

An Introduction to Evidence-Based Practices

April 2014



Table of Contents

Acknowledgments.....	1
A Brief History of the Evidence-Based “Movement”	2
Evidence-Based Medicine.....	2
The Evidence Based Practices Movement in Criminal Justice	3
Where Does Evidence Come From?.....	5
What is Effectiveness?	5
What are Scientific Methods?	6
Randomized Controlled Trials	6
Quasi-Experiments and Non-Experiments.....	7
What is not Scientific Evidence?	8
How Much Evidence is Enough?.....	8
Systematic Review and Meta-Analysis.....	9
Summary	10
Resources for Identifying EBPs.....	10
Implementing EBPs	12
Adapting EBPs for Local Use	13
What if there is No Evidence?	13
Summary.....	15
References Cited	16
Appendix: Evidence-Based Practices Resources.....	18

Acknowledgments

This briefing was prepared by Stan Orchowsky, Ph.D., Research Director for the Justice Research and Statistics Association. We wish to thank Tammy Woodhams and our partners at the National Criminal Justice Association, and Lesley Buchan and Ed Banks at the Bureau of Justice Assistance.

This project was supported by Award No. 2010-DJ-BX-K176 awarded by the Bureau of Justice Assistance, Office of Justice Programs, U.S. Department of Justice. The opinions, findings and conclusions or recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Justice.

The movement toward the use of evidence-based practices (EBPs) has been sweeping the criminal justice community in recent years. The purpose of this briefing paper is to provide policymakers with an introduction and overview of the key concepts and issues associated with the identification and use of EBPs in criminal justice. The briefing provides a brief history of the evidence-based movement, discusses what is meant by evidence and where evidence comes from, identifies sources for information on EBPs, discusses issues associated with implementing EBPs, and addresses the question of what to do when there is no evidence for a particular program or practice.

A Brief History of the Evidence-Based “Movement”

Evidence-Based Medicine

Today’s evidence-based movement has its origins in the field of medicine, where an initial interest in the *safety* of treatment was eventually joined by an equal interest in the *efficacy* of treatment. Beginning in the mid-1800s, parallel trends involving the increased use of scientific methods, statistical analysis, and discoveries from the natural sciences increased interest in distinguishing between effective and ineffective medical treatments based on patient outcomes (Office of Technology Assessment, 1976). Still, it took the better part of a century for the medical community to accept the importance of using empirical evidence to determine which treatments were safe and effective.

In 1938, the Federal Food, Drug, and Cosmetic Act was passed, requiring that the safety of new drugs be demonstrated by scientific investigation before marketing was allowed. The Act was amended in 1962 to add the requirement that efficacy as well as safety be demonstrated for drugs (Office of Technology Assessment, 1976). Despite steady advances over the decades, as recently as 40 years ago it was still possible for a British medical researcher and epidemiologist to create a stir in the medical community by asserting that most medical treatments being used by practitioners were not based on any valid evidence of effectiveness. In his 1972 book, *Effectiveness and Efficiency: Random Reflections on Health Services*, Archibald Cochrane argued that health services should be evaluated on the basis of scientific evidence, rather than on anecdotes, opinion or tradition (Przybylski, 2008). Four years later, the U.S. Office of Technology Assessment (OTA) issued the first of several reports supporting Cochrane’s thesis. In a 1976 report to Congress, for example, the OTA stated that “only 10 to 20% of all procedures used in present medical practice have been proven by clinical trial; many of these procedures

may not be efficacious” (Office of Technology Assessment, 1976, p. 7). Shortly thereafter, the medical community began assembling evidence on effective interventions drawn from rigorous studies and disseminating it in a way that practitioners could easily access and apply (Przybylski, 2008). This was facilitated by the development, between 1992 and 1996, of a series of 19 clinical practice guidelines sponsored by the Agency for Health Care Policy and Research (now the Agency for Healthcare Research and Quality) (Eddy, 2011).¹ In 1993, the Cochrane Collaboration (www.cochrane.org) began in the United Kingdom, with the goal of identifying and synthesizing evidence about effective clinical practices in medicine (Eddy, 2011).

The Evidence Based Practices Movement in Criminal Justice

Just a few years after Cochrane published his critique, Robert Martinson issued his now infamous synthesis of research in corrections (Martinson, 1974), followed by a book by Lipton, Martinson, and Wilks (1975), both of which seemed to lead to the conclusion that “nothing works” in rehabilitating offenders.² In the 1980s, numerous reviews were conducted to rebut Martinson, along with research into the effectiveness of alternative ways of preventing crime (Welsh, 2007). This included a series of publications by Canadian psychologist Paul Gendreau and his colleagues with titles such as “Effective Correctional Treatment: Bibliotherapy for Cynics” (1979) and “Treatment in Corrections: Martinson was Wrong” (1981).

