

What Did We Learn From the Evaluation of Project Greenlight?



JAMES A. WILSON

Russell Sage
Foundation

I. Introduction

It is well-established that people who end up in prison or jail have significant deficits in human and social capital. Nearly 75 percent of prison inmates have not completed high school, 83 percent have serious issues with drugs or alcohol, and up to 16 percent are estimated to be seriously mentally ill, with larger proportions reporting mental health problems of varying severity.¹ They typically have poor employment histories and work skills, have few financial and social resources, and may not have the knowledge needed to negotiate complex public-assistance bureaucracies. They overwhelmingly exhibit some degree of antisocial attitudes and beliefs, and tend to associate with peers who hold similar attributes. These attributes develop over the course of a lifetime and are not easily undone.

Substantial deficits in prosocial attributes are a major reason that the rehabilitation of incarcerated individuals is so difficult to achieve. Further declines in human capital and exacerbation of existing antisocial attributes are expressed as part of the reason in the recent literature arguing that incapacitation without rehabilitation is likely criminogenic.² The challenge of reentry programming is to increase an individual's chances of success, partly by helping them to avoid future involvement in the criminal justice system. But how does one develop an intervention that can address, often in a relatively short period of time, the many deficits that have accrued over a lifetime? What types of programs can, at the very least, help reduce criminal behavior and, at the best, help an ex-offender become a productive member of society?

A great deal of progress has been made in the last three decades since the nothing-works era, which deemphasized rehabilitation as a rationale for incarceration. Still, the empirical literature on correctional interventions can often be quite puzzling. Program evaluations of the same or very similar programs in different contexts often show a great deal of variation in the results.³ In some cases, interventions that exhibit strong positive effects in one instance may show no effect in another or, in more extreme cases, be associated with negative effects—actually increasing criminal behavior to a significant degree.⁴ As a result, one of the central tasks of criminological researchers and pro-

gram evaluators is to better understand why, for whom, and under what conditions correctional programs work.

In 2002–2005, I was the principal investigator on an evaluation of Project Greenlight, a prison-based reentry program developed by the Vera Institute of Justice that was associated with negative program effects. In this article, I review the basics of the program and the results of the evaluation, along with more recent evidence, and then discuss what I think the evaluation says about correctional practices and programs.

II. Project Greenlight—The Basics

Developed by the Vera Institute of Justice in 2000–2001, Project Greenlight was a fairly comprehensive program that operated in conjunction with the New York State Department of Correctional Services and the New York State Division of Parole. Greenlight was structured as an intensive eight-week intervention offered in the two months before individuals were released from prison and into the community.

The foundation of Greenlight was a modified version of the Reasoning and Rehabilitation (R&R) program, one of the most commonly evaluated cognitive-behavioral programs. The strongest evidence consistently suggests that interventions based on cognitive-behavioral foundations display the most robust effects in terms of reducing recidivism.⁵ Cognitive-behavioral programming finds its theoretical underpinnings in a very simple notion: Behavior, including behavior that leads to criminal offending, is largely learned. Thus, if criminal behavior is primarily a product of poor socialization or under-socialization, then those individuals can be appropriately socialized—that is, prosocial behavior can be learned.

Program developers condensed the standard six-month R&R program model to fit the eight-week Greenlight structure and increased class sizes to twenty-six from a recommended maximum of eight to twelve. In addition to the core R&R program, Greenlight also included a number of other program elements that had empirical or anecdotal support for reducing recidivism, including practical skills training, drug education and relapse prevention, job skills acquisition, housing assistance, family reintegration sessions, and familiarization with conditions

Federal Sentencing Reporter, Vol. 24, No. 1, pp. 76–79, ISSN 1053-9867 electronic ISSN 1533-8363.
©2011 Vera Institute of Justice. All rights reserved. Please direct requests for permission to photocopy or reproduce article content through the University of California Press's Rights and Permissions website, <http://www.ucpressjournals.com/reprintInfo.asp>. DOI: 10.1525/fsr.2011.24.1.76.

and expectations of parole supervision. Program participants worked with a case manager to develop a detailed release plan that was shared with the supervising parole officer. Finally, the program connected with local community-based service providers to inform participants about local nonprofit and governmental services before their release.

