Plaza Sésamo in Mexico: An Evaluation

by Rogelio Diaz-Guerrero, Isabel Reyes-Lagunes, Donald B. Witzke, and Wayne H. Holtzman

New Mexican field study of the impact of* Plaza Sésamo *on young children brings fresh evidence of gains in learning skills but also some surprises.

Since its creation in 1968, *Sesame Street* has been translated into many languages, but often with little alteration of the visual images in the film itself. Then in Mexico in 1971, a completely new production of *Sesame Street*, particularly adapted to Latin American culture, was developed. It was called *Plaza Sésamo*. Educators, psychologists, psychiatrists, and other specialists cooperated together in planning and conducting evaluative studies of a formative nature to assist the producers of *Plaza Sésamo* in developing the program. This series of formative investigations (4) laid the groundwork for embarking upon more extensive experiments to determine the impact of *Plaza Sésamo* upon learning in young children.

The first of these experiments was with preschool children in daycare centers in Mexico City (3). A total of 221 three-, four-, and five-year-old children from three different lower-class daycare centers were equally divided by age and sex and were randomly assigned to experimental and control groups. Children in the experimental groups watched *Plaza Sésamo* programs for 50-minute periods, five days a week, until the entire series of 130 programs had been broadcast—a total of six months of continuous viewing. At the same time, children in the control group were viewing cartoons and other non-educational TV programs on a different broadcast channel in a separate room. The television sets

Rogelio Diaz-Guerrero is the Director and Isabel Reyes-Lagunes and Donald B. Witzke are the Instituto Nacional de Ciencias del Comportamiento y de la Actitude Publica, A.C., Mexico City. Wayne H. Holtzman is Director of the Hogg Foundation, University of Texas. The research was supported in part by Ford Foundation Grant 730-0206. A more detailed report can be obtained from Dr. Diaz-Guerrero, Georgia 123, Mexico 18, D.F.

were placed on a pedestal so that every child could see and hear with no difficulty.

Since *Plaza Sésamo* was a relatively new program, none of the control children had ever seen it before. Strenuous efforts were made to prevent the control children from viewing *Plaza Sésamo* on another channel when it was broadcast each evening from 6:00 to 7:00 p.m. Further investigations toward the end of the experiment revealed that only a handful of control children viewed any *Plaza Sésamo* broadcasts when they were absent due to sickness. In no case did it appear that the experimental design had been compromised in any way.

The impact of *Plaza Sésamo* upon the children who viewed it was evaluated by a series of individually administered tests given in both the experimental and control groups at three points in time: (1) *pretest*—immediately prior to the exposure to *Plaza Sésamo* or the control films; (2) *during treatment*—seven weeks after beginning the experiment; and (3) *posttest*—at the end of the experiment. An interval of seven weeks between pretest and during-treatment testing sessions was chosen because the earlier evaluations of *Sesame Street* used this size interval (1, 2).

Daily ratings were also made by trained observers to measure the degree of attention exhibited by the children in the experimental group. These ratings were only made in two of the three daycare centers and were based on ratings of six children randomly selected each day. Over the period of 130 programs, every child was rated a number of times in these two daycare centers. The number of times each child was absent was also recorded for all three daycare centers.

Dropouts were relatively few. Of 221 children in the initial sample, 173 completed the experiment. No discernible bias due to dropouts could be discovered.

Nine individual tests were employed to measure the amount of learning for each child over the six-months period. Three of these tests—General Knowledge, Numbers, and Letters and Words—are criterion measures of skills specifically taught in the *Plaza Sésamo* programs. Five other tests—Relations, Parts of the Whole, Ability to Sort, Classification Skills, and Embedded Figures—are indirectly related to *Plaza Sésamo* but are not specifically criterion measures. The ninth test, Oral Comprehension, has no relation to the stated goals of *Plaza Sésamo* although it measures an important cognitive ability related to school readiness among preschool children.

