CHAPTER 10

PROGRAM EVALUATION*

*This chapter was written by Betsy A. Fulton, a Research Associate for the American Probation and Parole Association in Lexington, Kentucky.

INTRODUCTION

Evaluation evokes an array of feelings from fear to excitement, from resistance to support. Evaluation can be perceived as a threat to the status quo or an opportunity for change and growth. It can be very complex or amazingly simple. This chapter is designed to assuage the anxiety associated with evaluation by introducing its benefits and key components. It will provide readers with a basis for further exploration of methods for monitoring and evaluating teen court programs. Within this chapter, "program evaluation" will be discussed as a one-shot process implemented to carefully examine program goals, activities, and outcomes. "Performance-based measurement" will be introduced as a method for the ongoing monitoring and evaluation of program goals, activities, and outcomes. While many of the same issues apply to both, variations in their design and implementation have different implications for teen court programs. Ideally, agencies would combine both methods to carefully examine program effectiveness.

Evaluation can be perceived as a threat to the status quo or an opportunity for change and growth.

Upon conclusion of this chapter, readers will be able to

◆ discuss existing literature on teen court evaluations;

◆ understand the importance of evaluating their teen court program;

◆ use key questions to guide evaluation planning;

◆ discuss the problems with using recidivism as the primary measure of program effectiveness;

◆ discuss the advantages of using alternative outcome measures to assess program effectiveness;

◆ identify performance-based measures for the ongoing monitoring and evaluation of their teen court program; and

◆ initiate the development of a management information system that meets the program's informational needs.

WHAT IS KNOWN ABOUT TEEN COURT EFFECTIVENESS?

A review of literature and program materials suggests that there is a substantial amount of anecdotal information on the effectiveness of teen courts, but very little research-based information. As Haaga and Reuter (1995) point out, this lack of a research base and evidence of success is typical for programs aimed at primary and secondary prevention. Reasons for scanty evaluations seem to be the difficulty associated with attributing reductions in antisocial behavior to a single prevention program, and the intangible nature of the targets of prevention, such as changing the social environment or attitudes of potential delinquents (Haaga and Reuter, 1995).

The limited research base is not for a lack of trying — 55.9 percent of the respondents to the American Probation and Parole Association's (APPA, 1994) teen court survey reported conducting program evaluations. Typically, these evaluations focus on counting clients and events rather than attempting to ascribe offender change and success to the program or a specific component. This chapter is designed to take agencies beyond counting activities (which, by the way, are important to establishing program integrity) to measuring results as a means of determining program effectiveness.

Four teen court evaluations that attempt to ascertain program effectiveness in terms of offender and volunteer changes and outcomes were available for review. They include evaluations from

◆ Teen Court Arlington (Texas);

♦ Kentucky's Teen Court;

♦ Cumberland County Teen Court Program (North Carolina); and

♦ Southside Youth Council Teen Court (Indiana).

These research reports are summarized in Appendix G. While limited in rigor and scope, these evaluations revealed the following important findings:

♦ In the Texas program, comparison of the rates of success, as measured by recidivism, between teen court participants and a matched group of nonteen court participants indicated a statistically significant relationship between teen court and success (75 percent of the teen court participants were successful versus 64 percent of the nonteen court participants) (Hissong, 1991).

♦ In the North Carolina program, comparisons of overall recidivism rates indicated that a matched sample of preprogram juveniles had lower rates of recidivism than a sample of teen court participants. However, when the analysis controlled for age and offense type, no significant differences existed in recidivism rates (North Carolina Administrative Office of the Courts, 1995).

♦ The Texas and Indiana programs reported higher success rates for older juveniles (i.e., 16-year-olds in Arlington and 14- to 18-year-olds in Indiana. The Indiana evaluators imply that the moral development of younger juveniles may not be advanced enough to learn from teen court participation (Hissong, 1991; McCullough et al., 1995).

♦ In contrast, the North Carolina study found that older juveniles had higher rates of recidivism. The study also found that the type of offense seemed to be related to recidivism rates, with juveniles committing nonproperty and nonviolent offenses (e.g., driving, weapon, minor controlled substance, alcohol, public disturbance, pyrotechnic)

having higher rates of recidivism. The evaluators noted that older juveniles tended to commit these types of offenses and implied that the higher rates of recidivism for older juveniles may be more a function of offense type than age (North Carolina Administrative Office of the Courts, 1995).

♦ The evaluation on the Kentucky program found that teen court participants showed increases in favorable attitudes toward authority figures (Kentucky Administrative Office of the Courts, 1994-95).

♦ The Texas evaluation results indicated that program effectiveness waned after one year, as evidenced by increased rates of recidivism (Hissong, 1991).

Supplementary information from the APPA (1994) teen court survey suggests high success rates for teen court participants. Respondents reported a range of recidivism rates from 0 to 75 percent, with an average of 5 percent recidivism across programs. This limited research base suggests that teen court has the potential to deter further criminal activity, particularly among older juveniles, and that teen courts may want to examine their selection criteria in terms of the types of offenses committed. But no conclusive evidence exists regarding teen court effectiveness.

WHY EVALUATE?

Besides being the "right thing to do," there are three very practical reasons to evaluate teen courts.

