

警务工作评估 系统回顾研究

主编◎刘建宏

Policing: a Systematic Review of
Intervention Programs

 人民出版社

学术顾问◎付子堂

康拜尔合作组织刑事司法研究报告系列

警务工作评估 系统回顾研究

主编◎刘建宏

Policing: a Systematic Review of
Intervention Programs

责任编辑:张立
版式设计:石笑梦
责任校对:陈艳华

图书在版编目(CIP)数据

警务工作评估系统回顾研究/刘建宏 主编. —北京:人民出版社,2016.8
ISBN 978-7-01-016462-5

I. ①警… II. ①刘… III. ①公安工作-研究 IV. ①D035.3

中国版本图书馆 CIP 数据核字(2016)第 168071 号

警务工作评估系统回顾研究

JINGWU GONGZUO PINGGU XITONG HUIGU YANJIU

刘建宏 主编

人民出版社 出版发行
(100706 北京市东城区隆福寺街 99 号)

北京汇林印务有限公司印刷 新华书店经销

2016 年 8 月第 1 版 2016 年 8 月北京第 1 次印刷

开本:787 毫米×1092 毫米 1/16 印张:23

字数:390 千字

ISBN 978-7-01-016462-5 定价:62.00 元

邮购地址 100706 北京市东城区隆福寺街 99 号
人民东方图书销售中心 电话 (010)65250042 65289539

版权所有·侵权必究

凡购买本社图书,如有印制质量问题,我社负责调换。

服务电话:(010)65250042

序言一(英文版)

David P.Farrington

What works to reduce crime? How should offenders be dealt with so that they do not reoffend? What methods of preventing crime are most cost-effective? These are all questions to which citizens, as well as government officials, policy makers, practitioners, researchers, teachers and the news media deserve good answers. All such persons should have ready access to the most rigorous and up-to-date evidence on the effectiveness of interventions designed to reduce crime and offending. The best evidence on what works should be quickly accessible to those who need it.

Much practice in crime and justice, as in fields like medicine and education, is based on long-term traditions and clinical experience. Although tradition and experience often provide the only guidance for criminal justice practitioners, there is a growing consensus among scholars, practitioners, and policy makers that crime control practices and policies should be grounded as much as possible in the results of scientific research. Support for evidence-

based policy in criminal justice may be seen as part of a more general trend toward the use of scientific research for establishing rational and effective practices and policies in many fields. This trend is perhaps most prominent in the health professions where the idea of evidence-based medicine has gained strong government and professional support.

A central component of the movement toward evidence-based practice and policy is the reliance on systematic reviews of prior research and evaluation studies. The Campbell Collaboration Crime and Justice Group is an international network that aims to produce and make accessible the best evidence on what works in crime and justice. This network of scholars, policy makers, practitioners and others from around the world is preparing rigorous systematic reviews of high-quality research on the effects of criminological interventions.

These systematic reviews are being maintained and updated in light of new studies, insightful criticisms, or new methodological developments. They are readily accessible on the Internet; see www.campbellcollaboration.org. Through international collaboration, the Campbell Crime and Justice Group (CCJG) aims to ensure that relevant evaluation studies conducted all over the world will be taken into account in its systematic reviews, and that the evidence from such reviews will be made accessible globally through language translation and worldwide dissemination.

Characteristics of Systematic Reviews

What are systematic reviews? These are reviews that use rigorous methods for locating, appraising and synthesizing evidence from prior evaluation studies. They contain methods and results sections, and are reported with the same level of detail that characterizes high quality reports of original research. Other features of systematic reviews include:

1. *Explicit Objectives*. The rationale for conducting the review is made clear.
2. *Explicit eligibility criteria*. The reviewers specify in detail why they included certain studies and rejected others. What was the minimum level of methodological quality for inclusion in the review? Did they consider only a particular type of evaluation design such as randomized experiments? Did the studies have to include a certain

type of participant such as children or adults? What types of interventions were included? What kinds of outcome data had to be reported in the studies? All criteria or rules used in selecting eligible studies are explicitly stated in the final report.

