



SOCIAL POLICY RESEARCH
ASSOCIATES

The Effort to Implement the Youth Offender Demonstration Project (YODP) Impact Evaluation: Lessons and Implications for Future Research

Final Report
January 2008

Prepared for:

U.S. Department of Labor
Employment and Training
Administration
Division of Research and
Evaluation
200 Constitution Avenue, NW
Room N-5637
Washington, D.C. 20210

DOL Contract DOLJ051A20203
SPR Project No. 1148

1330 Broadway, Suite 1426
Oakland, CA 94612
Tel: (510) 763-1499
Fax: (510) 763-1599
www.spra.com

Prepared by:

Kate Dunham and Andrew Wiegand,
Social Policy Research Associates
and
Charles Michalopoulos,
MDRC

This page intentionally left blank.

Insert blank page here when making double-sided copies

DISCLAIMER

This report was prepared for the U.S. Department of Labor, Employment and Training Administration, Office of Policy Development and Research, by Social Policy Research Associates. Since contractors conducting research and evaluation projects under government sponsorship are encouraged to express their own judgment freely, this report does not necessarily represent the official opinion or policy of the U.S. Department of Labor.

This page intentionally left blank.

Insert blank page here when making double-sided copies

CONTENTS

Acknowledgements	1
Introduction	2
Background of YODP	3
YODP Sites Chosen for Impact Evaluation	4
Proposed Evaluation Design	9
Challenges to Implementing the Random Assignment Design	11
Challenges to Grantees' Program Design/Implementation.....	11
Concerns About the Legality of the Proposed Design	17
Political Concerns	20
Alternative Design.....	22
Comparison to Other Random Assignment Studies	23
Random Assignment Studies Involving the Courts.....	23
Large-Scale Random Assignment Studies Conducted by MDRC	27
Random Assignment Studies Funded by DOL	29
Lessons.....	31

This page intentionally left blank.

Insert blank page here when making double-sided copies

Acknowledgements

We would like to thank the numerous individuals who provided significant assistance in our efforts to conduct the impact evaluation of the Youth Offender Demonstration Project (YODP). Chief among these were our Department of Labor (DOL) Project Officers during the project, including Kerri Vitalo, Lekesha Campbell, and Dan Ryan. Additionally, we would like to thank other key staff in the Division of Youth Services at DOL who provided suggestions and assistance throughout the evaluation effort, including David Lah, Richard Morris, and Gregg Weltz.

In addition, we are grateful for the diligent and dedicated efforts of the staff of the YODP grantees themselves. In particular, we would like to thank our primary contacts (generally, though not always, the Directors of the Youth Offender project) for each project, including Dennis Torbett and J.D. Johnson in Nashville, Leina'ala Nakamura in Honolulu, Kevin O'Rourke in Brockton, Linda Womack and Robert Lewison in San Diego, Chekemma Fulmore-Townsend in Philadelphia, and Tony Martinez and Mary Ellen Boyd in New York City. Their efforts to coordinate with and solicit the participation of members of the judiciary and others in the justice system were greatly appreciated throughout the project. Similarly, we would like to thank the judges and other justice system staff with whom we were able to meet, many of whom expressed great willingness to work with the study to identify feasible ways in which they could participate. We also would like to thank Dr. David Altschuler of Johns Hopkins University for his valuable assistance as a subcontractor for this evaluation effort.

Finally, we would like to thank the many staff at both Social Policy Research Associates and MDRC involved in the initial efforts for the evaluation who provided extremely beneficial guidance, assistance, and feedback. Additionally, many of these staff also provided thoughtful comments on earlier versions of this report. Specifically, at MDRC we would like to thank Dan Bloom, Alison Rebeck Black, Sara Muller-Ravett, Marjorie Singer, Fred Doolittle, and Sharon Rowser. At SPR, we would like to thank Jo-Ann Adefuin, Eloisa Dellosa, Vinz Koller, Dennis Rojas, and Ron D'Amico.

We thank all these individuals for their important contributions to this evaluation effort. Any errors or omissions in the final document are the responsibility of the authors.

Introduction

In the summer of 2005, the Department of Labor Employment and Training Administration (DOL ETA) issued a request for proposals (RFP) to conduct an Impact Study of the Youth Offender Demonstration Program (YODP). The YODP evaluation team selected by DOL was comprised of Social Policy Research Associates (SPR), MDRC, Decision Information Resources (DIR), and Johns Hopkins University. As detailed in the RFP, the intended study was a random assignment evaluation to address a number of important questions about how to work most effectively with youth offenders to prevent recidivism and increase their employment and earnings. To try to answer these questions, the study would have had courts in six different jurisdictions agree to use random assignment methodology to assign youth to a control group in which they received only standard incarceration or assign youth to one of two study groups as follows: study group 1, in which youth were incarcerated but received aftercare services from a YODP grantee, or study group 2, in which youth received services from a YODP grantee in lieu of being incarcerated.

To implement the random assignment study required that the courts allow the random assignment process to determine which youth would be incarcerated and which would be assigned to a YODP grantee (essentially to be set free), instead of relying on the judge's discretion. A similar study took place in the Wayne County Juvenile Court in Detroit, Michigan in the mid 1980s. This study randomly assigned more than 500 youths to either a control group in which youth were assigned to state incarceration or to a study group in which youth received intensive supervision as an alternative to incarceration. This study was only successfully implemented due to a set of very specific circumstances (described in detail below) that did not exist within the six YODP sites.

Despite the evaluation team's best efforts, the evaluation team was unable to convince the courts and YODP programs in any of the six selected jurisdictions to accept the random assignment methodology. As a result, DOL opted not to proceed with the evaluation. This paper describes the evaluation team's efforts in persuading sites to implement the random assignment design, discusses the range of reasons why these efforts were unsuccessful, and explores what lessons can be drawn from this experience and other studies that have employed random assignment to aid the implementation of future random assignment evaluations. The paper begins with an overview of the six YODP grantees selected by DOL to participate in the evaluation and of the random assignment evaluation plan as outlined in the evaluation team's proposal. Next, the paper describes the various challenges encountered by the evaluation team in implementing the original random assignment design. Following this, the paper summarizes a number of other studies that have successfully implemented random assignment and draw comparisons and

contrasts of these studies to the intended design in the YODP evaluation. Finally, the paper concludes with a discussion of lessons learned from this effort and several recommendations that may assist DOL in implementing future random assignment evaluations.

Background of YODP

The Youth Offender Demonstration Project (YODP) was established by the U.S. Departments of Labor (DOL), Justice (DOJ), and Health and Human Services to provide services to youth offenders, those who are at risk of court involvement or those who are involved with gangs. YODP was part of a larger Youth Offender Initiative: Reentry Grant Program, which was designed to assist communities in implementing a comprehensive reentry approach to help youth offenders released from secure confinement to successfully transition back to their communities. YODP programs were to provide services and programs that prepared young offenders, gang members, and at-risk youth, ages 14 through 24, for positive integration into their communities and for long-term employment.

One of the primary purposes of YODP was to test various local approaches to linking the juvenile justice system with the workforce development system and community-based organizations that provide youth with job training, education, and supportive services such as adult mentoring. Though the local approaches may have varied considerably, they all reflected YODP's emphasis on partnership building as a primary tool for strengthening locales' existing strategies for reducing delinquent behavior among youth.

YODP was initially funded in Program Year 1998 by a \$13.1 million Congressional set-aside in DOL's Pilot and Demonstration budget for programs addressing the needs of youth involved, or at risk of becoming involved, with the juvenile justice system. In September 1998, DOL's Employment and Training Administration and DOJ's Office of Juvenile Justice and Delinquency Prevention announced solicitation SGA/DAA 98-015, which ultimately resulted in funding 14 governmental entities' local youth offender demonstration projects in May 1999 for a period of 24 months. The first six months of project operations were slated for planning, and the remainder for implementation. In June 2001, DOL funded ten of these initial 14 projects for an additional 12 months, and awarded grants to nine new sites for a period of 30 months.¹ In 2002,

¹ The last six months were allotted for follow-up data collection.

DOL provided a third round of funding, awarding 29 additional grants to new sites to implement YODPs that also were intended to operate for 30 months.²

YODP Sites Chosen for Impact Evaluation

In June 2004, DOL provided extensions to six of the sites selected as part of the third round of funding so that they could participate in the Impact Evaluation of YODP. These sites included programs in the following locations: Brockton, MA, Bronx, NY, Honolulu, HI, Nashville, TN, Philadelphia, PA, and San Diego, CA.

Table 1: Features of YODP Grantees

Name of Grantee Agency	Ages (Number) Served	Primary Services	Links to Justice System
Brockton Area Private Industry Council (BAPIC)	14-24 (75)	Aftercare services and alternative sentence; job readiness, subsidized and unsubsidized work experience, educational and recreational services	Mass. Dept. of Youth Services, Brockton Juvenile Court, Plymouth County House of Corrections
Wildcat Service Corporation (Bronx)	14-24 (119)	Alternative sentence; education, tutoring, mentoring, job readiness and recreational activities	Bronx Supreme (adult) Court
City of Honolulu Department of Community Services	16-24 (109)	Small aftercare program that was on hold; case management and job readiness training	None active
Home Builder's Institute (Nashville)	18-24 (140)	Aftercare and alternative sentence; 10-12 week vocational training in the construction industry, literacy, remedial education, and career development activities	Davidson County Drug Court (adult), TN Dept. of Children and Family Services
Philadelphia Workforce Dev. Corp./ Philadelphia Youth Network	14-24 (135)	Aftercare program; reentry services to youth on parole	Philadelphia Family (juvenile) Court
YCHANGE (Metro United Methodist Urban Ministry – San Diego)	18-24 (63)	Partial aftercare program; job readiness training, internships, educational instruction, and peer-to-peer mentoring	Adult Court Probation Officers, Court Schools

² These grants were intended to operate on a similar schedule as the initial 14 grantees, with six months for planning and 18 months for implementation. The final six months were intended to allow the grantees to collect follow-up data.