1. Evaluation improves the capacity of teen courts to successfully compete for limited public funds and support. As the demand for public accountability increases, demonstrated results are more often becoming the basis for allocating financial and human resources. Haaga and Reuter (1995) suggest that a lack of evidence regarding program effectiveness may contribute to the limited allocation of federal

and state resources to programs of a preventive nature such as teen courts. Teen courts must arm themselves with information and the capacity to demonstrate their value. In these financially stringent times, the foremost consideration of politicians and policymakers is "How much is it going to cost?" Second to that is "Is it worth it?" Teen courts find themselves competing with detention and probation for the limited juvenile justice purse, and juvenile justice as a whole is competing with every other publicly funded program from education to social services. In this competitive environment, teen courts must be able to "sell" themselves as effective and efficient programs.

In addition to competing for financial resources, teen courts also must compete for human resources in the form of volunteers. People are more likely to volunteer for a program that can demonstrate positive results or at least demonstrate a commitment to achieving program goals through the ongoing monitoring and evaluation of program activities.

Evaluation demonstrates a commitment to improved practices.

---

2. Evaluation promotes agency and community accountability. As stated in Chapter 4, the balanced approach that underlies many teen court programs calls for a shared responsibility between the juvenile justice system and the community for the control and reintegration of offenders (Bazemore and Umbreit, 1994). Evaluation demonstrates a commitment to improved practices. It highlights positive outcomes, uncovers ineffective practices, and guides agencies to explore alternative methods for achieving organizational goals. Key information about program struggles revealed through evaluation may elicit support and assistance for those improvements. In this manner, the teen court becomes accountable to the community for effective and efficient practices, and the community becomes

accountable to teen court and youth for support and assistance.

3. Evaluation creates a learning environment and contributes to organizational growth. Evaluation can be very threatening to program developers and personnel: What if the results are unfavorable? This fear creates resistance to evaluation. If, however, an organization adopts the view that "feedback," rather than "proof," is the objective of program evaluation, disappointing results (a) become an opportunity to examine alternative methods and strategies, and (b) open the door to new challenges and continued organizational development. Evaluation can assist program administrators in identifying staff and volunteer training needs and in maximizing resources. It provides a vision — a logical, well-planned pathway through the change process — and it leads to a healthy, vital organization.

KEY QUESTIONS TO GUIDE PROGRAM EVALUATION

The evaluation process will be simplified if considered upfront, during the program planning stage. Thinking about the type of information to be maintained and how the program's effectiveness will be determined will simplify data collection procedures and help ensure program integrity. Several key questions require consideration when developing processes and procedures for program evaluation.

What Are the Purposes of Program Evaluation?

While seemingly obvious, the general purpose of the evaluation should be discussed and clearly stated. Typically, program evaluation is conducted to determine

◆ the extent to which goals and objectives have been met;

◆ the extent to which the program was implemented as designed;

♦ the program's impact on youth, the agency, and the community; and

♦ how the program can be improved to achieve the desired goals.

Who Should Conduct the Program Evaluation?

To minimize the risk of biasing the program evaluation, an objective outside evaluator is recommended; this is not essential, however, and should not deter agencies from conducting their own program evaluation. Local universities are rich resources for evaluation expertise and interest. Professors and graduate students in criminal and juvenile justice programs may welcome the opportunity to design and conduct a program evaluation on teen court for little or no cost. Other potential sources of evaluators include someone with research expertise who works within the same governing agency as teen court (e.g., city, county, or state government) and volunteers with research expertise. If an outside evaluator is used, agency personnel should view themselves as customers with certain needs and expectations. They must be able to specify what information they hope to gain through program evaluation, what resources are available for evaluation, and potential barriers to evaluation. The evaluator's role is to design the evaluation to ensure the integrity of the information within the identified agency constraints.

Who Will Be Affected by the Evaluation?

Teen courts involve and affect many individuals and agencies. An analysis of who may be affected by the evaluation will allow evaluators to head off barriers to evaluation. All those affected (e.g., juvenile court personnel, representatives from other community agencies, volunteers, youth, and advisory board members) should be involved in designing the program evaluation. In this way, their concerns can be addressed, and they will know what to expect

from the evaluation process. Involving key stakeholders in the design also will gain their support for required data collection.

Considering the extent to which community members are involved in teen court development and operations, evaluations should be designed with extensive input from the community. Community members' definitions of success may include measures beyond the teen courts' interests. By measuring what is important to citizens, teen courts demonstrate commitment to the community and sustain community interest and involvement.

It is recommended that an evaluation task force be established to plan and implement the evaluation in conjunction with the evaluator. The remaining questions can be used to guide the discussion and activities of the task force.

What Are the Specific Research Questions to Be Answered through Program Evaluation?

The evaluation task force should carefully examine what they hope to learn from program evaluation and develop clear and concise research questions. The questions influence the kind of information to be gathered, the means of gathering the information, and the options for analysis. For example, an evaluation of teen court programs may seek to answer the following research questions related to program process:

♦ What are the characteristics of youth referred to teen court?

♦ To what degree did youth volunteers fulfill their responsibilities?

♦ What is the rate of referral and participation in the alcohol education program?

The purpose of the above types of research questions is to determine if the program is being implemented as designed. Research questions

related to program outcomes may include the following:

♦ What percentage of restitution ordered during fiscal year (FY) 1996 was collected?

♦ To what extent did teen court educate youth on the legal system?

♦ What percentage of youth completed their sentence as ordered?

♦ What impact did the teen court program have on the juvenile court docket and probation caseloads?

♦ What impact did participation in teen court (i.e., as defendants and volunteers) have on youth attitudes toward alcohol abuse and crime?