3. *The search for studies is designed to reduce potential bias.* There are many potential ways in which bias can compromise the results of a review. The reviewers must explicitly state how they conducted their search of potential studies to reduce such bias. How did they try to locate studies reported outside scientific journals? How did they try to locate studies in foreign languages? All bibliographic data bases that were searched should be made explicit so that potential gaps in coverage can be identified (and reviews can be replicated).
4. *Each study is screened according to eligibility criteria, with exclusions justified.* The searches always locate many citations and abstracts to potentially relevant studies. Each of the reports of these potentially relevant studies must be screened to determine if it meets the eligibility criteria for the review. A full listing of all excluded studies and the justifications for exclusion should be made available to readers.
5. *Assembly of the most complete data possible.* The systematic reviewer will generally try to obtain all relevant evaluations meeting the eligibility criteria. In addition, all data relevant to the objectives of the review should be carefully extracted from each eligible report and coded and computerized. Sometimes, original study documents lack important information. When possible, the systematic reviewer will attempt to obtain this from the authors of the original report.
6. *Quantitative techniques are used, when appropriate and possible, in analyzing results.* Although there is still some confusion about the meaning of these terms, it is useful to distinguish between a systematic review and a meta-analysis. A meta-analysis involves the statistical or quantitative analysis of the results of prior research studies. Since it involves the statistical summary of effect sizes and their correlates, it requires a reasonable number of intervention studies that are sufficiently similar to be grouped together. For example, there may be little point in reporting a weighted mean effect size based on a very small number of studies. Nevertheless, quantitative methods can be very important in helping the reviewer determine the average effect size of a particular intervention and in what circumstances and with what types of

people it works best.

A systematic review may or may not include a meta-analysis. For example, a reviewer may only find a few studies meeting the eligibility criteria. Those studies may differ just enough in the operational definition of the intervention or in the way they were conducted (etc.) to make formal meta-analysis inappropriate and potentially misleading. It is important not to combine apples and oranges in calculating a weighted mean effect size.

7. *Structured and detailed report.* The final report of a systematic review is structured and detailed so that the reader can understand each phase of the research, the decisions that were made, and the conclusions that were reached. In principle, it should be possible for an independent scholar to replicate both the review and the results.

The Campbell Collaboration

At a meeting in Philadelphia attended by over 80 persons from 12 different countries, the Campbell Collaboration was inaugurated in February 2000, to prepare, maintain and make accessible systematic reviews of research on the effects of social, educational, and criminological interventions. At that February 2000 meeting, the Campbell Collaboration established a Crime and Justice Group (CCJG) and a Steering Committee to coordinate the work of this Group. The original Steering Committee consisted of 10 persons from 7 different countries. I was appointed as the first Chair, and I managed to secure some funding from the British Home Office to support a part-time coordinator (Anthony Petrosino) for three years. The first meeting was held in Paris in April 2000, coinciding with a meeting of the Board of Directors of the International Society of Criminology (ISC). This was partly to emphasize the international remit of the CCJG and partly because four members of the Steering Committee were on the ISC ' s Board of Directors. It was agreed that the institutional home of the CCJG would be at the University of Pennsylvania, supported by Lawrence Sherman, who was the President of the ISC at the time.

At the first meeting, it was decided to commission reviews on 15 key topics such as the effectiveness of boot camps, child skills training, CCTV, neighborhood watch, and hot spots

policing. The aim was to select narrowly defined topics where there was likely to be only a small number(e.g.20–50) of high quality evaluations, nevertheless covering a wide range of criminological interests in total. Instead of waiting for researchers to propose topics, the CCJG proactively approached well-known scholars to do the first few reviews. This ensured that key topics were covered, that results could be speedily obtained, and that a good reputation was established quickly. Those who undertook systematic reviews were asked to agree to the following requirements:

1. A commitment to undergo a rigorous editorial review process not only from researchers but also policy makers, practitioners, and citizens to ensure that the review meets high scientific standards and is also written to be understandable to non-academic audiences.
2. A commitment to maintain transparent and open review processes so that users can comment and criticize each stage of the review, from its proposal through to its completion.
3. A commitment to use the most rigorous search methods available to ensure that all relevant studies are considered for inclusion or exclusion and not just those reported in easily accessible journals and books.
4. A commitment to cover literature from around the world and not just the English-speaking world.
5. A commitment to code and computerize key features of each evaluation study reviewed(so that anyone accessing the review can organize the studies according to such features as sample size, design, or effect size).
6. A commitment to explicitly report the final review so that readers can understand decisions made at each stage, justifications for those decisions, and how conclusions were reached.
7. A commitment to make the review available to broader audiences than readerships of peer-reviewed academic journals through electronic publication and dissemination into policy, practice, and media outlets.
8. A commitment to update the systematic review to incorporate new evidence, respond to criticisms, or use more advanced methods, on a regular basis.

Doing a Systematic Review

The first step in conducting a systematic review for the Campbell Collaboration Crime and Justice Group is to submit a proposed title to the Managing Editor, who is currently Charlotte Gill. This is refereed and, if accepted, is registered in the Campbell Collaboration Library of Systematic Reviews. The main reason for not accepting a proposed title would be overlap or duplication with an existing title. The second step is to complete and submit a draft protocol (a detailed description of how the systematic review will be completed). This should include the background to the review (hypotheses tested, operational definitions of interventions and outcome variables), the objectives of the review, strategies for searching the literature, selection criteria for including or excluding studies, and strategies for data extraction, coding, and analysis. This is also refereed and revised in the light of the referees' comments, before it is published on the Campbell website. David Wilson currently acts as Editor-in-Chief of the CCJG refereeing process.

The third step is to complete the systematic review. This is also refereed (by the Campbell Methods group as well as by criminologists) and again is likely to require revisions before it is published in the Campbell Library. These rigorous refereeing processes are designed to ensure that the published reviews are of high quality. Once a review is published on the Campbell website, anyone can post comments on it, and authors are encouraged to update the review every three years or so. Campbell Collaboration policy is that each review should have at least two authors, in order to facilitate tests of the reliability of inclusion/exclusion decisions and coding of key features of evaluation studies.

The information that is extracted and coded from each included study should comprise at least the following: principal investigators, full citations to all evaluation reports, funding, publication dates, design of the study, characteristics of experimental units (for example, age and gender of participants, prior crime rates of areas), sample size, hypotheses tested, interventions, implementation details, how extraneous variables were controlled so that it was possible to disentangle the impact of the intervention, who were the

program delivery personnel, what were the control conditions (since it is rarely possible to have a truly untreated control group), who knew what about the intervention (since double-blind trials are desirable), measurement of outcome variables (for example, official records and/or self-reports of crime), before and after measures of offending, length of follow-up period, and measures of effect size and variability of effect size. Authors of reviews are encouraged to code all variables independently, so that the reliability of coding can be assessed.

Decisions about what studies to include in a systematic review can be highly controversial, because they involve assessments of methodological quality. Authors of excluded studies may feel that their research has been negatively assessed or even “cast into outer darkness” (as one person has told me). I hope that Campbell Collaboration reviews will lead to an improvement in the quality of the primary evaluation research. In general, randomized experiments have the greatest internal validity. However, randomized experiments that evaluate criminological interventions are relatively uncommon. If Campbell Collaboration reviews were restricted to randomized experiments, they would be relevant to only a small fraction of the key questions for policy and practice in criminology. Therefore, for topics where there are few or no randomized experiments, reviewers also select high quality quasi-experimental evaluations for inclusion. The aim is to reach the most defensible conclusions based on the best available research.

Campbell Reviews

These five volumes present the conclusions of 36 reviews completed for the Campbell Crime and Justice Group. The five volumes cover policing, intervention and prevention, juvenile delinquency, corrections, and drugs. In general, the reviews show that many types of criminological interventions are effective.

In the area of policing, for example, hot spots policing (policing interventions targeting very small areas) is notably effective in reducing crime and disorder. Problem-oriented policing is similarly effective. Also, “pulling levers” focused deterrence strategies, that involve communicating costs and benefits to targeted offenders, are effective in reducing crime, and

crackdowns on gun carrying are effective in reducing gun crime. DNA testing is generally effective in increasing police clearance rates. However, there is insufficient evidence to draw conclusions about the effectiveness of counter-terrorism strategies.

