

Florida State University Libraries

2011

Evidence on the Effectiveness of Juvenile Court Sanctions

Daniel P. Mears, Joshua C. Cochran, Sarah J. Greenman, Avinash S. Bhati and Mark A. Greenwald



PRINT VERSION CITATION: Mears, Daniel P., Joshua C. Cochran, Sarah J. Greenman, Avinash S. Bhati, and Mark A. Greenwald. 2011. "Evidence on the Effectiveness of Juvenile Court Sanctions." *Journal of Criminal Justice* 39(6):509-520.

PRE-PRINT VERSION

Evidence on the Effectiveness of Juvenile Court Sanctions*

Daniel P. Mears, Joshua C. Cochran, Sarah J. Greenman,
Avinash S. Bhati, and Mark A. Greenwald

* Direct correspondence to Daniel P. Mears, Ph.D., Mark C. Stafford Professor of Criminology, Florida State University, College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, e-mail (dmears@fsu.edu), phone (850-644-7376), fax (850-644-9614). Joshua C. Cochran, M.S., doctoral candidate, Florida State University, College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, phone (850-644-4050), fax (850-644-9614), e-mail (jcochran@fsu.edu). Sarah J. Greenman, M.A., doctoral student, University of Maryland's Department of Criminology and Criminal Justice, 2220 LeFrak Hall, College Park, MD 20742, phone (301-405-4699), fax (301-405-4733), e-mail (sgreenma@umd.edu). Avinash Bhati, Ph.D., President, Maxarth, 509 Cedar Spring Street, Gaithersburg, MD 20877, phone (301-789-2966), fax (301-880-7152), e-mail (abhati@maxarth.com). Mark A. Greenwald, M.S., Chief of Research and Planning, Florida Department of Juvenile Justice, 2737 Centerview Drive, Tallahassee, Florida 32399, phone (850-921-0706), fax (850-922-2992), e-mail (mark.greenwald@djj.state.fl.us). Partial support for developing this paper came from grant #2010-JF-FX-0620, awarded by the U.S. Office of Juvenile Justice and Delinquency Prevention. Points of view in this paper are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice or of the Florida Department of Juvenile Justice.

BIOGRAPHICAL SKETCHES

Daniel P. Mears, Ph.D., is the Mark C. Stafford Professor of Criminology at Florida State University's College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, phone (850-644-7376), fax (850-644-9614), e-mail (dmears@fsu.edu). He conducts basic and applied research on crime and justice. His work has appeared in *Criminology*, the *Journal of Research in Crime and Delinquency*, and other crime and policy journals and in a book, *American Criminal Justice Policy* (Cambridge University Press 2010).

Joshua C. Cochran, M.S., is a doctoral candidate at Florida State University's College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, phone (850-644-4050), fax (850-644-9614), e-mail (jcochran@fsu.edu). His research interests include theory, international comparative analyses of criminal justice systems, and prisoner reentry.

Sarah J. Greenman, M.A., is a doctoral student at the University of Maryland's Department of Criminology and Criminal Justice, 2220 LeFrak Hall, College Park, MD 20742, phone (301-405-4699), fax (301-405-4733), e-mail (sgreenma@umd.edu). Her research examines procedural justice, victims, law enforcement, and court processing.

Avinash S. Bhati, Ph.D., is the President of Maxarth, 509 Cedar Spring Street, Gaithersburg, MD 20877, phone (301-789-2966), fax (301-880-7152), e-mail (abhati@maxarth.com). He has conducted research on a crime and justice programs and policies and has published in *Criminology*, the *Journal of Quantitative Criminology*, the *Journal of Criminal Law and Criminology*, and other prominent journals. He serves on the editorial board of the *Journal of Quantitative Criminology*.

Mark A. Greenwald, M.S., is the Chief of Research and Planning at the Florida Department of Juvenile Justice, 2737 Centerview Drive, Tallahassee, Florida 32399, phone (850-921-0706), fax (850-922-2992), e-mail (mark.greenwald@djj.state.fl.us). He conducts research and evaluation on Florida juvenile justice processing, programming, practices, and policies, and has published many reports as well as in *Youth Violence and Juvenile Justice*.

Evidence on the Effectiveness of Juvenile Court Sanctions

ABSTRACT

Purpose: The past decade has been witness to a proliferation of calls for evidence-based juvenile court sanctions—including various programs, interventions, services, and strategies or approaches—that reduce recidivism and improve mental health, drug dependency, and education outcomes. At the same time, an emerging body of work has identified “proven,” “evidence-based,” “best practice,” or, more generally, “effective” efforts to achieve these outcomes. Even so, grounds for concern exist regarding the evidence-base for these and other sanctions.

Methods: This paper describes the heterogeneity of sanctioning within juvenile justice and argues that, despite substantial advances in research, the heterogeneity severely delimits the generalizability of evaluations to date. It also raises questions about how much is in fact known about the effectiveness of many juvenile justice sanctions.

Conclusion: Extant research offers grounds for optimism. Even so, explicit articulation of the limitations of this research and the need for studies that examine external validity is important for developing evidence about “what works” in juvenile justice. Implications for research and policy are discussed.

KEYWORDS: juvenile court justice sanctions evidence-based effectiveness

INTRODUCTION

As envisioned by the founders of the juvenile justice system, delinquent youth should receive sanctions that both hold them accountable and rehabilitate them (Tanenhaus 2004; Bernard and Kurlychek 2010; Feld and Bishop 2011a). In recent decades, states increasingly have placed a greater premium on accountability, largely by imposing more severe sanctions (Howell 2009; Fagan 2010). States also have continued to promote—in theory if not in practice—the juvenile court’s broader and central mandate of providing for the “best interests” of youth (Butts and Mears 2001; Scott and Steinberg 2008; Merlo and Benekos 2010). Alongside of these trends has been an increased emphasis on evidence-based policies, programs, and practices (Sherman et al. 2002; MacKenzie 2006; Weiss et al. 2008; Drake et al. 2009; Greenwood and Turner 2011a-b).

Such changes suggest the need to assess the state of evidence for the broad range of sanctions that have always existed in juvenile court and that, if anything, have become even more wide-ranging in recent decades as states and local jurisdictions have devised ever-new sets of approaches for handling delinquent youth (Howell 2009; Feld and Bishop 2011a). This need is greater than in the adult justice system given the juvenile court’s mandate to create sanctions that serve not only to punish youth but also to promote the “best interests” of youth (Krisberg and Austin 1993; Butts and Mears 2001; Tanenhaus 2004). This mandate has resulted in a heterogeneous set of sanctions—including various punishments, programs, interventions, services, and strategies or approaches—that are supposed to simultaneously reduce recidivism and improve mental health, drug dependency, education, and other outcomes.

The scholarly community has been responsive to this situation. For example, many studies, meta-analyses, and reviews have been conducted that discuss different types of sanctions and identify them as being “proven,” “evidence-based,” “best practices,” or, more generally, “effective” (see, e.g., Sherman et al. 2002; MacKenzie 2006; Farrington and Welsh 2007; Lipsey and Cullen 2007; Drake et al. 2009; Howell 2009; Lipsey 2009; Greenwood and Turner 2011a-b; Henggeler and Schoenwald 2011; Krisberg 2011; MacKenzie and Freeland 2011). Indeed, one

recent review opined that the state of evidence is sufficient to conclude: “If [a] program does not include any of the proven strategies or follow the principles listed [in our review], then some of those strategies and principles must be added to the program for it to be considered ‘evidence-based’” (Greenwood and Turner 2011b:741).

However, many of these studies and reviews focus primarily on assessing the internal validity of a given sanction or intervention. That is, they examine whether the methodological rigor of extant studies allow for defensible claims that particular sanctions or interventions caused particular effects (Shadish et al. 2002; Farrington 2003). Viewed in this light, clear warrant for optimism exists. For example, many sanctions, including various programs and interventions, are described as “evidence-based” in the sense that at least some methodologically rigorous studies have showed that they produced, for the populations and in the settings where they occurred, positive impacts. What remains less clear, however, is the external validity—that is, the generalizability of an identified causal relationship across different populations and settings (Cook and Campbell 1979)—of these same sanctions. Put differently, can we expect the evaluated sanctions to produce comparable results with other populations and/or in other places?

This situation stems in part from debates within juvenile and criminal justice and, more broadly, social and health policy research, about the relative priority of internal versus external validity (McCall and Green 2004; Rothwell 2005; Steckler and McLeroy 2008; Sampson 2010). Historically, greater emphasis has been placed on the former, but the arguments put forth by Cronbach (1982) and others have led to greater attention to external validity (see, for example, Bardach 2004; Sampson 2010; Chen et al. 2011). For example, although there is no disputing the need for studies that establish whether a given sanction or intervention causes an effect, such work tells us little about whether these effects can be expected among populations or places that differ from those in extant studies. To illustrate, an evaluation might find that a drug court in a large city is effective in reducing offending among hard-core drug addicts who come from predominantly African-American communities (Gottfredson et al. 2003). However, such an evaluation leaves unanswered the question of whether this effect can be expected among a rural

drug-using offender population with different or less serious drug addiction problems or who have different racial, ethnic, or cultural backgrounds. For policymakers and justice system officials and practitioners, this question is critical. They typically want to know whether a particular approach will be effective for the populations or places in their jurisdiction or locale.

Unfortunately, even when we can have confidence in the internal validity of a set of studies that have evaluated a particular sanction, program, or intervention, we can not be confident in the external validity of the studies' findings. This problem is, if anything, greater in juvenile justice, as compared to criminal justice, because of the many sanctions and interventions available in juvenile court (Howell 2009; Greenwood and Turner 2011a-b; Krisberg 2011; MacKenzie and Freeland 2011). Specifically, the variability in sanctioning across jurisdictions raises questions about the extent to which estimated effects from particular studies generalize to settings that operate with different sets or amounts of sanctions, that use their sanctions for different populations, or that operate under different social, cultural, or financial conditions.