♦ What impact did teen court have on alcohol abuse among youth in the local school and community?

Both process and outcome measures are necessary to adequately evaluate a program. Agencies looking to prioritize their evaluation resources should first focus on process issues to obtain fundamental feedback on whether the program is being implemented according to specifications. Only then will outcomes be meaningful. More will be said about process and outcome measures, and the importance of both, later in this chapter. It is recommended that agencies limit the scope of evaluation to four to six well-defined research questions. Broad scope evaluations are difficult and very costly.

What Evaluation Design Will Be Used?

Time, cost, and expertise must be considered when selecting a research design. Agencies need to balance the desire for information with feasibility. Table 10-1 provides an overview of the strengths and weaknesses of the primary design options available.

What Data Need to Be Collected?

The agency must identify what data need to be collected to answer the research questions. Generally, offender data are needed in five key areas. These are

♦ identifying information (e.g., name, case number, age, sociodemographic information);

♦ intake/assessment information (e.g., criminal history, substance abuse background and needs, family history, and status);

♦ case objectives (e.g., specific sentencing objectives and time frames);

♦ case activities (e.g., teen court participation, sentence imposed, services provided, treatment participation); and

♦ case outcomes (e.g., changes in alcohol use, knowledge and skills, or family relationships; violations or new arrests; successful completion of sentence).

This list serves only as an example. Each agency must determine carefully what information is needed to plan and improve their teen court program. An important guideline is: only collect data that will be analyzed, reported, and used to modify and improve program operations.

What Data Sources Are Available?

Having accessible data is probably the most important reason for formulating an evaluation strategy during the program development stage. Data collection can be streamlined and simplified if forms and methods of program documentation are devised upfront. If, however, this is not possible, agencies can determine (1) what data are available and where; and (2) what data need to be generated and how. Examples of data sources include

♦ referral forms;

Table 10-1: Trade-Offs in Evaluation Designs

Trade-Offs in Evaluation Designs				
Evaluation Method	Steps	Strengths	Weaknesses	Best Used When
Process Evaluations:				
Planned program design vs. actual implementation.	Measure over or under targets.	Least costly.	Measures processes, not results. Targets may have no relationship to program effectiveness.	You need to know if the processes are being delivered according to design.
Outcome Evaluations:				
1. Before/after comparison.	Measure criteria (e.g., youth attitudes and behavior) before and after program.	Low cost/low expertise needed.	Low credibility; difficult to link inputs (e.g., teen court proceedings, sentence imposed) to outcomes (e.g., changes in youth attitudes and behaviors) because of internal and external influences.	Time and money are limited; criteria stable over time.
2. Time/trend projections of pre/post program.	Measure criteria over several intervals and project future trends.	Moderate costs and expertise.	Extreme variations may falsely imply a trend.	Historical data available; trend apparent.
3. Cohort comparisons (or quasi-experimental design).	Measure changes in similar groups with one group assigned to teen court and the other group assigned to standard procedures.	Low-moderate cost/time if data available; otherwise moderate-high cost/time.	Difficulty in finding matched groups raises validity issues.	Comparison group is similar to program group; Randomized evaluations are impossible.
4. Randomized.	Youth are randomly assigned to teen court or other program; identical groups are then compared.	Robust and systematic method.	High cost/time; very difficult; ethical considerations.	Where individuals will receive some type of program services; where determining program effectiveness is critical.

Source: Hatry, Winnie, and Fisk, 1981

- intake/assessment information;

- standardized pre/post instruments (e.g., to measure knowledge about the legal system or substance abuse before and after program participation, to measure attitudes toward authority or substance abuse);

- surveys (e.g., to gain information about the teen court experience from defendants, volunteers, community members);

- teen court dockets (e.g., number of defendants, sentences imposed);

♦ case documentation (e.g., activities, including payment of restitution, performance of community service, and participation in services; and outcomes, including violations, new arrests, type of termination); and

♦ juvenile court and police records.

Many of these data sources may be maintained by the teen court itself or by collaborating agencies. Others may need to be created for the purpose of collecting missing data needed to assess program effectiveness. Please see Appendix G for sample data sources.

Who Is Responsible for Collection, Coding, Management, and Analysis?

The type of data to be collected, the availability of a management information system, staffing configurations, and the current program design will influence the division of these responsibilities. In clearly defining these responsibilities, it is important to involve representatives from all levels of the organization and volunteers. Agencies can avoid duplication of efforts by using simplified data collection instruments and coding systems (e.g., 1 = restitution; 2 = community service; 3 = letter of apology). Simple data collection forms are particularly important when youth volunteers are responsible for recording the data. Furthermore, it is important to develop concise policies and procedures for data collection and analysis and to update them as needs and responsibilities change. Broad-based involvement and streamlined procedures will increase buy-in and commitment to the data collection process.

What Procedures Will Be Followed for Reporting the Information?

It is one thing to obtain data; it is quite another to explain data in a way that is both technically correct and useful. Key stakeholders within the teen court program should be informed of agency and program outcomes on a regular basis. An honest, straightforward approach to reporting outcomes is essential. It is always best to control information from the inside, rather than to leave its interpretation to someone who knows little about the issues associated with teen court.

