CrossMark

Information retrieval in systematic reviews: a case study of the crime prevention literature

Lisa Tompson[1] · Jyoti Belur[1]

Abstract

Objectives A defining feature of a systematic review is the data collection; the assembling of a meticulous, unbiased, and reproducible set of primary studies. This requires specialist skills to execute. The aim of this paper is to marshal tacit knowledge, gained through a systematic search of the crime prevention literature, to develop a 'how-to guide' for future evidence synthesists in allied fields.

Methods Empirical results from a recent systematic search for evidence in crime prevention are supplied to illustrate key principles of information retrieval.

Results Difficulties in operationalizing a systematic search are expounded and possible solutions discussed. Empirical results from optimizing the balance between sensitivity and precision with the criminological literature are presented. An estimation of database overlap for crime prevention studies is provided to guide other evidence synthesists in streamlining the search process.

Conclusions A high-quality search will involve a substantial time investment in honing the research question, specifying the precise scope of the work, and trialing and testing of search tactics. Electronic databases are a lucrative source of eligible studies, but they have important limitations. The diversity of expression across the criminological literature needs to be captured by the use of many search terms—both natural language and controlled vocabulary—in database searches. Complementary search tactics should be employed to locate eligible studies without common vocabulary. Grey literature should be ardently pursued, for it has a central role in the crime prevention evidence base.

✉ Lisa Tompson
l.tompson@ucl.ac.uk

✉ Jyoti Belur
j.belur@ucl.ac.uk

[1] Department of Security and Crime Science, University College London, 35 Tavistock Square, London WC1H 9EZ, UK

🦢 Springer

Introduction

Systematic reviews are a contemporary method in the evolution of evidence synthesis;
aiming to synthesize findings from primary studies to build a composite picture of the
evidence in a field. They supplant traditional narrative reviews, common throughout the
20th century, which have been extensively criticized for having opaque methods and
potentially unreliable conclusions (Glass et al. 1981). The defining feature of a
systematic review is that it uses explicit and accountable methods (Gough et al.
2013). The method is grounded in the principle that decision points in the research
process should be reported in enough detail so they are transparent and replicable by
others, and findings can be scientifically verified. By embracing wide-ranging and
comprehensive search tactics, through which multidisciplinary academic and grey
literature is pursued, and having "firm editorial criteria" (White 2009, p. 53) that are
explicitly documented, systematic reviews mitigate the possibility of bias.

When appropriate, and possible, primary study data located via systematic search
methods are statistically aggregated by meta-analytic techniques. Crime and criminal
justice studies have been slower than other fields—for example, psychology and
medicine—to embrace meta-analytic approaches for knowledge accumulation objec-
tives (Wells 2009). However, the last decade has seen a tremendous increase in
application of this aggregative method; in 1995, Petrosino reported that there had been
13 meta-analyses of crime prevention programmes to date (Petrosino 1995), in 2014
this number exceeded 184 (Bowers et al. 2014a).

A transparent search strategy is a distinctive characteristic of systematic reviews and
is intended to avoid subjectively choosing the sample of studies for synthesis. Subjec-
tivity in the search and selection process—synonymous with data collection—creates
the risk of producing unreliable or unrepresentative findings (Brunton et al. 2013).
Hence, objectivity is essential for calculating unbiased estimates of effect in quantita-
tive meta-analysis, as well as thematic configuration in qualitative evidence synthesis.

Although information retrieval is a specialist skill (Rothstein and Hopewell 2009),
relatively little scholarly attention has been paid to how primary studies are retrieved
while synthesizing evidence in the fields of criminology, criminal justice, and crime
prevention. Pioneering work by Petrosino (1995) is one of the few attempts to describe
in detail the data collection phase in a crime prevention meta-analysis. Whilst compre-
hensive, and still relevant for many of the principles of searching, this work is dated,
since the technological tools used to search and locate relevant records for evidence
synthesis have rapidly changed in the last two decades.

This paper is intended as an update of Petrosino's 1995 methodological note on
information search and retrieval. It documents the experience of applying principles of
information retrieval espoused by the Campbell Collaboration (Hammerstrøm et al.
2010) to a research project assembling the evidence base in crime prevention. In doing
so, we elicit the practical difficulties involved in systematically searching for evidence
in the current digital era and present possible solutions for overcoming them.

We begin by describing the nature and background of the research project that will be used to illustrate our arguments. Subsequently, we present our experiences of applying the principles of information retrieval in the area of crime prevention and discuss our key findings. The final section presents our conclusions and some guidelines to aid future researchers conducting systematic reviews in allied fields. This is the first attempt of its kind to develop a 'how-to' guide to conduct systematic searches in crime prevention and can be extended to criminology and criminal justice studies.

Overcoming bias

The introduction of conscious and unconscious bias into the systematic review process can take many forms. Conceptually, the most important form of bias is selection (or sampling) bias, in which the sample of studies selected fundamentally differs from the wider population of studies. In evidence synthesis nomenclature this is known as *publication bias*, and has been acknowledged in the academic literature for four decades (Greenwald 1975). It might be apt to rename this 'easier to find bias' since studies that are more visible are more likely to feature in a systematic review.

Strong, positive effect sizes—that is, in support of a hypothesis—are attractive to the academic community, to journal editors, and to funding bodies. They are therefore likely to be published more (publication bias), published more rapidly (time lag bias), cited more often (citation bias), published in multiple publications and outlets (multiple publication bias), and published in English (language bias) (Alderson and Green 2002).

Studies published in academic outlets (such as journal articles) are most visible to other researchers. Empirical research has established that published studies are not representative of all the research findings that are generated (Lipsey and Wilson 1993; Rothstein and Hopewell 2009; Smith 1980)—a problem commonly referred to as the 'file drawer' problem (Rosenthal 1979). That is, findings that produce neutral or negative results often languish in the bottom of filing cabinets, rather than get written up for publication in peer-reviewed journals (Dickersin 2005; Wilson 2009). If publication bias favors studies with positive results, then the aggregation or synthesis of these skewed results can culminate in a review that overestimates the effect of an intervention. Basing policies or practices on such erroneous effects can lead to a wastage of resources and potentially worsen a problem.

In criminology and criminal justice substantial research is conducted and produced by practitioners, or funded by bodies that receive no benefit in publishing findings in peer-reviewed journals. If written up at all, practitioner generated research is published outside of the control of academic publishers—known as *grey (gray) literature* (Auger 1998). Omitting grey literature from a systematic review can be costly, for Smith (1980) found that academic publications reported a higher programme effect than comparative studies in the grey literature.

A high-quality systematic review, then, aims to minimize bias in the sampling of studies for synthesis. In fields like medicine, *exhaustive searching* is the 'gold standard', as the point of such systematic reviews is to aggregate the findings of quantitative data (DeLuca et al. 2008). Here, researchers strive for a "sufficient sample of studies for unbiased aggregation" (Brunton et al. 2013, p.108), which involves identifying studies in the universe of studies that fit the inclusion criteria. Crucially,

however, it is rarely possible to know the limits of the universe of the population being studied (i.e., the sampling frame).

