*SOCIAL EVALUATION RESEARCH: THE EVALUATION
OF TWO POLICE PATROLLING STRATEGIES*

JOHN F. SCHNELLE,¹ ROBERT E. KIRCHNER,
M. PATRICK MCNEES, AND JERRY M. LAWLER

MIDDLE TENNESSEE STATE UNIVERSITY, NASHVILLE METROPOLITAN
POLICE DEPARTMENT, AND RUTHERFORD COUNTY GUIDANCE CENTER

In most social evaluation research it is difficult to achieve the degree of experimental rigor possible in an applied behavioral study. This study illustrates how the evaluation researcher can increase experimental rigor in the analysis of social interventions. In the first evaluation, a variation of the time-series design that offered maximum experimental control given the limitations of the situation, was employed to evaluate the effects of a specialized home-burglary police patrol. This design revealed that no effects could be attributed to the patrol. In the second evaluation, a multiple baseline-like design was possible in determining the effects of a police walking patrol. This design revealed that the patrol produced an increase in crime reporting but not in arrests. Social interventions often occur in a manner that allows varying degrees of experimental analysis. The evaluation researcher must attain optimal experimental analysis given the limitations of each social intervention.

DESCRIPTORS: community setting, social evaluation research, multiple baseline, walking patrol, time-series analysis

The importance of applying quasi-experimental and experimental design technology to the evaluation of legal reforms and social interventions that affect large numbers of people has been emphasized by a number of authors (Campbell, 1969; Glass, Willson, and Gottman²). Campbell³ in particular has argued eloquently for an experimenting society that utilized experimental and quasi-experimental logic to make decisions about the impact of social programs.

The design most favored in such evaluation research is familiar to the field of applied behavior analysis, *viz.* the time-series design (Gottman, 1973; Jones, Vaught, and Reid⁴; Risley

and Wolf, 1972). This design involves repeated measures before, during, and after intervention. Its primary strength is its applicability to many types of field situations that preclude the use of any design requiring strong treatment control. For example, Glass, Tiao, and Maguire (1971) evaluated the effects of a change in German divorce laws on divorce-rate data retrieved from archival records. In that situation, there was no possibility of randomly assigning the law change to different sections of the country, nor was it possible to stagger sequentially the intervention in multiple-baseline style. The most powerful form of analysis available was the time-series design.

Schnelle and Lee (1974) demonstrated the application of a time-series design to the evaluation of a prison disciplinary intervention. The

¹Reprints may be obtained from John F. Schnelle, Box 438, Middle Tennessee State University, Murfreesboro, Tennessee 37130. The authors wish to thank Dr. Gene V. Glass of the Laboratory for Educational Research, University of Colorado, for the direct technical assistance and instruction he provided in statistical analyses.

²G. U. Glass, U. L. Willson, and J. M. Gottman. *Design and analysis of time series experiments*. Unpublished manuscript, laboratory of Educational Research, University of Colorado, 1972.

³D. T. Campbell. *Methods for the Experimenting Society Evaluation Research Paper Series*. Northwestern University, Evanston, Illinois, 1971.

⁴R. R. Jones, R. S. Vaught, and J. R. Reid. *Time-series analysis as a substitute for single subject analysis of variance designs*. Paper presented at American Psychological Association Convention, Montreal, 1973.

disciplinary policy was directed toward inmate rule infractions and was applied to the entire prison at one time. There was no possibility of reversing the intervention, randomly selecting control groups, nor of sequentially applying the disciplinary technique to multiple baselines. In short, time-series technology offered the only means of retrieving evaluation data.

The primary disadvantage of the time-series design when compared to several stronger time-series design variations is the possibility of the existence of historical validity threats. Campbell and Stanley (1966) defined historical validity threats as "specific events occurring between the first and second measurement in addition to the independent variable." Reversal and multiple-baseline designs are variations of the time-series design that minimize such historical possibilities by controlling when treatment is administered and to whom treatment is delivered. With the latter designs, historical threats are rendered improbable by the fact that either the treatment is applied more than once and/or control time-series lines are present.

