JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1972, 5, 53-64 NUMBER 1 (SPRING 1972)

MODIFICATION OF BEHAVIOR PROBLEMS IN THE HOME WITH A PARENT AS OBSERVER AND EXPERIMENTER[1]

R. VANCE HALL, SAUL AXELROD[2], LUCILLE TYLER, ELLEN GRIEF, FOWLER C. JONES, AND ROBERTA ROBERTSON

UNIVERSITY OF KANSAS

Four parents enrolled in a Responsive Teaching class carried out experiments using procedures they had devised for alleviating their children's problem behaviors. The techniques used involved different types of reinforcement, extinction, and punishment. One parent increased the frequency of the wearing of an orthodontic device during five daily time checks by making an immediate monetary payoff contingent on wearing the device. A second parent increased the number of points earned for doing daily household tasks by providing back-ups for which the points could be exchanged. The parents of a 4-yr-old boy decreased the frequency of whines, cries, and complaints by removing social attention when such behavior occurred. A mother decreased the duration of time it took for her 5-yr-old daughter to get dressed by making permission to watch television contingent on dressing within 30 min of the time she got up in the morning. Brief reversals of contingencies were used to show causal relationships between the procedures used and the changes in behavior. Checks on the reliability of measurement were made by persons present in the home.

Most behavior modification experiments have been conducted by sophisticated researchers in institutional, classroom, or laboratory settings. Some studies, however, have employed parents as therapists for their childrens' behavior problems and were carried on in the home environment (Hawkins, Peterson, Schweid, and Bijou, 1966; Zeilberger, Sampen, and Sloane, 1968; Wahler, 1969). The subjects for these experiments exhibited sufficiently severe behavior difficulties that they were referred to a psychological clinic. The children were then evaluated and their parents were given suggestions on the type of behavior modification procedures that might alleviate their childrens' problems. The data collection tasks were conducted by the experimenters or trained observers in the children's homes.

The present studies, in contrast, were originated and conducted by individuals (the third, fourth, fifth, and sixth authors) whose main exposure to operant conditioning principles was a Responsive Teaching course in which they were enrolled.

A more detailed description of the course that follows the Responsive Teaching Model is given in Hall and Copeland (1971). Essentially, however, participants in the Responsive Teaching course were enrolled for 3 hr of credit. The class met for a 3-hr session once each week for 16 weeks (one semester). Lectures, films, quizzes, and discussion groups of about 10 persons led by a graduate student leader were used to present basic information on recording and measurement, applied behavior analysis research designs, learning theory principles, and examples of studies carried out by researchers and by previous class members. The basic course content can be found in the *Behavior Management Series* (Hall, 1971). The participants carried out their

[1]This research was partially supported by Grant HD 03144 from the National Institute of Child Health and Human Development, Bureau of Child Research and Department of Human Development, University of Kansas. Reprints may be obtained from R. Vance Hall, Juniper Gardens Children's Project, 2021 North Third Street, Kansas City, Kansas 66101.

[2]Formerly a post doctoral trainee on Grant HD 00183 from the U.S. Public Health Service, Department of Health Education and Welfare, National Institute of Child Health and Human Development, now at the University of Connecticut.

studies, aided by the leader assigned to their group, during the semester in which they took the course. The number of participants in the classes varied from 40 to more than 70.

The studies presented were carried out in the home environment, dealt with relatively mild behavioral difficulties, and required no special apparatus. In each case, the experimenters served as the primary observer.

EXPERIMENT I

Subject and Setting

Jerry first started wearing an orthodontic device when he was 8 yr old. The dental mechanism consisted of a removable head band held in place by a plastic band around his neck. Although the recommended wearing time was approximately 12 hr a day, Jerry used the device only a few hours daily. After 2 yr, the orthodontist reported little improvement in Jerry's condition. A move to a new area resulted in a second type of removable orthodontic device that also contained two bands. Again, Jerry used the device less often than recommended. After 8 yr, four dentists, and approximately $3300 in dental fees, Jerry's orthodontic condition was essentially unchanged.

Observations

Five times a day, at varying intervals, Mrs. T. observed Jerry to determine whether or not he was wearing his orthodontic device. On week days the first of these observations occurred just before breakfast. One came after breakfast but before Jerry left for school in the morning. One came after Mrs. T. arrived home from school in the afternoon but before the evening meal. One came shortly after supper and one came within 30 min of bedtime. On weekends, checks were made before breakfast, between breakfast and lunch, after lunch but before supper, shortly after supper and within 30 min of bedtime. The exact times of the checks varied somewhat and Jerry did not know exactly when the checks would be made. If both bands were in place, Mrs. T. placed a "+" on a recording sheet. If not, she reported a "0". Mr. T. conducted reliability checks each weekend. On signal from Mrs. T. both observed whether or not the device was in place. Mrs. T. recorded a "+" or "0" on the sheet. Mr. T. then told her whether he had scored the observation as "+" or "0", Mrs. T. then noted Mr. T's observation on the chart. In all cases agreement was 100%.

Experimental Phases

Baseline. Before experimental manipulations, the frequency with which Jerry wore the orthodontic device was noted for an eight-day period. During this phase Jerry was not told the purpose of the observations. As depicted in Figure 1, the mean baseline rate was 25%.