The survey of crime prevention evidence syntheses

As part of the UK Government's commitment to evidence-led policy-making, the 'What Works Centre for Crime Reduction' (WWCCR) was formed in 2013. This centre, hosted by the College of Policing (CoP), is supported by a 'commissioned partnership programme', delivered by a consortium of universities. A vital component of the research programme is a project to systematically identify the evidence base for interventions with a crime reduction focus. The overarching aim of the research programme is to inform the decision-making of practitioners and policy-makers.

Our research question for this project was set by the mission of the WWCCR: *to identify the best available evidence on approaches to reducing crime (and the potential savings to the police service, their crime reduction partners and the public).*[1] We defined 'best available evidence' as the findings from systematic reviews (anchored in any methodological philosophy) or meta-analysis evidence syntheses. 'Approaches to reducing crime' was translated into the broadly defined 'crime prevention', which required a measured crime reduction or prevention outcome (for more details on the inclusion criteria see Bowers et al. 2014b). Our overall aim was to search for evaluations of interventions in all relevant fields that might have a crime prevention outcome. To our knowledge, this was the first time such an ambitious task had been attempted.

To identify evidence on crime prevention we adopted principles of high-quality information retrieval commonly used by evidence synthesists (see more below). In this tradition, we created a review protocol that outlined the search strategy to be used (Bowers et al. 2014b). For quality assurance, this was reviewed by a panel of external experts and the core team at CoP. The search strategy articulated in the protocol stated that we would (1) review lists of known systematic reviews in crime prevention; (2) perform keyword searches of electronic databases; (3) review 'grey literature' reports from government, research and policing organizations; (4) check the reference lists of relevant studies; and (5) consult with experts. The literature searches were performed at the authors' institution and the UK National Police Library, with additional support from information specialists in the US and Australia.

Overall, the search tactics yielded a list of around 17,500 research articles, which were imported into the EPPI Reviewer 4 software tool.[2] This web-based research synthesis tool assisted in automatically identifying many of the duplicates, had functionality for coding and data extraction and provided a means of tracking allocated workloads. Once duplicate records were removed just under 14,000 articles remained. An extensive two-stage screening process was then employed; the first stage sifting the articles on title and abstract for potential candidate studies, with the second-stage screening the shortlisted full-text articles for eligibility against a set of inclusion criteria (Bowers et al. 2014b). Concomitantly with the second stage, the articles that met the inclusion criteria ($n=320$) were 'light coded' to extract basic information about each evidence synthesis such as study characteristics,

---

[1] http://college.pressofficeadmin.com/component/content/article/45-press-releases/593
[2] See http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3396 for more information.

population, outcomes, and methods used.[3] A description of the 'evidence map' produced from these light codes is available in Bowers and Tompson 2014.

The search strategy

We first reviewed materials from organizations with a well-established pedigree in systematic review methods. Chiefly, this included the internationally renowned Campbell Collaboration (Petrosino et al. 2001). We drew heavily from their comprehensive guidance document on searching for studies by Hammerstrøm and colleagues (2010), and viewed their specialist SPECTR databases and YouTube channel of conference sessions. We garnered further best practice from another known center of excellence—the UK University of York's Centre for Reviews and Dissemination.

Our search strategy was designed to follow procedures identified in best practice, but the process was also iterative to balance sensitivity and precision (Hammerstrøm et al. 2010). Sensitivity relates to casting a wide enough net so that all relevant studies are retrieved, whereas precision refers to the proportion of relevant studies to all studies retrieved. Crucially, precision and sensitivity are incompatible aims; highly sensitive strategies result in extensive searches that generate a lot of leads to irrelevant studies (false positives) (White 1994). High-precision strategies necessitate low sensitivity, and result in a higher number of false negatives (relevant studies not identified). Balancing the two is vital.

As the objective was to amass the evidence base in crime reduction, our protocol was intentionally sensitive. This meant that we chose many different sources and generated a wide range of keywords to apply in electronic searches. It is well known that the knowledge base across criminal justice is disparate and fragmented (Farrington et al. 1986; Hammerstrøm et al. 2010), and that studies on crime prevention span many fields. To ensure that we were not limited to our own disciplinary boundaries (and thereby liable to introduce bias into the selection of studies—see Brunton et al. 2013), wide-ranging and multi-disciplinary sources were selected. The feedback garnered from circulating the protocol assisted in the development of our sources and keywords. As principal sources of primary studies, the search process was started by conducting preliminary searches of relevant electronic bibliographic databases.

Electronic bibliographic databases

Each electronic bibliographic database usually focuses on a discipline, collection of disciplines, or a commercial publisher, and within this purview a collection of research records are held. Some are subject area (e.g., criminal justice) or discipline (e.g., social science) specific, others are multi-disciplinary (e.g., SCOPUS and Web of Science) and contain millions of records. Since no single database contains all extant research records, it is essential to discover the scope of each database by examining their particular focus. While Google Scholar is a popular source for validating search results

---

[3] More detailed coding is now being undertaken in accordance with the 'EMMIE' evidence appraisal tool (Johnson et al. 2015). This tool was under development when the 'light coding' was done. 337 reviews were originally identified, but greater scrutiny of these resulted in the attrition of 17 studies, largely due to reviews being identified as duplicate publications (i.e., the review authors wrote multiple publications on the same topic), and a few not precisely meeting the inclusion criteria.

from electronic databases, the limitations of its functionality mean that it has yet to usurp professional literature retrieval tools (Boeker et al. 2013).

Each record in a database has standard fields (such as Citation details) and indexed fields that can be used in a search. Oftentimes an abstract will be contained within a record. A few databases hold the full text documents for each record, but this is not typical. Detailed overviews of the databases relevant to criminology and criminal justice focused research questions are available in Reed and Baxter (2009) and Hammerstrøm et al. (2010).

Selecting appropriate databases to search is important, but usually constrained by access rights. At an early stage of our project, we carried out a scoping exercise that determined where the key journals likely to publish systematic reviews on crime prevention topics were indexed. These journals were derived from lists of known systematic reviews and meta-analyses in the criminal justice field (Petrosino 1997; Weisburd and Farrington 2015; Wells 2009). We also called upon information specialists in other countries to search databases not available to us (CINCH and Social Sciences Full Texts).