In 1980, the University of Chicago Press began publishing an annual volume entitled *Crime and Justice: A Review of Research*, which included reviews of existing literature on specific topics (although without considering the strength of the research designs or characterizing the effectiveness of individual programs and initiatives).³

Throughout the 1980s and early 1990s, the criminal justice researchers who undertook the task of summarizing what was known about effective programs were concerned with describing what the evidence showed about what types of interventions were effective. There was no systematic effort to identify specific programs that were shown to be effective, nor to rate the quality of the studies that led to their conclusions regarding effectiveness. This changed in

¹ The term “evidence-based” dates to this time period, when it appeared in the title of a 1992 article by David Sackett and his colleagues published in the *Journal of the American Medical Association* (“Evidence-Based Medicine: A New Approach to Teaching the Practice of Medicine”).

² Although many have argued that Martinson’s point was that the poor quality of the available evidence led to the conclusion that researchers could not say definitively “what works” in corrections.

³ *Annual Reviews*, a nonprofit organization, began publishing annual summaries of the literature in biochemistry in 1932, and quickly branched out to over 40 areas, including psychology (since 1950), sociology (since 1975), and law and social science (since 2005).

the mid-1990s with two different efforts to identify specific programs that were effective and to objectively assess the methodological quality of each of the studies supporting conclusions about “what works.”

In 1996, the Center for the Study and Prevention of Violence (CSPV), at the Institute of Behavioral Science, University of Colorado Boulder, designed and launched a national initiative to identify and replicate youth-focused violence, delinquency and drug prevention programs that have been demonstrated as effective. The project, initially called Blueprints for Violence Prevention, identifies prevention and intervention programs that meet a strict scientific standard of program effectiveness. The project initially identified 10 model programs and published detailed descriptions of the programs and the evaluation results. The results of this effort were to identify programs that the scientific evidence showed are effective, and to provide detailed information about these programs so that they could be replicated by others.

In 1996 Congress required the Attorney General to provide a "comprehensive evaluation of the effectiveness" of Department of Justice grants to assist state and local law enforcement and communities in preventing crime. This was the culmination of a long-standing interest on the part of Congress in the evaluation of crime prevention initiatives (Sherman, 1997). In 1972, for example, the Omnibus Crime Control and Safe Streets Act of 1968 was amended to require evaluations of local assistance grants, and the 1988 Anti-Drug Abuse Act Byrne Grants program limited funding to projects of “proven effectiveness” as demonstrated by program evaluation (Sherman, 1997).

In the 104th U.S. Congress, the Senate approved a bill that would have required up to three percent of funds for some local assistance programs to be targeted for evaluation of those programs. The House version of the bill did not include the evaluation set-aside, and the Conference Committee agreed to fund a comprehensive evaluation instead (Sherman, 1997). Congress required that the research for the evaluation be "independent in nature," and "employ rigorous and scientifically recognized standards and methodologies” (Sherman, 1997). The result was a report completed by Dr. Lawrence Sherman and his colleagues at the University of Maryland, an early and highly visible effort to identify EBPs in criminal justice by reviewing research and evaluation studies (Sherman, Gottfredson, MacKenzie, Eck, Reuter, & Bushway,

1997). The Maryland study was one of the first criminal justice efforts to “score” the evaluation studies it reviewed based on the strength of the scientific methods used.⁴

With the establishment of the Internet and the widespread availability of high-speed access to the Web, agencies and organizations began to develop online resources for identifying evidence-based practices in criminal and juvenile justice. These resources included the Office of Juvenile Justice and Delinquency Prevention (OJJDP)’s Model Programs Guide, established in 2000; the Office of Justice Programs’ CrimeSolutions.gov website, established by OJP in 2011; and the BJA-funded What Works in Reentry Clearinghouse, established in 2012. Each of these resources is discussed in greater detail in the section on “Resources for Identifying EBPs.”

Where Does Evidence Come From?

What do we mean when we use the term “evidence?” When we talk about evidence, we mean information about the effectiveness of a program, set of practices, or policy initiative that is generated using established scientific methods. The Office of Justice Programs (OJP) “considers programs and practices to be evidence-based when their effectiveness has been demonstrated by causal evidence, generally obtained through high quality outcome evaluations,” and notes that “causal evidence depends on the use of scientific methods to rule out, to the extent possible, alternative explanations for the documented change.”⁵ Below we will examine two of the key

What is Effectiveness?

