

EMMIE and the What Works Centre for Crime Reduction

Progress, challenges, and future directions for evidence-based policing and crime reduction in the United Kingdom

Aiden Sidebottom

Nick Tilley

Abstract

The United Kingdom has made significant investments in promoting and facilitating evidence-based policy and practice. In this chapter our focus is on the What Works Centre for Crime Reduction (WWCCR) and in particular EMMIE. EMMIE is a framework denoting five categories of evidence that are important to inform policy and practice decision-making: **E**ffect, **M**echanism, **M**oderator, **I**mplementation and **E**conomics. As part of the WWCCR, EMMIE was used to populate a toolkit for police and others with crime prevention responsibilities to draw on in deciding what to do to address crime problems (the “Crime Reduction Toolkit”). Across the range of interventions for which systematic reviews had been undertaken, the toolkit summarizes the quality of both the available evidence and findings in relation to the five domains comprising EMMIE. An appraisal of systematic reviews using EMMIE revealed substantial gaps in the crime reduction evidence base, whose repair will require primary studies and reviews that draw on a wide range of studies using diverse methods.

Citation

Sidebottom, A. and Tilley, N. (2022). EMMIE and the what works centre for crime reduction: Progress, challenges, and future directions for evidence-based policing and crime reduction in the United Kingdom. In Piza, E. and Welsh, B. (eds.) *The Globalization of Evidence-Based Policing: Innovations in Bridging the Research-Practice Divide* (pp.73-91). London, UK: Routledge Press.

Introduction

The notion of evidence-based policy and practice has emerged as a unifying theme for many applied disciplines. It has been characterized as both a “movement” (Welsh, 2019) and a “new paradigm” (Greenhalgh, Howick & Maskrey, 2014), inspired by developments in medicine in the 1990s (Guyatt et al., 1992) and extending to diverse fields such as development, education and, the focus of this chapter, policing and crime reduction. The basic idea of evidence-based policy and practice is disarmingly simple. It describes an approach in which policy and practice decisions combine professional judgment, a consideration of ethics, and reliable research evidence. Indeed, the public might be forgiven for assuming that such an approach describes conventional ways of working: on what other basis are decisions made? And yet, in crime reduction as with other policy areas, studies have repeatedly found inconsistencies in what is known from research and what is done in practice: decisions are more ordinarily governed by intuition, precedent, and popular sentiment, together with the continuance of practices shown to be ineffective, inefficient, and even harmful (see Lum, 2009; Welsh & Rocque, 2014). Evidence-based policy and practice calls for a better integration of scientific and tacit knowledge. Proponents argue that the net effect will be improvements in efficiency, effectiveness, and the minimization of harm (Sherman, 2013).

The United Kingdom (UK) government has made significant investments in efforts to enable global scientific evidence to play a greater role in shaping policy and practice. In 2013, the What Works Network, designed to “improve the way government and other organizations create, share and use high quality evidence for decision-making,” was launched.¹ The What Works Network was initially formed of nine independent What Works Centres, each charged with gathering together the best available evidence and making it available in a digestible form for potential users. Each What Works Centre covers a different policy area—from “health and social care” to “educational achievement” and “improving quality of life of older people”—which collectively account for an estimated £250 billion of annual public expenditure in the UK. As noted by Gough et al. (2018: 5), the “What Works Network represents one of the first attempts to take a national approach to prioritizing the use of evidence in public policy decision-making.”

This chapter focuses on one member of the What Works Network: The What Works Centre for Crime Reduction (WWCCR) based in the College of Policing, the professional body for policing in England and Wales responsible for setting the standards for police recruitment, training, and development. We describe the rationale, methods and findings from a three-year research program led by University College London which supported the early work of the WWCCR (for a fuller discussion see Hunter, May & Hough, 2017). Our involvement in that research program related mainly to a systematic effort to identify, appraise, standardize, and

distill research evidence for the purposes of informing policy and practice. The process and outputs of that exercise form the central thrust of this chapter. Our focus is twofold. First is the development and application of EMMIE, an acronym denoting five kinds of evidence that are important to crime reduction decision-making. Second is the Crime Reduction Toolkit, a web-based tool structured around EMMIE and designed to summarize and make accessible research evidence on crime reduction interventions.

The remainder of this chapter is formed of four sections, with each building on that which precedes it. We begin by describing the elements of, and rationale for, the EMMIE framework. We then discuss how EMMIE was drawn on in populating the Crime Reduction Toolkit. Next, we provide an overview of what has been found about the adequacy of the crime reduction evidence base for informing policy and practice through the systematic application of the EMMIE framework to a large sample of systematic reviews. We conclude with some suggestions about what is needed to improve the evidence base in policing and crime reduction, focusing on the role of primary studies, systematic reviews, and practitioners and policymakers.

EMMIE and evidence-needs

Basing decisions on evidence requires an evidence base. It also requires that evidence be accessible and relevant to those intended to use it. A key primary task in our work

to support the WWCCR was therefore that of summarizing what was known about the effectiveness of crime reduction interventions and, crucially, presenting our findings in a form that would help improve policy and practice.

Our starting point was to identify what we see as the evidence-needs of policymakers and practitioners with crime reduction responsibilities. We identified five, which we have referred to using the acronym, 'EMMIE' (Johnson et al., 2015). Needs are not necessarily the same as wants. The components of EMMIE describe forms of evidence in the absence of which decision makers are liable to fail to meet their objectives, to produce unwanted unintended net outcomes, or to use their resources ineffectively or inefficiently. Of course decision makers may not always realize what evidence can help them come to better conclusions. Indeed, convincing decision makers of the need for evidence in the interests of improvement has lain behind the whole movement towards evidence-based policy and practice. The contribution of EMMIE is to identify a broader range of evidence-needs than is often realized.

EMMIE refers to Effects (What is known so far about the net effects of the intervention?); Mechanisms (What is known so far about how the intervention's positive and negative effects are brought about?); Moderators (What is known so far about the contextual conditions for the activation of mechanisms producing positive and negative effects?); Implementation (What is known so far about the practicality of

introducing and operating the intervention?); and Economics (What is known so far about the costs, cost-effectiveness and cost–benefit outcomes of the intervention?).

