BEHAVIOR THERAPY 35, 561–585, 2004

An Evaluation of the Summer Treatment Program for Children With Attention-Deficit/Hyperactivity Disorder Using a Treatment Withdrawal Design

ANDREA M. CHRONIS
GREGORY A. FABIANO
ELIZABETH M. GNAGY
ADIA N. ONYANGO
WILLIAM E. PELHAM, JR.
ANDY LOPEZ-WILLIAMS
ANIL CHACKO
BRIAN T. WYMBS
ERIKA K. COLES
KAREN E. SEYMOUR

The State University of New York at Buffalo

The summer treatment program (STP) is a comprehensive intervention for attention-deficit/hyperactivity disorder (ADHD) that combines several empirically validated, behavioral treatment components in order to address multiple areas of impairment. The current study utilized a BAB treatment withdrawal design to evaluate the effectiveness of the STP. Participants included 44 children diagnosed with ADHD and comorbid disruptive behavior disorders who were enrolled in the STP. During the 6th week of the STP, behavior was measured during a planned period where all behavioral treatment components were withdrawn. Treatment was then reinstated in its

Andrea Chronis and Karen Seymour are now at the Department of Psychology, University of Maryland, College Park.

During the preparation of this report, Dr. Chronis was supported by grants from the National Institute of Mental Health (1 R03 MHO70666-1) and McNeil Consumer and Specialty Pharmaceuticals, manfacturer of Concerta. Gregory Fabiano was supported by National Research Scientist Award (NRSA) predoctoral fellowship (MH64243-01A1). Dr. Pelham was supported by grants from the National Institute of Alcohol Abuse and Alcoholism (AA11873), the National Institute on Drug Abuse (DA12414), the National Institute of Mental Health (MH 53554, MH62946), and the National Institute of Neurological Disorders and Stroke (NS39087).

Address correspondence to Andrea Chronis, Ph.D., University of Maryland, Department of Psychology, College Park, MD 20742; e-mail: achronis@psyc.umd.edu.

005-7894/04/0561–0585$1.00/0

entirety. Across measures of behavior, academic functioning, and teacher, counselor, and child ratings, substantial behavioral deterioration occurred during the withdrawal period, and behavior returned to previous levels upon the reinstatement of the STP treatment components. Results support the efficacy of the STP as an intervention for ADHD across multiple domains of impairment, including classroom and peer functioning.

Behavioral interventions have been widely employed for more than 30 years to successfully treat attention-deficit/hyperactivity disorder (ADHD; Pelham, Wheeler, & Chronis, 1998). Such interventions typically include outpatient-based clinical behavior therapy (i.e., parents and teachers are taught to implement behavioral strategies) or direct contingency-management strategies (i.e., behavioral interventions are implemented directly in the child's natural settings — such as classrooms, in the home, or on the playground — by paraprofessionals, consulting professionals, or expert teachers; Pelham & Murphy 1986). Contingency-management strategies usually result in greater behavioral effects than clinical behavior therapy (Pelham et al., 1998). Arguably, such intensive strategies are necessary to provide the foundation for comprehensive, chronic care needed to effectively treat a child with ADHD (American Academy of Pediatrics, 2001), as well as to make clinically meaningful changes in functional domains that predict later outcome (e.g., peer relationships; Coie & Dodge, 1998).

One example of comprehensive treatment for ADHD that intensively addresses important domains of impairment is the children's summer treatment program (STP; Pelham, Fabiano, Gnagy, Greiner, & Hoza, in press; Pelham, Greiner, & Gnagy, 1997; Pelham & Hoza, 1996). The STP combines evidence-based, contingency-management treatment components, including a token or point system, social reinforcement (i.e., praise), effective commands, time-out, and a daily report card (DRC) (for a review, see Pelham et al., 1998). These treatments are implemented across recreational and academic settings to improve children's peer relationships, interactions with adults, academic performance, and self-efficacy. Concurrently, parents are instructed in behavior management using well-established parent training programs; blinded, placebo-controlled, stimulant medication assessments are conducted when indicated; and follow-up treatment is provided in the child's school. An extensive treatment manual describes the program (Pelham, Grenier, et al., 1997), which has been refined and developed for more than 20 years.

Pelham and Hoza (1996) reported on 258 boys with ADHD, between the ages of 5 and 12, who participated in the STP. Before and after the STP, parents completed a standardized symptom rating scale (i.e., the Disruptive Behavior Disorder Rating Scale; Pelham et al., 1992), and an impairment rating scale (Fabiano et al., 1999). Results indicated statistically significant reductions in parent ratings of symptoms and impairment. In addition, parents, STP teachers, and STP counselors completed improvement ratings at the end of each summer (see also Pelham, Gnagy, et al., 2000). The overwhelming majority of children were rated by parents, teachers, and counselors as at least "somewhat" improved in important functional domains, such as rule-

following, classroom productivity, sports skills, and self-esteem — with many being rated as "much" or "very much" improved. Overall, the results of the Pelham and Hoza report indicated large behavioral effects of the STP reported by multiple raters across important domains of functioning. Results were similar across a variety of demographic, diagnostic, and socioeconomic factors (e.g., children exhibiting comorbid aggression, living in single vs. two-parent households, from diverse economic backgrounds).

Similar results were obtained across a variety of domains and measures in the STP conducted as part of the Multi-Modal Treatment Study of Children with ADHD (MTA; MTA Cooperative Group, 1999; Wells et al., 2000). In a subsample of the MTA sites, comparisons of pre- and posttreatment ratings, point system data, DRCs, improvement ratings, and parent satisfaction ratings revealed very large effects of behavioral treatment for both medicated and unmedicated children during the STP (Pelham, Gnagy, et al., 2000). The effects of the behavioral program alone were so large that incremental medication effects were obtained on only 5 measures out of more than 80 during the STP. Clearly, the effects of the STP appear to be large and are evident across raters, but the results from both of these studies are limited because the STP was not compared to a no-treatment control condition.

Given that the overall STP treatment package appears to produce or contribute to behavioral improvements (August et al., 2001; Pelham, Gnagy, et al., 2000; Pelham & Hoza, 1996), several studies have been conducted to determine the incremental benefit of individual treatment *components* in the STP treatment package. One such study investigated the incremental benefit of including time-out at varying levels of intensity in the STP (Fabiano et al., 2004). Time-out conditions (i.e., no time-out, a 5-minute, 15-minute, or time-out in which time was added or removed contingent upon behavior) were randomized on a weekly basis within subjects. Results indicated that the presence of any time-out condition significantly reduced the frequency of aggressive, destructive, and noncompliant behaviors beyond the effects of the other standard STP treatment components.

Similarly, the classroom-based components of the STP (i.e., DRC, point system, time-out) were evaluated in two studies using a crossover design, in which behavioral treatment components were implemented during some weeks but not others (Carlson, Pelham, Milich, & Dixon, 1992; Pelham et al., 1993), and behavioral treatment was crossed with medication. In these studies, the classroom behavioral components resulted in clinically and statistically significant improvements across a variety of measures (e.g., observations of disruptive behavior, teacher ratings, and classroom rule violations) relative to weeks when the behavioral classroom management procedures were not in place. The results of these studies strongly indicate that the STP classroom management procedures resulted in large behavioral effects. However, the effects of behavior modification were not significant for all measures and were smaller in size than the effects of high doses of medication. Results of the behavioral manipulation in these studies may have been limited by the

fact that the behavior modification procedures were continuously implemented during other times of the day (e.g., recreational activities, transitions from one activity to another). Therefore, carryover from the activities other than the classroom (e.g., baseball games) in which behavioral interventions remained intact may have influenced results found in the classroom setting by diluting the difference between the behavioral conditions.

These studies were replicated and extended by Kolko, Bukstein, and Barron (1999), who manipulated both medication and behavior modification in classroom and recreational activities during a summer treatment and enrichment program modeled after the STP. Participants included children with ADHD and comorbid disruptive behavior disorders who were referred for partial hospitalization. Within this sample, 75% of children were African American, and 44% of their families received welfare. Results of this study suggested that both medication and behavior modification had unique and incremental effects that differed across settings and individual children. Specifically, behavior modification improved oppositional behavior in both settings, and improved inattention/hyperactivity, positive behavior, and peer conflicts in the classroom. Behavioral intervention resulted in incremental effects beyond the effects of medication on negative behavior in the recreational setting. Medication improved symptoms of inattention and overactivity in both settings, and oppositional behavior and peer conflicts in the recreational setting. There were *no* incremental effects of medication beyond behavioral intervention in the classroom. Thus, the results of this study conducted by an independent research group support the effects of a modification of the STP intervention in both recreational and classroom settings for a racially and economically diverse sample.

