

# Youth Violence and Juvenile Justice

<http://yvj.sagepub.com>

---

## **Evaluability Assessment in Juvenile Justice: A Case Example**

James O. Finckenauer, Satenik Margaryan and Mercer L. Sullivan

*Youth Violence and Juvenile Justice* 2005; 3; 265

DOI: 10.1177/1541204005276267

The online version of this article can be found at:  
<http://yvj.sagepub.com/cgi/content/abstract/3/3/265>

---

Published by:



<http://www.sagepublications.com>

On behalf of:



[Academy of Criminal Justice Sciences](#)

**Additional services and information for *Youth Violence and Juvenile Justice* can be found at:**

**Email Alerts:** <http://yvj.sagepub.com/cgi/alerts>

**Subscriptions:** <http://yvj.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations** <http://yvj.sagepub.com/cgi/content/refs/3/3/265>

## EVALUABILITY ASSESSMENT IN JUVENILE JUSTICE

### A Case Example

James O. Finckenaer  
Satenik Margaryan  
Mercer L. Sullivan

*School of Criminal Justice, Rutgers University–Newark*

*An evaluability assessment of three programs of the Juvenile Justice Commission of the State of New Jersey was undertaken. Operation Hook-Up, the Hudson County YAP Challenge Program, and the Monmouth County Bullying Prevention Project were assessed. This included interviewing program personnel, site visits, reviewing the record-keeping systems, and meeting with the staff of the Juvenile Justice Commission and of the programs. This assessment enabled the researchers to acquire firsthand knowledge of the programs that led to the development of tailored evaluation designs for two programs. This article describes the programs, the assessment, and the bases for our recommendations.*

**Keywords:** *evaluability assessment; evaluation; juvenile justice; New Jersey*

Evaluation, and specifically policy and program evaluation, is often talked about in public policy circles. Indeed, it is a kind of sine qua non for policy “wonks.” This is especially so with respect to criminal justice policy. Talking about it and actually doing it are, however, quite different things. This is because supporting evaluation—with all that entails—is akin to supporting motherhood and apple pie. Who could possibly be opposed to doing rigorous assessments of what we are doing, especially when we are doing it with the taxpayers’ money? Well as it turns out, a lot of people are opposed. In other words, they support evaluation in the abstract, but not with respect to specific program A or B, in which they might have some vested interest in seeing A or B survive and grow. Or, they support

---

**Authors’ Note:** We wish to thank the staff of the Juvenile Justice Commission of the State of New Jersey, the Hudson County YAP (Youth Advocacy) Program, the Monmouth County Bullying Prevention Project, and Operation Hook-Up for their willingness to spend time with the researchers and their openness in discussing individual programs. Points of view expressed in the article are those of the authors and do not represent the official position or views of the Juvenile Justice Commission of the State of New Jersey, nor opinions of the staff of the Hudson County YAP Challenge Program, the Monmouth County Bullying Prevention Project, and Operation Hook-Up.

*Youth Violence and Juvenile Justice*, Vol. 3 No. 3, July 2005 265-275

DOI: 10.1177/1541204005276267

© 2005 Sage Publications

only an evaluation that produces the results they want. Evaluations are unfortunately subject to being politicized and otherwise manipulated in ways and to purposes not envisaged by the evaluators.

Given these and other realities about evaluation—that they are indeed costly, time-consuming, difficult to do with respect to maintaining the integrity of the original design, can be vastly unpopular if they do not produce the desired results, and so on—we describe an approach that helps maneuver through this mine field. In our scenario, we make two assumptions as reflected in our answers to the following questions: (a) Can you evaluate anything in the criminal justice policy arena? Yes, we believe you can literally evaluate anything at least in some way, at some level, and at some cost. (b) Should you therefore evaluate everything? No, you should not. How then should you determine what you evaluate and what you do not? The answer is to do an evaluability assessment, and that is the subject of our article.