What do we mean by the “effectiveness” of a program? In criminal justice, we tend to conceptualize effectiveness in one of several ways: reducing crime (in the case of policing interventions), reducing recidivism (correctional interventions), or reducing victimization/revictimization (prevention/victim-based interventions). For example, a program or intervention targeting probationers or parolees is considered effective if it reduces the likelihood of the individual committing another crime.⁶ There may be other indicators of effectiveness for such a program, but reducing recidivism is usually considered the “bottom line.”

⁴ The “methodological rigor” rating used in the study was based on a scale adapted from one used by the Center for Substance Abuse Prevention in their 1995 study of the effectiveness of substance abuse prevention efforts, which was the precursor to the National Registry of Prevention Programs (NREPP).

⁵ Crimesolutions.gov glossary (www.crimesolutions.gov/glossary).

⁶ Once program effectiveness is conceptualized, it must also be operationalized; that is, we must specify the specific operations/measures that will be used to define the concept. For example, there are many ways to define recidivism:

What are Scientific Methods?

The other key term used in OJP's EBP definition is "scientific methods." There are several key components of evidence that is produced using such methods. In particular, scientific evidence is:

- **objective:** it is observable by others, it is based on facts (rather than thoughts or opinions), and it is free of bias or prejudice that might be caused by personal feelings;
- **replicable:** it can be observed by others using the same methods that were used to produce the original evidence;
- **generalizable:** it can be applied to individuals and groups other than those who were involved in producing the original evidence.

In general terms, scientists (criminal justice researchers or program evaluators) assure that their evidence is *objective* by using precise, unambiguous measures to assess concepts such as recidivism. They assure that evidence is *replicable* by maintaining transparency of the methods they use to collect the information: explaining in detail what they collected and how they collected it, and subjecting their findings to assessment and review by their peers by presenting them at professional conferences and publishing them in refereed journals. *Generalizability* is more difficult to ensure, and usually results from gathering evidence from a representative sample of the kinds of people (offenders) about whom we are interested in forming conclusions.

Randomized Controlled Trials

The hallmark of the scientific method is experimentation. This means comparing two groups: those who receive the intervention (treatment group) and those who do not (control group). The outcomes or measures of effectiveness of interest (for example, recidivism) are compared for the two groups to determine if they are in the hypothesized (expected) direction. For example, if drug courts are effective, then we would expect that probationers seen in drug courts would be expected to have lower recidivism rates than a control group of probationers who appear in regular courts.

The key to ensuring, as the OJP definition states, that we can rule out alternative explanations for observed differences between the groups is that the groups must be the same on

rearrest, reconviction, and reincarceration. There are also different ways that we might measure, or obtain information on, these: police reports, court records, or even self-reports by perpetrators.

all factors other than the intervention. For example, if the drug court probationers are all first time offenders while the regular court offenders all have lengthy criminal histories, then we would expect to see differences in recidivism that are unrelated to the type of court in which they are seen. The best way to ensure the equivalency of the two groups is through random assignment; that is, individuals are assigned to the groups by the researcher/evaluator in a random manner such that each person has an equal chance of ending up in the experimental or control group. This is the best way to ensure that the two groups are equivalent on all factors except the one of interest (in our example, amount of supervision). These designs, known as randomized controlled trials (RCTs), provide confidence that observed differences are due to the intervention, and reduce the likelihood that evaluators will falsely conclude that the intervention being studied is effective. This is what is meant by “causal evidence.”

Quasi-Experiments and Non-Experiments

Randomized controlled trials (RCTs), are often referred to as the “gold standard” for producing evidence. However, there are a number of questions in criminal justice that cannot be easily addressed using RCTs. For example, to determine the effect of sentence length on recidivism, we cannot randomly assign offenders to receive different sentences. Interventions at the community level are also difficult to evaluate using RCTs (for example, determining the effectiveness of a county-based comprehensive domestic violence intervention program). In fact, it can be difficult to persuade any decision-maker (like a judge or program manager) to suspend their usual placement criteria in favor of random assignment to a particular program or intervention.⁷

In cases where RCTs are not feasible, other methods of designing evaluations may be employed that provide some assurance that observed differences are due to the intervention under study and not other factors. These designs, known as quasi-experimental designs, vary in terms of their level of sophistication and their ability to control for possible differences between the groups, other than the intervention, that might produce outcomes. For example, when assessing a program with limited capacity, an evaluator might employ a “waiting list” as a comparison group. The waiting list would consist of individuals who are eligible for the program but have not been admitted due to space considerations. Since those on the waiting list are

⁷ Ethical issues, legal considerations, and cost are additional factors that make implementing RCTs difficult or impractical.

eligible for the program, they should be similar in most respects to those actually in the program. It would thus be reasonable for the evaluator to expect that any observed differences in outcomes are due to the program itself, and not to other differences between the two groups. However, the evaluator cannot be certain of this, since the individuals were not assigned randomly to the two groups. It is for this reason that evidence produced by quasi-experimental designs is not considered as strong or as compelling as evidence from RCTs.⁸

Some evaluations may not manage to use quasi-experimental designs, but may rely on simple measurement of outcomes. For example, an evaluation of a rape awareness campaign may question women in the community about their knowledge of rape and prevention methods at the end of the campaign. This can be considered a “non-experimental” design, since it is not comparing outcomes of different groups or even of the same group at different times. Using this type of non-experimental design, any observations of knowledge cannot be unambiguously attributed to the campaign itself. This is because the women in the community who are questioned may have received other information, been exposed to a variety of situations, or had any number of experiences during the campaign, all of which would be unknown to the evaluator, that might have affected their knowledge of rape and prevention methods. Thus little weight would be given to any evidence of effectiveness produced by this type of assessment.

What is not Scientific Evidence?

Given the characteristics of the scientific method discussed earlier, it should be obvious that there are many types of information that might be collected in an evaluation that would not rise to the level of “scientific evidence.” In particular, opinions, testimonials, and anecdotes are not evidence of effectiveness in and of themselves. For example, a survey of probation and parole officers that shows positive attitudes about an offender reentry program is not evidence, by itself, of program effectiveness.⁹

How Much Evidence is Enough?

The discussion above suggests that there are levels of evidence, and evidence from some evaluations should be given greater weight than evidence from others because it of higher quality. The question arises, then, of how to consider the quantity of evidence. How much

⁸ This may be particularly true in those areas where evidence from RCTs is already available. For example, a quasi-experimental design that shows that a drug court is not effective in reducing recidivism will not count for much when weighed against the positive evidence of effectiveness produced by a number of RCTs of drug courts.

⁹ Although this might be important information that could be put to good use by the program in question.

evidence is needed to consider a specific program to be “evidence-based?” For example, what if a program has been assessed by only one RCT that showed positive outcomes? Should it be considered evidence-based? What if another program has been assessed by two or three quasi-experiments that have shown positive outcomes? Should that program be considered evidence-based? What about a third program where some evidence shows positive outcomes and other evidence shows no outcomes (or even negative outcomes)?

Unfortunately, there is no single satisfactory answer to the questions posed above. As we will see in the discussion of resources, different sources of information on EBPs handle the question of how the quality and quantity of evidence should be balanced differently. However, in recent years researchers and evaluators have focused less on single evaluations and more on examining the magnitude and consistency of the evidence produced by multiple studies of specific programs and initiatives.

Systematic Review and Meta-Analysis

One method that is used to examine multiple studies is conducting a systematic review. Systematic reviews are usually conducted by subject matter experts, as is the case with resources such as CrimeSolutions.gov. In these cases, formal criteria are used to assess the quality of available evaluations of a particular area (the “evidence base”), and conclusions are reached about the effectiveness of that intervention based on application of these criteria by the reviewers.

A second approach to identifying EBPs involves using a statistical technique known as “meta-analysis.” Meta-analyses use statistical methods to combine the results of multiple evaluations of a specific intervention to assess whether, when combined, they show positive program outcomes. Meta-analysis produces an average “effect size” for a particular outcome. For example, a meta-analysis of drug courts would review all available experimental and quasi-experimental evaluations of these programs, looking at outcome measures such as recidivism. Some studies may have shown large decreases in recidivism, others small decreases, and still others no decreases or even increases in recidivism. The meta-analysis would statistically combine these outcomes to produce an average recidivism reduction that could be attributed to drug courts. The statistical significance of this average recidivism reduction could be tested to determine if drug courts in general seem to be effective in reducing recidivism. The average recidivism reduction could also be used to compare outcomes produced by drug courts to those

of other types of criminal justice interventions, perhaps as part of a comparative cost analysis (see, for example, Drake, Aos, & Miller, 2009).

In 2009, Mark Lipsey published a meta-analysis of 548 studies of delinquency interventions published between 1958 and 2002 (Lipsey, 2009). Based on the results of this meta-analysis, Lipsey and his colleagues have developed the Standardized Program Evaluation Protocol (SPEP), a tool that assesses programs by rating how closely their characteristics correspond to those programs shown to be effective at reducing recidivism in the meta-analysis (Lipsey, Howell, Kelly, Chapman & Carver, 2010). The SPEP assesses juvenile justice programs the type of service the program provides, the treatment amount (duration and contact hours), treatment quality, and youth risk level.

