

Informing Citizen Security Policy: An Evidence Gap on Policing Interventions

Protocol

Authors:

Sergio Britto Lima, Citizen Security Division, sbrittolima@iadb.org

Camilo Acosta, Knowledge and Learning Division, camiloac@iadb.org

Indira Porto Gutierrez, Knowledge and Learning Division, indirap@iadb.org

This version: May 22, 2025

Last updated: October 7, 2025

1. Background

Robust evidence and targeted knowledge are essential for enhancing the effectiveness and impact of development projects. The Inter-American Development Bank (IDB) Group, a key development partner in Latin America and the Caribbean (LAC), views the active promotion of mechanisms that encourage the generation and utilization of relevant knowledge at both portfolio and program levels as an institutional priority. As such, the IDB has incorporated the mandate to improve feedback between operations and knowledge into the IDB Group's Institutional Strategy 2024-2030.

To achieve this goal, the IDB is developing mechanisms that incentivize the generation and strategic use of impact evaluation-related knowledge that can be leveraged to increase overall development effectiveness of Bank-supported interventions, initiatives, and projects. This involves learning from both successful and unsuccessful interventions to foster the continuous improvement of the Bank's performance and impact capacity. Given the vast quantity of research and variations in quality, employing cutting-edge methods to synthesize research is more vital than ever. Under such context, Evidence Gap Maps (EGMs) rise as a useful tool, as they visually demonstrate areas where evidence is concentrated and where it is lacking, thereby guiding research and knowledge generation efforts, as well as allowing for agile

consultations that can better inform project design, based on the identification of interventions that have actually worked to generate significant improvements in a given set of problems and outcomes of interest.

In the case of the Citizen Security Division (IFD/CIS) line of work, policymakers and experts alike demand access to evidence regarding which programs, initiatives, or interventions have been effective to reduce crime and violence and improve overall security standards—as well as insights into those that have proven ineffective or even produced detrimental effects, contrary to their initial objectives and aspirations. Over the past few decades, the evidence-based citizen security movement has gained momentum, leading to a significant increase in the amount of evidence available to practitioners, policymakers, and the general public. Unfortunately, however, most of that academic production has been concentrated in the developed world, and more specifically, in Anglo-Saxon countries, such as the US, the UK, and, to a lesser extent, Australia and Canada.

To promote evidence-based security and justice policies in Latin America and the Caribbean, the IDB launched, in 2023, the “Security & Justice Evidence-Based Platform,” the first online repository that consolidates information about evidence-based security and justice practices in the two main languages of the LAC region — Spanish and Portuguese. As of September 2025, the platform had catalogued approximately 91 types of solutions¹ and 700 evaluated cases². Within the area of policing, specifically, the platform has identified 135 evaluated cases, organized into 20 types of solutions.

Given that the evidence generated worldwide in the field of citizen security varies significantly in terms of quality, sophistication, and scientific rigor, the IDB has undertaken efforts to refine the information sources included in the platform. It is in that context that IFD/CIS decided to take efforts towards the construction of the Division’s first-ever Evidence Gap Map.

Building on the existing *Security & Justice Evidence-Based Platform*, the EGM systematically synthesizes and visually presents the evidence,

¹ In the terms proposed by the Evidence-Based Platform, a “solution type” refers to categories that brings together a set of initiatives, practices or types of interventions that share the same strategic focus (“what” they intend to do and/or problems they intend to solve or mitigate) and the same forms of action and approach (“how” they intend to address problems and advance their purposes). Examples of solution types are: [“community policing”](#), [“hot spots policing”](#), and [“problem-oriented policing”](#), among others.

² In the terms proposed by the Evidence-Based Platform, a “evaluated case” consists of examples of practical application of each type of solution included in the platform. They correspond, in general, to specific initiatives or programs through which a set of activities are implemented, developed and executed in an articulated manner, seeking to mobilize a set of resources (physical, human, financial or technological) in favor of the realization of a common objective and purpose and the achievement of a set of clearly identified and specified results. Examples of “evaluated cases” are the [“High Operational Dedication Program”](#) (PADO, in the Spanish acronym), which was a case of hot spots policing application in Uruguay, and the [“Pelotas' Pact for Peace”](#)), which was a case of [focused-deterrence](#) application in southern Brazil.

identifying both covered areas and topics, as well as existing gaps in the current research landscape. As such, the EGM will highlight areas that require further investigation, guiding future research priorities while providing valuable insights to ultimately enhance the development of effective and evidence-based public policies and programs.

2. Study objectives and questions

2.1 Objectives

The main objective of the EGM is to contribute to enhancing overall development effectiveness in the field of citizen security by compiling, cataloging, organizing, and synthesizing state-of-the-art empirical evidence, while presenting it in a visually appealing, user-friendly format to facilitate its use by policymakers, practitioners, and academics. Additionally, as secondary objectives, the EGM aims to inform the IDB's own programmatic agenda in citizen security, utilizing cutting-edge evidence to guide strategic selectivity and enhance its overall impact. Furthermore, it seeks to inform the IDB's knowledge agenda by closing identified gaps in the Latin American and Caribbean empirical literature.

In a broader context, the EGM looks to promote scientific research, expand the culture of rigorous evaluation, and elevate the role of knowledge as a central input of the policymaking process in the field of citizen security and justice. As such, in the same way that the *Evidence-Based Platform* seeks to contribute to the systematic process of improving public policies, the EGM will constitute a critical tool for researchers, policymakers, and public servants to gather the most technical, recent, consistent, and rigorous evidence regarding the effectiveness of citizen security solutions and interventions led or supported by police agencies.

2.2 Research questions

The EGM is dedicated to establishing the extent and characteristics of the existing empirical literature on a wide range of policing initiatives and strategies, as well as identifying what it says regarding their effectiveness, considering the main developmental challenges related to the field of citizen security and justice at large, and to the work of police agencies, specifically. More specifically, this EGM was developed to answer the following questions:

- What is the extent and the characteristics of empirical evidence on citizen security interventions led or supported by police agencies in the LAC region, specifically, and elsewhere around the globe?
- What impact evaluations and systematic reviews exist that can inform the effectiveness of these types of interventions?

- What new or underexplored areas/topics within the field of policing and citizen security should be prioritized for primary research?
- What does the evidence say about the effectiveness of each program or intervention in achieving its intended outcomes?

3. Methods

3.1 Framework development and scope

The Intervention-Outcome (I-O) framework from this EGM leverages existing knowledge and information consolidated in the *Security & Justice Evidence-Based Platform*, which was the main source for building the EGM. In addition, another key inputs are the [Citizen Security and Justice Sectoral Framework \(SFD\)](#)³, which includes a series of interventions covered by the Bank in its operations and the most common outcomes studied in the literature, and IFD/CIS's own project portfolio, which relies on both internal and external evidence to substantiate the selection of interventions and solutions included in each loan operation/project.

Regarding the outcomes of interest, the variables included in the EGM were identified using the same list that served as the basis for developing the *Evidence-Based Platform* (available [here](#)). This list was generated while taking into consideration the different types of violence (especially “interpersonal” violence) established in the [World Health Organization \(WHO\) World Report on Violence and Health](#) and the types of crime established in the international classification of crimes proposed by the [United Nations Office on Drugs and Crime \(UNODC\)](#). In addition to these references, the areas of action and strategic focuses covered by the IDB's Citizen Security and Justice SFD, and the primary outcome variables included in internationally renowned platforms, which were used as sources of information for the development of the Platform's Evidence Bank⁴.

In turn, the list of intervention types is largely based on a consolidation of several “solution types” presented in the *Evidence-Based Platform*. It is important to note, though, that some intervention types were not included from the platform due to its strict criteria for including/creating new “solution types”. The main reason for this is that, according to the [Evidence-Based Platform methodology](#), for a new solution type to be included in the platform's evidence-bank, it is necessary that a specific systematic review (with

³ The SFDs are knowledge documents ideally updated every three to four years. – that provide a synthesis of the main development challenges and should summarize which interventions work in particular contexts and identify knowledge gaps to guide future research efforts. SFDs are intended to be a flexible guide and reference source for the IDB Group's work on a particular topic.

⁴ Most notably the definition of the outcome variables list considered the variables included in [NIJ's Crime Solutions Platform](#).

or without meta-analysis) focused on such intervention is identified, and that this systematic review meets the minimum methodological requirements defined by the platform.

For the purpose of developing an EGM, applying such restrictive criteria would have excluded important interventions. For example, Women's Police Stations (WPS) would not have been incorporated, even though six rigorous studies (one RCT and five quasi-experimental evaluations) have assessed their effectiveness, five of them in the LAC region. Under the rule used by the Evidence-Based Platform, these studies would have been omitted simply because no systematic reviews exist to date that specifically synthesize the evidence on WPS. To avoid such gaps and ensure comprehensiveness, the EGM defined intervention types more broadly: they had to be led or supported by police agencies and, according to their Theory of Change (ToC) or Logical Framework, be explicitly intended to influence the outcome variables of interest.

3.2 Criteria for including or excluding studies (PICOS)

3.2.1. Population:

For the creation of the EGM, the priority target population includes all countries within Latin America and the Caribbean (LAC) countries,⁵ regardless of any specific demographic or social factors such as race, ethnicity, culture, language, gender, sexual orientation, religion, income level, or educational background.

While this was defined as the priority population scope, finding high-quality causal evidence on policing interventions in LAC remains challenging due to the region's institutional and data constraints ([Serrano-Berthet, 2023](#)). As such, and following what had been done previously for the very Evidence-Based Platform itself, the scope of the EGM was broadened, to allow for the inclusion of residents of any country either from within or from outside the region, while also taking efforts to increase the mapping of evidence from LAC countries, first and foremost, and, to a lesser extent, developing (low, middle, and high-middle income) countries, as documented in the following sections.

⁵ For Caribbean countries, the following nations were included: Antigua and Barbuda, Bahamas, Barbados, Cuba, Dominica, Grenada, Haiti, Jamaica, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, and Trinidad and Tobago. Non-autonomous British, American, French, and Dutch territories were not prioritized. In turn, for Central and South American countries, the following ones were considered as part of the priority "targeted audience": Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, Surinam, Uruguay, and Venezuela. The only exclusion, in terms of priority, in this case, is the non-autonomous territory of French Guiana.

3.2.2. Interventions:

The intervention framework used in the EGM is primarily based on the categories proposed by [Abt and Winship \(2016\)](#), who categorize citizen-security interventions and approaches as either “Place-Based” (which target the geographic locations where violence is disproportionately concentrated), “People-Based” (which focus on the individuals and/or groups who are disproportionately linked to crime and violence), or “Behavior-Based” (which concentrate on behavioral traits that are likely to trigger violence). To those categories, the IDB team added a fourth one, “Institutional Capacity-Based”, which encompasses intervention types that did not easily fit into one of those previously mentioned categories mainly since they were instead focused on improving internal systems, capabilities, technologies, and methods within the agencies, as a way of creating better conditions for the implementation of the place-, people-, and behavior-based approaches.

These four broad categories reflect the primary focus of the EGM and capture the general logic of where, how, or on whom the policing action is concentrated and are further explained below:

- **Behavior-based** interventions include those that concentrate on behavioral patterns that may be likely to trigger violence. The main objective of those interventions is to discourage and change such behaviors through deterrence, diversion, directed assistance, or increased awareness. Examples include focused deterrence targeting open-air drug markets ("pulling levers") and drug- and alcohol-related crimes and violence prevention, deterrence, and control programs, among others.
- **Institutional capacity-based** interventions include those that focus on the strengthening of core management systems, as well as on organizational processes, practices, and competencies, including personnel training, accountability and integrity systems strengthening, organizational structure optimization, data and analytics capabilities, and cross-agency coordination. The primary objective is to enhance the underlying conditions within the police agency and enhance operational efficiency and overall institutional effectiveness.
- **People-based** interventions include those that focus on the individuals and groups who disproportionately perpetrate violence (repeat/prolific offenders) or that are disproportionately affected by it (repeat victims). These interventions often involve monitoring, deterrence, and provision of support services, sometimes in partnership with community members, to provide and foster alternatives to criminal life pathways. Focused deterrence targeting violent individuals is

an example of people-based intervention focused on repeat offenders. In contrast, second-responder programs are an example of people-based intervention focused on revictimization prevention.

- **Place-based** interventions include interventions that target the geographic locations where violence and crime are disproportionately concentrated, or localities that concentrate key vulnerabilities and risk factors that are conducive to higher levels of crime and victimization. The primary objective is to reduce crime opportunities by disrupting the routine activities and environmental cues that facilitate and/or incentivize wrongdoing and antisocial behavior. Examples of place-based interventions are community policing, hot spots policing and broken-windows policing, among others.

Within these broad categories, interventions are further grouped into second-level subcategories, such as “group and/or armed violence prevention programs” or “community-level crime, violence, and disorder prevention strategies”. Each of these subcategories brings together types of interventions, solutions or strategies with common features, which makes it possible to classify and compare interventions that may differ in design or implementation but share a similar conceptual core.

Finally, at the third level, each subcategory includes the specific types of interventions that have been evaluated in the literature (e.g., *Body-worn cameras*, *hot spot policing*, etc.)⁶. This three-tiered structure allows for consistent classification while also capturing variation in how individual interventions are designed and implemented. Annex 1 presents a description of each intervention.

