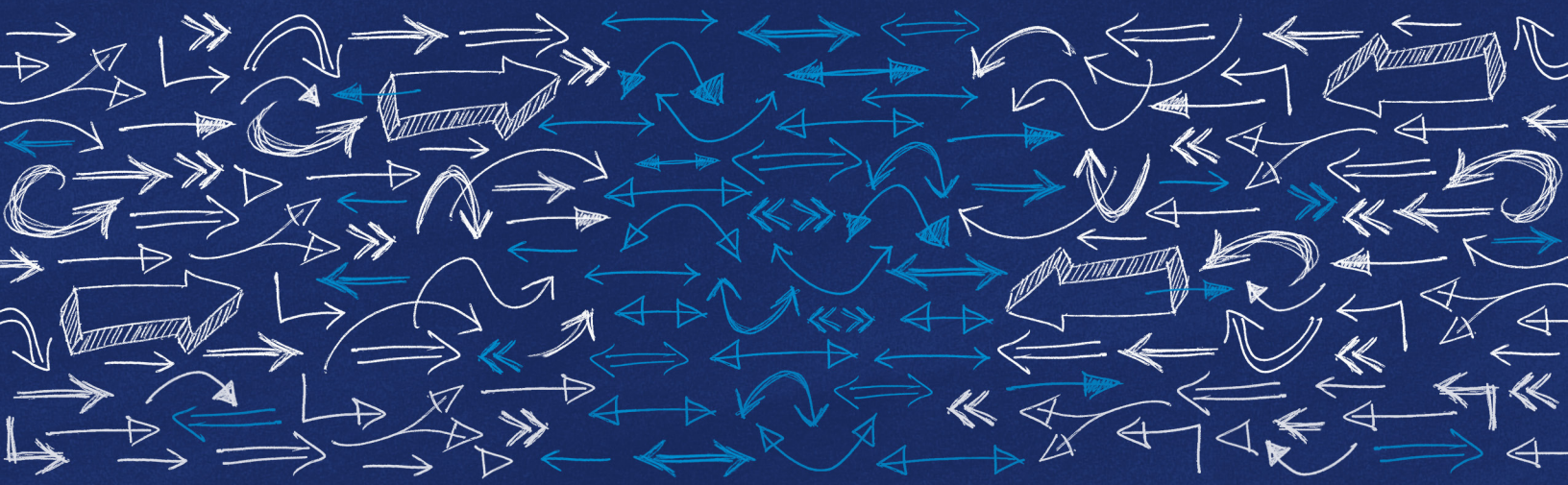
Center for Juvenile Justice Reform
• working across systems of care • georgetown university • gppi policy matters

Improving the Effectiveness of Juvenile Justice Programs

A New Perspective on Evidence-Based Practice



Mark W. Lipsey · James C. Howell · Marion R. Kelly
Gabrielle Chapman · Darin Carver

Improving the Effectiveness of Juvenile Justice Programs

A New Perspective on Evidence-Based Practice

December 2010

Mark W. Lipsey, Peabody Research Institute, *Vanderbilt University*

James C. Howell, Criminologist, *Pinehurst, North Carolina*

Marion R. Kelly, Consultant, *Richmond, Virginia*

Gabrielle Chapman, Peabody Research Institute, *Vanderbilt University*

Darin Carver, Juvenile Justice Programming Consultant, *META L.L.C.*

Contents

Preface 1

I. Introduction 5

II. The Cyclic History of Criminal Justice Treatment and Punishment Philosophies 7

A. From Rehabilitation to Punishment 7

B. Popularity of Deterrence Philosophies 8

C. Return to Rehabilitation. 8

D. Recent Policy Developments 9

III. Prevention and Intervention Programs for Juvenile Delinquency 11

A. The Key Role of Behavior Change Programs for Juvenile Offenders. 12

B. The Spectrum of Programs and the Challenge of Taking Effective Programs to Scale 12

C. Doubts about Whether Many Programs Used in Practice Are Actually Effective 13

IV. Evidence-Based Practice: More Than One Approach 17

A. Direct Evaluation of the Effects of the Program as Implemented 17

B. Model Programs with Evidence Certified by a Credible Source 18

C. Best Practice Guidelines Based on a Meta-analysis of Research Findings 19

V. Meta-analysis of Research on the Effects of Intervention Programs for Juvenile Offenders. 21

A. Analysis of the Findings of 548 Evaluation Studies of Delinquency Interventions 22

B. Program Characteristics Associated with the Greatest Effects on Recidivism 22

1. Risk Level of the Juveniles 23

2. Therapeutic versus Control Treatment Philosophies 23

3. Generic Program Types and Embedded Model Programs 25

4. Amount and Quality of Service 27

C. Effective Juvenile Justice Programs: Implications for Practice 28

VI. The SPEP: Evidence-Based Practice Guidelines 29
A. The Standardized Program Evaluation Protocol for Assessing Juvenile Justice Programs 29
B. The Experience of State Juvenile Justice Systems with the SPEP Tool 32
1. North Carolina 32
2. Arizona 33
C. Lessons Learned from the North Carolina and Arizona SPEP Projects 34
D. The Challenge of Evidence-Based Practice for Service Providers 35

VII. Integrating Evidence-Based Practice into Juvenile Justice Systems 37
A. A System Reform Framework 37
1. Prevention Tier 38
2. Intervention and Graduated Sanctions Tier 39
B. The Essential Tools: Risk Assessment, Needs Assessment, Case Management Plan 40
C. Examples of a Comprehensive Continuum of Prevention and Graduated Sanctions 42
1. San Diego's Comprehensive Strategy 42
2. Orange County's Comprehensive Strategy 43
3. Missouri's Comprehensive Strategy 44

VIII. Practical and Policy Considerations in Implementing Juvenile Justice Reform 47
A. Needed Improvements in Juvenile Justice Systems 47
B. The Challenges of Change and Sustainability 48

IX. Recommendations 51
A. To Juvenile Justice System Administrators 51
B. To Legislators 51
C. To Juvenile Justice State Advisory Groups 51
D. To Judges 51
E. To Treatment Providers 52

X. Conclusion 53

References 55

Preface

As a society we want our children to be healthy, safe, happy, fulfilled, and connected to others in a loving, positive manner—and as parents we do whatever we can to ensure those outcomes for our children. Those who work in the social services share the same goals for the children, youth, and families they serve.

Unfortunately, though individual workers do their best in this regard, they are too often significantly challenged by the systems within which they do their work to achieve the outcomes we want for our children. Appropriate and effective services may not be available, it may not be possible to match a youth's needs to the services that are available, and there may not be a way to determine if the services that are available are effective. These challenges are not the result of a lack of knowledge. We now have the knowledge to do this work more effectively; indeed, the research that we have in hand today far exceeds our knowledge base as little as 5 to 10 years ago. Research sponsored by the Office of Juvenile Justice and Delinquency Prevention, the National Institute of Justice, and the Bureau of Justice Statistics (all within the U.S. Department of Justice), the U.S. Department of Health and Human Services, and a number of foundations has helped to grow our knowledge. We now have research on best practices for juvenile justice–involved youth and the policies that support the practices. We find this reflected in the increased use of evidence-based practices and programs, in the growth of the science of risk and protective factors and criminogenic factors and characteristics, and in the development and use of validated risk and needs assessment instruments. We have learned about the importance of advancing our work on an ecological platform, serving youth closer to home, and better connecting youth to family, school, community, and pro-social peers while utilizing a strength-based approach. The true challenge is not, therefore, a lack of knowledge of what works, but rather is in translating the robust body of knowledge into practice.

This is what the framework presented in this paper is designed to do. By bringing together the work of Dr. James "Buddy" Howell and his colleagues on the Comprehensive Strategy for Serious, Violent, and Chronic Juvenile Offenders (Comprehensive Strategy) and the creation by Dr. Mark Lipsey of the Standardized Program Evaluation Protocol (SPEP), based on his groundbreaking meta-analyses of juvenile justice research, the framework presented in this paper is poised to meet one of the greatest challenges we have in juvenile justice practice today: how to bring together in a coherent manner the advances in knowledge noted above.

To demonstrate the need for a new approach, contemplate this scenario and whether it sounds familiar. A juvenile justice director is delighted to identify a number of "gold standard" programs that could be used to benefit his or her clients, whether found in the Blueprints for Violence Prevention developed by Dr. Delbert Elliott, or in OJJDP's Model Programs Guide, or in the Substance Abuse and Mental Health Services Administration's National Registry of Evidence-based Programs and Practices. The new programs are implemented with as much fidelity as possible in light of budget constraints and workforce limitations, while at the same time local programs that do not have rigorous evidence of success are diminished. Outcomes may improve for the clients who experience these gold standard programs, although their replication may be uneven with mixed levels of effectiveness. And their reach may be limited due to the expense associated with their implementation and resistance from providers who are reluctant to replace their current programs with new ones. Moreover, the programs are implemented in silos, disconnected from a systemwide quality assurance approach and a continuum of effective services to meet the needs of youth. Despite these challenges, the use of gold standard programs is viewed by many as a magic bullet, and in some instances, states are mandated to fund only these programs—resulting in reductions in funding for

local programs that may have measures of effectiveness, but that do not have rigorous evaluation studies.

The authors of this paper suggest that we can do better at translating knowledge into practice without wavering in our commitment to evidence that supports our policies and practices. I can best bring to life this need to do better through a story I was once told about a lecturer who was addressing an audience about how we decide to assume risk in our lives. He posed three hypothetical questions to a volunteer in the audience. First, he asked the volunteer to imagine that there was a steel construction I-beam, 15 feet long, 6 inches high, and 6 inches wide, lying in front of the podium and offered the volunteer $50 to assume the risk of walking across it. The volunteer indicated that she would assume the risk.

The second hypothetical presented a situation in which the I-beam had been lengthened to 30 feet and was located across a gorge that dropped 250 feet to a bed of rocks. Offered $100 to assume the risk of walking across the I-beam, the volunteer declined.

Presenting the third hypothetical, the lecturer kept the circumstances the same as in the second scenario, except for one significant difference. In this situation, the lecturer had one of the volunteer's children on one side of the gorge and was holding the child by the hand, over the edge of the gorge. The volunteer was on the other side of the gorge, and unless she crossed the I-beam, the lecturer would drop her child. The lecturer offered the volunteer $200 to walk across the I-beam. The volunteer hesitated for a long moment before responding, "Which one of my kids have you got?"

I am sure that any amusement you might find in this story may reflect the fact that you are a parent who has had "one of those days" with your kids—or that you were one of those kids! I share this story, however, not merely as an amusement, but to amplify a point. You are unlikely to ever meet an individual who says that he or she does not care about kids. All of us truly want what is best for children on some level. But the way that concern is expressed may vary a great deal. The woman in my story was being asked a very clear question: exactly what steps are you willing to take to help a child? What kind of priority do children have for you when the going gets tough, when there are choices to be made? And her answer revealed what may be an even harder question—which children are you willing to help?

The truth of the matter is that the vast majority of parents would do whatever it would take to get across that I-beam—in fact, virtually every adult would do whatever it would take to save that child. But as a society, perhaps through our benign neglect, we don't do whatever it takes, and kids to one extent or another are falling into the gorge. Our challenge is to take those extraordinary efforts that individual workers are willing to make and embed them into systems that operate efficiently, effectively, and fairly in meeting the needs of youth who come in contact with them—systems that make it possible for workers to grab our children by the hand and not let them fall.

The framework presented in this paper will help juvenile justice systems around the country reform their systems in this way. The overarching frame for the approach is to construct juvenile justice systems that are aligned along a continuum of care, from prevention to early intervention and then to more significant system involvement as needed. Incorporated into that continuum are the fundamental elements of valid risk and needs assessments, the matching of the level of risk and need to the appropriate service, and then ensuring that the services provided are effective at improving outcomes for the children and youth placed in them. By embedding Lipsey's SPEP in the Comprehensive Strategy framework, the approach presented in this paper allows us to maximize the use of the research we have while not getting stuck in the box of evidence-based programs more narrowly defined by the "gold standard" of program effectiveness. Instead, the SPEP allows juvenile justice agencies to compare their current services to best practices shown in the research to improve outcomes for juvenile justice–involved youth. This is done via an automated and ongoing process of quality improvement across virtually all services juvenile justice agencies provide. Though the use of gold standard programs is encouraged, if appropriate for the needs of the youth served, the SPEP approach allows for the retention of local programs and provides a systemwide quality assurance mechanism across the continuum of care.

The SPEP approach can be used as an overlay with any existing juvenile justice system. Though some will have to increase their commitment to the use of validated assessment instruments, research-informed programs and practices, and outcome measurements, the beauty of the approach and the timing of its introduction is that this is the direction that the juvenile justice field has been heading. This movement has been led by the likes of Terry Thornberry, David Huizinga, Rolf Loeber, Del Elliott, Rico Catalano, David Hawkins, Barry Krisberg, John Wilson, Peter Greenwood, Clay Yeager, and others, let alone Mark Lipsey and James "Buddy" Howell, the primary authors of this paper.

The juvenile justice field has been living in an evidence-based and outcome-driven world for the past decade, but has been missing the operating platform that would bring the various evidence-based pieces together. The framework presented in this paper provides this platform and facilitates this work going to the next level of implementation and performance. It provides the balanced and coherent framework of the Comprehensive Strategy with the quality assurance mechanisms of the SPEP. Indeed, it is my belief that the introduction of this framework and its adoption by juvenile justice agencies across the country will be one of the greatest advances in the juvenile justice field over the past several decades and into the next. It will facilitate an appropriate balancing of prevention and intervention while making it possible to create the greater levels of system accountability and performance that our knowledge now allows us to achieve.

I am delighted, therefore, to author the preface to the paper that will introduce this next generation of work. I thank all of the authors for their insight and vision in advancing the juvenile justice field and helping to improve the life outcomes of the children, youth, and families whose lives we touch. They have helped us make sure that, as a field, we are able to cross every I-beam that we confront in our work and grab every child and youth by the hand to help them lead healthy, safe, happy, and fulfilled lives.

Shay Bilchik
Research Professor/Center Director
Center for Juvenile Justice Reform
Georgetown Public Policy Institute
Georgetown University

I. Introduction

Juvenile justice systems in the United States have long struggled with the inherent tension between their role in meting out punishment for violations of law and their role as an authoritative force for bringing about constructive behavior change in the wayward youth who commit those violations. Our view is that the overarching and intertwined goals of juvenile justice should be ensuring public safety—protecting the public from any additional harm caused by juvenile offenders—and altering the life trajectories of those juveniles to not only reduce further criminal behavior but to improve their chances to prosper as productive citizens. Attaining those goals requires the capability to control behavior in the short term and the means to induce self-sustaining behavior change that will persist after youth are no longer under court supervision.

Juvenile justice systems have longstanding methods for controlling behavior, such as community supervision and custodial care, though these are not always used as efficiently and effectively as possible. Effective programming to reduce recidivism and produce other positive outcomes, however, has been more problematic. Juvenile justice systems make use of many treatment programs, but, in most cases, the effectiveness of those programs is difficult to determine and largely unknown. An increasing body of research evidence addresses this problem, but the findings of that research have not been well integrated into most juvenile justice systems. Translation of research into practice is always a challenge, but it has been exacerbated in this instance by overly narrow conceptions of how research should be used to inform juvenile justice practice.

This paper introduces a framework for major juvenile justice system reform—the integration of a forward-looking administrative model with evidence-based programming. The administrative model is organized around risk management and risk reduction aimed at protecting the public by minimizing recidivism. Evidence-based programming is organized around services that moderate criminogenic risk factors and enhance adaptive functioning for the treated offenders. Placements are guided by a disposition matrix that supports individualized disposition plans and is organized around the risk levels and treatment needs of offenders as assessed by empirically validated instruments. An array of effective programs is supported that provides sufficient diversity to allow matching with offenders' needs. This array of programs is integrated with a continuum of graduated levels of supervision and control so that offenders can be stepped up the ladder and placed in more highly structured program environments if behavior worsens and stepped down when there is improvement. Such a system is consistently forward-looking in basing program placements and supervision levels upon objective risk and needs assessments and in constructing case management plans focused on improving future behavior rather than punishing past behavior.

II. The Cyclic History of Criminal Justice Treatment and Punishment Philosophies

A. From Rehabilitation to Punishment

During most of the twentieth century, state sentencing policies were primarily offender oriented and based on a rehabilitative model of individualized sentencing (Tonry, 2009; Warren, 2007). Beginning in the 1960s, the national crime rate sharply increased. At the same time, evaluations of rehabilitative interventions were interpreted as showing that "nothing works" (Lipton, Martinson, and Wilks, 1975; Martinson, 1974), and these claims cast a shadow over therapeutic criminal and juvenile justice policy and practice (Tonry, 2004). In light of these developments, the federal government and many states turned to offense-based sentencing policies and embraced more punitive measures. The assumption that rehabilitative treatment was ineffective persisted throughout the 1970s and 1980s until scholars used advanced analytical tools to examine the evidence more closely (Cullen, 2005). The results refuted Martinson's negative assessment and showed that rehabilitative programs, if implemented well, can substantially reduce recidivism (Cullen, 2005; Lipsey and Cullen, 2007).

The assumption that rehabilitative treatment was ineffective persisted throughout the 1970s and 1980s until scholars used advanced analytical tools to examine the evidence more closely.

The pendulum swing from treatment to punishment filtered down from criminal justice to the juvenile justice system (Howell, 2003b). Two compelling images in the 1990s buttressed policies that enhanced punishment for juvenile offenders. A professor of politics and public affairs at Princeton University, John Dilulio, created and popularized the concept of juvenile super predators (1995). He coined this term to call public attention to what he characterized as a "new breed" of offenders, "kids that have absolutely no respect for human life and no sense of the future....These are stone-cold predators!" (p. 23). Elsewhere, Dilulio and his co-authors described these young people as "fatherless, Godless, and jobless" and as "radically impulsive, brutally remorseless youngsters, including ever more teenage boys who murder, assault, rob, burglarize, deal deadly drugs, join gun-toting gangs, and create serious disorders" (Bennett, Dilulio, and Walters, 1996, p. 27).

In addition, Dilulio and Wilson contrived another scary image in their prediction that a new wave of juvenile violence would occur between about 1995 and 2010, which they based in part on a projected increase in the under-18 population (Dilulio, 1996, 1997; Wilson, 1995). The dire warnings of a coming-generation of super predators that helped to promote punitive policies rested on three assumptions: that the relative proportion of serious and violent offenders among all juvenile delinquents was growing, that juvenile offenders were becoming younger and younger, and that juveniles were committing more and more violent crimes.

None of these assumptions proved to be correct. Various researchers debunked the super predator myth and doomsday projections (Howell, 2003b, 2009; Males, 1996; Snyder, 1998; Snyder and Sickmund, 2000; Zimring, 1998). Examination of the evidence by these researchers revealed that a new wave of super predators did not develop, nor did a general wave of juvenile violence occur. However, there was a sharp increase in adolescent and (mostly) young adult homicides in the late 1980s and early 1990s (Cook and Laub, 1998). But even at the height of that trend (1993), "only about 6 percent of all juvenile arrests were for violent crimes and less than one-tenth of one percent of their arrests were for homicides" (McCord, Widom, and Crowell, 2001,

p. 33). Furthermore, only very small increases were seen in victimization and self-report measures of crime, mostly in nonserious offenses (Howell, 2003b).

Crime control policies had already changed, however. Martinson's (1974) negative assessment of the effectiveness of rehabilitation continued to provide one impetus. In addition, the mass media, politicians, and law enforcement characterized youth crime as an epidemic of gun violence and crack cocaine, a claim that further fueled support for more punitive sentencing policies (Brownstein, 1996; Reeves and Campbell, 1994). Racial conflict during this period may also have contributed to the "get-tough" policies in both the juvenile and criminal justice systems, as some scholars have argued (Feld, 1999; Tonry, 2009; Tonry and Melewski, 2008).

B. Popularity of Deterrence Philosophies

The get-tough movement included increased emphasis on deterrence and a decline in rehabilitative approaches. Juveniles believed to have fulfilled Dilulio's characterization as super predators were thought to be beyond redemption; jailing and imprisonment was the presumed answer. "Just deserts" advocates promoted the use of punitive laws, policies, and practices in the juvenile justice system, including "three strikes" laws, determinate sentences, longer sentences, electronic monitoring, drug testing, shock incarceration, and other such measures (Howell, 2003b). Rehabilitation programs often were abandoned, whereas boot camps, Scared Straight programs, detention centers, and juvenile correctional facilities increasingly populated the nation's landscape (Howell, 2003b; Males, 1996; Roush and McMillen, 2000). Juvenile courts designated larger proportions of juveniles as serious and violent offenders and incarcerated more juveniles (Snyder and Sickmund, 2006).

Juvenile courts designated larger proportions of juveniles as serious and violent offenders and incarcerated more juveniles.

Such policies and practices, which deemphasize prevention of juvenile crime and rehabilitation of juvenile offenders, became common in the juvenile justice system through new state legislation. By the end of the 1990s, all the states had enacted laws to make their juvenile justice systems more punitive or to transfer more juveniles to the criminal justice system and confine them in adult prisons (Howell, 2009, pp. 288–90).

C. Return to Rehabilitation

Although many state legislatures rewrote their juvenile codes to endorse punitive objectives in the 1990s (Torbet and Szymanski, 1998), nearly all of them maintained wording that upheld the juvenile justice system's traditional rehabilitative mission (Bishop, 2006; Tanenhaus, 2002, 2004). Moreover, there continued to be public support for a rehabilitative approach to dealing with juvenile offenders despite assumptions to the contrary by some observers. As Cullen (2006) noted, "the notion that the American public is opposed to the treatment of juvenile offenders is a myth" (p. 665). A 2001 national survey, for instance, found that 80 percent of the sample of adults thought that rehabilitation should be the goal of juvenile correctional facilities and that more than 9 in 10 favored a variety of early intervention programs, including parent training, Head Start, and after-school programs. "The legitimacy of the rehabilitative ideal—especially as applied to youthful offenders—appears to be deeply woven into the fabric of American culture" (Cullen, 2006, p. 666).

