



Desarrollo y evaluación de programas de seguridad ciudadana en América Latina

**Protocolo para la prevención
del delito a partir de la
evidencia**

Lawrence W. Sherman

**Universidad de Cambridge y
Universidad de Maryland**

**Banco
Interamericano de
Desarrollo**

Instituciones para el
Desarrollo
(IFD)

**NOTA TÉCNICA
IDB-TN-436**

Julio 2012

Desarrollo y evaluación de programas de seguridad ciudadana en América Latina

**Protocolo para la prevención del delito a partir de
la evidencia**

Lawrence W. Sherman

Universidad de Cambridge y
Universidad de Maryland



Banco Interamericano de Desarrollo

2012

Catalogación en la fuente proporcionada por la
Biblioteca Felipe Herrera del
Banco Interamericano de Desarrollo

Sherman, Lawrence W.

Desarrollo y evaluación de programas de seguridad ciudadana en América Latina : Protocolo para la
prevención del delito a partir de la evidencia / Lawrence W. Sherman.

p. cm. — (Nota Técnica del BID ; 436)

Incluye referencias bibliográficas.

1. Crime prevention—Latin America. 2. Public safety—Latin America. I. Banco Interamericano de
Desarrollo. Sector de Instituciones para el Desarrollo. IV. Título. V. Serie.

IDB-TN-436

<http://www.iadb.org>

Las opiniones expresadas en esta publicación son de los autores y no reflejan
necesariamente los puntos de vista del Banco Interamericano de Desarrollo, de su Junta
Directiva o de los países que ellos representan.

El uso comercial no autorizado de los documentos del Banco está prohibido y puede ser
sancionado de acuerdo con las políticas del Banco y/o las leyes aplicables.

Copyright © 2012 Banco Interamericano de Desarrollo. Todos los derechos reservados.
Puede reproducirse libremente para fines no comerciales.

Contacto: Lawrence.Sherman@crim.cam.ac.uk

+1-267-269-1757

Resumen

Este protocolo está destinado a todas aquellas personas que trabajen con miras a reducir el nivel de criminalidad y mejorar la justicia en América Latina, pero contiene principios que pueden usarse en cualquier parte del mundo. Esos principios pueden resumirse en la frase “prevención del delito a partir de la evidencia”, un proceso en virtud del cual se utiliza evidencia fehaciente sobre hechos delictivos y su prevención como base de las teorías y los programas para fomentar la seguridad ciudadana. El término “evidencia” se define ampliamente en el sentido de observaciones objetivas y sistemáticas de todo tipo, es decir que abarca mucho más que los detalles forenses de un caso de delincuencia. Constituyen evidencia los datos obtenidos mediante métodos científicos para observar y pronosticar cualquier tipo de verdad, lo que incluye hechos sobre salud, educación, delincuencia y justicia. Algunos de los 10 principios de prevención del delito son: el delito debe medirse en forma precisa y confiable mediante sistemas auditados exhaustivamente; el delito debe clasificarse de una manera que resulte útil para la prevención; y el delito debe analizarse en diversas unidades o categorías: delincuentes, redes criminales, víctimas, microzonas (“zonas calientes”), comunidades, horarios y días de la semana, entre otras.

Códigos JEL: Z18

Palabras clave: evaluación de impacto, Seguridad Ciudadana, costo-beneficio, prevención del crimen basada en la evidencia, eficiencia y efectividad.

Marco para la planificación de innovaciones de seguridad ciudadana

Introducción

¿De dónde provienen las buenas ideas? Para lograr que América Latina sea una región más segura, definitivamente se necesitan buenas ideas. Preguntarse cómo encontrarlas debe ser el punto de partida de cualquier iniciativa con miras a aumentar la seguridad ciudadana (Johnson, 2010). Esta pregunta viene acompañada de un segundo interrogante: cómo diferenciar las buenas ideas de otras no tan buenas (o malas) (Sherman, 2011). Sólo después de haber respondido a esas dos preguntas, se podrá resolver una tercera cuestión más vital: *cómo poner en práctica las buenas ideas para construir sociedades más seguras*.

Hay quienes sostienen que las buenas ideas provienen de la inspiración, o incluso de la intuición, mientras que otros creen que surgen de un análisis sistemático. Según el premio nobel Daniel Kahneman (2011), ambas opiniones son correctas. La gente piensa de las dos maneras, a menudo al mismo tiempo. Una de estas formas de pensar es *rápida* y la otra *lenta*, y están controladas por partes distintas del cerebro: lo que los psicólogos llaman “sistema I” y “sistema II”. El sistema I es automático y simple; no se tiene la sensación de estar controlando el proceso de pensamiento. El sistema II es complejo y metódico; requiere concentración y, muchas veces, cálculos. Cuando nos enamoramos, usamos el sistema I; si queremos construir una catedral, usamos el sistema II. Pero el amor también puede inspirar buenas ideas para construir una catedral.

Durante los últimos 500 años en Occidente, e incluso antes en Asia, los pensadores han intentado separar las dos formas de razonamiento. Desde los trabajos de Francis Bacon y René Descartes hasta la obra *Blink, The Power of Thinking without Thinking* de Malcolm Gladwell (2005), se han alabado los méritos del pensamiento metódico por sobre los peligros del automático. Tal como sostiene Gladwell, a veces un juicio instantáneo da mejores resultados que una decisión elaborada, pero el autor también documenta los muchos desastres que se han producido por confiar únicamente en el sistema I prestando muy poca atención al sistema II de pensamiento. Para lograr buenos resultados en un problema complejo, se necesita tanto un buen criterio basado en la experiencia como un análisis detallado de los hechos y las teorías de causalidad.

Propósito

Este protocolo proporciona un marco para mejorar la seguridad ciudadana mediante el sistema II de pensamiento. En otras palabras, está concebido para complementar el buen criterio de líderes experimentados con análisis sistemáticos que respalden ese criterio. El método que utiliza está ampliamente probado y el concepto es muy poco original. Lo original es el intento de conectar las prácticas de seguridad ciudadana con uno de los inventos más exitosos de la historia de la humanidad: la ciencia.

Utilizar un método científico no implica el uso de delantales de laboratorio, microscopios o tubos de ensayo. La ciencia no es más que una forma sistemática de organizar la evidencia relacionada con cualquier teoría sobre cómo funciona el mundo y extraer conclusiones a partir de ellas. Pero no está en absoluto exenta de polémica. Ha sido ampliamente adoptada y también atacada, alabada y castigada, según los caprichos de quienes detentaran el poder en cada momento. Regímenes democráticos y dictatoriales por igual han rechazado evidencia fehaciente porque se oponían al saber del sentido común o los juicios instantáneos. La importancia del pensamiento intuitivo en el voto cívico ha disuadido incluso a los líderes más inteligentes de basarse en evidencia concreta. Sin embargo, en lo que se refiere a la prevención del delito, cualquier sociedad puede en segundos llegar a la conclusión de que los métodos tradicionales no son lo suficientemente buenos; es necesario probar métodos nuevos y diferentes como innovaciones.

Entonces, ¿qué es lo que este protocolo puede aportar a los vastos conocimientos (y las intensas emociones) que muchos ciudadanos y líderes ya tienen acerca de las innovaciones en materia de justicia y prevención del delito? La respuesta se resume en una sola palabra: comprobación. La validez de una idea no puede demostrarse mediante la lógica ni el instinto. La única manera de separar las buenas ideas de las malas es ponerlas a prueba en el mundo real.

La mayoría de las innovaciones en prevención del delito se ensayan, pero no se comprueban (Sherman, 2011). Es decir, las personas que tienen fe en la innovación quieren ensayarla poniéndola en práctica, pero se niegan a comprobarla por comparación directa. Sostienen que comparar la nueva idea con otra más antigua no es ético, porque ya saben que la idea nueva funciona mejor. Este argumento suele usarse en medicina a fin de avalar

nuevos “remedios milagrosos” (Millenson, 1997), sólo para descubrir que resultan en desgracias como el nacimiento de niños sin miembros (por ejemplo, debido a la talidomida) o un enorme desperdicio de dinero (por ejemplo, por amigdalectomías). En la larga historia de la medicina, el buen criterio de médicos con experiencia se fue complementando cada vez más con requerimientos legales relativos a la *comprobación* de la seguridad y la eficacia de cualquier medicamento antes de lanzarlo al mercado.

Una verdadera comprobación debe estar concebida de manera tal que la innovación sometida a ella pueda fracasar. Sin embargo, como dicen los líderes policiales ingleses con ironía, muchas veces las innovaciones están “condenadas al éxito”. Ensayar algo sin comprobarlo es contrario a la ciencia, ya que una verdadera comprobación debe establecer claramente la *manera* en que una hipótesis cualquiera puede ser refutada (Popper, 1959). Las hipótesis irrefutables son el fundamento de la fe, la esperanza y otras cosas buenas, pero no la forma en que los seres humanos en general han resuelto problemas complejos como el delito.

Este protocolo establece los pasos que cualquier persona puede seguir con miras a desarrollar y comprobar innovaciones para fomentar la seguridad ciudadana usando la mejor evidencia disponible. Incluye el uso de un marco teórico para explicar *cómo y por qué* las innovaciones funcionan o no. También abarca el rol de la implementación y mediciones para el cumplimiento de las tareas de las que dependen los buenos resultados. Sobre todo, indica cómo utilizar el principio de *comparación justa* para llegar a conclusiones a partir de la evidencia.

En el protocolo se utiliza la sigla DRIVER para marcar los seis pasos de una innovación a partir de la evidencia:

D corresponde al diagnóstico de la naturaleza y las causas directas del problema de delincuencia

R a la revisión de las prácticas existentes para encarar dicho problema

I a la implementación de las prácticas revisadas según fueron concebidas

V al valor agregado y la relación costo-beneficio que en teoría tiene la innovación

E a la evaluación de la implementación y la teoría de la innovación

R a la reflexión sobre nuevas formas de revisar y comprobar las prácticas.

Los pasos del DRIVER se describen en detalle en la próxima sección y, más adelante, en una lista de comprobación para cada paso. A fin de comprender su lógica, resulta útil partir de los 10 principios que se enumeran en el resumen del inicio y que se describen en detalle a continuación.

10 principios de prevención del delito a partir de la evidencia

1. El delito debe *medirse* en forma precisa y confiable mediante sistemas auditados exhaustivamente. Una buena medición es tan importante en la lucha contra el crimen como en el pronóstico del tiempo, la producción de alimentos, el crecimiento económico y muchas otras cuestiones de las sociedades complejas. Durante los últimos dos siglos, se desarrollaron dos sistemas importantes para medir el delito. Uno se centra en contabilizar las denuncias efectuadas a la policía; el otro, más reciente, consiste en utilizar encuestas de victimización.

La contabilización de los delitos informados a la policía siempre está incompleta debido a la decisión de los ciudadanos de no reportar ciertos crímenes. Además, es posible que la policía cuente menos delitos que los que se cometen en realidad, por presiones políticas para que las tasas de criminalidad parezcan menores. Las encuestas de victimización constituyen un elemento valioso para verificar que esas amenazas no afecten la validez de las mediciones. Pero no pueden reemplazar a la policía en la recopilación de datos detallados sobre cada delito, como la ubicación, la fecha y la hora exactas.

Los países más seguros del mundo, desde Japón hasta Suecia, hacen gran hincapié en que la policía registre en formato digital los detalles de todos los crímenes y delitos denunciados, lo cual hace que los informes sean más útiles a la hora de realizar un análisis criminológico. En el Reino Unido se ha dado incluso más importancia al control de los informes policiales, con puestos especiales para “encargados del control de registros” que periódicamente revisan la forma en que se registran los delitos en las comisarías. Sin embargo, ese país aún no logra proporcionar datos precisos sobre la ubicación de cada delito; un conocido sitio Web (www.police.uk, en inglés), que supuestamente se encarga de eso, contiene crasos errores en los mapas del delito (Palmer y Caddick, 2010).

La tecnología más avanzada para marcar los lugares donde se cometen delitos proviene de los sistemas de posicionamiento global por satélite (GPS), que la policía ya utiliza en algunos países en desarrollo (como India), pero no en la mayoría de los países del G-20. La adopción tanto del GPS como de las cámaras con micrófonos que ahora la policía puede portar para grabar las entrevistas con víctimas y testigos de un delito puede hacer que la policía ya no tenga que volver a hacer jamás un informe. En cambio, puede descargarse la grabación digital de los datos del video y del audio de cada turno de servicio y enviarse a los encargados especializados del control de registros delictivos, quienes pueden avanzar rápidamente en las grabaciones a fin de encontrar y codificar los reportes de delitos tomados por los agentes con la ayuda de una lista de comprobación que se puntea durante las entrevistas.

2. El delito debe *clasificarse* de una manera que resulte útil para la prevención. La herramienta más importante para la medición del delito la conforman los sistemas de información por episodio (IBRS, por sus siglas en inglés), en lugar de registrarse solamente totales por categoría y jurisdicción policial. Introducidos en Estados Unidos por el FBI en la década de 1980, este tipo de sistemas permite clasificar y analizar los delitos de diversas formas. Esta flexibilidad brinda a la policía y a los encargados de la formulación de políticas la posibilidad de adaptarse a los cambiantes patrones de criminalidad, reorganizando la forma en que se miden los delitos en función de necesidades operativas. Esta idea es diametralmente opuesta al concepto tradicional de utilizar el código penal como único sistema de clasificación. Además, proporciona la capacidad de contar con la información oportuna y precisa sobre el delito que ha distinguido a iniciativas policiales basadas en la evidencia, tales como el sistema COMPSTAT implementado en Nueva York y otras ciudades (Bratton, 1998).

¿Por qué no es conveniente utilizar únicamente el código penal para clasificar los delitos? La respuesta está en la gran diversidad de hechos que abarca cada categoría legal. Por ejemplo, en el caso de los homicidios hay muchos tipos de armas que pueden usarse para matar a alguien, pero la mayoría de los usos de esas armas no causan la muerte. Por lo tanto, el *comportamiento* de dispararle a alguien puede clasificarse en distintas categorías de delito según si la bala no alcanza a la persona a la que se le apuntó, si la lesiona pero no

la mata o si efectivamente le causa una lesión fatal. Los códigos penales ubican los disparos que no causan la muerte en la misma categoría legal que las peleas a golpes de puño y los apuñalamientos. No hay forma de realizar un seguimiento de las tendencias sobre el total de delitos con armas de fuego usando únicamente el código penal. Sin embargo, las armas de fuego pueden ser (y muchas veces son) el principal problema que debe tratarse para disminuir la cantidad de lesiones en delitos violentos, tal como demostraron las investigaciones llevadas a cabo en Bogotá y Cali (Villaveces et al., 2000).

La vigilancia policial orientada a los problemas (Goldstein, 1990; Braga y Bond, 2008) y otras estrategias ambientales de prevención del delito pueden requerir incluso más detalles sobre los actos delictivos. Estos detalles pueden incluir las condiciones de iluminación, la cercanía a tabernas y los horarios en los que esas tabernas están abiertas, la ubicación de la autopista más cercana y las condiciones climáticas. Ninguna de estas características puede usarse para clasificar los delitos de una manera que sugiera mejores prácticas de prevención, a menos que se cree un registro computarizado por cada episodio. No se trata de un requisito burocrático; los IBRS constituyen la fuente principal de inteligencia para la actividad estratégica.

3. El delito debe *analizarse* en diversas unidades o categorías: delincuentes, redes criminales, víctimas, microzonas (“zonas calientes”), comunidades, horarios y días de la semana, entre otras. Una vez que se han recopilado datos sobre cada episodio, se los puede organizar y depurar de acuerdo con muchas formas de orientar las prácticas de prevención del delito. Una forma habitual de hacerlo es por delincuente. Sin embargo, son pocos los delincuentes que actúan solos; en la mayoría de los delitos, hay coautores, o delincuentes que forman parte de redes sociales de actividad criminal. Algunas veces esas redes reciben el nombre de “pandillas”, pero con mayor frecuencia se trata de formas desestructuradas de contacto intermitente, un contacto similar al que tiene lugar a través de Facebook. Estas redes pueden también incluir a víctimas del delito, en especial hombres jóvenes que están en mayor riesgo de sufrir lesiones o la muerte por actos de violencia. Otras víctimas pueden no tener conexión social alguna con los victimarios, pero convertirse repetidamente en blanco de delitos en formas que pueden evitarse mediante tareas de prevención.

