*EVALUATING MUNICIPAL POLICY:
AN ANALYSIS OF A REFUSE-PACKAGING PROGRAM¹*

TREVOR F. STOKES AND STEPHEN B. FAWCETT

THE UNIVERSITY OF MANITOBA AND THE UNIVERSITY OF KANSAS

The effect of a municipal enforcement program to improve residents' packaging of refuse was analyzed in a multiple-baseline design across two areas of a city. The enforcement program involved instruction concerning refuse-packaging regulations, collection only of appropriately packaged refuse, and feedback notices to residents concerning the reasons for noncollection of their inadequately packaged refuse. Both the number of violations and the percentage of residences violating each day were markedly reduced during the enforcement program. Furthermore, sanitation workers considered that the packaging of refuse and the safety and efficiency of refuse collection had improved. This study served as a pilot evaluation of a policy change in the city sanitation department, and was functional in determining the direction of that policy.

DESCRIPTORS: municipal policy, public policy, refuse packaging, littering, trash, municipal services, consultants, evaluation research, multiple baseline, consumer satisfaction, social validation, neighborhood residents, adults.

The application of experimental and quasi-experimental methodologies in evaluating the effects of public policy changes has been the focus of some recent discussion (Campbell, 1969; Riecken, 1975; Weiss, 1972*a*). Prominent examples of policy evaluation include the examination of welfare used as a loan (Baldus, 1973), the Connecticut crackdown on speeding (Campbell and Ross, 1968), the British prosecution of persons driving under the influence of alcohol (Ross, Campbell, and Glass, 1970), changes in prison policy (Schnelle and

Lee, 1974), and police patrolling strategies (Schnelle, Kirchner, McNees, and Lawler, 1975). These post-hoc evaluations (Frank and Smith, 1976) were conducted either subsequent to or independent of decisions regarding policy formulation and implementation. The pilot testing of proposed policy changes before their formal adoption has been identified as another type of evaluation research (Campbell, 1969; Salasin, 1973). The New Jersey negative income tax experiment (Kershaw, 1972) is one example of the pilot testing of possible changes in federal social policy. Indeed, Weiss (1972*a*) characterized the contribution of evaluation research as providing a data basis for policy decisions.

The delivery of municipal services such as sanitation, recreation, transportation, and police and fire protection, requires numerous policy decisions. While the evaluation of the effectiveness of local services (The Urban Institute, 1974) and the productivity of municipal workers (National Commission on Productivity, 1973, 1975) have been topics of recent study, proposed changes in city service operations are not often subjected to formal pilot testing before implementation.

¹This research was supported in part by PHS Training Grant HD00183 to the first author, and in part by the Institute of Public Affairs and Community Development, University of Kansas. Special thanks are expressed to Merrill Stokes for invaluable assistance in data collection. Appreciation is also extended to Philip Bohlander and his colleagues in the Sanitation Workers Association for their dedicated help in the planning and conduct of this research. Buford Watson, City Manager, and Mike Wildgen, Assistant City Manager, City of Lawrence, ably assisted in the implementation of this research. Reprints may be obtained from either T. F. Stokes, Department of Psychology, University of Manitoba, Winnipeg, Manitoba, Canada, R3T 2N2, or S. B. Fawcett, Department of Human Development, University of Kansas, Lawrence, Kansas 66045.

The present study examined the effectiveness of one proposed change in the operation of a city sanitation service, *i.e.*, the modification of refuse packaging requirements. This research resulted from consultation with officials of the Sanitation Worker's Association on various problems within the sanitation department. Issues of concern to the workers and city management included the packaging of refuse by city residents, the quality of sanitation services, the safety of equipment and work procedures, job security, and worker morale. The inadequate packaging of refuse by city residents was a matter of particular concern to the workers. Therefore, this problem was selected for investigation. The authors received a mandate from the city government to conduct a pilot test of the proposed change in city policy, a municipal enforcement program for refuse packaging.