### **What Is Evaluability Assessment?**

Evaluability assessment tests the readiness for evaluation. A rationale for doing so was described in an internal National Institute of Justice (NIJ; 2002) evaluation strategy paper:

It is not always possible to determine the impact of a project. Sometimes there is no clear connection between what program operators say they want to accomplish and what they actually do. Spending some modest funds to interview program staff, identify key program elements, and assess the regularity of program efforts helps . . . develop design options and often protects against wasteful evaluation investments. (p. 2)

Is it worth committing the resources to actually do an evaluation? In addition, will decision makers be likely to use the information produced?

An evaluability assessment should clearly precede any formal evaluation effort. Rossi and Freeman (1985) define evaluability assessment as “a set of procedures for planning evaluations so that stakeholders’ interests are taken into account in order to maximize the utility of the evaluation” (p. 60). The purpose is to determine whether a program is evaluable and to what extent (Twain, 1983). In addition, evaluability assessment often is used to determine whether evaluation is likely to contribute to improved program performance and be a catalyst for program change (Thurston, Graham, & Hatfield, 2003; Wholey, 1994). As indicated above, this includes taking into consideration the commitment of program administrators and other key decision makers to use whatever information might be supplied by an evaluation. It is important to the success of the formal evaluation to achieve the cooperation of administrators, staff, and other stakeholders in the program. Among the aims of an evaluability assessment are to produce a systematic description of the program, to pinpoint the key question or questions to be addressed by any evaluation, and if warranted, to propose a formal evaluation.

Evaluability assessment is frequently utilized in studies of health services (Thurston et al., 2003; Thurston & Potvin, 2003). Criminal justice researchers have recently started utilizing evaluability assessments not only as a preliminary step to a formal evaluation but also as a tool of improving program performance as well. For example, Matthews, Hubbard, & Latessa (2001) used an evaluability assessment as a tool for improving correctional

programming. Poulin, Harris, and Jones (2000) used an evaluability assessment to determine how a Philadelphia delinquency prevention program actually defined success.

In this article, we describe an evaluability assessment of three programs operating under the auspices of the Juvenile Justice Commission of the State of New Jersey. The three programs (all in New Jersey) are called Operation Hook-Up, the Hudson County YAP Challenge Aftercare Program, and the Monmouth County Bullying Prevention Project. Let us begin with a brief description of the Commission and of the three programs.

### **The New Jersey Juvenile Justice Commission**

The programs we assessed represent contracted services for juveniles and operate under the oversight of the New Jersey Juvenile Justice Commission (NJ JJC). The Juvenile Justice Commission (JJC) was created by statute in December of 1995 to centralize reform efforts in the juvenile justice system of New Jersey (NJ JJC, 2002). Before its inception, the state's juvenile justice system was sprawled across three different departments in the executive branch of the state—the Departments of Law and Public Safety, Human Services, and Corrections, in addition to various private agencies. The JJC was created to centralize authority for planning, policy development, and service provision within the juvenile justice system (Keller, 1995).

### **The Programs**

The JJC approached the Rutgers University School of Criminal Justice in the fall of 2002 about the possibilities of a research collaboration. Following a series of meetings, the JJC was asked to nominate several of its programs for which they had a particular interest in evaluation. That interest could be, and indeed was, driven by a number of factors—refunding decision pending, significant investment already made, novelty of the approach, the politics of juvenile justice in New Jersey, and so on. Although the JJC was requesting an evaluation, we proposed to first do an evaluability assessment of the three nominees, and that idea was accepted.

#### *The Hudson County YAP (Youth Advocacy) Challenge Aftercare Program*

The YAP Challenge Program provides case management and advocacy services to youth returning to the community from confinement in the state training school. These services are intended to be over and above what is provided by parole. The clients who are selected for the Challenge Program come from among the pool of training school parolees returning to Hudson County. Hudson County, just across the river from New York City, is one of the most urbanized counties in New Jersey. The youth chosen for YAP are so selected because they are judged to be at particularly high risk for failure and return to confinement, through either a violation of parole or rearrest. This criterion is an important consideration in thinking about an evaluation.