Cada vez hay más evidencia de que el enfoque tradicional en las *personas* puede ser menos útil en la prevención del delito que el enfoque en los *lugares*. Aparentemente, la concentración de delitos en microzonas o “zonas calientes” de actividad criminal (Sherman, Gartin y Buerger, 1989) es más intensa que la concentración por delincuente (Weisburd, Sherman y Maher, 1993). Analizar el crimen y el delito por unidades geográficas más pequeñas puede revelar que, incluso en comunidades con altas tasas de criminalidad, en la mayoría de los lugares no se cometen actos delictivos. Esto significa que puede ser un desperdicio utilizar recursos tales como los que se emplean para la vigilancia policial en calles en las que no ha habido asesinatos en 10 años. En un análisis reciente por microzonas de los asesinatos cometidos en Londres en un período de 10 años se halló que el 42% de los miles de homicidios se había cometido en apenas el 6% de los muchos miles de unidades geográficas pequeñas en que se dividió la ciudad para el análisis (Jackson, 2011). Asimismo, Weisburd et al. (1993) informaron que más de la mitad de los crímenes y delitos cometidos en Seattle en más de 15 años se habían producido en apenas un 5% de los segmentos de calles.

Estas altas concentraciones de actos criminales en ciertos lugares pueden combinarse con la concentración por persona. Wikstrom et al. (2010) descubrieron que incluso los delincuentes frecuentes eran mucho más propensos a delinquir en “zonas calientes” que en zonas con bajos índices de criminalidad. Cuando se llevan estos datos a un plano temporal, resulta sencillo pronosticar qué días de la semana y en qué horarios hay más probabilidades de que los delincuentes más riesgosos cometan un delito contra las víctimas de mayor riesgo.

Sin embargo, en ciertos tipos de crímenes la materia más útil para el análisis puede seguir siendo el delincuente. La violencia doméstica es un ejemplo primordial de la necesidad de contar con bases de datos específicas de delincuentes, aunque las bases de datos sobre víctimas también pueden resultar útiles. El valor de ambas clases de datos reside en su capacidad para construir modelos predictivos que ayuden a la policía y a los agentes a establecer prioridades en el caso de los delincuentes o víctimas que más probablemente se vean envueltos en homicidios o daños graves.

4. Deben establecerse *prioridades* de prevención entre los tipos de unidades según el grado de *daño* ocasionado por las distintas clases de delitos. Siempre que los presupuestos son limitados, es necesario tomar decisiones sobre la manera en que se concentrarán los recursos. Estas decisiones establecen prioridades que muchas veces se distorsionan cuando se agrupan todos los tipos de delitos juntos. Muchos gobiernos, incluidos el de Estados Unidos y el del Reino Unido, informan la absurda cifra de “delitos totales” denunciados por año. Según esa métrica, un homicidio es lo mismo que un hurto, y una violación es lo mismo que un robo en una tienda. Ningún código penal sugiere eso. En muchos países hoy existen pautas para el dictado de sentencias con recomendaciones muy precisas sobre la cantidad de días de prisión (incluso cero) que debe tomarse como punto de partida de cualquier condena. Esa cantidad de días o meses en prisión puede usarse como una base razonable sobre la cual determinar el daño ocasionado por cada tipo de delito (Sherman, 2010).

El concepto de daño ocasionado por un delito es un concepto racional, del sistema II, que armoniza con los puntos de vista políticos y emocionales que corresponden al sistema I. El tiempo de encarcelamiento recomendado puede usarse como patrón de evaluación del daño con el fin de establecer prioridades transparentes para las inversiones en prevención. Además, puede servir como base sobre la cual redistribuir la “cartera” de inversiones; por ejemplo, entre prisión y libertad condicional, medidas correctivas y control policial, o situación de riesgo en la infancia y rehabilitación de adultos.

Con el objeto de que el enfoque sea transparente, puede usarse un índice del daño causado por el delito (CHI, por sus siglas en inglés), que multiplica el nivel de daño que ocasiona cada delito por la cantidad de delitos de ese tipo. Con una herramienta de esta naturaleza, es posible reemplazar un confuso conjunto de estadísticas criminales según las cuales algunos delitos aumentan y otros disminuyen, porque el índice brinda un único “calculo final” para saber si la actividad delictiva ha ocasionado más daño en un año que en el anterior, en una comunidad que en otra, o en un grupo de exconvictos que en otro. Al utilizar un índice de este tipo en lugar de mediciones tan rudimentarias como el porcentaje de delincuentes que vuelven a ser arrestados o condenados (reincidencia), es posible hacer que las evaluaciones de impacto de las estrategias de prevención sean mucho más informativas.

Los resultados calculados mediante un CHI son los únicos que realmente equivalen al análisis de pérdidas y ganancias de un entorno comercial. Mientras que en los informes comerciales los ingresos siempre se expresan en una única divisa, en los de la justicia penal no es así: cada tipo de delito es, en efecto, una “moneda” distinta. Nadie sumaría pesos, euros y dólares para saber cuántas unidades de “dinero” tiene; lo normal sería elegir una única divisa en la cual expresar todas las cantidades de dinero, que luego pueda convertirse a cualquier otra. Eso es exactamente lo que un CHI puede aportar a la prevención y al análisis del delito.

5. Una *minoría fuerte*, o una pequeña proporción de todas las unidades de conducta criminal, ocasiona la mayor parte del daño de la mayoría de los tipos de delitos.

Independientemente de cómo se mida el daño ocasionado por el delito, lo cierto es que está muy concentrado en un segmento pequeño de cualquier distribución (Sherman, 2007). Ya sea entre delincuentes o víctimas, lugares o momentos del día, el riesgo de delitos en el futuro varía ampliamente. No puede representarse mediante una curva acampanada (distribución normal) que abarque todas las unidades en riesgo. Prácticamente todos los análisis del delito por cualquier unidad muestran una *curva de potencia*, la forma de palo de hockey que identificó por primera vez el economista Vilfredo Pareto para describir la riqueza y los ingresos.

Esta distribución es evidente y potencialmente polémica. Es evidente, por ejemplo, que las zonas con altas tasas de criminalidad sufren mucho más daño a causa del delito que las demás. Sin embargo, es polémico proponer que las zonas con bajas tasas de criminalidad reciban menos atención que las que tienen una tasa alta, ya que posiblemente los contribuyentes de las zonas con bajas tasas de criminalidad paguen más impuestos y entonces exijan mayor protección. Es difícil explicar que corren un riesgo mayor, donde sea que vayan en su sociedad, si a las zonas con altas tasas de criminalidad se las priva de inversiones suficientes en prevención del delito.

Concentrarse en las minorías fuertes puede aportar beneficios mucho más allá de las zonas inmediatas en las que esas minorías viven o cometen delitos. Cuando se reduce la violencia en sus epicentros, toda la sociedad resulta beneficiada, incluso aunque no se preste atención a las zonas periféricas de menor criminalidad. Para explicar este concepto a

la población en general, se necesita liderazgo, valentía y habilidad. Es una idea formulada con el sistema II que a la vez puede incorporarse al entendimiento intuitivo del sistema I. “¡Desde luego que se necesitan más policías donde los tiroteos son más habituales!” El carácter evidente de afirmaciones como esta bien puede contrarrestar la idea de “desnudar a un santo para vestir a otro” implícita en la redistribución de impuestos según las necesidades. Quizá los mejores argumentos para explicarlo sean los conceptos de eficiencia y eficacia.

6. La eficiencia de la prevención del delito puede ser mayor cuando los recursos se concentran en las unidades de minorías fuertes, identificadas mediante pronósticos formulados a partir de su comportamiento anterior. Este principio deriva de la lógica pura y, como muchos otros argumentos de ese tipo, depende de un supuesto fundamental sobre hechos empíricos. El supuesto es el siguiente: si la misma práctica reduce el daño por delitos en la misma proporción en todas las unidades al mismo costo, es en las unidades con mayor nivel de criminalidad donde se puede obtener un mayor rendimiento total de la inversión en reducción de daños. Por ejemplo, si contar con un auto de policía para vigilar una zona cuesta US\$50 por hora, y el nivel de criminalidad es un 50% más bajo cuando hay vigilancia policial que cuando no la hay, se prevendrán más delitos (y más daño) por cada dólar invertido en una zona con alto nivel de criminalidad que en una zona con pocos delitos. Al menos, eso es lo que se supone.

La evidencia que avala este supuesto no es sólida, pero sí coherente. Ratcliffe et al. (2011) demostraron que, con la misma cantidad de vigilancia policial a pie por metro cuadrado en Filadelfia, la disminución del delito fue mayor en las rondas realizadas en las zonas en las que había un mayor nivel de criminalidad, y menor a medida que los niveles de criminalidad descendían. Desde ya, la validez de este hallazgo sólo es limitada debido al tamaño reducido de la muestra, y la prueba no se ha repetido. El principio 6 abarca mucho más que actividades de patrullaje. Pero se ha registrado al menos un descubrimiento similar, el de Olds et al. (1997), en un programa de visitas de enfermeras altamente calificadas a madres primerizas y sus bebés, en el que las reducciones del maltrato infantil más eficaces en función de los costos tuvieron lugar en las familias de mayor riesgo. De manera similar, en un programa de justicia restaurativa para delincuentes juveniles en

Australia, hubo una mayor reducción de la reincidencia entre quienes cometían delitos con más frecuencia (Woods, 2010).

Este principio (y su supuesto fundamental) tal vez pueda respaldarse mejor mediante una evaluación sistemática de toda la evidencia posible. Hasta que se complete tal evaluación, tanto la teoría como los hechos respaldan la eficacia de concentrarse en las minorías fuertes.

7. La mejor forma de determinar la *eficacia* de la prevención del delito es mediante pruebas de campo que comparen las tasas de criminalidad entre personas, zonas u otras unidades de análisis bajo distintos programas. Distinguir la eficacia de la eficiencia es sumamente importante. La eficiencia sólo es posible cuando las prácticas son eficaces; es decir, cuando producen los resultados que se pretendía lograr con ellas, como prevenir el delito. Un programa ineficaz nunca puede llevarse a cabo con eficiencia. No obstante, un programa eficaz puede estar ineficientemente asignado. Incluso puede no “funcionar” con algunos tipos de personas, o hacer que el delito aumente en ciertos segmentos de la población y disminuya en otros (Sherman, 1992). La única forma de determinar la eficacia de una iniciativa de prevención del delito es comprobarla en condiciones reales.

¿En qué consiste una “comprobación”? La mejor definición de “comprobación” es que se trata de una comparación de los resultados en cuanto a la actividad criminal en unidades similares, con y sin el uso de la práctica que se está poniendo a prueba. Las unidades, vale aclarar, pueden ser delincuentes, víctimas, lugares o días de la semana. La práctica que se comprueba puede compararse con otra práctica o con no hacer absolutamente nada. La utilización de un “grupo de control” en un experimento tiene lugar únicamente a fin de comparar resultados. Para sostener que una evaluación de impacto demuestra que una práctica “funciona”, es necesario por definición realizar una comparación: ¿“funciona” en comparación con qué?

La elección del término de comparación es una parte vital del desarrollo de ideas en prevención del delito. Por ejemplo, cuando se comparó el monitoreo electrónico de delincuentes en libertad condicional en Estados Unidos con los casos en los que no se utilizaba ese sistema, los hallazgos pueden haber indicado que el monitoreo ofrecía escasos

beneficios (Renzema y Mayo-Wilson, 2005). Pero cuando se lo comparó con la cárcel en Buenos Aires, los resultados demostraron que el monitoreo lograba que el nivel de criminalidad fuera menor por un costo más bajo (Di Tella y Schargrodsky, 2010).

La forma más clara de comprobar una práctica de prevención es compararla con no hacer absolutamente nada. Esta claridad proviene del hecho de conocer exactamente cuál es el costo de llevar a cabo la práctica que se está comprobando y saber que el costo de la comparación es, claramente, “cero”. La falta de acción puede ser más aceptable desde el punto de vista ético si en las unidades en cuestión no se ha hecho nada hasta el momento. En cualquier caso, si se dispone de recursos limitados para una nueva idea, será imprescindible utilizar algún tipo de proceso de selección. Esto llevará a un enfrentamiento entre los principios de “necesidad” y de “evidencia”, pero sólo a corto plazo. Ofrecer un programa a todos aquellos que más lo necesitan posiblemente signifique que el programa no prospere por falta de evidencia sobre su eficacia. En cambio, si el programa se ofrece a la mitad de la población que más lo necesita, podrán obtenerse evidencia válida de su eficacia en función de los costos. A largo plazo, todos salen beneficiados de saber qué funciona y qué no. Como mínimo, una comprobación controlada de este tipo puede servir para descubrir si un programa bueno en teoría resulta perjudicial en la práctica (McCord, 1978; McCord, 2003).

Las comparaciones con otras prácticas también pueden producir comprobaciones claras. En este caso, la claridad proviene de la asignación imparcial de una u otra práctica a cada unidad. Cuando el Departamento de Policía de Mineápolis decidió asignar al azar a sospechosos de violencia familiar para que fueran arrestados o recibieran orientación psicológica, el hecho de que la asignación de una u otra práctica se hiciera de forma aleatoria contribuyó a disminuir el nivel de parcialidad en la comparación de la reincidencia entre los dos grupos de más de 100 sospechosos (Sherman y Berk, 1984).

La parcialidad podría haberse presentado, por ejemplo, si se optaba por arrestar a los sospechosos que eran más hostiles hacia la policía y que tal vez tuvieran más probabilidades de reincidir en la supuesta violencia. O si se arrestaba a quienes ya tenían antecedentes policiales en vez de aquellos que nunca habían sido arrestados. En ese experimento, la asignación al azar no fue perfecta, pero hoy la tecnología moderna permite

reducir la parcialidad de una forma mucho más eficaz mediante plataformas de asignación aleatoria a través de Internet (Ariel y Sherman, 2012).

Las comprobaciones por comparación mediante la asignación al azar en general se denominan “experimentales”, mientras que aquellas en las que la asignación no es aleatoria suelen llamarse “cuasiexperimentales”. Estos diseños se han calificado con una puntuación de 1 a 5, en la que el 5 corresponde a experimentos con asignación aleatoria y el 1 a una correlación observada. Un diseño de nivel 2 consiste simplemente en una comparación de antes y después del mismo grupo de unidades; es decir, se comparan los resultados de delincuencia previos a la introducción de la práctica con los resultados posteriores a ella.

¿Cuál es el inconveniente de realizar una comparación de antes y después? El problema es la falta de control. Los resultados podrían cambiar, o no, por muchos motivos además del uso de las prácticas sometidas a comprobación. El hecho de no contar con un grupo de comparación o de control significa que no hay forma de calcular qué habría pasado si no se hubiera incorporado la nueva práctica, ya que, en materia de delitos, muchas veces las tendencias fluctúan naturalmente o por casualidad. Una comparación de antes y después puede indicar que un programa no funcionó cuando en realidad sí funcionó, o viceversa. Se toman demasiadas decisiones multimillonarias a partir de este tipo de comprobaciones, algo que jamás se permitiría en otras áreas, como la medicina o la agricultura. El delito es lo suficientemente importante para merecer una comprobación con estándares más rigurosos que las comparaciones de antes y después (Campbell y Stanley, 1966).

Tres ejemplos de evaluaciones de seguridad ciudadana en países de América Latina ilustran de diversas maneras el valor de utilizar comparaciones. Uno de ellos (Villaveces et al., 2000) es la evaluación de impacto del programa policial para prohibir la portación de armas en personas y automóviles durante los fines de semana largos en Bogotá y Cali. Ya que los homicidios por armas de fuego aumentaban en dichos fines de semana, el personal médico que lidiaba con los tiroteos tenía un interés en ver si la acción policial podía reducir los daños. La proacción consistió en establecer controles policiales de carretera donde chequear a personas y automóviles en busca de armas ocultas, lo cual teóricamente funcionaría de la siguiente manera: el chequeo en busca de armas conduce a una menor portación de armas de fuego, lo cual lleva a que se reduzca la densidad de armas de fuego en lugares públicos, lo cual a su vez hace que se produzcan menos tiroteos por impulso, ya

que hay menos armas disponibles. Al comparar la violencia en los fines de semana largos antes y después del comienzo del programa, y al realizar comparaciones con otros fines de semana largos en los que el programa no estaba en operación, los autores estuvieron en condiciones de estimar el modesto pero significativo impacto del programa en la prevención de homicidios.

