

Blueprints Selection Criteria

Model

Programs

The success of a community's violence prevention efforts will depend, in large degree, upon the preventive interventions used. That is why it is imperative to identify approaches that have been proven effective. Although a program model can rarely, if ever, be proven to be superior to all others, a particular model elicits greater confidence after its theoretical rationale, goals and objectives, and outcome evaluation data have been carefully reviewed. Although various scholarly reviews have identified exemplary programs, the methodological standards used in evaluating program effectiveness can vary. A few of these scholarly reviews have explicit standards, and one even scores each program evaluation on its methodological rigor, but for most the standards are variable and seldom made explicit. **The standard for the claims of program effectiveness in most of these reviews is very low.** Of those with explicit standards, Blueprints programs have the highest standards and meet the most rigorous tests of effectiveness in the field. There are several important criteria to consider when reviewing program effectiveness. **Three of these criteria are given greater weight: evidence of deterrent effect with a strong research design, sustained effect, and multiple site replication.** Blueprints model programs must meet all three of these criteria, while promising programs must meet only the first criterion.

Evidence of Deterrent Effect with a Strong Research Design

This is the most important of the selection criteria. Relatively few programs have demonstrated effectiveness in reducing the onset, prevalence, or individual offending rates of violent behavior. **The Blueprints Advisory Board accepts evidence of deterrent effects for three key indicators -- violence (including childhood aggression and conduct disorder), delinquency, and/or drug use -- as evidence of program effectiveness.** Providing sufficient quantitative data to document effectiveness in preventing or reducing the above behaviors **requires the use of evaluative designs that provide reasonable confidence in the findings (e.g., experimental designs with random assignment or quasi-experimental designs with matched control groups).** Most researchers recognize random assignment studies (randomized trials) executed with fidelity as providing the highest standard of program evaluation. Random assignments offer the most compelling evidence that study results are due to the intervention rather than to preexisting differences between experimental and control groups and/or other threats to internal validity, such as maturation, selection bias, and testing effects. In these studies, assignment to experimental or control conditions is determined solely by chance, and the likelihood of differences being attributed to the assignment process can be assessed.

When random assignment cannot be used, the Advisory Board considers studies that use control groups matched as closely as possible to experimental groups on relevant characteristics (e.g., gender, race, age, socioeconomic status, income) and studies with control groups that use statistical techniques to control for initial differences on key variables. As carefully as experimental and control groups are matched, however, it is impossible to determine if the groups may vary on some characteristics that have not been matched or controlled for and that are related to program outcome. Random assignment, therefore, is believed to be the most rigorous of methodological approaches.

Research designs vary greatly in quality, particularly with respect to several key aspects: sample size, attrition (loss of study participants over time), and measurement issues. At a minimum, the following issues need to be addressed: **(1) Sample sizes must be large enough to provide statistical power to detect effects.** It is more difficult to detect statistically significant differences between groups when small sample sizes are used. **(2) Attrition,** or loss of study participants, may be indicative of problems in program implementation or may be a failure to locate subjects during a follow-up period. Attrition is dangerous, particularly because it can compromise the integrity of the original randomization or matching process. It reduces confidence that the original sample and final sample are comparable and that the final experimental and control comparisons reflect only treatment effects. **(3) Tests to measure outcomes must be administered fairly, accurately and consistently to all study participants.** For example, the use of inconsistent measures over time may produce less reliable test scores. The instruments which are used to measure outcomes should be demonstrated to be reliable and valid.

School-based Evaluations. Evaluations of school-based programs, with schools as the unit of analysis, typically require multiple schools per condition to perform a main effects analysis with sufficient power to detect effects. Since meeting this criterion requires a complex evaluation which is very costly, it would eliminate most existing school-level evaluations from consideration in the Blueprints Series. Therefore, **school-based evaluations that use experimental or quasi-experimental designs with relatively few schools, but more than one in each condition, will be considered in the Blueprints Series if they meet an additional burden of proof. They must demonstrate consistency across effects and across replications with multiple measures from different sources. The theoretical rationale should be well developed, and there should be a rigorous evaluation of theory with evidence that the results are consistently in line with the expectations (i.e., there are changes in the risk and protective factors which mediate the changes in outcomes). Outcomes should be robust, with at least moderate effect sizes.** Evidence that the benefits of the program outweigh the costs are helpful. Our decision to accept this level of proof is driven totally by the state of current research, and it should not be assumed that this standard of proof is desirable. Evaluations with multiple schools is most desirable and should be encouraged among funders and researchers.