How data are reported should be determined by focusing on the audience for whom it is intended. Sharing both positive and negative outcomes will earn greater respect and credibility with all audiences. The amount and format of information should be carefully considered. The usefulness of long, comprehensive reports is most likely limited to agency personnel who are directly affected by their contents. Legislators, judges, and funding sources may prefer receiving only information that impacts their decision-making process. The information should be concise and, wherever possible, in the form of graphs, tables, or charts. Information presented in a usable, reader-friendly format is more likely to gain the desired attention and support.

Evaluation results, good or bad, must be shared with community members, particularly volunteers. They may have (1) insights into why the desired objectives were not achieved, and (2) ideas for program improvements. Seeing the positive results of their time and energy, community members will be motivated to stay involved.

How Will the Information Be Used?

Collecting data and measuring results is a critical first step toward understanding the issues associated with teen courts. But unless this information is acted on, the teen court program will remain stagnant. Improving programs and practices is a continuous, long-term process of testing, modifying, and retesting. Feedback from evaluation facilitates this process. Data can be used to drive program modifications, to identify staff training needs, to justify new programs and services, and to reallocate financial and human resources. In short,

comprehensive and accurate data provide a credible source of information and knowledge about the effectiveness of programs and practices.

PERFORMANCE-BASED MEASUREMENT

The value of program evaluations or studies cannot be overstated. However, they require a great deal of time and expense and only provide outcomes for a specified period of time. Performance-based measures provide agencies with organizational feedback that drives program improvements and are a continuous process for monitoring and evaluation, as opposed to a single-point-in-time assessment. Many of the same issues and questions discussed above can be used to guide the development of a performance-based measurement strategy.

Performance-based measures provide agencies with organizational feedback that drives program improvements and are a continuous process for monitoring and evaluation...

Performance-based measurements provide a systematic method for collecting and reporting data that reflect a teen court's values and make outcomes tangible. By focusing on alternative outcomes specifically linked to program components, interventions, and attitudinal and behavioral change, conclusions can be drawn about which aspects of the program lead to the ultimate goals of most juvenile justice programs: community protection and reduced recidivism. The following sections will discuss problems associated with recidivism as the primary

outcome measure for juvenile justice programs, explore the utility of alternative outcome measures to evaluate program effectiveness, and introduce the APPA's model for developing agency-specific performance-based measures.

Rethinking Recidivism

There is a high degree of consensus about the ultimate goal of any juvenile justice program. From the perspective of taxpayers, juvenile justice practitioners, academicians, legislators, and judges, public safety is the agreed-upon bottom line. But too often a program's impact on public safety is judged solely on *reduced recidivism*, and that bottom line becomes blurred because of the heavy burden this creates. According to APPA's (1994) teen court survey, recidivism is the primary focus of most teen court evaluations. Other outcomes of teen court, such as general deterrent effects or improved youth competencies, contribute to public safety and should be included in these judgments. Recidivism takes an agency from point A to point Z without much consideration of what occurs in between. It is difficult for an agency to take responsibility for, and be judged by, a single outcome. Furthermore, there are many problems associated with recidivism as the primary outcome measure.

The first problem with recidivism as the sole outcome measure is that numerous definitions are applied to the term "recidivism" (see Figure 10-1). Different definitions can produce radically different figures from the same data. Second, there is tremendous variance in the amount of time involved in recidivism studies. Coupled with the many operational definitions of recidivism, this time variance makes it nearly impossible to compare research results.

Figure 10-1: Definitions of Recidivism

What Is Recidivism?

◆ Any new arrest?

◆ New felony arrests only?

◆ Any new conviction?

◆ New felony conviction only?

◆ Any new commitment of 60 days or more?

◆ A new prison commitment only?

◆ New technical violations?

◆ A technical violation that results in incarceration?

◆ An arrest for the same crime?

◆ Any arrest in which the offender was fingerprinted?

◆ A new arrest that resulted in incarceration of the offender?

◆ Incarceration in a prison?

◆ Arrest for a misdemeanor offense?

◆ Violent felonies?

An extensive review of correctional evaluations revealed these diverse definitions of recidivism. Please see Boone, H.N. (1994, Winter). An examination of recidivism and other outcome measures: A review of the literature. *Perspectives*, 18(1), 12-18.

Third, as Petersilia (1993) points out, recidivism is a measure of postprogram behaviors over which the juvenile justice system has little control. She notes, "Schools do not follow up their graduates to see if they slip back into ignorance or fail to hold a job after leaving school" (p. 14). There are many outside the educational system who feel that one measure of schools' performance *should* be how many students become employed after graduation. They also fault the educational system for failing to produce results such as graduates who can read, write, and hold down jobs that pay more than minimum wage. Likewise,

categorically dismissing postprogram recidivism as one of several measures of outcome is tempting, but problematic. Customers of the juvenile justice system and other programs designed to impact delinquency (e.g., the public) may or may not agree that there are other more important performance measures, but discounting recidivism out of hand ignores the importance of their expectation — to be protected from delinquent youth.

The fourth problem is that recidivism rates are influenced by many internal and external factors (Waldo and Griswold, 1979; Maltz and McCleary, 1977). Increased or decreased activity by law enforcement agencies or a change in judicial philosophy could have an impact on recidivism rates. A "get tough on youth crime and drugs" campaign will increase the number of new arrests. A new judge may be lenient and take no action for noncompliance with teen court orders. Given either scenario, it is difficult to determine whether the change in recidivism rates was due to changes in the behavior of offenders or to changes in police and judicial actions.