METHOD

Participants and Setting

The study examined the refuse-packaging behavior of persons living in 183 residences in Lawrence, Kansas, a community of 50,000. These residences were arbitrarily divided into two experimental areas, one on either side of a major city street. There were 94 private residences in the west, and 89 in the east, typically single-family, multiple-bedroom homes with basements and garages. The west was a middle-class neighborhood where home values ranged from \$15,000 to \$30,000, with a mean of \$21,000. The east was a mixed neighborhood of lower and upper middle-class areas. Home values ranged from \$15,000 to \$115,000, with a mean of \$49,000. Prior permission to conduct this study and involve the 183 residences was obtained from the Lawrence City Commission, the elected representatives of the community.

The experimental areas were both part of the same collection route, but did not constitute the entire route. Refuse was left by resi-

dents either at the curb in front of their house or in the alley behind their house, depending on the particular part of the route, and was collected each Tuesday and Friday. The sanitation collection crew consisted of a truck driver and two laborers who loaded refuse into the back of the truck. The four men who took part in this study ranged in age from 22 to 29 yr; three were high-school graduates, one was a university graduate.

Observation

Before the sanitation crew collected refuse, packaging violations were counted each collection day in the experimental areas. Observations occurred at the same time each day. An observer was driven by car through each section according to a predetermined route. The car stopped at each residence to allow counting of violations. Occasionally, the observer got out of the car at a residence where the type or number of violations was not clearly determined by observation from the car. Observation in the west commenced at 8:15 a.m., and in the east at 9:00 a.m. Approximately 25 min was required to score the packaging violations in each section. The sanitation crew started working the west at approximately 8:30 a.m., and started the east at approximately 10:00 a.m. Observation of individual residences was always concluded before arrival of the crew.

The presence or absence of refuse to be collected was recorded for each residence. Furthermore, if refuse was out, that residence was scored for adherence to the packaging standards. The type and number of violations were scored for residences violating the regulations.

Violation types were predetermined by the city ordinance covering sanitation service. However, the precise wording of the violations adopted for the present study incorporated the concerns of the workers and city management personnel and were defined so as to facilitate reliable observation.

There were six categories of packaging violations:

1. *Edge of container (can or bag) too far from the collection point:* more than 1.8 m from the street curb or alley line. One violation each container. Curbside collection is more efficient because of the easy access to refuse containers.

2. *Overfilled refuse containers:* refuse extending more than 22.5 cm above the top rim of a can. One violation each can. Overfilled containers are a safety hazard to crewmen while carrying containers to the truck. Protruding objects of small diameter are particularly dangerous and likely to cause face and eye injuries. Overfilled containers also cause litter problems.

3. *Untied or torn plastic refuse bags.* One violation each untied bag; when bag tied and there is a hole or tear in the bag of more than 2.5 cm in length, one violation each bag. The use of refuse bags with at least a 2-mil thickness and not containing polyvinyl chloride (Environmental Protection Agency, 1974, 1976) are recommended because they allow efficient one-way pickup and are easy to handle. However, handling difficulty increases if these bags are not tied because they are difficult to pick up and carry, and liable to spill their contents, subsequently requiring additional time to collect the spillage. Similiar problems are also caused by torn bags that easily rupture when lifted.

4. *Unacceptable refuse containers:* cardboard boxes, paper bags, pasteboard or fiberboard barrels, and wicker baskets. One violation each container. These containers are especially troublesome if they become wet, when they become unstable and difficult to carry, and often fall apart and spill refuse on the street.

5. *Uncut or untied yard trimmings:* trimmings not tied in bundles less than 45.7 cm in diameter and less than 1.2 m long. One violation, any number of occurrences. Secured bundles are easier to pickup and provide fewer safety hazards. Bundles longer than 1.2 m do not easily fit in the back of the truck.

6. *Loose litter, trash, or garbage* greater than 5 cm in diameter and within 1.8 m of refuse container. One violation, any number of occurrences.

The reliability of observation was assessed on four days (3, 5, 9, and 13) by having the driver of the observation car independently score each residence at the same time as the regular observer. Reliability of the scoring of residences was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100. An agreement was scored if both observers agreed to the categorization of a residence. The reliability of scoring the total number of violations, and the number of each type of violation was calculated by dividing the smaller frequency by the larger frequency and multiplying by 100. These comparisons yielded the following percentages of interobserver agreement: scoring of violating residences, 98.3%; non-violating residences, 96.9%; residences without refuse, 98.6%; total number of violations each day, 94.4%; and, total number of each type of violation each day, 1, 89.1%; 2, 93.2%; 3, 91.2%; 4, 93.1%; 5, 100%; 6, 84.1%.