The program model for the intervention is an adaptation of an approach developed by a contractor, a nonprofit organization that has implemented this model for youth who are high risk in a number of settings in several U.S. states for a period of several years. However, this particular implementation seemed novel in targeting youth who were extremely

high risk—the parolees who were the highest risk returning in the county. Program activities involve meeting with youth and attempting to help them become involved in appropriate, prosocial activities, such as finding jobs, enrolling in and maintaining attendance at education and training programs, and getting access to substance abuse, mental health, and physical health services. Program staff does not itself supply these services but rather acts as advocates and liaisons with the providers of the services. Program staff also works and meets regularly with parole officials responsible for the community supervision of the youth.

#### *Operation Hook-Up*

Operation Hook-Up is a relatively new program just beginning to serve clients. The very first client was enrolled in August 2003 and was the first of a projected total of 25 clients to be served during the course of a 1-year contract period. There are two distinctive elements to this program, one deriving from the program model, the other from the nature of the partnership between the organization operating the program and the JJC that provides the funding.

The program model calls for the program to support a successful transition from the Stabilization and Reintegration Program, or so-called juvenile boot camp, back into the community from which the youth come, in this case the city of Paterson, New Jersey. A unique aspect of this particular organizational partnership is the fact that the organization providing the services is a faith-based organization, Pilgrimage Outreach Ministries. Faith-based initiatives in the United States have been much encouraged by the Bush Administration. The Pilgrimage Outreach Ministries is, in this case, receiving financial support directly from the state government. This is a new kind of partnership in the State of New Jersey and is part of a growing movement across the country toward developing such partnerships. The potential benefits and obstacles involved in such partnerships are not well understood. Potential benefits may derive from the community base and shared values of a faith community. On the other hand, the relative lack of experience and developed organizational structure may pose obstacles. In addition, attempts to work across the historical separation of church and state may pose challenges, either politically or in the form of conflicts between the shared values and established practices within the faith community, and the operational practices and regulations of state government. These are all issues to be considered in the evaluability assessment.

#### *The Monmouth County Bullying Prevention Project*

The third program we looked at was the Bullying Prevention Project in Monmouth County. Monmouth County provides considerable variation in its demographic and socioeconomic makeup. Areas vary from fairly old inner cities such as Asbury Park and Neptune, to quite wealthy communities on the oceanfront and in the outlying suburbs, many of which are bedroom communities for New York City. The project itself is being provided by a private nonprofit organization, Prevention First. It is partially funded by the Federal Office of Juvenile Justice and Delinquency Prevention under the Title V Incentive Grant for Local Delinquency Prevention Programs.

One of the intriguing characteristics of this project is that it has quite explicit theoretical grounding. It is based upon the work of Olweus (1978, 1993) on bullying and, more specifically, on the Bullying Prevention Program—a Blueprint Model Program from the Cen-

ter for the Study and Prevention of Violence, at the University of Colorado at Boulder (Olweus, Limber, & Mihalic, 1998). The model's approach to bullying prevention is characterized as being a "universal intervention for the reduction and prevention of bully/victim problems."

Bullying Prevention is being implemented in four school districts of Monmouth County that have been identified as being the most at risk—Asbury Park, Keansburg, Long Branch, Neptune, and Red Bank. These school districts represent only 9% of the juvenile population of the county; however, they collectively account for 43% of the admissions to the County Youth Detention Center (Prevention First, 2002).

### **Methods Employed in the Evaluability Assessment**

Our assessment included seeking answers to a host of questions. In seeking these answers, we employed a variety of approaches. These included interviewing program personnel and representatives of the JJC. We made at least one site visit to each program. We collected and analyzed program documents, such as funding proposals, published brochures, administrative manuals, annual reports, minutes, and any existing completed evaluations. In addition, we examined the record-keeping systems of the programs to determine what would be the need for additional data collection if a formal evaluation were to ensue.