Since other factors affect recidivism data, it is fallacious to conclude that nonrecidivism demonstrates rehabilitation or success, or that recidivism demonstrates failure (Waldo and Griswold, 1979). Recidivism must be examined within the context of changes in program practices and policy shifts within the jurisdiction. Internal and external threats to validity, such as those described above, could be reduced with a commitment to incorporating rigorous experimental designs into the evaluation of juvenile and criminal justice innovations.

The fifth and final problem related to recidivism as an outcome measure is that it is currently viewed as an all-or-nothing measure. The dichotomous measure does not allow for partial successes. If the time between positive alcohol tests or the number of days in school increases,

is this not a partial success? If the severity of the crimes committed by a youth is reduced (e.g., from driving under the influence to trespassing), is this not a partial success? Success or failure should not be determined by one single incident. To measure an ultimate goal (i.e., reducing recidivism) in the short-term, without looking at intermediate variables and outcomes, is somewhat unreasonable. The next section discusses the importance of examining alternative outcome measures that more accurately portray the nature and accomplishments of teen court programs.

Alternative Outcome Measures

John DiIulio (1992), in *Rethinking the Criminal Justice System: Toward a New Paradigm*, argues for juvenile and criminal justice institutions to expand the use of outcome measures beyond crime rates and recidivism. He argues, "Crime rates and recidivism are not the only, or necessarily the best, measure of what criminal justice institutions do" (p. 1). Indeed, there are numerous intermediate outcomes that more clearly gauge and illustrate the business of teen courts. A focus on recidivism would overlook the very activities that define the program. Teen courts provide treatment and services, educate youth on the legal system, and impose sentences designed to promote youth accountability to victims and the community. By measuring the *outcomes* of these specific activities, teen courts can better assess the effectiveness of various activities and program components. Furthermore, until efforts are made to disentangle these activities and components, teen courts will be unable to determine what it is that leads to behavioral change and ultimately to a reduction in recidivism.

"Crime rates and recidivism are not the only, or necessarily the best, measure of what criminal justice institutions do"

Using the problem of alcohol abuse as an example, the practicality of alternative outcome measures becomes clear. Several studies

suggest a correlation between alcohol abuse and delinquent behavior (National Institute of Justice, 1991; Hawkins et al., 1987; Dembo et al., 1990). A primary goal of teen court may be to reduce the level of alcohol abuse among the youth population in the school or community. How, then, can teen courts determine if this goal is being achieved? Recidivism rates alone provide very limited information about how effectively the problem of alcohol abuse is being addressed. Rather, the following types of research questions could guide agencies in this determination:

♦ What percentage of youth involved in teen court have an identified alcohol problem?

♦ What percentage of these offenders were referred to a substance-abuse awareness program? Outpatient treatment?

♦ What percentage of these offenders completed their treatment assignment?

♦ As measured by a pre/post assessment instrument, to what extent did youth attitudes toward alcohol use change?

♦ As measured by self-reports and collateral information obtained from family and schools, did the level of alcohol use change?

A number of additional research questions could be proposed. Clearly, this information would be beneficial to an agency attempting to reduce the level of alcohol abuse.

The importance of such questions becomes obvious when there is an established relationship between these factors and delinquency. Research has shown that certain factors correlate strongly with delinquency, including poor and inconsistent family practices and relationships, school failure, negative peer associations, and limited cognitive development (Hawkins et al., 1987; Schinke, Botvin, and Orlandi, 1991). Several studies have revealed a positive correlation between increased participation in drug and alcohol treatment and success on probation and parole (Anglin and Hser, 1990;

Leukefeld and Tims, 1988) and a relationship between improved in cognitive functioning and reduced recidivism rates (Ross, Fabiano and Diemer-Ewles, 1988).

As this research suggests, if these aspects of human development are changed, criminal and delinquent behavior is likely to change. Thus, by shifting the research to measure these intermediate outcomes (i.e., offender change), teen courts can begin to assess the effectiveness of a particular program or component, learn from successes, and fine-tune these programs. Therein lies the primary value of intermediate measures; they test (confirm/reject) assumptions about different elements of the theoretical models that underlie interventions.

It is through this type of exploration that recidivism and other traditional measures of success can ultimately be impacted. As organizations with a mission of public safety, teen courts, like other juvenile justice components, must be accountable for recidivism rates and play a more active role in developing and implementing policies and practices related to reduced recidivism. The measurement of intermediate outcomes simply facilitates this role and makes recidivism rates more meaningful within the context of teen court activities designed to meet diverse goals. The next section introduces a model of performance-based measurement to assist teen courts in identifying alternative outcomes that better reflect program goals, activities, and components.

Developing and Implementing Performance-Based Measurements

APPA's model for performance-based measurement (Boone and Fulton, 1995) provides a framework for developing agency-specific performance-based measures (Figure 10-2). The development of a comprehensive performance-based measurement strategy requires the examination (or development) of

◆ values inherent in the agency/program;

◆ an agency mission statement;

◆ goals of the agency/program;

◆ activities performed to accomplish the goals; and

◆ measures for determining how well the activities are being performed and what impact they are having.