Table 1: Intervention framework

Category	Intervention Category	Interventions	Listed in Evidence-Based Platform?
Behavior-based	Drug-related crimes and violence prevention, deterrence, and control programs	Drug Resistance Education Programs (DARE)	Yes
		Focused deterrence targeting open-air drug markets	Yes
		Street-level drug law enforcement	Yes
		DUI/DWI-focused policing	Yes
	Public mobilization police strategies	Communication practices for law enforcement	No
	Tough-on-crime police strategies	Stop, question, and frisk (SQF)	No
		Zero tolerance policing	No

⁶ As such, third-level categories are equivalent to the “types of solutions” presented in the Security & Justice Evidence-based Platform.

		Kingpin strategy against criminal groups	No
	Strategies to regulate and curb firearm ownership and use	Enforcement of firearms regulation	Yes
		Firearms-focused policing	Yes
Institutional capacity-based	Applied technologies for crime prevention/detection/clearance	Video surveillance systems (CCTV)	Yes
		License plate reader/electronic fencing systems	Yes
		Gunshot detection technology	No
		Real-time crime centers and/or fusion centers	No
	Police control and accountability strengthening and violence reduction programs	Police body-worn cameras (BWCs)	Yes
		Strengthening internal and external control/oversight mechanisms	No
		Procedural Justice	Yes
		De-escalation training	No
	Police reform programs	Modernization and improvement of police academies	No
		Results-oriented management	Yes
		Human resources management practices	No
		Police infrastructure renewal, expansion or reallocation	No
		De-policing strategies	No
		Defund the Police (DTP) strategies	No
	Disruptive/investigative police strategies	Enforcement of administrative police power and supervision of illegal markets	No
		Intelligence-led policing	No
		Criminal investigation practices improvement and strengthening.	Yes
		Advanced forensic techniques and technologies	Yes (A & B)
	Strengthening diagnostic and/or policymaking capacity strategies	Problem-oriented policing (POP)	Yes
		Criminal analysis improvement and strengthening	No
		Public security observatories	No
People-based	Group and/or armed violence prevention programs	Community-based violence interruption programs ("Cure Violence")	Yes
		Violent groups demobilization	No
		Comprehensive gang intervention	No
		Focused deterrence targeting violent groups	Yes
		Focused deterrence targeting violent individuals	Yes
	Police-led revictimization/recidivism prevention programs	Second responders' programs	Yes
		Women's police stations (WPS)	No
		Integrated systems for risk assessment and preventing revictimization	No
		Police-led juvenile diversion programs	Yes
Place-based		Youth curfews	Yes

	Community-level crime, violence, and disorder prevention strategies	Comprehensive territorial interventions	No
		Encouraging the use of private security equipment	Yes (A & B)
		Disorder policing ("broken windows")	Yes
		Community policing	Yes
		Neighborhood watch	Yes
		Crime prevention community councils	No
	Geographically focused policing initiatives	Proximity policing	No
		Hot spots policing (HSP)	Yes

3.2.3. Comparators:

Regarding what the interventions will be compared against, the fundamental principle is to include impact evaluations or studies that examine the causal relationship between an intervention (the treatment) and the observed impact (the outcome). In this sense, impact evaluations typically define a control group, which serves as a counterfactual; hence, the comparators are the control groups defined in the studies. In addition, the comparison depends on the study, and it can include minimal intervention, treatment, pre-post, business as usual, among others. Staggered treatments or pipeline controls are also allowed.

3.2.4. Outcomes:

Given the preventive approach that the IDB promotes, and following the [methodology used by the Evidence-Based Platform](#), the outcomes included in the EGM are categorized, in a first level, into three broad "primary", "secondary" and "tertiary" problems (Table 2).

- **Primary** problems refer to the main impacts pursued by projects in the field of citizen security and linked to the incidence and/or prevalence of crime, violence and disorder. These phenomena include, for example, lethal crimes, property crimes, criminal recidivism, and urban disorder, while also encompassing issues of a more subjective nature, such as the perception of safety.
- **Secondary** problems refer to the main variables that determine and/or condition the incidence/prevalence of primary outcomes. They can be classified into two main subtypes: "Risk factors," which are variables, events, and traits that increase the probability of crime occurrence (or the vulnerability to victimization, such as drug abuse); and "protective factors," which are variables, events, and characteristics that contribute to the development of individual, familial, or community resilience

and, therefore, decrease the likelihood that individuals become either victims or perpetrators of criminal and/or violent acts (e.g., strong social-emotional skills) ([Fisher & Lab, 2010](#)).

- **Tertiary** problems refer to institutional performance variables and institutional/governmental capacity, either to prevent, control and/or mitigate the incidence and prevalence of primary problems (e.g., “clarification of crimes”) or to have a decisive impact on secondary problems (e.g., “coverage, sufficiency and/or adequacy of public services for the protection and support of victims of violence”).

At a second level, each of these broad categories is subdivided into more specific outcome groups that cluster related phenomena, for example, lethal crime, violent crime, gender-based violence, or institutional trust and legitimacy.

Finally, at a third level, each of the aforementioned groups/clusters are further subdivided into the individual measurable indicators or constructs that directly relate to the outcome variables that were actually reported in each systematic review and impact evaluation included in the EGM (e.g., homicides, robberies, interpersonal injuries, trust in police agencies, etc.).

This three-tiered classification ensures both comparability across studies and the ability to capture the wide range of direct, indirect, and institutional effects of policing interventions. For a description of the outcome categories and outcomes, see Annex 2.

Table 2: Outcome framework

Category	Outcome Category	Outcome	Included in Evidence-Based Platform? ⁷
Primary	Crime & delinquency – Multiple crime/offense types	Crime and delinquency (multiple types)	Yes
		Crime displacement	No
		Diffusion of beneficial effects	No
	Disorder-related crimes and misdemeanors	Traffic accidents	Yes
		Disorder and public disturbance	Yes

⁷ The list of outcome variables of interest (or “problems”) of the EGM and the Evidence-Based Platform are based on the same original list (available [here](#)), which was based on the different types of violence (especially “interpersonal” violence) established in the [World Health Organization \(WHO\) World Report on Violence and Health](#) and the types of crime established in the international classification of crimes proposed by the [United Nations Office on Drugs and Crime \(UNODC\)](#). The difference between these two lists (manifested in the “No’s” listed in this last column) is explained by the fact that the platform’s list refers to those for which at least one evaluated case and/or solution type that was associated with it at the time of the Platform’s launch (April/2023). As such, the additional items (listed here with a “Yes” in the last column) refer mainly to the new references added in the EGM, and that were not originally presented in the Evidence-Based Platform (see section 3.4).

	Drug and alcohol-related crime and violence	Driving under the influence of substances	Yes
		Drug-related offenses	No
		Violence related to alcohol or drugs	Yes
	Gender-based violence	Femicides	No
		Domestic and intimate partner violence against women	No
		Violence against women	Yes
	Group violence	Presence and/or territorial control by violent groups	No
		Gang violence	No
	Lethal crime	Homicides	Yes
		Robbery-related homicides	No
	Perception of safety	Fear of crime and/or perceived insecurity	Yes
	Police abuse, lethality or victimization	Police abuse and/or lethality	Yes
		Resistance to arrest	No
		Victimization of police officers	No
		Contempt of authority	No
	Property crime	Property crimes (multiple types)	No
		Extortion	No
		Theft	Yes
		Robbery	Yes
	Recidivism	Repeat offenses/recidivism	Yes
	Violent crime	Assault or interpersonal injuries	Yes
		Sexual assault	Yes
		Violent crimes	Yes
		Gun violence	Yes
	Youth violence	Juvenile delinquency	Yes
		Victimization of youth	Yes
Secondary	Risk/protective factors associated with crime and violence	School attendance/performance	Yes
		Association with delinquent peers and/or engagement in risky behaviors	Yes
		Availability of firearms	No
		Socioemotional skills	Yes
		Alcohol abuse	Yes
		Drug abuse	Yes
		Collective efficacy	Yes
		Prevalence of a violent conflict resolution culture	Yes
		Mental health issues	Yes

		Socioeconomic vulnerability	Yes
Tertiary	Institutional performance	Access to public services	Yes
		Capacity to predict domestic violence recidivism	Yes
		Police operational efficiency	No
		Efficiency and/or effectiveness of the criminal justice system in holding offenders accountable	No
		Crime clearance rate	Yes
		Citizen satisfaction with police services	Yes
	Institutional trust and legitimacy	Trust and/or perceived legitimacy of police agencies	Yes
		Trust and/or perceived legitimacy of state institutions	No
		Propensity to report crimes	No

3.2.5. Study Design

The study design eligibility criteria is defined below, drawing on commonly accepted standards for impact evaluations ([Gertler et al., 2016](#)) and systematic reviews ([Waddington et al., 2012](#)).

The EGM will include only quantitative effectiveness literature, focused on impact evaluations and systematic reviews using attributional, causal designs to evaluate the effects of a clearly defined development intervention delivered in a real-world setting, rather than focusing solely on natural or market-based occurrences or controlled laboratory experiments without a discernible intervention component. Therefore, we will exclude studies primarily designed to determine the extent to which a specific technique, technology, treatment, procedure, or service works under ideal conditions, rather than attempting to answer a question relevant to the roll-out of a large program (i.e., lab-in-the-field).

We will only include studies that implement at least one of the following study designs widely used to evaluate intervention effectiveness ([Aloe et al., 2017](#); [Reeves, Wells, and Waddington, 2017](#)):

- A. Prospective studies that allocate participants to treatment and control groups using random assignment or quasi-experimental methods:
 - a. Randomized controlled trials (RCTs), with assignments at the individual, household, community, or other cluster level, and quasi-RCTs using prospective methods of assignment (such as alternation).

- b. Natural experiments with clearly defined intervention and comparison groups, which exploit natural randomness in implementation assignment by decision makers (e.g., public lottery) or random errors in implementation.
- B. Quasi-experimental designs where treatment arms are created without random assignment:
 - a. Regression discontinuity designs (RDD), either sharp or fuzzy designs, and other derived methods (i.e., kink RDD, differences in discontinuity).
 - b. Instrumental variables (IV).
 - c. Endogenous treatment-effects models, endogenous switching regression, and other methods synonymous with the Heckman two-step model.
 - d. Difference-in-differences (DID), two-way fixed-effects (TWFE), high-dimensional fixed effects, and two-way Mundlak regressions (TWM).
 - e. Interrupted time series (ITS) models, with or without a contemporaneous comparison group. An ITS model should include pre-intervention outcome data for a minimum of three time periods.
 - f. Weighting and matching approaches which control for observable confounding, including non-parametric approaches (e.g., statistical matching, covariate matching, coarsened-exact matching, propensity score matching) and parametric approaches (e.g., propensity-weighted multiple regression analysis).
 - g. Synthetic control methods, including their extensions: synthetic differences in differences, and generalized or augmented synthetic control methods.

In panel datasets, additional estimation strategies are often employed to address time dynamics, autocorrelation, and endogeneity, particularly when outcomes are persistent over time. These strategies include random effects models, feasible generalized least squares, and dynamic panel estimations. While these methods support inference, they only yield causal effects when combined with exogenous variation, valid instruments, or a robust identification strategy. The same applies to gravity models. Therefore, unless a clear identification strategy is provided by one of the methodologies above, these studies should be excluded.

Observational studies, evaluations, and case studies that do not meet the methodological conditions described above, such as before-after studies without a comparison group or cross-sectional studies using designs that do

not adequately address issues of selection bias or confounding, will not be included. Finally, we will also exclude the following study types: qualitative studies, feasibility studies, acceptability studies, and studies that examine willingness-to-pay for goods, services, processes, and business models. We acknowledge that the study types excluded from this map may contain valuable information; however, the focus of this EGM is to map existing rigorous evidence of intervention effectiveness.

On the other hand, this EGM also includes systematic reviews. A systematic review is a synthesis of research evidence on a particular topic obtained through an exhaustive and systematic identification of relevant studies and using widely accepted scientific strategies to minimize errors associated with appraising the design and results of studies. Reviews that have included study designs or methods not eligible for this map will be included if at least one eligible study design is included and reports results for at least one relevant intervention and one relevant outcome. Systematic reviews do not need to include a meta-analysis to be included in the map, since meta-analysis is often unsuitable when interventions are highly heterogeneous.

3.2.6. Other eligibility criteria:

a. Time frame:

There is no defined time frame for this EGM.

b. Language

Studies and reviews should be available in English, Spanish, French or Portuguese, which are the four official languages of the Inter-American Development Bank (IDB).

c. Publications:

This EGM includes both published and unpublished studies / “grey” literature, i.e., those that are not published in peer-reviewed scientific journals, such as books, handbooks, doctoral theses, master dissertations, among others. The inclusion of grey literature minimizes publication bias, which is the tendency only to publish articles that present positive evidence. Finally, this EGM accepts published articles in working paper series or institutional reports from organizations such as the IDB, IMF, World Bank, CAF, OEA, UN agencies, governmental agencies, and 3ie, or those posted on the *Evidence-Based Platform*.

