

Evaluation for Performance:

Toolkit for Title IV Safe and Drug-Free Schools Programs

2005 (2nd edition)

Available Online: http://www.michigan.gov/mdch/0,1607,7-132-2941_4871_4878-15022--,00.html

James M. O'Neill, Ph.D.

Associate Professor, Psychology Department
Madonna University, Livonia, MI;
O'Neill Consulting, Chelsea, MI

2003 (1st edition) co-authors:

Judith M. Pasquarella, M.P.A.

(former) Manager, Education Section
Office of Drug Control Policy
Michigan Department of Community Health

Henry J. Hastings, J.D., Ed.D.

(former) Acting Director, Michigan Institute for Safe Schools and Communities
College of Education, Michigan State University

This project was supported by the Office of Drug Control Policy (ODCP), Michigan Department of Community Health, with funds from the U.S. Department of Education (USDOE), Title IV, Safe and Drug-Free Schools and Communities Act. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the ODCP or USDOE. The toolkit may be freely reproduced; however, citation of the source is appreciated.

Suggested reference:

O'Neill, J.M., Pasquarella, J.M., and Hastings, H.J. (2005). Evaluation for performance: Toolkit for Title IV safe and drug-free schools programs (2nd edition). State of Michigan, Department of Community Health, Office of Drug Control Policy.

Chapter 1

Owner's Manual: Everything You Wanted to Know about Evaluation but Were Afraid to Ask

What does it mean to evaluate a program?

There is no widely agreed-upon definition of evaluation. A commonly used, but fairly intimidating definition is the “systematic application of scientific methods to assess the design, implementation, improvement or outcomes of a *program*.”² For our purposes, evaluation is the systematic collection of information to assess the “performance” of your program and its participants (i.e., the students). The performance of your program is influenced by many factors, including the quality of program planning, training, implementation and participation. Obviously, program performance also should influence student performance; if the performance of your program helps to alleviate or reduce student drug use and/or violence, you can conclude that your program was successful.

The focus on performance in evaluation is actually at the heart of the Principles of Effectiveness (PoE), from the *No Child Left Behind Act of 2001, Title IV, Part A: Safe and Drug-Free Schools and Communities Act*. The PoE articulate a minimum standard which, if met, will provide a greater likelihood that a prevention program is effective. In fact, four of the PoE focus directly on evaluation; the others concentrate on programs and parental participation.

Title IV, Part A Principles of Effectiveness Related to Evaluation

Needs Assessment:

- “assessment of objective data regarding the incidence of violence and illegal drug use in the...schools and communities to be served” [Section 4115(a)(1)(A)];
- “analysis of...the prevalence of risk factors...; protective factors, buffers, assets, or other variables in schools and communities” [Section 4115(a)(1)(D)]

Process and Outcome Evaluation:

- “refine, improve, and strengthen the program, and to refine the performance measures.” [Section 4115(a)(2)(B)];
- “assess (the program’s) progress toward reducing violence and illegal drug use in schools” [Section 4115(a)(2)(A)]

Is evaluation of prevention programs similar to evaluation of academic programs?

Evaluation of prevention programs and evaluation of academic programs are more similar than different. The following are similarities and differences:

Similarities: In both prevention and academic evaluation, you should:

- **Assess needs, process and outcome**, though the terms in education and prevention may vary slightly. For example, school improvement efforts based upon the North Central Association endorsement protocol includes development of a “student profile” (“needs assessment” in prevention), the selection and implementation of “strategies” and “professional development” efforts (process evaluation) and “assessments” of achievement using standardized and locally developed tests (outcome evaluation).
- **Conduct periodic evaluation over time.** The goals of both prevention and academic achievement are realized over time, and evidence of their accomplishment is better captured in multi-year assessments rather than one-time assessments. For example, a follow-up assessment of youth who completed the program in the previous year will help you determine if any changes were sustained over time.

- **Include assessment of “benchmarks.”** In education, benchmarks are incremental steps that can lead to mastery of a certain content area. In prevention, the progress made toward the long-term goal of reducing drug use and violence is monitored via “*risk factors*” and “*protective factors*,” which are also known as *intermediate outcomes*. Risk factors place the student at greater risk for problem behaviors, whereas protective factors help to buffer the effects of risk factors or directly reduce the potential for problem behaviors.