Program administrators and staff were asked to describe their programs in detail. This give-and-take process permits the researchers to get a sense of just what the staff persons in the program believe they are doing. This may or may not be consistent with what program documents indicate. You can also discern where there might be problems, things about which the staff might be defensive, and just how well actual practice is adhering to the plan. The site visits and review of program documents provide information on the size and characteristics of the actual or projected target population, and on project goals and intervention activities.

The three programs differed widely in the extent of available information. Because it was being partially federally funded and had to meet certain funding requirements, the Bullying Prevention Program had already conducted a wide range of internal evaluation-type activities. Thus, evaluation reports as well as multiple funding applications existed for the program. The YAP Challenge program had a few funding proposals as well as a written program description, but no evaluation information as such. Operation Hook-up was the most bereft of information, having only a brief description of the program that had been submitted to the JJC. In the latter case, we were very dependent on verbal descriptions of the program from JJC staff and Operation Hook-up staff.

The final step of the evaluability assessment was a joint meeting with all the program stakeholders and JJC representatives. Here we shared the results of our assessment and presented recommendations and a formal evaluation plan.

### **Findings**

As a result of the joint meeting, and consistent with our recommendations, an agreement was reached to proceed with an impact evaluation of the Challenge program, and a small-scale qualitative process evaluation of the Operation Hook-Up. Our recommendation, which was accepted, was not to undertake a formal evaluation of the Monmouth Bully-

ing Prevention program. We thus concluded with different recommendations for each of the three programs. In the following we describe how and why we came to these conclusions.

*The Hudson County YAP (Youth Advocacy) Challenge Aftercare Program*

Interviews with the program's director and conversations with JJC staff indicated that a significant change in program operations had occurred beginning in September 2002. At that point, a new director was hired who began to revamp operations. This raised the possibility that the intervention in the early period of the program was, in some sense, not the same as that implemented post-September 2002. This posed the issue of whether all individual outcomes could and should be combined, or whether outcomes for the latter period should be examined separately.

Next, information obtained from the single parole officer who had been responsible for the Challenge Program parolees since the beginning of the program indicated a high rate of program failure among those who had completed the program. Of 29 youth who had been discharged from the program and classified as either positive or negative discharges, 20, or almost 70%, were classified as negative, indicating either violation of parole or rearrest on new charges. Obviously, failure rates of this magnitude in an intervention sample would seem to limit at the outset the amount of improvement over a control or comparison group that could be expected.

We nevertheless recommended that an evaluation proceed in the face of these failure rates for several reasons. First, although the failure rates were high, they were not necessarily proof that the program was having no positive effects. Recidivism rates in most studies are typically very high. Whether the YAP Challenge rates are abnormally high is difficult to gauge. Juvenile recidivism rates have been studied far less than adult rates. In addition, the Challenge Program serves, by design, a very high-risk group of parolees. It is at least conceivable that a similar group of young parolees, with comparable risk scores, might have even higher recidivism rates.

Second, as mentioned above, there were indications that the Challenge Program was implemented differently during different time periods. And a very preliminary "eyeball" analysis of the flow of the parole records over time did indeed suggest a higher success rate in the more recent time period.

Absent the random assignment of an experimental design, the next most powerful research design makes use of a control group. Thus, to conduct a formal outcome evaluation of YAP Challenge, we recommended choosing an appropriate control or comparison group, whose members would be as similar as possible to those entering the Challenge Program. Fortunately, there were two viable possibilities for accomplishing this.

The evaluation design and work plan we proposed took account of all the above considerations. The basic design was as a quasi-experimental design using a within-Hudson County control group including violation of parole and rearrest (if possible) as outcome indicators. The evaluation proposed to assess failure versus success, and time to failure. It was further proposed to then assess costs and benefits based on the recidivism measures.

To aid in the interpretation of the quantitative outcome data, we proposed further to conduct a process evaluation to identify factors that might be related to program effectiveness. This effort would concentrate on the latter stages of the program because it was believed that that period better represented the current state of operations, and also because there had been staff turnover, and there was a greater likelihood of being able to interview



more recent and current staff members. This phase of the work included conducting one focus group and interviewing selected individual staff members. We also outlined a process for extending the process evaluation to program participants and, if deemed feasible, to a control group of nonparticipants who were also on parole in Hudson County.

