Article

Patterns

Uncovering social-contextual and individual mental health factors associated with violence via computational inference

Highlights

- Classification of violence using sociocontextual and individual mental health factors
- Study of appetitive, consequentialist, and impulsive DoVs
- Confessed DoVs in a large sample of Colombian ex-members of illegal armed groups
- Neural networks and machine learning identified the top factors associated with violence

Authors

Hernando Santamaría-García, Sandra Baez, Diego Mauricio Aponte-Canencio, ..., Jonathan Levy, Jean Decety, Agustín Ibáñez

Correspondence

agustin.ibanez@gbhi.org

In Brief

The study of human violence calls for methodological innovations. Here, we examined historical records for a large sample of ex-members of illegal armed groups in Colombia (N = 26,349) and combined deep learning and machine learning methods to identify the most relevant factors (>160) associated with different confessed domains of violence (DoVs). Results showed that accurate DoV classification required a combination of both social-contextual and individual mental health factors. The results support the development of computational approaches for multidimensional assessments of confessed DoV.

Santamaría-García et al., 2021, Patterns *2*, 100176
February 12, 2021
https://doi.org/10.1016/j.patter.2020.100176

CellPress

Patterns

CellPress
OPEN ACCESS

Article

Uncovering social-contextual and individual mental health factors associated with violence via computational inference

Hernando Santamaría-García,[1,2,18] Sandra Baez,[3,18] Diego Mauricio Aponte-Canencio,[4,5,17] Guido Orlando Pasciarello,[6,7] Patricio Andrés Donnelly-Kehoe,[6,7] Gabriel Maggiotti,[8] Diana Matallana,[1] Eugenia Hesse,[9,10] Alejandra Neely,[11] José Gabriel Zapata,[4] Winston Chiong,[12] Jonathan Levy,[13,17] Jean Decety,[14] and Agustín Ibáñez[9,10,11,15,16,18,19,*]

[1]Doctorado de Neurociencias, Departamentos de Psiquiatría y Fisiología, Pontificia Universidad Javeriana, Bogotá, Colombia
[2]Centro de Memoria y Cognición Intellectus, Hospital Universitario San Ignacio, Bogotá, Colombia
[3]Universidad de los Andes, Bogotá, Colombia
[4]Universidad Externado de Colombia, Bogotá, Colombia
[5]Agencia para la Reincorporación y la Normalización (ARN), Bogotá, Colombia
[6]Multimedia Signal Processing Group–Neuroimage Division, French-Argentine International Center for Information and Systems Sciences (CIFASIS)–National Scientific and Technical Research Council (CONICET), Rosario, Argentina
[7]Laboratory of Neuroimaging and Neuroscience (LANEN), INECO Foundation Rosario, Rosario, Argentina
[8]Asapp, Buenos Aires, Argentina
[9]Cognitive Neuroscience Center (CNC), Universidad de San Andrés, Buenos Aires, Argentina
[10]National Scientific and Technical Research Council (CONICET), Argentina
[11]Latin American Institute for Brain Health (BrainLat), Center for Social and Cognitive Neuroscience (CSCN), Universidad Adolfo Ibáñez, Santiago de Chile, Chile
[12]UCSF Weill Institute for Neurosciences, San Francisco, CA, USA
[13]Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya (IDC), Israel
[14]University of Chicago, Chicago, IL, USA
[15]Universidad Autónoma del Caribe, Barranquilla, Colombia
[16]Global Brain Health Institute (GBHI), University of California San Francisco (UCSF), San Francisco, CA, USA
[17]Department of Neuroscience and Biomedical Engineering, Aalto University, Finland
[18]These authors contributed equally
[19]Lead Contact
*Correspondence: agustin.ibanez@gbhi.org
https://doi.org/10.1016/j.patter.2020.100176

THE BIGGER PICTURE We assessed a comprehensive group of social-contextual and individual mental health factors to classify confessed acts of violence committed in the past among a large sample of Colombian ex-members of illegal armed groups (N = 26,349). We used a novel data-driven approach to classify subjects based on four confessed domains of violence (DoVs) and including two groups, (1) ex-members who admitted violent acts and (2) ex-members who denied violence in each DoV, matched by sex, age, and education stage. We found that accurate classification required both social-contextual and individual mental health factors, although the social-contextual factors were the most relevant. Our study provides population-based evidence on the factors associated with historical assessments of violence and describes a powerful analytical approach. This study opens up a new agenda for developing computational approaches for situated, multidimensional, and evidence-based assessments of violence.

Mainstream: Data science output is well understood and (nearly) universally adopted

SUMMARY

The identification of human violence determinants has sparked multiple questions from different academic fields. Innovative methodological assessments of the weight and interaction of multiple determinants are still required. Here, we examine multiple features potentially associated with confessed acts of violence in ex-members of illegal armed groups in Colombia (N = 26,349) through deep learning and feature-derived

machine learning. We assessed 162 social-contextual and individual mental health potential predictors of historical data regarding consequentialist, appetitive, retaliative, and reactive domains of violence. Deep learning yields high accuracy using the full set of determinants. Progressive feature elimination revealed that contextual factors were more important than individual factors. Combined social network adversities, membership identification, and normalization of violence were among the more accurate social-contextual factors. To a lesser extent the best individual factors were personality traits (borderline, paranoid, and antisocial) and psychiatric symptoms. The results provide a population-based computational classification regarding historical assessments of violence in vulnerable populations.

INTRODUCTION

Violence is a ubiquitous human phenomenon[1–3] that has a dramatic impact on the global economy, health, and the stability of countries.[4] Research has identified several factors related to the use of violence in civil war settings,[5] including social-contextual (social,[6] political,[7] and cultural)[8] and individual mental health factors (psychological determinants,[9] physical health,[2] personality,[10] and protective factors such as the ability to cope with stress and well-being).[9] Although previous studies have assessed predictors of violence in different samples,[11,12] the research lacks a combined assessment of contextual and individual measurements' interactions via novel machine learning methods to assess historical data related to violence in civil war settings.

The Colombian conflict has been pervasive during the past 50 years, with devastating societal and environmental consequences.[13] This conflict has resulted in 7,265,072 victims of forced displacement, 363,374 deaths, 167,809 victims of enforced disappearance, and 11,140 victims of anti-personnel mines.[14] Here, we analyzed data from a national study assessing the potential predictors of violence (PPVs) associated with different domains of violence (DoVs) in a civil war setting. Individuals were required to answer whether they committed violence moved by different motives, referred to in the literature as consequentialist,[15] appetitive,[16] retaliatory,[17] or impulsive[16] violence (i.e., DoV, see below for the theoretical background). In addition, individuals responded as to whether they had committed violence following all types of DoV (hereafter "global violence"). Participants were also required to answer questions assessing a large group of PPVs (these were based on previous reports and theoretical models, see below). The PPV assessment consisted of 162 questions exploring social-contextual and individual mental health factors that could be associated with violence. In each DoV classification, we assessed two different groups: ex-members who declared violent acts and participants who denied violence in each DoV, matched by sex, age, and educational level.

We analyzed data from a large sample of ex-members of Colombian illegal armed groups (N = 26,349, representing more than 90% of all individuals who were demobilized in the Colombian conflict from 2003 to 2012; see Table 1). These individuals participated in collective or individual demobilization processes and entered programs of transitional justice for reincorporation into civilian life (2003–2012). The total sample was recruited over 4 years and included Colombian ex-members of illegal groups who participated in collective or individual demobilization processes from 2003 to 2012. Specifically, 69.9% of the sample engaged in a collective demobilization, and 30.1% demobilized individually (Table 1). All participants belonged to guerrilla forces (the Revolutionary Armed Forces of Colombia [*Fuerzas Armadas Revolucionarias de Colombia*; FARC], the National Liberation Army [*Ejército de Liberación Nacional*; ELN], and other guerrilla forces) or paramilitary groups (the United Self-Defense Forces of Colombia [*Autodefensas Unidas de Colombia*]). Within the framework of transitional justice, the Agency for Reintegration and Normalization (*Agencia para la Reincorporación y Normalización*; ARN) led a process to assess social-contextual and psychophysical factors in ex-members of illegal armed groups. To this end, the ex-members answered a comprehensive questionnaire and a semistructured interview designed by ARN, studying different PPVs and DoVs. This questionnaire was based on previous research as detailed later (Experimental Procedures) and applied by trained evaluators from seven sites across the country for 4 years (2010–2013). All participants gave their voluntary signed informed consent at the beginning of the survey and confirmed their acceptance to participate in the ARN assessment, endorsing the study's goals. The study was approved by the relevant institutional review boards (IRBs) (see S1).

A large body of work has indeed examined the risk factors and determinants of different types of violence, including interpersonal violence, civil war violence, gender-based violence, and intimate partner violence (for reviews, see Facel et al. and Capaldi et al.).[9,18] The study of the determinants of violence has mainly been conducted through epidemiological methods and has focused on one set of determinants, either individual mental health or sociocontextual, with few studies analyzing both of these factors simultaneously.[9,19–23]

On one hand, individual factors usually include post-traumatic stress disorder (PTSD), anxiety, depression problems, psychosis, and substance abuse disorders.[9,23–28] In addition, personality trait dysfunctions, including paranoid,[10,29] antisocial,[29,30] borderline,[30,31] narcissistic,[10,30] and dependent traits,[10,30] are the most prevalent mental disorders associated with violence. On the other hand, different social-contextual determinants have been linked to different types of violence, such as early adverse childhood experiences,[32] reduced educational achievements,[33] past disruptive behaviors,[28] witnessing violence,[9] social network influences,[34] poor political participation, and reduced access to social resources.[7,35] The study of factors associated with violence has also been assessed in ex-combatants and individuals exposed to armed conflict.[21,25,36–41] Those studies have found that early exposure to adverse experiences, symptoms of PTSD, disruptive behaviors during childhood and adolescence, and impulsivity are consistently associated with more appetitive forms of violence rather than more reactive

Table 1. Demographic data for each dataset

Dataset	Participants with DoV	Participants without DoV	p
Global violence (n)	2,117	2,117	
Age (mean (SD))	32.28 (7.69)	32.27 (7.69)	0.99
Sex (F:M)	218:1,899	221:1,896	0.88
Educational level (years)	6.5 (0.9)	6.4 (1.1)	0.76
Domains of violence (DoVs)			
Consequentialist DoV (n)	4,035	4,035	
Age (mean (SD))	32.80 (7.94)	32.76 (7.89)	0.84
Sex (F:M)	420:3,615	419:3,616	0.97
Educational level (years)	6.2 (1.2)	6.4 (1.1)	0.46
Appetitive DoV (n)	4,035	4,035	
Age (mean (SD))	32.87 (7.62)	32.88 (7.65)	0.95
Sex (F:M)	587:3,448	586:3,449	0.97
Educational level (years)	6.9 (1.5)	6.7 (1.7)	0.56
Retaliatory DoV (n)	4,035	4,035	
Age (mean (SD))	32.58 (7.57)	32.57 (7.54)	0.96
Sex (F:M)	585:3,450	551:3,484	0.27
Educational level (years)	6.9 (0.9)	6.1 (1.6)	0.63
Impulsive DoV (n)	4,035	4,035	
Age (mean (SD))	33.96 (7.61)	33.97 (7.61)	0.98
Sex (F:M)	617:3,418	623:3,412	0.85
Educational level (years)	6.1 (1.9)	6.3 (1.8)	0.34

forms.[36,38,39,41–43] Most of the previous studies have partially assessed a small, theory-driven set of individual and social-contextual determinants of violence. However, few studies to our knowledge have assessed the combination of these factors to investigate their association with historical data related to violence in ex-members of illegal armed groups. This approach may help to track the presence of multiple potential factors associated with violence and their interplay.[3,44]

Machine learning methods refers to a set of statistical techniques that learn from large and potentially noisy datasets and help to elucidate the most important factors to reach a prediction.[45] Models obtained using these procedures are automatically tailored and situated to the relevant population and can be fitted without imposing additional statistical loads.[46] Previous meta-analyses have shown limited statistical robustness of studies using assessments of violence based on a small subset of risk factors.[9,20,22,23] Machine learning procedures can complement studies using structured assessment, including multiple interactions between a comprehensive group of potential risk factors. Recent studies in psychiatry and in communities exposed to violence have started to navigate in this direction.[47–52] The present computational learning methods can help to elucidate associations between complex variables and numerous interactions between them.[53]

The interaction of contextual and individual factors associated with violence in the Colombian civil war is not well understood.[54] Previous studies in other populations are not straightforwardly applied to Colombian settings,[9,20,23,36,39,41] present different types (intimate partner violence, interpersonal or civil war types of violence)[9,18] or forms of violence assessment (i.e., appetitive or impulsive reactive),[9] and usually consider a preset number of tested predictions based on classical statistical methods.[9,11,21,22,26,32,42,55–58] These antecedents call for specific studies to better understand the specific factors associated with historical assessments of confessed acts of violence during the Colombian conflict. The current study differs from previous reports assessing risk factors associated with violence regarding a novel methodology (deep neural network [DNN][59] models and subsequent random forest classifiers [RFCs],[60,61] Figure 1) to evaluate simultaneously the complex interactions between a large number of features (risk and preventive social contextual/individual factors) associated with different confessed acts of violence. Although previous studies have described most of these factors, we develop a computational approach combining these multiple features into the classification of different DoVs. Systematic revisions and meta-analyses have weighed the importance of social and individual mental health variables as risk factors of violence.[9,62] But to our knowledge, no other experimental study has analyzed the combined interactions between a large number of features potentially associated with different DoVs via computational learning.

In this work, we use the term "prediction" following its conventional meaning in data science: as a part of inference based on the power of a variable set to predict another dependent variable set, the inference based on these two variables is considered predictive.[63] This meaning should not be confounded with the actual or ontological "prediction" of an individual's likelihood of violence. Our computational approach should also be considered with caution,[64] and any risk of stigmatization or extrapolation to other populations should be explicitly prevented (see Discussion). Moreover, our design is retrospective (historic confessed acts of violence) and based on a Colombian population. Consequently, our approach is not designed to predict future violent acts or extrapolate the results to other sociocultural settings. Thus, our main question is whether computational learning models can find which self-reported social-contextual or individual factors better classify retrospective DoVs based on confessed acts of violence during the Colombian conflict.

Considering this background, our study assessed factors potentially associated with the confessed act of violence in a comprehensive way, by including a survey designed to assess an extended evaluation of social-contextual and individual mental health factors (162 factors) in a large sample of ex-members of illegal armed groups in Colombia.[15,54] In addition, our study contributes a new array of evidence to the understanding of factors associated with historical measures of violence during the civil war in Colombia by introducing a novel computational approach, which could help to deal with data multidimensionality and multiple predictors.