Against this backdrop, the goal of this paper is to contribute to research and policy discussions aimed at improving the effectiveness of the juvenile justice system. We first describe the heterogeneity in juvenile court sanctions. Second, we argue that the considerable heterogeneity in juvenile court sanctioning has profound implications for identifying “best” or “effective” approaches to reducing recidivism and achieving the other myriad objectives of juvenile courts. Specifically, it significantly delimits the potential generalizability of evaluations of juvenile justice sanctions. At the same time, it creates substantial challenges to creating a body of evidence about the extent to which various sanctions can reasonably be expected to consistently produce anticipated effects among diverse populations and places. We conclude by identifying directions for future research and policy.

HETEROGENEITY AMONG TYPES OF JUVENILE SANCTIONS

According to the historical mission of the juvenile court, youth should receive individualized

sanctions that reflect their unique needs, problems, and social contexts (Tanenhaus 2004; Sanborn and Salerno 2005; Fagan 2010). There may, of course, be disjunctures between this ideal and actual practice (Emerson 1969; Bortner 1982; Feld 1999; Kupchik 2006; Mears 2010; Welsh et al. 2010). Nonetheless, it is the ideal that arguably accounts for the considerable heterogeneity within and across juvenile court sanctions. Among other things, and in contrast to the adult court, it reflects the juvenile court's mission of providing not only a punishment response to offending but also a social welfare response guided by the notion of *parens patriae* (i.e., that "state as parent") (Simon 1995; Feld 1999; Tanenhaus 2004). It also reflects, again in contrast to the adult court, an interest in informal processing to arrive at sanctions—including various interventions, programs, and services—that will serve the "best interests" of youth. The focus on individualized sentencing means that youth may receive any of a range of different sanctions. As Snyder and Sickmund (2006:106) have noted, "most juvenile dispositions are multifaceted" and typically include, among other things, "some sort of supervised probation."¹

Despite the multiple and varied nature of the sanctions that youth may receive, national juvenile court processing statistics use a small handful of categories to depict the sanctions that youth receive: dismissal or release, other sanction, probation, or placement in a residential or custodial facility (Snyder and Sickmund 2006:177). These sanction categories have heuristic utility in providing a general perspective on the types of sanctions that youth receive. However, as will be discussed in the next section, there is considerable variation within each category. In addition, youth may simultaneously receive multiple sanctions. That makes it difficult to isolate the net effect of any one general type of sanction, much less the effect of any specific example of a type of sanction (e.g., placement in a specific type of low-risk residential treatment facility versus placement in a specific type of high-risk facility). Before discussing these issues, we first draw on national data to describe the general types of sanctions that exist in juvenile court and to highlight the fact that many of these sanctions result from informal processing. In the next section, we then describe the heterogeneity not only among these different types of sanctions but also within them and, in turn, the implications of such heterogeneity.

Insert table 1 about here

In 2007, the latest year for which national statistics exist, 1,666,100 delinquency cases were referred to juvenile court (Knoll and Sickmund 2010). As shown in table 1, of all these cases (excluding transfers), 68 percent (1,123,300) received a sanction; the rest (534,300) were dismissed. It might reasonably be assumed that formal processing—wherein a delinquency petition is filed with the court—would constitute the primary avenue through which sanctions arise and that, perforce, arise in situations where the offender is found guilty or, in the language of the juvenile court, “adjudicated delinquent.” In reality, formal processing and informal processing (in which no petition is filed with the court) contribute equally to the sanctioning of youth (Mears 2011). For example, formal processing contributed to just half (52 percent) of all sanctioned delinquency cases. The other half (48 percent) stemmed either from formal processing that resulted in no adjudication of youth as delinquent or from informal processing.² Notably, informally processed youth may receive similar sanctions or interventions to what formally processed youth may receive—including probation, restitution, community service, and referral for services (Snyder and Sickmund 2006:177)—though the amount and quality of these may vary. For example, with informal agreements, probation staff may monitor a juvenile’s compliance with the agreement but may provide little direct supervision.

The national statistics illustrate that juvenile court sanctions and interventions are heterogeneous, if only in the way in which they arise. For example, there are three broad categories of youth on probation: those who receive probation through informal processing (253,800 youth); those who receive probation through formal processing and yet were not adjudicated delinquent (61,100 youth); and those who receive probation through formal processing and were adjudicated delinquent (327,400). The important consideration here, however, is not simply that youth arrive at probation through different pathways. It is also that the conditions of probation and the intensity and quality of supervision, as well as services,

treatments, and programming, may vary as well. Put differently, to speak of “probation” is in fact to refer to a category that itself encompasses considerable heterogeneity.

The national statistics illustrate that “other sanctions,” as a general category, comprise almost 30 percent (332,400) of all sanctions administered by juvenile courts. Here, again, this category obscures considerable heterogeneity in the types, amount, and quality of punishments, treatments, services, and programming that youth receive.

Only 13 percent of all sanctioned delinquency cases are placed in some type of custodial facility and fewer than 1 percent are judicially transferred to adult court. The relatively small percent of sanctioned youth who receive custodial placements underscores that any discussion of juvenile court sanction effectiveness ideally should examine a range of sanctions and not just those that typically are viewed as the “deep end” of the juvenile justice system. That said, incarceration stands as one of the most prominent types of sanctions that juvenile courts provide and so merits attention. Here, again, heterogeneity is the rule rather than the exception. Custodial facilities can, for example, consist of detention centers, shelters, training schools, ranches, camps, and group and foster homes; they can be state or local, public or private, large or small, or low to maximum security; they can be specialized (e.g., focusing primarily on sex offenders or murderers); they can emphasize rehabilitation more so than punishment or punishment more so than rehabilitation; they can be overcrowded or under capacity; and so on (Hockenberry et al. 2011; Krisberg 2011). And, of course, all of these dimensions may intersect with one another. For example, there can be overcrowded facilities that are small and those that are large. Among overcrowded small facilities, there may be those with highly professional approaches to providing supervision and treatment and those that are less professional. The permutations are many and constitute the norm rather than the exception.

In short, sanctioning heterogeneity is ubiquitous in juvenile court and is so even when we focus on the small set of categories used to create national portraits of the disposition of delinquency cases. There is heterogeneity in whether cases are informally or formally processed and there is heterogeneity in the types of sanctions received (regardless of the informal or formal

mode of processing). In addition, and as will be further discussed, there is heterogeneity within each type of sanction (e.g., probation, “other,” custodial placement); there also is variation over time and within and across states and jurisdictions in the use of the different types of sanctions. As we will argue, such heterogeneity severely delimits the external validity of the effects identified in many studies of juvenile court sanctions, including various programs and interventions. That is, it is unclear whether comparable effects will arise with different populations or in different settings or conditions.

HETEROGENEITY WITHIN TYPES OF JUVENILE SANCTIONS

The juvenile court, as envisioned by its founders, is to be guided both by individualized sanctioning, involving dual punishment and social welfare emphases, and by a progression in sanctioning. Stated differently, there should be a graduated continuum of sanctioning, or a “disposition ladder,” one where youth receive the least restrictive and invasive sanction possible (Sanborn and Salerno 2005:368; see, generally, Holsinger and Latessa 1999; Howell 2009; Cooley 2011). More restrictive and invasive sanctions should be reserved for youth who more frequently engage in delinquency. The idea is that fewer restrictions may be more effective for first-time or infrequent offenders; the restrictions may deter more effectively, allow for more or better rehabilitation opportunities, and cost less. Whatever the merits of these assumptions, one result is a highly diverse set of sanctions, some administered as stand-alone sanctions and others administered in addition to or subsequent to other sanctions.

Table 2 presents a matrix of the typical types of sanctions used in juvenile court; they flow roughly from the least restrictive and invasive (e.g., probation) to the most restrictive and invasive (e.g., placement in custodial facilities).³ The listing, which expands on the categories used in national portraits of delinquency case processing (Snyder and Sickmund 2006), is far from exhaustive. Even so, it presents the types of dispositions typically found in most juvenile courts (see, e.g., Emerson 1969; Bortner 1982; Sanborn and Salerno 2005; Lopez and Russell

2008; Howell 2009). The table cannot, perforce, include every example of each type. The diversity of “regular probation” is illustrative. The precise amount and quality of supervision, the number of meetings with probation officers, the extent to which probation officers attempt to provide rehabilitative assistance—these and more may vary greatly within and across jurisdictions. There can be additional diversity within subtypes as well. For example, many specialized courts, such as teen and drug courts, are highly variable in their structure and operations (Butts et al. 2002; Butts and Roman 2004).

For these reasons, there exists no simple way to capture the full range of within-sanction category variation. Nonetheless, below we briefly describe some of the variability within each type of sanction. Any classification scheme necessarily obscures some variation that may exist between instances of a particular category within the classification scheme. Table 2 serves to highlight this issue and to illustrate the marked heterogeneity that can exist within any given type of sanction category. The first column presents the highest level of detail, the second a moderate level of detail, the third a relatively low level of detail, and the fourth the least amount of detail, as conveyed by the presence of only two types of sanctions (placement in a facility of some type versus dismissal or receipt of some non-commitment sanction, such as probation).

Insert table 2 about here

Diversion

We begin first with diversion. This category, one used in many studies, can include a wide panoply of programs, treatments, services, and interventions. Some may be viewed less as sanctions than as prods to obtain needed assistance, but many others may be viewed as sanctions because the juvenile court retains jurisdiction over the case, may require participation in the diversion effort, or may allow for supervision and for penalties if participation is viewed as inadequate (Klein 1979; Blomberg 1980; Polk 1984; Roberts 2004; Mears 2011).

There are many ways of classifying the types of diversion. Many of the programs and interventions listed in Table 3 and discussed further below may be viewed as instances of diversion. Klein (1979), for example, has described four categories of diversion treatment and services: counseling, advocacy, opportunity enhancement, and skill development. Here, for illustrative purposes only, Table 2 depicts two broad sub-categories of diversion: (1) delinquency prevention efforts aimed at providing various types of services and treatments that target criminogenic factors, and (2) child welfare efforts aimed more broadly at addressing any of a diverse range of needs that youth may have. Regardless of the classification system used, the fact remains that diversion programs encompass a large array of efforts. As Klein (1979) has emphasized of his own classification system: “These four categories . . . do not exhaust the repertoire of treatment strategies available to diversion and deinstitutionalization programs, but they do indicate the diversity of activities available” (p. 169).