In addition to the reversal and multiple-baseline designs, there are three other variations of the simple time-series design that enhance interpretative power by providing a measure of control for historical threats. The first variation is labelled by Cook and Campbell⁵ as an interrupted time series with nonequivalent dependent variables. Historical interpretations are minimized by collecting time-series data for some dependent variables that should be affected by treatment and for others that should not. The nonequivalent dependent variables should be conceptually related. Ross, Campbell, and Glass (1970) applied this design in the evaluation of a British breathalyzer crackdown, a concentrated effort to keep drunk drivers off the roads and hence hopefully to reduce automobile accidents.

The simplest analysis of the "Breathalyzer Crackdown" would be to compare measures of pre-intervention accidents to measures of postintervention accidents. Even if there were a change in accidents at the time of the intervention, such a time-series analysis would not separate the effects of the intervention from historical effects such as weather conditions, safer cars, newspaper publicity, *etc.*

Ross *et al.* (1970) noted that pubs in Britain are open only during certain hours. Thus, separate time-series data could be collected on traffic accidents occurring when pubs were open and when they were closed. (The importance of the distinction between accidents when pubs are closed or open is that historical interpretations of a decrease in fatalities should apply to all fatalities, irrespective of the time of day in which the accident took place.) Thus, if a change in accidents took place at the time of the intervention during open hours but not during closed hours, the evidence is strong that the intervention and not historical factors produced the change.

The second time-series variation supplements the simple time-series design with data from a nonequivalent, no-treatment control group. The major difference between the latter design and the multiple-baseline design is that control-group data are eventually treated in multiple-baseline designs, but not in the time-series variation here discussed.

Lawler and Hackman (1969) illustrated this design in their evaluation of participative decision-making concerning bonus money for work attendance. In this study, one set of workers made a group decision to award themselves extra money for work attendance. The effect of this decision and bonus was evaluated by collecting time-series data on attendance before and after the decision. A second group of men who did not engage in participative decision making were awarded extra money, and served as a control for historical threats.

Campbell and Ross (1968) evaluated a crackdown on speeding in Connecticut by comparing

⁵T. D. Cook and D. T. Campbell. *The design and conduct of quasi-experiments and true experiments in field settings*. Unpublished manuscript, Northwestern University, 1975.

change in time-series fatality measures in Connecticut to time-series fatality measures in several adjoining states. Once again, if change occurred at the point of intervention in Connecticut but not at the same time in the other states, it becomes unlikely that historical factors account for the change. At the very least, controls are present for historical factors common to all the states.

A still stronger variation of the previous two time-series designs results when treatment is applied to several time-series lines simultaneously while control series data are collected on either conceptually related dependent measures or on no-treatment control groups. Thus, in the speeding crackdown, if several states had simultaneously initiated speeding crackdowns, then Campbell and Ross (1968) could have evaluated changes in all treated states against changes in states that did not initiate the crackdown. In such a situation, historical interpretation of change are further minimized and treatment effects are replicated.

The third variation of the time-series design is the multiple-baseline design, which involves sequential intervention on all time-series baselines. Like the second design described, treatment effects can be interpreted by comparing pretreatment time-series data to posttreatment time-series data and by comparing treated baselines to untreated baselines. Finally, in both designs there are replications of treatment effects.

The major advantage of the multiple-baseline design seems to be that the comparability of the independent baselines is demonstrated only in the multiple-baseline design. In other words, if some baselines are treated and some are not, there is always the chance that the untreated control baselines are "dead" measures, and hence insensitive to any type of influence, either treatment or historical. If this premise is true, then untreated baselines are not adequate controls for the treated baselines in the control series design.

To this point, several time-series designs have been described that are applicable to the evaluation of a wide range of social interventions. The

remainder of this paper demonstrates the application of two of these designs to the evaluation of two police patrol strategies.

In neither patrol intervention was the evaluator given the opportunity to design the evaluation research strategy freely. Instead, the methods by which the intervention had to be introduced were first specified and the evaluation task was to design the most powerful research strategy given these limitations.

The first study evaluated the effect of police saturation patrolling, in which several critical patrol zones were intervened into simultaneously. The design involved comparing time-series data on the multiple treated zones to both nonequated control groups and nonequivalent dependent measures. The second study evaluated the effects of a police walking patrol that was introduced into patrol areas at separate times. A multiple-baseline design was employed in the latter situation.