Based on the results of these studies, the STP appears to be an effective contingency-management intervention for children with ADHD. However, only one controlled study (Kolko et al., 1999) compared the entire, multicomponent STP package to a control condition. To extend the existing literature on the STP by evaluating the entire multicomponent program in a *controlled* fashion, a treatment withdrawal study was conducted. Using a BAB, within-subjects design, behavioral treatment components of the STP package were withdrawn and reinstated in order to determine whether their removal would result in deterioration of child behavior.

Methods

Participants

Participants included 44 children[1] enrolled in the 2000 STP at the State University of New York at Buffalo. Participants were recruited from newspaper and

[1] One child was presented previously in a published case study (Chronis et al., 2001); however, the behavioral treatment manipulation described herein comprised less than one-third of the case study.

radio advertisements and mailings to local pediatricians, primary care physicians, mental health professionals, community schools, university employees, and families who called the Center for Children and Families at the State University of New York at Buffalo seeking services. All children met *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*; APA, 1994) criteria for ADHD, any subtype, based upon a semistructured *DSM-IV* parent interview (Pelham, 2002) and parent and teacher rating scales. Thirty-five of the children met *DSM-IV* criteria for ADHD–combined type; 8 met criteria for predominantly-inattentive subtype and 1 met criteria for the predominantly-hyperactive/impulsive subtype. In addition, 24 of the children (54.5% of the sample) met *DSM-IV* criteria for oppositional/defiant disorder (ODD) and an additional 13 children (29.5%) met criteria for conduct disorder (CD). The vocabulary and block design modules of the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991), and the Wechsler Individual Achievement Test Screener (WIAT; Wechsler, 1992) were administered to assess whether co-occurring learning problems were present.

Following the assessment, a 12-page consent form that provided a detailed description of each aspect of the STP was carefully read to and reviewed with the parents. On the consent form, it was explained that the various components of the treatments would be implemented in varying degrees during the summer. Specifically, parents were told that for a brief period of time, the behavior modification components would be withdrawn in order to learn how well each child was responding to the program and to provide the staff with additional information that could be used to make treatment recommendations. Parents were assured that behavior modification components would be reintroduced once it was determined that these treatment components were necessary to maintain the behavior of their child or a substantial portion of children in his/her group. Parents were encouraged to ask questions as they arose during the consent procedures, and these procedures were audiotaped.

Participants ranged in age from 6 to 13 years ($M = 10.2$, $SD = 1.96$) and had an estimated IQ above 80. The sample was 95% Caucasian and 90% male. As an index of socioeconomic status, parental education level (taking the higher of the two education levels for two-parent households) was high school for 11% of the sample, partial college/technical school for 25% of families, college/university for 27% of the families, and graduate-level training for 36% of the parents. Twenty-eight (63.6%) children lived with both parents, and 16 (36.4%) lived predominantly in single-parent households. Table 1 presents participant characteristics.

Twenty-five children were unmedicated throughout the study period, whereas 14 of the children were steadily medicated throughout the study period (11 and 3 children received stable doses of a stimulant medication and antidepressant, respectively). Five additional children were undergoing clinical, placebo-controlled medication evaluations during the study period or began an evaluation following the study period (Pelham & Hoza, 1987); only days on which these children *did not* receive active medication were used for

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR PARTICIPANT CHARACTERISTICS

Item	*M*	*SD*
Age in years	10.2	1.96
Full scale IQ[a]	108.9	20.9
Reading achievement[b]	106.7	14.9
Arithmetic achievement[b]	104.3	15.1
Spelling achievement[b]	101.4	14.8
DSM-IV items endorsed by parents or teachers		
Inattention	8.4	1.2
Hyperactivity/impulsivity	7.4	1.9
Oppositional/defiant	5.8	2.2
Conduct disorder	2.0	1.9
Abbreviated Conners Rating Scale–Parent[c]	18.0	4.6
Abbreviated Conners Rating Scale–Teacher	16.7	5.8
IOWA Conners Teacher Rating Scale Inattention-Overactivity[d]	10.0	3.0
IOWA Conners Teacher Rating Scale Oppositional-Defiant[d]	5.9	4.7
Disruptive Behavior Disorders Parent Rating Scale[e]		
Inattention	2.2	0.5
Hyperactivity/impulsivity	1.8	0.6
Oppositional/defiant	1.5	0.6
Conduct disorder	0.3	0.3
Disruptive Behavior Disorders Teacher Rating Scale[e]		
Inattention	2.2	0.7
Hyperactivity/impulsivity	1.8	0.6
Oppositional/defiant	1.2	0.8

[a] Wechsler Intelligence Scale for Children — 3rd ed. (1991).
[b] Wechsler Individual Achievement Test (1992).
[c] Goyette, Conners, & Ulrich, 1978.
[d] Loney & Milich (1982).
[e] Pelham, Gnagy, Greenslade, & Milich (1992).

this investigation. Children who were unmedicated were compared with children who were medicated on IOWA Conners (Loney & Milich, 1982; Pelham, Milich, Murphy, & Murphy, 1989) and Impairment Rating Scale (Fabiano et al., 1999) scores completed upon initial assessment into the STP to determine whether differences in ADHD symptomatology or impairment existed between these two groups. There were no significant differences between groups on IOWA Conners Teacher Rating Scale scores or Impairment Rating Scale scores. On the IOWA Conners Parent Rating Scale, parents rated medicated children as more severe than unmedicated children on the Inattention/Overactivity ($M = 11.3, SD = 2.2$ versus $M = 9.5, SD = 2.4$ for medicated and unmedicated children, respectively, $t[42] = -2.29, p < .05$) and Oppositional/

Defiant scales ($M = 9.6, SD = 4.2$ versus $M = 7.2, SD = 2.7, t[42] = -2.21$, $p < .05$, for medicated versus unmedicated children, respectively).

Design and Setting

The STP is an 8-week, comprehensive, manualized behavioral treatment program that combines behavior modification, sports skills training, social skills training, and problem-solving skills training in an integrated package (Pelham et al., in press; Pelham, Greiner, et al., 1997; Pelham & Hoza, 1996). Throughout the day, 12 children of similar ages and/or developmental levels were supervised by one lead counselor with previous STP experience and four undergraduate counselors, who provided high rates of praise and immediate feedback regarding positive and negative behavior and associated point gains and losses. Two of the 9 hours each day were spent in an academic classroom, staffed by a certified teacher and an aide. Children also spent 1 hour each day in an art classroom, staffed by an art teacher and one or two aides. The remainder of the day was spent in group-based recreational activities (softball, soccer, and basketball skill drills and games, and swimming). Parents attended a weekly parent training group to learn effective strategies for improving child behavior in the home setting (Cunningham, Secord, & Bremner, 1997). Clinical supervision of counselors and classroom staff was provided by an advanced doctoral student in clinical psychology (A.M.C.) and a Ph.D.-level psychologist (W.E.P.).

Treatment components described below were implemented in their entirety for the first 5 weeks of the 8-week program. During the 6th week, each of the four groups was randomly assigned to a 2-day period in which all behavior modification techniques were withdrawn. Half of the children had the procedures removed on Monday and Tuesday, with the procedures to be reinstated Wednesday, and half had the procedures in place on Monday and Tuesday and removed for Wednesday and Thursday. It was planned that individual children whose behavior deteriorated to extreme levels before the 2-day withdrawal had elapsed would have individual behavior modification components (DRC, time-out, individualized programs) reinstated immediately.

For two groups, the withdrawal of behavior modification was implemented for the 2-day period (one child from these two groups had behavior modification reinstated after 1 day due to severe negative behavior). For the two other groups (the youngest and oldest groups), the withdrawal was terminated early due to extreme disruptive and dangerous behavior (i.e., after 5 hours for the youngest group and after 6 hours for the oldest group). Data for these children were prorated when necessary for the remainder of the day using behavior frequencies from the portion of the day that behavior modification was not in place.