Despite having certain programmatic similarities, these sites also varied on several key characteristics, such as the programs and services offered (aftercare or alternative sentencing), the ages and number of youth offenders served, and the connections between the sites and local justice systems. This section provides a short description of each of the YODP programs and the intensive aftercare and alternative sentencing services they were providing at the outset of the evaluation. Additionally, Table 1 displays summary features for these six sites.

Brockton

The Brockton, Massachusetts YODP was operated by the Brockton Area Private Industry Council (BAPIC). As of March 2005, the site had served 75 youth offenders, 39 of whom were ages 14-17 at the time of enrollment, and the remaining 36 were ages 18-24. Although at the start of the evaluation, the Brockton YODP was operating both an alternative sentencing and an aftercare component as needed for the study, program managers were concerned that they would not be able to enroll the required number of juvenile offenders to meet the study's requirements. To ensure a sufficient sample of youth for the study, DOL required grantees to enroll 100 youth offenders in each program component, with another 200 enrolled in the control group; however, BAPIC staff estimated that they would only be able to enroll 85-90 juvenile offenders per year in the study.

In the fall of 2005, when the evaluation began, the program offered an alternative sentencing program for youth offenders, providing these youth with recreational activities, job readiness training, subsidized and unsubsidized work experience and educational programs. The youth were transported to and from the program and the Brockton Day Reporting Center where they were mandated to report.

The majority of youth participating in this alternative sentencing program were juvenile offenders (17 and under) referred by the Massachusetts Department of Youth Services (DYS), with which the program had a written agreement. These youth had already been committed to DYS care and DYS' Area Director then determined, based on a risk and needs assessment, that these youth were appropriate for the YODP alternative sentencing program. The Brockton Juvenile Court had also referred to BAPIC a few youth who had committed minor infractions and would not have been incarcerated even if the YODP program was unavailable. No formal agreement was ever developed between BAPIC and the court system.

The program also provided an aftercare program for youth offenders consisting of an individual assessment, case management, and services such as education, training, and mentoring. Although the program was amenable to serving all age groups through this component, in practice, the vast majority of the aftercare program participants were young adult offenders (18-24) released from the Plymouth County House of Corrections (an adult detention facility), with

only a few released from a juvenile facility run by DYS. One reason so few youth ages 17 and under participated in the aftercare program is that these youth typically are released when they “age out” of the juvenile justice system. As a result, the justice system often has little control of their post-release activities.

The Bronx

The YODP program in the Bronx was run by Wildcat Service Corporation. According to its quarterly report through March 2005, this site had served 83 youth offenders ages 14 to 17 and 36 ages 18-24.

Wildcat’s YODP alternative sentencing program provided youth offenders with a number of services, including alternative education (through an on-site charter school), tutoring, mentoring, case management, recreational activities, job readiness training, career exploration, subsidized work experience, community service assignments, and job search, placement and retention services. Most of the youth offenders in this program were mandated to participate by a single judge of the Bronx Supreme (adult) Court³ who saw only those individuals assigned to the adult system either because of their age or the seriousness of their crimes.⁴ If this judge determined that one of the youth before him for sentencing seemed appropriate for the YODP program, a Wildcat case manager would interview the youth and, if deemed a good fit for the program, the judge would mandate that the youth participate in the Wildcat program as an alternative to being incarcerated.

Although Wildcat developed an MOU with the New York Office of Children and Family Services to develop an aftercare component, by the time the evaluation commenced, there had been no activity as a result of this agreement. Thus, no youth had been enrolled or even referred, and, as a result, there was no specific set of services in place to be provided to these youth.

Honolulu

The Honolulu YODP program was operated by the City of Honolulu’s Department of Community Services. As of March 2005, this site had served 72 16-17 year-old offenders and 37 18-24 year-olds. The Honolulu YODP program operated out of Honolulu’s Youth Services Center, and provided all youth with case management and job readiness training. Other

³ In New York State, youth must be 15 or younger to be considered a juvenile by the court system.

⁴ Many of these individuals are termed by New York state law as Youthful Offenders. This means that the individual was a) convicted of a crime when the defendant was at least 16 and less than 19 years of age; b) eligible to have the conviction replaced with a non-criminal (youth offender) adjudication because of the nature of the crime and the youth’s prior record; and c) adjudicated a youthful offender through the discretion of a Supreme Court judge who believed that such adjudication was in the interests of justice.

<http://www.nyc.gov/html/djj/html/cases/html>

components received by many participants included: vocational training, tutoring, anger management, substance abuse treatment, recreational activities, alternative education and community service.

At the time of the evaluation, the Honolulu YODP program did not operate a service component that involved youth being enrolled in the YODP as an alternative to incarceration. Instead, the program primarily provided services to youth who had been sentenced to probation and were then referred by their probation officer or found the program on their own. To facilitate these referrals from juvenile probation officers, the site developed a written agreement with the juvenile probation system.

The Honolulu site also operated a very small juvenile aftercare program for a few youth sentenced to the Hawaii Youth Correctional Facility, Hawaii's primary juvenile incarceration facility. However, in the few months prior to the start of the evaluation, these aftercare services had ceased due to serious problems at the correctional facility that resulted in the facility being inaccessible to those not on staff.

Nashville

The Nashville YODP program was operated by the Home Builder's Institute (HBI), a national organization with headquarters in Washington, D.C. The HBI YODP site provided youth ages 18 and over with hands-on construction-related vocational training for 10-12 weeks in an open-entry, skill-based exit model. In addition, the HBI site provided participants with literacy instruction, educational remediation services, mentoring, career development activities, substance abuse and mental health services, and community service opportunities.

Prior to the beginning of the evaluation, the HBI program had established both an alternative sentencing and an aftercare program for offenders, but only for young adults, ages 18-24. These programs primarily received referrals from a residential drug treatment facility run by the Davidson County Drug Court (DC4) and the Tennessee Department of Children and Family Services.

As of March 2005, this program had served 140 youth ages 18-24. HBI expressed concern that, as an apprenticeship-based training program, their service model might not be well suited to serving those under 18. Despite this concern, for reasons that are described below, in the month or two before the evaluation began, the HBI site began enrolling youth ages 16-17 who were referred by juvenile probation officers, and by August 2005, they had enrolled 11 such youth.

Philadelphia

The Philadelphia site's YODP program was originally operated by the Philadelphia Workforce Development Corporation (PWDC). By the time the evaluation began, however, the program was being operated by the Philadelphia Youth Network (PYN), though several of the staff that had been involved with the program at PWDC also moved to PYN simultaneously. As of March 2005, this program had served 72 youth offenders ages 14-17 and 63 ages 18-24.

By the summer of 2005, this site had developed an extensive aftercare program for youth offenders delivered through a network of community-based centers, with YODP funding representing only a small part of this effort. Under this program, youth received either intensive or standard aftercare based on an assessment administered as youth exited incarceration. Those whose access indicated they had a high likelihood of reoffending received intensive aftercare services, including a reintegration worker who assisted the youth's probation officer, while those deemed to be at lower risk of reoffending received standard reentry services. To implement this aftercare program, the site had developed partnerships with both juvenile and adult local courts and probation systems. Prior to the evaluation, the program had employed a liaison to the Philadelphia Family (juvenile) Court to facilitate aftercare referrals, although this liaison left just before the evaluation began and was not immediately replaced.

At the onset of the evaluation, the program did not operate an alternative sentencing program. Although the program was negotiating with the Philadelphia Family Court for an alternative to incarceration model that would be run through two "discipline schools," this was only in early discussions, and no specifics had been formalized.

San Diego

San Diego's YODP program, entitled YCHANGE, was operated by Metro United Methodist Urban Ministry. By the end of March 2005, this site had enrolled 63 youth ages 18-24. As in Nashville, only in the months preceding the evaluation had this site begun to enroll youth under the age of 18. By August 2005, it had enrolled 24 youth younger than 18.

The YCHANGE program focused heavily on education and workforce development services and included job readiness training, internships, remediation and alternative education classes, tutoring, mental health services and case management. In addition, the program had a particularly large and strong peer-to-peer mentoring program. The program received many of its referrals for the program from adult probation officers and court schools. While some of these youth had never been incarcerated, others were referred to YCHANGE as they were being released from incarceration. Thus, this approach was serving as the aftercare component for the program.

At the time the evaluation began, this site had not implemented an alternative sentencing component. Although the program had explored with the juvenile courts the possibility of serving as an alternative sentencing provider, the courts had insisted the program add a community service component before formalizing this, which the program was planning to do through Americorps.

Proposed Evaluation Design

The heart of the proposed design for the Evaluation of the Impact of YODP centered on random assignment methodology. This design called for youth offenders, ages 14-24, in each site to be randomly assigned to one of three groups: 1) commitment to a residential facility (incarceration) with standard aftercare (the control group); 2) commitment to incarceration but with aftercare through a YODP program (study group 1); and 3) assignment to an alternative to incarceration in which the youth would participate in the YODP program rather than be incarcerated (study group 2). The process of assigning youth offenders to one of these groups was intended to take place at the point of sentencing by a judge. However, instead of having the judge make the decision about whether to sentence a youth to incarceration or an alternative sentencing program, a random assignment procedure was to be introduced that would determine to which of the three groups youth were assigned.