Figure 10-2: A Model for Developing Performance-Based Measures
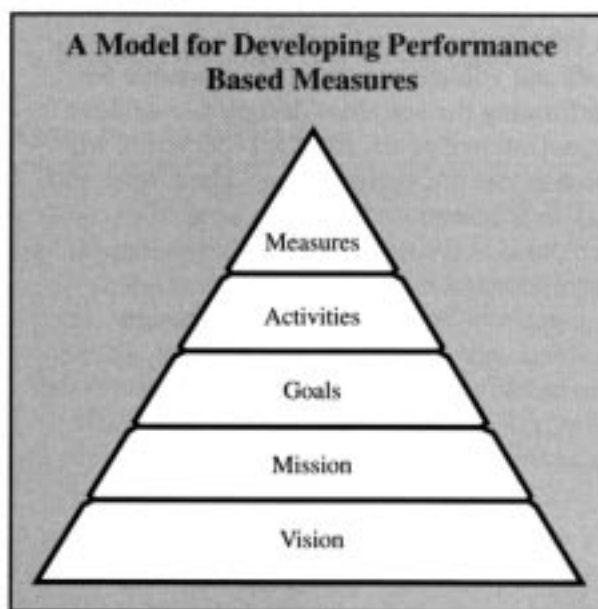
Source: (Boone & Fulton, 1995)

Alignment of these key organizational practices enhances an agency's chances for achieving desired results. The basic premise of the model is that performance-based measures should provide internal and external feedback at the policy, program, and staff levels about the relationships among values, goals and objectives, practices, and results. To the degree that performance measures are not integrated as part of standard business practices, the feedback will be less credible, less useful, and even can be contrary to an organization's objectives.

Involving Key Stakeholders

Staff and organizations will resist performance-based measures because such measures are threatening and represent change. Evaluation, in any form, can be discomforting. After the fact, favorable evaluations are warmly received; but few see negative feedback as an opportunity to learn. Involving a representative cross section of staff and community members in selecting process and outcome measures helps decrease normal fears and resistance to evaluation.

Of critical significance is the involvement of front-line personnel, volunteers, and supervisors in this developmental process. It is front-line staff and volunteers who are responsible for performing the activities designed to achieve organizational goals, and the supervisors who must assess this performance. Their input and buy-in is essential. Involving front-line personnel and volunteers can change their perceptions of this process from one that is threatening to one that offers opportunity. By inviting, *and valuing*, personnel input, agencies can identify process and outcome measures that truly reflect their values, mission, goals, and accomplishments (see Figure 10-3). A

performance-based measurement system that provides personnel with information and feedback on matters important to them will gain their commitment to the necessary practices of data collection and data compilation. Most importantly, it will gain their commitment to the results.

Identifying Agency-Specific Measures

As stated earlier, there are two types of performance-based measures (1) process measures (Was the program implemented as designed?), and (2) outcome measures (Did the program or practices achieve the desired results?). Both process and outcome measures are needed to assess program effectiveness. Examining processes helps to explain why such effects were produced and how processes can be modified to produce desired outcomes (Blalock, 1990). By controlling process, agencies can control outcomes. Processes can be examined through observation of program activities, interviews, and case audits (Harris, 1991). Outcome measures are needed to assess a program's immediate, intermediate, and ultimate impact. Rather than measuring how many youth attended a five-week teen court training session,

Figure 10-3: Critical Issues for Guiding Stakeholder Input

Guiding Stakeholder Input
Critical Questions

♦ What are the primary goals of the agency/program?

♦ What activities are conducted to achieve these goals?

♦ What is the theoretical basis for program activities?

♦ What is currently measured to demonstrate the success or failure of these activities or the overall agency/program?

♦ What perceptions are there about current measurements of program success or failure?

♦ What measurements would be a fair test of agency/program operations? Staff performance?

♦ What resources are needed to implement these measures?

outcome measures would assess changes in the youths' level of knowledge of the legal system or improvements in conflict-resolution skills.

Outcome measures are needed to assess a program's immediate, intermediate, and ultimate impact.

For example, suppose a teen court operates within a juvenile justice agency with the following value statement: "We believe in being sensitive to the needs of crime victims." In fact, the teen court was established as a primary mechanism for holding first-time youthful offenders accountable to victims. The teen court's purpose is: "The teen court program will provide education, services, and sanctions to youthful first-time offenders to promote accountability to victims and the community." Program goals include (1) increasing youth awareness regarding the effects of their behavior on victims and the community, and (2) restoring and compensating victims and the community for damage caused by the crimes of youthful offenders.

Activities supporting the first goal of increasing youth awareness may include participation in a five-week session on law-related education and attendance at a victim impact panel. Law-related education is designed to teach youth how their behavior impacts individuals and systems within the community and to foster social responsibility. Victim impact panels are designed to personalize the offense and to make offenders understand the pain and suffering their delinquent behavior causes individuals.

Process measures are needed to describe the extent to which these services are actually being delivered and to search for explanations of success, failure, and change (Harris, 1991). Process measures for goal #1 may include the following:

♦ **Rates of attendance at victim impact panels.** Outcomes of the victim impact panels are only meaningful once it has been determined that youth are, in fact, attending the panels. Rates of attendance can be collected through simple recordkeeping. Possible reasons for low attendance rates (See Figure 10-4) should be examined. They could be due to individual problems experienced by youth, such as transportation problems or basic refusal to attend, or due to a lack of communication between the teen court volunteers and defendants. Problems could be resolved by facilitating transportation for youth, by providing them with written notice of the panel a week ahead of time, or by imposing sanctions for noncompliance. Once attendance rates improve, outcomes also should improve.

Figure 10-4: Attendance Rates

Attendance Rates

Objective: 95% of all youth ordered to attend a victim impact panel during 1996 will attend as ordered.