Summary

To summarize, the identification of a program, practice or policy as evidence-based requires scientific evidence regarding its effectiveness. Stronger evidence is derived from randomized controlled trials and quasi-experiments, which help to ensure that observed positive outcomes are due to the intervention itself and not other factors. Evidence derived from multiple studies, combined either by expert assessment or by means of meta-analysis, should be weighted more heavily than evidence derived from a single evaluation.

Resources for Identifying EBPs

As noted previously, there are a number of Web-based resources available for identifying EBPs in criminal justice and related fields. A few selected resources are worth mentioning here; a more comprehensive list of Web-based resources can be found in the Appendix.

In criminal justice, the premier resource is CrimeSolutions.gov (www.crimesolutions.gov). Established by OJP in 2011, CrimeSolutions.gov provides information on 270 programs in a number of areas of criminal justice including corrections, courts, crime and crime prevention, drugs and substance abuse, juveniles, law enforcement, technology and forensics, and victims and victimization. Programs are rated as “effective,” “promising,” or “no evidence.”¹⁰ Each program’s rating can be based on one study or more than one study, and this is indicated in the rating. Ratings are assigned by program experts using a standardized protocol known as the Program Evidence Rating Instrument.

¹⁰ As of this writing, 27% of programs on the site are identified as effective, 61% as promising, and 12% as showing no effects.

According to the website, one of the reasons OJP created CrimeSolutions.gov is to “encourage justice practitioners to replicate programs with a track record of success, when it is reasonable and feasible to do so. Replicating programs that have been shown to work and that fit a community’s needs has the potential to save valuable time and resources compared to implementing untested programs that may or may not address the same problems as effectively.”

The Office of Juvenile Justice and Delinquency Prevention (OJJDP) established the Model Programs Guide (MPG) in 2000. The MPG was originally developed as a tool to support the Title V Community Prevention Grants Program, and was expanded in 2005 to include substance abuse, mental health and education programs. The MPG contains over 200 juvenile justice programs in the areas of prevention, immediate sanctions, intermediate sanctions, residential, and reentry. Programs are rated as either “exemplary,” “effective,” or “promising” based on the conceptual framework of the program; the program fidelity; the evaluation design; and the empirical evidence demonstrating the prevention or reduction of problem behavior, the reduction of risk factors related to problem behavior, or the enhancement of protective factors related to problem behavior. Ratings were established by a peer review panel, and are now based on the same rating instrument used by CrimeSolutions.gov.

The What Works in Reentry Clearinghouse (<http://whatworks.csgjusticecenter.org>) is a BJA-funded initiative established by the Council of State Governments in 2012 and designed to provide information on evidence-based reentry interventions. The site contains information about 56 initiatives in six focus areas (brand name programs, employment, family-based programs, housing, mental health, and substance abuse). Interventions are rated on a five-point scale: strong or modest evidence of a beneficial effect; no statistically significant findings; and strong or modest evidence of a harmful effect. The ratings were made by experts using standardized coding instruments.

Outside of the criminal justice arena, an important resource for EBPs is the Substance Abuse and Mental Health Services Administration’s (SAMHSA) National Registry of Evidence-based Programs and Practices (NREPP). NREPP (<http://nrepp.samhsa.gov>) includes almost 300 interventions in the areas of mental health and substance abuse treatment, substance abuse prevention, and mental health promotion. Independent reviewers assess studies in each area on the quality of research and on readiness for dissemination (which includes the availability of

implementation materials, availability of training and support resources, and availability of quality assurance procedures).

Even from this brief summary of available resources, we can see that different organizations and agencies take different approaches to identifying EBPs. Users should review the information provided on the websites carefully to determine what criteria and procedures are used to identify EBPs. In particular, users should be aware of the number of studies that support a particular program or practice, and whether these studies used RCTs or quasi-experimental designs. The Blueprints for Healthy Youth Development website provides a list of 500 youth programs rated on at least one of six federal or private organization EBP websites¹¹, including CrimeSolutions.gov and the OJJDP MPG (see www.blueprintsprograms.com/resources.php).