We followed a combination of theory- and data-driven computational approaches to assess the most relevant PPVs in determining DoVs. Thus, we first designed, organized, and included a large set of PPVs based on the previous research highlighted

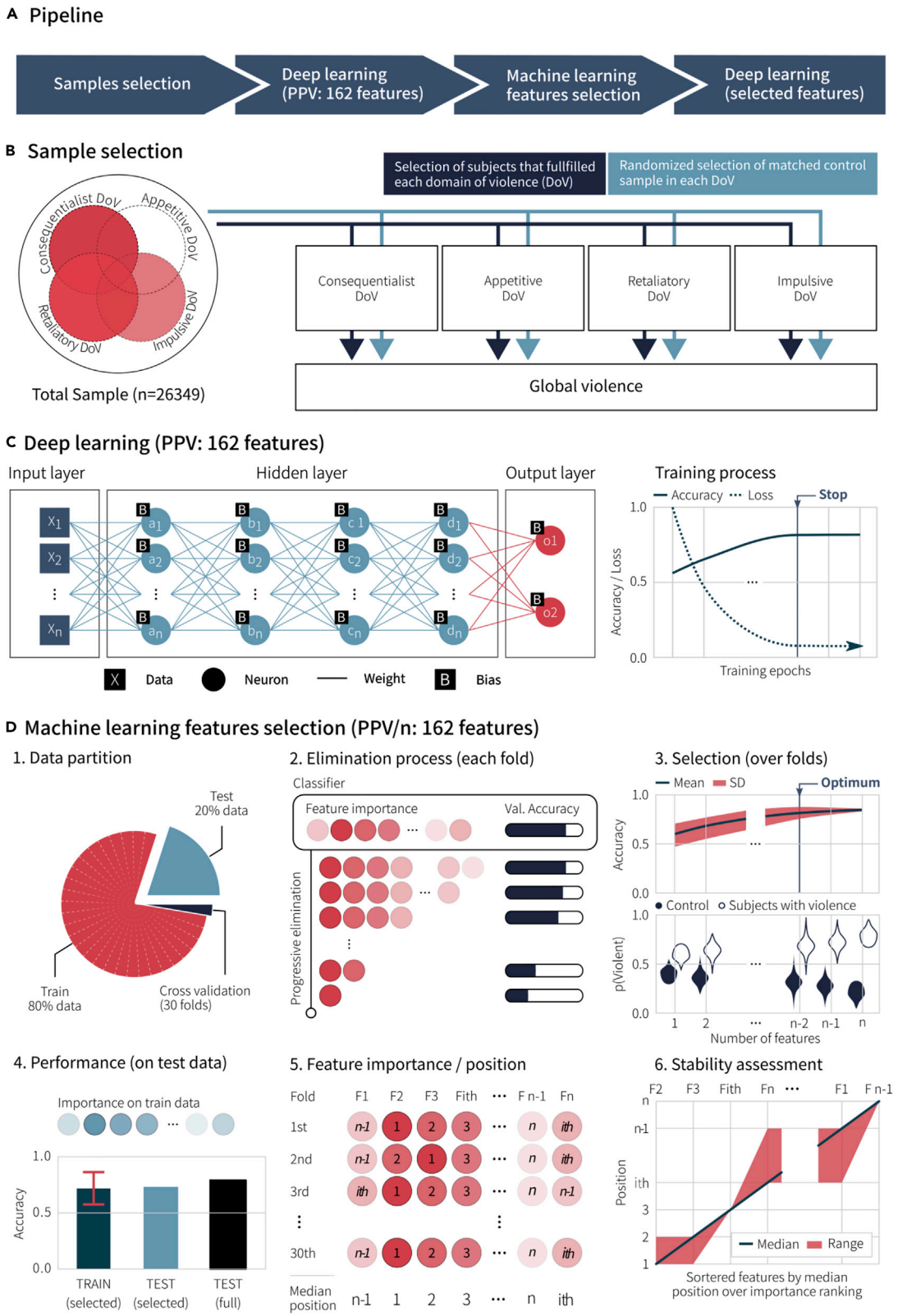


Figure 1. Study and data analysis design

(A) The pipeline of the computational data-driven approach following four steps: sample selection, the initial deep-learning neural networks procedure (DNN), feature selection with machine learning methods, and the second DNN.

(legend continued on next page)

Table 2. Deep learning neural networks results with the full set of features and after feature selection

	Accuracy (%)	Error	Number of iterations
Full set of features			
Global violence	96.24 ± 0.01	0.08 ± 0.01	2,060
Consequentialist DoV	92.30 ± 0.05	0.56 ± 0.02	6,426
Appetitive DoV	91.53 ± 0.05	0.39 ± 0.01	23,999
Retaliatory DoV	92.94 ± 0.05	0.55 ± 0.02	31,999
Impulsive DoV	87.53 ± 0.03	0.10 ± 0.02	1,651
Selected set of features			
Global violence	97.06 ± 0.02	0.25 ± 0.01	2,999
Consequentialist DoV	90.01 ± 0.06	0.63 ± 0.06	5,440
Appetitive DoV	92.03 ± 0.04	0.47 ± 0.04	1,499
Retaliatory DoV	92.22 ± 0.06	0.62 ± 0.06	1,042
Impulsive DoV	87.66 ± 0.01	0.12 ± 0.01	5,833

The accuracy of the validation set, the training error, and the number of iterations are shown for every dataset. Validation accuracy and training errors are expressed as the mean percentage ± SD. DoV, domain of violence.

above, including social-contextual and more individual mental health factors based on theoretical models. Afterward, we applied a data-driven approach that included DNN models[59] and random forest procedures[60,61] to robustly select the best PPVs in determining each DoV. The usage of DNNs and machine learning procedures is particularly helpful in dealing with wide data (i.e., a high number of PPVs).[59,65,66] These procedures involve minimal assumptions about the data-generating systems;[67,68] and they are able to perform statistical inferences with data gathered in uncontrolled experimental scenarios, in the presence of an extensive number of variables, and with non-linear interactions.[45,46] The combination of machine learning methods increase the robustness of predictions (see Makridakis et al. and Altman and Krzywinski).[67,45] Previous approaches have combined initial deep-learning analysis with subsequent procedures (e.g., random forest classification) to highlight the main features predicting the outcomes hiding in deep learning.[69,70] This approach has been classically used in different fields, including neurocognitive studies (see Hutzler for a review)[71] or genetic studies assessing massive data to identify particular loci that more accurately predict a clinical outcome.[72,73] Moreover, similar frameworks have been implemented in studies assessing inpatient violence (i.e., Menger et al.).[50] At the technical level, the combination of data science techniques involving deep learning and machine learning has proven robust to identify main predictors, enhance classification, and provide the top combination of features to classify outcomes.[70,74]

The first step in our study was setting a DNN to test if all PPVs are useful for reaching an accurate classification of each DoV. The DNN is a useful method to reveal the extent to which a group of features is relevant to track an outcome.[75,76] Afterward, we implemented a random forest procedure to delve into the structure of data and variable interactions and to track the most relevant features to predict different DoVs. Random forest procedures can handle large numbers of variables in large datasets and are a robust method for assessment of variable importance in comparison with classical linear regression models.[65,66] Finally, we ran a new DNN using only the best selected features captured by the random forest procedure to determine whether the non-linear interaction of these selected features can improve the classification of each DoV (see Figure 1 and Experimental Procedures).

RESULTS

Global violence analyses

Global violence refers to the expression of violent acts associated with all consequentialist, appetitive, retaliatory, and impulsive DoVs. Ex-members who declared that they had committed violent acts related to the four DoVs were included in this category. The initial DNN revealed that the full set of PPVs predicted global violence with high accuracy (validation set accuracy 96.24%; see Table 2 and Figures 2 and S3). These accuracy values represent the percentage of individuals adequately classified as individuals who admitted (or not) violent acts in each

(B) Sample selection procedure. Five datasets were generated: one dataset for global violence (participants who acknowledged the four types of domains of violence [DoVs] and one dataset for each DoV (consequentialist, appetitive, retaliatory, and impulse). Each DoV dataset included participants who presented each DoV and a group of control participants who did not show the DoVs.

(C) DNN. A schematic description of the multilayer DNN illustrates the procedure for tracking whether the total number of potential predictors of violence (PPVs) was useful to classify the DoV in each dataset (left). The right shows a schematic representation of accuracies obtained by the DNN, the x axis shows the number of iterations (epochs) required to obtain a particular accuracy, and the y axis shows the percentage of accuracy obtained by the DNN across epochs. To that end, each dataset was split into training and validation sets comprising 80% and 20% of the data, respectively, and the latter set was used to measure performance (right). The DNN was run twice: at the first stage to test whether the full set of factors could predict each DoV and after the machine learning methods to evaluate whether each DoV could be predicted with a reduced number of features.

(D) Subprocess of machine learning methods. We used a random forest classifier and a progressive feature elimination procedure to identify the factors that more accurately predict each DoV. For this purpose, each dataset was split into training and validation sets comprising 80% and 20% of the data, respectively, and the training set was split into 30 folds for cross-validation (D1). D2 depicts how the predictors were selected. It also reveals the number of features that reached good accuracy in predicting each DoV. D3 shows the mean validation accuracy over the folds. The optimal number of features was defined visually to improve the classifier's performance. The top shows a schematic representation of the accuracies obtained by the classifier. In this graph, the x axis shows the number of predictors included in the training process; the y axis shows the percentage of accuracy obtained. In the bottom graph, the x axis graphs the number of features included in the training process, and the y axis depicts the probability of being classified within each DoV (p(Violence)). D5 shows the importance of the features and their position in the ranking for each cross-validation iteration. D6 shows the stability of the features. In this graph, the x axis shows a kind of schematic representation of features organized according to their median position in the importance ranking. The y axis depicts the position of each feature in the full set of features (162 predictors) over the cross-validation iteration. The final step is shown in D4, where the classifier performance using the full set of features and the selected set are compared.

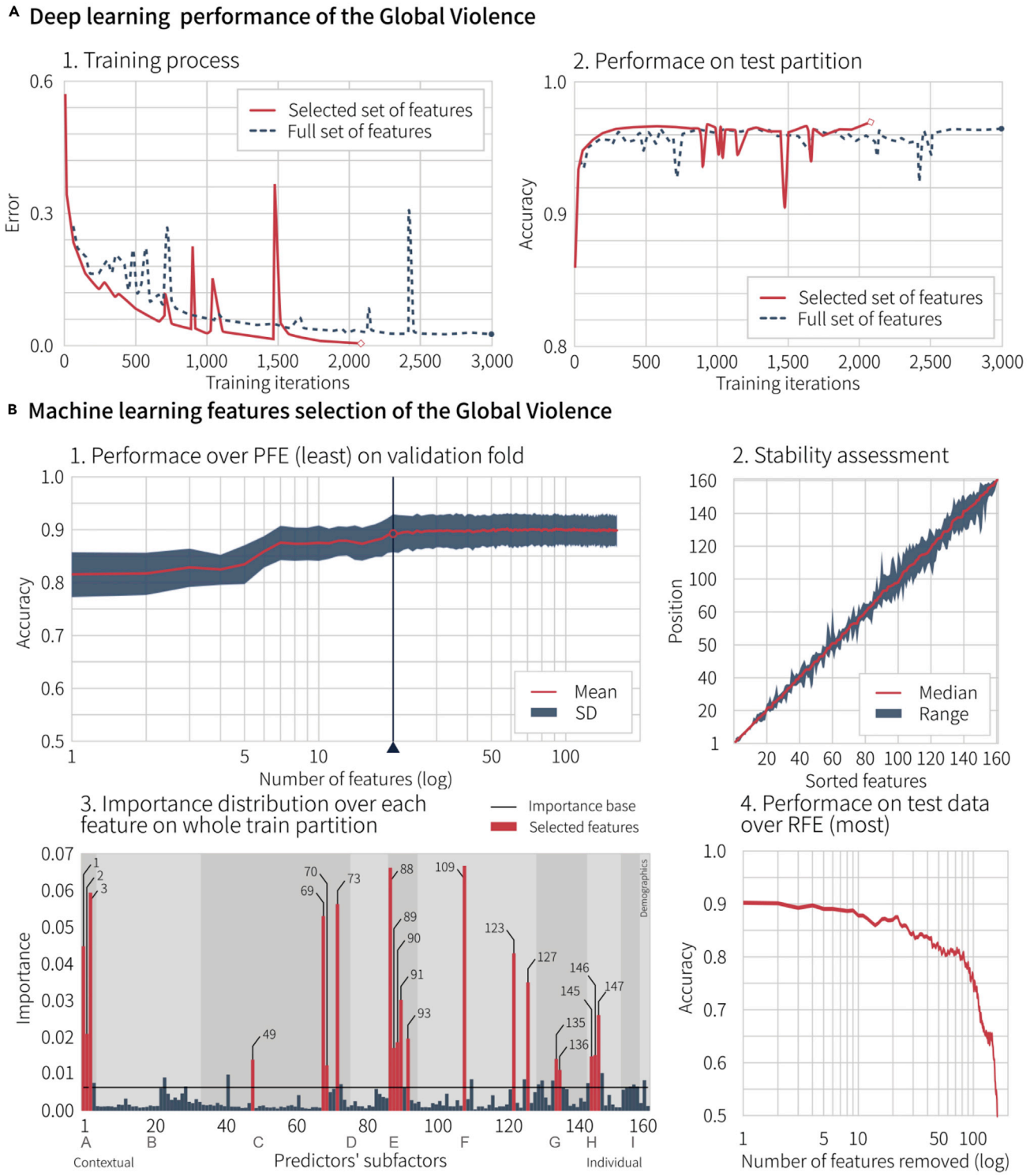


Figure 2. Classification of global violence based on potential predictors

(A) Deep-learning neural networks (DNN) procedure. Training error (A1) and performance (A2) on test partition following a DNN procedure on the dataset of the global DoV. A1 and A2 show the error and accuracy across iterations for the full (dotted blue) and selected (continuous red, after machine learning feature selection) set of features. The x axis in A1 shows the number of iterations of DNN. The y axis shows the proportion of error in the training of the DNN. The x axis in A2 depicts the number of iterations of DNN. The y axis shows the accuracy of DNN on the test partition.