3.3 Defined search strategy for the first EGM

The *Evidence-Based Platform* is the starting point of the search strategy and constitutes the main source of references for the EGM (251 studies, or 54% of

the papers included in the EGM come from the Platform). As documented correctly in the [platform's methodological annex](#), its search strategy began with a comprehensive process aimed at mapping multiple citizen security and justice evidence repositories worldwide.

Through this extensive search process, the IFD/CIS team identified approximately 40 repositories dedicated to consolidating evidence for various preventive solutions, interventions, and evaluated cases. However, upon a thorough assessment, these repositories were regarded as very heterogeneous in terms of structure, degree of sophistication, and the level of scientific rigor with which each one dealt with the “curation” of the evidence they found. Based on that, and upon the careful assessment carried out by the IFD/CIS team,⁸ a total of **seven** digital platforms were selected to be used as official sources for the Evidence Bank.

Of these seven reference platforms, the ones most relevant to the subfield of policing are: i. [Crime Reduction Toolkit](#) (for systematic reviews and meta-analysis); ii. [Campbell Collaboration](#) (for systematic reviews and meta-analysis); iii. [Crime Solutions](#) (for systematic reviews, meta-analysis, and experimental and quasi-experimental studies), iv. [Evidence-Based Policing Matrix](#) (only for experimental and quasi-experimental studies)⁹. All these platforms share the common feature of having a precise, rigorous, and transparent methodology for reviewing and including/excluding references and cases. For all these platforms, the mapping of papers to populate the *Evidence-Based Platform's* evidence bank had its cutoff in August 2020.

Moreover, the construction of the platform's evidence bank also relied on two systematic reviews developed by the IDB and/or associated consultants, one focused on programs centered on reducing homicides and robberies in Brazil ([Kopittke, 2019](#)), and another one on programs targeting the reduction of robberies and homicides in Latin America ([Kopittke, 2022 \[mimeo\]](#)). Each of these systematic reviews employed its own methodology and consulted a specific set of databases, defined by a specific set of keywords that guided the search, and included specific inclusion criteria and a standardized screening process, which can be accessed through their respective links (above). Lastly, there was an additional phase focused on closing existing gaps in the *Evidence-Based Platform*. More information about this process can be found in the [Technical Note on the “IDB's inventory”](#).

⁸ This assessment considered not only the scope of these repositories, but also, and primarily, the quality of the methodology used to map, screen, evaluate, and classify the practices and programs.

⁹ These are the sources that are more relevant for the field/area of policing interventions. However, the Evidence-Based Platform also consulted and extracted information from the following platforms, in addition to the previously mentioned ones: [Blueprints for Healthy Youth Development](#), [Social Programs that Work](#), and [California Evidence-Based Clearinghouse for Child Welfare](#).

In addition to the references collected directly from the Evidence-based Platform, an expedited search process was conducted, with the specific objective of increasing the EGM's representativeness regarding the intervention types that were not included in the Evidence-Based Platform¹⁰ at the time of the development of the EGM (see *Table 1: Intervention framework*). For this purpose, the team involved in developing the EGM executed a five-stage process.

First, the [IDB's Citizen Security Division's Sectoral Framework Document](#) (SFD) was consulted, and all relevant references (impact evaluations focused on policing interventions) were collected from there for later screening through the process defined by the present protocol (detailed in the following sections).

Second, the team also reviewed and extracted references from 37 IDB loan proposals of operations that had components related to policing interventions and that were approved by the IDB Board of Executive Directors between 2006 and 2024.¹¹ This review aimed to identify any relevant impact evaluations (or systematic reviews) cited as evidence to support the programs' vertical logics and/or theories of change.

Third, the team conducted a quick search, based on a restricted set of keywords ("impact"; "evaluation" or "assessment"; the name of each intervention type, such as "Women Police Stations" (WPS)¹² and the type of outcome most likely to be connected to each intervention, based on its general theory of change / logical framework (such as "domestic violence", in the case of the WPSs, for instance). This search was conducted in the following databases:

- CAF Publications - <https://www.caf.com/en/topics/r/research-for-development/publications/> y <http://scioteca.caf.com/discover>
- Cochrane Library - <https://www.cochranelibrary.com/>
- Google Scholar (*first 5 pages*) - <https://scholar.google.com/>
- IDB Publications - <https://publications.iadb.org/en>

¹⁰ As previously explained (see section 3.1), the main reason for this exclusion was the inexistence, at the time of the development of the Evidence-based Platform, of systematic reviews focused on such intervention (or "solution") types.

¹¹ These are the codes for the loan operations that were consulted: AR-L1074; AR-L1255; BR-L1187; BR-L1331; BR-L1343; BR-L1385; BR-L1387; BR-L1497; BR-L1546; BR-L1547; BR-L1590; BR-L1590; CH-L1142; CR-L1031; CR-L1137; EC-L1098; EC-L1294; EC-L1298; ES-L1025; GY-L1042; HO-L1063; HO-X1021; HO-G1244; HO-L1187; HO-G1251; HO-L1227; JA-L1009; JA-X1003; JA-X1006; JA-L1043; JA-X1008; JA-L1074; PE-L1224; PN-L1003; PN-X1011; PR-L1077; TT-L1003; UR-L1062; UR-L1112; UR-L1194. Further information on these programs can be found at the [IDB's project webpage](#).

¹² In some cases, the search included not solely the name/title used in the EGM for a given intervention type, but also other expressions commonly used to refer to it. That is the case, for instance, of the "Focused deterrence targeting violent groups" interventions, which are elsewhere referred to as "group violence intervention" (GVI), "pulling levers strategy", "ceasefire", or "conditional repression".

- IMF Working Papers - <https://www.elibrary.imf.org/view/journals/001/001-overview.xml>
- Institute of Labor Economics (IZA) Discussion Paper Series - <https://www.iza.org/publications/dp>
- International initiative for Impact Evaluation (3ie) - <https://www.3ie-impact.org/evidence-hub/publications>
- OEA Publications - <https://www.oas.org/ext/en/security/crime-prevention-network/Resources/Digital-Library/category/citizen-security>
- World Bank Publications - <https://www.worldbank.org/en/research>
- USAID Publications - <https://www.usaid.gov/innovation-technology-research/research>

Fourth, an informal consultation was conducted, involving all members from IFD/CIS between November 2023 and September 2025. These experts sent a series of recently published papers they encountered in their daily work through the Division's communication channels, and the EGM team was actively collecting and incorporating these additional references to the EGM original (pre-screening) list of academic references.

Lastly, during the final stages of EGM development (July 2025 – August 2025), the team extracted references from other systematic reviews that focused specifically on Latin America and the Caribbean, which had already been screened and accepted into the EGM. In total, four systematic reviews were considered and reviewed for mapping further impact evaluation studies that were later screened and, depending on their compliance with the established criteria (detailed in the following sections), included in the EGM ([Abt & Winship, 2016](#); [Silva, 2018](#); [Kopittke & Ramos, 2021](#); and [Cano et al, 2024](#)).

In total, 217 papers were mapped through these five additional steps (46% of all references included in the EGM). In the future, all these additional references will be subjected to the standard screening and methodological evaluative process described in the Evidence-Based Platform's methodological annex, and, if approved, will be included in the referred website, thereby diminishing the discrepancy in the content presented in the EGM and the Evidence-Based Platform.

3.4 Proposed search strategy for updates to the EGM

The field of citizen security, including the subfield of policing, is constantly evolving. This applies not only to the development of new solutions, interventions, and programs, but also to the body of academic research that supports them. For this reason, an EGM remains relevant only if it is regularly updated. Considering this context, this section defines the criteria for future updates to CIS's Policing EGM. In that case, the search strategy must be

based on a keyword list that will serve as the primary input for the teams responsible for the EGM. The following keywords should be incorporated in the search strategy:

- English: *delinquency, criminality, crime, felony, misdemeanor, infraction, offense, violence, homicide, violent death, violent crime, femicide, feminicide, battery, assault, property crime, robbery, theft, mugging, burglary, urban disorder, disorderly conduct, vandalism, disturbance (or breach) of the peace, nuisance, drug-related crime, gangs, factions, group violence, armed violence, gender-based violence, intimate partner violence, domestic violence, sexual violence, juvenile delinquency, fear of crime, perception of insecurity, intervention, project, policy, program, initiative, evaluation, impact, effect, efficacy, effectiveness, experimental, quasi-experimental.*
- Portuguese: *criminalidade, crime, delito, contravenção, infração, ofensa, violência, homicídio, morte violenta, crimes violentos, feminicídio, femicídio, agressão, assalto, crime contra a propriedade, roubo, furto, desordem urbana, vandalismo, perturbação do sossego, violação da ordem pública, crimes relacionados às drogas, gangues, facções, quadrilhas, grupos violentos, violência armada, violência de gênero, violência por parceiro íntimo, violência doméstica, violência sexual, agressão sexual, delinquência juvenil, medo, sensação de insegurança, intervenção, projeto, política, programa, iniciativa, avaliação, impacto, efeito, eficácia, efetividade, experimental, quase-experimental.*
- Spanish: *delincuencia, criminalidad, crimen, delito, falta, infracción, ofensa, violencia, homicidio, muerte violenta, delitos violentos, feminicidio, agresión, asalto, rapina, delitos contra la propiedad, robo, hurto, desorden urbano, vandalismo, perturbación de la tranquilidad, alteración del orden público, delitos relacionados con las drogas, bandas, facciones, pandillas, grupos violentos, violencia armada, violencia de género, violencia de pareja, violencia doméstica, violencia sexual, violación, juvenil, miedo, sensación de inseguridad, intervención, proyecto, política, programa, iniciativa, evaluación, impacto, efecto, eficacia, efectividad, experimental, cuasiexperimental.*

Moreover, in addition to the new references that come to be accepted by the *Evidence-Based Platform*, the search process for updating this EGM shall be conducted through the following databases:

- ArXiv - <https://arxiv.org/archive/econ>
- CAF Publications - <https://www.caf.com/en/topics/r/research-for-development/publications/> y <http://scioteca.caf.com/discover>

- Center of Economic Performance (CEPR) - <https://cepr.org/publications/discussion-papers>
- CESifo Network - <https://www.cesifo.org/en/publications/cesifo-working-papers>
- Cochrane Library - <https://www.cochranelibrary.com/>
- Criminal Justice Abstracts - <https://www.ebsco.com/products/research-databases/criminal-justice-abstracts>
- DOAJ - <https://doaj.org/>
- Education Resources Information Center (ERIC) - <https://eric.ed.gov/>
- Elsevier's Science Direct - <https://www.sciencedirect.com/>
- Google Scholar - <https://scholar.google.com/>
- IDB Publications - <https://publications.iadb.org/en>
- IMF Working Papers - <https://www.elibrary.imf.org/view/journals/001/001-overview.xml>
- Institute of Labor Economics (IZA) Discussion Paper Series - <https://www.iza.org/publications/dp>
- International initiative for Impact Evaluation (3ie) - <https://www.3ie-impact.org/evidence-hub/publications>
- JSTOR - <https://www.jstor.org/>
- National Bureau of Economic Research (NBER) - <https://www.nber.org/research>
- OEA Publications - <https://www.oas.org/ext/en/security/crime-prevention-network/Resources/Digital-Library/category/citizen-security>
- PsycINFO - <https://www.apa.org/pubs/databases/psycinfo>
- PubMed - <https://pubmed.ncbi.nlm.nih.gov/>
- Project MUSE - <https://muse.jhu.edu/>
- Redalyc - <http://www.redalyc.org/home.oa>
- RePEc/ EconPapers - <https://econpapers.repec.org/>
- SageJournals - <https://journals.sagepub.com/>
- SciELO - <https://www.scielo.org/>
- ScienceDirect - <https://www.sciencedirect.com>
- SSRN - <https://www.ssrn.com/index.cfm/en/>
- SpringerLink - <https://link.springer.com>
- UNICEF Publications - <https://www.unicef.org/reports>
- World Bank Publications - <https://www.worldbank.org/en/research>
- USAID Publications - <https://www.usaid.gov/innovation-technology-research/research>

Future updates of this EGM should be led by the Citizen Security Division and incorporate the keywords and databases listed previously. For these updates, new articles added to the *Evidence-Based Platform* must be reviewed. Additionally, it is recommended that a comprehensive systematic search be conducted, leveraging the previously mentioned keywords and

large-scale bibliometric databases, such as those listed above. The team responsible could also leverage the expertise of information specialists.

3.5. Screening approach

All studies included in the current EGM were put through the screening process based on the six conditions described below:

A.1 – Language:

This criterion ensures that only studies written in English, Spanish, Portuguese, or French are included, as these are the languages covered by the screening team and are the four official languages of the Inter-American Development Bank. Studies in other languages are excluded due to feasibility constraints in assessing their content.

A.2 – Intervention:

To be included, a study must examine at least one specific policy, program, or intervention aligned with the intervention categories defined in the EGM's framework. General crime trends, theoretical discussions, or studies lacking a defined intervention are excluded.

A.3 – Method:

This criterion assesses whether the study employs a rigorous causal inference method, such as those discussed in the section 'Study Design'.

A.4 – Year:

Only studies published in 2000 or later are eligible, in line with the "credibility revolution" in empirical research. Exceptions may be made if a sector expert validates a study's inclusion criteria and methodological rigor.