Behavior Modification Procedures

In the STP, a token economy (i.e., point system) was in place continuously throughout the day, in which children received immediate behavioral feed-

back and points that were exchanged for daily and weekly reinforcers (e.g., recess, field trips, social honors). Children earned points for appropriate behaviors (e.g., following rules, helping) and lost points for inappropriate behaviors (e.g., interrupting, teasing). Children were also assigned time-out for more serious behaviors, including intentional aggression, intentional destruction of property, and repeated noncompliance. A point system was in effect in the classrooms as well. Children began each of three classroom segments and the art classroom with a set number of points and lost points each time they violated one of the seven classroom rules; they earned points for work completion and accuracy. Children received DRCs, on which clinical staff members identified individual target behaviors and goals. Parents were taught in behavioral parent training classes to provide the children with home rewards based on their DRC performance. In addition, when the standard behavioral components were insufficient to produce behavioral change for a child, an assessment of the problem behavior was conducted and an individualized program was developed to target the child's specific problem areas.

Behavior Modification Withdrawal Procedures

During the 2-day withdrawal period, the children did not earn and lose points based upon their behavior, and privileges that had been contingent on behavior were provided to all children noncontingently (i.e., regardless of their behavior). Children continued to receive feedback about their behavior (e.g., "That's interruption"), but no points were associated with this feedback. Staff members were also instructed to refrain from providing praise for positive behavior. Morning social skills training groups were replaced with a discussion of sports skills that followed a similar format (e.g., staff-directed discussion, role-plays). Activity rules were not reviewed prior to each activity; however, children were given feedback when they broke a rule (e.g., "You stepped out of the transition line"). Children did not receive time-out for serious negative behaviors, but were told to leave the classroom if their behavior was so disruptive that the teacher was unable to instruct the class. In such cases, children were instructed to continue their work at a desk in the hallway outside of the classroom until they regained composure and were immediately returned to the classroom. Furthermore, children who engaged in dangerous or destructive behavior were physically guided or physically managed to ensure the safety of themselves and others. Finally, the children did not receive DRCs. Parents were not given feedback about their children's behavior at the end of the treatment day, and were instructed to refrain from providing home rewards or punishments based on STP behavior.

Prior to the beginning of this withdrawal period, counselors and teachers explained to children that these behavioral treatment components (e.g., the point system, DRC, time-out) would not be in place for 2 days. It was explained that they would still be told when they were behaving appropriately or inappropriately, but that they would not earn or lose points, that time-out would not be a consequence for serious negative behavior, and that all reinforcement (e.g.,

privileges, rewards) would be provided independent of behavior. Children were told that the reason they would not have any of these contingencies was to see how well they could behave.

Treatment Integrity/Fidelity

All staff members participated in an intensive, 2-week training program that consisted of lectures on behavior modification, small group activities, and role-play of STP recreational and classroom activities. Trainees were required to demonstrate competency in the procedures for specific techniques (e.g., time-out, physical management). Supervisors (A.M.C. and A.N.O.) and research staff completed treatment integrity and fidelity (TIF) forms and reliability observations regularly throughout the program to ensure that staff members were implementing the treatment as intended, both during the standard program and the behavior modification withdrawal conditions. Supervisors provided staff members with daily feedback on TIF observations. In addition, all staff members completed weekly point system reliability quizzes to reinforce the accurate classification of behaviors. Specifically regarding this study, it was explained to staff members that behavior modification components were being withdrawn to evaluate their effectiveness in managing child behavior and the extent to which the treatment had beneficial effects for individual children.

Dependent Measures

The frequencies of point system behaviors prior to, during, and after the withdrawal of behavior modification were compared (see Pelham, Greiner, et al., 1997, for a description of the point system). The behaviors included in the point system are commonly identified as targets of treatment for children with ADHD, and have been used in many studies of stimulant medication in the context of the STP. The following behavioral categories are derived from this point system: (1) following activity rules; (2) noncompliance; (3) interruption; (4) complaining; (5) conduct problems (lying, stealing, destruction of property, and aggression); (6) negative verbalizations (verbal abuse to staff, teasing peers, and swearing); and (7) rule violations. For each category, the daily total of behaviors exhibited by each child was averaged across days within the behavioral treatment condition. Reliability of point system dependent measures was calculated by having independent observers code behaviors of a subset of children over entire days, each week within each group. Reliability for each of the behavioral categories was determined by computing Pearson correlations between the group counselors and the independent observer across children; average Pearson *r*s were 0.88. These measures have been shown to be reliable and sensitive to treatment effects in our previous studies (e.g., Pelham, Arnoff, et al., 1999; Pelham, Gnagy, et al., 1999; Pelham et al., 2001, 2002).

In the classroom setting, the children were assigned individual seatwork at the appropriate academic level, on which they worked independently for 30 minutes daily. Assignments included reading, language arts, and arithmetic,

as appropriate, according to each child's academic needs. Productivity (percentage of assigned seatwork completed) and accuracy (percentage of completed work that was accurate) in seatwork tasks were recorded daily. Rates of following activity rules in the three classroom segments were used as dependent measures of behavior in the classroom setting. Reliability observations were conducted on classroom rule violations, and the correlation was 0.82.

As a measure of how well each child behaved with regard to individualized target behaviors, the percentage of DRC goals met was calculated for each child each day of the assessment. This has been shown to be a sensitive, idiographic index of treatment response (Pelham et al., 2001, 2002).

Counselors and teachers also completed a series of nine questions daily. The first set of questions referred to how effective they found the behavior management techniques they used that day, how successful they were in getting each child to complete his or her tasks, how successful they were in completing their own tasks, and how effective they felt overall. These ratings were made on a 7-point Likert scale, ranging from 0 (*less effective/successful*) to 6 (*more effective/successful*). The second set of questions referred to how frustrating and stressful their interactions with each child were, with scores ranging from 0 (*less frustrating/stressful*) to 6 (*more frustrating/stressful*). The final set of questions referred to how pleasant their interactions with each child were, how much they liked each child, how much the child liked "camp," and how well each child got along with peers on that particular day, with scores ranging from 0 (*less pleasant/liked less/got along worse with peers*) to 6 (*more pleasant/liked more/got along better with peers*). Responses in each of these three domains were averaged together to produce three aggregate ratings each for counselors and teachers.

Research assistants also asked the children similar questions on the day prior to the withdrawal, during the withdrawal period, and for 2 days after the reinstatement of behavior modification. Children rated how much they liked "camp," how well they got along with other children, and how well they behaved. These ratings were made on a 7-point Likert scale, ranging from 0 to 6, with lower scores being positive and higher scores being negative.

Finally, as an index of the effectiveness of the STP, parents, counselors, and teachers completed domain-specific improvement ratings at the end of the program (e.g., Pelham, Gnagy, et al., 2000). Because the domains of improvement were intercorrelated, and not every child had problems in each domain rated, only the overall improvement ratings are reported. The results of these ratings from the end of the program are reported to provide information on the overall program effectiveness and to supplement the behavioral measures and ratings collected at the time of the treatment withdrawal.

Results

For all analyses, average daily rates of behavior were computed for three time periods: (1) the 2 weeks prior to the behavior modification withdrawal

(referred to as Treatment 1), (2) the 2-day withdrawal period (referred to as Withdrawal), and (3) the 2 weeks after the withdrawal period (referred to as Treatment 2). Because some children were participating in medication evaluations or were absent on some days, the number of days for the Treatment 1 periods ranged from 2 to 8 ($M = 6.4$), and days for the Treatment 2 period ranged from 1 to 11 ($M = 8.3$). For each set of measures, multivariate analyses of variance were then conducted with behavior modification (Treatment 1, Withdrawal, and Treatment 2) as a repeated measure and medication status (unmedicated, medicated) as a between-subjects factor. Pairwise tests were conducted between Treatment 1 and Withdrawal, and between Withdrawal and Treatment 2, when significant effects of behavior modification were found. Comparisons between Treatment 1 and Treatment 2 were also conducted to determine whether children returned to prewithdrawal rates of behavior following the reinstatement of behavior modification procedures.

Effect sizes (d) for the withdrawal were calculated by subtracting the mean for the Withdrawal period from the pooled means of the Treatment 1 and Treatment 2 periods, and dividing the difference by the pooled standard deviation of the Treatment 1 and Treatment 2 periods. Point system scores were standardized so that negative d values reflect deterioration of functioning following removal of the behavioral intervention. The pooled means and standard deviations from Treatment 1 and 2 were used as baseline rather than computing the effect of treatment relative to Withdrawal because we were measuring the impact of treatment withdrawal (i.e., the degree of deterioration observed). Because variability was greater upon reinstatement of treatment, we pooled means and standard deviations from Treatment 1 and 2 to provide a more conservative estimate of effect size.

