

“Think Aloud”: A Program for Developing Self-Control in Young Aggressive Boys¹

Bonnie W. Camp,² Gaston E. Blom,³ Frederick Hebert,
and William J. van Doorninck

University of Colorado School of Medicine

“Think Aloud” was designed as a training program to improve self-control in 6- to 8-year-old boys. It involved modeling and verbalization of cognitive activity to foster use of verbal mediation skills in dealing with both cognitive and interpersonal problems. It was hypothesized that this training would lead to improvement in test performance and teacher ratings of classroom behavior in hyperaggressive boys. Twelve aggressive second grade boys participated in daily, 30-minute, individual sessions for 6 weeks. Normal and aggressive control subjects received no intervention. Teachers rated both trained and untrained aggressive boys as improving in aggressive behaviors but they rated the experimental group as showing improvement on a significantly larger number of pro-social behaviors. The pattern of performance on cognitive tests also changed significantly in the experimental group. On pretest, their pattern differed from normals and resembled the aggressive control group, while on posttest their pattern resembled normals and differed from aggressive controls. Suggestions were made concerning additional refinements needed in the program, but overall results indicated potential value in the present approach for providing assistance to aggressive boys in the early grades.

Manuscript received in final form November 10, 1976.

¹This investigation, from the Departments of Pediatrics and Psychiatry, University of Colorado Medical School, was supported in part by a Research Scientist Award No. MK2-47 356 from the National Institute of Mental Health and by Grant No. NEG 003-0029 from the National Institute of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the NIE, and no official endorsement by the National Institute of Education should be inferred.

We are grateful to Mary Ann Bash and Margaret Simmons for their assistance in carrying out the program and to the Denver Public Schools for their continued cooperation.

²Address all correspondence to Bonnie W. Camp, University of Colorado School of Medicine, Container C250, 4200 East Ninth Avenue, Denver, Colorado 80262.

³Gaston E. Blom is now at Michigan State University, East Lansing, Michigan.

Impulsivity and difficulty maintaining sustained response inhibition are characteristics which may contribute to both poor achievement and aggressive behavior problems in children. Several previous programs have demonstrated that these characteristics can be decreased in the testing situation with training in verbalization of problem-solving strategies (Bem, 1967; Meichenbaum & Goodman, 1971; Palkes, Stewart, & Kahana, 1968). Except for Meichenbaum and Goodman, however, previous investigators selected children for these studies on the basis of test behavior only and not on the basis of deviant behavior. It is not clear, therefore, whether similar results would be obtained in a deviant population.

There are also questions concerning whether training may be expected to affect behavior outside the training or testing situation. Meichenbaum and Goodman attempted to measure effects of their program on classroom behavior but found no significant differences between children in their training program and an attention-control group. Though disappointing, their results were readily understandable. Their program was brief, consisting of four sessions over 2 weeks; they selected children on the basis of assignment to the remedial class rather than on behavioral criteria; and their training program consisted entirely of visual materials and impersonal tasks. In addition, it was not clear what rationale guided their selection of classroom behavior to be examined.

Thus cognitive behavior modification appears to have promise as a method for improving behavior and test performance in aggressive children but several features of previous programs need modifying to increase the likelihood of effects with this group. The present study reports on development of such a program along with results of initial trials with hyperaggressive second grade boys. The specific purpose was to determine whether test performance in hyperaggressive boys could be altered by cognitive behavior modification training procedures and whether evidence of impact on behavior outside the training sessions could be demonstrated.

METHOD

Subjects

As part of a larger study, 85 regular first and second grade classroom teachers in the Denver Public Schools rated all boys in their classes on Miller's (1972) School Behavior Checklist (SBCL) between October 1974 and January 1975. The SBCL is a 96-item checklist containing a wide variety of deviant and prosocial behaviors which have been factored into seven scales. SBCLs were scored according to Miller's norms and a T score of 70 (+2 SD above the mean) on the Aggressive scale was used as the basis for identifying aggressive boys. Out of 832 teacher ratings, 115 boys met the criteria of $T > 70$ on the Aggressive

scale. Parent permission for testing was obtained on 52. All second grade boys in this group ($N = 24$) were selected for participation in the present study.