### *Operation Hook-Up*

Because this program had enrolled only one client at the time of our assessment, it was obviously not going to be possible to conduct a formal outcome evaluation for some time. We decided, however, that it would be possible, and so recommended that a process evaluation of the start-up and implementation phases of the project be conducted. Often referred to as a formative evaluation in the research literature, this type of research looks not at whether the program works, in terms of measured outcomes, but rather what the goals of the program are, how they evolve in the early phases as the various stakeholders take initial program ideas and translate them into a set of program practices, and what kinds of opportunities and obstacles arise along the way. After a certain period of time, a new program should be expected to have developed a distinctive approach, including a stated mission and a concrete set of procedures for carrying it out. When the program has settled down, it then makes sense to consider a formal outcome evaluation. The latter will be greatly enhanced by the findings of the formative or process evaluation. The process evaluation thus documents the creation of programs and provides a structure for designing a subsequent outcome evaluation that uses measures appropriate to what the program is really trying to accomplish.

Qualitative methods are most appropriate for process evaluations. By observing program operations directly and interviewing the stakeholders at various intervals, an evaluation team can systematically track the evolution of the program in its community and within its organizational environment over time. We proposed to embed a researcher in Operation Hook-Up, who would record observational and interview notes and develop a coding scheme that would provide the basis for analysis and preparation of a report on the implementation of the program and the opportunities and obstacles encountered.

This kind of evaluation would ask two kinds of questions related to the two distinctive features of the program. One set of questions has to do with the program model: Is the program able to recruit participants, to engage them in planning for release, and to provide contact and services during the 8-week aftercare cycle and subsequently? The other questions have to do with the evolving nature of the innovative direct partnership between a faith-based organization and state government: Do the two entities appear to work together effectively? If not, what are the problems and how do people deal with them? Does the faith-based organization appear to provide unique kinds of resources, and, if so, what are they? How do they work? Are there particular kinds of challenges faced by small, community-based organizations (faith based or otherwise) attempting to partner directly with state government? Are there examples of how such problems can be worked out?

To address these questions, we proposed to observe and interview a number of different kinds of stakeholders in the process. These include the juveniles themselves and their family members, correctional officials, program providers, and other members of their faith-based community. Interviewees also include officials of state government who provide support and monitoring for the project and who are responsible for managing this new kind of partnership within the larger context of state government.



*The Monmouth County Bullying Prevention Project*

After our site visit to this program and a review of program literature, our assessment was that it would be premature at this time to proceed with an outcome evaluation. A combination of program-specific structural, methodological, and contextual problems determined our decision not to propose a formal evaluation of the Bullying Prevention Project.

First, the program administrators are currently in the process of adapting an established and reputable program model to local circumstances. That process of adaptation, in our judgment, needs to advance much further before it would be appropriate to assess it. The Olweus et al. (1998) model that is being used here was originally developed for children significantly older than those participating in the program in Monmouth County. This age difference has consequences for program content and for the adaptation of the evaluation measures that would be needed to assess program outcomes. Thus, the education curricula, the evaluation methods, the questionnaire items, and the outcome measurements used in the Olweus model, all need to be adapted for use with the early elementary school students in Monmouth County.

This is not exactly an uncommon situation, as it can be expected that every program model must be adapted and refined to some degree before being implemented in a new environment. It is just that in this case, the adaptation is quite substantial—particularly with respect to the target population. We believe that the program needs time to attempt to complete this adaptation process before being critically evaluated.