EVALUATION I: THE EFFECTS OF SPECIALIZED HOME-BURGLARY SATURATION PATROLS ON BURGLARY RATES

It has been said that police burglary patrols have three functions: (1) to guard businesses and residences, (2) to watch suspected burglars, and (3) to move quickly and catch burglars on the scene (Hanna, 1974). Additionally, the *Crime-Specific Handbook* (Hanna, 1974) recommends that police sometimes shift normal patrol forces into target zones to provide saturation coverage for particularly troublesome areas.

The Kansas City Preventive Patrol experiment (Kelling, Pate, Deckman, and Brown⁶) attempted to determine the effects of varying patrol strategies and indicated the possibility that saturation patrolling would have no effect

⁶G. L. Kelling, T. Pate, D. Deckman, and C. E. Brown. *The Kansas City Preventive Patrol Experiment: A Summary Report*. Police Foundation, 1909 K Street, N.W., Washington, D.C.

on crime rates. In five "reactive" beats, vehicles entered the beat only in response to calls for service. In five control beats, the normal patrol level of one car per beat was continued. In five "proactive" beats, two or three marked cars patrolled in each beat. No effects on reported crime were found due to increasing or decreasing patrol strength in the beats.

The present investigation dealt with the effects of increasing patrol strength in particular zones, and focused on one particular type of crime, namely home burglaries. Differences between the present investigation and the Kansas City experiment are that (a) patrol strength in the present study was much higher than in the Kansas City experiment, (b) the specialized patrol focused only on home burglaries, (c) unmarked cars and plainclothes personnel were used in the target zones, and (d) before the intervention, the specialized burglary patrol received specific information about people with an arrest history for burglary.

METHOD

Setting and Subjects

This investigation was carried out in Davidson County, Tennessee. Nashville, with a population of 449,000, is the largest city in Davidson County. For police purposes, Davidson County is divided into three geographic areas called "sectors". The sectors are divided into a total of 32 zones (14, 10, 8 zones in each of the three sectors). The total Metropolitan Police force (sworn police officers) numbers 862. Thirty-five patrolmen were assembled into one group from various branches of the department to form a specialized burglary patrol.

Patrol preparation. After the 35 men had been established as a specialized burglary patrol, they were briefed on the specifications of their assignment, which included (a) the assignment of each man to one of three zones, (b) only residential burglaries during the second shift (8 a.m. to 4 p.m.) were targeted, and (c) patrolmen would work as two-man teams. Additionally,

each patrolman was given a book containing the identity, pictures, and information on individuals previously arrested for burglary.

Deployment and patrolling procedures. The 35 patrolmen selected for the burglary patrol were assigned to three zones (zone 31, zone 71, and zone 91) thought by the police to be particularly "troublesome". Zone 31 was 2.96 miles from zone 71 and 4.23 miles from zone 91. Zone 71 was 6.77 miles from zone 91. Before the burglary patrol was initiated, two patrolmen (one car) covered each zone. During the saturation patrol, eight to 15 men patrolled each of the three target zones during the second shift. The variance in numbers was due primarily to court appearances by the patrolmen.

On January 23, 1974, the burglary patrol saturated the three target zones and was terminated on March 1, 1974.

Data collection. Information was gathered on home-burglary rates for the three shifts in the target zones and home-burglary rates for three randomly chosen comparison zones. Burglaries are defined by the Nashville Metropolitan Police Department as reported instances in which the following two elements are present: (a) evidence for forced entry, and (b) citizen report of missing property resulting from the forced entry. The crime monitoring system used by the Nashville Metropolitan Police Department contains inherent reliability checks that ensure accurate recording of citizens report of burglary.

A burglary complaint is initiated when a citizen calls the central police telephone number. A police radio operator receives the complaint and dispatches a squad car to the scene. All incoming calls to the police department are automatically tape recorded. This tape is periodically checked to ascertain how well citizen calls are received by the complaint clerk. The patrolmen arriving on the scene must write a general report describing the situation. In the case of burglary reports, specific mention must be made of the method of entry into the burglarized environment and a description of stolen property. These general reports are turned in to a centralized

point at the end of a shift, and, based on the information specified in the general report, the incident is checked for proper classification as to the type of crime committed. Two people are responsible for making the classification independently. All situations in which forced entry was evident and in which something is reported to be missing are classified as burglaries. Periodically, classified reports are checked by the police Inspections Section for consistency and accuracy of the classification system.

