

**EVALUATION DESIGN
AND SURVEY METHODS REPORT**

by

Wesley G. Skogan

Institute for Policy Research
Northwestern University
Evanston IL 60208

5 December 1994

This project was supported by Grant #91-DB-CX-0017 awarded by the Bureau of Justice Assistance, Office of Justice Programs, U.S. Department of Justice. Points of view or opinions contained within this document are those of the author and do not necessarily represent the official position or policies of the US Department of Justice.

RESEARCH DESIGN

Data collection for the evaluation of CAPS began in the Spring of 1993, before the program began. To prepare for the evaluation, 1990 Census tract data were used to select large, mainly contiguous sections of the city which closely matched the demography of the five newly-announced prototype areas. These "comparison areas" were used to represent "what would have happened" in the prototype districts if there had been no CAPS program, for the program was not put in motion there until the end of the prototyping period. At various points in the evaluation we conducted sample surveys in the comparison areas, distributed questionnaires to police officers serving there, and observed conditions on samples of city blocks. The results provided a baseline against which to compare similar data from the prototype districts. In the parlance of evaluation, this was therefore a "quasi-experimental" study. The emphasis should properly be on "quasi," for a number of reasons. The comparison areas were imperfectly matched with the prototype districts, so if the program was received differently among various social groupings the population mix of the area might explain the apparent magnitude of the program's effect. So too could unique local events, political factors, and other neighborhood-level forces that influenced either the prototype or comparison areas, but not both. Because the number of cases involved is small, not the hundreds of "treatment and control" subjects used in many true experiments, differences between the prototype and comparison areas probably were important in determining how CAPS was received in different places.

The comparison areas were selected to include parts of at least three police districts, so that we could adjust at least some of our data for any independent community policing efforts that district commanders might decide to put in place during the course of the first year of the program. For this reason the comparison areas are significantly larger than the prototype districts. During the course of the evaluation we monitored events in the comparison areas to ensure that no large-scale efforts resembling the CAPS program were set in motion. Where we have data that can only be compared at the district level, the best match between the prototypes and police districts is: Districts 04 and 05 for Englewood and Austin; Districts 11 and 13 for Marquette; Districts 08 and 16 for Morgan Park; and Districts 19 and 20 for Rogers Park.

The matching factors for selecting the comparison areas were race and ethnicity, home ownership, and the percentage of residents living in buildings of 10 units or more. Because none of the CAPS prototypes included large blocks of public housing, we avoided those areas of the city. Table 1 presents Census figures that describe the extent of the initial match. The principal difficulty in matching areas to the Morgan Park District was the relatively high proportion of black home owners in that police district. The

Marquette District includes one of the only two areas of Chicago with a high proportion of Hispanic residents, making it difficult to find suitably matching census tracts. Rogers Park presented a challenge by its diversity on virtually every dimension.

The final evaluation and comparison areas are presented in Map 1. Based on this demographic matching the prototype and comparison areas proved to have extremely similar 1992 crime rates. For the evaluation, three prototype areas (Morgan Park, Rogers Park, and Marquette) have distinct comparison areas, while Austin and Englewood (which closely resemble one another) share one large control area. This was done to reduce survey costs.

SAMPLE SURVEYS

Relatively large Wave 1 surveys were conducted in the five prototype and four comparison areas of the city in the Spring of 1993, before the program began. Four areas — Morgan Park, Rogers Park, and their comparison areas — were re-surveyed in June, 1994. The remainder were re-surveyed in September, 1994. The initial areas were resurveyed first because elements of the CAPS program — principally community organizing efforts and some official beat meetings — were to begin in districts included in the comparison areas during the Summer. It appeared that CAPS activities were well advanced in Morgan Park and Rogers Park, and that the program had had a reasonable trial during the 14 months between its inauguration and the re-interviews. Interviews were conducted in the remaining five areas (three districts and two comparison areas) about 17 months into the program. Some official elements of the CAPS program were put in place citywide in July, 1994, but confusion about the implementation of the distinction between beat teams and rapid response teams, and delays in naming District Advisory Committees meant that there were, in our judgement, no visible program elements in place in the remaining comparison areas before the second round of interviews were conducted there in September.

In addition, we conducted independent city-wide surveys in the Spring of 1993 and 1994, to monitor program awareness and visibility, and assessments of the quality of police service by Chicago residents.

All of our survey work was conducted by telephone. This was principally a budgetary matter, but in addition reflected some real safety issues involved in fielding personal-interview surveys in Chicago's higher crime areas. By telephone we were able to recontact sample numbers many times (as many as 22 times), enabling us to locate and interview mobile and frequently-not-at-home individuals. We were also able to conduct interviews during the daytime and evening hours, for the convenience of respondents, and to break off the interview and resume it later when events intruded in the lives of busy respondents.

The price we paid was to exclude from consideration non-telephone households, which can be common in the poorest parts of Chicago. Table 2 summarizes 1990 Census data on the proportion of households in each of the city's 870 census tracts that had a telephone that year. It documents the strong links between race, poverty, housing, family structure, and the accessibility of people for telephone surveys. It also illustrates how dramatically phone access is restricted in higher-crime areas. Among the prototypes, household telephone access was fairly restricted (10-19 percent of households did not have a telephone) in Englewood, Austin and the Southern end of Marquette. It was very restricted (more than 20 percent of households did not have a phone) in the Northern end of the Marquette District. In Morgan Park and its comparison area, and in Rogers Park and most of its control area (with the exception of Uptown), telephone access rates stood at 92 percent or better.

Table 2
Citywide Correlates of Household Phone Ownership

pct. home owners	.50	robbery rate	-.38
pct. black	-.33	burglary rate	-.35
pct. Hispanic	-.21	vandalism rate	-.41
pct. female heads	-.64	gun crime rate	-.44
pct. less than high school education	-.55	public housing in the tract	-.30
pct. on public aid	-.67	average age of buildings	-.34
pct. income under \$15,000	-.66	average height of buildings	-.19
average family size	-.52	pct. of households with a car	.59

SOURCE: calculated from original 1990 Census data.

Panel Survey: Wave I

The first round of resident surveys was conducted by telephone between March 14 and May 7, 1993, prior to the official inauguration of the program. It was conducted by Northwestern University's Survey Research Laboratory (NUSL). Because we were targeting relatively small geographical areas that did not match telephone company switching areas, it was necessary to develop a sampling plan that reached the target areas in a cost-effective fashion. To accomplish this, half of the sample numbers for each area were selected at random from an up-to-date reverse telephone directory. Because they were listed numbers, we could ensure that those households lay within the targeted areas. The remaining sample of telephone numbers was selected by randomly scrambling the last digits of the listed numbers. Households responding to those numbers were then asked a brief series of geographical screening questions to ensure that they lived within the targeted areas. In the

end there were 1,294 completed randomly-sampled interviews and 1,278 listed-number interviews, for a total of 2,572 interviews. A total of 9,540 telephone numbers were processed in various ways to achieve this total. A total of 103 interviews were conducted in Spanish. To decrease nonresponse error, an extensive effort was made to deal with initial interview refusals and partial interviews. In all, 13 percent of the 2,572 completions came through this process. Persons 18 years of age or older were eligible to participate in the survey, and respondents were selected using a randomizing technique.

Figure 1 depicts the ultimate disposition of the sample telephone numbers selected for the survey. The cooperation rate for the survey (completions vs refusals and partial interviews) was 59 percent.

The representativeness of those who were successfully contacted and interviewed is difficult to judge, especially in light of the rapidly changing nature of many of the communities that were surveyed. Table 3 presents a rough comparison between 1990 Census data for the five prototype policing areas and the results of the Wave I survey. We calculated the Census figures by aggregating thousands of block-level parcels to match police district boundaries, and they are doubtless subject to error. In addition, our survey combined self-identified race with ethnicity, while the Census treats those as separate dimensions (for example, a significant number of Hispanics are classified by race in the Census as African-American).

Table 3 indicates that the Wave I survey, like many, underrepresented those who rent rather than own their home. This was most notable in Rogers Park (District 24). The survey was not far off the mark with regard to race, except in Morgan Park (District 22). The undercount of female-headed families was largest in Englewood, where the level was the highest according to the 1990 Census.

Table 3
1993 Wave I Survey vs 1990 Census Demographics

District	pct own home <u>census survey</u>	pct black <u>census survey</u>	pct Hispanic <u>census survey</u>	pct female head <u>census survey</u>
07	37 48	99 96	0 0	32 21
10	32 40	36 39	58 49	23 19
15	30 50	95 94	1 0	29 23
22	78 81	61 52	1 0	14 12
24	28 45	17 11	14 10	8 6

SOURCE: evaluation surveys and original calculations from 1990 Census data.

Figure 2 presents two comparisons which illustrate the "match" of respondents representing the prototype and comparison areas, on two key Wave 1 survey measures. The first is the extent to which residents of the various areas thought police in their area did a good job or very good job (as opposed to a fair or poor job) "... in dealing with the problems that really concern people in your neighborhood?" The second panel in Figure 3 depicts the percentage of respondents in each area who thought police were somewhat or very fair (as opposed to somewhat unfair or very unfair) "... when dealing with people in your neighborhood?" Figure 2 also presents 1993 citywide (English-speaking) data for the same measures. In these instances it is apparent that residents of the comparison areas for Districts 7, 10 and 15 were somewhat more positive about the police before the program began, while those in the comparison area for the 24th District were somewhat less positive. The match for the 22nd District was good in both instances. However, since the statistical analyses utilized in this evaluation take into account pre-program levels of support for the police and other factors, it is only if there is an interaction between levels of pre-program support and the program (say, if all respondents were exposed to the program but only already "pro-police" respondents were favorably impressed) that the differences illustrated in Figure 2 would be important.

Figure 3 examines another issue, the impact of the mixed-mode sampling frame used for the survey. It is generally accepted that listed number samples are somewhat biased in comparison to randomly-generated telephone samples. This is in part because of the large percentage and skewed demographics of of unlisted

residential households (about 50 percent of the total in Chicago), and in smaller measure due to the dated quality of printed directories in a defining a sampling frame for contacting highly mobile populations. Figure 3 illustrates some features of the samples we contacted via RDD and those who were selected randomly from directory listings. There were few differences between the two groups on many key measures used in the analysis, including the extent of personal victimization, fear of crime, contacts with police, and others not illustrated here. Telephone directory numbers were somewhat more likely to lead to older respondents, and they were much less likely to reach Hispanic households.

Panel Survey II

A second round of interviews was conducted with respondents to the first resident survey. Twelve items were added to the Wave I questionnaire, and 8 were deleted. A total of 1506 respondents from the original nine study areas were reinterviewed between June and September, 1994, either in English or Spanish. Respondents who were recontacted but who had moved since the first interview were screened to determine if they were still living in

the target area. If they were not, they were questioned about their old area if they had moved only recently; otherwise, they were excluded.

While there are many advantages to panel surveys, they typically suffer from significant levels of panel attrition. Given Americans' high mobility, which is compounded in higher crime areas and in places experiencing high levels of immigration, the capacity of telephone surveys to successfully recontact survey respondents after an extensive period of time is always limited. We anticipated this difficulty, but drawing a new telephone sample of the type we used to target respondents in small areas for Wave I would have cost an additional \$30,000, which was beyond the scope of our budget.

Across all areas we successfully contacted and reinterviewed 59 percent of the original survey respondents. The largest source of difficulty in finding them (12 percent of the original respondents) was that another party was using the telephone number and could not direct us to its original user. Eleven percent of the sample telephone numbers were non-working at the time of our second call, and two percent led only to an answering machine. Slightly over one percent of those we interviewed in 1993 were recontacted but had moved out of the study area some time ago, and were not reinterviewed. Four percent of the original respondents could not be interviewed because whoever answered the phone refused to cooperate, and the selected respondent never turned out to be available to be interviewed in another four percent of cases.

Table 4
Panel Reinterview Rates and Total Completions
for Evaluation Areas

Reinterview Rates	Prototype Districts		Comparison Areas*	
	Rate	Number	Rate	Number
07 Englewood	60	177	57	155
10 Marquette	51	170	55	141
15 Austin	55	166	57	155
22 Morgan Park	73	217	65	166
24 Rogers Park	58	178	56	136
All Districts	59	908	58	598

* Districts 07 and 15 share the same comparison area. It is only included once in the "All Districts" line.

Table 4 presents reinterview rates for the evaluation areas, and examines some of the strongest correlates of panel attrition. Reinterview rates varied significantly by area, ranging from 51 to 73 percent. Except for Morgan

Park, whose comparison area reinterview rate differed by 8 percentage points, rates for each pair of prototype and comparison areas were quite similar. As a result of this match, in the combined sample (summarized in the last row in Table 4) there was virtually no difference in panel attrition between the program and comparison areas.

Several demographic and social factors were strongly linked to panel attrition. These are summarized in Table 5. There was an eight percentage point difference between reinterview rates for blacks and whites (65 percent and 57 percent, respectively), and the smaller number of Wave I Hispanic respondents (283) were recontacted much less frequently. Younger respondents and those with less than some college education were less likely to be reinterviewed. Statistically, women were significantly more likely to be found than were men, but the difference between the sexes was not very large. There was a distressingly large gap in reinterview rates for home owners and renters, and higher income households were more likely to be successfully recontacted. Responses to a question about likely Wave I respondents were to be living in their neighborhood in a year were also good predictors of attrition, and attrition rates were higher among people who had already moved fairly frequently. Finally, interviewer judgments of the interest that respondents showed in the Wave I interview were significantly linked to Wave II completion rates. In addition to respondent interest, factors such as their cooperativeness at Wave I, interviewer assessments of the accuracy of their answers to factual questions, and their response to a question concerning their willingness to be recontacted in the future were also linked to attrition. Respondents who were initially selected from reverse directories were easier to reinterview than those initially selected via RDD.

In a multivariate analysis (which is not shown) differences between whites and blacks disappeared. Net of other demographic factors, Hispanics were less likely to be recontacted, as were younger, less educated, male renters.

Table 5
Panel Reinterview Rates for Social Groups

Selected Wave I Correlates of Panel Attrition and Percent Successfully Reinterviewed					
Race		Age		Education	
whites	65	18 to 26	41	no high school	54
blacks	57	27 to 40	54	high school	56
Hispanics	46	41 to 87	66	college	67
Others	60				
Income		Housing		Gender	
under \$20,000	54	renter	49	male	56
over \$20,000	64	owner	67	female	60
How likely to living here in one year?		Length of residence		Interest in the first interview?	
very unlikely	33	under 2 years	43	very interested	63
somewhat likely	56	2 to 7 years	55	somewhat interested	53
somewhat likely	58	8 to 20 years	64	not interested	36
very likely	65	20 years plus	66		

City-wide Surveys

City-wide interviews were conducted to monitor program awareness and visibility, and to monitor changes in the public's assessments of the quality of police service in Chicago. The questions were appended to an ongoing Chicago Area Survey Project (CASp) conducted each Spring by NUSL, so sample sizes and the number of questions that could be included were shaped by the design of the over-all survey. The surveys were conducted by telephone using RDD samples. Call-backs and other checks were made to verify the accuracy and completeness of the interviewing. One respondent 18 years of age or older was interviewed in each responding household. They were randomly selected from among all eligible respondents in the household using a standard random respondent selection technique.