Implementing EBPs

One of the keys to being able to take advantages of resources that provide lists of EBPs is being able to successfully implement the programs or practices. This is known as “implementing with fidelity.” As the CrimeSolutions.gov website notes:

If you want to replicate a successful program, you have to plan carefully and pay attention to details to accurately reproduce critical program elements that often include specific procedures, personnel qualifications, and client characteristics. The best way to get similar positive results from these programs is to replicate them with fidelity—using the same procedures, with the same kinds of people, and in the same kinds of settings (www.crimesolutions.gov/about_tips.aspx).

Unfortunately, it is often difficult to obtain details about the programs assessed on these various websites. Much of the research and evaluation reviewed on these sites is published in journals, where detailed program descriptions are not available. In fact, detailed program information or implementation manuals may not be available from any source, unless the program is what is sometimes called a “name brand” program (in which case implementation materials will be available for a fee). As noted earlier, SAMHSA’s NREPP includes a readiness for dissemination component that includes an assessment of the availability of implementation materials. This is obviously useful information for those deciding whether to adopt a particular program for their own use.

¹¹ This includes the 44 youth programs on the Blueprints’ own website that it finds to be “model” or “promising.” The list provides information on which website(s) rate which programs, so users can easily identify programs rated by multiple sites.

Adapting EBPs for Local Use

It is often the case that a program cannot be adopted for use directly, but must be *adapted* to fit a particular set of circumstances before it can be used. There may be a variety of reasons that one may choose to adapt a program, including differences in target population (age, rural vs. urban) and potential barriers to implementation such as time, money or resources. Most websites offer caution in adapting EBP programs, advising that key program components should be implemented with fidelity. However, as noted previously, it can be difficult or impossible to identify which program elements must be implemented exactly and which can be changed (and how) without affecting positive outcomes.¹²

In recent years, knowledge about how best to implement programs and practices has been increasing rapidly. One of the leading organizations in this “implementation science” movement has been the National Implementation Research Network (NIRN). The NIRN website (<http://nirn.fpg.unc.edu/>) provides a wealth of information on implementation. Those interested can begin with a comprehensive report produced by NIRN that summarizes what is known about implementation research (Fixsen, Naoom, Blase, Friedman & Wallace, 2005).

What if there is No Evidence?

While many readers of this briefing may be able to identify a program that suits their needs from one of the EBP resources listed above, others may find themselves in a different situation. Some may be interested in implementing a program which has not yet been subjected to rigorous evaluation. Others may be already funding or implementing “homegrown” programs that have not been evaluated. Still others worry about whether there will be room for innovation when an evidence-based approach is adopted. What should be done when there is no evidence of program effectiveness?

The basic answer to this question is that programs and policies should be based, to the extent possible, on theories and concepts that are supported by research. If programs are consistent with established theories of behavioral change, for example, and are implemented

¹² For more information, view JRSA’s Webinar *Making “What Works” Work for You: Evidence-Based Components and Adaptation* at www.jrsa.org/njiec/trainings-presentations.htm.

using (to the extent possible) core components of evidence-based programs (e.g., that high risk offenders receive more services than low risk offenders), we would expect them to be successful. On the other hand, programs or interventions that are based on questionable assumptions about behavior change that do not employ best practices would not be expected to show positive effects.

One example of a recent program that was considered innovative at the time it was implemented (and has received considerable national attention since) is Hawaii's Opportunity Probation with Enforcement (HOPE) program. Begun by Judge Steven Alm in 2004, the program responds to probation violations (such as testing positive for drug use) with immediate sanctions, usually a few days in jail. Evaluations have shown positive outcomes as a result of this approach.

While the HOPE intervention appeared to have been created by Judge Alm rather spontaneously (and thus could be considered an innovative program), the program in fact has a strong theoretical basis. Swift and certainty of punishment have been long established as effective principles in criminal justice. As one evaluation of HOPE explains, "the basic tenets of the HOPE program (the use of clearly articulated sanctions applied in a manner that is certain, swift, consistent, and parsimonious) are well supported by prior research (Hawken & Kleiman, 2009, p. 9)." Unfortunately, the history of criminal justice programming offers many examples of innovative programs and initiatives that were not well supported by prior research, and therefore doomed to failure.¹³

For many years, evaluators have been preaching the importance of specifying program goals and objectives, tying these explicitly to program activities, and measuring both the implementation of the activities and the corresponding outcomes. These are known as program "logic models" because they spell out the logic that connects what the program is doing to the outcomes it expects to produce. A solid program, even one that is not directly supported by scientific evidence, should be able to make a compelling case for how what it is doing is expected to result in positive changes (lower recidivism, fewer probation violations, etc.).

¹³ Boot camps and the "Scared Straight" program for juveniles are examples of initiatives where there was no compelling theory or research supporting the principles of behavioral change that presumably underlay the program activities.