One critical implication of this diversity is that evaluations of any one instance of, say, a delinquency prevention diversion sanction or a child welfare diversion sanction, may not generalize to other types of diversion efforts (Mears 2011). Indeed, the very existence of such categories suggests that we should be careful in generalizing about “diversion” efforts in general and instead should focus on types of diversion efforts, and, in turn, the relative effectiveness of specific manifestations of a given type of diversion. For example, skill development—one of Klein’s (1979) four categories of diversion—can “include tutoring programs, sex and grooming information programs, job training, role-playing, and paraprofessional training” (p. 169). In addition, various scholars classify some sanctions differently. They may, for example, treat informal probation as a type of diversion effort whereas others may not (Roberts 2004; Sanborn and Salerno 2005). The sum result is a delimiting of the extent to which we can generalize about the effects identified in any particular evaluation. The potential constraint on the generalizability of such evaluations is compounded by the likelihood that any given diversion effort may target a distinct population of youth, such as status offenders.

A related implication is that the different types of diversion used in a given jurisdiction

contribute to defining “business as usual” sanctioning there. Put differently, the diversion efforts define, at least in part, the baseline against which any alternative or new type of sanctioning is judged to be effective. Accordingly, any generalization about the effect of a new sanction necessarily implies that the effect may be expected only in jurisdictions that operate with similar baseline diversion practices and not necessarily in those with different baselines. For example, a particular type of informal probation program may be effective in jurisdictions that rely heavily on intensive probation only and that have few diversion efforts, but may be ineffective in jurisdictions that rely on a diverse range of diversion efforts. We return to this issue below.

“Other”

“Other” sanctions constitutes a category that also encompasses many types of possibilities, some of which might be classified as diversion efforts or probation. They may, for example, include fines, community service, and specialized courts. In each instance, there may be further variation. Fines may vary greatly (Bouffard and Muftić 2008), the types and amount of community service may vary greatly (Galaway 1988; Douglas and Hunninen 2008), and specialized courts, too, may vary (Mears et al. 2010). Consider drug courts, which first emerged in 1989 and now exceed 2,100 in number (Huddleston et al. 2008). These alternatives to traditional court processing emphasize many different dimensions, such as the integration of drug treatment services with case processing, provision of drug treatment and other needed services, drug testing, ongoing judicial monitoring of cases, and rapid application of rewards for positive behavior and punishments for negative behavior. Such elements may exist in a variety of amounts and combinations, so much so that there now exist a number of different drug court conceptual frameworks, including ones developed by the National Drug Court Institute, RAND, Temple University, and the Urban Institute, that aim to help make sense of exactly what drug courts are and what they do (Butts and Roman 2004). As with other types of specialized alternatives (e.g., mental health courts and teen courts), the generalizability of evaluations of

these “other” sanctions is limited by the variability in the types of specialized courts that exist, the populations that they serve, and the settings—and baseline sanctioning practices—that exist in the locales where the courts reside.

Probation

Probation has been described as the workhorse of the juvenile justice system (Torbet 1996). Indeed, and as depicted in table 1, of the 1,123,300 delinquency cases that received a sanction in 2007, 57 percent (642,300) resulted in probation. Thus, probation, more so than “other” sanctions or placements in custodial facilities, constitutes the primary sanction that most youth referred to the juvenile court receive and for that reason bears particular attention (Ward and Kupchik 2010; Mears 2011). Here, again, variability within the category of probation is far more the norm than the exception.

At its most general, probation allows youth to remain in the community but they must report to a probation officer and comply with certain conditions. Probation can and does, however, encompass a diverse set of arrangements, as Sanborn and Salerno (2005:364-368) have shown in their analysis of state statutes and practices. For example, the length of probation can vary and, in many instances, can be indefinite. Youth may return home but, should problems be present there, instead may be sent to relatives or elsewhere. A variety of conditions can be imposed, including curfews and suspension of a driver’s license and requirements that youth attend school or seek employment, pay fines or provide some type of restitution of various amounts, and/or provide community service of various amounts. Youth on probation, including both delinquents and status offenders, also can be placed on house arrest or home detention for variable amounts of time. In addition, they can be placed in regular detention as a form of “shock probation.” Not least, youth can be placed on regular or more intensive supervision, with the latter involving more frequent visits to and contact with probation officers. Notably, state statutes typically not only stipulate the types of conditions that are allowed but also leave open considerable leeway

for individual jurisdictions to devise their own conditions, such as requirements to receive or participate in various types of programming or services, as long as these are “reasonable.”

Juxtaposed against these potential sources of variability is the fact that jurisdictions may vary greatly in the extent to which they use different types of probation. At a minimum, they will vary in the amount of the different types that they can use. Urban jurisdictions, for example, typically may have more options than rural and suburban jurisdictions.

This situation has given rise to the creation of different types of probation classification systems. Sanborn and Salerno (2005), for example, have classified the different types into a five-level continuum: nonreporting probation, regular probation, day treatment programs, intensive probation, and house arrest. By contrast, Lopez and Russell (2008:384), in their analysis of a southwestern jurisdiction, described five distinct categories of probation: early intervention, school probation, standard probation, intensive probation, and treatment (e.g., oversight of youth in residential treatment facilities).

There is, ultimately, no correct classification system, and certainly not one that is universally agreed upon by researchers or the practitioner community. For illustrative purposes here, table 2 presents four general categories: (1) regular probation; (2) regular probation with treatment as a central focus; (3) intensive supervision; and (4) intensive supervision with treatment as a central focus. In each instance, there may be variability in the amount or quality of supervision. For “2” and “4,” there also may be variability in the amount and quality of treatment. In each instance, there may be differences in the extent to which a social work or welfare philosophy is emphasized and, concomitantly, a punishment or control philosophy is emphasized (see, e.g., Ward and Kupchik 2010).

The categories, when considered in light of the above list of probation types or conditions, serve the heuristic purpose simply of underscoring that even when using a relatively small set of categories, one ends up with very different sets of baseline conditions for assessing the impact of any new or different way of processing youth. For example, an urban jurisdiction that emphasizes control-oriented types of probation (e.g., house arrest) might find that a new effort

that emphasizes counseling and drug treatment services produces greater reductions in recidivism. However, this effect might not be as likely to occur in a suburban jurisdiction that, say, strongly emphasizes a range of supervision and rehabilitative strategies.

The question of the amount of supervision, treatment, fines, time on probation, and so on itself creates considerable variability within types of probation. What, for example, is the effect of relatively small fines (\$50) or large ones (\$1,000) and how might these effects vary if the fines were incurred as part of a constellation of other conditions of probation? A parallel situation exists when we focus on the quality of any of the various types of probation. That is, for any given type of probation, the precise design and quality of delivery may vary greatly. Community service, for example, might involve relatively little by way of meaningful activities or oversight, or, conversely, it might involve a service to a particular victim or community that is more meaningful and is carefully supervised by a probation officer.

Commitment

In the parlance of the juvenile court, youth are not imprisoned; rather, they are “committed” to or “placed” in custodial or secure detention facilities. Regardless of the terminology used, it remains the case that the “deep end” sanction for many youth is confinement in some type of secure facility for extended periods of time, sometimes until or even after (in the case of blended sentencing laws) the end of the juvenile court’s official jurisdiction (Snyder and Sickmund 2006; Feld 2009). The variability among these facilities is, as noted above, considerable. They may be state-run, county-run, or privately run; they may be small (e.g., holding 10-15 youth) or large (e.g., holding 200 or more youth); they may consist of varying custody levels (e.g., low, medium, or high) or emphases (e.g., therapeutic community, boot camp, wilderness camp); and they may include community-based residential facilities, independent living programs, and group and foster homes (Krisberg 2011:757; see, generally, MacKenzie and Freeland 2011).

The heterogeneity is reflected further by the various names given to juvenile justice

institutions (e.g., camps, treatment centers or facilities, correctional facilities, training schools, youth services centers, villages, schools, institutes, houses, cottages, academies). The diversity in terminology is so great as to make it “impossible for outsiders to understand how serious or harsh a commitment disposition is based on only the name of the youth’s destination” (Sanborn and Salerno 2005:371). Facilities also may vary in the types of youth placed in them and, as with probation, in their design, in how well they are operated, and, not least, the amount and quality of various treatments and programming (Mears and Barnes 2010; MacKenzie and Freeland 2011; Yazzie 2011). In addition, youth within any given facility may vary in their exposure to other facilities. One youth may serve the bulk of his or her commitment in, say, a low-risk residential facility, while another may spend time in low-risk, medium-risk, and high-risk facilities. Arguably, then, each experienced qualitatively different types of incarceration. In each instance, the youth, upon release, may enter different types of supervision conditions. There is, in short, no single “confinement” or prison-like experience but instead a multitude of possible sanction configurations.

Transfer to Adult Court

The juvenile court has, since its inception, provided for the possibility of transferring some youth to the adult court, but in recent decades the options and mechanisms for doing so have increased greatly (Mears 2003; Tanenhaus 2004; Sanborn and Salerno 2005; Feld 2009). Fewer than 1 percent of referred youth are judicially transferred to adult court each year; it is unknown how many are transferred through other mechanisms, such as prosecutorial transfer (Snyder and Sickmund 2006; Howell 2009). Regardless of the mechanism, however, the diversity of mechanisms for transfer, the diversity of youth populations subject to transfer, the diversity of sentences resulting from transfer (e.g., youth may or may not be convicted in adult court and they may or may not receive a prison sentence) all make it difficult to arrive at any clear assessment of the impact of transfer “in general” relative to the various dispositional alternatives

in place in specific states and jurisdictions (Mears 2010; Feld and Bishop 2011b).