At the beginning of this project, 23 aggressive second grade boys from 15 schools were still living in the project area. They were randomly assigned to an experimental group (Agg-Exp) of 12 and a control group (Agg-Con) of 11. The Agg-Con was reduced to 10 before group assignment was announced because the school social worker felt too much service was being provided.

The normal group consisted of boys who had no T score of 60 or higher on any SBCL scale. In the larger study, a total of 375 boys were so identified and parent permission for initial testing was received on 193. As each aggressive boy was selected for the larger study, a normal boy was selected randomly from those normal subjects in the same age group and census tract of residence. This resulted in a total of 24 second grade boys who received initial testing. Twelve of these were selected at random from schools where Agg-Exp boys were located to participate as normal controls (Norm-Con) for follow-up with no treatment other than regular classroom instruction.

Tests Administered

The complete battery of tests administered within 2 months prior to initiating the program is described in a previous paper (Camp, in press). For purposes of pretest-posttest comparisons in the present study, the following tests were repeated within 3 weeks after the program ended: Block Designs, Object Assembly, and Maze subtest from the Weschler Intelligence Scale for Children – Revised (WISC-R) along with recording of private speech;⁴ the Reading test from the Wide Range Achievement Test (WRAT); Auditory Reception from the Illinois Test of Psycholinguistic Abilities; Kagan's Matching Familiar Figures Test (MFF) with recording of private speech. These tests were selected because previous studies had shown that they contributed to a pattern of differences between aggressive and normal boys (Camp, in press) or because they had been used by Meichenbaum and Goodman (1971).

In addition to the above tests on which both pre- and posttest scores were available, an abbreviated version of the Preschool Interpersonal Problem-Solving Test (PIPS) (Shure & Spivack, Note 1) was administered as a posttest only. This was selected because it was expected to be influenced by the social training dialogues included in the treatment program.

⁴Recording and scoring of speech during the testing session is described more fully in Camp (in press). Word play and outer-directed, nonsocial speech were scored as Immature; inner-directed, self-guiding speech and inaudible muttering were scored as Mature; comments directed toward the examiner and tension-releasing verbalizations were scored Social.

Teacher Ratings

The complete SBCL was obtained originally in the fall of 1974 and again just prior to spring vacation in 1975. Correlations between these two administrations were in the range .82 to .84. Teachers completed items on two scales, Low Need Achievement and Aggression, again at the end of the training program. In addition to completing the checklist in the usual fashion, teachers were asked in this final administration to indicate for each item whether the child was "worse," "no change," or "improved" since spring vacation. Items on the two scales were presented in the same order as on the original SBCL. Teachers were unaware that two scales were involved and no mention was made of how or whether the items were to be separated.

Program

For this study, a training program entitled "Think Aloud"⁵ was designed for use with children in daily, 30-minute, individual sessions extending over 6 weeks.

The procedures used in training were very similar to those described by Meichenbaum and Goodman (1971) in placing heavy emphasis on modeling of cognitive strategies and concentrating on developing answers to the following four basic questions: What is my problem, What is my plan, Am I using my plan, and How did I do? (See Meichenbaum & Goodman, 1971, for a rationale regarding the choice of this group of questions.) To engage the child in reacting to all features of modeling (speech and action), the program used the "copycat" game initially. Then "copycat" was faded and the child was encouraged to verbalize his own strategy and eventually to fade the problem analysis and strategy planning to a covert level.

Problem-solving content included both cognitive, impersonal problems such as used by Meichenbaum and Goodman (1971) and interpersonal problem-solving games as described by Shure and Spivack (1974). In addition to the visual materials used by Meichenbaum and Goodman — e.g., Raven Matrices, PMA Perceptual Speed — we added auditory verbal tasks which would require blocking the first impulsive association and reasoning to a solution — e.g., a complex version of the Simon Says game, use of riddles, verbal justification of answers, and following complex instructions in a semantic conditioning task.