It is not surprising, therefore, that against the punitive trend in juvenile justice there was a countervailing trend toward embracing options that include both punishment and rehabilitation as central guiding tenets (Butts and Mears, 2001; Mears, 2002). Specialized courts—including drug, gun, domestic violence, and mental health courts—illustrate this contrary trend and can be found in both the juvenile and criminal justice systems. Bishop (2006) observed in her review of the three years of legislative actions from 2003 to 2005 that "efforts are underway to mitigate or even abandon punitive features [of juvenile laws enacted in the past decade and] to address the treatment needs of most juvenile offenders" (p. 660; see also Butts and Mears, 2001). For example:

- Some legislation aimed at improving individualized treatment for committed offenders was adopted (Mississippi, South Dakota, and Wyoming).
- Provision for mental health assessment and treatment was passed in four states (Connecticut, Idaho, Virginia, and Washington).
- Several states enacted laws to establish teen courts and other diversion programs.
- Four states (Colorado, Indiana, Ohio, and Michigan) passed legislation to provide drug treatment and several others included drug treatment in other initiatives.
- The Mississippi legislature phased out its boot camps.
- Illinois established monetary incentives for counties to reduce commitments to state institutions.
- Colorado and South Dakota enacted measures to separate juvenile offenders transferred to the criminal justice system from incarcerated adults.
- Connecticut enacted a measure to gradually raise the age of juvenile court jurisdiction from 16 to 18 by 2010. (The North Carolina legislature is also considering a similar measure.)
- Three states (Florida, Pennsylvania, and Washington) adopted evidence-based programming in juvenile corrections. (North Carolina, Oregon, and Tennessee legislatures have since enacted a similar requirement.)

D. Recent Policy Developments

Recent changes have been motivated by a variety of factors, including economic ones. Large budget deficits have caused some states to rethink high juvenile confinement rates. A few states have found it necessary to reduce funding for community programs in order to maintain the confinement infrastructure. Both Texas and Ohio have new legislation that prohibits the confinement of misdemeanants while other states, notably Mississippi, South Carolina, and Washington State, have struggled to meet the core requirement for deinstitutionalization of status offenders of the federal Juvenile Justice and Delinquency Prevention (JJDP) Act (Kelly, 2010). In some instances policy decision making is propelled by litigation, which is used more often now than in the past, both with regard to conditions of confinement as well as disputes about fundamental fairness and the quality of justice for juveniles before the courts. Between 2000 and 2007, 20 CRIPA investigations were made of 23 juvenile justice facilities in more than a dozen states (U.S. Department of Justice, 2007).[1]

Perhaps the two most progressive policy reforms of recent years are the drive for evidence-based practice, which focuses on effective treatments, services, and supports for children and families, and the effort to establish systems of care to address the infrastructure of funding and linkages between services and programs.

Perhaps the two most progressive policy reforms of recent years are the drive for evidence-based practice, which focuses on effective treatments, services, and supports for children and families, and the effort to establish systems of care to address the infrastructure of funding and linkages between services and programs. These themes have been embraced in educational, mental health, and child welfare services policy reforms, as well as in juvenile justice systems. The development of the Office of Juvenile Justice and Delinquency Prevention's (OJJDP) Comprehensive Strategy for Serious, Violent, and Chronic Juvenile Offenders in the 1990s parallels the emergence of systems of care in other human service fields. Title V of the JJDP Act codified the importance of community planning and collaboration in delinquency prevention programming.

The Blueprints for Violence Prevention initiative was an early adopter and innovator on the theme of evidence-based programs. Cost-benefit analyses conducted by the Washington State Institute for Public Policy[2] in

[1] *Since its enactment in 1980, the Civil Rights of Institutionalized Persons Act (CRIPA), 42 U.S.C. § 1997a et seq., has allowed the Civil Rights Division of the U.S. Department of Justice to investigate possible civil rights violations pertaining to persons in publicly operated institutions and to bring consequent legal actions against state or local governments (Blalock and Arthur, 2006).*

[2] *The Washington State Institute for Public Policy is an applied research group of the Washington State legislature.*

both juvenile justice and child welfare emanated from legislation in Washington State that limits funding to evidence-based practices. The institute's groundbreaking series of cost-benefit studies identified evidence-based public policy options for juvenile justice and demonstrated how investments in these options could decrease incarceration, save taxpayer dollars, and lower recidivism rates. These studies changed the policy conversation in Washington State from one focused on base funding levels to one focused on funding cost-effective evidence-based practice (Greenwood, 2010).

The proposed Federal Youth PROMISE (Prison Reduction through Opportunities, Mentoring, Intervention, Support, and Education) Act (H.R. 1064/S.435) is targeted to communities facing the greatest youth gang and crime challenges, enabling them to develop a comprehensive response to youth violence through coordinated prevention and intervention services. The act would mandate OJJDP to develop standards for evaluation of juvenile delinquency and criminal street gang prevention and intervention approaches carried out under the PROMISE Act. It would also create incentives for communities to establish intervention plans that include a broad array of evidence-based prevention and intervention programs. In addition, the bill would establish a National Center for Proven Practices Research. This center will collect and disseminate information to professionals and the public on current research and other evidence-based and promising practices related to prevention and intervention for juvenile delinquency and criminal street gang activity.

It is within the context of the renewed attention to rehabilitation and the associated recent policy developments that this paper is presented. In the following sections, the effectiveness of the rehabilitative approach will be further explored, as will the challenges of taking effective programs to scale. Different approaches to evidence-based practice will be discussed, including a particular focus on the use of meta-analysis as a way to identify best practices from the analysis of many individual studies. The findings of the most comprehensive meta-analysis done on juvenile justice services will be presented and a method for using that knowledge to improve current juvenile justice programs will be explored. This approach to establishing evidence-based practice will then be placed within the framework of OJJDP's Comprehensive Strategy for Serious, Violent, and Chronic Juvenile Offenders, which focuses on the use of structured decision-making tools, such as risk and needs assessments, and a continuum of graduated levels of supervision and control integrated with effective behavior change programs. Coupling a method to improve juvenile justice programs and services with the Comprehensive Strategy approach provides a holistic framework to ensure that a juvenile justice system is operating in an effective and efficient way to improve outcomes for the youth in its care.

III. Prevention and Intervention Programs for Juvenile Delinquency

Dealing effectively with juvenile delinquency involves two distinct but overlapping endeavors—prevention and intervention—each of which has somewhat different purposes and requires the efforts of somewhat different agencies and actors. For present purposes, we define prevention as community-based activities aimed at helping youth avoid delinquent behavior and consequently coming into contact with the juvenile justice system. Prevention programs are mainly developed and implemented by schools, social service agencies, mental and public health agencies, and the like. Juvenile justice agencies, of course, are also often involved, along with law enforcement, but the focus of prevention efforts is on youth who may be at risk for delinquent behavior but have not yet been referred to juvenile justice agents for response to an alleged delinquent offense.

Much is known about effective prevention programs from research and practice and the question of how to optimize such programs for cost-effective impact on juvenile behavior is worthy of careful consideration.

Prevention is an essential part of an effective strategy for addressing juvenile delinquency in any community. Indeed, if it were completely successful, there would be no need for a juvenile justice system and, even when only partially successful, it produces better outcomes for the affected youth, the community, and the juvenile justice system. Much is known about effective prevention programs from research and practice and the question of how to optimize such programs for cost-effective impact on juvenile behavior is worthy of careful consideration. That topic goes beyond the scope of this paper, however. Here we focus on the interaction of the juvenile justice system with alleged juvenile offenders who are presented to agents of that system for a response, typically by law enforcement or school personnel.

Some such cases are not accepted into the juvenile justice system, e.g., if the alleged offense is very minor or if there is little evidence that it actually occurred. Beyond that point, we will refer to the response of the juvenile justice system to such cases as intervention. For instance, cases may be formally accepted but immediately diverted out of the juvenile justice system with or without conditions. For our purposes then, diversion is an intervention. Juvenile justice interventions involve two components—a supervisory component and a treatment component—though either may be minimal or nonexistent for some interventions. The supervisory component consists of some structure for monitoring or controlling the youth's behavior, e.g., probation supervision, day reporting, electronic monitoring, nonsecure residential facilities, and secure custodial institutions. The treatment component consists of activities or services provided within the supervisory structure that are intended to facilitate positive behavioral changes that will endure after supervision has ended, e.g., counseling, victim-offender mediation, cognitive-behavioral therapy, academic instruction, vocational training, and the like.

The supervisory component constrains the juvenile's freedom to act and access to social and personal amenities to some degree or another, with incarceration representing the most extreme form. On the one hand, this impedes additional delinquent behavior and thus provides a means of protecting the public from a juvenile viewed as an immediate threat to persons, property, or self. On the other hand, such loss of freedom and access is aversive and thus constitutes punishment for the instigating offense. When the level of control goes beyond what is proportionate to the offender's risk to public safety, the supervisory component is being used punitively.

A growing body of research documents the key role of the treatment component in reducing the subsequent criminal behavior of juvenile offenders and the minimal or even negative effects of punitive interventions.

A. The Key Role of Behavior Change Programs for Juvenile Offenders

Although more and more evidence has emerged demonstrating that certain forms of treatment for juvenile offenders are effective (Cullen, 2005; Lipsey, 2009; Lipsey and Cullen, 2007), policy questions concerning the appropriate balance between treatment versus punishment continue. However, recent systematic research reviews reveal three very important findings that should inform future debate. First, for juvenile offenders in general, the juvenile justice supervisory apparatus of probation and court monitoring, group homes, correctional facilities, and the like has, at best, only modest favorable effects on subsequent recidivism and some evidence shows modest negative effects (Lipsey and Cullen, 2007; Petrosino, Turpin-Petrosino, and Guckenburg, 2010). Second, deterrence-oriented programs that focus on discipline, surveillance, or threat of punitive consequences (e.g., prison visitation Scared Straight–type programs, boot camps, and intensive probation supervision) on average have no effect on recidivism and may actually increase it (Lipsey, 2009). Third, many "therapeutic" programs oriented toward facilitating constructive behavior change have shown very positive effects—even for serious offenders (Lipsey, 2009; Lipsey and Cullen, 2007; Lipsey and Wilson, 1998). If reducing the subsequent criminal behavior of offenders with its associated benefits for public safety is the goal, the implications of these findings are that (1) juvenile offenders with low risk for reoffending should be diverted from the juvenile justice system; (2) offenders with moderate or high risk for reoffending should be subject to the minimal level of supervision and control consistent with public safety and be provided with appropriate, effective therapeutic services; and (3) subjecting juvenile offenders to punishment beyond that which is inherent in the level of control necessary for public safety is likely to be counter-productive to reducing recidivism.

B. The Spectrum of Programs and the Challenge of Taking Effective Programs to Scale

An increasing body of research on the effectiveness of treatment programs for juvenile offenders is available to practitioners, and the collective findings of that research have identified many effective programs and provided considerable detail about their key characteristics. There are two main reasons for this expansion of knowledge. First, the program evaluation base has expanded significantly, providing a deeper and more detailed body of empirical evidence about the effects of programs for juvenile offenders. Second, the quantitative technique of meta-analysis emerged and has been applied extensively to juvenile justice programs. Meta-analysis allows researchers to analyze and synthesize the characteristics of programs and the effects of those programs in a systematic, replicable manner. It also enables them to examine a wider range and larger number of program evaluation studies in an integrated fashion than was possible in the past.

Many meta-analytic reviews have been conducted on particular programs or types of programs for juvenile offenders, such as boot camps (MacKenzie, Wilson, and Kider, 2001), cognitive-behavioral therapy (Landenberger and Lipsey, 2005), prison visitation (Petrosino, Turpin-Petrosino, and Buehler, 2003), family therapy (Latimer, 2001), drug court (Wilson, Mitchell, and MacKenzie, 2006), victim-offender mediation (Nugent, Williams, and Umbreit, 2004), Multisystemic Therapy (Littell, Popa, and Forsythe, 2005), and the like. Other meta-analyses have examined multiple programs over a broad range of program types in order to compare the effectiveness of different programs for reducing the recidivism of juvenile offenders (Aos et al., 2001; Andrews, Zinger, and Hoge, 1990; Lipsey and Wilson, 1998; Lipsey, 2009). This extensive research synthesis work on a large and growing body of evaluation studies of treatment programs for juvenile offenders has identified many programs and program types that produce significant reductions in recidivism along with positive effects on such other outcomes as school attendance, family and peer relationships, employment, and mental health symptoms.

Research that tells us what works to address a particular behavior problem, however, is only a beginning point. Implementing those programs in an existing service system, while retaining their effectiveness, is yet another matter. At present, we know relatively little about the effects of taking evidence-based programs to scale in public health and related areas of mental health, education, welfare, and criminal justice. Nor do we know a great deal about how to do so in a way that attains the same positive outcomes observed in the research studies. Efforts to implement programs proven in research on a larger scale in other domains have, at best, produced uneven results.

For example, there have been major shortcomings in achieving high fidelity with evidence-based substance abuse and violence prevention programs in community settings (Fagan et al., 2008). "Delivering interventions in a manner congruent with the theory, content, and methods of delivery specified by program developers is important, yet communities often fail to achieve implementation fidelity outside of efficacy trials" (p. 257). In schools, two national assessments found poor implementation for many delinquency and violence prevention programs that the schools attempted to adopt (Gottfredson and Gottfredson, 2002; Hallfors and Godette, 2002).

The mental health field has also been challenged to deliver effective, evidence-based programs dating back to Knitzer's (1982) call for a system of care (SOC). Knitzer and Cooper (2006) recently assessed progress in developing systems of care at the 20-year mark. Their assessment is that although system-level effects with SOCs have been good, individual outcomes have not. Notable system-level effects include reduced reliance on residential placements and hospitalizations and increased use of intensive community-based services. Yet "a recent study demonstrated consistent adherence to SOC principles in initiative sites but no improvements in children's outcomes and no advantage in improved outcomes compared with non-SOC sites using services that embodied similar principles" (Knitzer and Cooper, 2006, p. 671). These observers also note that evidence-based care implemented in community-based settings has produced less promising effects than were found in the supporting research studies. "Early data show that evidence-based treatments are being applied in the field with varying degrees of consistency and fidelity. Familiarity with empirically supported practices varies, but even where practitioners received on-the-job training, systematic implementation was not assured" (pp. 673–74).

C. Doubts about Whether Many Programs Used in Practice Are Actually Effective

A number of widely recognized prevention or intervention programs intended to reduce antisocial or illegal behavior have proven to be ineffective in well-designed studies.

A number of widely recognized prevention or intervention programs intended to reduce antisocial or illegal behavior have proven to be ineffective in well-designed studies. Although it is perhaps the most widely recognized of all delinquency prevention programs, the Drug Abuse Resistance Education program (D.A.R.E.) is not effective (Rosenbaum, 2007). D.A.R.E. is one of the most poignant examples of a program initially presumed to be effective that continued to be used despite strong empirical evidence to the contrary. More than 30 evaluations were made of D.A.R.E., yet it operated for 25 years before its negative results were accepted.

Prison visitation programs are another example of an initially attractive program that was later found to be ineffective. This approach was invented during the moral panic over juvenile delinquency in the late 1970s (Finckenauer and Gavin, 1999) when a group of inmates at New Jersey's Rahway State Prison, known as the Lifers' Group, created what later became known around the world as the Scared Straight program. Also known as "juvenile awareness," the program brought young minor offenders into the prison and subjected them to shock therapy consisting of threats, intimidation, and aggressive persuasion techniques. The idea was to literally scare them away from delinquency, to scare them straight. Many writers and producers for the print and broadcast media were enamored of it because of its simplicity and intuitive appeal. As many as 12,500 youth visited the Lifers each year. However, as Finckenauer and Gavin

(1999, pp. 85–93) reported, empirical evidence of the effectiveness of the Rahway State Prison Scared Straight program was lacking from the beginning. Evaluations of other Scared Straight–type programs were mixed but generally showed negative results (pp. 129–39; see also Petrosino, Turpin-Petrosino, and Finckenauer, 2000).

Curfew laws are another approach to reducing juvenile crime and victimization that have not been supported by empirical research (Adams, 2007). Adams' conclusion was based on at least a dozen research studies that include, for example, a national study that examined the effects of new curfew laws in 57 large cities (McDowall, Loftin, and Wiersema, 2000). That investigation found that the introduction of juvenile curfew laws was not followed by reductions in juvenile arrests in any serious crime category. The researchers noted that "any impacts of the laws were small, and they applied only to a few offenses" such as burglary, larceny, and simple assault (pp. 88–89).

Neither juvenile nor adult boot camps have proven to be effective according to a comprehensive meta-analysis.

Neither juvenile nor adult boot camps have proven to be effective according to a comprehensive meta-analysis (Wilson, MacKenzie, and Mitchell, 2005). When boot camps are designed as paramilitary regimens, research shows that boot camps and other forms of disciplinary programs increase recidivism by about 8 percent, on average (Lipsey, 2009). About the only positive thing that can be said about boot camps is that the inmates in them view their environment as being more therapeutic than traditional juvenile reformatories (MacKenzie, Wilson, Armstrong, and Gover, 2001), which may say more about the reformatories than the boot camps. Any advantage that boot camps confer, however, appears to be offset by the potential in boot camps for psychological, emotional, and physical abuse of youngsters—particularly for children with a history of abuse and family violence.

Moreover, research has not supported the effectiveness of large, congregate, custodial juvenile corrections facilities for rehabilitating juvenile offenders. Studies have shown that in large, typically overcrowded correctional facilities, both treatment opportunities and effectiveness of service delivery are diminished, and that larger facilities are more likely than smaller ones to be crowded (Snyder and Sickmund, 2006, p. 223). Large facilities with little treatment programming in states such as California and Texas have been accompanied by very high recidivism rates (Blackburn et al., 2007; Ezelle, 2007; Lattimore et al., 2004; Trulson et al., 2007). Custodial concerns tend to override concerns about the delivery of treatment services in these settings, and program quality suffers (Roush and McMillen, 2000).

Similarly, it has been found that the most restrictive out-of-home placements for mental health treatment, including psychiatric hospitalization and placement in residential treatment centers, are not effective for most child and adolescent offenders (Burns et al., 1999; Knitzer and Cooper, 2006; U.S. Department of Health and Human Services, 2001). Inpatient hospitalization is the least effective of all (U.S. Department of Health and Human Services, 2001, p. 171); indeed, it may do more harm than good in many cases (Weithorn, 1988).

These examples demonstrate the potential for a number of widely used programs, adopted with good intentions, to be ineffective for reducing subsequent delinquency and, more troubling, to actually be harmful—that is, to increase rather than decrease delinquency. None of the programs or approaches described above was subjected to rigorous evaluation research before it was implemented at scale. Their attractiveness was based on their intuitive appeal, not on credible evidence of effectiveness. Had such research been conducted and attended to by the respective decision makers, it is likely that at least some of these programs would never have been implemented or, at least, not implemented as widely. Limiting investment to programs and approaches shown to be effective by research in pilot and demonstration projects prior to implementation would not only avoid the often considerable waste of human and financial resources associated with supporting ineffective programs, but also reduce the potential for harm to the juveniles subjected to those programs.

It is recognition of this history in juvenile justice, and similar histories in other service areas, that has largely

driven the evidence-based practice movement—the idea that the effectiveness of the treatments, services, and programs provided to those in need should have been demonstrated in credible research prior to widespread use. Though this movement has received more lip service than action to date, it is notable in the juvenile justice field that at least six states—Florida, North Carolina, Oregon, Pennsylvania, Tennessee, and Washington—have adopted legislation requiring evidence-based programming.

...much of the research that is needed to support evidence-based practice in juvenile justice is already available.

As summarized in the previous section of this paper, there is an evidence base that identifies effective programs for juvenile offenders. Thus, much of the research that is needed to support evidence-based practice in juvenile justice is already available. As also discussed, however, the availability of research evidence is not in itself sufficient for taking the effective programs it identifies to scale in a way that retains their effectiveness when they are widely implemented in routine practice. This paper now turns to further consideration of evidence-based practice with a discussion of what constitutes evidence for that purpose and how to translate it into practice.

IV. Evidence-Based Practice: More Than One Approach

Three main approaches can be used to translate research evidence on effective programs into practice for everyday use by practitioners and policymakers. The first approach is direct evaluation of each individual program used in practice to confirm its effectiveness and, if it is found ineffective, to use that evidence to improve or terminate it. A second is to implement with fidelity a program from a list of model programs certified by an authoritative source as having acceptable evidence of effectiveness. A third approach is to implement a type of program that has been shown to be effective on average by a meta-analysis of many studies of that program type, but to do so in the manner that the research indicates will yield that average effect or better.

A. Direct Evaluation of the Effects of the Program as Implemented

The form of evidence about effectiveness that is most specific to a program as it is actually practiced is an impact evaluation of that specific program conducted in situ. With accompanying process information about the nature of the services delivered and of the juveniles receiving those services, such research can assess program effects on selected outcomes in relation to the way the program is implemented. To provide the most valid results, an impact evaluation must use a control group of comparable juveniles who do not receive the program, preferably assigned randomly to program and no program conditions.

When well executed, an impact evaluation of this sort will provide a credible assessment of the effects of a program for juvenile offenders on their recidivism and any other measured outcomes of interest. The process component, in turn, will help identify implementation problems that may need to be corrected to improve the effects and subgroups of offenders who respond more or less positively to the program. A fully developed impact evaluation, therefore, not only assesses the program effects on the intended outcome but provides the basis for diagnosing any shortfalls in those effects as well as guidance for program improvement.