Un segundo ejemplo de este tipo se acerca bastante al ideal de la prueba controlada al azar. Di Tella y Schargrodsky (2010) comparan casos graves de delincuentes sentenciados en Buenos Aires por jueces cuyas preferencias diferían en cuanto a la forma de sancionar los delitos. Algunos jueces optaban por encarcelar a los criminales, mientras que otros elegían el arresto domiciliario con monitoreo electrónico. Como los casos se asignaban a los distintos jueces mediante una fórmula numérica aleatoria, es probable que otras posibles causas de cualquier diferencia en la repetición del delito afectaran por igual a ambos grupos y por lo tanto hubiera que controlar esas diferencias. Los autores observaron que en realidad los delincuentes condenados al arresto domiciliario mostraban tasas de repetición del delito inferiores que las de aquellos sentenciados a prisión (un hallazgo compatible con otros estudios que demuestran que las sentencias a prisión aumentan las tasas de criminalidad más que el hecho de no condenar a prisión al mismo tipo de delincuentes). Esta evidencia es muy robusta y resulta útil para el desarrollo de las políticas de sentencias en los países latinoamericanos.

El tercer ejemplo es menos riguroso, pero también demuestra un buen uso de las comparaciones controladas. Soares y Viveiros (2010) evaluaron la adopción de una estrategia de integración del tipo COMPSTAT para diferentes fuerzas policiales en Minas Gerais, Brasil. El programa de gestión policial se moldeó sobre la base del sistema COMPSTAT desarrollado por el Departamento de Policía de la Ciudad de Nueva York en los años noventa (Bratton, 1998). Los autores compararon las tasas de criminalidad en las zonas donde la policía de Minas Gerais había adoptado el sistema con las de otras zonas en las cuales esto no se había efectuado. Mediante el uso de una comparación de antes y después entre las zonas en las que se había adoptado el sistema y aquellas en las que no se lo había adoptado, durante los mismos períodos, los autores pudieron concluir que el programa había reducido los delitos a la propiedad pero no los homicidios.

El tercer ejemplo contiene una útil enseñanza en cuanto a cómo evaluar un programa. Cualquier evaluación que se fuerce para fines comparativos luego de los hechos se ve limitada en las clases de diferencias que pueden compararse. Es difícil asegurarse de que las manzanas estén siendo comparadas con manzanas y no con naranjas; esto es: las comparaciones de zonas con y sin programas se realizan mediante el uso de unidades que tienen similares problemas delictivos, demografía, etc. Los autores hicieron lo mejor que pudieron con lo que tenían, pero cuando el desarrollo de un programa y su evaluación van de la mano, puede haber una mejor correspondencia entre las unidades experimental y de comparación. También puede prestarse una mayor atención a detalles como asegurarse de que las zonas de comparación no sean adyacentes a las zonas experimentales, con lo cual estarían sujetas a un desplazamiento potencial del crimen o efecto de arrastre. Pero aun así hay formas de hacer que incluso las evaluaciones más modestas de asignaciones aleatorias sean controladas con mayor rigor.

Aun cuando se disponga de una cantidad pequeña de unidades, hay formas de incrementar la validez de los efectos causales estimados. El uso de estos principios permite que los estudios planificados puedan realizar comparaciones mucho más confiables, incluso aunque no puedan elevarse al nivel de una prueba aleatoria. Siempre existen motivos por los cuales no se pueden realizar pruebas altamente controladas de este tipo. Cuando las comprobaciones se realizan en zonas enteras u otras unidades grandes, por ejemplo, es posible que no haya una cantidad de unidades suficiente para garantizar la imparcialidad con una asignación totalmente azarosa. En esos casos, una posibilidad es agrupar 20 o 30 unidades de a pares lo más estrechamente posible. Luego, dentro de cada par, puede asignarse al azar la práctica sometida a comprobación a una unidad, pero no a la otra (véase, por ejemplo, Braga y Bond, 2008). Entonces, podrá considerarse que la diferencia promedio de los resultados de cada par equivale al efecto que la práctica tiene sobre el delito, en relación con la práctica (o la falta de intervención) con la que se la comparó.

Muchas veces, los experimentos de asignación aleatoria se rechazan por sus costos. Sin embargo, en el Reino Unido y Estados Unidos, recientemente se ha demostrado que es posible llevar a cabo buenas comprobaciones con asignación aleatoria por un costo inferior al de otros tipos de evaluaciones (Sherman, Strang y Ariel, 2011) y aportar evidencia más

fehaciente. Esto también significa que dicha evidencia podrá explicar mejor los resultados, en especial en lo referido a la teoría de cómo y por qué la práctica funciona o no.

Por último, la base de cualquier comprobación es una idea clara de la meta que se pretende alcanzar mediante la práctica o el programa de prevención del delito. Nunca se hicieron ensayos para determinar si la penicilina curaba la poliomielitis, la diabetes o el cáncer, sino que los estudios siempre se centraron en la curación de infecciones. Sin embargo, en materia de criminalidad, muchas iniciativas de prevención se justifican mediante el argumento de que previenen diversos tipos de delitos, sin explicar claramente cómo funcionan en teoría.

8. La teoría de toda práctica de prevención del delito debe especificar una secuencia lógica mediante la cual esa práctica genera *productos* en forma de acciones cuantificables que disminuyen los *resultados* en materia de delitos o injusticia, explicando *cómo* y *por qué* se espera que la práctica prevenga el delito. Las personas a menudo se muestran escépticas cuando se afirma que algo funciona (o no) para prevenir el delito, aun cuando se cuente con evidencia obtenida en buenos experimentos. Antes de aceptar la conclusión, necesitan comprender por qué eso funciona. La verdad es que, como los equipos de fútbol, los experimentos no siempre pueden explicar los resultados con razones convincentes; sencillamente se limitan a informar los resultados con sinceridad para que otros los interpreten. Tanto en el contexto de los deportes como en las evaluaciones de impacto, las grandes teorías satisfactorias, como las de “liderazgo” o “compromiso con el equipo”, son difíciles de probar.

Sin embargo, lo que cualquier experimento puede hacer es describir quién debe ocuparse de hacer qué, tal como ocurre en los equipos de fútbol. Si el arquero no impide que la pelota entre en el arco, el equipo perderá; si la policía no realiza los patrullajes, el experimento no mostrará una reducción del delito. Toda práctica de prevención del delito debe distinguir entre los *productos* en forma de acciones que determinadas personas deben realizar y los *resultados* que se espera obtener a partir de esos productos. Y lo que es aún más importante: tanto los productos como los resultados deben ser cuantificables.

Los productos se miden con más facilidad que los resultados; en la mayoría de los casos, su medición puede hacerse a través del recuento de las unidades que se entregarán.

En un experimento, se aumentó el patrullaje en algunas zonas calientes del delito (mientras que en otras no) y estudiantes de posgrado cronometraron cuántos minutos de vigilancia policial hubo en 100 zonas (Sherman y Weisburd, 1995). Estas observaciones permitieron calcular la proporción entre los minutos de patrullaje policial en el grupo experimental y los minutos de patrullaje en el grupo de control. Dicha proporción fue el principal indicador del nivel de producción de vigilancia policial que, según la hipótesis del experimento, tendría efecto sobre la tasa de criminalidad (y, de hecho, lo tuvo). Con el doble de patrullaje en las 50 zonas del grupo experimental que en las 50 zonas del grupo de control, el experimento reveló una clara diferencia en la dosis de vigilancia policial. Luego, pudo demostrarse que la diferencia en el nivel de criminalidad era directamente proporcional a la diferencia en el patrullaje.

En la actualidad, no hace falta tanta mano de obra para realizar las mediciones. Los radios portátiles de los agentes de policía en general cuentan con dispositivos de GPS incorporados. Esta medición pasiva de su ubicación en todo momento permite que un solo analista de datos pueda reconstruir la cantidad de minutos de vigilancia policial en cada zona del experimento a un costo muy bajo. Sin embargo, lo esencial en todo experimento es medir la diferencia de producción entre el grupo experimental y el grupo de comparación.

A veces los productos son múltiples y consecutivos, y es preciso medir cada paso por separado. Por ejemplo, en un programa extracurricular de deportes, el primer producto que debe medirse es el tiempo que el programa está abierto al público: no el horario *programado*, sino el tiempo real de funcionamiento, medido de manera independiente por alguien que no tenga relación alguna con la ejecución del programa, o bien mediante tecnologías pasivas tales como circuitos cerrados de televisión. El segundo producto podría ser la inscripción en el programa: ¿cuántos niños se han anotado para asistir al programa cinco veces por semana? El tercer producto podría ser la participación real: en promedio, ¿cuántos niños realmente asistieron al programa de deportes cada día? ¿Cuánto tiempo se quedaron?

Los productos de esta iniciativa de prevención del delito mediante un programa de deportes podrían concordar con la teoría criminológica de las “actividades de rutina” (Cohen y Felson, 1979): cuanto más tiempo estén los niños de alto riesgo bajo la supervisión de un adulto, menor es la probabilidad de que comentan un delito o sean

víctimas de uno. Esta teoría depende de buenas mediciones de la manera en que los niños ocupan su tiempo. No se trata de un requisito burocrático, sino de evidencia esencial. Si se rompe alguno de los eslabones de la cadena siguiente, no habrá motivos para esperar que el programa prevenga el delito:

Centro de deportes abierto → niños inscriptos → participación de los niños → menos oportunidades para el delito → menos delitos cometidos por niños o en contra de ellos.

Debe tenerse en cuenta que esta teoría es únicamente acerca de los niños inscriptos en el centro de deportes, no sobre la tasa de homicidios en la ciudad o el país donde el centro está ubicado. Las personas que proponen actividades de este estilo muchas veces sugieren una gran falta de conexión entre la provisión del programa y un resultado probable. Detallar la teoría, con todos los productos, es una buena forma de encaminarse hacia una teoría que explique cómo y por qué el programa debería modificar las tasas de criminalidad. Medir los productos puede revelar si la *teoría* en sí era errónea o si simplemente no hubo entrega de productos (o esta no alcanzó el umbral del nivel mínimo). Además, permite mantener los resultados estrechamente relacionados con la teoría sobre la forma en que los productos los modifican, en lugar de medir los resultados aislados de los productos.

En la teoría mencionada sobre el centro de deportes, el eslabón que quedó sin medir es la cantidad de oportunidades para el delito. Este es un claro ejemplo de una “caja negra” de causalidad: aunque se entreguen todos los productos, puede que igualmente no sepamos si los niños inscriptos en el programa tuvieron menos oportunidades de cometer delitos o ser víctimas de un delito. Entonces, si el programa no previene el delito, podría ser por ese motivo, o bien porque el programa simplemente no funcionó como se pensaba. De todas maneras, más allá de que este tipo de eslabones clave se midan o no, cualquier conexión entre la entrega de productos y las diferencias en los resultados de criminalidad sigue siendo una contribución importante al conocimiento. En la prevención del delito a partir de la evidencia, no se permite que lo mejor sea enemigo de lo bueno. Si se mide la mayoría de los productos, su entrega es suficiente para probar la forma básica de la teoría.

9. La ética de una práctica de prevención del delito no depende de su éxito en reducir el delito, sino de si la práctica a) respeta los derechos humanos y b) mantiene un equilibrio proporcional entre los daños causados por coacción y los daños causados por delitos. Este es un principio que debe aplicarse a cualquier iniciativa de prevención del delito, ya sea pública o privada. Es esencial destacar su importancia en el marco claramente utilitarista (Beccaria, 1764; Bentham, 1781) que ofrece este protocolo. Aunque una práctica sea eficaz según algunas mediciones, no puede considerarse aceptable si viola los derechos humanos o infringe conceptos éticos fundamentales. Algunos policías creen que torturar a sospechosos de delitos es una forma eficaz de solucionar los problemas de criminalidad, pero, desde el punto de vista ético, esa hipótesis no puede comprobarse. Incorporar alternativas a la tortura, como la recolección de muestras de ADN en las escenas del crimen, puede ser una forma éticamente aceptable de intentar reducir esta práctica. Pero siempre debe mantenerse en claro cuál es la línea que distingue lo que “funciona” de lo que “es ético”.

Algo que no está tan claro, y que muchas veces se convierte en una cuerda floja por la que deben caminar los gobiernos, es el concepto de balance de daños. Muchas prácticas de la justicia penal invaden la libertad de formas que se saben dañinas, como el arresto o el encarcelamiento. Un ejemplo excelente lo constituyen los métodos de detención y cacheo por parte de la policía, en los que se detiene y registra a personas inocentes en zonas de alto nivel de criminalidad con el objetivo de evitar muertes por heridas de bala (Villaveces et al., 2000). Esos métodos son claramente invasivos. Sin embargo, su validez desde el punto de vista ético tal vez dependa por completo de la relación costo-beneficio que generen (Sherman y Rogan, 1995). Si mediante esas operaciones policiales es posible salvar vidas sin herir ni causar daños permanentes a nadie, muchos jueces y votantes dirían que son éticas, mientras que otros tal vez estén en desacuerdo.

En esos casos, la línea que divide lo correcto de lo incorrecto no puede ser clara, pero puede hacerse más nítida con evidencia obtenida en buenas evaluaciones de impacto.

10. Las mejores pruebas para desarrollar y mejorar prácticas de prevención del delito a nivel local deben provenir tanto de evaluaciones de impacto internacionales como de análisis criminológicos locales de todo tipo, que incluyan evaluaciones de iniciativas usadas en el pasado y pronósticos sobre los patrones de criminalidad futuros. ¿De dónde provienen las buenas ideas? Proviene de cualquier parte y de todas partes, a nivel mundial y local; de la intuición del sistema I y del razonamiento del sistema II. Las publicaciones científicas sobre prevención del delito son numerosas en todo el mundo, y aumentan cada día más. En lo cotidiano, también pueden surgir buenas ideas repentinamente, pero de formas menos accesibles o transparentes. Tal vez lo más importante sea la evidencia sistemática sobre los problemas de criminalidad a nivel local, que señalan las prioridades y los objetivos principales sobre los que se puede actuar para reducir el daño ocasionado por el delito.

Muchas innovaciones en materia de prevención del delito se ponen en práctica sin que nadie jamás revise las publicaciones científicas mundiales. El peligro de trabajar de esta manera no es que siempre se esté reinventando la rueda, que es algo útil sin importar de dónde provenga la idea. El peligro es reinventar un neumático pinchado, que nunca funcionó como rueda en ninguna comprobación anterior de ninguna parte del mundo. Es posible que un programa que parece una buena idea según el juicio instantáneo del sistema I ya se haya sometido a comprobación metodológicamente con métodos del sistema II, y que los resultados hayan indicado que la idea es no sólo ineficaz, sino también perjudicial.

Los programas que intentan “aterrorizar” a jóvenes en riesgo son un ejemplo primordial de la importancia de revisar las evaluaciones de impacto internacionales (Petrosino, Turpin Petrosino y Buehler, 2002). Siete evaluaciones distintas han demostrado que, en promedio, llevar a jóvenes preadolescentes a visitar las cárceles no los “aterroriza”, sino que, por el contrario, aumenta su participación en actividades delictivas. Si alguien tiene la idea de comenzar una práctica de este tipo, debe advertírsele que primero revise las evaluaciones de impacto internacionales para conocer la evidencia (véase, por ejemplo, http://www.campbellcollaboration.org/reviews_crime_justice/index.php, en inglés).

Al buscar en fuentes internacionales, pueden encontrarse muchos otros programas que tal vez sean perfectos para las condiciones locales de criminalidad. El reciente desarrollo de una matriz para conectar evaluaciones de estrategias policiales con problemas criminológicos específicos es una indicación de la creciente facilidad con que se pueden realizar dichas búsquedas (véase <http://gemini.gmu.edu/cebcp/matrix.html>, en inglés).

Sin embargo, tal como se explica en la siguiente sección, el paso más importante para desarrollar soluciones eficaces a nivel local es realizar un análisis sistemático de los problemas de criminalidad locales. La capacidad de realizar un análisis de este tipo nos lleva nuevamente al primer principio de prevención del delito a partir de la evidencia, y a nuestro primer paso del modelo DRIVER. Pero, antes de dar ese paso, proporcionaremos dos elementos clave: algunas definiciones y un breve manual de uso.

Innovaciones, prácticas, programas y protocolos

En este protocolo se han usado varios términos casi como sinónimos. Sin embargo, es importante señalar algunas diferencias entre ellos.

Una *innovación* puede definirse como cualquier práctica o programa que no se ha utilizado en una ciudad o un país en el pasado reciente.

