SUPPRESSION OF SELF-STIMULATION: THREE ALTERNATIVE STRATEGIES

SANDRA L. HARRIS AND SHARLENE A. WOLCHIK

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

Four boys with autistic-like behavior were treated for self-stimulatory behavior with three different treatment procedures—time-out, differential reinforcement of other behavior (DRO), and overcorrection. All four boys showed a rapid response to the overcorrection procedure. Three boys demonstrated some evidence of decrement in responding with time-out. During the DRO procedure, one showed a modest decrease, two showed no change, but one exhibited a consistent increase in responding under this condition. A multiple baseline applied to one of the subjects failed to reveal any generalization of suppression from one setting to another. A strong but not perfect relationship was found between a frequency and a duration measure of self-stimulation. There was some evidence of negative side effects for one boy during overcorrection and for another during time-out. None of these negative side effects was enduring. There was also some indirect evidence that overcorrection facilitated appropriate play.

DESCRIPTORS: Self-stimulation, autism, overcorrection, time-out, DRO, autistic children

Self-stimulation or stereotyped behavior is repetitive bodily movement which serves no apparent purpose in the external environment. The etiology and function of self-stimulation remains debatable (Berkson, 1973). Nonetheless, it is a common occurrence among institutionalized retarded and autistic patients, with two thirds of them reported to spend some time engaged in such stereotyped motions (Berkson & Davenport, 1962).

In addition to any cosmetic objections that one might raise about the unattractive appearance of the behavior, a more serious concern is that self-stimulation may sometimes (Koegel & Covert, 1972; Koegel, Firestone, Kramme, &

Dunlap, 1974), but not always (Klier & Harris, 1977), interfere with learning or performance. Because self-stimulation may be potentially harmful to the patient as learner, a number of attempts to develop techniques for suppression of this response have been made. For example, Pendergrass (1972) used a 2-min time-out to suppress self-stimulation in two retarded children. Laws, Brown, Epstein, & Hocking (1971) reduced self-stimulatory behavior by instructing a teacher to turn away when his children self-stimulated and to reinforce appropriate behavior when they attended. Similarly, Azrin, Kaplan, & Foxx (1973) reduced self-stimulation in nine retarded adults by teaching them appropriate use of work and play materials.

In addition to such painless procedures as time-out and reinforcement of incompatible behaviors, other researchers have used physical punishment such as a slap (Koegel & Covert, 1972; Koegel et al., 1974) and shock (Baumeister & Forehand, 1972) to suppress self-stimulation. One of the most innovative approaches to treatment of this problem was Foxx and Azrin's (1973) use of overcorrection to reduce stereo-

Support for this research came in part from a grant by the Charles and Johanna Busch Memorial Fund and from the National Institute of Mental Health (MH 29897) to the first author. The authors thank Lenore Kopelovich, Maria Tagliavia, Russell Bolyard, and Andrea Smith, the teachers, for their cooperation. Jan S. Handleman provided extensive supervision of the corps of observers. Peter E. Nathan made useful editorial comments. Requests for reprints should be sent to Sandra L. Harris, Psychology Department, Douglass College, New Brunswick, New Jersey 08903.

typic mouthing and head weaving in retarded children. Other researchers have reported similar success with various forms of overcorrection (Doke & Epstein, 1975; Epstein, Doke, Sajwaj, Sorrell, & Rimmer, 1974; Wells, Forehand, Hickey, & Green, 1977).

Hobbs and Goswick (1977), in their review of research on the treatment of self-stimulation, noted that there have been few direct comparisons of the efficacy of different procedures. They also commented that most data were collected only for a short time each day or for several sessions over a few days. They identified the failure of most studies to examine generalization to other settings as an additional methodological concern. The purpose of Experiment 1 in the present study was to examine the utility of three different techniques—time-out, differential reinforcement of other behavior (DRO), and overcorrection with the same self-stimulating children. In Experiment 2, generalization from one treatment setting to another was explored. Experiment 3 examined the relationship between two measures of self-stimulation: frequency and duration. Each of these experiments lasted for several weeks and involved most of the child's entire school day.

GENERAL EXPERIMENTAL PROCEDURE

Subjects

The subjects were four boys enrolled at the Douglass Developmental Disabilities Center, Rutgers University. All had been diagnosed as autistic by outside agencies, were severely retarded in cognitive development, and had a high frequency of self-stimulatory behavior. All were without functional speech and were asocial upon admission.

James, 6 years, 8 months old, had been at the Center for 1 year prior to the study. His score on Rimland's (1971) E-2 Scale was +1 (classic autism is defined as +20 or greater). At the time of the study, James would remain seated and

cooperative for as long as 20 min without a break. When adult supervision was withdrawn, his activity might become frenetic and he would run about tossing objects and whirling around. This could be terminated with a firm command. James used some manual signs to communicate basic needs and had little speech. He could feed himself, was toilet trained, and could partially dress himself. He was being taught speech, manual signing, and pre-academic skills during the study.