The main advantage of direct evaluation is that the results apply in a very specific manner to that particular program as practiced.

The main advantage of direct evaluation is that the results apply in a very specific manner to that particular program as practiced. The evidence base supporting its effectiveness, therefore, is not drawn from studies done elsewhere and extrapolated to that particular program with the associated questions about how well that evidence applies. The disadvantages of direct evaluation as an approach to evidence-based practice, however, are considerable. First, impact evaluation requires resources, technical expertise, and favorable conditions with regard to the ability to create a control group and collect the desired process and outcome measures. For a juvenile justice system that uses many programs, mounting a credible impact evaluation of each would be prohibitively expensive. In addition, the specificity of the evaluation can also be a limitation. If the program changes in any significant way, e.g., through program improvement efforts or natural changes in its clientele, the results of the evaluation may no longer apply. Providing evidence that the altered program is still effective then requires a new impact evaluation.

As a practical matter, direct impact evaluations are generally conducted only for innovative or relatively unique programs that have not already been tested, or to replicate the findings of such evaluations when those programs are applied in different circumstances. Indeed, it is research of this sort that provides the evidence for model program

assessment or meta-analysis. A juvenile justice system would not likely undertake impact evaluations for all the programs it makes use of, but it might do so for a promising, innovative "home-grown" program as part of a commitment to evidence-based practice. It would also often be wise to conduct an independent evaluation on a model program when it is first implemented in a particular jurisdiction to be sure that it is as effective in the local circumstances as it was where the original research on it was conducted (see Barnoski, 2002 and 2004a, for example).

B. Model Programs with Evidence Certified by a Credible Source

The model programs approach entails selecting a recommended program from a list of research-supported programs and implementing it locally with fidelity to the program developer's specifications for how the program is to be delivered. In this approach, the recommended programs, typically called "model" or "exemplary" programs, are identified through a process of program-by-program reviews of the research. The programs typically considered for such reviews are specific brand-name programs that can be separately identified in the research literature, e.g., Functional Family Therapy, Aggression Replacement Therapy, and Multidimensional Treatment Foster Care. Well-known examples of model program catalogues relevant to juvenile justice include the University of Colorado Blueprints for Violence Prevention project (Mihalic et al., 2001) and the Office of Juvenile Justice and Delinquency Prevention's Model Programs Guide (http://www2.dsgonline.com/mpg/).

The research that supports the effectiveness of model programs is specific to the individual programs in the form in which they were delivered in that research. To implement such a program, therefore, means following the same protocol that defined the program in those research studies. This is typically provided in a program manual or similar written materials and may be accompanied by training materials, instruments for assessing the fidelity of implementation, and the like. In addition, an organizational infrastructure has been developed for some model programs to provide training, technical assistance, and materials to agencies that wish to implement the program. Thus practitioners and policymakers who want to adopt one of these programs typically are able to obtain descriptive information about how the program is supposed to be delivered and may also be able to access support for implementation from the developer or an associated organization.

As an approach to evidence-based practice, the "evidence" part of the model programs strategy consists of evaluation studies judged credible by some set of designated reviewers that demonstrate that a particular program had positive effects in the circumstances represented in those studies. Typically this demonstration is based on only a few studies at best, often only one or two. The fact that the program was shown to be effective in those instances nonetheless establishes that it has the potential to produce good effects, and that in itself distinguishes it from most of the programs currently in use with juvenile offenders. In order to have some assurance that those effects can be replicated in local applications, the program must be implemented the same way as was done in the research studies, that is, with fidelity to the program protocol and with similar juvenile participants. When local implementations depart from that protocol and target population, they also depart from the supportive evidence and thus diminish the expectation that outcomes comparable to those found in the research will follow.

The major advantages of the model program approach are the assurance from prior research that the identified programs have the potential to be effective and the availability of protocols to be followed to replicate the effects found in the research.

The major advantages of the model program approach are the assurance from prior research that the identified programs have the potential to be effective and the availability of protocols to be followed to replicate the effects found in the research. The disadvantages largely revolve around the requirement for strict adherence to the prescribed protocol. Obtaining licenses and training to properly implement the brand-name model programs

that are most available may be costly compared to local programs that are viewed as effective but which lack supporting research evidence. Local providers may also find it difficult to modify or abandon their established practices to adopt a model program "by the book," and they often resist or make their own adaptations to the program with the associated compromises to fidelity.

More generally, however, there are many challenging issues associated with translating an evidence-based program into routine practice in a way that closely replicates the relevant circumstances of the original research. As a result, the desirable program effects on delinquency and subsequent offending found in the research studies often are attenuated when those programs are scaled up for general application (Dodge, 2001; Karoly et al., 1998; Welsh, Sullivan, and Olds, 2010). There are numerous ways this can happen. First, as a practical matter, it may not be possible to restrict the scaled-up program to the same population represented in the research. In real-world settings, the program is likely to serve a more heterogeneous population than was used in the research studies. In addition, the service infrastructure for delivering the program is likely to be weaker than that organized by the program developer when conducting the evaluation research. It seldom is the case that sufficient resources—from trained service providers to public funds for personnel and capital expenditures—are available in everyday practice settings to fully meet the requirements of a model program when it is rolled out at scale. The expansion of a program beyond the relatively controlled circumstances of the research trials and the close supervision of the program developer make it challenging to maintain the critical program features that underlie its success.

To provide assurance that model programs implemented locally are, in fact, effective as delivered in that context, the best strategy is to adopt them first on a pilot basis and evaluate their effects before expanding them. Such evaluations should closely examine fidelity to the program protocol as well as outcomes. This was the approach taken by the Washington State Institute for Public Policy when several research-based programs were implemented in Washington State. The results showed that these programs did produce positive effects when implemented locally under real-world circumstances, but only when the programs were competently delivered in accordance with the developers' specifications (Barnoski, 2002, 2004a).

C. Best Practice Guidelines Based on a Meta-analysis of Research Findings

The model program approach to evidence-based practice focuses on distinct individual named programs and the research specific to each of them. Virtually all the prevention and intervention programs used with youth, however, also fall into more generic categories that distinguish the different types of programs. For example, Functional Family Therapy (FFT) for juvenile offenders is an instance of a broader program type generally referred to as family therapy. Similarly, Aggression Replacement Training is an instance of the cognitive-behavioral therapy type of program for offenders. Other widely recognized generic program types include mentoring, social skills training, behavior management, individual counseling, group counseling, parent training, and the like. These generic program categories are not restricted to named programs; they also contain home-grown programs that have not been packaged for broader dissemination but are nonetheless recognizable instances of one of the common program types.

When we turn to the research on the effectiveness of a particular type of program, such as family therapy, we often find many studies. So, although the evidence base specifically for, say, FFT may consist of only a few studies, dozens of studies have been done on family therapy programs with juvenile offenders. The FFT studies are there, as are the studies of other brand-name family therapy programs, but there are also many studies of home-grown family therapy programs that are fundamentally similar to the brand-name programs, though varying in the particulars. This larger body of evidence about the effectiveness of a program type not only involves more variation in the program particulars, but also in the program setting, personnel, and characteristics of the juveniles served. When that evidence shows generally positive effects, therefore, that is a finding that

is in many ways more robust than the findings of the few studies supporting one model program—that is, it is less specific to the small set of distinct circumstances in which the program was tested.

On the other hand, because there is a broad range of programs within a type, there is also more variability in the findings of the research on those programs. Some family therapy programs studied in some circumstances show much larger effects than others. Indeed, some studies show no effects or even slightly negative effects. Though the average effect might be positive, there are both more effective and much less effective programs distributed around that average. If we are to use that evidence to guide practice, we need to know which characteristics distinguish the more effective programs. If the research shows a systematic pattern that allows those characteristics to be identified, we can use that information to construct best practice guidelines that describe the characteristics of the most effective versions of the programs of a given type. Thus we might discover that family therapy programs are effective on average, but the ones that produce better than average effects are characterized by, say, a certain number of contact hours with the family and periodic individual sessions with the juvenile.

The technique for extracting and analyzing information about intervention effects and the characteristics of the interventions producing those effects from a body of qualifying research is called meta-analysis (Borenstein et al., 2009; Lipsey and Wilson, 2001). In a meta-analysis, a database is developed by trained coders using a structured coding protocol to extract information from eligible study reports. Studies are eligible for inclusion based on explicit criteria and are collected through an extensive literature search. In the case of program evaluations, the key data elements are statistical estimates of the treatment effects, known as effect sizes. Effect sizes represent the magnitude of the difference between the mean value on the outcome variable (e.g., recidivism) for the individuals receiving intervention and that for a comparable group not receiving the intervention. Effect sizes are standardized in a way that makes them comparable across studies. Meta-analysis techniques are well established and widely used to provide systematic syntheses of intervention research in education, social welfare, public health, and medicine as well as in juvenile and criminal justice (Cooper, Hedges, and Valentine, 2009).

Some of the model program lists—the OJJDP Model Programs Guide, for example—include generic program types mixed in with brand-name programs when one or more meta-analysis has shown average positive effects. Typically, however, these lists do not include further differentiation of the characteristics that distinguish the most and least effective programs of that type. Meta-analysis has more capability than has been utilized in these applications to not only identify program types with generally positive evidence of effectiveness, but also to identify the characteristics of the programs of that type shown by the evidence to be the most effective. Systematic compilation of that information allows practitioners and policymakers to determine which program types are supported by credible evidence. Such information can also be used to generate best practice guides for implementing or improving programs of those types in ways that align with the most effective variants of those program types. In this manner, evidence-based practice can be extended beyond brand-name model programs to those many local and home-grown programs that are more generic instances of program types whose effectiveness is adequately supported by research.

Systematic compilation of that information allows practitioners and policymakers to determine which program types are supported by credible evidence.

We will shortly present a more detailed description of one particular set of best practice guides for juvenile delinquency programs that are based on meta-analysis and illustrate their use. But first we must provide a fuller account of what has been learned about such programs from meta-analysis to make clear the source of the information for those best practice guides.

V. Meta-analysis of Research on the Effects of Intervention Programs for Juvenile Offenders

Dozens of meta-analyses have been conducted on evaluations of the effects of programs on the recidivism of juvenile offenders (Lipsey and Cullen, 2007). Almost all of these, however, have had a somewhat limited scope. They have focused on one type of program or program area (e.g., boot camps, cognitive-behavioral therapy, behavioral programs), or one type of offender (e.g., sex offenders), or a single named program (e.g., Multisystemic Therapy). The results of this work have been very informative for the respective topic areas and have generally confirmed the effectiveness of rehabilitative treatments for offenders. Nonetheless, it is difficult to piece such meta-analyses together into an overall picture of current knowledge about the nature of the most effective programs.

Rather than focusing on a predefined kind of program or offender, an alternate approach is to collect and meta-analyze all the available research on the effects of intervention with juvenile offenders, sorting it according to the types of interventions found, whatever they may be. Though a daunting task, this approach makes it possible to investigate certain important issues that are otherwise difficult to address. Examination of the full body of research on delinquency programs in a single meta-analysis, for instance, allows an integrated analysis of the comparative effectiveness of different program types and approaches. A meta-analysis of, say, cognitive-behavioral programs may demonstrate that they have positive effects on recidivism while another meta-analysis shows that family counseling also has positive effects. But which programs are most effective and for whom and under what circumstances? Answers to those questions are especially critical for practitioners interested in using the most effective programs applicable to their situations.

Such comparative assessments are not easy to make across different meta-analyses. The task is not as simple as determining which ones show the largest mean effect size. Effect sizes are influenced by variation in the subject samples and settings used in the primary studies, by the research methods applied in those studies, and by the procedures employed by the meta-analyst in representing and analyzing the intervention effects. Under these circumstances, simple comparisons of summary effect sizes can be misleading. Within an integrated meta-analysis, however, common procedures can be applied and statistical controls used to help level the playing field in a uniform manner so that comparative effectiveness can be better assessed.

Effect sizes are influenced by variation in the subject samples and settings used in the primary studies, by the research methods applied in those studies, and by the procedures employed by the meta-analyst in representing and analyzing the intervention effects.

Another advantage of a comprehensive meta-analysis of programs is the opportunity it provides to search for generalizations about the factors associated with effective programs. Useful guidance for practitioners does not come solely from lists of the programs and types of programs shown by research to have positive effects. It also comes from identification of the factors that distinguish the most effective programs and the general principles that characterize “what works” to reduce the recidivism of juvenile offenders and improve other outcomes.

Only one attempt has been made to conduct a comprehensive meta-analysis of the findings of all the available research on the effects of interventions with juvenile offenders. This was an effort begun by Mark Lipsey in the mid-1980s and continued, with periodic updates, to the present day. The results of this program of meta-analysis research have been reported in numerous publications over the years (e.g., Lipsey, 1992; Lipsey

and Wilson, 1998; Lipsey, 1999a, 1999b). The most recent analysis (Lipsey, 2009) is the most comprehensive in terms of both the number of studies included in the database and the scope of the factors investigated. We turn now to a summary of the findings of that analysis.

A. Analysis of the Findings of 548 Evaluation Studies of Delinquency Interventions

The database for Lipsey's (2009) comprehensive meta-analysis of the effects of delinquency interventions consisted of 548 studies that spanned the period from 1958 through 2002. These studies represented all the intervention research that could be located through an extensive search for published and unpublished reports of research that met the following key criteria:

1. The research was conducted in an English-speaking country and reported in English.
2. The juveniles studied were between 12 and 21 years of age.
3. The program's effect was measured on at least one delinquency outcome variable (e.g., rearrest, reconviction, return to court supervision, and so forth).
4. The outcomes of the target intervention program were directly compared to those of a control group of similar juveniles who did not receive the intervention.

Trained coders read each study that met these criteria and, using a computerized coding scheme, extracted information that described each study on a large set of variables divided into the following categories:

- Characteristics of the study methods
- Characteristics of the juvenile samples
- Level of juvenile justice supervision and control (e.g., diversion, probation, incarceration)
- Type of intervention or program applied
- Amount and quality of service
- Statistical effect size for the magnitude of the intervention effect on subsequent offending

The key variable in this analysis, of course, is the effect size; it indicates whether the study found that the intervention reduced subsequent offending and by how much. Across all 548 studies, the mean intervention effect was positive (reduced recidivism) and statistically significant. The magnitude of this effect was modest but not trivial, representing a one-year rearrest rate about six percentage points lower for the treated juveniles relative to the control juveniles. This overall average tells us very little about the effectiveness of the interventions, however, because there was enormous variability in the observed effects across the studies. Some of the effect sizes were very small, virtually zero, and even negative, while others were quite large. The 75th percentile effect size, for instance, represented a reduction of about 24 percentage points in the reoffense rate while the 90th percentile effect size represented a reduction of more than 40 percentage points. The most important question, then, is what are the programs or program characteristics that produce the large effects.

One set of variables related to the magnitude of the intervention effects is that representing the methodological characteristics of the studies, for example, the way recidivism was measured and the quality of the design for creating comparable treatment and control groups. In order to minimize any confusion in the analysis between the influence of these differences and those of the substantive characteristics of interest, the methodological variables were statistically controlled in all analyses. Further analysis was then done to isolate as much as possible the relationships between the recidivism effects and the characteristics of the intervention programs and the juveniles to whom they were applied. The details of these analyses and the methods used are described in detail in Lipsey (2009).

B. Program Characteristics Associated with the Greatest Effects on Recidivism

Given the overall finding in Lipsey's (2009) meta-analysis that some interventions show relatively large positive effects on the juveniles that participate, what do these successful programs look like? That was the

key question for this meta-analysis. Four programmatic aspects were found to be most relevant when considering what works best for reducing subsequent offense rates. These programmatic components and characteristics are described below.

1. Risk Level of the Juveniles

It is possible that some juveniles are generally more responsive to intervention programs than others and thus show larger effects across a wide range of program types. The analysis showed that there was little overall difference in effects associated with the demographic characteristics of age, gender, and ethnicity. The one characteristic of the juveniles receiving the interventions that did show an overall relationship was risk for delinquency as indexed by the nature and extent of prior offenses and the reoffense rates of the matched untreated controls. Interventions applied to high-risk delinquents, on average, produced larger recidivism reductions than when those interventions were applied to low-risk delinquents.

This finding can be understood rather easily in terms of the latitude for improvement among different risk groups. High-risk juveniles by definition are likely to have high reoffense rates and thus have the most room for improvement if they receive an effective intervention. Low-risk juveniles, on the other hand, have little likelihood of reoffending even without intervention and thus have little room for improvement. Especially notable is that this relationship with risk extended to the very highest risk samples found among the research studies—that is, there was no indication that there were juveniles whose risk level was so high that they did not respond to effective interventions.

The juvenile justice supervision status of the juveniles in these studies (e.g., diversion, probation or community supervision, incarceration) was, of course, strongly related to their risk characteristics. The juveniles under higher levels of supervision tended to be the higher risk cases. When risk level was statistically controlled, however, no difference was found in the overall effectiveness of the intervention programs in the different supervision categories. Thus a type of program that was effective for juveniles under probation or parole supervision in the community was equally effective when applied to juveniles in secure residential facilities once the fact that effects are generally larger for higher risk juveniles is taken into account.

In practical terms, juvenile justice systems will generally get more delinquency reduction benefits from their intervention dollars by focusing their most effective and costly interventions on higher risk juveniles and providing less intensive and costly interventions to the lower risk cases.

In practical terms, juvenile justice systems will generally get more delinquency reduction benefits from their intervention dollars by focusing their most effective and costly interventions on higher risk juveniles and providing less intensive and costly interventions to the lower risk cases. Moreover, they can expect similar benefits from their intervention programs for juveniles at a given risk level whether they are treated and supervised in the community or in residential facilities.

2. Therapeutic versus Control Treatment Philosophies

Not surprisingly, the meta-analysis found that the type of program was rather strongly related to its effects on reoffense rates—some programs are simply more effective than others, all else being equal. Given the great diversity of program types that appear in the research, however, describing and categorizing them in meaningful ways is a challenge. Lipsey (2009) found that one important distinction had to do with the overarching philosophy of the program. “Philosophy” in this context means the global approach to altering juvenile behavior taken by the program. From this perspective, two broad program philosophies could be distinguished. The first featured external control techniques for suppressing delinquency and included three categories:

- Programs oriented toward instilling discipline (e.g., paramilitary regimens in boot camps)

- Programs aimed at deterrence through fear of the consequences of bad behavior (e.g., prison visitation programs such as Scared Straight)
- Programs emphasizing surveillance to detect bad behavior (e.g., intensive probation or parole supervision)

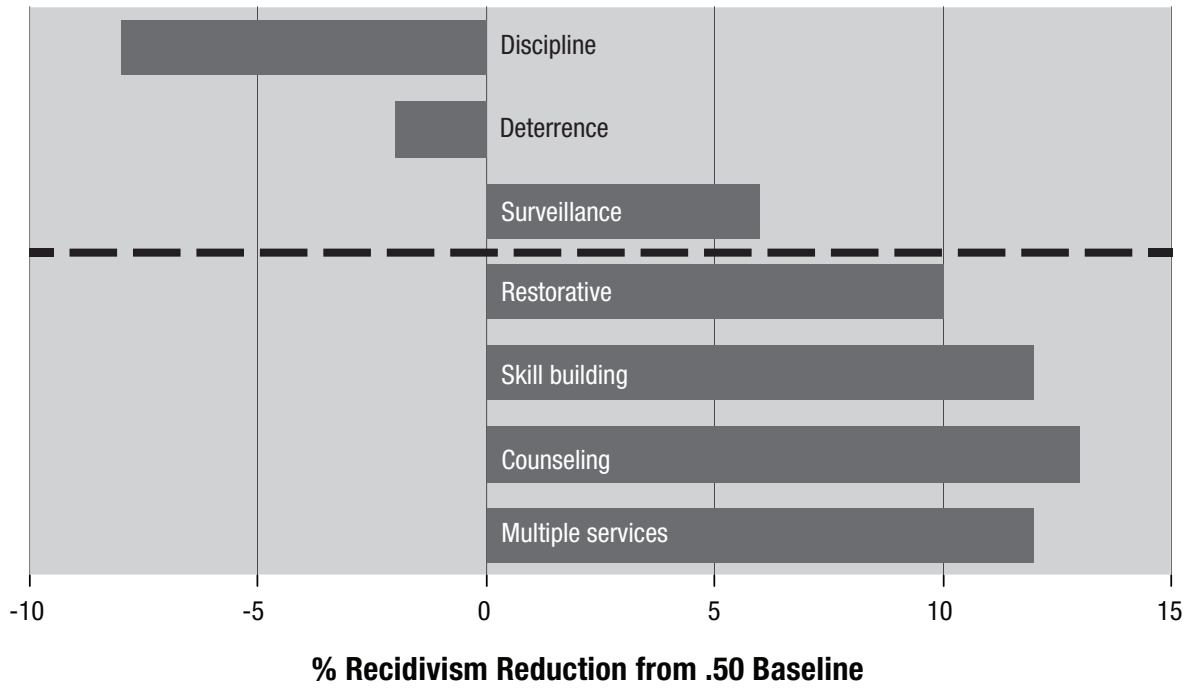
A contrasting philosophy involves attempts to bring about behavior change by facilitating personal development through improved skills, relationships, insight, and the like. This therapeutic philosophy included the following categories of programs:

- Restorative (e.g., restitution, victim-offender mediation)
- Skill building (e.g., cognitive-behavioral techniques, social skills, academic and vocational skill building)
- Counseling (e.g., individual, group, family; mentoring)
- Multiple coordinated services (e.g., case management and service brokering)

When the mean effects on reoffense rates were compared for the programs associated with these two broad approaches, the programs with a therapeutic philosophy were notably more effective than those with a control philosophy. Figure 1 shows the effects for the program categories within each of these philosophies. The zero (0) point indicates no program effect while positive values represent reductions in recidivism and negative values represent increases in recidivism. As can be seen, the programs in two of the control categories on average had negative effects.[3] The third category, programs relying mainly on surveillance, showed positive effects, but smaller ones than for any of the therapeutic program categories. This category includes mainly intensive probation programs, which often have significant counseling components by probation officers. They may thus represent a mix of control and therapeutic strategies.