(legend continued on next page)

DoV based on a certain number of PPVs. Then, we tracked the feature independence of the 162 factors before completing the down-selection process by controlling redundancy and stability. Feature redundancy factors due to high correlations between factors reduce the probability of a factor of being selected, affecting the stability of the feature selection process.[77] To track these potential issues, we ran successive iterations with different subsamples and verified the probability of each predictor being selected and the stability of that selection. Following these procedures, the results did not reveal redundancy issues in selecting factors, and the stability of selected predictors after iterations was high, as shown in Figure 2B2.

The RFC[60] used in the progressive feature elimination (PFE) analysis with the full PPVs reached 89% accuracy (SD 3.5%, sensitivity 91%, specificity 88%, area under the receiver operating characteristic curve (AUC) = 0.96, see Table 3 and Figure 2). The PFE revealed similar performance using only 20 PPVs of the full set. After a threshold of 88% accuracy, the inclusion of more features did not improve the classification performance (Figure 2); in addition, this analysis showed excellent stability for the first 20 features in the ranking (Figure 2). *Post hoc* analysis revealed that each of the individual predictors showed (by itself) poor predictive accuracy. By contrast, the combination of 20 predictors reached high predictive accuracy (above 88%, Figure 2). A new recursive feature elimination analysis confirmed the absence of single variables being able to predict global violence (Figure 2). A second RFC[63] using only the 20 selected predictors reached an accuracy of 88% (sensitivity of 90% and specificity of 87%; Table 3).

The machine learning analyses revealed that the most critical group of PPVs for global violence (Figure 2) were contextual subfactors. Notably, the essential contextual predictors were the mean scores of social network adversities, such as the experience of violence, normalization of violence at the first and second stages of life, and membership identification, and the mean scores of all contextual subfactors (at all stages of the life trajectory). Specifically, the essential contextual predictors were (1) the mean score of the social network adversities subfactor at each of the three stages of the life trajectory (in the first, second, and third stages, ranked 1st, 5th, and 18th, respectively); (2) seven items measuring the membership identification subfactors (ranked 2nd, 4th, 8th, 12th, 13th, 14th, and 19th); (3) the mean score of the normalization of violence subfactor at the first and second stages of the life trajectory (ranked 6th and 9th); and (4) the mean scores of all subfactors of the contextual factors at each of the three stages of the life trajectory (ranked 3rd, 7th, and 11th). In addition, to a lesser extent, global violence was also predicted by some individual subfactors, including paranoid, borderline, and traits and symptoms of manic episodes and post-traumatic stress. Specifically, the most relevant individual predictors were (1) antisocial, paranoid, and borderline traits (ranked 10th, 15th, and 16th, respectively) and (2) symptoms of affective exaltation episodes and PTSD (ranked 17th and 20th, respectively) (see Figures 2 and S1 and Table S3).

After performing the feature selection process, we tracked correlations between factors and DoVs, as revealed by Figure S2. We controlled the presence of possible correlations and collinearity between PPVs and DoVs using the variance inflation factor (Figures S2A–S2E). Following these analyses, only three contextual predictors (the average scores of all subfactors of the individual factors, ranked 3rd, 7th, and 11th) were collinear and significantly correlated with global violence (all variance inflation factor indexes above 1) after Sidak correction (see Figure S2). To assess the extent to which the presence of significant correlations determined predictive accuracy, we ran a new PFE analyses while discarding step by step each one of those factors. The results of these analyses revealed similar accuracies for the two databases (i.e., 89.1% with 20 predictors and 89% with 17 predictors). We also assessed each of those predictors' predictive accuracy using a decision tree classifier (the accuracy of none of the predictors exceeded 55%). The final DNN ran with only the 20 best predictors of global violence maintained a high accuracy (validation set prediction 97.06%; see Figure 2 and Tables 2 and 3).

Predictors of each DoV

Consequentialist DoV

The initial deep learning step yielded 92.3% accuracy (Table 2). The PFE analyses showed good stability for the first 20 features in the ranking, with an accuracy of 74%, a sensitivity of 73%, and a specificity of 73% (AUC = 0.81). Similar to global violence, social-contextual subfactors were better predictors of consequentialist violence than individual mental health subfactors. In particular, among the group of social-contextual factors, the most important predictors were (1) the mean scores of all subfactors of the social-contextual subfactors at the three stages of the life trajectory (ranked 1st, 2nd, and 3rd), (2) two items of the membership identification subfactor (ranked 4th and 6th), (3) normalization of violence (beliefs regarding violence subfactor, ranked 8th and 9th), and (4) the mean score of the social network adversities subfactor at the first and second stages of the life trajectory (ranked 5th and 10th). Among the group of individual mental health subfactors, the most important predictors were (1) the personality traits subfactor (antisocial, paranoid, borderline, dependent, and narcissistic traits, ranked 6th, 11th, 13th, 18th, and 20th, respectively) and (2) the mental health disorder

(B) Feature selection using machine learning techniques. Mean accuracy and standard deviation of the classifier over the progressive feature elimination (PFE) (B1) and the median position of features along with the importance ranking (B2) over the 30-fold cross-validation using the training partition (B1). The x axis in B1 shows a log representation of the number of features entered in the PFE. The y axis depicts the accuracy. The x axis in B2 depicts the features organized according to their median position along with the importance ranking. The y axis shows the position of each predictor of global violence over the cross-validation iterations. The optimal number of features is marked by the red line. B3 shows the importance and distribution of the predictors. The x axis in B3 depicts the place of each predictor in the full set of predictors. The y axis depicts the importance of each feature in relation to all groups of features in classifying global violence. The selected features are colored in red and numbered. In particular, the results show that the most important group of PPVs associated with global violence are contextual factors (social network adversity and beliefs regarding violence subfactors) and, to a lesser extent, a few individual factors. B4 shows the classifier's accuracy over the recursive most important feature elimination on the test partition. This graph reveals the stability of the predictors in classifying the DoV, as only after the elimination of 100 PPVs does the classifier's accuracy fall. The x axis depicts a logarithmic representation of the number of features removed from the PFE, and the y axis shows the accuracy of the test partition.

Table 3. Progressive feature elimination results

	Accuracy (%)	AUC	Sensitivity	Specificity	Precision	Recall	F1
Full set of features							
Global violence	89	0.96	0.91	0.88	0.87	0.91	0.89
Consequentialist DoV	75	0.82	0.74	0.75	0.75	0.74	0.75
Appetitive DoV	79	0.87	0.81	0.78	0.77	0.81	0.79
Retaliatory DoV	77	0.86	0.78	0.77	0.76	0.78	0.77
Impulsive DoV	66	0.72	0.70	0.64	0.57	0.7	0.62
Selected set of features							
Global violence	88	0.96	0.902	0.87	0.86	0.9	0.88
Consequentialist DoV	74	0.81	0.72	0.73	0.73	0.72	0.73
Appetitive DoV	78	0.86	0.80	0.77	0.75	0.8	0.78
Retaliatory DoV	76	0.85	0.77	0.76	0.76	0.77	0.76
Impulsive DoV	66	0.71	0.67	0.65	0.62	0.67	0.65

Accuracy of the random forest classifier on the test set and the area under the ROC curve (AUC) are shown. The sensitivity and specificity values are detailed for every dataset. DoV, domain of violence.

symptoms subfactor (affective exaltation episodes and psychotic, depressive, post-traumatic disorder and anxiety symptoms, ranked 14th, 15th, 16th, 17th, and 19th, respectively, Table S3). The final deep-learning analysis with the reduced group of features reached an accuracy of 74% (Tables 2 and 3, Figures 3B1–3B6 and S3).

Appetitive DoV

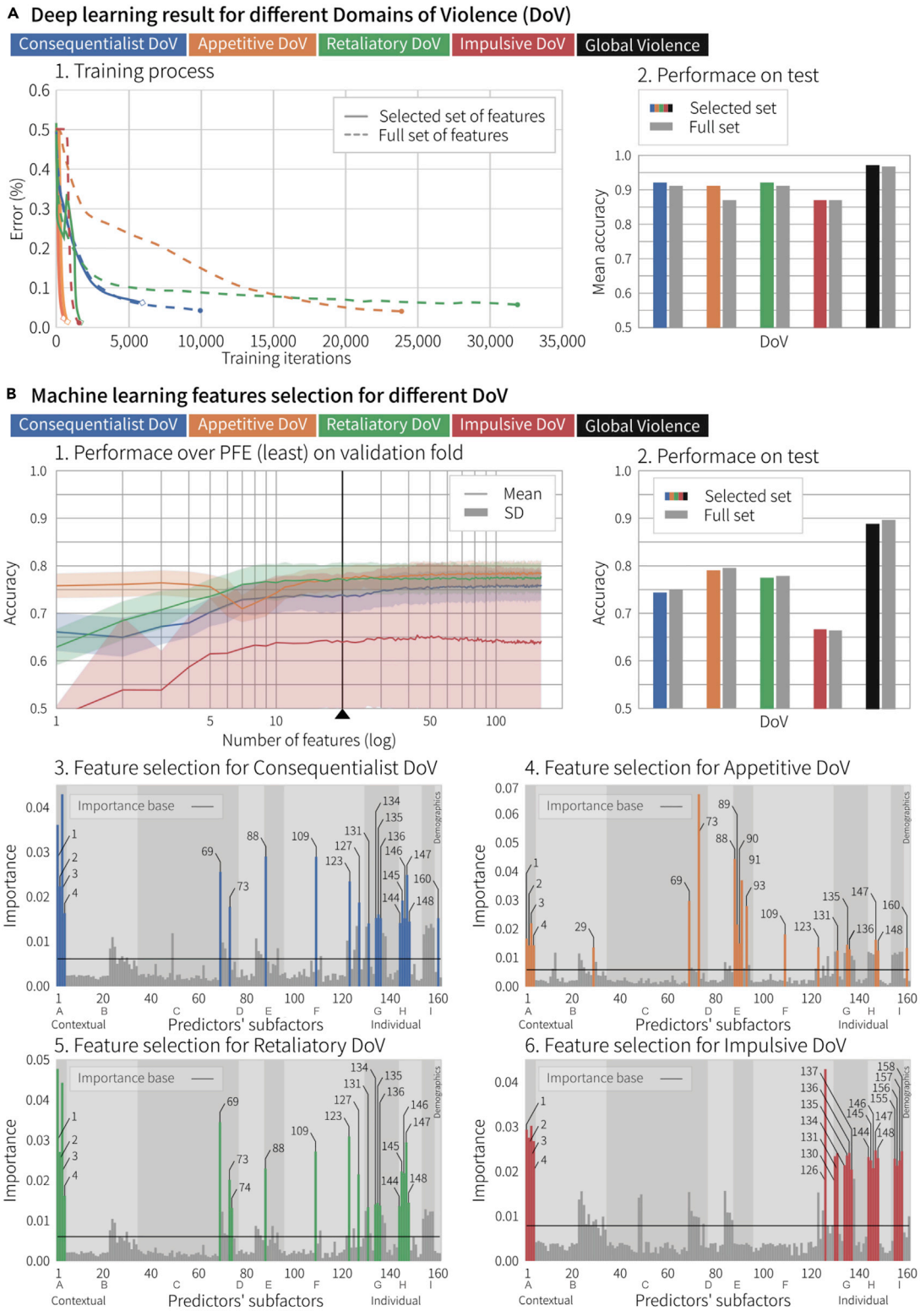
The initial deep learning step yielded an accuracy of 91.5% (Table 2 and Figure 3A). The PFE reached an accuracy of 78%, a sensitivity of 80%, and a specificity of 77% and AUC = 0.86. This analysis revealed that the best predictors of this DoV were also social-contextual factors. Specifically, the best predictors were (1) the mean score of the items of the social network adversities subfactor at each of the three stages of the life trajectory (ranked 1st, 4th, and 17th), (2) five items of the membership identification subfactor (ranked 2nd, 3rd, 5th, 7th, and 9th), (3) the mean score of normalization of violence at the first and second stages of the life trajectory (beliefs regarding violence subfactor, ranked 8th and 13th), (4) the mean scores of all social-contextual subfactors at each stage of the life trajectory (ranked 6th, 11th, 12th, and 14th), and (5) the mean score of difficulties in accessing educational and occupational resources (social vulnerability subfactor). Individual mental health factors were also identified as important predictors: (1) affective exaltation episodes, PTSD and anxiety symptoms (mental disorder symptoms subfactor, ranked 15th, 18th, and 20th, respectively) and (2) antisocial and dependent traits (personality traits subfactor, ranked 10th and 20th, respectively, Table S3). The final deep learning step using only the selected group of features reached a homologous classification accuracy level, as it was achieved by the full set of predictors (92.3% versus 91.5%, respectively, Tables 2 and 3, Figures 3B1–3B6 and S3).

Retaliatory DoV

The initial deep learning reached an accuracy of 92.9% (Table 2). The PFE analyses reached an accuracy of 76%, a sensitivity of 77%, a specificity of 76%, and an AUC of 0.85. As in the previous DoV, retaliatory violence was best predicted by social-contextual rather than individual mental health factors. Among the social-contextual factors, the best predictors were (1) the mean scores of all social-contextual subfactors at the three stages of the life trajectory (ranked 1st, 2nd, and 5th), (2) the mean score of the items of the social network adversities subfactor at the three stages of the life trajectory (ranked 3rd, 12th, and 17th), (3) the mean score of normalization of violence at the first and second stages of the life trajectory (beliefs regarding violence subfactor, ranked 4th and 11th), and (4) the mean score of the membership identification subfactor at the first and second stages of the life trajectory (ranked 7th and 10th). Among the group of individual mental health subfactors, the best predictors were (1) the personality traits subfactor (antisocial, paranoid, borderline, dependent, and narcissistic traits, ranked 6th, 8th, 9th, 14th, and 19th, respectively) and (2) the mental disorder symptoms subfactor (psychotic symptoms, affective exaltation episodes, post-traumatic disorder, and anxiety symptoms, ranked 15th, 16th, 18th, and 20th, respectively, Table S3). The final deep learning step using only the selected group of features reached an accuracy of 92.3% (Tables 2 and 3, Figures 3B1–3B6 and S3).