Data elements: Number of youth ordered to attend during 1996, number of youth who attended as ordered.

Formula: (Number of youth who attended as ordered ÷ number of youth ordered to attend) x 100.

Example: 112 youth were ordered to attend a victim impact panel during 1996. 82 youth attended as ordered.

(82 ÷ 112) x 100 = 73%. **Objective was not achieved.**

♦ **The extent to which the law-related education curriculum is being delivered as designed.** Flaws in delivery may explain limited gains in youth knowledge or skills. Curriculum delivery can be assessed through observation and a standardized rating form. Problems with the delivery of the curriculum may stem from uncooperative youth, time

limitations, or lack of skills on behalf of the instructor. Disciplinary actions can be developed to address the issue of uncooperative youth, schedules can be adjusted to allow for sufficient time for delivery, and training can be provided to enhance the skills of the instructor. These problems can be overcome if program personnel are aware that the problems exist; process measures provide a mechanism for uncovering the problems. In this case (Figure 10-5), the objective for curriculum delivery was achieved; theoretically, outcomes should be favorable. If not, other potential problems, such as the learning level of youth attending the program or the theory and philosophy underlying the curriculum, should be explored.

Figure 10-5: Curriculum Delivery Ratings

Curriculum Delivery Ratings

Objective: During 1996, 90% of the law-related education instructors observed will receive an above average rating for curriculum delivery.

Data elements: Total number of law-related education instructors observed, number of instructors receiving an above average rating.

Formula: (Number of instructors receiving an above average rating ÷ total number of instructors observed) x 100.

Example: 8 instructors were observed and rated. All 8 received above average ratings.

$(8 \div 8) \times 100 = 100\%$. **Objective was achieved.**

Outcome measures for goal #1 may include the following:

♦ **Knowledge and skills gains from law-related education.** A key objective of law-related education is to teach youth about individual rights and responsibilities. A pre/post assessment instrument could be used to assess changes in knowledge (Figure 10-6). Basic skills such as interpersonal communication and conflict resolution often are included in law-related education curricula. Changes in skill levels could be assessed through observation of role-playing or interactions among students. A failure to achieve stated objectives should lead to a modification of instructional techniques or lesson plans or to a reexamination of the target population participating in the program.

Figure 10-6: Extent of Knowledge Gain

Extent of Knowledge Gain

Objective: During FY 1996, 80% of student test scores for law-related education will increase by at least 10 points from the pretest to the posttest.

Data elements: Total number of students, number of students who increased their test scores by at least 10 points.

Formula: (Number of students who increased their test scores by at least 10 points ÷ total number of students) x 100.

Example: During FY 1996, 128 students participated in law-related education. 106 students increased their test scores by at least 10 points.

$(106 \div 128) \times 100 = 83\%$. **Objective was achieved.**

♦ **Rates of alcohol use.** Victim impact panels consisting of victims or survivors of drunk driving accidents are sometimes used with youth arrested for driving under the influence and other alcohol-related offenses. If victim impact panels are effective in increasing offender awareness and in changing attitudes, it follows that the negative behavior should decrease among offenders attending a panel

— and possibly among larger populations because of teen court's potential deterrent effects. Rates of alcohol use among program participants (defendants and volunteers) or school and community populations could be measured by self-reports, information obtained from collateral sources, new alcohol-related arrests, or school incident reports. If the rate of alcohol use does not decrease, reasons must be explored. Is it because youth are not attending the panels? Is it because the victims selected for participation on the panel are not effectively communicating their message? Is it because the youth have more serious alcohol problems requiring treatment? Or is it because the strategy itself is not effective? Having results such as those in Figure 10-7 will guide program personnel through an examination of the problem and lead to program improvements.

Figure 10-7: Rates of Alcohol Abuse

Rates of Alcohol Abuse

Objective: The rate of alcohol abuse among victim impact panel participants will not exceed 10% during a one-year followup period.

Data elements: Number of panel participants, number of youth participants using alcohol.

Formula: (Number of youth participants using alcohol ÷ total number of panel participants) x 100.

Example: During 1996, 82 youth participated in a victim impact panel. Self-reports and collateral information indicated that 21 youth participants used alcohol during the one-year followup period.

(21 ÷ 82) x 100 = 26%. **Objective was achieved.**

The same process could be used to identify performance-based measurements for goal #2, restoring and compensating victims and the community. Possible process measures for goal #2 include the percent of victim impact statements completed and the extent to which restorative sentences (i.e., restitution, community service) are imposed. Possible outcome measures for goal #2 include the proportion of restitution collected, the number of community service hours performed, or the extent of victim satisfaction with the teen court program.

As exemplified in Figures 10-4 through 10-7, both process and outcome measures should be stated as specific objectives to be achieved within a predetermined time frame. It is essential that the objectives be realistic given the teen court's resources and target population.

Linking Employee Evaluation to Performance-Based Measurement

The seven principles of results-oriented government (Figure 10-8) illustrate the importance of linking employee evaluation to performance-based measurement. Programs that do not pay careful attention to closely aligning employee evaluation with process and outcome requirements should not expect to see the program implemented as designed or the desired goals achieved. Using results-oriented measures such as those listed above, supervisors can establish specific performance standards for program personnel. For example, an employee evaluated as "exceeding expectations" should meet the following criteria or performance standards:

♦ Law-related education sessions conducted by employee were rated as above average or outstanding during quarterly observations (process).