[3] *All estimates of the mean reoffense effect sizes have been adjusted for methodological differences between the studies.*

Figure 1. Mean recidivism effects for the program categories representing control and therapeutic philosophies



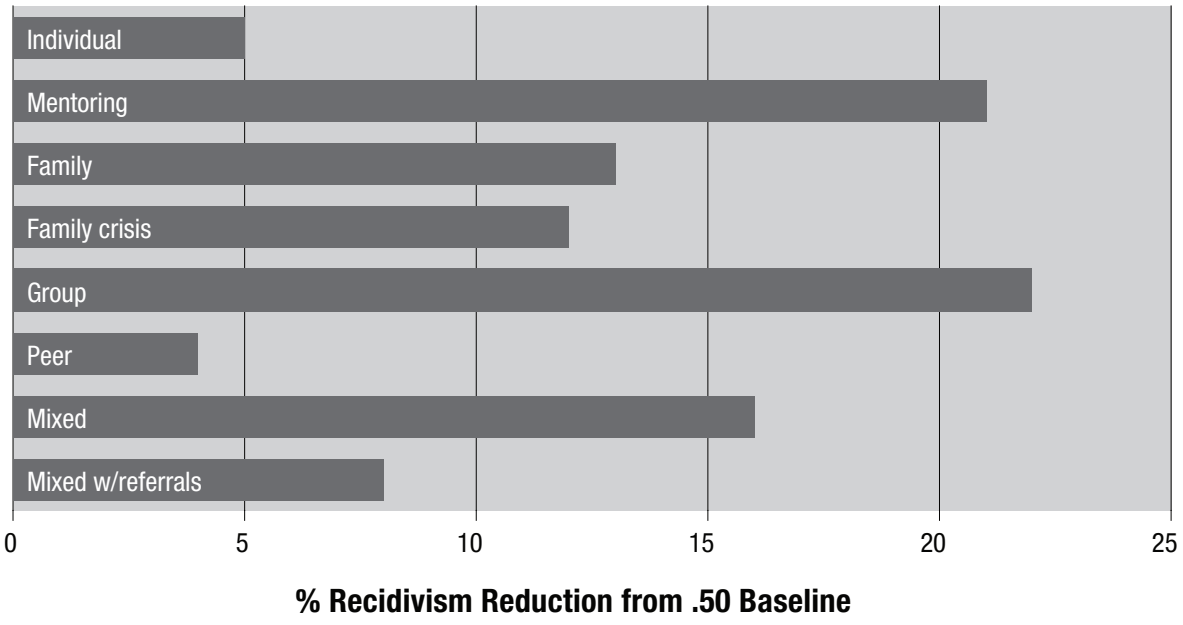
For purposes of guiding juvenile justice systems toward effective programs, the advice that follows from this portion of the meta-analysis is straightforward. To optimize the effects on recidivism and other outcomes, programs from the therapeutic categories should be favored and those from the control categories should be avoided as much as possible.

3. Generic Program Types and Embedded Model Programs

Within each of the program categories identified above as representing the control and therapeutic philosophies, programs were classified into subcategories according to their generic program type. For example, in the subcategory of counseling programs within the therapeutic philosophy, different kinds of counseling can be distinguished that vary in their effects on reoffense rates. Figure 2 shows the mean effects for the major generic types of counseling. Though they all show positive effects, the largest effects appeared for group counseling and mentoring programs. Similar variation across the generic program types was seen in the other therapeutic program categories. Behind counseling, the next largest category was skill-building programs (figure 1). Figure 3 shows that all the program types in the skill-building category also had positive effects, but behavioral programs (e.g., behavior contracting) and cognitive-behavioral programs had the largest mean effects.

Embedded within many of these generic program types are specific brand-name model programs that have been included in the evaluation research covered in the meta-analysis. These generally show positive effects on recidivism, as we would expect. However, they do not necessarily show notably better effects than the no-name programs of the same type. For example, Functional Family Therapy (FFT) and Multisystemic Therapy (MST) are both included in the generic program type labeled "family counseling." Figure 4 shows the distribution of statistical effect sizes found in evaluation studies of 29 family counseling programs. The larger effect sizes on the right-hand side of this distribution describe the most positive

Figure 2. Mean recidivism effects for the generic program types within the counseling category



effects on recidivism, that is, recidivism reductions. The effect sizes for recidivism outcomes found in the studies specifically of FFT and MST respectively are color coded and labeled.

As can be seen in Figure 4, most of the programs of the family counseling type had positive effects on recidivism. It is also the case that the four studies of FFT and the four studies of MST showed positive effects. The effects for those model programs, however, show variation, with some larger and some smaller, just as the other family counseling programs do, including the no-name ones. Moreover, the effect size estimates from the FFT and MST studies fall well within the range of the other family programs in this collection. Indeed, some no-name programs produced effects even larger than those found for the model programs.

In this example, we see that the model programs are indeed effective, and thus deserve their designation as evidence-based programs. At the same time, there is evidence for the effectiveness of family counseling programs as a generic type, so it is not unreasonable to say that family counseling programs are also evidence based. However, some of the studies of family counseling programs showed near zero or even negative effects, so a careful specification of the family programs that are evidence-based would also include whatever characteristics distinguish those on the high end of the effect distribution.

This portion of the meta-analysis has important implications for juvenile justice practice. First, the selection of the type of program (family counseling, social skills, mentoring, and so forth) is consequential. As long as the program type matches the needs of an offender, the largest potential effects on recidivism can be expected from the program types that showed the largest average effects in the research studies. Second, when a specific program of any given type is being selected, a model program should generally be a good choice, provided that one is available and can be implemented with fidelity. A third implication, however, is that local programs of that same type would also be expected to be effective if they are implemented in an appropriate manner. Effective implementation in these cases means ensuring that the program has the distinguishing characteristics of similar programs found in the research to have above average effects and that it avoids the characteristics of those found to have negligible or negative effects. Other than

Figure 3. Mean recidivism effects for the generic program types within the skill-building category

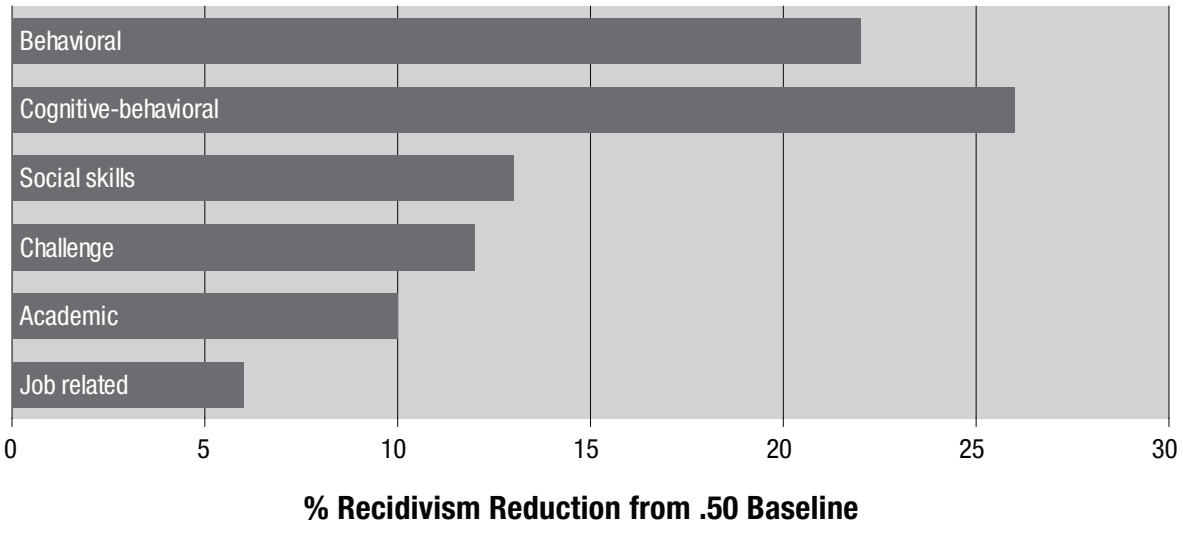
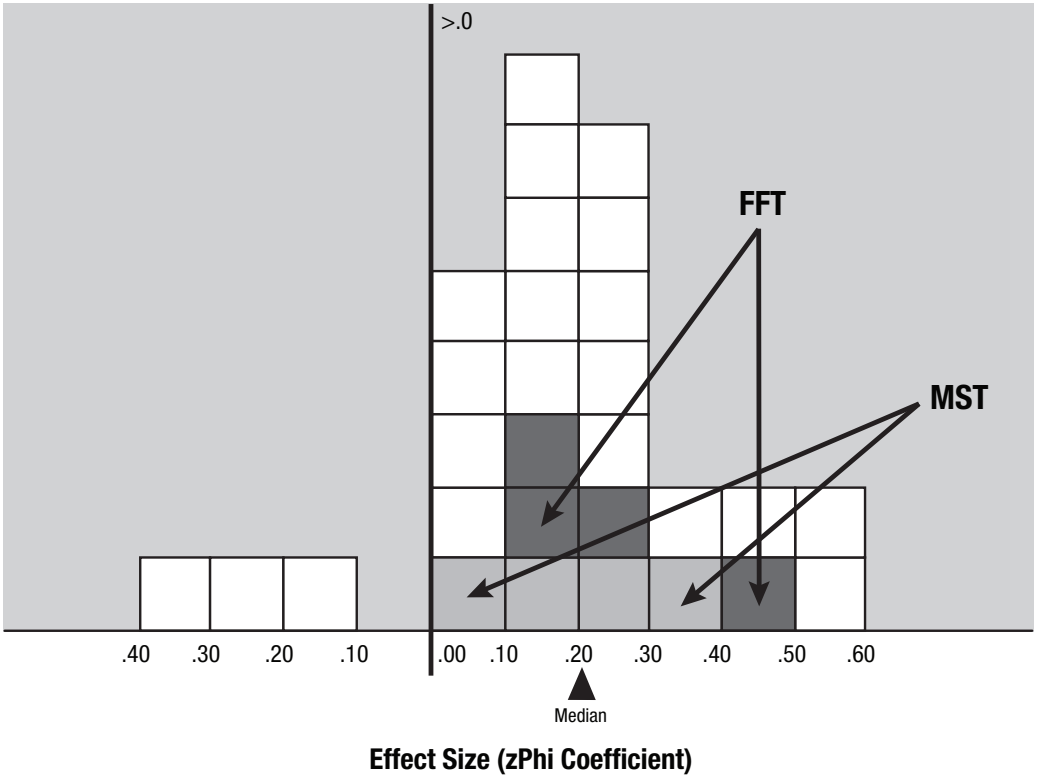


Figure 4. Effect sizes for family counseling interventions with those for FFT and MST identified



the risk level of the juveniles served, these distinguishing characteristics have mainly to do with the amount and quality of service provided, as described below.

4. Amount and Quality of Service

The final factor that was related to the magnitude of intervention effects on reoffense rates dealt with the way the program was implemented. The variables represented in that factor are simple, but important ones. First, a sufficient amount of the program service must be provided. This aspect is like the dose of a medicine—if the amount taken is too small, it is unlikely to have the expected effect. Of course, beyond a certain point, a larger dose does not necessarily improve the outcome. For each program type, recidivism reductions were associated with the duration of the service (days from start to termination) and total contact hours of service the juvenile received. To obtain at least the average effect on recidivism for that program type, the program duration and hours of contact must at least reach the average values for the programs of that type included in the meta-analysis.

Second, the quality of the program implementation was an important feature related to the magnitude of the effects. This aspect was not well reported in the research studies providing data for the meta-analysis, but indications of problems such as high dropout rates, staff turnover, poorly trained personnel, incomplete service delivery, and the like were associated with smaller effects. Also, when the program developer was involved in the delivery of the

program, and thus had a direct role in ensuring that it was delivered appropriately, the effects were larger. We conclude, therefore, that fidelity—that is, delivery of the program as intended to all recipients—is one of the keys to program success, as indeed we might expect.

C. Effective Juvenile Justice Programs: Implications for Practice

The meta-analysis found no factors other than those described above that were significantly associated with the intervention effects on reoffense rates. Furthermore, those factors, taken together, had a large enough relationship with the program outcomes to account for a substantial portion of the differences across programs in the magnitude of their effects on recidivism. No doubt there are many other features that relate to the success of particular programs implemented in particular circumstances that were not captured in this analysis. This small set of rather broad factors, however, goes a long way toward distinguishing the programs shown in the research studies to produce large enough effects on recidivism to have practical value in juvenile justice applications from those with negligible or even negative effects.

It is a fortunate finding that so much of the effectiveness of these programs can be accounted for by such a small number of quite straightforward factors. This means that close attention to these factors in the selection and implementation of programs for juvenile offenders can provide reasonable assurance that those programs will be effective for reducing recidivism. Moreover, the extensive research represented in the 548 controlled studies in the meta-analysis makes these factors evidence based. Thus one of the several defensible definitions of evidence-based programs is that they match the profile on this set of factors that the meta-analysis has shown to be characteristic of the most effective programs. That profile, to summarize, prescribes the following:

- Target high-risk cases. In particular, provide the most effective programs possible to the highest risk cases. Effective programs applied to low-risk cases will have small effects so it is not cost-effective to provide more than minimal, low-cost services to such cases.
- Use programs that take a therapeutic approach to changing behavior by focusing on constructive personal development. Minimize programs based on a control or deterrence philosophy.
- Favor those program types that have shown the largest effects in research studies when matching programs to the needs and problem areas of the juveniles served.
- Implement the selected programs well. Monitor each program to ensure that it is delivered as intended and that all the juveniles assigned to it receive at least an amount of service that corresponds to the average reported in the evaluation research on that type of program.

What the meta-analysis results tell us is that programs that more closely match this profile should be more effective.

What the meta-analysis results tell us is that programs that more closely match this profile should be more effective. This profile thus constitutes a simple set of best practice guidelines for juvenile justice programs. It also provides a basis for evaluating existing programs. Those programs that most closely match this profile are better programs in the sense that, on the basis of the available research, we expect them to have better effects. Those that fall short are not expected to be as effective, but the parts of the profile on which they fall short can be used to guide their improvement.

One approach to making the guidance for selecting, implementing, and improving juvenile justice programs more useful at a practical level is to incorporate it in an instrument that allows each local program to be rated according to how closely it matches the best practice profile derived from the meta-analysis. Such an instrument has been developed, field tested, and validated. It is described in the next section.

VI. The SPEP: Evidence-Based Practice Guidelines

The juvenile justice field needs a more efficient and holistic way to use the tremendous body of research now available to inform program practice. The extent of that research is sufficient to allow nearly the entire spectrum of juvenile justice programs to operate on an evidence-based platform. Although brand-name model programs may be implemented as part of that platform, local programs may also be supported by evidence of effectiveness, or may be enhanced in ways that align them with that evidence. Not all programs that are practiced locally may be of a type for which there is a research base, however. If they are to be used as part of an evidence-based platform, those programs must be separately evaluated with research conducted directly on them.

To translate the guidelines for effective programs that are derived from his meta-analysis into practical form, Lipsey developed the Standardized Program Evaluation Protocol (SPEP), a tool for comparing local juvenile justice programs to what has been found to be effective in the research. This was done in a straightforward manner. Each of the factors found in the meta-analysis to be importantly related to program effectiveness is represented in the SPEP and associated with a certain maximum number of points to provide a score. The number of points associated with each factor is derived directly from the statistical models used in the meta-analysis to predict program effects on recidivism. Those factors with stronger predictive relationships are assigned proportionately more points than those with relationships that are not as strong. Where appropriate, target values are set based on the median values found in the corresponding research, e.g., for service duration and number of contact hours.

The ratings on each factor in this scheme for a particular local program are derived empirically—they are not simply reflections of someone's judgment on what the program is doing. For instance, ratings of the risk level of the juveniles served must be based on the results of a valid risk assessment instrument or equivalent data. Ratings of the amount of service must be based on management information system data that report the service received by each juvenile. The SPEP instrument and the basis for the ratings it involves are described more fully in the next section.

A. The Standardized Program Evaluation Protocol for Assessing Juvenile Justice Programs

The Standardized Program Evaluation Protocol is a tool for comparing juvenile justice programs to what has been found to be effective in the research. More specifically, the SPEP creates a metric by assigning points to programs according to how closely their characteristics match those associated with the best recidivism outcomes for similar programs as identified in Lipsey's large (2009) meta-analysis of evaluation studies (described in the previous section). Although the SPEP is focused on recidivism, the programs found in the meta-analysis to be effective for reducing recidivism also had positive effects on other outcomes such as family and peer relations, mental health symptoms, and school attendance.

The SPEP is configured so that the maximum overall score is 100 points. Each of the ratings on the key effectiveness factors represented in the SPEP has a maximum value assigned in proportion to the strength of that factor for predicting recidivism effects in the statistical models used in the meta-analysis. Thus the maximum rating possible for the primary service type is larger than that for the risk level of the juveniles because the meta-analysis showed that, though both factors were independently related to recidivism effects, the primary service type was more strongly related. The key factors associated with program

effectiveness found in the meta-analysis and the basis for rating them in the SPEP are as follows.

Type of program. The SPEP covers only program types that take a therapeutic approach, as defined in the program categories used in the meta-analysis (e.g., family counseling, mentoring, cognitive-behavioral therapy, vocational training). The relative effectiveness of each program type for reducing recidivism that was found in the statistical analysis was used to categorize program types as having, on average, high, medium, or low effects on recidivism, keeping in mind that even the low program types nonetheless have positive average effects. The total number of points—which represents the proportionate contribution of program type to predicting recidivism effects—is distributed across these categories so that the maximum number of program type points goes to those in the high category with discounted scores given to program types in the medium and low categories.

To determine which program type a local program represents, and thus what its SPEP score is on that factor, descriptive information about the nature of the services it provides must be examined. That information is compared with the descriptions in a glossary of program types that was developed from the descriptions provided in the corresponding research studies included in the meta-analysis. The local program is then identified with regard to the program type it represents and, depending on whether that program type is classified as having low, medium, or high effectiveness, the corresponding SPEP rating is assigned. If a program does not match any of the program types in the glossary, it means that insufficient research exists for estimating the effectiveness of that type of program.

Many programs involve combinations of services that may represent different program types. In those cases, primary and supplementary services are distinguished and, if the supplementary services are of a different type from the primary service, but of a type shown to be effective in the research, bonus points are awarded for it.

Amount of treatment. Service amount is divided into duration and total contact hours, with the latter receiving somewhat more points in light of its slightly stronger relationship to outcomes. Service duration is assessed as the time (e.g., number of weeks) between the date of service intake and the date of service termination for each juvenile with a closed case who was served by the program over the period of time to which the SPEP is applied (e.g., SPEP ratings might be made annually). Similarly, total contact hours are assessed as the number of hours of direct exposure each juvenile had to substantive program activities. In both cases, these values must be determined from actual service records, not estimated subjectively.

The SPEP ratings for these service dimensions assign a greater or lesser proportion of the points available for amount of service according to the proportion of the juveniles served with service duration or contact hours that reach or exceed specified target values. Those target values are set at the average found in the corresponding research studies for programs of that type. This is based on the assumption that, if the amount of service provided at least reaches the average reported in the respective research studies, the program should attain at least the average effects on recidivism found for that program type.

Quality of treatment. The quality of the treatment implementation is the most difficult SPEP factor to rate on the basis of actual program data. This factor, as it is represented in the research studies and analyzed in the meta-analysis, refers to the extent to which the program was implemented as intended for every juvenile recipient. Such information is not generally collected as part of the management information or client-tracking systems used by juvenile justice agencies and may have to be developed in order to support full SPEP ratings. Drawing on the representation of this factor in the research studies, we identify the key dimensions of implementation quality as (1) a written protocol describing the intended service, (2) provision of training on the intended service for those delivering it, (3) a regular procedure for monitoring service to assess whether it is being delivered as intended, and (4) a procedure for taking corrective action when service delivery strays from what is intended. Note that these are not dimensions of clinical quality, which may be important but are not captured well in the research on which the SPEP is based. Rather, these are organizational matters that can be assessed in terms of the operating procedures established and maintained by the provider delivering the program being rated.

Youth risk level. Risk level is assessed for each youth based on a valid risk assessment instrument or equivalent data, e.g., prior offense history and current problem behavior. Within the risk range of the juveniles served by the programs in the juvenile justice system, a target value is set for that system which represents sufficiently high risk for such juveniles to be a priority for effective treatment. The SPEP risk points are then assigned in relation to the proportion of juveniles at that risk level or higher who are served by the program being rated.