What is the rationale for EMMIE? Those making decisions about whether to adopt an intervention have an interest in knowing its track record. What net effects has it had (broadly defined), and how have these been comprised in terms of both positive and negative effects? These findings will not necessarily be directly transferable to the decision maker’s specific situation, but an intervention that has tended to produce net positive effects clearly tells in its favor and likewise those that have been found repeatedly to produce net negative effects tells against it. This focus on net effects has been the main concern of randomized controlled trials (RCTs) and their quasi-experimental counterparts, a form of research design that has grown considerably in number in policing and crime reduction in recent decades (Neyroud, 2017), albeit that the strict standards used in ideal RCTs (including triple blinding) are rarely met (Sidebottom & Tilley, 2020).

Crime reduction interventions, like most others, rarely have consistent effects in all conditions and for all who are subject to the intervention. Moreover, they can have negative as well as positive effects that again vary by condition and by target. In view of this uncertainty, decision makers need to be able to draw on evidence that goes beyond simply an intervention’s net effects. These explain why we included mechanisms and moderators as evidence-needs in EMMIE. Knowing how interventions work (mechanisms), when, where, and amongst whom they are applied

(moderators or context) is crucial to understanding how the particular conditions faced by the decision maker matter for selecting or rejecting a particular intervention (Pawson & Tilley, 1997). It can also inform decisions about the modification and/or scaling-up of interventions.

Putting crime reduction interventions into practice is often complex and difficult (see Homel & Brown, 2017). There can, for example, be legal, organizational, ideological, occupational, cultural, and political hurdles to overcome. Decision makers therefore need to know what has been found about the conditions that enable or obstruct the implementation of any intervention they are contemplating and what has been learned about how it can be put in place. They need this to decide whether to adopt an intervention and, if so, how it can best be introduced in the conditions in which it is being applied. Many intervention failures are the result of implementation failures (Ekblom, 2010). Hence, evidence relating to implementation is crucial to improve decision-making in policing and crime reduction.

Finally, all decision makers have limited resources, which can be put to different uses. Decision makers therefore have an obvious interest in being cost-effective—delivering an intervention efficiently so that the unit cost of activity, output, and outcome are not higher than necessary. They need to make sure that the net benefits of the intervention are higher than the net costs. They need to have evidence that allows them to gauge whether the allocation of their finite resources is

achieving the best utilities overall. All this requires the collection and analysis of evidence relating to economic costs and benefits (Manning et al., 2016).

As will become clear from the discussion ahead, EMMIE was devised specifically with systematic reviews of crime reduction interventions in mind. This form of research has come to be seen as synonymous with evidence-based policy and practice (Bullock, 2020). When done well, systematic reviews are generally considered to be the “best” of the “best available evidence” to support decision-making (Neyroud, 2018). In this vein, the Crime Reduction Toolkit described here took as its unit of analysis systematic reviews of crime reduction interventions. However, despite this focus on systematic review evidence, it warrants mention at this point that the elements encapsulated in EMMIE are considered to have wider applicability to other domains and research designs. For example, EMMIE has been applied elsewhere for reviews concerned with child social care (Brand et al., 2019) and has clear relevance to the design needs of primary evaluation studies (Tilley & Westhorp, 2020).

EMMIE and the crime reduction toolkit

As outlined above, EMMIE was devised to draw attention to the evidence-needs of policymakers and practitioners. Our next step was to determine the extent to which these needs have been met. This involved two stages. First was the development of a coding instrument to assess the coverage and quality of systematic review evidence

across the five dimensions of EMMIE. Second was the application of this coding instrument to systematic reviews in crime reduction, which were identified using systematic search methods. In this section we discuss each stage in turn.

Systematic reviews, like all forms of research, are variable in their quality, completeness and, as a result, their utility for supporting decision-making. Several evidence appraisal tools have been developed in an effort to distinguish “better” from “worse” systematic review evidence, most notably AMSTAR (Shea et al., 2007) and ROBIS (Whiting et al., 2016). Informed by previous appraisal tools and in consideration of those features specific to research on policing and crime reduction,² we devised a coding instrument to extract and appraise the coverage and methodological quality of systematic review evidence across the five dimensions of EMMIE (Tompson, Bowers, Johnson & Belur, 2015).³ The coding instrument contains 105 items organized into six sections. The first section relates to descriptive information such as review author(s) and date of publication. The remaining five sections correspond to the five elements of EMMIE. In these five sections of the coding instrument, items fall into two broad categories. ‘EMMIE-E’ items refer to the findings themselves. Here information is simply taken from that which is reported in the systematic review being assessed (e.g. sample size, source(s) of data and so on). ‘EMMIE-Q’ items assess the *quality* of evidence. To do this, a bespoke scoring system was devised to assess the reliability and validity of findings for each element of EMMIE (Tompson et al., 2015; Johnson et al., 2015). These items are then

summed together to produce an overall quality rating for each element of EMMIE, scored using a five-point scale from 0 to 4, with 4 denoting the highest quality evidence, as indicated in Box 5.1. Each systematic review could thus score a maximum of 20 points (4 points per element of EMMIE). The basic assumption underlying the EMMIE-Q scale is that, all things being equal, systematic reviews that adequately attend to more of the issues of importance will be of more use to policymakers and practitioners.

Box 5.1

Rating scales for the methodological quality of evidence for each element of EMMIE (EMMIE-Q scores)

Effect

0 = Most forms of bias that could influence the study conclusions remain

1 = Although the review was systematic, many forms of bias that could influence the study conclusions remain

2 = Although the review was systematic, some forms of bias that could influence the study conclusions remain

3 = The review was sufficiently systematic that many forms of bias that could influence the study conclusions can be ruled out

4 = The review was sufficiently systematic that most forms of bias that could influence the study conclusions can be ruled out

Mechanism

0 = No reference to theory (i.e. a simple black box)

1 = General statement of assumed theory underpinning the intervention

2 = Detailed description of theory drawn from prior work

3 = Full description of theory of change and testable predictions generated from it

4 = Full description of theory of change and robust analysis of whether this operated as expected

Moderator

0 = No reference to relevant contextual conditions

1 = Ad hoc description of possible relevant contextual conditions

2 = Tests of the effects of contextual conditions defined post hoc using readily available variables

3 = Theoretically grounded description of relevant contextual conditions

4 = Collection and analysis of relevant data relating to theoretically grounded contextual conditions

Implementation

0 = No account of implementation or implementation challenges

1 = Ad hoc comments on implementation or implementation challenges

2 = Concerted efforts to document implementation or implementation challenges

3 = Evidence-based account of levels of implementation or implementation challenges

4 = Evidence-based account of implementation or implementation challenges and specification of what would be necessary for faithful replication elsewhere

Economics

0 = No mention of costs (and/or benefits)

1 = Only direct or explicit costs (and/or benefits) estimated

2 = Direct or explicit and indirect and implicit costs (and/or benefits) estimated

3 = Marginal or total or opportunity costs (and/or benefits) estimated

4 = Marginal or total or opportunity costs (and/or benefits) by bearer/recipient estimated

The next stage in our research was to apply the EMMIE coding instrument. To do this we performed a systematic search for systematic reviews in crime reduction (for a fuller discussion see Tompson et al., 2020). This involved keyword and hand searches of all major criminology and criminal justice databases and forward and backward citation tracking. To be eligible for inclusion, studies had to: (1) employ systematic review and/or meta-analytic methods; (2) report an outcome measure that demonstrated a quantifiable impact on crime (as opposed to related behaviors such as aggression); (3) be published between 1975 and 2015; and (4) be written in English.