According to the intended design, youth offenders in each site were to be randomly assigned to one of the three groups for a period of approximately one year until 200 youth had been assigned to the control group and 100 to each of the YODP alternative sentence and YODP aftercare groups within each site. Although for the purposes of estimating the impacts of the various treatments, it was critical to include as many youth as possible in the random assignment process, it was also clear from the outset that, in each jurisdiction, judges would need to exclude a wide range of offenders whose crimes were either too serious to be considered for an alternative to incarceration or were so minor that they would not be considered for incarceration. This fact made it necessary that, to participate in the study as initially designed, the six YODP grantees would experience a substantial shift in the types of youth they would be serving. As will be seen below, however, most did not understand this would be the case.

With this design in mind, SPR and its subcontractors submitted a proposal that called for conducting the evaluation as DOL had laid out in the RFP. In particular, the RFP called for the evaluation to begin enrolling and randomly assigning youth by January 2006, six months after the contract was awarded. The proposal noted many of the challenges that were likely to be encountered, and described how ambitious the effort would be. The proposal relied heavily on the fact that the grantees selected had all agreed to participate in the random assignment evaluation and to secure the cooperation of judges and other justice system representatives from

both the juvenile and adult courts in their area. With this assurance, the proposal developed sought to implement the intended design to examine the impacts and cost effectiveness of the various treatments.

According to the proposal, once youth were assigned to one of the three groups, they would then complete their sentence, either incarceration with standard or YODP aftercare or the alternative sentencing program operated by the YODP grantees. Other than this treatment difference, however, the youth across the three groups would differ only by chance in their characteristics, motivation, and all other factors of potential interest, due to the random determination of their sentence.

To assist judges and courts in correctly implementing the random assignment process, the evaluation team planned to conduct several visits to each site to explain the evaluation, answer questions about it, and design procedures that minimized the burden to grantees and court staff. Additionally, the evaluation included reimbursement for up to 20% of a court administrative staff person's time to help coordinate the required study procedures and complete the necessary paperwork. Finally, there were also plans to hold three meetings involving key staff and court personnel from each site to discuss challenges related to implementing random assignment and develop strategies for overcoming those challenges.

To assess the impacts of the program and treatments, the evaluation planned to measure a number of key outcomes for the participants of each group. These outcomes included measures of recidivism, employment, earnings, educational attainment (GED or H.S. diploma), risk-taking behavior and youth development. In addition to examining how these outcomes varied between the three groups, the evaluation also planned to examine how impacts of the program and treatments varied by site and by subgroups of youth, such as by age, race/ethnicity, gender, and category or number of offenses.

To collect data on these outcomes, the evaluation planned to conduct a survey of youth in each of the three groups at three separate times. The first, or baseline, survey was to have been administered at the point of random assignment, while the two follow-up surveys were to be administered 18 and 36 months later primarily using telephone interviews. The evaluation also planned to collect data from court and grantee records and obtain data on arrests and other recidivism from state and federal databases.

To be able to understand why differences between groups or sites might have occurred, the original evaluation design also called for two-day process visits to the grantees twice while random assignment was taking place. These visits would have examined program and service delivery characteristics at each site to identify variations in the treatments afforded youth,

documented the services provided, and examined other site characteristics, including the partnerships and leadership of each grantee. These visits would also have assessed organizational coordination and partnerships between programs and justice systems, social service and health care systems, educational systems and workforce development systems.

To examine whether any observed differences between groups might be related to community contexts, the evaluation also planned to conduct three visits to the communities in which the grantees operated. These visits were to be timed to correspond to the various surveys of youth so that we would better understand the community contexts in which they were living as they provided information about their outcomes.

Finally, the evaluation planned to conduct a cost effectiveness analysis that would have examined the differences in costs and cost effectiveness between standard incarceration and YODP alternative sentence and aftercare programs.

In sum, the original design for the YODP Impact Study described above was complex, especially in light of the diversity of the sites. Consequently, during the process of trying to implement this design, the evaluation team faced a number of critical challenges. The next section of this paper describes these challenges and their impact on the success of the study.

Challenges to Implementing the Random Assignment Design

The evaluation team encountered a number of challenges that made implementing the proposed design for the study very difficult and, ultimately, the evaluation team was unable to identify any sites that were able to do so. In some cases, the challenges were unique to each site, while in others the challenges were common across them. This section describes these challenges and discusses how they affected sites' and courts' willingness and ability to implement the proposed design. To do so, the major challenges are grouped into three distinct areas: (1) those that affected the program design and implementation grantees had established; (2) concerns about the legality of implementing the study; and (3) political concerns that affected how the study could be implemented. Each of these is discussed in turn below.

Challenges to Grantees' Program Design/Implementation

Changes in Eligible Age Range

The first significant challenge to the project emerged during an initial phone call for the evaluation, in which DOL announced that the study should focus on youth adjudicated in the juvenile court system. This meant that the age of eligible youth would be restricted from the original range of 14-24 (which had been specified in the RFP for the evaluation) to a more

narrow range of 14 to, at most, 18. The actual eligible age range because of this decision varied across the sites, as the age of majority (the age at which an individual is treated as an adult in each state court system) varied across the states in which sites were operating. Hence, in New York, in which the age of majority is 16, this decision had the effect of narrowing the eligible age range for youth to those ages 14 or 15. In most other sites, the eligible age range as a result of this decision was 14 to 17.

This narrowing of the eligible age range was more troublesome for some sites than others. For example, in New York, although the YODP program had been working primarily with 16 year olds, because the age of majority in New York is 16, these youth had been referred to the program through the adult court. Thus, these youth no longer would be eligible to participate in the study, rendering the relationship the program had established with this court (and judge) effectively meaningless for recruiting youth into the evaluation. While this did not, in itself, mean that the evaluation could not be implemented in this site, it did mean that the grantee would have to, at the very least, radically change how it recruited youth into the program in order to meet the requirements of the study. In particular, it meant the grantee had to establish relationships with a different court system and judges.

Similarly, the YODP grantee in Nashville used a model of pre-apprenticeship in the homebuilding industry, which relied heavily on youth who were 18 years and older because of state and local restrictions on the age at which an individual can work in this industry and due to requirements for youth under 18 to be in school. Until the change in the eligible age range, the grantee had been working closely with a judge from the adult court to obtain referrals; this relationship, too, would have been rendered irrelevant using the revised age range. In order to effectively serve the population referred from the juvenile (or family) court, the grantee would have been required to establish a new relationship with juvenile court authorities and to modify its services to accommodate state and local laws regarding those under age 18.

Specifically, the restriction to a younger age range would have necessitated that the program in Nashville be able to provide formal schooling opportunities to those under age 18. Prior to the beginning of the study, the grantee had not needed to provide high school-based educational instruction because of its heavy focus on those 18 and older. Since state law requires those under 18 to participate in an accredited school, however, serving the revised age range meant that the program had to either develop a partnership with an accredited school (that would coordinate with the vocational-based time demands placed upon these youth) or become an accredited school itself. The site chose the latter approach, but recognized this was a lengthy process during which they would be limited to serving youth in the revised age range that either had already received their GED or were 18 years old. This led to a serious concern over the number of youth

that would fit these parameters, independent of whether a court would be willing to allow them into the study.

Similarly, the grantee in San Diego had been working primarily with older youth, as the administrators of this program felt that other agencies in the community had been serving younger youth effectively. The move to restrict the eligible age range to those youth in juvenile court meant that San Diego would need to reformulate much of its program to better suit a younger target group. While the site was amenable to doing so, this meant that the services to be provided were still being developed as the study was underway. Further, there was some concern over whether they were duplicating some services that already existed in the community; it was this concern that had spurred them to focus on older youth prior to the onset of the study.

In each of these sites, then, the program model that was highlighted in securing the grant to participate in the study was going to undergo substantial transformation in order to participate based on the revised age range. Hence, whatever strength these programs had as a YODP model (and, indeed, they were selected as grantees for the evaluation presumably based largely upon the strength of their model) may have been somewhat undermined as a result of this change. As will be described below, the strength of the model is a critical determinant in the ability to “sell” random assignment to those stakeholders who must agree to participate (in this case, courts); asking sites to make changes from the model that was originally selected as a strong one may have hindered the ability to convince judges and other justice representatives of the utility in participating in the evaluation and of the benefits that could be expected for the youth participating in it.

In at least two other sites, Honolulu and Brockton, the grantees also expressed concern over the restriction in the eligible age range, because there are relatively few adjudicated youth of any age in these communities. Thus, any narrowing in eligibility simply meant a further reduction in the number of youth that could be recruited into the evaluation; as a result, these grantees feared they would not be able to achieve the necessary number of youth given the revised eligible age range. The final site, Philadelphia, had already primarily been working with youth from the juvenile justice system. As a result, this revision posed minimal challenges for this site.

These concerns were first expressed during introductory phone calls held with each site, and reiterated during the initial visits. During the initial site visits, the evaluation team met only with representatives of the juvenile court system because of the restricted age range. Following these visits, the evaluation team expressed to DOL a strong preference for expanding the potential pool of eligible youth back to its original 14-24 range. In consultation with DOL, it was agreed that sites could opt to work with youth in the adult or the juvenile courts, but ideally would not work

with both. Thus, grantees were asked to identify judges and other justice representatives that they would like to work with in *either* the adult court or the juvenile court.

This approach, too, led to concerns on the part of some grantees that working with only a single court system may hinder the efforts to recruit a sufficient number of youth into the study. Further, in at least two sites, grantees that had initially preferred to work with youth referred from the adult court system had, since the restriction to younger youth had been enacted, worked to make contacts with juvenile court judges. These sites were hesitant to break the ties they had established with these judges and work solely with adult court judges. As a result, grantees asked DOL for the ability to work with either juvenile or adult courts, or both, in their effort to solicit cooperation in the evaluation. After a series of discussions, DOL agreed to allow grantees and the evaluation team to proceed in this manner. Once this decision had been made, grantees and the evaluation team were thus able to meet with staff and judges from either or both of the court systems.