Una *práctica* es casi cualquier cosa que alguien haga con honesta confianza en que puede prevenir algunos delitos. Una práctica no necesariamente debe formar parte de un programa organizado; puede ser simplemente algo que las personas hacen en forma individual o colectiva. Un padre que le da una nalgada a su hijo por golpear a su hermano puede creer que de esa forma está previniendo el delito al imponer límites a la violencia (aunque cierta evidencia sugieren que castigar a los niños con nalgadas en realidad la promueve, ya que demuestra la aceptación por parte de los padres de la violencia como forma de solucionar conflictos) (Straus, 1996). Un maestro que enseña a sus alumnos a leer puede creer que eso ayudará a prevenir el delito porque fomenta el alfabetismo y aumenta las oportunidades de conseguir un empleo en una economía cada vez más globalizada. Un enfermero que denuncia a un padre por maltrato infantil puede estar previniendo el delito porque rescata al niño de una familia violenta. Un agente de policía que arresta a una

persona por un acto menor de violencia familiar (como una bofetada) puede creer que eso reducirá el riesgo de violencia en el futuro.

Un *programa* puede definirse como un conjunto de prácticas organizadas que en general implican la colaboración de dos o más personas, casi siempre en el contexto de una organización formal. Los programas pueden incluir las políticas de la policía y organismos de justicia penal, escuelas, y organizaciones recreativas, culturales y de promoción de la salud.

Un *protocolo* es un conjunto de principios y prácticas para llevar a cabo un programa. El término se utiliza mucho en el ámbito de la medicina para especificar los pasos que deben seguirse en una cirugía, el control de epidemias, un parto u otros procedimientos habituales. Con la esperanza de que el desarrollo de programas de seguridad ciudadana pueda volverse un procedimiento habitual en todo el mundo, este protocolo se propone hacer que dicho desarrollo pase de los palpitos del sistema I a los análisis formales del sistema II.

Cómo utilizar este protocolo

Hay dos formas de usar este protocolo. Una consiste en tomar un principio de prevención del delito a partir de la evidencia por vez, y desarrollar innovaciones que ayuden a aplicarlo a nivel local. La otra forma es seguir el proceso DRIVER completo de principio a fin. A menos que haya buenas mediciones del delito, puede que el proceso DRIVER completo no sea posible. Por lo tanto, es necesario establecer formas adecuadas de tomar mediciones antes de intentar utilizarlas en un proceso DRIVER completo. Una vez que esas formas de medición estén implementadas, podrán ejecutarse varios procesos DRIVER en paralelo.

Marco del proceso DRIVER

El proceso DRIVER: desarrollo por medio de la evaluación

DRIVER es un protocolo para desarrollar, comprobar y perfeccionar prácticas de prevención del delito en un proceso de mejora continua de la calidad a partir de la evidencia (Deming, 2000). La sigla, que en inglés remite a la palabra “conductor”, traza una analogía entre la importancia de la evidencia para prevenir el delito y para conducir un automóvil. Es posible, pero no recomendable, conducir un vehículo sin un tablero con información sobre la velocidad y los niveles de combustible, aceite y agua, o sin un parabrisas a través del cual pueda verse lo que hay alrededor. De la misma manera, es posible, pero no recomendable, prevenir el delito sin contar con información clara y esencial.

Los seis pasos del proceso DRIVER

Diagnóstico del problema específico que debe resolverse

El proceso DRIVER comienza con el diagnóstico de las causas y concentraciones de los patrones de criminalidad de interés. El primer paso para el diagnóstico es el *desglose*; es decir, dividir grandes problemas en partes más pequeñas y manejables. El delito suele dividirse en las siguientes categorías: crímenes y delitos violentos, delitos contra la propiedad, delitos relacionados con drogas y disturbios. También es posible realizar la división desde otros ángulos; por ejemplo, delitos que involucran a menores o a mayores de 18 años, con víctimas de sexo femenino o de sexo masculino, cometidos por delincuentes con o sin antecedentes, con o sin uso de armas, cometidos por personas que consumen o que no consumen drogas, cometidos o no por pandillas, o cometidos por pandillas vinculadas a cárteles de narcotráfico o grandes organizaciones criminales.

No hay una única forma correcta de diagnosticar los patrones de delincuencia en un país o comunidad. Las numerosas opciones disponibles sólo son más o menos útiles y adecuadas. Pero más allá de la opción que se elija, hay determinadas herramientas básicas que aumentarán la comprensión de las dimensiones y las trayectorias futuras del problema,

tanto por parte de los especialistas como por parte del público. Las *tendencias* en materia de criminalidad a lo largo del tiempo, y en distintas categorías, proporcionan información valiosa sobre el futuro inmediato y de más largo plazo en cuanto a las amenazas a la seguridad ciudadana. Los *patrones* de criminalidad por zona, horario o modus operandi (por ejemplo, robos en viviendas por distracción, o asaltos con allanamiento de morada para robar vehículos o las llaves de vehículos) pueden mantenerse constantes incluso aunque las tendencias de cantidad o frecuencia aumenten o disminuyan. Los *pronósticos* sobre personas que pueden cometer asesinatos u otros daños severos son cada vez más posibles gracias al avance de la minería de datos y las grandes bases de datos sobre antecedentes policiales (Berk et al., 2009).

Estas herramientas brindan diversas opciones para realizar un diagnóstico que establezca *prioridades*, con decisiones sobre cómo definir los *objetivos* a los que apuntar. En este contexto, el término objetivos se refiere, en primer lugar, a los tipos de unidades de análisis que se abordarán, como zonas con delitos frecuentes o delincuentes peligrosos en general. En segundo lugar, elegir un objetivo significa enumerar las unidades específicas que incluye cada categoría, como determinadas intersecciones de calles, jóvenes en riesgo en zonas con altas tasas de delincuencia, o ciertos delincuentes peligrosos que acaban de salir de la cárcel.

A fin de lograr la máxima eficiencia, el diagnóstico puede apuntar a unidades de las “minorías fuertes” en cada categoría de diagnóstico: aquellas unidades en las que los delitos son más numerosos y más graves. Otra posibilidad es orientarlo hacia las unidades más “alcanzables”: aquellas en las que hay más posibilidades de obtener resultados exitosos con el menor esfuerzo. El diagnóstico debe terminar con la fijación de *objetivos* concretos que sean lo más precisos posible. Es mucho más útil establecer las dimensiones precisas de la actividad delictiva que debe prevenirse que apuntar al delito en general o, incluso, a categorías específicas de delitos. Por ejemplo, si cada vez hay más robos a bancos, el objetivo podría ser reducir los robos a bancos. Pero si en algunos bancos hay más casos de disparos a empleados o clientes durante los asaltos, el objetivo podría ser reducir las lesiones y muertes producidas en los robos a bancos. Cada uno de esos objetivos podría requerir prácticas de prevención distintas.

Revisión de las prácticas para hallar soluciones innovadoras para cada problema

Una vez identificados y seleccionados todos los objetivos, se puede seguir un rumbo distinto dentro del proceso DRIVER para cada uno de ellos. El punto de partida de ese rumbo es la revisión de las prácticas y los programas existentes que pueden apuntarse a cada objetivo. La revisión puede hacerse para una categoría de objetivos (por ejemplo, bancos en los que hay asaltos varias veces por año) o para cada uno de los elementos de una categoría (por ejemplo, la sucursal del Banco XXX ubicada en Avenida Principal 1405).

Más allá de cuál sea el objetivo, la revisión de las prácticas existentes debe hacerse con información obtenida mediante una búsqueda de ideas a nivel local y mundial. Una búsqueda a nivel mundial de evaluaciones de impacto de medidas de seguridad bancaria, por ejemplo, proporcionará una serie de estudios que cualquier ciudad puede tener en cuenta (véase la lista de comprobación de este protocolo). Del mismo modo, una revisión de los datos locales también puede indicar algunos caminos para revisar las prácticas de prevención.

Si en los bancos en los que suelen producirse robos es menos probable que haya guardias armados que en los que no hay tantos asaltos, para intentar alcanzar el objetivo de disminuir la cantidad de robos a bancos, se podría hacer que la policía se reúna con los gerentes de los bancos para sugerirles que contraten a guardias armados. Si, por otro lado, en los bancos en los que hay guardias armados se producen menos robos, pero cuando estos ocurren hay *más* tiroteos y personas heridas, el objetivo del proceso DRIVER de reducir las heridas de bala podría llevar a recomendar la disminución de la cantidad de guardias armados.

Este ejemplo debería servir para ilustrar la forma en que puede conectarse la evidencia local con la mundial. Una vez completada la evaluación de las prácticas existentes, las prácticas mundiales y las consecuencias de los patrones y tendencias recientes en materia de delitos, el proceso DRIVER debe llevar a tomar una decisión sobre el programa o la política que se utilizará para facilitar el logro de los objetivos concretos establecidos.

Implementación de la práctica revisada

Cualquier intento de generar un cambio en organizaciones comunitarias o gubernamentales requiere un plan de implementación, que debe incluir lo siguiente:

- Una secuencia cronológica que indique qué se hará.
- Una identificación clara de quién hará qué, cuándo y con qué frecuencia.
- Uno o más indicadores que deban registrarse para medir si las tareas asignadas (productos) se están llevando a cabo y, de ser así, en qué medida.
- Una asignación clara de la responsabilidad por la recopilación de las mediciones de la implementación.
- Una estructura para que los líderes del proyecto puedan controlar periódicamente dichas mediciones.

Valor agregado y relación costo-beneficio que en teoría tiene la innovación

Antes de poner en práctica una revisión innovadora de las prácticas o programas, de acuerdo con el proceso DRIVER también debe planificarse la teoría de la revisión. A partir del plan de implementación, en el análisis del valor agregado debe calcularse la eficacia que tendrá la innovación en función de los costos.

Por ejemplo, hacer que ante el delito de portación de armas deba imponerse una pena de prisión podría llevar a que la población carcelaria se triplicase, o incluso se multiplicase por 10. Mediante un análisis de los datos judiciales pertinentes, podría calcularse el costo del cambio propuesto en el presupuesto del sistema penitenciario. También podría estimarse la cantidad de delitos con uso de armas que podrían prevenirse únicamente mediante la inhabilitación de los portadores de armas. Un recuento de los delitos con armas cometidos por personas con condenas anteriores por portación de armas podría cumplir bien ese propósito, y mejoraría en gran medida las meras conjeturas hechas hasta el momento. Si la proyección estimara que la práctica sólo prevendría dos delitos con armas cada 100 presos por año y el costo de mantener a una persona en prisión durante un año fuera de US\$25.000, el costo de prevenir dos delitos con armas sería de

US\$2,5 millones, lo que equivale a US\$1,25 millones por delito. Si sólo en uno de cada cuatro delitos con armas una persona resulta herida, el costo de prevenir una herida de bala sería de US\$5 millones.

Incluso aunque no sea posible calcular con precisión el valor de prevenir una herida de bala, una proyección de este tipo serviría como referencia para comparar la propuesta con la relación costo-beneficio de otras propuestas. También serviría como base para repetir el análisis con los datos reales de la implementación de la innovación, si es que se decide proceder tras evaluar el valor agregado y el costo-beneficio que la innovación tiene en teoría.

Evaluación de la implementación y la teoría de la innovación

Una vez tomada la decisión de poner en práctica una innovación, el paso siguiente del proceso DRIVER consiste en evaluar la implementación y la teoría de la innovación. La evaluación puede realizarse de manera constante en sesiones de control de la implementación. Y lo que es más importante: puede proporcionar a los líderes del proyecto de prevención del delito una advertencia temprana de que la innovación no se está implementando tal como se planificó. En ese caso, los líderes podrán decidir si es preferible suspender la innovación, hacer correcciones sobre la marcha en la gestión o el personal, o continuar con mensajes claros al personal correspondiente sobre la necesidad de una mejor implementación.

En el caso del experimento de vigilancia policial en zonas calientes de Mineápolis (Sherman y Weisburd, 1995), el aumento exponencial de la razón entre la cantidad de vigilancia policial en las zonas experimentales y las zonas de control hizo que se estableciera el requisito de que los policías presentaran registros diarios del tiempo que pasaban en las zonas que se les habían asignado. A su vez, esos registros se compararon con las mediciones independientes de las horas de presencia policial en las zonas asignadas. El resultado de este control permanente desde un comienzo fue el éxito de la implementación y un impacto importante en los tipos de delitos a los que se apuntaba.

La teoría de cualquier innovación de prevención puede expresarse mediante conexiones causales tanto medidas como no medidas. En el caso del estudio de Sherman y

Weisburd (1995), la teoría establecía que una mayor presencia policial aumentaría la amenaza percibida en general para disuadir el delito. Parte de la teoría dependía de tener una mayor presencia policial, lo cual se midió. Lo que no se midió fue si los delincuentes en potencia realmente percibían la amenaza o si al menos decidían por intuición (sin razonarlo) que no querían cometer delitos en zonas donde la presencia policial se había hecho mucho más visible.

Otra dimensión que no se midió de la teoría de la vigilancia policial en zonas calientes fue la hipótesis de que los crímenes disuadidos en una zona surgen en otras. Si bien Weisburd et al. refutaron sustancialmente esa hipótesis, fue prácticamente imposible hacer observaciones en ese sentido a medida que la innovación progresaba. Esto dejó a la cúpula policial con cierto grado de duda acerca de si la concentración de la vigilancia en zonas calientes realmente estaba disminuyendo la actividad criminal en la ciudad en general o si únicamente la estaba trasladando a otras áreas.

De todas maneras, con el objetivo claro de reducir determinados delitos en las zonas calientes del grupo experimental en comparación con los mismos tipos de delitos en zonas calientes similares (asignadas al azar) en las que no se aumentó la vigilancia, la comprobación de la teoría no pudo refutar su lógica. La innovación se implementó, la teoría funcionó y la estrategia se continuó, al menos en el largo plazo, si no en el corto. Muchas otras ciudades han adoptado la estrategia desde entonces, así como también países de todo el mundo.

Reflexión sobre nuevas formas de revisar y comprobar las prácticas

Más allá del éxito que tenga una innovación, siempre hay espacio para mejorar. En aras del progreso continuo, una experiencia exitosa debería llevar a la misma pregunta que un fracaso: ¿cómo se puede hacer un mejor trabajo? Formularse esta pregunta puede conducir directamente al comienzo de todo proceso DRIVER: una evaluación global de los últimos hallazgos de investigación, las nuevas tecnologías o los más novedosos programas e ideas que se hayan ensayado a nivel local o en otros lugares.

Esto es exactamente lo que ha sucedido en las dos décadas posteriores al éxito del primer experimento de vigilancia en zonas calientes: se identificaron numerosas

oportunidades para perfeccionar y ampliar la teoría de la innovación. En Mineápolis, posteriormente se realizó un nuevo diagnóstico en el que se identificó que el problema consistía en predecir y prevenir la aparición de nuevas zonas calientes, en lugar de vigilar las zonas ya existentes. En Jacksonville, Florida, se compararon dos estrategias de vigilancia en zonas con alto nivel de criminalidad y se descubrió que la vigilancia preventiva era menos eficaz a largo plazo que las estrategias de resolución de problemas mediante vigilancia policial orientada a los problemas (Taylor, Koper y Wood, 2011). En Filadelfia, se hizo una comprobación de las rondas de vigilancia a pie en zonas calientes, en virtud de la cual se hallaron reducciones en el nivel de criminalidad, cierta evidencia de desplazamiento, y una reducción neta de la actividad delictiva tras descontar de los beneficios el desplazamiento calculado (Ratcliffe, et al., 2011).

Un proceso DRIVER que no muestre ninguna mejora en un problema de criminalidad podría ser el resultado de la falta de un grupo de control. Es poco probable que todas las innovaciones puedan acompañarse con evaluaciones de impacto con asignación al azar. En la sección sobre el desarrollo de programas basados en la evidencia se explican los diseños cuasiexperimentales para la medición del impacto, pero incluso estos pueden resultar demasiado complicados para muchas organizaciones sin la ayuda de las universidades locales. Por lo tanto, si el enfoque está puesto en una tendencia delictiva que no disminuye, la forma más simple de avanzar es volver a analizar los delitos cometidos para detectar cualquier cambio en los patrones de criminalidad. Si en el nuevo análisis se observa que ha habido modificaciones en la actividad delictiva, la siguiente pregunta será si el cambio se ajusta a la teoría de la innovación. Si la respuesta es negativa, tal vez sea mejor volver a la etapa de diagnóstico y buscar una estrategia completamente nueva. En cambio, si los patrones de criminalidad cambiaron, tal vez sea preferible perfeccionar la innovación tal como se la desarrolló.