A.5 – Outcome:

The study must evaluate at least one relevant outcome as defined in the EGM outcome framework. Studies that focus solely on theoretical mechanisms, monitoring indicators, or non-relevant outcomes are excluded.

On the other hand, the screening process is structured in three sequential stages:

Stage 1: Initial screening

Evaluate A1 (Language), A2 (Intervention), and A3 (Method).

- If any of these are marked 0 (does not meet criteria), the study is excluded.
- If all are marked 2 (meets criteria), proceed directly to Stage 2 (full-text review).

Stage 2: Additional criteria screening (Year)

- Studies scoring 0 or 1 at A4 (Year) are flagged for sectoral review (e.g., by domain expert or lead reviewer) to decide whether to retain or exclude.
- Studies scoring 2 proceed to Stage 3.

Stage 3: Outcomes

- Studies that have passed the previous screening stages and report outcomes relevant to the intervention-outcome framework will be included.
- Studies for which the relevance of outcomes to the framework cannot be determined at earlier stages will undergo full-text screening to assess criterion A6 (Outcomes relevant to the framework) and confirm their eligibility.

4. Data extraction

4.1 Procedures

Data extraction was carried out using structured Excel spreadsheets, adapted from formats developed by 3ie, and customized to the policing EGM framework. For all the included impact evaluation studies and systematic reviews, extraction in this first pilot phase was conducted by the sector specialist, who in some cases consulted with KLD for clarification. No machine learning or text-mining tools were used.

The following information was extracted to produce the map. All the variables will be available on the interactive platform and will serve as filters for users:

- *Bibliographic information*: author, year, title, publication type, journal, DOI, abstract, publication URL, continent, country and language,
- *Study design and methodology*: evaluation design and evaluation method

- *Intervention information*: for each study, the three levels of interventions were extracted according to the intervention framework.
- *Outcome information*: for each study, the three levels of outcomes were extracted according to the outcome framework.
- *Effectiveness information*: for each study, the five levels of effectiveness were defined (Effective; Inconclusive; Mixed Evidence; Ineffective; Harmful).¹³
- *Custom fields*: Evidence-based platform tag, IDB study tag, study language tag, and journal quartile ranking (when relevant).

Additionally, other variables were extracted and linked to the study's results. These include findings such as reported effect estimates and the effectiveness of each intervention in relation to outcomes and effects (when available). These variables are documented in detail in the data dictionary accompanying the complete dataset, which includes impact evaluations and systematic reviews. They will be accessible through the IDB's open data catalogue.

4.2 Critical appraisal of systematic reviews

All systematic reviews included in the EGM were appraised for quality and methodological rigor using the SURE checklist (adapted to the EGM context by Zie), applying a fatal flaw rule. Under this rule, if any of the critical SURE criteria were not met, the review was automatically downgraded to the low-confidence category. If an important mitigating factor arises, the reviewer may decide to continue the appraisal. Coding was performed in Excel with fields tailored to capture both required appraisal elements and custom indicators. Based on these tools, reviews were classified into low-, medium- or high-confidence categories, following predefined criteria around study design, transparency, reporting standards, and risk of bias. Annex 3 includes this checklist. The detailed appraisal for each review will be published in an

¹³ "Effectiveness levels" classify impact evaluation findings based on the direction of the empirical relationship between the independent (i.e., the intervention being evaluated) and dependent (i.e., outcome variables that the intervention seeks to impact) variables. In sum, the effectiveness level reflects the general effect (or change) that the given intervention had on an outcome measure from before a program is implemented to the follow-up period, once the treatment group is compared to the effects observed in the control group / counterfactuals. More specifically, interventions showing positive, statistically significant effects are deemed "*effective*", whereas those with negative significant effects are considered "*harmful*", and those with no statistically significant differences between the effects observed for treatment and control groups are labeled "*ineffective*". "*Inconclusive*" is a category that is applied to those cases where the authors of a given paper reported major shortfalls in the evidence basis presented in a given impact evaluation (due to methodological or experiment implementation pitfalls, such as high friction or cross-group contamination, for instance), and/or when a systematic review concluded that the identified studies were insufficient to determine the effectiveness of a given intervention. Lastly, the classification as "*mixed evidence*" indicates that the study found contradictory effects (e.g., the intervention works in some population or contexts, but not in others, to a specific subset of outcomes, but not others, etc.).

accompanying Excel file, providing transparency and enabling users to understand the basis for confidence judgments.

5. Analysis and reporting

The EGM will be made available as an interactive tool, accessible to both internal and external audiences who have the link. The visualization of this first version of the EGM was used using the tools available by 3ie. This format will facilitate exploration of the evidence base and enable users to identify both clusters of rigorous research and persistent gaps. Future visualizations may be developed within the IDB.

The analysis of the Evidence Gap Map may be conducted at a later stage by the Citizen Security Division.

The full dataset of the EGM will be available to download from the IDB Open Data Catalog. This website will also include the data dictionary, this protocol, and the appraisal results of all systematic reviews.

6. Engagement and communication plan

The Citizen Security Division and the Knowledge and Learning Division will jointly lead the dissemination of the EGM. Both divisions will be responsible for engaging with relevant stakeholders and ensuring that the results are shared in a way that maximizes their policy and operational relevance.

The EGM on policing will be launched together with the EGM on transport during the IDB Knowledge Days, an internal event organized by the Knowledge and Learning Division, scheduled for October 9, 2025. This event will provide an opportunity to present the maps, discuss preliminary insights, and engage with colleagues across the Bank on their potential applications. Additional presentations to stakeholders may be organized after the launch to disseminate the findings further and encourage the use of the tool in programming and policy dialogue.

Together with these EGMs, the IDB Knowledge and Learning Division will also publish a user and a technical note, which will guide audiences on the relevant steps needed to use and develop EGMs. These materials will be available on the IDB Publications Catalog.

Annex 1: Interventions categories descriptions

This annex presents the complete set of intervention categories and interventions included in the policing EGM framework. For each first-level category, the relevant second-level subcategories and their corresponding specific interventions (third level) are listed, along with brief descriptions.

1. Behavior-based interventions

Drug-related crimes and violence prevention, deterrence, and control programs

This category includes different types of interventions aimed at addressing drug markets, drug trafficking, and other drug-related crimes. This category includes the following types of intervention: Focused deterrence targeting open-air drug markets, Street-level drug law enforcement and DUI/DWI-focused policing.

- Drug Resistance Education Programs (DARE): program where police visit schools to warn children about the harms of drugs and teach them refusal skills. It typically includes: (i) information about drugs and effects; (ii) fear arousal, stressing risks of use; (iii) moral appeals to raise awareness of drugs' social harms; and (iv) affective education, which promotes self-esteem, responsibility, and resistance to peer pressure.
- Focused deterrence targeting open-air drug markets: This type of focused deterrence aims to reduce drug trafficking and related violence in specific communities. The strategy, known as "drug market intervention," considers that the drug problem is linked to drug markets and involves integrated work between the police, the Public Prosecutor's Office, the Judiciary, and the Prison System, in addition to health services and social policies, combining deterrence with community mobilization.
- Street-level drug law enforcement: Police interventions aimed at suppressing street drug trafficking can involve two types of tactics: reactive and proactive. In the first case, police agencies suppress known drug sale points to reduce the use and availability of illicit drugs. This type of intervention involves raids, overt policing, and police investigation to arrest as many people involved in the illegal drug trade in a given territory. Proactive strategies may include problem-oriented policing (POP), community policing, and hotspot policing, among other proactive and preventive policing approaches.
- DUI/DWI-focused policing: Policing of alcohol consumption by drivers (DUI) aims to increase police presence (and/or visibility) to increase the perception and real risk of identification and arrest by the police of drivers driving under the influence of alcohol. In some countries, such

as Brazil, this act is generally considered involuntary manslaughter under criminal law. That is, it is a criminal offense in which someone's death results from the offender's actions due to negligence, lack of skill, or recklessness—even without intent to kill.

Public mobilization police strategies

This category encompasses different initiatives related to communication practices and campaigns carried out by police agencies, whether for transparency purposes, to encourage changes in behavior and social mores to reduce risks and/or vulnerability to crime and violence, or to promote greater societal engagement and community participation in public safety promotion efforts and programs undertaken by police forces.

- Communication practices for law enforcement: It includes the development and implementation of a communication strategy aimed at making security agencies communicate with the media (and the public in general) in a more professional manner, with the aim of fostering the promotion of legitimacy and citizen trust in these agencies, as well as their greater engagement in strategic actions to reduce and prevent crime and violence.

Strategies to regulate and curb firearm ownership and use

This category combines regulatory measures and policing strategies that share the common goal of curbing circulation, deter the (mis)use, and discourage de acquisition of firearms to enhance overall public safety. Regulations include licensing, background checks, restrictions, and registration to limit access and ensure responsible ownership. Policing strategies target illegal possession and use through intelligence-led operations, hotspot enforcement, and disruption of supply.

- Enforcement of firearms regulation: These are initiatives aimed at regulating and restricting civilian possession and carrying of firearms to reduce opportunities for "motivated actors" to access these weapons and thus reduce violent and lethal crime rates. Firstly, these initiatives can involve community strategies structured from a preventive character and a typical public health approach, which includes public campaigns and communication efforts to increase information, training, and promote the safe storage of firearms through campaigns and dissemination.
- Firearm-focused policing: Firearm-focused policing seeks to increase repression and seizure of illegal firearms. It can be implemented by a single agency or integrated among patrol, investigative, forensic, and intelligence units, targeting both groups specialized in arms robbery and individuals at high risk of armed violence. This type of intervention

increases integration between different police agencies and contributes to the implementation of saturation strategies in micro-territories with high circulation of people with illegal firearms.

Tough-on-crime police strategies

This category encompasses different types of “mano dura” policies designed to curb crime, violence, and insecurity by reducing criminal statistics and improving perceptions of safety through deterrence or incapacitation. They often include zero tolerance and aggressive policing, “stop, question and frisk,” and incapacitation tactics, as well as the “Kingpin strategy,” which seeks to weaken criminal organizations by arresting or neutralizing their leaders.

- Stop, question, and frisk (SQF): Stop-question-and-frisk interventions, or “stop-and-frisk,” refer to a preventive practice in which police officers temporarily detain, interrogate, and search civilians and suspects on the streets for weapons, drugs, or other contraband items. This assumes that the police have a reasonable suspicion that a crime has been, is being, or is about to be committed by the suspect.
- Zero tolerance policing: Zero-tolerance policing (ZTP) is a strategy that either or both encompass: i. the application of relentless order maintenance and aggressive law enforcement with the aim to reduce minor offenses and misdemeanors; and/or, ii. addressing more serious crime through incapacitation techniques, related to, for instance, (un-targeted) arrests, increased incarceration and harsher legal punishments.
- Kingpin strategy against criminal groups: The ‘kingpin’ strategy, also referred to as “leadership removal” strategy, focuses on dismantling criminal groups by apprehending and/or neutralizing their leaders. This strategy has long been at the heart of the War on Drugs, and assumes that ‘cutting off the head of the snake’ is the best way to incapacitate the body by incapacitating the management and leadership structures that enabled key activities, including production, transportation, distribution and financial management.

2. Institutional capacity-based interventions

Applied technologies for crime prevention, detection, and clearance

This category covers the application of innovative technologies with the objective of strengthening institutional capacity to optimize crime detection processes, increase efficiency and reduce response times and/or increase the effectiveness of initiatives aimed at preventing and solving crimes. The following types of interventions are included in this category: license plate readers/electronic fencing systems, video surveillance systems, gunshot detection technology, and real-time crime centers/fusion centers.

- Video surveillance systems (CCTV): This strategy entails the installation of CCTV cameras at key locations to deter crime, enhance perceptions of safety, and support investigations, especially for property crimes. When linked to 24/7 monitoring in Command and Control Centers (CCOs), they allow authorities to track “hot spots,” coordinate patrols, and manage emergencies. Integration can also support traffic control and municipal services like waste collection, lighting, or tree maintenance.
- License plate reader/electronic fencing systems: It includes the implementation of electronic fencing systems through the installation of vehicle plate readers at all city entrances and exits and key points such as bridges and tolls. In addition to producing an alert when previously registered vehicles pass through, the system can generate intelligence analysis on the behavior of vehicle entries or exits in the city by cross-referencing this information with criminal occurrences to support the identification and tracking of vehicles used by criminals and identifying cloned vehicles.
- Gunshot detection technology: It concerns an audio monitoring system that detects gunfire. Various sensors are installed in a specific area of communities with high homicide and shooting rates. These systems are equipped with technologies capable of pinpointing the exact location of gunfire and sending this information and the audio of the shots to the central (a Command and Control Center, or a Monitoring Center, for example), facilitating the prompt response by police forces. The system can also be integrated with video cameras that can move toward the source of the gunfire.
- Real-time crime centers/fusion centers: It includes the construction and/or strengthening of Command and Control Centers (CCO); Real-Time Crime Centers; and/or Fusion Centers. It may involve the construction and equipping of physical spaces where these centers will be installed, the development, and implementation of IT solutions, and relevant equipment and devices (e.g., body cameras, video surveillance cameras, gunshot audio monitoring, etc.). It can also involve process and governance architecture projects, and the review and/or establishment of standardized procedures for decision-making, and human resource management.

