[redacted]

Summary of Findings

The Impact of Policing on Social Disorder

by
Wesley G. Skogan
Center for Urban Affairs and Policy Research
Northwestern University
Evanston IL 60208 USA

Prepared under Grant 92-IJ-CX-0008 from the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view or opinions in this document are those of the author and do not necessarily represent the official position of the U.S.Department of Justice.

The Impact of Policing on Social Disorder

What is the impact of policing? Does a traditional response to community problems — intensive crackdowns in troubled areas — actually provide an effective solution? Can a widely-heralded alternative — community policing — live up to the expectations of its supporters? The answers are not clear. There have been relatively few systematic evaluations of policing programs that were effectively implemented. Other policing tactics have not been evaluated, and some of the goals of community policing — such as to alter departmental cultures — are subtle and difficult to assess. This report describes several policing projects that were examined carefully. Data from those evaluations can be used to compare the impact of community policing programs with those of intensive enforcement programs, and to benchmark both of them against conditions in comparison areas that represent normal levels and styles of policing.

One target these programs had in common was social disorder (cf, Skogan, 1990). Social disorder is signaled by bands of teenagers deserting school and congregating on street corners, prostitutes and panhandlers soliciting for attention, public drinking, vandalism, the verbal harassment of women on the street, street violence, and open gambling and drug use. Communities beset by disorder can no longer expect people to act in civil fashion in public places. Some social disorders are clearly illegal, and community residents can hope to get the police interested in those problems. But violators of other widely-approved standards of public conduct are not so clearly breaking the law. A great deal of disorderly behavior falls into such ambiguous legal categories as "disturbing the peace,"

"loitering," and "vagrancy." Many disorders do not have individual victims, despite their collective consequences. While these disorders often lead to complaints that the authorities "do something," the source of the public's concern is often the anticipation of illegal behavior or the possible consequences of growing disorder for the community, rather than a specific criminal incident. Because of the tenuous legal status of such complaints, and the fact that many disorders are not conventionally defined as serious problems even if they are illegal, getting the attention of the police or other municipal agencies can be difficult. Albert Reiss (1985) captured the flavor of disorderly conditions lying near the boundaries of the law when he dubbed them "soft crimes."

This report assesses the impact of several special policing programs on social disorder. In almost every case the programs described here were evaluated using a quasi-experimental research design. Each program was conducted in a different area, while another matched area was designated as a comparison area where no new policing programs were begun. Surveys of residents were conducted in the target and comparison areas before the programs began and again after they had been in operation for a period ranging from ten months to two years. A variety of other kinds of data were collected as well, and the actual implementation of the program was monitored in all the cities. Reports have appeared describing the individual projects; they are listed in the Citations. This report draws together some of the data on which they were based and describes a new analysis combining the results in one "meta-evaluation" of the programs.

Home Visits

Home visit programs (also known as "door-to-door" visits) are intended to gather information about neighborhood problems from citizens who have not called the police, as part of a problem-solving effort. They also are used to spread awareness of special police programs, and to introduce area residents to community beat officers. In Oakland, California, officers went door to door in an experimental policing area, introducing themselves to residents. Their job was to inform people in the target neighborhoods of the department's new emphasis on drug enforcement, to give them pamphlets on crime and drug programs, and conduct brief interviews asking about neighborhood problems. Their goal was to make contacts which might lead to useful information, alert the community to the drug problem, and perhaps deter potential offenders to their presence and visibility in the community. These door-step interviews were conducted in about 60 percent of the households in the target areas; about 50 percent of those interviewed indicated that drugs were a major problem in their community. In another section of Oakland home visits were combined with an intensive drug enforcement effort this is described below.

In Birmingham, Alabama, officers made home visits in order to pass out crime and drug prevention pamphlets and conduct interviews with area residents. They developed a questionnaire that ask residents about neighborhood crime problems and the whereabouts of drug trafficking. They eventually completed interviews at 60 percent of the occupied housing units in their target area.

Houston, Texas' home visit program was to help patrol officers to become more familiar with the residents of their areas and to learn about neighborhood problems. Officers were freed from routine patrol assignments for part of each daily shift to make household visits. Typically they would introduce themselves, explain the purpose of the visit, and inquire about neighborhood problems. They recorded these on a citizen contact card, along with the name and address of the person they interviewed. Officers left personal business cards, indicating that if there were further problems they should be contacted directly. The contact cards formed a mailing list for newsletters. During the ten-month evaluation, officers talked to approximately 14 percent of the adult residents of the area, and to about 45 percent of area merchants. About 60 percent of the people that were interviewed had something to complain about. Conventional crimes were most frequently mentioned, but about one-quarter of the residents mentioned a problem which falls into the disorder category, including disputes among neighbors, environmental problems, abandoned cars, and vandalism. The officers took numerous actions in response to problems they identified during these visits.

Houston's Community Organizing Response Team (CORT) attempted to create a local crime prevention organization, by first identifying a group of residents who would work regularly with them to define and help solve neighborhood problems. To test the CORT concept, officers first conducted their own door-to-door survey of the neighborhood. They members questioned approximately 300 residents about problems which they felt merited police attention, and whether they might be willing to host meetings in their homes. The survey told them a great deal about

the nature of area problems, and resulted in invitations to hold such meetings. Thirteen neighborhood meetings were held, each attended by 20-60 people. At these meetings CORT members identified a group of leaders who met regularly with their commander to discuss community problems and devise solutions involving both the police and residents. The group eventually held elections and formed committees, and by the end of the evaluation period had sixty official members. During the evaluation period special newsletters were mailed each month to all residents who had been contacted in the survey or who had participated in an activity. The CORT program tested the ability of police departments to assist in the development of community self-help organizations.

Storefront Offices

Small police substations have been opened in a number of cities to provide a locus for decentralized, neighborhood-oriented programs. They are a visible sign of police commitment to the community, and evaluation surveys indicate that they are widely recognized. Madison, Wisconsin, attempted to develop a "customer orientation" in providing police services by radically restructuring the police department and the way in which it was managed. To reform the organization, an innovative management structure was put in place that emphasized teamwork and employee participation in decision making. Police were to work as teams to identify and solve problems, with their managers working for them to secure the outside assistance and resources that they required to carry out their plans. A decentralized police substation was opened to experiment with these ideas. Officers there worked flexible hours and took responsibility for managing their own activity. They developed a plan for "value added

policing" that called for spending more time on calls for service and followup contacts with victims. They responded to most of the calls for service that originated from the area, and attempted to analyze them to identify community problems. Interviews with police officers revealed that, compared to those assigned elsewhere, officers in the experimental district saw themselves working as a team, that their efforts were being supported by their supervisors and the department, and that the department was really reforming itself. They were more satisfied with their job and more strongly committed to the organization. They were more customer oriented, believed more firmly in the principles of problem-solving and community policing, and felt that they had a better relationship with the community.

The Birmingham storefront was instituted in the evaluation's planned comparison area after eleven people were shot in there in a short period, just after the beginning of the research project. In response to community demonstrations, a police substation was opened, staffed 24 hours per day by eight police officers. They greatly increased the visibility of police in the community. The substation unit assisted in a clean-up of the public housing project which dominated the area.

Houston's neighborhood police substation provided a place for people to meet with police. Officers took crime reports and gave and received information from the public, and some community meetings were held there. Officers assigned to the station were freed from routine patrol for much of their daily shift. The substation was their base of operations for getting acquainted with neighborhood residents and business people, identifying and helping solve local problems, seeking ways of delivering better

service to the area, and developing programs to draw the police and community closer together. The staff quickly developed programs which extended into the immediate neighborhood, including holding a series of large community meeting, conducting special patrols in area trouble-spots, and devising a truancy program. On five occasions during the evaluation period the station staff distributed program newsletters throughout the neighborhood. The station provided a direct test of several aspects of community policing. It provided the officers who ran it a great deal of management autonomy, and flexibility in allocating their own time and effort. They responded by developing community-oriented programs, including a variety of new ways in which police and citizens could meet and exchange information and discuss their priorities.

Foot Patrols

There has been a great deal of interest in using foot patrols as a tool for community policing. Two versions of foot patrol were tested in Baltimore, Maryland. Foot officers were assigned to walk through test areas approximately 25 hours each week. They concentrated on busy commercial areas and recognized trouble spots, and talked frequently with residents, business owners, and people on the street. In one area the officers put additional stress on law enforcement and order maintenance; they spent much of their time dispersing groups of youths on street corners and looking for drug transactions and other legal infractions. In two other areas police officers were assigned to work as "Ombudsmen" with neighborhood residents to solve local problems. They walked foot patrol, attended community meetings, and spent a great deal of time talking to merchants and residents

about local problems. They developed a questionnaire which measured what residents thought were the most serious problems in the area, what caused them, and what could be done to solve them. Officers were to record how they had reacted to each problem, and their handing of them was reviewed by their supervisors. The officer serving one area was aggressive in his approach to possible drug dealers, broke up groups loitering on the street, and gave many traffic tickets. The officer in the other target area spent more time meeting with area residents, working to solve juvenile problems, conducting a neighborhood clean-up campaign, and organizing a block watch program. Baltimore's Ombudsman policing program included elements of both foot patrol and home visits, and it will be analyzed as an example of each.

Intensive Enforcement

In one area of Newark, New Jersey, police attempted to suppress crime and street disorder using traditional enforcement tactics. They conducted extensive "street sweeps" to reduce loitering and public drinking, drug sales, purse snatching, and street harassment by groups of men who routinely gathered along commercial streets in residential areas of the city. Congregating groups were broken up by police warnings and large-scale arrests. Foot patrol officers walked the areas in the evening; they were to become familiar with local problems, establish relationships with local merchants, disperse unruly groups, and ticket illegally parked cars. Special efforts were made to enforce traffic regulations in the area using radar units and by making frequent traffic stops to check for alcohol use. Random roadblocks were set up to check drivers' credentials, check for drunken driving, recover stolen vehicles, and arrest drives with

outstanding tickets and arrest warrants. There was also an attempt to clean up area parks and vacant lots, and to deliver better city services. This program tested the ability of the police to reassert their faltering authority, demonstrate that they controlled the streets of Newark, and crack down on forms of disorder thought to lead to serious crime.

In Oakland, a special drug enforcement unit conducted traditional police operations in one experimental area. They went undercover to make buy-bust arrests, and they used informants to buy drugs and identify distributors. They also mounted an aggressive, high-visibility program of stopping and searching motor vehicles, and conducting field interrogations of groups of men whenever they gathered in public places. The team was extremely active, made a large number of arrests, and apprehended a number of major drug traffickers in the target area. In another area intensive enforcement was coupled with a home visit program.

Like Oakland, Birmingham formed a special drug enforcement unit to crack down on open drug dealing in dilaudid and cocaine. The team concentrated on undercover operations. They made a series of videotaped purchases from street dealers, and then returned to the target area to make warrant arrests. Officers also posed as dealers and made videotaped drug sales to outsiders who were driving into the target area to make drug purchases. Throughout, they paid careful attention to the legality of their activities, to ensure that their cases could be successfully prosecuted. Ten officers were involved in this program for a six-month period, and made a number of arrests.

Denver, Colorado, mounted an intensive enforcement program that focused on drug sales and use in that city's public housing developments. As in many cities, uniformed police and officers on narcotics assignments in Denver avoided working in public housing areas: they scorned the residents, and could make more impressive seizures of drugs and money in better-off areas. To deal with this, the department formed a special six-officer Narcotics Enforcement in Public Housing Unit (NEPHU) to signal recognition of the importance of drug problems in public housing and to focus new energy and resources on those problems. The City promised to cooperate with the Denver Housing Authority, tenant advisory councils, and community members, in a program combining undercover enforcement, visible patrol, improved tenant management, public education, resident-based crime prevention, and community involvement. However, NEPHU actually only conducted aggressive undercover tactics. They developed a number informants (drug dealers or users who could be "spun" in exchange for money and relief from prosecution) who made controlled buys while under observation, and occasionally conducted their own buy-bust purchases of drugs. They used this information to secure warrants for searching apartments and making arrests; in the course of a year they made 176 drug-crime arrests, one-third of which were in the two projects that were monitored during the evaluation.

A Multi-Program Area

In one area of Newark, police implemented a variety of community policing and intensive enforcement efforts. The evaluation tested the ability of an ambitious multi-intervention program to effect crime and fear of crime in an extremely diffi-

cult area. Police opened a substation that took crime reports, distributed crime prevention information, gathered complaints about city services for referral to other municipal agencies, and answered questions. Local groups held meetings in the station during the evening, and about 300 people used the substation each month. Police also conducted home visits in the area, visiting homes and filling out brief questionnaires concerning neighborhood problems. The teams also distributed crime prevention information, told residents about block watch programs, and advertised the substation. During the course of the evaluation they questioned residents of 50 percent of the homes in the area. The sergeant supervising the team reviewed the questionnaires, and either his team dealt with the problems that residents identified or he passed them on to the special enforcement squad for their attention. The team also organized a neighborhood clean-up program and distributed a police newsletter. Finally, a special squad targeted street disorder in the area.

Summary of the Programs

Figure 1 classifies the 18 neighborhood programs that were fielded in these seven cities by their central elements. This classification reduces the community policing and intensive strategies described above to four general categories that were used to examine their impact. Houston's community organizing team gained a great deal of visibility through their initial survey of their target area, and from follow-up meetings that they held in resident's homes, so they are classed in the home visits category. Several of the more complex programs are

Community Policing Strategy Elements by City

	Comparison Areas	Foot Patrol	Home Visits	Store-front Office	Intensive Enforce-ment
Houston	✓		✓	✓	
Newark	✓	✓	✓	✓	✓
Denver					✓
Birmingham			✓	✓	✓
Oakland	✓		✓		✓
Baltimore	✓	✓	✓		
Madison	✓			✓	

NOTE: some neighborhood programs had more than one strategic element and are classified in more than one column.

classed under more than one strategy heading, reflecting the mix of activities that were involved. Baltimore's Ombudsman program involved elements of both foot patrol and home visits, so the two Ombudsman program areas will be treated as having been exposed to both strategic elements of community policing. The areas in Baltimore that were set aside just for foot patrol will be classified only in that category. The multiple program area of Newark was the home for foot patrol, door-to-door visits and a storefront office, as well as the target of an intensive enforcement campaign; as a result, it will be classed as having been subjected to all of those strategies. Madison's Experimental Police District was more than just a storefront office; it was the headquarters for a thorough-going experiment in the decentralized management of full-service police team. However, from the perspective of the public it was principally a local outpost, and it will be classified with other storefront programs.

Figure 1 also indicates which evaluations included comparison areas. These were areas (in Madison, the remainder of the

city) in which normal levels and modes of policing were maintained during the evaluation period. There were eight comparison areas in these evaluations. In one case (Birmingham, Alabama) the research design broke down when the comparison area was the subject of a wave of random violence and shootings shortly after the evaluation began; a neighborhood police substation was opened in that area. There was no comparison area in Denver.

Program Implementation

A more difficult task is to classify these programs by the extent to which they were implemented. As the brief program descriptions presented above indicated, they were not all strong programs; implementation problems plagued even these closely monitored projects. They were faced with three challenges. Some were disbanded in the face of rising calls for service, in order to restore t--aditional service levels. Others were discredited by mid-level managers who resented their loss of authority to lower ranking personnel. And some failed to endure because they did not succeed in changing the organizational culture of the department.

In particular, successful community policing programs must not ignore the "911 problem." Since the volume of telephone calls to big-city departments skyrocketed in the mid-1970s, police commitment to respond to these calls as quickly as possible has absorbed the resources of many departments. In effect, many departments are being managed by the thousands of citizens who call the police, not by their commanders. In the home visit area of Birmingham, pressure to respond quickly to calls for service at a time when the police district was understaffed led

to the abandonment of the problem-solving aspects of the program. Although officers had completed a large number of interviews, no effort was made to follow up on the information that was gathered. It was envisioned that they would do team-oriented problem solving with the information that they gathered, but a rise in calls for service in their area of the city came at the same time that the Christmas holiday season left the district understaffed. Under pressure to respond to the resulting deterioration in police response to 911 calls, officers who were to conduct the community policing program were reassigned to traditional patrol. The district commander responsible for devising the program was punished with an undesirable assignment for letting responses to calls for service slacken because of his commitment to the community policing experiment. At about the same time, Houston's city-wide community policing effort was halted following charges that police had allowed responses to calls for serviced to deteriorate, because (it was charged) officers were being freed from this responsibility in order to carry out the program. The program had powerful enemies among lieutenants and other mid-level supervisors in the department; the Chief of Police was soon fired, and little remains of her program.

In other cities, community policing has floundered in the face of the crime-fighting culture of traditional departments. In Oakland there was little enthusiasm for community policing among officers assigned to the program. While a few hard-working officers carried out the most easily monitored task – making home visits and conducting interviews – they did nothing to follow up on the information that they gathered. Their immediate supervisor dismissed the effort as "social work," and did nothing ensure that the community policing program developed in the Chief's

office actually was implemented in the field. None of the intended problem-solving policing was ever accomplished, and nothing was done with the information gathered in the door-step interviews. One officer conducting home visits actually quit his job because he was so frustrated by the lack of support for his efforts. In Baltimore, officers pulled from routine assignments to replace the foot patrol officers while they were on vacation were unenthusiastic about the assignment. One of the Ombudsman police officers preferred giving out traffic tickets to interviewing citizens and attending meetings. Denver's antinarcotics team was unable to secure the cooperation of the department's Patrol Division to maintain high-visibility patrols, and other units pulled out of the area altogether once NEPHU was formed. Survey measures indicate that police visibility declined once the program began, as did resident's reports of being stopped by police while in the area. The team was also unable to coordinate their efforts with the local Housing Authority. There were conflicts between them and Authority personnel over tenant eviction policies; project managers faced a conflict between demands that they strictly manage tenants and the requirement that they keep all units occupied; and NEPHU officers kept all their plans secret from development managers and ignored the participation of the Housing Authority Security Director, who they believed did not understand "real police work." The team resolved the conflict they perceived between doing "hard-nosed" narcotics work and community involvement or participation by giving up on the latter. In fact, they thought that residents were the problem, rather than a solution for it; rather than communities needing defending, they viewed the projects as hostile territory.

Thus, it is clear that differences in the extent of program implementation threaten to mask the effect of the various strategic elements of community policing that have been described here. To account for this, a measure was developed of the extent of implementation of each neighborhood program. This measure was included in all of the statistical analyses which follow. The measure was to reflect how adequately each program was staffed for the problem it addressed, how well organized and focused it remained in practice, and if it was appropriately conducted. This definition of implementation was posed in a brief questionnaire sent to four other evaluators involved in the original projects.[1] They were asked to rate each of the programs that they felt informed about as a weak, moderate, or strong implementation. After excluding the ratings of one discrepant evaluator a consensus implementation measure was derived by averaging their scores.

Measuring Social Disorder

Disorders violate widely shared norms about public behavior. However, it is the nature of disorder that it usually does not appear in official police statistics. Many disorders do not have individual victims, despite their collective consequences, others are not illegal, and it can be difficult to get the police interested in many which are because they are not conventionally defined as "serious crime." In this circumstance, surveys of neighborhood residents provide one of the best means of assessing

[1] In addition to the author, implementation ratings were contributed by Sampson Annan, Lawrence Sherman, Craig Uchida, and Mary Ann Wycoff.

the extent of the problem, and for judging if a program makes any progress against it.

In the surveys, respondents were first read an introductory statement. This was followed by a list of a variety of neighborhood problems for them to rate. The index of social disorder was constructed from responses to ten of them, which are listed in Figure 2.

Figure 2
Survey Measures of Social Disorder

Now I am going to read you a list of some things that you may think are problems in this neighborhood. After I read each one, please tell me whether you think it is a big problem, some problem, or no problem here in this development.

1. *Groups of people hanging around in the neighborhood?*
2. *People drinking in public places?*
3. *People saying insulting things or bothering people as they walk around the neighborhood?*
4. *Organized gangs?*
5. *Drug sales or use?*
6. *Disruption around schools; that is, youths hanging around making noise, vandalizing, or starting fights?*
7. *Truancy; that is, kids not being in school when they should be?*
8. *Cars being vandalized—things like windows or radio aerials being broken?*
9. *People being attacked or beaten up by strangers?*
10. *People being attacked or robbed?*

Variations in the wording of many of these questions were fairly slight. Items 6 and 7 were used alternately in various surveys. Item 9 was used in 3 cities, and item 10 in three others. Using these alternatives as substitutions, the social disorder score used here was based on eight component measures.

Assessing Program Impact

The statistical analysis examined the impact of living in an area where special policing activities took place. In areas where multiple programs were fielded respondents were treated as having been subjected to each of them. For example, respondents in the area in Oakland in which both home visits and intensive

enforcement programs were conducted were "tagged" as being exposed to both efforts. Strength of program implementation was measured by the three-point ranking described above. The analysis pooled the surveys from all seven cities and 27 areas, and probed for program effects after controlling for the extent of social disorder problems each respondent's neighborhood before the programs began. This approach had several advantages over examining the projects individually. First, it increased the size of the sample on which the statistical analysis is based. This should increase the precision with which various program effects can be identified. This approach should also increase the generality of the findings. Every type of program was fielded in more than one city, and in total 21 different target neighborhoods were involved. The pooling approach also enables us to probe the effects of the various components of multiple-strategy programs, like those conducted in neighborhoods in Newark and Oakland. Respondents there were compared statistically with those in other areas who were exposed to just one program element, or different mixes of elements. Respondents from the six comparison areas that were utilized in these evaluations were be included as well. They served the same valuable purpose they did in the original projects; they provided "baseline" respondents who were not exposed to any of the programs. In addition, each analysis was replicated with the addition of five demographic characteristics of each respondent: age, sex, education, home ownership, and length of residence. These were factors also potentially associated with perceived levels of neighborhood disorder that could not be affected by the programs. Race was not controlled, for it was highly associated with several neighborhoods and programs; a separate analysis of program effects by race will be described below.

Table 1 presents the results of this analysis. It presents unstandardized regression coefficients linking exposure to each program, and the measure of strength for each program, to neighborhood social disorder. The table omits coefficients for the pretest measure and, in the right-hand columns, for the five demographic control factors.

Table 1
Regression Analysis of Wave 2 Social Disorder

Program Type			With Demographics	
	coeff.	signf.	coeff.	signf.
intensive enforcement	.20	.00	.18	.00
home visits	-.06	.00	-.06	.00
storefront offices	-.06	.00	-.06	.00
foot patrols	.07	.00	.10	.00
program strength	-.05	.00	-.04	.00
R-squared	.37		.38	
N of cases	4577		4509	

NOTE: excludes coefficients for the pre-intervention measure and five demographic control factors.

Table 1 indicates that two community policing strategies — door-to-door visits and storefront offices — were significantly associated with lower levels of social disorder. On the other hand, exposure to intensive enforcement programs and foot patrol was associated with higher levels of social disorder afterward, compared to the other areas. All of these were net of the effect of the strength of program implementation, which was also significant.

Were these programs differentially effective against particular kinds of disorder? To explore this issue, the analysis described above was replicated using the individual components of the social disorder scale score. The results were similar, and virtually uniform. All of the significant coefficients associated with intensive enforcement and foot patrol were positive (suggesting that disorder went up, net of other factors), while all of the coefficients associated with storefront offices, door-to-door visits, and strength of program implementation were negative.

Differential Impact by Race

This leaves the issue of race. A report on Houston's community policing experiments concluded that black respondents living in that city's program areas did not benefit measurably from the programs that were fielded there. They were significantly less likely than whites to be aware of the programs, or to recall coming into contact with them. They were also less likely to change their views of the police or perceptions of levels of neighborhood problems (Pate, et al, 1986). Such a differential distribution of program benefits could create significant political problems for proponents of community policing, if they prove to be robust across cities and programs.

The data examined here is not ideal for considering the impact of race. Most of the Hispanics involved in these projects lived in just two cities: Houston (where they made up 20 percent of those interviewed) and Denver (56 percent). The communities involved in policing programs in Birmingham and Newark were almost completely black, and in Oakland 86 percent of the survey

respondents were black. Respondents to the Baltimore surveys were 51 percent black and 49 percent white, but this was a consequence of selecting an even mix of homogeneously black and white project neighborhoods. Almost every respondent in the Madison evaluation (97 percent) was white, on the other hand. As a result, the distribution of programs across the racial mix of neighborhoods included here was not uniform: blacks and Hispanics were over-represented in intensive enforcement areas. Otherwise, however, all groups of program and comparison-area respondents were 42-48 percent black, and black and Hispanic respondents lived in both strong and weak program areas. As a result, it is useful to use these data to probe for differential program effects.

The analysis indicates that two programs which did not appear to have much over-all impact — foot patrol and intensive enforcement — were indeed related to lower levels of social disorder among black respondents, but not among whites. Among whites, Wave 2 social disorder scores tended to raise rather than fall; among blacks they fell significantly. None of the analyses testing for differential program effects for Hispanics were significant. There were no significant racial differences in the impact of either storefront offices or home visits; the benefits of those programs which were documented above were uniform across racial groups.

While far from definitive, the results suggest that these programs did not particularly rebound to the disadvantage of racial minorities. These results could be observed using other methods. The same pattern could be observed in simple Wave 1-Wave 2 change scores: the change scores for whites were posi-

tive (higher at Wave 2) among those exposed to foot patrol and intensive enforcement, but they were negative (lower at Wave 2) for blacks. The change scores were large and negative for all groups living in home visit and storefront areas.

Conclusions

Implementation. The projects described above illustrate the difficulty of sustaining innovative policing programs. In Baltimore, one of the Ombudsman officers preferred to give traffic tickets rather than work with the community. In Oakland there was little enthusiasm for community policing among officers assigned to the program, and they did nothing to follow up on the information that they gathered. In Birmingham, pressure to respond quickly to calls for service at a time when the police district was understaffed also led to the abandonment of the problem-solving aspects of the home visit program. Houston's city-wide community policing effort ground to a halt following charges that police there had allowed responses to 911 telephone calls to deteriorate. Denver's drug enforcement team found it impossible to coordinate their efforts with the city's Housing Authority despite their focus on public housing, and they never followed through on their commitment to involve the community in their efforts. The pressure to maintain traditional service levels, the resistance of mid-level managers, and the failure of departments to change their organizational culture, all contributed to the partial failure of these new programs.

Impact. Proponents of community policing in particular must develop better answers to the question, "Does it work?" The evidence to date is mixed. The most consistent finding of the

evaluations reviewed here is that community policing improves popular assessments of police performance (cf, Skogan, 1993). While this is certainly an accomplishment – especially because it seems to affect all racial groups – it is vulnerable to the argument that this is merely a triumph of public relations. Rarely is there good evidence that crime has been reduced by these programs. For example, there is no good evidence that foot patrol – the most thoroughly evaluated form of alternative policing – effects the crime rate, although other benefits of this tactic are clear. The analysis presented here suggests that neither foot patrol nor intensive enforcement reduced levels of social disorder afflicting white or Hispanic residents of the five cities in which they were tested.

Equitable Policing. The evaluation of community policing in Houston found that the way in which several programs were run favored the interests of racially dominant groups and established interests in the community. The Houston experience illustrated that policing by consent can be difficult in places where the community is fragmented by race, class, and lifestyle. If, instead of trying to find common interests in this diversity, the police deal mainly with elements of their own choosing, they will appear to be taking sides. It is very easy for them to focus community policing on supporting those with whom the get along best and share their outlook. Critics of community policing are concerned that it can extend the familiarity of police and citizens past the point where their aloofness, professionalism, and commitment to the rule of law can control their behavior.

This reanalysis of seven policing field experiments did not find this to be a general problem. The benefits of storefront

offices and home visits were significant, and they were enjoyed by community members of all races. Interestingly, it was white non-Hispanics who were most unlikely to see a reduction in social disorder, either in foot patrol areas or in places targeted for special enforcement efforts.

CITATIONS

Baltimore

Pate, Antony and Sampson O. Annan. 1989. *The Baltimore Community Policing Experiment: Technical Report*. Washington, DC: The Police Foundation.

Pate, Antony. 1989. "Community-Oriented Policing in Baltimore." In Dennis J. Kenney (ed.), *Police and Policing*. New York: Praeger, 112-135.

Denver

Annan, Sampson and Wesley G. Skogan. 1993. *Drug Enforcement in Public Housing: Signs of Success in Denver*. Washington, DC: The Police Foundation.

Skogan, Wesley G. and Sampson Annan. 1993. "Drug Enforcement in Public Housing," in Davis, Lurigio and Rosenbaum (eds.), *Drugs and the Community*. Springfield, IL: Charles C. Thomas, 162-174.

Skogan, Wesley G. and Sampson Annan. 1993. "Narcotics Enforcement in Public Housing," in MacKenzie and Uchida (ed.), *Drugs and the Criminal Justice System*. Newbury Park, CA: Sage Publications, in press.

Houston and Newark

Bayley, David H. and Jerome Skolnick. 1986. *The New Blue Line*. New York: The Free Press.

Pate, Antony M., Mary Ann Wycoff, Wesley G. Skogan and Lawrence W. Sherman. *Reducing Fear of Crime in Houston and Newark: A Summary Report*. Washington, DC: The Police Foundation.

Sherman, Lawrence W. 1992. "Attacking Crime: Policing and Crime Control." In Michael Tonry and Norval Morris, Eds. *Modern Policing*. Chicago: University of Chicago Press, 159-230.

Skogan, Wesley G. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities*. New York: The Free Press.

Wycoff, Mary Ann and Wesley G. Skogan. 1986. "Storefront Police Offices: The Houston Field Test." In Dennis Rosenbaum, Ed. *Community Crime Prevention*. Newbury Park, CA: Sage, 179-201.

Oakland and Birmingham

Uchida, Craig, Brian Forst and Sampson Annan. 1990. *Modern Policing and the Control of Illegal Drugs: Testing New Strategies in Two American Cities*. Washington, DC: The Police Foundation (Final Technical Report).

Uchida, Craig, Brian Forst and Sampson Annan. 1992. *Modern Policing and the Control of Illegal Drugs: Testing New Strategies in Two American Cities*. Washington, DC: National Institute of Justice; US Department of Justice (Summary Report).

Madison

Wycoff, Mary Ann and Wesley G. Skogan. 1992. *Quality Policing in Madison: An Evaluation of Its Implementation and Impact*. Washington, DC: The Police Foundation (Final Technical Report).

Other

Reiss, Albert J. Jr. 1971. *The Public and the Police*. New Haven: Yale University Press.

Skogan, Wesley G. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities*. New York: The Free Press.

Skogan, Wesley G. 1993. "The Impact of Community Policing on Neighborhood Residents: A Cross-Site Analysis." In Dennis P.Rosenbaum (ed.), *Community Policing: Testing the Promises*. Newbury Park, CA: Sage Publications, in press.

The Impact of Policing on Social Disorder

TECHNICAL APPENDIX AND TABLES

Evaluation Surveys

Except in Madison, all of the interviews were conducted during personal home visits by trained interviewers. Respondents were selected by first listing all of the residential addresses in each program or control area. Sample addresses were then selected from the list using a randomly chosen interval; this approach was employed to ensure that respondents did not live next door to one another, but were spread evenly throughout the area. Individual respondents were then selected from eligible household members using a Kish table. Persons 19 years of age and older were eligible for selection. Multiple callbacks (usually up to eight) were made to locate selected respondents. Because the entire City of Madison was involved in the evaluation of the Experimental Police District, a different sampling technique was employed there. Sample households were selected using an area probability sampling frame maintained by the Wisconsin Survey Research Laboratory; half of those selected lived in the part of Madison served by the station, while the other half were scattered throughout the remainder of the city. (The sample excluded residents of university dormitories and student group quarters, and several census tracts inhabited largely by students.) The first wave of interviews were conducted in person; at that time respondents were asked for their telephone number, and they were reinterviewed later by telephone.

As Table A-1 indicates, the completion rate for the original surveys ranged from 58 percent (Oakland) to 84 percent (Birmingham). The reinterviews were attempted 8-24 months later, with an overall success rate of 71 percent. Oakland may be the worst case presented in Table A-1, especially in light of the relatively brief period (9 months) between the waves of interviewing. In Oakland the leading reason for failing to conduct a Wave 2 reinterviews (13 percent) was that no one was at home at the sample addresses after up to 8 visits. In 9 percent of cases the original respondent had moved, and in 4 percent the dwelling unit was vacant. Only 3.7 percent of those originally interviewed refused to be reinterviewed, and in 2 percent of cases the people living at the sample address claimed to not know the respondent who was being sought.

Table A-1
Interview Rates and Sample Sizes, by City

City	Percent Inter-viewed	Elapsed Months	Percent Reinterv-iewed*	Number of Areas	Range for Areas	Panel N
Houston		11	75	4	72-81	1294
Newark		11	75	3	73-77	960
Denver	76	12	75	2	74-76	390
Birmingham	84	8	75	3	71-82	438
Oakland	58	9	64	4	58-68	502
Baltimore	81	14	76	6	66-76	599
Madison	77	24	62	2	56-69	727
Total/Range	58-84	8-24	71	24	56-82	

An analysis of attrition was conducted using logistic regression to predict which respondents who were interviewed at Wave 1 were successfully reinterviewed at Wave 2. This identified factors that were independently related to attrition. Many of the factors that were associated with being reinterviewed reflected residential stability: home ownership, lengthy residence in the neighborhood, and having children. People who were satisfied with their neighborhood and who had not recently been a victim of crime were more likely to be reinterviewed as well. So were older respondents and women, and respondents who were rated as cooperative and interested by interviewers during the first interview. Dropouts were more likely to speak poor English, have less education, or be a student. Not surprisingly, net of all of these factors, the number of months that elapsed between the first interview and the second wave of interviewing was also associated with attrition.

Table A-2 summarizes the final count of two-wave panel interviews in each of the program and control areas examined here.

Table A-2
Number of Panel Interviews, by Area

Houston		Oakland	
storefront	330	door-to-door	112
door-to-door	389	enforcement	129
organizing	284	both progs.	167
control area	291	control area	94
Newark		Birmingham	
enforcement	307	door-to-door	163
multiple	323	storefront	144
control area	330	enforcement	131
Denver		Baltimore	
enforcement	203	ombudsman	105
enforcement	187	ombudsman	82
		control area	120
Madison		foot patrol	76
stationhouse	340	foot patrol	114
control area	387	control area	102

Strength of Implementation

The strength of implementation measure was to reflect how adequately each program was staffed for the problem it addressed, how well organized and focused it remained in practice, and if it was appropriately conducted. This definition of implementation was posed in a brief questionnaire sent to four other evaluators involved in the original projects: Sampson Annan, Lawrence Sherman, Craig Uchida, and Mary Ann Wycoff. They were asked to rate each of the programs that they felt informed about as a weak, moderate, or strong implementation. Only two raters (including the author) scored all 18 programs; the remainder rated between 4 and 8 programs. As Table A-3 indicates, 4 of the 5 raters were in broad but far from complete agreement about the relative rankings of the programs. The ratings of Evaluator Five were highly discrepant, and were thus discarded. A consensus implementation score was derived by averaging and rounding the scores of the remainder for each program. The final score ranged from "weak" (scored 1) to "strong" (scored 3).

Table A-3
Agreement (Correlation) Between Implementation Ratings

	Evaluator One	Evaluator Two	Evaluator Three	Evaluator Four
Evaluator Two	.38			
Evaluator Three	.26	.21		
Evaluator Four	.51	.58	-	
Evaluator Five	.00	-.85	-	-.34

Note: "-" indicates the pair did not share any overlapping ratings.

Table A-4 indicates the strength of implementation score for each project.