Point System Measures

The analysis of point system measures showed a significant effect of behavior modification, $F(16, 27) = 5.67, p < .001$, but no effect of medication status or the interaction (Fs < 1.7). The effect of behavior modification was significant ($p < .015$) for all measures, and pairwise tests showed significant changes on all measures from Treatment 1 to Withdrawal and from Withdrawal to Treatment 2. Substantial effects were found on most point system behaviors in the recreational setting, with the largest effects found on noncompliance, negative verbalizations, and conduct problems. Table 2 presents means, standard deviations, and effect sizes for the behavioral manipulation. As the Table illustrates, children followed rules less and exhibited greater frequencies of negative behaviors without the intensive STP behavioral intervention in place.

When the intervention was reinstated, the children returned to levels commensurate with their Treatment 1 levels for most behaviors. Significant differences were found between Treatment 1 and Treatment 2 on conduct problems and negative verbalizations ($p < .05$). There were no significant effects of medication status or Behavior Modification $\times$ Medication interactions

TABLE 2
MEANS, STANDARD DEVIATIONS, AND EFFECT SIZE (d) FOR BEHAVIORAL MEASURES AS A FUNCTION OF BEHAVIOR MODIFICATION

	Treatment 1	Withdrawal	Treatment 2	d[a]
Percentage following activity rules	60.31 (18.28)	47.99 (25.48)	62.87 (21.53)	−.068
Frequency of rule violations	22.54 (24.63)	94.13 (201.45)	28.66 (44.76)	−1.98
Noncompliance	2.00 (1.97)	26.34 (46.59)	2.78 (4.52)	−7.38
Interruption	9.41 (10.85)	34.24 (50.30)	10.99 (17.34)	−1.71
Complaining	4.47 (4.66)	19.79 (44.27)	5.08 (7.76)	−2.42
Conduct problems	0.36 (0.63)	7.38 (23.68)	1.38 (3.09)	−3.50
Negative verbalizations	4.21 (4.61)	75.37 (180.78)	7.92 (16.31)	−6.63
Percentage following class rules				
Seatwork	73.20 (26.72)	30.23 (33.91)	71.42 (31.19)	−1.45
Peer tutoring	72.96 (22.01)	41.82 (37.17)	69.67 (27.07)	−1.20
Computer	74.78 (22.11)	55.34 (35.69)	77.50 (21.11)	−0.96
Percentage seatwork completed	72.70 (22.08)	43.20 (38.48)	74.50 (26.72)	−1.25
Percentage seatwork correct	89.31 (9.47)	77.12 (28.99)	91.31 (6.80)	−1.62
Percentage positive daily report card	70.82 (10.37)	32.08 (27.91)	73.95 (15.63)	−3.10

[a] d = Effect sizes calculated by subtracting the mean for the Withdrawal period from the pooled means of the Treatment 1 and Treatment 2 periods, and dividing by the pooled standard deviation of the Treatment 1 and Treatment 2 periods.

found on any of the point system behaviors from Treatment 1 to Treatment 2. Figure 1 depicts composite behavior category frequencies at Treatment 1, Withdrawal, and Treatment 2. All patterns of point system behavior frequencies were consistent with this illustration.

Classroom Measures

As with the point system measures, the analysis of the classroom measures showed a significant main effect of behavior modification, $F(8, 35) = 11.82$, $p < .001$, but no effect of medication status or the interaction (Fs < 1). The effect of behavior modification was significant ($p < .001$) for all measures, and pairwise tests showed significant changes on all measures from Treatment 1 to Withdrawal and from Withdrawal to Treatment 2 and no differences between Treatment 1 and Treatment 2. The magnitude of effects on classroom measures was consistently large, particularly for rule following (see Table 2).

The measure of seatwork accuracy was analyzed in a separate analysis of variance because data for this measure were missing if a child did not complete any work. Five children were excluded from this analysis because they did not complete *any* work during Withdrawal. Results from this analysis par-

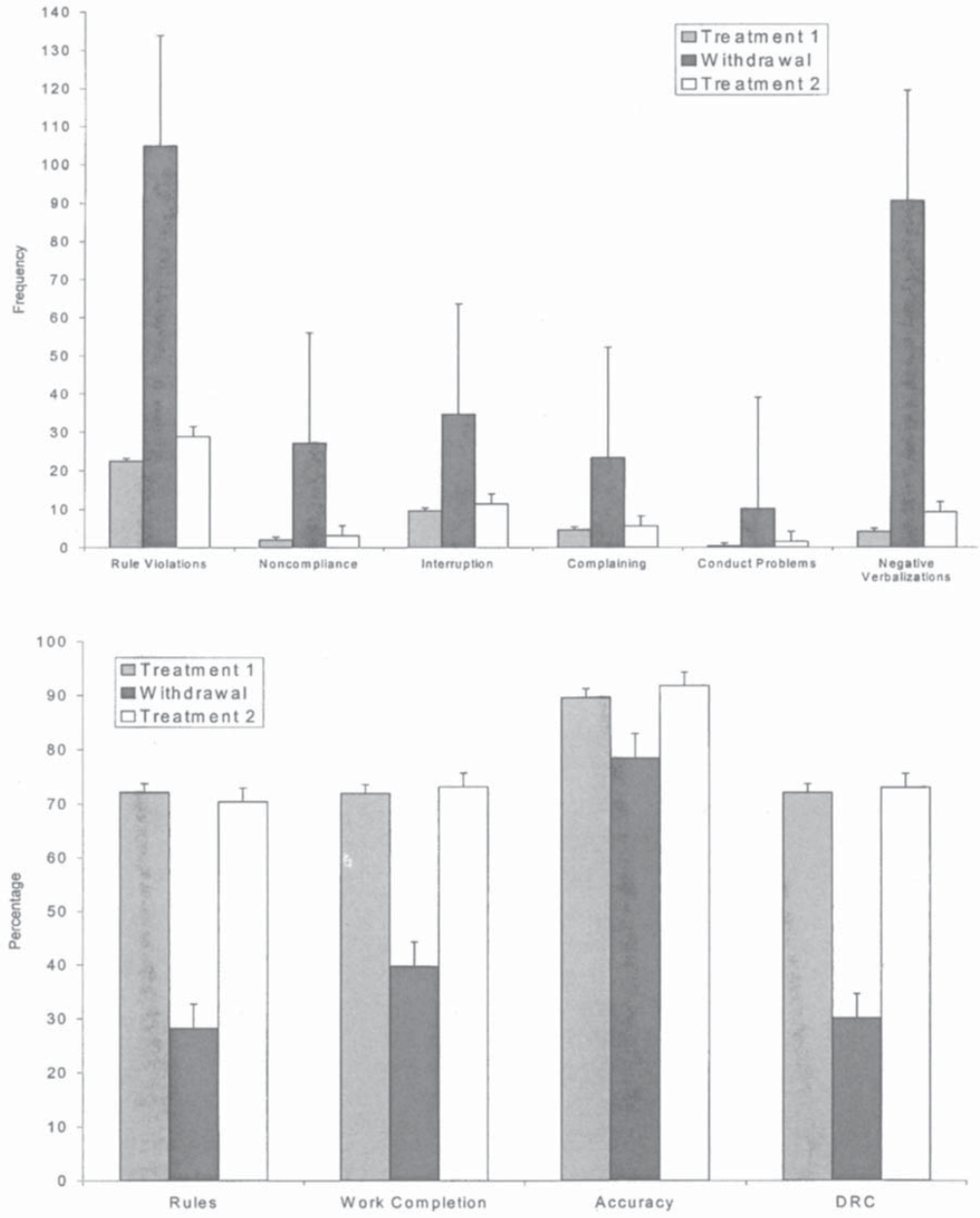

FIG. 1. Behavior at Treatment 1, Withdrawal, and Treatment 2.

alleled those of the other classroom measures, with a significant main effect of behavior modification, $F(2, 76) = 7.57, p < .01$, and no effect of medication status or the interaction. When the classroom intervention was reinstated, the children returned to levels commensurate with their Treatment 1 levels. Table 2 presents means, standard deviations, and effect sizes for classroom measures.