James's self-stimulation was initially defined as "any repetitive movement of hands in air (hands held over head, stretched in front or to side of body, or extended toward floor and waved about), repetitive movement of hands so fingers are pushed into ear, palm, or finger. Repetitive rubbing, stroking, or tapping hair, eyes, or fingers. Clapping hands." If James switched from clapping to any other form of self-stimulation or vice versa, this change marked the onset of a new episode. Otherwise, a 3-sec interval of nonself-stimulation marked the onset of a new episode. In the second phase of James's treatment, after winter vacation, the definition of self-stimulation was expanded to include bobbing of the head in a backward and forward motion two or more times. If he were asked a question and responded by nodding his head, this was not considered self-stimulation.

Matt was 5 years, 5 months old and had been at the Center for 2 years prior to the study. His score on Rimland's E-2 Scale was -25. Matt was a cooperative child who followed a variety of simple commands and had a fair range of basic self-help skills. He often sought physical contact but was indiscriminate about the source of this affection. Left unsupervised, he would quietly self-stimulate. He was being taught speech and pre-academic skills during the study.

For Matt, self-stimulation was defined as "repetitive (two or more times) rubbing, stroking, or tapping motion of hands or fingers against other hand, fingers, or body part (e.g., mouth, knee), clothes, another person's body or clothes, or an object (e.g., desk)." An interval of 3 sec

of nonself-stimulatory behavior marked the onset of a new episode.

Alan was 6 years, 9 months old and had been at the Center for 2 years. His score on Rimland's E-2 scale was +21. Alan was a "hypoactive" child who required constant prodding to respond. He could dress, feed, and toilet himself as well as speak, but did so only on demand. At the time of the study, he was being taught speech and pre-academic skills.

Alan's self-stimulation was defined as "any repetitive (two or more times) turning of objects in hand, crumbling food between fingers, stroking objects (e.g., paper, clothing, toys) or body parts (e.g., arm, finger), and tapping fingers together." Three sec of nonself-stimulation separated each discrete episode.

Karl, 7 years, 3 months old, had been at the Center for 7 months prior to the study. His score on Rimland's E-2 Scale was -13, thus suggesting that like Matt and James he did not resemble the classically autistic child. His expressive vocabulary was nil, but he had a moderately good receptive vocabulary and was capable of discriminating the letters of the alphabet. He could sustain 15 min of good sitting and established spontaneous eye contact. He was typically compliant in work but not in less structured situations, where he engaged in a variety of self-stimulating and aversive behaviors such as "sniffing" people. He could feed himself, was urine but not bowel trained, and could partially dress himself. He was being taught verbal imitation, concepts of size and shape, and self-help skills during the course of the study.

Karl's self-stimulation was defined as "two or more movements of hand(s) making contact with an object or body part." A 3-sec or longer pause constituted the end of an episode.

For all four boys, work activities consisted of verbal tasks such as learning noun, verb, and adjective labels, answers to simple questions, and so on. There were nonverbal tasks such as pointing to letters, numbers, or objects or matching to sample. Play activities for Alan included stringing beads, rolling a ball back and forth

with his tutor, assembling five-piece wooden puzzles, and coloring with crayons and paper. Karl's play activities were assembling five-piece wooden puzzles, using blocks to build towers and bridges, coloring with crayons and paper, molding play dough into simple shapes, and pushing a wooden needle through a sewing block.

James and Matt took part in Experiment 1, Alan in Experiment 2, and Karl in Experiment 3.

Setting

The Douglass Developmental Disabilities Center is a day school for children with autistic behavior. It enrolls 25 children who range in age from 4 to 12 years. Each is provided an individualized program carried out by tutors under professional supervision. The subjects in this study were three of four boys in a single class and one boy, Karl, from among four in another room. All data collection took place during their routine activities.

The structure of both classrooms was similar. All sessions took place in a classroom approximately 12 ft by 20 ft (3.7 m by 6.1 m). The children were in different parts of the room to minimize distraction as each worked on his individual lessons. There were cabinets, toy chests, and tables around the room as well as some folding screens to provide privacy. Each child had his own desk with two small chairs, one for the child and one for his tutor. On one wall was a one-way mirror which permitted observation of virtually the entire room.

Tutors

Advanced undergraduates in psychology, speech, and special education, each of whom worked a day a week at the Center for course credit, implemented the treatment. Each child had a tutor each day. The tutors were aware that a study was in progress but were told about each phase only as it was implemented.

Teachers

From September to November, James's, Matt's and Alan's classroom was supervised by a special education teacher with a master's degree and several years of experience. In December, a newly graduated teacher with a bachelor's degree took over the class. Karl's class was supervised by two graduate students in clinical psychology, one of whom was present each day.

Observers

Reliability data for Experiments 1 and 2 were collected by six senior psychology majors who had worked at the Center at least 1 year as tutors and were receiving course credit for their research experience. Karl's corps of observers was a second group of five undergraduates and several staff members.