To increase the likelihood that techniques developed in the training sessions would generalize to interpersonal situations, the training program designed by Shure and Spivack (1974) for kindergarten children was adapted for a slightly

⁵For information concerning availability of the manual for the "Think Aloud" Program, contact the senior author.

older child. Their program consists of a carefully sequenced series of games identifying emotions, thinking about how people have different likes and dislikes and learning to gather information about other people, determining antecedents to an emotion, considering what might happen next in various situations, and evaluating fairness of outcomes. These are followed by games which pose the problem of finding several alternative solutions to interpersonal situations, anticipating consequences, and evaluating outcomes. In addition to formal games and dialogues, Shure and Spivack presented a general problem-solving approach which could be incorporated into conversations in problem situations as they arose naturally. This approach, especially as it dealt with encouraging development of several alternative plans, solutions, and possible outcomes, was incorporated into the cognitive portion of the program as well.

The program introduced easy cognitive problems first, then preliminary social games beginning with the fifth day. Thereafter, cognitive and social problems were intermixed.

Procedure

A manual and script were prepared which incorporated the previously described material and procedures into a 6-week curriculum. The procedures were rehearsed by the two teachers who were assigned to work with the children. Preliminary role-playing, observation, and coaching during trial runs with non-project children also served to standardize program delivery. In addition, tape recordings were made of each session with each child in the project. These were reviewed regularly by the senior author and served as a basis for monitoring and supervising conduct of the program. Although no formal ratings of fidelity to the program were made, these procedures produced as much homogeneity in program delivery as seemed feasible. The 12 boys in the Agg-Exp group were divided into two groups based on geographical location in the city. One of two teachers was randomly assigned to work with children in each group beginning the week after spring vacation. Children were introduced to the program by being asked to figure out what game the teacher was playing as they walked down the hall. The teacher began copying the child's words and actions until he recognized what she was doing. After switching roles and further play, she explained that they were going to use the "copycat" game to learn how to "think aloud" so it could help them solve problems. At the end of the first session she explained to the child that they would be getting together every day at the same time to learn more about how to use "Think Aloud."

At various opportune times during the 6 weeks, the teacher suggested that thinking out loud could help in the classroom and asked the child to think of ways he could use thinking out loud in doing his schoolwork or in getting along with others.

At the end of the 6-week training period, posttesting was completed by graduate students otherwise uninvolved in the project and teacher checklists were obtained again with the modifications described previously.

RESULTS

Test Data

One objective of the present study was to determine whether participation in the Think Aloud program would alter test performance in treated aggressive (Agg-Exp) boys so that their pattern of scores would resemble normal boys (Agg-Con). Information pertinent to this question was obtained in two ways. One was through analysis of data collected in a pretest-posttest design and one through data collected in a posttest-only design. In addition, analysis of pretest-posttest data was performed in two ways, one through univariate analysis of covariance on individual test scores and one through analysis of discriminant scores derived from discriminant function of analysis.

Univariate analyses of covariance, using the pretest score as covariate, were performed on data reported in Table I. Results of two a priori planned comparisons are also reported in Table I. One comparison involved examining the differences between the Agg-Exp group and the Agg-Con group. The second comparison examined differences between the Agg-Exp and the Norm-Con group.

Tests on which the Agg-Exp group were significantly different from Agg-Con included mazes, reaction time on the MFF and Salkind's (Note 2) Impulsivity score from the MFF.⁶ Reading achievement, prorated Performance IQ, and scores for Immature and Social Speech showed a trend toward the predicted differences. Tests on which the Agg-Exp group remained similar to the Agg-Con group and different from the Normal-Controls were Object Assembly and Relevant and Other Speech on the MFF. The Agg-Exp group also showed a trend toward significantly greater Inefficiency than Normals.