The uncertainty concerning the eligible age range from the beginning of the study had an impact on the likelihood that judges and other stakeholders would be willing to take part in the study. Due to the programmatic changes that seemed to be necessary as the age range was shifted, it was difficult for programs to state clearly and concisely during meetings with stakeholders exactly what services would be provided to youth assigned to work with them. Further, in some sites, the successes that programs could otherwise have pointed to among their participants were rendered irrelevant, because the individuals would have been ineligible under the latest requirements for the evaluation or because they were drawn from a population that would not appear before the “current” judge(s).⁵

The period of uncertainty during the first several months of the study likewise prevented grantees from using YODP funds to provide services to youth since these resources were intended for youth who would be part of the random assignment study.⁶ This further limited the grantees’ ability to develop and implement new programs that could be used to convince judges to participate in the study.

⁵ This was especially true in New York. As noted above, the judge in New York that had a lengthy relationship with the grantee would not have been able to participate in the study under the revised age range. Thus, despite having referred many youth that he considered to demonstrate the success of the program, he would have been unable to refer youth to the program to participate in the study. The youth that would have been eligible under the new rules appeared before other judges in different courts.

⁶ Although DOL later did inform grantees that they could use YODP funds to serve youth during this period, several grantees remained uncertain about whether they could do so and, if so, which youth could be served.

While grantees and evaluation team staff tried to convince judges of the potential benefits by noting that another judge had willingly sent youth to the grantee, this had far less effect than if the judges with whom grantees in some sites had longstanding relationships were able to participate. In effect, then, in several sites, the initial meetings with the evaluation team and judges and other justice stakeholders provided little opportunity to “sell” programs that had been selected because of their strong program design and excellent outcomes with the types of youth who would be part of the study. Instead, grantees were forced to ask judges or attorneys what services would be needed to adequately provide for these youth. While there were no sites in which stakeholders refused to participate solely because of this, it is clear that the changing eligibility range and, thus, the changing programmatic elements of the grantees, played a role in lessening the likelihood that the study could be implemented.

Changes in Program Design

Since the grantees and the evaluation team responded to very different requests for proposals, other discrepancies between the programs grantees had been operating under the Youth Offender grant and the requirements of the evaluation meant that some grantees had to substantially change their programs to meet the evaluation requirements. For example, the six grantees selected to take part in the impact evaluation responded to a Solicitation for Grant Applications (SGA) that identified the eligible population as 14 to 24 year old young offenders, gang members, and youth at-risk of connecting with the Juvenile Justice system. Hence, many of the grantees had established programs that enrolled many youth who were not adjudicated. Thus, the requirement of the evaluation (specified in the RFP) that only adjudicated youth be included in the study meant that these grantees would no longer be able to enroll many of the youth they had become accustomed to serving (or, at least, that these youth would not count toward the evaluation requirements) and some would, thus, need to modify their programs once the evaluation was underway. For some grantees, such as the one in Hawaii, it meant that the evaluation represented the first time the grantee had made contact directly with judges.⁷

Further, each of the grantees had been operating a YODP for several years prior to being selected for participation in the impact evaluation. During this time, the requirements did not include establishing an alternative sentencing program or an aftercare program, the two intended treatments under the impact evaluation. Although these components were specified as part of the SGA to which grantees responded to operate programs for the impact evaluation, these were relatively new conditions for the grantees. As such, at the onset of the evaluation, several of the grantees had either not established at least one of these treatments or it was still early in

⁷ This grantee had established contacts with representatives of the probation department, but referrals were only made once it had been determined a youth would not be incarcerated.

development. For example, the YODP grantee in New York had been operating what they (and the cooperating judge) perceived to be a successful alternative sentencing program for youth from the criminal (adult) court, but they had made no effort by the time the evaluation was getting underway to develop an aftercare program. To do so required developing a relationship with New York's Office of Children and Family Services (OCFS), which had not happened before the evaluation team began meeting with the grantee. While there is no reason why this relationship could not be developed fruitfully, the grantee had not taken such steps and, as a result, there was little to persuade OCFS staff of the benefits of the program since it had yet to be established. Similarly, the grantee in Philadelphia had not developed an alternative sentencing program by the time the evaluation began. In fact, only in Nashville and Brockton had the grantees been operating both an aftercare and an alternative sentencing component prior to the onset of the evaluation, though in both cases the number of youth enrolled was quite small. Hence, the grantees (and, thus, the evaluation team) in several sites were compromised in their ability to convince judges or attorneys of the benefits of the program, because the program itself was either entirely undeveloped or very early in implementation.

Misunderstanding of the Requirements of Random Assignment

A further challenge encountered early in the evaluation was that most YODP grantees were not very clear on what random assignment required. Thus, even in those cases in which grantees had sought and secured agreement from judges and other justice officials, typically this agreement was based on a vague or, in some cases, inaccurate understanding of random assignment generally, and the intended evaluation design specifically. For example, in Philadelphia, representatives from the grantee and several justice agencies had been developing a plan to assess all incarcerated youth based on their likelihood of re-offending. These representatives had agreed that this assessment tool would be the primary mechanism by which youth would enter the study. For example, those incarcerated youth who were least at-risk for re-offending (approximately 70 percent of the youth, according to respondents) would be assigned to receive standard aftercare services upon their release from a residential facility, while youth deemed to be at greater risk for re-offending (the remaining 30%) would receive YODP services as part of an intensive aftercare program. The grantee in Philadelphia believed this would meet the requirements of the evaluation because it ensured some youth would receive standard services while some youth received intensive aftercare services. However, such an assignment mechanism is almost the opposite of random assignment, since youth are assigned based on their perceived risk levels and the two groups are known to differ on this characteristic.

Though the concept of random assignment was better understood among other grantees, few, if any, understood that the nature of the design meant there would be substantial changes in the youth that would be referred to the grantee. Specifically, any youth who were to participate in the study must

have committed crimes that were serious enough to warrant consideration for incarceration, but not so serious as to necessitate it, simply because by the nature of random assignment, both incarceration and the alternative to incarceration must have been acceptable for all youth in the study. The narrowing of the eligible pool would have caused a shift in the eligible youth for many of the grantees, but few understood that in the initial study conversations. As with the shift in the eligible age range, most grantees expressed their amenability to shifting to meet the needs of the evaluation, but the fact that it represented a shift may have made it more difficult to sell the grantees, and YODP, to courts and other justice officials already wary of allowing a random assignment of adjudicated youth. Further, the narrower range of eligibility caused concern among several grantees that they would be unable to sustain their staffing levels and service models because the rate of youth entering the study would be substantially slower than had been anticipated.

Concerns About the Legality of the Proposed Design

An additional barrier was the judges themselves. Despite the fact that judges had formally or informally agreed to participate in the study in several sites, during the initial meetings it became clear that none fully appreciated what a random assignment design entailed. Most had signed on to the evaluation believing that it simply meant that some youth would receive services from the grantee while others would not. Judges had assumed, or been told by the grantees, that the decision about whether a youth received YODP services would be theirs to make. Thus, their agreement to participate was not based on an accurate understanding of what the study would require. In each of the sites, then, the evaluation team explained to judges and attorneys how youth would be randomly assigned to the three experimental conditions. Upon learning of what would be required to participate, several judges asserted categorically that they could not participate. In Brockton, for example, the juvenile court judge listened to the evaluation team and, immediately upon the completion of the initial presentation, he said assertively that the design was “not acceptable.” Specifically, he said that the problematic part of the design was using the term “random.” According to this judge, the law does not allow “random” processes to dictate the treatment one receives.⁸ A judge in another site proclaimed that his “kids are not animals” that can be experimented with. Neither of these judges were amenable in any way to participating in the study as designed.

Judges in three sites were somewhat more receptive to the study, but noted that by participating they would be violating the judicial code of ethics and, thus, could not participate. One of these judges asked if there had ever been a study that allowed random assignment of the sort intended

⁸ As noted below, this judge also had substantial concerns about whether it was ethical to deny services to some youth.

in this evaluation. When informed that there had been a single such study that had occurred under quite unique circumstances,⁹ the judge was interested but did not see how similar circumstances applied in her situation. Several judges asked if the evaluation team had found anyone willing to go along with the intended design, so that they could speak with them to identify how the judge had been able to justify it legally.¹⁰ Thus, it was clear that judges in several sites were interested in participating if they could, but were searching for a way in which to legally implement the intended design. Their concern centered on the issue of allowing a random process to determine whether youth were incarcerated or not, which they instead believed was legally the role of the judge.

Other key stakeholders also expressed significant reticence to participate because of concerns about the legality of the design. For example, in one site, a public defender noted that he would be subject to disbarment were he to allow his defendants to participate in the study as designed. Attorneys in multiple other sites expressed similar concerns, though with somewhat less specificity or rhetoric. Each of these individuals' reactions, as well as those of several of the judges described above, was based on a concern that youth who otherwise would not have been incarcerated might have been randomly assigned to this disposition as a result of the evaluation. According to these individuals, this would be a clear violation of due process rights of the youth, as their individual liberties would be taken away based on a random determination rather than a reasoned consideration of the facts of their case.

The evaluation team (and DOL) had understood this prior to beginning the study and, hence, had been clear that the intended design would not mean that some youth who otherwise would not have been incarcerated would now be because of the study. To counteract the fear expressed by judges and defense attorneys, the evaluation team emphasized repeatedly that the evaluation would not result in youth being incarcerated if their crime(s) did not warrant it. Specifically, because it would have been unethical (and illegal) to consider randomly assigning youth who otherwise would have received an alternative to incarceration to be confined instead, the eligible pool could only include those youth who, without the study, would have been incarcerated. Thus, youth participating in the study would not be harmed by their participation (by being incarcerated when they otherwise would not have been), and could only benefit from participating (by gaining access they otherwise would not have to services that were potentially

⁹ Barton and Butts, "Viable Options: Intensive Supervision Programs for Juvenile Delinquents." *Crime and Delinquency*, Vol. 36 No. 2, April 1990, 238-256. This study is discussed in further detail below.