Table A-4
Program Implementation by City

	Strength of Program Implementation		
	Weak	Moderate	Strong
Houston			storefront door-to-door organizing
Newark		enforcement	multiple (foot, storefront & door-to-door)
Denver	enforcement (Quigg Newton)	enforcement (Curtis Park)	
Birmingham		door-to-door	storefront enforcement
Oakland	door-to-door		enforcement enforcement & door-to-door
Baltimore		foot patrol (two areas) ombudsman (Callaway)	ombudsman (Ellwood Pk)
Madison		storefront	

This classification of program implementation cross-cuts many of the community policing strategies being examined here. As Table A-5 indicates, by this measure there were weak, moderate and strong enforcement and door-to-door programs in these neighborhoods. There were no weak foot patrol or storefront programs, but their implementation did vary between moderate and strong in different communities. The variation reported in Table A-5 indicates that controlling statistically for strength of implementation will affect each of

the major program categories, increasing our confidence that the resulting estimates of program effect have indeed taken this factor into account.

Table A-5
Program Elements by Program Strength

	Implementation Strength		
	weak	moderate	strong
foot patrol		✓	✓
door-to-door	✓	✓	✓
storefront		✓	✓
enforcement	✓	✓	✓

It is not uncontroversial to employ measures of program strength in the analysis of evaluation data. In particular, it is not justified whenever the subjects of an experiment can "self-select" themselves into varying levels of treatment. If this is possible, the measured strength of treatment can be confounded with the very selection biases that true and quasi-experiments were developed to control. For example, if the treatment is a benefit, the most informed, organized, aggressive, and motivated subjects will likely receive more of it, if they can. If this is the case, the internal validity of conclusions drawn from the data is suspect (Mark, 1983). However, in this case program strength appears to reflect factors that were exogenous to the neighborhoods involved. As outlined above, these included city-wide rises in calls for service, disgruntled middle managers, departmental cultures ill-suited for the programs, and departmental politics. The sole exception may be the neighborhood police office in Birmingham, which was opened (in a planned control area) in response to public pressure following a rash of shootings. Otherwise, community residents simply received whatever service the department was capable of mounting; they were living where they were before the programs were announced, and there is little reason to suspect that self-selection by subjects is related to what happened to the level of program effort in their communities.

Measuring Social Disorder

Because this report is based on a reanalysis of existing surveys, there is no neat bundle of survey questions pointing to the extent of neighborhood social disorder. The mix of available questions varied somewhat from city to city, and there were minor variations in the way in which individual survey questions were worded. In this context, the best strategy was to select a subset of survey items referring to social disorders that were (a) available for multiple, overlapping sets of cities, and (b) substantially correlated with one another and single factored.

Table A-6 presents the correlation between each pair of component measures. All were positive and within the range of normal inter-item correlations for items drawn from surveys. In addition, a pairwise factor analysis was conducted on this set of measures. They were single factored, with the first factor explaining 48 percent of their total variance. That there was no difficulty in factoring the pairwise correlation matrix (eg, it was mathematically internally consistent), is further evidence that the items co-vary in consistent fashion. The auto vandalism measure had the lowest association with the others, based on its communality (.36) which is the variance in this component explained by the factor. However, vandalism is widely associated with the concept of disorder and its inclusion was judged to

increase the face validity of the overall measure. Because of the measurement strategy employed here, only 1843 respondents from Houston, Newark and Baltimore were presented and answered all eight of the measures going into the final scale score; among this subgroup, the reliability of the summary score (assessed using Cronbach's Alpha) was .84.

All of these analyses suggested that a simple additive scale was the best way to represent social disorder in this report. To create the summary disorder measure, responses to these component measures were combined by adding them together, then the resulting total was divided by the number of items which each respondent answered. The items thus made an equally weighted contribution to the total score, which is standardized to fall within a 1-3 range in every case. The scale score was fairly normally distributed, with a skew of .45 (good) and a kurtosis of -.86 (a bit flat, which does not present a problem for OLS regression). The 305 respondents answering fewer than four of the component items were dropped; they constituted about 3 percent of the initial of 4910 panel respondents.

Table A-6
Correlations Among Social Disorder Measures

	drink	gangs	groups	insults	drugs	school	vandals	assault
drink	-							
gangs	.40	-						
groups	.62	.39	-					
insults	.41	.37	.40	-				
drugs	.50	.43	.49	.31	-			
school	.44	.34	.39	.35	.42	-		
vandals	.35	.28	.33	.29	.35	.33	-	
assault	.40	.42	.37	.36	.43	.36	.40	-
Scale Score	.77	.67	.75	.64	.74	.69	.66	.69
N	3150	3022	3165	3141	3356	4194	4347	4283

NOTE: pairwise correlations for Wave 1 data. Number of cases is for the first wave of the panel survey; the number of cases for the second wave is approximately the same.

Variations in the wording of many of these questions were fairly slight. In Oakland respondents were asked about "drug buyers and/or sellers on streets or street corners?" The Houston survey asked about "groups of people hanging around on corners or in streets?" The Baltimore questionnaire referred to "people drinking in public places like on corners or in streets?"

Items 6 and 7 in the main report were used alternately in various surveys. School-related disorder was assessed by item 6 in three cities (Oakland, Birmingham, and Madison). Responses to item 7 were used in three other cities (Houston, Newark, and Baltimore). Both items were asked in Denver, so the results of the two there were averaged for inclusion in the disorder measure. The correlation between the two questions in Denver was +.57, and among the 479 respondents there who answered both questions, 85 percent of those who thought truancy was not a problem in their neighborhood also reported that disruption around schools was also not a problem. Item 9

was used in Houston, Newark and Baltimore, while Item 10 was used in Birmingham, Oakland and Madison. As a result of these substitutions, the social disorder scale score used here was based on eight component measures. Table A-7 summarizes the distribution of the eight component measures making up the disorder scale for the seven cities.