Police control and accountability strengthening and violence reduction programs

This category includes programs and interventions aimed at increasing the degree of oversight and control over police agencies, with the objective of ensuring higher levels of propriety, compliance and integrity, as well as

strengthening accountability mechanisms. The following types of interventions are included in this category: Strengthening of police internal affairs departments, Strengthening of external control/oversight mechanisms, Community safety councils, Procedural justice, Police training focused on de-escalation of the use of force, and Police body-worn cameras (BWCs).

- Police body-worn cameras (BWCs): Body cameras (sometimes referred to as "body-worn cameras") are audiovisual recording devices that attach to the uniforms of security agents and transmit real-time images and sound to monitoring centers (in addition to allowing recording). The use of this tool has a dual purpose. First, it aims to improve the relationship between agents and citizens, deterring potential episodes of abuse or violence. On the other hand, it serves the mission of protecting police officers against false allegations.
- Strengthening internal and external oversight mechanisms: It includes the implementation of mechanisms and solutions aimed at recording and publishing data on the use of force, as well as the institutional strengthening of internal affairs units to act in the analysis, investigation, resolution, and/or referral of complaints of crimes and administrative infractions committed by security agents.
- Procedural justice programs: These are programs designed to improve police legitimacy during interactions with citizens through training and implementing approaches based on the principles of "procedural justice," including standardizing procedures and techniques for the proper use of force in various types of intervention. These efforts also aim to contribute to increasing population adherence to a particular program in which community engagement is fundamental for the program to work.
- De-escalation training: Encompasses the development and implementation of training on de-escalation tactics targeted at (i) employing alternatives to use of force; (ii) advancing police–citizen communication strategies, (ii) safely responding to occurrences while protecting the safety of the individuals involved in those cases, as well as of law enforcement officers and the public; (iii) defusing situations involving armed or unarmed persons, and who may be experiencing a mental health or other crisis; and (iv) de-escalating community conflicts, disputes, and disagreements.

Police reform programs

This category includes different measures and initiatives that are usually proposed in comprehensive police reform programs, which include changes in the basic/central administration and management systems, as well as in the

structural aspects of these agencies, involving the following types of intervention: Results-based management models, modernization and qualification of police academies, human resources management practices, and renovation, expansion or re-rationalization of infrastructure.

- Modernization and improvement of police academies: It refers to the structuring, development, and implementation of programs and actions aimed at modernizing and qualifying police academies (in the case of state governments) and civil/municipal guards (in the case of municipalities). The initial training of new agents and the continuous training of the staff is a fundamental element for reducing violence rates and valuing public security professionals.
- Results-oriented management: Based on the management model originally conceived by the Compstat experience in New York, results-based management programs in citizen security advocate for establishing performance goals and incentive mechanisms aimed at generating relevant results in public security (e.g., reducing homicides or robberies).
- Human resources management practices: It encompasses the different practices that are part of the people management system, covering the entire life cycle of employees (agents, analysts, etc.) in police organizations, thus encompassing the policies of selection, allocation, (re)distribution, promotion, compensation, motivation and financial or non-financial incentives for productivity and/or effectiveness at an individual or team level.
- Police infrastructure renewal/expansion: It includes different efforts related to expanding, reforming and/or rationalizing the distribution of the main infrastructures through which police agencies organize their territorial distribution and offer services to the population, as in the case of police stations or battalions. It also includes initiatives aimed at strengthening/expanding the information and communication technology (ICT) infrastructure and/or promoting further digitalization of police agencies, with the aim of boosting operational efficiency and/or productivity levels.
- De-policing strategies: De-policing refers to an observed pattern of police disengaging due to a discretionary and sharp reduction in police proactive activities; such as patrolling, interaction with citizens, inquiries, stops, and apprehensions.
- Defund the Police (DTP) strategies: De-policing refers to an observed pattern of police disengaging due to a discretionary and sharp reduction in police proactive activities; such as patrolling, interaction with citizens, inquiries, stops, and apprehensions.

Disruptive/investigative police strategies

This category covers institutional strengthening initiatives aimed at increasing the capacity of law enforcement agencies to investigate and solve crimes, as well as their ability to affect and reduce specific illicit/illegal markets. This category includes the following types of interventions: Administrative police enforcement and illicit market monitoring, Intelligence-led policing, Improving and qualifying criminal investigation practices, and Advanced forensic techniques and technologies (including DNA banks and ballistic analysis systems, for example).

- Enforcement of administrative police power/illegal markets supervision: Also called the “market reduction approach” (MRA), this strategy uses police intelligence and administrative powers to curb the trade of stolen goods and reduce incentives for property crimes. It includes institutional efforts to train and equip security forces to monitor and disrupt illegal markets, as well as the reform of municipal regulations and efforts to strengthen inspections. The goal is to cut commercialization opportunities and limit criminal networks’ access to illicit profits.
- Intelligence-led policing: Intelligence-led policing (ILP) is a managerial model of law enforcement that seeks to put criminal intelligence at the forefront of police decision-making. ILP is a practice and policing model built around risk assessment and management that seeks to leverage technological advances in data collection and analysis to generate “intelligence” inputs that can be practically used to guide police strategies, tactics, and operations.
- Criminal investigation practices improvement: SIIt involves adopting practices, techniques, approaches, and technologies, as well as standardizing relevant procedures, protocols, and management models to leverage the investigative capacities of police agencies. The goal is to increase these agencies’ efficiency and effectiveness in conducting investigative processes promptly and timely, both in terms of increasing the proportion of cases solved (“clearance rate”) and reducing the times involved in resolution.
- Advanced forensic techniques and technologies: It concerns the improvement of technical-scientific police practices regarding the production, legally speaking, of technical evidence to support crime resolution, especially in cases involving crimes against life. It may involve advanced techniques, such as the creation and maintenance of a genetic profile bank and the application of DNA testing in police investigation, multi-biometric identification systems, ballistic profile databases, and the application of frontier technologies in digital forensic science.

Strengthening diagnostic and/or policymaking capacity

This category encompasses different policing strategies aimed at increasing the capacity to identify, detect, analyze and understand phenomena associated with the incidence of crime and violence, as well as the prevalence of risk and protective factors and relevant social determinants, with the purpose of leading, based on this in-depth diagnosis, to the formulation of more effective policies, programs and actions. Including: Problem-Oriented Policing (POP), Improvement and qualification of criminal analysis, and Public Safety Observatories.

- Problem-oriented policing (POP): Problem-oriented policing (POP) is a proactive policing model that seeks to establish a process through which security problems are identified, analyzed, and prioritized in-depth, to define priority actions capable of addressing their main determinants, generating a sustainable solution to that problem. Typically, the implementation of POP is carried out using the SARA Method (an acronym in English for Scanning, Analysis, Response, and Assessment).
- Criminal analysis improvement and strengthening: Criminal analysis is the systematic application of methods to deepen existing knowledge on crime incidence/prevalence, and criminal dynamics. It involves analyzing data on offenders, locations, times, and methods to identify patterns and threats. Insights help agencies solve crimes, detect prolific offenders, select evidence-based tactics, prioritize micro-territories for patrol, address community problems, and plan resources more efficiently. This makes it a key tool for proactive and strategic policing.
- Public security observatories: Public security observatories are centers dedicated to monitoring, analyzing, and disseminating relevant data in public security (monitoring crime indicators, official data, academic research, and government budgets, for example). They can also act in (or sponsor) specific studies or victimization research. Observatories produce their analyses and generate their reports, infographics, seminars, and meetings to subsidize strategic actions for crime and violence control and prevention that may be presented and potentially absorbed by the competent authorities.

3. People-based interventions

Group and/or armed violence prevention programs

This category includes different strategies that have in common the objective of reducing violence, especially, but not only, armed violence caused by organized crime groups (gangs, gangs, cartels, etc.). The following types of intervention are included in this category: the different modalities of focused deterrence, "cure violence" strategies, armed group demobilization strategies, and comprehensive gang intervention programs.

- Community-based violence interruption programs (Cure Violence): The "Cure Violence" Program is based on a concept originating from epidemiology by addressing violence as an infectious disease, as it presents three key characteristics: clustering (occurs in specific places and times), self-replication (seems to multiply "autonomously" if not interrupted), and presence of epidemic waves (concentration of a significant volume of occurrences within a short period).
- Violent groups demobilization: It refers to a strategy of prioritizing the efforts and resources of the Public Prosecutor's Office to prosecute chronic criminal activities.
- Comprehensive gang intervention: It is a strategy that the Office of Juvenile Justice and Delinquency Prevention (OJJDP/DoJ) of the United States Department of Justice recommends as a solution to American municipalities to address the problems generated by gangs and violent groups involved in illegal activities and/or groups that exert illegal (and violent) territorial dominance over socially vulnerable communities.
- Focused deterrence targeting violent groups: The strategy of focused deterrence on violent groups (also known as violent group intervention) refers to a type of intervention in which strategic, tactical, and operational alliances are established between different institutions and agencies of the security and justice systems to deter certain groups responsible for a disproportionate volume of crimes and violence in a given locality.
- Focused deterrence targeting violent individuals: The strategy of focused deterrence on highly violent individuals refers to a type of intervention in which strategic, tactical, and operational alliances are established between different institutions and agencies of the security and justice systems to deter certain individuals responsible for a disproportionate volume of crimes and violence in a given locality.

Police-led revictimization/recidivism prevention programs

This category includes police initiatives for people who have been victimized, in order to reduce harm and, above all, prevent re-victimization. These may focus on the victims, on their care, and on the provision of services to prevent further episodes of violence, or they may focus on the aggressors, with the aim of deterring them from reoffending and preventing them from becoming "prolific offenders". Including: second response programs, specialized police stations for women, integrated systems for risk assessment and re-victimization prevention, and police-led juvenile diversion programs.

- Second responder programs: It includes periodic visits by multi-sectoral teams (municipal and/or police officers, social workers, health professionals, and public defenders) to the homes of women in situations of domestic and family violence, to verify compliance with urgent protective measures provided and/or to preventively attend to domestic violence occurrences/complaints, aiming to deter acts of violence or threats against women and reduce (re)victimization.
- Women Police Stations (WPS): Women's police stations are police stations specializing in crimes with female victims and are focused on specific types of crime related to gender-based violence, such as psychological violence, and family/domestic violence, as well as on specific types of threats and sexual violence. Some units are also connected to and might divert the victims to other public/social services such as financial help, counseling, and specialized medical care/assistance.
- Integrated risk assessment systems: These initiatives create integrated systems to assess and reduce the risk of revictimization of women facing domestic or intimate partner violence. They use shared databases across agencies (health, security, justice, and social assistance) and risk forms to identify and analyze predictive factors. In some cases, algorithms are used to classify risk levels, triggering alerts and preventive actions like restraining orders, electronic monitoring of aggressors, institutional shelter, or inclusion in police visit services.
- Police-led juvenile diversion programs: They comprise a set of strategies that the police can implement as an alternative to criminal prosecution of young people. Redirection programs focus on inserting young people who have committed minor offenses into social prevention programs instead of presenting them to the criminal justice system.

4. Place-based interventions

Community-level crime, violence, and disorder prevention strategies

These are initiatives implemented at the community level to prevent crime, violence, and urban disorder by addressing risky behaviors and situational conditions. Interventions include youth curfews, comprehensive territorial actions, use of private surveillance equipment, firearms regulation, and targeted policing, disorderly policing (“Broken Windows”), community policing, and neighborhood watch programs that strengthen local capacities and improve social control.

- Youth curfews: They refer to programs for enforcing curfew laws, that is, keeping adolescents and young people (generally under 17) in the

domestic environment at certain times (especially at night). The purpose of these interventions is to reduce these youths' exposure to circumstances most conducive to committing crimes (or being victimized). In several cases, these laws grant the police the authority to stop and question young people, and require them to return home or face fines or sanctions, as appropriate.

- Comprehensive territorial interventions: This preventive approach focuses on vulnerable neighborhoods, mobilizing territorial agents—such as social workers, mediators, and municipal staff—to resolve local disputes, reduce disorder, and address basic needs like lighting or waste, in many cases in partnership with police agencies. At the same time, it seeks to build community capacity by strengthening local organizations, supporting councils, and training “focal points.” Beyond reducing disorder, these interventions promote state legitimacy and indirectly reduce crime and violence.
- Encouraging the use of private security equipment: This strategy combines prevention with rapid response to burglaries in homes and businesses. It can be done in a proactive fashion, when security forces identify vulnerable areas and advise residents and shopkeepers on adopting protective measures such as reinforced doors, alarms, access control, or electronic tags; or in a reactive fashion, when the effort is directed at ensuring a quick initial response to calls related to burglaries to ensure the best possible results regarding that occurrence (identifying suspects, protecting victims, accessing key witnesses).
- Disorder policing (“broken windows”): The programs focusing on disorder and nuisance policing, also called “broken windows” policing or public order maintenance policing, are based on the premise that by reducing disorder and the incidence of minor offenses in a given community, it is possible to improve the quality of life and citizens' sense of security and reduce the likelihood of more serious crimes.
- Community policing: In its most basic version, this approach seeks to ensure that the same police officers are permanently assigned to a specific geographical area, so they can become familiar with local problems and residents, thus promoting greater collaboration (and effectiveness). These programs emphasize and prioritize community involvement in the policing process to develop partnerships between the police, community members, and civic organizations to establish and advance priorities, as well as define and monitor the tactics employed.
- Neighborhood watch: Neighborhood watch, or “community surveillance,” creates local groups that partner with police to deter crime such as disorder and burglary. Initially designed to provide “extra eyes and ears” for law enforcement, these programs rely on residents reporting suspicious activity and improving household security. Today,

they are often combined with community policing or CPTED strategies, boosting local social capital and reducing vulnerability to crime and violence.