ADDITIONAL HETEROGENEITY

As the above discussion highlights, the sanctioning heterogeneity in juvenile court is difficult to overstate. Notably, however, few evaluations take account of this heterogeneity. For example, as illustrated in column 2 of table 2, a study might examine two broad categories of diversion rather than the multiplicity of specific diversion efforts that exist; it might examine two broad categories of “other” sanctions or two broad categories of probation (regular and intensive); or it might divide placements into two categories (low or medium risk versus high or special risk). More frequently, evaluations use the column 3 categories (e.g., diversion, other, probation, and placement). Still others study the effects of placement (in general) as compared to non-custodial sanctions (column 4). It is, of course, the categories in columns 3 and 4 that are frequently used to characterize case processing dispositions in juvenile court nationally (Snyder and Sickmund 2006) and in many state or local jurisdictions.

The more typical evaluation situation does not actually focus on the above types of comparisons. Rather, it consists of one in which a particular sanction—for example, a particular diversion program, “other” sanction, type of probation, or custodial facility—is compared to a counterfactual scenario. That is, recidivism, mental health, drug dependency, education, and other outcomes among youth subject to this particular sanction are compared to what would have happened to the youth in a particular locale had they not received the particular sanction. Notably, this counterfactual scenario itself is highly variable across jurisdictions and results from a diverse set of sanctioning practices as used on a diverse youth population.

There is, in fact, even greater variability in juvenile court sanctioning than the above discussion indicates. There are, for example, many sanctions that may be characterized as programs or interventions. They frequently are, or are a part of, the sanction that a youth receives. Indeed, the many types of programs and interventions identified in meta-analyses and

reviews as “effective” approaches to reducing delinquency and recidivism may serve as the sanction or, alternatively, as part of the sanction imposed by juvenile courts. Table 3, for example, lists examples of what Greenwood and Turner (2011a) term “proven delinquency prevention/recidivism” “programs,” “strategies,” and effectiveness, as well as the types of juvenile justice punishments and interventions examined in Lipsey’s (2009) meta-analysis. Many of the listed programs, strategies, and interventions in fact have been developed as ways to provide more effective sanctioning that achieves the multiple goals of the juvenile court, including punishment, accountability, and rehabilitation. In the discussion that follows, we argue that this heterogeneity in sanctioning profoundly limits the external validity of studies to date and in turn the state of evidence about “what works.”

Insert table 3 about here

IMPLICATION OF JUVENILE SANCTION HETEROGENEITY: LIMITED EXTERNAL VALIDITY OF THE “WHAT WORKS” LITERATURE

To this point, we have highlighted the considerable heterogeneity in juvenile sanctions. The heterogeneity stems from variability between broad types, or categories, of sanctions and also from within them. It stems, too, from the amount and quality of exposure to each type and from various combinations of exposure to the different types of sanctions. It also stems from the mode of processing—informal versus formal—that led to a given type of sanction. It stems from potential variability across settings (e.g., urban versus rural, public versus private) in a given sanction’s design and implementation or in the population that it serves. And, not least, it stems from the broader social, cultural, and organizational context within which the sanction is administered. It is this context that provides the basis for defining “business as usual” sanctioning and, in turn, for assessing the relative effectiveness of a given sanction.

This heterogeneity substantially delimits the external validity of the bulk of extant juvenile

sanction evaluations, including meta-analyses and reviews. First, extant evaluations invariably have examined particular sanctions and done so in the absence of any clearly defined or agreed upon classification system for categorizing juvenile justice sanctions. Accordingly, studies typically are limited to being able to claim that a particular program in a particular setting appeared to generate better outcomes than some general constellation of sanctions in that same setting and that collectively constituted the de facto, “business as usual” pattern of sanctioning.

Second, a particular sanction’s effects occur relative to that of the full set of alternatives—the “business as usual” practices—in a given jurisdiction. This counterfactual situation varies across jurisdictions. Accordingly, an identical sanction involving identical youth might appear more effective or less effective in another jurisdiction, depending on the counterfactual conditions that exist there. For example, in one jurisdiction, we may find that a given type of sanction, which may include specific programs or interventions, is “effective” but in another jurisdiction we may find that it is not. It is the same sanction in each instance, but its effectiveness is not absolute; rather, it is relative to the full range of sanctions that exist and the extent to which they are used.

Third, and related to the above point, any identified effectiveness of a particular sanction is relative to an average effect across the full range of sanctions that, in a particular jurisdiction, constitute the “business as usual” conditions that the youth otherwise would have received. Perhaps, for example, one of the “business as usual” sanctions, involving placement in program X, reduces recidivism by 10 percent relative to the other “business as usual” sanctions. The effectiveness of program X would be obscured by combining the outcomes for youth exposed to it with the outcomes of youth from the other “business as usual” programs, many of which in fact may be highly ineffective. Most evaluations, including randomized experiments, do not address this situation. Instead, they examine individuals who were placed in a new type of sanction (or program or intervention) and compare their outcomes with those of youth who typically receive some “business as usual” sanction.

This approach would be unproblematic if the latter type of sanction was only of one type. In reality, though, and as the discussion above highlighted, youth in juvenile court can and do

experience a myriad of sanctions and these in turn typically are implemented variably and in differing amounts for different sub-populations of youth (see, e.g., Lowenkamp et al. 2010; Mulvey et al. 2010; Olson 2010; Welsh et al. 2010; Cooley 2011; Farrell et al. 2011). Consider, for example, the variable amount and quality of programming, within and across juvenile courts and justice systems and, at the same time, the variable amount and quality of programming that addresses cultural and race and ethnic differences among youth, the gender-specific needs of youth, their mental health needs, and so on (Mears and Aron 2003; Zahn et al. 2009; Mears et al. 2010; Sharkey et al. 2010; Gagnon and Barber 2011). In short, the counterfactual situation in most counties likely differs greatly from one place to another and thus affects the size of effect that a study, whether using a quasi-experimental or experimental design, will identify.⁴

Taken together, these considerations suggest that there is, in fact, limited evidence to date about the external validity of evaluations of juvenile court sanctions, including the many programs and interventions that constitute these sanctions. The external validity of these studies is further hampered by the fact that many evaluations are not of representative efforts. Rather, they are of programs and interventions that have been adopted in areas where the most need arose and where special attention was given to sustained, high-quality implementation. The sum result is a set of estimated effects that exceed what likely would occur in contexts where less need exists and where implementation is of a lower quality (Mears 2010).

To be certain, many excellent studies exist that have high internal validity. Indeed, among the best, most rigorous studies, there clearly is evidence that particular programs and interventions appear to reduce recidivism among youth referred to juvenile court. Whether these effects can be expected to occur when these programs and interventions are implemented elsewhere is, however, largely unknown. Indeed, even when several studies have been conducted of a particular type of program, such as Multi-Systemic Therapy (MST), and even when the studies suggest that the program is effective (cf. Littell 2006, 2008), basic questions about external validity remain with respect to whether identified or different effects could be expected if the program were adopted in different types of jurisdictions (e.g., rural versus urban)

or other populations, or, more generally, settings that not only employ different sets of sanctions in varying amounts but also vary with respect to social, cultural, or organizational context.

INTERNAL VALIDITY RECONSIDERED

Questions about the external validity of published evaluations of various juvenile justice sanctions also raise questions about the internal validity of these studies. Consider, for example, the question of what exactly is being captured by meta-analyses and reviews that identify a range of juvenile justice sanctions, including various programs and interventions, as being effective (see table 3; see, generally, Butts and Mears 2001; Lipsey and Cullen 2007; Howell 2009; Lipsey 2009; Greenwood and Turner 2011a-b; Henggeler and Schoenwald 2011; Krisberg 2011; MacKenzie and Freeland 2011). It of course could be the case that the sanctions in these studies actually are effective, at least for the “treated” populations and places. Alternatively, it is possible that the effects do not result per se from the specific sanctions or interventions.

Why? It could be that the effects instead arise from the following conditions. First, some communities may have demonstrably limited or poor sanction options; second, they may be sufficiently motivated, perhaps because of real or perceived increases in juvenile crime, to pursue a better option; and, third, they may devise or adopt an approach that fits their local needs and populations. In so doing, these communities may give considerable care to the new approach, creating new services or adapting some type of “evidence-based” program identified by research to local needs and contexts. The approach might have a theme (e.g., an emphasis on drug treatment) but typically might feature several different activities collectively designed to address the targeted population. Indeed, many juvenile justice sanctions, especially those identified as effective or “best practices,” employ multi-modal responses to juvenile offending.

In this situation, there are two distinct causal forces that may exist. On the one hand, there are the sanction-specific elements implemented by the jurisdiction (e.g., drug treatment, family counseling, frequent contact with a probation officer). In this scenario, an identified effect

reflects the specific design of the sanction. On the other hand, there are the more general, process- or context-specific elements. There is, for example, the dissatisfaction with current sanctioning options, the perceived need to take action, the good will and energy of juvenile justice personnel and possibly others (e.g., families and community members, personnel from other agencies) involved in the new effort, and the care and attention spent thinking about how best to address the needs and issues of a particular youth population. Under either scenario, an evaluation might well identify an impact of a particular sanction. In that sense, it would have high internal validity. However, it would likely fail to identify the possibility that the underlying cause of a particular effect may have had less to do with the specific features of a given sanction (including any given program or intervention that is part of or consists of the sanction) than with more general, process-related dimensions.

In this sense, the evaluation lacks internal validity because it cannot identify what exactly is producing the observed effect—the intervention or the process-related factors represented by the intervention. The implications of this situation are significant. Among other things, if the causal effect arises from the process dimensions, it would suggest that it is not, as asserted by Greenwood and Turner (2011b:741), that a jurisdiction must adopt specific elements of identified “proven” programs or strategies to be viewed as evidence-based. Rather, the jurisdiction would want to consider adopting processes similar to those that led to the creation or adoption of these programs or strategies (e.g., careful assessment of the needs of a given population of youth and the limitations of the sanctions currently in use for them, developing collaborative efforts within and across different parts of the juvenile justice system).