♦ Learning among the employee's law-related education students increased by an average

of 10 points as measured by pre- and posttests (outcome).

Figure 10-8: Seven Principles of Results-Oriented Government

> ### Seven Principles of Results-Oriented Government
>
> 1. What gets measured get done.
>
> 2. If you don't measure results, you can't tell success from failure.
>
> 3. If you can't see success, you can't reward it.
>
> 4. If you can't reward success, you're probably rewarding failure.
>
> 5. If you can't see success, you can't learn from it.
>
> 6. If you can't recognize failure, you can't correct it.
>
> 7. If you can demonstrate results, you can win public support.

Source: Osborne and Gaebler, 1993, pp. 146-155

Since promoting responsibility and enhancing knowledge among volunteers is a key objective of teen court, similar performance standards could be used to evaluate volunteer performance. For example, to hold volunteers accountable for goal #2, restoring victims and the community, performance standards could include the following: (1) a victim impact statement will be completed in 80 percent of the cases in which there is an identified victim; and (2) 90 percent of the community service ordered will be completed within 90 days of the sentencing date.

Performance standards assist personnel and volunteers in staying focused on goals and results and facilitate discovering innovative methods for achieving results. If the rate of community service completion is low, perhaps the teen volunteers could help identify factors

contributing to the problem and brainstorm about ways to increase compliance with these orders.

Performance standards assist personnel and volunteers in staying focused on goals and results and facilitate discovering innovative methods for achieving results.

People are an organization's greatest resource. Given the proper learning environment and structured feedback on meaningful performance criteria, staff and volunteers will work to improve outcomes and achieve desired goals.

MANAGEMENT INFORMATION SYSTEMS

An automated management information system is essential to efficient data collection and analysis. A state-of-the-art management information system can reduce paperwork, maintain data in an organized fashion, and provide quick access to information. The ideal system will allow collaborating agencies to share and exchange information. However, a multiuser information system requires decisions about ownership of records and responsibilities for updating and maintaining records. Procedures also must be developed to ensure the confidentiality of youth records. The following recommendations may assist with developing a management information system that meets program needs:

♦ **Establish a committee to guide the implementation of an automated management information system.** By involving key staff and community members in the decision-making process, buy-in can be enhanced and resistance to change can be minimized. Furthermore, a steering committee can provide valuable information on agency needs and operating procedures.

♦ **Consult a computer systems expert to examine agency needs, assist with the**

preparation of a request for proposals, and review vendor bids. This will ensure agency needs are addressed in all hardware and software purchases. Local universities and volunteers are potential sources for this type of assistance.

♦ **Carefully evaluate a number of management information system hardware and software options.** Computer hardware and software prices vary greatly from vendor to vendor. It is important to "shop around." For example, in one agency a number of vendors were asked to provide bids on identical specifications. The difference between the lowest and highest bid was almost double. By examining several hardware and software options, the steering committee and administrators can select the best system for the agency at a fair price.

♦ **If finances and expertise allow it, develop a program- or agency-specific management information system.** Three options exist for the software that maintains, organizes, and retrieves program data. One option is "off-the-shelf," generic software packages available for juvenile justice agencies. The choices are limited, but such packages may meet agency needs. A second option is "public domain" software developed with support from federal, state, or local funding. Public domain software may be less expensive, but most likely there still will be charges for support, the cost of copying the program, and printing the necessary program documentation. The third, and best, option is to hire a computer programmer to develop a system that meets the specific informational needs of the agency. Again, sources for computer programmers may include universities, the teen court's governing agency, and volunteers. To allow for monitoring and tracking events throughout a youth's involvement in teen court, specific

programming recommendations include (1) developing a relational data base to avoid data redundancy and to retain appropriate histories of chronological events (i.e., the system should be able to identify and relate intake information on a defendant from one data screen with termination information on that defendant in another data screen); and (2) developing a data base that is offender-based rather than offense- or incident-based.

♦ **Evaluate management information system capabilities periodically.** Computer technology is changing rapidly. Additional hardware or software purchases could make the system more effective and efficient.

CONCLUSION

This chapter provided teen courts interested in examining the effectiveness of their programs with a framework for systematically monitoring and evaluating program activities and outcomes. Agencies interested in conducting a large-scale program evaluation can use the key questions listed to guide their interaction with selected evaluators to ensure that their informational needs are being addressed through a credible research design. By applying the model of performance-based measurement, agencies can overcome the complexities of evaluation and identify immediate, intermediate, and ultimate outcomes that more accurately reflect the values and practices of teen court.

Although evaluation is viewed as the final phase in program development, it should not be confused with the end. Evaluation results must guide program improvements and modifications; otherwise, why bother? If properly implemented, program evaluation and a system of performance-based measurements will keep teen courts at the vanguard of juvenile justice programming.

CHECKLIST FOR EVALUATION

Have teen court program organizers or staff —

❑   Clarified motivations for evaluation?

❑   Examined resources available for evaluation?

Human _____

Financial _____

Technological _____

❑   Developed mechanisms for obtaining input to evaluation utilizing a performance-based measurement process?

State research questions (should be clearly related to program goals and activities).	Identify process measures that will indicate the degree of program implementation.	Identify outcome measures that will indicate degree of goal achievement.	Determine what data are available to answer research questions or to support measures.	Determine what additional data collection processes must be developed.

❑   Clarified responsibilities for data collection and analysis?

❑   Determined how information will be reported and to whom?

❑   Developed an action plan for implementing results?