Impulsive DoV

The initial deep-learning analysis reached an accuracy of 87.5% (Table 2). The PFE analyses reached an accuracy of 66%, with a sensitivity of 67%, a specificity of 65%, and an AUC of 0.71. This DoV was better predicted by social-contextual subfactors, but compared with the other domains, an increased number of individual mental health subfactors were found. Among the group of social-contextual subfactors, the best predictors were (1) beliefs about the value of complying with laws and juridical regulations (beliefs regarding violence subfactor, ranked 1st) and (2) the mean scores of all social-contextual subfactors at the three stages of the life trajectory (ranked 2nd, 3rd, and 4th). Among the individual mental health factors, the best predictors were (1) personality traits subfactors (antisocial, dependent, narcissistic, paranoid, and borderline traits, ranked 6th, 11th, 13th, 16th, and 18th, respectively), (2) mental disorder symptoms subfactors (symptoms of anxiety, PTSD, affective exaltation, depression, and psychosis, ranked 8th, 9th, 10th, 12th, and



A. Mean Contextual Predictors, B. Social Network Adversities, C. Social Vulnerability, D. Political Context, E. Membership Identification, F. Normalization of Violence, G. Mental Symptoms, H. Personality Traits, I. Protective Aspects.

Figure 3. Classification of each domain of violence based on potential predictors

(A) A1 shows the training error and performance on the test partition following a DNN procedure on the dataset for each DoV, including the consequentialist (blue), appetitive (orange), retaliatory (green), and impulsive (red). A2 shows a comparison of the accuracies in predicting each DoV based on the DNN run with a full set

(legend continued on next page)

17th, respectively), and (3) protective aspects subfactor (self-acceptance, self-projection, physical integrity, environmental adaptation skills, and socioemotional skills, ranked 7th, 14th, 15th, 19th, and 20th, respectively, see Table S3). The final deep-learning analysis with the reduced group of features reached an accuracy of 88.6% (Tables 2 and 3, Figures 3B1–3B6, and S3).

Linear associations among PPV and each particular type of DoV

As reported for global violence, we assessed the PPVs significantly and linearly correlated with each DoV score. This pattern of correlations did not reach significant levels after Sidak correction and revealed no collinearity between PPVs and DoVs (Figures S2A–S2E).

DISCUSSION

Although previous reports have assessed factors associated with violence in different large samples,[11,12] we combined an innovative methodological approach (two-stage computational approach) with multiple social-contextual and individual factors in a large sample of ex-members of an illegal armed group. This was a simultaneous assessment of a sizable number of factors (more than 160 features) potentially associated with historical confessed acts of violence, including reported social-contextual and individual factors (comprising both risk and protective factors). To our knowledge, no other study (excluding systematic reviews or meta-analyses) has simultaneously analyzed the importance of a similar number of potential factors associated with archival data related to multiple DoVs, and including both potential risk and protective variables. Results within the tested dataset revealed that contextual conditions are stronger factors associated with DoV than the individual dispositions.

Global violence

Our results suggest a set of social-contextual factors that are relevant factors associated with historical measures of violence in Colombian civil war settings. By using a PFE, we analyzed the number of predictors needed to maintain a classification of global violence similar to those obtained analyzing the full set of PPVs. To this end, we ran a group of secondary analyses with a reduced number of factors (selected by the PFE) and assessed the extent to which those factors upheld the accuracy values compared with those obtained with the complete group of factors. Furthermore, we followed this procedure to determine the most influential factors associated with each DoV and order them according to its predictive value. A similar procedure has been followed in previous studies.[78–80] In our dataset, the analysis of the global violence dataset revealed a selected group of 20 PPVs reaching a performance similar to those observed with the full set of PPVs. After a threshold of 88% accuracy, additional features did not improve the classification performance (Figure 2).

In our dataset, as well as in previous research, past social network adversities[44] and reduced access to social resources[2] were partially associated with measures of violence. Arguably, stigmatization, discrimination, social exclusion, and exposure can be associated with a cycle of violence (i.e., the association between early threatening experiences and violence). This cycle has been previously related to genetic-epigenetic vulnerability,[81] cognitive and affective self-regulation,[81] and primary attachment difficulties.[82] Membership identification was also a relevant factor in the classification in our dataset. Affective-symbolic groupthink facilitates and intensifies violent group activities.[15,83] Sacred group values and collective causes reinforce group identity and promote violence.[6] Reduced access to social resources and a strong identification to the ideals of a social group have been postulated as strong predictors of rebellion and anti-state aggression.[84] Normalization of violence is another known factor associated with violence in civil war settings in other reports.[2,8,44] Moreover, violence can be used to exhibit power,[1] propel social mobility,[85] and vindicate honor.[86] Violence normalization is grounded in the social inculcation of violent roles.[1,5,85,86] Regarding individual mental health subfactors, our results parallel previous evidence showing that personality types (paranoid, antisocial, borderline, narcissistic, and dependent)[9] and psychiatric disorders (manic episodes, psychosis, anxiety, depression, and PTSD)[9,62,87,88] are associated with violence in civil war settings. Previous research has shown that personality traits including distrust, poor empathy, disregard for norm compliance, risky behaviors, emotional instability, emotional dysregulation, social isolation, attachment anxieties, externalization symptoms, and executive social-emotional impairments also influence violent behaviors.[9,62,87–89]

Specific DoVs

Social-contextual factors show high accuracy to classify all DoVs but exhibit a special relevance to characterize measures of the consequentialist DoV: 12 of the best 20 PPVs were social-contextual factors. This pattern of results coincided with previous studies showing utilitarian decisions influenced by different social factors, including social class,[90] political participation,[91] and cultural norms that normalize violence.[92] The consequentialist DoV was also classified by individual PPVs indexing personality traits (paranoid, borderline, dependent, and narcissistic traits) and affective disorders. A major tendency to

of factors or a reduced number of factors. The x axis in A1 depicts a logarithmic representation of the number of features entered in the progressive feature elimination (PFE). The y axis represents the accuracy of the PFE. The y axis in A2 depicts the mean accuracy of the DNN using a full set of predictors and a selected set of predictors in each DoV.

(B) Machine learning feature selection for each DoV. Mean accuracy and standard deviation of the classifier over the PFE (B1). The x axis in B1 depicts a log representation of the number of features entered in the PFE. The y axis in B2 represents the mean accuracy using a full set of predictors (gray) and a selected set of predictors (colored) in each DoV. The optimal number of features is marked by the black column. B3–B6 show the distribution of the importance of the features of each DoV (consequentialist, blue; appetitive, orange; retaliatory, green; and impulsive, red) on the whole training partition. The x axes in B3–B6 depict the place of selected predictors in the full set of predictors. The y axes in B3–B6 represent the importance of each feature in relation to all groups of features in predicting each DoV. The selected features are colored and numbered.

follow utilitarian decisions has been predicted by personality traits (mostly persons with individualistic rather than collective traits)[91] and by emotion regulation and cognitive control mechanisms.[93]

Appetitive and retaliatory DoVs shared a similar group of factors underlying the classification. Most of the top 20 PPVs for both DoVs were social-contextual factors, including the presence of social adversities, the normalization of violence, and membership identification. This pattern confirms previous studies showing major appetitive and retaliatory forms of violence in individuals who had early social adversities, as well as cultural and contextual acceptance of violence.[16] Furthermore, both DoVs were predicted by the presence of antisocial, borderline, and narcissistic traits, as well as by the presence of affective and anxiety disorders. Appetitive and retaliatory forms of violence were also highly prevalent in individuals who exhibited more egoistic personality traits and poor emotion regulation.[36]

In the case of the impulsive DoV, the most recurrent features associated with this historical measure of violence were the individual mental health factors (more than 15 of the 20 best PPVs). Crucially, reduced protective aspects associated with ability to cope with stress, self-projection, and emotional skills (potential protective aspects) were relevant factors for this DoV classification, confirming previous results.[94,95] Individuals with poor coping styles and reduced emotional regulation resources are more likely to exhibit aggression and violent behaviors in the presence of stressful situations.[96]

Insights from computational approaches in the study of violence

Our study reveals that a specific combination of contextual and individual factors is more accurate than any isolated factor in classifying subjects presenting positive historical measures of violence across all DoVs. Moreover, relevant information for classification accuracy is concentrated in a combined set of PPVs. By performing a before-and-after feature selection comparison of the DNN performance, we confirmed the quality of the chosen feature set (i.e., no predictive power was lost by discarding other features). Thus, a combination of social network adversities, normalization of violence, high membership identification, and particular personality traits and mental symptoms seems to be a factor associated with historical confessed acts of violence from the Colombian civil war in different DoVs. We ruled out the possibility that our results were only a consequence of collinearity or simple correlations between PPVs and DoVs (Figure S2). Furthermore, our results are not derived from a simple summation of positive responses associated with PPVs. They revealed low predictive classification values when we analyzed linear associations between PPVs and DoVs. Moreover, the pattern of results for each DoV revealed a different set of PPVs, and the inclusion of more predictors beyond the PFE threshold did not improve the classification accuracy in any DoV.

Our study adds novel information to the study of factors associated with violence. First, we included a large number of social-contextual and individual mental health determinants of violence. Second, few studies have simultaneously evaluated different DoVs, including consequentialist, appetitive, retaliatory, and reactive violence. Third, some studies have used machine learning procedures to assess risk factors of violence in psychiatric patients[50,51] and in individuals exposed to violence.[47] However, in our study, we used machine learning methods to weigh a large set of social-contextual and individual mental health factors associated with confessed acts of violence in a susceptible sample of ex-members of illegal armed groups of the Colombian conflict, one of the most protracted armed conflicts across the world.[37,54]

Computational learning procedures can provide complementary data-driven tools with translational relevance.[97] For instance, in ex-members of illegal armed groups, machine learning methods have been applied to predict specific appetitive aggression and levels and post-traumatic stress profiles.[40] Thus, machine learning and data mining methods may be useful to create insights into past patterns of violence from large datasets in ex-members.[15] However, several steps should be performed (e.g., replication and confirmation of main predictors in independent samples, validation in different contexts, and evaluation of machine learning procedures to predict future behavior) before designing translational applications.[98]

Although our results highlight the potential use of computational learning to identify patterns associated with violence in ex-members of illegal armed groups, the use of artificial intelligence systems in legal and criminal settings should be considered with extreme caution, especially in decision-making processes. This is relevant regarding our approach, designed to classify past (historical) confessed acts of violence but not future violence. A recent call,[99] although designed for the prediction of future events, is also relevant for our study. Such call emphasizes the importance of considering trust calibration and uncertainty in machine learning. Beyond uninterpretable deep learning results, the machine learning analysis includes trust calibration as it allows one to provide explanations for the main predictors (providing interpretability) and requires understanding of the system's capabilities (interpretability) and reliability (uncertainty estimates).[99] Our procedures followed these recommendations at the current dataset level by (1) reducing uninterpretable results (deep learning) with identification of main features (random forest and machine learning PFE); (2) implementing calibration, feature stabilization, and multiple accuracy metrics; (3) avoiding the misidentification of probabilities of individual classification with ontological causality between predictors and outcomes; and (4) connecting the data-driven results with previous theoretical and empirical evidence favoring interpretability and readability. Any interpretation of the present results should consider these analytical and conceptual restrictions.

Limitations and further research

Our study has important limitations and opens up a new an agenda for further research. Our assessment included some factors recollecting events across a lifespan continuum. Previous studies have discussed the potential problems associated with the use of self-reports to track past events.[9,100] However, other reports have shown the relevance of tracing retrospective risk factors related to mental health outcomes,[101] social determinants of health,[102,103] and risk factors related to violence.[20,25] Our results revealed that the life stage in which social-contextual

factors occurred was relevant to improving the computational predictions.

Another potential limitation of our study is the use of self-report scales for the PPV assessment. Although it could bias the ex-members' answers, it did not affect the capacity of the data to effectively determine DoV as it is revealed by the high classification accuracies reached by our procedures. In addition, our results did not reveal a biased selection of a particular type of determinant of violence. Therefore, we did not find explicit clues of specific (i.e., non-generalized or restrictive) bias in the pattern of results. Furthermore, reports of PPV and DoV were collected in an anonymous and protected clinical setting during the reintegration process, and individuals were informed that their responses in this assessment did not have an impact on their legal processes. Those considerations suggest that individuals did not bias their responses in a specific way. However, future studies should include other variables less dependent on self-report to assess potential determinants of violence.

In addition, although we observed high accuracies in the classification of all DoVs, these domains were tracked with categorical outcomes associated with confessed acts of violence. We were limited to such measures because particular details on the intensity and type of violence committed by participants were protected by confidentiality agreements. Nonetheless, the level of accuracy obtained by our approach was high. Future assessment should confirm our results by using less obtrusive measurement strategies, such as experiments or randomized response techniques, and by assessing different degrees of violence. The current state of the art of studies on violence in civil war settings calls for the development of more ecological research approaches.[104] New research initiatives in civil war settings may benefit by assessing interactively different sources of violence.[105]

In this study, we have analyzed each DoV in isolation, to improve the determination of specific contributions to each domain (except in global violence, where the four DoVs were considered). Although previous studies suggest that individuals tend to exhibit a particular DoV over others, a restricted set of individuals show violence associated with mixed DoVs.[16] In our study, the individuals who presented two or three DoVs were excluded to avoid multiclass classification issues.[106] Future studies should assess the possible overlap and interactions between DoVs and try to establish accurate determinants of individuals with violence due to mixed motives. The comparison of features between DoVs is out of the scope of this work. However, future studies should assess systematic comparisons of features associated with each type of DoV by implementing other normalization and multiple comparison classifications among the different datasets. Furthermore, future studies should assess the existence of complex interactions and hierarchies between risk factors and different DoVs. Finally, future studies should also compare DNN methods and machine learning with classical statistical approaches to assess a large combined set of predictors in large samples. Their methodological approach, which is out of the scope of the current paper, would be useful for other works related to methodological comparisons among techniques.

An additional limitation concerns the selected architecture and training process of DNNs. No validation set was used during the training process. The large number of iterations needed for the DNNs to converge could be related to the sigmoid activation functions selected in all units rather than in the rectified linear units. Similarly, the mean squared error choice as a loss function (in contrast to cross-entropy) can result in ever-decreasing rates and slow down training.[107] However, the neural networks were not aimed to obtain the best-optimized test but to (1) test whether it was possible to classify between subjects using the full set of data and then (2) examine the models' classification using the random forest selected features.