Sustained Effects

Although one criterion of program effectiveness is that it demonstrate success by the end of the treatment phase, it is also important to demonstrate that these program effects endure beyond treatment and from one developmental period to the next. Designation as a Blueprints program requires a **sustained effect at least one year beyond treatment, with no subsequent evidence that this effect is lost.** Unfortunately, many programs that demonstrate initial success fail to show long-term maintenance of the effects after the intervention has ended. Depending on whether effects are immediate or delayed, the full impact of an intervention or treatment may not be realized at the end of treatment. Significant improvement may be realized over time, or a decay or decline may result. For example, if a preschool program designed to offset the effects of poverty on school performance (e.g., Head

Start) demonstrates its effectiveness when children start school, it is also important to demonstrate that these effects are sustained over a longer period of time. Unless this protective effect is sustained through high school, it is unlikely to have an impact during this critical period when problem behavior is at its peak: the effect must be sustained if it is to help adolescents maintain a successful life course trajectory. Although programs that have specifically failed to produce a sustained effect do not qualify for the Blueprints model or promising categories, programs that have not yet demonstrated long-term effects (because sufficient time has not yet elapsed or follow-up analyses were never planned) may be considered as promising programs.

Multiple Site Replication

Replication is an important element in establishing program effectiveness and understanding what works best, in what situations, and with whom. Some programs are successful because of unique characteristics in the original site that may be difficult to duplicate in another site (e.g., having a charismatic leader or extensive community support and involvement). Replication establishes the strength of a program and its prevention effects and demonstrates that it can be successfully implemented in other sites.

Programs that have demonstrated success in **diverse settings (e.g., urban, suburban, and rural areas) and with diverse populations (e.g., different socioeconomic, racial, and cultural groups)** create greater confidence that such programs can be transferred to new settings. As communities prepare to tackle the problems of violence, delinquency, and substance abuse, knowledge that a specific program has had success in various settings with similar populations adds to its credibility.

Some projects may be initially implemented as a multisite single design (i.e., several sites are included in the evaluation design). When this occurs, the evaluation should check for overall main effects and sources of variation across sites. **Becoming a Blueprints model program requires at least one replication with demonstrated effects.** This criterion does not need to be met to qualify as a promising program.

Additional Factors

In the selection of Blueprints model programs, two additional factors are considered: whether a program conducted an analysis of mediating factors and whether a program is cost effective.

Analysis of Mediating Factors. **The Blueprints Advisory Board looks for evidence that change in the targeted risk or protective factor(s) mediates the change in violent behavior. This evidence clearly strengthens the claim that participation in the program is responsible for the change in violent behavior, and it contributes to our theoretical understanding of the causal processes involved.** In its reviews of different programs, the Advisory Board has discovered that many programs reporting significant deterrent "main effects" have not collected the data necessary to complete an analysis of mediating factors.

Costs versus Benefits. **Program costs should be reasonable and should be less or no greater than the program's expected benefits.** High price-tag programs are difficult to sustain when competition is high and funding resources low.

Implementing expensive programs that will, at best, have small effects on violence is counter-productive. Although outcome evaluation research established that Blueprints programs were effective in reducing violence, delinquency, and drug use, very few data were available initially regarding the costs associated with replicating these programs.

Two recent cost-benefit studies involving Blueprints programs -- the RAND Corporation Study and a study by the Washington State Institute for Public Policy -- suggest that these programs are cost-effective (Greenwood, Model, Rydell, & Chiesa, 1996; Washington State Institute for Public Policy, 1998, 2001).

The selection criteria identified above establish a high standard, one that has proved difficult for most programs to meet, thus explaining why there are only 11 Blueprints programs. This high standard reflects the level of confidence necessary, however, for recommending that communities replicate these programs with reasonable assurances that they will prevent violence. **The Blueprints model programs are not intended to be a comprehensive list of programs that work, but rather reflect a selection of programs with strong research designs for which we have found good evidence of their effectiveness in delinquency, violence, or substance abuse prevention and reduction.** There is no implication that programs not on this list are necessarily ineffective. Chances are that there are a number of good programs that have just not yet undergone the rigorous evaluations required to demonstrate effectiveness. But our evaluations have also revealed that many programs are ineffective, and a few are iatrogenic (i.e., harmful). Without evaluations, we just don't know. It is in the best interests of our children to evaluate, so we can have confidence that what we are doing for them actually helps. As time goes on and new research findings are published, CSPV hopes to add to this list other credible, effective programs which communities can use confidently. CSPV will also continue to follow evaluations of Blueprints programs to refine our knowledge of their effectiveness for specific populations and over longer periods of time.

Retrieved May 12, 2004 from:

<http://www.colorado.edu/cspv/blueprints/model/criteria.html>