Procedures

During Experiment 1, the following procedures were in effect only during work periods. During Experiment 2, they were expanded to encompass work and play. During Experiment 3, they were in effect only for play periods. All observation periods lasted 15 min. Any deviations in this schedule were noted in exact minutes; sessions less than 10 min were automatically discarded as too brief to be stable. Only a few of the hundreds of sessions were actually discarded. All of Karl's sessions ran the full 15 min. Nevertheless, on one day because of illness and another because of snow he did not complete a full five sessions, and these days were discarded from the analysis. Number of sessions a day varied somewhat for the other boys because of bad weather, illness, etc. James's total number of sessions averaged 9.4 per day (range 3 to 12), Matt averaged 9.6 per day (range 4 to 12), Alan averaged 2.8 (range 2 to 3) play sessions and 6.9 (range 2 to 8) work sessions a day.

During work, boy and tutor faced one another on small chairs. Play sessions for Alan took place

on the floor, those for Karl at a small table. All sessions were held in the classroom while other children worked on their individual programs, and sessions were evenly distributed from 10:00 a.m. to 2:00 p.m. each day.

Baseline. During baseline, the tutor noted each discrete episode of self-stimulation on a data sheet. The tutor followed regular classroom procedures. On those relatively infrequent occasions when self-stimulation coincided with a correct response, it was arbitrarily decided to reinforce the correct response. Although contrary to clinical practice, this minimized the decision making by the tutor and insured consistency—in essence, the tutors were to ignore self-stimulation. The boys were never restrained or given other attention for self-stimulation, nor were they reinforced for "quiet hands" or "nice sitting."

The pace of the work sessions was readily determined by the tutors. They presented trials at frequent intervals, and the boys had several opportunities a minute to be reinforced for correct academic responding. Since play activities did not occur at a pace determined by the tutor, intervals of reinforcement and prompting were preprogrammed for Karl in Experiment 3. Every 20 sec, the tutor regarded Karl's behavior and either (a) reinforced appropriate play if it were occurring or (b) prompted with physical and verbal prompts if play were absent.

DRO. During DRO, the boy was socially reinforced after every second academic trial if not engaged in self-stimulation. The precise sequence of these events was as follows: The tutor reinforced the academic response if correct, and then recorded whether or not the response was correct. As the tutor turned back to the child, if he were not engaged in self-stimulation, the boy was praised for "quiet hands," "nice playing with blocks," "good sitting." In Experiment 1, his arms and hands were gently stroked by the tutor. In Experiments 2 and 3, touching was deleted but food reinforcement was added.

The DRO procedure for Experiment 3 deviated somewhat from this pattern in that the ac-

tivities involved play rather than work and, thus, were not marked by discrete trials. Instead of using every second academic trial as the marker for reinforcement, every 20 secs the tutor would (a) reinforce appropriate play if it were occurring or prompt it if it were not; and (b) 5 sec later if the child was not engaged in self-stimulation, he was praised and given food reinforcement with some specification of the appropriate behavior (e.g., "Nice using your hands to play").

Overcorrection. Immediately following the onset of each episode of self-stimulation, the tutor said a stern "No hand play" and briskly guided the boy's hands through an exaggerated hand clap with the arms stretched over the child's head for 10 sec. As soon as this exercise was completed, the tutor returned to the work program and presented a new trial. Praise for correct responses remained in effect. The children received no reinforcement for nonself-stimulatory behavior. In Experiment 3, Karl received a prompt to play or was reinforced for having returned to play 20 sec after the end of the overcorrection process.

Time-out. Immediately following the onset of each episode of self-stimulation, the tutor said sternly "No hand play" and turned her/his head away from the child for 10 sec. Because the tutor could not see the child during this interval, self-stimulation was not recorded by the tutor during time-out episodes. During James's, Alan's, and Karl's sessions, the observers continued to watch the child and maintained a separate record of self-stimulation during the time-out episode. These data were not kept for Matt.

After the 10-sec interval passed, the tutor said "OK," and he/she turned back toward the child to present a new trial. If the child were sitting quietly as the tutor looked at him, a new trial was presented. If he were self-stimulating, the time-out process was repeated. In the case of Karl, the boy was prompted to play at the end of the 10-sec interval if self-stimulation were absent. The 20-sec interval preceding his next prompt or reinforcement for appropriate work began at this point.

EXPERIMENT 1

Subjects

James and Matt served as subjects.

Experimental Design

A reversal design was used, with a period of baseline preceding the implementation of each treatment phase. Each posttreatment baseline continued until frequency of self-stimulation returned to pretreatment levels. Each treatment phase continued until there was clear evidence of suppression of self-stimulation or 5 days had passed, whichever was sooner. The sequence of treatment procedures was randomly assigned. For Matt, the order was baseline, time-out, baseline, overcorrection, baseline, DRO, baseline, DRO. For James, the order was baseline, DRO, baseline, overcorrection, baseline. After winter vacation, this sequence resumed with baseline, time-out, baseline, overcorrection, baseline, overcorrection. Two factors required this extension of James's participation in the study after vacation: One was a long delay in returning to pretreatment levels of responding after overcorrection was first applied; the other was the appearance of a new form of self-stimulation after vacation.

Reliability

Reliability data were evenly distributed across all conditions for both boys. For Matt, data were collected on 24 of the 49 days of the study. A total of 177 of the 458 individual sessions were observed. For James, data were collected on 48 of the 92 days of the study. A total of 275 of the 871 individual sessions were observed.