Data from all first and second grade subjects who participated in the larger study ($N = 93$) were utilized to derive a discriminant function score for pretest data. However, the discriminant function analysis used only those test variables on which both pre- and posttest measures were available. Coefficients derived from this analysis of pretest scores were then applied to posttest data available

⁶ Salkind (Note 2) has described a method for deriving an Impulsivity and an Efficiency score from the MFF. Average reaction time and total errors are first converted to z scores. The Impulsivity score (positive values represent more impulsivity) is derived by the formula $\frac{z_{\text{error}} - z_{\text{time}}}{2}$. The Efficiency score (positive values represent inefficiency) is derived by the formula $\frac{z_{\text{error}} + z_{\text{time}}}{2}$.

Table I. Pre- and Posttest Means and *F* Values for Planned Nonorthogonal Contrasts on Adjusted Means^a

Tests	Pretest group			Posttest group			<i>F</i> contrasts	
	A-E	A-C	N-C	A-E	A-C	N-C	A	B
	<i>N</i> = 12	<i>N</i> = 10	<i>N</i> = 12	<i>N</i> = 12	<i>N</i> = 10	<i>N</i> = 12	(A-Exp vs. A-Con)	(A-Exp vs. N-Con)
SBCL-AGG	71.0	75.2	46.0	69.2	74.2	47.3	n.s.	6.54 ^b
SBCL-LNA	58.9	57.7	42.4	57.6	57.1	41.4	n.s.	n.s.
WISC-R tests								
Block design	7.9	9.2	9.0	9.2	9.7	9.7	n.s.	n.s.
Object assembly	9.3	9.2	10.3	10.0	10.4	12.1	n.s.	3.38
Mazes	8.7	10.0	10.3	13.0	10.0	11.3	6.96 ^b	5.77 ^b
PIQ	90.7	95.8	98.6	104.6	99.7	106.3	3.51	n.s.
ITPA								
Auditory reception	29.0	34.7	38.7	32.6	35.8	33.7	n.s.	n.s.
WRAT	2.1	2.6	2.7	2.5	2.7	3.3	4.08	n.s.
WISC-R private speech								
Immature	3.8	12.7	4.8	5.2	3.1	1.9	3.89	n.s.
Mature	30.7	30.2	18.4	8.1	8.0	8.0	n.s.	n.s.
Social	10.8	9.5	7.3	3.4	5.7	1.8	2.99	n.s.
MPF								
Inner speech	1.3	.5	.6	.6	.8	1.2	n.s.	n.s.
Other speech	2.9	1.9	.4	13.0	11.0	3.9	n.s.	5.76 ^b
Relevant speech	2.5	1.9	.8	7.2	6.7	2.7	n.s.	4.72 ^b
Irrelevant speech	.5	.7	.2	.6	1.0	.3	n.s.	n.s.
Time (average)	10.1	9.7	11.5	17.1	10.6	11.2	6.39 ^b	5.89 ^b
Errors (total)	18.7	19.3	14.9	12.6	14.2	11.2	n.s.	n.s.
Impulsivity	.43	.58	-.36	-.31	.35	.05	4.69 ^b	n.s.
Efficiency	.03	.04	-.23	.30	-.05	-.25	n.s.	3.86

^a*F* values with $p > .10$ are reported as n.s.

^b $p < .05$.

^c $p < .01$.

on subjects in the present study. In this fashion, each subject in the present study received both a pretest discriminant score and a posttest discriminant score. Table II shows the standardized discriminant function coefficients for variables used in the present study in order of importance. In this table, the larger the value, the greater the weight placed on this score, irrespective of the sign. The positive and negative signs associated with the centroids indicate whether the sign associated with each discriminant score results in weighting toward the aggressive or the normal group.