The first city-wide survey was conducted from May 2 to June 16, 1993. A total of 460 respondents living in the City of Chicago were interviewed, along with 593 suburban respondents. Our special policing questions were included only in the Chicago version of the questionnaire. A total of 1597 sample telephone numbers were processed in various ways. Sample numbers were called up to 28 times to identify their status. The completion rate from all eligible households was 60 percent. The cooperation rate of the survey (completions versus refusals and partial interviews) was 67 percent. To decrease nonresponse error, an effort was made to recontact households that initially refused to be interviewed or broke off the interview while it was in progress. Almost 30 percent of those cases were completed, and in all, about 20 percent of the completions came through this process.

An important feature of this survey is that it was conducted only in English. Unlike our panel and 1994 city-wide surveys, Spanish language interviewing was not possible. As a result, all of the city-wide comparisons between 1993 and 1994 that are presented in this report exclude Spanish-language interviews conducted in 1994, to increase the comparability of the two data points.

A second survey was conducted between April 24 and June 5, 1994. A total of 1,314 residents of the City of Chicago were interviewed via RDD sampling; our police-related questions were not included on the questions administered to suburban residents. Two versions of the survey were fielded. One sample of 711 respondents was administered the full CASP questionnaire which integrated our policing supplement. Another 603 respondents were administered a shorter police-and-demographics version which dropped many of the omnibus items included in the full metropolitan area study.

Sample numbers were re-dialed up to 30 times to determine their status. Like most RDD surveys, a large number of the randomly-generated sample telephone numbers (27 percent) proved to be non-working. An additional 13 percent were non-residential, and almost 5 percent reached only an answering machine and no person could be contacted. Whoever answered refused to participate in the survey at about 8 percent of the sample numbers, and 4 percent of the randomly selected respondents refused to cooperate. Another 6 percent of numbers led to selected respondents who were then never available to be interviewed. Interviews were conducted in both English and Spanish, so the number of language-related noncompletions was very small. To decrease nonresponse error, an extensive effort was again made to deal with initial interview refusals and partial interviews. In all, about 18 percent of the completions came through this process. Overall, the survey cooperation rate (completions vs refusals) was 67 percent, and the completion rate for known eligible households was 60 percent.

Survey Weighting

Surveys that interview only one respondent per household under-represent respondents living in multiple-adult households. In addition, RDD samples over-represent respondents living in households with multiple telephone numbers. Both of these factors are sources of unequal selection probabilities for different respondents. It is common to account for unequal selection probabilities by weighting the data to correct for both factors when it is being analyzed. However, while weighted data offer better descriptive representations of the populations from which they were drawn, they are unsuited for analyses which rely on tests of statistical significance. Explorations of the panel and city-wide surveys indicated that the results of

descriptions drawn from weighted and unweighted data generally differed only at the decimal place, so only unweighted results and appropriate tests of significance are presented here.

Table 1
Selected Demographics of Evaluation Prototype and Comparison Areas

1990 Census	percent black	percent Hispanic	percent under 18	percent own home	percent large buildings
Englewood					
District 07	99	1	34	37	11
Comparison 07*	92	6	37	33	28
Marquette					
District 10	36	58	36	29	6
Comparison 10	37	46	34	23	16
Austin					
District 15	95	1	32	30	22
Comparison 15*	92	6	27	33	28
Morgan Park					
District 22	61	1	25	78	6
Comparison 22	50	11	28	61	8
Rogers Park					
District 24	16	14	21	28	56
Comparison 24	10	23	19	24	51

* NOTE: Districts 07 and 15 share the same comparison area.

This is not a blank page.

Map 1 CAPS and Evaluation Control Areas

goes here.