The presumed motivation of ex-members to reintegrate into society raises questions about a potential tendency for participants to emphasize the presence of some factors regardless of the degree to which these factors were present, and thus artificially inflate the role of these features. The emergence of *post facto* rationalizations of violent and illegal acts can play a potential role.[108] However, although this potential bias could misdirect the ex-members' answers, we did not find a pattern of results causing us to think this bias had selectively affected a single set of features. Moreover, this risk may be attenuated, as the ARN explicitly informed participants that their answers would not affect any legal or reintegration processes. Similarly, participants would minimize their engagement in violent actions, perhaps tending to report only those already known to have been committed by the legal system. However, these influences will reduce or abolish only the features' power in the machine learning pipeline to accurately classify the different DoVs. In contrast, our results suggest that, despite these potential self-report unknown biases, the data collected are robust enough to classify the DoVs with high accuracy. Despite these considerations, our data cannot be considered causal in any sense. The role of moralization of violence triggered by reconciliation and rehabilitation processes is beyond the scope of this work. Future studies using more implicit measures to assess individuals' attitudes toward past violent or illegal acts may be better designed to evaluate this question.

The present results should be considered with caution, especially regarding ethical challenges, including the risk of stigmatization or extrapolation to other populations.[64] Our data are relevant for population-based assessment factors associated with historic confessed acts of violence in the context of the Colombian conflict. In this country, complex origins of violence have long historical roots in land possession conflicts, governmental difficulties in dealing with social inequalities, and political confrontations among liberals and conservative parties (since 1946). In addition, it has been promoted by disputes between state forces and the guerrillas, the incursion of paramilitary forces, and drug trafficking. Future studies should also assess the role of political mediators of violence, such as cultural practices related to political threats affecting the experience of armed conflicts. Political ideology seems to have a role in mediating the expression of violence in Colombian conflicts.[109] Thus, further studies should analyze additional ideology and political factors. These specific antecedents and the way they could shape the interplay between PPVs and DoVs prevent any extrapolation to different sociocultural settings. Future cross-cultural research should shed light on the comparability of findings across different manifestations of violence in other scenarios.

Our results do not resolve whether the best predictive factors in classifying violence could be used to predict future violent acts. Our approach is not aimed to establish causality. Furthermore, our results are not targeted at ontologically categorizing participants as violent or non-violent individuals, as this approach could increase stigmatization. Prediction, as stated above, is considered only from the outcomes (variables) being statistically predicted and not from an ontological prediction of violence. Similarly, the results do not imply that the selected predictors should be used in countering violence (i.e., considering an individual more prone to violence when he or she presents some of the top determinants of violence). Our results provide only the first, yet important, steps toward understanding the complexity of interactions between multiple factors associated with past violence. Further studies should assess our approach's robustness to characterize future expressions of violence in new samples and new populations using different methodologies, including longitudinal studies. Also, further studies could examine whether the PPVs change during the lifespan.

Conclusions

Our results provide the first successful population-based classification of different DoVs based on Colombian historical data and call for the development of computational approaches for situated, multifactorial, and evidence-based factors associated with violence in vulnerable participants. In brief, a set of specific social contextual factors, in addition to individual mental health measures, seems to be associated with a greater classification of different DoVs.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Agustín Ibanez (agustin.ibanez@gbhi.org).

Materials availability

There are restrictions to the availability of materials and answers of this study due to the ethical requirement to ensure participants' confidentiality, but they are available from the lead contact on reasonable request.

Data and code availability

The datasets and codes supporting the current study have not been deposited in a public repository because they refer to protected information related to reintegration processes but are available from the lead contact on request. The datasets generated by each DoV and the Python scripts for each data-driven procedure are available upon request from the lead contact.

Participants

The total sample was recruited over 4 years and included 26,349 Colombian ex-members of illegal groups who participated in collective or individual demobilization processes from 2003 to 2012. Specifically, 69.9% of the sample engaged in a collective demobilization and 30.1% demobilized individually (see Table 1). All participants belonged to guerrilla forces (FARC, ELN, and other guerrilla forces) or paramilitary groups (the United Self-Defense Forces of Colombia, *Autodefensas Unidas de Colombia*). Most of the ex-members participated in collective or individual demobilizations in the context of the normative development of transitional justice in Colombia.[14] As part of the legal demobilization process, each participant individually gave a full, voluntary deposition of the group's activities, responsibilities, or crimes that may have involved violent acts. All offenses were documented and confirmed by legal records but were not included in the database due to the confidentiality restrictions of transitional national justice. Consequently, only the categorical scoring of DoVs was included in this study. In addition, in the demobilization processes, the ex-members of illegal armed groups answered a comprehensive questionnaire and a semistructured interview studying different PPVs and DoVs to detect any potential social-contextual and psychophysical risks in ex-members of illegal armed groups. All participants gave their voluntary signed informed consent at the beginning of the survey and confirmed their acceptance to participate in the ARN assessment, endorsing the study's goals. Participants were informed that their responses would be anonymous and that their answers would not be used to affect or modify their legal processes. They were also informed that they could refuse to participate in the survey at any stage of the research. Data included in this study were revised and approved by the ARN IRB, as part of the formal demobilization processes of ex-members of illegal armed groups. After that, Colombia's Externado University and the ARN signed an agreement to carry out the research on the data collected, which also included research goals and explicit accomplishments concerning previous ethics and legal approvals. The IRB of Externado University also approved the proposed research. The national survey designed by ARN had two aims: (1) an assessment of the mental health and quality of life of ex-members of illegal armed groups (this work was not related to the goals of the current study) and (2) an evaluation of the potential trajectories associated with violence by assessing social and mental health factors associated with confessed acts of violence in ex-members of illegal armed groups. Our work is the first report of this second aim. Two authors of the present study participated in the design of the original ARN survey.

Assessment of domains of violence

As a part of the extensive questionnaire implemented by the ARN, the participants were required to report the presence or absence of different DoVs. In particular, the participants were required to answer whether they had committed aggressive and violent acts, including the typical acts that occurred during the Colombian conflict (agreement to commit crimes, attacks against property, homicide, extortion, and kidnapping[14]) within different DoVs. Ex-members were included in a DoV if they declared that they had committed violent acts in that domain. A group of questions was used to determine participants' inclusion in the following four DoV categories.

Consequentialist DoV

This refers to a form of violence promoted by a utilitarian ("the ends justify the means") principle. This behavior is grounded in the moral notion that the only things that matter in determining the ethical rightness of an action are its consequences.[91] Aggression associated with utilitarian reasons seems to be more rooted in controlled, effortful, and conscious cognitive mechanisms rather than automatic and unconscious cognitive processes.[91] In addition, this type of behavior could also be associated with empathy skills, political decision-making processes, and personality traits.[15,110] Ex-members were included in this category if they declared that they had committed violent acts to pursue the group's strategic goals.

Appetitive DoV

This refers to a form of aggressive behavior accompanied by feelings of pleasure in the perpetration of violence. This behavior can be sustained in a cycle of positive reward that leads to perpetuating aggression. Appetitive violence is displayed intentionally and it is associated with domination and intimidation.[16,36] Ex-members were included in this category if they declared that they had experienced pleasure or enjoyment from perpetrating illegal violent acts.

Retaliatory DoV

This concerns a form of violence that happens when someone feels that he or she has been wronged and decides to take justice into their own hands and return the grievance. Retaliatory violence usually takes place when conflict escalates to the point of harm and usually is manifested in the form of cathartic behavior.[111] This DoV usually emerges in response to negative emotions.[17] Ex-members were included in this category if they declared that they had committed violent acts for vengeance or in retaliation against insults.

Impulsive DoV

This refers to a reactive form of violence that occurs in a sudden and unpredictable way and is activated in response to aversive stimuli or toward a perceived or imagined provocation.[16] This reactive form of aggression leads to a reduction of aversive emotional arousal associated with anger or fear and is associated with impulsivity and weak emotional and cognitive control

mechanisms.[96] Ex-members were included in this category if they declared that they had committed violent acts as a consequence of episodes of impulsivity or anger.

Global violence

This refers to the expression of violent acts associated with all four previous DoVs. Although this pattern of violence is unconventional, previous studies have reported that some individuals could exhibit an ultraviolent pattern of behavior with high ratios of recidivism mediated by utilitarian, active, but also reactive forms of violence.[16] Ex-members were included in this category if they declared that they had committed violent acts associated with the four DoVs.

The analysis of these DoVs is based on a theoretical background (see Elbert et al., Chester and DeWall, and Balash and Falkenbach),[16,17,110] but also supported empirically, as shown by a principal component analysis (PCA). PCA is useful to identify main components underlying a large set of variables.[112] PCA is robust in creating single scores and vectors that reduce a dataset's multidimensionality to identify, on a data-driven basis, the existence of different categories across data.[113,114] We used PCA to confirm on a data-driven basis the extent to which each theory-driven DoV can also be composed of different components. The PCA using the total PPVs (162 factors) reduced the multidimensionality of the total PPVs to two major components. We assessed in each component a t test (Bonferroni corrected) using the Euclidean distance between each two DoVs (see Table S5). The results of these analyses revealed that each DoV differed statistically from the other DoVs and from individuals who denied any DoV.

Assessment of potential predictors of violence

Social-contextual factors

The items of this factor comprised five different subfactors: (1) social network adversities,[44] (2) social vulnerability,[2] (3) political context participation,[35] (4) membership identification,[15,83] and (5) normalization of violence.[2,8,44] The aforementioned subfactors evaluated different stages of the participants' life by including items about situations that occurred before the participants joined the armed group (first stage), during the time that they were integrated in the armed group (second stage), and during the time after they left the armed group and accepted the reintegration processes (third stage). A detailed description of the items in each subfactor is provided in Table S1.

Contextual subfactor: social network adversities

Threatening social experiences. The participants were asked about their social exclusion experiences, including stigmatization, discrimination, or social exclusion, as well as the antecedent of being a victim of abuse by family relatives.[44]

Exposure to experience of violence. The participants were asked whether they had been victims of direct violence, whether their close relatives had been victims of violence, and whether they had witnessed a death or severe injury.[44]

Contextual subfactor: social vulnerability

Access to socioeconomic resources. We assessed social class restrictions, including the access to facilities that ex-members experienced in their life trajectory.[2]

Access to educational and occupational resources. The quality and sufficiency of educational or professional training resources were assessed. For instance, this subfactor included questions on access to socioeconomic resources and whether the participants considered it normal to have debts or financial problems. The relevance of formal studies for status mobility or job skill learning to support an honest living[2] was also assessed.

Social-contextual subfactor: political context

Political situations that impel social conflict. This subfactor included the evaluation of political guarantees and opportunities, the freedom to express political ideas, and access to political participation. In addition, we assessed the presence of political culture environments that promoted corruption and authoritarianism and evaluated conditions that restricted free speech and expression.[7,35]

Social-contextual subfactor: membership identification

Group attachment and membership values are factors associated with violent behavior and aggressiveness.[115] To assess this factor, we evaluated the network and types of attachment that the participants had with other members of illegal groups, commanders and comrades, as well as the degree of identification with the activities performed by the group.[15,83]

Social-contextual subfactor: normalization of violence

This subfactor included items assessing the beliefs, attitudes, and acts that normalize the actions of armed groups, including the following beliefs.

Beliefs that normalize armed group life. We assessed the presence of cultural patterns that idealize the dynamics of combat, life in the armed group, the war environment, and the naturalization of armed conflict in the communities of the participants,[2,8,44] even in the time before they joined the armed group.

Beliefs and attitudes toward compliance with juridical regulations. We evaluated beliefs and attitudes in contradiction with adjustments to normative and legal frameworks. The participants were assessed for their opinions regarding and motivations for complying with community and juridical norms. In addition, we scrutinized different problems in the generation of a legal income.[2,8,44]

Individual factors

The items in this group corresponded to three subfactors: (1) mental disorder symptoms, (2) personality traits, and (3) quality of life. The mental disorder and personality trait subfactors were assessed following the model of classification of the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition, text revision (DSM-IV-R). The quality of life subfactor was assessed using an instrument previously used in populations with antecedents of violence.[94]

The individual factor included three subfactors: mental disorder symptoms, personality traits, and individual protective factors.

Individual subfactor: mental disorder symptoms

To track mental symptoms, the participants were assessed with regard to whether they had presented symptoms of the most prevalent mental disorders (following the criteria of the DSM-IV-R) previously associated with violence, including depression,[88] anxiety,[9,88] affective exaltation episodes,[9,62] primary psychotic disorders,[9,31,87] PTSD,[9,81] substance abuse,[9,87] intermittent explosive disorder,[116] and pathological gambling.[117] Each symptom was assessed using yes/no questions referring to the criteria of the major symptoms of each mental disorder. The ARN survey assessed mental health symptoms following a previously standardized instrument (the Composite International Diagnostic Interview, CIDI).[118] This instrument is a comprehensive, fully structured interview designed to be used for the assessment of mental disorders according to the definitions and criteria of the DSM-IV.[119,120] The CIDI has been shown to be reliably implemented in surveys to assess the prevalence, severity, and burden of mental problems in large datasets[119–122] as performed in the current study. Nevertheless, our study did not pretend to provide a formal confirmation of a diagnosis of a mental disorder, as this requires an individualized assessment with an experienced professional in a clinical setting.

Individual subfactor: personality traits

All participants were also assessed for the most salient personality traits associated with violence, following the model of personality clusters described in the DSM-IV-R. This model includes three personality clusters (A, B, and C). The paranoid, schizoid, and schizotypal personality traits are integrated in cluster A; the borderline traits, antisocial narcissistic traits, and histrionic traits compose cluster B; and the dependent, obsessive, and avoidant personality traits are grouped in cluster C.[123] In this study, we explored the presence of personality traits previously associated with violence, including paranoid,[10,29] antisocial,[29,30] borderline,[30,31] narcissistic,[10,30,124] and dependent personality traits.[10,30] Each participant was asked whether a particular trait described his or her personality well. Paranoid traits included distrust and suspicion of others, the misinterpretation of ambiguous remarks as threatening, and retaliation; antisocial traits included poor empathic skills, a disregard for norm compliance, deceitfulness, and appetite for risk. Borderline traits included emotional instability, feelings of emptiness, and externalized behaviors, including property damage. Narcissistic traits included a sense of grandeur, the need to be recognized, poor empathic skills, and utilitarianism. Finally, dependent traits included attachment anxieties, social isolation, and reactive emotional dysregulation.[89]

Individual subfactor: protective factors

Following a well-designed instrument,[125] the participants answered a group of questions regarding protective aspects that can prevent violence, including the degree of self-acceptance (ability to accept one's perceived limitations and possibilities),[94] self-projection (skills for generating a life plan and prospection),[126] environmental adaptation skills (resources for adapting to contextual changes),[94] social and emotional skills (abilities to regulate emotional responses and integrate

into social groups),[94] and determinants of physical integrity (resources to satisfy basic health needs and the presence of physical symptoms).[127] Here, we summarize the five components of this subfactor. A detailed description of the items that compose each aspect is provided in Table S2.