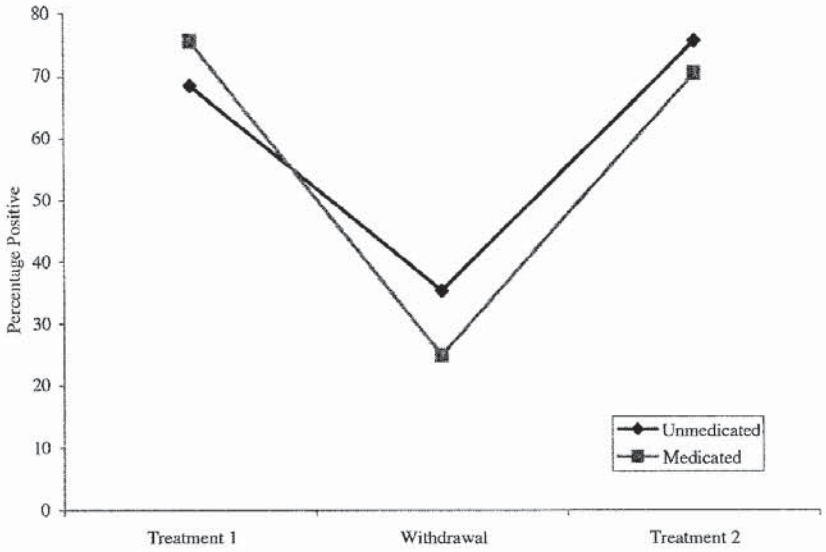

FIG. 2. Percentages of individualized target behaviors achieved on the daily report card as a function of medication and behavior modification.

DRC Targets

On the days when the behavioral treatment was not in place, children were significantly less likely to meet the individualized behavioral goals (i.e., DRC targets), $F(2, 41) = 48.94, p < .001$. The magnitude of these effects was very large (effect size = −3.10). Medication status did not produce a significant main effect, but the interaction between the factors was significant, $F(2, 41) = 3.53, p < .05$. Figure 2 illustrates the interaction and shows that medicated children had higher baseline DRC percentages than unmedicated children, but deteriorated substantially more in Withdrawal. The two groups showed similar behavioral improvement during Treatment 2, although the unmedicated group had a slightly higher DRC percentage. Notably, Figure 2 illustrates that the pattern of behavior for medicated and unmedicated children was similar, with immediate deterioration upon Withdrawal and return to Treatment 1 levels upon the introduction of Treatment 2.

Staff Ratings

The counselor ratings of their feelings of successfulness and effectiveness with the children paralleled the child behavioral measures. The analyses showed a significant effect of behavior modification on counselor effectiveness/successfulness, $F(2, 41) = 55.49, p < .001$, with no significant effect of medication status or the interaction. Similarly, ratings of frustrating and stressful interactions with the children showed a significant behavior modification effect, $F(2, 41) = 46.52, p < .001$, with no effect of medication status or the interaction. Finally, counselor ratings of pleasantness of child interactions, how much the children liked "camp," how much they liked the chil-

TABLE 3
MEANS, STANDARD DEVIATIONS, AND EFFECT SIZES (d) FOR STAFF AND CHILD RATINGS AS A FUNCTION OF BEHAVIOR MODIFICATION

	Treatment 1	Withdrawal	Treatment 2	d[a]
Counselor ratings				
Effectiveness/successfulness	5.12 (.45)	3.27 (1.69)	5.17 (.62)	3.48
Frustration/stress	.73 (.42)	2.52 (1.84)	.67 (.53)	3.79
Pleasantness/liking	5.03 (.48)	3.55 (1.15)	5.05 (.60)	2.76
Teacher ratings				
Effectiveness/successfulness	4.70 (.68)	2.81 (1.89)	4.81 (.74)	2.75
Frustration/stress	.82 (.62)	2.67 (2.03)	.84 (.78)	2.63
Pleasantness/liking	4.84 (.51)	3.40 (1.25)	4.93 (.65)	2.57
Child ratings				
How much did you like camp today?	1.74 (1.60)	1.96 (2.24)	1.37 (1.99)	.22
How well did you get along with the other kids today?	2.10 (1.31)	3.12 (2.19)	1.78 (1.71)	.78
How well did you behave today?	1.57 (1.59)	1.99 (2.15)	1.17 (1.44)	.41

Note. Counselor and teacher ratings were on a scale of 0 (*not at all*) to 6 (*very much*). Child ratings were on a scale of 0 (*very much*) to 6 (*not at all*).

[a] d = Effect sizes calculated by subtracting the mean for the Withdrawal period from the pooled means of the Treatment 1 and Treatment 2 periods, and dividing by the pooled standard deviation of the Treatment 1 and Treatment 2 periods.

dren, and how well the children got along with their peers also showed significant effects of behavior modification, $F(2, 41) = 72.83, p < .001$, with no effect of medication status or the interaction. Counselors rated themselves as more effective and successful when behavioral procedures were in place than when they were withdrawn (see Table 3). Likewise, counselors were more frustrated and stressed by their interactions with the children during the behavioral withdrawal period; they found the children more pleasant and liked the children more when behavioral procedures were in place. Finally, they reported that the children liked camp more and got along better with peers in the presence of the behavioral treatment. In all cases, planned comparisons demonstrated significant differences between Treatment 1 and Withdrawal and between Withdrawal and Treatment 2; differences between Treatment 1 and Treatment 2 counselor ratings were nonsignificant. The effect sizes in Table 3 indicate the magnitude of effects on counselor ratings was very large. Comparisons between counselor ratings at baseline and upon reinstatement revealed no significant effects of behavior modification or Behavior Modification $\times$ Medication interactions.

Ratings from the teachers showed the same effects as counselor ratings (see Table 3). All ratings showed significant and consistently large effects of behavior modification: $F(2, 41) = 61.84, p < .001$, for effectiveness/successfulness ratings; $F(2, 41) = 45.67, p < .001$, for frustration/stress ratings; and

$F(2, 41) = 69.46, p < .001$, for pleasantness/liking ratings. No effects of medication status or the interaction were found for any of the teacher ratings. Planned comparisons between teacher ratings at Treatment 1 and Withdrawal and Withdrawal and Treatment 2 were significant for all teacher ratings.

Child Self-Ratings

Three children were excluded from the analysis of child self-ratings due to missing data (e.g., the child went home before the research assistant was able to administer questionnaires). There was a significant effect of behavior modification on child ratings of how well they got along with peers, $F(2, 38) = 11.79, p = .001$. The results of planned comparisons suggested that differences between Treatment 1 and Withdrawal and between Withdrawal and Treatment 2 were significant, while differences between Treatments 1 and 2 were not. Overall, children reported getting along better with their peers when behavior modification was in place. Additionally, child self-ratings of how well they behaved showed a near-significant effect of behavior modification, $F(2, 38) = 2.90, p < .10$, such that children reported being slightly better behaved when behavior modification was in place. Child ratings of how much they liked camp were not affected by behavioral treatment.

Improvement Ratings

Improvement ratings from the end of the summer indicated that the vast majority of counselors, teachers, and parents rated virtually all of the children as at least somewhat improved, with counselors endorsing the largest magnitude of improvement, followed by parents, then teachers. Counselors rated 100% of the children as at least somewhat improved (36% somewhat improved, 55% much improved, 9% very much improved), parents rated 94% of the children as at least somewhat improved (3% no problem in this area, 3% somewhat worse, 60% somewhat improved, 34% very much improved), and teachers rated 89% as at least somewhat improved (11% no change, 66% somewhat improved, 18% much improved, and 5% very much improved).

Discussion

The current study used a BAB, treatment-withdrawal design to provide support for the efficacy of the STP treatment package for children with ADHD. Results demonstrated that the removal of the behavioral treatment components resulted in substantial increases in child misbehavior. The removal of these behavioral components also resulted in reductions in staff members' perceived successfulness/effectiveness and the pleasantness of adult-child interactions, and increases in staff frustration/stress levels. The reinstatement of the behavioral system resulted in significant improvements, virtually back to baseline levels. These results document that even though the structure, staff-to-child ratio, behavioral feedback, supervision, and sports skills training that comprise the background for the STP intervention were retained, children

displayed a substantial deterioration in behavior when contingency-management procedures were removed. The magnitude of the impact of withdrawing the behavioral treatment was very large across multiple dependent measures in multiple domains and settings, with the largest effects found on noncompliance, negative verbalizations (e.g., interruption, teasing, complaining), and conduct problems (e.g., aggression, destruction).

Very large effects of behavior modification were also found on DRC performance ($d = -3.10$). The DRC is the primary measure of individual child functioning included in this study because it provides an *idiographic* measure of how well each child behaved in the areas that were deemed most clinically important for him or her (e.g., Pelham et al., 2002). This finding suggests that the removal of behavioral treatment components resulted in deterioration in the very behaviors that were most impairing for each child, and that the reinstatement of these behavioral components resulted in recovery in these areas.