Formas en las que *no* deben diseñarse los programas

Una buena forma de desarrollar la mejor práctica es estudiar las prácticas malas. A los deportistas, muchas veces los entrenadores les explican cómo *no* deben correr o patear la pelota. A los cantantes, los entrenadores de la voz les indican la forma en que *no* deben

cantar una nota aguda; a los cirujanos, se les enseña cómo no hacer un corte; y a los arquitectos, se les explica cómo un mal diseño puede hacer que un edificio se venga abajo. En muchos campos, las investigaciones demuestran que, para poder realizar bien una tarea compleja, hacen falta 10.000 horas de práctica (Gladwell, 2005b; Kahneman, 2011: 238). Sin embargo, muchas personas que crean programas de seguridad ciudadana no tienen la ventaja de contar con experiencia en el área, y mucho menos con un instructor que les indique qué es lo que no deben hacer.

El objetivo de esta sección es cumplir ese rol de “instructor” para quienes diseñan programas de seguridad ciudadana con poca o ninguna capacitación para esa tarea. Aquellos lectores que ya hayan cometido los errores que se mencionan deben saber que la intención de este texto no es en absoluto faltar el respeto a su trabajo. Este análisis se proporciona únicamente con la intención de ayudar a todas las personas a diseñar mejores programas, con una mejor comprensión del proceso DRIVER.

Pero antes de comenzar, algunos comentarios sobre la información. Durante 10 años, en la Universidad de Pensilvania, el proceso DRIVER fue el tema de un curso de 15 semanas y 45 horas de clase con una cantidad importante de lecturas. En el curso se formaron equipos de cinco estudiantes de alta calificación graduados en administración pública y criminología. A cada equipo se le asignaba un objetivo en el cual concentrar el proceso DRIVER; por ejemplo, lograr una reducción de los casos de maltrato infantil o violencia con armas. Luego, se le pedía que preparara una presentación de 15 minutos que abarcara los seis pasos del proceso DRIVER. A pesar de todas las charlas, lecturas y debates, los equipos cometían algunos errores frecuentes, que se tomaron como base para esta sección.

Además, se incorporó información de evaluaciones de cientos de propuestas presentadas ante entidades estatales y federales estadounidenses con el objeto de obtener fondos para proyectos de prevención del delito, y algunas propuestas similares presentadas ante distintas competencias y programas de financiamiento en países de América Latina. Por último, esta sección también refleja errores comunes de órganos de justicia penal de Europa, Asia, Asia del Sur, Australia y América que el autor ha observado desde 1970.

Primero el programa, después el problema

Sin duda, el error más frecuente es inventar un programa y después buscar un problema que sirva para justificarlo. Este es particularmente el caso cuando la invención no se crea a partir de una evaluación de la evidencia o la teoría sobre las causas y la prevención de un determinado problema de delincuencia. El diseño del programa surge en un chispazo de intuición, sin otro fundamento para determinar sus posibilidades de éxito. Este tipo de ideas pueden provenir de lo que los medios de noticias británicos llaman “gobierno de sofá” en referencia a las ideas que surgen cuando las personas (primeros ministros o jefes de policía) conversan informalmente sentadas en un sillón.

EJEMPLO: El programa de Educación Preventiva contra el Consumo de Drogas (DARE, por sus siglas en inglés). Por su nombre, parecería ser que el programa apuntaba a reducir el consumo de drogas, pero en realidad consistía en enviar policías a las escuelas para que hablaran en persona con niños preadolescentes. El programa se extendió como reguero de pólvora en Estados Unidos posiblemente porque la imagen de un oficial de policía uniformado en un aula se veía muy bien en televisión. Incluso se extendió a otros países. Mientras tanto, distintas evaluaciones de impacto cuasiexperimentales indicaban que el programa no tenía ningún efecto en el consumo de drogas por parte de los niños que lo recibían.

Entonces, comenzó a debatirse si las evaluaciones de impacto se estaban llevando a cabo correctamente o si el período de seguimiento era lo suficientemente prolongado. El trasfondo de estos debates eran los indicios de que el programa tal vez no evitaba el consumo de drogas, pero en cambio lograba con mucho éxito que los niños confiaran en la policía y admiraran a los oficiales. En otras palabras, el compromiso con el programa en sí mismo desencadenó un proceso de búsqueda de objetivos que pudieran lograrse con él, en lugar de comenzar con un objetivo concreto y buscar el mejor programa para alcanzarlo. Ante las evaluaciones de impacto, el alcalde de Salt Lake City, Utah, decidió cancelar el financiamiento que la ciudad otorgaba al programa DARE, tras lo cual los seguidores del programa intentaron separar al alcalde de su cargo.

Irónicamente, en el mismo período, el tema de la confianza en los oficiales de policía cobró más importancia teórica y política (National Research Council, 2004). Es perfectamente posible que el programa DARE resulte eficaz para ampliar la legitimidad de las instituciones policiales, lo cual puede ser un factor independiente que contribuya a que las personas respeten la ley (Tyler, 1990). Si el programa DARE hubiera sido concebido para mejorar la actitud de los niños hacia los oficiales de policía, bien podría haber sido un programa exitoso. Pero el punto sigue siendo el mismo: el proceso DRIVER debe comenzar con un problema que haya que resolver, no con un programa que se desee promover.

Primero el programa, después la teoría sobre su valor agregado

Otro error es concebir un programa o una práctica de manera intuitiva, y luego esforzarse para formular una teoría que explique cómo y por qué el programa debería aportar valor al logro de los objetivos de prevención del delito.

EJEMPLO: Programas de arte y recreación, como la propuesta de la Liga Atlética de la Policía estadounidense conocida como “baloncesto de medianoche”, que fue rechazada por el Congreso en la década de 1990. Como en el caso del programa DARE, la imagen de jóvenes participando en actividades que fomentan las relaciones sociales es intuitivamente atractiva. Siguiendo un juicio instantáneo del sistema I, puede recibir mucho apoyo. Sin embargo, quienes proponen estos programas tienen dificultades a la hora de especificar cómo es que pueden reducir el nivel de criminalidad. El argumento es que el tiempo que los niños pasan “fuera de la calle” y bajo la supervisión de adultos puede hacer que no se metan en problemas. Sin embargo, ese argumento depende de la cantidad de horas por día que los niños estén supervisados, y la cantidad de días por semana que puedan asistir a los programas. También depende de que los programas no generen enfrentamientos entre personas (o grupos) que provoquen episodios de violencia fuera del ámbito de supervisión.

Otra teoría que suele sugerirse luego de que se lleva a cabo este tipo de programas es que la recreación mejora la autoestima de los niños y su confianza en sí mismos. Sin embargo, no hay evidencia:

- 1) De que la recreación mejore la autoestima de los niños, en especial en situaciones competitivas en las que la mayoría de los participantes pierde ante unos pocos ganadores.
- 2) De que, aunque la recreación logre que los niños tengan más confianza en sí mismos, esa sensación sea un factor de protección contra el delito.

El proceso DRIVER funciona mejor si se comienza con una teoría criminológica establecida (y comprobada) y luego se planifican actividades a partir de esa teoría. Un ejemplo es la teoría del control (Hirschi, 1969). Esta teoría sostiene que es menos probable que un niño se convierta en delincuente si tiene lazos más fuertes con la sociedad y los valores morales tradicionales. Estos lazos pueden fortalecerse mediante los siguientes elementos:

- *Apego* a familias y personas que respetan la ley.
- *Compromiso* con las normas y las instituciones sociales (la escuela, un empleo).
- *Participación* en actividades convencionales.
- *Creencia* en la importancia de estos elementos.

Esta teoría implica que, en realidad, es posible que no sea mejor para los niños estar en entornos con un enfoque puesto en sus necesidades o con relaciones de poder horizontalistas, donde posiblemente no se aprecien los valores morales tradicionales. En cambio, sugiere que quizá sea más importante que los niños pasen tiempo en ámbitos centrados en los adultos o como mínimo en actividades muy estructuradas directamente dirigidas por adultos con un control riguroso.

Entonces, comenzar con la teoría del control podría hacer que se prefieran orquestas sinfónicas antes que actividades deportivas o coros antes que actividades plásticas. Cualquier actividad grupal con control estricto y espíritu de inclusión puede fomentar mayor apego, compromiso, participación y creencia en estos elementos que actividades individuales más competitivas o solitarias. El proceso DRIVER puede partir de este tipo de

“traducción” de la teoría a la práctica para elegir prácticas que se ajusten a la teoría, y no al revés.

Varios programas sin vínculos lógicos entre sí

En ocasiones, un gran problema de delincuencia genera una gran respuesta. Por ejemplo, el Banco Interamericano de Desarrollo (BID) muchas veces recibe pedidos de financiamiento para paquetes de programas de prevención del delito que apuntan a mejorar la seguridad ciudadana. Sin embargo, por lógica, es más probable que esos paquetes funcionen si los componen programas que se refuerzan y complementan entre sí.

EJEMPLO: Supongamos que un país tiene una tasa de homicidios muy alta, y numerosos secuestros tanto de ciudadanos locales como de ejecutivos extranjeros que están de visita por negocios. Supongamos que el país puede gastar o pedir prestados US\$300 millones. ¿De qué forma *no* debería planificarse un paquete de programas en los cuales invertir los fondos?

- Decidir en primer lugar cómo dividir el dinero entre organismos de poder.
- Decidir en primer lugar cómo dividir el dinero para beneficio supuesto de los niños, las mujeres o lo que Raúl Zaffaroni, miembro de la Corte Suprema de Justicia argentina, denomina “víctimas-héroes”, asociados más a la política simbólica que a la seguridad.
- Elegir entre distintas propuestas los programas que, por intuición o juicio instantáneo del sistema I, “suenan” más viables.

Si se utiliza cualquiera de esos métodos para crear un paquete de programas, es probable que el resultado disperse los recursos y desperdicie dinero de dos formas a la vez. Una de las formas de desperdiciar dinero consiste en invertirlo en programas que no funcionarán, uno por uno. La otra es perder la oportunidad de atender cuestiones institucionales profundas de las estructuras de prevención del delito a nivel nacional.

Si los problemas principales son los homicidios y los secuestros, el proceso DRIVER debe concentrarse en esos dos problemas. Si los fracasos de la policía ocupan un lugar preponderante en ambos tipos de delito, el análisis puede examinar las posibles causas de la corrupción y las fallas policiales: salarios insuficientes, pocos requisitos en cuanto a educación al momento de la designación de cargos, una diferencia muy marcada entre los funcionarios de alto nivel y la mayoría de los policías uniformados, falta de responsabilidad externa, o grandes alianzas o amenazas en relación con los principales cárteles de narcotráfico.

Si la policía es una preocupación central para la seguridad ciudadana, tiene más sentido pensar en un programa amplio de reformas de esa institución que en un conjunto de programas desconectados que ofrecen pequeños beneficios a cada sector de la población. Si, en cambio, el problema principal lo constituyen los cárteles de narcotráfico, el objetivo del proceso DRIVER podría ser un paquete de programas para reducir el daño que provocan. Más allá de que los problemas sean grandes o pequeños, concretos o amplios, el proceso DRIVER siempre debe comenzar con un análisis que determine de dónde proviene el daño mayor. Como un misil termoguiado, el proceso DRIVER debe ir en busca de los daños para hallar el peligro mayor y convertirlo en su primera prioridad.

Programas sin modelos lógicos

Algunos líderes se oponen a las evaluaciones de impacto porque no quieren someter el modelo lógico de sus programas a un escrutinio pormenorizado. Y es por eso mismo que el requisito de una evaluación de impacto es tan importante para cualquier innovación. Un modelo lógico, o teoría sobre la forma en que debería funcionar el programa, es esencial para explicar la relación entre una innovación y su objetivo. Sin embargo, muchas innovaciones de prevención del delito no identifican sus objetivos, y mucho menos un modelo lógico para cumplirlos.

EJEMPLO: Se creará una nueva academia de policía. En el pedido de financiamiento, no se indica quiénes serán los alumnos, qué se enseñará ni cuál es el propósito de la academia. Se proporciona un único indicador de rendimiento: la cantidad de casos de delincuencia

resueltos por policías graduados de la academia. Sin embargo, no se menciona de qué manera se investigan los delitos ni si se asignan distintos tipos de casos a los graduados de la academia nueva en comparación con los de la academia anterior. Ni siquiera está claro si los investigadores resuelven los delitos en forma individual, en lugar de trabajar en equipos que tal vez incluyan tanto a graduados de la academia nueva como de la anterior.

Una nueva academia de policía puede ser una muy buena idea como parte de un proceso DRIVER que apunte a crear una institución policial más eficiente, justa y menos corrupta. Ni siquiera es necesario que haya un objetivo específico de reducción del delito si la intención del proceso es mejorar el control policial en sí mismo. Para mejorar la seguridad ciudadana, no es necesario que haya una conexión directa entre cada innovación y un tipo de delito en particular. En cambio, lo que se necesita es una gran claridad en cuanto a las causas y los efectos. La teoría de que una nueva academia de policía mejorará la integridad y competencia de la policía es un modelo claramente lógico. El modelo puede tener numerosos objetivos o indicadores, pero la idea principal es perfectamente clara, sin importar la complejidad de los detalles. Esa transparencia se contrapone al objetivo vago y opaco de simplemente crear una nueva academia sin modelo lógico alguno.

Programas sin datos locales

Es un gran error concebir una innovación sin antes diagnosticar el problema a nivel local. Lamentablemente, muchos programas de seguridad se conciben sin acceso a datos criminológicos locales. Puede que sus diseñadores sólo tengan la posibilidad de consultar anécdotas y notas en los periódicos. En estas situaciones, ese no es un obstáculo para resolver un problema importante, sino que *es* el problema más importante.

EJEMPLO: Una escuela propone desarrollar un programa extracurricular para los niños de una zona con alto nivel de criminalidad, con el objetivo de disminuir el riesgo de que participen en delitos. La escuela no tiene ningún dato sobre la cantidad de estudiantes, por grupo etario, que ya fueron acusados de algún delito. No tiene la capacidad de utilizar esos datos para orientar sus esfuerzos o al menos asegurarse de que los niños delincuentes no estén concentrados en determinados grupos. Y lo que es peor, tampoco tiene la capacidad

de medir la delincuencia entre los niños que completaron el programa ni entre quienes aún están asistiendo. Por lo tanto, no hay forma de conectar esta idea aparentemente buena con la prevención del delito.

Podría argumentarse que el propósito de la escuela no es mejorar la accesibilidad ni la precisión de los datos sobre criminalidad. Asimismo, también podría sostenerse que las escuelas no tienen nada que ver con la delincuencia. Sin embargo, la necesidad de contar con datos precisos y utilizables sobre el delito es tan grande que a toda institución que se ocupe del delito puede requerírsele que comience por la obtención de información. Si se trabaja sobre una base sólida de datos criminológicos precisos, incluso manteniendo prudente respeto por los derechos de privacidad, puede construirse una oportunidad ilimitada para la prevención del delito.

Programas sin evidencia anterior sobre los problemas

Un tema relacionado con la falta de datos a nivel local es la falta de datos mundiales. La delincuencia está cambiando rápidamente en todo el mundo como resultado de la tecnología y la comunicación. Si el problema criminológico de interés no se ha producido nunca en una localidad, tal vez pueda obtenerse información de otras zonas en las que sí se produjo. Si el problema no ha ocurrido nunca en ningún lugar, el desafío será mayor. En ese caso, la mejor respuesta sería seguir buscando evidencia de patrones similares.

EJEMPLO: Las “movilizaciones relámpago” (también conocidas por su nombre en inglés: “flash mob riots”) para generar disturbios llegaron al Reino Unido en 2011. Nunca antes se habían visto de esa manera: cientos o miles de personas a las que se les indicaba que se reunieran en un determinado lugar para realizar saqueos o incendios. Las instrucciones llegaban por mensajes de texto, Twitter y otros medios de comunicación instantáneos a través de dispositivos de comunicación digital personales. La policía no estaba preparada para lidiar con estas multitudes que se trasladaban por Londres más rápido de lo que se demoraba en hacer llegar suficientes efectivos al lugar. El asunto se debatió en el Parlamento, donde se consideró la posibilidad de usar camiones cisterna y balas de goma. Sin embargo, se presentó muy poca evidencia sobre quiénes participaban en las

movilizaciones o cuál podía ser la dinámica de estos problemas de orden público en rasgos más generales.