Social Reinforcement. During Baseline$_1$, Mrs. T. noticed that she was giving Jerry attention, in the form of reprimands, when the bands were not in place. During the second phase of the study, Jerry's mother did not refer to the orthodontic device when her son was not wearing it, but praised him if the bands were in place when she made the five daily time sample checks. For the nine days of contingent social reinforcement, the orthodontic device was in place 36% of the time.

Delayed Monetary Payoff. Dissatisfied with Jerry's progress during Social Reinforcement, Mrs. T. explored the effectiveness of paying her son money when he was wearing the dental device. Mrs. T. told Jerry that each time he was checked and the bands were in place he would receive 25 cents. If the bands were not in place he would lose 25 cents. The results were marked on a kitchen calendar after each observation, with the exchange of money taking place at the end of the month. For the 15 days of Delayed Monetary Payoff, the mean rate of wearing the apparatus increased to 60%.

Immediate Monetary Payoff$_1$. Although the frequency of appropriate behavior increased to more than twice the Baseline$_1$ rate during Delayed Monetary Payoff, Jerry's mother attempted to achieve further gains. During the Immediate

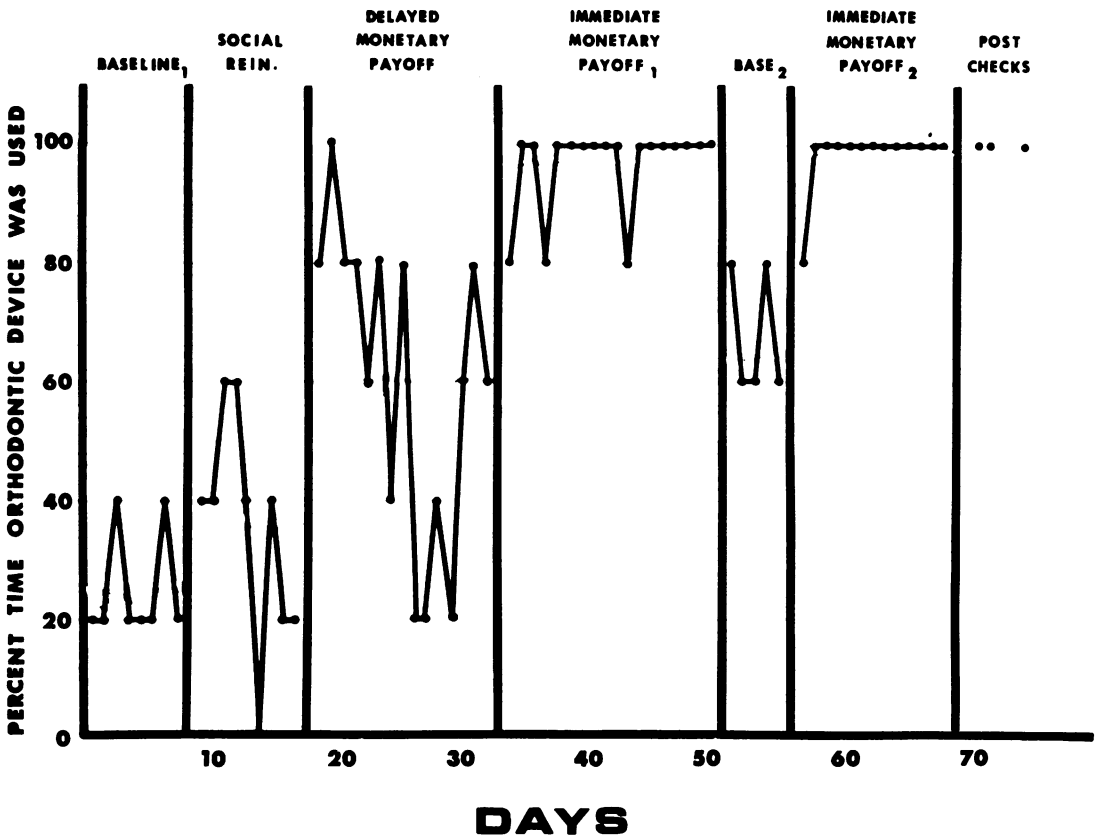

Fig. 1. A record of the percentage of time a teen-age boy used an orthodonic device. Measurements were taken five times a day at varying intervals. *Baseline*$_1$—before experimental manipulations. *Social Reinforcement*—ignoring subject when he was not wearing the device and praising him when he was. *Delayed Monetary Payoff*—paying youngster 25 cents when he had apparatus in place during an observation, and charging him 25 cents when it was not in place. Monetary exchange took place at end of month. *Immediate Monetary Payoff*$_1$—same as Delayed Monetary Payoff except that money was exchanged immediately after each check. *Baseline*$_2$—reinstatement of Baseline$_1$. *Immediate Monetary Payoff*$_2$—reinstatement of Immediate Monetary Payoff$_1$. *Post Checks*—periodic checks after termination of formal experiment.

Monetary Payoff$_1$ phase, Mrs. T. and Jerry made the 25-cent exchange immediately after each of the five daily observations. For the 18 days of this contingency the mean rate increased to 97%.