Not surprisingly, staff members working with these children reported more frustration, stress, negative interactions, and negative feelings for children when behavior modification was withdrawn. These effects were consistently quite large. Although behavioral interventions are often perceived as overwhelming and effortful on the part of the individuals that are asked to implement them (e.g., Fabiano et al., 2002), counselor and teacher ratings obtained in the current study clearly suggest the opposite — behavioral improvements that resulted from the effective use of behavioral techniques serve to *reduce* stress and frustration. This finding is consistent with treatment outcome studies that report reductions in parental stress and depression, as well as high levels of parental satisfaction, following behavioral parent training (e.g., Anastopoulos, Shelton, DuPaul, & Guevremont, 1993; Pelham, Burrows-MacLean, Gnagy, Fabiano, et al., 2003; Webster-Stratton, 1990).

Finally, when the children's counselors, parents, and teachers were asked to rate each child's improvement following the STP, virtually every child was rated as at least somewhat improved in overall functioning. Such a finding is important because ADHD is now commonly conceptualized as a *chronic* condition (American Academy of Pediatrics, 2001). Effective treatments for ADHD will be those that teach adaptive skills, reduce problematic behaviors that cause impairment in daily life functioning, and educate parents and others charged with the child's care to cope with the challenges that accompany an ADHD diagnosis over the course of development. Therefore, improvement in daily life functioning and reducing functional impairment are the important, socially valid standards against which treatment outcome must be judged.

Another interesting finding was that for most dependent measures, frequencies of negative behaviors and staff ratings returned to prewithdrawal rates; however, the negative verbalization and conduct problem composites were significantly higher when behavioral treatment was reinstated. The brief withdrawal and reinstatement mimics an inconsistently applied behavioral treatment. Although these differences were small and did not appear to be clinically significant, this finding highlights the need to implement behavioral treatments

on a consistent basis in order to maintain optimal outcomes (Chronis et al., 2001; Pelham & Fabiano, 2000). This may be the case for oppositional or conduct problems, such as talking back or aggression, in particular, as reflected by these behavior categories. These issues have obvious implications for situations in which children interact with multiple teachers or between multiple households in which behavioral treatments are not being implemented consistently.

These findings can also add to the current discussion on the appropriateness of group treatment for children with behavior problems. Dishion, McCord, and Poulin (1999) have recommended that children with conduct problems not be treated in group settings, because "deviancy training" may occur. Based on the results of the present study, in which the majority of the participants had comorbid ODD or CD, that supposition may need to be qualified. In the context of a *regular summer camp or unstructured group therapy program* where behavioral contingencies are not in place, such deviancy training may occur. In this study, this phenomenon was evidenced by the substantial increases in negative verbalizations and conduct problems and reinforcement for these behaviors from peers during the Withdrawal condition that led to early reinstatement of the behavioral intervention for two of the four groups, as well as the failure to completely reverse to Treatment 1 levels in the Treatment 2 condition with regard to negative verbalizations and more serious conduct problems. However, prior to the Withdrawal, there was no evidence of deviancy training, and the reinstatement of behavioral interventions during Treatment 2 returned the rates of these problem behaviors to near pre-Withdrawal levels. These results inform clinicians that deviancy training does not appear to be a concern in group treatments like the STP that *utilize evidence-based, contingency management interventions* to target negative behaviors.

Notably, the effects of the STP treatment package were evident regardless of medication status. In fact, medicated children showed the most extreme response to the withdrawal of behavior modification — during the Withdrawal period, they exhibited an average of 17.6 conduct problems per day, compared to only 2.6 conduct problems in the unmedicated group. Clearly, medication alone was insufficient to control their behavior. The most parsimonious explanation for this outcome is that the children in this study who were medicated during the summer were the children who exhibited the most severe behavior problems (this is supported by the parent intake ratings on the I/O and O/D factors of the IOWA Conners Parent Rating Scale). Alternatively, it could be speculated that children who were unmedicated may have learned more skills for dealing with frustrations and challenges by virtue of the fact that they had to negotiate such challenges without the aid of medication. Overall, these findings suggest that intensive behavioral treatment had incremental benefit over medication for children who exhibited the most severe forms of problem behavior, and highlights the importance of *consistent* implementation of behavioral interventions for children whether they are medicated or not.

The results of this investigation should be placed within the context of previous studies conducted within the STP. For example, a number of medication studies (e.g., Pelham et al., 1990; Pelham, Aronoff, et al., 1999; Pelham, Gnagy, et al., 1999; Pelham & Hoza, 1987; Pelham et al., 2002) have illustrated large effects of medication across the STP activities. However, behavioral treatment was not manipulated in these studies. The fact that medicated children's behavior in the present study deteriorated as much as the behavior of unmedicated children when behavioral treatment components were withdrawn suggests that a substantial portion of the behavioral improvement attributed to medication in these studies may have been due to the STP intervention itself. Because medication was not manipulated in the current study, we were unable to determine the relative effects of behavior modification and medication. In the one study where children were randomly assigned to participate in the STP either steadily medicated or unmedicated (Pelham, Gnagy, et al., 2000), children improved on the vast majority of measures when contingency management was implemented, and there were few differences between medicated and unmedicated groups. In light of the current study, we can speculate that in the Pelham, Gnagy, et al. (2000) study, even though behavioral treatment was not manipulated, the STP was responsible for the majority of the participants' improved behavior, and that the medicated group obtained little benefit from the adjunctive medication. Future studies that include a larger sample and manipulate medication and behavior modification within subjects are necessary to further investigate the relative effects of medication and behavior modification within the context of the STP.

These promising findings regarding the efficacy of the STP should be considered in light of several methodological limitations. First, a true baseline of behavior in the STP setting was not obtained prior to the initial introduction of contingency management components. As a result, the no-treatment component could only be examined following the intensive behavioral intervention. Had we included an initial period without behavior modification at the beginning of the program (i.e., a true baseline) to which we could compare the effects of the intensive intervention, we could more clearly interpret the effects. With the current design, the withdrawal of the active treatment components provided a sharp contrast to the intensive STP intervention. That is, children were accustomed to consistent positive and negative consequences for their behavior in the STP. During the Withdrawal period, they received privileges noncontingently that had previously been contingent upon their behavior. It is entirely possible that some portion of the deterioration in child behavior in the absence of behavioral treatment components reflected "contrast effects." Perhaps children responded to this change by "taking advantage of the situation" and engaging in high rates of misbehavior. Moreover, had a longer withdrawal period been possible and not clinically contraindicated, we may have seen a decrease in elevated rates of negative behavior once the novelty of the situation wore off. This phenomenon could then be understood as "extinction bursting," in which a dramatic increase in negative behavior is

observed shortly following the removal of a reward before the negative behavior begins to weaken (Kalish, 1981). Contrary to this speculation, another study conducted in a subsequent STP using similar methodology that included two, week-long withdrawal periods indicated that behavior *worsened* with successive days in the no-behavior-modification condition and upon the second withdrawal for many children (Coles et al., in press; Pelham, Burrows-MacLean, Gnagy, Fabiano, et al., 2003). Further, there were individual differences in the rate at which behavior improved during reinstatement (Coles et al., in press). The substantial behavioral changes during the Withdrawal condition are clinically important, whether a contrast effect or not, as such a socially inappropriate response to changes in contingencies may be viewed as an additional area of impairment for children with ADHD. To address these methodological issues, future studies of this type should include an initial baseline period and should withdraw the behavioral intervention for longer periods of time to test whether an increase in negative behavior is maintained. These studies should also investigate the length of time it takes children to regain improvements in behavior following disruptions in treatment (Coles et al., in press).

It is important to note that many widely cited studies of stimulant medication effects conducted in the STP setting have the same methodological limitations described above (e.g., rapidly alternating days of placebo versus medication). However, the manipulation of medication and behavior modification in this manner is different in that medication studies permit the use of a placebo, so that children, parents, and treatment providers may remain blind to treatment condition. The removal of behavior modification is far more obvious to the children and staff. However, behavioral frequencies utilized as primary outcome measures in the STP are objective, and careful measures of adherence to the treatment protocol were utilized to ensure that staff interactions with children remained consistent across conditions (aside from differences outlined in the protocol), thus minimizing the potential influence of staff expectations on these results.