Formal data collection did not begin until every tutor and observer had reached at least 80% agreement on performance one day. During training, tutors were given feedback on their performance by observers. They also received accuracy feedback throughout the study. Tutors were aware that reliability data would be taken but were not told which sessions would be ob-

served. The use of one-way mirrors with a sound system permitted this covert collection.

During baseline, overcorrection, and DRO, the observers and tutors collected data on frequency of self-stimulation for each work interval. During time-out for James, the observers were required to collect additional information on the frequency of self-stimulation during the 10-sec time-out episodes.

Reliability was calculated by dividing the number of agreements during each 15-min interval by the number of agreements plus disagreements, then multiplying this number by 100.

Results

Reliability. Overall interrater reliability for the frequency of Matt's self-stimulation during 15-min sessions was 82.54%. Percent agreement scores for each condition were: baseline 88.56% (range 72.3% to 95.2%), DRO 87.9% (range 86.8% to 91.8%), time-out 77.8% (range 71.7% to 89.3%), and overcorrection 75.9% (range none). On only three days did reliability fall below 80%. For James, overall reliability was 84.85%. Percent agreement for each condition were: baseline 82.08% (range 54.6% to 100%) DRO 79.19% (range 77.5% to

91.8%), time out 90.82% (range 84.4% to 98.4%) and overcorrection 87.29% (range 65.3% to 100%). On 11 days the reliability fell below 80%.

Matt's performance. Nine days of pretreatment baseline yielded a mean of 9.8 self-stimulatory responses per session (Figure 1). The introduction of 4 days of time-out resulted in a marked decline to a mean of 5.1. During 7 days of the subsequent baseline, the frequency of Matt's self-stimulation rose to a mean of 10.7. The application of overcorrection then brought on an immediate, striking decline ($\bar{X} = 2.7$). With 6 days of baseline afterwards, Matt's rate of self-stimulation again reached pretreatment levels ($\bar{X} = 9.5$). Unexpectedly, the subsequent DRO led to an increase in self-stimulation ($\bar{X} = 13.5$). Then, during a 9-day baseline, there was a decline to a mean of 12 responses per interval. The final 5 days of this period approximated the pretreatment baseline ($\bar{X} = 10.8$). To reconfirm the unexpected increase resulting from DRO, this treatment was reintroduced; once again it yielded an increase in self-stimulation behavior ($\bar{X} = 13.8$).

James's performance. Pretreatment baseline for James indicated a mean of 12.4 self-stimulatory responses per session (Figure 2). This rate

Fig. 1. Matt's mean number of self-stimulatory responses per session across all treatment conditions.

Fig. 2. James's mean number of self-stimulatory responses per session across all treatment conditions in Phase 1 and Phase 2.

of responding appeared essentially undisturbed by DRO ($\bar{X} = 10.84$) and a return to baseline ($\bar{X} = 11.2$). The subsequent use of overcorrection, however, resulted in an immediate decline ($\bar{X} = 2.5$). During the following 29 days of baseline, responding rose to a mean of 7.8, well below pretreatment levels.

Because of our inability to achieve a stable reversal and because an additional form of self-stimulation had appeared over winter vacation, James's definition of self-stimulation was revised and a new baseline established. Using this new definition, the mean frequency of self-stimulation was found to be 28.46. The introduction of time-out had only slight effect on this performance ($\bar{X} = 31.06$). Return to baseline ($\bar{X} = 24.12$) was then followed by sharp drop in responding when overcorrection was applied ($\bar{X} = 9.72$). This time a reversal was obtained in 6 days ($\bar{X} = 22.5$). The final application of overcorrection yielded an average of 5.2 responses.

The data kept by the observers during time-out provided information unavailable from the tutor's records. Since the observers had recorded self-stimulation during time-out episodes as well as the rest of the interval, their data encompassed the entire session whereas the tutor's data reflected only a portion of the session.

A comparison of those data kept by tutors during time-out and the additional records kept

by the observers showed a modest difference between the two. Data that included responses during the time-out episodes (observers) yielded a mean of 38.68; those based on the exclusion of time-out episodes (tutors) produced a mean of 31.06. Since these differences were not great and the pattern of responding each day was identical for the two sets of data, it was decided to use the data excluding time-out episodes as these were based on a larger number of observations.

EXPERIMENT 2

Subject

Alan served as subject.

Experimental Design

A multiple baseline, reversal design was used. Data were collected during both work and play sessions throughout the study. Each of the three treatment procedures was introduced during a work period while Alan remained in baseline during play. Each treatment phase was preceded by a 5-day baseline period during work and continued until there was clear evidence of suppression or 5 days had passed, whichever was sooner. Time-out and DRO were randomly assigned, and overcorrection was used as the final treatment because of the difficulty experienced in obtaining a reversal with James. In the final phase, overcorrection was used in both work and

play. The sequence of conditions for Alan was baseline, DRO, baseline, time-out, baseline, overcorrection, baseline, overcorrection (work and play).