Finally, in the case of burglaries, the general reports are turned in to the Detective Division, which is given seven days to make a supplementary investigation and report. This supplementary report is designed to check the information in the first general report and to add further information. In the case of minor burglary loss, the supplementary investigation may be conducted by telephone. The officer making the initial general report has no way of knowing which detective will make the supplementary report. Based on the supplementary report, the reported crime may be reclassified. These internal checks ensure that there is little chance of either fabricating a report of burglary that did not occur, or of neglecting to record a reported burglary.

RESULTS

The upper three panels of Figure 1 show the weekly home-burglary rates for the second shift in the three target zones before, during, and after the burglary patrol. In all three target zones, initiation of the burglary patrol did not decrease burglaries substantially in those zones. Employing a time-series analysis computer program provided by Gene V. Glass, the data were analyzed statistically. There were no significant changes in the level or slope of the time-series lines at the point of intervention. The change in level T tests were: $T = -1.92$, $T = 1.55$, $T = 0.52$, $D.F. = 55$, $p > 0.05$. The means of the pre-intervention phase and intervention phase for the three target zones were 2.8, 2.9; 4.0, 3.4; and 1.9, 2.0, respectively.

The middle three panels of the figure illustrate the home-burglary rates in the target zones for a shift other than when saturation patrol was present.

Again, no significant changes in the level or slope of the time-series lines were found. Level change T tests were: $T = 0.40$, $T = 0.14$, $T = -0.78$, $D.F. = 55$, $p > 0.05$). The means of the pre-intervention phase and intervention phase were 2.3, 2.0; 1.9, 2.0; and 1.4, 1.6, respectively.

The lower three panels of the figure show home-burglary rates for the three randomly chosen control zones. No significant changes in the level or slope of burglaries occurred during the saturation of the target zones. Level change T tests were: $T = 0.70$, $T = 1.62$, $T = 0.35$, $D.F. = 55$, $p > 0.05$. The means of the pre-intervention and intervention phases were 2.4, 3.0; 2.06, 3.6; and 1, 0.3, respectively.

The time-series statistical analysis is necessary to test for changes in either the slope or level of data collected over time because of the typical interdependency nature of the time-series measures. This interdependency, unless corrected, would lead to deflated various estimates and hence spuriously large T or F ratios if such ratios were calculated by the usual methods. Time-series analysis leads to an adjusted T ratio in the sense that the time-series scores are transformed into scores with uncorrelated error terms. These transformed scores can then be tested for differences in level or slope between phases. The time-series analysis proceeds by first identifying the basic dependency nature of the time-series measures by means of a correlogram. A lag 1 correlation shows the correlation between each measure and the measure one step removed from that measure. Lag 2 correlations show the relationship between each measure and the measure two steps removed, *i.e.*, observation 1 correlated with observation 3, and so on to all possible correlational lags within a time series. A typical type of dependency relationship as revealed by such correlogram analysis would be an autoregressive relationship, *i.e.*, the next value in the

Fig. 1. Upper three panels show frequency of residential burglary reports for the three target zones and patrol shift in which the saturation patrol took place. The middle three panels illustrate frequency of residential burglary reports in the three saturated patrol zones during a shift in which the saturation patrol was not present. The last three panels show frequency of residential burglaries for three randomly selected comparison zones that did not receive the saturation patrol.

time series is predicted by a constant times the previous score plus unpredicted noise. Next, the time-series program performs a complete normal theory, linear least-squares analysis for all specific values of Φ and Θ , which are the specific parameters describing the basic generating function of the time series. The computer program defines the best generating function of the time-series line by showing the specific values of Φ and Θ for which the residual sum of squares is minimized. At this point, T test for significant changes in level or slope, given the basic generating function of the time series, can be calculated and are calculated by the program.

Linear regression curve fitting is better known in the social sciences than is the latter process but is not recommended for time-series data because of: (1) the underlying relationship of time-series data is usually not linear and (2) the assumption that the repeated observations are in-

dependent samples cannot be met. The generating function procedure used in the present time-series analysis clarifies the manner in which the time-series data were generated by taking into consideration the dependency nature of the series.