III. Research Design and Results

The Department of Correctional Services identified incarcerated individuals who met all the criteria for participating in the Greenlight intervention, then research staff assigned them to one of three groups: Greenlight, Transitional Services Program (TSP, the Department of Correctional Services standard release program), or Upstate (UPS). Approximately nine to ten weeks before release, Greenlight and TSP participants were transferred from correctional facilities throughout New York State to the Queensboro Correctional Facility in New York City and randomly assigned to one of the two programs at facility intake.⁶ Individuals in the UPS group were released directly from upstate facilities to their home communities in New York City.

The evaluation of outcomes after one year revealed surprising results.⁷ Greenlight participants performed worse than TSP and UPS participants on every measure of recidivism examined: arrests, felony arrests, and parole revocations. In nearly all differences between Greenlight and the two controls, these differences were statistically significant. Furthermore, for both felony arrests and revocations, the TSP group performed worse than the UPS group (although the differences were not significant).

These findings were quite perplexing. How did an intervention that took two years to develop, was positively perceived by staff and participants, and drew on elements with empirical and anecdotal support in reducing recidivism, end up making things worse? And how did both the Greenlight intervention and the standard Department of Correctional Services program result in more detrimental outcomes than UPS, which provided no programming or services at all?

Improper program implementation is often identified as the culprit in evaluations that yield nonsignificant, or even negative results.⁸ Obviously, if a program isn't implemented correctly, it's not likely to achieve the desired outcomes. Problematic implementation can take any number of forms, but the net result is likely to be the same: already limited resources being wasted to achieve null—or, even worse—negative outcomes. Thus, it was important to examine whether Greenlight had been poorly implemented.

Another possibility was that failure rates might shift over time and that the program effects would simply take longer to play out—meaning that perhaps the one-year evaluation was too short. To consider this possibility, we collected new data and extended the analysis from twelve to thirty months.

In addition, we also wanted to examine the results more closely by looking at participants' risk level. Why focus on risk? Targeting higher-risk offenders makes both theoretical (the risk principle) and practical sense.⁹ By definition, higher-risk individuals tend to be characterized by greater deficits in prosocial attributes. They may have more substance abuse problems, poorer social skills, greater impulsivity, stronger antisocial attitudes and beliefs, more problematic peer associations, or some combination of all these—all of which can be targeted through the appropriate intervention. Because low-risk individuals tend to have fewer of these negative attributes, intervening with higher-risk individuals is considered likely to produce greater reductions in recidivism. In addition, targeting limited correctional resources to areas in which the greatest reductions in recidivism can be achieved is likely the best use of those resources. Although negative effects were the overall result of the one-year evaluation of Greenlight, given the literature on risk, we wondered whether these effects might be differentially distributed by participants' risk level. If true, higher-risk participants might show positive effects, with the negative outcomes concentrated among lower-risk participants.

In the follow-up analysis over thirty months, we did find that the effects were differentially distributed by participants' risk level.¹⁰ However, the effects were distributed in a direction counter to that suggested by the risk literature. Low-risk Greenlight participants exhibited moderate positive effects when compared with the TSP group, and moderate-risk and high-risk participants exhibited much larger negative effects. The follow-up analysis also clarified that the UPS group released directly into the community with no pre-release programming performed better than both of the groups transferred to the Queensboro facility.

We were now left with two sets of anomalous findings to explain. First, why did both of the groups that were transferred to Queensboro and received pre-release programming have much worse outcomes than individuals released directly into the community with no programming at all? Second, contrary to what the literature suggests, why did the moderate-risk and higher-risk Greenlight participants do so much worse and the low-risk Greenlight participants do so much better?

