The Prevention of Alcohol and Drug Involvement II: Evaluating Preventive Interventions

La Prevención de la Adicción al Alcohol y Drogas II: Evaluando Intervenciones Preventivas

Raymond P. Lorion
Ohio University

This paper focuses on issues related to the design, conduct and interpretation of studies of the effectiveness of interventions to prevent involvement with alcohol, tobacco, and other drugs (ATOD). As explained in the preceding paper (Lorion, 1998), a preventive intervention should be built on a solid foundation of epidemiological, etiological and developmental work and component field-testing. In asserting that position, I do not mean to suggest that we need to know everything about the disorder(s) or problem(s) whose prevalence we wish to reduce. We must, however, know enough to create a theory of the problem and its complementary theory of the solution. After working on these issues for nearly 25 years, I have come to respect the challenge we are pursuing. I now truly appreciate the complexity with which emotional and behavioural problems evolve over time and across a series of successive transactions between individuals and the settings, circumstances and experiences they encounter. The developmental sciences generally, and developmental psychopathology particularly, are beginning to reveal the nature and mechanisms of such processes. Gradually, we are beginning to model pathogenic processes and thus beginning to understand the course of emotional and behavioural development. From that knowledge we must derive the material from which our intervention is to be built.

I have long argued that the success of any preventive intervention depends on the completeness of its informational base (Lorion, 1983; Lorion, Myers, Bartels & Dennis, 1994; Lorion, Price & Eaton 1989). This position is consistent with the National Academy of Science's Institute of Medicine's endorsement of the "preventive intervention research cycle" (Mrazek & Haggerty, 1994). It is also reflected in the NIMH's (National Institute of Mental Health, 1996) five phases of preventive intervention research. In the prior paper, I refer to these phases as a "basic recipe for any prevention program". This paper examines how the recipe applies to the evaluation of preventive interventions. To do so, I will go through each of the following phases:

1. Define the problems of interest and study their extent.
2. Study risk and protective processes that influence the (non) development of these problems.
3. Develop and assess the efficacy of preventive trials to change the risk and protective factors and thereby influence problem incidence and prevalence.
4. Conduct large-scale trials of demonstrably effective prevention programs.
5. Facilitate program diffusion and evaluation.

Program evaluation is relevant to each phase because evaluation should be an integral part of planning, conducting and disseminating the intervention.

Simply stated, the primary evaluation task for each phase is to determine that its informational requirements are met. If they are, then proceeding to the next phase is reasonable. If they are not, the decision to proceed in the absence of the necessary information must be made carefully and thoughtfully. Is enough known about the problem to translate into a scientifically sound intervention design? Do we know who is at risk? Can we gain access to them? Can we implement the intervention when, where and in the manner necessary? The re-
cord of success for interventions to reduce the prevalence of ATOD involvement remains rather disappointing to this point. I would argue that past failures reflect the costs of saving time, saving money or arguing that the seriousness of the problem justifies moving ahead in spite of significant gaps in knowledge. I would also argue that doing so has costs the field considerably in terms of our credibility, our self-confidence and public faith in our ability to control ATOD problems. It has also resulted in the waste of scarce funds and resources. I believe that we would be much further along in meeting our challenge had we not taken past “short-cuts”. Admittedly, my comments provide a map for the route. I recognize, however, that some steps will need to be taken with only partial information. I hope, however, that at least we acknowledge what information we lack and incorporate that lack into planning for the intervention’s implementation and evaluation.

Defining the problem

As explained in the preceding paper, I use the term involvement to refer to the range of substance related problems including use, abuse, dependence, acquisition and (particularly in the case of drugs) distribution. I noted that substances include ATOD and believe that all substances, including alcohol and tobacco, should be considered across the life span. Since I focus on youth, I am not concerned about the issue of legality, since for them, all of these substances are illegal in the United States. Tobacco is not supposed to be sold to those younger than 18; alcohol is not supposed to be available to or consumed by those below 21. Given their role as gateways to other substances, the prevalence of their use among youth, and their consequent risks to individual health and welfare, these substances must be given first priority in targeting preventive efforts. Ideally those efforts will reduce the number of youth beginning to use each of these substances. At the very least, preventive efforts must seek to reduce use insofar as possible. By contrast, alcohol and tobacco are legally available to adults. Rather than focus on the onset of their consumption, therefore, we may need to design indicated interventions for those adults who seek to reduce or end consumption and thereby avoid the health and behavioural consequences of continued use.