Figure 2 shows the average daily burglary related arrests for the three experimental and three control zones. Burglary related arrests were defined as arrest for first-, second-, and third-degree burglary, receiving and concealing stolen property, and possession of burglary tools. Receiving and concealing stolen property was counted as a burglary related offense because in Nashville, a burglary charge is often reduced to receiving and concealing stolen property when investigators or prosecutors do not feel that they have adequate evidence to press for a guilty verdict. The frequency of arrests for burglary related crimes was collected for all shifts from the

Fig. 2. Frequency of arrests for three saturated zones *versus* three control zones.

three saturated zones and the three comparison zones for 15-weeks before the saturation patrol and for all weeks during the saturation patrol. Separate time-series analyses were conducted for the combined weekly arrest data from the experimental zones and for the combined weekly arrest data from the three comparison zones. There was a significant increase in the level of arrests for the saturated zones ($T = 2.81$, D.F. = 16, $p < 0.05$) but not for the comparison zones ($T = 0.61$, D.F. = 16, $p > 0.05$). There were no changes in the slopes of the time-series lines. Once again, the tests were accomplished by time-series analysis.

DISCUSSION

In this evaluation, development and use of a specialized home-burglary patrol produced no changes in burglary rates in three zones selected for saturation. The inclusion of data on comparison shifts and comparison zones indicates no overall increase in burglary rates at the point that the target zones were saturated. The lack of change in the comparison shifts and zones rules out the possibility that the burglary patrol actually suppressed what would have been an increase in burglary rates. In other words, the effects of the patrol on burglary rates in the target zones were not obscured by general increases in burglary rates. Visual inspection of city wide time-series data on burglary rates reaffirmed the comparison zone data, which showed no increase in burglary reports.

By looking at only several data points before and several data points after the intervention in each saturated zone, police officials were led to believe that the saturation patrolling produced a decrease in burglaries. This false belief is revealed by the fact that three articles were published in Nashville newspapers quoting police spokesman that the saturation patrol reduced burglaries. In fact, there was a statistical shift downward in all saturated zones over a short time span on either side of the intervention that can be largely accounted for by regression arti-

facts. In other words, the police intervened near a time period when burglaries were near a peak in all treated zones, and thus a statistical shift downward can be predicted. Campbell and Erbebacher (1970) and Campbell and Stanely (1966) provide other examples of how regression artifacts can lead to misleading conclusions about treatment outcomes.

The no-effect finding of this study gives additional support to the contention stimulated by the Kansas City patrol experiment (1974) that various commonly used patrol strategies have little if any effect on crime. The results of the Nashville evaluation do not provide support for some recommended practices to reduce burglaries, such as: that patrols should provide saturation coverage for troublesome burglary areas and that patrol officers be provided with lists, pictures, and information on suspects frequenting the patrol areas (Hanna, 1974).

In discussing these results a cautionary point should be noted. The results do not indicate that the specialized burglary patrol had no effect on people's lives in the target zone or the city. Also, generalizations to other types of crime or patrolling strategies should not be made at this point. However, the results do indicate that the specialized burglary patrol did not serve its expressed function: to reduce burglaries in the target zones.

It is also interesting that the saturation patrol led to an increase in burglary arrests. The fact that arrests increased provides direct evidence that the police were actively involved in the saturated zones but also raises several unanswered questions about the relationship between arrests for specific crimes and the frequency of these crimes.

EVALUATION II: THE EFFECTS OF POLICE WALKING PATROL ON REPORTED CRIME

A fashionable innovation in police department law enforcement work is the walking patrol. Even though there are variations in differ-

ent cities, the basic idea of the walking patrol is to put police officers in personal contact with the citizens they serve (O'Connor and Vanderbosch, 1967; Whesehand and Cline, 1971).

The walking patrol in Nashville, Tennessee, involved police officers located in a centralized "storefront" office of two patrol zones from 9 a.m. to 5 p.m. The introduction of the walking patrol occurred sequentially, being started in one area some six months before being introduced in the other. This sequential introduction, made necessary by manpower limitations and cost factors, provided the opportunity to evaluate the effects of patrol by multiple-baseline technology.

METHOD

Subjects and Setting

The walking patrol experimental areas were picked by the Nashville Housing Authority as being most appropriate and most in need of the walking patrol. The Nashville Housing Authority requested that patrol area 110 be worked with first because of greater need. Statistics supported the contention that area 110 experienced more crime than the other area chosen.