Second, during our meeting with the Prevention First staff, we posed questions about the dosage and intensity of the antibullying activities being undertaken in the elementary schools. The notion of the so-called strength of the dosage has been borrowed by program evaluators from the medical field and basically refers to the question of how much program is actually being delivered to the target population. Just how intense is it? In Monmouth County, although the program has conducted various activities throughout the four school districts (see program description above), the number of such activities differs by school and sometimes by classroom as well. The program representatives indicated during our meeting that it is the teachers and other school staff, for example, bus drivers, custodians, and cafeteria workers, who bear the brunt of implementing the bullying prevention activities, after orientation and training by Prevention First personnel. These teachers, as well as the auxiliary school staff, are already heavily taxed and, in some cases, overwhelmed (in the view of program administrators) by other state-mandated requirements. Consequently, the Prevention First representatives were pessimistic about the quality and quantity of the bullying prevention activities actually taking place in the schools. It seemed clear that the strength of the dosage of program activities in any given school varied, perhaps considerably, depending on the dedication of the particular school's personnel to bullying prevention, and the requirements of their other academic and supervisory duties.

In our view, these issues of the dosage and intensity of the delivered services would threaten the internal validity of any assessment of the program impact. An evaluation that did not consider these issues would thus produce invalid results that would most likely be negative. In sum, any evaluation would come up with ambiguous results regarding the impact of this program on reducing the incidence of bullying and improving the overall school climate.

Under normal circumstances, one could argue that these might seem like reasonable grounds for a process evaluation that would document concerns and problems regarding the integrity with which the program is being implemented. Normally that would be true. How-

ever, given the current political climate in the State of New Jersey (as well as nationwide in the United States with respect to school violence), it was our judgment that the state policy makers might not be ready for any evaluation that produced ambiguous or less-than-glowing results. This is because bullying in schools has become a highly politicized issue, not only in New Jersey but elsewhere as well. The fact that the results would be invalid would be lost or disregarded, and the attention would focus on the probably bogus negative findings that the program did not work.

After such deadly school violence incidents as the Columbine High School shooting, several U.S. states have introduced laws mandating schools to have programs and plans for dealing with all minor and major incidents of violence and bullying. In 2001, for example, the Washington State Senate passed legislation requiring schools to adopt strict policies against harassment, intimidation, and bullying (Verhovek, 2001). New Jersey followed suit by enacting its own bullying prevention legislation in 2002. That law mandates that school boards create a policy against bullying—prohibiting harassment, intimidation, or bullying on school property, school buses and school functions (Sullivan, 2002). The law also requires that school policies include procedures for reporting and investigating bullying incidents. Thus, in such a concentrated political environment, any evaluation of a bullying prevention program would have the potential for receiving a considerable amount of attention and publicity. Given the tendencies of the media and of political officials to simplify and sometimes sensationalize issues, a program evaluation that did not necessarily provide a clear answer to the question of whether the program works might be overinterpreted and prove fatal to the program's continued existence. Premature results that could be interpreted in a negative manner might end a promising but not yet fully realized effort such as the Monmouth County Bullying Prevention Program.

## Conclusion

The outcome evaluation of the YAP Challenge program has been completed. The findings suggest that although the program does not significantly reduce recidivism among juvenile offenders, the program model indicates that collaborative relationships between state government and private, community-based agencies, particularly between parole and outside aftercare service providers, are worth considering. The formative evaluation efforts of Operation Hook-Up are still in progress.

The evaluability assessment described here was carried out in a reasonably quick fashion for relatively little money. Even so, it produced useful answers to the questions about readiness for evaluation, whether it would be worthwhile to commit resources to a formal outcome evaluation, and with respect to the likelihood that decision makers would use the forthcoming results. Of particular interest here is the possibility that, in two cases, it could be anticipated that there might be particular political interest in the results. In one because it is a faith-based initiative, and in the other because of the spotlight on school violence in general and bullying in particular. In both cases, the evaluability assessment helped sensitize us to the possibility that any evaluation might become politicized and/or otherwise manipulated. This result alone was sufficient to justify the effort.