- Crime prevention community councils: It includes the establishment and/or strengthening of community safety councils as mechanisms for social participation and external control of security policies. In this sense, it may involve the creation and/or training and empowerment of these spaces for dialogue between society and government representatives, from the perspective of co-management of public security policies, whether in proposing, negotiating, deciding, implementing, or overseeing public policies.

Geographically focused policing initiatives

This category encompasses a set of types of intervention that, despite their differences in form, approach, or policing strategy, have in common the deployment of concentrated action in a specific territory, which can be very restricted (as in the case of hot spot policing) or broader (as in the case of proximity policing). Through this geographically concentrated action, the aim is to reduce crime, violence, disorder and/or the feeling of insecurity.

- Proximity policing: Proximity policing programs aim to increase the presence and visibility of police officers, whether in conflict areas or specific localities, neighborhoods, or regions characterized by high crime, violence, or disorder; always to significantly reduce these indicators. Unlike the community policing model, proximity policing does not necessarily presuppose the structuring of methods and spaces for dialogue and agreement between the police and the local community.
- Hot spots policing (HSP): It is a strategy of preventive patrolling in specific "hot spots" of the city, that is, urban micro-territories that concentrate high crime rates (whether specific intersections, blocks, or corners, or specific public spaces such as train and bus stations or a square), to reduce the crime rates observed in these locations. This policing strategy advocates for proactive and preventive action in these locations through targeted patrols that focus on the critical hot spots identified by the police through spatial analysis techniques.

Annex 2: Outcomes categories and outcomes descriptions

This annex provides the full descriptions of the second-level outcome categories and their corresponding specific outcomes (third level) as included in the policing EGM framework. The outcomes are organized across three levels: Primary, Secondary, and Tertiary.

1. Primary outcomes

Crime & Delinquency – Multiple crime/offense types

This category encompasses various types of criminal activity. It is important to note that this category applies to those specific cases in which systematic reviews (or individual impact evaluation studies) evaluated the effectiveness of a particular (type of) intervention on a wide range of offenses, in a generalized way (through a consolidated index, for example), without necessarily presenting individual results for each typology.

- Crime and delinquency (multiple types): Category that encompasses multiple types of criminal activities. It includes both specific sub-types and cases in which systematic reviews (or impact evaluation studies) assessed the effectiveness of a given (type of) program across a broad range of crimes in a generalized manner (e.g., using a composite index), without necessarily presenting individual results for each crime type.
- Crime displacement: Crime displacement effects (also called "spill-over effect") refer to cases where the implementation of a given crime-prevention program or interventions unintentionally shifts crime rather than eliminates it. That is, instead of stopping offending altogether, it may simply cause offenders to change where ("geographical displacement"), when ("temporal displacement"), or how (crime type - or tactical -- displacement) they commit a crime.
- Diffusion of beneficial effects: In criminology, "diffusion of beneficial effects" occurs when a crime-prevention program reduces crime not only in targeted areas but also in surrounding ones. Benefits may extend across space, time (lasting after the intervention ends), or crime types (reductions in offenses not directly targeted). For example, hot spot policing can decrease crime in both intervention zones and nearby areas.

Disorder-related crimes and misdemeanors

This category includes variables related to urban disorder or breaches of the peace that reduce quality of life or harm perceptions of security in a given

area. It covers acts of social incivility (violations of norms of coexistence), misdemeanors/public order offenses (e.g., unauthorized activities, visual, physical, or noise pollution), and vandalism (damage or destruction of monuments, public, or private property).

- Traffic accidents: Category that covers different types of traffic accidents (e.g., collisions, crashes, rollovers, pedestrian run-overs, impacts with fixed objects, etc.). It includes both specific sub-types and cases where systematic reviews (or impact evaluations) assessed the effectiveness of a certain (type of) program over a wide range of automobile accident outcomes in a generalized way (e.g., using a composite index), without presenting individual results for each accident type.
- Disorder and public disturbance: This category refers to a disruption of the peace, order, or tranquility of a public place, encompassing a range of behaviors from minor disturbances that cause alarm, or create danger to others, deteriorate the urban fabric, reduce quality of life, and/or negatively affect the population's perception of safety.

Drug and alcohol-related crime and violence

This category encompasses both the different criminal acts associated with the possession, manufacturing, and/or distribution of illegal narcotics, as well as the different acts of violence that have in common the fact of having been committed under the influence of alcohol or drugs.

- Driving under the influence of substances: Driving under the influence of alcohol or other psychoactive substances is an offense that endangers not only the driver's life but, more importantly, the lives of others. In some countries, such as Brazil, this act is generally considered involuntary manslaughter under criminal law. That is, it is a criminal offense in which someone's death results from the offender's actions due to negligence, lack of skill, or recklessness—even without intent to kill. Brazilian law may also classify this as "eventual intent" when the offender's conduct demonstrates acceptance of the risk of causing death.
- Drug-related offenses: Category that includes violations of laws governing the manufacture and/or distribution of psychoactive substances deemed illegal under a given legal system, as well as various acts of violence that share the common factor of having been committed under the influence of drugs.
- Violence related to alcohol or drugs: Category that includes various acts of violence that share the common characteristic of having been committed under the influence of alcohol or drugs.

Gender-based violence

This category covers crimes against women and girls, including both lethal violence (e.g., homicide, femicide) and non-lethal violence (e.g., bodily injury, rape). It applies to cases assessing specific forms of violence as well as studies or reviews that evaluate programs' effects on violence against women more broadly, such as through consolidated indices, without reporting separate results for each type.

- **Femicides:** Femicide (also referred to as femicide) is defined as the intentional killing of women and girls with a gender-related motivation, which can be driven by stereotyped gender roles, discrimination against women and girls, unequal power relations between women and men, or harmful social norms. It refers, thus, to the gender-based killings of women and girls.
- **Violence against women:** Category encompassing crimes committed against the female population (adult women or girls). This broad category may include both lethal violence (e.g., homicides, femicides) and non-lethal violence (e.g., bodily harm, rape). It covers both specific sub-types and cases in which systematic reviews (or impact evaluations) assessed the effectiveness of a program across a wide range of forms of violence against women in a generalized way (e.g., through a composite index), without presenting results for each type individually.
- **Domestic and intimate partner violence against women:** Domestic violence refers to a type of violence, most often directed against women ("gender-based violence"), committed by family members or others living in the same household. It may involve acts or omissions that cause physical, sexual, or psychological harm, as well as moral or material damage to the victim. This category also includes what is commonly known as intimate partner violence (IPV).

Group violence

This category, which is also commonly referred to as collective or intergroup violence, refers to acts of violence committed by a group or collective of individuals against another group or individuals to achieve political, social, or economic goals.

- **Presence and/or territorial control by violent groups:** The presence of criminal groups and gangs that systematically use violence and/or engage in criminal practices—as a way to resolve conflicts, dominate territories, or generate economic profit (e.g., drug trafficking, control of local economic activities, extortion, robbery)—is a significant factor in the prevalence of crime and violence. Territorial control by these groups often leads to rights violations and abuse, even in monopolized

criminal markets. The risk of lethal violence increases with inter-group conflicts.

- **Gang violence:** Category encompassing various acts of violence committed by groups of individuals who regularly plan and execute crimes for obtaining political, social, or economic goals, involving the use of violence. This includes situations where the perpetrator is a member of a gang or violent group and uses or threatens to use force against the victim—whether the victim belongs to the same group, a rival group, or is unaffiliated.

Lethal crime

Covers all crimes that result in death and are committed with the intent to kill. Included in this category are murder and non-negligent manslaughter, theft-related homicides, and aggravated assault (battery followed by death).

- **Homicides:** This category refers to a type of crime in which the perpetrator intentionally takes another person's life. While used as a general term, for the purpose of this EGM, "homicides" solely refer to lethal criminal acts, with no necessary distinction between cases with or without the intent to kill. It thus potentially includes cases which are elsewhere dealt with separately, such as "murder" and "manslaughter".
- **Robbery-related homicides:** Refers to a crime combining two offenses: robbery (taking property through force) and homicide (intentional killing of the victim). Although not specifically defined as such in all penal codes around the globe, this type of offense is recognized in several legal systems and is referred to in the international literature as "robbery-related homicides".

Perception of safety

This category covers the subjective phenomenon of individuals' perception of their safety and/or their fear of becoming victims of some type of violence or crime.

- **Fear of crime and/or perceived insecurity:** Fear of crime, or perceived insecurity, is the subjective feeling of anxiety or vulnerability about becoming a crime victim, regardless of actual crime rates. It is shaped by personal experiences (e.g., victimization), social influences (e.g., media, neighborhood reputation), and environmental cues (e.g., poor lighting, vandalism). This fear often changes behavior—avoiding places, limiting activities, or adopting protective measures.

Police abuse, lethality, or victimization

This category includes both episodes of brutality and lethality committed by police officers against citizens (suspected or not of criminal activity) as well

as assaults and attempts on the lives of police officers committed by "civilians".

- Police abuse and/or lethality: Encompasses episodes in which a law enforcement officer deviates (intentionally or due to incompetence) from the appropriate level of force against a suspect or offender, causing harm and/or, ultimately, death.
- Resistance to arrest: Refers to acts in which a citizen resists or opposes the execution of a legal act with violence or threats against the law enforcement officer performing it.
- Victimization of police officers: Gathers episodes involving lethal attacks against the life of a police officer (whether on or off duty).
- Contempt of authority: The crime of contempt refers to the act of disrespecting or insulting a public official—here, a police officer—in the performance of their duty or because of it, through offensive, defamatory, or slanderous speech, threats, obscene gestures, etc.

Property crime

This category encompasses various types of property crime, i.e., criminal actions aimed at stealing or damaging the property of another person or organization. It covers both violent (robbery) and non-violent (theft/larceny) property crimes, in their various forms: Burglary (Breaking and entering-B&E), street robbery/crime, auto-theft, etc.

- Property crimes (multiple types): Category encompassing different types of property crime—i.e., criminal acts aimed at harming another person's or organization's property. It includes both specific sub-types and cases where systematic reviews (or impact evaluations) assessed the effectiveness of a program across a broad set of property crimes using a composite index, without reporting results for each type individually.
- Extortion: Extortion is the act of obtaining something—often financial gain or high-value assets—through force, threats, or intimidation, often involving abuse of power or authority.
- Theft: Category encompassing different types of theft (i.e., stealing without violence), regardless of how the crime was carried out (e.g., pickpocketing, shoplifting, residential theft, auto theft, etc.).
- Robbery: Category encompassing different types of robbery (i.e., theft involving serious threat or violence), regardless of how the act was committed (e.g., mugging, burglary, breaking and entering, carjacking, etc.).

Recidivism

This category groups different types of criminal recidivism, regardless of age group (juvenile/adult), sex (female/male), or specific type of crime (violent

crimes, sexual crimes, crimes against property, etc.).

- Repeat offenses/recidivism: Category that includes various forms of criminal recidivism. It covers both specific sub-types and cases where systematic reviews (or impact evaluations) assessed the effectiveness of a program on recidivism in a broad range of crimes using a composite index, without reporting results for each specific type.

Violent crime

Violent crimes are defined as those that involve the use of force or the threat of force against another person, resulting in injury or death. They are characterized by acts that inflict physical harm or the potential for harm, and are considered serious offenses with significant legal consequences. Examples include homicide attempt, assault, rape, and sexual assault.

- Assault or interpersonal injuries: Category that includes incidents, records, or self-reports of physical aggression and non-lethal bodily injury.
- Sexual assault: Category covering legal definitions of sexual assault, including situations where victims are coerced into sexual acts through violence or serious threats.
- Violent crimes (multiple types): Category encompassing a set of violent crimes in which the perpetrator uses or threatens to use force. It includes both specific sub-types and broader cases where evaluations assessed the impact of a program on various violent crimes using a composite index, without disaggregated results.
- Gun violence: Category covering various types of criminal activity involving firearms, including gunshot injuries, non-lethal firearm wounds, and firearm threats.

Youth violence

This category encompasses all crimes committed against children, adolescents, and young people, including episodes of violence committed against them, as well as so-called "juvenile delinquency", i.e., the commission/perpetration of criminal acts by these same demographic groups.

- Juvenile delinquency: Category encompassing various forms of youth-perpetrated crime. It includes both specific sub-types and broader cases where evaluations assessed program effects on youth crime using a composite index, without individual results.
- Victimization of youth: Category covering crimes committed against children, adolescents, youngsters, and youth. It includes both specific types and cases where systematic reviews or evaluations assessed the

effect of a program on a broad range of violence against minors using a composite index, without disaggregated outcomes.