En los meses siguientes a los disturbios de agosto, se reunieron más evidencia sobre las personas arrestadas y sus antecedentes policiales. Sin embargo, se recopilaron poca evidencia sobre el comportamiento de las “movilizaciones relámpago”, el que aparentemente se había utilizado durante años para organizar fiestas más que disturbios. Pensar planes de prevención para este tipo de comportamientos es una tarea compleja, pero sólo puede facilitarse con investigaciones más globales sobre la forma en que dicho comportamiento se desarrolla.

Programas sin evidencia anterior sobre las soluciones

En términos generales, es desaconsejable implementar un programa sin evidencia alguna sobre la probable eficacia de la solución que aportaría para un problema determinado. Tal vez la solución no se haya comprobado en un experimento aleatorio controlado, pero de todas maneras puede que haya cierta evidencia al respecto en las publicaciones sobre evaluaciones a nivel mundial. Ignorar esa evidencia puede llevar a la adopción de programas ineficaces o incluso perjudiciales.

EJEMPLO: Personal de respuesta secundaria para el seguimiento de episodios de violencia familiar. Esta solución complementa la respuesta de la policía a incidentes de violencia familiar en el momento en que se están produciendo. Cuando la policía acude a un hogar durante una emergencia, en general no cuenta con habilidades especiales para lidiar con el conflicto. En Estados Unidos se implementaron varios programas a partir de los cuales, entre una y dos semanas después de la llamada de emergencia, se enviaba a un orientador o trabajador social junto con la policía al hogar donde se había producido el incidente. Estos programas se evaluaron en varios experimentos aleatorios sólidos, como uno que se llevó a cabo en Miami con una población de origen étnico mayoritariamente latino. Los resultados muestran que esas visitas únicamente hicieron que aumentaran las llamadas a la policía, sin evidencia alguna de que se hubiese reducido la cantidad de lesiones ni de que se hubiese obtenido cualquier otro beneficio. Cualquier país o ciudad que

adoptara un programa de este estilo estaría ignorando evidencia concreta de que, como mínimo, el dinero invertido se estaría desperdiciando. En el peor de los casos, el programa podría incluso hacer que haya más violencia en el hogar.

Programas con una “caja negra”: sin medición de los pasos clave de la implementación

Tal vez la peor manera de concebir un programa es no tener una forma directa de medir si mínimamente se lo está llevando a cabo. Incluso con una buena evaluación de impacto, es difícil saber cómo interpretar los datos sobre los *resultados* si se desconoce la información sobre los *productos*. Esto hace que quienes financiaron el proyecto no puedan saber con certeza si el programa no funcionó o si ni siquiera se llevó a cabo. ¿El remedio no logró curar la enfermedad o la enfermera nunca le dio el remedio al paciente? (Otra posibilidad: la enfermera le dio el remedio al paciente, pero este nunca lo tomó).

EJEMPLO: Se implementa un programa de terapia cognitivo-conductual (TCC) en una cárcel. En una evaluación de impacto, se estudian las tasas de reincidencia de los delincuentes que salen de la cárcel y se observa que estas descienden en el año siguiente a la implementación del programa de TCC, en comparación con el año anterior. Pero no se proporciona ningún dato sobre la cantidad de detenidos que realmente comenzó la TCC ni mucho menos sobre cuántos completaron las 12 horas de sesión recomendadas en la revisión de evidencia de evaluaciones anteriores. No se sabe si el porcentaje de detenidos que recibieron la TCC fue del 100% o del 5%. Ningún proceso DRIVER puede encontrar sentido alguno en un programa así de “exitoso”.

Programas “sin resultados”: sin medición del impacto en los problemas

Tal vez el error más frecuente en el diseño de un programa sea no analizar ninguna información sobre los resultados obtenidos de modo que se pudiera establecer una conexión específica entre dichos resultados y el programa en cuestión. Aunque haya datos precisos sobre los productos, que indiquen cuántas personas recibieron qué cantidad de sesiones del programa y en cuánto tiempo, es posible que no haya información sobre la cantidad de

delitos que esas personas cometieron más adelante o la cantidad de delitos de los que fueron víctimas. Entonces, el proceso DRIVER no podrá llevar a ninguna conclusión acerca de si se debe continuar, detener o perfeccionar el programa.

EJEMPLO: Una entidad de servicios sociales contrata a “asistentes sociales desvinculados” de entre 25 y 35 años para que trabajen en la calle entablando lazos de amistad con jóvenes que participan en pandillas. Cada uno de los trabajadores pasa una gran cantidad de tiempo con un grupo de jóvenes de entre 10 y 30 años que viven en una zona determinada. Se registra detalladamente con quiénes se encuentran, qué actividades de fomento social organizan, qué lugares visitan e incluso qué negocios (legales) emprenden. Sin embargo, la entidad no recibe ningún dato sobre las actividades delictivas de los jóvenes. Los asistentes sociales tampoco llevan un registro de los arrestos de los que puedan enterarse directamente por boca de los jóvenes. Sin esa información, el proceso DRIVER no puede completarse.

Programas “sin comparación”: sin medición del valor agregado

El error más frustrante que puede cometerse en un programa es el de hacer todo bien salvo una cosa. En virtud del proceso DRIVER, podría seleccionarse un programa elaborado a partir de la evidencia que tenga una buena teoría. Pero si no se realiza una comparación con un grupo similar en el que el programa no se aplique (o en el que se aplique otro programa), las conclusiones a las que se llegue podrían ser totalmente erróneas.

EJEMPLO: Un programa para delincuentes juveniles podría medir los productos que intenta lograr; por ejemplo, los encuentros entre orientadores y niños. Podría medirse un resultado antes y después del programa, como la frecuencia con que se arresta a los niños. Podría hacerse cualquier cosa salvo calcular el valor agregado: la diferencia real que el programa generó en cualquier tendencia criminológica. Incluso puede que parezca que el programa falló cuando en realidad tuvo éxito. Si las tasas de arresto de los niños participantes en el programa subieron, se podría suponer que iban a subir de cualquier modo, y que el programa no pudo evitar ese aumento. Sin embargo, con un grupo de

comparación adecuado, la evidencia podrían indicar que los arrestos aumentaron mucho menos que entre un grupo de niños similares que no participaron en el programa.

Encontrar un grupo de comparación puede resultar difícil, pero aun si no hay uno disponible, pueden usarse valores de referencia para proyectar una comparación virtual. Incluso es posible utilizar a los niños inscriptos en el programa como su propio grupo de control, comparando el desempeño real con el desempeño previsto (Sherman, 1998). Pero más allá de la manera en que se lo calcule, un método justo de comparar la población intervenida con un caso “contrafáctico” sin el programa es esencial en el proceso DRIVER.

Desarrollo de programas a partir de la evidencia

¿Qué es la evidencia?

La evidencia consiste en observaciones recopiladas de manera sistemática para respaldar una conclusión. El uso de ese tipo de observaciones en el desarrollo de programas de seguridad ciudadana ha sido bastante escaso en comparación con el uso de mejores prácticas para la planificación en el área de los negocios o la salud. En esas y otras áreas se requiere una gran cantidad de evidencia correspondiente al sistema II de pensamiento para sostener grandes inversiones. Las políticas de seguridad ciudadana siempre se orientaron mucho más al sentido común, teorías formuladas con el sistema I, y las costumbres y prácticas tradicionales. La diferencia entre el desarrollo de programas de seguridad ciudadana a partir de la evidencia y lo que se hacía anteriormente es, en gran medida, un aumento exponencial de la necesidad de respaldar las conclusiones con evidencia fehaciente; es decir, evidencia que pueda resistir otras interpretaciones y teorías que podrían debilitar la validez de la conclusión.

Las dos conclusiones más importantes a las que se debe llegar en los programas de seguridad ciudadana requieren evidencia en materia de *pronósticos* y de *prevención*. Ambas cuestiones constituyen terrenos muy propicios para el enfrentamiento entre el sistema I y el sistema II, y son también esenciales para el proceso DRIVER. Y ambas pueden comenzar con iniciativas sistemáticas de describir patrones y tendencias de criminalidad, y poblaciones de delincuentes.

Patrones y tendencias

Un patrón es cualquier distribución repetida de episodios, como el hecho de que el sol salga todos los días y se ponga todas las noches. Los patrones de criminalidad pueden tener numerosas variables, desde el modus operandi (asesinatos con cuchillos en peleas en tabernas) hasta el móvil (asesinatos por cuestiones de honor de mujeres musulmanas que infringen la ley de la Sharia) o los datos demográficos (robos a bancos cometidos por hombres ancianos en sillas de ruedas). Una tendencia es un tipo de patrón que consiste en un cambio, aumento o disminución, en la frecuencia de los episodios o las tasas a lo largo del tiempo. En virtud de los patrones y las tendencias de criminalidad, es posible elegir los problemas que van a tratarse o desglosar problemas grandes en objetivos clave. El hecho de que la mitad de todos los asesinatos de un año se haya cometido en tres barrios en los que no había habido asesinatos el año anterior, por ejemplo, es tanto un patrón como una tendencia. Puede impulsar diversos análisis y generar ideas sobre formas de evitar nuevos asesinatos. También puede permitir la realización de pronósticos.

Pronósticos

Los pronósticos son predicciones acerca de si algo va a suceder o no, como si va a llover o no (un resultado binario), o del grado en que algo va a ocurrir, como la cantidad de lluvia que caerá (un resultado continuo). Los ejemplos podrían perfectamente implicar si va a haber un asesinato este fin de semana en esta ciudad o cuántos asesinatos habrá. También podrían considerar si un determinado delincuente va a cometer otro delito cuando salga de la cárcel o bien cuántos delitos cometerá.

Los pronósticos de este estilo, formulados con el sistema I, se basan en la experiencia y en “pálpitos”, o incluso en estudios de caso detallados. Los pronósticos formulados con el sistema II se realizan a partir de análisis estadísticos de enormes bases de datos que son esenciales para lograr pronósticos precisos de acontecimientos poco frecuentes (Meehl, 1954). En cinco décadas de lucha entre el enfoque “clínico” del sistema I y el enfoque “estadístico” del sistema II, los pronósticos estadísticos han

derrotado constantemente a los clínicos (Kahneman, 2011). Sin embargo, son muchas las entidades de justicia penal que confían en pronósticos clínicos, mientras que sólo algunas han adoptado métodos estadísticos.

Los pronósticos estadísticos no sólo son más precisos, sino que además pueden funcionar en varios niveles. Pronosticar en qué barrio es posible que se cometa un asesinato tal vez lleve a evaluar qué delincuentes en particular presentan mayores probabilidades de cometer el próximo homicidio. Los análisis estadísticos tanto de datos sobre personas como de datos sobre barrios permiten realizar pronósticos en varios niveles con claras repercusiones en las estrategias de prevención.

Prevención: la escala de métodos científicos de Maryland

Una vez que se pronostica que en una unidad hay un alto riesgo de que se produzca un problema de delincuencia, puede intentarse intervenir con programas de prevención a fin de reducir ese riesgo. La eficacia de la intervención es una cuestión totalmente independiente de la exactitud del pronóstico. Sin embargo, la única forma de saber si la intervención funcionó es comparar las unidades en las que se la implementó con las unidades en las que no (este es el séptimo de los 10 principios enumerados en el resumen).

Cuando se realiza una comparación, esta debe ser lo más justa (imparcial) posible. Lo ideal es que, en promedio, la única diferencia entre las unidades en que se aplica la intervención y las demás sea justamente esa: si hubo intervención. La aproximación más cercana a ese ideal en general se obtiene cuando las comparaciones se realizan mediante experimentos con asignación al azar, conocidos también como ensayos controlados aleatorios (ECA). Las iniciativas de prevención evaluadas de esa forma se ubican en la parte superior de la escala de métodos científicos de Maryland (Sherman, 1997), una escala de cinco niveles con la que se califica el grado de “validez interna” de los diseños utilizados con mayor frecuencia para evaluar iniciativas de prevención del delito. Los cinco niveles de la escala son los siguientes:

Nivel 1. Correlación entre un programa de prevención del delito y una medición de la actividad delictiva o los factores de riesgo.

Nivel 2. Secuencia temporal entre el programa y un resultado claramente observado en cuanto a la delincuencia o los riesgos, o presencia de un grupo de comparación sin demostrar que dicho grupo es comparable con el de tratamiento.

Nivel 3. Comparación entre dos o más unidades de análisis, una con el programa y la otra sin él.

Nivel 4. Comparación entre varias unidades con y sin el programa, controlando la presencia de otros factores, o con un grupo de comparación no equivalente en el que sólo se observan diferencias menores.

Nivel 5. Asignación al azar y análisis de unidades comparables en el grupo del programa y el de comparación.

En general, la mejor evidencia disponible es la que se obtiene con un diseño de investigación que tenga la calificación más alta posible en la escala de 1 a 5. El motivo es que cuanto mayor sea la calificación en la escala, más claridad habrá acerca de si el programa de prevención ocasionó alguna diferencia entre las unidades a las que se apuntó y las de comparación, o si la apariencia de causalidad en realidad era engañosa porque la verdadera causa de la diferencia fue otro factor no controlado.

Al evaluar las comprobaciones anteriores de cualquier programa, hay una mayor certeza de que el resultado sea correcto en los diseños de mayor nivel. Sin embargo, la certeza no es lo único que debe tenerse en cuenta. Podría decirse que es más importante elegir un programa que tenga un gran efecto, o un alto grado de prevención por monto de dinero invertido, que elegir un programa que claramente esté generando un efecto débil (baja rentabilidad de la inversión). Tal como sostiene la organización The Campbell Collaboration, “la magnitud del efecto es clave”. Es posible que programas que han tenido un gran efecto en comprobaciones anteriores sean la mejor opción aunque la certeza sobre esos efectos sea menor. Incluso puede que sea así, por ejemplo, aunque las comprobaciones anteriores no hayan sido “estadísticamente significativas”, siempre y cuando los resultados hayan ido en la dirección correcta y hayan tenido gran magnitud.

En ese contexto, la expresión “no ser estadísticamente significativo” debe interpretarse sólo como “no es claro” más que como “es débil”. Hay muchos efectos

importantes cuya claridad puede ser dudosa; en ese caso, la magnitud del efecto debe ser el indicador principal en la decisión de seleccionar o no el programa para usarlo en el ámbito local. Todo esto está sujeto a algunas otras consideraciones, que se detallan a continuación.

Comprobaciones anteriores de programas similares

La elección de un programa de prevención es, en efecto, un pronóstico de que el programa funcionará. El mejor fundamento para dicho pronóstico lo constituyen las comprobaciones anteriores del mismo programa o de otros similares. Una vez que, como parte del proceso DRIVER, se ha analizado la información local, es fundamental buscar en Internet evidencia de comprobaciones anteriores de cualquier programa que pueda llegar a implementarse en una jurisdicción.

La mejor evidencia sería una revisión sistemática de experimentos de asignación aleatoria reiterados en los que se haya comprobado el mismo programa o la misma práctica con los mismos grupos de comparación. Los experimentos reiterados brindan una mayor *validez externa* (o capacidad de generalización), que es la confianza en que lo que funcionó en un lugar funcionará también en otro.

Si no hay ninguna revisión disponible de este tipo, un único experimento con asignación aleatoria puede aportar evidencia más fehaciente que varios cuasiexperimentos. Un buen experimento puede ser más convincente que varios cuasiexperimentos menos concluyentes porque tiene una mayor *validez interna*, ya que en ese único estudio se habrán descartado más explicaciones alternativas que amenazaran su conclusión.

Si no es posible hallar ningún ensayo aleatorio, la mejor fuente de evidencia que sigue serán los cuasiexperimentos reiterados. Este tipo de estudios son más alentadores cuando todos o la mayoría de sus resultados apuntan en la misma dirección, incluso si los efectos no son estadísticamente significativos. Las posibilidades de que varios resultados que apuntan en la misma dirección (por ejemplo, a una reducción del delito) sucedan por casualidad son bastante bajas. Esto significa que un *patrón* de resultados puede ser significativo desde el punto de vista estadístico aunque los hallazgos individuales de un estudio no lo sean.

Comprobaciones anteriores de soluciones para problemas similares

Si no es posible hallar comprobaciones de soluciones para un determinado problema, el desarrollo de programas puede hacerse a partir de la evidencia más similar que exista. Una comprobación de una solución para un problema similar con una población distinta, o en una etapa diferente del proceso de justicia penal, puede generalizarse razonablemente para aplicarse al problema de interés.