¹⁰ Although at one point we believed a juvenile court judge in Nashville had agreed, and was willing to speak with judges in other sites, ultimately this agreement fell through, and we could not identify any judges that were willing to agree.

beneficial, or even by receiving an alternative to incarceration). While this approach alleviated the concerns of judges and attorneys described above, it presented two other significant challenges.

First, it substantially reduced the number of youth eligible for the study, especially for those in the juvenile justice system. According to the juvenile court judges with whom we spoke, incarceration is the disposition of last resort for youth; a youth's crimes must be numerous or particularly egregious to even merit consideration for incarceration in most districts (and certainly in the six sites with which we were working). Thus, the eligible pool was narrowed greatly by the fact that the intended design virtually required that only those youth whose crimes would merit the possibility of incarceration be included. Further, youth whose crimes were so severe that they simply couldn't be considered for an alternative to incarceration also would have been excluded from participation in the study. Narrowing the eligible pool of youth from both extremes posed a real threat to the ability of sites to recruit sufficient numbers of youth into the study.¹¹ Indeed, representatives in at least two sites noted that recruiting the desired number of youth using the intended design would require at least two full years, rather than the one specified as part of the RFP for the evaluation, while representatives in Hawaii and Brockton, the sites with the fewest adjudicated youth, were concerned that recruiting sufficient numbers of youth would take several years under this approach. Even in New York, the site with the most adjudicated youth, there was substantial concern about their ability to enroll the requisite numbers of youth. In 2005, for example, the family court in the county in which the grantee operated, Bronx County, had incarcerated a total of only 363 youth. This is below the 400 youth required for the study, and does not even factor in that many of these youth would not have been eligible for the study (due to the seriousness of their crimes, refusal to take part, etc.). Hence, this restriction posed a substantial barrier to implementing the intended design.

The second concern raised by the proposed design was one of public safety. Since all youth eligible for the study must otherwise have been incarcerated (to alleviate concerns of due process violations), implementing the study would have meant that some youth that met this criterion would be released into their communities instead. Given that incarceration is a disposition used only in a small percentage of juvenile cases, it tends to be used when there is a perceived threat, either to the individual or to her/his community. Hence, implementing the

¹¹ A second major concern of this approach was that it greatly minimized the external validity of the study, in that any observed impacts of the YODP services could be generalized only to those youth whose crimes were within the very narrow range such that they were serious enough to be incarcerated, but not too serious to merit consideration of an alternative to incarceration. Judges' estimates of the percentage of youth who fell into this range varied between "two or three" percent to "no more than ten percent." While this has little effect on whether the design *could* be implemented, it does substantially affect the generalizability (and, thus, the utility) of any findings were it to be implemented.

intended design would have required judges, attorneys, and the YODP grantees to overlook the perceived threat involved in releasing some portion of these youth into their communities. Doing so could be viewed as a serious ethical violation, and led directly to a third set of challenges encountered in implementing the proposed design: political concerns about the evaluation and its implications.

Political Concerns

As noted in the previous section, one challenge to implementing the proposed design was that narrowing the eligible pool to those who otherwise would have been incarcerated raised significant concerns about public safety. The study design required that 100 youth, who otherwise would have been incarcerated in each of the six communities, be randomly assigned to an alternative sentence. As noted above, juvenile court judges typically view incarceration as the disposition of last resort for the youth in their courts. Given this, the notion that they could, instead, use a random process to allow 100 of these youth to be allowed back into their communities immediately raised grave concerns for these judges.

Further, this concern was heightened in some of the communities in the study that had experienced highly publicized events in which an adjudicated youth was released into the community and subsequently committed additional crimes.¹² For example, in Nashville (and Tennessee generally), the Department of Children’s Services (DCS) is responsible for overseeing those youth who are incarcerated in Tennessee. Under this system, juvenile court judges refer youth who are to be incarcerated to DCS, which then assigns youth to an appropriate facility.¹³ In the months leading up to the beginning of the evaluation, however, DCS had been under a barrage of criticism in the media and from political representatives because several youth that had been released into the community committed further crimes, including homicide in one well-publicized instance. Hence, representatives from DCS were understandably reticent to enter into an agreement that would require they surrender their authority to determine whether a given youth would be incarcerated. At one point during our discussions with the Nashville site, key representatives from DCS had agreed to participate in the study seemingly despite these concerns. Their understanding was that the juvenile court judge would refer youth to them with the order that they be placed into the study,¹⁴ obviating

¹² Although we were able in the proposal to anticipate some of the potential problems to be faced as a part of the evaluation, the site-specific political issues described here and below were generally unknown in advance, in part because they occurred only as the evaluation was beginning.

¹³ Those youth who are to be placed on probation are given their probation terms directly by the judge.

¹⁴ In contrast, the judge had agreed to participate on the assumption that she would refer youth to be incarcerated by DCS, which would then determine that some youth were eligible for the study.

any culpability that DCS itself might have. When DCS stakeholders realized, in a joint meeting with the juvenile court judge, that they would still be responsible for identifying the youth to be eligible for the study, they quickly reversed their initial position and informed the evaluation team and the grantee that they could not participate because of the political implications if a youth recidivated while in the study.

Representatives in other sites expressed similar concerns about the political implications. For example, one judge noted that the study seemed fine in the abstract, but then it would “be implemented, something bad will happen, and it will be my name in the paper. You researchers won’t be mentioned.” This comment conveys the general sentiment expressed by nearly all judicial representatives in the six sites that the possibility of political fallout from a youth recidivating while in the study greatly outweighed any potential advantage of learning about the impact of the YODP services through a rigorous evaluation. Ultimately, this was a critical concern that played a primary role in the inability to implement the study as intended.

There was a second political concern that created a significant challenge in some of the sites. Specifically, as argued below, one of the critical components in successfully implementing a random assignment evaluation in any context is that there should be an excess of potential participants, so that individuals are not denied services to which they otherwise would have access were it not for the study restrictions. This situation did not exist in at least two of the sites in the sample. For example, as noted above, there are relatively few youth that are incarcerated in Brockton. The judge in this county believed that all youth should have the opportunity to benefit from the services of the grantee, not just those randomly assigned to receive these services. When he noted this, the grantee acknowledged that they could serve all incarcerated (and probated) youth with the funding they had received.¹⁵ Thus, assigning only certain youth to receive them effectively meant withholding solely because of the evaluation services to which they would otherwise have had access. The following section discusses how this situation makes random assignment exceedingly difficult and potentially unethical and note how, in the single study in which a design similar to the one intended in this evaluation was implemented, there was a clear and state-mandated excess that enabled random assignment to be used to determine which individuals received services.

An additional challenge was implementing a random assignment evaluation in the context of other reforms in the judicial system that had been ongoing since well before the evaluation. This

¹⁵ Note that the grantee originally received the funds only on the condition that they would implement a random assignment impact study. Hence, it is not entirely accurate that the funding was available to serve all youth; rather, since the funding had been awarded, the grantee was able to speculate that all youth potentially eligible could have been served, independent of the requirements tied to that funding.

was most prevalent in Philadelphia in which stakeholders in the juvenile justice system had been working to develop an assessment tool to classify incarcerated youth into low versus high risk of reoffending.¹⁶ As described above, the grantee mistakenly believed this would allow them to comply with the requirements of the random assignment study. Despite the fact that this arrangement would not allow for the implementation of the intended design, the grantee (and other stakeholders) were adamant about not tampering with this assessment tool, because of the delicate political balance they had achieved in designing and implementing it. Thus, the grantee insisted that any random assignment procedure would have to be implemented in conjunction with this tool. Such political considerations made it exceedingly difficult to identify any way in which the intended design could be implemented, especially in combination with the other challenges discussed above.

Alternative Design

Grantees and court staff in several sites did suggest they would be able to implement a modified version of the evaluation design. This alternative design would have eliminated or reduced many of the concerns expressed above and, as a result, would have been feasible to implement, according to the grantees and court staff in these sites. The alternative would have allowed judges to determine whether a youth would be incarcerated or not, exactly as happens typically. Once this determination had been made, however, youth in each of the two groups (those incarcerated and those not) would have been randomly assigned to either receive YODP services or not. Hence, one-half of the incarcerated youth would have been assigned to receive standard aftercare services, while the other half received intensive aftercare services through a YODP grantee. Similarly, one-half the youth who had been given an alternative to incarceration would receive standard probation supervision while the remaining half would have been assigned to work with the YODP grantee as part of their probation.

This alternative design is similar to at least one other current evaluation of the justice system in which random assignment is being employed after the courts have determined whether or not an individual should be incarcerated. This evaluation is examining the impacts of specific residential care and reentry services on youthful offenders in Florida.¹⁷ In this study, youth (ages 16 to 18) who had been sentenced to residential care in four regions in Florida were then randomly assigned either to enter the Avon Park Youth Academy (APYA) for their care or to

¹⁶ A similar assessment tool had been developed in San Diego, and stakeholders were similarly reluctant to abandon its use as a classification tool for determining what services youth would be provided or mandated to receive.

¹⁷ National Council on Crime and Delinquency. 2007. *Avon Park Youth Academy/STREET Smart Program Evaluation: Preliminary Findings*.

one of several alternative providers that routinely provide such care in Florida.¹⁸ Note that this is directly comparable to one of the two options that would have been available had the alternative design for the YODP evaluation been adopted – youth were randomly assigned either to receive services from the APYA or from a standard care provider only *after* a judge had sentenced them to residential care. Somewhat similarly, the ongoing Transitional Jobs Reentry Demonstration, funded by the Joyce Foundation and DOL, is randomly assigning prisoners who are being released from prison in four Midwestern cities either to have access to a transitional jobs program or only to receive basic job search assistance. Although in this study random assignment occurs only after an individual is being released from incarceration, and thus occurs outside the justice system, it does further illustrate that an experimental evaluation can be carried out if the point of random assignment occurs after an individual is sentenced. Individuals who have been in prison are then randomly assigned either to receive the treatment of interest or not.

