



711 Westchester Avenue, White Plains, New York 10604
Tel: 914.421.2525 Fax: 914.683.6998

Conducting Valid Evaluations of LifeSkills Training and Other Prevention Programs

Many school districts and agencies conduct local evaluations of prevention programs such as LifeSkills Training (LST) in order to assess the effectiveness of the program with their students or clients. Unfortunately, resources for evaluation purposes are often limited, making it difficult for evaluators to attend to a number of important issues that will ensure high quality evaluations. The following text outlines some basic considerations that are essential in program evaluation. Failing to attend to any of these considerations can lead to misleading or inaccurate evaluation results.

Key Ingredients in Study Design

Control Group. The use of alcohol, tobacco, and other substances typically increases during the early years of adolescence, as do the related risk factors. This is true even among students who receive a prevention program. Therefore, the best way to examine prevention effects is to compare changes over time in those who receive a prevention program with those in a control group who did not receive a prevention program. Within a control group design, it is often observed that substance use and related risk factors *increase less* in the intervention group compared to the control group. In other words, the program was effective in preventing the usual amount of increase in a particular problem behavior.

Equivalent Groups at Pre-test. The intervention and control groups should be equivalent in size, demographics, and risk behaviors, as much as possible before the intervention starts. Demonstrating pre-test equivalence between the intervention and control groups helps to establish that any observed behavioral differences after the intervention are due to the prevention program and not to pre-test differences. If such pre-test differences do exist, they can be controlled for in the statistical analyses.

Random Assignment. The best designed evaluation studies are those that attempt to minimize pre-test differences by using random assignment to intervention or control conditions. It is best to randomize entire units (such as schools) to a condition in order to prevent students in the control group from being exposed to part of the prevention program (e.g., contamination across conditions). It is helpful to group or “block” these units prior to random assignment on variables that may be associated with substance use or other risky behavior. In school-based studies, for example, one might block schools in terms of average pre-test smoking levels; schools with the highest average smoking levels would be randomly assigned to a condition, then schools with moderate smoking scores would be assigned, etc. Studies that use this type of randomized block design are more likely to have equivalent intervention and control groups at the pre-test relative to studies that do not.

Sufficient Number of Participants. Evaluation designs with small numbers of participants often have inadequate statistical power to detect significant changes across groups or over time. Furthermore, with small samples, any one student’s answers are more likely to affect the overall findings; therefore the results are more likely to reflect idiosyncratic responding and cannot be readily generalized to other groups of students.

Standardizing Data Collection

Standardized Data Collection Procedures. Data collection procedures should be documented in a written protocol to ensure that data are collected in a standardized way. Data collection protocols should include an emphasis on the confidential nature of the data being collected in order to enhance the truthfulness of responses. It is best if someone from outside the school collects the data in order to make students feel comfortable answering the questions honestly. Collecting biological or biochemical indicators of behavior (e.g., carbon monoxide breath samples in smoking prevention studies) with the survey data can help to enhance the validity of the self-report data.

Trained Data Collectors. The training and experience of the individuals administering the questionnaire can affect the quality of the evaluation data that is collected. Data collectors should be carefully selected and thoroughly trained to properly administer the evaluation instruments. They should be prepared to respond to any questions that arise during the course of administering the questionnaires.



711 Westchester Avenue, White Plains, New York 10604
Tel: 914.421.2525 Fax: 914.683.6998

Conducting Valid Evaluations of LifeSkills Training and Other Prevention Programs (p.2)

Tracking Implementation Fidelity. Program implementation fidelity should be carefully monitored as part of a comprehensive process evaluation. If a program is not implemented fully or accurately (for example, in the correct sequence), then it may not be effective. In fact, the argument has been made by leading prevention scientists that local evaluations of evidence-based prevention programs such as LST should direct limited resources to process measures (i.e., whether the program was delivered with high fidelity) rather than to outcome measures (i.e., student substance use).

Maximizing Participant Retention

Retention of Participants. Data collected from participants after the intervention is completed should include all (or a large percentage) of the original sample of participants. Participants should be assigned an ID number in order to enable the tracking of their responses over time while maintaining student confidentiality.

Differential Attrition. Participants who drop out of interventions are often at higher risk than those that remain in the program. Thus, if high-risk participants in the control condition drop out at a higher rate than those in the intervention condition, the control group may appear to have improved more than the intervention group, particularly if these dropouts were engaging in risky behaviors at the pre-test measurement. Attrition of students should be examined and shown to be similar across intervention and control groups. Furthermore, attrition of high-risk youth should be examined and shown to be similar across intervention and control groups.

Analyzing the Data

Level of Analysis. The primary focus in determining the effectiveness of a prevention program is typically on change at the individual level. However, if schools were randomly assigned to the intervention and control conditions, rather than the individuals, then the statistical analysis should take this into account. Either the multilevel nature of the data should be considered or the analyses should be done at the school level. In the latter situation, there is a considerable loss of statistical power and only very large effects will be noted.

Clustering of Students. A second concern is that students within schools tend to cluster together with respect to their substance use and other risk and protective factors. For example, the behavior among students within a school is often more similar than the behavior of students across schools. This is because friendship groups share similar activities and interests (including substance use). These groups occur naturally within schools' boundaries and can facilitate social learning regarding a wide variety of behaviors and contribute to similarities in normative beliefs regarding problem behavior. Thus, the clustered nature of the data should be taken into account in the statistical analyses.