Certain details of the SPEP rating scheme must be tailored to the particular juvenile justice system using it, the programs that are offered, and the nature of the data and data systems that are available. Figure 5 presents an example of a SPEP form for summarizing the ratings that would be generated for a particular program, in this case one serving youth on probation. This form depicts the rating categories and illustrates the proportionate points that are available in each. Keep in mind that the actual ratings are made on the basis of program information and service data for each program; the SPEP form summarizes the results of the rating process and provides a format for adding up the ratings to obtain a total score. Because of the way the rating dimensions are defined and the points are allocated in relation to the meta-analysis results, the

Figure 5. Example of a SPEP form for summarizing the ratings for a local program

Standardized Program Evaluation Protocol (SPEP) for Services to Probation Youth

	Possible Points	Received Points
Primary Service: High average effect service (35 points) Moderate average effect service (25 points) Low average effect service (15 points)	**35**	
Supplemental Service: Qualifying supplemental service used (5 points)	**5**	
Treatment Amount: Duration: % of youth that received target number of weeks of service or more 0% (0 points) 20% (2 points) 40% (4 points) 60% (6 points) 80% (8 points) 100% (10 points)	**10**	
Contact Hours: % of youth that received target hours of service or more 0% (0 points) 20% (3 points) 40% (6 points) 60% (9 points) 80% (12 points) 100% (15 points)	**15**	
Treatment Quality: Rated quality of services delivered: Low (5 points) Medium (10 points) High (15 points)	**15**	
Youth Risk Level: % of youth with the target risk score or higher: 25% (5 points) 50% (10 points) 75% (15 points) 99% (20 points)	**20**	
Provider's Total SPEP Score:	**100**	**[Insert Score]**

total score for a particular program can be interpreted as a measure of how closely the key characteristics of that program match the profile of characteristics shown in the meta-analysis to be most strongly associated with effects on recidivism. Viewed from a diagnostic and program improvement perspective, low ratings on any of these factors identify aspects of a program that should make the greatest difference in its effectiveness if they were to be improved. The SPEP is thus designed not only to evaluate each program against an evidence-based best practice profile, but to provide guidance for improving programs that fall short in that evaluation.

B. The Experience of State Juvenile Justice Systems with the SPEP Tool

An initial version of the SPEP tool was implemented in North Carolina in 2001, tailored to the prevention and court supervision programs funded through the North Carolina Department of Juvenile Justice and Delinquency Prevention.

An initial version of the SPEP tool was implemented in North Carolina in 2001, tailored to the prevention and court supervision programs funded through the North Carolina Department of Juvenile Justice and Delinquency Prevention. In 2006 it was adopted by the Arizona Juvenile Justice Services Division for application to state-funded programs for juveniles on probation. In addition, a project to apply it in Tennessee to programs in residential facilities for juvenile offenders was recently launched. These applications have led to a number of refinements in the SPEP scheme, and much has been learned about the best way to incorporate it into state-level juvenile justice systems as well as some of the challenges involved in that process. Most important, these projects provided an opportunity to conduct validation studies of the SPEP that tested the relationship between SPEP program ratings and recidivism outcomes for the juveniles served by the rated programs.

1. North Carolina

With the passage of the 1998 Juvenile Justice Reform Act, North Carolina became the second state (after Washington) to mandate that only effective services for juvenile offenders would be eligible for state funding. The act required the North Carolina Department of Juvenile Justice and Delinquency Prevention (DJJDP) to ensure that this mandate was implemented and to evaluate programs funded through the state's Juvenile Crime Prevention Councils as a condition of continued funding. However, the state did not provide any funds to DJJDP for such a statewide evaluation. When DJJDP officials learned of the work underway to develop the SPEP from Lipsey's meta-analysis, they decided to try it out as an evaluation tool. By scoring the DJJDP-funded programs against evidence-based guidelines, administrators could both assess the effectiveness of the programs and comply with the state mandate to fund only effective services.

The North Carolina SPEP Project was initiated in October 2001. After an initial development phase, the project progressed to pilot testing in selected rural and urban counties, followed by statewide rollout in 2006. The research team and the North Carolina DJJDP staff were able to classify almost all of the state-funded prevention and court supervision programs into the categories of program types for which there was sufficient research to develop a SPEP rating scheme. Only one type of program used by DJJDP, called Guided Growth, could not be classified into any of the primary service categories with sufficient research to be included in the SPEP.

DJJDP had a client-tracking system in place that routinely received data from service providers about the nature and amount of service provided to each juvenile. Information from that system was adapted to provide the input data for rating the SPEP factor related to the amount of service received by the juveniles in each program. DJJDP also had a validated risk assessment instrument in place to provide the data needed to rate the SPEP factor on the risk level of the juveniles served. Drawing on these data sources, it was possible to produce SPEP scores for programs statewide electronically through an automated system that remains in place.

DJJDP already had excellent offender management tools (a disposition matrix, a validated risk assessment instrument, an excellent needs assessment instrument, and a disposition grid) that it had used effectively to assign offenders to appropriate placement options and reduce admissions to the state's Youth Development Centers by 68 percent (North Carolina Department of Juvenile Justice and Delinquency Prevention, 2008). Adding the SPEP tool to DJJDP's repertoire permitted a statewide evaluation of the presumptive effectiveness of the community-based programs the state was increasingly using for juvenile offenders.

To assess the validity of the overall SPEP scores and the ratings on each of the factors that contribute to those scores, a recidivism analysis was conducted using data for juveniles served by 50 SPEP-rated prevention programs and 113 SPEP-rated programs for juveniles under court supervision in the community (Lipsey, Howell, and Tidd, 2007). Risk-adjusted recidivism rates for these juveniles were generated with statistical models that predicted recidivism based on risk and prior delinquency history. These models were used to estimate the recidivism rate of the juveniles under conditions where they all had the same initial risk for recidivism. Given equal risk, it was expected that actual recidivism rates would be lower for juveniles served by programs with high SPEP ratings than those served by programs with low SPEP ratings. These analyses found that the SPEP scores were moderately correlated with the risk-adjusted recidivism rates, with larger relationships found for the court supervision cases than for the prevention cases.

2. Arizona

Former OJJDP Deputy Administrator Rob Lubitz was familiar with the North Carolina SPEP Project. After his appointment as director of the Juvenile Justice Services Division (JJSD) in Arizona, Mr. Lubitz championed SPEP implementation across all Arizona court services programs. JJSD staff began implementing the SPEP rating scheme for their contract service providers in five pilot counties in the fall of 2006. Follow-up activities were then aimed at prompting providers to plan program improvements that would elevate their SPEP scores. Subsequently, the SPEP has been expanded to all JJSD-funded programs in the state.

Based on information obtained during the contracting process and by direct contact with the providers, it was possible to classify nearly all the Arizona programs as representative of therapeutic program types that were included in the research that supported the SPEP and thus to apply the SPEP ratings to them. The exceptions were brief behavior-specific programs—short educational programs on topics related to juvenile behavior problems—for which there is insufficient research on which to base a SPEP. The Arizona JJSD has a well-developed data system that includes detailed risk assessment scores capable of supporting an especially differentiated SPEP risk rating.

Obtaining data about the duration of service and number of contact hours by the respective programs for each juvenile proved more challenging.

Obtaining data about the duration of service and number of contact hours by the respective programs for each juvenile proved more challenging. JJSD did not collect service information in this form, but the financial records did identify the billable service units provided to each juvenile. JJSD staff was able to use that data to establish service start and end dates and to convert the service units into contact hours. There was no reasonable source for quality of service ratings for each program, however, so the JJSD staff embarked on a project to develop a rating scheme based on information that could be required as part of the contracting process, supplemented by site visits as needed.

JJSD's experience with the SPEP implementation stimulated some additional innovations. To oblige providers to attend to the SPEP ratings for their programs as a programmatic diagnostic tool, JJSD staff developed a format for program improvement plans that had to be filed by each provider during the contracting cycle. The SPEP experience also highlighted the importance of matching programs with offender needs and motivated JJSD staff to develop an improved needs assessment instrument.

Two validation studies of the SPEP scores were conducted in Arizona: an initial one in the five-county pilot test (Lipsey, 2008) and a second with data from the statewide implementation (Redpath and Brandner, 2010). In both cases the actual recidivism rates for the juveniles served by each program were compared with the rates predicted for them based on their risk profiles and prior delinquency history. In both studies, the ratings on the individual SPEP factors were correlated with the difference between actual and predicted recidivism—when the SPEP ratings were higher, the actual recidivism was lower relative to predicted recidivism. For the 18 programs in the pilot counties with the highest overall SPEP scores, the actual 6- and 12-month recidivism rates for the juveniles served averaged about 12 percentage points lower than predicted. For the 48 programs with lower scores, the difference between actual and predicted recidivism rates was a negligible one percentage point. Among the 90 programs in the subsequent statewide study, juveniles in programs with the highest overall SPEP scores had recidivism rates that averaged about 5 percentage points lower than predicted; juveniles in programs with lower SPEP scores had average recidivism rates about 4 percentage points higher than predicted.

C. Lessons Learned from the North Carolina and Arizona SPEP Projects

The projects in North Carolina and Arizona demonstrated that the SPEP could be implemented statewide and used routinely to assess juvenile justice programs according to how closely their characteristics match evidence-based best practice profiles. Most important, the studies conducted in these states showed that the SPEP scores for the rated programs were related to the recidivism rates of the juveniles served by those programs. Juveniles with equal risk for recidivism had lower recidivism rates when served by programs with high SPEP scores than when served by programs with lower SPEP scores. The SPEP scheme appears to be working as expected and shows encouraging empirical validity as a guide to effective programming for juvenile offenders.

Some of the other lessons learned from the North Carolina and Arizona SPEP projects that have implications for other juvenile justice systems in which the SPEP might be implemented include:

- If the North Carolina and Arizona juvenile justice systems are typical, the overwhelming majority of juvenile justice programs can be classified and evaluated using the SPEP if appropriate service and risk data are available.
- The initial SPEP scores for the programs in these states were relatively low despite the fact that they have juvenile justice systems strongly oriented to treatment and rehabilitation. These SPEP scores indicated that most programs had considerable room for improvement. The greatest shortfall indicated by the SPEP in these states was in the amount of service provided.
- Careful matching of programs with treatment needs is a relatively new frontier in juvenile justice. Though the SPEP does not rate this directly, program managers in both North Carolina and Arizona recognized that better matches would lead to greater program effectiveness. Embedded in this matching is a better understanding of risk reduction by providers and a better understanding of treatment protocols by court and juvenile justice personnel.
- Program administrator and staff turnover are significant impediments to successful SPEP implementation. When SPEP champions in the juvenile justice agencies departed or staff members familiar with the SPEP were shifted into other positions, SPEP implementation was disrupted. To sustain SPEP implementation, teams need to be trained and committed to the process rather than only a few individuals.
- Considerable systemwide improvements can be made without great cost by making key adjustments in the management of offenders, e.g., making better use of risk and needs assessment tools and the SPEP allows administrators to target higher risk offenders with effective programs matched to their needs, decrease reliance on expensive residential placements, and focus efforts on improving the outcomes of existing programs and services.

- Administrators may have myriad reasons for being reluctant to embrace evidence-based programs, but resistance to changing established practices may well be the predominant one. Strong management leadership is essential to generate support, and extensive training and technical assistance must be provided.

At the operational level, implementation of the SPEP for program evaluation and improvement requires commitment and adaptation by juvenile probation and correctional services.

At the operational level, implementation of the SPEP for program evaluation and improvement requires commitment and adaptation by juvenile probation and correctional services. Use of the SPEP needs to be institutionalized via policy and procedural directives or manuals, desktop guides, and the like to ensure consistent application. To support continued commitment, it would be best if the SPEP were implemented in a manner that produced objective evidence of progress, such as a reduction of the dynamic risk and need levels of offenders or a reduction in recidivism due to the proper matching of offender risks and needs to effective rehabilitative services. One method for obtaining such evidence is to readminister risk and needs assessment instruments throughout the life of a case. Procedures may also be needed for presenting program results in staff case-planning meetings and judicial reviews.

D. The Challenge of Evidence-Based Practice for Service Providers

Although the SPEP instrument can be used by juvenile justice administrators to assess the expected effectiveness of the treatment programs they use and to guide improvement in those programs, it is the providers of those programs who must respond if the evidence embedded in the SPEP is to influence their practice. The idea of evidence-based practice seems relatively simple: Have the providers of services to juvenile offenders use therapeutic programs that have been shown in research to reduce recidivism, and implement those programs the same way they were implemented in the research that found the best outcomes. Unfortunately, the challenges associated with realizing this idea can be quite complex.

For the most part, clinicians—including delinquency practitioners—are ambivalent about the role science should play in the interactions that occur between therapist and client. This is one of the reasons why research has failed to make its way consistently into those interactions (APA Task Force on Evidence-Based Practice for Children and Adolescents, 2008). Therapists tend to rely on a mixture of good intentions, some theory, practical wisdom, and—depending on how long they have been engaged in this difficult work—the use of specific techniques guided by their experience. This is often referred to as “treatment as usual.” Unfortunately, a number of studies have demonstrated that usual care is at best uneven and, at times, harmful (Knitzer and Cooper, 2006; Weisz et. al., 2005).

The prevailing interpretation of evidence-based programming as the use of model programs developed and evaluated elsewhere presents further challenges. With little understanding of the difficulties associated with quality implementation, practitioners who adopt these programs frequently find their efforts poorly supported. The clinician (or worse, the clinician’s supervisor) may have attended a presentation on an evidence-based program at a conference or read a book written by the program developer. That experience is then followed by the clinician’s attempt to use what was heard or read, perhaps as interpreted by a supervisor. This is not only a constricted view of how to go about evidence-based practice, but an underestimation of the power of inertia in clinical practice (Miller et al., 2006). The Washington State experience (Barnoski, 2004a) has left little doubt that effective use of “off-the-shelf” program models requires significant start-up costs, great care, and strong sustained ties to the original program developers (Carver, 2004).

Beyond practitioner ambivalence about research and emphasis on model programs as the way to move research into practice, other dilemmas are readily apparent. The short list of problems includes:

- Matching the distinctive program requirements that were established in research settings to the real-life organizational constraints faced by practitioners, e.g., large caseloads, little supervision, and resource limits on the types, frequency, and duration of services.
- Sustaining the program over time in the face of clinician and management turnover.
- Finding out if the time and effort required to implement the evidence-based program produced the desired effect, despite limited capacity to track outcomes.

Any approach to evidence-based practice that avoids the difficulties associated with adopting model programs would likely be more readily embraced by the practitioner community. In this regard, the underlying premise of the SPEP is attractive—that matching current programming to the characteristics shown by research to distinguish effective programs is a form of evidence-based practice. This perspective not only widens the practitioner's understanding of evidence-based practice, but also addresses several of the challenges identified above. Rather than calling for a wholesale exchange of the services being provided, the SPEP allows many practitioners to use the same set of clinical tools currently in place, though it does challenge them to use those tools with the right clients, the right frequency and duration, and stringent quality control monitoring. Comparatively speaking, the message of the SPEP is that "It does not take a magic bullet program to impact recidivism, only one that is well made and well aimed" (Lipsey, 2009, p. 145).

It is important to note that though the SPEP adds practicality, implementing it will necessarily require some degree of change from treatment as usual.

It is important to note that though the SPEP adds practicality, implementing it will necessarily require some degree of change from treatment as usual. As such, practitioners will greet it with some of the same natural ambivalence about research described above. Clinicians never like to hear that even a portion of what they are doing may be ineffective. Also, if a model program with three or four research studies showing effectiveness is viewed as out-of-touch with real practice, prescribed changes to treatment strategy based on meta-analysis may seem even further distanced from the clinician's day-to-day work. For the SPEP to be successful, careful attention must be paid to how these changes are presented to practitioners and practical suggestions must be provided for implementing them. Long-time practitioners know from their own experience that delinquency reduction is complex and that there is no magic bullet (Carver, 2005). What they may be able to learn from the SPEP is that relatively manageable adjustments to their work can yield significant benefits for both clients and communities.

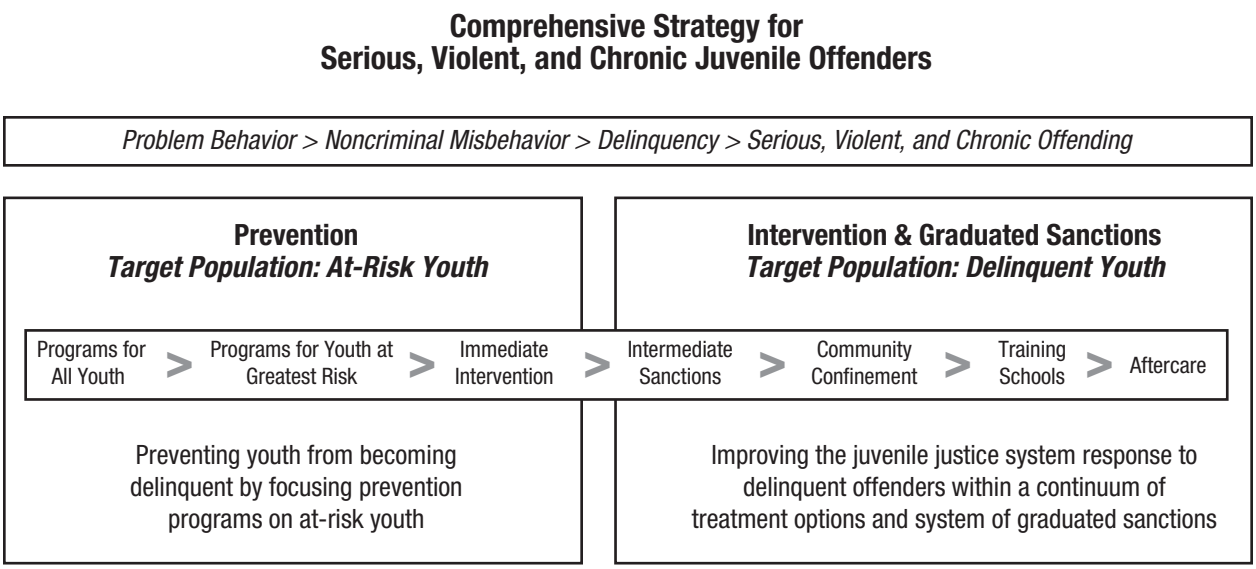
VII. Integrating Evidence-Based Practice into Juvenile Justice Systems

The SPEP provides a critical tool for effective management of juvenile justice systems that has been missing. However useful it may be for identifying effective programs and guiding improvement of ineffective ones, it is only one piece of the juvenile justice puzzle. To make its strongest contributions to efforts to reduce delinquency, the SPEP must be used as part of a comprehensive strategy for optimizing the effectiveness of the juvenile justice system.

A. A System Reform Framework

The Comprehensive Strategy (CS) for Serious, Violent, and Chronic Juvenile Offenders (Wilson and Howell, 1993, 1994) is a framework for guiding state and local system reforms to address juvenile delinquency in a cost-effective manner. The CS is a two-tiered system for responding proactively to juvenile delinquency (Figure 6). In the first tier, delinquency prevention, youth development, and early intervention programs are relied on to prevent delinquency and reduce the likelihood that at-risk youth will appear in the juvenile justice system. If those efforts fail, then the juvenile justice system, the second tier, must make proactive responses by addressing the risk factors for recidivism and associated treatment needs of the offenders, particularly those with a high likelihood of becoming serious, violent, and chronic offenders. At the same time, supervision proportionate to the risk to public safety posed by the respective offenders must be applied. In the Comprehensive Strategy framework, the supervision and control component is referred to as sanctions, a term also used in this fashion in many juvenile justice systems.

Figure 6. The Comprehensive Strategy for Serious, Violent, and Chronic Juvenile Offenders



Sources: Wilson & Howell (1993, 1994); Howell (2003a, 2003b, 2009)

The CS framework consists of a spectrum of program options sufficient to address the diverse treatment needs and risk profiles of the target juveniles as well as a continuum of graduated sanctions sufficient to exercise the control necessary to ensure both public safety and the participation of youth in the assigned programs. More specifically, the CS framework is structured around six levels of parallel program interventions and sanctions, moving from least to most restrictive, plus aftercare for youth released from secure facilities:

- Community primary prevention programs oriented toward reducing risk and enhancing strengths for all youth
- Focused secondary prevention programs for youth in the community at greatest risk but not involved with the juvenile justice system or, perhaps, diverted from the juvenile justice system
- Intervention programs tailored to identified risk and need factors, if appropriate, for first-time minor delinquent offenders provided under minimal sanctions, e.g., diversion or administrative probation
- Intervention programs tailored to identified risk and need factors for nonserious repeat offenders and moderately serious first-time offenders provided under intermediate sanctions, e.g., regular probation
- Intensive intervention programs tailored to identified risk and need factors for first-time serious or violent offenders provided under stringent sanctions, e.g., intensive probation supervision or residential facilities
- Multicomponent intensive intervention programs in secure correctional facilities for the most serious, violent, and chronic offenders
- Post-release supervision and transitional aftercare programs for offenders released from residential and correctional facilities

1. Prevention Tier

Though prevention is not the focus of this paper, it is an important part of a truly comprehensive strategy for addressing juvenile delinquency. The prevention component of the CS framework consists of the two initial program levels of the continuum: primary prevention and secondary prevention. In this framework, primary prevention refers to universal prevention programs, meaning that all youth are recipients in a community-wide program or a program provided to all youth in local school classrooms, community centers, and the like. Secondary prevention programs target children in the community with identified risk factors for delinquency and related adverse outcomes. These may be pre-delinquent youth who have not yet appeared in the juvenile justice system and who receive school- or community-based programs. Or these may be youth referred to the juvenile justice system for minor offenses but judged to be sufficiently at risk to warrant services and be diverted to community- or school-based prevention programs.

Use of a research-based risk and protection framework within the public health model helps structure the delinquency prevention enterprise in communities.

Use of a research-based risk and protection framework within the public health model helps structure the delinquency prevention enterprise in communities. The public health model is familiar to practitioners because of its widespread application in the health arena. Juvenile delinquency and other child and adolescent problem behaviors share many common risk and protective factors (Durlak, 1998; Loeber and Farrington, 1998). Thus prevention programs oriented toward reducing risk and enhancing protective factors can have beneficial effects for ameliorating a range of adverse outcomes. These programs can be successfully promoted by providing community members with training and technical assistance in risk-protection assessment and strategic prevention planning. For instance, the Life Skills Training program has demonstrated success as an approach to preventing tobacco, alcohol, and marijuana use (Botvin, Mihalic, and Grotpeter, 1998).

2. Intervention and Graduated Sanctions Tier

The intervention and graduated sanctions component of the Comprehensive Strategy consists of the last four levels of the overall CS framework in which treatment programs are combined with levels of supervision or control appropriate to the nature of juveniles' offenses and their risk for reoffending.