Benchmarks = Risk and Protective Factors	
Examples of Risk Factors²	Examples of Protective Factors^{2,3}
<ul style="list-style-type: none"> • Availability of drugs/firearms • Community norms favoring drug use/violence • Low neighborhood attachment • Family conflict • Favorable parent behaviors/involvement in drugs/violence • Lack of commitment to school • Alienation and rebelliousness • Friends who engage in problem behaviors 	<ul style="list-style-type: none"> • A resilient temperament • Positive social interaction • Intelligence • Positive relationships (with parents, teachers, peers) that promote close bonds • Beliefs and clear standards about good health and the dangers of drugs and violence.

Differences: Evaluation in prevention is different than in academics in at least four important ways, all of which involve assessment:

- **Assessment of something that did not happen.** In academics, student assessments are suppose to reveal gains in knowledge, skills and behaviors. However, when SDFS prevention programs are evaluated, the goal is to demonstrate that drug use, violent behavior, and related attitudes have been averted, but otherwise would have occurred if the prevention program had not been implemented. Consequently, prevention evaluation systems must be designed to monitor trends in attitudes and behaviors prior to their projected increases based upon needs assessment information.
- **Assessment of positive behavior versus harmful or illegal behavior.** Academic evaluation mostly involves measurement of knowledge and skills that are desirable and therefore something students are motivated to report or demonstrate. On the other hand, there are serious negative consequences (e.g., arrest) for violence and ATOD use, so students may underreport or otherwise hide such behaviors and related attitudes, producing less reliable results for your program. Strategies to minimize this problem are covered in Chapter 4.
- **Assessment of permanent versus fluctuating behavior.** In academics, an increase in student knowledge and skills is a gradual process and results in fairly permanent change (e.g., once you learn addition, it’s hard to forget). However, problem behaviors such as violence and ATOD use may emerge or diminish quickly, especially during certain age periods. Therefore, assessment of problem behaviors should include multiple years within the sensitive age period(s) for the targeted problem behavior(s).
- **Data collection.** For academic evaluation, the collection of data (e.g., tests and quizzes) is typically conducted by classroom teachers or other personnel in the normal course of the school day. Prevention programs, however, may involve non-school staff (e.g., the evaluator) or require extracurricular data collection procedures (e.g., administration of a survey). Therefore, planning and cooperation are especially important.

Why is evaluation so important?

The most important reason for evaluation is that it helps to monitor whether progress is being made in creating a safe and healthy learning environment, which ultimately promotes academic achievement. Prevention programs – and their evaluation – should be viewed as a cornerstone of

the entire learning process. Persistently dangerous schools (or other negative learning environments) diminish the learning environment and compromise academic performance.

Are there more reasons to evaluate your program? There are several, including the following, all of which lead directly to increased program performance:^{4, 5}

- **Improving your program:** Seeing what worked and what didn't can help you make changes.
- **Decision-making:** Evaluation helps you make informed decisions about the program's future direction.
- **Accountability:** Evaluation helps you demonstrate that the program objectives were met and funds were spent appropriately and efficiently. That means money is used for services to the maximum number of students.
- **Providing information to stakeholders** (students, teachers, administrators, parents, community, funders) to recognize and reinforce the progress made in creating safe and healthy learning environments.
- **Providing information to the prevention field:** Evaluation can provide new information about service delivery. This may be useful to staff and participants or others who wish to replicate your program in their communities.
- **Getting more money for your program:** Good evaluation helps you gain support for your program, which in turn provides a better chance for additional funding. Besides the Federal Government, funds may be available from State and local agencies and private organizations.
- **Demonstrating that prevention works in Michigan schools and communities:** It allows ODCP to conduct regional and state-wide evaluations in order to demonstrate whether school-based prevention works in Michigan.

Testimonial
Evaluation has been extremely useful in determining which components of our programs make the greatest impact. This helps especially when budgets are tight and we're forced to narrow our focus for our target audiences."
<i>Steve Sukta SDFS Coordinator, Jackson ISD</i>

Why is evaluation sometimes perceived negatively?