Generating search terms for electronic databases

Keyword searches of electronic databases require the generation and testing of individual search terms, and thereafter the compilation of a search syntax combining all the terms. Generating search terms requires a two-fold approach: the creation of a list of *natural language* search terms, and the creation of a list of *controlled vocabulary* terms. Controlled vocabulary is the manual assignment of descriptors to a research record, and is done to link studies that might not share the same natural language vocabulary. They are supposed to 're-express' the content of documents with standard indices, often referred to as subject headings (White 1994). (Anticipated controlled vocabulary for the current study can be found in the author keywords for this paper.) The multidisciplinary nature of criminology and criminal justice studies results in idiosyncratic and scattered vocabulary. For this reason, to reduce false negatives, all relevant synonyms and appropriate subject headings should be used in a search (Reed and Baxter 2009).

The research question is the orientating force for search term generation. The question, therefore, should be expressed clearly and precisely, and subsequently be broken down by major ideas (concepts) or components of the question. A list of terms and their synonyms can then be developed which articulate these individual components. In our survey of crime prevention evidence syntheses we needed to identify *crime* terms, *prevention* terms and *research design* terms.[4] Defining crime prevention required a considered approach; we chose to break it down into crime *and* prevention, which afforded the opportunity to define crime types of interest. We compiled a list of these from UK government websites,[5] and augmented this with terms for more general

---

[4] We considered using frameworks (like PICO/T and SPIDER, Cooke et al. 2012) developed in clinical fields that assist in structuring a research question and its accompanying inclusion criteria. However, none of these provided a satisfactory fit with our focal research question.

[5] Including http://webarchive.nationalarchives.gov.uk/20130128112038/http://www.justice.gov.uk/downloads/legal-aid/eligibility/list-of-criminal-offences.pdf and

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/116226/user-guide-crime-statistics.pdf

criminal behavior. We then considered international equivalents—for example 'delinquent' and 'offender' for the term 'criminal'.

To refine our keywords, and to assist in generating controlled vocabulary, we took advice from information specialists who recommend using known (potentially) eligible studies to establish a baseline of search terms (Hammerstrøm et al. 2010; White 1994). These were harvested from the reference lists of existing meta-analyses and systematic searches in criminal justice and crime prevention (Petrosino 1997; Weisburd and Farrington 2015; Wells 2009). This search tactic is known as *backwards reference checking*, which is often undertaken towards the end of a search strategy. In our experience, however, this is fruitful to do at the outset, for it produces a set of studies that can be used in various ways to benchmark the relevance of syntaxes developed for electronic database searches.

We looked up the potential candidate studies (n=199, 63 of which were eventually appraised as eligible; see Table 3) to ascertain how they were recorded in electronic databases. In doing so, we found a large variance in the controlled vocabulary used to index the research records. Prominently, we found that records indexed in separate databases held by the same commercial provider[6] had different controlled vocabulary. This chimes with the warnings issued in the information specialist literature; index terms are considered relatively robust in medical databases, but are inconsistently applied in other databases (Hammerstrøm et al. 2010). So whilst controlled vocabulary appears more objective, it is not infallible; it is often produced by postgraduate students without domain expertise who subjectively interpret the study.

Further, we used database thesauri whenever available to generate broader and narrower terms associated with crime and crime prevention. These thesauri organize controlled vocabulary in a database into an alphabetical listing or hierarchical system so that a term can be used to search for other related terms. When searching in some databases, one can select clusters of controlled terms to appear in the search; this is referred to as 'exploding' a subject heading (index term) so that its subordinate terms are also used. Thesauri are database-specific, therefore controlled vocabulary needs to be tailored for each database searched and particular care must be taken if multiple databases are searched simultaneously through a commercial provider interface.[7]

The final stage of generating search terms related to the research design component of our inclusion criteria—evidence syntheses. This stipulated that a study had to either be a systematic review and/or a meta-analysis. As there are no developed reporting guidelines for systematic reviews in crime prevention (see Sidebottom and Tilley 2012), there are many search terms that could be used by scholars to describe an evidence synthesis. For example, 'scientific review' and 'bibliographic review' are,

---

[6] For example the ProQuest interface holds the following databases: Applied Social Sciences Index and Abstracts (ASSIA), Criminal Justice Periodicals, Education Resources Information Centre (ERIC), International Bibliography of the Social Sciences (IBSS), National Criminal Justice Reference Service (NCJRS), Sociological Abstracts (SA) and ProQuest theses and dissertations.

[7] Whilst it is possible to search multiple databases simultaneously through a commercial provider (e.g., in our case, ProQuest), the strengths and weaknesses of such an approach must be considered. The strengths are that it can reduce the time expended to search and export the results, and can automatically remove duplicate records, which assists in the early screening phase. The limitations are that non-universal controlled vocabulary cannot be applied, nor a specific numbers of records from each database easily identified. In our experience, the forfeit of database-specific index terms in a simultaneous search was not compatible with our highly sensitive search strategy and so we opted for searching each ProQuest database separately.

perhaps, two of the less obvious terms used. For this reason, we consulted search filters designed to identify systematic reviews in other disciplines,[8] and added terms we felt might be used across the social sciences and public health literature.

Testing search terms

It is prudent to test search terms before embarking on searching electronic databases (Petrosino 1997). Dip sampling some results may reveal missing search terms. Additionally, estimates of the likely search hits obtained can be compared to available resources. This testing can be likened to the 'piloting' stage of a primary study, whereby the feasibility, time expended and adverse effects can be estimated. Some search terms in their first incarnation may be too broad, and return an unmanageable number of hits in a database search. For example, one of the search terms in our first list was 'review', representing the research design we were interested in. When this term was used to search the title, abstract, and keywords in the large multi-disciplinary electronic database SCOPUS, over three million hits were returned. Other imprecise terms were 'drugs' and 'disorder', which retrieved an unwieldy number of medical studies.

To increase the precision of our keywords, we incorporated speech marks, wildcards, and proximity operators into our search terms. Speech marks are simply used to treat the words within a phrase as related, such as "calls for service" and "bodily harm". Larger databases typically have functionality for inserting wildcards into search terms. These take the form of left-hand truncation (e.g., *crime), right-hand truncation (e.g., crime*), or both. Wildcards can also be inserted in the middle of words to account for international spelling variations (if the database does not automatically search for these); for example, 'offen?es'. These can be applied with ascending sophistication depending on the functionality of the database.

Furthermore, proximity operators offered in many databases help increase the precision of terms. For instance, a search for studies that focus on offenders who have committed firearms offences can be expressed in many ways. Ignoring synonyms here for the sake of clarity, the two key search terms are *firearms* and *offenders*. Using these with the Boolean operator AND yields an unmanageable number of search hits when applied to the title and abstract fields. Using a proximity operator (e.g., NEAR or WITH) allows the association between the words to be specified within given parameters. For example, the following syntax specifies that the first word (with wildcard for plurals) needs to be located within five words of the second (also with wildcard)— "firearm? NEAR/5 offender?".

We performed sensitivity analysis on each search term with the aim of generating the most precise version to be included in the final search. This involved trialing different combinations of the wildcards and proximity operators, and empirically testing the results to ascertain if the known studies (as derived through the backwards searching described above) would be captured.