EJEMPLO: Hasta el momento no se ha realizado ningún experimento con asignación aleatoria que consista en amenazar a los miembros de pandillas con arrestarlos por infracciones menores, una estrategia que en inglés se conoce como “pulling levers” (literalmente, “accionar palancas”, en referencia a que se mueven todos los recursos disponibles para asegurar la disuasión de los delincuentes). Sin embargo, sí se ha hecho un experimento de amenazas específicas a personas en libertad condicional a las que se les realizan análisis periódicos para detectar si consumen drogas mientras se encuentran en libertad condicional. Ese experimento (llamado “Proyecto HOPE”) evaluó el uso de la detención inmediata por una noche la primera vez que la persona daba positivo en un análisis toxicológico, dos noches la segunda vez, tres la tercera, y así sucesivamente. El efecto general de la asignación aleatoria a ese tratamiento fue una gran disminución en la cantidad de arrestos, en comparación con personas en circunstancias similares de libertad condicional con quienes no se usó esa práctica (Hawken y Kleiman, 2009). Este experimento es una evidencia directa del problema de delincuencia en las personas que están en libertad condicional. Sin embargo, también puede ser una evidencia indirecta de los efectos de las amenazas de arresto inmediato de los miembros de una pandilla en caso de que cualquiera de ellos cometa un asesinato (Weisburd, Sherman y Maher, 1993).

El proceso iterativo de análisis y diseño

El desarrollo de programas a partir de la evidencia es *iterativo*: un ir y venir entre evidencia y conclusiones, nuevos datos y nuevas acciones. El modelo DRIVER depende enteramente de la evidencia que tenemos hoy, y sus conclusiones pueden cambiar si hay evidencia

nueva en el futuro. Desde el punto de vista científico, una evidencia siempre está incompleta, pero siempre puede ampliarse. Ese concepto es diametralmente opuesto a un punto de vista legal o ideológico que esté arraigado y no acepte modificaciones por cuestión de principios. El único aspecto en el que el desarrollo de programas a partir de la evidencia puede admitir esa intransigencia es en lo referido a derechos humanos y valores éticos inconfundibles. Todas las demás conclusiones sobre prácticas de prevención del delito deben estar abiertas al cambio como resultado de nueva evidencia. De lo contrario, todavía creeríamos que la tierra es plana.

Conexión entre la seguridad y el desarrollo económico

El proceso DRIVER iterativo para mejorar la seguridad ciudadana es igual al proceso de desarrollo económico a partir de la evidencia. Ambos dependen de teorías sobre lo que funciona, que se comprueban con un mayor o menor grado de imperfección a través de la evidencia. Ambos utilizan descripciones sistemáticas de los problemas; por ejemplo, quién está desempleado (o es víctima de un delito), dónde y por qué. Y ambos pueden resolverse mediante programas similares.

El conocimiento científico de los efectos del desarrollo económico sobre el delito todavía es escaso. Hay varios hallazgos sorprendentes que indican que el delito crece al mismo tiempo que la economía, y que los programas de empleo no logran prevenirlo. Sin embargo, también hay evidencia sólida de que el delito disminuye cuando desciende la desigualdad en el nivel de riqueza e ingresos. La tasa de homicidios, en particular, está relacionada con el coeficiente de Gini de desigualdad de los ingresos. Lo mismo sucede con la tasa de asesinatos de ciudadanos por parte de la policía (Sherman y Langworthy, 1979). Puede que el aumento del empleo al mismo tiempo que el de la desigualdad sea de poca ayuda en la lucha contra el crimen, pero que un fortalecimiento del empleo a largo plazo logre fomentar una reducción del nivel de desigualdad y, por consiguiente, de los delitos violentos.

Tal vez una razón más importante para establecer una conexión entre la seguridad y el desarrollo económico sea el daño que la delincuencia genera en la economía. En Estados Unidos la evidencia indica que el costo de la delincuencia constituye el 17% del producto

interno bruto (PIB), lo que incluye los costos directos del delito así como también los fondos invertidos en justicia y seguridad: un recordatorio llamativo sobre el papel de la delincuencia como impuesto oculto sobre el crecimiento económico (Ludwig, 2006). Una visión del desarrollo económico a largo plazo tal vez requiera que cada país garantice la seguridad para la realización de negocios, en especial si se trata de una economía en vías de globalización.

Protocolo DRIVER: lista de comprobación

Esta lista de comprobación es una guía paso a paso para el desarrollo de programas de seguridad ciudadana a partir de la evidencia. Está organizada en una serie de preguntas que deben responderse muy brevemente, pero con precisión. Puede usarse con diversos propósitos, como registrar las decisiones tomadas en una serie de reuniones o los puntos clave de acuerdos sobre cómo delinear un programa propuesto. También puede fundamentar pedidos de financiamiento para un programa o permitir que las entidades financieras evalúen la solidez de una propuesta en la cual invertir. Quizá quienes puedan sacar mayor provecho de esta lista sean las organizaciones que se ocupan tanto de desarrollar como de ejecutar programas.

Existe evidencia sólida de que las listas de comprobación pueden mejorar el rendimiento en ámbitos diversos, como la prevención de accidentes aéreos, errores quirúrgicos y accidentes nucleares (Gawande, 2009). Usar esta lista de comprobación puede hacer que se encuentren más y mejor evidencia que permita desarrollar prácticas y programas de prevención del delito eficaces en función de los costos. Tal como señala Malcolm Gladwell en su reseña del libro *The Checklist Manifesto*, de Atul Gawande:

Gawande comienza por señalar la diferencia entre los errores por ignorancia (errores que cometemos porque no sabemos lo suficiente) y los errores por ineptitud (errores que cometemos por no aplicar correctamente lo que sabemos). En el mundo moderno, afirma, el fracaso se produce por el segundo de esos tipos de errores.

Al igual que el proceso DRIVER, la lista de comprobación sigue un ciclo clásico de investigación-acción para la solución de problemas, que fue propuesto por primera vez por el psicólogo Kurt Lewin (1958). Al finalizar cada ciclo, puede ser útil comenzar una nueva lista de comprobación. Para problemas de suma urgencia, tal vez resulte provechoso actualizar el proceso DRIVER todos los años ya que, de esa manera, se ayuda a garantizar que el proceso se mantenga al día con la evidencia a nivel tanto local como mundial.

Paso 1. Diagnóstico del problema específico que debe resolverse: concentraciones y causas

1.1. ¿Cuál es el problema específico que debe resolverse? (por ejemplo: homicidios e intentos de homicidio entre pandillas).

1.2. ¿Qué categorías del problema tienen mayor relevancia para cualquier intervención? (por ejemplo: determinadas pandillas).

1.3. ¿A qué categorías en particular debe dárseles más prioridad y por qué? (por ejemplo: las tres pandillas que cometen más asesinatos).

1.4. ¿Qué datos cuantificables existen para el análisis? (por ejemplo: archivos policiales sobre homicidios y agresiones).

1.5. Si no hay ningún dato cuantificable directamente relacionado con el problema, ¿qué datos se acercan más a la medición del problema definido? (por ejemplo: registros judiciales de condenas por homicidio e intento de homicidio).

1.6. Análisis de las minorías fuertes: ¿qué porcentaje del problema está concentrado en qué porcentaje de las unidades que se mencionan a continuación?

- El __ % de los delincuentes identificados cometió el __% de los delitos.
- El __% de las víctimas identificadas sufrió el __% de los delitos.
- El __% de los delitos se produjo en el siguiente horario: de __:00 a __:00.
- El __% de los delitos se produjo en los siguientes días de la semana: ____, ____ y ____.
- El __% de los delitos se produjo en el __ % de (calles, microzonas).

1.7. Datos demográficos de los delincuentes identificados o condenados.

De sexo masculino: __%

Menores de 18 años: __%

Menores de 25 años: __%

Mayores de 40 años: __%

Con condenas anteriores: __%

Con arrestos anteriores: __%

Exconvictos: __%

Del grupo étnico más importante: __%

Que residen en zonas con alto nivel de pobreza: __%

De clase económica alta: __%

1.8. Datos demográficos de las víctimas de _____.

De sexo masculino: __%

Menores de 18 años: __%

Menores de 25 años: __%

Mayores de 40 años: __%

Con condenas anteriores: __%

Con arrestos anteriores: __%

Exconvictos: __%

Del grupo étnico más importante: __%

Que residen en zonas con alto nivel de pobreza: __%

De clase económica alta: __%

1.9. Móviles (describa hasta tres de los móviles más frecuentes, por ejemplo: rivalidad entre pandillas).

1. _____

2. _____

3. _____

1.10. Método o tecnología (medio) (por ejemplo: armas, fraude, corrupción de la policía).

1.11. ¿Qué patrones de respuesta se proporcionan actualmente mediante alguno de los siguientes recursos?

- Vigilancia policial
- Procesamiento
- Fianza
- Condenas
- Sentencias
- Programas penitenciarios

1.12. ¿Qué programas de prevención se están llevando a cabo actualmente, si es que se lleva a cabo alguno?

- Gobierno
- Familia
- Escuelas
- Participación en el mercado laboral legal
- Sistemas de seguridad en lugares (por ejemplo, hogares de las víctimas)
- Sistemas de seguridad en situaciones (por ejemplo, al entrar a las tabernas)

1.13. ¿Qué señalan las publicaciones mundiales sobre las macrocausas del problema?

(por ejemplo: desigualdad de ingresos, debilidad de las instituciones estatales, densidad de armas, zonas con alta concentración de pobreza).

1.14. ¿Qué señalan las publicaciones mundiales sobre las microcausas del problema?

(por ejemplo: competencia en el narcotráfico, extorsión a empresas, bebidas alcohólicas a precios bajos, redes familiares de delincuentes).

1.15. ¿Qué tendencias se midieron sobre el problema y qué indican?

- Tendencia nacional en un período de entre cinco y 10 años (por ejemplo: un 37% de aumento de la delincuencia)
- Tendencia local en un período de cinco años (por ejemplo: un 44% de aumento de la delincuencia)

1.16. ¿Cuál es el motivo más probable de las tendencias recientes?

1.16.a. ¿Qué evidencia hay de esa teoría, si es que hay alguna?

1.17. ¿En qué mapas se muestran las concentraciones geográficas del problema? (resuma y adjunte).

1.18. Describa la cadena de causalidad: mencione todos los eslabones de la cadena de causalidad que contribuyen, simultáneamente o en algún tipo de secuencia, a la incidencia reiterada de estos delitos (por ejemplo: amplia disponibilidad de armas de fuego). (Agregue todos los renglones que sean necesarios).

1. _____
2. _____
3. _____
4. _____
5. _____

1.19. Rotura de la cadena: ¿cuál de estos eslabones de la cadena de causalidad haría más probable que se redujera el problema si su conexión se rompiera?

1.20. Eslabón más frágil: ¿qué eslabón de la cadena de causalidad podría romperse con más facilidad (aunque no sea el que tiene mayor impacto)?

1.20.a. ¿Por qué sería más fácil romper ese eslabón de la cadena de causalidad que otros?

1.21. ¿Cuál es el objetivo al que apunta la iniciativa que se propone usted desarrollar? (por ejemplo: reducir la portación de armas en público).

1.22. Mencione toda organización judicial penal, gubernamental o privada a la que haya presentado este diagnóstico y resuma la respuesta obtenida.

Paso 2. Revisión de programas con los que sería posible resolver cada problema

2.1. Búsqueda en Internet de fuentes en inglés.

2.1.a. ¿Qué palabras clave utilizó?

2.1.b. Marque todas las bases de datos en las que haya buscado información.

www.crimesolutions.gov _____

www.campbellcollaboration.org _____

www.homeoffice.gov.uk _____

www.ncjrs.gov _____

www.google.com _____

Google Scholar en Google.com _____

2.2. Búsqueda de fuentes en español

2.2.a. ¿Qué palabras clave utilizó?

2.2.b. Marque todas las bases de datos en las que haya buscado información.

www.google.com _____

(Confeccione una lista de las demás fuentes que haya consultado).

2.3. Selección de descripciones de programas pertinentes: describa uno o más programas que haya encontrado durante las búsquedas en Internet y que parezcan más prometedores para utilizarlos en su ciudad.

2.4. Selección de evaluaciones de impacto pertinentes: ¿cuáles de los programas prometedores que halló en Internet cuentan con evaluaciones de impacto que hayan alcanzado al menos el nivel 3 de calificación en la escala de métodos científicos de Maryland? (véase más arriba la sección “Prevención: la escala de métodos científicos de Maryland”).

2.5. Evaluaciones sistemáticas: ¿alguno de los programas que parecían prometedores fue sometido a reiteradas evaluaciones de impacto resumidas de forma sistemática en una reseña o un meta-análisis? Si es así, explique brevemente los resultados.

2.6. Adaptación de programas comprobados en otros lugares a las condiciones locales: ¿qué programas encontró en otros lugares que (en teoría) parecen apropiados para usarlos a nivel local, si es que encontró alguno?

2.7. ¿Qué programa encontró que pueda modificarse para utilizarlo en el ámbito local, y qué modificaciones le haría?

2.8. Creación de un programa nuevo. Si ninguno de los programas evaluados o descritos en otros lugares parece adecuado, describa una iniciativa nueva que desee implementar que, al parecer, no se haya intentado en ningún otro lugar hasta el momento.

2.9. Desarrollo de un modelo lógico. Describa, en teoría, la cadena de causalidad entre el programa que propone implementar y el resultado cuantificable que espera lograr.

(Por ejemplo: puestos de vigilancia policial para disminuir la portación de armas y los homicidios con armas).

1. Identificación de microzonas con altas tasas de crímenes con armas → 2.

Inspección policial al azar en busca de armas en vehículos o a personas que entran en la zona → 3. Menor portación de armas en esas zonas → 4. Menor cantidad de disparos en esas zonas → 5. Menos muertes por heridas de bala en esas zonas.

1. _____ → 2. _____ → 3. _____ → 4. _____ → 5. _____

2.10. Comprobación de la idea nueva mediante experimentos mentales. Describa los resultados de cualquier “experimento mental” que haya realizado con cualquier grupo focalizado (*focus group*) en el que se haya conversado sobre el modelo lógico explicado anteriormente.

2.11. Decisión provisoria. Describa los elementos principales del programa que se implementará, si se determina que es eficaz en función de los costos para reducir el problema de criminalidad.

Paso 3. Implementación del programa propuesto

3.1. Del modelo lógico al plan de acción: resuma, en orden cronológico, los pasos que deben seguirse para implementar la iniciativa propuesta utilizando el cuadro que se proporciona a continuación (este contiene un ejemplo que debe borrarse a fin de dejar lugar para su propuesta). Agregue tantas columnas y filas como sean necesarias.

Tarea	Parte responsable	Comienzo	1er. trimestre	2do. trimestre	3er. trimestre	4to. trimestre
Trazar mapa de zonas de delitos con armas	ONG “Fin de la violencia con armas”	X	X	X	X	X
Seleccionar policías	Enlace de la policía	X				
Capacitar a los policías	ONG y academia de policía	X				
Asignar los policías a los puestos de vigilancia	Grupo de análisis criminológico de la policía y ONG	X				
Comprar chalecos antibalas para los policías	ONG	X				
Hacer cacheos por la noche			X	X	X	X
Hacer cacheos de madrugada			X	X	X	X
Hacer cacheos durante el día			X	X	X	X

Evaluar el progreso y los inconvenientes en revisiones mensuales de la implementación			X	X	X	X
---	--	--	---	---	---	---

3.2. Del plan de acción a la medición de las acciones planificadas: enumere todas las mediciones de acciones (productos) que se evaluarán todos los meses (o con más frecuencia).

3.3. De la medición de las acciones a la gestión: describa qué debe hacerse si las mediciones indican que el programa no se está implementando tal como se lo planificó.

3.4. Previsión de obstáculos: ¿qué obstáculos para la implementación se prevén, y cuál es el plan para superarlos?

3.5. Prevención de obstáculos: ¿qué obstáculos previstos podrían evitarse si el programa se planificase de otra manera?

Paso 4. Valor agregado y relación costo-beneficio que en teoría tiene el programa propuesto

4.1. ¿Cuántos delitos se cometieron el año pasado en la población a la que se apunta? Anote el mejor cálculo de la cantidad de delitos totales que se cometieron en la zona o entre las personas a las que apuntará el programa propuesto.

Cantidad de delitos

Fuente de la información

4.2. Proyección de una magnitud razonable del efecto del programa propuesto: ¿qué porcentaje de reducción en el volumen del problema de delincuencia se espera que el programa genere en la población intervenida en un período de un año?

__%

4.3. Cálculo de la cantidad total de delitos prevenidos: aplique el porcentaje de reducción al total de delitos denunciados el año pasado para determinar la cantidad total de delitos que se evitarán de acuerdo con esa reducción durante el primer año del programa.

Se prevendrán __ delitos de (complete con los tipos de delitos) _____.

4.4. Proyección del valor del beneficio que aportará esa magnitud de efecto: con todos los datos sobre costos que estén disponibles, determine el valor de evitar un delito (véase Cohen, 1979; Cook y Ludwig, 2002; Ministerio del Interior del Reino Unido--Home Office, 2001). Incluya información sobre el costo por delito de los siguientes conceptos, y sume los costos de todas las áreas.

Tratamiento médico de la víctima	_____
Días laborales perdidos	_____
Indemnización a víctimas de delitos	_____
Investigación	_____
Procesamiento	_____
Encarcelamiento	_____
Otros costos (enumérelos)	_____
Costo total por delito	_____

4.5. Proyección del costo anual del programa propuesto: enumere los costos del programa, tanto los externos como los costos operativos existentes, y súmelos todos para calcular el total por año.

<u>Concepto</u>	<u>Monto</u>
_____	_____
_____	_____
_____	_____
_____	_____
Otros	_____
Costo total	_____

4.6. Proyección del costo del programa propuesto por delito prevenido: divida el costo total del programa por la cantidad de delitos prevenidos.

$$\text{Costo anual del programa} = \frac{\text{_____}}{\text{delitos prevenidos}} = \text{_____}$$

4.7. Relación costo-beneficio: divida el costo del programa por delito prevenido por el costo de cada delito, y exprese el resultado como una razón entre el costo del delito y el del programa.

$$\text{Costo del programa por delito} / \text{costo de cada delito para la sociedad} = \text{_____ a _____}$$

4.8. ¿Es la propuesta eficaz en función de los costos?

Responda afirmativamente si el costo por delito cometido excede el costo por delito prevenido.

Responda negativamente si el costo por delito cometido es inferior al costo por delito prevenido.

4.9. Paso siguiente.

4.9.a. Si la propuesta no *va a ser* eficaz en función de los costos, dé por terminado este proceso DRIVER y comience uno nuevo.

4.9.b. Si la propuesta *va a ser* eficaz en función de los costos, continúe con los pasos 5 y 6.

Paso 5. Evaluación del programa propuesto

5.1. ¿La evaluación se realizará mediante un ensayo controlado aleatorio (ECA)? En caso afirmativo, consulte el siguiente sitio para obtener un ejemplo de protocolo experimental y el formulario que debe seguirse a fin de preparar un protocolo para el proyecto propuesto: <http://www.crim.cam.ac.uk/research/experiments/beck-crimport1.0.pdf> (en inglés). El protocolo CRIM-PORT reemplaza completamente esta parte de la lista de comprobación. Si no se usará un ECA, complete el resto de la sección.

5.2. ¿Cuáles son las unidades de análisis a las que se apunta? (Marque una opción).

Ciudades _____

Comunidades o barrios _____

Manzanas _____

Intersecciones de calles _____

Personas (delincuentes, víctimas, agentes de policía) _____

5.3. ¿Cómo se seleccionarán los casos?

5.4. ¿Cómo se elegirán los casos que recibirán la innovación?

5.5. ¿Cómo se elegirán las unidades de comparación que no recibirán la innovación?

5.6. ¿Cómo se medirán en la evaluación los recursos invertidos?

5.7. ¿Cómo se medirá la implementación (los productos) tanto en las unidades intervenidas como en las de comparación?

5.8. ¿Cómo se medirán los resultados tanto en las unidades intervenidas como en las de comparación?

5.9. ¿Cómo se compararán los resultados de las unidades intervenidas con los de las de comparación?

Ahora está usted listo para implementar el programa propuesto.

5.10. Resultados del análisis: a partir de la evidencia del período de comprobación inicial, ¿qué resultados pueden informarse mediante el plan de análisis descrito anteriormente?

Razón de la implementación de los productos: entre las unidades intervenidas y las de comparación =

Razón del costo de implementación de los productos: entre las unidades intervenidas y las de comparación =

Razón del impacto de los resultados: entre las unidades intervenidas y los casos contrafácticos =

¿En qué dirección apunta el efecto (las unidades intervenidas estuvieron mejor o peor)?

¿Cuál es la probabilidad estadística de que el efecto se haya producido por casualidad? (P = __)

Paso 6. Reflexión sobre el programa al final del período de comprobación

Este paso está destinado a implementarse al final de un período de comprobación inicial. Algunas innovaciones pueden comprobarse rápidamente, mientras que para otras tal vez haga falta más tiempo. Esto depende, en parte, del volumen de delitos y de la naturaleza de la práctica o del programa en cuestión. Asimismo, este paso puede ejecutarse en cualquier momento en que surjan dudas sobre la viabilidad del programa o la necesidad de hacer correcciones sobre la marcha.

6.1 Implementación: en términos generales, ¿el programa se implementó tal como se lo planificó?

6.1.a. Si la respuesta es que no, ¿qué elementos no se están implementando?

6.1.b. ¿Cuáles son los motivos principales de las fallas de implementación?

6.1.c. ¿Es posible corregir las fallas de implementación? Si lo es, ¿de qué manera?

6.2 Costos: ¿los costos del programa están dentro del presupuesto?

6.2.a. Si la respuesta es que no, ¿qué costos se están excediendo del presupuesto?

6.2.b. ¿Cuáles son los motivos principales por los que se excedió el presupuesto?

6.2.c. ¿Es posible corregir el exceso? Si lo es, ¿de qué manera?

6.3. Personal: ¿el programa tiene el personal adecuado?

6.3.a. Si la respuesta es que no, ¿qué personal debe agregarse?

6.3.b. ¿Cuáles son los motivos principales por los que no hay personal disponible?

6.3.c. ¿Es posible corregir la falta de personal? Si lo es, ¿de qué manera?

6.4. Impacto: ¿el programa está teniendo el impacto deseado (en la medida que sea)?

6.4.a. Si la respuesta es que no, ¿por qué?

6.4.b. ¿Es posible corregir la falta de impacto? Si lo es, ¿de qué manera?

6.5. Eficacia en función de los costos: ¿el nivel de impacto es suficiente para que el programa sea eficaz en función de los costos?

6.5.a. Si la respuesta es que no, ¿cuál es el motivo principal por el que no se está logrando un impacto suficiente?

6.5.b. ¿Cuál es la causa principal del bajo rendimiento del programa?

6.5.c. ¿Es posible corregir el problema del nivel de impacto? Si lo es, ¿de qué manera?

6.6. Paso siguiente: ¿qué se hará a continuación? (Marque una opción).

6.6.a. Se continuará el programa sin modificaciones.

—

6.6.b. Se suspenderá el programa y se comenzará un nuevo proceso DRIVER. ___

6.6.c. El programa se continuará tras algunas modificaciones.

(Describa los cambios). _____

Referencias

- Ariel, Barak, Lawrence Sherman y Jordi Vila. 2012. "Random Assignment without Tears: How to Stop Worrying and Love the Cambridge Randomizer." *Journal of Experimental Criminology* 8(2): 193-208.
- Beccaria, Cesare. 1764. *On Crimes and Punishment*. Londres. Disponible en: http://www.constitution.org/cb/crim_pun.htm.
- Bentham, Jeremy. 1781. *An Introduction to the Principles of Morals and Legislation*. Disponible en: <http://www.utilitarianism.com/jeremy-bentham/index.html>.
- Berk, Richard A., Lawrence Sherman, Geoffrey Barnes et al. 2009. "Forecasting Murder within a Population of Probationers and Parolees: A High Stakes Application of Statistical Learning." *Journal of the Royal Statistical Society, serie A, Statistics in Society*, 172 (parte 1): 191-211.
- Braga, Anthony A. y Brenda J. Bond. 2008. "Policing Crime and Disorder Hot Spots: A Randomized, Controlled Trial." *Criminology*, 46: 577-607.
- Bratton, William. 1998. *The Turnaround: How America's Top Cop Reversed the Crime Epidemic*. Nueva York: Random House.
- Campbell, Donald y Julian Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Cohen, Laurence y Marcus Felson. 1979. "Social Change and Crime Rates: A Routine Activity Approach." *American Sociological Review*, 44: 588-608.
- Cook, Philip y Jeins Ludwig. 2002. *Gun Violence: The Real Costs*. Oxford, MA: Oxford University Press.
- Di Tella, Rafael y Ernesto Schargrotsky. 2010. "Criminal Recidivism after Prison and Electronic Monitoring." Cambridge, MA: National Bureau of Economic Research. Disponible en: http://www.pazciudadana.cl/docs/ext_20101129105830.pdf.
- Gawande, Atul. 2010. *The Checklist Manifesto*. Nueva York: Henry Holt.
- Gladwell, Malcolm. 2005a. *Blink: The Power of Thinking without Thinking*. Boston, MA: Little, Brown and Company.
- . 2005b. *Outliers*. Boston, MA: Little, Brown and Company.
- Goldstein, Herman. 1990. *Problem-Oriented Policing*. Nueva York: McGraw-Hill.
- Hawken, Angela y Mark Kleiman. 2009. "Managing Drug Involved Offenders with Swift and Certain Sanctions: Evaluating Hawaii's HOPE." Washington, D.C.: Instituto Nacional de Justicia de Estados Unidos. Disponible en: <https://www.ncjrs.gov/pdffiles1/nij/grants/229023.pdf>.

- Hirschi, Travis. 1969. *Causes of Delinquency*. Berkeley, CA: University of California Press.
- Home Office. 2011. *Revisions Made to the Multipliers and Unit Costs of Crime used in the Integrated Offender Management Value for Money Toolkit*. Londres: Ministerio del Interior. Disponible en: <http://www.homeoffice.gov.uk/publications/crime/reducing-reoffending/IOM-phase2-costs-multipliers>.
- Jackson, Mark T. 2011. Tesis de maestría. Institute of Criminology, University of Cambridge, Police Executive Program.
- Johnson, Stephen. 2010. *Where Good Ideas Come From: The Natural History of Innovation*. Nueva York: Riverhead Books.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Nueva York: Farrar, Straus and Giroux.
- Lewin, Kurt. 1958). *Group Decision and Social Change*. Nueva York: Holt, Rinehart and Winston.
- Ludwig, Jens. 2006. Testimonio ante el Senado de Estados Unidos, Committee on the Judiciary, 19 de septiembre. Publicado en http://www.judiciary.senate.gov/hearings/testimony.cfm?id=e655f9e2809e5476862f735da11bb8b4&wit_id=e655f9e2809e5476862f735da11bb8b4-1-3.
- McCord, Joan. 1978. "A Thirty-Year Follow-Up of Treatment Effects." *American Psychologist*, 33: 284-289.
- . 2003. "Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs." *ANNALS of the American Academy of Political and Social Science*, 587: 16-30.
- Meehl, Paul. 1954. *Clinical versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. Mineápolis: University of Minnesota Press.
- Millenson, Michael. 1997. *Demanding Medical Excellence: Doctors and Accountability in the Information Age*. Chicago: University of Chicago Press.
- National Research Council. 2004. *Fairness and Effectiveness in Policing: The Evidence*. Washington, D.C.: National Academies Press.
- Olds, David, John Eckenrode, Charles R. Henderson, Jr. et al. 1997. "Long-Term Effects of Home Visitation on Maternal Life Course and Child Abuse and Neglect: Fifteen-Year Follow-Up of a Randomized Trial." *JAMA*, 278(8):637-643.
- Palmer, Ian y Andrea Caddick. 2010. Artículo presentado ante la British Society of Evidence-Based Policing. Manchester, noviembre.

- Petrosino, Anthony, Carolyn Turpin Petrosino y John Buehler. 2002. *“Scared Straight” and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency*. Oslo: Campbell Collective. Disponible en: www.campbellcollaboration.org.
- Popper, Karl. 1959. *The Logic of Scientific Discovery*. Nueva York: Basic Books.
- Ratcliffe, Jerry, Travis Taniguchi, Elizabeth R. Groff y Jennifer D. Wood. 2011. “The Philadelphia Foot Patrol Experiment: A Randomized Controlled Trial of Police Patrol Effectiveness in Violent Crime Hotspots.” *Criminology*, 49: 795-831.
- Renzema, Marc y Evan Mayo-Wilson. 2005). “Can Electronic Monitoring Reduce Crime for Moderate to High-Risk Offenders?.” *Journal of Experimental Criminology*, 1: 215–237.
- Sherman, Lawrence. 1992). *Policing Domestic Violence: Experiments and Dilemmas*. Nueva York: Free Press.
- . 1997. “Thinking about Crime Prevention.” En: Lawrence Sherman et al., *Preventing Crime: What Works, What Doesn't, What's Promising?* Informe presentado al Congreso de Estados Unidos. Washington, D.C.: Departamento de Justicia de Estados Unidos. Disponible en: <https://www.ncjrs.gov/works/chapter2.htm>.
- . 2007. “The Power Few Hypothesis: Experimental Criminology and the Reduction of Harm.” *Journal of Experimental Criminology*, 3(4). 52.
- . 2010. “Less Prison, More Police and Less Crime: How Criminology Can Save The States From Bankruptcy.” Conferencia presentada ante el Instituto Nacional de Justicia de Estados Unidos, 21 de abril. Washington, D.C.: Instituto Nacional de Justicia. Disponible en: <http://www.nij.gov/multimedia/presenter/presenter-sherman/>.
- . 2011. “Criminology as Invention.” En: Mary Bosworth y Carolyn Hoyle, *What Is Criminology?* Oxford, MA: Oxford University Press.
- y Dennis P. Rogan. 1995. “Effects of Gun Seizures on Gun Violence: Hot Spot Patrols in Kansas City.” *Justice Quarterly*, 12(4).
- y Robert Langworthy. 1979. “Measuring Homicide by Police Officers.” *Journal of Criminal Law and Criminology*, 70(4): 546-560.
- y Richard A. Berk. 1984. “The Specific Deterrent Effects of Arrest for Domestic Assault.” *American Sociological Review*, 49(2): 261-271.
- , Patrick R. Gartin y Michael E. Buerger. 1989. “Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place.” *Criminology*, 27:27-55.
- y David Weisburd. 1995. “General Deterrent Effects of Police Patrol in Crime Hot Spots: A Randomized, Controlled Trial.” *Justice Quarterly*, 12(4): 635-648.

- , Heather Strang y Barak Ariel. 2011. “Experiments without Grants: How to Conduct Low-Cost Randomized Trials.” Artículo presentado ante la American Society of Criminology, Division of Experimental Criminology.
- Soares, Rodrigo R. e Igor Viveiros. 2010. “Organization and Information in the Fight against Crime: An Evaluation of the Integration of Police Forces in the State of Minas Gerais, Brazil.” Documento de discusión Nro. 5079 del IZA. Bonn: IZA.
- Straus, Murray. 1996. “Spanking and the Making of a Violent Society.” *Pediatrics*, 98: 837-842.
- Taylor, Bruce, Christopher Koper y Daniel J. Woods. 2011. “A Randomized Controlled Trial of Different Policing Strategies at Hot Spots of Violent Crime.” *Journal of Experimental Criminology*, 7: 149-181.
- Tyler, Tom. 1990. *Why People Obey the Law*. New Haven: Yale University Press.
- Villaveces, Andrés, Peter Cummings, Victoria E. Espitia et al. 2000. “Effect of a Ban on Carrying Firearms on Homicide Rates in 2 Colombian Cities.” *JAMA*, 283: 1205-1209.
- Weisburd, David, Lawrence Sherman y Lisa Maher. 1993. “Contrasting Crime-Specific and Crime-General Theory: The Case of Hot Spots of Crime.” En: Freda Adler y William S. Laufer (eds.), *Advances in Criminological Theory*, vol. 4, pp. 45-70. Nuevo Brunswick, NJ: Transaction.
- Wikstgrom, Per-Olof H. y Vania Ceccato, Beth Hardie y Kyle Treiber. 2010. “Activity Fields and the Dynamics of Crime: Advancing Knowledge About the Role of the Environment in Crime Causation.” *Journal of Quantitative Criminology*, 26 (1): 55–87.
- Woods, Daniel J. 2010. Disertación doctoral. Filadelfia, PA: Universidad de Pensilvania, Departamento de Criminología.