You are probably familiar with – or have experienced first-hand – the ways in which evaluation can be viewed as trivial, counterproductive and even threatening. Evaluation reports can be treated as irrelevant and used only to fill shelf space. Because program evaluation requires funding, time and technical skills, it can be perceived as diverting limited program resources from schools and students. Teachers and administrators may be concerned that evaluation activities will interfere with time spent on other initiatives. Program coordinators and directors may fear that their program will not withstand the scrutiny of evaluation; but beware: it's conceivable that the program is more harmful than helpful. In addition, the collection of data from multiple sources often requires cooperation with various community groups, each of whom may have different or even conflicting agendas for the evaluation.

Alternatively, evaluation is viewed as essential, productive and less threatening when you do the following:

- **Promote collaboration and consensus.** Collaboration includes key stakeholders - administrators, teachers, school staff, parents, and students, for example - who have respect for one another's roles and equal partnership in the evaluation process.⁶ Collaboration also allows you to participate in a more comprehensive evaluation involving efforts from multiple sectors (e.g., school, police, and community) and funding sources, which

You'll Pay Later. . .	
A critical factor in collaboration is stakeholder consensus on the needs and intended purposes and outcomes of the program. Without such agreement, the evaluation – and eventually the program - will be viewed as a waste of precious time and resources.	

minimizes duplication of efforts and allows your community to demonstrate better that prevention works.

- **Provide information about effective evaluation.** When stakeholders appreciate the benefits of good evaluation, they become good consumers and advocates of evaluation results. Good evaluation of prevention programs, like other programs, has evolved over the past 25 years from an emphasis on process evaluation *or* outcome evaluation to a more productive shift toward evaluating both processes *and* outcomes.
- **Focus on program refinement.** The stark reality is that prevention – in addition to treatment and interdiction – is needed to reduce the unacceptably high rates of violence and ATOD use in our society. And a number of research-based programs and strategies have been developed to address these problems. The focus of evaluation should be to monitor program *fidelity* and continuously refine these programs, ensuring that they match the needs of the *target population*. For example, less than ideal outcomes should lead to a reconsideration of the process (e.g., types of programs and their implementation) that helped produce such outcomes, as well as a review of the needs for and goals of the program.
- **Include multiple indicators of program efforts and outcomes.** Reliance on one indicator (e.g., the number of fights) to assess student outcomes places too much importance on that indicator and creates a myopic view of the value of the program.⁷ The use of multiple indicators (known as *triangulation*) will provide a more comprehensive picture of program value, with fewer “blind spots.” Also consider the inclusion of intermediate outcome indicators, such as risk and protective factors, to serve as benchmarks on the road toward *long-term outcomes*. Not all coordinators have the resources to include multiple indicators, but it is very important to consider all feasible indicators to assess program efforts and outcomes.
- **Recognize that evaluation is already part of your routine** – As professionals, all of us engage in evaluation activities as part of our efforts to improve the lives of our youth. We check on the progress of our lesson plans, for example, and monitor students’ progress through exercises and exams. Evaluating prevention programs is no different – the key is to make it part of your regular routine, so that the effort remains manageable and critical opportunities to collect data are not missed.

Example
Multiple indicators for fighting
<ul style="list-style-type: none">• Referral reports• Suspension reports• Student self-reports• Classroom observations• Playground observations

What if my evaluation results are negative?

Usually, the concern about negative results lies in whether an intended *outcome performance goal* was achieved, for example, whether students indeed exhibited less violent attitudes or fewer violent incidents. Although negative results can be demoralizing, they shouldn’t be minimized or ignored, but rather seen as an opportunity for coordinators and other stakeholders, in consultation with ODCP, to reflect upon the program and related factors that may have accounted for the unwanted results. This reflection process will not only help you, but your learning community of other coordinators and ODCP as well.

It is also important to realize that negative results can emerge even after years of successful programming and outcomes, which might be attributable to changes in the student population, and/or shifts in community attitudes toward ATOD and violence, among other reasons.

Negative results could be due to one or more of the following reasons, all of which should be considered in your efforts to improve program outcomes in future years:

“You used the wrong car.” Programs should be grounded in *scientifically based research* demonstrating their effectiveness. Because such programs have a track record of success, their use increases the likelihood that your students also will benefit. Of course, the utilization of a scientifically based program doesn’t guarantee positive results, especially if the program is not implemented with fidelity and/or is a poor fit with the needs of your targeted population.