Analyses of variance and covariance were carried out to see whether or not the *Plaza Sésamo* viewers did significantly better than the children who watched only cartoons. In general, children in the experimental group showed greater gains in test performance over the six-month period than did those in the control group (see Table 1).

The main results can be summarized as follows:

1. Regardless of age-group, the children who watched *Plaza Sésamo* for six months did significantly better on at least four of the nine criterion tests than did the control children who only watched cartoons during this period.

2. The greatest increases for *Plaza Sésamo* viewers occurred on the three tests most closely related to the stated goals of *Plaza Sésamo*—General Knowledge, Numbers, and Letters and Words.

3. Oral Comprehension, the test unrelated to *Plaza Sésamo*, also revealed significantly greater gains for the *Plaza Sésamo* viewers than for the control children in all three age-groups.

4. The four- and five-year-olds showed the greatest gains from watching *Plaza Sésamo*, while the three-year-old experimental children failed to differ significantly from the control children on five of the nine tests. Test-retest correlations across the six months showed satisfactory stability (.41–.59) for four of the tests—General Knowledge, Numbers, Embedded Figures, and Oral Comprehension—but showed instability for the remaining five tests (correlations of .08 to .25). In spite of weaknesses in these five tests, positive results favoring the *Plaza Sésamo* viewers were obtained for them just as they were obtained for the more reliable tests.

5. Although the most rapid gains of the experimental children over the control groups occurred in the first seven weeks of viewing *Plaza Sésamo*, the gap between the experimental and control groups continued to grow throughout the six months.

6. Within the combined experimental group, the degree of attention to *Plaza Sésamo* correlated positively (as high as .49) in six of the nine posttest measures, indicating that children who attend regularly to the *Plaza Sésamo* program gain more than children whose attention wanders.

7. Experimental children with a large number of absences did less well on the post-treatment test battery than did children who attended regularly.

The outcome of this first experiment indicates that significant gains are made in a number of cognitive and perceptual areas by preschool children who watch *Plaza Sésamo* for the entire six months of the program.

Table 1: Experimental-minus-control differences in adjusted posttest mean scores showing favorable impact of Plaza Sésamo

	3-year-olds	4-year-olds	5-year-olds
General Knowledge	2.7	7.3***	4.8***
Numbers	4.4**	7.8***	6.2***
Letters and Words	2.9**	4.5***	5.1***
Relations	.5	1.6**	.7
Parts of the Whole	.5	.6	1.6*
Ability to Sort	.5	3.0***	3.2***
Classification Skills	2.0**	3.2***	2.3*
Embedded Figures	3.7	2.0*	.5
Oral Comprehension	3.1*	5.3**	2.4*

[a] The adjusted posttest scores were regressed on the initial scores in order to adjust for the initial differences and provide an estimate of the net gain over the period of the study.

* $p = .05$
** $p = .01$
*** $p \leq .001$

This first experiment was sufficiently promising to justify a large-scale summative evaluation under field conditions to see how well the positive findings concerning* Plaza Sésamo *hold up when the viewing conditions are not so rigorously controlled.

The basis for designing this field study was provided by a demographic study carried out by the INCCAPAC research staff. An index of socioeconomic status based upon both the education and occupation of the father was used to select children in two social classes: (1) lower-class, unskilled, or semiskilled workers with incomplete primary education; and (2) blue-collar, skilled workers, and office employees with six to twelve years of education. A sample of rural children drawn from economically deprived families in three small villages in the central valley of Mexico constituted a third major category.

From 50 daycare centers surveyed within the Health Department of Mexico City, 12 were selected as having large numbers of preschool children, good representation of lower class and blue-collar families, and adequate facilities for viewing. From the 15 rural communities visited, three villages were selected. Complete surveys were made of all families in these villages in order to obtain the necessary information for the research and the consent of parents for their children to participate in the experiment. In a similar manner, home visits and interviews were conducted with all parents in the urban daycare centers.