IV. Making Sense of the Unusual Findings

Why did those in the UPS group, who were not transferred and who did not receive any programming, do better than individuals in the Greenlight and TSP groups?¹¹ The most plausible explanation is related to the involuntary relocation and coerced programming that the Greenlight and TSP groups were subjected to right before release. The Greenlight program designers thought that bringing participants back to their home community would help the reintegration process by exposing them to local community-based service providers. However, the empirical evidence asserts—and most people will generally recognize—that when individuals are embedded in a

community (even a prison community), have social networks, and have a semblance of control over their lives, they are likely to have more positive outcomes. After all, these goals were exactly what the Greenlight program aimed to achieve—to embed ex-offenders in a community, provide social networks, and give them a semblance of control over their lives. Doing these things is, in a very fundamental way, what reintegration means.

However, by forcibly relocating individuals right before release, what the program might have done was uproot them from a familiar community in which they felt secure and force them into an unknown, perhaps somewhat confusing community without enough time to adjust to the situation. Furthermore, once in the new community, inmates were forced to participate in programming under the threat of loss of good-time release. Sufficient empirical evidence indicates that people who experience involuntary moves and lack a sense of control over their lives experience such negative emotional states as anger, frustration, and increased stress. Furthermore, all of these emotional states have been tied to poorer life outcomes of various sorts, including criminal behavior. This explanation seems the most plausible reason why the UPS group, whose lives were not disrupted right before the end of their sentence, did so much better than the two groups that were transferred.

But why did the higher-risk Greenlight groups do so much worse than the TSP groups, except those classified as low risk? Poor program implementation does not seem to explain the differences in outcomes in this case; else, why would low-risk participants benefit? One component of program implementation is responsivity—that is, targeting an intervention to the learning styles and abilities of the participants—a factor that needs to be taken into account.

The risk principle says that the most intensive interventions should be targeted to higher-risk offenders where the greatest reductions in recidivism might be achieved. In this case, a program already considered intensive was made even more comprehensive and intensive. One might initially think that making a program more comprehensive would make it even more effective. However, making a program more intensive does not necessarily make it appropriate.

By definition, low-risk inmates are likely to have better attention spans and verbal skills, display less impulsivity, and generally hold more attributes that make them receptive to programming. In short, they're likely to be less disruptive and typically better students. Taking these factors into account implies that lower-risk individuals were more likely to easily absorb and process the compressed Greenlight program, especially in the short time frame in which it was presented. Those individuals identified as higher risk have fewer of the attributes that make them receptive to programming of this type. By condensing the program and adding additional program elements, the program may have been too intensive for the abilities of

higher-risk inmates, especially when presented in such a short period of time.

V. What Were the Major Lessons?

Our experience with Project Greenlight taught us lessons both specific to corrections and more generally about program development. In corrections, the project findings raised two important issues. First, our analysis raised questions about the potential negative consequences of prison transfers, especially when they occur right before release. Although we found no empirical studies about the effects on inmates of transferring them from one location to another, empirical evidence from nonprison studies strongly suggests negative effects for individuals who experience involuntary moves. The possibility that such findings also apply to incarcerated individuals is reasonable enough to at least warrant further investigation.

Second, our analysis adds to the literature on correctional interventions and risk. Individuals of differing risk levels tend to have varying needs and disparate learning styles, making it difficult to develop a single, broad-based program that will be equally effective for all. The outcome also suggests that some intensive interventions may be too intensive for moderate-risk and higher-risk inmates, especially if delivered in a short period of time. This conclusion is consistent with the responsivity principle—that is, programming needs to be delivered in a manner consistent with the abilities, learning styles, and motivations of the individuals involved. Perhaps more important, our findings also suggest that programming can be delivered to low-risk inmates and still achieve significant reductions in recidivism. Prior research indicates that cognitive behavioral programming has shown larger effects with lower-risk inmates than with higher risk, although such differences are negligible.¹² Assessing risk and needs was an afterthought in the Greenlight intervention and likely was one of the underlying program flaws that led to the results described here.