A repeated measures analysis of variance was used to evaluate hypotheses regarding change in the pattern of test scores. The hypothesized treatment effects were expected to result in a significant interaction between time of discriminant score and treatment group with the Agg-Exp and Agg-Con groups differing from Normals on pretest and Agg-Con differing from Agg-Exp and Normals

Table II. Standardized Discriminant Function Coefficients Derived from Analysis of Test Patterns in Aggressive and Normal Boys

Measure	Coefficient
MFF – IRREL speech	.46
MFF – average response time	-.44
WISC – IMM speech	.40
WISC – mazes	-.36
ITPA – auditory reception	-.32
MFF – errors	.30
WISC – object assembly	-.15
WRAT – reading grade	-.15
WISC – social speech	-.09
WISC – mature speech	-.05
MFF – revelant speech	.05
WISC – block design	-.02
Centroid	
Aggressive	.49
Normal	-.49

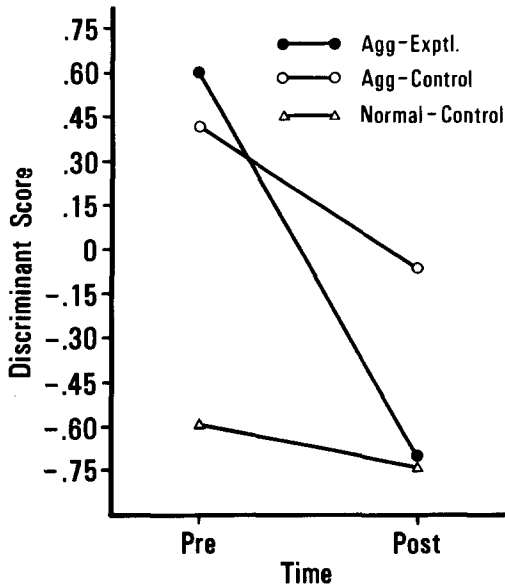


Fig. 1. Pre- and postdiscriminant scores for Aggressive-Experimental, Aggressive-Control, and Normal-Control groups.

on posttest. The pre- and posttest centroids for subjects in the present study were .60 and $-.69$ for Agg-Exp, .42 and $-.03$ for Agg-Con, and $-.58$ and $-.73$ for Norm-Con. These are presented graphically in Figure 1.

Results of the analysis of variance confirmed all hypotheses. Significant differences were observed for pre- versus postdiscriminant scores, $F(1, 30) = 14.47, p < .001$; for Treatment Group, $F(2, 30) = 3.80, p < .05$; and the Discriminant Score \times Treatment Group interaction, $F(2, 30) = 4.29, p < .05$. The mean square error for testing contrasts within the interaction was .447. These contrasts showed a significant difference on pretest between Norm-Con and the Aggressive groups combined, $F(1, 30) = 21.74, p < .001$, and nonsignificant differences between the two aggressive groups. On posttest, significant differences were observed between Agg-Con and Agg-Exp, $F(1, 30) = 5.32, p < .05$, while differences between Agg-Exp and Norm-Con were nonsignificant.

The PIPS test was given as a posttest only. According to Shure and Spivack, emotionally and behaviorally disturbed children tend to offer fewer different solutions to the problems while engaging in more repetitive talk, irrelevant answers, etc. Group means for these different scores are present along with results of a one-way analysis of variance in Table III. Agg-Exp boys gave significantly more solutions than either Agg-Con or Norm-Con boys and they showed a trend toward surpassing Agg-Con in the proportion of solutions to total output. However, whatever advantage they may have had in this regard was probably offset by the fact that it appears to have been gained through their use of a higher proportion of aggressive solutions than either Agg-Control or Normal Controls. Thus, the training program seems to have loosened their tongues but failed to assist them toward developing enough constructive alternatives.

Table III. Comparison of Posttreatment PIPS Scores in Three Groups

Category	Group means			F contrasts	
				A	B
	A-E <i>N</i> = 12	A-C <i>N</i> = 10	N-C <i>N</i> = 13	(A-E vs. A-C)	(A-E vs. N-C)
Total solutions	8.8	7.2	6.4	5.66 ^a	14.55 ^b
Total other talk	5.5	7.6	5.5	n.s.	n.s.
Sol/Sol + talk	.659	.509	.551	3.65 ^c	n.s.
Agg Sol/Tot Sol	.241	.057	.038	8.26 ^b	11.50 ^b

^a $p < .05$.

^b $p < .01$.

^c $p < .10$.