Our initial searches returned nearly fifteen thousand documents, which, when subject to detailed screening, left 337 reviews that met our inclusion criteria (see Bowers, Tompson, & Johnson (2014) for an analysis of these 337 reviews). Eligible reviews fell into two broad categories: those which related to a specific crime reduction intervention (e.g. CCTV) and those which related to a specific population or behavior (e.g. youth offending or drunk-driving). As mentioned above, given that interventions constituted the unit of analysis for the Crime Reduction Toolkit, application of the EMMIE coding instrument began by focusing only on those systematic reviews which related to single crime reduction interventions (n = 70). Following the categorization system used by Weisburd, Farrington and Gill (2017), these 70 reviews related (in descending order) to the following broad areas of crime reduction: correctional interventions (n = 12), situational interventions (n = 7), developmental and social prevention (n = 5), sentencing and deterrence (n = 5), restorative justice (n = 4), other (n = 4), community interventions (n = 3), policing and partnerships (n = 3), drug treatment (n = 2) and publicity (n = 1). The most common category of reviews was defined as mixed (n = 21), combining elements of two or more categories of crime reduction intervention, such as Policing and Situational interventions or Community and Drug Treatment. Each review was independently coded by two researchers using the EMMIE appraisal instrument.

Figure 5.1 is a screenshot of the Crime Reduction Toolkit.⁴ It shows that for each intervention there are five columns denoting the five elements of EMMIE. In

each column, information is presented on both the EMMIE-E and EMMIE-Q scores. By way of illustration, consider the first entry: “After school clubs.” EMMIE-E, which denotes the presence or absence of evidence, is represented by a series of symbols. The symbols in columns four and five are blank. This indicates that in the assessed systematic review(s), no information was provided on Implementation and Economics. The columns relating to Effect, Mechanisms, and Moderators are shaded, indicating that evidence pertaining to these elements of EMMIE was reported. For Effect, the combination of crosses and ticks indicates the strength and direction of conclusions drawn from the review(s) on the net positive and negative effects of, here, after school clubs (details on the meaning of different combinations of ticks and crosses are provided on the toolkit homepage). The horizontal bar beneath each symbol relates to the assessed quality of evidence per element of EMMIE (i.e. the EMMIE-Q score), as set out in Box 5.1. One shaded square denotes a one rating and a full shaded bar denotes the top rating of 4. Clicking on each intervention takes the user to a brief narrative summary based on the systematic reviews and organized according to EMMIE.

Intervention ▲▼	Impact on crime ▲▼ Effect	How it works ▲▼ Mechanism	Where it works ▲▼ Moderator	How to do it ▲▼ Implementation	What it costs ▲▼ Economic cost	
After school clubs	X✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Aggression Replacement Training (ART) for reducing antisocial behaviour	X✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Alcohol pricing	✓✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Alley gating	✓✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Boot camps	X✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Brief interventions for alcohol use disorders within criminal justice settings	✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦
Car breathalyser lock	✓✓ ██████	⚙️ ██████	📍 ██████	❓ ██████	£ ██████	✦

FIGURE 5.1: The Crime Reduction Toolkit

The adequacy of the evidence base for crime reduction policy and practice

Application of the EMMIE coding instrument to 70 systematic reviews provided the material with which to populate the Crime Reduction Toolkit, and to do so in a

standardized way oriented towards the needs of crime reduction decision makers. It also allowed us to chart the trends and gaps in the assessed systematic reviews. This speaks to a key question mentioned previously: the extent to which the crime reduction evidence base addresses the needs of policymakers and practitioners. It is to these trends and gaps that we now turn.

Before proceeding, it is important to acknowledge a methodological decision, built into our research brief from the start, which may have biased the nature of the evidence base we uncovered. In populating the Crime Reduction Toolkit, the starting point was to identify systematic reviews that focused on drawing together findings that estimated net effects. As reported above, to be included in our sample a systematic review had to report an outcome measure that demonstrated a quantifiable impact on crime. Simply put, our search criteria were effects oriented. In adopting this strategy it is possible that we may have neglected reviews that did not meet the conventional levels of internal validity. Moreover, those conducting the reviews, whose findings we summarize in EMMIE terms, are likely to have ignored the many studies that were not oriented to the production of net effects findings with high levels of internal validity. Indeed, the conventionally found attrition patterns from the large numbers of studies initially identified to the normally very small number of experimental and quasi-experimental studies that pass muster in internal validity terms and which form the basis of the synthesis, likely overlooks what could be learned of MMIE from excluded studies.

Following this, what we found, perhaps unsurprisingly, is that the systematic reviews we did identify and code tended to be weak in their treatment of mechanisms, moderators and implementation. More surprisingly, however, they were almost universally weak in relation to evidence on economics. The surprise follows because the net effects concerns of most systematic reviews are well suited to monetizing benefits (using standard costs of crimes of the sort provided in Heeks et al., 2018) which can be compared to intervention costs data, provided these have been collected.