Phase I requires that we define the problem, i.e., make explicit what behaviour(s), problem(s) or disorder(s) we will target through our intervention program. General terms such as “use” or “abuse” are not adequate to meet that requirement. Recall that it requires not simply defining the problem but also studying its extent. A lack of specificity at the outset of the program design will compromise the integrity of matching program intent with its recipients and ultimately the likelihood that the intended outcomes can be demonstrated. At the risk of being tedious, let me review with you the complexity of a concept as seemingly simple as use. Depending on its meaning, rather different segments of the population will be targeted and the intervention’s effectiveness will be measured in terms of quite different outcomes.

Is the program designed to prevent initial use or continued use? Assume that the intervention is designed to reduce the onset of substance use. Logically, therefore, one must target those with no prior history of use of the substance to be avoided. In the case of tobacco and most other substances, that criterion appears relatively clear and its confirmation depends solely on the reliability and validity of whatever procedure one uses to determine history. In the case of alcohol and prescribed medications, however, distinctions must be made between socially sanctioned use and its prohibited counterpart. Wine consumed as part of religious or family ceremonies is generally distinguished from its consumption under other circumstances.

Medications used as prescribed must be distinguished from their non-prescribed use. To select an appropriate sample for an intervention to reduce the prevalence of onset (i.e. to reduce incidence), one must generally target children. My own research with low-income, minority children, revealed use of alcohol and tobacco in children as young as nine years old (i.e. Flores-Fah, Lorion & Jakob, 1997). Were we intent on reducing onset, it would be necessary to confirm the validity of our self-report measure and, in a cross-sectional design, survey large numbers of children to determine the proportion of those who have never used a substance and those who have. Those proportions represent the “base rate” of use in a target population. I recommend that such information be obtained cross-sectionally. In that way, the proportion of non-users at each age can be determined as can the critical point in development at which the transition from use to non-use is most likely to occur in a population of interest. If, for example, there is a substantial increment in the proportion of users between grades
four and six, placement of the intervention in grades four and five can be justified. Information about transition points also provides insight into when one is likely to see an intervention effect. In this case, one would expect that fewer sixth graders would report use following participation in the intervention during the prior year or two.

"Use" however, might refer instead to continuation rather than onset. In such instances, the targets will be individuals who have tried a substance one or more times, rarely if ever repeat that experience. Alternatively, the target may be individuals who consume the substance regularly but with sufficient temporal spacing between use to avoid dependence or addiction. It should be clear that each of these distinctions would alter the targeting of an intervention, its contents and its criteria for success.

You may have noted the term *instead* when I referred to other meanings of use. I do not intend to suggest that one must restrict an intervention to a single group or to a single, narrow definition of "use". What I do intend is to make explicit the need to be explicit about one’s intervention goal and most importantly to know exactly who is receiving the intervention. Consider the differences in the subgroups of “users” described thus far. Imagine if they, in unknown or diverse proportions across cohorts, were recruited to participate in an intervention to “prevent alcohol use”. What criterion would one apply to determine program success? Unless documented at the individual level, comparing the proportion of users and non-users across time could be meaningless. Users could have stopped and non-users could have started and the data would suggest no change. Hardly an accurate interpretation of what happened.

Without going into the details, I will point out that similar precision is necessary if the intervention’s goal is to reduce the prevalence of dependence, addiction, or the problematic consequences of substance use (i.e. driving under the influence, academic failure, unintended pregnancy, violence). In each case, we need to define the criteria by which the presence or absence of the outcome can be confirmed and determine the precision with which we can apply the criteria.

**Risk Assessment**

As noted, phase I requires that we define the problem(s) and study its (their) extent. Thus, for each substance of interest, we must apply epidemiological methods to acquire information about the proportion of users and the nature of their use across time. That information, in turn, will inform decisions about the focus, targets, procedures and outcomes of the intervention. To do so, however, the information must go beyond confirming the existence of a problem. Phase II involves acquiring information about the problem(s)’s distribution throughout the population or subgroup of interest and identifying demographic (i.e. age, gender, socio-economic level, race), individual (i.e. emotional status, personality characteristics) and situational (i.e. familial style, academic achievement, neighborhood) factors associated with the presence or absence of the problem(s). Those factors associated with an increase in incidence or prevalence are labeled “risks”; those associated with a reduction are labeled “protective factors.” Both should be considered when assessing the extent of a problem.

The importance of this aspect of developing the body of information relevant to the design of an intervention is that it offers the major solution to what has been referred to as the “base rate problem”. For example, the fact that most of the problem(s) about which we are concerned occur in a relatively small portion of the population (Lorion, Price & Eaton, 1989). Use of tobacco and alcohol by primary graders (K-3), for example, occurs in less than 5% of that population. Prevalence rates double around grade 4 and continue to increase thereafter. If onset of use is the focus of the intervention, therefore, one will need extremely large samples to document reduction if the primary grade population at large is targeted. Integrating information about known risk and protective factors allows for the inclusion of risk enhancement procedures into recruitment strategies. Simply stated, risk enhancement involves maximizing the likelihood that the problem of concern will occur in the intervention sample.

Epidemiologists use two bio-statistical procedures within a case-control design to assess and index the strength of a risk or protective factor. In case control studies, individuals with a confirmed disorder (i.e. cases) are compared with individuals who do not display the disorder. “Relative risk” refers to the incidence rate of the problem in those exposed to the risk relative to the incidence rate in the non-exposed group. As Lilienfeld and Lilienfeld (1980) note, the incidence rate can be estimated by comparing the cross products of the entries provided in Figure 1. Also referred to as the “odds ratio”, the relative risk indicates how many times more likely
one is to experience the problem if the risk is present than if it is not. By calculating the relative risk for combinations of risk factors, as well as protective factors, one can determine, which particular subgroups are most likely to become problematic if intervention is not provided. Youth who have a record of academic failure, physical or sexual abuse, a family history of an addictive disorder and live in a single-parent household, for example, may be at substantially heightened risk for early alcohol or tobacco use.

Were risks and protective factors not correlated as highly as most are the calculation of relative risk for sets of characteristics would be simple. One would need only to sum the relative risk of each characteristic. Although it would provide a rough index, it is likely to overestimate, in many cases substantially, the relationship between combinations of risk and protective factors and the outcome to be avoided. Accurate calculation of relative risk would inform one how broadly or narrowly to cast the recruitment net. The lower the base rate, the narrower the set should be to maximize that those most at risk will be included within the study. If the base rate, for example, is 10% and one can derive a profile of those at risk which enhances that risk by a factor of three or four, the base in the resulting sample would approach 30-40%.

Given the number of risk and protective factors which have been identified for emotional, behavioural and substance disorders, how does choose among them? Which should be included in the combination. “Attributable risk” estimates the “maximum proportion of a disease that can be attributed to a characteristic or risk factor; alternatively, it is considered the proportional decrease in the incidence of a disease if the entire population were no longer exposed to the suspected etiological agent” (Lilienfeld & Lilienfeld, 1980, p. 217). In other words, how much would the incidence of alcohol use among adolescents be reduced, for example, if they assumed that few of their peers used the substance? If few people in the population share the risk factor or characteristic and its relative risk is rather low, little advantage is to be gained by targeting those with such risk. By contrast, neutralizing a risk factor (or set of such factors) which occurs in a substantial portion of the population and which has a high relative risk, has the potential for substantial impact on incidence and prevalence.

Few preventive interventions thus far have been designed on the basis of relative and attributable risk analyses. Such studies are viewed generally as too expensive and too lengthy. In the absence of their findings, however, a substantial amount of time, money and effort may be invested in an intervention which accomplishes its goal, for example, to reduce the risk factor(s) but makes little difference on problem’ scale in the population. With that in mind, I ask you which is the better investment? Assuming that we can design an effective preventive intervention without highly specific information about its risk factors is like assuming that we can understand and travel around a foreign country without a map, without directions and without even knowing the language. That approach would only for small countries with little geographic variation, few roads that are incredibly well marked, and no variation among its people. Even if we could do it, it doesn’t sound very interesting and most would wonder why bother?

**Focussing on Processes**

Phase II requires that we study not simply factors but the etiological processes to which they contribute. In the preceding paper, I spoke briefly about recent advances in the developmental sciences, which are enhancing our understanding of the genesis of emotional and behavioural states. We are beginning to understand transactional processes and the systemic ways in which individuals and environments influence each other through series

<table>
<thead>
<tr>
<th>Characteristic (case)</th>
<th>With disorder</th>
<th>Without disorder (Control)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>present</td>
<td>A</td>
<td>B</td>
<td>A + B</td>
</tr>
<tr>
<td>absent</td>
<td>C</td>
<td>D</td>
<td>C + D</td>
</tr>
<tr>
<td>Total</td>
<td>A + C</td>
<td>B + D</td>
<td>A + B + C + D = N</td>
</tr>
</tbody>
</table>

* Based on Lilienfeld & Lilienfeld, 1980.

Perhaps their finding of greatest relevance to the prevention sciences are the concepts of equifinality and multifinality. The first, equifinality, helps us to understand inconsistencies across attempts to link antecedent conditions or individual characteristics and specific outcomes. It appears that comparable outcomes may result from diverse antecedents. A history of abuse, for example, has been linked to many emotional, behavioural and addictive problems. Similarly, poverty appears to be a risk factor for a diversity of problems. The second, multifinality, links single antecedents with multiple outcomes. This may explain the seeming disparity across studies of risk factors for delinquency or alcohol dependence.

Over the past few years, my research team and I have been attempting to understand these developmental processes. As we examine our data concerning antecedents to outcomes ranging from involvement in violence to early alcohol and other drug involvement, we believe that we have found a commonality across these multiple antecedents. If we are right, then one explanation for equifinality might be what we have termed “developmental asynchrony”, for example, an encounter which exceeds and individual’s developmental readiness to cope with it. Consistent with work by Lazarus and others on the nature of stress and coping, asynchronous situations overtax the system. Examples of diverse antecedents include early physical development in girls, assigning parental responsibilities and expectations to a child in a single parent family and living in a violent and dangerous home or neighborhood. It also appears to us that the diversity of outcomes resulting from a common antecedent may reflect social norms concerning acceptable or facilitated forms of expressing negative affect, rejection of authority, etc.

Clearly far more work needs to be done before we understand the classificatory schemas which determine equifinality and the psychosocial processes reflected in multifinality. Recognition of both processes, however, seems quite relevant to the design and evaluation of preventive interventions. Insofar as equifinality is concerned, the diverse antecedents may, in fact, guide us in the components of risk enhancement algorithms and thereby improve our chances of including within an intervention a substantial portion of those likely to experience a particular outcome.

In a similar manner, identification of the diverse outcomes linked to a specific risk factor may enable us to broaden the array of dependent measures used to assess an intervention’s outcomes. Ideally, an intervention will reduce substantially an outcome; alternatively, it may consistently and reliably reduce somewhat the prevalence of a number of undesirable outcomes. By combining both of these strategies, it may be possible (and necessary) to design interventions with sufficient attributable risk among the targeted antecedents to confirm the programs’ effects across their associated diverse outcomes. Pragmatically, such an approach will require new models of risk assessment and multivariate models for the assessment of outcomes.

Preventive Trials

Phases I and II provide the information necessary for deciding on the specific outcomes to be effected and identifies their associated antecedents. In combination with available etiological and developmental theory, the ingredients can be assessed in terms of their adequacy for organizing a “theory of the problem” and relating that to a “theory of the solution”. For the latter theory to mean anything, however, one must concretize the elements of that theory into intervention components. These components include the design of procedures for gaining access to and recruiting the appropriate sample, obtaining information about their risk status and history of substance involvement and involving them in the activities and information which define the intervention. Ideally, this portion of the intervention’s design will be continually monitored within a process evaluation paradigm. In most instances, this paradigm will include some form of management information system developed during the initial phases of the program’s evolution and continued throughout its efficacy trials and the initial stages of its dissemination phase.

Interventions develop through a series of stages. Each of the components, for example, must itself be designed, applied in test situations and refined to maximize its approximation of its intent. When developing risk detection or outcome measures, for example, my team proceeds through a sequence of qualitative studies. First, we examine and critically
analyse the conceptual, empirical and clinical literature to learn as much as possible about the problem(s) of interest. If possible, we will find opportunities to observe the problem in vivo and to interview individuals who present the problem, who observe its manifestations or who work with its victims. These interviews are most helpful in refining our understanding of how the problem develops and how it can be prevented.

Focus interviews with adolescent alcohol and other drug users, for example, changed our assumptions about the process of onset. Frequently, the image of the initial episode of use contains an “innocent” non-user whose intention to avoid use is overcome by pressure from one or more peers. Our findings suggest that for some adolescents, a process occurs over time within which the non-user develops curiosity about and interest in the substance and enters situations in which opportunities for use exist. Early in the process, involvement is limited to observation and decision-making. We speculate that this is a point at which an individual may be very responsive to information about the avoidance of use, which is accurate. Presumably, at this point, the non-user is conducting an informal study to validate the contradictory messages provided about alcohol, tobacco and other drugs. As noted in the preceding paper, for example, crack users have told me of their conviction that they could resist the substance’s addictive quality since they have been able to do so for months before they became regular users, dependent and then addicted.

When involved in the development of measures, we involve representatives of the intended respondent group in the refinement of questions and response options. In the past, for example, we interviewed 4th graders individually to obtain their input about an intended screening approach. We asked them to read each item aloud (confirming its readability), to tell us its mean (confirming its comprehensibility) and, in many cases, to suggest other, more ways appropriate ways to state the item. Frequently, we need to repeat this process several times before the scale is finalized. The second review regularly includes input from others (i.e. teachers or parents) involved with our intended program recipients. The point of this activity is to maximize both the face and content validity of the measure. Once developed, it is field-tested to gain information about its structural characteristics, its distributional qualities and its psychometric characteristics (i.e. reliability and validity).

The same effort is required in developing informational packets or presentations. If activities are to be involved, they must be tested in the field to confirm their age-appropriateness, their time demands, their cultural acceptability and their capacity to retain the participant’s interest and attention. Conceptually and empirically established links between activities and outcomes may matter little if considered boring, demeaning or stigmatizing.

Once together and tested under limited and controlled conditions, the components of the intervention must be tested within an efficacy protocol, commonly referred to as a preventive trial. Ideally, these trials are designed as experimental tests of the intervention. As such, an appropriate sample is recruited and pre-data are collected using the dependent measures designed or selected for the study. If possible, the sample is then randomly divided into intervention and non-intervention groups. In many cases of community-based trials, the unit of assignment may be a school, a neighborhood or even an entire community. As you are probably aware, random assignment is an ideal design component rarely achieved in reality. Whether this reflects the impossibility of doing or rather than assumed impossibility of doing so should be determined directly in the setting in which the efficacy trial is to occur. In my experience, if adequately explained and discussed during the initial negotiations with the setting, random assignment of schools, for example, to intervention or control status is often possible. This likelihood increases when the control setting is assured the opportunity to implement the intervention once the efficacy trial is completed. Since most such studies involve a limited series of consecutive cohorts, such assurance can usually be fulfilled within a three year period. If random assignment is not possible, the pre-test data may allow for the evaluation to be carried out within a matched groups or matched settings protocol. If such is not possible, then one must resort to quasi-experimental methods to maximize the scientific validity of the efficacy results.

As important as the assignment of individuals and settings to intervention conditions may be, it is equally important that appropriate manipulation checks be incorporated within the study’s process evaluation components. A comprehensive and accurate record of the activities carried out and information provided within the intervention must be available. Similar records should be obtained
PREVENTION OF ALCOHOL AND DRUG INVOLVEMENT

from the control settings in order to confirm that contamination across settings has not occurred. It is most important in the early stages of the efficacy trial to ensure the intervention's fidelity and consistency across settings. Too many promising programs have appeared weak because rather than an intervention, a series of related although distinct interventions were, in fact, being assessed.