The next step was translating the search syntax to each database. Typically this requires identifying non-universal controlled vocabulary and technical conventions.

---

[8] Search filters are clusters of words which have been derived to yield a high precision of studies with a particular research design. A valuable repository of these is available at the ISSG Search Filters Resource— https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home

Careful reading of the FAQs and search tips provided by each database reveals appropriate application of wildcards and proximity operators alongside limitations that searchers should be mindful of. Not all databases alert a searcher when they misapply these operators, and therefore adhering to technical conventions of each database can protect against unsuitable results being returned.

Building up search syntaxes

Electronic database fields can be searched with individual search terms, but the most powerful types of searches apply a *search syntax*, which is a collection of search terms interspersed with Boolean operators (see below). The chief operators referred to here are 'AND' and 'OR'; the 'NOT' operator should be applied with caution due to the instability of the results it can return (Hammerstrøm et al. 2010).

As noted previously, a review's research question will steer the search strategy and syntax used. The area of interest usually lies at the intersection of multiple concepts— be they population, intervention, comparator, or outcome. Our search had three primary concepts: crime, prevention, and an evidence synthesis research design. Figure 1 illustrates these as a Venn diagram: the shaded area represents our area of interest— where we wished to locate relevant studies. The Boolean operator AND is used to retrieve studies that fall within the overlap between such concepts.

Recall that researchers may use different words to describe the same concepts— especially across fields studying similar phenomena from different ideologies. This diversity of expression is essential to capture in a search, which is the reason synonyms for each of the concepts from the research question are used. These synonyms are combined with the Boolean OR operator. A useful schematic from Hammerstrøm (no date) is adapted in Fig. 2 to demonstrate how the OR and the AND operators are
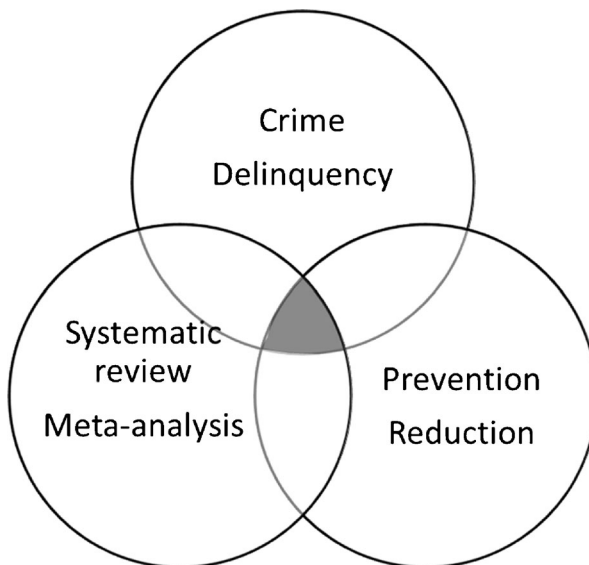

Fig. 1 A Venn diagram illustrating the locale of the studies of relevance to the survey of crime prevention evidence syntheses

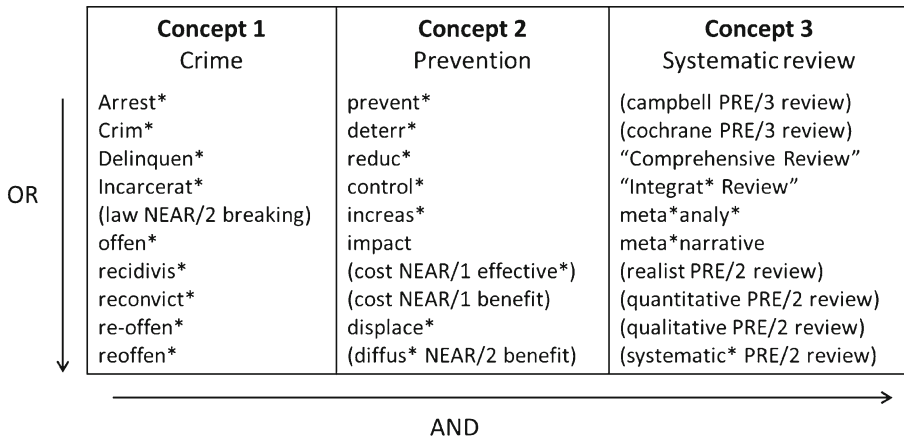| Concept 1 Crime	Concept 2 Prevention	Concept 3 Systematic review
Arrest* Crim* Delinquen* Incarcerat* (law NEAR/2 breaking) offen* recidivis* reconvict* re-offen* reoffen*	prevent* deterr* reduc* control* increas* impact (cost NEAR/1 effective*) (cost NEAR/1 benefit) displace* (diffus* NEAR/2 benefit)	(campbell PRE/3 review) (cochrane PRE/3 review) "Comprehensive Review" "Integrat* Review" meta*analy* meta*narrative (realist PRE/2 review) (quantitative PRE/2 review) (qualitative PRE/2 review) (systematic* PRE/2 review)

OR

AND

Fig. 2 Schematic illustrating how search terms are combined with Boolean operators across the different concepts of a research question (adapted from Hammerstrøm (no date))

combined in a search.[9] This shows that search terms that represent the *same* concept are combined with OR, whereas search terms that represent *different* concepts are combined with AND. Applying Boolean operators in an appropriate structure is critical to a well-performing search.

The preceding discussion segues to the logic of combining natural language and controlled vocabulary terms in a search. Usually an electronic database will permit the searcher to 'build' search syntax incrementally, so that each step of the search can be viewed alone, and in combination with other steps. Our search, with the typical sequence of steps, is illustrated below. This emphasizes that each concept is broken down into natural language and controlled vocabulary terms; and these are searched individually before being combined with OR (steps 3, 6, and 10). Once the search terms from the distinct concepts are appropriately fused, they are then combined with each other using AND (steps 7 and 11)—resulting in the final search syntax.

1. Natural language terms for crime types
2. Controlled vocabulary terms for crime types
3. #1 OR #2
4. Natural language terms for prevention outcome
5. Controlled vocabulary terms for prevention outcome
6. #4 OR #5
7. #3 AND #6
8. Natural language terms for evidence synthesis research design
9. Controlled vocabulary terms for evidence synthesis research design
10. #8 OR #9
11. #7 AND #10

---

[9] Note that this illustrative example uses a handful of our search terms; the full search syntax can be found in the Appendix.

Refining search results

Once the search syntax was formulated for each database, we considered how the search results might be refined. Common ways of doing this include restricting the results by time period or by language. We chose to restrict our search to after 1975—the first documentation of systematic review methods (Gough et al. 2013). Further, due to resource constraints, our protocol specified a restriction to studies published in English—a common but risky strategy (as it incurs publication bias, see Brunton et al. 2013).