The current study manipulated only two intensities of behavior modification—the absence of any behavioral treatment components versus the presence of an intensive, multimodal program. There remains a need to evaluate varying intensities of behavioral interventions. For example, behavior modification offered in the community typically involves parent training with or without school interventions and/or social skills training. The STP is an intensive, multimodal program that involves a substantially greater amount of training and resources to implement. Thus, an important question involves the extent to which the STP results in greater benefits than these more widespread, less intensive approaches. The MTA study utilized an intensive, behavioral intervention including an STP, and showed positive effects of behavioral treatment (see discussion below). Whether a less intensive behavioral treatment would have resulted in the same changes was not assessed (Pelham, 1999). We are currently conducting a study that addresses many of the methodological

limitations of the present investigation (Pelham, Burrows-MacLean, Gnagy, Coles, et al., 2003). In this study, three levels of behavior modification intensity are crossed with three doses of medication. Specifically, groups of children receive each of three behavior modification conditions for 3-week periods during the course of the summer: (1) high-intensity, (2) low-intensity, and (3) no behavior modification. Preliminary results show dose-response curves of behavior modification similar to those obtained with medication.

The current study represents an important step in establishing the *efficacy* of the STP package in a controlled manner. Additional studies of the STP in community settings and by independent research groups are needed to establish its *effectiveness*. Thus far, the Kolko et al. study lends support to the effectiveness of the STP when conducted by an independent research group with a diverse group of children, and like the current study, found that behavior modification improved behavior in both recreational and classroom settings. However, Kolko and colleagues found incremental effects of behavior modification beyond the effects of medication only in the recreational setting (not in the classroom). Additional studies must be conducted by independent research groups to determine whether STP effects are, in fact, comparable to those found in studies conducted by Pelham and colleagues, and to examine variables that may contribute to the magnitude of effects across settings (e.g., treatment integrity and fidelity). The MTA study, which included the STP, demonstrated that the STP could be exported to six academic sites that had never conducted the program. Exportability to nonacademic settings is the next step.

Given that the paraprofessionals directly implementing the STP in academic settings are typically undergraduate students, who are trained and supervised by senior staff members, it is reasonable to expect that community agency staff members can also be trained to implement the STP. In fact, the STP is currently offered at a number of nonacademic mental health agencies in North America. Staff members at these agencies attend an intensive training program conducted by Dr. Pelham's staff in Buffalo and Pittsburgh and are provided with regular consultation. Although the effectiveness of the STP offered in these community settings has not yet been systematically evaluated, parent satisfaction and improvement ratings appear comparable to data collected in our university-based programs (Pelham et al., in press; Pelham, Manos, & Janakovic, 2000). Such *effectiveness* studies are of utmost importance in establishing the transportability of academic-based programs into community settings (Pelham, Manos, & Janakovic, 2000).

Finally, some discussion of the ethical considerations in utilizing a treatment-withdrawal design is necessitated. We decided to utilize the current design to maximize the amount of treatment that children received — that is, in this study, children were only without intensive treatment for a 2-day period at most. A multiple baseline design was too unwieldy in an 8-week program — that is, it would force us to refrain from providing the maximum

level of treatment intensity for some time — and would necessarily confound the amount of treatment received (i.e., time in the program) with the addition of behavioral treatment components. Still, there was some concern about the withdrawal of behavior modification for entire groups of severely impaired children. Necessary precautions were in place prior to the withdrawal of treatment components to ensure that treatment would be promptly reinstated for children for whom the withdrawal was clinically contraindicated. At the same time, the manipulation allowed an assessment of individual children's responses to behavior modification components and the extent to which behavioral gains generalized to times when contingencies were not in place.

In summary, the STP resulted in substantial behavioral effects across multiple measures of functioning in important domains and across multiple settings. The effect sizes of the STP are comparable to or surpass the effects of medication from other studies conducted in the same setting. This report *clearly* documents the efficacy of the STP and provides very strong evidence that behavior modification procedures are effective for treating important areas of impairment in children with ADHD. This conclusion is in contrast to several prominent publications that state that behavior modification is not effective or not as effective as medication for treating ADHD (e.g., Jadad et al., 1999; MTA Cooperative Group, 1999). Of course, at the same time, findings obtained in the present study are similar to those found in studies of stimulant medication and suggest that the effects of the STP were not maintained for even a day without the continued presence of behavior modification. This finding may help to explain the somewhat smaller effects of behavior modification relative to medication reported at endpoint in the MTA study, in which outcome measures were collected 4 to 6 months after intensive behavior modification and behavioral consultation had ended, but while medication was active and doses were at their highest levels (Pelham, 1999). Indeed, ongoing behavioral parent training and follow-up behavioral interventions in community schools are considered crucial components of the STP treatment package if maintenance of acute treatment gains is expected. We have long argued that behavioral interventions for a chronic disorder such as ADHD must be continued at some level of intensity in all settings in which a child is impaired over the long term in order to attain optimal outcomes (Chronis et al., 2001; Pelham & Fabiano, 2000). Indeed, 1-year follow-up in the MTA study reveals that the behaviorally treated group maintained their gains far better than the medicated groups, with far fewer medicated children (MTA Cooperative Group, 2004). Since the MTA behavioral condition included parent training, school intervention, and the STP, this supports the effectiveness of intensive and comprehensive behavioral treatment. In summary, intensive behavior modification programs such as the STP are powerful interventions for improving the functioning of children with ADHD, and should therefore be considered a core treatment component of any ADHD intervention.

References

Abramowitz, A. J., Eckstrand, D., O'Leary, S. G., & Dulcan, M. K. (1992). ADHD children's responses to stimulant medication and two intensities of a behavioral intervention. *Behavior Modification, 16*, 193–203.

American Academy of Pediatrics. (2001). Clinical practice guideline: Treatment of the school-aged child with attention-deficit/hyperactivity disorder. *Pediatrics, 108*, 1033–1044.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anastopoulos, A. D., Shelton, T. L., DuPaul, G. J., & Guevremont, D. C. (1993). Parent training for attention-deficit hyperactivity disorder: Its impact on parent functioning. *Journal of Abnormal Child Psychology, 21*, 581–596.

August, G. J., Realmuto, G. M., Hektner, J. M., & Bloomquist, M. L. (2001). An integrated components preventive intervention for aggressive elementary school children: The Early Risers Program. *Journal of Consulting and Clinical Psychology, 69*, 614–626.

Carlson, C. L., Pelham, W. E., Milich, R., & Dixon, J. (1992). Single and combined effects of methylphenidate and behavior therapy on the classroom performance of children with ADHD. *Journal of Abnormal Child Psychology, 20*, 213–232.

Chronis, A. M., Fabiano, G. A., Gnagy, E. M., Wymbs, B., Burrows-MacLean, L., & Pelham, W. E. (2001). Comprehensive, sustained behavioral and pharmacological treatment for ADHD: A case study. *Cognitive and Behavioral Practice, 8*, 346–359.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 779–862). New York: John Wiley & Sons.

Coles, E. K., Pelham, W. E., Gnagy, E. M., Burrows-MacLean, L., Fabiano, G. A., Chacko, A., et al. (in press). Treatment response to a comprehensive behavioral intervention. *Journal of Emotional and Behavioral Disorders*.

Cunningham, C. E., Bremner, R., & Secord-Gilbert, M. (1997). *COPE: The community parent education program: A school-based family systems oriented course for parents of children with disruptive behavior disorders*. Hamilton, ON: COPE Works.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755–764.

Fabiano, G. A., Pelham, W. E., Gnagy, E. M., Kipp, H., Lahey, B. B., Burrows-MacLean, L., Chronis, A. M., Onyango, A. N., & Morrisey, S. (1999, November). *The reliability and validity of the children's impairment rating scale: A practical measure of impairment in children with ADHD*. Poster presented at the annual meeting of the Association for Advancement of Behavior Therapy, Toronto, Ontario.

Fabiano, G. A., Pelham, W. E., Manos, M. J., Gnagy, E. M., Chronis, A. M., Onyango, A. N., Lopez-Williams, A., Burrows-MacLean, L., Coles, E. K., & Meichenbaum, D. L. (2004). An evaluation of three time-out procedures for children with attention-deficit/hyperactivity disorder. *Behavior Therapy, 35*, 449–469.

Fabiano, G. A., Pelham, W. E., Pisecco, S., Evans, S. W., Manos, M. J., Caserta, D., Hannah, J. N., & Johnston, C. (2002, November). *A nationally representative survey of classroom-based, behavior modification treatment for ADHD*. Poster presented at the annual meeting of the Association for Advancement of Behavior Therapy, Reno, NV.

Hoza, B., Pelham, W. E., Sams, S. E., & Carlson, C. L. (1992). An examination of the "dosage" effects of both behavior therapy and methylphenidate on the classroom performance of two ADHD children. *Behavior Modification, 16*, 164–192.

Jadad, A. R., Boyle, M., Cunningham, C., Kim, M., & Schachar, R. (1999). *Treatment of attention-deficit hyperactivity disorder: Evidence report/technology assessment no. 11*. Rockville, MD: Agency for Healthcare Research and Quality.