Teacher Ratings

A second objective in the present study was to determine whether teacher ratings would indicate significant improvement in the Agg-Exp boys. These ratings were analyzed both in terms of *T* scores on the LNA and Aggressive scales of the SBCL and ratings of change on individual items. Pre and post *T* scores on the LNA and Aggressive scales are included in Table I. Analysis of covariance on the post-test scores showed no significant differences between Agg-Exp and Agg-Con while Agg-Exp and Normals differed only on the Aggressive scale. This amount of stability is not surprising since these scores are fairly gross measures and the checklist was not prepared as a dependent measure.

Teacher ratings of improvement, on the other hand, should be more sensitive to smaller degrees of change than would be required to alter a *T* score. Groups were therefore compared on the average number of items improved on the LNA and Aggressive scales. It was predicted that Agg-Exp would improve more than either Agg-Con or Norm-Con on both of these scales. In this instance, however, it was also of interest to determine whether Agg-Con showed more or less improvement than Norm-Con as it was to compare Agg-Exp results with the other two groups. The Tukey test recommended by Winer (1971) was used to evaluate differences between all pairs of means. On the Aggressive scale, the average number of items improved was 9.7 for Agg-Exp, 10.9 for Agg-Con, and .75 for Normal-Control. For the Tukey test, a difference of 4.87 was necessary to reject the hypothesis of no difference between these means at $p < .05$. By this criterion, the two Agg groups did not differ from each other but did differ significantly from the Normal-Controls.

On the LNA scale, however, the Agg-Exp group showed significantly more improvement than the Agg-Control group, which in turn did not differ significantly from the Normal-Control group in number of items improved. Here the mean number of items improved was 10.3, 3.9, and 1.7 for the Agg-Exp, Agg-Control, and Normal-Control groups, respectively. A difference of 4.57 was required for rejecting the null hypothesis at the .05 level of confidence.

DISCUSSION

The purpose of the present study was to evaluate the impact of 6 weeks with the Think Aloud program on test performance and teacher ratings of behavior in young aggressive boys. Specifically, it was hypothesized that the treated boys would show changes in the direction of becoming more like normal boys and less like untreated aggressive controls. This hypothesis was most clearly supported by results of the teacher ratings of improvement in prosocial behavior and by changes in the pattern of performance on a battery of cognitive tests.

The pre- and posttest findings are consistent with the results reported by Meichenbaum and Goodman (1971) following a similar brief program for boys selected only on the basis of test impulsivity. Although some of the tests contributing to the discriminant score involved activities similar to those used in the training program (WISC-R Mazes, Matching Familiar Figures), others did not (WRAT Reading). Furthermore, the private speech measures obtained during testing could be considered measures of spontaneous generalization from the training program to the testing situation.

The demonstration of improved prosocial behavior in the classroom is encouraging, even though the treated group did not differ significantly from the untreated group in reduction of aggressive behavior. However, serious questions could be raised about the validity of teacher ratings both as a measure of change and as a measure of treatment effect in the present study.

Although some have questioned whether teacher ratings accurately reflect the behavior they are purportedly measuring, others have found them to be useful in evaluating effects of treatment in disturbed children (Conners, 1972; Werry & Sprague, 1970; Werry, Sprague, & Cohen, 1975). Furthermore, teacher reports are a form of data which is important in its own right. In contrast to the limited time samples often involved in obtaining observational data, teachers have the advantage of prolonged daily contact which provides a large time sample on which to base their report, particularly in regard to infrequent behaviors. In addition, there are significant correlations between independent behavior observations and teacher report ($r = -.38$) when it involves frequent behaviors readily observed by both teachers and observers, as in the case of items of the LNA scale (Camp & Zimet, 1974) versus observations of off-task behavior.

In the present study, a more serious question concerns the validity of the teacher ratings as a measure of change. Since teachers knew whether the child being rated was in the program or not, one could propose that the expectation of behavioral change with the program could result in general ratings of improvement. The fact that there is a difference in results of their ratings on the two scales argues against this explanation.