Table 5.1 shows our findings in terms of evidence quality for each dimension of EMMIE in the 70 single-intervention systematic reviews we coded. Several themes are apparent. Given the previously mentioned selection methods, the rather better scores for Effect are to be expected, with 56 percent of identified reviews scoring a 3 or 4 in terms of evidence quality, indicating that many or most forms of bias could be confidently ruled out. Internal validity of the studies was generally high. However, only one in seven reviews scored a 3 or 4 for Mechanism. The remaining reviews lacked a theory of change for the intervention under review, much less a tested one. Much the same goes for Moderators and the specification of contextual variations that might affect intervention impact, again only one in seven reviews stated any theoretical grounds for expecting variation in outcome patterns. Implementation was covered slightly better with one in five reviews providing empirical evidence of what was done by way of implementation and of the hurdles that had to be overcome. In regard to economics, the findings revealed widespread inattention, with only one of

the 70 reviews scoring a 3 and none a 4. What all this indicates is that none of the identified systematic reviews that promised potential adequacy in the measurement of effect size had provided a sufficient breadth of evidence to meet the needs of decision makers contemplating the intervention in question. We think this is a major finding.

Q-score	Effect	Mechanism	Moderator	Implementation	Economics
0	4	9	10	11	64
1	13	36	21	30	5
2	14	15	29	16	1
3	25	8	5	12	0
4	14	2	5	1	0
Total	70	70	70	70	70

Source: Tompson, L. et al. (2020). How Strong is the Evidence-Base for Crime Reduction Professionals? *Justice Evaluation*.

Table 5.1: Evidence quality for 70 single-intervention systematic reviews in EMMIE terms

An EMMIE-informed research agenda for evidence-based policing and crime reduction

Our use of EMMIE to appraise the coverage and quality of evidence in a large sample of systematic reviews revealed two main findings. In relation to evidence regarding the effectiveness of interventions, the quality of systematic review evidence was found to be generally high. In relation to the other aspects of EMMIE, evidence was found to be seldom reported and, where it was, the quality of evidence was generally low. Systematic review evidence on the cost-effectiveness of crime reduction interventions was noticeably sparse.

It is perhaps unsurprising that we found the evidence base in policing and crime reduction to be stacked mainly with studies concerned with establishing the effectiveness of an intervention. As indicated previously, this may be partly a consequence of our own search strategy. It may also reflect the dominant narrative of evidence-based policy and practice, which has become synonymous with the shorthand “what works?” as opposed to the less catchy but more nuanced and EMMIE-compliant question of “what worked for whom, how and under what circumstances?” (Pawson & Tilley, 1997). To be clear: robust evidence on the effects of an intervention forms a key component of evidence-based policing. Practitioners and policymakers are clearly better equipped to make decisions about whether or not to invest in, modify, or abandon an intervention if evidence is available on its effects and the size of those effects. The generally high levels of internal validity in those systematic reviews assessed using EMMIE is therefore clearly a positive and important finding. And yet, we remain persuaded that the evidence provided by

rigorous studies assessing effect sizes of interventions at particular times and places, though useful of course, does not capture all the evidence-needs of decision makers in policing and crime reduction. Even a series of discrete studies, we believe, is insufficient. They show that a given intervention *can* be effective but not that it will always be so. Decision makers are faced with specific problems in specific situations. They need evidence that will help them do better than they would without the evidence. As we argued earlier, past track records have a role to play, but ideally more is needed if evidence-based policing is to make good on the promise of delivering improvements in effectiveness and efficiency. The MMIE elements of EMMIE highlight additional evidence-needs which we argue are important to decision makers. The reviews we examined as part of our contribution to the WWCCR show that so far systematic reviews have not done well in providing the full range of evidence needed to improve decisions, even when they provide technically strong measurements of effect size.

In this final section we set out some ideas for expanding effectiveness evidence in the interests of improving policing and crime reduction. We have three generic suggestions in relation to this. The first is that primary studies of the effectiveness of interventions be broadened to encompass all components of EMMIE. The second is that evidence reviews be broadened to include studies that may not meet the conventional quality thresholds applied in systematic reviews that are focused only on effect sizes of interventions, but which have something useful to say

about the MMI and E of an intervention. The third is that decision makers being invited to make greater use of research evidence be alerted to how they can best draw on evidence relating to all the components of EMMIE.

Primary studies

Evaluation is complex. Different studies tend to address different questions (Tilley & Westhorp, 2020). Some focus mainly or only on implementation. Some treat interventions as black boxes and assess net effects without concern for how these effects have been brought about. Some focus on equity in allocating resources or on outcomes. Some focus on the intervention theory, either by developing a logic model or by devising and testing “context-mechanism-outcome pattern configurations” (Pawson & Tilley, 1997). Some evaluations are oriented to theory development and testing, some to accountability for the funding allocated to an initiative, some to cost-effectiveness, some to cost–benefit calculations, and some to generalization. The notion that evaluations can be put on some kind of hierarchy, with a gold standard to which all should aspire, misses the variations in purpose: methods used are oriented to purposes.

EMMIE summarizes what a comprehensive evaluation, meeting all decision maker needs, would look like. Of course, not all evaluations will need to address all evidence-needs. If the potential of an intervention to create net intended outcomes is already established, as is the case with interventions such as hotspots policing (see

Braga, Turchan, Papachristos & Hureau, 2019) or focused deterrence (see Braga & Weisburd, 2012), the next questions might, for example, relate to how to implement it effectively, how to deliver it at the lowest cost, to understand the subgroups for which it is ineffective (or harmful) and so on (Sidebottom & Tilley, 2020). We believe that this sort of cumulative research agenda, which is oriented towards the different elements of EMMIE and likely involving multiple methods of the kind Donald Campbell (after whom the Campbell Collaboration was named) came to embrace (Campbell, 1987), would be useful in building an improved evidence base for policing and crime prevention.

Reviews

We recognize that the commissioning of EMMIE-compliant primary evaluations may be expensive. It is therefore with more broadly-based systematic reviews that we think greater progress is likely to be made. Such reviews would synthesize evidence using more inclusive criteria than are traditionally applied and be oriented to assembling evidence that speaks to all the components of EMMIE, not just effects.

This is not standard practice. As alluded to previously, most systematic reviews begin with a wide and comprehensive search for eligible studies, but fetch up with a very small number (sometimes none) that pass methodological muster (see e.g. Perkins et al., 2017). It's as if nothing useful can be learned from those studies tossed onto the spoil heap. There are, of course, clear advantages to strict sifts that eliminate

the vast majority of potentially eligible studies in that the residue tends to be small and manageable and the extraction of data for meta-analysis relatively straightforward. Selecting from the vast number of other studies to identify the subset that speaks usefully to MMIE is tricky. Indeed, we believe there is more work to be done on how to objectively and reliably extract and synthesize information that relates to MMI and E. Currently the best that can be advocated are searches that begin with convenience and stop when no more is being learned from the studies read, a version of analytic induction—the method developed by Znaniecki (1934) and famously used by Becker (1953) in his sample of those learning to become marijuana users. Our own efforts, relating to systematic reviews on the use of alley gating to prevent domestic burglary (Sidebottom et al., 2018) and security tags to reduce shop theft (Sidebottom et al., 2017; Sidebottom & Tilley, 2018) began with studies selected for their robustness in measuring effect size but then strayed beyond these in an opportunistic way to try to cover information on MMI and E as best we could. Box 5.2 shows what we did in relation to the systematic review of security tags.