A subsequent section of this report briefly discusses the research questions designs of this type can answer, and identifies the specific research questions the alternative design developed for YODP would have allowed the evaluation to answer.

Comparison to Other Random Assignment Studies

The earlier sections described what happened in the attempts to conduct a random assignment evaluation of YODP, how the grantees and court staff in the six sites viewed the process, and the challenges encountered in trying to implement the intended design. This provided some insights into how the evaluation might have been conducted differently to maximize success in those sites. This section examines other random assignment evaluations for additional clues about what helped make such studies successful. In doing so, the following three broad categories of projects are discussed: 1) those that successfully carried out random assignment in a court setting, 2) large-scale random assignment studies conceived and funded by DOL, and 3) large-scale random assignment studies conducted by MDRC. This is not a comprehensive set of such studies, but they are fairly representative, and in the case of studies involving the criminal justice system, the list contains the closest parallels to what was attempted in YODP.

¹⁸ In this evaluation, youth who received residential care through APYA also were provided reentry services, such as community support, and educational and vocational services, through a program called STREET Smart. Both APYA and STREET Smart were funded in part by the U.S. DOL.

Random Assignment Studies Involving the Courts

There have been a number of random assignment studies of the effects of services for prisoners or others involved with the criminal justice system, but there have been few studies in which randomization determined whether someone was incarcerated.

Barton and Butts

As discussed in the introduction, Barton and Butts (1990) conducted a study in which youth in the Wayne County Juvenile Court system were randomly assigned to incarceration or to one of three alternatives to incarceration, including an in-home program operated by the juvenile court and two programs run by private agencies. All three programs involved intensive probation services using small caseloads and frequent client-worker contact. Youth recommended by the juvenile court for incarceration were screened for eligibility, and those with very violent offenses, with a documented history of psychiatric disturbance, or with no potential home in the community were screened out. Those assigned to the programs were placed on probation while their incarceration orders were rescinded.

The circumstances behind this study were somewhat unique. In 1983, in response to the growing number of incarcerations, Michigan limited the number of Wayne County youth who could be committed to the state for supervision and placement to 500 per year. Since Wayne County had committed approximately 700 youth in the previous year, judges in the county were forced to commit substantially fewer youth than they had previously.

In a discussion with one of the study's authors, Jeff Butts, Dr. Butts described a lengthy and often difficult process that required numerous visits and discussions to convince judges to carry out random assignment. The primary approach was to use data from the court system to show judges that there was substantial overlap in the characteristics of those who were committed and those who were not. This convinced judges that they were sometimes uncertain whether incarceration was the correct choice, and therefore that random assignment might be ethically used in some cases to provide an opportunity to learn about the benefits of the alternatives. In short, the researchers were able to couple the unique circumstances of the overcrowding in state juvenile facilities with the judges' acknowledged uncertainty in some cases.

Drug Courts

Drug courts have also occasionally been studied using random assignment, but it is unclear how often individuals in control groups were incarcerated in these studies. In evaluating the Baltimore City Drug Court, for example, researchers at the University of Maryland convinced judges to randomly assign individuals who had been convicted of a drug offense to a drug court

or treatment as usual,¹⁹ but it is not clear from their study how often the usual treatment consisted of incarceration. Clients included circuit court and district court cases supervised by Parole and Probation as well as less serious district court cases to be processed by the Alternative Sentencing Unit. To get judges to agree to let them use random assignment, they used a procedure that was close to one of the alternatives developed to persuade judges to participate in YODP. In particular, randomization results were given to the judge as a recommendation and the judge was allowed to deviate from the recommendation in adjudicating the case. In most cases, however, the recommendation from random assignment was followed, in part because the judges had agreed to participate in the study.

Domestic Violence

Randomization has also been used to study the effects of arrest on spousal abuse cases. The seminal study was conducted in Minnesota in the early 1980s. In that study, police responding to domestic violence complaints were randomly given one of three actions to take: arrest the offending party, order the party out of the residence, or advise the couple on solving their problems.²⁰ A comparison of outcomes for 314 cases assigned to the three groups found that arrests reduced the likelihood of subsequent violence by 50 percent.²¹

Although the results of the Minnesota experiment were promising, the original authors and criminological scholars urged replication. The National Institute of Justice (NIJ) therefore sponsored a set of replications, which were eventually carried out in five sites.²² Each replication involved experimental comparisons of arrest and alternatives and measured victim safety using police records and victim interviews. The replication project is called the Spouse Assault Replication Program (SARP).

To choose sites in the replication, NIJ solicited proposals, which set some guidelines for the sites, such as requiring random assignment to arrest and other alternatives and collecting information on outcomes from both police records and victim interviews. Seventeen law

¹⁹ Gottfredson, Denise C. and M. Lyn Exum. 2002. "The Baltimore City Drug Treatment Court: One-Year Results from a Randomized Study." *Journal of Research in Crime and Delinquency*. 39(3): 337-356.

²⁰ Although this study did involve random assignment to determine whether an individual would be arrested, it is not clear whether those who were arrested were also incarcerated. Hence, while conceptually somewhat similar to the Barton and Butts study, and the intended YODP design, this study may not have involved a random process determining whether or not an individual was incarcerated.

²¹ Sherman, Lawrence W. and Richard A. Berk. 1984. "The specific deterrent effects of arrest for domestic violence assault." *American Sociological Review* 49(2): 261-272.

²² Maxwell, Christopher D., Joel H. Garner, and Jeffrey A Fagan. 2001. *The Effects of Arrest on Intimate Partner Violence: New Evidence from the Spouse Assault Replication Program*. Washington: National Institute of Justice.

enforcement agencies applied to be in the study, and five sites were chosen (Charlotte, Colorado Springs, Dade County, Milwaukee, and Omaha). Each replication required the officers to report whether the case was eligible for the study before they were told the result of random assignment. The new studies also tended to include larger samples (only 296 in Omaha, but 638 in Charlotte, 906 in Dade, 954 in Milwaukee, and 1238 in Colorado Springs) and broader samples (e.g., same-sex couples and female offenders). Overall, the results indicated that arrest reduced the proportion of victims reporting at least one additional act of violence by 25 percent but did not significantly or substantially reduce the number of incidents reported to the police.²³

Child Support

In response to the Family Support Act of 1988's goal of enforcing child support obligations among noncustodial parents, MDRC undertook the Parents' Fair Share (PFS) demonstration, in which fathers with child support arrears were randomly assigned to a special program designed to help increase their earnings and their involvement with their children.²⁴ Those who were not assigned to the special program received standard treatment, which involved incarceration in some cases.

Since child support cases are dealt with in civil court, these fathers would have been incarcerated only as a result of a contempt citation, which meant they would be released as soon as they complied with the court order. Of particular importance as a comparison with the effort to study YODP through random assignment, the fathers were not being incarcerated for a fixed length of time or for a long time. In five of the seven study sites, the noncustodial parent would be released when they paid some amount that was set to be within their means or when they spent enough time in jail to convince the judge they could not make a payment. In a sixth site, the child support enforcement agency could institute a criminal contempt proceeding that would involve more serious jail time, while in the final site the court often set definite times for release.

Even these sanctions were rarely used. In three of the seven study sites, contempt was rarely invoked because the process was more labor and resource intensive than the sites could afford.

²³ This seeming contradiction is due to the fact that those reporting at least one additional act reported a larger number of acts, on average.

²⁴ Miller, Cynthia and Ginger Knox. 2001. *The Challenge of Helping Low-Income Fathers Support Their Children: Final Lessons from Parents' Fair Share*. New York: MDRC. Doolittle, Fred and Suzanne Lynn. 1998. *Working with Low-Income Cases: Lessons for the Child Support Enforcement System from Parents' Fair Share*. New York: MDRC.

Even when warrants were issued, enforcement was given a low priority and noncustodial parents were more often arrested on other charges, such as traffic violations.

Although incarceration is not a major piece of the standard treatment in these cases, conducting random assignment still required cooperation from judges. In contrast to YODP, judges' cooperation was sought before a site was chosen for the study. This was done in several steps. First, when sites responded to a request for proposals to be in the study, they had to show a commitment from the local child support enforcement agency and court system, even though it was not clear what those agencies were committing to. This is similar to what happened in YODP, in which sites submitted letters from local courts along with their applications to be in the study.

The next stage of site selection in PFS narrowed the set of potential sites to those that would conduct a pilot study to demonstrate that random assignment could be carried out, that individuals were being treated in accordance with their assignment, and that the PFS program was being well run. Since random assignment could not be carried out without the cooperation of the courts, project staff met with court personnel before sites were chosen for the pilot phase. It is not clear, however, whether any sites were excluded from the study because judges would not agree to random assignment.

A critical difference between the sites selected for inclusion in the YODP study and those in the PFS study was this second stage, during which sites initially selected for the study had to demonstrate their commitment and ability to participate in the random assignment evaluation by successfully conducting a pilot study of the procedures. This second stage of selection ensured the study would only proceed in sites that could successfully implement the required procedures.