From the evaluation of state and local crime prevention programs that was mandated by the U.S. Congress, Sherman et al. (1997, pp. 10-20) concluded that "not every grant [their study focused on grant-funded programs] requires an evaluation." Their argument was that

Absent the resources and the skill needed for achieving the statutory definition of an evaluation as an impact assessment, the requirement that all crime programs be evaluated has resulted in few being evaluated. Spending adequate funds for strong evaluations in a few sites is far more cost-effective than spending little amounts of money for weak evaluations in thousands of sites. (Sherman et al., 1997, chap. 10, p. 20)

They thus recommended that evaluation funds should be conserved for impact evaluations that meet a reasonable degree of scientific certainty. We have taken this advice to heart. Our own evaluability assessment efforts pointed out that even when there is a desire and willingness to conduct an evaluation, considerations of the scientific value of evaluation efforts and the feasibility of a rigorous evaluation design, the possibility of political manipulation of the results of the evaluation, and cost-effectiveness should play a paramount role in deciding to undertake a full program evaluation.

## REFERENCES

- Keller, S. J. (1995, November 27). Streamlining juvenile justice. *The New York Times*, p. 1.
- Matthews, B., Hubbard, D. J., & Latessa, E. (2001). Making the next step: Using evaluability assessment to improve correctional programming. *The Prison Journal*, 81(4), 454-472.
- National Institute of Justice. (2002). *NIJ's evaluation strategy: What is the current evaluation state of practice at NIJ?* Unpublished manuscript.
- New Jersey Juvenile Justice Commission. (2002). *Introduction*. Retrieved April 28, 2004, from [www.state.nj.us/lps/jjc/info\\_intro.htm](http://www.state.nj.us/lps/jjc/info_intro.htm)
- Olweus, D. (1978). *Aggression in the schools: Bullies and whipping boys*. Washington, DC: Hemisphere Publishing.
- Olweus, D. (1993). *Bullying in school: What we know and what we can do*. Oxford, UK: Blackwell.
- Olweus, D., Limber, S., & Mihalic, S. (1998). *Book nine: Bullying prevention program*. Boulder: University of Colorado at Boulder.
- Poulin, M. E., Harris, P. W., & Jones, P. R. (2000). The significance of definitions of success in program evaluation. *Evaluation Review*, 24(5), 516-536.
- Prevention First. (2002). *Title V continuation application: Bullying Prevention Project*. Ocean, NJ: Author.
- Rossi, P. H., & Freeman, H. E. (1985). *Evaluation: A systematic approach* (3rd ed.). Beverly Hills, CA: Sage.
- Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., & Bushway, S. D. (1997). *Preventing crime: What works, what doesn't, what is promising*. Washington, DC: U.S. Department of Justice.
- Sullivan, J. (2002, September 15). Worth noting: The next capital offense may be hair in the inkwell. *The New York Times*, p. 3.
- Thurston, W. E., Graham, J., & Hatfield, J. (2003). Evaluability assessment: A catalyst program change and improvement. *Evaluation and the Health Professions*, 26(2), 206-221.
- Thurston, W. E., & Potvin, L. (2003). Evaluability assessment: A tool for incorporating evaluation in social change programs. *Evaluation*, 9(4), 453-469.
- Twain, D. (1983). *Creating change in social settings: planned program development*. New York: Praeger.
- Verhovek, S. H. (2001, March 11). Can bullying be outlawed? *The New York Times*, p. WK2.

Wholey, J. S. (1994). Assessing the feasibility and likely usefulness of evaluation. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 15-39). San Francisco, CA: Jossey-Bass.

***James O. Finckenaauer, Ph.D.**, is a professor in the School of Criminal Justice at Rutgers University and the former director of the International Center of the National Institute of Justice. His research interests include juvenile justice, organized crime, crime and justice in the former Soviet Union, and the politics of criminal justice.*

***Satenik Margaryan** is a Ph.D. student in the School of Criminal Justice at Rutgers University. Her main research interests include comparative penal reform and juvenile justice. She has a MPA from John Jay College of Criminal Justice, City University of New York, and MA in criminal justice from Rutgers University.*

***Mercer L. Sullivan, Ph.D.**, an anthropologist, is currently an associate professor in the School of Criminal Justice at Rutgers University. He has written extensively about delinquency and youth crime, the male role in teenage pregnancy and parenting, community development efforts in inner-city neighborhoods, and adolescent violence.*