Procedure

The primary deviation from the general procedure during baseline was that data were collected during play as well as work. During play, Alan's tutor helped him with blocks, puzzles, coloring, and stringing beads. If self-stimulation continued for more than 15 sec during play, the tutor discontinued the act by gently separating Alan's hands or removing the object. His hands were put on his thighs, and he was prompted to resume playing. No comment was made about his self-stimulation. If he was engaged in self-stimulation at the onset of a work trial, the tutor gently removed the object and positioned his hands without comment.

Overcorrection and time-out were identical to Experiment 1; DRO differed in that Alan's hands were not stroked or patted. He was, however, given food reinforcement. The elimination of physical contact was prompted by the observed increase in self-stimulation by Matt under this condition. The addition of food was used to heighten the reward value of this procedure.

Reliability. Reliability data were collected every day Alan was in the study during at least one work and one play session. A total of 198 of the 271 work and 86 of the 110 play sessions were observed. Tutors were aware that reliability data were being collected but did not know which sessions were observed because of one-way mirrors.

In addition to reliability data for frequency of self-stimulation, the observers were responsible for obtaining duration data on the total length of time Alan engaged in self-stimulation. The myriad of clinical demands precluded collection of these data by the tutors. Reliability data were obtained for the duration measure on 43 of the 288 sessions. These encompassed 16 of the 39 days that duration data were collected. To measure duration, the observer started a stop-

watch when self-stimulation began and stopped it at the end of each episode. The cumulative time on the watch at the end of the session reflected total self-stimulation for that period. Reliability was calculated by dividing the number of seconds of self-stimulation recorded by Observer 1 by the total for Observer 2 and multiplying that number by 100.

Results

Reliability. Overall average interrater agreement for the frequency of Alan's self-stimulation during play sessions was 89.49%. During baseline, the average was 82.06% (range 33.3% to 96%) and during overcorrection 96.92% (range 75.3% to 100%). There were 12 days below 80%. The percentage agreement during all work conditions was 85.43%. Across conditions reliability figures were baseline 81.95% (range 70% to 91.7%), DRO 84.72% (range 81% to 88.7%), time-out 86.63% (range 79.4% to 95.3%) and overcorrection 88.41% (range 75% to 100%). There were 10 days below 80%.

Alan's performance—Work. A mean pretreatment baseline of 14.7 self-stimulatory responses per work session was established over 5 days (Figure 3). The introduction of DRO ($\bar{X} = 15.5$) and a return to baseline ($\bar{X} = 14.2$) had no impact on this responding. Time-out led to a 3-day increase in self-stimulation, followed by 2 days at near baseline level. The mean for time-out was 18.3. Following a return to baseline ($\bar{X} = 13.9$), overcorrection was introduced. It produced a decline in response rate ($\bar{X} = 4.2$). A return to baseline brought frequency back to a mean of 10.0; this rate dropped to a mean of 1.0 during the final overcorrection period.

As in Experiment 1, a comparison of time-out data with time-out episodes included ($\bar{X} = 18.68$) and excluded ($\bar{X} = 18.34$) yielded virtually identical results and suggests that Alan rarely engaged in self-stimulation during a time-out episode.

Alan's performance—Play. The extended baseline during play failed to reveal any system-

Fig. 3. Alan's mean number of self-stimulatory responses per session across all treatment conditions during work and play.

atic effect of work treatment on play sessions (Figure 3). Mean responses during play varied little as a function of treatment in effect during work. There was, however, a downward trend in responding which began during the third work baseline and continued throughout the study.

The introduction of overcorrection during play brought self-stimulation sharply down to a mean of 2.5 as compared to an average of 10.4 in the 5 days preceding this treatment.

Relationship between frequency and duration of self-stimulation. A Pearson product moment correlation between number of self-stimulatory behaviors during a session and a number of seconds of self-stimulation yielded a correlation of .50 ($p < .005$) during play and .71 ($p < .005$) during work. Thus, although not perfect, there was a strong and highly significant relationship between the two measures.

EXPERIMENT 3

Subject

Karl served as subject.

Experimental Design

A reversal design was used, with a period of baseline preceding the implementation of each treatment phase. Each posttreatment baseline continued until frequency (but not necessarily duration) of self-stimulation approximated pre-treatment levels. Each treatment phase continued until responding appeared stable, but for no less than 5 days. The sequence of DRO and time-out was randomly assigned, and overcorrection was used as the final treatment because of the difficulty in obtaining a reversal with James. The sequence of conditions for Karl was baseline, DRO, baseline, time-out, baseline, overcorrection, baseline. Overcorrection and its subse-

quent baseline were briefer than desired because of the end of the school year.

The procedures followed in Experiment 3, resembled closely those of Experiment 1. The primary deviation was that data were collected during play rather than work. An initial attempt at a multiple baseline like that of Experiment 2 was foiled by Karl's low rate of self-stimulation when his tutors asked frequent questions. Preliminary observation revealed that self-stimulation was substantial only in less structured situations, and ethical demands precluded our altering his work sessions to increase self-stimulation.