2. Secondary outcomes

Risk/protective factors associated with crime and violence

This category refers to key risk factors that increase the vulnerability of individuals, groups, or communities to involvement in violence or crime, either as victims or perpetrators. Risk factors may appear at different levels: individual (e.g., drug abuse), family (e.g., domestic violence), community/geographic (e.g., presence of organized crime), social (e.g., tolerance of gender-based violence), or economic (e.g., crises and sudden income loss).

- **School attendance/performance:** Low school attendance limits students' learning and skill development, undermining academic performance and cognitive growth. Poor performance is a strong predictor of dropout, especially when combined with absenteeism or grade retention. Youth "Not in Education, Employment, or Training" (NEET) are particularly vulnerable, as NEET status is consistently linked to a higher risk of delinquency-related behaviors.
- **Association with delinquent peers and/or risky behaviors:** Evidence shows that among individual risk factors for youth violence, one of the most prominent is having delinquent or antisocial peers or being part of a gang or violent group. This significantly increases the likelihood of future criminal behavior and victimization.
- **Availability of firearms:** Refers to the availability of firearms in a given community, especially in the case of those produced illegally and/or legally purchased but diverted to the illegal market (through theft, diversion from security institutions, or international trafficking), and/or acquired without proper regulation (e.g., imposition of background checks). Greater firearm circulation increases access by potential offenders, which is a major risk factor for crime, especially violent crime.
- **Socioemotional skills:** Socioemotional skills include emotional awareness, emotion management, and interpersonal skills. They comprise the knowledge, skills, and attitudes needed to understand, express, and regulate emotions and make responsible decisions. Systematic reviews indicate that strong socioemotional development is a protective factor against delinquency, while antisocial behavior, low self-control, impulsiveness, and tolerance for peer misbehavior increase the risk.
- **Alcohol abuse:** Early alcohol use is a significant risk factor for youth violence, as well as for victimization. Among adults, the abusive use of alcohol is regarded as a key predictor that increases the likelihood of

several criminal behaviors, such as child maltreatment, violence against women, and sexual abuse, among others.

- **Drug abuse:** Early use of psychoactive substances is a major risk factor for violent behavior in youth, as well as for victimization. Among adults, the abusive use of drugs is regarded as a key predictor that increases the likelihood of several criminal behaviors, such as child maltreatment, violence against women, and sexual abuse, among others.
- **Collective efficacy:** Collective efficacy relates to social capital—a set of norms, rules, and trust-based networks that facilitate cooperation in a community. Areas with social disorganization and weak community ties are more vulnerable to violence and crime, especially when combined with poverty, lack of services, or control by criminal organizations. Social disorganization weakens informal control and fosters distrust, creating a favorable environment for crime.
- **Prevalence of violent conflict resolution culture:** The presence of a culture where conflicts are resolved through violence is a risk factor for higher societal violence, especially reactive or emotionally charged violence. Weak community capacity for conflict resolution and lack of mediation policies are predictors of increased lethal crime.
- **Mental health issues:** The prevalence of mental health issues is a relevant risk factor—not only for crimes or non-instrumental violence by individuals with psychological disorders but also for violence against women, particularly in cases involving depression or other mental health conditions.
- **Socioeconomic vulnerability:** Recent studies show that sudden crises or job losses increase the likelihood of individuals committing both economically motivated and violent crimes. This is especially true for groups more vulnerable to liquidity shocks—e.g., young workers or those with unstable jobs and low education levels.

3. Tertiary outcomes

Institutional performance

Category that encompasses a wide range of measures, including crime clearance rates, asset recovery, response times, seizures, and community satisfaction indexes, with an eye to assessing the efficiency and efficacy/effectiveness of a police agency in fulfilling its mission to ensure law and order and protect citizens while maintaining legitimacy within the community.

- **Access to public services:** Limited access to services—due to physical, geographic, linguistic, economic, or other barriers—hinders vulnerable groups' ability to reduce risks of victimization or involvement in crime. For example, lack of access to justice limits peaceful conflict resolution, while poor service quality (e.g., delays, lack of courtesy) undermines

trust in institutions and discourages use of formal mechanisms.

- Capacity to predict domestic violence recidivism: The use of forensic techniques in support services targeting women can help the timely identification of victims of violence and/or domestic abuse. With advances in statistical and computational methods, it is now possible to use intelligent algorithms (machine learning) to predict the risk of revictimization. When leveraged properly, this can enable security, health, and social services to act preventively, protecting high-risk women and avoiding repeated victimization.
- Police operational efficiency: This category includes various indicators related to the efficiency of police agencies in preventing and controlling crime and violence, thus encompassing measures such as response times, seizures, processing times, reporting accuracy, and so on. Overall efficiency, as well as the coverage and quality of police services, are regarded as key factors in assessing a State's institutional capacity to address and sustainably reduce crime.
- Efficiency/effectiveness of the criminal justice system: This includes indicators of criminal justice institutions' capacity, such as processing times, conviction rates, and judicial congestion. Weak accountability fosters impunity, encouraging offender behavior and undermining public trust. This erodes cooperation with judicial institutions, creating a vicious cycle that diminishes institutional legitimacy and fuels crime and violence.
- Crime clearance rate: Weaknesses in forensic systems and investigative procedures lead to low clearance rates for homicides. This contributes to a sense of impunity, which reinforces criminal behavior and erodes public trust in the police, reducing cooperation and feeding a cycle of crime and violence.
- Citizen satisfaction with police services: Dissatisfaction with policing models and with the quality of service (e.g., speed and courtesy) undermines trust in the police and their perceived legitimacy, distancing them from the community. This weakens public willingness to collaborate, further eroding security and effectiveness—since police performance often depends on public cooperation to report crimes and act as witnesses.

Institutional trust and legitimacy

Institutional trust refers to the level of confidence the public has in the police as a whole, in terms of both its capacity to fulfill its institutional mandate and mission, as well as to do so through actions that are legitimate and fair. This trust is crucial for law enforcement's effectiveness, as it encourages public cooperation and compliance with the law.

- Trust and/or perceived legitimacy of police agencies: Declining public trust in the police and their legitimacy reduces the likelihood that

individuals will cooperate with or turn to the police for conflict mediation or protection. This sets off a vicious cycle that fuels crime and erodes perceptions of safety.

- Trust and/or perceived legitimacy of state institutions: Trust in the state refers to citizens' belief in government and public institutions' fairness, legitimacy, and efficiency. High trust fosters law compliance and cooperation, strengthening social order. Low trust, by contrast, reduces compliance and cooperation, heightens fear and extralegal behaviors, and creates a feedback loop where crime erodes trust, weakening crime-prevention efforts.
- Propensity to report crimes: The propensity to report crimes—especially to the police—is often used as a proxy for social trust. One of the main factors influencing the decision to report is the belief that doing so will lead to action and have positive outcomes, both individually (e.g., asset recovery) and collectively (e.g., improved policies, reduced recidivism).

Annex 3

Checklist for making judgments about how much confidence to place in a systematic review of effects (adapted version of SURE checklist)^[1]

Section A: Methods used to identify, include and critically appraise studies

<p>A.1 Were the criteria used for deciding which studies to include in the review reported?</p> <p>Did the authors specify:</p> <p>A.1.1 Types of studies</p> <p>A.1.2 Participants/ settings/ population</p> <p>A.1.3 Intervention(s)</p> <p>A.1.4 Outcome(s)</p> <p><i>Note.</i> This information cannot be determined by looking at the types of studies included, because some eligible populations, designs, interventions, and outcomes might not have been examined in the studies.</p>	<p>Yes Partially No</p> <p>Coding guide - check the answers above YES: All four should be yes NO: All four should be no PARTIALLY: Any other</p>
<p>A.2 Was the search for evidence reasonably comprehensive?</p> <p>Were the following done:</p> <p>A.2.1 Language bias avoided (no restriction of inclusion based on language)</p> <p>A.2.2 No restriction of inclusion based on publication status</p> <p>A.2.3 Relevant databases searched: at least one database that includes grey/unpublished literature, as well as either: (a) for health, at least two relevant comprehensive subject databases (such as PubMed/MEDLINE, EMBASE and CENTRAL),^[2] or (b) for social sciences, at least two relevant comprehensive subject databases (such as IDEAS) and one comprehensive general database (such as EconLit, PsychInfo, Scopus)</p> <p>A.2.4 Reference lists in included articles checked</p> <p>A.2.5 Authors/experts contacted</p> <p><i>Notes.</i> When authors do not mention limitations on language or publication status, code Yes. The use of "published" often simply means released (e.g., "studies published between 1990 – 2010") and not necessarily that studies were excluded based on publication status; do not code No simply because the authors use "published" in this way. When authors do not mention that reference lists were searched or experts contacted, code No. If authors were only contacted for study results data, code No. Checking reference lists of review articles does not fully meet A.2.4 requirement (code <i>Partially</i>) but is a mitigating factor. Grey literature typically means research that is not published in sources such as books or journal articles. The following databases include grey literature: Academic Search Complete (includes many conference proceedings), CAB Abstracts, searches conducted using CADATH checklist, clinicaltrials.gov, Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Library, Embase (includes 3.6m+ conference abstracts), Google, Google Scholar, Healthcare Management Information Consortium (HMIC), IDEAS/RePEc, National Technical Information Service (NTIS), OpenSIGLE/OpenGrey, PsycEXTRA, Scopus (includes ~10m conference papers). If you identify additional sources, please notify the DEP team. Searching websites of relevant governmental agencies and non-governmental organizations can also identify grey literature. Note that MEDLINE/PubMed, a comprehensive database of journals, does not include grey literature: "For indexing in MEDLINE, NLM currently selects publications that it considers to be journals."; see also Citrome L. Beyond PubMed: Searching the "Grey Literature" for Clinical Trial Results. <i>Innov Clin Neurosci.</i> 2014;11(7-8):42-46. EBSCO and OVID are platforms, not databases. If an author only reports searching "EBSCO" without identifying the databases searched, code <i>Partially</i></p>	<p>Yes Partially No Can't tell</p> <p>Coding guide - check the answers above: YES: All five should be yes PARTIALLY: Relevant databases and reference lists are both reported NO: Any other</p>
<p>A.3 Does the review cover an appropriate time period?</p> <p><i>Is the search period comprehensive enough that relevant literature is unlikely to be omitted?</i></p> <p><i>Note.</i> If the authors do not report the search period, check the publication date of the earliest included study. If the study was published before 1990 this can be coded Yes.</p>	<p>Yes Can't tell (only use if no information about time period for search) No Unsure</p>

	<p>Coding guide: YES: Generally this means searching the literature at least back to 1990 NO: Generally if the search does not go back to 1990 CAN'T TELL: No information about time period for search</p> <p>Note: With reference to the above – there may be important reasons for adopting different dates for the search, e.g. depending on the intervention. If you think there are limitations with the timeframe adopted for the search which have not been noted and justified by the authors, you should code this item as a NO and specify your reason for doing so in the comment box below. Older reviews should not be downgraded, but the fact that the search was conducted some time ago should be noted in the quality assessment. Always report the time period for the search in the comment box.</p>
--	---

<p>A.4 Was bias in the selection of articles avoided?</p> <p>Did the authors specify: A.4.1 Independent screening of full text by at least 2 reviewers A.4.2 List of included studies provided A.4.3 List of excluded studies provided</p> <p>Notes. For A.4.1, independent screening means that both screeners screened all full-text without knowing what the other screener decided (that is, one screener and one verifier does not meet criterion). If the authors note two screeners and do not use the word “independent” but mention a third reconciler to resolve differences, assume independence. Other acceptable methods include (a) the use of machine learning approaches (e.g., priority classifiers), provided a portion of machine excluded studies are checked or (b) double screening until an acceptable level of reliability (at least .85) is reached, with a percentage of subsequent coding being checked to protect against coder drift. If authors report double screening a small portion of studies, but do not report their inter-rater reliability, code No. When authors do not mention whether independent screening was conducted by at least two reviewers, code No. Single screening at title and abstract is acceptable.</p> <p>The list of excluded studies does not need to include studies whose abstracts were screened out as ineligible. Because journals often have word count limits, reviews published in journals do not need to have a list of excluded studies and are coded <i>Not Applicable</i>.</p>	<p>Yes Partially No</p> <p>Coding guide: YES: All three should be yes, although reviews published in journals are unlikely to have a list of excluded studies (due to limits on word count) and the review should not be penalised for this. PARTIALLY: Independent screening and list of included studies provided are both reported NO: All other. <u>If a list of included studies is provided, but the authors do not report whether or not the screening</u></p>
---	--