After 500 interviews with parents of daycare children, a startling fact was revealed: only two percent of the urban children had never seen *Plaza Sésamo*. While the investigators were aware that it might be difficult to find good control cases who had never seen *Plaza Sésamo* since the program had been broadcast for two years prior to the beginning of this second experiment, the extent to which *Plaza Sésamo* had reached into every household in Mexico City had not been realized prior to this survey. No alternative was available other than to design an experiment for these urban children which would use as control cases children contaminated by uncertain amounts of prior viewing of *Plaza Sésamo*. Fortunately, *Plaza Sésamo I*, the initial version of the program, had been discontinued in January of 1974, providing a six-month period just prior to the experiment in which no *Plaza Sésamo* programs would appear. In any event, random assignment of children to experimental and control groups within each daycare center would equalize prior exposure.

During May and June of 1974, pretest data were obtained on 1,113 cases of four- and five-year-old children in the rural and urban settings. The basic research design can be divided into two major phases according to the kind of treatment administered. In Phase I, one-half of the children viewed *Plaza Sésamo* while the other half watched cartoons. In Phase II immediately thereafter, about one-half of the original experimental group watched a new version of *Plaza Sésamo* (*Plaza Sésamo II*) while the other part of the experimental group looked at cartoons. The original control group was divided in a similar manner so that one-half of them were viewing *Plaza Sésamo II* and the other half continued to watch cartoons for the second six months of the study.

Sixteen tests were employed for evaluating the individual children at three

points in the experiment:(1) *pretest*—the beginning of the study; (2) *posttest I*—between Phase I and Phase II; and (3) *posttest II*—at the end of the experiment immediately following Phase II. Of the nine tests employed in the first experiment, only Parts of the Whole was dropped from the test battery. Letters and Words was divided into two tests with a number of new items to improve the precision of measurement. From Robert Klein's work with Guatemalan children, four new test scores were added—Elimination, Object Naming, Object Recognition, and Understanding Sentences. Because of the widespread prior exposure to *Plaza Sésamo* for all of the children in the experiment, a new test for measuring the degree of prior knowledge of *Plaza Sésamo* was developed by selecting the 15 most popular characters from the program and developing new test materials that yield a score for recognition and a score for naming of the characters. And finally, a simple motor test of active-passive coping style was experimentally developed and added to the battery.

Unlike the first experiment, three problems were encountered which could only be partially overcome, resulting in some compromise of the original design.

1. One-third of the children dropped out of the study in Phase I, the first six months. Nearly one-half of this attrition was due to the fact that on September first, when the school year began, many parents enrolled their children prematurely in the first grade, an illegal practice that is nevertheless widespread. This extensive attrition occurred equally in both experimental and control groups. During Phase II, the second six months of the study, only eight percent more attrition occurred.

2. On June 2, one month before the experiment began, *Plaza Sésamo I* resumed broadcast without warning in spite of earlier assurances that the program would not start until July. Therefore some children in both experimental and control groups saw *Plaza Sésamo I* at home for several weeks before the experiment began.

3. Minor contamination of control cases may also have occurred by an unexpected recycling of *Plaza Sésamo I* during the Christmas holidays and by commercial exploitation of *Plaza Sésamo* characters in the marketplace. Of course, experimental and control children were equally exposed to these influences.

For the above reasons, special analyses were carried out using control cases that had been purified by elimination of all children who could recognize any of the *Plaza Sésamo* characters other than the two most popular ones in the *Plaza Sésamo* Exposure Test devised specifically for this purpose. Analyses were also carried out on the entire control sample separately for Phase I and Phase II of the experiment. As in the first experiment, analyses of covariance were done separately for the four-year-olds and five-year-olds. Comparisons of experimental and control groups were also made separately for the lower-class and blue collar urban children and for the rural children. Highlights of the findings follow:

Rural children. The most striking outcome of the study is the ineffectiveness of *Plaza Sésamo* to improve significantly the cognitive development of rural

preschool children. Neither four- nor five-year-olds, with or without purified control groups, showed any significant differences between viewing *Plaza Sésamo* or cartoon in either Phase I or Phase II of the experiment. These disappointing findings cannot be attributed to contamination of control cases although inadvertant viewing by controls may have played a minor role in the rate of learning in some cases. Rural children showed significantly poorer attendance than urban children, a factor that may be of some importance in the negative results obtained. Further studies are underway to determine the influence of this factor. In fairness to the producers of *Plaza Sésamo*, it should be emphasized that rural children never were the target population of *Plaza Sésamo* programming.