Data Collection

The primary purpose of this replication was to examine more closely the relationship between frequency and duration of self-stimulation. Hence, two separate measures of self-stimulation were obtained. The frequency data adhered to procedures followed in Experiments 1 and 2. Because their clinical responsibilities prohibited the tutors from maintaining a continuous record of self-stimulation, a second person was added to collect duration data. This observer was responsible for recording self-stimulation on a continuous 10-sec observe-and-record basis. He/she was cued to time intervals by a timer which emitted a tone through an earphone. When reliability data for the duration measure were collected, a second observer was able to plug into the same device. The tone was audible only to the observers. Two kinds of observers were used in Experiment 3. Duration observers were present during every session to collect these data; reliability observers were used to check the reliability of both frequency and duration data.

The observer who recorded frequency of self-stimulation was also responsible for obtaining a measure of density of reinforcement to determine whether DRO contained a richer schedule of reinforcement for incompatible behavior than did the other conditions. Every 10 sec, in addition to recording presence of self-stimulation, the duration observer noted whether the child

had received either primary (food) or secondary (praise) reinforcement for refraining from self-stimulation or for playing appropriately. For reinforcement to be considered directed toward incompatible behavior, the tutor had to specify the nature of the behavior and include the word "hands" (e.g., "Good using your hands to play" or "Nice quiet hands"). Reinforcement for task behavior similarly had to specify the behavior (e.g., "Nice coloring").

Reliability data. Reliability data were collected on 34 of the 46 days Karl was in the study with one session for the frequency measure and one for the duration measure as well as the measurement of reinforcement density. Unlike the previous experiments, data collection was not covert because of a breakdown in the sound system. Thus, tutors and duration observers were aware that reliability data were being collected, although all sat in positions which precluded their seeing each other's records.

A total of 230 sessions were conducted (46 days, 5 sessions per day). Reliability data were scattered over all five play activities and evenly distributed across conditions. Reliability for duration and density measures was calculated separately for each by dividing the number of agreements for 10-sec intervals by the number of agreements plus disagreements, then multiplying that number by 100.

Results

Reliability. Overall average interrater reliability for frequency of Karl's self-stimulation was 90.64%. During baseline the average was 91.7% (range 71% to 100%), during DRO the average was 86.57% (range 68% to 93%), during time-out the average was 85.14% (range 45% to 100%), and during overcorrection the mean was 100%. On 7 of the 34 days, the reliability fell below 80%.

Overall average interrater reliability for the duration of Karl's self-stimulation was 93.7%. During baseline the average was 92.6% (range 80% to 100%), during DRO the average was 93.75% (range 87% to 99%), during time-out

Fig. 4. Karl's mean frequency and mean duration of self-stimulation responses per 15-min session across all treatment conditions.

the average was 92.43% (range 75% to 99%), during overcorrection the mean was 99% (range 96% to 100%). On only 1 day did reliability fall below 80%.

The reliability data for density of reinforcement were similarly high with primary reinforcement of task averaging 100% across all conditions, secondary reinforcement of task 97.41%, primary reinforcement of incompatible behavior 99.7%, and secondary reinforcement of incompatible behavior 99.7%. On only 1 day did any of these fall below 90%.

Karl's performance. The two measures of Karl's self-stimulation, duration and frequency, are depicted in Figure 4. The mean pretreatment baseline for duration of self-stimulation was 37.27 and that for frequency was 27.4 over 6 days. The introduction of DRO led to a very modest decline in frequency ($\bar{X} = 20.15$) but little change in mean duration ($\bar{X} = 33.05$) over 8 days. There does, however, appear to

have been some initial decline in duration followed by an increase and leveling off. The return to baseline brought a rise in frequency to pretreatment levels ($\bar{X} = 29.29$) and a more substantial rise in duration ($\bar{X} = 43.50$) over 6 days. Timeout led to a decrease in both measures (frequency $\bar{X} = 17.9$; duration $\bar{X} = 28.19$) over the 8 days of treatment. The return to baseline brought the average for frequency back to 21.00 and that for duration to 46.4 over a 10-day period. This baseline was longer than the others because of the decline in frequency between Days 6 and 9. During overcorrection, the mean for both measures declined sharply (frequency $\bar{X} = 3.5$; duration $\bar{X} = 3.7$) over 5 days. A brief 3-day return to baseline saw frequency immediately return to a mean of 17.87 and duration to 34.60.

A comparison of duration data during timeout with time-out episodes included ($\bar{X} = 28.19$) and excluded ($\bar{X} = 19.28$) showed a

clear difference, with Karl self-stimulating a good deal while being ignored. As was the case with James, the pattern of responding on a day-to-day basis was identical for the two sets of data. In Karl's case, it was decided to use the data including time-out episodes since these reflected his total performance.

Frequency of reinforcement. The data regarding density of reinforcement are clear cut. Karl was never offered primary or secondary reinforcement for incompatible behaviors in any condition except DRO. During DRO, primary reinforcement averaged 14.18 and secondary reinforcement 14.28 in each 15-min session. Task reinforcement occurred across all conditions. There was no instance of primary reinforcement for play. The data for secondary reinforcement suggest a steady decline from the start of the study until overcorrection was instituted and an abrupt increase noted. Thus, the means for secondary reinforcement of play were baseline (16.03), DRO (13.38), baseline 2 (11.73), time-out (8.6), baseline 3 (8.6), overcorrection (20.64), and baseline 4 (14.73).