Box 5.2

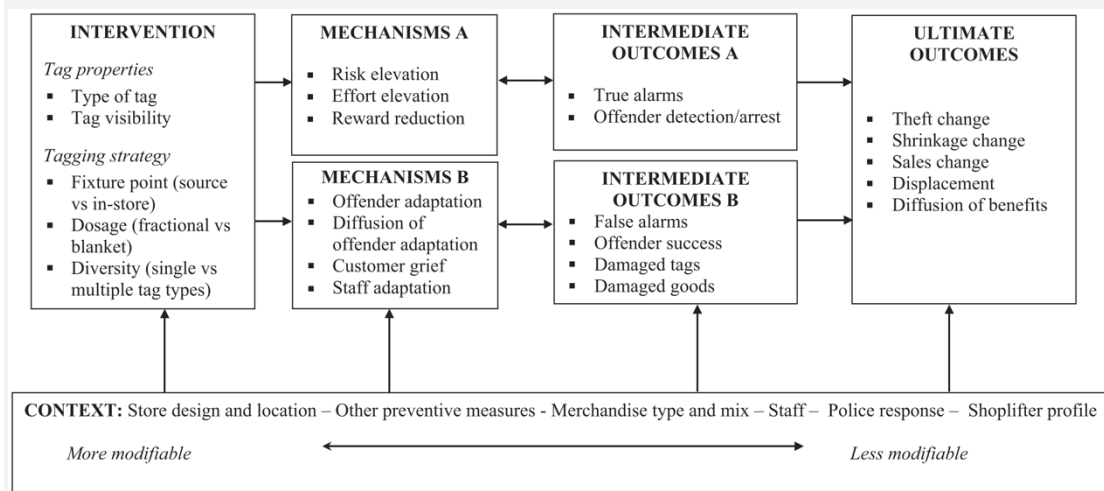
An EMMIE-informed review of tagging to prevent shop theft

We performed systematic keyword searches of the published and unpublished literature, hand-searched relevant journals and conducted forward and backward citation searches to identify studies that provided evidence on the use of tags to reduce

theft in shops. Eligible studies were then assessed and retained on the basis that they reported evidence on one or more of the five elements of EMMIE. In addition, we approached retailers and were able to assemble several internal reports of quasi-experiments conducted by one large retailer to assess the effectiveness and cost-effectiveness of tags they were contemplating for their stores. This extended the range of gray literature on which we could draw in assessing effect sizes. It also revealed something of the decision-making that goes into implementing tags (for example, the trade-off between reduced losses and costs). We also visited retailers to obtain a better idea of how tagging was being implemented in practice. In this way, drawing on a wide range of pre-existing studies using diverse methods, we were able to build a theory of tagging successes and failures and the conditions providing for those successes and failures.

The figure below shows the theory of tagging that emerged from the review as we undertook it. The four columns depict a causal sequence running left to right. The first (intervention) highlights key considerations regarding the type, prevalence, and installation of tags. The second (mechanism) refers to the causal processes through which tags produce their effects. Columns three and four identify intermediate and ultimate outcome measures we might expect to see following the activation of said mechanisms. In each case both positive and negative effects are presented, reflecting the mixed results reported in the literature. Bridging these columns is the bottom panel which sets out those conditions thought to moderate the activation of tagging

mechanisms. These are organized according to the extent to which they might plausibly be modified by retailers (and related parties) in the interests of loss prevention: those factors towards the left-hand side are generally more amenable to change than those on the right.



The synthesis of quantitative findings on net effects with observational studies, our own observations, and discussions with those involved in undertaking tagging, informed by some general, tested theory in situational crime prevention, enabled the production of an EMMIE-type review. This attended not only to the net-effect findings of experimental studies but also to mechanisms, moderators, and implementation issues alongside an understanding of the economic considerations employed in practice by retailers deciding whether to use tagging and if so how.

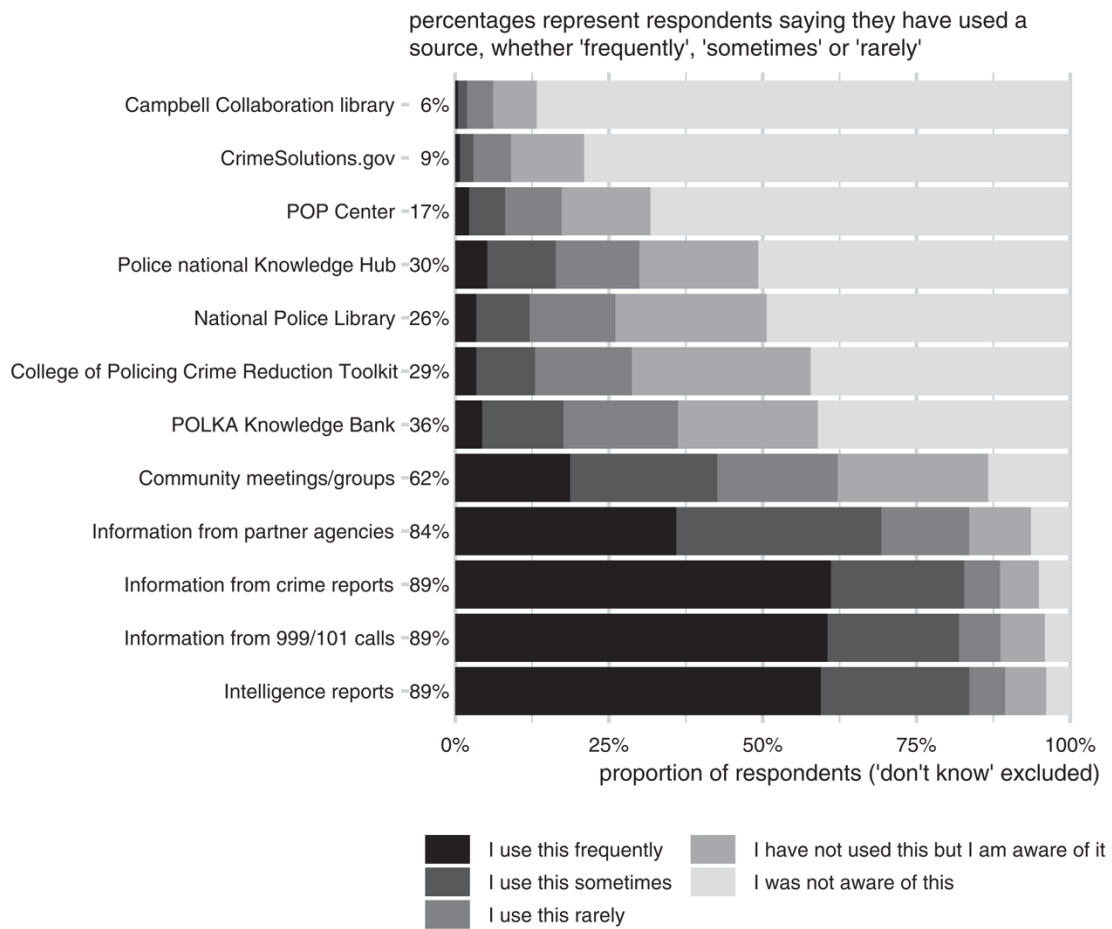
Decision makers and evidence

Persuading policymakers and practitioners in all fields to pay attention to evidence and to do so in ways that improve performance is notoriously difficult (Nutley, Davies & Walter, 2002). Even amongst enthusiastic advocates of evidence-based medicine, there are concerns that what has been done in some cases may have led to unintended deteriorations in care, where results of RCTs have been applied thoughtlessly in a kind of ready reckoner way (Greenhalgh et al. 2014). Policing and crime reduction are no different. Studies have identified numerous barriers to the routine use of research evidence in practice, ranging from a lack of familiarity with research evidence (Tompson et al., 2017), variations in receptivity to research evidence (Telep, 2016), and a tendency to privilege experience and the “craft” of policing over research evidence (Weisburd & Neyroud, 2011).

The challenges are substantial. How do we get the right evidence to the right potential users and equip them to make use of that evidence in ways that will improve the chances they make better decisions? The answer does not appear to be one of simply supplying the evidence. There are now numerous resources which variously seek to package and present research evidence in the interests of guiding practice—the Crime Reduction Toolkit described herein is but one such example. Figure 5.2 shows the findings of a recent survey of police officers and staff in England and Wales from a sample of 20 police forces (Sidebottom et al., 2020). It makes sobering

reading. When asked about the sources of information they typically consult when trying to better understand and respond to issues of crime and disorder, only 6 percent of respondents reported frequently making use of the Campbell Collaboration library, for example. By comparison the Crime Reduction Toolkit embracing EMMIE fared better, but still not well. It appears that police officers and staff, in England and Wales at least, much more commonly turn inwardly to their own sources of information than to research-based depositories.

Revealed rather than reported consultation of the WWCCR provides a slightly more optimistic picture, at least with regard to the direction of travel. Figure 5.3 shows the trend in monthly hits for the Crime Reduction Toolkit from inception to mid-2020. The numbers remain modest, but they are gradually increasing. However, consultation of the Crime Reduction Toolkit does not mean it was used or whether when used it was used sensibly. Indeed, evidence is currently lacking on the extent to which the Crime Reduction Toolkit has effected positive change in policing and crime reduction (Hunter, May & Hough, 2017).



74% of respondents answered these questions

Figure 5.2: Sources of information drawn on when problem-solving (n = 3,064)
 Source: Sidebottom et al. (2020)

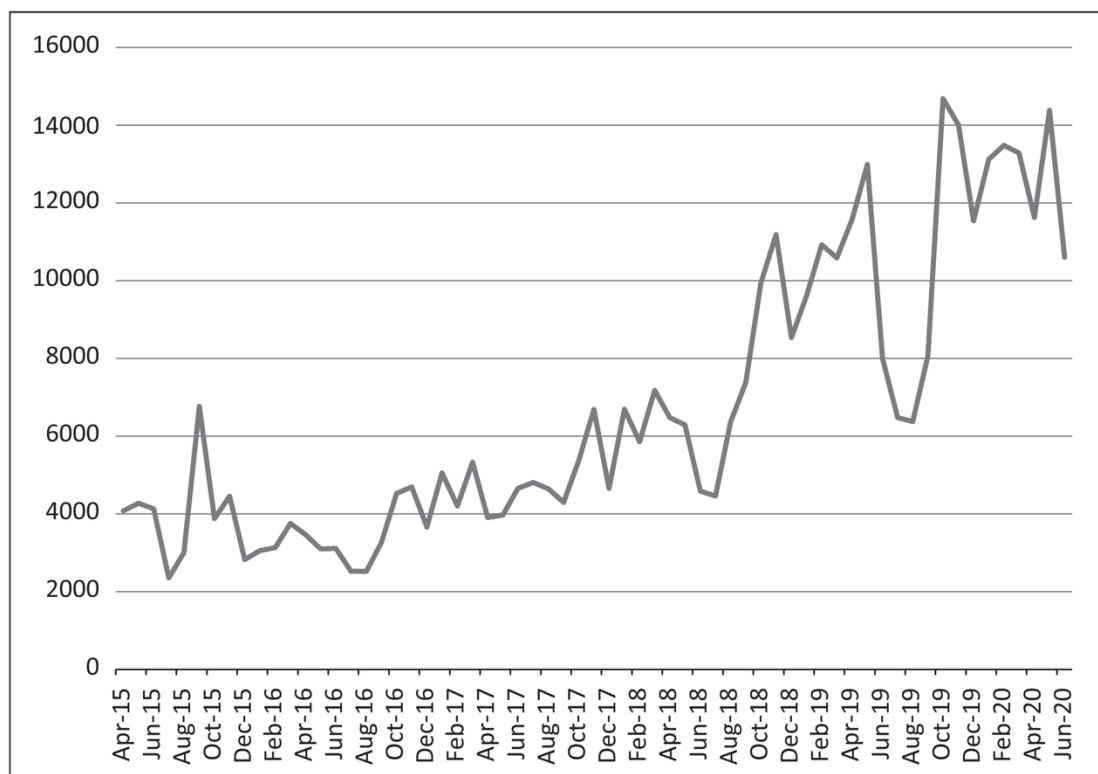


Figure 5.3: Monthly hits on the Crime Reduction Toolkit (April 2015 to June 2020)