Urban lower-class four-year-olds. Generally negative results were also obtained in the Phase I comparisons of experimental and control groups among the lower-class four-year-old children in Mexico City. At the end of six months, the gains made in test scores by *Plaza Sésamo* viewers were identical to children who watched only cartoons. After twelve months, however, the lower-class four-year-olds who watched *Plaza Sésamo* for an entire year did slightly better on the second posttest measures than did the children who watched only cartoons. Only General Knowledge yielded a posttest difference that was significant beyond the .01 level, although Embedded Figures, Numbers, and Classification Skills were significant at the .05 level. The difference in posttest means for General Knowledge was 4.7 in favor of the *Plaza Sésamo* viewers. None of the other criterion measures showed any difference between the viewers and non-viewers.

Posttest scores were also compared for children who watched *Plaza Sésamo* during Phase I and then viewed cartoons during Phase II (the EC group), as well as the children who watched cartoons for the first six months followed by exposure to *Plaza Sésamo* in the last six months (the CE group). When compared with the children who remained in the control group for the entire six months (the CC group), the CE group did significantly better on the final posttests for General Knowledge and Numbers but not for any of the other cognitive tests. No significant differences were found in comparing the children who saw *Plaza Sésamo* for the entire 12 months (the EE group) with the EC group, with the exception of the *Plaza Sésamo* Exposure Test which only indicates that the viewers were more familiar with the *Plaza Sésamo* characters than the non-viewers. It is apparent that, as a group, the lower-class four-year-olds yielded results only slightly better than the negative ones obtained for the rural children.

Urban blue-collar four-year-olds. At the end of six months, the experimental children did significantly better than controls on General Knowledge, Numbers, Words, and Ability to Sort. After 12 months of viewing *Plaza Sésamo* or cartoons, the particular criterion tests proving significant were slightly different although the overall effect was essentially unchanged. The magnitude of the difference in gain between experimental and control cases remained very small—only 2.7 for General Knowledge at the end of the experiment. Continued exposure to *Plaza Sésamo* in Phase II did very little, if anything, to improve

the cognitive achievement of these children beyond what had already been achieved in Phase I.

Urban lower-class five-year-olds. Only General Knowledge, Numbers, Letters, and Words showed a slightly significant gain ($p = .05$) for the experimental groups when compared with the controls over the first six months of the study. At the end of 12 months, even these tenuous gains disappeared. Children in the control groups, even when purified for possible contamination, did just about as well on the posttest measures at the end of the experiment as did the experimental groups who watched *Plaza Sesamo* for 12 months.

Urban blue-collar five-year-olds. Letters, Words, Object Naming, and Object Recognition were the only criterion measures that showed any differences between experimental and control groups at the end of Phase I. With the exception of Words, even these differences disappeared at the end of Phase II. As in the case of the other urban children, the gains for five-year-old blue-collar viewers are disappointing when compared with the gains of non-viewers.

While no consistent treatment effects were revealed by any of the criterion tests, individual tests such as General Knowledge or Numbers did occasionally show greater gains for *Plaza Sésamo* viewers than for non-viewers. Since many of the criterion tests were positively intercorrelated, a new measure, the Content Learning score, was developed by combining individual scores on General Knowledge, Numbers, Letters, Words, Ability to Sort, and Classification Skills. The reliability and comprehensiveness of this Content Learning score were considerably higher than the individual tests. The estimated reliability for four-year-olds was .83, while that for five-year-olds was .91, sufficiently high to justify increased confidence in the use of the Content Learning score for measuring the outcomes of the different treatment procedures. Even with this improved criterion measure, in only one of the six groups was there a statistically significant difference between the children who viewed *Plaza Sésamo* for 12 months and the children who only viewed cartoons throughout the experiment.