Relationship between frequency and duration of self-stimulation. Visual examination of the frequency and duration data show a general consistency. Nevertheless, there were some discrepancies as when duration appeared to rise slightly during the last days of DRO but frequency remained stable and during the baseline after time-out when duration rose quite substantially but frequency declined. With overcorrection, both measures dropped sharply. These data are consistent with the less precise measure of Experiment 2, which suggested a highly significant but imperfect relationship between the two measures.

DISCUSSION

The single most consistent finding from the present study was that overcorrection led to an immediate and dramatic decline in self-stimulation for all four boys. For James, this effect was so potent that in spite of the brevity of the treatment period, he failed to return to pretreatment

baseline responding after a full month. The procedure used for overcorrection differs substantially from that of Foxx and Azrin (1973) and should more precisely be called a derivative of overcorrection. It is important to note that the duration was very brief, only 10 sec, and, thus, far less than reported in most studies. For example, Epstein et al. (1974) used a 2.5-min procedure and Foxx and Azrin's (1973) procedure took 5 minutes. The primary differences between the present study and earlier studies may be that the boys were guided through the clapping motions for the full 10 sec, whereas in other studies subjects were encouraged to go through the generally more complex activity themselves. The short duration of the present method makes it practical for classroom and home application.

It might be misleading to label the present procedure as "positive practice," since the activity employed is not one that is especially desirable. We did not wish the boys ultimately to clap their hands over their heads instead of self-stimulating. Positive practice would more correctly have required them to write with a pencil, roll a car, or engage in another constructive activity. Wells et al. (1977) have demonstrated the feasibility of such an approach.

Hobbs and Goswick (1977) commented on the possibility that new forms of self-stimulation may appear when using overcorrection (Epstein et al., 1974; Rollings, Baumeister, & Baumeister, 1977). One of our children, James, confirmed this expectation. His head-bobbing behavior appeared after overcorrection had largely suppressed his other forms of self-stimulation. This was the only negative side effect, and when the definition of self-stimulation for him was expanded to encompass this behavior, it, too, responded to treatment.

The effects of time-out were variable for the boys. James showed no response at all. For Matt and Karl, time-out clearly had a suppressive effect. Matt's rate of responding declined on the second day the procedure was in effect and remained below baseline for 3 days. Likewise, both the frequency and duration of self-stimulation

declined below baseline levels for Karl. Alan's data are perhaps the most interesting. Time-out for him led to an initial increase in self-stimulation followed by a decrease. This is consistent with the informal observations of his tutors, who reported that Alan appeared more frustrated during the onset of time-out than during any other phase of the study. He seemed to be actively working to gain attention by kicking and hitting the tutors. This may have signified the value that attention from adults held for him.

The effectiveness of time-out as a treatment procedure is less clear than overcorrection. For some children, such as Matt, Karl, and perhaps Alan, the approach appears effective, although not as immediate in its impact as overcorrection. For others, such as James, there are some data suggesting that time-out may actually increase self-stimulation at least during time-out episodes. For those children for whom attention or access to other aspects of classroom life are of relatively little value, time-out may simply offer the opportunity to engage in self-stimulation unhindered by other demands. The accurate assessment of social reinforcement would be important in predicting the value of time-out for treating self-stimulation. It should also be noted that for none of the children did time-out eliminate self-stimulation.

DRO had no apparent effect on James or Alan and only a minimal impact on both the frequency and duration of Karl's self-stimulation. James and Alan were eligible for reinforcement once every 45 to 60 sec; Karl, on a richer schedule, every 20 sec. His schedule of reinforcement was much denser during DRO with primary and/or secondary reinforcement for incompatible behavior occurring about once every minute as opposed to zero in all other conditions. A richer schedule would have made it difficult to continue with the boys' normal routines. Matt, contrary to our expectations but consistent with Foxx and Azrin's (1973) experience, showed a distinct increase in responding with DRO. Although able to replicate this finding with him, we lacked sufficient time to do a more fine-

grained analysis of the response before the school year ended and he left for another program.

In addition to evaluating the effect of the three procedures upon self-stimulation, the present research explored the extent to which treatment effects generalized to another setting. An examination of Alan's behavior during play sessions demonstrated little generalization. Had the effect of overcorrection generalized to play, one would have expected to see a decrease in responding during that phase of the play baseline. There was none. This lack of generalization is a familiar problem with autistic youngsters (Harris, 1975). Had some "biological need" required him to compensate for decreased opportunity to self-stimulate during work, there might have been an increment in responding during play. This was not evident.

Experiment 3 attempted to address a serious methodological flaw in the preceding studies by simultaneous collection of frequency and duration data. Alan's data had shown a highly significant but not perfect correlation between the two measures. Theoretically, Alan could have engaged in a full 15 min of continuous self-stimulation and earned a frequency count of 1. In order to minimize this likelihood, his episodes of self-stimulation were disrupted at specific intervals. Matt and James self-stimulated in short bursts so that this was less of an issue with them. Nevertheless, recognizing the inherent limits of frequency data, we undertook Experiment 3 to explore this relationship.