“You didn’t use and/or follow a roadmap.” Negative results commonly originate from poor planning or poor execution of the program. Better planning and/or execution should be considered for all elements of essential programming, such as buy-in from staff and administration, staff training, implementation fidelity, and valid and reliable evaluation measures.

“You hit unexpected roadblocks.” Even if you used the right car and followed a good roadmap, there may be unforeseen roadblocks to successful outcomes. For example, the program may have been received poorly by the students, or students did not participate as planned. It’s also possible that an event in the district, such as a recent lapse in enforcement of rules against violence, may have weakened the traction of your prevention message. Because unanticipated roadblocks are possible, it’s critical to detect them early and take steps to overcome them.

Testimonial
We found negative results with our violence prevention program despite a solid needs assessment, program and implementation. Post-program follow-up revealed that the students felt that the program was for “losers.” Once we improved the image of the program, we found that more students participated and we achieved our goal of reduced violent behavior.
<i>John Belaski former SDFS Coordinator, Kent ISD</i>

“Your destination wasn’t realistic.” If you found negative results despite using the right car for your needs and following a good roadmap without incurring significant roadblocks, you should consider whether your original goals were on track. Was the timeline for change too short? Was the targeted behavior too resistant to change? Were the participants more at-risk than originally believed? Finding answers to these questions usually involves a more thorough needs assessment that includes not only a general review of objective *risk factors* and *protective factors* in the student population, but also reactions from trusted program participants and other students to determine their unmet needs and concerns. In addition, a more thorough needs assessment may reveal a need to refine or replace your current program or, more generally, your prevention approach (e.g., *universal* vs. *indicated* programs). Finally, negative findings might suggest weak points in prevention theory that need to be examined and shared with others.

What is periodic evaluation?

Periodic evaluation, a term used in the Title IV Principles of Effectiveness, is the best way to ensure that you are on the right road toward helping students.

The word “periodic” is essential because if you wait too long before evaluating progress in your prevention efforts, you might “pay later” by missing some critical opportunities to improve process or show positive changes in students.

How long is too long to wait before you evaluate? Based upon the experiences of several successful SDFS Coordinators, assessment of progress made toward goals should be conducted twice yearly and reported to ODCP as part of a mid-year report and year-end report. The purpose of each report is to “check your gauges,” which involves measuring and reporting progress made in your program processes and

Testimonial
"Our district has made great progress in evaluating our ODCP-funded projects, thanks largely to our efforts to identify problems early in the evaluation process and seek technical assistance from ODCP."
<i>Yolanda Williams SDFS Coordinator Flint Community Schools</i>

outcomes. The purpose of a mid-year report is to check your gauges regarding the process evaluation, such program planning, implementation and participation, as well as any obstacles encountered. In addition, the mid-year report provides you and ODCP the opportunity to identify needs for technical assistance in programming and/or evaluation.

The year-end report is a second effort to check your gauges regarding process evaluation, as well as an opportunity to evaluate progress in outcomes made during the program year, to ensure you are on the right road to reach your outcome performance goal(s). Examples of completed process reports and outcome reports are provided in Appendix D, along with blank SDFS mid-year and year-end report forms.

Note: Unless you are a one-year grantee, the year-end or final report does not necessarily have to include outcome results each year, especially if your outcome performance goal bridges across multiple years. However, at a minimum, outcome results should be reported *every two years* for each outcome performance goal of your program. Why two years, at most? Most programs should produce at least some changes in attitudes or behaviors within a two-year period, especially scientifically based programs. Wait any longer and you might miss some important changes in attitudes/behaviors or, worse yet, waste precious time traveling down the wrong road without realizing it.