The possibility that unreliability in the scale influenced results must also be considered. Though reliability has not been established for the rating form on which teachers rated improvement, there is no reason to believe that unreliability would be greater for one group of aggressive boys than another. Unreliability in the measures would have tended to decrease the likelihood of finding significant differences. Yet the magnitude of differences on the Aggressive scale was so small it is difficult to believe that the two groups actually differed on this scale. Differences on the prosocial scale were large enough to be significant despite possible unreliability.

It therefore seems reasonable to conclude that the Think Aloud program was powerful enough to produce significant improvement in both test performance

and classroom behavior in aggressive boys. At the same time, the program failed to produce changes in all areas examined and in some instances may have had a negative effect. The points where the program seemed to fail was in channeling the voluminous verbal output of aggressive boys and in improving efficiency. In designing the program we failed to recognize the extent to which chatter, silliness, and inappropriate verbal activity might actually interfere with goals of the program. There were handled by ignoring them rather than by trying to attack them directly. Similarly, aggressive solutions to interpersonal problems were considered along with nonaggressive solutions, and no effort was made to consider these categorically "bad." Rather, they were evaluated in light of the consequences along with other kinds of solutions.

Apparently this approach was not enough, given the current program. It is of course possible that more time with the program as it stands could have brought the children to a more appropriate level of functioning with regard to quality of verbal output, efficiency, and social appropriateness of their thinking. Another approach, however, would be to structure the program differently so that silliness, for example, is attacked directly and the negative consequences of aggressive solutions are emphasized more.

And finally, even if one accepts the teacher ratings as valid, the present design does not allow us to determine whether the observed results, especially the ratings of improved classroom behavior, are attributable to the type of program or to increased individual attention. At this point in our research, however, it was felt that the first priority was to determine whether any alteration in behavior or test performance could be demonstrated in hyperaggressive boys. Further research should help to clarify the specific aspects of the program which are most responsible for the observed changes and to offer less questionable data in support of the "real life" changes. Results of the present study are nevertheless encouraging and suggest that further efforts to develop this approach for use by school personnel may be quite fruitful.

REFERENCE NOTES

1. Shure, M. B., & Spivack, G. *The PIPS test manual*. Unpublished manuscript, Hahnemann Medical College and Hospital, Philadelphia, Pennsylvania.
2. Salkind, N. J. *Errors and latency on the MFFT: A reassessment of classification strategies*. Paper presented at the Society for Research in Child Development, Denver, Colorado, April 1975.

REFERENCES

- Bem, S. L. Verbal self-control: The establishment of effective self-instruction. *Journal of Experimental Psychology*, 1967, 74, 485-491.
- Camp, B. W. Verbal mediation in young aggressive boys. *Journal of Abnormal Psychology*, in press.

- Camp, B. W., & Zimet, S. G. The relationship of teacher rating scales to behavior observations and reading achievement of first-grade children. *The Journal of Special Education*, 1974, 8, 353-359.
- Conners, C. K. Pharmacotherapy of psychopathology in children. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood*. New York: Wiley, 1972.
- Meichenbaum, D. H., & Goodman, J. Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology*, 1971, 77, 115-126.
- Miller, L. C. School behavior check list: An inventory of deviant behavior for elementary school children. *Journal of Consulting and Clinical Psychology*, 1972, 38, 134-144.
- Palkes, H., Stewart, M., & Kahana, B. Porteus maze performance of hyperactive boys after training in self-directed verbal commands. *Child Development*, 1968, 39, 817-826.
- Shure, M. B., & Spivack, G. *A mental health program for kindergarten children: A cognitive approach to solving interpersonal problems*. Philadelphia: Community Mental Health/Mental Retardation Center, Department of Mental Health Sciences, Hahnemann Medical College and Hospital, 1974.
- Werry, J. S., & Sprague, R. L. Hyperactivity. In C. G. Costello (Ed.), *Symptoms of psychopathology: A handbook*. New York: Wiley, 1970.
- Werry, J. S., Sprague, R. L., & Cohen, M. N. Conners' teacher rating scale for use in drug studies with children - an empirical study. *Journal of Abnormal Child Psychology*, 1975, 3, 217-229.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1971.