ELSEVIER

Evaluation and Program Planning 28 (2005) 247–256

EVALUATION and PROGRAM PLANNING

www.elsevier.com/locate/evalprogplan

Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (D.A.R.E.) program

Sarah Birkeland, Erin Murphy-Graham, Carol Weiss*

Harvard Graduate School of Education, Harvard University, 467 Gutman, Appian Way, Cambridge, MA 02138, USA

Abstract

D.A.R.E. is the most popular school-based drug abuse prevention program in the U.S., but evaluations have found that positive effects on students' knowledge, attitudes, and behavior (often observed right after the program) fade away over time. By late adolescence students exposed and not exposed to the program are indistinguishable.

Some school districts ignore the evidence and continue to offer D.A.R.E. In our study of 16 school districts, we found two persuasive reasons: (1) Evaluations generally measure drug use as the main outcome, but school officials are skeptical that any low-input short-term program like D.A.R.E. can change adolescents' drug-taking behavior. (2) Evaluations often do not often report relationships between cops and kids. Improvement in these relationships is a main reason for many districts' continued implementation of D.A.R.E. Districts also mention other understandable although more problematic rationales for keeping D.A.R.E.

Keywords: Evaluation impact; Knowledge utilization; Prevention; Substance abuse; Education; D.A.R.E.

1. Introduction

When evaluations are conducted according to scientific canons, evaluators assume that people should pay attention (e.g. House, 1980; Patton, 1997; Pawson & Tilley, 1997; Rossi & Freeman, 1993). When decision makers do not listen to the evidence provided by good evaluations of programs, evaluators assume that they are deficient in something. Perhaps the fault is lack of attention: they did not hear the results. Perhaps the fault is lack of understanding: they did not grasp the message. Or most likely, the fault is purposeful avoidance: an over commitment to their program coupled with unwillingness to hear contrary news (Cohen, 1979; House, 1993; Lindblom, 1990; Lindblom & Majone, 1988; Reimers & McGinn, 1997). When policy makers do the exact opposite of what the evaluation suggests, such as cling to a program that evaluation has repeatedly found wanting, their sins appear magnified. Yet it is possible that under some circumstances, they have good reasons for doing so.

The counter-intuitive message of this paper is that there may be cases when ignoring evaluation evidence makes sense. This unexpected conclusion emerged from a study of the influence of evaluations of the D.A.R.E. program on school district decisions about which drug abuse prevention program to run in their schools. This study, supported by the Robert Wood Johnson Foundation, centered on districts' responses to evaluation findings and the competing influence of other factors on their decisions.

By all accounts, D.A.R.E. is the most popular school-based substance abuse prevention program in the United States. Statistics provided by D.A.R.E. America show the program, developed in 1983 by the Los Angeles Police Department and the Los Angeles Unified School District, was being used by more than 80% of school districts in America by 2001.[1] In 2001, George W. Bush became the third sitting President to issue a Proclamation for a National Day for D.A.R.E., and 2003 was the 13th consecutive year such a proclamation was signed.[2]

* Corresponding author. Tel.: +1 617 495 4144; fax: +1 617 496 3095.
E-mail address: carol_weiss@gse.harvard.edu (C. Weiss).

0149-7189/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.evalprogplan.2005.04.001

[1] D.A.R.E. is also now used by at least one school district in over 40 other nations (D.A.R.E. America website at www.dare-america.org).

[2] D.A.R.E. is also recognized in the popular culture; for example, the football movie 'Any Given Sunday' had a scene showing the team owner giving a check for $200,000 to the local D.A.R.E. program.

However, ask anyone you meet on the street what he or she knows about D.A.R.E. and they will likely tell you that it does not work. Evaluation evidence, showing that D.A.R.E. is not effective in preventing adolescent drug use, has been widely covered in the news media and popular press. The D.A.R.E. program has been evaluated many times across a variety of contexts. Evaluations have included a number of long-term randomized experiments with large samples and long follow-up periods. These studies report consistent findings: no statistically significant effect for D.A.R.E. on self-reported drug use.

In 2000 we began the Study on Decisions in Education: The Case of D.A.R.E. We wanted to find out what influenced the decisions that school districts made about drug abuse prevention and specifically about whether or not to implement the D.A.R.E. program. What role did evaluation evidence play? We found that the evaluation evidence, through a variety of channels, did influence decision making about the D.A.R.E. program in the majority of districts we studied. However, a handful of districts were dismissive and wary of evaluation evidence. Despite the bad press D.A.R.E. received, they were determined to continue the program. In this paper we discuss their reasons.

We found that individuals in these districts were dismissive of evaluation evidence for several reasons. First, they had never expected D.A.R.E. alone to prevent adolescent drug use; therefore the news that it did not was no surprise. Second, they believed that evaluators 'missed the boat,' focusing their studies on the wrong outcome measures. The most valuable outcome of D.A.R.E., according to these respondents, is the relationships it fosters among police, families and schools. Yet most evaluation studies neglect that outcome. Finally, decision makers valued their personal experience with the program as more convincing than scientific evidence. They believed that their program was unique and their D.A.R.E. officer exceptional. Some of this might sound like ex post-rationalization; nevertheless, several important lessons for evaluators can be drawn from the case of D.A.R.E.

2. Background and research context

2.1. The D.A.R.E. program and evaluation evidence

D.A.R.E. was developed in Los Angeles in 1983 to bring police officers into elementary school classrooms (usually 5th or 6th grade) for about an hour a week for one semester. The officers provide information about drugs and the consequences of their use, and they teach means for resisting peer pressure to use drugs, concepts of self-confidence, and decision making skills. D.A.R.E. was embraced by school districts around the country. When we began the study, most school districts were implementing the D.A.R.E. program in one form or another, and most still do. Although no reliable data exist on numbers of districts, D.A.R.E. America reports that 70–80% of school districts run the program.[3]

However, a growing body of evaluative evidence has shown that the D.A.R.E. program is ineffective in preventing drug use among adolescents. In the early 1990s, the National Institute of Justice funded an influential meta-analysis of recent, rigorous evaluations (Ennett, Tobler, Ringwalt, & Flewelling, 1994). The researchers found minimal effects for D.A.R.E. in preventing adolescent drug use (Ennett et al., 1994). The authors concluded, 'D.A.R.E.'s limited influence on adolescent drug use behavior contrasts with the program's popularity and prevalence' (Ennett et al., 1994: 1399).

Subsequent studies reported similar findings: although several studies showed positive effects on knowledge and attitudes, they did not show statistically significant effects for D.A.R.E. on self-reported drug use. Such was the case with the randomized experiments conducted in Illinois (Rosenbaum, Gordon, & Hanson, 1998), Colorado (Dukes, Stein, & Ullman, 1997) and Kentucky (Clayton, Cattarello, Anne, & Bryan, 1996). In the latter three cases, the evaluations were conducted by well-known investigators with considerable experience conducting evaluations of drug prevention programs. Another well-known investigator, Denise Gottfredson, conducted a comprehensive review of school-based programs designed to reduce delinquency or drug use. She concludes, 'Evaluations show that as it is most commonly implemented, D.A.R.E. does not reduce substance abuse appreciably' (1997:16).

The evidence that students exposed to D.A.R.E. fared no better than students without D.A.R.E. was widely disseminated. Interviews with evaluators aired on the CBS and ABC nightly news. Study results appeared in scores of national and local newspapers including the New York Times and the Boston Globe, and in weekly periodicals such as The Chronicle of Higher Education and US News and World Report. The dissemination was so widespread that by 2000, even people without any connection to education or health care, without children in the schools, knew the gist of the evaluation findings. D.A.R.E. America claimed that the majority of school districts continued to use their program, despite the negative evaluation results.

This apparent contradiction seemed a strategic opportunity for understanding the considerations driving school decision making. If decision makers were not attending to good evidence, what was going on? Why were school districts seemingly unreceptive to the increasingly clear evidence that students exposed to D.A.R.E. in the 5th or 6th grades were no more likely than other students to stay clear of drugs as teenagers?

[3] Cited on D.A.R.E. America website at www.dare-america.org.

2.2. Ignoring evaluation evidence

Evaluators generally agree that decision makers do not usually put evaluation findings directly into use (Caplan, 1977; Husen, 1994; Weiss, 1980, 1982). They learn from the evaluations, they take them into account in setting priorities and planning, they get background knowledge and ideas from the evaluations, but they do not regularly use them as a basis for immediate decisions (Caplan, 1977; Knoor, 1977; Patton et al., 1977; Weiss & Bucuvalas, 1980). Recently Henry and Mark (2003) and Kirkhart (2000) have suggested that discussions of the consequences of evaluation should be phrased in terms of the 'influence' of evaluation rather than its 'use,' echoing Weiss's discussions two decades earlier (1980, 1982).

However, some writers on evaluation point out that when evaluators take special steps to enhance the influence of their study, they often have good effects. Patton (1997), for example, has emphasized 'utilization-based evaluation,' a plan for conducting studies in ways that give decision makers a strong say in how the questions are defined and the study is managed, as well as a major investment in disseminating results. His definition of effect is 'intended use by intended users,' and his work suggests that evaluators can often have more influence than is true in the general run of evaluation experience.

Moreover, there are cases in which evaluation has had important consequences. The influence has been true not only for local-level evaluations but also for national and international studies. Well-known instances of influential evaluations include the STAR experiment on class size in Tennessee (Mosteller, 1995), the Perry Preschool Program (Berutta-Clement, Schweinhard, Barnett, Epstein, & Weikart, 1984), and the Mexican welfare program PROGRESA (Kruger, 2002).

But why do not decision makers ordinarily apply evaluation results directly to decisions? Is that not why they requested evaluation in the first place-to help them make better decisions about the intervention they run? The answer is: not always. Sometimes evaluations are undertaken to satisfy requirements for receiving a grant. Sometimes evaluations are imposed on programs by superordinate authorities, such as federal departments, while decisions about how the program runs are made locally. But even when decision makers have themselves requested evaluation, often with an intention to apply results to decisions, things often go awry. Researchers of evaluation use (or evaluation influence) have found a number of obstacles that can block the path.

Lipton (1992) found that the quality of evaluation studies is sometimes suspect, or even outright poor. Weiss (1998) notes poor conduct of studies, inadequate interpretation of results, and intrusion of evaluators' biases. Decision makers therefore, do not find the results worth attending to. Caplan (1977) found that evaluators and decision makers have different interests, languages, timing, and worldviews. He found the differences so wide that he dubbed them 'two communities.' Becker (1984) notes that evaluation results do not always point to obvious actions. Decision makers have to interpret the results in terms of the situation on the ground and figure out what to do about them. Petersilia (1987) states that before evaluation results become available, the situation under study may have changed. The results are no longer relevant. Husen (1994) is one of the scholars who blame ineffective dissemination for the lack of evaluation use. These and many other factors have figured prominently in the literature.

In our study we started from the premise that evaluation had not had direct consequences on decision making. By continuing D.A.R.E. in the face of the evaluation evidence, schools districts were ignoring the evidence. We found several reasons why. The first reason is intuitive; people in a number of school districts disagreed with the results. They explained their dismissal of evidence by arguing that the research could not be valid because what happens in one context is not generalizable to another. This finding is consistent with prior research (Weiss, 1980, 1993, 1999) which shows that decision makers are more likely to pay attention to results that confirm what they already believe (Weiss, 1980, 1993, 1999). When results contradict a pre-existing position the decision maker is more likely to dismiss these findings than to change her mind. Research can more easily strengthen pre-existing beliefs than alter them.

We also found that people dismissed evaluation evidence because it addressed only official program goals. The notion that evaluations should evaluate not only the official goals of the program but also be responsive to stakeholders' other interests and expectations has a long history in the evaluation literature (Rossi & Freeman, 1993; Weiss, 1993, 1999; Patton, 1997). Almost every text on evaluation advocates attention to the interests of audiences beyond the office that asks for the study. They variously mention program staff, clients, program managers, program funders, policy makers, and sometimes the larger community. It is a commonsensical idea that evaluators should attend to people's concerns, especially if the evaluator wants the results to have an influence on decisions.

Another reason for the negative pall of evaluation results is that studies have accepted bloated promises and political rhetoric as authentic program goals (Weiss, 1993). Many of the individuals we interviewed thought that the goals of D.A.R.E. measured by evaluators were inflated. It seemed unreasonable to expect D.A.R.E. to prevent drug use entirely. If evaluators want decision makers to pay attention to their results, programs should have more modest expectations and they should be evaluated against more reasonable goals.

Finally, decision makers valued the D.A.R.E. program for reasons the evaluations did not usually consider. Evaluation results are not likely to be persuasive to those for whom other values have a higher priority (Weiss, 1999). Using an example from our study that we will describe in

further detail below, if a decision maker thinks that the D.A.R.E. program will keep adolescents from experimenting with drugs, she will take the negative evaluation findings seriously. However, if she is satisfied that the D.A.R.E. program develops better relationships between children and the police, then drug use data mean less.

3. Data sources and methods

In seeking to understand why some people ignored evaluation evidence in the case of D.A.R.E., we chose to focus on a purposeful sample of 16 communities in four states. We first chose the four states—Colorado, Massachusetts, Kentucky, and Illinois—based on the fact that a large-scale evaluation of D.A.R.E.'s effectiveness had been conducted in each of those states. Within each state, we selected four towns or districts[4] in which to conduct research. We sought state-by-state consistency on two levels: first, we limited selection to communities with populations between 40,000 and 200,000; second, in each state we selected two communities that reported implementing the D.A.R.E. program at that time, and two that did not.

In 2001, we travelled to each of the sixteen communities to conduct in-person, semi-structured interviews with respondents. We used a snowball technique to identify respondents, by first contacting the school district administrator responsible for prevention, and then asking for the names of other people to interview. Respondents included school district personnel involved in selecting prevention programs and community members who were influential in decisions about prevention. Because D.A.R.E. is staffed by law enforcement personnel, and usually paid for by them too, we also interviewed police officials who were involved in implementing and making decisions about the program. We conducted brief follow-up telephone interviews with officials from each district in the early spring of 2003. In total we conducted 128 interviews.

The interviews were conducted by members of our research team, following a semi-structured protocol. They lasted between 45 minutes and 3 hours. We tape recorded and fully transcribed the interviews in almost all cases, except in those rare instances when respondents refused to be taped or the machinery malfunctioned.

Using our research questions and preliminary analysis as a basis, we developed an initial coding scheme for the interview data. We coded transcripts using the qualitative data analysis software package Atlas.ti, and checked for inter-rater reliability. Based on the coded data, we created a case study for each community, outlining themes in the decision making process and local-level officials' responses to evaluation evidence. We then examined the cases for cross-cutting themes. Finally, in order to verify those themes, we returned to the coded interview transcripts and gathered evidence in support of, and in contradiction to, our arguments.

A table listing the (pseudonymous) communities and the status of the D.A.R.E. program in each at the time we conducted the interviews appears below. You will notice that of the eight communities that were not implementing D.A.R.E. at the time we selected them, six had once implemented the program and decided to discontinue it. Of the eight that were implementing D.A.R.E. at the time we selected our sample, two dropped the program between our initial interviews and the follow up interviews we conducted in 2003. In another paper we explore communities' decisions to discontinue D.A.R.E. and the factors that influenced those decisions (Weiss, Murphy-Graham, & Birkeland, 2005) Table 1.

4. Findings

Having analyzed interview data concerning decisions about the D.A.R.E. program in sixteen communities, we have a much richer understanding of why so many seem to have ignored the evaluation evidence. Six of the eight districts that had D.A.R.E. in 2001 continue to use the program in spite of negative evaluation results. Even in

Table 1
Pseudonymous communities in our sample ($N=16$) listed by the status of the D.A.R.E. program at time of data collection

Status	Community name (pseudonym)
Never had D.A.R.E.	Carlsburgh, MA
	Hatsfield, CO
	Total, 2
Had once implemented D.A.R.E., but had dropped the program prior to first interview (2001)	Orchard Grove, MA
	Westview, CO
	Danville, IL
	Gardner, IL
	Marlboro, KY
	Princeton, KY
	Total, 6
Had D.A.R.E. in 2001, but dropped it by the follow-up interview (2003)	North Fork, KY
	Riverton, CO
	Total, 2
Still have D.A.R.E.	Clovertown, IL
	Dover, KY
	Trimble Falls, KY
	Cartersville, CO
	Hilltown, MA
	Cedar Point, MA
	Total, 6
	Grand total: 16

[4] In collecting the data for this paper, we interacted with officials from two domains: schools and police departments. In some cases, the police departments' jurisdictions were the same as the school districts'. For example, both might serve the population within the city limits of a small town. However, in other cases, the police served the town while the school district covered the entire county. Sometimes the implementation of D.A.R.E. was city-wide, sometimes district-wide, sometimes county-wide (and sometimes, of course, not at all.) Therefore, in this paper, we refer to the communities studied as 'towns or districts'—there are some of each.

communities in which the D.A.R.E. program was terminated—often due, in some part, to the negative evaluations—the decision to end the program was unpopular; many people we interviewed remained supportive of the program and skeptical of negative evaluation evidence.

In the cases of the eight communities that were implementing the D.A.R.E. program when we began this study in 2001, we found that most decision makers knew about negative program evaluations, though few had actually read the studies. Respondents had read about the studies in the newspaper or heard about them on TV. During several interviews, school and police officials showed us saved newspaper clippings that reported negative study findings. In general, however, they still supported the D.A.R.E. program. This was also true of many school and police officials even in the communities that had recently stopped implementing the D.A.R.E. program.

4.1. They never expected D.A.R.E. to prevent drug use

Negative evaluations of D.A.R.E.'s effectiveness at keeping kids off drugs were not surprising, nor particularly noteworthy, to many respondents. These school and police officials believed that the evaluations measured unrealistic, inflated goals. They shared a belief that no one intervention is strong enough to counter the drug pressures in society; the idea that a one semester, one hour a week curriculum would prevent future drug use in adolescents struck them as naïve. School and police respondents argued that D.A.R.E. could do some good by outlining the perils of drugs and urging students to take responsibility for making sensible decisions, but it was not likely to blot out all the other influences on young people. They pointed to the cultural appeal of drugs and the influence of peers, television, and even some families in maintaining a drug-friendly culture. Against all these powerful pressures, how could one reasonably expect 17 hours of instruction to push back the tide?

The school superintendent in a Massachusetts town that had just voted to discontinue D.A.R.E. summarized this view in expressing his disapproval of the decision:

> The one thing that we were very clear about, even if the national organization [D.A.R.E. America] wasn't, was that we never portrayed the D.A.R.E. program when we adopted it as something that was going to prevent kids from using drugs and alcohol. At best, you know, we would have presented all of our programs, in 10th grade and 8th grade and other programs as maybe helping out along the way, but again, it would have been a pretty lofty claim that when you spend a few hours in a classroom of the school you are going to change kids' behavior.

A school official in a Colorado district that was also discontinuing D.A.R.E. expressed a similar sentiment. The program was, in her view, a small part of a larger, community-wide effort.

> Part of the thinking with any drug and alcohol program here was that if D.A.R.E. were done as one very small component of a much larger community initiative, that it would be good value added. But I don't think anybody here ever saw it as *the* drug and alcohol education program. It's a tiny program when you think about it. It's a very minuscule part of the program.

Respondents asserted that substance abuse is a societal problem with deeply entrenched causes. A school official in a Massachusetts community that still uses D.A.R.E. said,

> When you have kids living in a culture where drugs are in their houses, where drugs are being used by their parents and their parents' friends on a recreational basis, where it's in television and in movies, it's very hard for them to stay on that slippery slope. I don't think the problem is the D.A.R.E. program, I think the problem really is the societal emphasis on drugs, and so, I don't think any program is going to sustain 100% or even 80% success, when you're dealing with a culture that just doesn't take it seriously.

A D.A.R.E. officer in Massachusetts expressed the belief that no matter what, some kids are going to use drugs, and no drug prevention program will be able to prevent drug use entirely. A school board member in the same town agreed that expecting D.A.R.E. to alter students' rates of drug use is a 'silly goal.' Moreover, she said, experimenting with drugs at some point during the teenage years is not a catastrophe for young people; the important thing is help keep them as safe as possible.

Local officials also chafed at the expectation that one curriculum could solve a problem they see as family and community-related. A long-time officer in Riverton's now defunct D.A.R.E. program reacted to media attention about D.A.R.E.'s negative evaluations. "They can tell me D.A.R.E. doesn't work. But my question to them is 'Is it D.A.R.E.? Is it parents? Is it community?' It's not a curriculum that's going to necessarily prevent a kid from drinking or using drugs... It's the whole community."

These decision makers saw no reason to discontinue the program, just because D.A.R.E.'s stated goals are inflated or unrealistic. However, the fact that the evaluations measured unrealistic goals was one of the reasons why participants did not take their findings seriously.

4.2. Evaluators 'missed the boat': D.A.R.E. works because of the relationships that it builds

In focusing heavily on the 'official' goal of the program, the prevention of drug use, the D.A.R.E. evaluations missed

one of its most important achievements. According to our interviews, what D.A.R.E. is really good for is building relationships that might not otherwise develop. D.A.R.E. fosters good personal relationships between students and their families on one side, and law enforcement officials on the other. The officers who implement the D.A.R.E. program, and the school officials who invite them, value those relationships.

A Massachusetts superintendent captured this widespread sentiment about the true value of the program:

> If you ask the question 'Does it reduce crime?' that is, 'Does it reduce the illegal use of drugs and alcohol?' apparently D.A.R.E. can't demonstrate that for various reasons. If you ask, 'Does it help kids understand their community better? Does it produce favorable relationships between police and kids?' all of the survey results—and a number of states have done this—publish positive results.

When his community voted to discontinue the D.A.R.E. program because of evaluation evidence, he was disappointed in that decision, calling it 'a tempest in a teapot.'

School-level respondents in several communities also remarked that the program has changed the way children view police. As a school official in a Colorado town where D.A.R.E. continues with great support explains,

> Police are often looked at as the bad guy, or the one that's going to come in and get you for being a bad guy, and I think that D.A.R.E. provides an opportunity for our young kids particularly to find out that officers can be a resource for protection, for answers for some questions, for direction and for care.

A school official in one Kentucky community said that he introduced the D.A.R.E. program with the goal of improving police officers' image: 'Kids assumed police were pigs.' He wanted to expose children to police officers in a consistent, positive way. More than anything, he wanted D.A.R.E. to help children see police officers as human beings. He is certain that D.A.R.E. has worked in this regard. He proudly commented, 'You never hear the word pig in school hallways now.'

These relationships encourage and rejuvenate police officials, those often in charge of finding budget money to support the program. As the Safe and Drug Free Schools coordinator in a Kentucky town explained.

> If you go into a school and it's a middle school where an officer hasn't been in a long time, and you're walking down the halls, I mean, kids are hanging out the doors. 'Hi Mr. Hack!' They just love it… They let an officer walk in and talk to them and they just —it's nice if the kids have a positive relationship with an officer and it's good for the officer too, I mean, they don't get positive things every day.

Several police officials we interviewed remarked upon D.A.R.E.'s success at improving their public images; they explained that D.A.R.E. has made them seem 'more human' in the eyes of children in the community. Children gain respect not only for their D.A.R.E. officers, but for other police officers as well. Police officers see this as a reason to continue implementing the program, even in the face of negative evaluation evidence. As an officer in one Massachusetts town that continues to implement the D.A.R.E. program explained, spending time in schools also helps police officers do their work. Because they get to know children in a casual, non-threatening environment, later interactions proceed more smoothly.

> I'll be out working something and I'll have to have my hat on and sometimes people don't recognize me with my hat, and I'll start confronting them and they start giving me a hard time. But then as soon as they realize I'm Officer Brown, 'oh, Officer Brown', and they're nice and they listen to what I have to say and stop what they're doing. I think some of [the other police officers] are like, 'oh, the kids wave with all fingers now instead of just one!'

A Kentucky officer who helped one of his D.A.R.E. students cope with a difficult family situation remarked that statistical analyses of rates of adolescent drug use can never capture the value of such interactions. 'Say that was the nightly news, that would be a human interest story. Not something that you put on a bar graph or a pie chart or anything like that.'

In addition to the relationships that D.A.R.E. fosters between police and children, our respondents report that D.A.R.E. has also strengthened ties between local schools and police. Bringing D.A.R.E. officers into schools has improved the communication and teamwork between police and school departments. A Kentucky D.A.R.E. officer reports that this is the most important benefit of the D.A.R.E. program: 'Without a doubt, the main benefit is the exposure that law enforcement has to the community and the kids, the positive relationship that is set up between the school district, the students and the law enforcement agencies.' A Massachusetts police chief agrees: 'One of the most important benefits and by-products is the relationship we have now with the school department. It couldn't be better… it really couldn't be better. If I need anything, I just have to pick up the phone.'

Several school officials also commented on the benefits associated with improved police–school cooperation. For example, a school official in Colorado explained that when her district had D.A.R.E., they saw its prevention message as secondary and believed its main benefit was improved police–school relations: 'Even though we touted loving

D.A.R.E., we really loved having the officers in the school. We saw lots and lots of gains in that. The most important benefits came from police–school relationships, and the clear prevention message of D.A.R.E. was merely a side benefit.' After the Columbine High School shootings of 1999, school officials in at least one Colorado district relied on the D.A.R.E. officers, who had solid relationships with the students, to help restore a sense of safety and calm. D.A.R.E. officers lined up outside the local high school door, creating a corridor for students to pass through as they entered the building for the first time after the shootings.

Local government reformers often advocate closer working relationships among city agencies, particularly agencies that serve the same clientele. With the D.A.R.E. program in place, schools and police report cooperating in serving children. In the eyes of some respondents, evaluations of the D.A.R.E. program 'missed the boat' because they failed to capture one of D.A.R.E.'s main benefits: improved relationships between kids and police officers and law enforcement agencies and schools. Because of this, the evaluations are less useful and relevant to those who participate in and manage the program.

4.3. Personal experience is more convincing than scientific evidence

As police and school officials across the sites in our study discussed their attitudes toward evaluations of the D.A.R.E. program, many described skepticism of the generalizability of research findings. D.A.R.E. may not have been effective in the communities in which the studies were conducted, they reasoned, but no two communities are the same. These school and police officials typically believed that D.A.R.E. was 'working' in their towns, and they were unwilling to allow the evaluations to change their minds. Some collected data themselves to confirm their hunches. For example, a school official in a Massachusetts town that maintained the D.A.R.E. program simply stated, 'I remember one of the articles [reporting D.A.R.E.'s ineffectiveness] was on the basis that they were reviewing kids in the 6th grade, and kids weren't retaining any of the teachings. Well, I don't find that to be true in Cedar Point.' His survey data convince him that drug use among local youth is slowly declining, and he finds that data more credible than the studies he has read. Another Cedar Point school official says that D.A.R.E. 'is not without controversy in parts of the country, but I just sort of see the way it's handled here in the city and what we do with it. I think it's a positive experience for the kids.'

An active community member in Riverton, Colorado opposed the town's decision to discontinue the program based on negative evaluation evidence. She, too, articulated skepticism about the findings of large-scale studies, saying, 'A lot of what works locally, works locally. And I think when you're looking at evaluations, you know, meta-analyses, et cetera, you have to see what's happening locally in relationship to those things.' Her admonition to 'see what is happening locally' was raised by many other respondents; their cautions about what to look at centered on two topics: how the curriculum is being implemented, and who is teaching it.

Respondents from school districts and police departments in all four states dismissed negative evaluation evidence about D.A.R.E. on the grounds that different officers teach the curriculum differently. A former Colorado D.A.R.E. officer remarked, 'Sometimes I would read criticism of the program, and I would look at the criticism, and it would almost look to me as if they're talking about something different. Because what they were criticizing was not what I was doing.' Despite the fact that D.A.R.E. America mandates fidelity to the curriculum, and all D.A.R.E. officers are rigorously trained for standardized delivery, he believes that negative results in other communities must be the result of differences in how the curriculum is taught. He is sure that as he teaches it, D.A.R.E. 'works.'

This D.A.R.E. officer also reported variation across the state in how different police departments implement D.A.R.E., saying

> I really think [the approach to D.A.R.E.] varies from department to department. You go to [a nearby town], who does kindergarten through high school, and they swear by this program. And other departments just do little pieces of it. But it's just different. It varies. It depends on where you go. So, a nationwide study, I don't think is really valid. I mean, you can't say what happens in Ohio, or some other place, is the same as what happens here, because it varies on how you present it.

He was quick to dismiss evaluation findings on the grounds that the differences in his community's approach to D.A.R.E. are the key to its success relative to that in other communities. Other respondents made similar assertions, noting differences in their communities' timing or style of delivery that set them apart from those studied in evaluations. For example, a school official in Cedar Point, Massachusetts, pointed to unusually strong norms of communication between schools and police as the reason D.A.R.E. works in his community. He reasoned that studies finding D.A.R.E. to be ineffective must have been conducted in communities in which schools and police were not communicating well.

Perhaps not surprisingly, respondents in different communities across all four states described their local D.A.R.E. officers as outstanding, clearly more invested in student success and skilled in delivering the curriculum than D.A.R.E. officers elsewhere in the country. One school official in Riverton explained why it was a mistake to discontinue the program on the basis of evaluations: "I think our community was kind of blessed in the sense that we had

really quality D.A.R.E. officers. And so we would try to tell that to folks, and say, 'You know, I don't know what Cleveland D.A.R.E. officers were like. I don't know how well trained they were, I don't know how well they presented the material to the class.'"

A D.A.R.E. officer in Kentucky believes that his personal commitment to keeping youngsters off of drugs made D.A.R.E. an effective prevention tool in his community. 'Do I think I saved kids from doing drugs? Yes. Do I think some of the other people teaching the D.A.R.E. program after I left the program did? No… I think they just did it, put their time in and went home… I think passion has a lot to do with everything.'

Participants' experiences in the program and the anecdotal information they have regarding its success are more convincing than the findings of evaluation. They are less likely to believe study results than what they have observed through first-hand experience in the program. Disagreeing with the findings, they argue that study results cannot be generalized to their context.

5. Discussion

The reasons that district officials offer for ignoring evaluation evidence appear to be a combination of rationality and rationalization. While D.A.R.E. does not do what it was marketed to do, it does bring real benefits to their communities—benefits they value even in the face of pressure to drop the program. By citing other advantages of the program, they are putting a good spin on pervasive negative research findings. Still, their decisions to continue implementing it are based in an assessment of the pros and cons, rather than simple ignorance. Most district-level decision makers knew about negative program evaluations and the widespread publicity those evaluations received, but many still supported the D.A.R.E. program. We learned that because these people did not agree with the research findings they were less likely to heed them. It would be naïve to expect that evaluation will change the minds of people who have considerable experience with the program, especially in the short-term.

Furthermore, it is hard to avoid the conclusion that evaluations might be more influential if they took better account of stakeholders' concerns, expectations and perspectives. We learn from this case that some programs that cannot prove positive effects on stated goals are valuable to communities for other reasons. When that is the case, community members are likely to ignore evaluation evidence. They may make decisions that look illogical or uninformed, when in fact their logic, and the information on which they are basing decisions, is clear and explicit. When evaluators conduct research—even sound, scientific research-without exploring the goals and values of program implementers, they may be disappointed to find that their work has little influence.

Our findings also raise the question of whether D.A.R.E.'s intended goal of entirely preventing adolescent drug use is inflated. Several individuals working with the program thought so. A number of individuals we interviewed frankly explained that they never expected D.A.R.E. to prevent drug use. They believed that some kids are going to experiment with drugs in their teens. Other kids will use drugs more regularly, and no drug prevention program will be able to prevent drug use entirely. Therefore, they prioritize other program benefits.

Police officers and school officials do see a value in the 'connectedness' that D.A.R.E. fosters between children and police and police and schools. One D.A.R.E. officer from Colorado expressed a concern that such benefits may not be quantifiable: 'I don't know that you can put numbers to it, you know, when we start talking about relationships… I don't know how you measure that, and how you put that on paper when it comes to justifying manpower.'

His perspective also illustrates that individuals derive their understandings and policy preferences from a variety of sources. In this case, personal experience with the program and anecdotal information about its impact were more convincing than evaluation studies. Evaluation evidence is not always convincing to those who have first-hand experience with a program. In many cases, D.A.R.E. officers and school principals relied on local, informal gauges of prevention to justify their continued implementation of the program.

We also learned that people ignored evaluation evidence because they did not believe that the findings were generalizable to their context. This is a reason that gives us pause. Crediting the evidence of one's own senses is reasonable and certainly widespread. Yet if evaluation and other forms of research are to have a beneficial influence on policy and practice, people need to understand and privilege sound scientific results. Evaluators have to explain that programs can achieve good things and save some kids, but research is about probabilities. With sound sampling procedures, evaluators can collect data that generalizes to the whole population from which the sample was drawn—within the limits of sampling error. Idiosyncratic cases balance out. Overall, sound evaluation demonstrates what the chances are that young people will be better off because of the program. Can all districts be like Lake Woebegone and have outcomes above average? We chose districts from states that had well-known evaluations. If evaluation shows that programs in that state had non-significant effects, will people pay attention whatever the quality of their own D.A.R.E. program and their own D.A.R.E. officers?

6. Conclusions and lessons for evaluators

New federal legislation such as the No Child Left Behind Act, which ties school-based drug prevention program

funding to scientific evidence by incorporating the Department of Education's Principles of Effectiveness, will make it more difficult for people to ignore evaluation evidence. Nevertheless, it is important to understand why evaluation evidence is dismissed so that evaluators can be responsive to these reasons.

We found that individuals in these districts were disinclined to follow evaluation findings for several reasons. The most compelling were the first two: (1) They believed that the evaluations focused on the wrong outcome measures. The most valuable outcome of D.A.R.E., according to these respondents, is the relationships it fosters among police, families and schools. Yet, so far as they knew, the evaluation studies neglect that outcome. Of course, their knowledge of the studies was superficial. What most people knew was the diluted message delivered by the media. Most of them had not read the original studies or any scholarly article about them.

(2) They had never expected D.A.R.E. to prevent adolescent drug use. It was too frail an intervention to alter adolescents' developmental strategies. Why would you expect one hour a week of instruction for one semester to change young people's knowledge and behavior four or five years down the road? They supported D.A.R.E. as part of a wider, community-based prevention strategy. The idea that D.A.R.E. alone could be an effective prevention intervention seemed naïve.

We find these reasons sensible. Evaluators need to look at outcomes that people on the local scene value. They cannot accept the grandiose goals enunciated by programs as the only reality. True, articulating grandiose goals is a technique used by program developers to market programs, but evaluators should know enough about the programs to use realistic outcome measures. As program people shift their goals over time, evaluations need to be flexible enough to study the revised goal outcomes. Findings about attainment of more modest goals may not get the same play in the press and on television, but these kinds of findings matter to many people. They make sense.

Our conclusions are based on 16 non-randomly selected school districts in four states and deal with one program. While suggestive, our findings cannot be the basis for firm 'lessons' either for decision makers or evaluators. Moreover, although we did not highlight the fact in this paper, six of the eight districts in the 'No D.A.R.E.' category had dropped the program prior to our study, and two of the eight districts that were running D.A.R.E. at the start of our study had abandoned it by 2003 (See Weiss et al., 2005). Evaluation evidence had percolated into the consciousness of key decision makers, and it had an influence. Not all the reasons for dropping D.A.R.E. had to do with evaluation, but evaluation was often a part of the reason. Evaluations were having an impact. However, given the depth of support for D.A.R.E. shown in our study, it is possible that the conversion to other programs will be half-hearted.

References

Becker, H. (1984). The case of Germany, experiences from the Education Council. In T. Husen, & M. Kogan (Eds.), *Research and policy: How do they relate?* (pp. 121–129). New York: Pergamon Press.

Berutta-Clement, J. R., Schweinhard, L., Barnett, W., Epstein, A., & Weikart, D. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age*. Ypsilanti, MI: High/Scope Press.

Caplan, N. (1977). A minimal set of conditions necessary for the utilization of social science knowledge in policy formulation at the national level. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 183–198). Lexington, MA: Lexington-Heath, 183–198.

Clayton, R., Cattarello, Anne, M., & Bryan, M. (1996). Johnstone the effectiveness of drug abuse resistance education (project D.A.R.E.): 5-Year follow up results. *Preventative Medicine*, *25*, 307–318.

Dukes, R. L., Stein, J. A., & Ullman, J. B. (1997). Long-term impact of drug abuse resistance education (D.A.R.E.): Results of a 6-year follow-up. *Evaluation Review*, *21*(4), 483–500.

Ennett, S. T., Tobler, N. S., Ringwalt, C. L., & Flewelling, R. L. (1994). Resistance education? A meta-analysis of Project D.A.R.E. outcome evaluations. *American Journal of Public Health*, *84*(9), 1394–1401.

Henry, G. T., & Mark, M.M (2003). Mark beyond use: Understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation*, *24*(3), 293–314.

House, E. R. (1980). *Evaluating with validity*. Beverly Hills, CA: Sage.

House, E. R. (1993). *Professional evaluation: Social impact and political consequences*. Thousand Oaks, CA: Sage.

Husen, T. (1994). Educational research and policy making. In T. Husen, & N. Postlethwaite (Eds.), *International encyclopedia of education* (pp. 1857–1864). Oxford: Pergamon Press, 1857–1864.

Kirkhart, K. E. (2000). Reconceptualizing evaluation use: An integrated theory of influence. In V. J. Caracelli, & H. Preskill (Eds.), *The expanding scope of evaluation use* (pp. 5–23). San Francisco: Jossey-Bass, 5–23.

Knoor, K. (1977). Policymakers' use of social science knowledge: Symbolic or instrumental?. In C. H. Weiss (Ed.), *Using social research in public policy making*. Lexington, MA: Lexington-Heath.

Kruger, A. (2002). Putting development dollars to use, south of the border. *New York Times*. Section: Economic scene. May 2.

Lindblom, C. E. (1990). *Inquiry and change: The troubled attempt to understand and shape society*. New Haven: Yale University Press.

Lindblom, C. E., & Cohen, D. K. (1979). *Usable knowledge: Social science and social problem solving*. New Haven, CT: Yale University Press.

Lipton, D. S. (1992). How to maximize utilization of evaluation research by policymakers. *Annals of the American Academy of Political and Social Science*, *52*(1), 175–188.

Majone, G. (1988). Policy analysis and public deliberation. In Robert B. Reich (Ed.), *The power of public ideas*. Cambridge, MA: Ballinger.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of children*, *5*(2), 113–127.

Murphy-Graham, E., Weiss, C. H., Birkeland, S. (Forthcoming). *An alternate route to policy influence: Evaluations affect D.A.R.E.* Forthcoming. American Journal of Evaluation.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new-century text*. Thousand Oaks, CA: Sage.

Patton, M. Q., Grimes, P., Guthrie, K., Brennan, N., French, B., & Blythe, D. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (Ed.), *Using social research in public policy making*. Lexington, MA: Lexington-Heath.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Thousand Oales, CA: Sage.

Petersilia, J. (1987). *The influence of criminal justice research*. Santa Monica, CA: Rand Corporation.

Reimers, F., & McGinn, N. (1997). *Informed dialogue: Using research to shape education policy around the world.* Westport, CT: Praeger.

Rosenbaum, D. P., Gordon, S., & Hanson, S. (1998). Assessing the effects of school-based drug education: A six-year multilevel analysis of project D.A.R.E.. *Journal of Research in Crime and Delinquency*, *35*, 381–412.

Rossi, P., & Freeman, H. (1993). *Evaluation: A systematic approach.* Newbury Park, CA: Sage Publications.

Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, diffusion, utilisation*, *1*, 381–404.

Weiss, C. H. (1982). Knowledge utilization in decision making: Reflections on the terms of the discussion. *Research in Sociology of Education and Socialization*, *3*, 17–41.

Weiss, C. H. (1993). Where politics and research evaluation meet. *Evaluation Practice*, *14*(1), 93–106.

Weiss, C. H. (1998). Improving the use of evaluations: Whose job is it anyway? *Advances in Educational Productivity*, *17*, 263–276.

Weiss, C. H. (1999). The interface between evaluation and public policy. *Evaluation*, *5*(4), 468–486.

Weiss, C. H., & Bucuvalas, M. (1980). *Social science research and decision-making*. New York: Columbia University Press.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: Evidence from a study of the Drug Abuse Resistance Education (D.A.R.E.) Program. *American Journal of Evaluation*, *26*, 12–31.