Any theory of the problem and theory of the solution must include precise information about the temporal aspects of the problem(s) to be addressed (Lorion, 1987, 1990). As I have explained elsewhere, a period of time passes between antecedent conditions and exposure to risk factors and the onset of pathogenesis. More time passes as the pathogenic process unfolds and the problem(s) begin to manifest. Temporal issues of most importance include when in the first period to stage the intervention and when thereafter to locate subsequent measurements of the presence or absence of pathogenesis. Epidemiologists refer to the incubation "period" as the period between infection and symptom manifestation. If infection cannot be avoided, the incubation period represents both an opportunity for intervention to arrest the process and the minimum interval before the intervention's efficacy can be confirmed. Recognition of this fact explained part of the complexity confronting AIDS researchers as they attempt to design effective preventive interventions. It also explains the difficulty in trying to reduce the occurrence of cross-generational child abuse given the lag between the risk (i.e. being a victim of abuse) and the opportunity to express that risk (i.e. once parenthood is achieved).

As we think about the diversity of ATOD related outcomes which might be targeted by preventive interventions, it is apparent that each has its own temporal demands and consequent implications for the design of the evaluation protocol. Although it is rarely included within an efficacy trial's design, I am increasingly convinced that estimations of what Mayer (personal communication, 1996) refers to as the "preventive fraction" ought to determined prior to the conduct of the intervention. Related to attributable risk, the "preventive fraction" refers to the maximum reduction in the prevalence of the problem(s) if the intervention operates as intended. I suggest that not as the minimum criterion for defining success but rather as an index whereby we can track the degree to which we have approached that maximum. Thus if the best one could hope for is a reduction of 1/3 in the prevalence and a 1/4 reduction is obtained, the intervention has achieved much of its potential effect.

An important reason for combining process and outcome procedures in the design of an efficacy trial is that it allows for consideration of program fidelity as a contributor to program outcomes. This is generally possible when the efficacy trial involves delivery of the intervention across multiple settings and one can compare the outcomes observed across settings and across levels of fidelity. Through such analyses, it becomes possible to speculate about the differential contribution of program components to the achievement of preventive goals.

Ideally, however, efficacy trials will involve multiple iterations of the intervention design, each of which builds on what is learned from the former trials. Through a series of such trials, the intervention evolves to the point that it meets expectations of achievable outcomes, cost-effectiveness and replicability. It is at that point that one moves to the next phase.

Effectiveness Trials

Phases IV and V relate to the design and conduct of large-scale evaluations of an intervention across multiple communities and varying degrees of program fidelity. At issue is whether the intervention is sufficiently robust to have its impact under real-life conditions and in the absence of the level of control characteristic of the efficacy trial. Effectiveness trials make available the intervention materials and, ideally, some degree of training in its procedures. Process evaluation continues in order to determine the limits to which one can deviate from the original protocol and retain acceptable levels of impact. Such information is particularly important as one refines both the material and training elements of the intervention.

Cross-site studies are increasingly important in my nation's efforts to organize a pool of demonstrably effective preventive interventions, which can be affordably obtained and implemented by communities. Participation in such studies typically exchanges access to materials and training for cooperation in the continuing collection of outcome data across settings, implementations, etc. Having participated in such studies both as a program implementer and as an evaluator, I can attest to their difficulties. In my estimation, they have not yet fulfilled their potential both because of the challenges
associated with carrying out a scientifically controlled study across dozens if not hundreds of settings. Commitments to cooperate notwithstanding, it is extremely difficult to obtain the information sought in a timely and complete manner. It is particularly difficult to convince settings that their continued will not depend on their performance and on their documented achievement of preventive goals. My colleagues who devote their lives to such endeavors have my sincere respect and best wishes.

Conclusion

Hopefully, my comments over the past two days in combination with those of your colleagues who have undertaken such efforts have provided you with insights and hope for the future. I can assure that I have truly optimistic about our capacity as social and behavioral and health scientists to reach our goals. What we know about development combined with gains in epidemiology, biostatistics and information processing has strengthened our hand enormously.

References