The walking patrol concept was first introduced in both patrol areas in a tenants' meeting. Within two months after these tenant meetings (which occurred six months apart in the two zones), the police department rented office space within the patrol zone and walking patrol started immediately afterwards.

Population density, socioeconomic levels, and age groupings were similar in both areas, each consisting of low-income government housing with a predominately black population. A history of poor police-community relations was evident in the two patrol zones. There were three white and three black officers in each patrol zone.

Measures. Since the police were primarily concerned with the patrol's effect on crime and arrests, the measures employed to evaluate the walking patrol were reported offenses and ar-

rests. Before the walking patrol, offenses were reported by one method only: citizens calling the city police number to report an incident. These telephoned incidents were classified by a complaint clerk and given immediately to a dispatcher, who dispatched a patrol car to investigate. The reports of the officer were written, classified as to the type of crime committed, and fed into an information retrieval source. All calls that were dispatched required a follow-up report by the investigating patrol officer. One hundred per cent of the calls resulted in matched follow-up reports. Arrest data were collected directly from arrest reports filled out by the officer at every arrest. These data were also computerized.

After the walking patrol was initiated, offense statistics were collected as usual by citizens' calls to a centralized police number. In addition, complaints could be taken directly to a patrolman working within the walking patrol zone. These complaints were written on a daily log by the walking patrolman and eventually fed into a computer system. The police radio operator and the dispatcher were not involved in the latter process. Arrest statistics were collected from arrest reports. There was no change in the method of writing arrest reports.

Reliability checks. Reliability checks of reported crime by two independent observers were not possible in this study. The two reporting sources who could be in error were the radio operator receiving the telephone complaint calls and the patrol officers receiving direct citizen complaints. The most probable recording error would be neglecting to write up a complaint, rather than writing up a complaint that did not actually occur, since a follow-up report, citing names, times, and addresses, was made for every recorded incident. These follow-up reports were often done by officers other than the officer who initially recorded the complaint information. Thus, a fake complaint write-up would be readily uncovered by a follow-up report. It was not possible for an officer to predict exactly when a follow-up would be done. Thus, the follow-up

reporting procedure served as a reliability check for writing up false complaints. Throughout this study, there was no evidence of a follow-up investigation that did not confirm the fact that a citizen actually made a complaint.

In addition, a random sample of people who reported offenses was called upon by a member of the police Inspections Section, who asked a series of questions related to how satisfied the complainant was with his interaction with the police. Once again, officers could not predict when such a consumer index would be taken, and there was no instance of report falsification in the consumer surveys.

That an officer might neglect to write up a complaint was possible but improbable, because if an officer neglected to write up an offense, then the citizen reporting the offense could complain to the police department's Internal Security Section about police inactivity in regard to the complaint. There was no change in the number of such citizen complaints during this study.

Reliability checks on arrests can be made only by checking the disposition of the arrest case. If there were a change in the number of arrests disallowed for technical reasons, or a change in arrest-conviction ratios, then the hypothesis that officers changed arrest procedures would be plausible (instrumentation threat). In this particular case, there was no evidence that officers changed normal arrest procedures. Unfortunately, with social measures such as arrest statistics, it is still not possible to disregard completely the possibility that subtle changes in arrest methods did occur.

RESULTS

The solid line in Figure 3 illustrates the offense reports taken over a 28-month and 34-month period before the walking patrol and for a six-month period in Zone 1 and a five-month period in Zone 2 after the walking patrol. Data collection in Zone 1 terminated six months after the walking patrol intervention because several new police interventions were undertaken in this zone.

It is obvious that reported offenses increased after the patrol was introduced in both patrol zones. Time-series analysis showed a significant change in the level of reported offenses in both zones ($T = 1.84$ and $T = 7.0$, $p < 0.05$), respectively. There were no changes in slope in either zone. The actual types of crime showing the most increase at the time of the walking patrol intervention were theft, simple assault, public drunkenness, and disorderly conduct. Major crime, such as murder, rape, and burglary, did not show such increases. The extent of this increase due to citizens reporting directly to the officers is illustrated by the dotted line after the walking patrol intervention. The solid line illustrates total offenses reported, and was computed by adding the number of complaints made directly to patrolmen in the zone to the number of telephoned complaints; there were no direct citizen reports to officers before the walking patrol.