The results of our experience with Karl show clearly that it is possible to decrease the frequency of a child's self-stimulation while actually observing an increase in the total amount of time he spends self-stimulating. While there was a generally consistent relationship between the two measures, there were several points of significant deviation—most specifically when duration increased and frequency remained stable in the later days of DRO and during the baseline after time-out when duration rose but frequency declined. The effective suppression during overcorrection brought both measures to

a rapid decline. It appears that while the frequency measure is a less accurate reflection of behavior change than duration, it nevertheless can provide clinical data of substantial value. Duration is to be preferred, especially for research, but frequency is not without informational value in clinical practice.

It is interesting to speculate about the changes in Karl's receipt of secondary reinforcement for play. The data reflect a steady decline in such reinforcement over time. If we take this as an indirect measure of performance, it suggests he was engaging in less appropriate play behaviors, regardless of the procedure in effect, until overcorrection was instituted. Perhaps he was growing bored. Whatever the reason, it is evident that the modest suppression of self-stimulation achieved by DRO and time-out had no beneficial effect on task performance. With overcorrection, however, there was an immediate and striking increase in secondary reinforcement. Karl was apparently engaging in considerably more play behavior during this time. Since the task remained the same, one would not think it likely he had simply become less "bored." Rather, there appears to have been an actively stimulatory effect from the overcorrection process. Such a finding needs to be examined with more direct measures of task performance in future research. If substantiated, this positive side effect would be additional support for the clinical value of overcorrection in the suppression of self-stimulation.

REFERENCES

- Azrin, N. H., Kaplan, S. J., & Foxx, R. M. Autism reversal: Eliminating stereotyped self-stimulation of retarded individuals. *American Journal of Mental Deficiency*, 1973, **78**, 241-248.
- Baumeister, A. A., & Forehand, R. Effects of contingent shock and verbal command on body rocking of retardates. *Journal of Clinical Psychology*, 1972, **28**, 586-590.
- Berkson, G. Visual defect does not produce stereotyped movements. *American Journal of Mental Deficiency*, 1973, **78**, 89-94.
- Berkson, G., & Davenport, R. K. Stereotyped movements in mental defectives 1: Initial survey. *American Journal of Mental Deficiency*, 1962, **66**, 849-852.
- Doke, L. A., & Epstein, L. H. Oral overcorrection: Side effects and extended applications. *Journal of Experimental Child Psychology*, 1975, **20**, 496-511.
- Epstein, L. H., Doke, L. A., Sajwaj, T. E., Sorrell, S., & Rimmer, B. Generality and side effects of overcorrection. *Journal of Applied Behavior Analysis*, 1974, **7**, 385-390.
- Foxx, R. M., & Azrin, N. H. The elimination of self-stimulatory behavior of autistic and retarded children by overcorrection. *Journal of Applied Behavior Analysis*, 1973, **6**, 1-14.
- Harris, S. L. Teaching language to nonverbal children with emphasis on problems of generalization. *Psychological Bulletin*, 1975, **82**, 565-580.
- Hobbs, S. A., & Goswick, R. A. Behavioral treatment of self-stimulation: An examination of alternatives to physical punishment. *Journal of Clinical Child Psychology*, 1977, **6**, 20-23.
- Klier, J. L., & Harris, S. L. Self-stimulation and learning in autistic children: Physical or functional incompatibility. *Journal of Applied Behavior Analysis*, 1977, **10**, 311. (Abstract)
- Koegel, R. L., & Covert, A. The relationship of self-stimulation to learning in autistic children. *Journal of Applied Behavior Analysis*, 1972, **5**, 381-387.
- Koegel, R. L., Firestone, P. B., Kramme, K. W., & Dunlap, G. Increasing spontaneous play by suppressing self-stimulation in autistic children. *Journal of Applied Behavior Analysis*, 1974, **7**, 521-528.
- Laws, D. R., Brown, R. A., Epstein, J., & Hocking, N. Reduction of inappropriate social behavior in disturbed children by an untrained paraprofessional therapist. *Behavior Therapy*, 1971, **2**, 519-533.
- Pendergrass, V. E. Time-out from positive reinforcement following persistent high rate behavior in retardates. *Journal of Applied Behavior Analysis*, 1972, **5**, 85-91.
- Rimland, B. The differentiation of childhood psychoses: An analysis of checklists for 2,218 psychotic children. *Journal of Autism and Childhood Schizophrenia*, 1971, **1**, 161-174.
- Rollings, J. P., Baumeister, A. A., & Baumeister, A. A. The use of overcorrection procedures to eliminate the stereotyped behaviors of retarded individuals: An analysis of collateral behaviors and generalization of suppressive effects. *Behavior Modification*, 1977, **1**, 29-46.
- Wells, K. C., Forehand, R., Hickey, K., & Green, K. D. Effects of a procedure derived from the overcorrection principles on manipulated and nonmanipulated behaviors. *Journal of Applied Behavior Analysis*, 1977, **10**, 679-687.

Received August 11, 1978

Final acceptance September 22, 1978