Self-acceptance. This component assessed the degree of one's self-acceptance as a valuable and pleasant human being and the ability to accept one's perceived positive and negative aspects as well as one's limitations and possibilities.[94]

Self-projection. This component examined the conditions, resources, and expectations of the person for building a life project and the skills of imagining him- or herself in the future with a concrete life plan.[94]

Environmental adaptation skills. This component assessed the personal resources, abilities, and past experiences that support adaptation to life changes, such as displacement, migration, or forced change from an urban to a rural area as a consequence of armed conflict.[94]

Socioemotional skills. This component assessed the ability to face stressful situations, such as the death of a relative or the presence of financial difficulties. In addition, to explore this factor, the participants were asked about their skills to regulate emotional responses as well as the presence of relevant and intimate and supportive relationships.[94]

Physical integrity. This component assessed respondents' resources to satisfy basic health needs. In addition, it assessed the presence of symptoms of physical challenges, including chronic disease, pain symptoms, and disabilities.[128]

Following previous procedures,[129,130] the answers to each social contextual question (see Tables S1, S2, and S3) were entered in analyses in a binary form. In addition, we performed a global score of questions assessing each social contextual factor, including (1) social network adversities, (2) social vulnerability, (3) political context, (4) membership identification, and (5) normalization of violence. Each independent question (ordinal variables) and the global scores of each social contextual factor (continuous variables) were assessed in RFC analyses of each DoV.

Data analyses

Computational approach

Computational data-driven procedure. We followed a combination of automatized analyses that included the implementation of DNN[59] and machine learning methods, particularly, random forest procedures[60,61] to increase the robustness of analyses in determining different DoVs. We implemented DNN to test to what extent the contextual and individual factors could reach an appropriate classification of violence in each DoV. Afterward, we employed a random forest procedure to select the group of best factors (contextual or individual) to determine each DoV. Finally, using a new DNN, we tested whether the group of selected factors could improve the accuracy of the classification of each DoV.

Computational methods

Sample selection procedure. The sample selection was used to generate a well-matched dataset to apply machine learning procedures. As the number of participants in DoV groups was different, we matched the group size of each DoV to reduce bias associated with differences in the sample number between datasets following a previous procedure.[61] Thus, we generated independent datasets (one for each DoV) including ex-members of armed groups who declared they had committed violent acts associated with each particular type of DoV. In each dataset, the individuals who reported violence were matched by sex, age, and education following a one-by-one procedure with a control group of ex-members who denied violence associated with the assessed DoV. The sociodemographic factors in the datasets were controlled for considering that the violence has shown to be modulated by those factors.[9] In each dataset, we included only individuals who declared one type of DoV. In consequence, no single subject pertaining to one DoV was also included in another dataset. However, the individuals who declared violence associated with all DoVs were included in the group of global violence. These individuals were not included in other DoV datasets. Moreover, individuals who declared violence associated with two or three types of DoV were not further included in the analyses to avoid specificity issues associated with multiclass classification problems.[106]

From the total sample, we generated an initial dataset (n = 2,117) of participants who fulfilled the four DoV criteria (global violence). A group of control participants (i.e., participants who denied violence associated with any DoV, matched by sex, age, and educational level, n = 2,117) was also selected. Thus, the dataset of global violence comprised 4,234 individuals. Then, we followed a similar procedure to generate four additional databases, one for each DoV. As the numbers of participants in the DoV groups was different, we matched the group size of each DoV to the smaller DoV group to reduce bias associated with differences in the sample number between datasets and the overestimation of individuals within one category.[61] Following this criterion, each one of the four DoV databases comprised 8,070 individuals, of which 4,035 individuals acknowledged one particular type of DoV and 4,035 control individuals denied violence associated with this specific DoV (Table 1 and Figures 1A and 1B). A single group of control subjects was used for comparison with the four DoV. Also, from the group of 4,035 controls, we additionally selected a pseudorandom sample of 2,117 individuals who were age, gender, and education matched with the individuals who acknowledged the four DoVs (global violence dataset).

Deep-learning neural network procedures

DNN allows a computational model consisting of multiple processing levels that learn data representations with multiple levels of abstraction.[59] This procedure has demonstrated high discrimination power in comparison with other machine learning approaches.[59] Thus, using a DNN implemented in Python (TensorFlow library),[131] we classified the participants based on different types of DoV. We initially included all PPVs of the database (n = 162). Then, after performing machine learning feature selection, we included only the selected relevant features (n = 20) in a second DNN to assess to what extent the classification of DoV is preserved and optimized with a selected number of features.

Features were normalized to a range of values between −1 and 1, so they all had a similar scale to avoid saturation on the activation function and make the gradient descent converge faster. The dataset was then split into training and test sets, which comprised 80% and 20% of the data, respectively (Figure 1C). The latter group was used to measure performance during the training process in every iteration. A similar procedure was used in the second deep-learning analysis ran with the selected relevant features (n = 20).

The DNN architecture was a fully connected neural network. Its major advantage is that no assumptions on the input are made. This neural network consisted of four hidden layers (Figure 1B) that contained the same number of nodes as the number of features included in the analysis (162 in the first analysis and 20 after feature selection). A non-linear sigmoid function was applied as the activation function on each hidden layer.

An Adam optimizer was used because it is an efficient extension of a stochastic gradient descent optimizer.[132] Individual adaptive learning rates for each node are updated based on average estimates of the first and second moments (mean and uncentered variance, respectively) of the gradients as learning unfolds. The algorithm's hyperparameters include learning rate, decay rate for the moving averages (β_1 and β_2), and ϵ (a small value that is fixed to avoid division by zero). Learning rates for the models were initially set at a standard value of 0.01[132] and decreased if the model failed to achieve a low error rate (<6%).[59,133] Decay rates and ϵ were initialized to default values ($\beta_1 = 0.9$ and $\beta_2 = 0.999$, $\epsilon = 10^{-8}$).[132]

Parameter initialization included assigning random values between 0 and 1 to the weights and zero values to the biases. However, for all datasets (global violence and the databases for each DoV) with all features and after feature selection, Xavier initialization[134] was applied. Parameters were initialized as random draws from a truncated normal distribution with mean 0 and standard deviation following Equation 1:

$$\sigma = \sqrt{\frac{2}{n}}, \qquad \text{(Equation 1)}$$

where n is the number of nodes in each layer. This procedure reaches a minimum of the cost function faster and more efficiently by keeping the variance constant from layer to layer. The weights were still random, but positive and negative values close to 0 were assigned to produce outputs that followed a similar distribution across all nodes. Table S4 shows the initialization of the learning rates and the initialization values of the weights for each dataset.

The cost function used to measure the error between the neural network's output and the actual target was the mean squared error, computed as shown in Equation 2:

$$MSE = \frac{1}{2}\sum_{i}^{n}\left(y_i - \hat{y}_i\right)^2, \quad \text{(Equation 2)}$$

where MSE is the mean squared error, y_i is the actual target, $\hat{y} = Sigmoid(x_i \times w_i)$ is the output, where x_i is the sample, and w_i is the weight. This measure reflects the average magnitude of the error, irrespective of its direction. Because the difference between the observed and the predicted values is squared, it heavily penalizes predicted values, which are very different from the observed ones.

As a regularization procedure for avoiding overfitting, a dropout approach[135] was employed in the fourth hidden layer with a keep probability of 0.5. The optimization procedure was iterated until the minimum error on the training set and the maximum accuracy on the validation set (the number of observations that were correctly classified) were reached (Figures 1A–1C and Table S4).

Progressive feature elimination procedure

To assess which features were the most relevant to the classification, we performed a PFE analysis.[136] These analyses were performed with a non-linear classifier, specifically an RFC implemented in Python's scikit-learn package,[136] with a fixed number of trees (1,000) and the recommended number of features (P) in each split, where P is the square root of the total number of features.[136] Thus, we used an RFC that systematically assessed ordinal and continuous variables.[137,138] Nominal variables without intrinsic ordering were not included in analyses, as they affect RFC performance.[137–139] The RFC iteratively calculates for each variable and split point the best fit for classification. In both ordinal and continuous variables, this procedure's result is a binary split, which is assessed using the same function. First, we left out 20% of the sample to measure the performance of the final selection. The remaining 80% was split into 30 stratified (balanced) folds. We trained 30 RFCs using 29 folds, leaving out one of them (validation fold) each time (Figure 1D1), following a similar approach used in a previous study.[136]

Random forest, feature selection procedure

We address a classification problem where the input is a set of N vectors in a D-dimensional space, where each dimension is a PPV. We consider a two-class classification problem: non-violent (nV) versus a certain DoV.

So, given a training set composed of N vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k, \ldots \mathbf{x}_N]$, where $\mathbf{x}_k \in R^D$, together with a corresponding vector of labels $\mathbf{y} = [y_1, y_2, \ldots y_k, \ldots y_N]$, where $y_k \in [nV,DoV]$, we build an RFC so that for any new sample $\mathbf{x}^* \in R^D$ we can predict a $y^* \in [nV,DoV]$.

To assess which features were the most relevant to the classification task, we generated a feature ranking and selected a number of top-ranked features to optimize the classifier performance. RFC intrinsically generates a feature ranking based on Gini impurity: a decision-tree measure that chooses the optimal threshold to split a certain feature. Each tree in the forest can compute how much the weighted impurity decreases due to the splits for each feature; then the impurity decrease from each feature can be averaged over the forest.[140–142]

Considering a feature $f \in R$ for which we have N samples tagged by $y_k \in [nV,DoV]$, where p_{nV} is the proportion as samples tagged as non-violent and $n_{DoV} = 1 - p_{nV}$; the Gini impurity index is shown in Equation 3:

$$GINx(f) = \sum_{k \in [nV,sdV]} p_k(1 - p_k). \quad \text{(Equation 3)}$$

Then the decrease in the Gini impurity in node n due to a split over feature f is shown in Equation 4:

$$GINx(f,n) = GINx(f) - g_L(n) \times GINx_L(f) - g_R(n) \times GINx_R(f), \quad \text{(Equation 4)}$$

where $g_L(n)$ and $g_R(n)$ are the proportion of samples at the left and right of the split in the node n. $GINx_L$ and $GINx_R$ are the impurity measured over the samples at left and right of the split.

Thus, the computes for the whole forest are shown in Equation 5:

$$\underline{\Delta GINx}(f) = \frac{1}{|n(f)|}\sum_{n \in n(f)} \Delta GINx(f,n), \quad \text{(Equation 5)}$$

where n(f) is the set of nodes in the whole forest in which the feature is involved. Finally, the average is normalized, so the sum of the feature importance is equal to 1 (see Supplemental information S2).

Each trained RFC ranked the features according to their importance index. Such importance reflects how much the average Gini impurity index decreased in the forest due to its use as a node in a tree. We used this ranking to progressively eliminate the features one by one, removing, at each step, the feature with the lowest importance (Figure 1D2). In each step, a new RFC was trained over the feature-reduced dataset, and its performance (accuracy) was assessed on the validation fold.

Using the mean validation accuracy over the 30 folds, we visually defined the optimal number of features (N), i.e., using more than N features failed to improve the classifier's performance (Figure 1D3). We used 30 folds to have a desirable number of mean values to statistically model the distribution of accuracies.[143] We expected the Nth first features (the most important features) to maintain their position through the trials (Figure 1D5). We visually compared how much a feature moved in the ranking with respect to the median position to assess the stability of the ranking of the features across the 30 folds (Figure 1D6).

The final selection was made by training a new RFC on the full set, sorting the features by their importance, and selecting the N most important. Following a previous procedure,[136] the final performance was estimated on 20% of the sample that we kept apart. The accuracy of using the N selected features should not be worse than that of using the full set of features (Figure 1D4).

Post Hoc *analysis (recursive feature elimination)*

We performed a recursive most important feature elimination analysis to verify the information distribution between the selected features. Because this procedure was designed to evaluate to what extent an accurate classification was dependent on a single feature or a minimum set of relevant information, we performed it only for the global violence dataset. The dataset was iteratively trained on the whole training partition (80%), and its performance was measured on the test partition (20%). A cross-validation procedure was used on the training partition. To this end, we trained 30 RFCs using 29 folds, leaving out one of them (validation fold) each time. We used this procedure following a similar approach.[61,144] We used 30 folds to obtain a desirable number of mean values to statistically model the distribution of accuracies.[145] The classifier started the process using the full set of features, and for each iteration, the most important features were removed from the dataset until only one was left. Finally, the process generated a test accuracy curve that showed how much the classifier's performance decreased due to feature removal. We expected to obtain a smooth descending curve (as shown in Figure 2B4), which means that information is not concentrated in a few features but is distributed among a selected set. A single train-test procedure was used at each step (80% for training and 20% for test) to assess the importance of the features in each iteration (Figure 1D).

We used confusion matrices to improve the visualization of the performance of the PFE in categorizing each DoV. Each row of the matrix represents instances in the actual class (violence versus non-violence), while each column represents instances in the predicted class (violence versus non-violence). We estimated two confusion matrices for each DoV to compare the performance of the PFE in predicting violence using the full set of predictors (PPV = 162 features) versus the selected predictors (PPV = 20 features) (Figures S1A–S1E).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.patter.2020.100176.

ACKNOWLEDGMENTS

A.I. is partially supported by CONICET, ANID/FONDAP (15150012), PICT (2017-1818 and 2017-1820), the Inter-American Development Bank, Alzheimer's Association GBHI ALZ UK-20-639295, Alzheimer's Association SG-20-725707, NIH/NIA R01 AG057234, Tau Consortium, and the Global Brain Health Institute. This work was supported by the Colombian Agency for Reincorporation and Normalization (*Agencia para la Reincorporación y la Normalización*, ARN). The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions.