Arrest frequencies did not change systematically at the time of the intervention in either patrol zone, and thus are not illustrated.

DISCUSSION

The walking patrol intervention produced an increase in reported crime, a large part of which was due to citizens making direct reports to patrolmen in addition to telephoning reports to the central police number. Thus, the increase in reported crime seems to be partially due to what Campbell (1969) called instrumentation factors. Instrumentation is defined by Campbell as changes in the calibration of a measuring instrument or changes in the observers or scores that may produce changes in the obtained measurements. Despite the fact that instrumentation changes can explain much of the increase in reported crime, the instrumentation in this case has applied significance, since citizens will seemingly report more crime if placed in close proximity with police officers. A survey by the Law Enforcement Assistance Administration (1974) indicated that unreported crime in several large cities exceeded actual reports of crime. The

Fig. 3. Solid lines indicate total reported crime in the two walking patrol zones. Dotted lines indicate direct citizen report of crime to patrolman. Direct citizen reporting was nonexistent before the intervention.

walking patrol is a logical way to increase crime reporting, even though the effect of the walking patrol on unreported crime measures was not directly evaluated in this study.

The present results were in part anticipated by Press (1972), who considered the possibility that an increased patrol strength in a precinct in New York City would produce an increase in reported crime. Likewise, Etzioni (1968) reported cases of increases in reported crime when police departments changed their reporting procedures.

It should be noted that there was a significant increase in crime reporting directly to police in both zones, but that there was a greater relative increase in telephone reports to police in Zone 2 than in Zone 1. The reasons for this difference are not apparent from this study and could logi-

cally be due to several factors, *i.e.*, the lower overall baseline crime reporting in Zone 2 as compared to Zone 1 or possibly publicity that took place in Zone 2 but not in Zone 1. Nevertheless, it is still possible to state that the walking patrol increased overall reporting of crime in both zones only at the time of intervention. In other words, the fact that the walking patrol had a replicable effect on crime reports is incontestable.

A final point that should be made clear is the important differences between the multiple-baseline design employed to evaluate the walking patrol and multiple-baseline analysis in which the evaluator controls the treatment variable. Cook and Campbell⁷ list a number of va-

⁷Cook and Campbell, *op. cit.*

lidity threats that are made probable by a lack of treatment control. Specific to the case of the walking patrol, it is conceivable that the area that the Nashville housing authority picked to undergo the intervention first, was at the time of the intervention experiencing large amounts of crime not reported because of citizen mistrust in the police who were immediately patrolling the area. This "historical" event could have contributed to the subsequent large increase in reports of crime only in that area and not in the comparison area. This validity threat possibility could be termed the interaction of selection and history and/or the interaction of selection and regression. In other words, perhaps the walking patrol would show dramatic change in crime reporting only if an area were immediately upset with its former patrol car officers or if crime reporting were suppressed to unusually low levels. If both the area and the time of intervention were picked randomly, then such interaction validity threats would be less probable. In addition, since both areas were not randomly picked from the total population of areas that potentially would have been amenable to walking patrol strategies, then questions can be raised about how the results from these two selected areas would generalize to other areas. In other words, the factors that led these areas to be selected could also have been the factors necessary for the increased reporting of crime that occurred when the walking patrol was initiated. The latter external validity problem would not be so obvious if the areas were randomly picked from the total population of possibilities.

GENERAL DISCUSSION

This report has demonstrated the application of the most powerful design technology possible in two applied situations that precluded a large degree of treatment control. In other words, the police department was not engaging in research when they conceptualized the patrol strategies, but instead were reacting to political and economic realities. No financial resources were set

aside to evaluate the effects of the patrol intervention.

These quasi-experimental evaluations produced the surprising conclusion (to the police) that (a) saturation patrol had no significant effect separate from regression artifacts; (b) the walking patrol produced an overall increase in the report of crime. In neither case was it necessary to fit the methods of initiating patrol intervention to a preconceived evaluation design to arrive at these conclusions.

The need for simultaneous intervention into several high-rate burglary zones was largely determined by public pressure on the police administration to reduce crime rates and by the existence of three zones that were simultaneously experiencing high burglary rates. The sequential intervention into two areas by the walking patrol was made necessary by the fact that the police could not afford to start the patrol in two areas at once. The selection of the walking patrol areas was determined by a history of poor police-public relations in the two zones and requests from the Nashville Housing Authority for action.