Municipal Enforcement Program

The enforcement program involved instructions and feedback about packaging and contingent collection of refuse appropriately packaged. Instructions consisted of a letter signed by the City Manager and the Superintendent of Sanitation mailed to each resident following the last day of baseline. The letter emphasized the City's intent to enforce the packaging standards: "The sanitation crew will be required to collect only refuse meeting the packaging standards . . . If part of your refuse is packaged correctly, and part packaged incorrectly, only the refuse meeting the standards will be collected . . ." Attached to this letter was a copy of the standards for packaging garbage and trash, which detailed the violation types and provided reasons why these violations reduced the safety and efficiency of the sanitation service. In addition, each letter contained a copy of a violation tag that would be left as a courtesy to residents if part of their refuse was not collected. Feedback consisted of violation notices distributed by the

sanitation crewmen riding on the back of the truck. These notices, checklists specifying the relevant violation type, were usually attached to the handle of the refuse can. Fifteen to 30 sec was required to fill in and leave a notice at a violating residence. Nonviolating residences did not receive written feedback. Contingent collection involved collection only of adequately packaged refuse. All refuse failing to meet the minimum standards was not collected until the packaging met the standards.

Experimental Conditions

The effect of the enforcement program was analyzed in a multiple-baseline design across the west and east neighborhoods.

Baseline. Baseline observation in both sections commenced shortly after the City Commission formally approved the study and directed that the enforcement procedures be implemented in the two test neighborhoods. Additional public notification of the particular neighborhoods involved and the details of the enforcement procedures to be followed (non-collection of inappropriately packaged refuse) appeared in the local newspaper's report of the Commission meeting, which also occurred before baseline. During baseline, the sanitation crew followed normal procedures, collecting refuse packaged correctly or incorrectly.

Enforcement program in the west. Following the fifth collection day, a notification letter was mailed to each resident in the west section. The enforcement program was instituted beginning the next collection day and remained in effect through Day 15. Baseline conditions remained in effect for the east section.

Enforcement program in the east. Following the tenth collection day, similar notification was provided to residents in the east section. The enforcement program was implemented as of Day 11, and remained in effect through Day 15.

Followup. Following Day 15, a letter was sent by the City to each of the 183 residences in the experimental area. This letter thanked the residents for their cooperation in the testing

of procedural changes in the Sanitation Department and informed them of the successful conclusion of the study. The sanitation crew was instructed to discontinue the enforcement procedures after Day 15 and resume the "normal" baseline conditions. Four weeks later, on Day 23, a followup observation of packaging violations was made.

Ratings by Sanitation Crew

In addition to the data obtained from direct observation, the sanitation crewmen who collected the refuse were asked to rate the experimental sections on various evaluative dimensions. To establish the social validity (Fawcett and Miller, 1975; Minkin, Braukmann, Minkin, Timbers, Timbers, Fixsen, Phillips, and Wolf, 1976) of the effects of the enforcement program, ratings were obtained on several Likert-type scales (1932): "Overall, how well packaged was the refuse in this section? How difficult was it to collect the refuse in this section? How safe was it to collect the refuse in this section?" These scales were administered by a Sanitation Workers' Association representative at the end of baseline (Day 5) and again at the end of the study (Day 15). Two crewmen contributed the ratings during baseline, and three crewmen contributed the ratings after the enforcement program. Because of staff turnover, only one crewman contributed ratings on both pre- and postscales. In addition, a consumer satisfaction rating was obtained from the crewmen after the completion of the study: "Overall, how satisfied are you with the packaging enforcement program?" There was perfect agreement between independent observers on the scoring of the rating-scale data.

RESULTS

Figure 1 shows the daily number of violations in the west and east sections. Before enforcement, there were means of 59.4 and 79.9 violations per day in the west and east respectively. After enforcement, the number of violations was markedly reduced to means of 9.7

Fig. 1. Daily number of packaging violations in the west and east neighborhoods.

Fig. 2. Daily percentages of residences violating the packaging regulations in the west and east neighborhoods, based on those residences with refuse out.

and 23.2 per day in the west and east respectively. These lowered numbers of violations were maintained at a followup observation four weeks after the study was completed.