For chronic offenders, who account for a disproportionately large amount of delinquency, their offending careers develop over time. Thus a continuum of programs aimed at different points along the life course has a much better chance of succeeding than a single intervention. Because certain risk factors operate at particular times in individuals' lives, a developmental perspective is necessary for constructing a full continuum of delinquency prevention and intervention programs. Early on, for example, programs may be needed that address family risk factors. In adolescence, peer influences are predominant, and the most appropriate programs may be those that buffer the effects of exposure to delinquent peer influences and the spread of delinquency and violence in adolescence. On the other hand, interventions that counter individual risk factors (e.g., mental health problems) and community risk factors (e.g., high-crime neighborhoods) may be needed all along the life course.

The collective effect of a well-constructed spectrum of programs is likely to be much greater than the impact of a single program, as illustrated in a RAND cost-benefit study of juvenile delinquency prevention and treatment programs.

By developing a continuum of integrated programs and sanctions, juvenile justice systems can match offenders' risk levels and treatment needs to appropriate services and supervision at any point of development of offender careers. The collective effect of a well-constructed spectrum of programs is likely to be much greater than the impact of a single program, as illustrated in a RAND cost-benefit study of juvenile delinquency prevention and treatment programs (Greenwood et al., 1996). The RAND researchers found that, if implemented statewide, a combination of four delinquency prevention and offender treatment programs could achieve the same level of serious crime reduction as California's "three strikes" law, which mandated imprisonment for the third strike. The researchers projected that these four programs would cost less than $1 billion per year to implement throughout California, compared with about $5.5 billion per year for "three strikes." Thus, at less than one-fifth the cost, the four programs could prevent more serious crimes than imprisonment would. As the RAND researchers noted, "Based on current best estimates of program costs and benefits, investments in some interventions for high-risk youth may be several times more cost-effective in reducing serious crime than mandatory sentences for repeat offenders" (Greenwood et al., 1996, p. 40).

More generally, the intervention and graduated sanctions tier of the CS calls for a proactive and balanced approach that integrates long-term delinquency prevention and short-term behavior supervision and control. This portion of the CS is based on the following core principles (Wilson and Howell, 1993):

- Immediate and effective intervention when delinquent behavior occurs to prevent delinquent offenders from becoming chronic offenders or committing progressively more serious and violent crimes. Initial intervention efforts, under an umbrella of system authorities (police, intake, and probation), should be centered in the family and other core societal institutions. Juvenile justice system authorities should ensure that an appropriate response occurs and act quickly and firmly if the need for formal system adjudication and sanctions is demonstrated.
- Identification and control of the small group of serious, violent, and chronic juvenile offenders who have committed felony offenses or failed to respond to nonsecure community-based rehabilitation services offered by the juvenile justice system. Measures to address delinquent offenders who are a threat to community safety may include placement in secure community-based facilities, training schools, and other secure juvenile facilities. Even the most violent

or intractable juveniles should not be moved into the criminal justice system before they age out of the jurisdiction of the juvenile justice system.

B. The Essential Tools: Risk Assessment, Needs Assessment, Case Management Plan[4]

The treatment programs integrated into the intervention and graduated sanctions tier of the CS framework must, of course, be effective for reducing recidivism or they will have little value to the juvenile justice system. As discussed earlier, the SPEP provides a tool for assessing the expected effectiveness of programs of a generic type that have been evaluated in credible research studies. Moreover, the SPEP can guide improvement for programs that fall short in that assessment. Although the SPEP gives juvenile justice administrators the opportunity to access the rich body of evidence on which programs work and which do not, administrators also need to know what works for whom, how to match programs to the risk level and needs of individual offenders, and how to go about moving effective programs into everyday practice.

All cases are not equal. Some offenders require substantial service intervention and supervision (high risk), others much less attention (low risk). Service interventions should address each youth's identified needs (individualized case plan). Juvenile offenders typically have multiple treatment needs in several developmental domains of their lives—family, school, peers, and so on. Several specific services may be needed to adequately address the array of presenting problems. Three tools—the main instruments of the Structured Decision-Making Model™[5]—are used within the framework of the CS to guide decisions on these matters. First, a risk assessment instrument is used to determine the level of sanctions needed to protect the public from a particular offender and the appropriateness of the most intensive services available. Next, a needs assessment instrument is employed to guide selection of treatment programs that have an appropriate focus. Then, to find the best match between offender and program, which is critical for effective rehabilitation, the results of the needs assessment must be used in tandem with the results of the risk assessment to place the offender in a particular supervision level and treatment program within that supervision level. This program placement is guided by a third tool, the case management plan.

Risk assessment. From the time of their creation, juvenile courts and correctional agencies have used some means of assessing offenders' risk levels. There are four basic approaches to risk assessment: staff judgments, clinical assessments, consensus-based assessments, and empirically derived assessments based on actuarial data (Gottfredson and Moriarty, 2006; Wiebush, 2002). Two of these are not reliable at all: informal staff judgments result in over-classification (i.e., too many false positives), and clinical assessments have been shown to be significantly less accurate than empirically derived assessments (Gottfredson and Moriarty, 2006; Grove et al., 1990; Grove and Meehl, 1996). Consensus-based risk assessments (i.e., assessments based on items agreed on by a group of agency staff) are also less accurate than empirically derived instruments (Gottfredson and Moriarty, 2006; Wiebush, 2000, 2002).

A valid risk assessment instrument is one that does what it purports to do—that is, it accurately distinguishes between youth according to the probability that they will subsequently engage in delinquent behavior (Wiebush, 2002). Research supporting the validity of risk assessments has increased dramatically in recent years (Gottfredson and Moriarty, 2006). One reason is that, with the growth of automated court and correctional record systems, large databases are now available to researchers for risk assessment studies. Risk assessment instruments have been validated on more than a dozen state juvenile populations and in other studies (Wiebush, 2002). In addition, risk assessment instruments recently have been validated for several serious violent offender subgroups, including felony recidivists (Barnoski, 2004b), first-time referrals versus second- and third-time referrals (LeCroy, Krysik, and Palumbo, 1998), and potential chronic

[4] The authors express appreciation to Dennis Wagner of the National Council on Crime and Delinquency for generously sharing information on best practices in structured decision making.

[5] Registered trademark of the National Council on Crime and Delinquency, all rights reserved. For more information: http://www.nccd-crc.org/crc/crc/c_sdm_about.html.

offenders among second-time offenders (Smith and Aloisi, 1999). Three risk assessment instruments have been validated for successful classification of offenders with regard to their likelihood of recidivating with violent offenses: in Maryland (Wiebush, Johnson, and Wagner, 1997), Missouri (Johnson, Wagner, and Matthews, 2001), and Virginia (Wiebush, Wagner, and Erlich, 1999).

Matching the most effective programs with the highest risk offenders yields the greatest reductions in recidivism, and using risk assessment tools with the SPEP provides a systematic approach to optimizing those effects.

Valid risk assessment instruments are used in a structured decision-making scheme within the CS framework to estimate the level of sanctions needed to protect the public from the threat posed by an offender and to identify those youth most appropriate for intensive programs aimed at reducing recidivism. Matching the most effective programs with the highest risk offenders yields the greatest reductions in recidivism, and using risk assessment tools with the SPEP provides a systematic approach to optimizing those effects.

Needs assessment. Needs assessments are used to determine the specific program interventions to be delivered within the designated custody or supervision level (Wiebush, 2002). A needs assessment is intended to do the following:

- Provide an overview of the level of seriousness of the juvenile offender's treatment needs
- Provide information that can assist professionals in developing a treatment plan to address the juvenile's needs
- Provide a baseline for monitoring the juvenile's progress
- Provide a basis for establishing workload priorities
- Aid agency administrators in evaluating resource availability throughout the jurisdiction and determining program gaps that need to be filled

Unlike risk assessments, needs assessments do not predict future behavior; thus they are not developed through empirical research. Instead, jurisdictions employ a consensus approach to identify and set priorities for the most important service issues. Local professionals are responsible for selecting the items to include in the needs assessment instrument. They are guided in this effort by existing state and federal laws (e.g., laws addressing special education services), research identifying effective and promising programs, and local philosophies about effective rehabilitation services. In the structured decision-making model, needs assessment results are used to adjust the placement of offenders in various risk levels (as recommended by risk assessment results). For example, a juvenile offender who is determined to be at medium risk and who has a very high treatment needs score might be placed in a program for high-risk juveniles to take advantage of the relatively intensive treatment services offered by the program.

Needs assessment instruments typically include items concerning offender needs in areas that correspond with risk factors for delinquency, including family functioning or relationships, school attendance and behavior, peer relationships (e.g., negative peer associations and gang involvement), and individual problems (e.g., substance abuse and emotional stability). Many instruments also include measures of health and hygiene, intellectual ability, and learning disabilities.

Case management plan. Youth risk and needs assessments are the primary tools for assigning cases to a level of supervisory control and for identifying the treatment interventions for the case management plan. Client risk and treatment needs are assessed to develop individualized case plans to reduce recidivism. Case plans are executed by delivering identified service interventions during case supervision. Case plans are a statement of an agency's intention to deliver future service interventions, and specific services and timelines are included to allow implementation of the plan to be reviewed and monitored. Periodic reassessments of treatment needs also help case managers monitor client progress and can indicate when adjustments might be needed in individual treatment regimens.

We next discuss three successful examples of Comprehensive Strategy implementation that illustrate several of the strategy's key principles. The first, the San Diego County Breaking Cycles program, shows how a community can integrate the prevention and graduated sanctions components of the Comprehensive Strategy. The second example, the Orange County, California, 8% Early Intervention Program, illustrates how a community can effectively target potential and identified serious and chronic juvenile offenders with a model continuum of sanctions and services. The third example, Missouri's statewide graduated sanctions approach, illustrates the effective use of structured decision-making tools.

C. Examples of a Comprehensive Continuum of Prevention and Graduated Sanctions

1. San Diego's Comprehensive Strategy

San Diego County was the first site to implement the entire Comprehensive Strategy. The strategy was developed and implemented in 1996–97 under the leadership of the Juvenile Justice Coordinating Council, part of the Children's Initiative of San Diego County, which provided coordination and staff support (www.thechildrensinitiative.org). San Diego's Comprehensive Strategy consists of two main components: prevention and graduated sanctions. These components are linked in an overall program called Breaking Cycles (Burke and Pennell, 2001). The Breaking Cycles program has three specific goals (Burke and Pennell, 2001, p. 27):

- To reduce the number of at-risk minors who become delinquent by involving them in a prevention program
- To improve the juvenile justice system through implementation of a system of graduated sanctions with a focus on community-based treatment
- To break the cycle of substance abuse and family problems that fosters crime and violence

The prevention component targets youth who have not yet entered the juvenile justice system but who evidence problem behaviors such as chronic disobedience to parents, curfew violations, repeated truancy, multiple attempts to run away from home, and drug and alcohol use. This secondary prevention approach is different from that of most community prevention strategies, which typically focus primary prevention programs on all youth. Five Community Assessment Teams (CATs) provide referrals and services to at-risk youth and their families by linking them with social supports in the community strategically located for easy access across the county. A team composed of a coordinator, probation officer, case manager, and other experts conducts comprehensive individualized youth and family assessments that guide the development of either a case management plan (for a long-term case) or a referral to community agencies (for a short-term case).

Youth in a second target group—those in juvenile court for delinquency involvement—access the graduated sanctions component of Breaking Cycles through a juvenile court commitment decision determined, in part, by a Probation Department screening committee. This determination is based on the offender's current offense and prior criminal history, as well as on the results of a risk assessment. A Breaking Cycles case plan is then developed for each youth. The case plan is family centered and strengths based. It is designed to promote accountability, rehabilitation, and community protection. Youth are assigned for variable lengths of program participation—90, 150, 240, or 365 days—depending on risk severity and treatment needs. The following continuum of placement options is used:

- Institutional placement (e.g., minimum-security custody)
- Community-based placement (e.g., day treatment in the Reflections Program)
- Home placement (e.g., the Community Unit)

Each of these intervention levels is linked with community programs and resources that carry out the comprehensive treatment plan.

Each of these intervention levels is linked with community programs and resources that carry out the comprehensive treatment plan. Most juvenile offenders begin their rehabilitative process in highly structured institutional

settings and are stepped down to lower levels of program structure and supervision (i.e., community-based and home placement, as shown above) as reassessments are made. Youth may also be stepped up from initial less restrictive placements or after having been stepped down to lower levels—again, depending on reassessment results. Services are linked to each of the three placement levels to provide youth and their families "with a fluid and seamless system of service delivery" (Burke and Pennell, 2001, p. 31).

Burke and Pennell (2001) conducted a process and outcome evaluation of the San Diego Comprehensive Strategy that encompassed both the secondary prevention component and the graduated sanctions component. They found that the prevention component succeeded in keeping most at-risk youth out of the juvenile justice system. Remarkably, fewer than 20 percent of the long-term CAT cases were referred to court, and only 7 percent of all long-term CAT cases subsequently were adjudicated delinquent (p. 6). Compared with other at-risk juveniles in the comparison group, long-term CAT clients were less likely to use alcohol and drugs and more likely to perform better in school after participating in the program.

Burke and Pennell also found that the graduated sanctions component was effective in keeping offenders from progressing to more serious delinquency. Regardless of commitment length, youth in the Breaking Cycles program were less likely than similar preprogram cases to have a court referral for a felony offense or to be adjudicated for a felony offense during the 18-month follow-up period. Breaking Cycles youth also were less likely to be committed to long-term state correctional facilities, less likely to be using alcohol or drugs, and more likely to be enrolled in school during the follow-up period.

In addition, Burke and Pennell (2001, p. 11) found that the San Diego juvenile justice system was more cost-efficient as a result of its implementation of the Comprehensive Strategy. This was attributed, first, to the targeting of appropriate youth for Breaking Cycles program intervention, and second, to the effectiveness of the program in reducing juvenile delinquency. In addition, the San Diego County Probation Department succeeded in intervening with offenders at early points in their delinquent careers and also in holding them accountable when their offending continued.

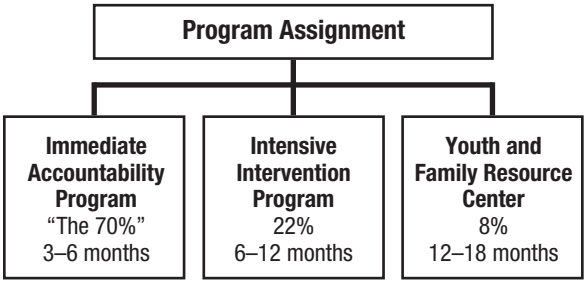
2. Orange County's Comprehensive Strategy

The Orange County, California, Probation Department has used the Comprehensive Strategy to develop a unique system of graduated sanctions and a parallel continuum of program options. The 8% Early Intervention Program is the first known implementation of the Comprehensive Strategy's graduated sanctions component (Schumacher and Kurz, 2000). The system began with development of a program for potential serious and chronic juvenile offenders that was based on the research finding that 8 percent of court referrals were the most chronic offenders, typically with five prior arrests. These offenders were identified as court wards who were 15.5 years old or younger at the time of their first or second court referral and who had at least three of four profile risk factors, which placed them at greatest risk of becoming serious chronic offenders. These offenders were identified at court intake through a risk assessment instrument and admitted to the 8% Early Intervention Repeat Offender Program.

The Orange County juvenile probation project also identified two other groups of offenders: a medium-risk group (22 percent of the total sample) and a low-risk group (the remaining 70 percent). The 22 percent group had one or two of the four profile risk factors, which placed them at medium risk of becoming serious chronic offenders. The low-risk group had none or only one of the four profile risk factors (Figure 7).

An interdisciplinary team of practitioners from throughout the county then developed a model continuum of juvenile justice services to manage all three groups simultaneously. The team used the Comprehensive Strategy to guide development of the intervention approach. Youth in the 70 percent low-risk group were assigned to the Immediate Accountability Program, those in the 22 percent medium-risk group were assigned to the Intensive Intervention Program, and those in the 8 percent high-risk group were admitted to the Early Intervention Youth and Family Resource Center's (YFRC) Repeat Offender Prevention Program. In addition, court-

Figure 7. Orange County model juvenile justice continuum



Source: Schumacher and Kurz, 2000; modified from Orange County, California, Probation Department, Model Continuum of Juvenile Justice Services

referred youth age 15.5 or older and placed in custody for 90 days or more were enrolled in the Challenge Program at the YFRC. The final component of the continuum was transitional aftercare (the Transitional Program).

The low-risk youth in the Immediate Accountability Program were supervised by volunteer probation officers who linked them with community-based programs and ensured that they met accountability (sanction) requirements. These offenders did not present any significant needs for intervention services. Youth in the medium-risk group were in the Intensive Intervention Program for a period of 6–12 months. Because these moderate-risk youth could have escalated to a higher risk status, they received intensive, integrated intervention and accountability services immediately upon program assignment. They were also subject to intensive supervision probation sanctions, along with a continuum of multiagency intervention services for them and their families at the YFRC. Youth in the high-risk group were assigned to the Repeat Offender Prevention Program for a period of 12–18 months. Like those in the medium-risk group, they received intensive integrated interventions and intensive supervision sanctions. However, they and their families also received a wide array of additional services, including:

- Intensive in-home family services
- Health screening, health education, and basic health services
- Substance abuse services
- Mental health services
- A full spectrum of on-grounds educational services

The Orange County program represents an impressive graduated sanctions system for probation services that combines immediate and intermediate sanctions with a continuum of treatment programs (Schumacher and Kurz, 2000, pp. 43–46). All wards of the court receive appropriate sanctions and services, based mainly on the results of risk and needs assessments. Offenders can be moved up and down the continuum of sanctions and program levels, depending on their progress in staying out of trouble and their success in treatment programs. The YFRC component of the program proved to be a valuable asset. Many of the 8% wards and their family members had such serious problems that they required brief periods of residential treatment. Parental problems such as child abuse, substance abuse, and criminal involvement were addressed. Follow-up data from the evaluation study showed that the 8% youth had fewer petitions for new law violations, and 8 out of 10 of them had either none or one new petition versus only 6 out of 10 for control cases (Schumacher and Kurz, 2000).

The Orange County system is a premier example of how jurisdictions should apply the Comprehensive Strategy's intervention and graduated sanctions framework. This system goes beyond most applications of the structured decision-making model by formally organizing distinct program structures for low-, medium-, and high-risk offenders. Any large jurisdiction would benefit from building a similar structure.

3. Missouri's Comprehensive Strategy

Missouri's statewide development of a continuum of graduated sanctions and services sets a very high standard for other states to follow. Guided by the Comprehensive Strategy, Missouri created a structured decision-making model that uses risk and needs assessments and a classification matrix: the Missouri Juvenile Offender Risk and Needs Assessment and

Classification System (Office of State Courts Administrator, 2002). A major goal of the state in establishing this classification system is to promote statewide consistency in the classification and supervision of juvenile offenders. The three tools of the Missouri system are as follows:

- An actuarial risk assessment tool, completed before court adjudication, that classifies youth into three categories: high, moderate, or low probability of reoffending. The risk assessment instrument has been validated (Johnson, Wagner, and Matthews, 2001)

Figure 8. Missouri risk and offense case classification matrix

Missouri Risk & Offense Case Classification Matrix

Offense Severity	Group 1 Offenses	Group 2 Offenses	Group 3 Offenses
Risk Level	Status Offenses Municipal Ordinances/ Infractions	Class A, B, & C Misdemeanors/ Class C & D Felonies	A† & B Felonies
Low Risk	A) Warn & Counsel B) Restitution C) Community Service D) Court Fees & Assessments E) Supervision	A) Warn & Counsel B) Restitution C) Community Service D+) Court Fees & Assessments E) Supervision	B+) Restitution C+) Community Service D+) Court Fees & Assessments E) Supervision F) Day Treatment G) Intensive Supervision H) Court Residential Placement I) Commitment to DYS
Moderate Risk	A) Warn & Counsel B) Restitution C) Community Service D) Court Fees & Assessment E) Supervision	*A) Warn & Counsel* *B) Restitution* *C+) Community Service* *D+) Court Fees & Assessments* *E) Supervision* *F) Day Treatment*	B+) Restitution C+) Community Service D+) Court Fees & Assessments E) Supervision F) Day Treatment G) Intensive Supervision H) Court Residential Placement I) Commitment to DYS
High Risk	A) Warn & Counsel B) Restitution C) Community Service D) Court Fees & Assessments E) Supervision	B+) Restitution C+) Community Service D+) Court Fees & Assessment E) Supervision F) Day Treatment G) Intensive Supervision H) Court Residential Placement I) Commitment to DYS	H) Court Residential Placement I) Commitment to DYS

† Mandatory certification hearings are required by statute for all Class A Felonies. In the event the juvenile is not certified, the juvenile officer should refer to this column of the matrix for classification purposes.
+ This symbol indicates options that should never be used as a sole option for youths who score in that cell, but only in conjunction with other options.

Source: Missouri Office of State Courts Administrator (http://www.courts.mo.gov/file/Classification%20Matrix%2012.20.00.pdf)

- A classification matrix that recommends sanctions and service interventions appropriate to the youth's risk level and most serious adjudicated offense
- A needs assessment instrument that recommends services that will reduce the likelihood of a youth's reoffending by reducing risk factors linked to recidivism

In addition, Missouri has developed a set of standards—Performance Standards for the Administration of Juvenile Justice—that help balance individual rights and treatment needs with public protection (Office of State Courts Administrator, 2000). These standards establish a common framework within which juvenile justice personnel can understand and assess the work of juvenile and family courts and enhance the courts' performance. They are "premised on the notion that court performance should be driven by core values of equity, integrity, fairness, and justice" (Office of State Courts Administrator, 2000, p. 2). The standards also include contact guidelines for high, medium, and low levels of supervision based on the placement of offenders in the classification matrix. Lastly, Missouri conducted a workload study to determine whether juvenile court staff was meeting intake and supervision performance standards (Johnson and Wagner, 2001). Overall, court staff met expected standards in 93 percent of the cases tracked.