	<i>has been done by 2 reviewers, then this section is downgraded to NO.</i>
<p>A.5 Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included?^[10]</p> <p>A.5.1 The criteria used for assessing the quality/ risk of bias were reported</p> <p>A.5.2 A table or summary of the assessment of each included study for each criterion was reported</p> <p>A.5.3 Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). "Sensible" is defined as a recognised quality appraisal tool/ checklist, or similar tool which comprehensively assesses bias (internal validity) in included studies Please see footnotes for details of the main types of bias such a tool should assess.</p> <p><i>Notes. Identified tools with sensible criteria include: Academy of Nutrition and Dietetics Quality Criteria Checklist, Cochrane Handbook, The Delphi List, Effective Public Health Practice Project (EPHPP) Quality Assessment Tool, Guide to Community Preventative Services Study Quality tool, Joanna Briggs Institute Checklists for RCT/QED, National Institutes of Health's Quality Assessment Tool for Controlled Intervention Studies (sometimes labelled NHLBI tool).</i></p> <p>Child Health Epidemiology Reference Group (CHERG) study design & quality standards, Grading of Recommendations Assessment, Development and Evaluation (GRADE) RoB criteria, (CHERG and GRADE provide a set of guidelines for synthesizing evidence from multiple impacts on an outcome. As part of these multi-step processes, RoB is assessed, but other dimensions are also assessed (such as consistency of results across all studies). For A5.3, what needs to be reported is the individual ratings for each study on design/quality standards (CHERG) or risk of bias (GRADE)).</p> <p>For case-control studies and cohort studies, the Newcastle-Ottawa Scale uses sensible criteria that are focused on risk of bias as does Methodological Index for Non-Randomized Studies (MINORS). Note that these designs typically are not as rigorous as RCTs or even QEDs.</p>	<p>Yes Partially No Not Applicable (to be used only if there were no eligible quantitative studies)</p> <p>Coding guide: YES: All three should be yes PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY. NO: Any other</p>
<p>A.6 Overall – how much confidence do you have in the methods used to identify, include and critically appraise studies?</p> <p><i>Use the guidance below to determine the overall score for section A, based on your answers to each of the questions in this section.</i></p> <p>High confidence applicable when the answers to the questions in section A are all assessed as 'yes'</p> <p>Low confidence applicable when any of the following are assessed as 'NO' above: not reporting explicit selection criteria (A1), not conducting reasonably comprehensive search (A2), not avoiding bias in selection of articles (A4), not assessing the risk of bias in included studies (A5)</p> <p>Medium confidence applicable for any other – i.e. section A3 is assessed as 'NO' or can't tell and remaining sections are assessed as 'partially' or 'can't tell'</p>	<p>Low confidence (limitations are important enough that the results of the review are not reliable)</p> <p>Medium confidence (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously if a better review cannot be found)</p> <p>High confidence (only minor limitations)</p>

Section B: Methods used to analyse the findings

<p>B.1 Were the characteristics and results of the included studies reliably reported?</p> <p>Was there:</p> <p>B.1.1a Independent data extraction by at least 2 reviewers</p> <p>B.1.1b Independent risk of bias assessment by at least 2 reviewers</p> <p>B.1.2 A table or summary of the characteristics of the participants, interventions and outcomes for each included study.</p>	<p>Yes No Partially Not applicable (e.g. no included studies)</p>
---	---

<p>B.1.3 A table or summary of the results of all the included studies</p> <p><i>Notes.</i> Independent extraction means that both extractors extracted all data without knowing what the other extractor decided (that is, one extractor and one verifier does not meet criterion). If the authors note two reviewers and do not use the word “independent” but mention a third reconciler to resolve differences, assume independence. When authors do not mention whether independent extraction was conducted by at least two reviewers, code No. Forest plots are an appropriate summary of the results, as is reporting that summarizes the findings by outcome domain.</p>	<p>Coding guide: YES: All three should be yes PARTIALLY: Criteria B.1.1 and B.1.3 are yes, but some information is lacking on B.1.2. No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO. NOT APPLICABLE: if no studies/no data</p>
<p>B.2 Are the methods used by the review authors to analyse the findings of the included studies clear, including methods for calculating effect sizes if applicable?</p> <p><i>Note.</i> An example of acceptable reporting: “fixed effects meta-analysis, with standardized mean differences for continuous outcomes and response ratios for dichotomous outcomes”</p>	<p>Yes Partially No Not applicable (e.g. no studies or no data)</p> <p>Coding guide: YES: Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don't need to say this explicitly. PARTIALLY: Some reporting on methods but lack of clarity NO: Nothing reported on methods</p>
<p>B.3 Did the review describe the extent of heterogeneity?</p> <p>B.3.1 Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies?</p> <p>B.3.2 Did the review discuss the extent to which there were important differences in the results of the included studies? (Note, This item is not about which specific factors might explain differences in the results - that is covered in section B6)</p> <p>B.3.3 If a meta-analysis was done, was the I^2, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?</p> <p><i>Notes.</i> Code B.3.1 No if analyses includes studies with implausibly different interventions, comparisons, or populations. If a narrative analysis, the authors need to have a rationale for why studies were not combined (such as interventions were too different) or Code B.3.1 as No. For meta-analyses, reporting a metric for heterogeneity is sufficient for B.3.2. For non-meta-analysis, mentioning heterogeneity in results is enough (for example, The impacts varied from X to Y or Study A found X and Study B found Y).</p>	<p>Yes Partially No Not applicable (e.g. no studies or no data)</p> <p>Coding guide: YES: First two should be yes, and B.1.3 should be yes if applicable PARTIALLY: B.3.1 is yes NO: Any other NOT APPLICABLE: if no studies/no data</p>
<p>B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the <u>primary question</u> the review addresses and the available data?</p> <p>B.4.1 How was the data analysis done? Descriptive only Vote counting based on direction of effect</p>	<p>Yes Partially No Not applicable (e.g. no studies or no data) Can't tell</p>

<p>Vote counting based on statistical significance Description of range of effect sizes Random effects meta-analysis Fixed effects meta-analysis Meta-regression Bayesian approaches Network meta-analyses (NMA) Other: specify Not applicable (e.g. no studies or no data)</p> <p>B.4.2 How were the studies weighted in the analysis? Equal weights (this is what is done when vote counting is used) By quality or study design (this is rarely done) Inverse variance (this is what is typically done in a meta-analysis) Number of participants (sample size – this was standard practice in early meta-analyses) Other: specify Not clear Not applicable (e.g. no studies or no data)</p> <p>B.4.3 Did the review address unit of analysis errors? Yes - took clustering into account in the analysis (e.g. used intra-cluster correlation coefficient) No, but acknowledged problem of unit of analysis errors No mention of issue Not applicable - no clustered trials or studies included</p> <p>Note on B.4.1: There should be a clear justification if fixed effects meta-analysis is used. A fixed effects model assumes one true effect size, and that the only differences are due to sampling error. This is highly unlikely in international development due to large variations in context, participants, implementation, etc., thus a random effects model is typically most appropriate when meta-analysis is used.</p> <p>For network meta-analysis (NMA), the review must (1) discuss why NMA is appropriate, (2) present a network diagram where the thickness of the lines reflects the number of studies for each direct effect, (3) provide information on inconsistency factors and global test for inconsistency, (4) provide a table with the relative effect between each pair of interventions, and (5) provide a ranking of interventions using rankograms and cumulative ranking plots. Authors should interpret these graphs carefully if inconsistency in the network is detected.</p> <p>Bayesian approaches can be used for both meta-analysis and NMA. Prior distributions are needed for the particular intervention being analysed.</p> <p>Note on B.4.3: Unit of analysis issues arise when the unit assigned is a cluster, such as a school, but the units analyzed are individual people, such as students. If the analysis does not account for this clustering, the standard errors will be too large and accordingly the estimated statistical significance will be too small. Studies can account for the clustering using an appropriate hierarchical linear model or a random effects econometric model (note that random effects meta-analysis does not fix this problem, which exists at the study level). A systematic review can address these errors by requiring that the study use the correct analysis or by adjusting results using an intra-class correlation (typically the ICC is given a default value)</p>	<p>Coding guide: YES: If appropriate table, graph or meta-analysis (or descriptive where meta-analysis not possible and authors report magnitude of effects for all included studies) AND appropriate weights AND unit of analysis errors addressed (if appropriate). PARTIALLY: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors not mentioned or not addressed (and should have been). NO: If descriptive OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph or meta-analyses. NOT APPLICABLE: if no studies/no data CAN'T TELL: if unsure (note reasons in comments below)</p>
<p>B. 5 Does the review report evidence appropriately?</p> <p>B.5.1 The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately B.5.2 Where studies of differing risk of bias are included, results are reported and analysed separately by risk of bias status</p> <p>Notes. Making clear which evidence is subject to low risk of bias can be accomplished in a table listing RoB for each study or by listing RoB for each study on each RoB criterion; that</p>	<p>Yes No Partially Not applicable</p> <p>Coding guide: YES: Both criteria should be fulfilled (where applicable) NO: Criteria not fulfilled</p>

<p>is, if A5.2 is Yes, then B5.1 is Yes (but the reverse is not true). Reporting only study design is not sufficient to meet B5.1. For B5.2, narrative analysis must group or report by RoB, it is not sufficient to simply report RoB of each study. If the SR does not use sensible criteria to assess RoB, then B5.1 is <i>No</i>.</p> <p><i>Note on reporting evidence and risk of bias:</i> For reviews of effects of 'large n' interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of 'small n' interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders).</p> <p>For B.5.1, This item examines whether the SR clearly identifies which studies have low/high RoB, so that the reader understands the strength of evidence supporting each impact (the reporting can be for individual studies or an outcome domain). This differs from A5.2 (which examines the reporting of RoB at the criterion level) and B5.2 (which requires overall analysis/reporting by RoB). An overall GRADE quality of evidence rating cannot be used to meet this requirement because the GRADE rating is based on RoB but also additional factors such as consistency of results, indirectness of evidence, imprecision, and reporting bias. However, if the SR reports the RoB dimension separately (typically labeled "study limitations" or "risk of bias") for each outcome domain, that fulfills this criterion. For similar reasons, the overall CHERG quality assessment does not fulfill this requirement.</p> <p>Item B.5.2 applies only when there are low risk of bias studies included in analyses. If all studies in an analysis are deemed some concerns or high risk of bias, this point is not applicable.</p>	<p>PARTIALLY: Only one criterion fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate) NOT APPLICABLE: No included studies</p>
<p>B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies?</p> <p>B.6.1 Were factors that the review authors considered as likely explanatory factors clearly described?</p> <p>B.6.2 Was a sensible method used to explore the extent to which key factors explained heterogeneity?</p> <ul style="list-style-type: none"> Descriptive/textual Graphical Meta-analysis by sub-groups Meta-regression Other 	<p>Yes Partially No Not applicable</p> <p>Coding guide: YES: Explanatory factors clearly described and appropriate methods used to explore heterogeneity PARTIALLY: Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been) NO: No description or analysis of likely explanatory factors NOT APPLICABLE: e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore heterogeneity of the results</p>
<p>B.7 Overall - how much confidence do you have in the methods used to analyse the findings relative to the primary question addressed in the review?</p> <p><i>Use the guidance below to determine the overall score for section B, based on your answers to each of the questions in this section.</i></p>	<p>Low confidence (limitations are important enough that the results of the review are not reliable)</p>

<p>High confidence applicable when all the answers to the questions in section B are assessed as 'yes'.</p> <p>Low confidence applicable when any of the following are assessed as 'NO' above: critical characteristics of the included studies not reported (B1), not describing the extent of heterogeneity (B3), combining results inappropriately (B4), reporting evidence inappropriately (B5).</p> <p>Medium confidence applicable for any other: i.e. the "Partial" option is used for any of the 6 preceding questions and/or B.2 and/or B.6 are assessed as 'no'.</p>	<p>Medium confidence (limitations are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously if a better review cannot be found)</p> <p>High confidence (only minor limitations)</p>
---	---

Section C: Overall assessment of the reliability of the review

<p>C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?</p>	<p>Additional methodological concerns (e.g., reviews by a single author)</p> <p>Robustness</p> <p>Interpretation</p> <p>Conflicts of interest (of the review authors or for included studies) – note issues in comment section</p> <p>Other</p> <p>No other quality issues identified</p>
<p>C.2 Are there any mitigating factors which should be taken into account in determining the reviews reliability?</p>	<p>Limitations acknowledged (note, this is not a sufficient reason to upgrade a score, but should be noted in the assessment summary if limitations are acknowledged)</p> <p>Strong policy conclusions drawn (including in abstract/ summary) in the absence of high-quality evidence</p> <p>Any other factors</p> <p>Note. A low confidence review cannot be upgraded by simply acknowledging the limitations.</p>
<p>C.3 Based on the above assessments of the methods how would you rate the reliability of the review?</p> <p><i>Coding guide:</i></p> <p>High confidence in conclusions about effects: high confidence noted overall for sections A and B, unless moderated by answer to C1</p> <p>Medium confidence in conclusions about effects: medium confidence noted overall for both sections A and B or that you have assessed medium for A or B and high for the other section.</p> <p>Low confidence in conclusions about effects: low confidence noted overall for sections A or B, unless moderated by answer to C1 or C2. For example, if there is only one reason A or B is low confidence and there is a relevant mitigating factor that makes that reason less problematic, this can be assessed as Medium Confidence (e.g., the screening/extraction was not independent (leads to low) but two people screened/extracted all studies (for example, one checked the other and they report an acceptable level of reliability)).</p>	

Note. There are two cases where an SR can receive High Confidence even though was assessed Medium Confidence on Section A and the only reason for Medium is because (1) authors were not contacted to identify additional studies; however, the literature search involved multiple website searches, which serves an equivalent function, and (2) authors did not cross-checked references in all included studies; however, the authors did crosscheck all references in other review articles (at least two), which serves an equivalent function.

^[1] Adapted from Supporting the Use of Research Evidence (SURE) Collaboration. SURE checklist for making judgments about how much confidence to place in a systematic review. In: SURE guides for preparing and using policy briefs. www.evipnet.org/sure