Baseline$_2$. Prior to a five-day reversal phase, Mrs. T. informed Jerry that he was making excellent progress in improving his mouth structure and that monetary exchange no longer seemed necessary. The mean rate of 64% that occurred during the Baseline$_2$ phase was greater than the rate during Baseline$_1$, Social Reinforcement, and Delayed Monetary Payoff, but below that during Immediate Monetary Payoff$_1$.

Immediate Monetary Payoff$_2$. Jerry was again paid immediately for wearing the orthodontic apparatus and lost money if the device was not in place. During the 13 days of this phase, the subject was wearing the apparatus on 98.5% of the checks.

Post Checks. On Day 68, Mrs. T. informed Jerry that money would be exchanged during each observation, but that checks would be made only occasionally. On Days 70, 71, and 74 single checks were made and in each case Jerry was wearing the orthodontic device. Observations were then made at intervals approximately two weeks apart. The bands were consistently in

place. Eight months after the study was initiated, the dentist indicated that great progress in Jerry's mouth structure had been achieved, and that it was no longer necessary to wear the apparatus.

EXPERIMENT II

Subject and Setting

For six months, Mrs. G. tried, without success, to get Antoinette, a 10-yr-old girl, to perform routine household tasks, such as cleaning her bedroom, sweeping the floor, and making her bed. Verbal reminders and unsystematic punishment produced little improvement in Antoinette's behavior.

Observations

Table 1 contains a list of the eight tasks and a brief definition of those tasks that Mrs. G. expected Antoinette to complete. Each day, Mrs. G. kept a record of the tasks Antoinette performed and the points she earned on a chart. During experimental phases the chart was posted on Antoinette's bedroom door. Checks were made each day just before the evening meal at about 6:00 p.m. Throughout the study, a neighbor made reliability checks every four to six days. The neighbor periodically made "visits" just before supper. She made an independent record of the points earned on a separate recording sheet and the records were then compared item by item. On several occasions, Mrs. G. made specific requests in the neighbor's presence for Antoinette to do odd jobs to check agreement on whether or not she complied. All 13 reliability checks resulted in 100% agreement.

Experimental Phases

Baseline$_1$. As depicted in Figure 2, baseline conditions were in effect for eight days. Mrs. G. explained to Antoinette that mother was extremely busy lately and that Antoinette, being "a big girl" was expected to share the responsibility for household tasks. Mrs. G. had devised the point system for the chores at this time but had not revealed it to Antoinette. During this period, Antoinette completed only two tasks for an average of 1.25 points per day.

Points. Mrs. G. described the point system to Antoinette. A graph was placed on the door of Antoinette's bedroom with a chart indicating the number of points that could be earned for each task. Every evening, the mother and daughter recorded the points that Antoinette earned that day. During the six days in which this contingency was in effect, Antoinette's average increased to 16.7 points per day. The range of completed tasks was from two to four a day.

Table 1

The examples of points Antoinette could earn each day for completing eight household tasks.

Task	Points
Bed made up—Covers straight, neat and smooth, sheets not visible, pillows covered, blankets folded, no other items on bed.	5
Clothes hung properly—Clothes hung straight on hangers in closet, no hangers on closet door.	5
Personal articles neatly placed—Top of dresser neat, articles arranged in symmetrical array, no powder spilled, etc.	5
Floor swept—No cement dust film on floor (Home was of cinder block construction in housing project where white powdery dust accumulated quickly).	5
Straighten and dust living room—Magazines on shelves, TV in place (not where A watched it on floor), table cleaned, books in place, no dust on furniture.	10
Kitchen duties—Wash or dry dishes, dishes in cupboard, towel on rack.	20
Bathroom duties—Towel on rack, soap in soap dish, lavatory clean and dry.	20
Odd jobs on request—Clean out car, bring in groceries, sweep off porch, etc.	5-20