For policymakers and practitioners, the situation directly raises questions that return us to the problem of external validity. Bardach (2004:209) has observed the following: “If the target-site environment introduces variations from what was analyzed in the source sites, the problem [of external validity] becomes much more difficult. You are extrapolating, not replicating” (see also Sampson 2010). In juvenile justice, target sites invariably will and do differ in their sanctioning. As discussed above, even a category as seemingly straightforward in meaning as “probation” can

encompass an enormous range of possibilities. In addition, many juvenile justice sanctions, including programs and interventions listed in meta-analyses and reviews, lack clear protocols and descriptions for how they are to be implemented (e.g., dose, quality, duration). And, again, the putative effectiveness of these sanctions may stem more from process-related rather than design-specific dimensions. The uncertainty about what produces the effect associated with many “evidence-based” juvenile justice sanctions, including various programs and interventions, thus ultimately raises questions as well about the external validity of the research on them.

JUVENILE JUSTICE SANCTIONS: THE STATE OF EVIDENCE

There have been substantial advances in recent decades in identifying particular programs, interventions, and strategies that “work” (see, e.g., Howell 2009; Lipsey 2009; Greenwood and Turner 2011a-b; Henggeler and Schoenwald 2011; Krisberg 2011; MacKenzie and Freeland 2011). One major benefit of this scholarship is the identification of many options that states and local communities might consider as they work to improve life outcomes—including not just reduced offending but also improvements in education, mental health, and other life domains—of youth who come into contact with the juvenile justice system. Even so, the state of evidence is far from definitive concerning these identified sanctions, programs, and interventions as well as the many that have received less attention (Mears and Barnes 2010; Olson 2010). This situation derives from at least four considerations.

The first consists of the fact that there exists no established set of criteria for how different juvenile justice sanctions, and various court-ordered programs and interventions, should be classified. The result is that while we may know much about some “trees” (e.g., particular programs and interventions), we know little about the “forest” (i.e., the full spectrum of sanctions employed in juvenile courts). Meta-analyses and reviews do not directly address this issue. Some meta-analyses, for example, employ reasonable classification schemes in the sense that they have face validity (see, e.g., Lipsey 2009; Cooley 2011; Greenwood and Turner 2011a-b).

But the classifications do not address the types of heterogeneity in sanctioning discussed above.

A related issue is that meta-analyses may code various sanctions, including any attendant programs or interventions, in multiple ways (see, e.g., Lipsey 2009:133). So, program X might well be coded as having five different dimensions. In turn, analyses are conducted that identify “the” effects of particular types of interventions, such as those that have dimension 1, those that have dimension 2, and so on. These characteristics then typically are examined using an additive model in which specific characteristics are assessed as having an effect “net” of other factors. However, interventions that have multiple characteristics by definition do not exert effects net of one factor or another. Rather, an effect, should one exist, stems from the multiplicity of factors.

Second, as discussed above, many times new types of sanctioning approaches, including particular programs or interventions, arise in contexts that themselves may account for an identified effect more so than anything intrinsic to the sanction itself. For example, in jurisdictions that face substantial juvenile crime problems and that employ few effective approaches to address them, there may emerge a critical mass of concern that leads various stakeholders (e.g., court personnel, child welfare agencies, schools, community members) to create or adopt carefully calibrated responses that include multiple types of efforts (e.g., greater accountability through more frequent visitation with probation officers, drug treatment, collaborations with school officials and child welfare agencies). Such conditions might lead to the implementation of any of a variety of efforts that could be equally effective. As such, it is these conditions more so than any one particular type of sanction that creates any observed improvements in outcomes measured in impact evaluations.

Third, as also discussed above, the external validity of many sanctions, including those examined in leading meta-analyses and reviews, remains largely unknown. That is not to say that the practices, programs, interventions, strategies, and principles identified by Lipsey (2009) and others (e.g., Howell 2009; Greenwood and Turner 2011a-b) would not be effective with diverse populations and in diverse settings and for diverse outcomes. Rather, it is that there simply is limited empirical evidence that addresses this issue (Olson 2010; cf. Farrington 2003).

Fourth, the state of evidence concerning the internal validity of the sanctions examined in these meta-analyses and reviews is not in and of itself “evidence” that objectively represents social scientific consensus. Accordingly, claims that there are “proven” recidivism programs, strategies, and principles risk overstating the state of science and of how many social scientists would characterize the research to date. Typically, for example, social science identifies empirical support for some relationship or another, or fails to find such support, but it does not in any absolute sense “prove” the relationship.

Fifth, the evidence concerning the effects of various sanctions on a diverse set of outcomes is scant. For example, the mission of the juvenile justice system in most states includes goals related to a range of outcomes, not just recidivism. State juvenile code purpose clauses highlight this fact. They usually emphasize at least one of the following, and usually more: (1) restorative justice aims (e.g., individual accountability to victims and the community), (2) aims consistent with the 1925 Standard Juvenile Court Act (e.g., the provision of “the care, guidance, and control that will conduce to [a juvenile’s] welfare and the best interest of the state”), (3) aims consistent with the 1969 Legislative Guide for Drafting Family and Juvenile Court Acts (e.g., providing for the “care, protection, and wholesome mental and physical development of children involved with the juvenile court”), (4) accountability and protection aims (e.g., using punishment both to create offender accountability and to promote deterrence), and (5) child welfare aims (Snyder and Sickmund 2006:98-99). Given this range of goals, the fact that there exists no body of evidence that systematically compares the relative effectiveness of various sanctions across a range of outcomes, and not just various measures of recidivism, is a caution against claims about the effectiveness of particular sanctions, programs, and interventions (Mears and Barnes 2010).

Sixth, there is also the fact that few rigorous studies have examined the critical question—one that is especially salient in recent years given the rise in juvenile incarceration throughout much of the past three decades—of the relative effectiveness of various incarcerative sanctions to other non-incarcerative sanctions. From their review of research on the effectiveness of adult incarceration, for example, Nagin et al. (2009:177) have noted: “A remarkable fact is that

despite the widespread use of imprisonment across democratic nations and the enormous expansion of the prison system in the United States, rigorous investigations of the effect of incarceration on reoffending are in short supply” (see also Cullen et al. 2011). Much the same can be said of juvenile incarcerative sanctions. As MacKenzie and Freeland (2011:793) found in their review: “There are few studies of residential placement versus community placement.” Such accounts suggest that incarceration has either no effect or a criminogenic effect (see, e.g., Nagin et al. 2009:178; Cullen et al. 2011:50S), but they suggest, too, the possibility that the effectiveness of any given facility may depend on what happens inside it and also the type and level of supervision upon release (see, e.g., MacKenzie and Freeland 2011:793). The larger finding, however, is the fact that few rigorous studies exist and those few that have been conducted do not address the heterogeneity in juvenile court sanctioning, including the variability within custodial types of sanctions.

CONCLUSION

In an era in which calls for evidence-based policy and practice have increased, there has emerged a corresponding focus on generating evidence about “what works” to reduce delinquency and to achieve other juvenile justice goals, such as improving the life chances of youth by targeting their strengths and addressing their mental health problems, drug use and addiction, educational deficits, and other areas of need (Farrington and Welsh 2007; Mears and Butts 2008; Allison et al. 2010). In response, scholars, in conjunction with the practitioner community, have conducted evaluations, meta-analyses, and reviews that highlight that effective juvenile justice interventions exist (Sherman et al. 2002; MacKenzie 2006; Howell 2009; Lipsey 2009; Greenwood and Turner 2011a-b; Henggeler and Schoenwald 2011; Krisberg 2011; MacKenzie and Freeland 2011).

Although this work has advanced knowledge about efforts to improve the lives of youth, it has not created a body of research that directly assesses the external validity of the diverse array

of sanctions in juvenile justice, including the spectrum of programs and interventions that typically are, or are part of, these sanctions. That is, there is strong evidence that some efforts have produced beneficial effects for the targeted populations and areas. Even so, there remains much about juvenile justice sanctioning that is unknown. For example, despite large-scale increases in juvenile incarceration, few studies exist that provide credible assessments of the impact of incarceration versus non-incarcerative sanctions (see, generally, Loughran et al. 2009; MacKenzie and Freeland 2011; Nagin et al. 2009).

In addition, and what has been argued here, we know little about whether identified effects associated with specific sanctions, including various programs and interventions, can be expected to occur with other populations or in other areas. This situation stems from many factors. Not least among them is the fact that the juvenile court can administer different sanctions, either through informal or formal processing, as well as multiple sanctions simultaneously, and these in turn can vary greatly in what they entail and in their amount and quality. For example, not all probation sanctions are equal with respect to the levels and duration of supervision, the various conditions that must be met, or the extent to which social services and programs or interventions are involved. Similarly, a multitude of residential placement facilities exist, with varying emphases on treatment and punitiveness, types and amounts of programs and services, quality and amount of staffing, and so on (MacKenzie and Freeland 2011). Diversion, as a general category, encompasses a similar level of heterogeneity (Klein 1979). In any given jurisdiction, then, any identified effects of, say, a particular intervention may well be specific to the constellation of efforts used for similar youth in that jurisdiction.

This heterogeneity arguably is a strength of the juvenile court (Guarino-Ghezzi and Loughran 2005; Howell 2009). It allows greater flexibility for individualized sanctioning and being responsive to the needs and capacities of the juvenile court and the local community as well as local views about how best to hold youth accountable and to help them. At the same time, it necessarily creates a highly variable approach to sanctioning that significantly limits the potential generalizability of evaluations of local efforts to other populations and places. In addition,

absent a systematic assessment of the external validity of the many sanctions, programs, and interventions used in juvenile justice, it remains unclear to what extent “evidence-based” efforts derive their impact from the elements specific to these efforts or to a more general set of factors—what Bardach (2004) has referred to as “basic mechanisms”—such as the presence of a community of practitioners and citizens committed to creating better outcomes for youth and to ensuring that any new effort is well-designed and well-implemented.