Table A-7
Distribution of Social Disorder Measures Across Cities

	drinking	gangs	groups	insults	drugs	school	vandals	assault
Houston	✓	✓	✓	✓	✓	✓	✓	✓
Newark	✓	✓	✓	✓	✓	✓	✓	✓
Denver	✓	✓	✓	✓		✓		
Birmingham					✓	✓	✓	✓
Oakland					✓	✓	✓	✓
Baltimore	✓	✓	✓	✓	✓	✓	✓	✓
Madison					✓	✓	✓	✓

An important question is how valid these survey-based measures are. One way to judge this would be to match respondents' impressions with carefully structured observations of the same areas, but that was not done. There is little research anywhere on the correspondence between perceived social disorder and independently observed neighborhood events. However, neighborhood conditions such as litter, graffiti, and building abandonment are easier to count and compare to residents' perceptions. Ralph Taylor and his colleagues (1985) have conducted extensive surveys and observational studies of neighborhoods in Baltimore. For one project they had student observers make carefully controlled counts of litter, graffiti and building abandonment in 66 areas. These observations were very substantially correlated with survey measures of the perceived extent of the same set of problems. This suggests that survey respondent's reports of local conditions can be used as fairly accurate measures of the "objective" conditions around them. The measures summarized above in Table A-4 all are substantially correlated (and in fairly similar fashion) with a set of criterion variables they theoretically should impact (cf. Skogan, 1990). These include indicators of neighborhood satisfaction (correlations ranging from -.27 to -.38) and fear of crime (range -.21 to -.33). This may be interpreted as evidence of their construct validity.

Measuring Program Exposure

The statistical analysis that follows examines the impact of living in an area where any of a number of special policing activities took place. Each respondent will be flagged by dichotomous measures indicating whether or not they lived in areas where foot patrols or door-to-door visits were conducted, storefront offices were opened, or intensive enforcement tactics were pursued. In areas in which multiple programs were fielded, respondents will be identified as having been subjected to all of them. For example, respondents in the area in Oakland in which both door-to-door visits and intensive enforcement programs were conducted will be "tagged" as being exposed to both efforts. Table A-8 summarizes the number of respondents that were exposed to each of the program elements. The correlations among the dichotomous program exposure measures ranged from -.05 to .34, and the correlation between them and the program strength ranged from .25 to .51. None of these presented significant

multicolinearity problems. Because control area residents are also the group which is "excluded" by the four program exposure measures, their scoring on the program implementation measure (zero) is effectively neutralized.

Table A-8
Number of Two-Wave Panel Respondents
Exposed to Each Strategic Element

City	Control Areas	Foot Patrol	Home Visits	Storefront Offices	Intensive Enforcement
Houston	291	-	673	330	-
Newark	330	323	323	323	630
Denver	-	-	-	-	390
Birmingham	-	-	163	144	131
Oakland	94	-	279	-	129
Baltimore	222	377	187	-	-
Madison	387	-	-	340	-
Total	1324	700	1625	1137	1447

Note: some respondents were exposed to multiple strategic elements, and are included in more than one column.

An Alternative Analysis: Before-After Change

A simple comparison of levels of perceived social disorder before and after these programs were implemented suggests that they all were associated with improving neighborhood conditions. Table A-9 examines before-after change scores that were calculated by subtracting each individual's post-implementation score from their earlier response. A positive value for this measure of change indicates that social disorder increased over that period, while a negative score indicates that it declined. Table A-9 presents average change scores for respondents exposed to each type of program, further divided by the strength of the program's implementation. There is also a summary analysis pooling all of the programs, which also divides respondents by the strength of the program in their area. A test of the significance of differences associated with program strength is presented in each instance; all of them were significant.

Table A-9 suggests that stronger programs had more beneficial results that did weaker ones, and that all of the strong programs were associated with lower levels of disorder. The weak door-to-door and enforcement programs were all associated with slight increases in disorder, and comparisons from left to right in Table A-9 reveal generally declining levels of disorder with increasing program strength. Where there are data for all three levels of program implementation (including for the over-all treatment measure), the differences in means presented in Table A-9 are also linear; there was no evidence of a significant deviation from linearity, moving from category to category. Among the strong programs it would be difficult to choose the most effective, for the declines in average disorder were fairly similar in each instance. The average change score for the six control groups is not given in Table A-9; it was .03, a positive value that indicates that average levels of social disorder tended to go up (very slightly) between the two waves of interviews in the control areas.

Table A-9
Difference Between Wave 1 and Wave 2 Social Disorder Scores
by Program Type and Strength of Implementation

	Strength of Program Implementation			
Program Type	Weak	Moderate	Strong	(Signif-icance)
foot patrols		.01	-.11	(.002)
storefronts		-.02	-.16	(.000)
door-to-door visits	.06	-.07	-.14	(.001)
intensive en-forcement	.07	-.05	-.11	(.000)
all programs	.07	-.03	-.14	(.000)

Table A-10 presents change scores for racial subgroups, for the various program exposure measures. Like the multivariate analysis of race by program interaction effects, it points to program effects among blacks which were not characteristic of whites, in two instances: in foot patrol and intensive enforcement areas.

Table A-10
Difference Between Wave 1 and Wave 2 Social Disorder Scores
by Program Type and Racial Group

	Change Scores for Racial Groups		
Program Type	Whites	Blacks	Hispanics
foot patrols	.16	-.04	-.02
home visits	-.14	-.05	-.22
storefronts	-.11	-.13	-.14
intensive en-forcement	.03	-.06	.03
control areas	.05	.01	.01

However, these simple change scores are confounded by the mix of programs that were fielded in some areas. For example, respondents exposed to intensive enforcement efforts in Denver, Birmingham, and Newark, are combined with residents of areas in Newark and Oakland who were also exposed to community policing programs. The foot patrol group includes residents of two neighborhoods in Baltimore that were served by new foot beats, but it also includes residents of Ombudsman neighborhoods in Baltimore, which featured foot patrol, and an area of Newark where foot patrols were combined with intensive enforcement, a storefront office, and door-to-door visits. It also does not take into account what we can learn from respondents who were not exposed to any new policing effort. To make use of all this sometimes overlapping information, it is necessary to define the analysis problem more carefully in a multivariate statistical model.

CITATIONS

Allison, Paul D. 1988. "Change Scores ad Dependent Variables in Regression Analysis." *Sociological Methodology*. San Francisco: Jossey-Bass, 93-114.

Mark, Melvin M. 1983. "Treatment Implementation, Statistical Power, and Internal Validity." *Evaluation Quarterly*, 7, 543-549.

Skogan, Wesley G. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities*. New York: The Free Press.

Taylor, Ralph, Sally Ann Shumaker and Stephen D. Gottfredson. 1985. "Neighborhood-level Links Between Physical Features and Local Sentiments: Deterioration, Fear of Crime, and Confidence." *Journal of Architectural Planning and Research*, 2, 261-275.