AUTHOR CONTRIBUTIONS

Conceptualization, A.I., H.S.-G., S.B., and D.M.A.; Methodology, H.S.-G., S.B., G.P., P.D.K.,G.M., and E.H.; Software, G.P., P.D.K., G.M., and E.H.; Validation, G.P., P.D.K., E.H., and H.S.-G.; Formal Analysis, G.P., P.D.K., G.M., E.H., S.B., and H.S.-G.; Investigation, D.M.A., J.G.Z., and the ARN; Resources, D.M.A. and J.G.Z.; Data Curation, H.S.-G., S.B., and A.N.; Writing – Original Draft, H.S.-G. and A.I.; Writing – Review & Editing, A.I., S.B., W.C.H., D.M., J.L., J.D., and H.S.-G.; Supervision, A.I.; Project Administration, H.S.-G., D.M.A., and A.I. All authors participated sufficiently in the work and approved the final version of the manuscript for submission.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 3, 2020
Revised: October 21, 2020
Accepted: November 30, 2020
Published: December 23, 2020

REFERENCES

1. Glowacki, L., Isakov, A., Wrangham, R.W., McDermott, R., Fowler, J.H., and Christakis, N.A. (2016). Formation of raiding parties for intergroup violence is mediated by social network structure. Proc. Natl. Acad. Sci. U S A *113*, 12114–12119.
2. Gomez, J.M., Verdu, M., Gonzalez-Megias, A., and Mendez, M. (2016). The phylogenetic roots of human lethal violence. Nature *538*, 233–237.
3. Heise, L.L., and Kotsadam, A. (2015). Cross-national and multilevel correlates of partner violence: an analysis of data from population-based surveys. Lancet Glob. Health *3*, e332–e340.
4. Wagner, Z., Heft-Neal, S., Wise, P.H., Black, R.E., Burke, M., Boerma, T., Bhutta, Z.A., and Bendavid, E. (2019). Women and children living in areas of armed conflict in Africa: a geospatial analysis of mortality and orphanhood. Lancet Glob. Health *7*, e1622–e1631.
5. Spinney, L. (2012). Human cycles: History as science. Nature *488*, 24–26.
6. Earl, J. (2013). Age and Social Movements. The Wiley-Blackwell Encyclopedia of Social and Political Movements (Wiley-Blackwell).
7. Barber, B.K., McNeely, C., Spellings, C., et al. (2012). Role of political factors in wellbeing and quality of life during long-term constraints and conflict: an initial study. Lancet *380*, https://doi.org/10.1016/S0140-6736(13)60199-3.
8. Lim, M., Metzler, R., and Bar-Yam, Y. (2007). Global pattern formation and ethnic/cultural violence. Science *317*, 1540–1544.
9. Fazel, S., Smith, E.N., Chang, Z., and Geddes, J.R. (2018). Risk factors for interpersonal violence: an umbrella review of meta-analyses. Br. J. Psychiatry *213*, 609–614.
10. Nestor, P.G. (2002). Mental disorder and violence: personality dimensions and clinical features. Am. J. Psychiatry *159*, 1973–1978.
11. Zagar, R.J., Busch, K.G., Grove, W.M., and Hughes, J.R. (2009). Can violent (re)offense be predicted? Review of the role of the clinician and use of actuarial tests in light of new data. Psychol. Rep. *104*, 247–277.
12. Zagar, R.J., and Grove, W.M. (2010). Violence risk appraisal of male and female youth, adults, and individuals. Psychol. Rep. *107*, 983–1009.
13. Bohorquez, J.C., Gourley, S., Dixon, A.R., Spagat, M., and Johnson, N.F. (2009). Common ecology quantifies human insurgency. Nature *462*, 911.
14. International, A. (2018). Amnesty International Report 2017/18. https://www.amnesty.org/en/countries/americas/colombia/report-colombia/.
15. Baez, S., Herrera, E., García, A.M., Manes, F., Young, L., and Ibáñez, A. (2017). Outcome-oriented moral evaluation in terrorists. Nat. Hum. Behav. *1*, 0118, https://doi.org/10.1038/s41562-017-0118.
16. Elbert, T., Schauer, M., and Moran, J.K. (2018). Two pedals drive the bicycle of violence: reactive and appetitive aggression. Curr. Opin. Psychol. *19*, 135–138.
17. Chester, D.S., and DeWall, C.N. (2016). The pleasure of revenge: retaliatory aggression arises from a neural imbalance toward reward. Soc. Cogn. Affect Neurosci. *11*, 1173–1182.
18. Capaldi, D.M., Knoble, N.B., Shortt, J.W., and Kim, H.K. (2012). A systematic review of risk factors for intimate partner violence. Partner Abuse *3*, 231–280.
19. Douglas, K.S., and Skeem, J.L. (2005). Violence risk assessment: getting specific about being dynamic. Psychol. Public Pol. L. *11*, 347.
20. Favril, L., Yu, R., Hawton, K., and Fazel, S. (2020). Risk factors for self-harm in prison: a systematic review and meta-analysis. Lancet Psychiatry *7*, 682–691.
21. Fazel, S., Chang, Z., Fanshawe, T., Långström, N., Lichtenstein, P., Larsson, H., and Mallett, S. (2016). Prediction of violent reoffending on release from prison: derivation and external validation of a scalable tool. Lancet Psychiatry *3*, 535–543.
22. Fazel, S., Singh, J.P., Doll, H., and Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. BMJ *345*, e4692.
23. Ramesh, T., Igoumenou, A., Vazquez Montes, M., and Fazel, S. (2018). Use of risk assessment instruments to predict violence in forensic psychiatric hospitals: a systematic review and meta-analysis. Eur. Psychiatry *52*, 47–53.
24. Cornaggia, C.M., Beghi, M., Pavone, F., and Barale, F. (2011). Aggression in psychiatry wards: a systematic review. Psychiatry Res. *189*, 10–20.
25. Hawton, K., Linsell, L., Adeniji, T., Sariaslan, A., and Fazel, S. (2014). Self-harm in prisons in England and Wales: an epidemiological study of prevalence, risk factors, clustering, and subsequent suicide. Lancet *383*, 1147–1154.
26. Papadopoulos, C., Ross, J., Stewart, D., Dack, C., James, K., and Bowers, L. (2012). The antecedents of violence and aggression within psychiatric in-patient settings. Acta Psychiatr. Scand. *125*, 425–439.
27. Pickard, H., and Fazel, S. (2013). Substance abuse as a risk factor for violence in mental illness: some implications for forensic psychiatric practice and clinical ethics. Curr. Opin. Psychiatry *26*, 349.
28. Witt, K., Van Dorn, R., and Fazel, S. (2013). Risk factors for violence in psychosis: systematic review and meta-regression analysis of 110 studies. PLoS One *8*, e55942.
29. Johnson, J.G., Cohen, P., Smailes, E., Kasen, S., Oldham, J.M., Skodol, A.E., and Brook, J.S. (2000). Adolescent personality disorders associated with violence and criminal behavior during adolescence and early adulthood. Am. J. Psychiatry *157*, 1406–1412.
30. Fossati, A., Barratt, E.S., Borroni, S., Villa, D., Grazioli, F., and Maffei, C. (2007). Impulsivity, aggressiveness, and DSM-IV personality disorders. Psychiatry Res. *149*, 157–167.
31. Huguelet, P., and Perroud, N. (2010). Is there a link between mental disorder and violence? Arch. Gen. Psychiatry *67*, 540.
32. Hughes, K., Bellis, M.A., Hardcastle, K.A., Sethi, D., Butchart, A., Mikton, C., Jones, L., and Dunne, M.P. (2017). The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health *2*, e356–e366.
33. Jakobsen, I.S., Fergusson, D., and Horwood, J.L. (2012). Early conduct problems, school achievement and later crime: findings from a 30-year longitudinal study. New Zealand J. Educ. Stud. *47*, 123.
34. Mooijman, M., Hoover, J., Lin, Y., Ji, H., and Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. Nat. Hum. Behav. *2*, 389–396.
35. Esteban, J., Mayoral, L., and Ray, D. (2012). Ethnicity and conflict: theory and facts. Science *336*, 858–865.
36. Augsburger, M., Meyer-Parlapanis, D., Bambonye, M., Elbert, T., and Crombach, A. (2015). Appetitive aggression and adverse childhood experiences shape violent behavior in females formerly associated with combat. Front. Psychol. *6*, 1756.

37. Baez, S., Santamaría-García, H., and Ibáñez, A. (2019). Disarming ex-combatants' minds: toward situated reintegration process in post-conflict Colombia. Front. Psychol. *10*, 73.

38. Kaplan, O., and Nussio, E. (2018). Explaining recidivism of ex-combatants in Colombia. J. Conflict Resolut. *62*, 64–93.

39. Köbach, A., Nandi, C., Crombach, A., Bambonyé, M., Westner, B., and Elbert, T. (2015). Violent offending promotes appetitive aggression rather than Posttraumatic stress—a replication study with Burundian ex-combatants. Front. Psychol. *6*, 1755.

40. Köbach, A., Schaal, S., and Elbert, T. (2015). Combat high or traumatic stress: violent offending is associated with appetitive aggression but not with symptoms of traumatic stress. Front. Psychol. *5*, https://doi.org/10.3389/fpsyg.2014.01518.

41. Nandi, C., Crombach, A., Bambonye, M., Elbert, T., and Weierstall, R. (2015). Predictors of posttraumatic stress and appetitive aggression in active soldiers and former combatants. Eur. J. Psychotraumatol. *6*, 26553.

42. Elbogen, E.B., Johnson, S.C., Newton, V.M., Fuller, S., Wagner, H.R., and Beckham, J.C. (2013). Self-report and longitudinal predictors of violence in Iraq and Afghanistan war era veterans. J. Nerv Ment. Dis. *201*, 872.

43. Maguen, S., Metzler, T.J., Bosch, J., Marmar, C.R., Knight, S.J., and Neylan, T.C. (2012). Killing in combat may be independently associated with suicidal ideation. Depress. Anxiety *29*, 918–923.

44. Fulu, E., Miedema, S., Roselli, T., McCook, S., Chan, K.L., Haardörfer, R., and Jewkes, R. (2017). Pathways between childhood trauma, intimate partner violence, and harsh parenting: findings from the UN Multi-country Study on Men and Violence in Asia and the Pacific. Lancet Glob. Health *5*, e512–e522.

45. Bzdok, D., Krzywinski, M., and Altman, N. (2017). Machine Learning: A primer (Nature Methods).

46. Bzdok, D., Altman, N., and Krzywinski, M. (2018). Statistics versus machine learning. Nat. Methods *15*, 233–234.

47. Blair, R.A., Blattman, C., and Hartman, A. (2017). Predicting local violence. J. Peace Res. *54*, 298–312.

48. Fernandes, B.S., Williams, L.M., Steiner, J., Leboyer, M., Carvalho, A.F., and Berk, M. (2017). The new field of 'precision psychiatry'. BMC Med. *15*, 80–87.

49. McIntosh, A.M., Stewart, R., John, A., Smith, D.J., Davis, K., Sudlow, C., Corvin, A., Nicodemus, K.K., Kingdon, D., Hassan, L., et al. (2016). Data science for mental health: a UK perspective on a global challenge. Lancet Psychiatry *3*, 993–998.

50. Menger, V., Scheepers, F., and Spruit, M. (2018). Comparing deep learning and classical machine learning approaches for predicting inpatient violence incidents from clinical text. Appl. Sci. *8*, 981.

51. Menger, V., Spruit, M., van Est, R., Nap, E., and Scheepers, F. (2019). Machine learning approach to inpatient violence risk assessment using routinely collected clinical notes in electronic health records. JAMA Netw. Open *2*, e196709.

52. Velupillai, S., Hadlaczky, G., Baca-Garcia, E., Gorrell, G.M., Werbeloff, N., Nguyen, D., Patel, R., Leightley, D., Downs, J., Hotopf, M., and Dutta, R. (2019). Risk assessment tools and data-driven approaches for predicting and preventing suicidal behavior. Front. Psychiatry *10*, 36.

53. Bzdok, D., Altman, N., and Krzywinski, M. (2018). Points of significance: statistics versus machine learning. Nat. Methods *15*, 233.

54. Reardon, S. (2018). Colombia: after the violence. Nature https://www.nature.com/immersive/d41586-018-04976-7/index.html.

55. Hulme, P.A. (2004). Retrospective measurement of childhood sexual abuse: a review of instruments. Child. Maltreat. *9*, 201–217.

56. SanSegundo, M.S., Ferrer-Cascales, R., Bellido, J.H., Bravo, M.P., Oltra-Cucarella, J., and Kennedy, H.G. (2018). Prediction of violence, suicide behaviors and suicide ideation in a sample of institutionalized offenders with Schizophrenia and other psychosis. Front. Psychol. *9*, 1385.

57. Snowden, R.J., Gray, N.S., Taylor, J., and MacCulloch, M.J. (2007). Actuarial prediction of violent recidivism in mentally disordered offenders. Psychol. Med. *37*, 1539–1549.

58. Yang, M., Wong, S.C., and Coid, J. (2010). The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools. Psychol. Bull. *136*, 740.

59. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436.

60. Diaz-Uriarte, R., and Alvarez de Andres, S. (2006). Gene selection and classification of microarray data using random forest. BMC Bioinformatics *7*, 3.

61. Donnelly-Kehoe, P.A., Pascariello, G.O., García, A.M., Hodges, J.R., Miller, B., Rosen, H., Manes, F., Landin-Romero, R., Matallana, D., Serrano, C., et al. (2019). Robust automated computational approach for classifying frontotemporal neurodegeneration: multimodal/multicenter neuroimaging. Alzheimers Dement (Amst) *11*, 588–598, https://doi.org/10.1016/j.dadm.2019.06.002.

62. Fazel, S., Lichtenstein, P., Grann, M., Goodwin, G.M., and Langstrom, N. (2010). Bipolar disorder and violent crime: new evidence from population-based longitudinal studies and systematic review. Arch. Gen. Psychiatry *67*, 931–938.

63. Cox, D.R. (2006). Principles of Statistical Inference (Cambridge University Press).

64. Bostrom, N., and Yudkowsky, E. (2014). The ethics of artificial intelligence. In The Cambridge Handbook of Artificial Intelligence, *1*, K. Frankish and W.M. Ramsey, eds., K. Frankish, ed. The Cambridge Handbook of Artificial Intelligence *(Cambridge University Press), pp. 316–334*.

65. Goin, D.E., Rudolph, K.E., and Ahern, J. (2018). Predictors of firearm violence in urban communities: a machine-learning approach. Health Place *51*, 61–67.

66. Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. Am. Stat. *63*, 308–319.

67. Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: concerns and ways forward. PLoS One *13*, e0194889.

68. Poldrack, R.A., Huckins, G., and Varoquaux, G. (2020). Establishment of best practices for evidence for prediction: a review. JAMA Psychiatry *77*, 534–540.

69. Chen, Z., He, N., Huang, Y., Qin, W.T., Liu, X., and Li, L. (2018). Integration of A Deep learning classifier with A random forest approach for predicting malonylation sites. Genomics Proteomics Bioinformatics *16*, 451–459.

70. Montavon, G., Samek, W., and Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. Digital Signal. Process. *73*, 1–15.

71. Hutzler, F. (2014). Reverse inference is not a fallacy per se: cognitive processes can be inferred from functional imaging data. NeuroImage *84*, 1061–1069.