Summary

For the last 40 years or so, the criminal justice field has been moving slowly but inexorably toward the use of scientific evidence to develop programs and interventions designed to prevent and reduce crime and victimization. There are now many resources that can provide funders and program managers with detailed information on evidence-based practices in almost all areas of criminal justice. Many questions and challenges remain regarding the implementation of these EBPs, and researchers and scholars are now turning their attention to these issues. It is clear, however, that we have reached a point in time where policymakers are demanding that programs and initiatives be supported by solid empirical evidence. With diminishing resources available for funding criminal justice issues, understanding how to identify and implement EBPs will be critical for decisionmakers in all areas of the justice system.

References Cited

- Drake, E.K., Aos, S., & Miller, M.G. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: implications in Washington State. *Victims and Offenders, 4*, 170-196.
- Eddy, D.M. (2011). The origins of evidence-based medicine – a personal perspective. *American Medical Association Journal of Ethics, 13*, 55-60.
- Fixsen, D.L., Naoom, S.F., Blase, K.A., Friedman, R.M. & Wallace, F. (2005). *Implementation research: a synthesis of the literature*. Tampa, FL: National Implementation Research Network, University of South Florida.
- Gendreau, P. (1981). Treatment in corrections: Martinson was wrong. *Canadian Psychology, 22*, 332-338.
- Gendreau, P., & Ross, R. R. (1979). Effective correctional treatment: bibliotherapy for cynics. *Crime and Delinquency, 25*, 463-489.
- Hawken, A., & Kleiman, M. (2009). *Managing drug involved probationers with swift and certain sanctions: evaluating Hawaii's HOPE*. Washington, DC: National Institute of Justice.
- Lipsey, M.W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders, 4*, 124–47.
- Lipsey, M.W., Howell, J.C., Kelly, M.R., Chapman, G., & Carver, D. (2010). *Improving the effectiveness of juvenile justice programs: a new perspective on evidence-based practice*. Washington, DC: Georgetown University Center for Juvenile Justice Reform.
- Lipton, D.S., Martinson, R. & Wilks, J. (1975). *The effectiveness of correctional treatment: A survey of treatment valuation studies*. New York: Praeger Press.
- Martinson, R. (1974). What works? - questions and answers about prison reform. *The Public Interest, 35*: 22-54.
- Office of Technology Assessment (1976). *Assessing the efficacy and safety of medical technologies*. Washington, DC: Author.
- Petrosino, A. (2013). Reflections on the genesis of the Campbell Collaboration. *The Experimental Criminologist, 8*, 9-12.
- Przybylski, R. (2008). *What works: effective recidivism reduction and risk-focused prevention programs*. Denver, CO: Colorado Department of Public Safety.

Sherman, L.W., Gottfredson, D., MacKenzie, D., Eck, J., Reuger, P., & Bushway, S. (1997). *Preventing crime: what works, what doesn't, what's promising*. Washington, DC: National Institute of Justice.

Sherman, L.W. (1997). Introduction: the Congressional mandate to evaluate. In Sherman, L.W., Gottfredson, D., MacKenzie, D., Eck, J., Reuger, P., & Bushway, S. *Preventing crime: what works, what doesn't, what's promising*. Washington, DC: National Institute of Justice.

Welsh, B.C. (2007). *Evidence-based crime prevention: scientific basis, trends, results and implications for Canada*. Ottawa: National Crime Prevention Centre.

Appendix: Evidence-Based Practices Resources*

Systematic Reviews and Program Ratings

Crime and Delinquency

CrimeSolutions.gov (www.crimesolutions.gov)

Established by the Office of Justice Programs in 2011, CrimeSolutions.gov provides information on 270 programs rated as “effective,” “promising,” or “no evidence.”

Model Programs Guide (MPG) (www.ojjdp.gov/mpg)

Established by the Office of Juvenile Justice and Delinquency Prevention in 2000, the MPG rates over 200 juvenile justice programs rated as either “exemplary,” “effective,” or “promising.”

What Works in Reentry Clearinghouse (<http://whatworks.csgjusticecenter.org>)

Established by the Bureau of Justice Assistance, rates 56 initiatives in six focus areas on a five-point scale: strong or modest evidence of a beneficial effect; no statistically significant findings; and strong or modest evidence of a harmful effect.

Education

Best Evidence Encyclopedia (<http://www.bestevidence.org>)

Created by the Johns Hopkins University School of Education's Center for Data-Driven Reform in Education with funding from the Institute of Education Sciences, this site classifies programs in math, reading, science, comprehensive school reform, and early childhood education as having strong, moderate or limited evidence of effectiveness.