To manage the numbers of the search results we applied further restrictions. These included the document type (e.g., excluding speeches, meeting papers, teaching materials, book reviews); the precise subject heading (e.g., excluding subjects such as "film studies" and "food science"); or, in the case of larger databases which permit visual analysis of the search results, exact keywords (e.g., in SCOPUS we excluded terms such as "heart arrest" and "prognosis"). For quality assurance, and to protect against subjective bias creeping into the search at this stage, two researchers agreed the removal of every term from the search results. We again tested the results against our list of known studies to check that the removal of terms did not have deleterious consequences.

Lastly, in the databases containing grey literature (see more below) we found few index terms that could be used to restrict the considerable search results. For the theses and dissertations, we reasoned that crime or security must be mentioned at least once in the document, and therefore refined the results with these keywords (using the full text as the search field).

Results of electronic database searches

Table 1 illustrates the 15 databases that were searched in our study, along with, for each, the total number of records retrieved, the number of records duplicated across multiple databases, the number of unique records, the number of relevant studies,[10] and, where available, the approximate number of records in each database at the time of the search. Visual inspection of Table 1 shows that over 15,000 records were identified across the 15 databases, with over a fifth of these being duplicates. The number of records that were returned from each database roughly corresponds with the overall database size—so that large interdisciplinary databases like SCOPUS and Web of Knowledge returned a substantial number of records. However, more modestly sized topic specific databases such as Criminal Justice Abstracts and National Criminal Justice Reference are the exception to this trend, returning many relevant and non-relevant records. It should be noted that PsycINFO was the database used to test the reference data set of studies against decisions made in the search strategy process, and therefore the proportion of relevant hits for this database could be an artefact of this process, rather than being the most fruitful in terms of precision.

The distribution of eligible records across the 13 databases we searched ($n=174$)[11] can be seen in Table 2. The 'total' column is the number of records in the eventual list

---

[10] Relevant studies were those that on first inspection fit our inclusion criteria. They were removed from the final sample of eligible studies if they were found to be duplicates of, or linked to, other studies.

[11] Two further electronic databases—Social Science Full Texts and CINCH—were unavailable to the authors and were searched by information specialists at other institutions.

Table 1 Studies found in individual databases for the survey of systematic evidence in crime prevention

Database	Total	Duplicates	Unique	N relevant	Database records
Applied Social Sciences Index and Abstracts (ASSIA)	585	169	416	22	
Criminal Justice Abstracts (CJA)	1901	582	1319	77	Almost 500,000
Criminal Justice Periodicals (CJP)	815	403	412	28	
CINCH	119	2	117	0	
Education Resources Information Centre (ERIC)	625	167	458	10	Over 1,000,000
International Bibliography of the Social Sciences (IBSS)	430	240	190	12	Over 2,628,800
National Criminal Justice Reference Service (NCJRS)	2719	330	2389	54	Over 200,000
ProQuest theses and dissertations	706	108	598	3	Almost 3,000,000
PsycEXTRA	371	24	347	7	Over 315,000
PsycINFO	765	308	457	84	Over 3,500,000
SCOPUS	3283	648	2635	91	Over 53,000,000
Social Policy and Practice	470	136	334	30	Over 300,000
Social Sciences Full Texts	228	150	78	3	
Sociological Abstracts	1089	162	927	18	Over 955,030
Web of Knowledge	1528	120	1408	32	Over 60,000,000
Total	15,634	3549	12,085		

of included studies, for each database. The 'unique' column refers to the number of studies, for each database, that were not retrieved from other databases. This shows that just under half (84 of 174 studies) were found in a single database, which reinforces that a thorough search cannot rely on one database alone. Table 2 also shows that, for this search strategy at least, there is some redundancy across the databases; for example, three databases (ASSIA, IBSS, and T & D) have no unique records. This suggests that it was unnecessary to search these databases to arrive at the eventual sample of eligible studies. However, with no empirically derived guidelines available this could not be anticipated prospectively. For other research questions, it is likely that the pattern of high and low uniqueness across databases would be different. We now turn to the complementary sources used in our search.

Grey literature sources

Systematic reviews that rely exclusively on electronic databases are likely to miss important studies, thus biasing the overall results of the synthesis. An imperative part of our search strategy was pursuing 'grey' literature. This was difficult: presumably grey literature is also known as 'fugitive literature' (Sechrest et al. 1979) because it is elusive and hard to track down via traditional means. Purposive searching for grey literature is demanding throughout the entire research process—it is more challenging to search for, because structured databases on grey studies are few. It is more resource intensive to retrieve documents (Reed and Baxter 2009), and it can considerably increase the final

Table 2 Database overlap for the studies found in the survey of systematic evidence in crime prevention

Database	Total	Unique	With ASSIA	With CJA	With CJP	With ERIC	With IBSS	With NCJRS	With PsycEXTRA	With PsycINFO	With SCOPUS	With SP&P	With SA	With T & d	With WoK
ASSIA	15	0	N/A	9	4	0	4	4	0	6	11	2	0	0	2
CJA	27	15	0	N/A	2	1	0	0	0	6	9	1	1	0	0
CJP	11	1	1	9	N/A	1	1	2	0	3	8	1	2	0	0
ERIC	5	2	1	1	0	N/A	1	1	0	1	0	0	0	0	0
IBSS	3	0	1	3	0	0	N/A	0	0	1	3	1	0	0	2
NCJRS	30	17	1	9	2	2	0	N/A	1	5	9	5	1	0	2
PsycEXTRA	3	2	0	0	0	1	0	0	N/A	N/A	0	0	0	0	0
PsycINFO	16	14	0	2	1	0	0	1	N/A	N/A	0	0	0	0	1
SCOPUS	32	15	1	4	0	0	0	1	1	3	N/A	2	1	0	7
SP&P	12	12	0	0	0	0	0	0	0	0	0	N/A	0	0	0
SA	7	1	0	5	4	0	3	4	0	3	4	1	N/A	0	1
T & D	3	0	0	0	2	0	0	0	0	0	0	0	3	N/A	0
WoK	10	5	0	0	0	0	0	0	0	1	3	2	0	0	N/A

84 unique studies out of a total of 174 studies

sample of studies for synthesis (Weisburd et al. 1990). Accordingly, the retrieval of
grey literature is now considered a distinct research specialization (Rothstein and
Hopewell 2009) and guidance strongly recommends collaborating with information
specialists (Brunton et al. 2013; Hammerstrøm et al. 2010; Reed and Baxter 2009). To
this end, we commissioned the services of a grey literature information specialist[12] with
extensive criminal justice knowledge to assist in searching for grey literature.

First, we searched for grey literature through the electronic database search. Some of
the databases (in Tables 1 and 2) contained grey literature as well as academic
publications—for example National Criminal Justice Reference Service and Criminal
Justice Abstracts. Others focus exclusively on grey literature, namely, PsycEXTRA,
ProQuest theses and dissertations and Social Policy and Practice. Structured keyword
searches, as outlined above, were used to retrieve grey literature from these sources.