Kalish, H. I. (1981). *From behavioral science to behavior modification*. New York: McGraw-Hill.

Kolko, D. J., Bukstein, O. G., & Barron, J. (1999). Methylphenidate and behavior modification in children with ADHD and comorbid ODD and CD: Main and incremental effects across settings. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 578–586.

Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in behavioral pediatrics. *Advances in Developmental and Behavioral Pediatrics, 3*, 113–147.

MTA Cooperative Group. (1999). 14-Month Randomized Clinical Trial of Treatment Strategies for Attention Deficit Hyperactivity Disorder. *Archives of General Psychiatry, 56*, 1073–1086.

MTA Cooperative Group. (2004). The NIMH MTA Follow-up: 24-month Outcomes of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder (ADHD). *Pediatrics, 113*, 754–761.

Pelham, W. E. (1999). The NIMH Multimodal Treatment Study for ADHD: Just say yes to drugs alone? *Canadian Journal of Psychiatry, 44*, 981–990.

Pelham, W. E. (2002). *Attention-deficit/hyperactivity disorder: Diagnosis, nature, etiology, and treatment*. Buffalo, NY: CTADD.

Pelham, W. E., Aronoff, H., Midlam, J., Shapiro, C., Gnagy, E., Waxmonsky, J., Chronis, A. M., Onyango, A. N., Forehand, G., Nguyen, A., Hauber, A., Barry, P., & Kaye, D. (1999). A comparison of Ritalin and Adderall: Efficacy and time-course in children with ADHD. *Pediatrics, 103*, E3.

Pelham, W. E., Burrows-MacLean, L., Gnagy, E. M., Coles, E. K., Wymbs, B. T., Chacko, A., Walker, K., Arnold, F., Keenan, J. K., Onyango, A. N., Fabiano, G. A., Hoffman, M. T., & Massetti, G. M. (2003). *A dose-ranging study of behavioral and pharmacological treatment for children with ADHD*. Poster presented at the annual meeting of the Association for Advancement of Behavior Therapy, Boston, MA.

Pelham, W. E., Burrows-MacLean, L., Gnagy, E. M., Fabiano, G. A., Coles, E. K., Tresco, K. E., Chacko, A., Wymbs, B. T., Wienke, A. L., Walker, K., & Hoffman, M. T. (2003). *Transdermal methylphenidate, behavioral, and combined treatment for children with ADHD*. Manuscript submitted for publication.

Pelham, W. E., Carlson, C., Sams, S. E., Vallano, G., Dixon, M. J., & Hoza, B. (1993). Separate and combined effects of methylphenidate and behavior modification on the classroom behavior and academic performance of ADHD boys: Group effects and individual differences. *Journal of Consulting and Clinical Psychology, 61*, 506–515.

Pelham, W. E., & Fabiano, G. (2000). Behavior modification. *Child and Adolescent Psychiatric Clinics of North America, 9*, 671–688.

Pelham, W. E., Fabiano, G. A., Gnagy, E. M., Greiner, A. R., & Hoza, B. (in press). Comprehensive psychosocial treatment for ADHD. In E. Hibbs & P. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice*.

Pelham, W. E., Gnagy, E. M., Burrows-Maclean, L., Williams, A., Fabiano, G. A., Morrissey, S. M., Chronis, A. M., Forehand, G. L., Nguyen, C. A., Hoffman, M. T., Lock, T. M., Fielbelkorn, K., Morse, E., Coles, E. K., Panahon, C. J., Steiner, R. L., Meichenbaum, D. L., & Onyango, A. N. (2001). *Once-a-day Concerta™ methylphenidate versus t.i.d. methylphenidate in laboratory and natural settings*. Retrieved June 7, 2004, from http://www.pediatrics.org/cgi/content/full/107/6/e105.

Pelham, W. E., Gnagy, E. M., Chronis, A. M., Burrows-MacLean, L., Fabiano, G. A., Onyango, A. N., Meichenbaum, D. L., Williams, A., Aronoff, H. R., & Steiner, R. L. (1999). A comparison of morning-only and morning/late afternoon Adderall to morning-only, twice-daily, and three times-daily methylphenidate in children with attention-deficit/hyperactivity disorder. *Pediatrics, 104*, 1300–1311.

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 210–218.

Pelham, W. E., Gnagy, E. M., Greiner, A. R., Hoza, B., Hinshaw, S. P., Swanson, J. M., Simpson, S., Shapiro, C., Bukstien, O., & Baron-Myak, C. (2000). Behavioral vs. behavioral

and pharmacological treatment in ADHD children attending a summer treatment program. *Journal of Abnormal Child Psychology, 28*, 507–526.

Pelham, W. E., Greenslade, K. E., Vodde-Hamilton, M. A., Murphy, D. A., Greenstein, J. J., Gnagy, E. M., Guthrie, K. J., Hoover, M. D., & Dahl, R. E. (1990). Relative efficacy of long-acting stimulants on children with attention deficit-hyperactivity disorder: A comparison of standard methylphenidate, sustained-release methylphenidate, sustained-release dextroamphetamine, and pemoline. *Pediatrics, 86*, 226–237.

Pelham, W. E., Greiner, A. R., & Gnagy, E. M. (1997). *Summer treatment program manual*. Buffalo, NY: Comprehensive Treatment for Attention Deficit Disorders, Inc.

Pelham, W. E., & Hoza, B. (1996). Intensive treatment: A summer treatment program for children with ADHD. In E. Hibbs & P. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (pp. 311–340). New York: APA Press.

Pelham, W. E., & Hoza, J. (1987). Behavioral assessment of psychostimulant effects on ADD children in a Summer Day Treatment Program. In R. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 3, pp. 3–33). Greenwich, CT: JAI Press.

Pelham, W. E., Hoza, B., Kipp, H. L., Gnagy, E. M., & Trane, S. T. (1997). Effects of methylphenidate and expectancy on ADHD children's performance, self evaluations, persistence, and attributions on a cognitive task. *Experimental and Clinical Psychopharmacology, 5*, 3–13.

Pelham, W. E., Hoza, B., Pillow, D. R., Gnagy, E. M., Kipp, H. L., Greiner, A. R., Waschbusch, D. A., Trane, S. T., Greenhouse, J., Wolfson, L., & Fitzpatrick, E. (2002). Effects of methylphenidate and expectancy on children with ADHD: Behavior, academic performance, and attributions in a summer treatment program and regular classroom settings. *Journal of Consulting and Clinical Psychology, 70*, 320–335.

Pelham, W. E., Manos, M., & Janakovic, F. (2000, August). Intensive behavioral treatment for children with ADHD. In J. R. Weisz (Chair), *Making evidence based treatments work in clinical practice*. Symposium conducted at the meeting of the American Psychological Association, Washington, DC.

Pelham, W. E., Milich, R., Murphy, D. A., & Murphy, H. A. (1989). Normative data on the IOWA Conners teacher rating scale. *Journal of Clinical Child Psychology, 18*, 259–262.

Pelham, W. E., & MTA Cooperative Group. (2003). *Parent and teacher evaluation of treatment in the MTA: Consumer satisfaction and perceived effectiveness*. Manuscript submitted for publication.

Pelham, W. E., & Murphy, H. A. (1986). Attention deficit and conduct disorders. In M. Hersen (Ed.), *Pharmacological and behavioral treatment: An integrative approach* (pp. 108–148). New York: John Wiley and Sons.

Pelham, W. E., Wheeler, T., & Chronis, A. (1998). Empirically supported psychosocial treatments for ADHD. *Journal of Clinical Child Psychology, 27*, 190–205.

Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *Journal of Abnormal Child Psychology, 18*, 479–492.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition: Manual*. San Antonio: The Psychological Corporation.

Wechsler, D. (1992). *Wechsler Individual Achievement Test: Manual*. San Antonio: The Psychological Corporation.

Wells, K. C., Pelham, W. E., Kotkin, R. A., Hoza, B., Abikoff, H. B., Abramowitz, A., Arnold, L. E., Cantwell, D. P., Conners, C. K., Del Carmen, R., Elliott, G., Greenhill, L. L., Hechtman, L., Hibbs, E., Hinshaw, S. P., Jensen, P. S., March, J. S., Swanson, J. M., & Schiller, E. (2000). Psychosocial treatment strategies in the MTA study: Rationale, methods, and critical issues in design and implementation. *Journal of Abnormal Child Psychology, 28*, 483–505.

RECEIVED: April 4, 2003
ACCEPTED: August 11, 2003