72. Abdollahi-Arpanahi, R., Gianola, D., and Peñagaricano, F. (2020). Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. Genet. Selection Evol. *52*, 12.

73. Kong, Y., and Yu, T. (2018). A deep neural network model using random forest to extract feature representation for gene expression data classification. Sci. Rep. *8*, 16477.

74. Nicholls, H.L., John, C.R., Watson, D.S., Munroe, P.B., Barnes, M.R., and Cabrera, C.P. (2020). Reaching the end-game for GWAS: machine learning approaches for the prioritization of complex disease loci. Front. Genet. *11*, 350.

75. Rusk, N. (2016). Deep learning. Nat. Methods *13*, 35.

76. Zhou, J., and Troyanskaya, O.G. (2015). Predicting effects of noncoding variants with deep learning–based sequence model. Nat. Methods *12*, 931–934.

77. Wang, H., Yang, F., and Luo, Z. (2016). An experimental study of the intrinsic stability of random forest variable importance measures. BMC Bioinformatics *17*, 60.

78. Kautzky, A., et al. (2017). Refining prediction in treatment-resistant depression: results of machine learning analyses in the TRD III sample. J. Clin. Psychiatry *79*, 16m11385.

79. Kharoubi, R., Oualkacha, K., and Mkhadri, A. (2019). The cluster correlation-network support vector machine for high-dimensional binary classification. J. Stat. Comput. Simulation *89*, 1020–1043.

80. Obermeyer, Z., and Emanuel, E.J. (2016). Predicting the future - big data, machine learning, and clinical medicine. N. Engl. J. Med. *375*, 1216–1219.

81. Nemeroff, C.B. (2016). Paradise lost: the neurobiological and clinical consequences of child abuse and neglect. Neuron *89*, 892–909.

82. Sariaslan, A., Larsson, H., D'Onofrio, B., Långström, N., and Lichtenstein, P. (2014). Childhood family income, adolescent violent criminality and substance misuse: quasi-experimental total population study. Br. J. Psychiatry *205*, 286–290.

83. Swann, W.B., Jr., Gomez, A., Huici, C., Morales, J.F., and Hixon, J.G. (2010). Identity fusion and self-sacrifice: arousal as a catalyst of pro-group fighting, dying, and helping behavior. J. Pers Soc. Psychol. *99*, 824–841.

84. Regan, P.M., and Norton, D. (2005). Greed, grievance, and mobilization in civil wars. J. Conflict Resol. *49*, 319–336.

85. Thrasher, J., and Handfield, T. (2018). Honor and violence : an account of feuds, duels, and honor killings. Hum. Nat. *29*, 371–389.

86. Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? Science *324*, 1293–1298.

87. Fazel, S., Langstrom, N., Hjern, A., Grann, M., and Lichtenstein, P. (2009). Schizophrenia, substance abuse, and violent crime. JAMA *301*, 2016–2023.

88. Fazel, S., Wolf, A., Chang, Z., Larsson, H., Goodwin, G.M., and Lichtenstein, P. (2015). Depression and violence: a Swedish population study. Lancet Psychiatry *2*, 224–232.

89. Berman, M.E., Fallon, A.E., and Coccaro, E.F. (1998). The relationship between personality psychopathology and aggressive behavior in research volunteers. J. Abnorm Psychol. *107*, 651–658.

90. Côté, S., Piff, P.K., and Willer, R. (2013). For whom do the ends justify the means? Social class and utilitarian moral judgment. J. Personal. Soc. Psychol. *104*, 490–503.

91. Everett, J.A.C., and Kahane, G. (2020). Switching tracks? Towards a multidimensional model of utilitarian psychology. Trends Cogn. Sci. *24*, 124–134.

92. Fok, L.Y., Payne, D.M., and Corey, C.M. (2016). Cultural values, utilitarian orientation, and ethical decision making: a comparison of U.S. And Puerto Rican professionals. J. Bus. Ethics *134*, 263–279.

93. Zhang, L., Li, Z., Wu, X., and Zhang, Z. (2017). Why people with more emotion regulation difficulties made a more deontological judgment: the role of deontological inclinations. Front. Psychol. *8*, 2095.

94. Garofalo, C., Holden, C.J., Zeigler-Hill, V., and Velotti, P. (2016). Understanding the connection between self-esteem and aggression: the mediating role of emotion dysregulation. Aggress Behav. *42*, 3–15.

95. Veronese, G., Pepe, A., Jaradah, A., Al Muranak, F., and Hamdouna, H. (2017). Modelling life satisfaction and adjustment to trauma in children exposed to ongoing military violence: an exploratory study in Palestine. Child. Abuse Negl. *63*, 61–72.

96. Davidson, R.J., Putnam, K.M., and Larson, C.L. (2000). Dysfunction in the neural circuitry of emotion regulation–a possible prelude to violence. Science *289*, 591–594.

97. Toh, T.S., Dondelinger, F., and Wang, D. (2019). Looking beyond the hype: applied AI and machine learning in translational medicine. EBioMedicine *47*, 607–615.

98. Skeem, J., and Lowenkamp, C. (2020). Using algorithms to address trade-offs inherent in predicting recidivism. Behav. Sci. Law. https://doi.org/10.1002/bsl.2465.

99. Tomsett, R., Preece, A., Braines, D., Cerutti, F., Chakraborty, S., Srivastava, M., Pearson, G., and Kaplan, L. (2020). Rapid trust calibration through interpretable and uncertainty-aware AI. Patterns (N Y) *1*, 100049.

100. Jolliffe, D., Farrington, D.P., Hawkins, J.D., Catalano, R.F., Hill, K.G., and Kosterman, R. (2003). Predictive, concurrent, prospective and retrospective validity of self-reported delinquency. Crim Behav. Ment. Health *13*, 179–197.

101. Moffitt, T.E., Caspi, A., Taylor, A., Kokaua, J., Milne, B.J., Polanczyk, G., and Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychol. Med. *40*, 899.

102. Barros, A.J., Ronsmans, C., Axelson, H., Loaiza, E., Bertoldi, A.D., França, G.V., Bryce, J., Boerma, J.T., and Victora, C.G. (2012). Equity in maternal, newborn, and child health interventions in Countdown to 2015: a retrospective review of survey data from 54 countries. Lancet *379*, 1225–1233.

103. Santamaría-García, H., Baez, S., Gómez, C., Rodríguez-Villagra, O., Huepe, D., Portela, M., et al. (2020). The role of social cognition skills and social determinants of health in predicting symptoms of mental illness. Transl. Psychiatry *10*, 165–186, https://doi.org/10.1038/s41398-020-0852-4.

104. Checkel, J.T. (2017). Socialization and violence: Introduction and framework (SAGE Publications Sage UK), pp. 592–605.

105. Gates, S. (2017). Membership matters: Coerced recruits and rebel allegianceMembership matters. J. Peace Res. *54*, 674–686.

106. Tharwat, A. (2018). Classification assessment methods. Appl. Comput. Inform. *Vol. ahead-of-print:No. ahead-of-print*, https://doi.org/10.1016/j.aci.2018.08.003.

107. Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., and Dehmer, M. (2020). An introductory review of deep learning for prediction models with big data. Front. Artif. Intell. *3*, 0–12, https://doi.org/10.3389/frai.2020.00004.

108. Manrique Rueda, G. (2018). Working in violence: moral narratives of paramilitaries in Colombia. Theor. Criminology *24*, 370–386.

109. Ugarriza, J.E., and Craig, M.J. (2012). The relevance of ideology to contemporary armed conflicts: a quantitative analysis of former combatants in Colombia. J. Conflict Resol. *57*, 445–477.

110. Balash, J., and Falkenbach, D.M. (2018). The ends justify the meanness: an investigation of psychopathic traits and utilitarian moral endorsement. Personal. Individual Differ. *127*, 127–132.

111. Bushman, B.J., Baumeister, R.F., and Phillips, C.M. (2001). Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding. J. Pers Soc. Psychol. *81*, 17–32.

112. Jolliffe, I.T., and Cadima, J. (2016). Principal component analysis: a review and recent developments. Philos. Trans. A Math. Phys. Eng. Sci. *374*, 20150202.

113. Costa, P.S., Santos, N.C., Cunha, P., Palha, J.A., and Sousa, N. (2013). The use of bayesian latent class cluster models to classify patterns of cognitive performance in Healthy ageing. PLoS One *8*, e71940.

114. Selzam, S., Coleman, J.R.I., Caspi, A., Moffitt, T.E., and Plomin, R. (2018). A polygenic p factor for major psychiatric disorders. Transl. Psychiatry *8*, 205–209.

115. Atran, S., and Ginges, J. (2012). Religious and sacred imperatives in human conflict. Science *336*, 855–857.

116. Coccaro, E.F. (2012). Intermittent explosive disorder as a disorder of impulsive aggression for DSM-5. Am. J. Psychiatry *169*, 577–588.

117. Roberts, A., Coid, J., King, R., Murphy, R., Turner, J., Bowden-Jones, H., Du Preez, K.P., and Landon, J. (2016). Gambling and violence in a nationally representative sample of UK men. Addiction *111*, 2196–2207.

118. Kessler, R.C., and Üstün, T.B. (2004). The world mental health (WMH) survey initiative version of the world health organization (WHO)

composite international diagnostic interview (CIDI). Int. J. Methods Psychiatr. Res. *13*, 93–121.

119. Kessler, R.C., Heeringa, S., Lakoma, M.D., Petukhova, M., Rupp, A.E., Schoenbaum, M., Wang, P.S., and Zaslavsky, A.M. (2008). Individual and societal effects of mental disorders on earnings in the United States: results from the national comorbidity survey replication. Am. J. Psychiatry *165*, 703–711.
120. Tamayo Martinez, N., Rincón Rodríguez, C.J., de Santacruz, C., Bautista Bautista, N., Collazos, J., and Gómez-Restrepo, C. (2016). Mental problems, mood and anxiety disorders in the population displaced by violence in Colombia; results of the National Mental Health Survey 2015. Rev. Colomb. Psiquiatr. *45* (*Suppl 1*), 113–118.
121. Gomez-Restrepo, C., Tamayo-Martínez, N., Buitrago, G., Guarnizo-Herreño, C.C., Garzón-Orjuela, N., Eslava-Schmalbach, J., de Vries, E., Rengifo, H., Rodríguez, A., and Rincón, C.J. (2016). Violence due to armed conflict and prevalence of mood disorders, anxiety and mental problems in the Colombian adult population. Rev. Colomb. Psiquiatr. *45* (*Suppl 1*), 147–153.
122. van der Westhuizen, C., Wyatt, G., Williams, J.K., Stein, D.J., and Sorsdahl, K. (2016). Validation of the self reporting questionnaire 20-item (SRQ-20) for use in a low- and middle-income country emergency centre setting. Int. J. Ment. Health Addict. *14*, 37–48.
123. Berg, J.M., Kennedy, J.C., Dunlop, B.W., Ramirez, C.L., Stewart, L.M., Nemeroff, C.B., Mayberg, H.S., and Craighead, W.E. (2017). The structure of personality disorders within a depressed sample: implications for personalizing treatment. Pers Med. Psychiatr. *1-2*, 59–64.
124. Lambe, S., Hamilton-Giachritsis, C., Garner, E., and Walker, J. (2018). The role of narcissism in aggression and violence: a systematic review. Trauma Violence Abuse *19*, 209–230.
125. Smith, S.D., Lynch, R.J., Stephens, H.F., and Kistner, J.A. (2015). Self-perceptions and their prediction of aggression in male juvenile offenders. Child. Psychiatry Hum. Dev. *46*, 609–621.
126. MacDonald, J.M., Piquero, A.R., Valois, R.F., and Zullig, K.J. (2005). The relationship between life satisfaction, risk-taking behaviors, and youth violence. J. Interpers. Violence *20*, 1495–1518.
127. Fishbain, D.A., Cutler, R.B., Rosomoff, H.L., and Steele-Rosomoff, R. (2018). Risk for violent behavior in patients with chronic pain: evaluation and management in the pain facility setting. Pain Med. *1*, 140–155.
128. Campbell, J.C. (2002). Health consequences of intimate partner violence. Lancet *359*, 1331–1336.
129. Fredrickson, B.L., Grewen, K.M., Algoe, S.B., Firestine, A.M., Arevalo, J.M., Ma, J., and Cole, S.W. (2015). Psychological well-being and the human conserved transcriptional response to adversity. PLoS one *10*, e0121839.
130. Manor, O., Matthews, S., and Power, C. (2000). Dichotomous or categorical response? Analysing self-rated health and lifetime social class. Int. J. Epidemiol. *29*, 149–157.
131. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems (Tensorflow.org).
132. Kingma, D.P. and Ba, J.L. in Proceedings of the 3rd International Conference on Learning Representations (ICLR).
133. Goodfellow, I.B., Yoshua, and Courville, A. (2016). Deep Learning (MIT Press).
134. Glorot, X., and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. J. Machine Learn. Res. Proc. Track *9*, 249–256.
135. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. J. Machine Learn. Res. *15*, 1929–1958.
136. Donnelly-Kehoe, P.A., Pascariello, G.O., and Gomez, J.C. (2018). Looking for Alzheimer's Disease morphometric signatures using machine learning techniques. J. Neurosci. Methods *302*, 24–34.
137. Cutler, A., Cutler, D.R., and Stevens, J.R. (2012). Ensemble Machine Learning (Springer), pp. 157–175.
138. Wright, M.N., and König, I.R. (2019). Splitting on categorical predictors in random forests. PeerJ *7*, e6339.
139. Altman, N., and Krzywinski, M. (2017). Ensemble methods: bagging and random forests. Nat. Methods *14*, 933–934.
140. Breiman, L. (2001). Random forests. Machine Learn. *45*, 5–32.
141. Breiman, L., Friedman, J., Stone, C.J., and Olshen, R.A. (1984). Classification and Regression Trees (CRC press).
142. Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. Machine Learn. *46*, 389–422.
143. Greenwood, J.A., and S.M.. (2012). Sample size required for estimating the standard deviation as a per cent of its true value. https://doi.org/10.1080/01621459.1950.10483356.
144. Kohavi, R. (1995) A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In Proceedings of the 14th International Joint Conference on Artificial Intelligence, Vol. 2, Montreal, 20-25 August 1995, 1137-1145.
145. Greenwood, J.A., and Sandomire, M.M. (1950). Sample size required for estimating the standard deviation as a per cent of its true value. J. Am. Stat. Assoc. *45*, 257–260.