What Works Clearinghouse (<http://ies.ed.gov/ncee/wwc>)

Developed by the Department of Education's Institute of Education Sciences, the Clearinghouse provides information in over 200 specific areas related to topics/outcome domains such as dropout prevention, early childhood education, and student behavior. For each intervention, the site provides an improvement index, an effectiveness rating, and an indication of the extent of the available evidence.

Health and Medicine

Cochrane Collaboration (www.cochrane.org)

The Cochrane Collaboration is a nonprofit organization that publishes systematic reviews related to healthcare. Over 5,000 reviews in over 30 areas of health and medicine are published online in the Cochrane Database of Systematic Reviews, including child health, mental health, and tobacco, drugs and alcohol dependence.

* Key articles and reports cited in the References section also serve as useful resources, including: Drake et al., 2009; Fixsen et al., 2005; Lipsey, 2009; Lipsey et al., 2010; Przybylski, 2008; and Sherman et al., 1997.

The Community Guide (<http://www.thecommunityguide.org>)

Established by the Department of Health and Human Services' Community Preventive Services Task Force, the Guide produces systematic reviews of effective programs in over 22 areas of health services, including violence, mental health, and alcohol abuse.

Mental Health and Substance Abuse

National Registry of Evidence-based Programs and Practices (NREPP)
(<http://nrepp.samhsa.gov>)

Established by the Substance Abuse and Mental Health Services Administration, NREPP includes almost 300 interventions in the areas of mental health and substance abuse treatment, substance abuse prevention, and mental health promotion.

Multiple Types of Social Interventions

Campbell Collaboration (www.campbellcollaboration.org)

An offshoot of the Cochrane Collaboration, the Campbell Collaboration publishes systematic reviews in the areas of crime and justice, education, social welfare, and international development. Almost 100 different reviews are available online.

Coalition for Evidence-Based Policy (<http://evidencebasedprograms.org>)

This nonprofit organization provides ratings of 45 programs in 12 areas, including crime/violence prevention, K-12 education, and substance abuse prevention/treatment. Programs are designated as "top tier" (those with evidence of sizeable, sustained effects on important outcomes based on randomized controlled trials) or "near top tier" (missing evidence of sustained effects).

Youth Development

Blueprints for Healthy Youth Development (www.blueprintsprograms.com)

Developed by the University of Colorado's Institute for Behavioral Science, the Blueprints website identifies 46 model and promising programs.

Promising Practices Network (PPN) (www.promisingpractices.net)

Developed and maintained by the Rand Corporation, PPN identifies programs that have been shown to improve outcomes for children. Programs are designated as "proven" or "promising." The site includes 28 proven programs and 58 promising programs.

Other EBP Resources

Center for Evidence-Based Crime Policy (<http://cebcp.org>)

Developed by the Department of Criminology, Law and Society at George Mason University, the Center provides a variety of resources related to evidence-based policing and other areas of criminal justice, including the translation of research to practice.

EPISCenter (<http://www.episcenter.psu.edu>)

Penn State's EPISCenter, supported by the Pennsylvania Commission on Crime and Delinquency, promotes the use of evidence-based delinquency prevention and intervention programs through research, advocacy, and technical assistance.

Evidence-Based Medicine Resource Center (<http://www.nyam.org/fellows-members/ebhc>)

This site, established by the Section on Evidence-Based Health Care of the New York Academy of Medicine contains references, bibliographies, tutorials, glossaries, and on-line databases to guide those embarking on teaching and practicing evidence-based medicine. It offers practice tools to support critical analysis of the literature and MEDLINE searching, as well as links to other sites that help enable evidence-based medical care.

National Implementation Research Network (NIRN) (<http://nirn.fpg.unc.edu>)

The University of North Carolina's NIRN provides information on implementation science and organizational change. NIRN conducts research and publishes information on how to effectively implement evidence-based programs on a national scale.

National Juvenile Justice Evaluation Center (NJJEC) (www.jrsa.org/njjec)

Developed by the Justice Research and Statistics Association in 2010 with funding from OJJDP, NJJEC's goal is to improve the evaluation capacity of states, tribes, and local communities and facilitate the use of evidence-based programs and practices in juvenile justice.

Washington State Institute for Public Policy (www.wsipp.wa.gov)

Created by the Washington legislature, WSIPP conducts research on evidence-based practices in education, criminal justice, welfare, and health. WSIPP is particularly known for its work in cost-benefit analysis, and the development of a methodology to calculate the costs and benefits of a variety of criminal justice initiatives.