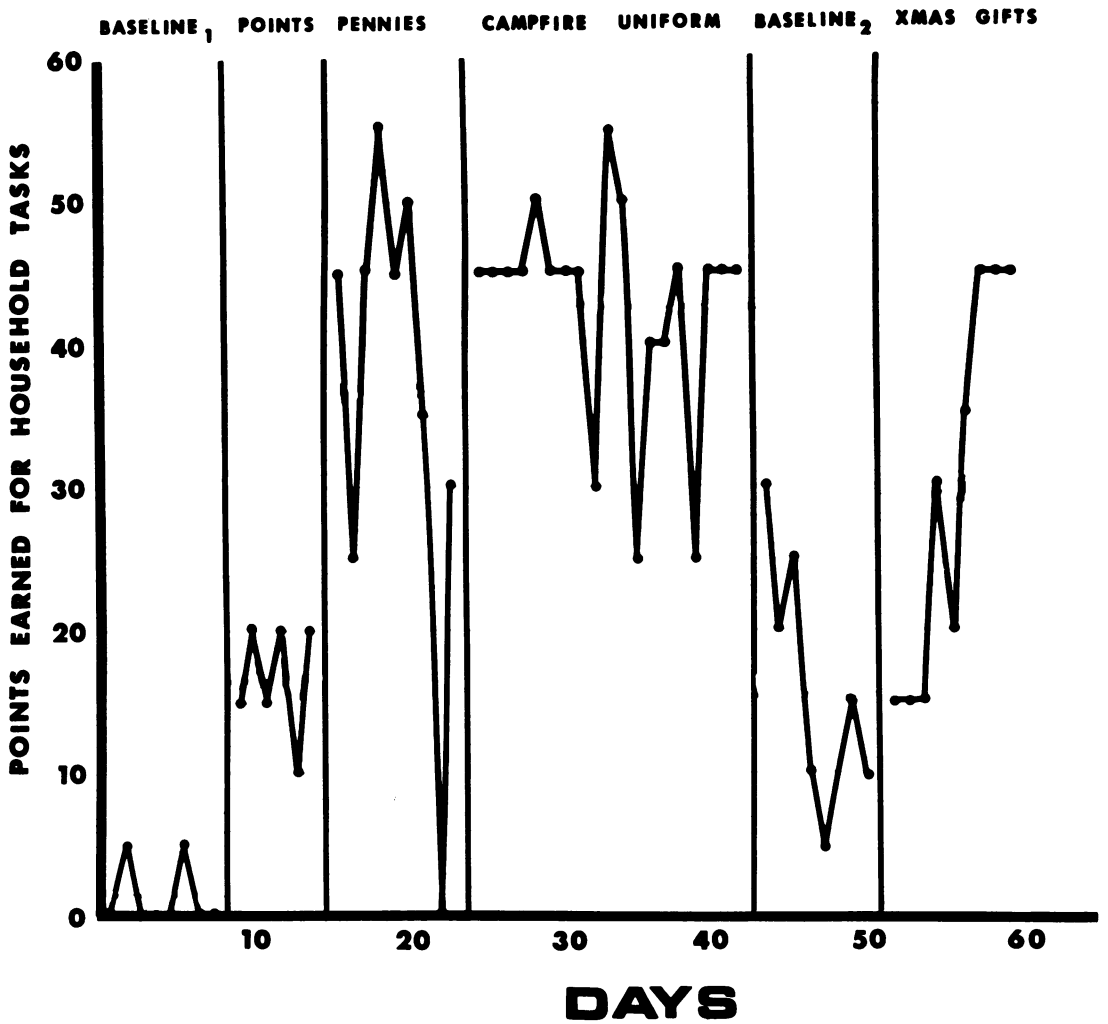

Fig. 2. The number of points a 10-yr-old girl earned each day for performing household tasks. *Baseline*$_1$—before experimental manipulations. *Points*—graphing the points the subject earned each day. *Pennies*—points were exchanged for pennies at the rate of one penny per point. *Campfire Uniform*—points were exchanged for blouse and skirt of campfire uniform at rate of 400 points per item. *Baseline*$_2$—reinstatement of Baseline$_1$. *Xmas Gifts*—points could be exchanged for money toward purchase of Christmas gifts, at rate of one penny per point.

Pennies. The Pennies stage of the experiment was the same as the Points stage, except that each point could be exchanged for one penny. During the nine days that this system was employed, the mean number of points increased to 36.7 with a range from 0 to 55 points.

Campfire Uniform. During the final day of the Pennies stage, Anoinette returned from school with a request form to join the Campfire Girls. Mrs. G. signed the form and told Antoinette that she could save the points she earned and exchange them for the blouse and skirt of the uniform that prospective Campfire girls were expected to purchase. Mrs. G. assigned a cost of 400 points to each item of the uniform. By the nineteenth day of this stage, Antoinette had earned in excess of 800 points and was able to purchase the blouse and skirt. The mean number of points earned was 42.4 per day.

Baseline$_2$. Baseline conditions were reinstated for eight days. Mrs. G. told her daughter that she was nearing the end of the graph paper and

that future reinforcement would not be available until "my next paycheck comes." The mean of 15.6 points for this period was greater than that during Baseline$_1$, but below the means during Points, Pennies, and Campfire Uniform phases and trending downward.

Xmas Gifts. Mrs. G. informed Antoinette that the points she earned could be saved toward the purchase of Christmas gifts, at the exchange rate of one penny per point. The mean number of points for the nine days of this phase was 29.4. This average was greater than that achieved during Baseline$_1$, Points, or Baseline$_2$, but below the means attained during the Pennies and Campfire Uniform stages. It should be noted, however, that the number of points Antoinette earned was increasing and they had in fact reached the level of the previous reinforcement stage in the last four days of the Christmas Gifts phase.

EXPERIMENT III

Subject and Setting

The subject was a 4-yr-old boy who, according to his parents' reports, whined and shouted at a high frequency. The experiment took place in Terry's home.

Observations

A record of Terry's verbalizations, which were of such pitch and loudness that the observer considered them to be whines or shouts, was kept by the subject's parents. Measurements were taken from approximately 9:00 a.m. to 9:00 p.m. On weekdays, Mrs. J. kept a daily record from 9:00 a.m. until 6:00 p.m., whereas Mr. J. noted the frequency from 6:00 p.m. until 9:00 p.m. Mr. J. recorded the data from 9:00 a.m. until 9:00 p.m. on weekends. Terry went to a morning nursery for 2 hr three times a week and occasionally was left with a babysitter. No records were kept during these intervals but since they occurred during all phases of the study there was no indication the data were significantly affected. Reliability checks were attained six times by having both parents record behaviors simultaneously for the entire day.