Large-Scale Random Assignment Studies Conducted by MDRC

This paper also examines several MDRC studies that used random assignment to identify whether the lessons learned from these studies could be applied both in understanding why the YODP evaluation could not proceed as originally designed and how the project could have operated differently to allow the intended design to be implemented. Since its inception, MDRC has been involved in a number of large-scale, multi-site random assignment studies that required gaining cooperation of local staff and program operators in conducting random assignment. The Parents' Fair Share project described above is one example. Other examples are the National Evaluation of Welfare-to-Work Strategies (NEWS), a study of welfare-to-work programs in seven locations; the Enhanced Services for the Hard-to-Employ demonstration project of programs in four sites; the Employment Retention and Advancement (ERA) project of 15 programs to help current and former welfare recipients in eight states; the Supporting Healthy

Marriage evaluation of marriage education programs for low-income families in seven sites; and the Evaluation of Child Care Subsidy Strategies (with Abt Associates as prime contractor) of four child care subsidy programs.

These and other studies conducted by MDRC share a number of features that are instructive in understanding some of the challenges encountered during the YODP evaluation. They each began with a concept – such as evaluating programs that promote employment retention and advancement among current and former welfare recipients or those that increase employment among hard-to-employ groups – but did not begin with a set of sites or programs. The initial step was to canvass the country looking for places that were running or could run interesting programs, were willing to carry out random assignment, and were large enough to support a random assignment evaluation. The evaluation team also provided technical assistance to sites during the early stages of implementation to strengthen their programs. Finally, in a number of cases, program funding required sites to agree to use random assignment to determine who would receive program services.

Since these projects share so many features, the next section focuses on one example, the Enhanced Services for the Hard-to-Employ Demonstration. Funded by the Department of Health and Human Services and the Department of Labor, the project's overarching goal was to use random assignment to study programs in six sites that were designed to increase employment among people with serious barriers to work. The project was extremely flexible in that the definition of hard-to-employ was quite broad as were the possible treatments that would be considered. In order to focus the study on important groups or programmatic approaches, the project commissioned about a dozen issue papers that discussed possible target populations (such as substance abusers) and promising programs or approaches. Based on these papers and other contacts, the evaluation team – which included both research and operations staff – identified about 50 potentially promising sites or programs around the country and conducted initial phone conversations with each site. That process identified about 15 sites where the potential for evaluation was strong enough to warrant visits and further discussions and where the intervention was based on a strong theory that suggested it would be able to increase employment.

Eventually four sites were chosen for the project and random assignment was carried out. In Philadelphia, long-term welfare recipients were assigned to two versions of programs designed to help them find work or to a control group. In Rhode Island, depressed Medicaid recipients were assigned to an intervention designed to encourage them to enter treatment or to a control group. In a study of the Center for Employment Opportunities in New York City, recently released prisoners were assigned to a program that included subsidized transitional work or to a less intensive program that provided resources to help them find work. In Kansas and Missouri,

families interested in oversubscribed Early Head Start programs were randomly assigned to receive an enhanced version of the program or to a control group.

In moving from 15 to 4 sites, sites were dropped from consideration for a number of reasons. One criterion was whether the site could provide a large enough sample for a random assignment study to have enough statistical power. Some sites were simply not big enough, either because their programs could not serve enough individuals or they could not recruit enough individuals to create a control group ethically, or both. Another important criterion was whether the program was strong enough or the difference from the control group large enough to warrant an evaluation. At least one site was dropped because it wanted to offer the control group so many services that the evaluation would have been testing a relatively minor differential. Finally, programs that were already running were favored over those that would have to be developed, although one of the four sites in the evaluation (Rhode Island) consisted of a new implementation of an approach that had been tried in a different location with a different target population.

While the original goal was to carry out random assignment in more sites, only four sites were included in the evaluation because only they operate programs worth evaluating. In those four sites, the evaluation is more intensive than originally planned, for example, including more follow-up surveys. It should be noted, however, that more time and money was needed to recruit and develop the four sites than originally budgeted. This is consistent with many studies that MDRC has conducted since it is not easy to find sites that are willing to carry out random assignment, that can carry out random assignment, and that are running programs that are strong enough to warrant evaluation. Finding or developing those programs is a time-consuming and intensive effort.

Random Assignment Studies Funded by DOL

Finally, site selection was reviewed in two large-scale random assignment studies funded and conceived by DOL: the National Job Training Partnership Act Evaluation and the Job Corps Evaluation.

National JTPA Evaluation

From 1982 to 1998, the Job Training and Partnership Act (JTPA) provided a major source of federal funding for job training for disadvantaged youth and adults. When nonexperimental evaluations of similar training programs funded under the Comprehensive Employment and Training Act (CETA) provided somewhat divergent results that depended on analytical and statistical decisions, DOL decided to evaluate JTPA using a national random assignment study.

Three features of JTPA distinguished it from other programs and complicated the selection of evaluation sites. First, service delivery areas (SDAs) were relatively independent and diverse because Congress wanted to reduce the federal government's role in implementing the programs. Second, governance of each program included several organizations, such as Private Industry Councils (PICs), which shared decision-making power and chose the administrative entity responsible for implementing the program. Finally, there were detailed regulations about who would be served by JTPA programs. Due to this structure, SDAs could not be required to participate in the evaluation and recruitment of sites was made more difficult by the complicated governance structure. The diversity of services also made it difficult to standardize treatment across sites.

At the beginning of the study, the evaluation team's goal was to choose a random sample of sites operating JTPA programs in order to get nationally representative impact estimates. However, the decentralized nature of the SDAs made this infeasible. Specifically, site selection began with a stratified random sample of JTPA programs, but none of the first 73 randomly chosen sites approached by MDRC agreed to participate. Of 228 sites that were contacted, 62% cited ethical or public relations problems, and 54% cited concerns that the control group would affect their ranking on performance standards used by DOL to assess each area's performance. After this experience, DOL changed the design to get a convenience sample of sites. Overall, fewer than 10% of sites contacted agreed to be in the study.²⁵

In order to obtain this convenience sample, the design team was confronted with numerous issues, including (1) how the sites should be selected and their participation gained, (2) how to minimize intrusiveness of the evaluation and random assignment, (3) how to choose the point of random assignment, (4) how to allocate participants between the program and control group, (5) how long the control group embargo should last, (6) how to define the treatment, and (7) how large the treatment and control groups needed to be to achieve adequate statistical power. In the end, the team was able to convince enough sites to take part in the study so that results from the National JTPA Evaluation are some of the most useful results available on the effects of training programs for men, women, and youth.

Job Corps

Another major evaluation of a DOL-funded program is Job Corps, which was conducted by Mathematica Policy Research, Inc. The evaluation of Job Corps has somewhat less relevance to YODP, however, because it was much more difficult for sites to be able to decline to participate in the study. Once the national director of Job Corps was convinced that the evaluation was

²⁵ Doolittle, Fred and Linda Traeger, 1990. *Implementing the National JTPA Study*. New York: MDRC.

worthwhile and that random assignment was acceptable, the evaluation team attended regional meetings during which it became clear there was substantial pressure on sites to participate if asked. Thus, rather than recruiting and persuading sites to participate, the evaluation team used these meetings instead to explain the project, to answer questions, and to determine the best way to handle random assignment. As the evaluation proceeded, however, it became clear that some participating sites were resistant to implementing the study. Indeed, one agency resisted implementing the procedures until they had exhausted appeals to their congressional representative even after random assignment had officially begun. The support of the study from the highest levels of DOL's Employment and Training Administration was able to overcome this resistance, however, and ensure the study could be implemented as initially designed.

Lessons

This summary of selected random assignment evaluations points to several lessons that might inform future DOL efforts to conduct random assignment studies.

- **Sites must understand random assignment before they are chosen for an evaluation.** In Parents' Fair Share, for example, a set of promising sites were visited so that random assignment could be explained. This was done both to make sure the sites were willing to carry out random assignment and, just as important, to make sure they understood what random assignment meant for them before agreeing to take part in the evaluation. In their study, Barton and Butts spent considerable time talking to judges in Detroit to get them to accept random assignment. In the Spouse Abuse Replication Project, results from the Minnesota experiment were widely understood and sites were chosen from a group that applied to take part in the replication study. By contrast, in YODP, it was clear that some sites did not understand the basic concepts of random assignment. Moreover, what random assignment meant changed in several ways after the evaluation began. As a result, some sites were not prepared to serve the individuals who would be included in a random assignment evaluation. For example, although the need to serve a middle group of youth that judges could see either incarcerating or not was understood by DOL and the evaluation team at the outset of the study, it was less well understood by grantees, especially because they were not limited in their use of funds to only youth who were within this group. This created numerous issues in trying to convince the courts that they should cooperate with evaluation procedures. One strategy in YODP might have been to choose the evaluator before the sites, allowing the evaluator to visit the sites and local court staff to assess the ability of prospective sites to carry out their part in the study and to explain random assignment and the needs of the evaluation.
- **Do not change the rules for the study after sites have been selected or recruited.** The strategy described above would also have reduced complications that stemmed from differences between the RFPs responded to by the grantees and the evaluation team. Had sites been recruited only after they fully understood

what was to be required of them, they would have had a greater likelihood of implementing the design as intended. Even if this had been the case, however, it is critical that the rules for the study not be changed after grantees are selected. In the case of YODP, the age ranges (and, hence, the courts) to be included were modified well after grantees had been selected, thereby minimizing or making irrelevant some of the foundation work these grantees had done with judges and other court staff. A critical lesson, then, is to establish a foundation for the evaluation that includes specific target groups, and not to alter this foundation unless grantees (and other key stakeholders) can argue persuasively for such a change.