Figure 8 shows Missouri's Risk and Offense Case Classification Matrix, developed by the state's Office of State Courts Administrator. The complete Missouri Juvenile Offender Classification System includes the following tools:

- An empirically validated risk assessment for estimating a youthful offender's relative likelihood of future delinquency
- The classification matrix (Figure 8), which links the level of risk with offense severity to recommend graduated sanctions
- A needs assessment for identifying the underlying psychosocial needs of youth
- A method for assessing juvenile offender adjustment to supervision, which incorporates a supervision reassessment form and a set of Web-based reports on the risk and need characteristics of youthful offenders (links are provided to each of these formal decision-making tools and written reports on system functions at the Missouri Office of State Courts Administrator: http://www.courts.mo.gov/page.jsp?id=233)

In more recent reforms, Missouri's Division of Youth Services now serves youth offenders who cannot be maintained in community settings in small, dormitory-style rehabilitation facilities close to their homes. In those facilities the focus is on (1) individualized and group treatments with a clear treatment model, (2) supervision, not correctional coercion, (3) skill building, (4) family partnership and involvement during confinement, and (5) aftercare. Only 9 percent of youth discharged from the Division of Youth Services were sentenced to adult prison within three years of release, and just 15 percent were reincarcerated within two years of release (Annie E. Casey Foundation, 2010).

VIII. Practical and Policy Considerations in Implementing Juvenile Justice Reform

For optimal performance, the entire juvenile justice system needs to operate on a research-informed, evidence-based platform. The Comprehensive Strategy with the Standardized Program Evaluation Protocol integrated as the central tool for ensuring evidence-based programming for juvenile offenders provides a holistic approach to ensure that a juvenile justice system and its service system are operating in a cost-effective way to improve outcomes for the youth in its care. Moreover, this strategy is neither burdensome nor cost prohibitive to implement, nor does it undermine the current service delivery system. In fact, the CS/SPEP framework enhances and enriches the current service delivery system by infusing evidence-based program improvements into existing services and programs. In this regard, the CS/SPEP supports a shift in how juvenile justice systems operate and how they are held accountable for accomplishing their mission.

A. Needed Improvements in Juvenile Justice Systems

The juvenile justice system has made enormous progress in program development and system reforms over the past twenty years. Nonetheless, many systems continue to struggle with achieving a balance of community-based versus institutionally based care and managing the tension between their rehabilitative versus public safety functions. In particular, juvenile justice systems in most states struggle with three challenges: (1) reducing reliance on incarceration, especially for minority youth; (2) building effective community-based programs for probation, reentry, aftercare, and parole systems to accommodate reductions in secure confinement; and (3) ensuring that effective programs are targeted to appropriate offenders in a way that will have optimal effects on recidivism. Meeting these challenges requires a clear view of the outcomes expected of the juvenile justice system, a well-developed plan for achieving them, and effective use of management tools for implementing that plan.

The results of the cyclical swings described earlier in this paper provide an additional context for reform. For children born in 2001, the Children's Defense Fund estimates that without appropriate intervention, current trends will result in one in every three black boys and one in six Hispanic boys being incarcerated at some point in their lifetime, a trajectory the organization has described as a "cradle-to-prison pipeline" (Edelman, 2009). The traditions of juvenile "correctional" work and its kindred relationship with adult correctional models in policy, practice, and staffing are barriers to the adoption of robust, developmentally appropriate systems of juvenile sanctions and treatment services. Gubernatorial appointments of juvenile justice administrators are often former prosecutors, law enforcement officers, or adult correctional administrators. Their affinity with the adult correctional tradition can lead to unduly harsh attitudes toward youth and negative attitudes toward rehabilitation. As reform-minded juvenile justice administrator Tim Decker observed about his experience as director of the Missouri Division of Youth Services:

> *Across the system there were entrenched organizational cultures. We protected both turf and the status quo instead of shared values and communication. Misperceptions related to the service offerings and strengths of others in the system were common. Both efforts faced a prevalent status quo bias and little faith in the possibility of a different approach to serving the youth. (Decker, 2010)*

Given these cultural traditions, it is understandable that punitive programs for juvenile offenders are commonplace despite considerable evidence of their lack of efficacy. This is especially true when such punitive strategies are used for female offenders, who often have a history of physical and sexual abuse (McGarvey and Waite, 2000; Rubin, 2000; Sedlak and McPherson, 2010). To quote Blueprints for Violence Prevention program founder Delbert Elliott, in his opening plenary remarks at the 2010

Blueprints conference: "To continue to place our kids into programs that we know don't work is unethical."

B. The Challenges of Change and Sustainability

> *All of the credible research shows that a continuum of evidence-based prevention programs for youth identified as being at risk of involvement in delinquent behavior, and intervention for those already involved, will greatly reduce crime and save much more than they cost when compared to the avoided law enforcement and social welfare expenditures. And the research reveals that these programs are most effective when provided in the context of a coordinated, collaborative local strategy involving law enforcement and other local public and private entities working with children identified as at risk of involvement in the criminal justice system. (Quoted with permission from Rep. Robert C. "Bobby" Scott, D-VA, Youth PROMISE Act White Paper: Fighting Juvenile Crime vs. "Playing Politics," 2009)*

Resistance to the adoption of evidence-based practice and systems of care is well recognized in the literature on program implementation (Fixsen et al., 2005; Paulson, Fixsen, and Friedman, 2004). Some of the barriers are practitioner based, while others are environmental, organizational, and systemic barriers that must be overcome for effective implementation. The juvenile justice field is rife with skepticism about the adoption of evidence-based practice. One common fear is that new evidence-based programs will siphon funding away from presumably effective, yet untested services. This is especially true of settings that do not have a cultural context of continuous quality improvement, an orientation toward being a learning organization,[6] or values related to excellence. Skepticism also emanates from beliefs that the local juvenile offender population is unique and that evidence-based programs "not invented here" are not responsive to local needs and populations. Moreover, current budgetary shortfalls have increased staff workloads, cut operating support, depleted training resources, and resulted in funding directed toward minimum compliance with required protections for juveniles. In the present climate, there is neither support for expanding reimbursement strategies (such as in state Medicaid plans) nor incentives for the adoption of evidence-based practice. Advocates for evidence-based practices must begin to address the funding barriers through active public comment, legislation, and consensus building.

Even when begun with vigor and commitment, change efforts often fizzle out. The foundations of the Comprehensive Strategy are community-based consensus building and a long-term commitment to systems improvement. It is the community consensus that creates the comfort to change practice and make the changes routine. Strong, consistent, adaptive agency leaders and champions who advocate in support of change appear to be key factors in achieving sustainability. These principles need to be recognized from the beginning and then form a continuing guide to implementation throughout the change process. A leader's personal, unwavering commitment to change tends to carry the day. A champion's ability to engage key public stakeholders in making a commitment to change is vital. Combined, these actions create positive affect and social support around the change process. Successful leaders and advocates inspire hope, excitement, camaraderie, and a sense of urgent purpose. They celebrate the sheer joy of creating something meaningful together.

Tim Decker argues, "Often we try to adopt new program ideas like evidence-based therapy without paying attention to the underlying but requisite change in culture. Effective change in organizational culture requires a strategic mindset of how to change the status quo. Cultural change, for example, is often driven by having the right people who share a set of beliefs and philosophies. In Missouri, we now operate on the belief that all youth desire to do well and succeed. This sounds simple, but it reflects a significant change from the old model" (Decker, 2010).

One step along the path of sustainability is the codification of evidence-based criteria in legislation and in policy. On

[6] *The attributes of "learning organizations" include strong leadership, an open and inclusive management culture, a stable resource base from which to launch process improvement, and transparent and accessible performance data.*

the federal level, the proposed Youth PROMISE Act and the Maternal, Infant, and Early Childhood Home Visiting Program, which is part of the enacted Affordable Care Act (P.L. 111–148), are landmark achievements for new investment in evidence-based programs and strategies. A number of states have also enacted a variety of approaches to codifying research-based practice. A Tennessee law (Tennessee Code, Chapter 525) is a good example. It establishes four levels of programming: evidence-based, research-based, theory-based, and pilot programs.

1. "Evidence-based" means a program or practice that meets the following requirements:
   a. The program or practice is governed by a program manual or protocol that specifies the nature, quality, and amount of service that constitutes the program; and
   b. Scientific research using methods that meet high scientific standards for evaluating the effects of such programs must have demonstrated with two (2) or more separate client samples that the program improves client outcomes central to the purpose of the program;
2. "Research-based" means a program or practice that has some research demonstrating effectiveness, but that does not yet meet the standard of evidence-based;
3. "Theory-based" means a program or practice that has general support among treatment providers and experts, based on experience or professional literature, may have anecdotal or case study support, and has potential for becoming a research-based program or practice; and
4. "Pilot program" means a temporary research-based or theory-based program or project that is eligible for funding from any source to determine whether or not evidence supports its continuation beyond the fixed evaluation period. A pilot program must provide for and include:
   a. Development of a program manual or protocol that specifies the nature, quality, and amount of service that constitutes the program; and
   b. Scientific research using methods that meet high scientific standards for evaluating the effects of such programs must demonstrate on at least an annual basis whether or not the program improves client outcomes central to the purpose of the program.

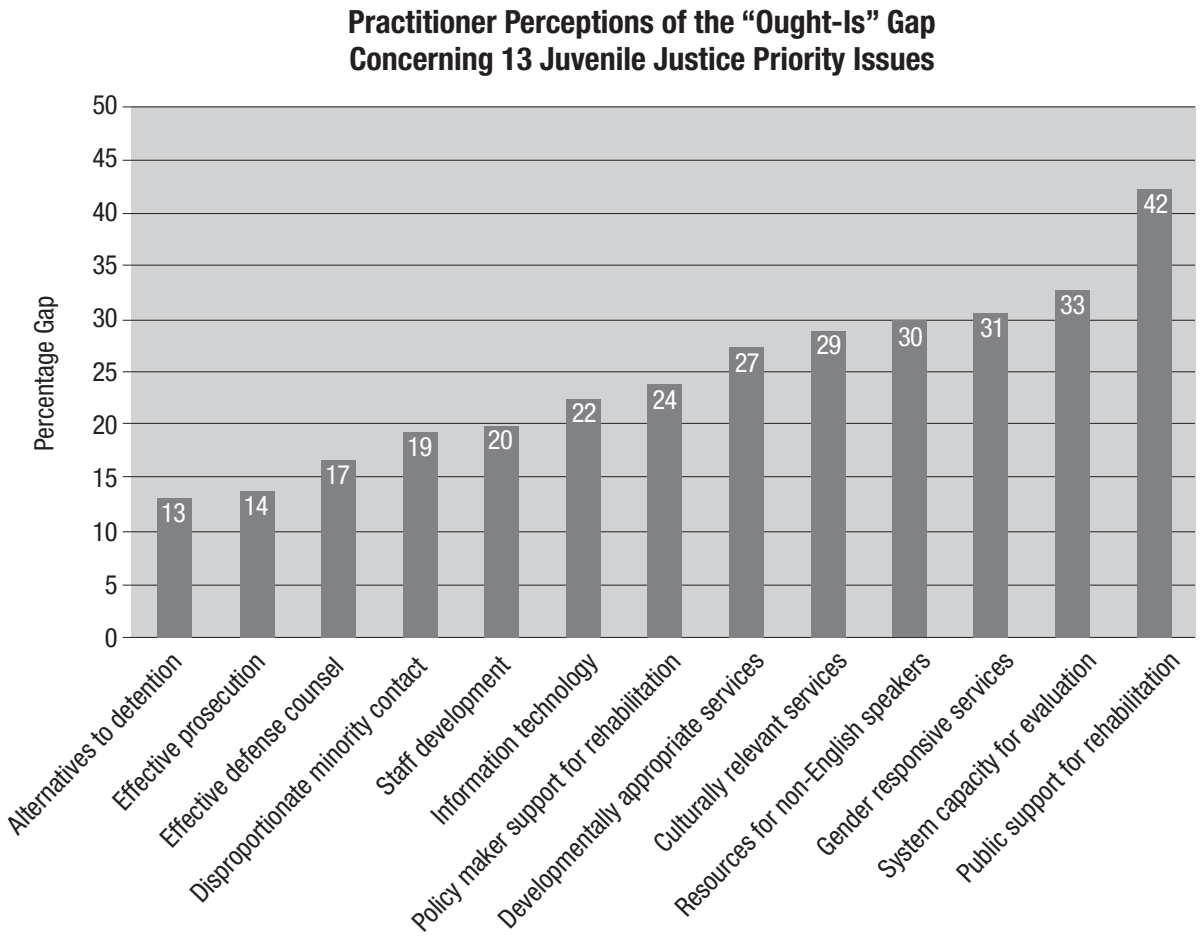
The largest "ought-is" gap that respondents perceived was with respect to the relative emphasis on rehabilitation.

An additional challenge that lies ahead is to reconcile the professional view of critical priorities for juvenile justice with the public support of offender rehabilitation. A 2007 national survey of juvenile court practitioners—which included judges, prosecutors, public defenders, and court administrators in each of the nation's 300 most-populated counties—measured how juvenile justice professionals view recent changes in policy and practice affecting the juvenile justice system (Mears et al., 2010). Respondents' views were elicited about what they believe should be critical priorities in the administration of juvenile justice and the extent to which there is a disjuncture between what should be and what is on 13 juvenile justice issues. The largest "ought-is" gap that respondents perceived was with respect to the relative emphasis on rehabilitation. Specifically, practitioners as a group strongly believe that public support for rehabilitation of young offenders should be a top priority in their jurisdictions but that the public does not give it sufficient support. The second largest gap pertained to system capacity for program and policy evaluation. Clearly, these are pressing needs in many jurisdictions.

The belief that the American public is not supportive of or is opposed to the treatment of juvenile offenders is a common misconception. Cullen (2006) noted that a 2001 national survey found that 80 percent of the sample of adults thought that rehabilitation should be the goal of juvenile correctional facilities, and that more than 9 in 10 favored a variety of early intervention programs, including parent training, Head Start, and after-school programs. "The legitimacy of the rehabilitative ideal—especially as applied to youthful offenders—appears to be deeply woven into the fabric of American culture" (p. 666). Numerous other public opinion polls also show that, for juveniles, the public believes that treatment is particularly important, especially early intervention programs (Cullen, 2006; Cullen et al., 2007). Overall, the public reports

being willing to pay for juvenile rehabilitation and early intervention programs (Nagin et al., 2006) and favors "a balanced approach, one that exacts a measure of justice, protects the public against serious offenders, and makes every effort to change offenders while they are within the grasp of the state" (Cullen, 2007, p. 721).

Figure 9. Practitioner perceptions of the "ought-is" gap concerning 13 juvenile justice priority issues



Source: Mears et al., 2010

IX. Recommendations

A. To Juvenile Justice System Administrators

- Build a forward-looking administrative model, a system organized around risk management that supports the development of individualized disposition plans for offenders. Placements should be guided by a disposition matrix. The program continuum should be populated with effective intervention programs and integrated with a graduated sanctions framework. Structured decision-making tools should be used to increase system capacity for (1) better matching of offender treatment needs with effective services in comprehensive treatment plans, (2) targeting of higher risk offenders, and (3) making improvements in prevention, court, and correctional programs across the entire continuum. Having these structured decision-making tools in place, along with an automated management information system, and efficiently using them is essential for effective systemwide implementation of an evidence-based system.

B. To Legislators

- Legislate mandatory evidence-based programming for all youth services. Although the federal and some state codes have begun to move in this direction, the effort needs to be diffused. Codified criteria ensure that only evidence-based and research-based program services are implemented statewide.
- Promote meritorious theory-based and pilot programs by providing limited funding, with the expectation that some of these will achieve the higher evidence-based standard.

C. To Juvenile Justice State Advisory Groups

- Improve cross-system coordination and collaboration and spur broader juvenile justice system reform by ensuring that the structural components of risk assessment and graduated sanctions are fully implemented as a platform for evidence-based program improvements, greater system accountability, and reduction of recidivism.
- Vigorously address the overuse of confinement, especially related to compliance with the deinstitutionalization of status offenders, separation of juveniles from adults, and disproportionate minority contact core protections of the Juvenile Justice and Delinquency Prevention Act.
- Guide jurisdictions toward implementation of evidence-based programming by incorporating guidelines for evidence-based programs in requests for proposals.

D. To Judges

- Serve as a driving force to draw together the various agencies and individuals who comprise the juvenile justice system, for the purpose of developing and implementing meaningful system reform. Reforms should be guided by the National Council of Juvenile and Family Court Judges' *Juvenile Delinquency Guidelines: Improving Court Practice in Juvenile Delinquency Cases* (Publication Development Committee, 2005). Instituting these best practices will support evidence-based programming.
- Serve as a positive influence with treatment providers by establishing clear expectations for application of evidence-based juvenile justice services within your jurisdiction. Treatment providers are highly cognizant that judges order juveniles to participate in specific

treatment programs. Judges who express a preference for referring youth to programs that are aligned with current recidivism reduction research will provide the motivation needed to get youth and families the quality services they deserve.

E. To Treatment Providers

- Recognize that the research matters to youth, families, and communities. Although a few influential therapists may extol the advantages of having the freedom to deliver individualized services without the influence of research, a host of juveniles and their families have already experienced the superior benefits of evidence-informed treatment from therapists willing to adhere to that evidence, and a host of juveniles and their families are likely still waiting for services with just such benefits.
- Learn the language of risk and risk reduction as applied to the juvenile justice system. Target and intensively serve those youth deemed to be high risk by validated juvenile justice risk assessment tools. Youth exhibiting the highest risk levels need the most intensive services for the longest duration. Youth with moderate and low risk should be offered a less intensive array of services.
- Begin now to clearly articulate in concise treatment service manuals the clinical protocols and procedures being used by clinicians. One of the strongest messages coming from the research is that fidelity—the quality with which the treatment is delivered—is crucial to successful outcomes. In most organizations the question remains, “fidelity to what?” As practices are better articulated, adherence and quality can be measured and improved.

X. Conclusion

This paper has presented a framework for juvenile justice system reform that is organized around evidence-based treatment programs for juvenile offenders integrated into a comprehensive strategy for deploying those programs in a cost-effective manner that maximizes effects on recidivism. As experience with the OJJDP Comprehensive Strategy on Serious, Violent, and Chronic Juvenile Offenders has shown, the use of structured decision-making tools coupled with a system of graduated sanctions and an array of effective services results in a juvenile justice system that addresses offenders' criminogenic risks while protecting public safety. However, if the treatment services provided within the CS framework are not effective, the success of the entire system is jeopardized. By embedding the SPEP within the CS framework, we have an evidence-based, validated tool for assessing the expected effectiveness of those services and guiding improvement when they fall short. As a result, the CS/SPEP framework offers the potential for substantial improvements in the way our juvenile justice systems operate.

References

Adams, K. (2007). Abolish juvenile curfews. *Criminology and Public Policy*, 6, 663–71.

Andrews, D. A., Zinger, I., and Hoge, R. D. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28(3), 369–404.

Annie E. Casey Foundation. (2010). *The Missouri Model: Reinventing the practice of rehabilitating youthful offenders. Summary report*. Baltimore: The Annie E. Casey Foundation.

Aos, S., Phipps, P., Barnoski, R., and Lieb, R. (2001). *The comparative costs and benefits of programs to reduce crime*. Olympia: Washington State Institute for Public Policy (http://www.wsipp.wa.gov/).

APA Task Force on Evidence-Based Practice for Children and Adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care*. Washington, DC: American Psychological Association.

Barnoski, R. (2002). *Washington State's implementation of Functional Family Therapy for juvenile offenders: Preliminary results*. Document No. 02-08-1201. Olympia: Washington State Institute for Public Policy (http://www.wsipp.wa.gov/).

Barnoski, R. (2004a). *Outcome evaluation of Washington State's research-based programs for juvenile offenders*. Document No. 04-01-1201. Olympia: Washington State Institute for Public Policy (http://www.wsipp.wa.gov/).

Barnoski, R. (2004b). *Assessing risk for re-offense: Validating the Washington State Juvenile Court Assessment*. Olympia: Washington State Institute for Public Policy (http://www.wsipp.wa.gov/).

Bennett, W., Dilulio, J., Jr., and Walters, J. (1996). *Body count*. New York: Simon & Schuster.

Bishop, D. M. (2006). Public opinion and juvenile justice policy: Myths and misconceptions. *Criminology and Public Policy*, 5(4), 653–64.

Blackburn, A. G., Mullings, J. L., Marquart, J. W., and Trulson, C. R. (2007). The next generation of prisoners: Toward an understanding of violent institutionalized delinquents. *Youth Violence and Juvenile Justice*, 5, 35–56.

Blalock, B., and Arthur, P. (2006). Advocates needed to safeguard rights of youth in DOJ conditions cases. *Youth Law News*, 27(4), 1–5.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley.

Botvin, G. J., Mihalic, S. F., and Grotpeter, J. K. (1998). *Life Skills Training: Blueprints for violence prevention, Book five*. Blueprints for Violence Prevention Series (D. S. Elliott, Series Editor). Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science.

Brownstein, H. (1996). *The rise and fall of a violent crime wave: Crack cocaine and the social construction of a crime problem*. Guilderland, NY: Harrow and Heston.

Burke, C., and Pennell, S. (2001). *Breaking Cycles evaluation: A comprehensive approach to youthful offenders*. San Diego: San Diego Association of Governments.

