

CONFLUENT EDUCATION AND EVALUATION RESEARCH

Jordan Horowitz and Joel H. Brown

One of the issues currently facing evaluation researchers is the need to address multiple levels of data collection, analysis, and interpretation using data and methods that are appropriately matched to those levels. For example, when evaluating community interventions an assessment of individual factors (e.g., academic achievement of students, socioeconomic status of families) fails to address many of the salient issues at the community level such as the sociocultural context within which the intervention is being implemented. In this chapter, we will:

- Show how the confluent education model can be used as a multilevel research tool providing a conceptual basis for evaluating human service programs.
- Show how this model has been applied in a recent drug prevention education evaluation.

Advances in Confluent Education, Volume 1, pages 113-141.
Copyright © 1996 by JAI Press Inc.
All rights of reproduction in any form reserved.
ISBN: 0-7623-0080-9

- Take a closer look at contemporary U.S. drug education in light of traditional and confluent evaluation models.

The unique contribution of this model is in making explicit those questions that will elicit a fuller understanding of the program under study than is currently available.

BACKGROUND: EVALUATION RESEARCH

At its simplest level, evaluation research is concerned with describing the structure of the program in question, the implementation of that program, and the extent to which the program has (or has not) achieved the stated goals. Additionally, the relationships among program structure, process, and outcome are explored. The ultimate goal of evaluation research, then, is to describe what service orientations work best under which conditions and for whom.

In a volume on assessing the quality of medical care programs, Donabedian (1980) defines the three areas of evaluation research: structure, process, and outcome. To relate them to a general evaluation of human services, we paraphrase them here:

1. Structure is, "the relatively stable characteristics of the providers of [services], of the tools and resources they have at their disposal, and of the physical and organizational settings in which they work."
2. Process, "the primary object of study is a set of activities that go on within and between practitioners and [service recipients]."
3. Outcome is, "a change in a [service recipient's] current and future status that can be attributed to antecedent [practices]."

Despite past differences among researchers over the appropriate focus of research in quality of care, these three components are not mutually exclusive. Each has meaning only in relation to the others; they are bound to one another by causal connections: structure leads to process resulting in outcomes (Donabedian, 1980; Wyszewianski, 1988).

THE CONFLUENT EDUCATION MODEL

In the confluent education model the evaluation researcher to makes explicit the implicit multiple dimensions of social interactive processes found in programs under study. Addressing multiple levels of analysis corresponding to three human interaction dimensions: (a) psychological aspects within the individual (intrapersonal); (b) interaction between individuals (interpersonal); and the social context of the organization (social contextual) one brings to the forefront the relationships within and among social processes and outcomes. In addition to examining program outcomes, the model emphasizes the social processes through which outcomes are produced; that is, program development and implementation (Goetz & LeCompte, 1984).

Making explicit those elements implicit in each dimension of a social system allows a new understanding to emerge; advancing both the methods and interpretation of evaluation research findings. Explicating each dimension supports attempts to understand how various aspects of people and social systems function singly and as part of a larger system. A confluent education approach toward program evaluation, enables us to provide clients with powerful and useful explanations of the hows and whys of program successes and/or failures. In practical terms, the confluent education model supports an analysis of the relationships within and among program participants and recipients at multiple levels.

Making explicit the elements of perception and interaction within and among the social contextual, interpersonal, and intrapersonal dimensions supports the emergence of a holistic view of program effectiveness. As described elsewhere in this volume, this conceptual framework has its roots in Gestalt theory. By taking a relative view along multiple dimensions, the confluent education model provides a framework for ensuring an evaluator considers the perceptions of those involved as essential to describing the program. Furthermore, since human perception is based on the integration of several elements and their interactions, confluent evaluators consider affective as well as cognitive statements important to evaluating a human service program.

Traditionally, for example, in qualitative evaluation research the cognitive domain of perception has been extensively mined while the affective domain is rarely examined beyond issues of client/

participant satisfaction. The value of assessing affective perceptions typically is overlooked and critical gaps regarding programs' potential for success remain unexplained. Along with integrating the intrapersonal, interpersonal, and social contextual factors, integration of cognitive and affective information is an essential aspect of completing the gestalt of program evaluation. We believe that to ignore the affective domain compromises the ability of evaluation researchers to fully understand program processes and outcomes.

INTEGRATING CONFLUENT EDUCATION AND EVALUATION RESEARCH

How might it look if researchers applied the confluent education model to traditional evaluation components? In Table 1 we pose questions that highlight elements of the program to be evaluated. These questions, operationally defined in an evaluation design, provide both researcher and client with a full understanding of the implementation and effectiveness of human service program(s). These questions also are designed to ensure that cognitive and affective information is available to the researcher.

We do not mean to suggest these are necessarily the only questions that may be relevant in any given cell. The converse might be asked (e.g., in the structure/social contextual cell *How does the program structure affect the social contextual dimension of human interaction?*). In some cases additional questions might arise (e.g., in the process/interpersonal cell *How do participants' interactions at the program development stage affect program implementation?*).

In the table, program participants are distinguished from the target population. The target population are those individuals for whom the services being provided are anticipated to have a positive effect. Participants are all those involved in developing and implementing the program including funding agents, developers, administrators, and service providers, as well as the target population.

In a confluent education evaluation model researchers seek to understand program structure, process, and outcome within and between each of the three human interactive dimensions. That is, the confluent education model holds that in order to locate overall program effects, we shuttle among emerging aspects within and

Table 1. Confluent Education and Evaluation Research Design

	<i>Structure</i>	<i>Process</i>	<i>Outcome</i>
<i>Social contextual</i>	What aspects of the social environment affect the program structure?	How does the social context affect program development and implementation?	How does the social context affect the target population?
<i>Interpersonal</i>	What aspects of social interaction affect the program structure?	How do program participants interact with each other and with the program?	How do program participants' interactions with each other affect the target population?
<i>Intrapersonal</i>	What is the program participants' perception of how the program is designed?	How does the program participant experience the program, as delivered?	How does the program participant's understanding and experience of the program affect the target population?

between the three dimensions. Eventually, a complete whole understanding of the program emerges. In some studies certain elements come to the foreground, while other aspects seem less important. In other studies many salient elements are present in each dimension. It is important to note that each evaluation aspect is not necessarily mutually exclusive. For example, there frequently is overlap between interpersonal aspects of structure and process. Furthermore, this overlap is not interpreted as confounding the findings of any one area. Areas in which aspects overlap may be the most fertile areas for relevant information.

RELIABILITY/VALIDITY AND PRACTICAL EVALUATION

Any time one examines the potential of a new methodological tool two issues emerge. First, to what extent are the methods underpinning the model reliable and valid? Second, is the model practical to use? The following section examines these two questions.

When examining an evaluation model the conscientious scientist examines reliability and validity issues. With one assumption in mind we do so here. We assume that a confluent evaluation model is best utilized when the researcher wants to explain program effects as they emerge directly from the data. This method is grounded in the data

more directly than in traditional evaluations, which often depend on interpolation. In qualitative research, for example, validity is established based on the extent to which the evaluation participants' social milieu is accurately described and accounted for. Here, systematic constant comparison is an essential process for establishing validity. Kirk and Miller (1986) point out that constant comparisons are particularly sensitive to discrepancies between the researchers' interpretation and meanings actually intended by study participants. Through systematic qualitative comparisons within and between interview data, validity is established. Here, the potential for establishing validity is especially enhanced because operationalizing evaluation questions like the ones above at multiple dimensional levels and in evaluation areas explicitly ensures multiple constant comparisons bound by similar phenomena. Multilevel and area comparisons relative to similar phenomena provide an increased ability to locate the most significant effects, where they are most likely occurring and how and why they are occurring. In the confluent evaluation model of systematic comparisons validity is enhanced.

To a large extent qualitative reliability depends upon validity. Normal questions of reliability might include, "Could these results be repeated if this study were administered in a similar fashion again?" (Kirk & Miller, 1986, p. 42). Such consistency assumes that the findings of the study were valid in the first place. In assessing reliability the dynamic aspect of the confluent model are especially important. Evaluators and clients alike can be apprised of assertions as they emerge, therefore they can be tested for reliability within the bounds of the current study. For example, in examining the process evaluation area an important assertion emerges. If the data are collected and analyzed in a stepwise fashion, the assertion can be tested on two other levels, post hoc and a priori. Post hoc, evaluators can examine process data as it relates to previously collected structure data. Consequently, early process and structure data can help the evaluator form reasonable a priori outcome hypotheses to be tested during the next step of the design and/or analyses. Also, this stepwise application of a confluent education model nicely lends itself to triangulation, confirmatory findings from multiple data sources (Jick, 1979). As you will see in our mixed method confluent evaluation example, by the time researchers were preparing to administer outcome survey data, the structure and process data pinpointed the most fruitful areas for outcome hypothesis testing.

We were able to hone our assertions so that fewer outcome assessments were necessary and sufficient to test key evaluation assertions. If rigorous systematic comparisons are taking place in a stepwise fashion, then it is likely that outcome findings from an alternative source will triangulate in support of earlier assertions.

A dynamic confluent evaluation model enhances reliability and validity in several important ways. First, through constant comparisons of data across multiple levels and dimensions, each bounded by similar occurrences. Second, through ongoing confirmation/disconfirmation of findings. Third, through the development of process and structure based a priori hypotheses that are then tested at the outcome level. When data are collected and analyzed in a stepwise fashion, each of these benefits can be accrued within the initial evaluation. Although this section has focused on qualitative research, and we will not further discuss reliability and validity (due to space limitations), many of these same premises and potential can be derived from a primarily quantitative confluent evaluation.

In addition to examining issues of validity and reliability when evaluating the utility of a confluent evaluation model, one might ask, "how practical is it?" After all, on its face it might appear that by comparing multiple dimensional levels and evaluation areas, we are calling for an excessive data collection and analysis procedure. Several researchers frame our response to these issues. Zelditch (1962) believed that in framing a research design, informational adequacy and efficiency are of utmost importance. Like validity, informational adequacy asks the evaluator to maximize the possibility that he or she will be able to adequately understand the social milieu.

In Zelditch's construct of framing a research design a most significant issue is one of efficiency, allowing the evaluator to collect informationally adequate data at the least cost (Marshall & Rossman, 1989). In fact, in the area of education there is an evaluation standard for efficiency:

The evaluation should be efficient and produce information of sufficient value, so that the resources can be justified.... An evaluation is cost effective if its benefits equal or exceed its costs. Benefits include... but are not limited to, publicly identifying... effective and ineffective programs projects, materials

and services; discovering how the monetary and nonmonetary costs of a program might be reduced without decreasing its services; and fostering understanding of activities and how they are perceived in a given setting and from a variety of perspectives (Sanders, 1994 p. 77).

Again, it is the dynamic nature of the confluent evaluation model that permits researchers to meet or exceed these evaluation standards. The nine questions asked in Table 1 allow for a great deal of flexibility, which in turn allows the researcher to be efficient. The flexibility arises from how each of the questions are operationally defined for each study, and how the data is to be collected and analyzed. For example, as noted earlier, depending on the program to be evaluated there may be overlap of dimensional levels. The key to efficiency is in making the operational definitions clear, for example, that certain data will be construed as reflecting multiple dimensional levels and evaluation areas. In a confluent model efficiency is achieved through making multidimensional subjectivity clear, and then linking it with rigorously defined and obtained outcome data. Multidimensional subjectivity is defined as rigorously obtained descriptors regarding evaluation phenomena (structure and process) at the intra- and interpersonal, and social contextual levels obtained from germane participants. Such definitions are not necessarily a priori. Sometimes multidimensional subjectivity is post hoc, and can elucidate an a priori hypothesis. For example, in a relatively low-budget evaluation, benefits over costs could be secured when data regarding perceived activities at the intrapersonal level in structure and process are collected in one collection opportunity. Given the need for a stepwise understanding of the data we recommend collecting outcome data on a separate occasion. While a pause between collection opportunities might at first seem inefficient, at least one other collection opportunity actually allows for an increase in efficiency because it makes explicit a step back for reflection to choose the appropriate operational definition(s) of outcomes. This reflection stage is an essential feature of the confluent education model. It is an explicit opportunity for the researcher to develop within-study awareness. It is this reflective stage, supported by valid and reliable subjective data, that provides the formative basis for outcome collection data and analyses. Again, a stepwise procession increases efficiency by allowing benefits to be understood during the study, thereby eliminating the often last minute need for

classic "hunts for significance" in outcome data. The key to efficiency is that in a stepwise way, the confluent evaluation model asks the researcher to make their rationale for collection explicit while in the throes of the study. Because researchers can look at the "big picture" while still in the study, the confluent evaluation model enhances efficiency without compromising validity or reliability.

For ease of presentation, the findings in the illustrative example are presented in an organized fashion based on the Table 1 matrix. We begin with the structural aspect of the social contextual dimension. We then address, in turn, the process and outcome aspects of this dimension before moving on to the structural, process, and outcome aspects of the interpersonal dimension and conclude with the intrapersonal dimension. (In discussing reliability and validity and evaluation standards, we have already described the process of how evaluators might proceed in a confluent evaluation.) By presenting results from an actual evaluation along with brief meta-comments on our thinking we will show how the most salient aspects of the findings come to the foreground, a clear progression of evidence unfolds, and a coherent picture emerges.

AN ILLUSTRATIVE EXAMPLE

To demonstrate the benefits gained from using the confluent education model in evaluation we turn to the field of preventing substance use and abuse among youth. We present a detailed account of results from a program evaluation in which each of the human interaction dimensions are considered, leading to the potential for program changes based on a comprehensive program description. Examples are taken from an evaluation of the California Drug, Alcohol, and Tobacco Education (DATE) programs. DATE represents the umbrella term for one of the largest school-based alcohol and drug education programs in the United States. From 1991-1994 an evaluation was conducted along three quantitative dimensions: cost, program implementation, and students' self-reported knowledge, attitudes, beliefs, and behaviors (Romero et al., 1993, 1994). We incorporate results from a survey of over 5,000 students in grades seven to twelve throughout California.

Additionally, in 1992 and 1993 a large scale qualitative evaluation component was undertaken (Brown & D'Emidio-Caston, 1995). We

include the findings from a subsample of nearly 400 educators, administrators, and community members. Most importantly, we use the findings from 40 focus groups conducted with close to 250 students interviewed regarding their knowledge, attitudes, beliefs, and behaviors regarding the drug education they received.

SOCIAL-CONTEXTUAL DIMENSION

In assessing the social context of a program the focus is on those aspects of the environment that influence the perceptions and/or behaviors of program participants. Program participants include those designing and delivering the program, as well as those receiving services.

Structure

What Aspects of the Social Environment Affect the Program Structure (e.g., Physical Plant, Staffing Patterns)?

Respondents from school and district administrators, faculty, and staff reported that the DATE program structure was negatively influenced by the methods of state fund disbursement and the limitations on its use. They related uncertainty about the availability of DATE program funds to their reluctance to hire personnel or plan additional program components. A majority of respondents at the district level, and many respondents at the school level, expressed the following concerns:

It is difficult for us to do advance planning. Last year we were told all the way through the summer, when we're supposed to have our budgets out and our planning done, that our tobacco money would be cut in half. Then, in the fall, after our budgets were done, we're told, well, it didn't get cut in half. So, it's just like we're looking right now, we're getting notified by TUPE [Tobacco Use Prevention Education Program] again that they may not give us any funding for next year. So, here we are in our planning time and we're having to sit and gamble and weigh the odds, are we going to have the money or are we not going to have the money? Then, at the same time, we come in and are told that you have to have the stamp by a certain time. Well, the proper way to spend money is to have planning time and something you can count on so that you put up a long-term, consistent program where you can go in and do something really worthwhile and not have to run out and spend

the money because you really only had it seven months that they told you what you're really going to get—how can you run a full-year program if you don't know what you're going to get (#228, pp. 1-2).

In evaluating the program this information comes to the foreground as creating a context for what types of programs may be implemented and how widely such services may be distributed across the state. These concerns lead to implementation of fiscally conservative services, of short duration, that can be applied to masses of students. We discern that in a bureaucratic structure, new programs are not likely to be initiated if funding is late, intermittent, or perceived as having a potential for being short lived. As it turned out, these administrators and practitioners were correct. Despite massive federal funding for these programs (over 1.6 billion dollars), program funding was diverted to different areas beginning in 1995. Here then, financial aspects of the social environment were found to affect the program structure.

Process

How Does the Social Context Affect Program Development and Implementation?

Another key aspect of the DATE program can be found in the funding application guidelines. Similar to federal guidelines, emphasis was placed on addressing the needs of at-risk students. Under the title *Philosophy and Purpose of DATE* it is stated:

Extensive research on risk factors offers a clear direction for prevention programs....In planning prevention programs, begin by reviewing the following list of risk factors and protective factors. Determine which risk factors are most significant in your school community. Then inventory the resources that might be available to you in reducing these risk factors and increasing protective factors. With this information you can formulate objectives and activities that are designed to deal with the most important problems facing your students (California Department of Education, 1991-1992 p. viii).

The application, itself, also emphasizes the importance of reducing risk factors for drug, alcohol, and tobacco use and other problem behaviors of youth. Department of Education officials proceeded to

identify 36 risk and 4 protective factors believed to be related to substance “abuse.” To gain DATE money, school districts showed how they would provide substance use and abuse prevention based on a risk factor model. The guidelines were clear: the development and implementation of prevention programs was to occur in the context of a risk and, to a lesser extent, a protective factor approach.

Thus, we find that the model to be used by people in schools is prescribed by the funding agency. This allows no opportunity for an approach developed from the bottom up. This becomes increasingly important as it is paralleled at the district level (discussed subsequently in the interpersonal area), precluding student involvement in program development.

Outcome

How Does the Social Contextual Environment Affect the Program Participants?

In the DATE evaluation, the fiscal and social context imposed by the state clearly had an effect on participants. Funding issues led to a great deal of concern about education in general, and the DATE program specifically:

I’ve been in this business 31 years and this is the most frustrating time that I’ve ever faced because I think we’ve got a lot of neat projects going and programs and I don’t know how to keep them running (#010, p. 8).

In the DATE program the risk factor model, which provided the context within which educators were to develop and implement prevention programs, encouraged a perception among a majority of educators that most students fit into this at-risk category.

We are addressing the risk factors that show up, with the idea that it’s real hard for me to point out which of our kids are not at risk (#014, p. 13).

From these brief quotes and myriad other data respondents reported how the fiscal and conceptual context affected the DATE program. Fiscally, participants were frustrated and reluctant to implement new and comprehensive programs. Conceptually, to gain funding participants spoke of a majority of students being at-risk for

substance abuse. Given the need among school districts for the resources DATE provided, adopting a risk factor approach was almost a necessity. Here, we have shown that both the structural and process related social contexts were linked with a contextual outcome that has implications for the DATE program along both the interpersonal and intrapersonal dimensions.

THE INTERPERSONAL DIMENSION

The interpersonal dimension is concerned with interactions among program participants. The reader is reminded that program participants include a number of groups beyond the target population (e.g., administrators, funding agents, program developers, et al.). Here we explore the effects of these interactions on the DATE program.

Structure

What Aspects of Social Interaction Affect the Program Structure?

When we looked at social interaction and program structure it became immediately apparent that the focal point was the structure of school districts, designed primarily for social interaction in a top-down or hierarchical way. Not surprisingly, the fact that these programs were school-based had a distinct effect on the structure of delivering drug, alcohol, and tobacco prevention programs. Following, respondents describe the DATE Program as being designed and implemented by few people who convey to others what services to deliver and how they should be delivered. This pattern of social interaction, also prescribed interpersonal relations among DATE participants:

Traditionally this district has had an extremely strong Board of Trustees, which sometimes because of their eagerness to do what is right and to influence decisions, moves over into what we call macro-management [sic], which causes site-level people and district-level people some frustration because that management from the board may be heavily influenced by the last person who talked to them (#379, p. 2).

In its prototypical state (of which we have many examples), the DATE program social network proceeds through many filters before a program is delivered to students: major elements of program development begin at a school board of trustees and/or the superintendent, are given to a district planner (either the person in charge of curriculum and/or the DATE coordinator), to the principal or vice principal, the program deliverers, and finally to the students. The job of those within the school is to comply with implementation of the DATE program prevention and intervention services. The hierarchical network structure constitutes the primary glue around which interpersonal interaction takes place.

Process

How do Program Participants Interact With Each Other and With the Program?

One of the key findings of the DATE evaluation was that the nature of the interactions between program providers and students was as important as the material being provided in determining program effects. Programs typically are delivered using teaching methods that are not effective in reaching students.

For example, recent research suggests that a didactic approach in which the program provider delivers a message to students with no opportunity for interaction is not effective (Tobler, 1992). This didactic approach also typically is noted as a problem by students:

I think the problem with education is this kind of education is that you're constantly being shoved down your throat it's so wrong, if you do it you're a terrible evil person, instead of just education saying I know some of you people do it, why do you do it, let's try and help you so you don't do it anymore (#530, pp. 17-21).

It's kind of like everybody knows that drugs and smoking are bad for you so it's not like a teacher can sit there and can pound it into you so you're not going to do it (#531, pp. 13-15).

Yeah, it's so boring, just like the other school classes, I mean, but we already know about it and that's not going to help us. We already know about it and it's not going to do anything (#531, pp. 28-29).

The credibility of the instructor also is called into question. In the survey of 5,045 students, statewide, it was found that 30 percent of

students disliked the people who deliver DATE services (23.4% disliked them *a lot*) and another 39.4% had a neutral attitude. This left less than 30 percent liking the service provider. Along the same lines, 40.9 percent of students reported not being affected *at all* by the people who deliver DATE services. Students overwhelmingly wanted to hear the “real story” from those who experienced it:

Yeah, but the health teacher doesn't really know, you know.

Oh yeah, the health teacher doesn't know, he's reading from the book.

Yeah, he's just reading from the book and if they had brought someone in that knew and that like went through it, I think it would be a lot better (#531, pp. 13-15).

Here, we also can note the importance of the affective response of students. This is important because research literature shows that neutral and/or negative affect toward someone who is trying to influence another's behavior is likely to lead to a lack of, or short-term, compliance, here construed as not using drugs because of the drug education students receive (e.g., Brown & Raven, 1995; Raven, 1965, 1983, 1993).

Non-classroom programs such as Red Ribbon Week, school assemblies, and contests also are universal. As one school district person noted:

When we do these things we have T-shirts we've made up that we give to the kids, number one as a reward for doing it for us, number two to promote what we are doing. We have poster contests, essay contests, and we give out a zillion awards to the kids (#275, pp. 13-14).

These activities, apparently designed specifically to address the students' affective response to prevention programs, fail because the students' affective response is tied to a desire for effective prevention practices and not simply more of the same. These activities also are not developmentally appropriate. In many of the high school focus groups students mentioned they wanted more harm reduction programs, smaller groups rather than assemblies, and more counseling. Despite receiving virtually the same programs the high school students' educational needs are different than, for example, the needs expressed by elementary school students. As students developed, most expressed the desire for more content of drug

information including both sides of an issue, delivered through an educational process that includes experiential panels and talks by those who have been through both substance use and substance abuse, in a manner that respects their growing sense of maturity.

Considering the negative affect linked to interactions among program participants, what kinds of outcomes can one expect?

Outcome

How do Program Participants' Interaction With Each Other Affect the Target Population?

Let us continue with findings from the student survey to examine this dimension. Students were asked the extent to which their decision to use or not use was due to the classes and activities in their schools. While 15.7 percent of the students responded "a lot" or "completely," nearly three times that amount, or 43 percent (2,169 of 5,045 students surveyed), responded *not at all*. Furthermore, students took responsibility for their own use with 58.5 percent responding to another item that their decisions regarding use were either *a lot* or *completely* due to themselves. Thus, we find the target population remained relatively unaffected by their interactions with program providers.

One of the more important findings of the DATE evaluation was that the target population often felt judged as deviant. Students described a concern that they or their peers were too often removed from the school system instead of served. Not surprisingly, the target population of students exhibit decreased attachment to school:

If it is shoved into you that you're a terrible person when you do this, you know, you kind of want to back away from the education process because they've already made a judgment upon you (#530, pp. 17-21).

This issue is at the heart of the failure of the risk factor model. Districts receive prevention funds based on the numbers of youth-at-risk; yet these youth eventually leave school or are removed from the system. The explicit message students describe receiving from educators is the desire to help them understand drug issues and help those who need it. However, by the time students reach high school they perceive an implicit message that runs counter to the explicit

message. On an affective level, students describe a negative response from school personnel that leaves many healthy adolescents feeling like deviants; this, simply for contemplating substance use choices and their potential range of responses. They are left wondering why their peers with substance abuse problems are being exited from the system without support or help. As noted by students:

They are not in this for helping you, they are in [it] for getting rid of the bad kids and just having all good kids at school.

If they suspect you of smoking or having drugs on you or whatever, if they see a kid like that in their school then, instead of suspending them and getting them out of school, why don't they help them? (#531, p. 21).

This system of removing youth in need of services instead of assisting them is noticed by some school district personnel, as well:

We still get rid of too many kids... those are the kids that the state of California and the United States of America have identified as their target population... the kids that are at risk the most, are the kids that are exited from the system, and they do not have access to the resources.... The kids that we need to keep in and provide resources to are the kids that we exit from the system (#558, p. 18).

In examining the interpersonal dimension, then, we find that the hierarchical structure of the school district is the defining factor in the way prevention services are delivered. It is a top-down system with little chance of student influence. The interpersonal process, framed within the risk factor model is antithetical to current research in teaching practices and is experienced by the students as ineffective and incredible. Finally, the result of these interactions are poor school bonding and exclusion of those most in need from the services that might be available to them.

THE INTRAPERSONAL DIMENSION

The intrapersonal dimension is concerned with the personal experiences and perceptions of program participants. Although important among all three dimensions, it is here that the respondents' affective statements (i.e., those concerned with feeling and experience) play an essential evaluative role.

Structure

What is the Program Participants' Perception of How the Program is Designed?

Many respondents described being frustrated with the hierarchical nature of planning and implementation of the DATE program. When asked how DATE programs should be developed and implemented, one respondent replied:

What should be is a program set up so the control of the situation, the choice of things to do, appropriate objectives, should be developed at site level. Yes, with a district coordinator, but it should be developed at site level, by the people who are requested to be a part of the program (#213, p. 9).

By saying "...it should be developed at the site level," this respondent was confirming the dominant view expressed by many program participants that the DATE program is not developed on-site (i.e., at the school). The prescriptive nature of these comments indicates the respondent was dissatisfied with the status quo and was telling us what he would like to see. This individual would like to see more control and interaction at the school site.

Another respondent's beliefs about the effects of legislative action on program design reveal conflicts between what needs to happen and what actually happens in program design:

...I think some of what legislature and folks who make the decisions about money do is they play the money game, they dangle the money out there and say "here's the latest program, the latest idea, who is going jumping through the hoops to get the money?" And, people go jumping through the hoops and they get the money and then they have got to implement the program based on that. Sometimes the issues are so much larger than that, that the money doesn't really cover it, so you end up doing much more than what is just involved in one targeted area. Yet, you're also the most vulnerable in the sense that if the legislature decides that this is not the latest issue and they want to go from drug and alcohol abuse to battered children or some other issue, then that is where the new money is going to be and now all the sudden everybody has got to switch hats. I think that's a serious mistake. We're talking about really providing services for kids with the issues of drug and alcohol abuse, children who are physically abused, sexually abused, etc. We are out to provide those services and stop playing the games about, "Well here's the biggest problem" (#009, p. 8).

Process

How Does the Program Participant Experience the Program as Delivered?

The hierarchical model of interaction coupled with the mandated assumptions about at-risk students engenders one primary type of program orientation this educator describes:

R: I'd like to think the message is drug abuse is life abuse, which is our slogan and everyone knows that slogan. I think the message is that drug, alcohol, and tobacco, substance abuse or substance use, is not acceptable...

I: So use equals abuse for any of those substances?

R: Yes (#292, p. 5).

In the data we repeatedly found examples of educators delivering their intrapersonal point of view that any substance use equals abuse. This educator experiences the program as necessitating one message, one point of view which is to be delivered to students. And how does the target population report experiencing the program?

When students described their individual experiences of the program the results were no surprise. Many middle school students described their own experience of the program in the following way:

R: I think it's nothing! It's exaggeration!

R: They lie to you so you won't do it!...

R: Oh, they lie to you so that you won't do the drugs! They think you're dumb!

I: Do you think that works?

R: No. [laughs]

I: Do you think that's what they really do?

R: Yeah, sometimes (#508 p.10)

Across students there was general doubt of the veracity of educators. This was associated with the beliefs that were brought to the process through the hierarchical nature of the interaction; namely that most students are at-risk and that any substance use equals abuse.

Another aspect of the experience of the program, as delivered, is the implementation of policy. A universal policy, noted in responses from elementary, middle, and high school students, is that students who use are suspended or expelled:

- (1) R: A couple of girls in my class right now they were smoking last year...[and] they got caught smoking in the rest room... and they got suspended for like three days (elementary school student, #562, p. 12).
- (2) R: All I know, from what I know they get expelled. That's about it (middle school student, #552, p. 10).
I: Have any of your friends been busted?
- (3) R: Yes. Right now they're up for expulsion right now. To be expelled and then they would have to go to another school, I guess (high school student, #506, p. 13).

This is in contrast to the general belief among students that the consequences of use were not supportive of students needing interventions for substance use related problems:

- (1) R: If they suspect you of smoking or having drugs on you or whatever, if they see a kid like that in their school then, instead of suspending them and getting them out of school, why don't they help them? (#531, p. 21).
- (2) R: Yes, because I do know, I had a friend who had a real rough time trying to go straight!
I: Cause there wasn't any help available?
R: There, at this school, no! (#593, p. 8).

Here, the target population experienced the program as punitive rather than supportive. Next, we see how these experiences translate into limited opportunities for significant effective impact.

Outcome

How Does the Program Participant's Understanding and Experience of the Program Affect the Target Population?

As one would expect, by this time, students are left feeling inadequate about themselves and distrustful of the education system.

When asked about how they feel about the programs they received this student responds:

Depressed. Because if he's about talking to us about drugs and alcohol and all these kind of things he should come out with those, you know, he should talk to us the right—you know, with the whole thing, not just say a little bit and then just leave the rest behind (#568, pp. 8-9).

Because this student felt that DATE service providers did not come out with “the whole thing,” he said that it made him feel “depressed.” Both the survey data and interview questions addressing affective issues (about how the DATE program made students feel) represent an explicit examination of intrapersonal outcomes which triangulate with one another. Many students described wanting more than “a little drug education” indicating that this was what they perceived they were receiving and they walked away dissatisfied.

Well, like I said, he'll be talking about something and then when you like ask for more information he like really doesn't want to come out like and tell us the whole thing (#568, pp. 8-9).

Furthermore, students are affected by what they experience as discomfort on the part of the service provider. That is, the nature of the interaction leads the target population to certain beliefs about those providing information. This, in turn, results in the student turning inward or to experimentation:

R: I also wanted to say that what I think that part of the problem seems to be is that any subject the teachers or anybody is really uncomfortable talking about it seems to all get pushed to the side and I think that more with drug and alcohol abuse, I think that it is something that people are very uncomfortable talking about.

I: You mean adults?

R: Yes, adults. I think that, especially, I guess, when you're in high school or junior high, also, when somebody says to you don't do this, because it's bad, you know, the automatic reaction of a teenager growing up is to say why is it bad? Everybody is curious about what this does, the way they're curious about anything and I think that by just saying that it's bad and just putting things in the curriculum is doing that without saying this is why, this is what can happen, and showing what can happen to you and presenting both sides realistically

and when you don't do that then people don't know what the consequences are and are more apt to stray (#530, pp. 17-21).

DISCUSSION

In the past, researchers have struggled with conceptual frameworks allowing them to examine somewhat limited aspects of social systems (Lewin, 1956). While valuable in showing general relations of individuals to groups, current social network models do not take into account many of the subtleties and complexities of relations within and among levels in the social system. Moreover, little if any of this research has been applied to evaluation research.

Through examination of school-based drug education, we illuminated the conceptual framework and process of conducting an evaluation grounded in the confluent education model. Research indicates that explicating multiple dimensions of social interactions offers a more complete basis for understanding the dynamics of a social network than has previously been considered (Knoke & Kuklinski, 1982). At the heart of this kind of evaluation is an examination of all three dimensions prescribed by the model: the intrapersonal, interpersonal, and social contextual. If performed in a stepwise fashion, within well-defined operational parameters, this model represents much more than simply asking nine basic questions. Confluent based evaluations are reliable, valid, and efficient; issues of concern to evaluators.

A confluent evaluation perspective and model is a necessity for other reasons related to the link between main tenets of confluent education, social policies, outcomes, and social problems. As we have shown, confluent education evaluators build into their methodology a dynamic view of the programs under study. In our evaluation example, we showed how several apparently unrelated factors, for example, a hierarchical organizational structure and a risk factor model are actually concomitant. Through a limited communication network (hierarchical), and the money to go along with it, a program with equally myopic possibilities (the risk factor view of adolescents and their substance use) for solving a problem, can be promulgated. In our example then, what are the implications of these concomitant relationships? First, the maintenance of a social and political dynamic between researcher and funder accompanied by the public

appearance that something is being done about the drug problem. When, for example, researchers comply with funders' requests to abandon the scientific method, to demonstrate the effectiveness of a model instead of test a model's effectiveness, they become confederates in a tainted process. Through an incomplete picture of programs and their effects, a second implication arises from the first: ineffective and at the same time, expensive, drug education programs. A confluent evaluation model asks the researcher to step back and reflect so that apparently disconnected factors can be connected. In short, for a society in which there are decreasing dollars with which to deal with social problems, the implications are unmistakable and significant. It is a contaminated relationship between we as researchers, policymakers, and the public which allows for highly visible, yet ineffective programs. This relationship is somewhat, if not highly, responsible for the exacerbation of social problems.

But it is difficult to really understand these implications unless we complete the gestalt; we cannot appreciate the implications of using a confluent model until we focus on these institutionalized patterns so often found in drug education policy and research. Next, we discuss how traditional research modalities linked with policy decisions, exacerbate social problems when two key aspects of the confluent model are ignored: gaining awareness and taking responsibility.

In the larger research effort from which the examples in this methodological paper were drawn, we showed how U.S. school-based drug education researchers neglected to evaluate student perceptions regarding the drug education they received (i.e., the intrapersonal dimension). We will show how ignoring this dimension has had detrimental effects on those very individuals targeted for these human services, thereby exacerbating the social problem of substance use.

In the 1970s and early 1980s in the larger social context of U.S. drug education there were attempts to understand and educate students on the differences between substance use and abuse (NIDA, 1981). Experimentation among adolescents, while not promoted, was seen somewhat as a dialectic of growth (Jessor & Jessor, 1977). Students were not merely recipients of drug education, they were viewed as participants in the development and implementation of such programs. For example, the National Institute of Drug Abuse (NIDA) recommended that teachers be able to, "assist students in learning how to weigh the consequences of possible decisions they could make on drug issues" (1975, p. 18). However, in the middle

1980s the social context in the United States shifted. Researchers and policymakers alike changed their perception of adolescents to one of deviants in need of clear and tough policies (Brown & Horowitz, 1993). For example, in 1989 the White House National Drug Control Policy stated, "unlike some previous school-based approaches, resistance training takes a firm moral stand that using drugs is wrong and should be resisted Any student caught selling or distributing drugs is immediately expelled" (p. 50). Adolescents as deviants received help most often in the form of discipline. When, as we have found in our research, this represents both policy and a view of adolescents is there any doubt that this linkage between policymaker and researchers will preclude the involvement of the target population, adolescents, in the prevention effort? This in turn relieved the prevention researcher of the responsibility to examine the nature of adolescent interaction, and how students perceived these programs. Here, awareness of the intrapersonal dimension was neglected, and the very target population, with whom researchers were ethically responsible to work, students, were virtually excluded from evaluations of such programs. As a consequence, we found that prevention research was based almost exclusively on the outcome of adolescent substance use, inadequately examined, yet most often explained by risk factors and peer pressure (Brown & Horowitz, 1993). These explanations, more indicative of the necessity for researchers to maintain relationships with funders than anything else, were not tied to predictive data.

A closer examination of the dominant researcher perspective over the past two decades reveals a paradox. Why were awareness and responsibility neglected then and now? The answer lies in further specifying the kinds of student data absent during that period in relation to researcher's perspectives. With the exception of the California DATE evaluation, the affective responses of students and prevention specialists have been largely ignored in the evaluation of prevention programs. In fact, affective responses are missing from many evaluations. House (1994) notes, "Research methodology depends primarily on the nature of the subject matter of the discipline, the content, the object of what one is trying to investigate" (p. 14). When the "object" of study and its subject matter are as affect-laden as in prevention education it is logical for the evaluation researcher to attend to the affective responses of the participants. Frequently though, the affective response is dismissed as too subjective to be of

value. Shinn addresses the subjectivity of responses, "Multiple sources yield multiple truths. The researcher's job is to understand and to model the individual and extraindividual sources of variance" (1990, p. 118). When the affective response is found to be generalizable and is shown to have an effect on outcomes, as well, the subjective perspective is a valid and reliable instrument. The paradox then: researchers are charged with evaluating the programs, yet in the case of prevention research, a highly volatile field, perhaps the most germane information, visceral affective responses of students toward these programs, are virtually ignored as a confound. Traditionally, when dealing with confounding variables, researchers have a choice, include the variable or exclude it. Over the past twenty years, it is precisely the exclusion of these subjective affective responses of the target population incorrectly identified as a confound to be excluded that has allowed the promulgation of ineffective drug education programs. If the student's voices were present in this research, then it would be difficult for researchers and policymakers to claim that they were unaware of their perceptions. As included in the DATE evaluation it becomes immediately evident that student's concerns about program ineffectiveness are real; if grounded in the data, their voices cannot simply be dismissed as teenage rebellion. Most importantly then, making explicit multilevel affective responses forces researchers to explain outcomes derived directly from the data. To the extent that valid and reliable systematic comparisons are made, is also the extent that the target population's affective responses are considered meaningful explanatory factors. Explaining outcomes then shifts from rationalizing expectations and political realities to the validity of accepting perception, and the similarities among many truths. In other words, here, awareness and responsibility means representing rigorously obtained affective perceptions and linking them with relevant outcomes. Sadly, in prevention education these multiple truths have been substantially absent.

Regarding outcomes, being aware and taking responsibility also means being rigorous on a different level. At the outcome level, while considerable descriptions of program structure and process are available, there is little evidence of the effectiveness of prevention programs in the research literature. In the case of drug prevention education, the major outcome is substance use rates; and they continue to climb among youth (Johnston, O'Malley, & Bachman,

1994), yet prevention programmers continue to do more of the same. If the proposed model were in use, there would be equal focus on outcomes. Poister (1982) called this approach *performance monitoring*. That is, "the most effective use of information gathered through monitoring procedures requires that comparison be made between the data found and some standard of effectiveness" (Posavac & Carey, 1985, p. 157). Are, for example, reductions in the rates of increase in use (Pentz et al., 1990) a good outcome? Standards of effectiveness describing good outcomes of prevention programs must be developed. These standards must be agreed upon by the research and prevention community and then programs and evaluators must be held to them. It should be noted, that we are not talking about an absolute invariant standard nor an absolutely variable standard. We are saying that an appropriate outcome standard could be devised, one that is based on the multiple truths bounded by similar phenomena as previously described. If that reasonable standard is not met, then we realistically examine the nature of the phenomena and if necessary, reasonably adjust our premises and standards. For example, in light of the American desire for substance use, unreasonable standards include a drug free state or nation by the year 2000. As we alluded to earlier, such contaminated outcome standards have been linked with the dynamic relationship between researchers and policymakers (Baizerman & Compton, 1992; Brown & Horowitz, 1993; Placier, 1993). While contaminated standards help people to maintain the dynamic system between researcher and funder, and keep the public satisfied that the social problem is being addressed, some research has revealed another effect of unreasonable outcome standards; yet again, such standards actually serve to exacerbate social problems. Publicly attractive but useless outcome standards (a drug free society by the year 2000) actually permit ineffective work to continue, while the social problem worsens. In the case of drug prevention education it is by poorly educating those who may be experimenting with substances and by excluding those who really need help (Brown & D'Emidio-Caston, 1995). In addition to ignoring critical research elements, it is now evident that maintaining contaminated outcomes exacerbates social problems.

We maintain that partly due to incomplete examination of the social system of U.S. school-based drug education policy and practices, programs shifted more toward that which did not work. Such programs were doomed because program evaluations failed to

address some critical issues. First, although they may have been part of the programs, affective intrapersonal descriptions of student perceptions of the programs were rarely taken into account. Related to this, the social context in which drug education took place also shifted without being included in evaluation research. Finally, because of their contaminated linkages with the political milieu, standards of effectiveness are unreasonable and unobtainable. Thus, when evaluation research failed to examine the full social system, and utilized unreasonable outcomes the ultimate result of their suggestions was unsuitable program shifts with potentially negative consequences for the target population.

Given the context of rising social and political concerns, an accompanying myopic research which supported those concerns but did not adequately explain the social problems, the ways in which such programs are shifted now seem plausible. And as the social conditions continue to worsen, creating different relationships and different ways of evaluating may become more palatable.

In sum, a confluent education evaluation model rigorously conducted makes multiple ways of making meaning explicit, thereby enlarging the scope of rational choice. Despite the "yang" of the contaminated relationships between researchers and funders, there is also a "yin." There may be a new readiness to embrace the use of such a model. The multilevel view presented here offers increased potential for validity, reliability, and efficiency, always an interest of evaluation researchers. At the same time, especially when it comes to substance use, an increasingly concerned public and body politic alike may be willing to embrace rigorous confluent based evaluation(s). An evaluation process which illuminates from multiple perspectives will come to represent the first steps toward being aware, taking responsibility, and making reasonably informed change. If we do not attempt to appreciate multiple ways of making meaning, then we have resolved ourselves to only one choice, the exacerbation of our social ills.

REFERENCES

- Baizerman, M., & Compton, D. (1992). From respondent and informant to consultant and participant: The evolution of a state agency policy evaluation. In A. M. Madison (Ed.), *Minority Issues in program evaluation* (pp. 5-16). *New Directions in Program Evaluation*, no. 53. San Francisco: Jossey-Bass.

- Brown, J. H., & D'Emidio-Caston, M. (1995). On becoming at risk through drug education: How symbolic policies and their practices affect students. *Evaluation Review*, 19(4), 451-492.
- Brown, J. H., & Horowitz, J. E. (1993). Deviance and deviants: Why adolescent substance use prevention programs do not work. *Evaluation Review*, 17(5), 529-555.
- California State Department of Education (1991-1992). *Healthy kids, healthy California: drug, alcohol, tobacco, education (DATE) district application for funding*. Sacramento, CA: California State Department of Education.
- Donabedian, A. (1980). *Explorations in quality assessment and monitoring: Vol. 1. The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Administration Press.
- Goetz, J.P., & LeCompte, M. (1984). *Ethnography and qualitative design in educational research*. Orlando, FL: Academic Press.
- House, E. R. (1994). Integrating the quantitative and qualitative. In W.R. Shadish, C. S. Reichardt, & S. F. Rallis (Eds.), *New directions for program evaluation: No. 61. The qualitative-quantitative debate: New perspectives* (pp. 13-22). San Francisco, CA: Jossey-Bass.
- Jessor, R., & Jessor, S.L. (1977). *Problem behavior and psychosocial development: A longitudinal study of youth*. New York: Academic Press.
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24, 602-661.
- Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (1994). *National survey results from the Monitoring the Future Study, 1994*. Ann Arbor, MI: Institute for Social Research.
- Kirk, J., & Miller, M. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: Sage Publications.
- Knoke, D., & Kuklinski, J.H. (1982). *Network analysis*. Thousand Oaks, CA: Sage Publications.
- Marshall, C., & Rossman, R. B. (1989). *Designing qualitative research*. Newbury Park, CA: Sage Publications.
- Pentz, M. A., Trebow, E. A., Hansen, W. B., MacKinnon, D. P., Dwyer, J. H., Johnson, C. A., Flay, B. R., Daniels, S., & Cormack, C. (1990). Effects of program implementation on adolescent drug use behavior: The midwestern prevention project (MPP). *Evaluation Review*, 14, 264-289.
- Placier, M. L. (1993). The semantics of state policy making: The case of "at-risk." *Educational Evaluation and Policy Analysis*, 15(4), 380-395.
- Poister, T. H. (1982). Performance monitoring in the evaluation process. *Evaluation Review*, 6, 601-623.
- Posavac, E. J., & Carey, R. G. (1985). *Program evaluation: Methods and case studies* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Raven, B.H. (1965). Social influence and power. In I.D. Steiner & M. Fishbein (Eds.), *Current studies in social psychology* (pp. 371-382). New York: Holt, Rinehart, Winston.
- Raven, B.H. (1983). Interpersonal influence and social power. In B.H. Raven & J.Z. Rubin (Eds.), *Social psychology* (2nd ed., pp. 399-444). New York: Wiley.

- Raven, B.H. (1983). The bases of power: Origins and recent developments. *Journal of Social Issues, 49*(4), 227-251.
- Romero, F., Carr, C., Pollard, J., Houle, D., Brown, J., Gaynor, J., Fleming, T., Flaherty, J., Martino, T., & Karam, R. (1993). *Drug, alcohol, and tobacco education evaluation (DATE): Second-year interim evaluation report*. Prepared by the Southwest Regional Laboratory for the California Department of Education. Los Alamitos, CA.
- Romero, F., Bailey, J., Carr, C., Flaherty, J., Fleming, T., Radio-Gaynor, J., Houle, D., Karam, R., Lark, M., & Thomas, C. (1994). *California programs to prevent drug, alcohol, and tobacco use among in-school youth: annual evaluation report*. Prepared by Southwest Regional Laboratory for the California State Department of Education. Los Alamitos, CA.
- Sanders, J. R. (1994). *The program evaluation standards*. Newbury Park, CA: Sage.
- Shinn, M. (1990). Mixing and matching: Levels of conceptualization, measurement, and statistical analysis in community research. In P. Tolan, C. Keys, F. Chertok, & L. Jason (Eds.), *Researching community psychology: Issues of theory and methods* (pp. 111-126). Washington, DC: American Psychological Association.
- Tobler, N.S. (1992). Drug prevention programs can work: Research findings. *Journal of Addictive Diseases, 11*(3), 1-27.
- Wyszewianski, L. (1988). Quality of care: Past achievements and future challenges. *Inquiry, 25*, 13-22.
- Zelditch, M. (1962). Some methodological problems of field studies. *American Journal of Sociology, 67*, 566-576.