Undoubtedly, a more powerful experimental technology could have been applied to both interventions if supplementary political and financial resources were available for a true experimental evaluation. In both cases, public opinion and victimization surveys would have been helpful. Also, in both cases it would have been more experimentally sound, and more costly, if the evaluator could have randomly determined which zones received treatment, and when the treatment was delivered.

It is the point of this report, however, that until the police are given resources necessary to place their operation on an experimental basis, an already existing evaluation technology can be used at little or no cost. This quasi-experimental technology can be applied in such a way that the usual procedures that social agencies employ in their interventions do not have to be changed. In other words, the evaluation can be largely unobtrusive.

The importance of the present study lies as much in the demonstrated application of the quasi-experimental strategy as in the information it provided about patrols, even though useful conclusions about patrols were produced. To make clearer statements about police patrols it would be necessary to have a more detailed description of what the police officers actually did in the patrols. Such research is now being conducted, and it must be said that this more refined patrol research would not have been politically possible without the preliminary research described in the present report.

This paper primarily illustrates that usable data can be immediately gained by the applied behavior analyst if efforts are directed toward the evaluation of normally occurring social interventions. It would seem that such unobtrusive experimental technology is a logical first step toward placing social agencies on a sounder experimental basis.

REFERENCES

- Campbell, D. T. Reforms as experiments. *American Psychologist*, 1969, **24**, 407-429.
- Campbell, D. T. and Stanley, J. C. *Experimental and quasi-experimental design for research*. Chicago: Rand McNally, 1966.
- Campbell, D. T. and Ross, H. L. The Connecticut Crackdown on speeding: time series data in quasi-experimental analysis. *Law and Society Review*, 1968, **8**, 33-53.
- Campbell, D. T. and Erbebacher, A. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful! In J. Hellmuth (Ed.), *Compensatory education: a national debate*, Vol. III of the disadvantaged child. Pp. 185-210. Reply to replies, Pp. 221-225. New York: Bruner/Mazell, 1970.
- Etzioni, A. "Shortcuts" to social change? *The Public Interest*, 1968, **12**, 40-51.
- Glass, G. U., Tiao, G. C., and Maguire, T. O. Analysis of data on the 1900 revision of German divorce laws as a time-series quasi-experiment. *Law and Society Review*, 1971, **4**, 539-562.
- Gottman, J. M. N-of one and N-of two research in psychotherapy. *Psychological Bulletin*, 1973, **80**, 93-105.
- Hanna, J. Crime-Specific Burglary Handbook: evaluates reviews of O. C. J. P. Projects. *Bulletin of Criminal Justice Planning*, 1974, **7**, 5-12.
- Law enforcement assistance administration, Trust and the police. *Time*, 1974, **103**, 124.
- Lawler, E. E. and Hackman, J. R. Impact of employee participation in the development of pay incentive plans: a field experiment. *Journal of Applied Psychology*, 1969, **53**, 467-471.
- O'Connor, G. W. and Vanderbosch, C. G. (Eds.) *The patrol operation*. Library of Congress, Washington, D.C., 1967.
- Press, S. J. Police manpower versus crime. In J. M. Tanur (Ed.), *Statistics: a guide to the unknown*. San Francisco: Holden-Day, 1972.
- Risley T. R. and Wolf, M. M. Strategies for analyzing behavioral change over time. In J. Nesselbroade and H. Reese (Eds.), *Life-span development psychology: methodological issues*. New York: Academic Press, 1972.
- Ross, H. L., Campbell, D. T., and Glass, G. U. Determining the social effects of a legal reform: the British "Breathalyzer" crackdown of 1967. *American Behavioral Scientist*, 1970, **13**, 493-509.
- Schnelle, J. F. and Lee, F. J. A quasi-experimental retrospective evaluation of a prior policy change. *Journal of Applied Behavior Analysis*, 1974, **7**, 483-496.
- Whesehand, P. M. and Cline, J. L. *Patrol operations*. Englewood Cliffs: Prentice Hall, 1971.

Received 3 February 1975.

(Final acceptance 10 May 1975.)