When only one parent was recording, cries, whines, and complaints were tallied on recording sheets, one of which was posted in the kitchen, one in the parent's bedroom. On weekend days when reliability checks were made, the mother tallied on the recording sheets. The father tallied with a pencil on a piece of paper he carried in his pocket. At the end of the day the tallies were totalled. The reliability index was determined by dividing the smaller observed frequency by the larger observed frequency. Agreement ranged between 75 and 100% with a mean of 85.5%. Interestingly, the experimental extinction procedures did not seem to affect reliability measures; the mean agreement during baseline sessions was 85% while that during experimental phases was 86%.

Experimental Phases

Baseline$_1$. The frequency of Terry's shouts and whines under "normal" conditions was recorded for 19 days. During this period, mother and father attended to their son's inappropriate verbalizations by either comforting him or ordering him to stop. Figure 3 indicates that the mean number of whines and shouts for the baseline period was 10.2 per day.

Extinction of Whining and Shouting$_1$. During the second phase of the study, Terry's parents ignored him when he whined or shouted. If he emitted an inappropriate verbalization, mother and father turned away from him and engaged in other activities. Whenever possible, they left the area entirely. For the 14 days of extinction, the mean number of whines and shouts decreased to 4.6 per day. The consistency of the effect was demonstrated by the fact that all 14 data points during this period were below the mean that occurred during Baseline$_1$.

Baseline$_2$. Baseline conditions were reinstated for three days. Mr. and Mrs. J. again attended to Terry when he whined or shouted. The mean for this phase increased to 8.7 verbalizations per day.

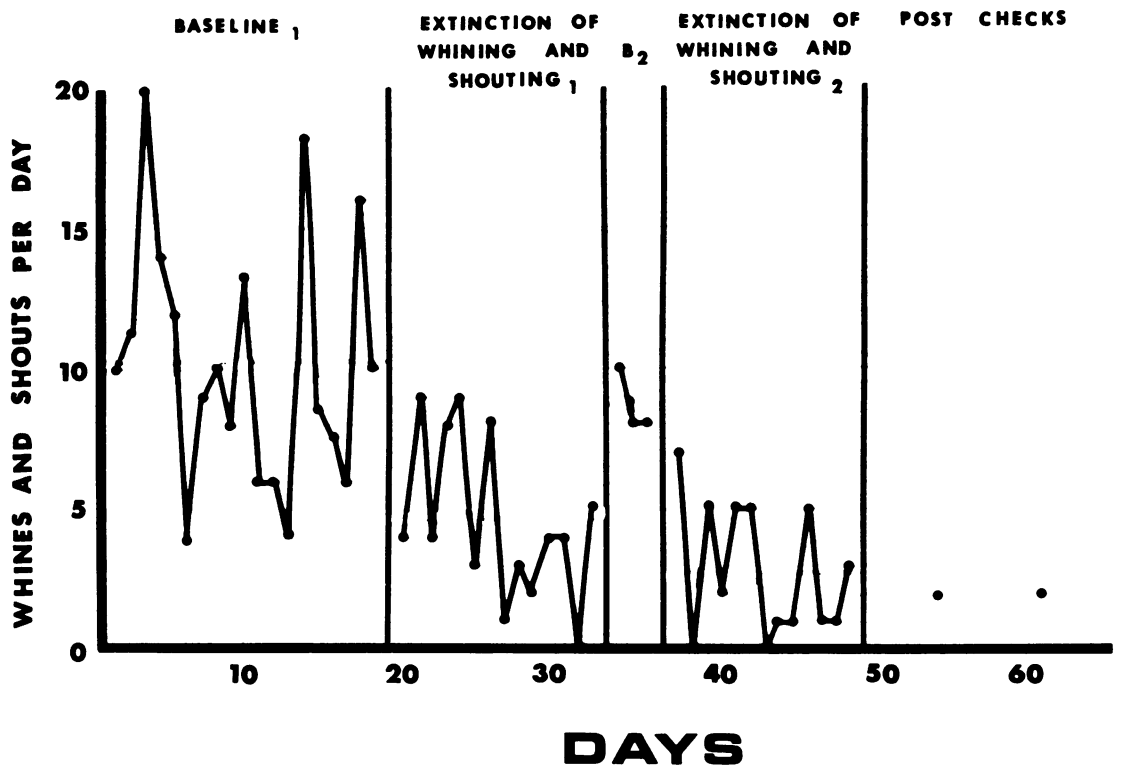

Fig. 3. The frequency of shouts and screams per day made by a 4-yr-old boy. *Baseline*$_1$—before experimental manipulations. *Extinction of Whining and Shouting*$_1$—ignoring inappropriate verbalizations. *Baseline*$_2$—reinstatement of Baseline$_1$. *Extinction of Whining and Shouting*$_2$—reinstatement of previous extinction procedure. *Post Checks*—observations of subject after formal termination of experiment.

Extinction of Whining and Shouting$_2$. For 13 days, Terry's parents again ignored his inappropriate verbal behaviors. The mean number of whines and shouts decreased to 2.8 per day. This average was below that attained during the previous extinction stage.

Post Checks. On the fifth and eleventh days of the experiment after the formal termination of Extinction of Whining and Shouting$_2$, post checks on Terry's behavior were made. During both days, a total of two whines and shouts was noted.