Viewed in this light, jurisdictions perhaps should not be adopting the types of efforts described in recent reviews (e.g., Greenwood and Turner 2011a-b) so much as focusing on better assessments of current efforts, creating greater community and court mobilization, devising sanctions that build on these assessments and community capacities, and emphasizing high-quality implementation (see, generally, Mears 2010; Olson 2010). The possibility that the process more than specific programs and interventions is what matters most would accord with the fact (1) that a diverse and large set of programs and interventions have been found to be effective in reducing delinquency across different levels of penetration into the juvenile justice system (Lipsey 2009:143), and (2) that ineffective programs and interventions have tended to be “those with poor or no theoretical basis, were poorly implemented, focused on punishment, deterrence, or control instead of rehabilitation, and emphasized the formation of ties or bonds without first changing the individual’s thought process” (MacKenzie and Freeland 2011:789).

What, then, is to be done? Several implications can be identified. First, in a related vein, researchers and the practitioner community should be cautious in generalizing the results of individual evaluations and also of meta-analyses and reviews. Among other things, the identified effects use “business as usual” comparisons that may not apply to other settings. Accordingly, cost-benefit assessments, which typically build on the estimated effects derived from studies of programs in a few select places (see, e.g., Cowell et al. 2010; Kliet et al. 2010) and not studies of particular programs as implemented in a diverse array of settings (see, however, Greenwood and Turner 2011a-b; Welsh and Farrington 2011), also may not apply to other settings. Caution does not require that researchers refrain from identifying that some

efforts appear to be effective in reducing delinquency in some settings and for some populations that operate within specific social, cultural, and organizational contexts. But it does mean that, at the same time, attention should be drawn to potentially important caveats about the external validity of studies. For example, studies should emphasize, where appropriate, the extent to which particular sanctions, including various programs and interventions, can be expected to produce similar effects with other populations or in different settings and whether outcomes other than delinquency are affected.

Second, and by extension, jurisdictions may want to attend not only to what the literature identifies as “evidence-based.” Rather, they may want to examine and capitalize on community resources and interests and the capacity and interest of diverse court practitioners to mobilize and create systemic changes that might yield greater aggregate improvements in juvenile justice as compared to investment in a few select special efforts.

Third, there is a need for the scholarly and practitioner communities to develop a consistent definition of what is meant by “evidence-based” (Davies et al. 2000; Mears 2010). The term can have a variety of meanings, and may reference, for example, any study that involves empirical analysis or any study that involves an experimental design. Of course, experiments stand as a conventional gold standard for what many would consider to be “evidence.” However, the trade-off with an experiment is a reduction in external validity. That situation is, as we have argued here, amplified in a context in which a multitude of sanctions create highly different baseline, or “business as usual,” conditions. “Evidence” could also mean that there exists empirical research that documents that a need exists for a particular type of intervention or that existing efforts are poorly implemented. As noted at the outset, some scholars view any program that includes some unspecified type or amount of “proven” strategies or principles as “evidence-based” (Greenwood and Turner 2011b:741). Without greater consistency in what is meant by evidence-based, the result will be the potential for misleading “labeling” that may confuse more than help and, worse, result in dropping potentially effective approaches for more expensive alternatives that may be no more effective. At the same time, if evidence-based is a term applied only to the

results of experimental designs, the risk is that virtually no efforts to achieve systems-level reforms could ever be justified as evidence-based since experimental designs of such efforts typically are not possible (Mears 2010).

Fourth, and building on the above recommendations, there is a need for what might be termed “real-time” monitoring of juvenile court sanctions—in all their diversity—and for identifying the populations and areas where these occur as well as the outcomes associated with them (Mears and Butts 2008). At present, such monitoring is minimal, and is largely precluded by data constraints, including the fact that “few states and counties collect any recidivism data on incarcerated youths” (Krisberg 2011:762). Such monitoring would not provide an impact evaluation of sanctions. It would, however, provide a foundation for determining the operations of the juvenile justice system and the outcomes associated with a wide range of sanctions.

In turn, this information could be used to provide an empirical basis for identifying potential areas of greatest need. Perhaps, for example, some jurisdictions appear to excessively use incarceration whereas others seem to excessively avoid it. There is no set criterion for determining “excessive,” but the ability to assign numeric values to court practices and to the outcomes associated with them could contribute to informed debates about areas where change is most needed. Monitoring sanction practices and associated outcomes also could contribute to efforts to understand where poor implementation of particular sanctions more so than a lack of appropriate sanctions is the problem. To illustrate, a sudden increase in recidivism of youth sanctioned to program X might more likely indicate an implementation problem that surfaced recently rather than a flaw in the program design. Not least, the information could contribute to efforts to conduct impact evaluations and to describe more precisely the sanctioning context that constitutes the baseline set of conditions from which an assessment of impact is made. Information about the full spectrum of sanctioning practices and outcomes associated with them also would facilitate efforts to increase understanding of the factors that give rise to observed effects (Cook 2005; Mears and Butts 2008; Welsh et al. 2011) and, in turn, to assess the external validity of various sanctions and the programs and interventions associated with them.

This last point bears elaboration. In recent years, there has emerged greater understanding in the medical community that efforts “to make meaningful comparisons of medical interventions, one must consider clinical heterogeneity of patient populations, intervention combinations, and outcomes” (Horn and Gassaway 2010:S17; see also Longford 1999). That insight has led to greater attention to “practice-based evidence” (PBE) research. A limitation of such research is that it does not provide for the strong internal validity associated with experiments. A benefit, however, is that, in contrast to many experiments, “PBE studies do not limit the types and numbers of interventions or artificially restrict patient variability” (Horn and Gassaway 2010:S17); they thus increase the external validity of their assessments. In so doing, they also allow for potentially more accurate assessments, ones that reflect “real world” conditions—which include patient, treatment, and outcome heterogeneity—of the comparative effectiveness of various interventions for specific populations under specific conditions and for specific outcomes (see also Heckman and Smith 1995; Nutley and Davies 2000; Lipsey 2003).

There exists, in short, a clear precedent from the medical field on which juvenile justice research can draw for devising PBE studies that systematically address youth, sanction, and outcome heterogeneity and that, in turn, allow for assessment of the external validity of various juvenile court interventions. Such studies also provide a mechanism by which to provide ongoing monitoring and assessments of the “best practice” interventions that they implement. Should expected benefits of these interventions fail to materialize, efforts can be taken to determine if faulty implementation accounts for the shortfall or if perhaps the interventions simply do not appear to be effective in the particular locales where they are adopted. That is not to say that PBE studies should be emphasized over experimental studies, only that both are needed if we are to have a body of evidence that can address concerns about both internal validity and external validity (Fleurence et al. 2010; see also Wells 2009; Welsh et al. 2011). Within the field of juvenile justice, there is in fact some movement toward evaluation efforts that are consistent with PBE approaches and that, at the same time, attempt to integrate such efforts with systems-level assessments (see, e.g., Lipsey et al. 2010).

Ultimately, if juvenile justice is to proceed in an evidence-based manner, it will not be enough to accumulate an increasingly large body of sanction-, program-, or intervention-specific evaluations or, by extension, meta-analyses of them that focus primarily on internal validity. Such efforts clearly are important and support arguments that, in general, sanctions and programming should be guided by principles of effective correctional intervention (Cullen and Gendreau 2000) and should, for example, focus on matching the treatment needs of youth with appropriate services, targeting high-risk offenders, and developing decisionmaking tools that allow for monitoring, assessing, and improving court sanctions and programs (Mears and Butts 2008; Lipsey et al. 2010). They also should focus on system-level efforts to improve juvenile justice (Mears et al. 2010; Bilchik 2011). However, precisely because of the heterogeneity of juvenile court sanctioning, there increasingly is also a need to address the issue of external validity. Specifically, there is a need to develop a body of research that identifies the types of sanctions—including various combinations of punishments, programs, treatments, services, and interventions—that are effective with specific populations in specific areas and that identifies the conditions under which specific effects, for particular outcomes, can be expected. Armed with such research, policymakers, practitioners, and scholars will be better positioned to devise efforts that improve the effectiveness of the juvenile justice system.

NOTES

¹ That insight has been emphasized by other scholars. Almost 50 years ago, for example, Morris (1966:645) observed that “probation and prison are not unities, in themselves, but are rather diverse bundles of treatment and custody methods.”

² These percentages are derived from national delinquency processing estimates provided in Knoll and Sickmund (2010:4).

³ It bears noting that arrest itself can be viewed as a sanction—especially given the discretionary nature of many arrests—one whose effects may well vary depending on characteristics of the youth, the type of offense, characteristics of the community, the perceived legitimacy of the police, and other such dimensions (Huizinga and Henry 2008).

⁴ The variability exists at other units of analysis as well, including states (Mears 2010) and countries (Penner et al. 2011).

REFERENCES

- Allison, K. W., T. Edmonds, K. Wilson, M. Pope, & A. D. Farrell. (2010). Connecting youth violence prevention, positive youth development, and community mobilization. *American Journal of Community Psychology, 48*, 8-20.
- Bardach, E. (2004). The extrapolation problem: How can we learn from the experience of others? *Journal of Policy Analysis and Management, 23*, 205-220.
- Bernard, T. J., & M. C. Kurlychek. (2010). *The cycle of juvenile justice*. 2nd edition. New York: Oxford University Press.
- Bilchik, S. (2011). Redefining the footprint of juvenile justice in America. In N. E. Dowd (Ed.), *Justice for kids: Keeping kids out of the juvenile justice system* (pp. 21-38). New York: New York University Press.
- Blomberg, T. G. (1980). Widening the net: An anomaly in the evaluation of diversion programs. In M. W. Klein & K. S. Teilmann (Eds.), *Handbook of Criminal Justice Evaluation* (pp. 572-592). Beverly Hills, CA: Sage.
- Bortner, M. A. (1982). *Inside a juvenile court: The tarnished ideal of individualized justice*. New York: Columbia University Press.
- Bouffard, J. A., & L. R. Muftić. (2008). The effectiveness of community service sentences compared to traditional fines for low-level offenders. *The Prison Journal, 87*, 171-194.
- Butts, J. A., J. Buck, & M. B. Coggeshall. (2002). *The impact of teen court on young offenders*. Washington, D.C.: Urban Institute.
- Butts, J. A., & D. P. Mears. (2001). Reviving juvenile justice in a get-tough era. *Youth and Society, 33*, 169-198.
- Butts, J. A., & J. Roman, Eds. (2004). *Juvenile drug courts and teen substance abuse*. Washington, D.C.: The Urban Institute.
- Chen, H. T., S. I. Donaldson, & M. W. Mark. (2011). Validity frameworks for outcome evaluation. *New Directions for Evaluation, 130*, 5-16.