Notes:

Chapter 2

Car Talk: Basic Evaluation Concepts

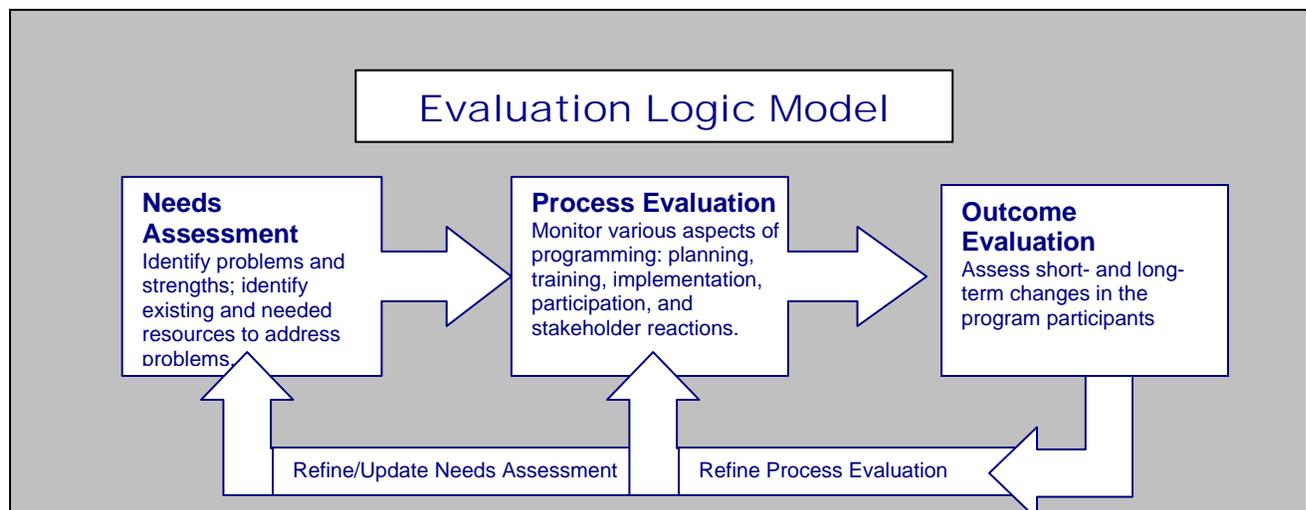
What is involved in evaluating a program?

As part of your approved ODCP grant proposal, you conducted a thorough *needs assessment* and wrote *outcome performance goal(s)*, which included the identified *program*, targeted population and expected level of attitudinal or behavioral change. The completion of these steps means that you've mapped out where you are, where you want to go, and how to get there – you're on the road to a successful evaluation!

Now, you're ready to conduct the other two types of evaluation: process evaluation and outcome evaluation. **Process evaluation** involves *monitoring various aspects of programming*: planning, training, implementation, participation, and stakeholder reactions. In car talk, process evaluation is similar to “monitoring the journey” of your programming to determine if you are following the map. programming is **Outcome evaluation** involves *assessment of participant changes* – intermediate and long-term – resulting from the program. In outcome evaluation, your focus is answering the question, “Are we there yet?” regarding participant outcomes (e.g., reduced violent behavior).

The following is a diagram of a method, a logic model, of how each type of evaluation – needs assessment, process evaluation and outcome evaluation – “flows” to create a comprehensive evaluation system. The logic model highlights three major points:

- **Evaluation begins with needs assessment.** The starting point for evaluation is a comprehensive needs assessment, which focuses the process and outcome evaluation. A poor needs assessment will lead to dead ends in programming and outcomes.
- **Outcomes are interpreted from processes.** The outcomes of a program are best understood in light of information collected from the process evaluation. For example, negative outcomes could be due to various “process” shortcomings, such as uneven staff training, poor implementation, low program fidelity, negative reactions from staff, or some combination of these.
- **Evaluation is continuous.** Information from the outcome evaluation is used to refine future needs assessment and process evaluation activities.



What's the difference between a process and outcome evaluation?

Process evaluation focuses on *program performance*, whereas outcome evaluation is concerned with *participant performance* resulting from the program. In other words, process evaluation is similar to assessing the “journey” of your program, from the planning and training to implementation and participation. In outcome evaluation, you ask the question, “Are we there yet?” – it involves assessing whether you’ve reached your intended “destination” or outcome performance goal(s) related to student attitudinal/behavioral change(s).