EXPERIMENT IV

Subject and Setting

Elaise, a 5-year-old preschool girl, had a tendency to take long periods of time to dress herself after waking each morning. Her mother's efforts in putting the clothes out the previous night and insisting that she get dressed more quickly were ineffective in changing Elaise's behavior.

Observations

Mrs. R. consulted her wrist watch and recorded on a chart the time to the nearest minute that Elaise arose and the time at which she became fully dressed. Arising meant leaving her bed by placing both feet on the floor to begin her daily activities after being called by her mother each morning. Fully dressed meant being clothed in the shoes, stockings, underclothes, dress/or play clothes the mother had laid out for Elaise the night before. Reliability checks were made on four occasions by an 8-yr-old sister and on one occasion by Elaise's aunt who was visiting in the home. Mrs. R. synchronized her watch with the kitchen clock used by the second observers. The sister gave a verbal report to Mrs. R. indicating the times when Elaise arose and

finished dressing. The aunt made an independent written record. All five checks resulted in 100% agreement.

Experimental Phases

Baseline$_1$. A record of the amount of time Elaise spent in dressing during all experimental sessions is shown in Figure 4. Under "normal" conditions, the range of times was from 1 hr 0 min to 6 hr 35 min. The mean rate for the 18 days of Baseline$_1$ was 3 hr 10 min.

Loss of TV Time$_1$. Beginning with Day 19, Elaise was required to finish dressing within 30 min after awaking. If she failed to meet the criterion, she was not permitted to watch television until 3:30 p.m. that day. During the 17 days this contingency was in effect, her average dressing time was 23 min. Only once did she miss her television privileges.

Baseline$_2$. The punishment criterion was removed for seven days. The mean duration of dressing time during this phase was 1 hr 26 min. This rate was greater than that occurring during the Loss of TV Time$_1$ stage but lower than the rate during Baseline$_1$.

Loss of TV Time$_2$. Elaise was again required to dress within 30 min after waking. Her mean dressing time for the seven days of this condition was 20 min. This average was 3 min less than the average during loss of TV Time$_1$.

DISCUSSION

One of the most frequently stated advantages of operant conditioning techniques is that the procedures are uncomplicated and can easily be applied in therapeutic situations. Nevertheless, a perusal of the pertinent literature indicates the

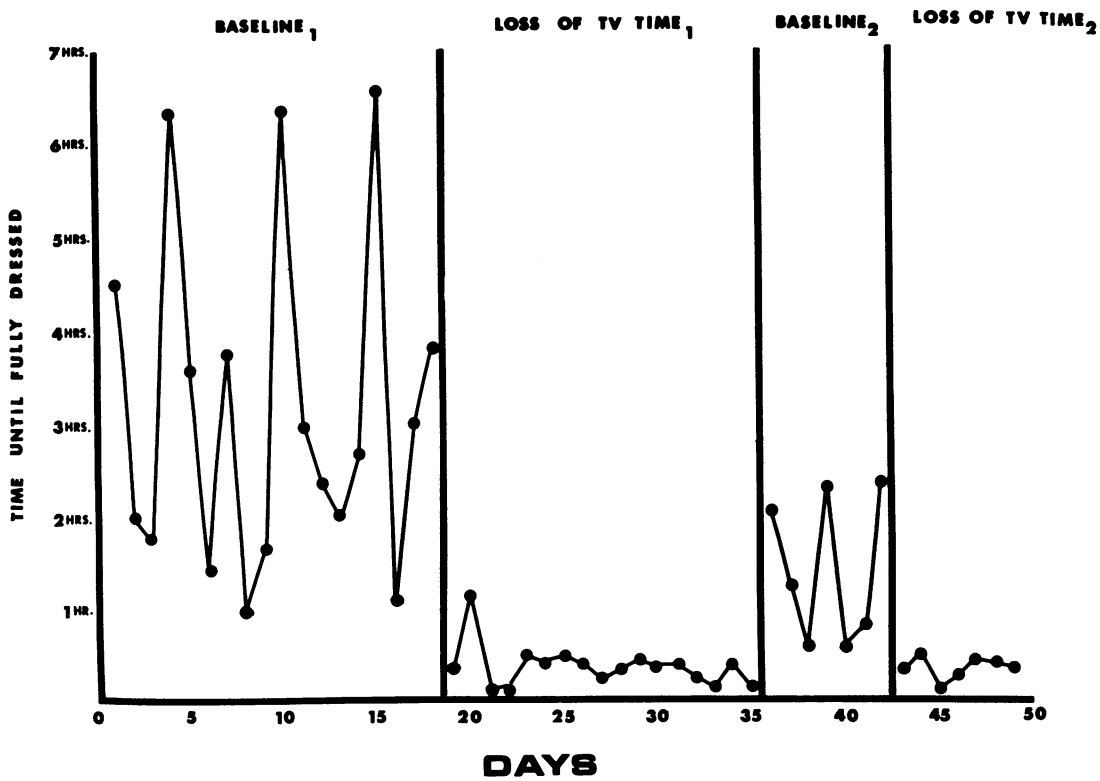

Fig. 4. Duration of time it took a 5-yr-old girl to get dressed after waking. *Baseline*$_1$—before experimental procedures. *Loss of TV Time*$_1$—subject was not permitted to watch television if she was not dressed within 30 min of arising. *Baseline*$_2$—reinstatement of Baseline$_1$. *Loss of TV Time*$_2$—reinstatement of Loss of TV Time$_1$.

presence of relatively few investigators of operant principles. One hypothesis for this finding might be that although the application of operant conditioning principles is simple, scientific investigation remains a complex task. The above studies, like those reported by Hall, Axelrod, Foundopoulos, Shellman, Campbell, and Cranston (1971) and Hall, Fox, Willard, Goldsmith, Emerson, Owen, Porcia, and Davis (1971) indicated, however, that individuals with relatively little training in operant techniques can devise and conduct behavioral experiments without unduly compromising scientific rigor. In each experiment, the investigator provided procedural manipulations (*i.e.,* a "reversal" design) that strengthened the validity of his findings. The experimenter served as the primary observer for his study, but also used a second observer to establish the reliability of his measurements.