Considering both sections together, the number of each category of violation showed the following reductions after the enforcement program was implemented: 1.8-m violations from a mean of 24.4 per day during baseline to a mean of 5.0 per day during enforcement; over-filled cans, 25.6 to 5.4; plastic bags, 14.2 to 5.8; inappropriate containers, 61.8 to 11.0; untied trimmings, 1.8 to 1.0; litter, 18.4 to 7.6.

Figure 2 shows the daily percentages of residences violating at least once. During the study, a mean of 45 of the 183 residences did not have their refuse out when violations were counted. In the west, a mean of 39.9% of residences with refuse out committed at least one violation per day before enforcement, whereas a mean

of 12.5% of residences violated per day during the enforcement program. Similarly, in the east, the mean percentage of residences violating decreased from 45.4% to 22.6%. Lowered rates were maintained at the followup observation.

Of the 183 residences studied, 27 (14%) never committed a violation during the 15 collection days (eight weeks) of the study. Of the 156 residences that committed at least one violation, 88% averaged fewer violations per day during the enforcement period than they had averaged per day during baseline; 6% showed no change; and 6% showed increased violations. Seventy-four of the 183 residences met the noncollection contingency: 45 residences experienced the contingency on one day, 17 on two days, seven on three days, four on four days, and one residence had refuse not collected on six different days. Of course, many residents observed the application of the enforcement

contingency to their neighbors. It should also be noted that with some residences, the *same* refuse was not collected on subsequent days. For example, for the residence where trash was left six times, the same violations were sometimes left for consecutive days before they were correctly repackaged and then collected. On six of the 10 days during which the enforcement program was in operation, the senior author drove immediately behind the sanitation truck to monitor the enforcement procedures. Approximately 90% of all violations were tagged and left during these days. When the tagging was assessed without the knowledge of the crew (*i.e.*, after the truck had passed), approximately 80% of the violations were tagged and left. Those occasions when the crew failed to tag and leave refuse were ignored both during the obtrusive and unobtrusive monitoring.

Resident complaints were also assessed. After the study was completed, the City's complaint records were examined and all complaints concerning missed or partial pickups were counted. Any complaints concerning the enforcement program would have been so categorized, along with other complaints about misses. The number of complaints was totalled for the month during which baseline observation occurred and the two months during which the enforcement program operated. A chi-square analysis showed that complaints in the experimental areas increased in the same proportion as in the other routes serviced by the same crew ($\chi^2(1) = 0.020, p > 0.05$) and in the same proportion as in the rest of the city ($\chi^2(1) = 3.478, p > 0.05$).

DISCUSSION

The enforcement program, which consisted of instruction, feedback, and contingent refuse collection, was effective in markedly reducing the number of violations of packaging regulations and the percentage of residences violating. Correlated with these objective data changes, the ratings by the sanitation crew established the social validity of the effects of the enforcement program. The differences between sections in baseline and enforcement levels of violations were mirrored by the men's ratings: refuse was rated as better packaged, and safer and easier to collect. The workers' ratings suggest that these procedures might be a valuable component in a comprehensive safety and efficiency program for sanitation workers.

Fewer than 10% of the test neighborhoods' residents complained to the sanitation workers or city officials about the requirements of the enforcement program when part of their refuse was not collected. This type of program will probably always be met with some dissatisfaction, especially when first implemented. In this context, it should be noted that the number of complaints in the study areas did not differ significantly from the rest of the city.

Fig. 3. Means of the sanitation crewmen's ratings of experimental neighborhoods with regard to adequacy of packaging, and difficulty and safety of collection, before and after the enforcement program.

Figure 3 summarizes the social validation data. Compared with the baseline evaluations, the sanitation crew rated the refuse in each section to be better packaged, easier to collect, and safer to collect after the enforcement program was implemented. The range of the ratings on any question never exceeded one point. The consumer satisfaction data showed that the men were very satisfied (mean of 7 on a scale of 7) with the enforcement program.