Burns, B. J., Goldman, S. K., Faw, L., and Burchard, J. (1999). The wraparound evidence base. In B. J. Burns and S. K. Goldman (Eds.), *Promising practices in wraparound for children with serious emotional disturbances and their families: Systems of care* (pp. 77–100). Washington, DC: American Institutes for Research, Center for Effective Collaboration and Practice.

Butts, J. A., and Mears, D. P. (2001). Reviving juvenile justice in a get-tough era. *Youth and Society*, 33, 169–98.

Carver, D. (2004). Using outcome management to guide practice in the treatment of mentally ill juvenile offenders: Lessons learned at Project Empower. Perspectives, *American Probation and Parole Association*, Summer, 33–39.

Carver, D. (2005). Adolescent drug abuse: Why family therapists are becoming counselor of choice. *Addiction Professional*, November, 22–28.

Cook, P. J., and Laub, J. H. (1998). The unprecedented epidemic of youth violence. In M. Tonry and M. H. Moore (Eds.), *Youth violence* (pp. 27–64). Chicago: University of Chicago Press.

Cooper, H., Hedges, L. V., and Valentine, J. C. (Eds.) (2009). *The handbook of research synthesis and meta-analysis*. New York: Russell Sage Foundation.

Cullen, F. T. (2005). The twelve people who saved rehabilitation: How the science of criminology made a difference. *Criminology*, 43, 1–42.

Cullen, F. T. (2006). It's time to reaffirm rehabilitation. *Criminology and Public Policy*, 5, 665–72.

Cullen, F. T. (2007). Make rehabilitation corrections' guiding paradigm. *Criminology and Public Policy*, 6, 717–28.

Cullen, F. T., Bose, B. A., Jonson, C. N. L., and Unnever, J. D. (2007). Public support for early intervention: Is child saving a "habit of the heart"? *Victims and Offenders*, 2, 109–24.

Decker, T. (2010). Posted by the author at http://www.governing.com/blogs/bfc/slow-change-powerful.html, August 2.

Dilulio, J. J., Jr. (1995). Arresting ideas. *Policy Review*, 74, 12–16.

Dilulio, J. J., Jr. (1996). They're coming: Florida's youth crime bomb. *Impact*, Spring, 25–27.

Dilulio, J. J., Jr. (1997). Jail alone won't stop juvenile super-predators. *The Wall Street Journal*, June 11, A23.

Dodge, K. A. (2001). The science of youth violence prevention: Progressing from developmental epidemiology to efficacy to effectiveness to public policy. *American Journal of Preventive Medicine*, 20(1S), 63–70.

Durlak, J. A. (1998). Common risk and protective factors in successful prevention programs. *American Journal of Orthopsychiatry*, 68, 512–20.

Edelman, C. W. (2009). Testimony on the Youth PROMISE Act before the Subcommittee on Crime, Terrorism and Homeland Security of the House Committee on the Judiciary, July 15.

Ezelle, M. E. (2007). Examining the overall and offense-specific criminal career lengths of a sample of serious offenders. *Crime & Delinquency*, 53, 3–37.

Fagan, A. A., Hanson, K., Hawkins, J. D., and Arthur, M. W. (2008). Implementing effective community-based prevention programs in the community youth development study. *Youth Violence and Juvenile Justice*, 6, 256–78.

Feld, B. C. (1999). *Bad kids: Race and the transformation of the juvenile court*. New York: Oxford University Press.

Finckenauer, J. O., and Gavin, P. W. (1999). *Scared Straight: The panacea phenomenon revisited*. Prospect Heights, IL: Waveland.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., and Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa: University of South Florida, Louis de la Parte Florida Mental Health Institute, National Implementation Research Network (FMHI Publication #231).

Gottfredson, D. C., and Gottfredson, G. D. (2002). Quality of school-based prevention programs. *Journal of Research in Crime and Delinquency*, 39, 3–35.

Gottfredson, S. D., and Moriarty, L. J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency*, 52, 178–200.

Greenwood, P. W. (2010). *Preventing and reducing youth crime and violence: Using evidence-based practices*. Sacramento: Governor's Office of Gang and Youth Violence Policy.

Greenwood, P. W., Model, K. E., Rydell, C. P., and Chiesa, J. (1996). *Diverting children from a life of crime: Measuring costs and benefits*. Santa Monica, CA: Rand.

Grove, W. M., Eckert, E. D., Heston, L., Bouchard, T. J., Segal, N., and Lykken, D. T. (1990). *Clinical vs. mechanical prediction: A meta-analysis*. Minneapolis: Department of Psychology, University of Minnesota.

Grove, W. M., and Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, 2, 293–323.

Hallfors, D., and Godette, D. (2002). Will the "principles of effectiveness" improve prevention practice? Early findings from a diffusion study. *Health Education Research*, 17, 461–70.

Howell, J. C. (Ed.). (1995). *Guide for implementing the comprehensive strategy for serious, violent, and chronic juvenile offenders*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Howell, J. C. (2003a). Diffusing research into practice using the comprehensive strategy for serious, violent, and chronic juvenile offenders. *Youth Violence and Juvenile Justice: An Interdisciplinary Journal*, 1, 219–45.

Howell, J. C. (2003b). *Preventing and Reducing Juvenile Delinquency: A Comprehensive Framework*. Thousand Oaks, CA: Sage.

Howell, J. C. (2009). *Preventing and reducing juvenile delinquency: A comprehensive framework* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Johnson, K., and Wagner, D. (2001). *Missouri multi-county juvenile officer workload study report*. Madison,WI: National Council on Crime and Delinquency.

Johnson, K., Wagner, D., and Matthews, T. (2001). *Missouri juvenile risk assessment re-validation report*. Madison, WI: National Council on Crime and Delinquency.

Karoly, L. A., Greenwood, P. W., Everingham, S. S., Houbé, J., Kilburn, M. R., and Rydell, C. P. (1998). *Investing in our children*. Santa Monica, CA: RAND.

Kelly, J. (2010). JJDP noncompliance rises in 2009. *Youth Today*, August.

Knitzer, J. (1982). *Unclaimed children: The failure of public responsibility to children and adolescents in need of mental health services*. Washington, DC: Children's Defense Fund.

Knitzer, J., and Cooper, J. (2006). Beyond integration: Challenges for children's mental health. *Health Affairs*, 25(3), 670–79.

Landenberger, N. A., and Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4), 451–76.

Latimer, J. (2001). A meta-analytic examination of youth delinquency, family treatment, and recidivism. *Canadian Journal of Criminology*, 43(2), 237–53.

Lattimore, P. K., Macdonald, J. M., Piquero, A. R., Linster, R. L., and Visher, C. A. (2004). Studying the characteristics of arrest frequency among paroled youthful offenders. *Journal of Research in Crime and Delinquency*, 41, 37–57.

LeCroy, C. W., Krysik, J., and Palumbo, D. (1998). *Empirical validation of the Arizona Risk/Needs Instrument and assessment process*. Tucson: LeCroy & Milligan Associates.

Lipsey, M. W. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartman, L. V. Hedges, R. J. Light, et al. (Eds.), *Meta-analysis for explanation* (pp. 83–127). New York: Russell Sage Foundation.

Lipsey, M. W. (1999a). Can intervention rehabilitate serious delinquents? *Annals of the American Academy of Political and Social Science*, 564, 142–66.

Lipsey, M. W. (1999b). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *Virginia Journal of Social Policy and the Law*, 6, 611–41.

Lipsey, M. W. (2008). *The Arizona Standardized Program Evaluation Protocol (SPEP) for assessing the effectiveness of programs for juvenile probationers: SPEP ratings and relative recidivism reduction for the initial SPEP sample*. Nashville: Vanderbilt University, Center for Evaluation Research and Methodology.

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., and Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3, 297–320.

Lipsey, M. W., Howell, J. C., and Tidd, S. T. (2007). *The Standardized Program Evaluation Protocol (SPEP): A practical approach to evaluating and improving juvenile justice programs in North Carolina. Final evaluation report*. Nashville: Vanderbilt University, Center for Evaluation Research and Methodology.

Lipsey, M. W., and Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber and D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 313–45). Thousand Oaks, CA: Sage Publications, Inc.

Lipsey, M. W., and Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Lipton, D., Martinson, R., and Wilks, J. (1975). *The effectiveness of correctional treatment: A survey of treatment evaluation studies*. New York: Praeger.

Littell, J. H., Popa, M., and Forsythe, B. (2005). Multisystemic Therapy for social, emotional, and behavioral problems in youth aged 10–17. *The Campbell Collaboration Library* (http://www.campbellcollaboration.org/frontend.asp).

Loeber, R., and Farrington, D. P. (Eds.). (1998). *Serious and violent juvenile offenders: Risk factors and successful interventions*. Thousand Oaks, CA: Sage.

MacKenzie, D. L., Wilson, D. B., Armstrong, G. S., and Gover, A. R. (2001). The impact of boot camps and traditional institutions on juvenile residents: Perceptions, adjustment, and changes in social bonds, impulsivity, and antisocial attitudes. *Journal of Research in Crime and Delinquency*, 38, 279–313.

MacKenzie, D. L., Wilson, D. B., and Kider, S. B. (2001). Effects of correctional boot camps on offending. *Annals of the American Academy of Political and Social Science*, 578, 126–43.

Males, M. A. (1996). *The scapegoat generation: America's war on adolescents*. Monroe, ME: Common Courage Press.

Martinson, R. (1974). What works? Questions and answers about prison reform. *Public Interest*, 35, 22–54.

McCord, J., Widom, C. S., and Crowell, N. A. (2001). *Juvenile crime, juvenile justice*. Washington, DC: National Academy Press.

McDowall, D., Loftin, C., and Wiersema, B. (2000). The impact of youth curfew laws on juvenile crime rates. *Crime & Delinquency*, 46(1), 76–91.

McGarvey, E. L., and Waite, D. (2000). *Profiles of incarcerated adolescents in Virginia correctional facilities: Fiscal years 1993–1998*. Richmond, VA: Department of Criminal Justice Services.

Mears, D. P. (2002). Sentencing guidelines and the transformation of juvenile justice in the twenty-first century. *Journal of Contemporary Criminal Justice*, 18(1), 6–19.

Mears, D. P., Shollenberger, T. L., Willison, J. B., Owens, C. E., and Butts, J. A. (2010). Practitioner views of priorities, policies, and practices in juvenile justice. *Crime & Delinquency*, 56(4), 535–63.

Mihalic, S., Irwin, K., Elliott, D., Fagan, A., and Hansen, D. (2001). *Blueprints for violence prevention* (Juvenile Justice Bulletin). Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Miller, W. R., Sorensen, J. L., Selzer, J. A., and Brigham, J. S. (2006). Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment*, 35, 25–39.

Nagin, D. S., Piquero, A. R., Scott, E. S., and Steinberg, L. (2006). Public preferences for rehabilitation versus incaraceration of juvenile offenders: Evidence from a contingent valuation study. *Criminology and Public Policy*, 5, 627–52.

North Carolina Department of Juvenile Justice and Delinquency Prevention. (2008). *2007 Annual report*. Raleigh: Author.

Nugent, W. R., Williams, M., and Umbreit, M. S. (2004). Participation in victim-offender mediation and the prevalence of subsequent delinquent behavior: A meta-analysis. *Utah Law Review*, 137(1), 137–66.

Office of State Courts Administrator. (2000). *Report on standards for the administration of juvenile justice*. Jefferson City, MO: Juvenile and Adult Court Programs Division, Office of State Courts Administrator.

Office of State Courts Administrator. (2002). *Missouri's juvenile offender risk and needs assessment and classification system: User manual*. Jefferson City, MO: Juvenile and Adult Court Programs Division, Office of State Courts Administrator.

Paulson, R., Fixsen, D., and Friedman, R. (2004). *An analysis of implementation of systems of care at fourteen CMHS grant communities*. Tampa: Louis de la Parte Florida Mental Health Institute, University of South Florida.

Petrosino, A., Turpin-Petrosino, C., and Buehler, J. (2003). Scared Straight and other juvenile awareness programs for preventing juvenile delinquency: A systematic review of the randomized experimental evidence. *Annals of the American Academy of Political and Social Science*, 589, 41–62.

Petrosino, A., Turpin-Petrosino, C., and Finckenauer, J. (2000). Well-meaning programs can have harmful effects! Lessons from experiments of programs such as Scared Straight. *Crime & Delinquency*, 46, 354–79.

Petrosino, A., Turpin-Petrosino, C., and Guckenburg, S. (2010). Formal system processing of juveniles: Effects on delinquency. *Campbell Systematic Reviews* (www.campbellcollaboration.org).

Publication Development Committee (2005). *Juvenile delinquency guidelines: Improving court practice in juvenile delinquency cases*. Reno, NV: The National Council of Juvenile and Family Court Judges.

Redpath, D. P., and Brandner, J. K. (2010). *The Arizona Standardized Program Evaluation Protocol (SPEP) for assessing the effectiveness of programs for juvenile probationers*. Phoenix: Arizona Supreme Court, Administrative Office of the Courts, Juvenile Justice Service Division.

Reeves, J. L., and Campbell, R. (1994). *Cracked coverage: Television news, the anti-cocaine crusade, and the Reagan legacy*. Durham, NC: Duke University.

Rosenbaum, D. P. (2007). Just say no to D.A.R.E. *Criminology and Public Policy*, 6, 815–24.

Roush, D. W., and McMillen, M. (2000). *Construction, operations, and staff training for juvenile confinement facilities* (Juvenile Accountability Incentive Block Grants Program Bulletin). Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Rubin, H. T. (2000). Teen Quest: Female-specific program services for Colorado's delinquent girls. *Juvenile Justice Update*, 6, 1–16.

Schumacher, M., and Kurz, G. (2000). *The 8% solution: Preventing serious, repeat juvenile crime*. Thousand Oaks, CA: Sage.

Scott, R. C. (2009). Youth PROMISE Act White Paper: Fighting juvenile crime vs. "playing politics." Washington, DC: U.S. House of Representatives, Office of Congressman Robert C. "Bobby" Scott. Access: http://www.bobbyscott.house.gov.

Sedlak, A. J., and McPherson, K. S. (2010). *Conditions of confinement: Findings from the survey of youth in residential placement* (Juvenile Justice Bulletin). Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Smith, W. R., and Aloisi, M. F. (1999). Prediction of recidivism among "second timers" in the juvenile justice system: Efficiency in screening chronic offenders. *American Journal of Criminal Justice*, 23, 201–22.

Snyder, H. N. (1998). Serious, violent, and chronic juvenile offenders: An assessment of the extent of and trends in officially recognized serious criminal behavior in a delinquent population. In R. Loeber and D. P. Farrington (Eds.), *Serious and violent juvenile offenders: Risk factors and successful interventions* (pp. 428–44). Thousand Oaks, CA: Sage Publications, Inc.

Snyder, H. N., and Sickmund, M. (2000). *Challenging the myths* (Juvenile Justice Bulletin). 1999 National Report Series. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Snyder, H. N., and Sickmund, M. (2006). *Juvenile offenders and victims: 2006 national report*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Tanenhaus, D. S. (2002). The evolution of juvenile courts in the early twentieth century: Beyond the myth of immaculate construction. In M. K. Rosenheim, F. E. Zimring, and D. S. Tanenhaus (Eds.), *A century of juvenile justice* (pp. 42–73). Chicago: University of Chicago Press.

Tanenhaus, D. S. (2004). *Juvenile justice in the making*. New York: Oxford University Press.

Tonry, M. (2004). *Thinking about crime: Sense and sensibility in American penal culture*. New York: Oxford University Press.

Tonry, M. (2009). Explanations of American punishment policies: A national history. *Punishment and Society*, 11, 377–94.

Tonry, M., and Melewski, M. (2008). The malign effects of drug and crime control policies on black Americans. In M. Tonry (Ed.), *Crime and justice: A review of research* (vol. 37) (pp. 1–44). Chicago: University of Chicago Press.

Torbet, P., and Szymanski, L. (1998). *State legislative responses to violent juvenile crime: 1996–1997 update* (Juvenile Justice Bulletin). Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Trulson, C. R., Marquart, J. W., Mullings, J. L., and Caeti, T. J. (2007). In between adolescence and adulthood: Recidivism outcomes for a cohort of state delinquents. *Youth Violence and Juvenile Justice*, 3, 355–77.

U.S. Department of Health and Human Services. (2001). *Youth violence: A report of the surgeon general*. Rockville, MD: Author.

U.S. Department of Justice. (2007). *Department of Justice activities under the Civil Rights of Institutionalized Persons Act: Fiscal year 2006*. Washington, DC: Office of the Attorney General, U.S. Department of Justice.

Warren, R. (2007). *Evidence-based practice to reduce recidivism: Implications for state judiciaries*. Williamsburg, VA: National Center for State Courts, National Institute of Corrections.

Weisz, J., Sandler, I., Durlak, J., and Anton, B. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist*, 60, 628–48.

Weithorn, L. A. (1988). Mental hospitalization of troublesome youth: An analysis of skyrocketing admission rates. *Stanford Law Review*, 40, 773–838.

Welsh, B. C., Sullivan, C. J., and Olds, D. L. (2010). When early crime prevention goes to scale: A new look at the evidence. *Prevention Science*, 11, 115–25.

Wiebush, R. G. (2000). *Risk assessment and classification for serious, violent, and chronic juvenile offenders*. Madison, WI: National Council on Crime and Delinquency.

Wiebush, R. G. (Ed.). (2002). *Graduated sanctions for juvenile offenders: A program model and planning guide*. Oakland, CA: National Council on Crime and Delinquency and National Council of Juvenile and Family Court Judges.

Wiebush, R. G., Johnson, K., and Wagner, D. (1997). *Development of an empirically based risk assessment instrument and placement recommendation matrix for the Maryland Department of Juvenile Justice*. Madison, WI: National Council on Crime and Delinquency.

Wiebush, R. G., Wagner, D., and Erlich, J. (1999). *Development of an empirically based assessment instrument for the Virginia Department of Juvenile Justice*. Madison, WI: National Council on Crime and Delinquency.

Wilson, D. B., MacKenzie, D. L., and Mitchell, F. N. (2005). Effects of correctional boot camps on offending. *Campbell Systematic Reviews* (http://campbellcollaboration.org/lib/download/3/).

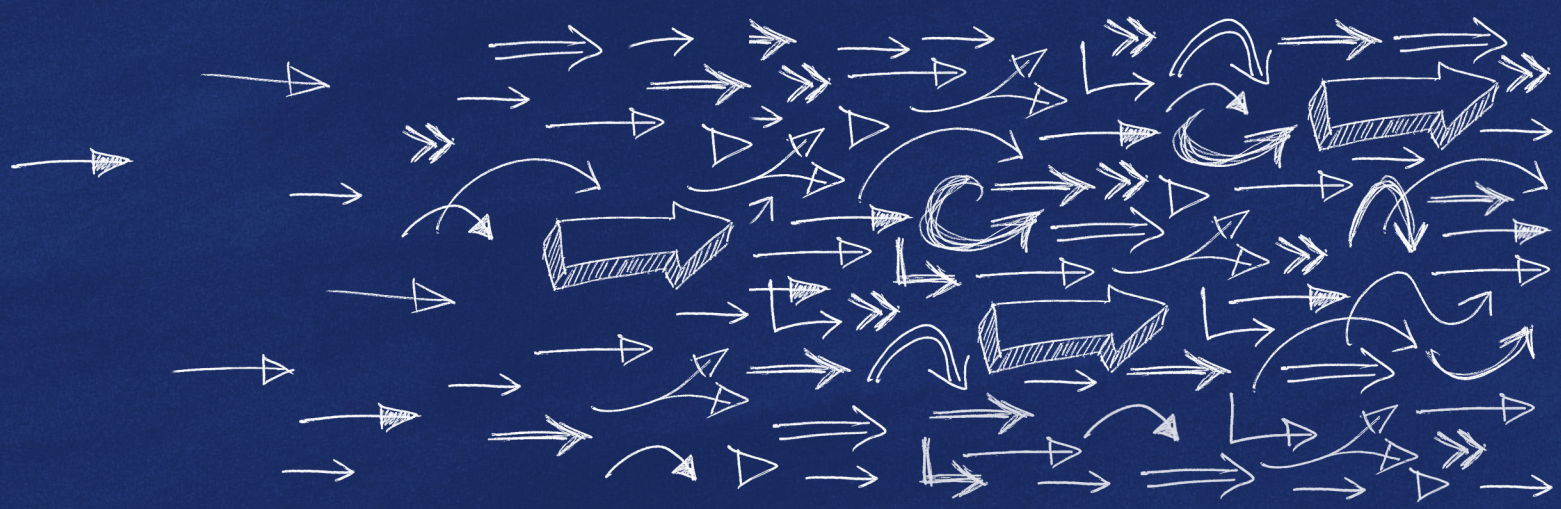
Wilson, D. B., Mitchell, O., and MacKenzie, D. L. (2006). A systematic review of drug court effects on recidivism. *Journal of Experimental Criminology*, 2, 459–87.

Wilson, J. J., and Howell, J. C. (1993). *A comprehensive strategy for serious, violent, and chronic juvenile offenders*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Wilson, J. J., and Howell, J. C. (1994). OJJDP's comprehensive strategy for serious, violent, and chronic juvenile offenders. *Juvenile and Family Court Journal*, 45, 3–12.

Wilson, J. Q. (1995). Crime and public policy. In J. Q.Wilson and J. Petersilia (Eds.), *Crime* (pp. 489–507). San Francisco: ICS Press.

Zimring, F. E. (1998). Toward a jurisprudence of youth violence. In M. Tonry and M. H. Moore (Eds.), *Youth violence* (24th ed., pp. 477–501). Chicago: University of Chicago.

Center for Juvenile Justice Reform
Georgetown Public Policy Institute
Georgetown University
Box 571444
3300 Whitehaven Street, NW, Suite 5000
Washington, D.C. 20057

http://cjjr.georgetown.edu
jjreform@georgetown.edu