There is more good news in the area of intervention and prevention. In general, improved street lighting is followed by a decrease in crime, and closed-circuit television is also effective in reducing crimes in certain settings (e.g. in car parks). Neighborhood watch is also effective in reducing crime. Anti-bullying programs in schools are also generally effective, although there are too few evaluations of interventions to prevent cyber bullying on the internet to draw firm conclusions about these. Generally, court-mandated interventions for individuals convicted of domestic violence are effective in reducing repeat violence according to official records, but there were too few studies of interventions to reduce cross-border trafficking to draw conclusions about these.

Results are more mixed in the area of juvenile delinquency. Early family and parent training programs are generally effective in reducing antisocial behavior and delinquency, and the same is true of mentoring and self-control programs. However, parental imprisonment, formal system processing of juveniles, and the “scared straight” program all have undesirable effects. It is important to know what does not work as well as what works.

There are again desirable and undesirable results in the area of corrections. Cognitive-behavioral programs for criminal offenders are generally effective, as are cognitive-behavioral programs for serious, violent and chronic juvenile offenders. Non-custodial sentences are generally more effective than custodial sentences in reducing recidivism. However, correctional boot camps and non-custodial employment programs are not effective in reducing recidivism, and there are too few rigorous cost-benefit analyses of sentencing to draw firm conclusions.

In the area of drugs, it is clear that many types of interventions are effective. Drug courts are particularly effective, as are incarceration-based drug treatment and drug substitution programs. Also, problem-oriented policing and community-wide policing approaches are effective in disrupting street-level drug markets and reducing drug use.

Conclusion

These five volumes provide the best available information about what works and what does not work in reducing crime. They should form the basis of wide-ranging coordinated evidence-based strategies for crime prevention and crime control.

序言一(中文版)

大卫·法林顿

怎样才能减少犯罪？如何处理罪犯，才能让他们不再犯？最有效的预防犯罪的方法是什么？这些问题都需要好好解答。对于普通民众、政府官员、政策制定者、实践工作者、教师学者以及新闻媒体来说，都需要有了解最严谨和先进的用以评估减少犯罪的干预项目有效性证据的途径，并且是能快速获得这些证据的途径。

同医药和教育界一样，很多犯罪和司法的实践都是以良久传统和临床经验为基础的。虽然传统和经验常常只用于指导刑事司法实践工作者，但是越来越多的学者、实践者以及政策制定者意识到犯罪控制的实践与政策也需要尽可能的科学研究结论的指导。支持基于实证的刑事司法政策，其实也是许多领域里运用科学研究来建立理性的行之有效的实践与政策大趋势的部分体现。这种趋势也许在医疗健康领域尤为突出，因为基于证据的医药研究已经获得了政府以及专业领域的大力支持。

对以往的调查以及效果评估研究的系统性回顾是这场基于证据的实践与政策制定的核心。康拜尔合作组织的犯罪与刑事司法小组是一个致力于评估犯罪预防以及

刑事司法政策有效性的国际性组织。具体来说,是由世界各地的相关学者、政策制定者、实践工作者以及其他专业人士对那些关于犯罪干预有效性的高质量研究进行严谨的系统性回顾。

这些系统性回顾基于新出现的研究、锐评,以及新方法的发展而及时改进和更新,并且在互联网上可以浏览网址 www.campbellcollaboration.org。通过国际合作,康拜尔合作组织犯罪与刑事司法小组(CCJG)力求能对那些来自世界各地的相关评估研究进行系统性回顾,而得出的相关证据通过翻译和全球宣传可被广泛运用。

系统性回顾的特征

系统性回顾是什么?所谓系统性回顾就是运用严谨的方法来定位、评估以及整合那些从先前评估研究中获得的证据。其包含了方法和结果两部分,并且以与原来研究相同的细节层次报告出来,而这些细节也正是原研究高质量的体现。系统性回顾的其他特征还包括:

1. 目标明确:评估的基本原理清晰。
2. 甄选标准明确:审阅者对遴选的研究报告有着明确的标准。入选的研究在其研究方法的质量上的最低要求是什么?他们只考虑了诸如随机试验一类的特定类型的评估研究吗?相应的研究必须包含特定的参与者,如儿童或者成年人?包含了哪些干预的类型?哪些数据结果需要报告?所有用于甄选研究的标准都会在最后的报告中清晰呈现出来。
3. 搜寻相关研究时须减少可能的偏倚。很多潜在的因素都会让偏倚影响评估的结果。审阅者须明晰地陈述其在搜寻研究时如何减少偏倚。他们是如何定位发表在领域外期刊上的研究报告的?他们是如何定位非英语研究报告的?所有用于遴选的文献数据库都须明确,这样潜在的偏差才能被甄别出来(回顾才具有可重复操作性)。
4. 每个研究入选或者排除都要依据相应的标准。文章的搜寻往往是通过定位引文和摘要来找到可能相关研究。我们需要对这些相关研究的报告进行筛选,判断其是否符合评估标准。而对于那些落选的研究,我们要给读者提供一份完整的清单,并说明落选原因。
5. 尽可能整合最完整的数据。通常系统性回顾者都会尝试获得所有符合要求的

评估研究,然后从符合标准的报告里提取所有与评估目标相关的数据,再进行编码和电脑处理。有时候,原研究报告会缺失某些重要信息。可能的话,评估者会尝试从原文作者处获得这些缺失的信息。

6. 适当的时候,定量分析方法也会用于结果分析。虽然系统回顾和元分析这两个概念的界定仍然有点模糊不清,但是将二者区分开来是有益处的。元分析方法包括了对前人研究的结果进行统计或者量化分析。由于牵涉了效应量大小和相关关系的统计汇总,所以对干预研究的数量有要求,而且这些研究要有足够的相似之处才能被整合。例如,如果找到的研究很少,那么在报告其加权平均效应量大小的时候可能就没有相应的点。然而,量化研究在评估者判断某个干预分析的平均效应量大小和在何种情况下对何种人最有效的问题上是大有裨益的。

一次系统性回顾可能包含了元分析,也可能不包含。例如,评估者只找到很少量符合要求的研究。这些研究可能在操作定义上或者操作过程等环节上有差异,而这些差异正好使得元分析不适用,或者产生误导。我们当然不该把风马牛不相及的东西整合起来,然后算出它们的加权平均效应量的大小。

7. 报告要层次分明而又详尽。最终评估报告需要有分明的层次和详尽的细节,这样读者才能了解研究的各个阶段,才能明白所做的决定和最终达成的结论。原则上,其他独立学者应该能重复该评估的操作,并得到相同的结果。

康拜尔合作组织

2000年2月,在费城一个由来自12个国家的80多名与会者参加的会议上,康拜尔合作组织诞生了,其目的是致力于对那些与社会、教育以及犯罪有关的干预研究效果的系统性回顾。在当年的会议上,康拜尔合作组织还成立了犯罪与刑事司法小组(CCJG)以及协调相关工作的指导委员会。指导委员会最初由来自7个国家的10人组成。我当时被任命为第一主席,并成功获得了英国内政部的资金支持,用以支付一位临时协调员(Anthony Petrosino)三年的工资。康拜尔合作组织第一次会议于2000年4月在巴黎召开,是与国际犯罪学学会(ISC)的董事会会议同时召开的。造成此情况的原因有两个,一方面是为了强调CCJG的国际性,另一方面是由于有4名指导委员会委员同时也是ISC的董事会成员。CCJG机构的大本营设立在宾夕法尼亚大

学,负责人是当时的 ISC 主席 Lawrence Sherman。

第一次会议上,我们决定就 15 个关键话题进行评估,如军事训练营、儿童技能训练、闭路电视、邻里监督和热点警务的有效性。这样是为了在明确评估对象的同时(虽然相关高质量的评估研究可能会较少,如只有 20—50 篇),又不失犯罪学研究范围上的广度。CCJG 并没有等着研究者来挑选主题,而是主动找到名声显赫的学者让其做前期少量的评估。这样在保证关键话题都有人回顾的同时,还能较快得到结论,与此同时也能迅速将招牌打响。康拜尔合作组织对系统回顾者提出了以下几点要求:

1. 不仅是研究者,包括政策制定者、实践工作者以及普通民众都需要对所有评估进行严谨的审核,这样才能确保评估不但符合高的科学标准,而且对于非专业人士来说也能通俗易懂。
2. 评估的所有步骤都要有所体现,这样读者才能针对各环节提出修改建议。
3. 所用的搜索研究方法必须是最严谨的,才能确保所有相关研究都被考虑到,而不只是那些在容易找到的杂志和书籍上的研究报告。
4. 搜索时不能局限于英文文献,要全面。
5. 要对评估研究里的所有关键特征进行编码和电脑处理(任何读到该回顾的人都能根据样本大小、研究设计或者效应量大小等特征对回顾有整体认识)。
6. 最终的回顾报告要条理清晰,这样读者才能理解不同阶段所做的决定、每个决定的理由,以及结论是如何获得的。
7. 通过电子刊物、政策和实践宣传,或者传媒帮助,让目标读者的范围更广一些,而不只局限于学术期刊里同侪审阅的那些人。
8. 要通过整合新证据,反馈评论,或者运用新方法对这些系统回顾进行定期更新。

如何进行系统回顾

康拜尔合作组织犯罪与刑事司法小组系统回顾的第一个步骤是向执行编辑提交拟题(现任执行编辑是 Charlotte Gill)。题目在审核后,如果予以采纳,就会被注册于康拜尔合作组织系统回顾图书馆。那些被否定了的拟题,最主要是因为它们和现存的题目有所交叉或者重复。

第二步是完成和提交草案(如何完成系统回顾的详尽描述)。其中包括了背景

介绍(假设检验、干预的操作定义以及结果变量)、评估目标、文献搜索策略,文章入选或落选的标准以及数据提取、编码和分析的策略。这个过程也需审核,在出版于康拜尔合作组织网站前还要根据评审的建议进行校订。当前是由 David Wilson 担任 CCJG 该步骤审核主编。

第三步就是完成系统回顾。这个步骤也需要审核(由康拜尔方法小组和犯罪学家负责),在出版前也需要相应的校稿。严谨的审阅是为了确保系统回顾的质量。回顾一旦发表,任何人都可以在康拜尔的网站对其做评论。康拜尔鼓励作者对所作回顾进行定期更新,周期为3年左右。康拜尔合作组织还规定每篇回顾至少要有两个作者,这样既方便检验研究入选/落选决定是否可靠,又能提高编码评估研究的关键特征时的效率。

每篇入选的研究里摘要和编码的信息至少应该包含以下内容:主要研究者,完整的引文情况,基金信息,出版日期,研究设计,实验单位的特征(如参与者的年龄与性别,区域里犯罪率的历史记录),样本量大小,假设检验,干预手段,操作细节,如何控制外扰变量以减小外界干扰,谁是该项目的执行人员,其他的控制条件是什么(因为很难做到完全排除对控制组的影响),何人了解干预项目的哪些情况(当然双盲检验最为理想),结果变量的测量(如官方记录和/或犯罪自我报告),犯罪行为的前测与后测,追踪研究的历时,以及效应量的测量与差异。回顾的作者应当独立编码所有变量,这样我们才能评估编码的可靠性。

选择什么样的研究进行系统回顾是极具争议性的,因为这里涉及了如何评价搜索方法质量的高低。那些落选的作者可能会觉得自己的研究被低估了,甚至认为是“被排斥”了(曾经一个作者对我如是说)。我希望康拜尔合作组织能够引导评估研究质量的提高。一般来说,随机试验最具有内部效度。然而,用于评估犯罪干预的随机试验相对来说并不常见。若康拜尔系统回顾只局限于那些随机试验,那么其所能涉及的犯罪学方面的政策和实践也会很受局限。因此,对于那些只有少量或者没有随机试验的研究主题,评估者也可以从中挑出质量高的类实验研究。如此一来,根据这样的搜索策略所得的研究结论才最有说服力。

关于康拜尔系统回顾丛书

您手中的这套丛书(共五册)将呈现康拜尔合作组织犯罪与刑事司法小组的 38