The conduct of policy research is not a simple activity (Riecken, 1975). As with many other research endeavors, the researchers cannot enter the experimental setting, collect data, and depart without regard to the political realities of that setting. Policy research necessarily involves close cooperation with both the policy makers and those persons involved in the procedural testing of proposed changes. The sanitation worker's association initiated the present research through their identification of the packaging problem and request for assistance from the authors. The City of Lawrence (through the office of the City Manager) and the workers' association provided the support and political insulation essential for the conduct of the research (Weiss, 1972*b*). The authors formulated the details of the municipal enforcement program in consultation with local residents, the sanitation workers, and City management personnel. They also developed the experimental procedure, design, and measurement system. Through the City Manager, a request to conduct a pilot testing of the possible policy change was submitted to the City commission. Such approval was given and the research was then conducted.

Following the study, the research findings were discussed with sanitation workers and City management personnel with the aim of formulating a proposal for city-wide implementation of the test procedures. The research data and an implementation plan similar to the procedures of this study were then submitted to the policy-making body, the elected City commission. After considering the authors' report, the commission adopted the recommendations: a new refuse-packaging ordinance was passed. This new ordinance incorporated the old specifications for refuse packaging, but also included the new enforcement provisions tested in the present study. The City management personnel were then instructed to implement the new policy.

In summary, the municipal enforcement program decreased the number of refuse-packaging violations in the test neighborhoods. The data formed the nucleus of a recommendation of

change in one area of public policy. Thus, the present study described an inexpensive and effective application of contingencies to ameliorate a municipal problem, and provided an example of the contribution of applied behavior analysis to the data-based determination of public policy.

REFERENCES

- Baldus, D. C. Welfare as a loan: an empirical study of the recovery of public payments in the United States. *Stanford Law Review*, 1973, **25**, 123-250.
- Campbell, D. T. Reforms as experiments. *American Psychologist*, 1969, **24**, 409-429.
- 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.
- Environmental Protection Agency. *Residential collection systems. Volume 1: Report Summary*. Publication No. SW-97c.1. Washington, D.C.: U.S. Government Printing Office, 1974.
- Environmental Protection Agency. *Decision-makers guide in solid waste management*. Publication No. SW-500. Washington, D.C.: U.S. Government Printing Office, 1976.
- Fawcett, S. B. and Miller, L. K. Training public-speaking behavior: an experimental analysis and social validation. *Journal of Applied Behavior Analysis*, 1975, **8**, 125-135.
- Frank, J. E. and Smith, R. A. Social scientists in the policy process. *The Journal of Applied Behavioral Science*, 1976, **12**, 104-117.
- Kershaw, D. N. A negative income tax experiment. *Scientific American*, 1972, **227**, 19-25.
- Likert, R. A technique for the measurement of attitudes. *Archives of Psychology*, 1932, No. 190.
- Minkin, N., Braukmann, C. J., Minkin, B. L., Timbers, G. D., Timbers, B. J., Fixsen, D. L., Phillips, E. L., and Wolf, M. M. The social validation and training of conversational skills. *Journal of Applied Behavior Analysis*, 1976, **9**, 127-139.
- National Commission on Productivity. *Opportunities for improving productivity in solid waste collection*. Publication No. CP73012. Washington, D.C.: U.S. Government Printing Office, 1973.
- National Commission on Productivity. *Employee incentives to improve state and local government productivity*. Publication No. CP75015. Washington, D.C.: U.S. Government Printing Office, 1975.
- Riecken, H. W. Social experimentation. *Society*, 1975, **12**, 34-41.
- 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.

- Salasin, S. Experimentation revisited: A conversation with Donald T. Campbell. *Evaluation*, 1973, 1, 7-13.
- Schnelle, J. F. and Lee, F. J. A quasi-experimental retrospective evaluation of a prison policy change. *Journal of Applied Behavior Analysis*, 1974, 7, 483-496.
- Schnelle, J. F., Kirchner, R. E., McNees, M. P., and Lawler, J. M. Social evaluation research: the evaluation of two police patrolling strategies. *Journal of Applied Behavior Analysis*, 1975, 8, 353-365.
- The Urban Institute and International City Management Association. *Measuring the effectiveness of basic municipal services: initial report*. International City Management Association, February, 1974.
- Weiss, C. H. *Evaluation research: methods for assessment program effectiveness*. Englewood Cliffs, New Jersey: Prentice-Hall, 1972. (a)
- Weiss, C. H. Where politics and evaluation research meet. *Evaluation*, 1972, 37-45. (b)

Received 12 April 1976.

(Final acceptance 18 November 1976.)