The urban lower-class four-year-olds showed slightly higher Learning scores after viewing *Plaza Sésamo* than did their control counterparts. Only one other positive outcome worthy of note was found in the analyses of the remaining two mixed treatment groups for the urban lower-class four-year-olds. The EC children and the CE children did significantly better than the CC children who watched only cartoons for the entire 12 months. For the urban lower-class four-year-olds, exposure to *Plaza Sésamo* for six months in either Phase I or Phase II of the experiment produced significantly greater improvement in the Learning score than exposure to neutral cartoons for the entire 12 months.

In summary, generally negative results concerning the value of exposure to *Plaza Sésamo* were obtained for all children except the urban lower-class four-year-olds. Among the lower-class four-year-olds, the 12 months gain in the Learning score for the *Plaza Sésamo* viewers is only slightly greater than the gains achieved by children who merely watched cartoons for the same length of time.

Results of the first rigorously controlled experiment are unequivocally clear. Children who watched *Plaza Sésamo* daily for six months under controlled

viewing conditions in this experiment showed significantly greater gain in cognitive development than did comparable children who watched only cartoons for the same period. By contrast, the larger field experiment under less well-controlled conditions but with a larger number of children, some of whom watched *Plaza Sésamo* for 12 months rather than only six, yielded generally negative results.

No single criterion test or combination of measures produced a consistently significantly greater gain for* Plaza Sésamo *viewers than non-viewers.

Why were the outcomes of these two experiments so markedly different? Of particular concern is the possibility that the experimental design in the field study was seriously compromised and weakened by factors beyond the control of the investigators. It is instructive to examine more critically the questions that arise concerning this large-scale summative evaluation of *Plaza Sésamo*.

Did uncontrolled viewing of Plaza Sésamo *ruin the experiment?* At the beginning of the experiment, nearly all of the urban children had been exposed to *Plaza Sésamo*, usually in casual or minor ways. But only a small fraction of the rural children had been previously exposed to the program. Whatever uncontrolled exposure occurred was distributed equally among both experimental and control groups. Random assignment of children to treatment groups preserved the basic integrity of the research design. In no analysis was there ever any significant difference found across treatment groups in the pretest performance on the criterion measures. If one assumes that a significant amount of measurable learning takes place with even casual, occasional exposure to *Plaza Sésamo* (an unlikely assumption in view of other available evidence in this experiment) and that further learning by systematic daily viewing beyond such occasional exposure is relatively minor (an even more implausible assumption, granted the validity of the first one), then—and only then—is the minor contamination of the urban children sufficiently serious to call into question the generally negative results obtained. Special analyses with purified control cases did not alter the essentially negative conclusions concerning the impact of *Plaza Sésamo* upon cognitive development.

Did selective attrition occur which biased the results against positive findings? The loss of one-third of the cases in the first six months of daily television or cartoon viewing was surprisingly large and serious. The greatest single cause of attrition was ambitious mothers who pushed their children illegally into the first grade. A special study of the dropouts versus those who remained indicated that on pretest measures the dropouts were significantly better performers. An important point to remember in this regard, however, is that the attrition occurred equally well among the experimental groups and the controls. The cases who remained in the study until posttest data were collected tended to be a slightly more homogeneous sample. This homogeneity may work in the direction of reducing somewhat the differential treatment effects but it can hardly be the primary cause of the negative outcome.

Did uncontrolled absences weaken the experiment? As in any field study, some children were frequently absent from the daycare and treatment groups while others attended regularly without any absence over the entire six months of the experiments. The absent children might view *Plaza Sésamo* at home, thus tending to reduce differences between the mean posttest scores for experimental and control groups. Separate studies indicated that those children who attended regularly in either the experimental or control group did better on the achievement tests than children who were frequently absent. The absent control children did not really profit from any incidental exposure to *Plaza Sésamo* while absent. Nevertheless, among rural children, absenteeism was sufficiently great to cast doubt upon the negative results. Additional analyses may reveal that rural children who were faithful in attendance did improve as a result of viewing *Plaza Sésamo*. Aside from the rural children, more likely than not, absenteeism had no major adverse effect upon the experimental outcome.