We think much more needs to be done better to understand how evidence is used and what can be put in place to extend reference to it in informing improved decision-making. This might involve rather radical rethinking of the collection, synthesis, and presentation of evidence. For example, this chapter has taken as its starting point the What Works Network established by the UK Government and the conceptions of evidence and evidence use assumed by them and reflected in the WWCCR and the Crime Reduction Toolkit. An alternative (or maybe complementary and additional) starting point is found in the popcenter website.⁵ At the time of

writing, the popcenter website contains 74 guides relating to specific crime and public safety problems the police are expected to address. These guides work through what the evidence suggests about the possibilities of effective responses to deal with presenting problems, and what needs to be considered in selecting and implementing those responses. These guides come closer to the problems that policymakers and practitioners are trying to address and the conditions in which they decide what to do. To our knowledge, however, there is as yet no formal methodology for determining better or worse specific problem guides.

Where universal measures are being considered for widespread problems (for example, car theft and security measures or malaria and vaccination), net across-the-board findings of the sort tested by RCTs are clearly crucial (albeit that the other elements of EMMIE are also relevant). Where problems and their contexts are more heterogeneous, problem-focused evidence may be more important. This suggests that what works for evidence collection and review is context dependent. This in turn suggests that evidence users need to be apprised of the kinds of evidence they should be drawing on.

Conclusion

The case for evidence-based policy and practice is both intuitive and compelling. Evidence-based development has improved livelihoods. Evidence-based medicine has saved lives. Against this backdrop, the rise of evidence-based policing and crime

reduction observed in recent decades now seems inevitable: it is what we have come to expect from doctors and dentists so why not those with responsibilities for crime control and public safety?

We have endeavored in this chapter to explain the background thinking to EMMIE and to describe efforts to apply it as an improved way of finding, synthesizing, and presenting evidence that can be used by decision makers to improve outcomes in policing and crime reduction. Our findings from a comprehensive effort to identify all systematic reviews of evidence using robust methods to measure effect sizes of interventions suggest that the evidence base for crime reduction interventions does not adequately meet the evidence-needs of policymakers and practitioners, especially with regard to mechanisms, moderators, implementation and economic value. Our discussion of these findings suggests that primary studies could usefully be broadened to look beyond robust measurement of the effect sizes of interventions using experimental and quasi-experimental methods. However, more yet might be achieved by synthesizing evidence from evaluation studies focusing on different questions and using different methods to provide EMMIE-based accounts of the promise of interventions that can be drawn on by decision makers. We are encouraged that the kinds of development we are advocating for policing and crime prevention broadly reflect some recent developments and debates in evaluation, evidence collection, and evidence use in healthcare (see Cowen & Cartwright, 2020).

In the interests of starting debate and informing future research, we finish with five precepts for EMMIE-informed, evidence-based policing and crime reduction. It:

1. informs choice of interventions in relation to local and national problems by robust measurements of past effectiveness
2. shows how interventions produce their intended and unintended outcomes so that attempted replications capture their active ingredients
3. sensitizes decision makers to conditions needed for a given intervention to produce intended effects while not producing negative side effects
4. alerts policymakers and practitioners to ways to implement interventions successfully and forewarns them of implementation hurdles, and
5. estimates costs and benefits to inform the allocation of scarce resources that could be put to alternative uses.

References

- Becker, H S (1953). 'Becoming a Marihuana User.' *Am. J. Soc.* 59: 235–242.
- Bowers, K., Tompson, L., & Johnson, S. (2014). Implementing information science in policing: mapping the evidence base. *Policing: a journal of policy and practice*, 8(4), 339–352.

- Braga, A. A., & Weisburd, D. L. (2012). The effects of focused deterrence strategies on crime: A systematic review and meta-analysis of the empirical evidence. *Journal of Research in Crime and Delinquency*, 49(3), 323–358.
- Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: an update of an ongoing systematic review and meta-analysis. *Journal of experimental criminology*, 15(3), 289–311
- Brand S., Morgan F., Stabler L., *et al* (2019). ‘Mapping the evidence about what works to safely reduce the entry of children and young people into statutory care: a systematic scoping review protocol.’ *BMJ Open* 9:e026967.
- Bullock, K. (2020). Research synthesis, systematic reviewing and evidence-based policing. In Fielding, N., Bullock, K. and Holdaway, S. (eds) *Critical reflections on evidence-based policing*. Abingdon, Oxon: Routledge.
- Campbell, D. (1987). ‘Guidelines for monitoring the scientific competence of preventive intervention research centres’ *Knowledge: Creation, Diffusion, Utilisation* 8(3): 389–430.
- Cowen, N., & Cartwright, N. (2020). Street-level Theories of Change: Adapting the Medical Model of Evidence-based Practice for Policing. In Fielding, N., Bullock, K. and Holdaway, S. (eds) *Critical reflections on evidence-based policing*. Abingdon, Oxon: Routledge.
- Ekblom, P. (2010). *Crime prevention, security and community safety using the 5Is framework*. Basingstoke, Hampshire: Springer.

- Gough D, Maidment C, & Sharples J. (2018). *UK What Works Centres: Aims, methods and contexts*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.
- Greenhalgh, T., Howick, J., & Maskrey, N. (2014). 'Evidence based medicine: a movement in crisis?' *Bmj*, 348.
- Guyatt, G. and 30 other members of the Evidence-Based Medicine Working Group (1992). 'Evidence-Based Medicine.' *JAMA* 268(17) 2420–2425.
- Heeks, M, Reed, S., Tafsiri, M. & Prince, S. (2018). *The Economic and Social Costs of Crime*, Second Edition, Research Report 99. London: Home Office.
- Homel, P. & Brown, R. (2017). Implementation: partnership and leverage in crime prevention. In Tilley, N. and Sidebottom, A. (eds) *Handbook of Crime Prevention and Community Safety*, second edition. Abingdon, Oxon: Routledge.
- Hunter, G., May, T., & Hough, M. (2017). An evaluation of the 'What Works Centre for Crime Reduction'. Final Report. Available at:
https://whatworks.college.police.uk/About/Documents/ICPR_Final_Evaluation_WWCCR.pdf
- Johnson, S., Tilley, N. & Bowers, K. (2015). Introducing EMMIE: An evidence rating scale for crime prevention policy and practice. *Journal of Experimental Criminology*. 11(3): 459–473.

- Lum, C. (2009). Translating police research into practice. *Ideas in American policing*, 11, 1–15.
- Manning, M., Johnson, S., Tilley, N., Wong, G. & Vorsina, M. (2016). *Economic Analysis and Efficiency in Policing, Criminal Justice and Crime Reduction: What Works?* London: Palgrave Macmillan.
- Neyroud, P.W. (2017). *Learning to Field Test in Policing: Using an Analysis of Completed Randomised Controlled Trials Involving the Police to Develop a Grounded Theory on the Factors Contributing to High Levels of Treatment Integrity in Police Field Experiments*. Doctoral dissertation, University of Cambridge.
- Neyroud, P.W. (2018). Systematic Reviews: “Better Evidence for a Better World. In R. Mitchell and L. Huey, (2018), *Evidence-Based Policing: An Introduction*. Bristol: Policy Press.
- Nutley, S., Davies, H., & Walter, I. (2002). *Evidence based policy and practice: Cross sector lessons from the UK*. Swindon, UK: ESRC UK Centre for Evidence Based Policy and Practice.
- Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*, London: Sage.
- Perkins, C., Beecher, D., Aberg, D., Edwards, P. & Tilley, K. (2017). Personal Security Alarms for the Prevention of Assaults Against Healthcare Staff. *Crime Science* 6: 11.

- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., ... & Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC medical research methodology*, 7(1).
- Sherman, L. W. (2013). The rise of evidence-based policing: Targeting, testing, and tracking. *Crime and justice*, 42(1), 377–451.
- Sidebottom, A. & Tilley, N. (2018). Towards a theory of tagging in retail environments. In V. Ceccato and R. Armitage (eds) *Retail Crime International Evidence and Prevention*. Switzerland: Springer.
- Sidebottom, A. & Tilley, N. (2020). Evaluation evidence for evidence-based policing: Randomistas and realists. In Fielding, N., Bullock, K. and Holdaway, S. (eds) *Critical reflections on evidence-based policing*. Abingdon, Oxon: Routledge.
- Sidebottom, A., Bullock, K., Armitage, R., Ashby, M., Clemmow, C., Kirby, S., Laycock, G. and Tilley, N. (2020). Problem-oriented policing in England and Wales 2019. Jill Dando Institute of Security and Crime Science, University College London.
- Sidebottom, A., Thornton, A., Tompson, L., Belur, J., Tilley, N., & Bowers, K. (2017). ‘A systematic review of tagging as a method to reduce theft in retail environments.’ *Crime Science* 6: 7.

- Sidebottom, A., Tompson, L., Thornton, A., Bullock, K., Tilley, N., Bowers, K. & Johnson, S. (2018). 'Gating Alleys to Reduce Crime: A Meta-Analysis and Realist Synthesis. *Justice Quarterly*. 35(1): 55–86.
- Telep, C. (2016). 'Police Officer Receptivity to Research and Evidence-Based Policing: Examining Variability Within and Across Agencies', *Crime and Delinquency*, 63(8), 976-999.
- Tilley, N., & Westhorp, G. (2020). 'Evaluation Research, Quantitative.' In P. Atkinson, S. Delamont, A. Cernat, J.W. Sakshaug, & R.A. Williams (Eds.), *SAGE Research Methods Foundations*. United Kingdom: SAGE Publications.
- Tompson, L. A., Bowers, K. J., Johnson, S. D., & Belur, J. B. (2015). EMMIE evidence appraisal coding tool [Dataset]. London, UK: UCL Department of Security and Crime Science. Accessed 20 December 2019. Available at <https://discovery.ucl.ac.uk/id/eprint/1462098/>
- Tompson, L., Belur, J., Morris, J., & Tuffin, R. (2017). How to make police–researcher partnerships mutually effective. In Tompson, L. and Knuttson, J. (eds.) *Advances in evidence-based policing* (pp. 175–194). Abingdon, Oxon: Routledge.
- Tompson, L., Belur, J., Thornton, A., Bowers, K. J., Johnson, S. D., Sidebottom, A., Tilley, N. & Laycock, G. (2020). How Strong is the Evidence-Base for Crime Reduction Professionals? *Justice Evaluation Journal*, 4:1, 68-97.

- Whiting, P., Savović, J., Higgins, J. P., Caldwell, D. M., Reeves, B. C., Shea, B., ...
& Churchill, R. (2016). 'ROBIS: a new tool to assess risk of bias in systematic
reviews was developed'. *Journal of Clinical Epidemiology*, *69*, 22–234.
- Weisburd, D. & Neyroud, P. (2011). 'New perspectives in policing: Police science—
Toward a new paradigm', Harvard Executive Session on Policing and Public
Safety, Washington, DC: National Institute of Justice.
- Weisburd, D., Farrington, D. P., & Gill, C. (2017). What works in crime prevention
and rehabilitation: An assessment of systematic reviews. *Criminology &
Public Policy*, *16*(2), 41–449.
- Welsh, B.C. (2019). Advocate: evidence-based policing for crime prevention. In
Weisburd, D., & Braga, A. A. (Eds.). (2019). *Police innovation: Contrasting
perspectives*. Cambridge, United Kingdom: Cambridge University Press.
- Welsh, B. C., & Rocque, M. (2014). When crime prevention harms: A review of
systematic reviews. *Journal of Experimental Criminology*, *10*(3), 24–266.
- Znaniecki, F. (1934). *The Method of Sociology*. Farrar and Rinehart, New York.

Acknowledgments

We would like to acknowledge the contributions of colleagues to research carried out
in support of the *What Works Centre for Crime Reduction* which is discussed here.

They include Jyoti Belur, Karen Bullock, Kate Bowers, Nigel Fielding, Shane

Johnson, Gloria Laycock, Amy Thornton and Lisa Tompson. We also thank Karen Bullock, Gloria Laycock, Lisa Tompson and the editors for commenting on earlier drafts of this chapter.

Notes

¹ www.gov.uk/guidance/what-works-network

² For example, issues of crime displacement or diffusions of crime control benefits are central to policing and crime reduction but are not featured in generic evidence appraisal instruments.

³ The EMMIE evidence appraisal instrument can be accessed here:

<https://discovery.ucl.ac.uk/id/eprint/1462093/>

⁴ <https://whatworks.college.police.uk/toolkit/Pages/Welcome.aspx>

⁵ <https://popcenter.asu.edu>