Conference proceedings can also be an invaluable source of grey literature. For
reasons unrelated to research quality, many studies that are presented at conferences are
never subsequently published (Rothstein and Hopewell 2009). Of those that are, they
may be published many months or years later than the conference (Scherer et al. 2007).
Whilst (for large conferences) searching can be a time-consuming business, it can
generate important leads to studies that may not otherwise be found, and is a good way
of horizon-scanning for ongoing studies. The abstracts from the American Society of
Criminology are indexed in Criminal Justice Abstracts and thus were picked up in our
keyword search of that database. These led to one eligible dissertation being identified.

The grey literature information specialist assisted in manual searches (called hand
searches despite these sometimes being conducted electronically) of publications from
25 international policing agencies and practitioner-oriented research organizations (see
Bowers et al. 2014b for more details), along with the large gray (grey) literature
database managed by Rutgers University. We also had help searching the UK National
Police Library catalogue and consulted wider grey literature resources.[13]

Supplementary sources

As previously mentioned, we employed backwards reference checking at the beginning
of the search. This is part of a wider form of lead-generation that can be done in either
direction: backward searches look at the references of eligible studies, forward searches
look at which authors have cited an eligible study in their own references (known as
'citation analysis'). The approach is akin to snowball sampling in primary studies,
whereby one study leads to several others, and each lead is followed until all are
exhausted. It is referred to as 'pearl growing' or 'pearling' (Cooke et al. 2012).

In previous studies, reference checking has been reported as a particularly fruitful
search tactic; for example Papaioannou et al. (2009) found 7 % of their eligible studies
in a social science systematic review through this method. This is because the citations
that authors choose "make substantive or methodological links between studies

---

[12] This specialist also acts as a Trial Search Coordinator for the Campbell Collaboration.

[13] The scholarly organization Grey Literature Network Service (www.greynet.org) has many such resources.
Open Grey (www.opengrey.eu) is a European initiative to provide open access to grey studies published in
Europe. Similarly, WorldCat catalogues many grey literature items. Other guides on information retrieval
helpfully include subject-specific databases and resources that can help synthesists locate relevant grey studies
(Hammerstrøm et al. 2010; Rothstein and Hopewell 2009).

explicit" (White 1994, pp.51–52). Such links are made independent of vocabulary and, accordingly, they remedy the limitations of relying on natural language terms and controlled vocabulary in electronic databases. Therefore they provide connections between studies on a given topic which might not have been obvious through other means (Pao and Worthen 1989).

Citation analysis is time-consuming, and it is wise to devise a means of prioritizing studies to be analyzed in this way. We performed citation analyses in Web of Knowledge on 'review of reviews' method studies, and those that had been initially 'light' coded as integrating multiple systematic reviews. These 35 studies generated 460 leads—395 were already identified, but the remainder mostly consisted of leads to grey literature. Six studies were eventually included in our survey through these means; and several others were linked to included studies (that is, they were the same study written up in a different publication).

During the course of searching, we called upon the so-called 'invisible college' (Cooper 1989) of researchers in crime prevention to help us identify eligible studies. In particular, the lead author contacted scholars via a range of electronic routes to chase 'leads' from conference abstracts and other studies that were mentioned to us by academic and practitioner contacts. The scientific community was especially helpful when sourcing the full texts of potentially relevant studies, and our grey literature information specialist utilized her well-developed network for this task. Researcher networks are fundamental to scientific communication and should be considered a source of studies in their own right—in particular ones that have never been published and are sitting in (electronic or physical) file-drawers. Unpublished studies are inaccessible to any other search tactic, and therefore contacting experts in relevant fields is an important component of any search strategy. Moreover, in the process of screening and coding the studies derived through our search, we noted prolific authors and consulted online lists of their publications. Relatedly, when sourcing book chapters we browsed other chapters in the books in the event they led us to other candidate studies (Reed and Baxter 2009).

Sources of candidate studies

Eventually, once duplicates and linked documents—that is, studies published in multiple places—were removed, 320 candidate studies were eligible according to our inclusion criteria. The sources of these studies are presented in Table 3 according to the chronology in which they were searched (duplicates found as the search progressed are not included). This shows that, as previously mentioned, backwards searches of lists of known systematic reviews in crime prevention was productive in producing a baseline of studies to work with. Electronic database searching was the most lucrative, netting 164 studies (if the dissertation lead from the conference proceedings is included). However, it should be noted that searching and screening the database search results was also the most time-consuming part of the process—it took a team of three full-time researchers approximately 6 months to search and screen all the electronic database hits (including the protracted stage of testing and refining the search terms), with two postgraduate students sourcing several hundred full texts for the second screening phase.

Table 3 Sources of candidate studies

Search tactic	N studies	% of eligible studies
Checking known lists of crime prevention research syntheses (backward reference list checks)	63	19.7
Electronic database searching	163	50.9
UK Police Library Catalogue	1	0.3
Hand searches	71	22.2
Forward reference list checks (citation analysis)	6	1.9
Conference proceedings	1	0.3
Serendipity	15	4.6
Total	320	99.9

Outside of electronic databases, searches for grey literature (the UK Police Library Catalogue and hand searches of practitioner websites), yielded a total of 72 studies; see Table 3. Forward searches based on a selection of eligible studies generated six studies. This is a smaller proportion than reported in searches for primary studies, which may have been influenced by our method of prioritizing studies for citation analysis. Lastly, serendipity led us to identify 15 studies as eligible for our final sample. These were studies found through our scouring of prolific authors' publication lists, skimming additional chapters in identified books, journal alerts and notifications from platforms that share academic publications.

The composition of our final sample in terms of publication types demonstrated an interesting trend. As might be expected, 192 studies, or 60 %, were journal articles, with smaller numbers of book chapters and books (18 and 12 respectively). A sizeable proportion (just over 30 %) of our eligible studies came from grey literature (85 reports and 14 dissertations). This is lower than the 48 % of unpublished sources that Wilson (2009) reported when examining 11 Campbell Collaboration systematic reviews, but we think it still noteworthy given that we were specifically seeking evidence syntheses. This demonstrates the value governments and other non-academic bodies place on evidence syntheses across the spectrum of crime prevention (see Bowers et al. 2014a), and is a trend that is likely to persist.

As a final note, the point made earlier about the lack of reporting guidelines in criminology generally, and crime prevention more specifically, is an important one (Sidebottom and Tilley 2012). This is amplified by the lack of structured abstracts and common vocabulary across the social sciences (Glanville et al. 2008). For these reasons, systematic searchers cannot rely on the key information being in the title alone. Underscoring this point, in our eventual database of evidence syntheses in crime prevention, of the 224 records that were coded as systematic reviews, only 89 had 'systematic review' in the title; of the 176 meta-analyses, only 92 had meta-analysis (or a variant of) in the title. Cross-disciplinary common vocabulary is hence sorely needed for high-quality evidence synthesis, and is a fertile area for future research efforts.