Were the criterion measures of performance as administered by the examiners appropriate for evaluating Plaza Sésamo? The 16 tests given individually to all children before, during, and after the field experiment were even more thorough than those employed in the first experiment. Many of the tests were identical while others were improved versions or new instruments that had been developed elsewhere. The examiners were well trained in testing young children. Psychometric studies indicated that the tests were indeed superior to those used previously. Side studies indicated no biases due to differences among the examiners. It should also be noted that all of the criterion tests differentiated age and social class significantly even though they failed to differentiate viewers from non-viewers in most cases.

Would other methods of statistical analysis reveal significant results where the current methods failed? Given the truly experimental nature of the research design in which children were randomly placed in either experimental or control groups at the beginning of the study, analysis of variance for posttest scores would probably have been just as adequate as the more refined covariance analysis. Since no pretest differences were found, the outcome of the analysis would be essentially the same in either case. One could also argue for a more sophisticated manner of treating gain scores. If test reliabilities were very low and if pretest differences were indeed significant, other procedures would be warranted. Under the circumstances, however, the present analyses are completely sufficient.

One must also keep in mind the fact that literally hundreds of F-ratios were computed in each of the major analyses where individual test variables rather than the composite Content Learning score were used. In such cases, occasional F-ratios appear significant when in fact such significance arises purely by chance alone. Setting a more stringent criterion would result in even less significance than was found. The basic fact remains that the absolute magnitude of the difference in gains between the experimental and control groups is very small relative to the large individual differences within groups and the massive nature of the daily exposure for 12 months to *Plaza Sésamo*.

What differences remain in the two experments that can explain the different outcomes? Granting that none of the above questions is sufficient to explain the marked differences in outcome for the two experiments, one important difference still remains that may explain the conflicting results. In the first experiment, six experimental children were selected at random every day in two of the daycare centers and were separated from the other children in a special room where they had a television set just for them. Two assistants sat at both sides of the viewers and recorded their attention for every segment of the program. This side investigation to determine the relationship of attention to achievement may well have had an important reinforcement effect upon the children in the experimental groups for the earlier study. A separate analysis of the initial scores and amount of gain for these two daycare centers as compared to the third where no study of attention was made, revealed that probably being selected for the attention study did not have a major effect upon the child's learning. But one can't be too sure. It is fairly obvious that in the earlier study, more adults attending to children were present at more times, creating a subtly different atmosphere in the first experiment than in the second.

Clearly, additional research is necessary to pin down more precisely the nature of such environmental variables influencing outcome. Undoubtedly, *Plaza Sésamo* can be made considerably more effective than it is in its standard broadcast form by incorporating some kind of adult reinforcement or guidance. Being exposed to *Plaza Sesamo* in a fairly large group of children or in a home with other distractions may not be sufficient in itself for optimum learning to take place. It may well be necessary, especially with disadvantaged children, to provide related forms of stimulation and reinforcement.

REFERENCES

1. Ball, S., and G. A. Bogatz. *The First Year of Sesame Street: An Evaluation*. Princeton, N. J.: Educational Testing Service, 1970.
2. Bogatz, G. A., and S. Ball. *The Second Year of Sesame Street: A Continuing Evaluation*, Vols. 1 and 2. Princeton, N. J.: Educational Testing Service, 1971.
3. Diaz-Guerrero, R., and W. H. Holtzman. "Learning by Televised *Plaza Sésamo* in Mexico." *Journal of Educational Psychology* 66, 1974, pp. 632–643.
4. Diaz-Guerrero, R., R. Bianchi Aguila, and R. Ahumada de Diaz. *Investigacion Formativa de Plaza Sésamo*. Mexico: Editorial Trillas, 1975.