political extremism*. Public Safety Canada. <https://osha.europa.eu/en/topics/riskassessment>
- Pressman, E. (2016). The complex dynamic causality of violent extremism: Applications of the VERA-2 risk assessment method to CVE initiatives. In A. J. Masys (Ed.), *Disaster forensics: Understanding root cause and complex causality* (pp. 249-269). Springer International Publishing. https://doi.org/10.1007/978-3-319-41849-0_10
- Pressman, E., & Davis, N. (2022). Violent extremism risk assessment and screening analysis: Applicability, challenges, and new developments. *Journal of Policing, Intelligence and Counter Terrorism*, 17(3), 301-313. <https://doi.org/10.1080/18335330.2022.2117567>
- Pressman, E., Duits, N., Rinne, T., & Flockton, J. (2016). *Violence Extremism Risk Assessment version 2 revised: A structured professional judgment approach* (unpublished manuscript).
- Pressman, E., & Flockton, J. (2010) VERA-2. *Violent Extremism Risk Assessment, version 2 manual* (unpublished manuscript).
- Pressman, E., & Flockton, J. (2012). Calibrating risk for violent political extremists and terrorists: The VERA 2 structured assessment. *The British Journal of Forensic Practice*, 14(4), 237-251. <https://doi.org/10.1108/14636641211283057>
- Pressman, E., & Flockton, J. (2014). Violent extremist risk assessment: Issues and applications of the VERA-2 in a high-security correctional setting. In A. Silke (Ed.), *Prisons, Terrorism and Extremism* (pp. 122-143). Routledge. <https://doi.org/10.4324/9780203584323-9>
- RAN. (2021). *Violent Extremism Risk Assessment, version 2-revised (VERA-2R)* Pressman, Rinne, Duits, Flockton (2016). Radicalisation Awareness Network (RAN). <https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran/collection-inspiring-practices/>

- [ran-practices/violent-extremism-risk-assessment-version-2-revised-vera-2r-pressman-rinne-duits-flockton-2016_en](#)
- Roberts, K., & Horgan, J. (2008). Risk assessment and the terrorism. *Perspectives on Terrorism*, 2(6), 3-10.
- Santos-Hermoso, J., González-Álvarez, J. L., Macía Vázquez, A., Pozuelo Rubio, F., & Chiclana, S. (2021). Jihadist extremism in Spanish prisons. Characterisation of the inmates linked to jihadist radicalisation and analysis of their behaviour in prison. *Terrorism and Political Violence*, 35(4), 733-753. <https://doi.org/10.1080/09546553.2021.1972978>
- Scarcella, A., Page, R., & Furtado, V. (2016). Terrorism, radicalisation, extremism, authoritarianism and fundamentalism: A systematic review of the quality and psychometric properties of assessments. *Plos One*, 11(12), Article e0166947. <https://doi.org/10.1371/journal.pone.0166947>
- Secretaría General de Instituciones Penitenciarias. (2018). *Instrumento de evaluación del riesgo de radicalismo violento* [Violent radicalism risk assessment instrument].
- Sumpter, C. (2020). Realising violent extremist risk assessments in Indonesia: Simplify and collaborate. *Journal for Deradicalization*, 22, 97-121.
- Teijón-Alcalá, M. (2023). La evaluación del riesgo de radicalismo violento en prisión: problemas de científicidad y teorización en el instrumento DRaVY [Assessing the risk of violent radicalism in prison: Problems of scientificity and theorization in the DRaVY instrument]. In A. Gil-Gil & E. Maculan (Eds.), *La ejecución de las penas por delitos de terrorismo* (pp. 367-402). Dykinson.
- Thijssen, G., Masthoff, E., Sijtsema, J. J., & Bogaerts, S. (2022). Understanding violent extremism: Risk and protective factors in a jihadi male detainee population in the Netherlands. *European Journal of Criminology*, 20(3), 973-995. <https://doi.org/10.1177/14773708221132887>
- Van der Heide, L., van der Zwan, M., & van Leyenhorst, M. (2019). *The practitioner's guide to the galaxy - a comparison of risk assessment tools for violent extremism*. International Centre for Counter-Terrorism - The Hague (ICCT). <https://doi.org/10.19165/2019.1.07>
- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. (1997). *HCR-20 assessing risk for violence: Version 2*. Mental Health Law & Policy Institute, Simon Fraser University.
- Weeks, D. (2021). Lessons learned from U.K. efforts to deradicalize terror offenders. *CTC Sentinel*, 14(3), 33-39.
- Wilson, C. M., Desmarais, S. L., Nicholls, T. L., Hart, S. D., & Brink, J. (2013). Predictive validity of dynamic factors: Assessing violence risk in forensic psychiatric inpatients. *Law and Human Behavior*, 37(6), 377-388. <https://doi.org/10.1037/lhb0000025>
- Wolfowicz, M., Litmanovitz, Y., Weisburd, D., & Hasisi, B. (2021). Cognitive and behavioral radicalization: A systematic review of the putative risk and protective factors. *Campbell Systematic Reviews*, 17(3), Article e1174. <https://doi.org/10.1002/cl2.1174>
- Wong, S. C. P., & Gordon, A. (2006). The validity and reliability of the Violence Risk Scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law*, 12(3), 279-309. <https://doi.org/10.1037/1076-8971.12.3.279>