- Cook, T. D. (2005). Emergent principles for the design, implementation, and analysis of cluster-based experiments in social science. *Annals*, 599, 176-198.
- Cook, T. D., & D. T. Campbell. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cooley, V. A. (2011). Community-based sanctions for juvenile offenders: Issues in policy implementation. *Criminal Justice Policy Review*, 22, 65-89
- Cowell, A. J., P. K. Lattimore, & C. P. Krebs. (2010). A cost-benefit study of a breaking the cycle program for juveniles. *Journal of Research in Crime and Delinquency*, 47, 241-262.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- Cullen, F. T., & P. Gendreau. (2000). Assessing correctional rehabilitation: Policy, practice, and prospects. In J. Horney (Ed.), *Policies, processes, and decisions of the criminal justice system* (pp. 109-175). Washington, D.C.: National Institute of Justice.
- Cullen, F. T., C. L. Jonson, & D. S. Nagin. (2011). Prisons do not reduce recidivism: The high cost of ignoring science. *The Prison Journal*, 91, 48S-65S.
- Davies, H. T. O., S. M. Nutley, & N. Tilley. (2000). Debates on the role of experimentation. In H. T. O. Davis, S. M. Nutley, & P. C. Smith (Eds.), *What works: Evidence-based policy and practice in public services* (pp. 251-276). U.K.: Policy Press.
- Douglas, T., & M. Hunninen. (2008). *Making things right: Meaningful service for juvenile offenders*. Pittsburgh, PA: National Center for Juvenile Justice.
- Drake, E. K., S. Aos, & M. G. Miller. (2009). Evidence-based public policy options to reduce crime and criminal justice costs: Implications in Washington state. *Violence and Victims*, 4, 170-196.
- Emerson, R. (1969). *Judging delinquents: Context and process in juvenile court*. Piscataway, NJ: Transaction.
- Fagan, J. (2010). The contradictions of juvenile crime and punishment. *Daedalus*, 139, 43-61.
- Farrell, J. L., D. W. Young, & F. S. Taxman. (2011). Effects of organizational factors on use of

- juvenile supervision practices. *Criminal Justice and Behavior*, 38, 565-583.
- Farrington, D. P. (2003). Methodological quality standards for evaluation research. *Annals*, 587, 49-68.
- Farrington, D. P., & B. C. Welsh. (2007). *Saving children from a life of crime: Early risk factors and effective interventions*. New York: Oxford University Press.
- Feld, B. C. (2009). *Cases and materials on juvenile justice administration*. 3rd edition. St. Paul, MN: West.
- . (1999). *Bad Kids: Race and the transformation of the juvenile court*. New York: Oxford University Press.
- Feld, B. C., & D. M. Bishop. (2011a). *The Oxford handbook of juvenile crime and juvenile justice*. New York: Oxford University Press.
- . (2011b). Transfer of juveniles to criminal court. In B. C. Feld & D. M. Bishop (Eds.), *The Oxford handbook of juvenile crime and juvenile justice* (pp. 801-842). New York: Oxford University Press.
- Fleurence, R. L., H. Naci, & J. P. Jensen. (2010). The critical role of observational evidence in comparative effectiveness research. *Health Affairs*, 10, 1826-1833.
- Gagnon, J. C., & B. R. Barber. (2011). Preventing incarceration through special education and mental health collaboration for students with emotional and behavioral disorders. In N. E. Dowd (Ed.), *Justice for kids: Keeping kids out of the juvenile justice system* (pp. 82-106). New York: New York University Press.
- Galaway, B. (1988). Restitution as innovation or unfilled promise. *Federal Probation*, 52, 3-14.
- Gottfredson, D. C., S. S. Najaka, & B. Kearley. (2003). Effectiveness of drug treatment courts: Evidence from a randomized trial. *Criminology and Public Policy*, 2, 171-196.
- Greenwood, P. W., & S. Turner. (2011a). Juvenile crime and juvenile justice. In J. Q. Wilson, & J. Petersilia (Eds.), *Crime and public policy* (pp. 88-129). New York: Oxford University Press.

- . (2011b). Probation and other noninstitutional treatment. In B. C. Feld & D. M. Bishop (Eds.), *The Oxford handbook of juvenile crime and juvenile justice* (pp. 723-747). New York: Oxford University Press.
- Guarino-Ghezzi, S., & E. J. Loughran. (2005). *Balancing juvenile justice*. 2nd edition. New Brunswick, NJ: Transaction.
- Heckman, J. J., & J. A. Smith. (1995). Assessing the case for social experiments. *The Journal of Economic Perspectives*, 9, 85-110.
- Henggeler, S. W., & S. K. Schoenwald. (2011). Evidence-based interventions for juvenile offenders and juvenile justice policies that support them. *Social Policy Report*, 25, 3-20.
- Hockenberry, S., M. Sickmund, & A. Sladky. (2011). *Juvenile residential facility census, 2008: Selected findings*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.
- Holsinger, A. M., & E. J. Latessa. (1999). An empirical evaluation of a sanction continuum: Pathways through the juvenile justice system. *Journal of Criminal Justice*, 27, 155-172.
- Horn, S. D., & J. Gassaway. (2010). Practice based evidence: Incorporating clinical heterogeneity and patient-reported outcomes for comparative effectiveness research. *Medical Care*, 48, S17-S22.
- Howell, J. C. (2009). *Preventing and reducing juvenile delinquency: A comprehensive framework*. 2nd edition. Thousand Oaks, CA: Sage.
- Huddleston, C. W., D. B. Marlowe, & R. Casebolt. (2008). *Painting the current picture: A national report card on drug courts and other problem-solving court programs in the United States*. Alexandria, VA: National Drug Court Institute.
- Huizinga, D., & K. L. Henry. (2008). The effect of arrest and justice system sanctions on subsequent behavior. In A. Liberman (Ed.), *The long view of crime: A synthesis of longitudinal research* (pp. 220-254). New York: Springer.
- Klein, M. W. (1979). Deinstitutionalization and diversion of juvenile offenders: A litany of impediments. *Crime and Justice*, 1, 145-201.

- Klietz, S. J., C. M. Borduin, & C. M. Schaeffer. (2010). Cost–benefit analysis of multisystemic therapy with serious and violent juvenile offenders. *Journal of Family Psychology, 24*, 657-666.
- Knoll, C., & M. Sickmund. (2010). *Delinquency cases in juvenile court, 2007*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.
- Krisberg, B. (2011). Juvenile corrections. In B. C. Feld & D. M. Bishop (Eds.), *The Oxford handbook of juvenile crime and juvenile justice* (pp. 748-770). New York: Oxford University Press.
- Krisberg, B., & J. F. Austin. (1993). *Reinventing juvenile justice*. Thousand Oaks, CA: Sage.
- Kupchik, A. (2006). *Judging juveniles: Prosecuting adolescents in adult and juvenile courts*. New York: New York University Press.
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders, 4*, 124-147.
- . (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals, 587*, 69-81.
- Lipsey, M. W., & F. T. Cullen. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science, 3*, 297-320.
- Lipsey, M. W., J. C. Howell, M. R. Kelly, G. Chapman, & D. Carver. (2010). *Improving the effectiveness of juvenile justice programs: A new perspective on evidence-based practice*. Washington, D.C.: Center for Juvenile Justice Reform.
- Littell, J. H. (2008). Evidence-based or biased? The quality of published reviews of evidence-based practices. *Children and Youth Services Review, 30*, 1299-1317.
- . (2006). The case for multisystemic therapy: Evidence or orthodoxy? *Children and Youth Services Review, 28*, 458-472.
- Longford, N. T. (1999). Selection bias and treatment heterogeneity in clinical trials. *Statistics in Medicine, 18*, 1467-1474.
- Lopez, V., & M. Russell. (2008). Examining the predictors of juvenile probation: Officers’

- rehabilitation orientation. *Journal of Criminal Justice*, 36, 381-388.
- Loughran, T. A., E. P. Mulvey, C. A. Schubert, J. Fagan, A. R. Piquero, & S. H. Losoya. (2009). Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology*, 47, 699-740.
- Lowenkamp, C. T., M. D. Makarios, E. J. Latessa, R. Lemke, & P. Smith. (2010). Community corrections facilities for juvenile offenders in Ohio: An examination of treatment integrity and recidivism. *Criminal Justice and Behavior*, 37, 695-708.
- MacKenzie, D. L. (2006). *What works in corrections? Reducing the criminal activities of offenders and delinquents*. New York: Cambridge University Press.
- MacKenzie, D. L., & R. Freeland. (2011). Examining the effectiveness of juvenile residential programs. In B. C. Feld & D. M. Bishop (Eds.), *The Oxford handbook of juvenile crime and juvenile justice* (pp. 771-798). New York: Oxford University Press.
- McCall, R. B., & B. L. Green. (2004). Beyond the methodological gold standards of behavioral research: Considerations for practice and policy. *Social Policy Report*, 15, 1-12.
- Mears, D. P., & L. Y. Aron. (2003). *Addressing the needs of youth with disabilities in the juvenile justice system: The current state of knowledge*. Washington, D.C.: The Urban Institute.
- Mears, D. P. (2011). The front end of the juvenile court: Intake and informal vs. formal processing. In B. C. Feld & D. M. Bishop (Eds.), *The Oxford handbook of juvenile crime and juvenile justice* (pp. 573-605). New York: Oxford University Press.
- . (2010). *American criminal justice policy: An evaluation approach to increasing accountability and effectiveness*. New York: Cambridge University Press.
- . (2003). A critique of waiver research: Critical next steps in assessing the impacts of laws for transferring juveniles to the criminal justice system. *Youth Violence and Juvenile Justice*, 1, 156-172.
- Mears, D. P., & J. C. Barnes. (2010). Toward a systematic foundation for identifying evidence-based criminal justice sanctions and their relative effectiveness. *Journal of Criminal Justice*,

38, 702-810.