In process evaluation, information is collected about *programming*, such as the quality of staff training, the delivery of the program, and reactions to the program. In outcome evaluation, information is collected about *participant changes* resulting from the program, such as the direction of change, the amount of change, the type of change (e.g., expected versus unexpected), and the degree to which changes were related to the program.

Example Process vs. Outcome Evaluation

Process evaluation focuses on the *program performance*, whereas outcome evaluation is concerned with *participant performance* that resulted from the program.

The following are examples of **process indicators**:

- Number and type of staff trained
- Number and length of lessons taught
- Number and type of participants served
- Suggested changes to program

The following are examples of **outcome indicators**:

- Amount of alcohol consumed in past 30 days
- Number of fights in the past 30 days
- % of students who approve of ATOD
- % of students who approve of fighting

How do I conduct a process and outcome evaluation?

Both process and outcome evaluation can be completed in a four-step process (see Table on following page). Each step is defined by a series of questions designed to steer you through a complete “economy model” process and outcome evaluation that will get you to your destination. The questions were developed in consultation with state agencies, local administrators, staff and coordinators, and will serve as the framework for your reports to ODCP.

Detailed information about process and outcome evaluation is presented in Chapters 3 and 4, respectively, and the four-step process with an accompanying Table is presented again in these chapters. The following is a basic overview of each step.

The purpose of Step 1 is to “focus the evaluation” by identifying performance questions to guide the process and outcome evaluation. The questions provided in Step 1 are essential for understanding and communicating the processes and outcomes for each program. However, you might want to add questions if you feel that monitoring other features is necessary.

Steps 2 and 3 involve the methods, such as measures (e.g., surveys), used to answer the questions in Step 1. Measures are similar to gauges on a car that provide information about the status of various components, such as fuel, mileage, oil, and tire pressure. The focus of Step 2 is to selecting the best and most convenient gauges for your program processes and outcomes, whereas step 3 involves collecting, organizing and summarizing the information collected from those gauges.

Step 4 is used to make program adjustments and ensure program sustainability. This is an often overlooked but critical step, because your program’s top performance and future are not guaranteed. For example, some fine-tuning and planning may be required to accommodate changing student needs and/or secure additional resources to meet those needs.

Evaluation Step	Process Evaluation Questions	Outcome Evaluation Questions
1. Focus on Performance: Use Performance Questions	<ul style="list-style-type: none"> • Were facilitators adequately trained to conduct the program or provide the strategy/service? • Have all planned activities been implemented in all intended classrooms/schools? Were they accomplished on schedule? If not, what remains to be done? • Were there any obstacles/challenges? If so, what steps were taken to remedy the problem(s)/obstacle(s)? • What were the reactions of the students, staff and administrators to the program? • What changes occurred in leadership or personnel? What effect did these changes have? 	<p><u>For each approved outcome performance goal/measure:</u></p> <ul style="list-style-type: none"> • Were the outcome data collected on schedule? If not, provide the reason(s) and a plan to collect the data. • Were the outcomes in the expected direction? • Did the outcomes meet or exceed the performance goal/measure? • Were the outcomes different for various groups (e.g., males vs. females)? • Were there unintended positive or negative outcomes? • How clearly were the outcomes attributable to the program?
2. Choose the Best Gauges: Select Indicators, Measures and Sources	<ul style="list-style-type: none"> • What process <i>indicators</i> will be measured to answer the performance questions? • What <i>measures</i> will be used? • What information <i>source(s)</i> will be used? 	<ul style="list-style-type: none"> • What outcome indicator(s) will be measured to answer the performance questions? • What measures will be used? • Are the measures reliable and valid? • What information source(s) will be used?
3. Check the Gauges - What Do They Say: Collect, Organize and Summarize Information	<ul style="list-style-type: none"> • Who will collect the data? When? • Who will enter/organize the data? When? • In what format(s) (numbers, words, graphs) will the data be summarized? • What are the answers to the performance questions in Step 1? • How and when will the results be reported to stakeholders? 	
4. Enhance Performance: Make Program Adjustments and Increase Sustainability	<ul style="list-style-type: none"> • How will the information be used to enhance the program while preserving <i>fidelity</i>? Is the selected program indeed the right program given the needs and <i>target population</i>? • How will the information be used to increase sustainability? 	

Notes: