



SLAPPING LIPSTICK ON THE PIG (PART II)

“Proving” that crime-control strategies work is laden with pitfalls

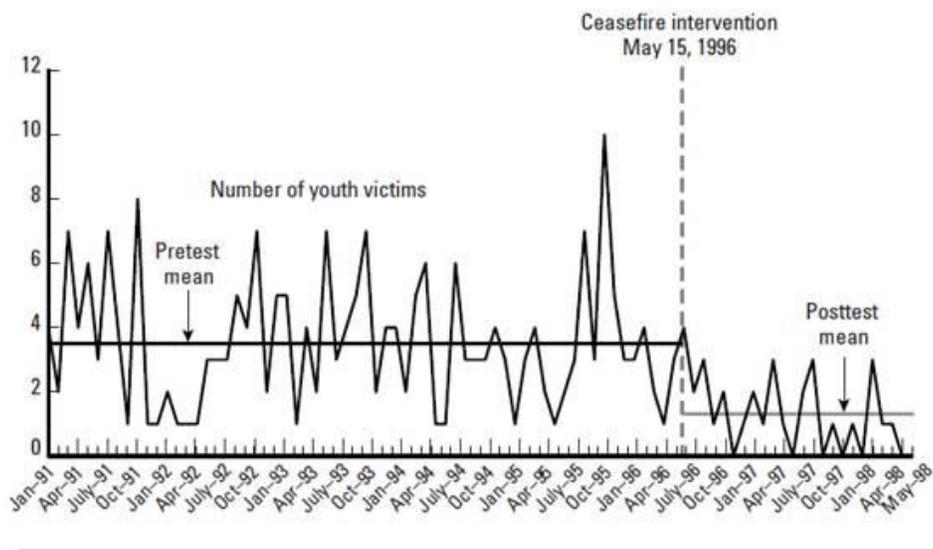
By Julius (Jay) Wachtel. In August 2005 the prestigious journal *Criminology & Public Policy* published “[Did Ceasefire, Compstat, and Exile Reduce Homicide?](#)”, an analysis by Richard Rosenfeld and two colleagues from the University of Missouri-St. Louis of three celebrated violence reduction programs: Boston’s Project Ceasefire and Richmond’s Project Exile (both discussed in [Part I](#)) and Bill Bratton’s Compstat, a program that began in New York City and spread throughout the U.S.

Each program was widely credited with success. But according to the authors none had been satisfactorily evaluated. Using sophisticated statistical techniques, they sought to determine whether declines in homicide in Boston, Richmond and New York City went significantly beyond drops that were being experienced elsewhere. Corrections were taken for police coverage, incarceration rate, level of cocaine use, population density and resource deprivation, the last a composite measure that includes factors such as poverty rate and male unemployment.

Their conclusions rattled more than a few cages. Once extrinsic factors were taken into account New York’s drop in homicide didn’t significantly exceed that of comparable areas. Compstat might be a terrific idea, but in this study it wasn’t demonstrably so. Richmond, on the other hand, easily passed the test, its adjusted 22 percent yearly decline in firearm homicide proving significantly better than reductions elsewhere.

Ceasefire proved to be a mixed bag. As this chart from the [Ceasefire report](#) illustrates, a steep and persistent decline in the number of youth gun homicide victims coincided with the project (pre/post-intervention means 3.5/1.3. See “Reducing Gun Violence: the Boston Gun Project’s Operation Ceasefire,” National Institute of Justice, September 2001, p.58)

Exhibit 2-2 Monthly Counts of Youth Homicide in Boston



Examining an extended three-year post-intervention period, Rosenfeld and his colleagues calculated that Boston enjoyed an adjusted 30-percent yearly drop in youth gun homicides, nearly twice the 16 percent yearly reduction reported in comparable cities. However, since the actual number of deaths was few, the 14 percent improvement wasn't enough to reach statistical significance. (Expanding the victim age range, thus increasing their number by only three yielded what the writers termed "marginal" significance.)

But let's not quibble. Boston Ceasefire posted impressive real-world results. That's to be expected. Police, probation officers and Federal agents served warrants, did stop-and-frisks, made drug and gun busts and conducted probation and parole checks. Yet, although the [NIJ Research Report](#) concedes that the program incorporated the "certainty, swiftness and severity of punishment" aspects of the deterrence model, the tendency has been to credit Ceasefire's success to its unique notification and warning aspects. That explanation has become so common that when discussions at the recent NIJ conference turned to the program one could be excused for thinking that there was no enforcement component at all. On the contrary: as the descriptive sections of the NIJ report make clear, police & probation efforts were very substantial. They were certainly so from the perspective of offenders, who are unused to concerted law enforcement measures, and particularly if they persist.

Teasing out just how much of Boston Ceasefire's fourteen percent gain came from locking people up and how much from everything else was impossible then, and it's impossible now. As one of Project Safe Neighborhoods' evaluators told the blogger, specifying the effects of, say, notifications is well-nigh impossible.

In 1999 the University of Illinois School of Public Health initiated [Chicago Ceasefire](#).



Don't be fooled by the "Ceasefire" label -- this is an unique approach. Street workers and "violence interrupters" prowled inner-city areas, identifying and counseling high-risk youth, mediating disputes and defusing situations that might lead to violence. Every effort was made to keep staff members independent and credible. Unlike Boston, there was no deployment of police, and while official tips about violence were welcomed, information only flowed one way.

A recent [NIJ evaluation](#) reports mixed results. Seven of Chicago Ceasefire's sites were matched with seven locations where the program was not in effect. Homicides fell significantly more than in the matched area at only one site (again, death counts were very small.) Other results were more promising. When compared to matched locations, four project sites experienced additional decreases of 17 to 24 percent in shots fired, and four demonstrated additional decreases of 16 to 34 percent in actual shootings.

Evaluating Chicago Ceasefire presents many challenges. There were other projects, including PSN, operating in and near Ceasefire sites. Assessors also raised serious doubts about the equivalency of the comparison sites. That's a potentially fatal flaw. High-crime locations such as those where Ceasefire was deployed tend to attract more policing. Without data on the nature and intensity of law enforcement activity, attributing improvements to program effects is risky.

There's another concern. Consider, for example, the far higher violence rates of [PSN vis-à-vis non-PSN cities](#). Of course, you say: that's how sites were selected in the first place! But extreme scores are unstable and apt to revert to more moderate levels for no discernible reason. If a generalized crime drop is already underway, precipitous changes could be easily misinterpreted. Absent a robust research design, bundling high-crime locales is just asking for trouble come evaluation time.

In “Knowing when to fold 'em: an essay evaluating the impact of Ceasefire, Compstat and Exile,” UCLA statistician [Richard A. Berk](#) gloomily concludes that unless programs are specifically designed to be rigorously evaluated doing so may be unwise.

What if random assignment, a strong quasi experiment, or a convincing analysis of observational data are not in the cards? Even if the policy questions are vital, it may be wise to throw in the hand. Suspect science, even the best that can be done under the circumstances, does long run damage to the credibility of all science. The position taken here is that under these circumstances, responsible researchers should withdraw until stronger studies are possible. It may even be possible to help make those stronger studies more likely.

Guiding the lily with unsupportable claims ultimately works to everyone's disadvantage. Yet public servants don't have the luxury not to decide, and their decisions must be based on *something*. Often that “something” is their best judgment, informed with hefty doses of real-world experience. Next week in the (hopefully) final part of this series we'll examine some promising real-world approaches to fighting crime and violence.