The problem behaviors in the above experiments are typical of those found in many homes. The difficulties were not sufficiently severe that the parents sought clinical assistance, but they were a source of family friction and could have led to greater problems if a solution were not discovered. An example of this notion was given in Experiment I. The orthodontist indicated that had Jerry used the dental device when treatments were first started, his mouth structure would have been corrected in less than a year, and his parents could have saved approximately $2100 in dental fees. The cost of the experiment was less than $30.

In Experiments I and II, different positive reinforcement techniques, including token and social reinforcement, increased the frequency of appropriate behaviors. In Experiment III, an extinction procedure resulted in a decrease in whining and shouting, and in Experiment IV, contingent punishment eliminated the excessive dressing time of a 5-yr-old girl. The studies were designed in such a manner that the experimenter (*i.e.,* parent) could apply the procedures and record the data without significantly upsetting his daily routine. This factor is crucial for studies conducted in a home environment because the investigator usually cannot devote his entire attention to the problem behavior. The fact that the investigators were able to employ resources that are already found in most homes increases the general applicability of the present experiments.

A frequent criticism of behavior modification studies is that although the subject often demonstrates a dramatic change in behavior, the procedures are applied for a relatively short period of time. In Experiment I, however, the study was continued for eight months, at which time the problem situation was completely corrected. The procedures in Experiment II were still being applied at the time this article was written, approximately 2 yr and a move to Texas after the home token system was originally instituted, although the original techniques were modified somewhat. Mrs. G. reported that the tasks required for reinforcement had changed and that Antoinette now helped choose the back-up reinforcers. When Mrs. G. had suggested they discontinue the procedures Antoinette had stated: "Mommy, I can just do a better job when I can see how I'm doing." Apparently the effects of the procedure favorably impressed others in the subject's environment. Mrs. G. reported that an older daughter observed the change in Antoinette's behavior, read several operant conditioning studies, and subsequently trained the family cat to sit up, shake hands, and roll over. In addition, a neighbor of Mrs. G., who originally performed reliability checks, reported success with similar behavior modification techniques with her six children. The implications are obvious if effective systems are being developed that can be adapted and put to practical use without the requirement for formal study or training.

The procedures for recording and for obtaining checks on the reliability of observations in these studies were devised by the parents who carried them out.

In three studies, 100% agreement was obtained and in the third experiment, an acceptable

level (85.5%) was achieved. The high reliability of observation may have been influenced by several factors. One factor may have been that the behaviors chosen were for the most part rather discreet and easily discriminated. In the first experiment, the parents agreed that it was easy to see whether or not the bands were in place. They either were or they were not. Mrs. G. stated that it was easy to make a similar decision about whether or not tasks were completed in Experiment II. They either were or were not done. One thing that helped to make this so was the great amount of fine white dust present in the housing project where they lived, which aided in discriminating as to whether sweeping and dusting had occurred. During the experimental phases, Antoinette participated in marking the chart, which was a further check on the accuracy of recording. In Experiment IV, Mrs. R. reported there was little trouble in knowing when Elaise had arisen for she arose quickly after being called and had to leave her bed either to get her clothes, which had been laid out, or to come to breakfast, watch television, or engage in other high-probability behaviors. It was also easy to see when she was fully dressed. During experimental phases, Elaise promptly informed her mother when she was dressed because to do so was in her own best interest. This of course, assisted her mother in noting the time promptly during these phases. In Experiment III, it was perhaps less easy to make such a discrimination. For one thing, the behavior was auditorally perceived rather than visually, tone and inflection were involved and event recording or a frequency count of transient events was involved, rather than direct measurement of a permanent product (See Hall, 1971) as was the case in the other three studies. These differences may have accounted for the somewhat lower, yet acceptable agreement found in Experiment III.

Another factor may have been related to the fact that the independence of the observers is somewhat open to question. Azrin, Holz, Ulrich, and Goldiamond (1961) pointed out the importance of independent reliability checks. In Experiment I, Mr. T. made a verbal report of his observation decision to Mrs. T. which she recorded after having recorded her own "+" or "0". It would have been "cleaner" had he made an independent simultaneous record in which he recorded "+" or "0" on his own sheet with the records being compared at the end of the day's observation. Nevertheless, the method used was an approximation, was pragmatic, and was better than no reliability check. In Experiment II, a more rigorous and acceptably independent procedure was used. In Experiment III, a good attempt at independent recording was made, although the records could have been affected by the fact that it may have been possible for one or the other of the parents to see when cries, whines, or complaints were being recorded or, during experimental phases, when extinction procedures were being carried out. The latter case would not seem to have been true, however, since similar percentages of agreement were found during the baseline and extinction phases.