- **Key stakeholders besides the grantees must understand random assignment.** In considering sites for Parents' Fair Share, the study team talked to judges and others in the court system to get their agreement to participate in the random assignment process before a site could be chosen for the study. Similarly, Barton and Butts spent considerable time and effort getting to know the judges and their procedures before they could convince them that random assignment was an ethical alternative given the state-imposed constraint on the number of incarcerations they could make. In SARP, the applicants to be in the study were local criminal justice systems; thus, the application process ensured the group making decisions about incarceration had agreed to random assignment. In YODP, by contrast, it was clear that court staff did not fully understand random assignment and had preconceptions about random assignment that precluded serious discussions about its use in most sites. If the YODP evaluator had been chosen before the sites, as recommended above, this would have allowed the evaluator to explain random assignment to local court staff before sites were selected.
- **Strong programs make it easier to sell random assignment.** In Parents' Fair Share, project staff indicated that judges were more likely to agree to the use of random assignment if they thought the program services provided a valuable alternative to their standard practice. Likewise, in YODP, the conversations with judges proceeded furthest by far in Nashville, where the YODP program model was already in place and easy to understand. By contrast, it was difficult to have meaningful discussions with court staff in the Bronx, Philadelphia, San Diego, and Honolulu. In each of those sites, court staff were clearly interested in promising sets of services for delinquent youth, but it was difficult to convince them the grantees were running strong programs because the grantees had not yet clearly defined interventions that were appropriate to a random assignment study. While their existing programs often focused on an at-risk group of youth, random assignment would require them to serve a middle group of youth who judges could reasonably see incarcerating or assigning to an alternative to incarceration. In addition, their existing programs had to be substantially changed to reflect DOL decisions about the target population that were made after the sites were chosen. In YODP, perhaps the schedule for beginning random assignment could have been delayed to allow the Technical Assistance provider to help the sites develop and strengthen their programs before meeting with the courts.

- **There must be an excess of potential participants so random assignment can be ethically carried out.** In Brockton, for example, the juvenile court judge thought the grantees' services were so valuable that he would not agree to random assignment because it would have meant denying their services to some youth who would otherwise have received them. Since Brockton is such a small community, the grantee believed it was able to serve all eligible youth with DOL resources, eliminating the possibility of an ethical use of random assignment. In JTPA, one of the primary reasons that sites were not interested in the evaluation was due to a lack of excess applicants for their slots. By contrast, in the study by Barton and Butts, this excess was created by the state mandate that placed a formal cap on the number of youth that could be incarcerated.
- **Program funding can be used to encourage sites to participate in random assignment evaluations.** Although the grantees' proposals required them to agree to use random assignment before they received program funding, the court systems were not required to make a similar promise. In Brockton, for example, the juvenile court judge saw no reason to agree to random assignment since he was already sending youth to the BAPIC for YODP services. Perhaps it would have been easier to convince judges to participate in the study if they believed program funding was contingent on this agreement and if the sites had been chosen only after judges understood this.

In a future study, DOL could implement these lessons learned in several different ways.

- **Choose the evaluation team before the grantees.** As mentioned above, one method of attempting to achieve these sometimes stringent goals is to choose the evaluation contractor before determining which sites will be included in the evaluation. The evaluation contractor would have had phone conversations with each of the potential YODP grantees to determine whether programmatic and sample considerations allowed random assignment to be a viable option in the site, followed up by visits to the most promising sites. This strategy might have eliminated from consideration sites such as Brockton, where there were too few potential study participants to fill a control group, or Philadelphia, where the assignment of youth to various disposition alternatives was conducted in a systematic way that seemed to preclude random assignment. Likewise, the evaluation contractor could have had initial discussions with court staff to explain random assignment, to determine whether local conditions made random assignment a possibility, and to try to convince the courts to use random assignment. Doing this before the sites are chosen has two advantages: 1) it allows the evaluation to exclude sites where court staff are unalterably opposed to random assignment, and 2) it allows the funder to use random assignment as a requirement for the court to have access to YODP programs (i.e., in locations where the demand for services exceeds the program's ability to provide them). The second advantage was actually in place in the YODP study, but because grantees had already received the funding and had been using it to serve youth who were not eligible for the evaluation, there was less clarity than could have been the case had funds been awarded only after procedures were put in place. Were this clearer to grantees and court staff, this might also have reduced ethical

concerns about denying existing services to a randomly determined group since the services would not have been available (either at all, if the program had not been funded, or to all individuals, since demand exceeds the ability to provide them) without the use of random assignment. Although one disadvantage of this approach is that evaluation sites might not be nationally representative, limiting the ability of the evaluation to draw inferences more broadly, because the YODP was a demonstration project it was already limited in this regard.

- **Require the courts to apply to be in the evaluation study.** Another possible alternative for YODP might have been to accept applications from the courts rather than the grantees. This was essentially the approach used in SARP, where criminal justice agencies applied to be in the replication study. Although it is possible that court personnel willing to use random assignment might have been in sites with relatively weak YODP programs, it is easier to provide technical assistance to strengthen programs than it is to overcome political or ethical concerns about random assignment.
- **Gain explicit prior agreement from the courts regarding the use of random assignment.** An alternative to having courts apply for the evaluation might have been to ask the sites to obtain explicit agreement from key court personnel on the use of random assignment before they were chosen for the impact evaluation. Although several of the grantees included a Memorandum of Understanding (MOU) with some court representatives or agencies as part of their grant applications, as noted above, there was significant confusion about what random assignment would actually entail such that these agreements meant little in practice. This approach would work best if the evaluation contractor were chosen prior to selecting the grantees/sites.

Even if each of these lessons had been in place for the YODP evaluation, however, it is unclear that the intended design could have been implemented. Although following these suggestions may well have resolved or avoided the programmatic challenges described above, the legal, ethical, and political challenges involved in randomly assigning individuals, especially youth, to incarceration or an alternative likely would still have been too great to implement the evaluation as originally intended. To the extent that such a design can be implemented, it seems apparent that it would require highly exceptional circumstances, such as those that existed to allow Barton and Butts to conduct their study. Absent such highly unusual circumstances, the challenges encountered during the YODP evaluation and identified in this paper are likely to be insurmountable.

Given this, DOL might want to enter evaluations such as YODP with more flexibility in terms of the research questions of primary interest. In the Hard-to-Employ evaluation, for example, finding sites was made easier by the broad range of programs that could be considered. In SARP, some variation across sites was allowed in the alternatives to arrest. In Parents' Fair Share, random assignment was carried out in two different ways, depending on whether potential study participants could be identified through administrative records. If anything, this type of

flexibility is even more crucial in studies involving the legal system, where ethical and legal issues involving random assignment are more difficult than in other studies.

Such flexibility likely could have influenced the implementation of the YODP evaluation, as well. For example, the overarching question to be answered by the intended design was whether adjudicated youth who receive an alternative to incarceration have better outcomes than youth who are incarcerated with standard or more intensive aftercare services. There is little debate that this is a question of great interest to the judicial community (and those who encounter it). The requirements to answer this overarching question, however, make implementing a study that can do so riddled with the challenges described earlier in this paper. Indeed, the stringent requirements of such a study are likely the primary reason only one such study has been conducted, and even that study was possible only under extraordinary and highly unique circumstances.

In contrast, an alternative design that several of the YODP sites had agreed they would be able to implement would have allowed DOL to identify for which youth (incarcerated, probated, and subgroups within these) YODP services are particularly effective. This alternative design was similar in most respects to the intended design. The sole difference is that judges would first have determined the proper disposition (or sentence, in criminal courts) for youth that appeared before them. Thus, youth would have been organized into two groups—those incarcerated and probated—and, within these groups, randomly assigned to receive YODP services or not. While this would not provide an answer to the question of whether alternatives to incarceration lead to better outcomes among youth, it would have allowed the evaluation team to identify whether YODP services had positive impacts on outcomes for youth in each of the two groups. For example, this alternative design would have identified the impacts of YODP services on youth who were given an alternative to incarceration; similarly, it would have revealed the impacts of YODP services on those youth on probation. Further, since nearly all youth offenders would have been eligible to participate under the alternative design, results from this study would have been generalizable to a far broader population of youth offenders than the intended design would have been.

Such results would have helped to determine whether the programs funded through YODP were worthwhile, and enabled DOL, if it opted to do so, to target future funds toward programs that targeted their services to youth that stood to benefit most from them. Additionally, finding that YODP services improved employment and reduced recidivism for certain youth would likewise have given those providing services to youth involved in the criminal justice system an effective program model to follow. As noted above, this alternative design is similar to at least one other current study conducted within the justice system.

Similarly, adopting the alternative design for the YODP evaluation would have required DOL to abandon a research question of critical interest – whether alternatives to incarceration provide

superior outcomes for youth (or similar outcomes at less cost). This would have resulted in an evaluation that identified whether YODP services were effective for youth that had been incarcerated or for youth placed on probation. Such an evaluation would have been conceptually similar to Avon Park study described above. In that study, preliminary results suggest that youth in the treatment group (i.e., those assigned to receive residential care through the Avon Park Youth Academy and reentry services through STREET Smart) have somewhat lower arrest rates in several categories in the one year following their reentry into the community.²⁶ While these results provide potentially extremely useful information about whether those who are committed to residential care should receive such services (or comparable ones) rather than standard residential care, they do not answer the broader question of whether APYA/STREET Smart services produce better outcomes than an alternative to incarceration.

In conclusion, the funder of any evaluation must determine whether the tradeoff of being able to answer many questions of interest at the expense of answering the single question of greatest interest is worth continuing the study. In the case of YODP, DOL decided against pursuing alternative designs. In future efforts, approaching random assignment evaluations with flexibility in the range of questions to be answered may well expand the number and types of programs that can be examined using this rigorous methodology and could increase the understanding of what types of programs are effective in working with a broad range of clients or participants.

²⁶ National Council on Crime and Delinquency. 2007. *Avon Park Youth Academy/STREET Smart Program Evaluation: Preliminary Findings*. Though not all the differences in mean arrest rates reach the level of statistical significance, subsequent regression analysis that controls for pre-assignment risk factors strengthen the observed effects, suggesting the APYA/STREET Smart programs significantly reduce the number of both felony and property arrests compared to standard residential care programs to which youth are sentenced in Florida.