- Mears, D. P., & J. A. Butts. (2008). Using performance monitoring to improve the accountability, operations, and effectiveness of juvenile justice. *Criminal Justice Policy Review, 19*, 264-284.
- Mears, D. P., T. L. Shollenberger, J. B. Willison, C. E. Owens, & J. A. Butts. (2010). Practitioner views of priorities, policies, and practices in juvenile justice. *Crime and Delinquency, 56*, 535-563.
- Merlo, A. V., & P. J. Benekos. (2010). Is punitive juvenile justice policy declining in the United States? A critique of emergent initiatives. *Youth Justice, 10*, 3-24.
- Morris, N. (1966). Impediments to penal reform. *University of Chicago Law Review, 33*, 627-656.
- Mulvey, E. P., C. A. Schubert, & C. A. Odgers. (2010). A method for measuring organizational functioning in juvenile justice facilities using resident ratings. *Criminal Justice and Behavior, 37*, 1255-1277.
- Nagin, D. S., F. T. Cullen, & C. L. Jonson. (2009). Imprisonment and reoffending. *Crime and Justice, 38*, 115-200.
- Nutley, S. M., & H. T.O. Davies. (2000). Criminal justice: Using evidence to reduce crime. In H. T. O. Davis, S. M. Nutley, & P. C. Smith (Eds.), *What works: Evidence-based policy and practice in public services* (pp. 93-116). U.K.: Policy Press.
- Olson, J. R. (2010). Choosing effective youth-focused prevention strategies: A practical guide for applied family professionals. *Family Relations, 59*, 207-220.
- Penner, E. K., R. Roesch, & J. L. Viljoen. (2011). Young offenders in custody: An international comparison of mental health services. *International Journal of Forensic Mental Health, 10*, 215-232.
- Polk, K. (1984). Juvenile diversion: A look at the record. *Crime and Delinquency 30*, 648-659.
- Roberts, A. R. (2004). The emergence and proliferation of juvenile diversion programs. In A. R. Roberts (Ed.), *Juvenile justice sourcebook: Past, present, and future* (pp. 183-195). New

York: Oxford University Press.

Rothwell, P. M. (2005). External validity of randomized controlled trials: "To whom do the results of this trial apply?" *Lancet*, 365, 82-93.

Sampson, R. J. (2010). Gold standard myths: Observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology*, 26, 489-500.

Sanborn, J. B., Jr., & A. W. Salerno. (2005). *The juvenile justice system: Law and process*. Los Angeles: Roxbury.

Scott, E. S., & L. Steinberg. (2008). *Rethinking juvenile justice*. Cambridge, MA: Harvard University Press.

Shadish, W. R., T. D. Cook, & D. T. Campbell. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Sharkey, J. D., J. B. Sander, & S. R. Jimerson. (2010). Acculturation and mental health: Response to a culturally-centered delinquency intervention. *Journal of Criminal Justice*, 38, 827-834.

Sherman, L. W., D. P. Farrington, B. C. Welsh, & D. L. MacKenzie, Eds. (2002). *Evidence based crime prevention*. London: Routledge.

Simon, J. (1995). Power without parents: Juvenile justice in a postmodern society. *Cardozo Law Review*, 16, 2501-2564.

Snyder, H. N., & M. Sickmund. (2006). *Juvenile offenders and victims: 2006 National Report*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Steckler, A., & K. R. McLeroy. (2008). The importance of external validity. *American Journal of Public Health*, 98, 9-10.

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

Torbet, P. M. (1996). *Juvenile probation: The workhorse of the juvenile justice system*. Washington, D.C.: Office of Juvenile Justice and Delinquency Prevention.

Ward, G., & A. Kupchik. (2010). What drives juvenile probation officers? Relating organizational contexts, status characteristics, and personal convictions to treatment and

- punishment orientations. *Crime and Delinquency*, 56, 35-69.
- Weiss, C. H., E. Murphy-Graham, A. Petrosino, & A. G. Gandhi. (2008). The fairy godmother—and her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, 29, 29-47.
- Wells, E. (2009). Uses of meta-analysis in criminal justice research: A quantitative review. *Justice Quarterly*, 26, 268-294.
- Welsh, B. C., & D. P. Farrington. (2011). The benefits and costs of early prevention compared with imprisonment: Toward evidence-based policy. *The Prison Journal*, 91, 120S-137S.
- Welsh, B. C., M. E. Peel, D. P. Farrington, H. Elffers, & A. A. Braga. (2011). Research design influence on study outcomes in crime and justice: A partial replication with public area surveillance. *Journal of Experimental Criminology*, 7, 183-198.
- Welsh, B. C., C. J. Sullivan, & D. L. Olds. (2010). When early crime prevention goes to scale: A new look at the evidence. *Prevention Science*, 11, 115-125.
- Yazzie, R. A. (2011). Availability of treatment to youth offenders: Comparison of public versus private programs from a national census. *Children and Youth Services Review*, 33, 804-809.
- Zahn, M. A., J. C. Day, S. F. Mihalic, & L. Tichavsky. (2009). Determining what works for girls in the juvenile justice system: A summary of evaluation evidence. *Crime and Delinquency*, 55, 266-293.

Table 1. Sanctions Resulting from Informal Processing versus Formal Processing of Delinquency Cases in the United States, 2007

	Number of Cases	Percent of all Cases
Informally processed		
Dismissed	313,200	
Probation	253,800	
Other sanction	173,100	
<i>Total</i>	<i>740,100</i>	44.4
Formally processed		
Not adjudicated delinquent		
Dismissed	221,100	
Probation	61,100	
Other sanction	49,100	
<i>Total</i>	<i>331,300</i>	19.9
Adjudicated delinquent		
Probation	327,400	
Other sanction	110,200	
Placed	148,600	
<i>Total</i>	<i>586,200</i>	35.2
Transferred to adult court	8,500	0.5
Total	1,666,100	100.0

Source: Knoll and Sickmund (2010:4).

Table 2. Heterogeneity Within and Across Types of Sanctions and Interventions

1. High Detail	2. Moderate Detail	3. Low Detail	4. Least Detail
<i>Dismissal</i>	<i>Dismissal</i>	<i>Dismissal</i>	<i>Non-commitment (including diversion, other, and probation)</i>
<i>Diversion=del. prevention</i> Various types <i>Diversion=child welfare</i> Various types	<i>Diversion=del. prevention</i> <i>Diversion=child welfare</i>	<i>Diversion</i>	
<i>Other=fine</i> Various amounts <i>Other=community service</i> Various types <i>Other=specialized court</i> Various types	<i>Other=fine/comm. service</i> <i>Other=specialized court</i>	<i>Other</i>	
<i>Probation=regular</i> Variable amount Variable quality Social work orientation Punishment orientation <i>Probation=regular with services or treatment</i> Variable amount Variable quality Social work orientation Punishment orientation <i>Probation=intensive</i> Variable amount Variable quality Social work orientation Punishment orientation <i>Probation=intensive with services or treatment</i> Variable amount Variable quality Social work orientation Punishment orientation	<i>Probation regular</i> <i>Probation intensive</i>	<i>Probation</i>	
<i>Placement=low risk*</i> <i>Placement=med. risk*</i> <i>Placement=high risk*</i> <i>Placement=special risk*</i> * May be community- based, state-run or private, or any of a range of specific types of facilities.	<i>Placement=low or medium risk</i> <i>Placement=high or special risk</i>	<i>Placement</i>	
<i>Transfer to adult court</i>	<i>Transfer to adult court</i>	<i>Transfer to adult court</i>	

Table 3. Heterogeneity Within and Across Types of Sanctions and Interventions

Greenwood and Turner (2011a:107-119)	Lipsey (2009:133-136)
<i>Proven Delinquency Prevention/Recidivism Programs</i>	<i>Juvenile Justice Supervision and Control</i>
Nurse Family Partnership	No supervision
Functional Family Therapy	Diversion
Multidimensional Treatment Foster Care	Probation or parole
Adolescent Diversion Project	Incarceration
Aggression Replacement Therapy	
Multisystemic Therapy	<i>Type of Intervention</i>
	Surveillance
<i>Proven Delinquency Prevention/Recidivism Strategies</i>	Deterrence
Cognitive behavioral therapy	Discipline
Behavioral programs	Restorative programs
Group counseling	Restitution
High school graduation	Mediation
Mentoring	Counseling and its variants
Case management	Individual counseling
Counseling/psychotherapy	Mentoring by volunteer or paraprofessional
Pre-K education for low-income families	Family counseling
Mixed counseling	Short-term family crisis counseling
Teen court	Group counseling led by a therapist
Family counseling	Peer programs
Social skills training	Mixed counseling
Challenge programs	Mixed counseling with referral for services
Family crisis counseling	Skill building programs
Mediation	Behavioral programs
Multiple coordinated services	Cognitive behavioral therapy
Skill building programs	Social skills training
Restorative justice for low-risk offenders	Challenge programs
Academic training	Academic training
Service broker	Job-related interventions
Sex offender treatment	Multiple coordinated services
Restitution	Case management
Mixed counseling with referral	Service broker
Job-related interventions	Multimodal regimen
Peer counseling	
Diversion with services	<i>Amount and Quality of Service</i>
Multimodal regimen	Amount of service
	Quality of implementation
<i>Proven Principles of Effectiveness</i>	
Integrity of treatment implementation	
Focus on high-risk youth	
Longer duration of treatment	
Well-established program	
Communities that Care	