Greenlight program developers had the best of intentions in trying to help ex-offenders return successfully to their communities. Intervening to change human behavior is a complicated, serious process that has the potential to produce positive behavioral change. However, because of the unintended consequences that are part and parcel of nearly every intervention, intervening can also make things worse. It seems important then, to realize that interventions intended to help can produce harmful, as well as beneficial outcomes. Thousands of programs are being carried out in the United States every day—in corrections, K–12 education, health and human services, and so on—with many of them, due to lack of rigorous evaluations, unclear about their effects. Our evaluation of the Greenlight intervention highlighted some unintended consequences and provided some explanations about why negative program effects occurred. In so doing, the failures of the program have provided greater insights into

some of the intricacies of correctional interventions and perhaps can inform future correctional practice.

Notes

- ¹ CAROLINE W. HARLOW, BUREAU OF JUSTICE STATISTICS SPECIAL REPORT: EDUCATION AND CORRECTIONAL POPULATIONS (U.S. Department of Justice, Jan. 2003); S. BELENKO, TRENDS IN SUBSTANCE ABUSE AND TREATMENT NEEDS AMONG INMATES (U.S. Department of Justice, 2002); Paula M. Ditton, Mental Health and Treatment of Inmates and Probationers, (U.S. Department of Justice, 1999).
- ² See S. N. Durlauf & D. S. Nagin, *Imprisonment and Crime: Can Both Be Reduced?* 10 CRIMINOLOGY & PUB. POL. 9 (2011).
- ³ M. W. Lipsey & Francis T. Cullen, *The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews*, 3 ANN. REV. L. & Soc. Sci. 297 (2007).
- ⁴ E.g., J. McCord, *Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs*, 587 ANN. AM. ACAD. POL. & Soc. Sci. 16 (2003); C. T. Lowenkamp & E. J. Latessa, *Increasing the Effectiveness of Correctional Programming Through the Risk Principle: Identifying Offenders for Residential Placement*, 4 CRIMINOLOGY & PUB. POL'Y 263 (2005).
- ⁵ D. A. ANDREWS & J. BONTA, *THE PSYCHOLOGY OF CRIMINAL CONDUCT* (4th ed. 2006).
- ⁶ The Greenlight evaluation used a strong quasi-experimental design to assign Greenlight and Transitional Services Program participants for the first few months until a true randomized design was instituted.
- ⁷ J. A. Wilson & R. C. Davis, *Good Intentions Meet Hard Realities: An Evaluation of the Project Greenlight Reentry Program*, 5 CRIMINOLOGY & PUB. POL. 303 (2006).
- ⁸ E.g., C. T. Lowenkamp, E. J. Latessa, & P. Smith, *Does Correctional Program Quality Really Matter? The Impact of Adhering to the Principles of Effective Intervention*, 5 CRIMINOLOGY & PUB. POL. 575 (2006).
- ⁹ E.g., ANDREWS & BONTA, *supra* note 4.
- ¹⁰ James A. Wilson & Christine Zozula, *Risk, Recidivism and (Re) Habilitation: Another Look at Project Greenlight*, PRISON J. (forthcoming 2012).
- ¹¹ One potential explanation for at least part of our findings is that despite our rather rigorous research design, some sort of selection bias plagued our assignment protocol. As an example, perhaps individuals who were lower risk somehow came to characterize the Upstate group and somewhat higher risk individuals came to characterize the Greenlight group. Although it appears that the Upstate group might have been somewhat lower risk, our descriptive and multivariate analyses suggest that differences in the risk levels between the groups did not account for the differences in success (or failure) rates we observed. We are therefore fairly confident in our research design.
- ¹² L. S. J. Tong & D. P. Farrington, *How Effective Is the "Reasoning and Rehabilitation" Programme in Reducing Reoffending? A Meta-Analysis of Evaluations in Four Countries*, 12 PSYCH., CRIME & L. 3 (2006).