Lessons Learned in the Multisite Violence Prevention Project Collaboration: Big Questions Require Large Efforts

Multisite Violence Prevention Project

Abstract

This paper summarizes some organizational, scientific, and policy lessons that have emerged in the formation and conducting of the collaboration of the Multisite Violence Prevention Project. We contend that these lessons are valuable for other collaborations and are important for furthering the utility of scientific efforts. A central contention is that large-scale efforts such as this collaboration are underused but are essential for efficient advancement of knowledge about preventing youth violence.

Introduction

Prevention is, by necessity, big science. It is an undertaking that is large in scope and grand in its ambitions.[1,2] Youth violence prevention is a good example. It requires complex theories that can incorporate multiple influences, multiple levels of causal factors, and elaborate considerations of the roles of time (development) and place (community and ecologic differences).[3–6] It is based in a multidisciplinary knowledge base and integrates multiple methods.[7,8] It focuses on patterns of prevalence and incidence in identifying risk and measuring change, emphasizing entire populations rather than specific individuals.[9,10]

Often, violence prevention requires multiple-component interventions with complicated implementation procedures directed at multiple levels of influence and that rest on integrating multidisciplinary information.[11,12] Violence prevention is also big science because it attempts to fulfill the expectation that elegant theory guide the work yet be simultaneously adaptable to the needs and demands of the "real world."[13] It is meant to be a careful, organized, scientifically verifiable approach to a multifaceted and situational problem and an approach that is meant to produce simple and clear direction for reducing this serious public health threat.[14] It must be grounded in the broad empirical literature on risk and protective factors, while at the same time addressing cultural and contextual issues relevant to the diverse populations of youth at which it is directed.[15] The careful, systematic effort needed to accomplish these goals occurs within the context of pressure from communities in urgent need of effective strategies for addressing this serious problem.[8]

Address correspondence and reprint requests to: Patrick H. Tolan, PhD, Institute for Juvenile Research, Department of Psychiatry, Families and Community Research Group, 840 S. Wood Street, Chicago IL 60612-7347. ptolan@psych.uic.edu or tolan@uic.edu. Multisite Violence Prevention Project (corporate authors) includes: Centers for Disease Control and Prevention, Atlanta GA: Robin M. Ikeda, MD, MPH; Thomas R. Simon, PhD; Emilie Phillips Smith, PhD; Le'Roy E. Reese, PhD (all Division of Violence Prevention, National Center for Injury Prevention and Control); **Duke University, Durham NC:** David L. Rabiner, PhD; Shari Miller-Johnson, PhD; Donna-Marie Winn, PhD; Steven R. Asher, PhD; Kenneth A. Dodge, PhD (all Center for Child and Family Policy except Asher [Department of Psychology]); **University of Georgia, Athens GA:** Arthur M. Horne, PhD (Department of Counseling and Human Development Services); Pamela Orpinas, PhD (Department of Health Promotion); William H. Quinn, PhD (Department of Child and Family Development); Carl J. Huberty, PhD (Department of Educational Psychology); **University of Illinois at Chicago, Chicago IL:** Patrick H. Tolan, PhD; Deborah Gorman-Smith, PhD; David B. Henry, PhD; Franklin N. Gay, MPH (all Institute for Juvenile Research, Department of Psychiatry); **Virginia Commonwealth University, Richmond VA:** Albert D. Farrell, PhD; Aleta L. Meyer, PhD; Terri N. Sullivan, PhD; Kevin W. Allison, PhD (all Department of Psychology).

Yet, little of the work in violence prevention is large in scope or implementation. Fiscal limitations, conventions of scientific practice in intervention research, and limitations in methods of measurement and analysis all impede research on the necessary scale.[9] Practical compromises must be made in the scope, the robustness, and the utility of research projects and, therefore, in the speed with which they are completed and in their applicability to public health practice. In fact, it may be argued that what limits the prevention of youth violence is this gap between the necessary scope and what can be accomplished given the limits of theory, methods, funding, and traditions of organizational practice in the conduct of science.

The current study, the Multisite Violence Prevention Project (MVPP), arose in response to this concern, essentially a gap between efficacy and effectiveness in research and application.[2,14] Efficacy research refers to carefully controlled, usually random assignment and relatively small-scale trials of interventions to determine whether they have a statistically significant effect. Effectiveness refers to attempts to evaluate or document the effect of interventions with proven efficacy in less-controlled conditions and with more diverse populations.[4]

To bridge the gap between the robustness of research in efficacy trials and the necessary scope of effectiveness evaluations, the MVPP[15] incorporated a scientifically rigorous method but moved beyond the traditional dictates of efficacy research and involved a very diverse population in a "real-world" setting, in this case a school. The resulting design and implementation features, as described in other articles in this supplement, grew out of this interest.

The Multisite Violence Prevention Study

Although there have been important advances during the past decade in theories and research on risk and prevention of youth violence, very few such prevention efforts in the nation have been based in this research.[12,17] This situation was also once true of other prevention efforts, such as substance abuse prevention, whereby interventions were suggested and implemented with little evidence to back their claims. The substance abuse field, however, has been more successful of late in promoting interventions backed by research. The field has moved from small-scale, quasi-experimental studies to large, randomized, multisite, long-term studies.[18,19] However, in violence prevention, considerable uncertainty remains about the viability of many efficacy efforts in practice (often referred to as "the real world").

Further, many gaps exist between issues of most concern to public health officials and those of most concern to researchers regarding violence prevention. A clear example is that most efficacy studies assign conditions and make comparisons at the individual level, usually comparing effects within a given setting. However, in reality, most programs are administered to an entire school or community.[20] Despite strong theoretical arguments for implementing violence-prevention programs at the larger social system level,[14] such interventions have rarely been evaluated.[10] The evaluations that have been conducted often do not meet such basic standards as control groups or repeated observations on outcome measures.[10,21]

It is unclear to what extent ignoring this mismatch between interests in efficacy and effectiveness lead to misleading results about the value of a given intervention (e.g., whether overestimating or underestimating its effectiveness). What is clear is that the gap leaves it to the consumer to judge how the setting characteristics of their implementation are to be decided, and it leaves only partially explained which aspects of a given program are critical to retain and which are adaptable in a "proven" program.[a]

There are precedent and models for such large-scale scientifically careful evaluations in prevention of substance use and in altering health behaviors.[22] Community-based efforts to prevent heart disease have also undergone large-scale, multisite evaluations. The Stanford

Five-City Project,[23] the Three-Community Study in California,[24,25] and the North Karelia Project in Finland[26] all used intervention and control groups to assess the effects on behavior and risk factors at the population level. The community was the unit of assignment in a large-scale intervention designed to limit drunk driving and related injuries. The evaluation randomly surveyed residents of three communities in different states who had been assigned to either a control or comparison group. In the area of violence prevention, the FastTrack intervention, for example, is a highly coordinated multisite study of a specific intervention program delivered in schools and designed to prevent conduct disorder and associated problems such as violence. [27,28] Others[11] have conducted replications of interventions across multiple sites at one time. Although these examples show that the gap can be bridged or at least considered, how attempting to undertake such large-scale work affects the scientific endeavor has not been extensively written about.

As noted, the Multisite Violence Prevention Project is an attempt to apply the scientific rigor that typifies an efficacy trial to a highly diverse population and setting (four cities, with variations in socioeconomic status distributions, ethnic groups, urbanicity, and residence location). In addition, program organization and condition assignment were made at a level more comparable to a typical application (e.g., at the school level). This approach was a significant departure from "business as usual" in violence-prevention research. This departure led to significant challenges to the presumptions of all involved and what we think may be some useful lessons for others. For example, this project required innovation in the work relationships among the sponsoring agency and the investigators, in how investigators worked together, and in how the project related to the participating communities. This bold step did not overcome, obviate, or circumvent the previously mentioned limitations of such endeavors, and it engendered some limitations that smaller-scale research can avoid. However, the effort is instructive in how to narrow the gap between basic efficacy knowledge and public health utility in reducing youth violence. In this paper we attempt to describe some of the lessons we learned in applying theories, methods, and procedures customary to prevention trials; in organizing measurement and implementation of the scientific principles; and, perhaps most telling, in understanding the organizational issues required in any collaborative prevention endeavor. We do not presume that we are the first or even the most acutely observant of these issues, nor do we assume that they are all generalizable to other large-scale efforts. We offer these lessons in hope that they may be helpful to others engaged in developing, funding, and conducting violence-prevention research.

Six Lessons Learned About Conducting Large-Scale Violence-Prevention Evaluation Research

1. Efficacy and effectiveness may not be as distinguishable as previously assumed, and it may take too long to undertake them sequentially

A first lesson arising from the project was that the lengthy time required to complete efficacy and effectiveness trials could be shortened without sacrificing quality research. Efficacy is usually differentiated from effectiveness first by its order: efficacy precedes effectiveness.[4] This "building block" approach to scientific progress rests on the traditional and orderly

---

[a]A "proven" program is typically an intervention with empirical evidence of efficacy. The debate is intensifying over how efficacy research should be used. Some argue for the more careful scientific study of implementation and the critical factors needed to maintain effects. Accordingly, they promote exacting fidelity to the original methods, staffing, and program characteristics of efficacious programs. The Blueprints for Youth Violence Prevention[14] is the strongest example of this approach. Others have argued that the value of efficacy research is in identifying "best practices"—characteristics that consistently, across interventions, differentiate efficacious from nonefficacious efforts. They promote synthesizing findings on individual programs, usually by way of meta-analyses, to identify critical characteristics that should guide local development of violence prevention.[38,42] At present, the question is left open to speculation because there have not been studies that can compare the two approaches. Absent such empirical direction, a preoccupation with the "right" approach for dissemination rather than how each can be used to improve prevention efforts seems misguided.

advance from theory to refined and controlled trials to judicious and limited decrease in controls. Once this progression is complete, effectiveness is presumed adequate for full-scale application.[2] This view dominates funding patterns and what is considered "good science" and, therefore, dictates largely which prevention efforts are likely to be undertaken and which evidence is considered most credible.

However, this approach can easily consume 5 to 10 years between the onset of the efficacy trial and the start of the effectiveness trial. It can easily take two decades to gain the sufficient evidence necessary for public health applications (i.e., going to scale). This period does not include the time invested in piloting, theory development, and funding solicitation that precedes an efficacy trial, nor the problems in determining generalization of the results beyond the trial population. Furthermore, this time frame is likely a best-case scenario, presuming that efficacy is established in the first trial and, similarly, that benefits are retained throughout the effectiveness trial.

Given these time frames, perhaps it is worth considering the approach taken in this study: large-scale studies that are efficacy trials in the sense that the interventions, although derived from prior smaller-scale efficacy studies, are not exact replicas. The intent is to incorporate critical elements of both efficacy and effectiveness studies. As in an efficacy trial, this effort includes intervention components that are well grounded in theory and that appear promising on the basis of previous research. Standardized manuals, close supervision, and fidelity checks are used to ensure consistency in implementation across interventionists and sites.[29–31] Other key elements of efficacy studies are also used, such as random assignment and the careful measurement of outcomes and mediating variables using data from multiple sources.[32] The scope of this study, however, is more typical of an effectiveness study and includes a level of diversity in the interventionists, setting characteristics, and participants rarely found in an efficacy study.[16,33] For this reason, the interventions from the prior efficacy trials were modified to be applicable across settings and populations and to be endorsed by a diverse group of researchers. The intent was to incorporate the diversity that is often eschewed in efficacy trials while testing approaches with prior efficacy evidence that often defines an effectiveness trial.

The scale of implementation and comparison of effects also requires that relations with schools and communities more closely approximate those faced in effectiveness trials or in general application. We engaged whole school systems in all but Chicago, where we worked with administration to identify schools serving poorer communities. Schools were enlisted with their consent and agreement to be randomly assigned to the project. In most cases, relationships already existed between research teams and the school systems, enabling us to build on a prior history of collaboration. However, neither schools nor parents participated directly in developing the intervention, except to provide feedback on the pilot. The prior history of collaboration and our genuine efforts to listen to feedback from teachers and parents created an environment of mutual respect and investment. Our approach required continued attention to school interests in serving all children, to the desire to offer services where resources were least, and to continued awareness of controversial research activities and research questions that might lead to consternation among teachers or parents. When dealt with straightforwardly and openly, the parents and teachers were committed to cooperation.

These administrative issues had to be incorporated into the project operation rather than being screened out. Also, by making the random assignment at the school level, the results (differences in prevalence rates between groups exposed to different interventions) more directly reflect the outcomes of interest to school officials, public health agencies, and scientists (Farrell AD, Meyer AL, Sullivan TN, Kung EM, Virginia Commonwealth University, Richmond VA, unpublished study, 2002). This level of assignment also permits one to consider

change in the ecologic context not just in individuals occupying that context. In addition, the result of moderating effects based on differences in conditions can more readily be examined, given that there are natural, or what might be termed meaningful, practical variations among those included in the intervention. Naturally occurring variations across schools also provide an opportunity to examine potential factors that may moderate intervention effects. Although experimental manipulation of such variables would provide stronger evidence of such moderator effects, it is hard to envision school systems allowing control over characteristics such as school size or structure and demographic characteristics of students. Just as important, even if such unusual control were accorded researchers, the generalizability of findings is likely to be limited.

We are not suggesting that findings of within-group differences are translatable simply because there are some variations in the characteristics of the trial setting. Rather, we are suggesting that some large-scale trials may be valuable, and not merely as replications of programs found efficacious in one setting, often under usually atypical conditions. Such trials may be more efficient because they do not require the decades of work to reach some applicable understanding of whether a given approach is useful for most schools or communities.[2]

2. The focus should be on changing the social ecology (including individuals within it not apart from it)

One of the challenges of this undertaking was how to reconcile the interest in school-level assignment with the budgetary and organizational requirements of such an effort.[33] We quickly recognized that an adequate sample size (of schools) for adequate statistical power, would eclipse the capabilities of most sites and the relatively large allocation of resources for this effort. For example, there are few school systems with enough middle schools to create a meaningful random assignment and to ensure that site and condition were not confounded. Also, engaging multiple school systems imposed substantially greater and complex organization of efforts regarding engagement and implementation. Even an effort of this size —unprecedented, we believe—faced difficulty in managing enough schools to gain adequate statistical power.

It was very tempting to consider abandoning the school-level assignment and ignoring (or at least minimizing the consideration of) the social ecology of the school and the community as factors in the effects found. However, it was clear that such an accommodation was a major limitation of the field. Practical concerns have driven the field to act on the assumptions that changing individuals and changing prevalence rates are the same, and that communities and other social ecologic groupings can be treated as moderators, but not critical aspects, of risk that might fundamentally affect the findings in the trials. However, as one of the collaborators on this project noted, "How can we know if we do not carry out a study that can address this question?"

Such an assumption cannot be tested without trials that include setting when determining how random assignment is to be accomplished. Developing the focus of this collaboration, and evaluating the implications of such a focus for research design and operational complexity, made apparent why such a scope of work is rarely undertaken. Yet, it also made apparent why it is critical that more violence-prevention efforts be organized into larger efforts that can produce findings that are more readily translatable to practice and policy. If violence prevention is to have its intended effect, we must focus on changing the developmental context of the school and other settings where children spend time as much as testing for changes in the children's cognition and behaviors.[3,34,35]

3. Violence-prevention research requires teams, but should they be dynasties, conglomerates, or initial public offerings (IPOs)[b]?

This undertaking also differed from others because the call for proposals required a capacity to engage in the science and the collaboration it required, and not simply be based on a specific research proposal.[16] The selected teams of investigators were chosen because of experience in violence-prevention research, because of a proven capability of engaging local schools and communities, and because of their willingness to collaborate with others, often unknown to them. Each site brought a cadre of investigators, as did the funding agency (the Centers for Disease Control and Prevention, CDC), creating a large group with central planning and decision-making interest.

It might be argued that violence-prevention research is differentiated from other types of social science and health research by the extent to which it is conducted by teams of investigators pooling their expertise. Of course, other prevention projects have also included teams of researchers (e.g., Fast Track,[27] Early Alliance Prevention Trial,[36] Life Skills Training,[17] to name a few), but our approach differs from those projects in that ours involved forming a team made up of groups of researchers who had not worked together prior to this. This involvement meant that devising the intervention design and implementation efforts had to emerge from consensus rather than out of a particular perspective, without losing the intent to base this research on intervention that has prior efficacy study.

Our project essentially took collaboration to a next step, bringing together established and newly formed teams to form a new entity, what we have come to refer to as "the corporation."[37] We use the term corporation to capture the more formal structures involved in the process than are evident in less formalized collaborations. The formal policies we set up, and initial attention to organizational details (see below) signaled our intent to move this effort to a new level. This structure also highlights how research groups are organized in determining how work is advanced. Not only does it highlight the need to reconsider the limited scope of many projects, but it also suggests that prevention researchers should give more consideration to how efforts are organized and whether this organization is consistent with the goals of the large-scale research.

Although many approaches to large-scale pooling of expertise are plausible, we often turned to past efforts in other collaborations for direction on the best ways to manage the project.[18,27] Thus, our effort was marked by an initial and extensive consideration of which organizational form was most advantageous and the limits and advantages of different organizations. This needed consideration may be an important lesson for other violence-prevention research efforts —organizational relationships are best planned rather than allowing them to become the unidentified force that constrains and often impedes the strongest science.

Larger-scale efforts can be differentiated in several important ways. These differences reflect variations in how the collaboration was formed, the nature of the relationship of the participants to each other, the extent to which the effort grows out of prior work of all involved, and other factors. These efforts can be characterized as dynasties (e.g., characterized by longstanding leadership with stable organizational structure and multiple joint projects over time), conglomerates (characterized by several strong and stable groups combined to address multiple questions through a large-scale effort), and IPOs (characterized by an organization of introducing the collaborators at the first meeting, noting the urgency of the impending deadline, identifying what each brings to the table and the collective risk and success opportunity, and

[b]These terms are used to connote and emphasize differences in how the groups are formed and the organizational atmosphere of the collaboration. They are not meant as literal or full descriptions of the organization nor as judgments of one as more functional or desirable than another.

emphasizing the need to produce a product that can be operational quickly, and to be ready to manage upcoming responsibilities fully). There is value in each approach, but more important, each has different strengths and limitations. Although dynasties provide stability of organization and shared theoretical perspectives, they can attenuate the consideration of diverse concerns. Conglomerates can incorporate several well-developed approaches, but they can be burdened by competition to focus on "our piece" at the cost of the overall effort. IPOs can juxtapose irreconcilable priorities and internecine struggles about direction and resource allocation, but they can force the resolution for the sake of "making this work." The lack of choice of collaborators means that many differences that might not occur among conglomerates or dynasties can become substantial. The differences are, however, often immediately apparent and, therefore, addressed as integral to the collaboration. The benefit to IPOs is that each priority and each suggested approach, activity, or idea must be argued persuasively to convince the others that it was worth pursuing. The dynasties and conglomerates often have greater affinity among the investigators but may have less opportunity to fully review assumptions.

The present effort was more akin to the IPO. Fortunately, the funding was organized to permit a year of planning and, therefore, some time to get to know one another as we worked on designing and implementing the project. It became evident early on that each team had its own established ways of doing things and priorities based on specific values, and that practices often differed. In many cases, these practices had evolved over years and had never really been questioned for justification. That we were "incorporated" forced us to work out differences, and perhaps that pressure led to more careful scrutiny of any given person's or group's perspectives and, ultimately, more refined programs, measurement strategies, and analytic approaches. Different teams of researchers often share information about their approach to various problems, but rarely at the level required to implement a complex project consistently across sites. The necessity of developing a common research design and the procedures for its implementation produced an environment in which each site was able to benefit from the experience of the others and provided opportunities to explore their "standard" practices. The more common approach to large studies—to award dynasties or form conglomerates—may appear more orderly, but it may not necessarily produce stronger studies.

A key element in the progression of such efforts is the role of the funding or sponsoring agency. In the present case, the CDC scientists engaged as full collaborators with substantial input and responsibilities for managing the study. The agency made clear its interest in facilitating the best science and practical results, and "dug in" to realize this goal, which seems unusual in violence prevention. We came to refer to them as the fifth site. Thus, it may also be that violence prevention and the scale of scientific work it requires necessitate a shift in funding agency– investigator relationships, including less-strict boundaries regarding roles.

4. Experimental design in youth violence prevention: refined methods in an unhygienic world

The project design entailed new challenges in applying methods that have guided the field for many years.[35,38] These challenges included the manner in which schools were assigned to condition, the level of standardization that could be achieved across sites, and analytic methods. Random assignment to condition is usually considered one of the most desirable characteristics in violence-prevention research. Differences by condition at post-test or follow-up can be attributed to the intervention and not to other differences in subjects assigned to any condition. However, this assumption of the benefit of random assignment depends substantially on achieving true randomness in assignment. This assumption is strained when the number of units randomly assigned is smaller and the ways in which they might differ are more complex, particularly along dimensions that are relevant to the intervention focus.[39] Schools can vary in many ways even if seemingly similar along some key dimensions. Also, it is rare to have enough schools located in one geographic area of one school system to meet the number of

units that the assumption of randomness requires. Often this rarity leads to built-in heterogeneity among the sample and, therefore, across conditions, as well as confounding school and location characteristics.[34] Therefore, researchers often resort to approximations of randomness assumptions or to necessary blocking (e.g., within site) prior to randomly assigning to condition. Yet, these efforts can impose other inequities. For example, in one setting in the present study, random assignment resulted in a disproportionate number of schools with large Latino student populations being assigned to the selective intervention condition. Even this large study could include only a relatively modest number of schools.[39]

However, the lesson we learned is that the value of such compromises should not always be based on the standards of random assignment. Although it may be easier to limit systematic differences by condition by including enough subjects, in our case, assuming that setting conditions were unimportant to understanding results proved to be a serious compromise. We were faced with the choice of randomly assigning schools across all four sites to conditions or randomly assigning schools within site. Thus, we have come to view the issue as less one of "which is better" and more of "which aspect of validity is more critical and how best to balance attention to both validity criteria." Projects of this scope, and of the scope we believe necessary to advance violence prevention, may be hampered by acting as though there is homogeneity when there is, in fact, always meaningful differences that remain unaccounted for simply through random assignment.

A related issue is deciding what are real or substantive differences and what are negligible variations. Managing fidelity across settings is complicated because of different staff and different initial perspectives of the collaborating groups. Moreover, there can easily be important variations across sites in conditions under which interventions might occur and the relationship to the schools and communities. These differences can also influence discussions about other design features, such as which aspects of the research methods or intervention implementation are considered most critical. Others can consider what is desirable or even considered essential by one site or one portion of the collaborators as nonessential, derivative, or impractical.

We came to view this question as, "what is a difference and what is a slight variation?" Reaching common understanding required extensive discussions with concordant patience that one's perspective would be considered. It required framing viewpoints with an eye to both the scientific and methodologic concerns and the practical and utilitarian value. It required an appreciation that perspectives that had been honed within a given site and a given set of research projects were going to need expansion, revision, and often a letting go of approaches considered precious. It required everyone to sift carefully through all the potential areas of focus and the almost endless areas of assessment and research that might be undertaken. We had to identify the most central or valuable questions for this project, and what was required to address these well, and (the most difficult) to determine which of the many personally valued emphases had to be left out despite their potential value. Finally, we had to grow as a collaborative group to recognize that when the basis for evaluating preventive benefits begins with a single site, single investigative team, and often a single supervisor and validater of "fidelity," there is much left unanswered about how implementation should incorporate adaptations. The process highlights that in most violence-prevention research, ensuing attempts to replicate may be burdened with many unimportant requirements that are difficult to distinguish from the essential requirements.

What this process revealed is that it may be unrealistic to believe that these requirements can be scientifically distinguished through a series of studies that systematically vary implementation requirements or that loosen criteria and control for fidelity. There may be value in undertaking intervention trials that instead include varied conditions and rest on a limited

set of necessary fidelity markers. When heterogeneity is incorporated into the initial test, the intervention robustness seems to be a bigger influence on effects than when it is screened out.

That this collaboration strained the reliance on random assignment and fidelity means that it also strained reliance on analytic methods that are usually applied in measuring outcome effects. There have been important advances in analysis of prevention trials during the past 10 years (including multilevel modeling, growth-curve modeling, multisource estimates of constructs, and mixing of categorical and continuous measurement). However, it is still the case that estimating effects and differentiating effect of the interventions with confidence may be unattainable when relying on any one method, and that the type of evaluation needed for this type of study may exceed the capabilities of any single method. That the design is hierarchical, longitudinal, and multivariate and represents comparison of multiple conditions is a great challenge for most of the currently available approaches. This collaboration helped to highlight the importance of including very sophisticated methodologists in organizing large-scale prevention trials. It also highlighted the tension between undertaking studies that can fit analytic models and attempting to fit existing analytic models to the research questions.

5. Developmental–ecologic theory and the design of youth violence prevention: well, we haven't quite figured that part of the theory out yet

Most youth violence-prevention theory rests on the recognition that prevention is modifying development that is based on ecologic risk factors.[7,9] However, what is meant by the developmental ecology, how it affects development (and risk in particular), and how it should be measured and considered in analyses are still relatively unspecified. For example, there is very limited theory or prior studies to guide how one might want to estimate effects on the ecology.[40] In estimating the effects of the interventions on the schools, is the ecologic effect simply the average of the subjects within that school, or should there be a more complex formulation such as differential effect on high-risk youth, or estimates from key informants, or changes in processes within that setting when dealing with violence?[40] How does one compare effects on persons who function as developmental influences on children, such as teachers, compared with change in children's behavior and attitudes? How should subgroup effects be interpreted? These limitations in measurement strategy are accompanied by limitations in measurement development and understanding of the value and biases of various methods of collapsing the data.[9]

Similarly, given that the sketchiness of theory about how ecologic conditions affect development and risk and in how to measure ecologic effect, modeling effects can be difficult and, by necessity, may need to be more exploratory than desirable for an experimental test. In some cases, basic operational defining of theoretical formulations may be needed, whereas in others specifying how different aspects of the ecology interrelate will be difficult to predetermine. In this collaboration, our intent was to view the school as an ecologic unit for intervention focus and for measuring effects, which led us to realize the inadequacy of most theories and the scant empirical basis to guide us. At times, the formulation of theory was undertaken barely preceding measurement development or intervention design. In some cases, there was a need to roughly formulate the operationalization of "affecting the ecology," but to be inclusive in regard to measurement to permit construct development work to occur when there was time (subsequent to the actual intervention trial).

This formulation suggests that violence prevention may be limited to the extent that we continue to conduct research that fails to emphasize the roles of local setting and organizational relationships that affect development and, therefore, the effect of and meaning of interventions. A contextual developmental and intervention theory may carry the illusion of elegance because context is ignored, but such elegance may impede a practical, useful understanding of how best to prevent youth violence. However, given the costs and elaborate theoretical schema needed

to incorporate ecologic characteristics into developmental risk studies and intervention trials, it may be that there is a need for the compromise of operationalization and measurement refinement as part of large-scale intervention studies.

In attempting to address explicitly the effect on the ecology, the collaboration also highlighted another important problem for the field: "theoretical models are like toothbrushes; everyone has one and no one wants to use any one else's." The act of contrasting, reconciling, letting go of, and meshing diverse theoretical perspectives was one of the most difficult processes of this collaboration, but it may also have been one of the most useful. Each idea and each assumption (what the proponent often offered as "everyone knows") was scrutinized critically. Maintaining the strong interest of all investigators necessary in carrying out the study created a pressure to be inclusive of many ideas that were not equally supported by prior research and in many cases were not reconcilable. This organizational need was countered by the recognition that the other scientists' support of a given approach or method was required to have it fully implemented and maintained across sites over the course of the collaboration. This stressful discussion, however, provided a very valuable process. To be retained, the advocate had to address multiple concerns and criteria and do so persuasively on the basis of strong science.

This process also highlighted limitations in many prior multicomponent prevention trials. Many emerged by adding on to an initial intervention that was of most interest to the developers. The theoretical basis for different components was often quite different in how well developed they were, and often each component was driven by a relatively simplistic and unidimensional risk theory, rather than an integrated multidimensional understanding of children's development and risk (e.g., a family intervention based on a theory of family processes related to risk, a child-focused intervention based on a different theory of child's cognitive processes related to risk, a separate theory of teacher behavior related to risk organizing that component). Admittedly, we did not achieve full reconciliation across components of theory. But the nature of this collaboration and the scope of the intended work did require relating the specific theories to an overall view of how risk occurs within a developmental ecology and was another influence on modification of several source interventions. This type of collaboration makes more apparent the need for theories that can account for risk processes across aspects of the social ecology and that permit testing of how changes in one aspect can affect those in another (e.g., how teacher training might affect a child's change in beliefs about aggression). Also, it promotes more consideration of the interrelationship of developmental influences, one of the key distinguishing characteristics of ecologic approaches. This examination also helps highlight why valid tests of prevention depend on assignment to condition at the level of the school or other ecologic unit and not the more typical individual, child, or family level within (ignoring) setting or unit. This lesson is a challenge that was greater than this collaboration was able to fully achieve, but it seems to only strengthen the need for such undertakings. They provide our best opportunity to develop multilevel theories of intervention effects and risk reduction.[9,41]

6. Organizing successful large-scale collaborations: impediments, challenges, and picking your headaches

Large-scale research carries with it larger-scale and more complex organizational issues than typify most prevention trials. Large-scale violence prevention endeavors demand a respect for the time, energy, and wisdom required to successfully manage complex organizations and budgets—two skills rarely emphasized in our training. Similarly, these are skills rarely considered in selecting participants in large-scale collaborations. Managing investigators of such a collaboration may find that they spend much more time managing than investigating. They may spend more time researching organizational studies and management science than prevention science. Collaborating investigators may wonder why there is much more talk about

budgets and publishing rules in many meetings than about the nuances of prevention theory. Such collaborations call for developing and following more-formal organizational roles and rules, which is inconsistent with the academic traditions of freedom and informality in relationships and roles among investigators. These features may seem alienating to the scientists involved and may be discounted too readily. They may be viewed as getting in the way of the work or constraining creativity and achievement rather than facilitating it, sustaining the organization, and permitting efficient decision making. There may be little opportunity to ensure full agreement on each matter or full consideration of each collaborator's view. There may be a frustrating diffusion of responsibility without imposed structures for duties and decision making, a structure that may seem bureaucratic and contrary to the typical working relationships in research.

Our approach was to formalize the decision making across sites through a structure that assigned working groups to specific areas of the study development and to overlay this structure with final decision making based on consensus among the site principal investigators (or lacking consent, majority rule). Working groups usually consisted of representatives of each site and the CDC to ensure that each site was represented and that expertise was spread across sites. We used weekly conference calls to conduct the exchanges and to make recommendations and decisions. We formed working groups for each intervention component, for measurement strategies and analytic planning, and for overall project management. In addition, we formulated formal publication and intellectual credit policies. We undertook centralized data processing, followed by highly structured data documentation. A central organizing principle was that our management structure was that of a "corporation," with tasks formally assigned and centrally integrated. Thus, one common issue was how efforts to promote and further our work in one area affected other areas, and then how to reconcile the relative value of each critical role assumed by senior management. It was also necessary to have final decision-making authority reside within a small group to ensure timely advances in the needed work.

Our example and these other types of collaborations can highlight the dependence of the scientific work on the organizational health of the collaboration.[37] It is not simply an expedient nicety to pay attention to organization, resource allocation, and corporate relationships—nor is it likely something that is met with easy agreement. However, it is a central, and perhaps the most essential, expression of the mission or purpose of the collaboration. It is only with thoughtful attention to these organizational issues that achieving the collaboration's scientific goals becomes possible. For example, a regular issue in such collaborations is to what degree to allocate resources for intervention implementation versus measuring effects. Often, there is a struggle to measure all the constructs of most interest without overwhelming the participants and, therefore, the intervention engagement. An initial focus on the corporate obligations and methods of decision making and related responsibility helps manage these issues more consistently and efficiently. Otherwise, one can find that the decision making is haphazard, unsustained, and rancorous. Conversations change unpredictably with whatever current crisis is before the group. For example, it is easy to fall into adding more measurement to address all the interests of the collaborating groups. Without strong organization, however, the result is an overly long assessment package that overwhelms the rest of the operation and seriously detracts from the scientific value of the study.

These tensions and the related need to focus on organization and management as intently as on the scientific content probably occur in most violence-prevention efforts, albeit less overtly than an IPO might impose. Similarly, the extent to which decision making can be idiosyncratic without due attention to the organization of the collaboration may be a more immediate and stronger example of the difficulty attaining consistency across studies in construct definition and measurement. Thus, collaborations may, because they make evident the importance of management organization for the research conducted, illustrate the need for more careful

consideration of how we conduct violence-prevention science. These collaborations can bring to cognizance issues such as how we allocate resources to the research versus the program implementation, how and what measures and constructs are considered critical, and how we ensure that the interventions and research are conducted as planned.

Next Big Steps in Large-Scale Youth Violence-Prevention Research

The lessons discussed here are not a complete rendering of the important lessons gained from this project; they are merely some that we have found most salient and believe may be most useful for others to consider in further research. As a reading of this article may suggest, large-scale collaborations face many problems and require many major compromises. These may seem large enough or the compromises too great to warrant further efforts. Scientists and funders may choose to retreat to the more controlled, and seemingly more valid, smaller-scale efforts that are the mainstream of our field. We hope, though, that we have also conveyed the inherent limitations of smaller-scale work. In doing so, we underscore the need to frame the judgment not as which approach is more controlled and, therefore, more scientifically useful (emphasizing internal validity), but rather as which approach and its inherent compromises can better move scientific knowledge forward and better guide practices and policies. In fact, our intent is not to pit these approaches against each other. Rather, it is to highlight the value of larger-scale efforts and to emphasize the complimentary relation they have. There has been an unbalanced prizing of small-scale but "clean" studies over the less refined but perhaps more directly applicable large-scale efforts (of which this study is an example). To the question "Are the potential gains and efficiency in translating intervention theory into usable practices worth the inherent pain and risk inherent?" we say yes. In part, this answer is because the alternatives have such inherent limitations for application. Unless the science is refocused to include the scale needed here while recognizing that the efficacy–effectiveness distinction may be unrealistic, the effect on violence rates will be limited. The unacceptable number of injuries and deaths will continue while we wait for the accumulated results of smaller-scale, systematic research. The extent of extrapolation and inference needed for implementation will not diminish. This multisite collaboration may be valuable not only because of the efficacy findings it produces but also because it adds to the lessons learned from other large-scale prevention efforts and offers direct lessons on how to undertake more effective violence-prevention research in large-scale collaborations that can make a difference.

This article ends with the same coda as most research studies. Although this approach and its examples have many limitations, it appears that the initial effort is promising; however, a more confident judgment will require more time and data on the experience. What seems clear is that such large-scale efforts are needed for advancing violence prevention.

References

1. Glynn TJ. Comprehensive approaches to tobacco use control. Br J Addict 1991;86:631–5. [PubMed: 1859931]

2. Price, RH.; Lorion, RP. Prevention programming as organizational reinvention: from research to implementation. In: Shaffer, D.; Philips, I.; Enzer, NB., editors. Prevention of mental disorders, alcohol and other drug use in children and adolescents. Rockville MD: U.S. Department of Health and Human Services; 1989. p. 97-123.

3. Farrell, AD.; Camou, S. School-based interventions for youth violence prevention. In: Lutzker, J., editor. Violence prevention. Washington DC: American Psychological Association; In press

4. Mrazek, PJ.; Haggerty, RJ. Reducing risks for mental disorders: frontiers for preventive intervention research. Washington DC: National Academy Press; 1994.

5. Stoolmiller M, Eddy JM, Reid JB. Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. J Consult Clin Psychol 2000;68:296–306. [PubMed: 10780130]

6. Tolan PH, Guerra NG, Kendall P. A developmental-ecological perspective on antisocial behavior in children and adolescents: toward a unified risk and intervention framework. J Consult Clin Psychol 1995;63:579–84. [PubMed: 7673535]

7. Kellam, SG.; Rebok, GW. Building developmental and etiological theory thorough epidemiologically based preventive intervention trials. In: McCord, J.; Tremblay, R., editors. Preventing antisocial behavior: interventions from birth through adolescence. New York: Guilford; 1992. p. 162-95.

8. Potter, L.; Mercy, J. Public health perspective on interpersonal violence among youths in the United States. In: Stoff, D.; Breiling, J.; Maser, J., editors. Handbook of antisocial behavior. New York: J Wiley & Sons; 1997. p. 9-27.

9. Tolan PH, Gorman-Smith D. What violence prevention research can teach us about developmental psychopathology. Dev Psychopathol 2002;14:713–29. [PubMed: 12549701]

10. Gottfredson, DC. Schools and delinquency. New York: Cambridge University Press; 2001.

11. Catalano, R.; Arthur, M.; Hawkins, D.; Berglund, L.; Olson, J. Comprehensive community- and school-based interventions to prevent antisocial behavior. In: Loeber, R.; Farrington, D., editors. Serious and violent juvenile offenders: risk factors and successful interventions. Thousand Oaks CA: Sage Publications; 1998. p. 248-83.

12. US Department of Health and Human Services. Youth violence: a report of the Surgeon General. Washington DC: Government Printing Office; 2001.

13. Rappoport, J.; Seidman, E.; Davidson, WS. Demonstration research and manifest versus true adoption: the natural history of a research project to divert adolescents from the legal system. In: Munoz, RF.; Snowden, LR.; Kelly, JG., editors. Social and psychological research in community settings. San Francisco CA: Jossey-Bass Publications; 1979. p. 101-44.

14. Elliot, D.; Tolan, PH. Youth, violence prevention, intervention and social policy: an overview. In: Flannery, D.; Hoff, R., editors. Youth violence: a volume in the psychiatric clinics of North America. Washington DC: American Psychiatric Association; 1998. p. 3-46.

15. Dodge KA. The science of youth violence prevention: progressing from developmental epidemiology to efficacy to the effectiveness to public policy. Am J Prev Med 2001;20(suppl):63–70. [PubMed: 11146262]

16. Multisite Violence Prevention Project. The Multisite Violence Prevention Project: background and overview. Am J Prev Med 2004;26(suppl):3–11. [PubMed: 14732182]

17. Tolan PH. Youth violence and its prevention in the United States: an overview of current knowledge. J Inj Control Safety Promotion 2001;8:1–2.

18. Botvin G, Baker E, Dusenbury L, Tortu S, Botvin E. Preventing adolescent drug abuse through a multimodal cognitive-behavioral approach: results of a three-year study. J Consult Clin Psychol 1990;58:437–46. [PubMed: 2212181]

19. Bodtvin G, Baker E, Dusenbury L, Botvin E, Diaz T. Long-term follow up results of a randomized rug abuse prevention trial. JAMA 1995;273:1106–12. [PubMed: 7707598]

20. Farrell AD, Meyer AL, Kung EM, Sullivan TN. Development and evaluation of school-based violence prevention programs. J Clin Child Psychol 2001;30:207–20. [PubMed: 11393921]

21. Howard KA, Flora J, Griffin M. Violence prevention programs in schools: state of the science and implications for future research. Appl Prev Psychol 1999;8:197–215.

22. Wagenaar, AC. Minimum drinking age and alcohol availability to youth: issues and research needs. In: Hilton, ME.; Bloss, G., editors. Economics and the prevention of alcohol-related problems. Bethesda MD: The Institute of Medicine; 1993. p. 175-200.

23. Fortmann SP, Flora J, Winkleby M, Schooler C, Taylor CB, Farquhar J. Community intervention trials: reflections on the Stanford Five-City Project experience. Am J Epidemiol 1995;142:576–86. [PubMed: 7653465]

24. Farquhar J, Maccoby N, Wood P. Community education for cardiovascular health. Lancet 1977;1:1192–5. [PubMed: 68285]

25. Farquhar J. The community-based model of life style intervention trials. Am J Epidemiol 1978;108:103–11. [PubMed: 707471]

26. Puska P, Tuomilehto J, Salonen J. Changes in coronary risk factors during comprehensive five-year community program to control cardiovascular diseases. Br Med J 1979;2:1173–8. [PubMed: 519351]

27. Group CPPR. Initial impact of the Fast Track prevention trial for conduct problems: I. The high risk sample. J Consult Clin Psychol 1999;67:631–47. [PubMed: 10535230]

28. Group CPPR. Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. J Consult Clin Psychol 1999;67:648–57. [PubMed: 10535231]

29. Meyer AL, Allison KW, Reese LE, Gay FN, Multisite Violence Prevention Project. Choosing to be violence free in middle school: the student component of the GREAT Schools and Families universal program. Am J Prev Med 2004;26(suppl):20–28. [PubMed: 14732184]

30. Orpinas P, Horne AM, Multisite Violence Prevention Project. A teacher-focused approach to prevent and reduce students' aggressive behavior: the GREAT Teacher Program. Am J Prev Med 2004;26 (suppl):29–38. [PubMed: 14732185]

31. Smith EP, Gorman-Smith D, Quinn WH, Rabiner DL, Tolan PH, Winn DM, Multisite Violence Prevention Project. Community-based multiple family groups to prevent and reduce violent and aggressive behavior: the GREAT Families Program. Am J Prev Med 2004;26(suppl):39–47. [PubMed: 14732186]

32. Miller-Johnson S, Sullivan TN, Simon TR, Multisite Violence Prevention Project. Evaluating the impact of interventions in the Multisite Violence Prevention study: samples, procedures, and measures. Am J Prev Med 2004;26(suppl):48–61. [PubMed: 14732187]

33. Henry DB, Farrell AD, Multisite Violence Prevention Project. The study designed by a committee: design of the Multisite Violence Prevention Project. Am J Prev Med 2004;26(suppl):12–19. [PubMed: 14732183]

34. Group MACSR. A cognitive-ecological approach to preventing aggression in urban settings: initial outcomes for high-risk children. J Consult Clin Psychol 2002;70:179–94. [PubMed: 11860044]

35. Shadish, WR.; Cook, RD.; Leviton, LC. Foundation of program evaluation. Newbury Park CA: Sage Publications; 1991.

36. Dumas JE, Prinz RJ, Smith EP, Laughlin J. The EARLY ALLIANCE Prevention Trial: an integrated set of interventions to promote competence and reduce risk for conduct disorder, substance abuse, and school failure. Clin Child Fam Psychol Rev 1999;2:37–53. [PubMed: 11324096]

37. Schilling RF, Schinke SP, Kirkham MA, Meltzer NJ, Norelius KL. Social work research in social service agencies: issues and guidelines. J Soc Serv Res 1988;11:75–87.

38. Tolan, PH.; Brown, CH. Methods for evaluating intervention and prevention efforts. In: Trickett, PK.; Schellenbach, C., editors. Violence against children in the family and the community. Washington DC: American Psychological Association; 1998. p. 439-64.

39. Hsu LM. Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. J Consult Clin Psychol 1989;57:131–7. [PubMed: 2647799]

40. Shinn, M. Mixing and matching: levels of conceptualization, measurement, and statistical analysis in community research. In: Tolan, PH.; Keys, C.; Chertok, F.; Jason, L., editors. Researching community psychology: issues of theory and methods. Washington DC: American Psychological Association; 1990. p. 111-26.

41. Cicchetti D, Toth SL. The role of developmental theory in prevention and intervention. Dev Psychopathol 1992;4:489–93.

42. Lipsey, MW.; Wilson, DB. Effective intervention for serious juvenile offenders: a synthesis of research. In: Loebe, R.; Farrington, D., editors. Serious and violent juvenile offenders: risk factors and successful interventions. Thousand Oaks CA: Sage; 1998. p. 313-45.