Discussion

Currently, scholars undertaking systematic searches of criminology and crime prevention literature have a limited reference point, for there is a dearth of empirical examples to guide them through the process. In articulating the experiences of applying principles of high-quality information retrieval to the field of crime prevention this paper goes some way to addressing this. In doing so, we have elicited and marshaled tacit knowledge, which can usefully inform the search process for future evidence synthesists. Furthermore, this paper provides results from the first empirical tests of optimizing the balance between sensitivity and precision with the criminological literature. In this discussion we summarize these key findings.

Systematic searching of the literature mirrors the wider research process in that it is not a linear process. A high-quality search will involve a substantial time investment in honing the research question, specifying the precise scope of the work, and trialing and testing of search tactics. Whilst it is good practice to publish the intentions of the search strategy prospectively in a protocol document, it is not uncommon for the search strategy to evolve once the process is underway (Reed and Baxter 2009). Indeed, refinement of the strategy through reflective practice assists in shrewd use of resources.

Time and practical constraints are critical contributory factors in systematic search decisions. While an exhaustive search—finding all relevant studies across all possible sources — is the common ambition, it should be emphasized that to do so would require unlimited resources. Pragmatically, White advises that "the point is not to track down every paper that is somehow related to the topic… the point is to avoid missing a useful paper that lies outside of one's regular purview" (White 1994, p.44). Crucially, it is about minimizing publication bias. As stressed throughout this paper, an optimal search synchronizes high sensitivity with high precision, and thus strikes a balance between identifying all the records of relevance to a research question, managed within the context of available resources. Tactics should therefore be devised with resources in mind, as each source can significantly increase the number of studies that need to be screened and retrieved (prior to the coding and synthesis stage of a review).

Studies on crime prevention topics (and criminology more generally) originate from a broad range of constituent fields. For this reason, systematic searches on topics in this area should employ a myriad of sources so that disciplinary boundaries and the limits of researcher knowledge do not bias the search. In addition, the multidisciplinary nature of crime prevention results in idiosyncratic and scattered vocabulary. Such language is not easily anticipated in searches of electronic databases, and thus care should be taken to generate a wide range of synonyms so that the diversity of expression may be accounted for.

On a related note, high-quality systematic searches recognize the value of using a range of search tactics, and do not rely solely on database searches. Many search tactics exist to mitigate the known problems with publication bias (for example backwards and forwards reference checking, hand searches and consultation of relevant annotated bibliographies). Pursuant to this, in the field of crime prevention it is vital to search widely for grey literature; that is, research findings that are published outside of academic outlets. It is worth acknowledging that grey literature retrieval is a specialist skill, which few academics have. Accordingly, we affirm the recommendation in the information retrieval guidance to work closely with information specialists in the field of interest (Hammerstrøm et al. 2010).

Our survey of crime prevention evidence syntheses cumulated a sample with just under a third of the studies being grey literature – a finding that is consistent with previous empirical enquiry (Wilson 2009). This is, however, an under-researched area and the quantification of grey literature in the criminology field is an important research agenda. For example, it would be useful to know what proportion of conference abstracts make it to final publication, the time it takes from conference to publication and the reasons for not publishing (i.e., weak or null results). This would facilitate an empirically based estimate of the impact of publication bias in the field.

In a related vein, methods to detect inflated estimates of effect in primary studies are now being applied in allied fields such as psychology (Francis et al. 2014). These tests are performed to estimate the probability that appropriately conducted experiments will produce as many 'successful' outcomes as reported in the literature. Quantification of potential bias in primary study evidence is an exciting development, and one the criminological community could apply to estimate the magnitude of published biased findings.

Lastly, it is important to recognize that each research question will require its own tailored approach. The focal topic of the systematic search will dictate the selection of the sources, the keywords and controlled vocabulary used in electronic database searching, and complementary search tactics. The overarching principle is to search widely, meticulously and judiciously so that as many studies are found within the scope of the available resources. Akin to research more generally, data collection is rarely perfect, but adopting these precepts can protect against biased results.

Appendix

Search syntax applied in PsycINFO database through the Ovid interface on December 17, 2013.

1. (abduct* or aggressi* or ("anti*social adj2 behavio*") or ("anti?social" adj2 "behavio*")) or arrest* or arson* or assault* or blackmail or "bodily harm" or burglar* or "calls for service" or convict* or counterfeit* or crim* or delinquen* or ((dr*nk or dangerous) adj2 driv*) or DUI or explosi* or firearm or fraud or homicide or incarcerat* or incest or infanticide or kidnapping or (knife adj2 crim*) or (law adj2 breaking) or manslaughter or (money adj2 laundering) or murder or offen* or prostitute* or "public disorder" or rape or recidivis* or reconvict* or re-offen* or reoffen* or riot* or robber* or shoot* or shoplift* or terrori* or theft* or unlawful or vandalism or violen* or weapon or wounding or (drug adj2 misuse) or (drug adj2 abuse) or (drug adj2 market) or (drug adj2 deal*) or (drug adj2 traffick*) or (drug adj2 supply) or (drug adj2 possess*)).ab,ti.

2. exp Crime/ or criminal behavior/ or antisocial behavior/ or juvenile delinquency/ or criminals/ or gangs/ or criminals/ or perpetrators/ or female criminals/ or male criminals/ or mentally ill offenders/ or crime/ or antisocial behavior/ or criminal behavior/ or juvenile delinquency/ or recidivism/ or sexual abuse/ or terrorism/ or violence/

3. #1 OR #2

4. (prevent* or deterr* or diver* or reduc* or control* or increas* or impact or (cost adj1 effective*) or (cost adj1 benefit) or displace* or (diffus* adj2 benefit)).ab,ti.

5. exp Crime Prevention/ or exp criminal rehabilitation/ or "costs and cost analysis"/ or resource allocation/ or recidivism/ or spatial frequency/ or temporal frequency/

6. #4 OR #5

7. #3 AND #6

8. (meta analy* OR meta?analy* OR meta*analy* OR (review adj5 evidence) OR (review adj5 literature) OR (review adj5 evaluation) OR (systematic* adj2 review) OR "comprehensive bibliography" OR "Integrat* Review" OR "comprehensive Review" OR (quantitative adj2 review) OR (qualitative adj2 review) OR (thematic adj2 analy*) OR meta*narrative OR (data adj2 extract*) OR (data adj2 integrat*) OR (narrative adj2 review) OR (realist adj2 review) OR (campbell adj3 review) OR (cochrane adj3 review)).ab,ti.

9. exp meta analysis/ or exp response to intervention/ or exp evidence based practice/

10. #8 OR #9

11. #7 AND #10

12. Limit #11 to (human and English language and year="1975–2014")

13. Limit #12 to (abstract collection or bibliography or chapter or dissertation or encyclopedia entry or journal article or publication information or reviews)

14. Limit #13 to (not "0100").md.

15. Limit #14 to (3236 criminal behavior & juvenile delinquency or 3380 rehabilitation or 3386 criminal rehabilitation & penology or 4210 civil rights & civil law or 4230 criminal law & adjudication or 4250 mediation & conflict resolution or 4270 crime prevention or 4290 police & legal personnel)

References

Alderson, P., & Green, S. (Eds) (2002). Cochrane collaboration open learning material for reviewers: Cochrane collaboration. Available at: www.cochrane-net.org/openlearning/. Accessed 15 Aug 2010.

Auger, C. P. (1998). *Information sources in Grey literature* (4th ed.). London: Bowker-Saur.

Boeker, M., Vach, W., & Motschall, E. (2013). Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC Medical Research Methodology, 13*(1), 131.

Bowers, K. J., & Tompson, L. (2014). Mapping the crime reduction evidence base: A descriptive analysis of the WP1 Systematic Review Database. DOI: 10.14324/000.wp.1462099. Available at: http://discovery. ucl.ac.uk/1462099/1/Work%20Package%201%20Research%20Map.pdf. Accessed 8 Jul 2015.

Bowers, K. J., Tompson, L., & Johnson, S. D. (2014a). Implementing information science in policing: mapping the evidence base. *Policing: A Journal of Policy and Practice., 8*(4), 339–352.

Bowers, K. J., Johnson, S. D., Tilley, N., Tompson, L., & Belur, J. (2014b). Protocol for work package 1: Identifying existing systematic reviews of crime reduction. Available at: http://discovery.ucl.ac.uk/ 1462098/1/What%20works%20Work%20Package%201Protocol%20May2014.pdf. Accessed 8 Jul 2015.

Brunton, G., Stansfield, C., & Thomas, J. (2013). Finding relevant studies. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (pp. 107–134). London: Sage Publications Ltd.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research, 22*(10), 1435–1443.

Cooper, H. C. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Beverley Hills: Sage.

DeLuca, J. B., Mullins, M. M., Lyles, C. M., et al. (2008). Developing a comprehensive search strategy for evidence based systematic reviews. *Evidence Based Library and Information Practice, 3*(1), 3–32.

Dickersin, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 11–34). Chichester: Wiley.

Farrington, D. P., Ohlin, L. E., & Wilson, J. Q. (1986). *Understanding and controlling crime*. New York: Springer.

Francis, G., Tanzman, J., & Matthews, W. J. (2014). Excess success for psychology articles in the journal science. *PLoS ONE, 9*(12). doi:10.1371/journal.pone.0114255

Glanville, J., et al. (2008). So many filters, so little time: the development of a search filter appraisal checklist. *Journal of the Medical Library Association: JMLA, 96*(4), 356–361.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. London: Sage.

Gough, D., Oliver, S., & Thomas, J. (Eds.). (2013). *An introduction to systematic reviews*. London: Sage Publications Ltd.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82*, 1–20.

Hammerstrøm, K. (no date). Searching for studies - the Campbell Collaboration. Available at: https://www. youtube.com/watch?v=2xX-Hlv1V3k. Accessed 6 Jan 2015.

Hammerstrøm, K., Wade, A. & Klint, A. (2010). Searching for studies: A guide to information retrieval for Campbell Systematic Reviews. Campbell Collaboration.

Johnson, S. D., Tilley, N., & Bowers, K. J. (2015). Introducing EMMIE: an evidence rating scale to encourage mixed-method crime prevention synthesis reviews. *Journal of Experimental Criminology.* doi:10.1007/ s11292-015-9238-7.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational and behavioral treatment: confirmation from meta-analysis. *American Psychologist, 48*, 1181–1209.

Pao, M., & Worthen, D. (1989). Retrieval effectiveness by semantic and citation searching. *Journal of the American Society for Information Science, 40*(4), 226–235.

Papaioannou, D., Sutton, A., Carroll, C., et al. (2009). Literature searching for social science reviews: consideration of a range of search techniques. *Health Information and Libraries Journal, 27*, 114–122.

Petrosino, A. J. (1995). The hunt for randomized experimental reports: document search and efforts for a "what works?" meta-analysis. *Journal of Crime and Justice, 18*(2), 63–80.

Petrosino, A. J. (1997). *"What works?" revisited again: A meta-analysis of randomized field experiments in individual-level interventions*. Newark: Rutgers University.

Petrosino, A., Boruch, R. F., Soydan, H., Duggan, L., & Sanchez-Meca, J. (2001). Meeting the challenges of evidence-based policy: the campbell collaboration. *The ANNALS of the American Academy of Political and Social Science., 578*(14), 14–34.

Reed, J. G., & Baxter, P. M. (2009). Using reference databases. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (pp. 74–100). New York: Russell Sage.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641.

Rothstein, H. R., & Hopewell, S. (2009). Grey literature. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (pp. 103–126). New York: Russell Sage.

Scherer, R. W., Langenberg, L., & von Elm, E. (2007). Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews*. (2):MR000005.

Sechrest, L., White, S. O., & Brown, E. D. (1979). *The rehabilitation of criminal offenders: Problems and prospects*. Washington, DC: National Academy of Sciences.

Sidebottom, A., & Tilley, N. (2012). Further improving reporting in crime and justice: an addendum to Perry, Weisburd and Hewitt. *Journal of Experimental Criminology, 8*(1), 49–69.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4*, 22–24.

Weisburd, D., & Farrington, D. (Eds) (2015). *Systematic reviews in crime prevention*. Springer. In press

Weisburd, D. L., Sherman, L., & Petrosino, A. (1990). *Registry of randomized experiments in criminal sanctions, 1950–1983*. Los Altos: Sociometrics Corporation, Data Holdings of the National Institute of Justice.

Wells, E. (2009). Uses of meta-analysis in criminal justice research: a quantitative review. *Justice Quarterly, 26*(2), 268–294.

White, H. D. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 41–56). New York: Russell Sage.

White, H. (2009). Scientific communication and literature retrieval. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis* (pp. 51–72). New York: Russell Sage.

Wilson, D. B. (2009). Missing a critical piece of the pie: simple document search strategies inadequate for systematic reviews. *Journal of Experimental Criminology, 5*(4), 429–440.

Lisa Tompson is a Senior Research Associate at the Department of Security and Crime Science at University College London. Her research interests focus on the use of spatio-temporal methods in crime analysis and crime prevention and systematic review methodology.

Jyoti Belur is a Senior Research Associate at the Department of Security and Crime Science at University College London and a former senior police officer in the Indian Police. Her research interests lie in policing theory and practice, organised crime, qualitative research methodology and systematic reviews.