In Experiment IV, a problem similar to that noted in Experiment I was evident. In this case, the age of the sibling who acted as a second observer may have influenced the parent to alter the recording procedure the sister used. In other instances, however, young observers have proven to be reliable observers and recorders of behavior (Hall, Cristler, Cranston, and Tucker, 1970). The fact that an aunt did carry out one entirely independent check using a similar recording technique does lend credence to the observational record.

Another factor common to all the studies is that in each case the primary observer was aware that a reliability check was in progress. Romanczyk, Kent, Diament, and O'Leary (1971) provided some evidence that agreement improves when the primary observer is aware that a check is being made, even though the difference in agreement may not be significant.

Some deficiencies have been pointed out in the rigor with which the observation and reliability

procedures were carried out in these studies. It should be emphasized, however, that at the time they were conducted some 2 to 3 yr ago they compared favorably along these dimensions with many studies carried out in more controlled settings.

Subsequent studies carried out by parents who have participated in the Responsive Teaching course have become increasingly sophisticated and rigorous along these dimensions. The long lead time between writing up, submitting, reviewing, revising, and publishing, however, means that by the time more sophisticated studies are published they too will be outdated. We would like to mention, however, that currently, whenever possible, in addition to the reliability checks carried out by the *"in vivo"* second observers, at least one additional check is carried out by the group leader of the Responsive Teaching class. It has been noted in our experience that high levels of agreement between observers and similar levels of behavior to those being reported have been consistently found.

Perhaps it would be well to note another kind of reliability as to the efficacy of the studies and of the procedures used, which was first pointed out by Amber Tribble (1968). Mrs. Tribble carried out a study in which she attempted to lose weight by decreasing her calorie intake. At the end of the study she pointed out that the decrease in the number of calories would, according to health authorities, result in the approximate weight loss she actually obtained. In Experiment I, a similar kind of validation occurred when Jerry's dentist, R. D. Boice, D.D.S., stated that Jerry's mouth structure had suddenly begun to improve once experimental procedures were begun and that by his estimate the parents could have saved several years and $2100 had they employed them 8 yr earlier. By the same token, in Experiment II, the fact that Mrs. G's neighbor, Mrs. Frances McGuire, decided to use the procedures with her own children is an indication at least that they were being carried out and were effective.

It is of further interest perhaps to note that both Mrs. G. and Mr. and Mrs. T. are continuing to use token reinforcement procedures to maintain appropriate behavior in their offspring over 2 yr after completion of their studies, even though the behaviors worked with have changed. In the other two cases, the original behaviors are no longer reported to be problems and similar procedures have not been systematically used, although Terry's parents report that they have been careful to use their attention to reinforce appropriate behavior, albeit unsystematically.

REFERENCES

Azrin, N. H., Holz, W. H., Ulrich, R., and Goldiamond, I. Control of the content of conversation through reinforcement. *Journal of the Experimental Analysis of Behavior,* 1961, **4**, 25-30.

Hall, R. V. Behavior Management Series. *Part I. The Measurement of Behavior, Part II. Basic Principles, Part III. Applications in School and Home.* H and H Enterprises, 1971.

Hall, R. V., Axelrod, S., Foundopoulos, M., Shellman, J., Campbell, R. A., and Cranston, S. S. The effective use of punishment in the classroom. *Educational Technology,* 1971, Vol. XI, **4**, 24-26.

Hall, R. V. and Copeland, R. E. *The responsive teaching model: A first step in shaping school personnel as behavior modification specialists.* A paper presented at the Third Banff International Conference on Behavior Modification, April, 1971.

Hall, R. V., Cristler, C., Cranston, S. S., and Tucker, B. Teachers and parents as researchers using multiple baseline designs. *Journal of Applied Behavior Analysis,* 1970, **3**, 247-255.

Hall, R. V., Fox, R., Willard, D., Goldsmith, L., Emerson, M., Owen, M., Davis, F., and Porcia E. The teacher as observer and experimenter in the modification of disputing and talking-out behaviors. *Journal of Applied Behavior Analysis,* 1971, **4**, 141-149.

Hawkins, R. P., Peterson, R. F., Schweid, D., and Bijou, S. W. Behavior therapy in the home: Amelioration of problem parent-child relations with the parent in a therapeutic role. *Journal of Experimental Child Psychology,* **4**, 99-107.

Romanczyk, R. G., Kent, R. N., Diament, C., and O'Leary, K. D. *Methodological problems in naturalistic observation.* Paper presented at the Second Kansas Follow Through Symposium on Applied Behavior Analysis, April, 1971.

Tribble, A. *Reduction of caloric intake through a self-imposed consequence.* Unpublished study, Education 115, University of Kansas, 1968.

Wahler, R. G. Oppositional children: a quest for parental reinforcement control. *Journal of Applied Behavior Analysis,* 1969, **2,** 159-170.

Zeilberger, J., Sampen, S. E., and Sloane, H. N. Modification of a child's problem behaviors in the home with the mother as therapist. *Journal of Applied Behavior Analysis,* 1968, **1,** 47-53.

Received 1 October 1970.
(Revised 21 October 1971.)